

Wprowadzenie

W ostatnich latach metody mikroekonometryczne zdobywają coraz większą popularność i uznanie badaczy. Jest to związane przede wszystkim z rozwojem technik gromadzenia i przetwarzania danych. Mikroekonometria to grupa metod przeznaczona do analizy mikrodanych. Współcześnie mikrodane, czyli dane jednostkowe, są gromadzone w wielu dziedzinach gospodarki takich jak finanse, ubezpieczenia, rynek pracy czy badanie zachowania gospodarstw domowych.

Jednym z obszarów analiz ilościowych prowadzonych na poziomie mikro jest analiza skuteczności działań podejmowanych przez instytucje publiczne. Dotyczy to przede wszystkim różnych form wsparcia takich jak szkolenia podnoszące kwalifikacje pracowników, szkolenia dla osób bezrobotnych czy programy wsparcia dla firm. Z reguły takie działania są skierowane do specyficznych grup odbiorców. Występujące zjawisko selekcji lub samoselekcji próby uniemożliwia przeprowadzenie ilościowej oceny skuteczności tych działań z wykorzystaniem klasycznych narzędzi ekonometrycznych (Strawiński 2007).

Metoda łączenia danych jest narzędziem pozwalającym na ocenę skuteczności podejmowanych działań o charakterze programów. Jest to grupa metod, wywodząca się z nauk eksperymentalnych i przeniesiona na grunt dziedzin nauki, w których prowadzenie w pełni kontrolowanych eksperymentów nie jest możliwe albo jest w znacznym stopniu ograniczone. Na przykład, nie jest możliwe przeprowadzenie eksperymentalnego badania skuteczności leku, który może zagrażać zdrowiu lub życiu (Szulc 2012).

Łączenie danych pozwala uzyskać odpowiedź na pytanie, „co by się stało, gdyby?”. Nie jest ono pytaniem nowym ani specyficznym

dla nauk ekonomicznych. W naukach eksperymentalnych badacze poszukują na nie odpowiedzi, dokonując serii eksperymentów. W dziedzinach, w których nie jest to możliwe, wykorzystuje się dostępny aparat matematyczny do nadania właściwości zbioru eksperymentalnego danym niepochodzącym z eksperymentu. Jedną z takich metod jest łączenie danych. Umożliwia ona przeprowadzenie analizy kontrfaktycznej. Jednym ze sposobów łączenia danych, który w ostatnich latach uzyskał uznanie badaczy, jest łączenie według prawdopodobieństwa (ang. *Propensity Score Matching*). Zastosowanie metody nie ogranicza się do nauk ekonomicznych, jest ona z powodzeniem wykorzystywana w medycynie, naukach politycznych czy naukach rolniczych.

Niniejsza książka jest pogłębionym studium własności statystycznych estymatorów metody *Propensity Score Matching*, stosowanych do szacowania przeciętnych efektów oddziaływania. Asymptotyczne właściwości estymatorów zostały już przez badaczy opisane, brakuje natomiast badań dotyczących prób o niewielkiej liczbie obserwacji. Niniejsza książka jest próbą wypełnienia tej luki. Efekty oddziaływania można szacować, wykorzystując wiele sposobów. Dwoma najczęściej wykorzystywanymi miarami są przeciętny efekt oddziaływania oraz przeciętny efekt oddziaływania wobec jednostek poddanych oddziaływaniu. Pierwsza wskazuje, jaki jest efekt działania programu wobec przeciętnej jednostki w populacji. Druga miara opisuje efekt działania programu wobec jednostek w tym programie uczestniczących. W kontekście szacowania tych wielkości zostaną ukazane właściwości statystyczne estymatorów uzyskiwanych z wykorzystaniem dwóch metod bazujących na oszacowaniach *propensity score*, czyli prawdopodobieństwa poddania oddziaływaniu. Wspomnianymi metodami będą łączenie danych według prawdopodobieństwa oraz ważenie odwrotnością oszacowania prawdopodobieństwa.

Pewną trudnością, która stała przed autorem, jest brak ugruntowanej terminologii w języku polskim dotyczącej metod łączenia danych. Jedyną znaną autorowi pozycją o charakterze akademickim, w której tematyka łączenia danych jest omawiana, jest podręcznik *Mikroekonometria* pod redakcją prof. dr. hab. Marka Gruszczyńskiego. Kilka lat temu, przed powstaniem pierwszego wydania podręcznika, próbo-

waliśmy uzgodnić jednolite tłumaczenia anglojęzycznych terminów. W wielu przypadkach się to udało, jednak drobne różnice pozostały. Przykładem takiej rozbieżności jest tłumaczenie na język polski pojęcia *matching estimation*. W podręczniku *Mikroekonometria* zaproponowano „estymacja przez dopasowanie” (rozdział 9 autorstwa Adama Szulca), ja proponuję „estymacja przez łączenie”. Obie propozycje są dosłownymi tłumaczeniem pojęcia *matching*. Oznacza ono zarówno dopasowywanie, dobieranie, jak i łączenie w pary. Wydaje mi się, że termin „łączenie” lepiej oddaje charakter pojęcia, gdyż jego istotą jest łączenie podobnych obserwacji, najczęściej w pary. Ponadto w ekonometrii słowo „dopasowanie” jest kojarzone ze stopniem dopasowania modelu do danych empirycznych. Aby uniknąć niejasności terminologicznej, na końcu niniejszej książki jest umieszczony załącznik, w którym zestawiono stosowane tłumaczenia na język polski terminów anglojęzycznych. Dodatkowo w celu uniknięcia zamieszania, w książce obok określeń polskich pojawiają się ich anglojęzyczne odpowiedniki.

Osobnego komentarza wymaga wyjaśnienie tytułu książki. Jego pierwszy człon jest w języku angielskim, nie ma bowiem utrwalonego w literaturze tłumaczenia go na język polski. Moje tłumaczenie „łączenie według prawdopodobieństwa” stara się oddać istotę metody, jednak nawet dla specjalistów w dziedzinie mikroekonometrii mogłoby być niezrozumiałe. W pracy oba pojęcia będą pojawiać się wymiennie, chociaż *Propensity Score Matching* częściej. Ta sama uwaga dotyczy pojęcia *propensity score* i jego tłumaczenia – „prawdopodobieństwo oddziaływania”.

Na książkę składają się trzy rozdziały. Zadaniem pierwszego jest wprowadzenie czytelnika w świat efektów oddziaływania i łączenia danych. Ma on charakter teoretycznego przeglądu najważniejszych prac z literatury dotyczącej analizy kontrfaktycznej i łączenia danych. Zaprezentowano fundamenty analizy kontrfaktycznej, które wywodzą się z prac Neymana (1923a i 1923b) i Fishera (1925) na temat randomizacji. Opisano dwa podstawowe efekty oddziaływania oraz warunki ich identyfikacji. Opierając się na fundamentalnej pracy Rosenbauma i Rubina (1983), omówiono metodę łączenia obserwacji oraz propozycję autorów wykorzystania *propensity score* do łączenia observa-

cji. W dalszej części rozdziału przedstawiono założenia i statystyczne właściwości metody *Propensity Score Matching* w dużych próbach. Fragment tej części bazuje na mojej rozprawie doktorskiej. Wartością dodaną jest zebranie w jednym miejscu wyników najważniejszych prac z ostatnich 15 lat. Rozdział zamyka prezentacja różnych algorytmów używanych do łączenia danych oraz narzędzi diagnostycznych pozwalających ocenić jakość uzyskanego łączenia. Omówiono różne warianty algorytmu dobierania najbliższego sąsiada – jego zaletą jest szybkość, a wadą niejednoznaczność. Innym algorytmem jest ważenie odwrotnością oszacowania prawdopodobieństwa, którego jakość zależy m.in. od jakości przyjętego modelu dla prawdopodobieństwa. Rozdział kończy prezentacja dyskusji naukowej, która toczyła się na początku XXI wieku i dotyczyła zagadnienia, czy wykorzystując metody łączenia według prawdopodobieństwa, można odtworzyć wyniki eksperymentalne.

W rozdziale drugim zaprezentowano wyniki badań symulacyjnych nad właściwościami estymatorów metody łączenia według prawdopodobieństwa w małych próbach. Własności metod statystycznych w małych próbach są kluczowe w praktycznym zastosowaniu narzędzi ekonometrycznych. W konkretnych problemach badawczych, np. w pomiarze efektu interwencji na rynku pracy, badacz zazwyczaj staje przed oceną ilościowego efektu oddziaływania na podstawie badania o charakterze pilotażowym. Cechą charakterystyczną badań pilotażowych jest ich niewielka skala, a zatem niewielka liczebność próby. Ważna jest więc informacja, czy właściwości statystyczne metod łączenia, które udowodniono, korzystając z twierdzeń granicznych, są również zachowane w próbach o liczebnościach zbliżonych do liczebności prób rzeczywistych. W literaturze brakuje ogólnych prac dotyczących własności estymatorów metody *Propensity Score Matching* w małych próbach. W rozdziale zaprezentowano wyniki własnych oryginalnych badań symulacyjnych. Struktura eksperymentalnego planu symulacyjnego jest analogiczna do moich wcześniejszych prac. Wykorzystany plan eksperymentalny oraz część prezentowanych wyników dla estymatorów przeciętnych efektów oddziaływania wobec jednostek poddanych oddziaływaniu zostały wcześniej opublikowane przeze mnie w artykułach. Cechą odróżniającą uzyskane wyniki od wcześniejszych rezultatów jest

cel badania oraz jego dużo szerszy zakres. Analizowano dwa sposoby szacowania efektów oddziaływania oraz wpływ różnych technik wymuszających spełnienie założeń metody *Propensity Score Matching* na statystyczne własności estymatorów dla efektów oddziaływania. Analizowano standardowe miary jakości: obciążenie, wariancję i pierwiastek błędu średniokwadratowego estymatora. Nowym elementem niewystępującym we wcześniejszych pracach są analizy rozmiaru procedury łączenia i mocy testów istotności estymatora oszacowanej wartości efektu oddziaływania.

Celem trzeciego rozdziału jest próba prezentacji zastosowań i właściwości statystycznych omawianych metod zastosowanych do analizy rzeczywistych zbiorów danych. W rozdziale zaprezentowano cztery przykłady. Pierwszy z nich jest oryginalnym badaniem pokazującym, w jaki sposób można wykorzystać dostępne dane pochodzące z Badania Aktywności Ekonomicznej Ludności (BAEL) do oceny skuteczności dodatkowych szkoleń podnoszących kwalifikacje zawodowe oraz ich wpływu na sytuację uczestniczących w nich pracowników. Pozostałe przykłady zostały zaadaptowane z literatury. Drugi z nich powstał na podstawie artykułu, którego jestem współautorem. W przykładzie są porównane relatywne korzyści dla gospodarstw wiejskich z wyboru strategii dywersyfikacji albo specjalizacji działalności. Trzeci przykład pochodzi z mojego własnego artykułu. Są w nim analizowane dane dotyczące amerykańskiego programu subsydiowanych miejsc pracy. Celem badania jest oszacowanie efektu brutto programu, czyli efektu dla beneficjenta bez uwzględniania kosztu programu. Ostatni przykład został zaadaptowany z artykułu Vanderberghe’a i Robina (2004). Celem artykułu jest porównanie wpływu umieszczania dzieci w szkołach prywatnych na ich osiągnięcia edukacyjne.

W każdym z przykładów podjęto próbę oszacowania przeciętnego efektu oddziaływania (ATE) oraz przeciętnego efektu oddziaływania wobec jednostek poddanych oddziaływaniu (ATT) metodą łączenia według oszacowanych wartości *propensity score* oraz metodą przeważania według odwrotności oszacowań *propensity score*. Wskazano na dwa ważne wyniki. Po pierwsze, wykorzystanie metod łączenia danych przyczynia się do zbilansowania rozkładów cech w grupie

eksperymentalnej i dobranej do niej grupie kontrolnej. Po drugie, nie istnieje jedna uniwersalna i zawsze skuteczna metoda wykorzystania wektorów prawdopodobieństwa oddziaływania w małych próbach.