

Marcin Szymkowiak

Podęjście kalibracyjne w badaniach społeczno-ekonomicznych

WYDAWNICTWO UEP



UNIWERSYTET
EKONOMICZNY
W POZNANIU

$D_4(d, w) = d^T G_4(g)$

Marcin Szymkowiak

Podejście kalibracyjne w badaniach społeczno-ekonomicznych

WYDAWNICTWO UEP



UNIwersytet
EKONOMICZNY
W POZNANIU

Poznań 2019

Komitet Redakcyjny

Szymon Cyfert, Elżbieta Gołata (przewodnicząca), Jacek Lisowski, Ewa Mahuszyńska, Eliza Szybowicz (sekretarz), Krzysztof Walczak, Ryszard Zieliński

Recenzenci

Mirosław Krzyśko, Mirosław Szreder

Projekt okładki

Ewa Wąsowska

Redakcja i korekta

Magdalena Kraszewska

© Copyright by Uniwersytet Ekonomiczny w Poznaniu
Poznań 2019



Ta książka jest udostępniana na licencji Creative Commons – Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 4.0 Międzynarodowe

ISBN 978-83-66199-23-1

eISBN 978-83-66199-89-7

<https://doi.org/10.18559/978-83-66199-89-7>

WYDAWNICTWO UNIWERSYTETU EKONOMICZNEGO W POZNANIU

ul. Powstańców Wielkopolskich 16, 61–895 Poznań

tel. 61 854 31 54, 61 854 31 55

www.wydawnictwo.ue.poznan.pl, e-mail: wydawnictwo@ue.poznan.pl

adres do korespondencji: al. Niepodległości 10, 61–875 Poznań

Skład: Łukasz Pańkowski

Druk: Zakład Graficzny Uniwersytetu Ekonomicznego w Poznaniu

ul. Towarowa 53, 61–896 Poznań, tel. 61 854 38 06, 61 854 38 03

Spis treści

Wstęp	7
1. Podejście kalibracyjne – zarys problematyki	15
1.1. Wprowadzenie	15
1.2. Podstawowe oznaczenia i definicje	16
1.3. Estymator Horvitz-Thompsona wartości globalnej	25
1.4. Uogólniony estymator regresyjny wartości globalnej	27
1.5. Rozwój teorii kalibracji	34
1.6. Wybrane zastosowania podejścia kalibracyjnego	45
1.7. Wnioski	48
2. Teoretyczne podstawy podejścia kalibracyjnego	49
2.1. Wprowadzenie	49
2.2. Estymatory kalibracyjne wartości globalnej – podejście minimalizujące funkcję odległości	50
2.3. Algorytm wyznaczania mnożników Lagrange’a w podejściu kalibracyjnym	68
2.4. Badanie symulacyjne – podejście minimalizujące funkcję odległości	78
2.5. Wnioski	87
3. Podejście kalibracyjne w ujęciu funkcyjnym	89
3.1. Wprowadzenie	89
3.2. Estymatory kalibracyjne wartości globalnej – podejście funkcyjne	90
3.3. Badanie symulacyjne – podejście funkcyjne	99
3.4. Wnioski	106
4. Podejście kalibracyjne wspomagane modelem	109
4.1. Wprowadzenie	109
4.2. Estymatory modelowo-kalibrowane wartości globalnej	111

4.3. Estymatory kalibracyjne wartości globalnej wspomagane modelem regresji LASSO i grzbietowej	120
4.4. Estymatory hybrydowo-kalibrowane wartości globalnej	122
4.5. Badanie symulacyjne – podejście kalibracyjne wspomagane modelem	126
4.6. Wnioski	131
5. Podejście kalibracyjne w badaniach z brakami odpowiedzi	134
5.1. Wprowadzenie	134
5.2. Braki odpowiedzi jako źródło błędów nielosowych	135
5.3. Kalibracja wag – podejście jednokrokowe	141
5.4. Kalibracja wag – podejście dwukrokowe	150
5.5. Podejście kalibracyjne w Badaniu Aktywności Ekonomicznej Ludności	154
5.6. Wnioski	169
6. Kalibracja zintegrowana w spisach mieszanych	171
6.1. Wprowadzenie	171
6.2. Idea spisów mieszanych na przykładzie NSP 2011	175
6.3. Kalibracja wag w NSP 2011	180
6.4. Teoretyczne podstawy kalibracji zintegrowanej	188
6.5. Propozycja kalibracji zintegrowanej w NSP 2011	202
6.6. Wnioski	214
Zakończenie	217
Bibliografia	221
Spis tabel	236
Spis rysunków	237
Spis algorytmów	239
Dodatek: Skrót i symbole	240
Summary: The calibration approach in socio-economic surveys	243

*Książkę dedykuję Żonie Kasi,
Synowi Jędrzejowi oraz Rodzicom*

Wstęp

W ostatnich latach obserwowany jest wzrost zapotrzebowania na informacje statystyczne dotyczące różnych zjawisk o charakterze społeczno-ekonomicznym. Popyt na aktualne i wiarygodne dane zgłaszany jest przez wiele instytucji na szczeblu państwowym, regionalnym i lokalnym. Są one niezbędne z punktu widzenia podejmowania odpowiednich decyzji przez władze rządowe i samorządowe, a także w kontekście monitorowania i oceny procesów społeczno-ekonomicznych w celu wzbogacenia wiedzy o otaczającej nas rzeczywistości.

Konieczność zaspokojenia potrzeb szerokiego grona odbiorców danych statystycznych wymusiła na organach statystyki publicznej dostarczanie informacji o wysokiej jakości, na którą składają się bardziej precyzyjne dane, o większym znaczeniu dla oceny analizowanych zjawisk oraz udostępniane na bieżąco (Witkowski, 2010). Statystyka publiczna, aby sprostać tym oczekiwaniom i wyzwaniom, podlega nieustannym przeobrażeniom, które obejmują wykorzystanie nowoczesnych technologii pozyskiwania i przetwarzania informacji oraz zaawansowanych metod badawczych, a także stosowanie danych pochodzących z różnych źródeł statystycznych. Dostrzegalne zmiany w zakresie organizacji badań statystycznych oznaczają konieczność wykorzystywania danych nie tylko ze spisów i badań reprezentacyjnych, ale w coraz większym zakresie z różnego rodzaju rejestrów i systemów administracyjnych. Większą rolę będą odgrywać również inne źródła w postaci internetu oraz tzw. big data (Beręsewicz i Szymkowiak, 2015; Beręsewicz, 2016; Daas i Puts, 2014; Daas, Puts, Buelens i van den Hurk, 2015; Szreder, 2015a; 2017; Zeelenberg, 2016).

Pomimo długiej tradycji wykorzystywania danych pochodzących ze spisów powszechnych, częstszego użycia rejestrów administracyjnych oraz innych źró-

deł danych, jak wspomniane dane z internetu czy big data, w dalszym ciągu jest oczywiste, że doniosłą rolę w statystyce publicznej odgrywają badania reprezentacyjne. Wynika to z kilku zasadniczych powodów. Po pierwsze dane pozyskiwane ze spisów powszechnych czy rejestrów administracyjnych pokrywają tylko pewien wycinek rzeczywistości. Istniejącą lukę w wielu innych obszarach, związanych z ubóstwem, rynkiem pracy, niepełnosprawnością, jakością życia, zdrowiem czy edukacją w dalszym ciągu wypełniają informacje pochodzące z odpowiednich badań reprezentacyjnych. Po drugie, jak zauważa Szreder (2016, s. 1–2), ważnymi przesłankami stosowania badań reprezentacyjnych są dostrzegalny postęp w rozwoju teorii próbkowania i wnioskowania statystycznego oraz powszechna świadomość niedostatków badań wyczerpujących. W nurt związany z ewolucją teorii estymacji wpisuje się również podejście kalibracyjne w rozumieniu zaprezentowanym w pracy Deville'a i Särndala (1992), któremu poświęcona jest niniejsza książka, a które w badaniach reprezentacyjnych odgrywa szczególną rolę. Kluczowym aspektem jest wreszcie relatywnie niski koszt ich realizacji, przetworzenia oraz udostępnienia danych.

Istotnym czynnikiem przemawiającym za wykorzystaniem badań reprezentacyjnych w procesie pozyskania odpowiedniej jakości danych jest także ich bogata historia i ugruntowana pozycja, wsparta przez odpowiednio wypracowaną metodologię. Jak podają Wright i Farmer (2000, s. iv), początków badań reprezentacyjnych można się doszukiwać już w XIX wieku, kiedy to wybitny matematyk francuski Pierre Simon de Laplace przekonał rząd francuski, aby na podstawie wylosowanej próby złożonej z okręgów administracyjnych zwanych gminami dokonać oszacowania liczebności populacji Francji. Wśród innych wczesnych zastosowań metody reprezentacyjnej, jak podają Rao i Fuller (2017, s. 146), należy wymienić pracę Kiaera (1897), który może być postrzegany jako pierwszy statystyk postulujący wykorzystanie badań częściowych opartych na próbie losowej kosztem badań wyczerpujących. Warto również wspomnieć o pracach Bowleya, który – jak podają Wright i Farmer (2000, s. iv) – wykorzystał losowanie systematyczne do oceny stopnia ubóstwa w mieście Reading w Anglii w 1912 roku.

W rozwoju metody reprezentacyjnej należy także uwzględnić pracę wybitnego polskiego matematyka i statystyka Jerzego Sławy-Neymana (1934), twórcy nowoczesnego pojęcia przedziału ufności, który wprowadził go do teorii testowania hipotez statystycznych i wytyczył kierunki rozwoju w obszarze badań reprezentacyjnych na dziesiątki lat. Wśród pozostałych, kluczowych opracowań, które odegrały fundamentalną rolę w rozwoju metody reprezentacyjnej i praktyce badań próbkowych, należy wymienić prace Cochran (1977), Deminga (1950), Hansena,

Hurvitza i Madowa (1953), Kisha (1965), a wśród polskich prace Brachy (1996), Pawłowskiego (1972), Steczkowskiego (1995), Wywiśla (1999; 2010) czy Zasepy (1962; 1972).

W rozwoju badań reprezentacyjnych szczególnie ważną rolę odegrała także kalibracja, której podwaliny zostały stworzone przez Deville'a i Särndala (1992). Jest ona techniką powszechnie wykorzystywaną w praktyce badań reprezentacyjnych, której idea polega na takim skorygowaniu wag wynikających z przyjętego planu losowania próby, aby odtworzyć znane wartości globalne wszystkich zmiennych pomocniczych. Mimo że teoria podejścia kalibracyjnego została sformalizowana we wspomnianej już pracy Deville'a i Särndala (1992), jej początków należy upatrywać w metodzie rakingu¹ zaproponowanej przez Deminga i Stephana (1940).

W tym miejscu należy podkreślić, że w książce zarówno termin kalibracja, jak i podejście kalibracyjne będą używane zamiennie, mimo że to drugie w rozumieniu autora ma nieco szerszy kontekst. Podejście kalibracyjne (ang. *calibration approach*) będziemy definiować jako cały proces estymacji nieznanymi parametrami w populacji generalnej bądź w odpowiednio zdefiniowanych subpopulacjach (domenach), który obejmuje dobór zmiennych pomocniczych, kalibrację wag, statystyczną ocenę uzyskanych wag, konstrukcję odpowiednich estymatorów kalibracyjnych i ich wariancji oraz analizę otrzymanych wyników. W wąskim rozumieniu kalibracja jest zatem jednym z etapów podejścia kalibracyjnego i oznacza korektę wag, tak aby odtworzyć znane wartości globalne zmiennych pomocniczych. W szerszym ujęciu obejmuje również pozostałe etapy opisanego procesu estymacji z wykorzystaniem estymatorów kalibracyjnych.

Kalibracja nie jest jedyną techniką statystyczną stosowaną w konstruowaniu wag, które mogą być wzięte pod uwagę w procesie uogólniania wyników. Można bowiem tutaj zastosować inne metody wyznaczania wag, w tym m.in. regresję logistyczną, gdzie wagi są odwrotnością prawdopodobieństwa określanego jako skłonność respondenta do udzielenia odpowiedzi na pytania zawarte w kwestionariuszu ankiety (Kalton i Flores-Cervantes, 2003). W wypadku kalibracji uzyskujemy jednak wagi, które odtwarzają znane wartości globalne zmiennych pomocniczych, co jest niezwykle pożądaną własnością w praktycznych zastosowaniach. Co więcej, inne znane metody konstruowania wag, takie jak wspomniany już raking czy estymacja typu GREG², są szczególnymi przypadkami kalibracji.

¹ Metoda korygowania liczebności w odpowiedniej tablicy kontyngencji z wykorzystaniem algorytmu iteracyjnego proporcjonalnego dopasowania, której celem jest odtworzenie znanych wartości brzegowych. Więcej informacji na temat tej techniki statystycznej można znaleźć w podrozdziale 1.5.

² Będzie ona przedmiotem rozważań w podrozdziale 1.4.

Współcześnie kalibracja jest techniką stosowaną w badaniach reprezentacyjnych zarówno przez krajowe urzędy statystyczne, jak i ośrodki badania rynku. Jej zastosowania nie ograniczają się przy tym do badań próbkowych, ale dotyczą także spisów realizowanych metodą mieszaną (Szymkowiak, 2014), badań pełnych bazujących na rejestrach i systemach administracyjnych (A. Wallgren i B. Wallgren, 2014) czy badań pełnych obejmujących całą populację, w których poważnym problemem są odmowy (Klimanek i Szymkowiak, 2017).

Warto nadmienić, że kalibracja jako metoda korygowania wag nie tylko jest wykorzystywana na potrzeby odtworzenia znanych wartości globalnych zmiennych pomocniczych, co ma się przyczynić do zachowania zgodności z wynikami innych badań statystycznych, ale odgrywa szczególną rolę w takich obszarach badawczych jak problematyka braków odpowiedzi i błędów pokrycia (Särndal i Lundström, 2005) czy zagadnienie wartości odstających związane z występowaniem skrajnie asymetrycznych rozkładów cech (Duchesne, 1999). Najnowszy nurt w podejściu kalibracyjnym stanowią jej zastosowania w badaniach, w których próby mają nielosowy charakter (Chen i in., 2017; Chen, Valliant i Elliott, 2018a; 2018b). Wynika z tego, że chociaż metody oferowane przez teorię kalibracji są już szeroko omawiane w literaturze przedmiotu, to nadal zauważalna jest potrzeba ich dalszego rozwoju. Dotyczyć to może wspomnianych już prób nielosowych, a także konieczności jej zastosowania w spisach realizowanych metodą mieszaną czy w badaniach, w których poważnym problemem są występujące braki danych. Odpowiednia korekta wag ma nie tylko się przyczynić do zachowania spójności ze strukturami znanymi z innych badań, takich jak spisy powszechne czy rejestry, ale także ma docelowo wpłynąć na redukcję obciążenia i poprawę jakości precyzji oszacowań na skutek istniejących błędów nielosowych.

Celem niniejszej pracy jest przedstawienie podejścia kalibracyjnego w bardzo szerokim ujęciu, z uwzględnieniem aktualnego dorobku naukowego, doświadczeń będących pokłosiem zastosowań tej metody w praktyce badań statystycznych oraz autorskich rozwiązań i propozycji w kontekście rozwoju statystyki publicznej w Polsce. Ukazano w niej różnorodność zagadnień obejmujących wykorzystanie kalibracji w badaniach z brakami odpowiedzi czy w spisach realizowanych metodą mieszaną. Analizie poddano również tzw. podejście funkcyjne w zagadnieniu wyznaczania wag kalibracyjnych oraz wspomagane modelem, z uwzględnieniem najnowszych prac i rozwiązań metodologicznych w tym zakresie. Tak sformułowany cel ogólny opracowania składa się z następujących celów szczegółowych:

- dyskusji podstawowych pojęć z zakresu metody reprezentacyjnej i podejścia kalibracyjnego,

- wskazania ewolucji teorii kalibracji wraz z obszarem jej zastosowań w praktyce badań statystycznych,
- konstruowania estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości i ograniczeń nakładanych na wagi wraz z dyskusją najważniejszych ich własności,
- budowania estymatorów kalibracyjnych wartości globalnej w kontekście podejścia funkcyjnego wraz z omówieniem kluczowych ich własności,
- budowania estymatorów modelowo-kalibrowanych wartości globalnej, w tym z uwzględnieniem modeli regresji grzbietowej i adaptacyjnego modelu regresji LASSO oraz hybrydowych, wraz z dyskusją ich najważniejszych własności,
- próby odpowiedzi na pytanie o możliwość wykorzystania podejścia kalibracyjnego na niższym poziomie agregacji przestrzennej w badaniach z brakami danych na przykładzie Badania Aktywności Ekonomicznej Ludności (BAEL),
- próby odpowiedzi na pytanie o możliwość zastosowania kalibracji zintegrowanej w spisach realizowanych metodą mieszaną na przykładzie Narodowego Spisu Powszechnego Ludności i Mieszkań z 2011 roku (NSP 2011).

Realizacji celów podporządkowana została struktura książki. Rozdział pierwszy ma charakter wprowadzający. Sformułowano w nim niezbędne pojęcia z zakresu metody reprezentacyjnej, kluczowe z punktu widzenia rozważanej teorii kalibracji. Dokonano także studiów literaturowych w świetle aktualnego stanu wiedzy, wskazano na rozwój podejścia kalibracyjnego oraz przedstawiono obszary najważniejszych zastosowań kalibracji. Uwzględniono przy tym własny dorobek autora w zakresie wykorzystania omawianej metody w różnych badaniach statystycznych.

Rozdział drugi poświęcono przedstawieniu podejścia kalibracyjnego w bardziej zmatematyzowanym ujęciu. Dyskusji poddano sposób konstruowania wag kalibracyjnych oraz estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości. Zostały przy tym omówione szczegółowo ich najważniejsze zalety oraz wady.

W rozdziale trzecim przedstawiono z kolei za literaturą przedmiotu inny sposób konstruowania wag kalibracyjnych, które uzyskuje się w tzw. podejściu funkcyjnym wykorzystującym koncepcję zmiennych instrumentalnych.

Rozdział czwarty poświęcono estymatorom kalibracyjnym wartości globalnej wspomaganym odpowiednim modelem. Oprócz przeglądu najważniejszych estymatorów prezentowanych w literaturze, uwzględniających również podejście hybrydowe oraz wspomagane modelem regresji grzbietowej i adaptacyjnym mo-

delem regresji LASSO, w badaniu symulacyjnym zaproponowano wykorzystanie nowych estymatorów hybrydowo-kalibrowanych, które w swej konstrukcji wykorzystują wspomniane modele regresji grzbietowej i LASSO.

W rozdziale piątym uwaga została skupiona na możliwości wykorzystania podejścia kalibracyjnego w badaniach statystycznych z brakami danych. Od strony teoretycznej przedstawiono uogólnienie metody konstruowania estymatorów kalibracyjnych wartości globalnej w podejściu jednokrokowym oraz dwukrokowym typu A i B rozważanej przez Särndala i Lundströma (2005) na przypadek dowolnie zdefiniowanej funkcji odległości. Zaproponowano również wykorzystanie podejścia kalibracyjnego w obszarze rynku pracy w przekroju podregionów, tj. na poziomie, dla którego brakuje oficjalnych szacunków publikowanych przez Główny Urząd Statystyczny. O jego znaczeniu niech świadczy to, że wśród podregionów mamy największe miasta w Polsce, dla których zawsze lokalni politycy i działacze gospodarczy dopominają się o informacje z zakresu rynku pracy. Zastosowanie i specyfikację związanych z tym problemów badawczych omawianej metody przedstawiono na danych z Badania Aktywności Ekonomicznej Ludności, w którym jednym z kluczowych problemów są występujące braki odpowiedzi.

Ostatni, szósty rozdział pracy zawiera rozważania dotyczące możliwości zastosowania kalibracji w spisach powszechnych realizowanych metodą mieszaną. Ten typ spisu był wykorzystywany w rundzie spisowej 2010–2011 w wielu państwach europejskich, w tym w Polsce (Gołata, 2018; Valente, 2010). W pracy przedstawiono sposób konstruowania wag kalibracyjnych w kontekście ludności faktycznie zamieszkałej zgodnie z wypracowaną w 2011 roku w spisie koncepcją korygowania wag. Wagi te zostały wyznaczone w Ośrodku Statystyki Małych Obszarów Urzędu Statystycznego w Poznaniu, a autor książki kierował pracami grupy metodologów odpowiedzialnych za proces kalibracji. W książce przedstawiono również propozycję konstruowania tzw. wag zintegrowanych w spisach mieszanych na przykładzie spisu powszechnego z 2011 roku. Należy podkreślić, że tego typu rozwiązanie nie zostało zastosowane w tym badaniu. Wyniki przeprowadzonej analizy wskazują, że podejście zintegrowane może nabrać szczególnego znaczenia w kontekście kolejnego spisu, który zostanie przeprowadzony w 2021 roku zgodnie z mieszanym modelem realizacyjnym. Na zakończenie dyskusji poświęconej zawartości każdego z rozdziałów warto zaznaczyć, że kończą się one podsumowaniami, w których ujęto najważniejsze wnioski.

Uwzględniając sformułowane cele oraz strukturę opracowania, należy podkreślić, że prezentowana praca ma charakter teoretyczny, empiryczny i aplikacyjny. W warstwie teoretycznej dokonano matematycznego opisu estymatorów kalibracyjnych wartości globalnej w ujęciu klasycznym, w podejściu funkcyjnym, wspo-

magany modelem, w badaniach z brakami odpowiedzi oraz w obszarze tzw. kalibracji zintegrowanej. Niektóre z rozważanych podejść miały przy tym autorski charakter, co należy postrzegać w charakterze novum metodologicznego.

W warstwie empirycznej głównym celem była analiza najważniejszych własności rozważanych estymatorów kalibracyjnych z uwzględnieniem oceny uzyskanych wag kalibracyjnych. Realizacji tego celu służyły odpowiednio zaprojektowane badania symulacyjne przedstawione w rozdziałach 2–4, w których wykorzystano rzeczywiste dane pochodzące z Badania Budżetów Gospodarstw Domowych (BBGD). Warto podkreślić, że w rozdziałach tych zastosowane zostały własne kody napisane w języku R (R Core Team, 2018), w których centralną rolę odgrywały pakiety **survey** (Lumley, 2004; 2017), **glmnet** (Friedman, Hastie i Tibshirani, 2010) oraz zmodyfikowana wersja funkcji *calibWeights* z pakietu **laeken** (Alfons i Templ, 2013). Modyfikacja polegała na dodaniu funkcji odległości w postaci sinusa hiperbolicznego i zastosowaniu pakietu do analizy macierzy rzadkich **Matrix** (Bates i Maechler, 2018) w celu przyspieszenia czasochłonnych obliczeń.

W warstwie aplikacyjnej dokonano zastosowania omawianych metod kalibracji w Badaniu Aktywności Ekonomicznej Ludności oraz w spisie powszechnym z 2011 roku, który został przeprowadzony metodą mieszaną. W tym celu ponownie skorzystano ze wspomnianych powyżej pakietów i funkcji języka R. Jedyne na potrzeby oceny precyzji oszacowań dla danych spisowych z wykorzystaniem otrzymanych wag zintegrowanych użyto programu SAS i procedury *survey-means*. Egzemplifikacja podejścia kalibracyjnego w bardzo ważnych badaniach statystycznych stanowi niewątpliwie novum poznawcze.

Książka powstała w ramach grantu naukowego: Estymacja pośrednia w zakresie badania niepełnosprawności na podstawie NSP 2011. Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS4/01472.

Podjęcie kalibracyjne – zarys problematyki

1.1. Wprowadzenie

W badaniach prowadzonych przez urzędy statystyczne i agencje rynkowe na całym świecie kalibracja jest powszechnie wykorzystywaną metodą korygowania wag wynikających z planu losowania próby. Jest ona stosowana w różnego rodzaju badaniach: spisach przeprowadzanych metodą mieszaną w celu uzyskania spójności pomiędzy najważniejszymi strukturami danych z różnych źródeł informacji zasilających spis (Szymkowiak, 2014), w badaniach reprezentacyjnych (Brick, 2013; Deville i Särndal, 1992; Haziza i Beaumont, 2017; Singh i Mohl, 1996), w badaniach pełnych bazujących na rejestrach administracyjnych (A. Wallgren i B. Wallgren, 2014) czy w badaniach ankietowych obejmujących całą populację (Klimanek i Szymkowiak, 2017).

W literaturze przedmiotu termin kalibracja funkcjonuje już od lat 90. XX wieku (Deville i Särndal, 1992). W ostatnich latach obserwowany jest jednak gwałtowny rozwój teorii w zakresie kalibracji oraz jej różnych zastosowań – w badaniach z brakami odpowiedzi czy błędami pokrycia (Cobben, 2009; Haziza i Lesage, 2016; Särndal i Lundström, 2005), w badaniach z wartościami odstającymi (Beaumont i Alavi, 2004; Duchesne, 1999; Ren i Chambers, 2002; Welsh i Ron-

chetti, 1998) czy w badaniach, w których wykorzystywane próby mają nielosowy charakter (Beręsewicz i Szymkowiak, 2018; Chen, 2016; Chen i in., 2017; 2018a; 2018b; Elliot, Resler, Flannagan i Rupp, 2010; Elliot i Valliant, 2017).

W niniejszym rozdziale sformułowano najważniejsze pojęcia z zakresu metody reprezentacyjnej, niezbędne z punktu widzenia rozważań w dalszych rozdziałach książki. Opisano dwa kluczowe estymatory wartości globalnej: Horvitz-Thompsona i uogólniony estymator regresyjny (ang. *generalized regression estimator* – GREG), które w badaniach próbkowych prowadzonych przez krajowe urzędy statystyczne odgrywają szczególną rolę. Omówiono ponadto rozwój podejścia kalibracyjnego oraz praktyczne zastosowania estymatorów kalibracyjnych, uwzględniających zarówno polskie, jak i zagraniczne doświadczenia.

Analizy prowadzone w tej części stanowią punkt wyjścia do rozważań przedstawionych w kolejnych rozdziałach niniejszej pracy, w których omówiono (wraz z egzemplifikacją) estymatory kalibracyjne wartości globalnej przy różnych założeniach: minimalizujące funkcję odległości, w podejściu funkcyjnym, wspomaganie modelem, uwzględniające braki danych czy kalibrację zintegrowaną w spisach realizowanych metodą mieszaną.

1.2. Podstawowe oznaczenia i definicje

Literatura z zakresu metody reprezentacyjnej jest niezwykle bogata. Dotyczy to zarówno pozycji zagranicznych, jak i polskich. Przedstawione definicje pochodzą przede wszystkim z pracy Cassela, Särndala i Wretmana (1977) oraz Särndala, Swenssona i Wretmana (1992). Można je również znaleźć w opracowaniach polskich autorów: Brachy (1996; 1998), Wywiata (1999; 2010) czy Żądły (2015), a wiele z nich także u Krzyński (2004).

Definicja 1.1. Liczbę obiektów populacji skończonej U oznaczamy przez N i nazywamy jej liczebnością.

Przykładami populacji są gospodarstwa domowe czy osoby w wieku 15 lat i więcej, będące członkami tych gospodarstw (Główny Urząd Statystyczny [GUS], 2018a), osoby stale zamieszkałe (zameldowane) na obszarze Polski bez względu na to, czy te osoby przebywały w kraju w czasie spisu, czy też były za granicą, oraz osoby przebywające czasowo (GUS, 2012), a także przedsiębiorstwa o liczbie pracujących 10 i więcej (Urząd Statystyczny w Poznaniu, 2017).

Interesującą nas cechę obiektów w populacji U oznaczać będziemy przez Y . Wartości cechy Y dla obiektów z tej populacji wynoszą odpowiednio y_1, \dots, y_N .

Zbiór wszystkich możliwych wartości tej cechy w populacji oznaczać z kolei będziemy przez \mathcal{Y} . Pojęcie cechy statystycznej jest tożsame z pojęciem zmiennej losowej w teorii prawdopodobieństwa. Zatem formalnie możemy przyjąć następującą definicję.

Definicja 1.2. *Cechą statystyczną nazywamy rzeczywistą funkcję Y określoną na zbiorze \mathcal{Y} , taką że $Y: \mathcal{Y} \rightarrow \mathbb{R}$.*

Przykładami cech w badaniach statystycznych mogą być przychody przedsiębiorstw o liczbie pracujących 10 i więcej, miesięczne wydatki gospodarstw domowych na określony produkt czy wykształcenie osoby. Zazwyczaj w badaniach prowadzonych przez krajowe urzędy statystyczne przedmiotem zainteresowania jest informacja na temat wielu różnych cech jednocześnie.

W wypadku populacji skończonych, dla których każdy element można poddać badaniu w takim rozumieniu, że od nas zależy, czy zmierzona będzie wartość cechy Y dla konkretnego obiektu, czy też nie, naturalne jest zdefiniowanie parametru jako pewnej funkcji wszystkich wartości cechy w populacji (Koronacki i Mielniczuk, 2018). Może to być na przykład średnia wartość cechy.

Definicja 1.3. *Niech $\mathbf{y} = (y_1, \dots, y_N)^T$, gdzie y_1, \dots, y_N są wartościami cechy Y w populacji. Parametrem populacji nazywamy pewną funkcję θ wektora \mathbf{y} , taką że $\theta = f(y_1, \dots, y_N) = f(\mathbf{y})$.*

Definicja 1.4. *Funkcją parametryczną nazywamy dowolną funkcję rzeczywistą $g(\theta)$ parametru θ .*

W badaniach statystycznych realizowanych przez krajowe urzędy statystyczne przedmiotem zainteresowania są funkcje parametryczne. Są to zazwyczaj:

- wartość globalna cechy Y :

$$\tau_Y = \sum_{k \in U} y_k, \quad (1.1)$$

- średnia cechy Y :

$$\mu_Y = \frac{\tau_Y}{N}, \quad (1.2)$$

- iloraz wartości globalnych cech X i Y :

$$R = \frac{\tau_Y}{\tau_X}, \quad (1.3)$$

– kwantyl rzędu α :

$$Q_{Y,\alpha} = \inf\{t \mid F_Y(t) \geq \alpha\}, \quad (1.4)$$

gdzie

$$F_Y(t) = \frac{\sum_{k \in U} I(t - y_k)}{N} \quad (1.5)$$

oraz

$$I(z) = \begin{cases} 1, & \text{jeżeli } z \geq 0, \\ 0, & \text{jeżeli } z < 0. \end{cases} \quad (1.6)$$

W szczególnym przypadku, gdy $\alpha = 0,5$, parametrem populacji jest tzw. mediana.

Do innych funkcji parametrycznych, które rzadziej są przedmiotem zainteresowania w rzeczywistych badaniach statystycznych, należy zaliczyć:

– wariancję cechy Y :

$$\sigma_Y^2 = \frac{1}{N-1} \sum_{k \in U} (y_k - \mu_Y)^2, \quad (1.7)$$

– minimalną wartość cechy Y :

$$Y_{\min} = \min\{y_1, \dots, y_N\}, \quad (1.8)$$

– maksymalną wartość cechy Y :

$$Y_{\max} = \max\{y_1, \dots, y_N\}, \quad (1.9)$$

– kowariancję cech X i Y :

$$\text{Cov}(X, Y) = \frac{1}{N-1} \sum_{k \in U} (x_k - \mu_X)(y_k - \mu_Y). \quad (1.10)$$

Przykładami parametrów cech w populacji, które są przedmiotem estymacji w badaniach prowadzonych przez krajowe urzędy statystyczne, mogą być: przeciętne miesięczne przychody na osobę w gospodarstwach domowych, liczba osób bezrobotnych, biernych i pracujących, liczba osób niepełnosprawnych biologicznie i prawnie w przekroju powiatów województwa wielkopolskiego czy mediana miesięcznego dochodu rozporządzalnego na osobę w gospodarstwach domowych w Polsce.

Ponieważ parametry w populacji nie są znane, podlegają estymacji na podstawie danych pochodzących z próby.

Definicja 1.5. *Próbką uporządkowaną nazywamy każdy ciąg $\underline{s} = \{k_1, \dots, k_{n'}\}$, gdzie $k_i \in U$, a n' jest liczebnością próby uporządkowanej. Zbiór wszystkich prób $\underline{\mathbf{S}}$ postaci \underline{s} nazywamy przestrzenią prób.*

Definicja 1.6. *Próbką nieuporządkowaną (próbą) nazywamy niepusty zbiór s , taki że $s \subseteq U$. Liczebność próby oznaczana jest przez n , a zbiór \mathbf{S} wszystkich prób postaci s nazywamy przestrzenią prób.*

Definicja 1.7. *Efektywną liczebnością próby nazywamy liczbę niepowtarzających się elementów w próbce.*

Definicja 1.8. *Rozkład prawdopodobieństwa $P(s)$ określony na przestrzeni prób \mathbf{S} nazywamy planem losowania. Dla każdej próby $s \in \mathbf{S}$ spełnia on warunki: $P(s) \geq 0$ oraz $\sum_{s \in \mathbf{S}} P(s) = 1$.*

Jednym z prostszych sposobów losowania jednostek do próby z populacji jest losowanie proste (ze zwracaniem bądź bez zwracania). W wypadku losowania prostego ze zwracaniem przestrzeń prób prostych (n -elementowych) z populacji N -elementowej U jest zbiorem o liczebności N^n . Plan losowania próby ze zwracaniem jest określony wzorem:

$$\forall_{s \in \mathbf{S}} P(s) = N^{-n}. \quad (1.11)$$

Natomiast w wypadku losowania prostego bez zwracania przestrzeń prób n -elementowych jest zbiorem o liczebności $\binom{N}{n}$. W takiej sytuacji plan losowania próby prostej bez zwracania określony jest następującym wzorem:

$$\forall_{s \in \mathbf{S}} P(s) = \binom{N}{n}^{-1}. \quad (1.12)$$

Powyżej omówione sposoby losowania próby są w praktyce rzadko wykorzystywane (Szreder, 2010a). Wynika to przede wszystkim z tego, że losowanie takie charakteryzuje się niższą precyzją oszacowań w porównaniu z innymi schematami, zwłaszcza w sytuacji gdy populacja jest zróżnicowana. W losowaniu tym nie wykorzystuje się ponadto żadnej informacji o badanej populacji. Jest ono również na ogół kosztowne z punktu widzenia przygotowania operatu losowania i pomiaru. Zdecydowanie częściej można się spotkać z losowaniem warstwowym, grupowym (zespołowym) czy dwustopniowym. Losowanie proste (zarówno ze zwracaniem, jak i bez zwracania) jest natomiast często wykorzystywane w symulacyjnych badaniach nad własnościami rozważanych estymatorów.

W wypadku wspomnianego już losowania warstwowego zakładamy, że $U = \bigcup_{c=1}^C U_c$, gdzie zbiory U_c są niepuste i określone są mianem warstw. Zakłada się ponadto, że $U_c \cap U_{c'} = \emptyset$ dla $c \neq c'$, gdzie $c = 1, \dots, C$. Liczebność każdej warstwy wynosi N_c , tj. $|U_c| = N_c$. W wypadku gdy z każdej z warstw losujemy niezależnie próbę prostą bez zwracania o liczebności n_c , plan losowania próby warstwowej bez zwracania wyraża się wzorem:

$$\forall_{s \in \mathbf{S}} P(s) = \prod_{c=1}^C \binom{N_c}{n_c}^{-1}. \quad (1.13)$$

W wypadku losowania grupowego zakładamy, że populacja U podzielona jest na G niepustych zbiorów U_h zwanych zespołami, tj. $U = \bigcup_{h=1}^G U_h$, $U_h \cap U_{h'} = \emptyset$ dla $h \neq h'$, gdzie $h = 1, \dots, G$, i $\forall_h |U_h| = M$. Spośród G wyróżnionych grup losuje się bez zwracania g grup, dla których obserwacji podlegają wszystkie elementy populacji. W takim przypadku plan losowania próby grupowej (zespołowej) określony jest wzorem:

$$\forall_{s \in \mathbf{S}} P(s) = \binom{G}{g}^{-1}. \quad (1.14)$$

W praktyce badań realizowanych przez krajowe urzędy statystyczne można się spotkać dość często z tzw. losowaniem dwustopniowym. Podobnie jak w wypadku losowania grupowego, zakładamy, że populacja U jest podzielona na G niepustych zbiorów U_h zwanych zespołami, tj. $U = \bigcup_{h=1}^G U_h$, $U_h \cap U_{h'} = \emptyset$ dla $h \neq h'$, gdzie $h = 1, \dots, G$, i $\forall_h |U_h| = N_h$. W losowaniu dwustopniowym na pierwszym stopniu losowania losuje się bez zwracania g grup spośród G grup. W drugim kroku (drugi stopień losowania) z N_h elementów z każdej wylosowanej na pierwszym stopniu grupy losuje się n_h jednostek. W takim wypadku plan losowania próby dwustopniowej wyraża się wzorem:

$$\forall_{s \in \mathbf{S}} P(s) = \binom{G}{g}^{-1} \prod_{h=1}^g \binom{N_h}{n_h}^{-1}. \quad (1.15)$$

Losowanie dwustopniowe jest wykorzystywane w wielu badaniach prowadzonych przez krajowe urzędy statystyczne. Na przykład, we wspomnianym już Badaniu Aktywności Ekonomicznej Ludności realizowanym przez Główny Urząd Statystyczny jednostkami losowania pierwszego stopnia są w miastach tzw. rejony statystyczne, a na wsi obwody spisowe. W niektórych wypadkach jednostki losowania tworzy się, łącząc ze sobą dwa lub więcej sąsiednich rejonów lub ob-

wodów. Z kolei jednostkami losowania drugiego stopnia są mieszkania. Jednostki pierwszego stopnia są przy tym losowane z zastosowaniem tzw. warstwowania, a podstawę podziału na warstwy stanowi podział Polski na województwa (GUS, 2018a).

Również w Badaniu Budżetów Gospodarstw Domowych Główny Urząd Statystyczny stosuje losowanie dwustopniowe, warstwowe, z różnymi prawdopodobieństwami wyboru na pierwszym stopniu. Jednostkami losowania pierwszego stopnia są rejony statystyczne lub zespoły rejonów, a na drugim stopniu losowane są mieszkania. Szczegółowy opis procedury losowania można znaleźć w opracowaniu GUS (2017a).

Losowanie dwustopniowe jest stosowane także przez Główny Urząd Statystyczny w Europejskim Badaniu Warunków Życia Ludności (EU-SILC) z różnymi prawdopodobieństwami wyboru na pierwszym stopniu. Jednostkami pierwszego stopnia są obwody spisowe. Na drugim stopniu losowane są z kolei mieszkania. Badaniu podlegają wszystkie gospodarstwa domowe zamieszkałe w wylosowanych mieszkaniach. Warto wspomnieć, że jednostki pierwszego stopnia są przed losowaniem warstwowane. Warstwami są województwa, a wewnątrz województw jednostki pierwszego stopnia są warstwowane według klasy miejscowości. Pełen opis losowania zastosowanego przez Główny Urząd Statystyczny w badaniu EU-SILC można znaleźć w publikacji GUS (2017b).

Definicja 1.9. *Mechanizm losowania, który realizuje plan wyboru jednostek populacji do próby, nazywa się schematem losowania.*

W badaniach statystycznych na potrzeby zastosowania określonych estymatorów oraz oceny ich wariancji wykorzystuje się tzw. prawdopodobieństwa inkluzji pierwszego i drugiego rzędu. Poniższa definicja określa prawdopodobieństwo inkluzji rzędu r , tj. prawdopodobieństwo wyboru do próby s jednostek populacji k_1, \dots, k_r , którego szczególnymi przypadkami są prawdopodobieństwa inkluzji pierwszego ($r = 1$) i drugiego ($r = 2$) rzędu.

Definicja 1.10. *Prawdopodobieństwem inkluzji rzędu r (ang. r -order inclusion probability) jest:*

$$\pi_{k_1, \dots, k_r} = \sum_{s \in A(k_1, \dots, k_r)} P(s), \quad (1.16)$$

gdzie $A(k_1, \dots, k_r) = \{s: k_i \in s, i = 1, \dots, r\}$.

Prawdopodobieństwa inkluzji pierwszego i drugiego rzędu dla wybranych planów losowania można wyznaczyć w dość łatwy sposób. Na przykład, dla loso-

wania prostego ze zwracaniem i bez zwracania prawdopodobieństwa te można wyrazić następującymi wzorami:

– losowanie proste ze zwracaniem:

$$\pi_k = 1 - \left(1 - \frac{1}{N}\right)^n, \quad (1.17)$$

$$\pi_{kl} = 1 - 2\left(1 - \frac{1}{N}\right)^n + \left(1 - \frac{2}{N}\right)^n, \quad (1.18)$$

– losowanie proste bez zwracania:

$$\pi_k = \frac{n}{N}, \quad (1.19)$$

$$\pi_{kl} = \frac{n(n-1)}{N(N-1)}. \quad (1.20)$$

Definicja 1.11. *Wagę wynikającą z planu losowania (ang. design weight) definiujemy jako:*

$$d_k = \frac{1}{\pi_k}. \quad (1.21)$$

Analogicznie wagę d_{kl} definiujemy jako:

$$d_{kl} = \frac{1}{\pi_{kl}}. \quad (1.22)$$

Wagi d_k oraz d_{kl} odgrywają istotną rolę w wyznaczaniu ocen punktowych estymatorów oraz ich wariancji dla różnych parametrów. Wagi d_k mają również kluczowe znaczenie w podejściu kalibracyjnym, gdyż podlegają odpowiedniej korekcji z wykorzystaniem zestawu zmiennych pomocniczych, tak aby były spełnione właściwe równania kalibracyjne. Szczegółowo proces ten opisano w rozdziale drugim niniejszej książki.

Jak wcześniej wspomniano, kluczowym elementem jest wykorzystanie próby, aby wyciągnąć wnioski dotyczące nieznanego parametru θ lub, bardziej ogólnie, pewnej funkcji g parametru θ (funkcji parametrycznej). W tym celu konstruuje się statystykę $\hat{\theta} = T(Y_1, \dots, Y_n)$, zwaną estymatorem punktowym, w taki sposób, aby wartość tej statystyki, zwana oceną punktową, była bliska wartości parametru θ (Krzyśko, 2004, s. 67).

Definicja 1.12. *Estymatorem parametru θ jest statystyka $\hat{\theta} = T(Y_1, \dots, Y_n)$. Oceną parametru θ nazywamy wartość estymatora $\hat{\theta}$.*

Naturalnym pytaniem, jakie pojawia się w tym miejscu, jest pytanie o to, jakiej statystyki powinniśmy użyć do oceny parametru θ . Punktem wyjścia do rozważań na temat jakości estymatorów jest liczbowa charakterystyka dokładności estymatora T , za którą przyjmuje się błąd średniokwadratowy. Definiuje się go w następujący sposób (Krzyśko, 2004, s. 68):

Definicja 1.13. *Błędem średniokwadratowym estymatora $\hat{\theta}$ (ang. mean squared error – MSE) nazywamy wyrażenie postaci :*

$$\text{MSE}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]. \quad (1.23)$$

Pierwiastek z błędu średniokwadratowego estymatora $\hat{\theta}$ (ang. root mean squared error – RMSE) oznaczają będziemy przez:

$$\text{RMSE}(\hat{\theta}) = \sqrt{\text{MSE}(\hat{\theta})}. \quad (1.24)$$

Przyjmując błąd średniokwadratowy $\text{MSE}(\hat{\theta})$ za miarę dokładności estymatora, można w klasie wszystkich estymatorów wprowadzić porządek częściowy zgodnie z poniższą definicją.

Definicja 1.14. *Estymator $\hat{\theta}_1$ jest lepszy od estymatora $\hat{\theta}_2$, jeżeli dla każdego parametru $\theta \in \Theta$:*

$$\text{MSE}(\hat{\theta}_1) \leq \text{MSE}(\hat{\theta}_2) \quad (1.25)$$

i chociażby dla jednej wartości θ spełniona jest nierówność ostra:

$$\text{MSE}(\hat{\theta}_1) < \text{MSE}(\hat{\theta}_2), \quad (1.26)$$

gdzie Θ jest przestrzenią parametrów.

Można pokazać, że najlepszy estymator w sensie powyższej definicji istnieje niezmiernie rzadko. Zwykle jest bowiem tak, że dla pewnych wartości parametru θ z dwóch estymatorów $\hat{\theta}_1$ i $\hat{\theta}_2$ lepszy jest np. estymator $\hat{\theta}_1$, a dla innych wartości parametru – estymator $\hat{\theta}_2$ (Krzyśko, 2004, s. 69).

Definicja 1.15. *Estymator $\hat{\theta} = T(Y_1, \dots, Y_n)$ parametru θ jest nieobciążony, jeżeli dla każdego $\theta \in \Theta$ spełniony jest poniższy warunek:*

$$E(\hat{\theta}) = \theta. \quad (1.27)$$

Definicja 1.16. *Obciążeniem estymatora $\hat{\theta}$ nazywamy różnicę postaci:*

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta. \quad (1.28)$$

Definicja 1.17. *Wariancją estymatora $\hat{\theta}$ nazywamy wyrażenie postaci:*

$$D^2(\hat{\theta}) = E\left(\hat{\theta} - E(\hat{\theta})\right)^2. \quad (1.29)$$

Można pokazać, że między błędem średniokwadratowym, wariancją i obciążeniem estymatora $\hat{\theta}$ parametru θ istnieje następująca zależność:

$$\text{MSE}(\hat{\theta}) = D^2(\hat{\theta}) + B^2(\hat{\theta}), \quad (1.30)$$

która w szczególnym przypadku dla estymatora nieobciążonego prowadzi do równości $\text{MSE}(\hat{\theta}) = D^2(\hat{\theta})$.

Definicja 1.18. *Średnim błędem szacunku estymatora $\hat{\theta}$ nazywamy wyrażenie postaci:*

$$D(\hat{\theta}) = \sqrt{D^2(\hat{\theta})}. \quad (1.31)$$

Definicja 1.19. *Względny średni błąd szacunku estymatora $\hat{\theta}$ (ang. relative estimation error – REE) nazywamy wyrażenie postaci:*

$$\text{REE}(\hat{\theta}) = \frac{D(\hat{\theta})}{|\hat{\theta}|} \cdot 100\%. \quad (1.32)$$

Miara ta w literaturze anglojęzycznej oznaczana jest także jako *CV* (ang. *coefficient of variation*) i nazywana współczynnikiem zmienności. Również w publikacjach Głównego Urzędu Statystycznego, przy podawaniu wskaźników precyzji estymacji, oznaczenie to jest często używane. Zgodnie z wytycznymi, jakie Główny Urząd Statystyczny przyjmuje podczas publikowania wyników z badania reprezentacyjnego, oszacowania, dla których $CV < 10\%$, można uznać za wiarygodne. Oszacowania, dla których CV przyjmuje wartości z przedziału 10–20%, należy interpretować ze szczególną ostrożnością. Z kolei do oszacowań, dla których $CV > 20\%$, należy podchodzić sceptycznie i powinny być one publikowane tylko w postaci zagregowanej, tj. na wyższym poziomie agregacji (GUS, 2013). Wskaźnik CV można wykorzystać także do wyznaczenia odpowiedniego przedziału ufności, który z określonym prawdopodobieństwem $1 - \alpha$ pokrywa prawdziwą wartość estymowanego parametru.

W dalszej części książki dla względnego błędu szacunku estymatora będziemy jednak używać oznaczenia REE.

Definicja 1.20. *Względny błąd średniokwadratowy estymatora $\hat{\theta}$ (ang. relative root mean squared error – RRMSE) nazywamy wyrażenie postaci:*

$$\text{RRMSE}(\hat{\theta}) = \frac{\text{RMSE}(\hat{\theta})}{\hat{\theta}}. \quad (1.33)$$

Miara ta opisuje, jaki jest udział błędu estymacji w wartości szacowanego parametru.

1.3. Estymator Horvitz-Thompsona wartości globalnej

W badaniach próbkowych prowadzonych przez krajowe urzędy statystyczne na całym świecie procesowi estymacji podlega wiele różnych parametrów. Jak wcześniej wspomniano, najczęściej są to wartość globalna, średnia, kwantyle czy bardziej złożone parametry takie jak iloraz dwóch wartości globalnych. Parametry te mogą się odnosić na przykład do liczby osób bezrobotnych, pracujących czy biernych zawodowo (wartość globalna), mediany bądź średniego przychodu przedsiębiorstw.

Estymacja parametrów odbywa się na podstawie próby odpowiednio wylosowanej ze skończonej populacji, pobranej zgodnie z określonym planem losowania. Istnieje możliwość uzyskania informacji na temat wybranych parametrów również ze spisów czy rejestrów administracyjnych, które w statystyce publicznej odgrywają szczególną rolę. Ze względu na incydentalny charakter spisów oraz ich ograniczony zakres informacyjny, badania próbkowe są ważnym źródłem danych na temat sytuacji społeczno-gospodarczej w kraju. Jednak spisy oraz rejestry administracyjne mogą stanowić bogate zasoby zmiennych pomocniczych, które odgrywają istotną rolę w procesie estymacji, również w kontekście rozważanego w pracy podejścia kalibracyjnego. Równie ważne są inne źródła, takie jak internet czy tzw. big data (na przykład dane pochodzące z Facebooka, Twittera, od operatorów telefonii komórkowej itd.), które w coraz większym zakresie wykorzystywane są przez krajowe urzędy statystyczne (Beręsewicz i Szymkowiak, 2015; Daas i Puts, 2014; Daas i in., 2015; Szreder, 2015a; Zeelenberg, 2016). Mimo że w statystyce publicznej są one w coraz większym stopniu wykorzystywane jako źródło cennych informacji, również w kontekście podejścia kalibracyjnego (Beręsewicz i Szymkowiak, 2018), nie są przedmiotem głębszych rozważań w niniejszej pracy.

W dalszej części książki analizie poddano problem estymacji wartości globalnej, która odgrywa kluczową rolę wśród wszystkich potencjalnych parametrów.

Opisano estymator Horwitza-Thompsona oraz uogólniony estymator regresyjny, który (jak pokazano w rozdziale drugim) jest szczególnym przypadkiem estymatora kalibracyjnego. Ujmując zagadnienie bardziej formalnie, założmy, że przedmiotem estymacji jest wartość globalna w populacji U zmiennej Y określona wzorem (por. wzór (1.1)):

$$\tau_Y = \sum_{k=1}^N y_k, \quad (1.34)$$

gdzie N oznacza liczebność populacji, a y_k to wartość badanej cechy Y dla k -tej jednostki populacji, przy czym $k = 1, \dots, N$. Zakładamy, że z populacji $U = \{1, \dots, N\}$ losujemy zgodnie z określonym planem losowania próbę s o liczebności n . Niech $\pi_k > 0$ oznacza prawdopodobieństwo inkluzji pierwszego rzędu (por. wzór (1.16)), a d_k wagę wynikającą z planu losowania próby zdefiniowaną jako $d_k = 1/\pi_k$ (por. wzór (1.21)). Przy przyjętych powyżej oznaczeniach wartość globalną (1.34) zmiennej Y można wyrazić alternatywnie jako $\tau_Y = \sum_{k \in U} y_k$ lub w skrócie $\tau_Y = \sum_U y_k$. Wartość globalna we wzorze (1.34) może się odnosić zarówno do zmiennej Y ciągłej (na przykład przychód przedsiębiorstwa), jak i do zmiennych dyskretnych. Na przykład, jeśli przyjmiemy, że Y jest zmienną dychotomiczną, dla której:

$$y_k = \begin{cases} 1, & \text{jeżeli osoba jest pracująca,} \\ 0, & \text{jeżeli osoba jest bezrobotna,} \end{cases}$$

gdzie $k = 1, \dots, N$, to wartość globalna (1.34) odnosić się będzie do liczby osób pracujących.

W badaniach statystycznych prowadzonych przez krajowe urzędy statystyczne do szacowania wartości globalnej wyrażonej wzorem (1.34) wykorzystuje się często tzw. estymator Horwitza-Thompsona (1952) określony wzorem:

$$\hat{\tau}_{\text{HT}} = \sum_{k=1}^n d_k y_k = \sum_{k \in s} d_k y_k. \quad (1.35)$$

Niewątpliwą zaletą tego estymatora jest jego prostota, a także to, że jest on nieobciążony, tj. $E(\hat{\tau}_{\text{HT}}) = \tau_Y$. Wynika to z następujących przekształceń:

$$\begin{aligned} E(\hat{\tau}_{\text{HT}}) &= E\left(\sum_{k \in U} \pi_k^{-1} y_k I(k \in s)\right) = \sum_{k \in U} \pi_k^{-1} y_k E(I(k \in s)) = \\ &= \sum_{k \in U} \pi_k^{-1} y_k \pi_k = \tau_Y. \end{aligned}$$

Wariancja tego estymatora wyraża się wzorem (Särndal i in., 1992):

$$D^2(\hat{\gamma}_{HT}) = \sum_{k \in U} \sum_{l \in U} \Delta_{kl}(d_k y_k)(d_l y_l), \quad (1.36)$$

gdzie $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$, a π_{kl} oznacza prawdopodobieństwo inkluzji drugiego rzędu, tj. prawdopodobieństwo wyboru do próby s elementów populacji k i l (por. wzór (1.16)).

Estymator wariancji (1.36) ma postać (Särndal i in., 1992):

$$\hat{D}^2(\hat{\gamma}_{HT}) = \sum_{k \in s} \sum_{l \in s} \frac{\Delta_{kl}}{\pi_{kl}}(d_k y_k)(d_l y_l). \quad (1.37)$$

Estymator Horvitz-Thompsona w literaturze przedmiotu określany jest również mianem bezpośredniego lub ekspansyjnego. Nie wykorzystuje on na potrzeby estymacji wartości globalnej, poza wartościami zmiennej Y , informacji na temat zmiennych pomocniczych. Charakteryzuje się jednak dużą wariancją, zwłaszcza w wypadku prób o małej liczebności bądź w sytuacji, gdy w odniesieniu do zmiennej Y występuje duża frakcja braków danych. Nie można go również wykorzystać dla tzw. prób zerowych, tj. w sytuacji, gdy w jakimś przekroju brakuje reprezentanta w próbie. Z tego względu w badaniach prowadzonych przez Główny Urząd Statystyczny może być on wykorzystywany do estymacji wybranych charakterystyk jedynie na poziomie całego kraju, regionu, a rzadziej województw.

1.4. Uogólniony estymator regresyjny wartości globalnej

Uogólnione estymatory regresyjne GREG stanowią klasę estymatorów, w których konstrukcji wykorzystywana jest informacja o zmiennych pomocniczych. Za protoplastę uogólnionych estymatorów kalibracyjnych można uznać już estymatory regresyjne rozważane przez Cochran (1977), aczkolwiek kompleksowy opis tej klasy estymatorów można znaleźć przede wszystkim w pracy Särndala i współautorów (1992). Są to estymatory wspomagane modelem (ang. *model-assisted approach*), których używa się na potrzeby oszacowania wartości teoretycznych $\hat{\gamma}_k$. W podejściu tym zakłada się istnienie pewnego modelu wiążącego zmienną objaśnianą z wektorem zmiennych objaśniających \mathbf{x}_k , tj.:

$$\mu_k = \mu(\mathbf{x}_k, \boldsymbol{\beta}), \quad (1.38)$$

gdzie $\mathbf{x}_k = (x_{k1}, \dots, x_{kJ})^T$ to wektor złożony z wartości zmiennych objaśniających dla k -tej jednostki, $k \in U$, J oznacza liczbę zmiennych objaśniających, a $\hat{\boldsymbol{\beta}}$ to wektor parametrów podlegających estymacji. Przyjmując, że $\hat{\boldsymbol{\beta}}$ jest oszacowaniem wektora parametrów $\boldsymbol{\beta}$, wartości modelowe \hat{y}_k można przedstawić jako:

$$\hat{y}_k = \mu(\mathbf{x}_k, \hat{\boldsymbol{\beta}}). \quad (1.39)$$

W estymatorach typu GREG wykorzystywany model μ , wspomagający proces estymacji i łączący zmienną objaśnianą z wektorem zmiennych objaśniających \mathbf{x}_k , może być różnej postaci. W szczególności można rozpatrywać tutaj model liniowy, modele nieliniowe, uogólnione modele liniowe oraz uogólnione modele liniowe z efektami mieszanymi (stałymi oraz losowymi). Biorąc pod uwagę różne postacie modelu μ , estymator typu GREG stanowi całą klasę potencjalnych estymatorów, które można wykorzystać na potrzeby estymacji wartości globalnej τ_Y .

Uogólniony estymator regresyjny $\hat{\tau}_{\text{GREG}}$ wartości globalnej (1.34) może być przedstawiony w następujący sposób (Särndal i in., 1992):

$$\begin{aligned} \hat{\tau}_{\text{GREG}} &= \sum_{k \in U} \hat{y}_k + \sum_{k \in s} d_k (y_k - \hat{y}_k) = \\ &= \sum_{k \in s} d_k y_k + \left(\sum_{k \in U} \hat{y}_k - \sum_{k \in s} d_k \hat{y}_k \right) = \hat{\tau}_{\text{HT}} + \left(\sum_{k \in U} \hat{y}_k - \sum_{k \in s} d_k \hat{y}_k \right). \end{aligned} \quad (1.40)$$

W praktyce jednym z najczęściej wykorzystywanych modeli w konstruowaniu uogólnionego estymatora regresyjnego GREG jest model liniowy. W tym wypadku wartości \hat{y}_k wyznaczone są ze wzoru:

$$\hat{y}_k = \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_s, \quad (1.41)$$

gdzie:

$$\hat{\boldsymbol{\beta}}_s = \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in s} d_k q_k \mathbf{x}_k y_k, \quad (1.42)$$

a q_k jest odpowiednio dobranym czynnikiem skalującym. W zastosowaniach przyjmuje się często $q_k = 1$. Nie jest to jedyny możliwy sposób ustalenia tej wagi. Dla jednej zmiennej pomocniczej x_k można przyjąć również $q_k = 1/x_k$.

W wypadku uwzględnienia modelu liniowego estymator typu GREG określony wzorem (1.40) ma następującą postać (Särndal, 2007):

$$\hat{\tau}_{\text{GREG}}^{\text{lin}} = \left(\sum_{k \in U} \mathbf{x}_k \right)^T \hat{\boldsymbol{\beta}}_s + \sum_{k \in s} d_k (y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_s). \quad (1.43)$$

Jest to tzw. liniowy uogólniony estymator regresyjny wartości globalnej. Estymator ten można zapisać również w następującej postaci:

$$\hat{\tau}_{\text{GREG}}^{\text{lin}} = \sum_{k \in s} w_k y_k, \quad (1.44)$$

gdzie:

$$w_k = d_k g_k, \quad (1.45)$$

$$g_k = 1 + q_k \boldsymbol{\lambda}^T \mathbf{x}_k, \quad (1.46)$$

przy czym

$$\boldsymbol{\lambda}^T = \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1}. \quad (1.47)$$

Uogólniony liniowy estymator regresyjny GREG ma następującą własność:

$$\sum_{k \in s} w_k \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k. \quad (1.48)$$

Powyższa równość wynika z następującego przekształcenia, którego punktem wyjścia jest zapis estymatora typu GREG w postaci (1.43) równoważnej w stosunku do (1.44), a następnie przyjęcie w miejsce y_k wektora \mathbf{x}_k^T :

$$\begin{aligned} \sum_{k \in s} w_k \mathbf{x}_k^T &= \left(\sum_{k \in U} \mathbf{x}_k \right)^T \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right) + \\ &+ \sum_{k \in s} d_k \left[\mathbf{x}_k^T - \mathbf{x}_k^T \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right) \right] = \\ &= \left(\sum_{k \in U} \mathbf{x}_k \right)^T. \end{aligned} \quad (1.49)$$

Stąd $\sum_{k \in s} w_k \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k$. Równość ta ma bardzo duże znaczenie praktyczne, gdyż orzeka, że wagi w_k w liniowym uogólnionym estymatorze regresyjnym odzwierciedlają znane wartości globalne wszystkich zmiennych pomocniczych wziętych pod uwagę w budowie modelu liniowego. Równość ta stanowi także fundament teorii kalibracji i pokazuje, że liniowy estymator typu GREG jest szczególnym przypadkiem estymatora kalibracyjnego wartości globalnej dla odpowiednio dobranej funkcji odległości. Będzie to przedmiotem szerszej dyskusji w rozdziale drugim niniejszej książki.

Mimo że liniowy uogólniony estymator regresyjny (1.43) jest w praktyce powszechnie wykorzystywany, istnieją również inne wersje estymatora (1.40) w zależności od przyjętego modelu, który służy do wyznaczania wartości teoretycznych \hat{y}_k . Na przykład, można w tym celu wykorzystać klasę uogólnionych modeli liniowych $g(\mu_k) = \mathbf{x}_k^T \boldsymbol{\beta}$, gdzie $g(\cdot)$ jest odpowiednio dobraną funkcją łączącą. W takim wypadku wartości \hat{y}_k wyznaczamy jako:

$$\hat{y}_k = \hat{\mu}_k = g^{-1}\left(\mathbf{x}_k^T \hat{\boldsymbol{\beta}}\right), \quad (1.50)$$

gdzie $g^{-1}(\cdot)$ oznacza funkcję odwrotną do $g(\cdot)$.

Jeżeli w charakterze funkcji $g(\cdot)$ przyjmimy funkcję logitową: $\text{logit}(\mu_k) = \mathbf{x}_k^T \boldsymbol{\beta}$, to wartości \hat{y}_k można przedstawić jako:

$$\hat{y}_k = \hat{\mu}_k = \frac{e^{\mathbf{x}_k^T \hat{\boldsymbol{\beta}}}}{1 + e^{\mathbf{x}_k^T \hat{\boldsymbol{\beta}}}}. \quad (1.51)$$

W takiej sytuacji mamy do czynienia z przykładem tzw. nieliniowego uogólnionego estymatora regresyjnego typu GREG. Uogólnionym estymatorom regresyjnym wspomaganym modelem regresji logistycznej poświęcona jest praca Lehtonena i Veijanena (1998). Omówieniu ich własności w wypadku losowania zespołowego jest z kolei poświęcony artykuł Kennela i Vallianta (2010). W dalszej części rozważać będziemy jednak uogólniony liniowy estymator regresyjny GREG postaci (1.43).

Asymptotyczną wariancję uogólnionego liniowego estymatora regresyjnego GREG (1.43) można uzyskać, wykorzystując metodę linearyzacji Taylora (Särndal i Lundström, 2005, s. 37):

$$D^2\left(\hat{\tau}_{\text{GREG}}^{\text{lin}}\right) = \sum_{k \in U} \sum_{l \in U} \left(\frac{d_k d_l}{d_{kl}} - 1 \right) e_k e_l, \quad (1.52)$$

gdzie:

$$e_k = y_k - \mathbf{x}_k^T \boldsymbol{\beta}_U \quad (1.53)$$

oraz

$$\boldsymbol{\beta}_U = \left(\sum_{k \in U} q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in U} q_k \mathbf{x}_k y_k. \quad (1.54)$$

Estymator wariancji estymatora typu GREG wartości globalnej może być przedstawiony w poniższy sposób:

$$\hat{D}^2\left(\hat{\tau}_{\text{GREG}}^{\text{lin}}\right) = \sum_{k \in s} \sum_{l \in s} (d_k d_l - d_{kl}) \hat{e}_k \hat{e}_l, \quad (1.55)$$

gdzie reszty \hat{e}_k można wyrazić jako:

$$\hat{e}_k = y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_s \quad (1.56)$$

a $\hat{\boldsymbol{\beta}}_s$ wyraża się wzorem (1.42).

Särndal i Lundström (2005, s. 37) proponują również, aby jako estymator wariancji estymatora typu GREG (1.43) wartości globalnej wykorzystać:

$$\hat{D}_{\text{mod}}^2(\hat{\tau}_{\text{GREG}}^{\text{lin}}) = \sum_{k \in s} \sum_{l \in s} (d_k d_l - d_{kl}) (g_k \hat{e}_k) (g_l \hat{e}_l), \quad (1.57)$$

gdzie g_k wyraża się wzorem (1.46).

Należy jednak zauważyć, że oszacowania wariancji z wykorzystaniem wzorów (1.55) i (1.57) na ogół nie będą się różniły zbyt mocno, gdyż czynniki g_k oscylują zazwyczaj wokół wartości jeden.

Wprowadzony w tej części uogólniony estymator regresyjny GREG można również wykorzystać na potrzeby estymacji wartości globalnej w domenach, na które jest podzielona populacja (Lehtonen, 2009; Myrskylä, 2007; Särndal i Lundström, 2005). Stanowić je mogą na przykład powiaty (przekrój przestrzenny) czy domeny wyróżnione na podstawie jakichś cech (płeć, grupy wieku itd.) bądź ich kombinacje. Ujmując powyższe zagadnienie formalnie, zakładamy, że populacja U składa się z D rozłącznych domen postaci U_1, \dots, U_D . Przyjmujemy też, że w każdej z D wyróżnionych domen przedmiotem zainteresowania jest wartość globalna $\tau_{Y_1}, \dots, \tau_{Y_D}$ zmiennej Y , gdzie:

$$\tau_{Y_d} = \sum_{k \in U_d} y_k, \quad d = 1, \dots, D. \quad (1.58)$$

Funkcję przynależności jednostki k do domeny d definiujemy w następujący sposób:

$$\delta_{dk} = \begin{cases} 1, & \text{jeżeli } k \in U_d, \\ 0, & \text{jeżeli } k \in U \setminus U_d. \end{cases} \quad (1.59)$$

Zakładamy, że $N_d = \sum_{k \in U} \delta_{dk}$ jest liczebnością domeny d , $d = 1, \dots, D$. W takim wypadku wartość interesującej nas cechy dla jednostki k należącej do domeny d możemy zapisać:

$$y_{dk} = \begin{cases} y_k, & \text{jeżeli } k \in U_d, \\ 0, & \text{jeżeli } k \in U \setminus U_d \end{cases} \quad (1.60)$$

lub $y_{dk} = \delta_{dk}y_k$. Wówczas wartość globalna (1.58) może być przedstawiona w następujący sposób:

$$\tau_{Y_d} = \sum_{k \in U} y_{dk}, \quad d = 1, \dots, D. \quad (1.61)$$

Uogólniony liniowy estymator regresyjny GREG wartości globalnej (1.61) w domenie d może być zatem zapisany jako:

$$\hat{\tau}_{\text{GREG},d}^{\text{lin}} = \sum_{k \in s} d_k g_k y_{dk}, \quad (1.62)$$

gdzie g_k zdefiniowane jest wzorem (1.46), a y_{dk} wzorem (1.60).

Estymator wariancji uogólnionego liniowego estymatora regresyjnego GREG można wyrazić wzorem (1.55), zastępując y_k przez y_{dk} . W konsekwencji estymator wariancji można zapisać jako:

$$\hat{D}^2\left(\hat{\tau}_{\text{GREG},d}^{\text{lin}}\right) = \sum_{k \in s} \sum_{l \in s} (d_k d_l - d_{kl}) \hat{e}_{dk} \hat{e}_{dl}, \quad (1.63)$$

gdzie reszty \hat{e}_{dk} można wyrazić wzorem:

$$\hat{e}_{dk} = y_{dk} - \mathbf{x}_k^T \hat{\beta}_{sd} \quad (1.64)$$

oraz

$$\hat{\beta}_{sd} = \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in s} d_k q_k \mathbf{x}_k y_{dk}. \quad (1.65)$$

Rozważane w tej części pracy uogólnione estymatory regresyjne nie są na ogół nieobciążone. Ich obciążenie jednak dąży do zera wraz ze wzrostem liczebności próby. Obciążenie jest małe również zazwyczaj w wypadku dobrze wyspecyfikowanego modelu łączącego zależność między zmienną objaśnianą a wektorem \mathbf{x}_k zmiennych objaśniających. Spełniona jest ponadto poniższa równość³:

$$\frac{(\hat{\tau}_{\text{GREG}} - \tau_Y)}{N} = O\left(n^{-1/2}\right). \quad (1.66)$$

W wypadku złej specyfikacji modelu lub gdy w próbie będą występować jednostki odstające, dla których wyznaczone reszty będą duże, zastosowanie uogól-

³ We wzorze (1.66) wykorzystano zaproponowaną w 1894 roku przez Paula Bachmannna tzw. notację O w analizie asymptotycznej. Notacji O poświęcona jest znakomita książka Grahama, Knutha i Patashnika (2006).

nionych estymatorów regresyjnych GREG skutkuje zazwyczaj znacznym przeszacowaniem lub niedoszacowaniem wartości globalnej. Istnieje również zagrożenie, że estymatory typu GREG mogą zwracać ujemne oszacowania. Jest to konsekwencją przybierania przez niektóre wagi g_k wartości ujemnych. Należy jednak podkreślić, że mimo wskazanych powyżej niedogodności, w wypadku istnienia silnej zależności między zmienną objaśnianą a wektorem zmiennych objaśniających estymatory te charakteryzują się niskim obciążeniem oraz wysoką precyzją. W odniesieniu do liniowego uogólnionego estymatora regresyjnego GREG spełniony jest również warunek (1.48), co ma duże znaczenie w praktycznych zastosowaniach przez krajowe urzędy statystyczne. Zapewniona jest bowiem zgodność rozumiana w ten sposób, że oszacowane na podstawie próby s wartości globalne wszystkich zmiennych pomocniczych x_k równe są znanym wartościom globalnym tych zmiennych w populacji. Warunek ten stanowi fundament teorii kalibracji. Jak pokazano w dalszej części książki, w rozdziale drugim, uogólniony liniowy estymator regresyjny GREG jest szczególnym przypadkiem estymatora kalibracyjnego dla odpowiednio dobranej funkcji odległości.

Estymatory typu GREG często są wykorzystywane w praktycznych badaniach prowadzonych przez krajowe urzędy statystyczne na całym świecie. Estevo, Hidioglou i Särndal (1995) wskazują, że Statistics Canada wykorzystuje je do szacowania wartości globalnych, średnich, odsetków itp. w 20 głównych badaniach z zakresu statystyki przedsiębiorstw oraz w wybranych badaniach społecznych, w tym w kanadyjskim badaniu siły roboczej będącym odpowiednikiem polskiego badania BAEL.

W literaturze przedmiotu można również znaleźć wiele zastosowań estymatorów typu GREG w badaniach społeczno-gospodarczych. Dehnel (2012) wykorzystowała uogólniony liniowy estymator regresyjny na potrzeby statystyki przedsiębiorstw. Wykorzystując dane z badania SP3 (Badanie Działalności Gospodarczej Przedsiębiorstw) oraz zbiory danych z systemu podatkowego Ministerstwa Finansów, w badaniu symulacyjnym porównała własności estymatora typu GREG z lokalnym estymatorem regresyjnym. Przedmiotem estymacji była suma wynagrodzeń brutto przedsiębiorstw w przekroju województw i rodzaju prowadzonej działalności gospodarczej (sekcja PKD). Kompleksową analizę wykorzystania estymatorów GREG na potrzeby statystyki gospodarczej wraz z omówieniem ich własności autorka przedstawiła też w innej monografii (Dehnel, 2010).

Z kolei Gołata (2004) dokonała oszacowania wybranych charakterystyk rynku pracy (m.in. liczby bezrobotnych lub pracujących) w przekroju powiatów województwa wielkopolskiego. W tym celu rozważała wykorzystanie estymatorów

konstruowanych w statystyce małych obszarów, w tym estymatora typu GREG. Również Żądło (2015) w badaniu symulacyjnym, wykorzystując rzeczywiste dane na poziomie gmin, dotyczące rynku pracy, dokonał porównania własności różnych wariantów estymatora GREG w odniesieniu do zmiennej, którą była liczba zarejestrowanych bezrobotnych.

Estymatory typu GREG zostały także wykorzystane w obszarze bardzo ważnego zjawiska społecznego, którym jest niepełnosprawność. Klimanek, Szymkowiak i Józefowski (2018) na potrzeby estymacji odsetka osób niepełnosprawnych prawnie i biologicznie w liczbie ludności w wieku 15 lat i więcej w przekroju powiatów województwa wielkopolskiego i odpowiednich kategorii wykształcenia zastosowali sześć różnych estymatorów z zakresu statystyki małych obszarów, w tym liniowy uogólniony estymator regresyjny GREG. Wykorzystując rzeczywiste dane z NSP 2011, pokazali, że przestrzenny rozkład odsetka osób niepełnosprawnych prawnie i biologicznie w domenach zdefiniowanych jako przekrój powiatu i poziomu wykształcenia na obszarze województwa wielkopolskiego nie jest równomierny. Wykazali również, że estymacja typu GREG charakteryzowała się niskimi średnimi błędami szacunku.

Na zakończenie tego punktu warto podkreślić, że w badaniach statystycznych prowadzonych przez krajowe urzędy statystyczne zaleca się, aby w wypadku estymacji dla małych obszarów w pierwszej kolejności brać pod uwagę, oprócz estymatora bezpośredniego i złożonego, estymator typu GREG. Dopiero wówczas, gdy wyniki uzyskane z wykorzystaniem tych trzech estymatorów dają nieakceptowalne szacunki (na przykład zbyt niską precyzję oszacowań), postuluje się, aby wykorzystać bardziej skomplikowane metody, które oferuje statystyka małych obszarów, bazujące na podejściu modelowym (Tzavidis, Zhang, Luna Hernandez, Schmid i Rojas-Perilla, 2018).

1.5. Rozwój teorii kalibracji

Podejście kalibracyjne, które jest przedmiotem rozważań w niniejszej książce, zostało formalnie opisane w artykule Deville'a i Särndala (1992). Praca ta była kamieniem milowym w rozwoju teorii estymacji, przede wszystkim w kontekście wyznaczania wag z uwzględnieniem zmiennych pomocniczych, które można wykorzystać w procesie uogólniania wyników na całą populację lub odpowiednie subpopulacje (domeny). Warto jednak zauważyć, że sama idea wykorzystania zmiennych pomocniczych oraz ich wartości globalnych na potrzeby estymacji znana już była w latach 40. XX wieku. W pracy Deminga i Stephana (1940) po-

kazano sposób konstruowania estymatora ilorazowego wykorzystującego technikę iteracyjnego proporcjonalnego dopasowania (ang. *iterative proportional fitting* – IPF) na potrzeby oszacowań wartości globalnych w dwuwymiarowych tablicach kontyngencji. Ideę tę można było również uogólnić w celu estymacji liczebności w tablicach kontyngencji większych wymiarów.

Z formalnego punktu widzenia podejście zaproponowane przez Deminga i Stephana (1940) odnosiło się do dwuwymiarowych tablic kontyngencji $t \times w$, gdzie t oznacza liczbę wierszy, a w liczbę kolumn w tablicy. W rozważanej metodzie Deming i Stephan założyli, że estymacji będą podlegać nieznane liczebności N_{ij} w dwuwymiarowej tablicy kontyngencji, przy czym $i = 1, \dots, t$ oraz $j = 1, \dots, w$. Przyjęli przy tym, że $N = \sum_{i=1}^t \sum_{j=1}^w N_{ij}$. Ponadto założyli, że wartości brzegowe $N_{i.} = \sum_{j=1}^w N_{ij}$ i $N_{.j} = \sum_{i=1}^t N_{ij}$, gdzie $i = 1, \dots, t$, $j = 1, \dots, w$, są znane. Wartości brzegowe mogą pochodzić na przykład z odpowiednich badań reprezentacyjnych, a same liczebności N_{ij} w tablicy kontyngencji są nieznane i muszą podlegać estymacji (tab. 1.1).

Tabela 1.1. Dwudzielcza tablica kontyngencji ($t \times w$)

Kategorie zmiennej X	Kategorie zmiennej Y				Ogółem
	Y_1	Y_2	...	Y_w	
X_1	N_{11}	N_{12}	...	N_{1w}	$N_{1.}$
X_2	N_{21}	N_{22}	...	N_{2w}	$N_{2.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
X_t	N_{t1}	N_{t2}	...	N_{tw}	$N_{t.}$
Ogółem	$N_{.1}$	$N_{.2}$...	$N_{.w}$	N

Założmy, że n oznacza liczebność próby s , a n_{ij} liczebność tej części próby s , która odpowiada komórce (i, j) . Liczebności n_{ij} w analizowanej tablicy kontyngencji mogą pochodzić z badania reprezentacyjnego, a także z innych źródeł, takich jak poprzedni spis czy odpowiedni rejestr administracyjny. Ze względu na to, że liczebności n_{ij} w końcowej tabeli na ogół nie będą się sumowały do odpowiednich liczebności brzegowych $N_{i.}$ oraz $N_{.j}$, należy dokonać ich odpowiedniej korekty (kalibracji) tak, aby spełnione były równania:

$$N_{i.} = \sum_{j=1}^w N_{ij}, \quad (1.67)$$

oraz

$$N_{.j} = \sum_{i=1}^t N_{ij}. \quad (1.68)$$

Oszacowania nieznanych liczebności N_{ij} w komórce (i, j) można otrzymać z wykorzystaniem estymatorów $n_{ij} \frac{N}{n}$, $n_{ij} \frac{N_{i.}}{n_{i.}}$ lub $n_{ij} \frac{N_{.j}}{n_{.j}}$, przy czym $n_{i.} = \sum_{j=1}^w n_{ij}$ i $n_{.j} = \sum_{i=1}^t n_{ij}$. Niestety, zaproponowane estymatory mają tę własność, że zsumowane oszacowania po wierszach lub kolumnach dwuwymiarowej tablicy kontyngencji niekoniecznie muszą odtwarzać znane wartości brzegowe w tablicy jednocześnie. Nie jest więc spełniony warunek spójności polegający na tym, że odpowiednie oszacowania w środku tablicy kontyngencji są zgodne z wartościami brzegowymi.

Przyjmując m_{ij} za poszukiwane oszacowania liczebności N_{ij} w odpowiedniej tablicy kontyngencji, w pracy Deminga i Stephana (1940) wstępnie rozważano minimalizację funkcji odległości:

$$\phi = \sum_{i=1}^t \sum_{j=1}^w \frac{(m_{ij} - n_{ij})^2}{n_{ij}}, \quad (1.69)$$

przy spełnieniu poniższych warunków:

$$\sum_{j=1}^w m_{ij} = N_{i.} \frac{n}{N}, \quad i = 1, \dots, t, \quad (1.70)$$

$$\sum_{i=1}^t m_{ij} = N_{.j} \frac{n}{N}, \quad j = 1, \dots, w - 1. \quad (1.71)$$

Zaproponowany przez nich algorytm IPF umożliwił co prawda spełnienie równań (1.70) i (1.71), nie minimalizował jednak funkcji odległości danej wzorem (1.69). Ostatecznie Ireland i Kullback (1968) pokazali, że liczebności \hat{m}_{ij} będące rozwiązaniem powyższego zadania optymalizacyjnego minimalizują tzw. funkcję odległości Kullbacka-Leiblera, określaną również mianem funkcji dyskryminacyjnej czy względnej entropii (por. wzór (2.20)). Opisana powyżej sytuacja, w której poszukuje się w tablicy kontyngencji nowych liczebności tak, aby odtworzone zostały znane wartości brzegowe, odnosi się do tzw. estymatorów zachowujących strukturę (ang. *structure preserving estimation* – SPREE). Warto wspomnieć, że zasygnalizowana powyżej metoda poszukiwania liczebności w tablicy kontyngencji, bazująca na algorytmie IPF, określaną jest mianem rakingu (ang. *raking ratio estimation*⁴). Jest ona wykorzystywana przez wiele krajowych urzędów statystycznych w badaniach próbkowych. Na przykład, Statistics Canada stosuje

⁴ W polskiej literaturze brak odpowiedniego tłumaczenia nazwy tej metody. Często przy jej opisie pozostawia się oryginalną nazwę.

ją w okresach międzypisowych na potrzeby uaktualniania odpowiednich tablic kontyngencji. W tym celu wykorzystuje zarówno dane pochodzące z ostatniego spisu, jak i z bieżących badań reprezentacyjnych (Binder i Théberge, 1988). Raking jako metoda estymacji był również przedmiotem zainteresowań wielu innych badaczy. Dotyczy to zarówno metod wyznaczania wariancji tego typu estymatora (Brackstone i Rao, 1979; Oh i Scheuren, 1987), jak i jego innych praktycznych zastosowań, m.in. w badaniu handlu detalicznego (Hidiroglou i Patak, 2006).

Metodę zaproponowaną przez Deminga i Stephana (1940) można uznać za początek podejścia kalibracyjnego stanowiącego przedmiot rozważań w niniejszej książce i określić mianem kalibracji zagregowanej czy estymacji typu SPREE. Dokonujemy bowiem kalibracji zagregowanych wartości w środku odpowiedniej tablicy kontyngencji, tak aby odtworzyć znane wartości brzegowe. Ciekawe zastosowanie estymatorów typu SPREE wraz z wyjaśnieniem idei kalibracji wartości zagregowanych w środku tablicy kontyngencji w celu odtworzenia oszacowań wartości brzegowych przedstawiają na przykładzie niepełnosprawności Szymkowiak, Józefowski i Klimanek (w druku). Z kolei Józefowski i Szymkowiak (2013) na przykładzie rynku pracy pokazują praktyczne wykorzystanie estymatorów typu SPREE dla trójwymiarowych tablic kontyngencji. Wskazują jednocześnie na podobieństwo konstruowania tego typu estymatorów do konstruowania estymatorów kalibracyjnych wartości globalnej.

Jak wspomniano na początku tego podrozdziału, za kluczową w rozwoju teorii kalibracji uważa się pracę Deville'a i Särndala (1992). Autorzy sformułowali podwaliny pod nową metodę, która obecnie jest często wykorzystywaną techniką w obszarze badań próbkowych. Praca ta stanowiła również impuls do rozwoju teorii kalibracji, która obecnie obejmuje m.in. zagadnienia związane z brakami odpowiedzi (Särndal i Lundström, 2005), wartościami odstającymi (Duchesne, 1999), podejściem wspomaganym modelem (Wu i Sitter, 2001) czy hybrydowym (Lehtonen i Veijanen, 2012; 2015; 2017). Deville i Särndal (1992) w swoim artykule pokazali bowiem możliwości korygowania wag d_k wynikających z planu losowania próby w nieco idealnej sytuacji, tj. abstrahując od występujących we wszystkich badaniach błędów nielosowych w postaci braków danych, wartości odstających czy błędów pokrycia. Sformułowane przez nich podejście kalibracyjne odnosiło się do korygowania wag wynikających z planu losowania próby, tak aby odtworzone zostały odpowiednie równania kalibracyjne i aby zminimalizować wartość właściwie dobranej funkcji odległości⁵. Pokazali, że dla różnie dobranych funkcji

⁵ W kolejnych częściach książki teoria kalibracji zostanie bardziej sformalizowana z wykorzystaniem odpowiedniego aparatu matematycznego.

odległości estymatory kalibracyjne wartości globalnej są asymptotycznie równoważne, a szczególnym przypadkiem estymatorów kalibracyjnych jest wspomniany już uogólniony estymator regresyjny GREG. Co więcej, dzięki zastosowaniu kalibracji można osiągnąć zgodność rozumianą w ten sposób, że odtwarzane są znane struktury w odniesieniu do kluczowych zmiennych pomocniczych oraz zwiększa się precyzja oszacowań.

Esteveo i Särndal (2000) pokazali następnie, że estymatory kalibracyjne wartości globalnej można również konstruować z wykorzystaniem tzw. podejścia funkcyjnego i koncepcji wektorów zmiennych instrumentalnych. Dla odpowiednio dobranych parametrów wskazali, że wiele znanych estymatorów kalibracyjnych (na przykład estymator GREG) jest szczególnym przypadkiem takiego podejścia. Zagadnieniu temu poświęcony został rozdział trzeci niniejszej książki.

W procesie wyznaczania wag kalibracyjnych nie wszystkie wartości globalne zmiennych pomocniczych mogą zostać dokładnie odtworzone. W takiej sytuacji można wykorzystać kalibrację z karami (ang. *penalized calibration*), w której wagi poszukuje się, wprowadzając dodatkowe kary na zmienne pomocnicze, w zależności od stopnia niespełnienia odpowiedniego równania kalibracyjnego. W takim podejściu można wymusić zatem, aby niektóre równania kalibracyjne dla pewnych kluczowych zmiennych były dokładnie odtworzone, podczas gdy dla pozostałych zmiennych pozostawia się pewien margines błędu. Problem braku odtworzenia niektórych wartości globalnych najczęściej jest związany z koniecznością spełnienia zbyt dużej liczby równań kalibracyjnych, tj. z sytuacją, gdy dysponujemy dużą liczbą zmiennych pomocniczych. Guggemos i Tillé (2010) określają tę sytuację jako zagadnienie „nadmiernej kalibracji” (ang. *over-calibration*). Idea estymacji kalibracyjnej w takim wypadku polega na tym, aby złagodzić warunki występujące w równaniach kalibracyjnych, nakładając pewną funkcję straty, która „karze” duże wartości różnic $\sum_{k \in S} w_k \mathbf{x}_k - \sum_{k \in U} \mathbf{x}_k$ wynikających z niespełnienia równania (1.48). Zagadnieniu temu poświęcony jest artykuł Bardsleya i Chambersa (1984) czy wspomniana już praca Guggemosa i Tillé (2010).

Niestety, w badaniach statystycznych rzadko mamy do czynienia z brakiem błędów nielosowych (głównie w postaci braków danych), co ma wpływ na jakość szacunków. W związku z tym, w ostatnich latach dużą wagę przywiązuje się do wykorzystania podejścia kalibracyjnego, w którym uwzględnia się bardziej realną sytuację, tj. występowanie w badaniu braków odpowiedzi. Kott (2006), Kott i Chang (2010), Lundström i Särndal (1999), Särndal i Lundström (2005) czy Thiabaudeau i Slud (2009) zaproponowali różne podejścia do wykorzystania kalibracji, aby skorygować obciążenie będące konsekwencją występowania w badaniu braków odpowiedzi. W tym miejscu zasygnalizowana zostanie jedynie możliwość

zastosowania podejścia kalibracyjnego w badaniach z brakami odpowiedzi, gdyż kwestii tej poświęcono odrębny rozdział książki.

Inny nurt w obszarze kalibracji obejmuje estymatory kalibracyjne wykorzystujące podejście modelowe (modelowo-kalibrowane), którego teoretyczne podstawy zostały sformułowane w pracy Wu i Sittera (2001) oraz Wu (2003). W podejściu tym zakłada się, że istnieje pewien model wiążący zmienną objaśnianą ze zmiennymi pomocniczymi x_1, x_2, \dots, x_J , a wagi kalibracyjne wyznacza się w taki sposób, aby poprzez minimalizację odpowiedniej funkcji odległości odtwarzały znaną liczebność populacji N oraz wartość globalną złożoną z wartości teoretycznych wynikających z przyjętego modelu. Estymatory modelowo-kalibrowane, ze względu na dużą elastyczność w wyborze modelu, mogą być przydatne w estymacji wartości globalnej zmiennej dychotomicznej, politomicznej czy ciągłej. W wypadku odpowiednio dobranego modelu obserwowalny jest wzrost precyzji estymatorów modelowo-kalibrowanych w porównaniu z estymatorami kalibracyjnymi nieuwzględniającymi w swojej konstrukcji modelu (ang. *model-free calibration*). Ich wadą jest jednak to, że dla różnych zmiennych objaśnianych należy każdorazowo budować odrębny model, co oznacza konieczność wyznaczania oddzielnych zestawów wag kalibracyjnych. Montanari i Ranalli (2009) zaproponowali metodę przewyciężenia tej niedogodności, tj. stworzenia jednego systemu wag kalibracyjnych, co ma zapewnić spójność dla wszystkich kluczowych zmiennych wziętych pod uwagę w procesie kalibracji. Odnosi się to zarówno do wszystkich zmiennych pomocniczych, jak i do zmiennych objaśnianych, dla których budowane są odpowiednie modele służące do wyznaczania wartości teoretycznych. Na zakończenie warto dodać, że wartości modelowe, do których wartości globalnych kalibruje się wagi wynikające z planu losowania, można wyznaczyć z wykorzystaniem modeli zarówno parametrycznych (Wu, 2003), jak i nieparametrycznych (Montanari i Ranalli, 2005), na przykład sieci neuronowych. Ciekawą koncepcję wykorzystania tzw. hybrydowych estymatorów kalibracyjnych, nawiązującą także do podejścia modelowego, zaproponowali Lehtonen i Veijanen (2015). W podejściu tym wagi kalibracyjne wyznaczone są w taki sposób, aby odtworzone zostały zarówno znane wartości globalne zmiennych pomocniczych, jak i wartość globalna złożona z wartości teoretycznych, uzyskana na podstawie odpowiednio zbudowanego modelu między zmienną objaśnianą a zmiennymi pomocniczymi. Estymatory modelowo-kalibrowane będą przedmiotem rozważań w dalszej części książki.

Warto również wspomnieć, że estymatory kalibracyjne mogą być wykorzystywane w badaniach statystycznych z wartościami odstającymi. Sytuacja taka dość często występuje w badaniach gospodarczych, w których rozkłady cech charakte-

ryzuje silna asymetria prawostronna (na przykład przychód przedsiębiorstw). Dotyczy to również badań społecznych takich jak BBGD czy EU-SILC, w których cechy, takie jak dochody czy wydatki gospodarstw domowych, także odznaczają się silną asymetrią prawostronną i występowaniem wartości odstających. Zagadnieniu konstruowania odpornych estymatorów kalibracyjnych, które biorą pod uwagę występowanie wartości odstających w próbie, poświęcona jest praca Duchesne'a (1999).

Problem konstruowania wag kalibracyjnych na różnych poziomach (na przykład gospodarstw domowych i należących do nich osób) jest z kolei przedmiotem tzw. kalibracji zintegrowanej. Ten rodzaj kalibracji jest szczególnie pożądany w badaniach gospodarstw domowych. Jej idea polega na tym, aby wszystkie osoby przypisane do danego gospodarstwa domowego miały te same wagi oraz aby zostały odtworzone odpowiednie równania kalibracyjne dla osób oraz gospodarstw domowych. Jest to szczególnie pożądana własność, gdy zachodzi potrzeba jednoczesnej estymacji pewnych parametrów dla osób oraz gospodarstw domowych. Bardzo proste rozwiązanie tego problemu zostało zaproponowane w pracy Alexandra (1987). W podejściu tym przypisuje się wagę odpowiedniej osoby (na przykład głowy gospodarstwa domowego) do całego gospodarstwa. Inne rozwiązanie tego problemu było rozpatrywane w pracy Lemaître'a i Dufoura (1987), a sprowadzało się do utworzenia zmiennych kalibracyjnych wykorzystywanych w procesie kalibracji na poziomie osób, którym odpowiadają średnie wartości zmiennych pomocniczych wyznaczanych w obrębie każdego gospodarstwa domowego oddzielnie. Ideę zaproponowaną przez Lemaître'a i Dufoura rozwinęli następnie Renssen i Nieuwenbroek (1997) na przypadek wielu badań, w których występują wspólne zmienne, oraz Neethling i Galpin (2006), którzy rozważali problem wyznaczania wag kalibracyjnych z wykorzystaniem tylko zmiennych pomocniczych na poziomie gospodarstwa domowego bądź gospodarstwa domowego i osób jednocześnie. W szczególności, w pracy Neethlinga i Galpina wykazano, że uwzględnienie zmiennych pomocniczych w procesie kalibracji zarówno na poziomie gospodarstw domowych, jak i osób poprawia dokładność i precyzję oszacowań w wypadku występowania zjawiska braków odpowiedzi. Warto również wspomnieć o pracach Estevao i Särndala (2003; 2006) oraz Kolenikova i Hammera (2015), w których zagadnienie kalibracji zintegrowanej było szczegółowo dyskutowane zarówno od strony teoretycznej, jak i praktycznej. Kalibracja zintegrowana będzie również przedmiotem rozważań w rozdziale szóstym książki, w kontekście spisów realizowanych metodą mieszaną.

W literaturze przedmiotu najwięcej uwagi poświęca się konstruowaniu estymatorów kalibracyjnych wartości globalnej bądź średniej. Rozwój teorii kalibracji

nie ograniczał się jednak wyłącznie do tych dwóch ważnych charakterystyk w populacji. W pracach Krapavickaitė i Plikusasa (2005) oraz Plikusasa (2001; 2003) proponuje się zastosowanie podejścia kalibracyjnego do szacowania ilorazu wartości globalnych w populacji (por. wzór (1.3)). Z kolei w pracach Plikusasa i Pumputisa (2007; 2010) rozważa się zastosowanie paradygmatu kalibracji do szacowania kowariancji w populacji (por. wzór (1.10)).

Wiele uwagi poświęca się również estymacji kwantyla rzędu α w populacji (por. wzór (1.4)) z wykorzystaniem podejścia kalibracyjnego. W tym miejscu warto podkreślić, że próby konstruowania estymatorów kwantyla rzędu α z użyciem zmiennych pomocniczych były już znane, zanim kalibracja stała się powszechnie wykorzystywaną metodą w badaniach próbkowych. Jedną z pierwszych wzmianek na temat wykorzystania informacji dodatkowych w szacowaniu kwantyli można znaleźć w pracy Chambersa i Dunstana (1986). W artykule tym autorzy, stosując podejście wspomagane modelem, rozważali metodę konstruowania estymatorów kwantyla rzędu α , bazując na estymatorze funkcji rozkładu z uwzględnieniem zmiennych pomocniczych. Metoda ta, jak wykazali później Chambers, Dorfman i Dunstan (1992), może jednak dawać gorsze oszacowania w porównaniu z klasycznymi estymatorami – w wypadku źle dobranego modelu opisującego rozkład analizowanej cechy w populacji. Do podobnych wniosków, na podstawie zastosowanego podejścia symulacyjnego, doszli Rao, Kovar i Mantel (1990). Jak wykazali, w razie niewłaściwego modelu opisującego rozkład badanej cechy w populacji, podejście modelowe może dawać gorsze wyniki w porównaniu z zastosowaniem klasycznych estymatorów kwantyli, których konstrukcja jest oparta tylko na informacjach pochodzących z próby.

Jednym z pierwszych badaczy, którzy zwrócili uwagę na możliwość zastosowania podejścia kalibracyjnego w konstruowaniu estymatorów kwantyla rzędu α , był Kovačević (1997). W zaproponowanej metodzie wykorzystał kalibrację do budowy estymatora dystrybuanty, opierając się na odpowiednich momentach zmiennych pomocniczych. Warto dodać, że podejście to znalazło następnie zastosowanie w badaniach panelowych europejskich gospodarstw domowych. Z ostatnich osiągnięć z zakresu estymatorów kalibracyjnych kwantyli rzędu α na szczególną uwagę zasługują prace Rena (2002). Jak podają Harms i Duchesne (2006) – był on pierwszym badaczem, który w kompleksowy sposób zastosował paradygmat kalibracji sformułowany przez Deville'a i Särndala (1992) w odniesieniu do estymatora kalibracyjnego dystrybuanty i kwantyla rzędu α . Dalsze prace z zakresu kalibracji nawiązywały do wspomnianego artykułu Rena. Na szczególną uwagę zasługuje przy tym metoda wyznaczania estymatorów kalibracyjnych kwantyli zaproponowana przez Harmsa i Duchesne'a (2006), bazująca na koncepcji tzw.

dystrybuanty interpolacyjnej. Szymkowiak (2009b) rozszerzył tę koncepcję na przypadek badań z brakami danych, przy czym dokonał również uogólnienia estymatora kalibracyjnego kwantyla rzędu α w sytuacji, gdy znane są kwantyle dowolnego rzędu zmiennych pomocniczych. Warto również wspomnieć o pracach, w których wykorzystano koncepcję modelowo-kalibrowanych estymatorów i regresji nieparametrycznej na potrzeby estymacji kwantyla rzędu α (Rueda, S. Martínez-Puertas, H. Martínez-Puertas i Arcos, 2006, 2007; Rueda, Sánchez-Borrogo, Arcos i Martínez, 2010).

Kalibracja jako technika korygowania wag d_k wynikających z planu losowania próby zazwyczaj jest wykorzystywana w badaniach prowadzonych w jednym okresie. Można ją jednak również zastosować w kontekście badań wielookresowych, w których estymacji parametrów dokonuje się w populacjach czy subpopulacjach w różnych okresach. Na przykład, może to dotyczyć wspomnianego już wcześniej kwartalnego Badania Aktywności Ekonomicznej Ludności. Zagadnieniu wykorzystania podejścia kalibracyjnego w badaniach wielookresowych poświęcona jest praca Żądły (2011; 2015).

Ciekawą koncepcję tworzenia systemu wag, wykorzystującą podejście kalibracyjne, jest tzw. metoda powtarzanego ważenia (ang. *repeated weighting method* – RWM). Wykorzystywana jest ona przez Statistics Netherlands w tzw. spisach wirtualnych. Jej celem jest każdorazowe tworzenie dla z góry zdefiniowanych tabel spisowych odpowiednich wag, tak aby zapewnić zgodność i spójność wszystkich tabel wynikowych, które powstają z wykorzystaniem różnych źródeł danych zasilających spis. W podejściu tym wagi są wyznaczane w sekwencyjny sposób, oddzielnie dla każdej tabeli spisowej. Stąd nazwa techniki, gdyż proces konstruowania wag jest powtarzany każdorazowo dla kolejnych tabel spisowych (de Waal, 2015; Knottnerus i van Duin, 2006; Lumiste, 2011). Särndal i Traat (2011) wskazują, że metoda powtarzanego ważenia stanowi pewną odmianę podejścia kalibracyjnego.

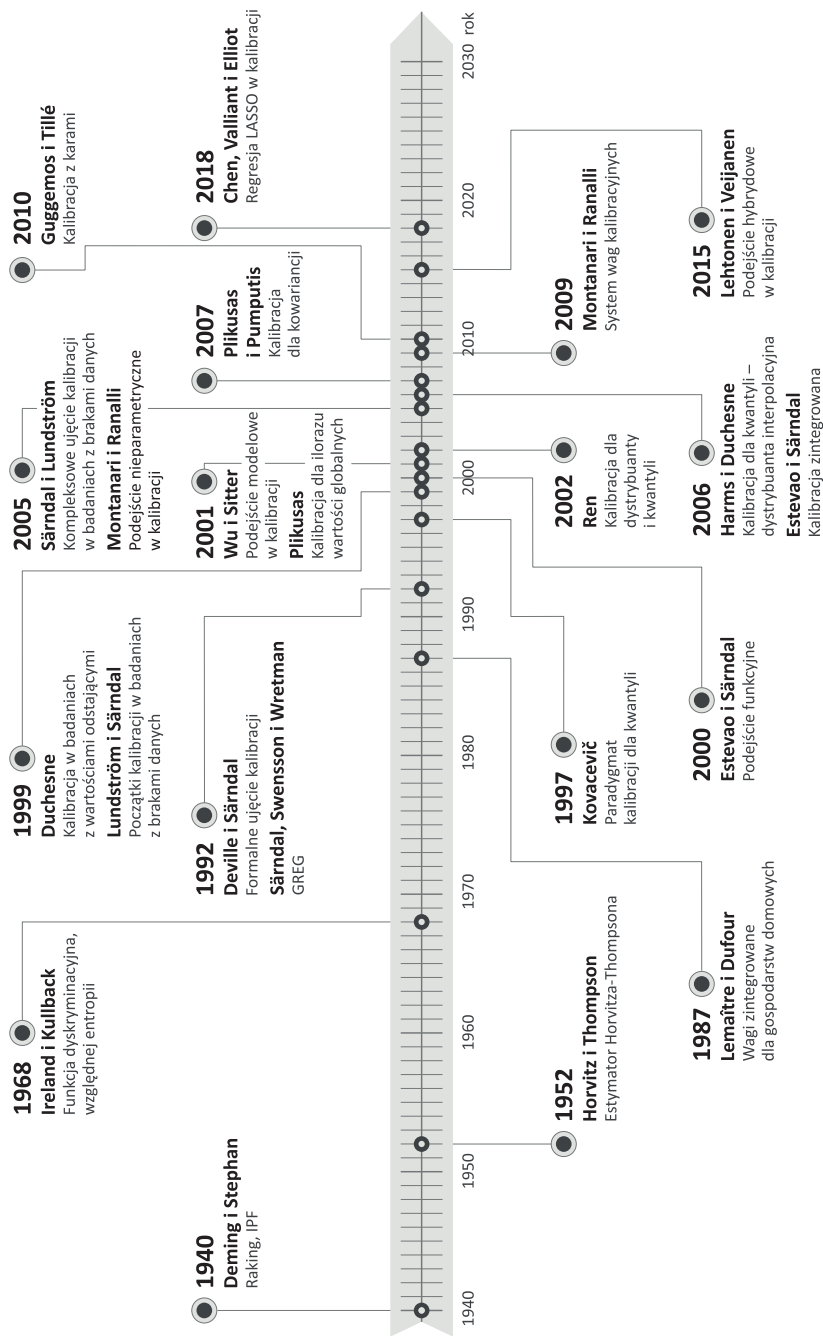
Na zakończenie rozważań poświęconych rozwojowi teorii kalibracji w badaniach statystycznych warto wspomnieć o możliwości jej wykorzystania w odniesieniu do prób nielosowych (na przykład danych pochodzących z internetu, portali społecznościowych czy szeroko rozumianych big data). Jak podają Rao i Fuller (2017), próby nielosowe są wykorzystywane w statystyce już od czasów Neymana, a ich szczególnym przypadkiem są próby kwotowe mające zastosowanie w badaniach marketingowych, głównie ze względu na konieczność ograniczania kosztów. Szreder (2010b, s. 171) zauważa ponadto, że w niektórych sytuacjach wykorzystanie prób nielosowych jest koniecznością. Może to wynikać z trudności w zapewnieniu każdej jednostce z populacji jednakowej możliwości dostania się

do próby (brak operatu losowania lub jego zła jakość). Konieczność ta może być również pochodną krótkiego czasu przeznaczonego na realizację badania.

Pierwsze zastosowania prób kwotowych można znaleźć w opracowaniach Sudmana (1966; 1976). Z kolei Moser i Stuart (1953) oraz Stephan i McCarthy (1958) byli pierwszymi badaczami, którzy dokonali porównania prób kwotowych z próbami losowymi. Należy jednak podkreślić, że ze względu na rosnące koszty badań, obciążenia respondentów i obserwowalne malejące wskaźniki udzielanych w badaniach odpowiedzi, próby nielosowe będą miały obecnie coraz większe zastosowanie (Elliot i Valliant, 2017). Dotyczyć to będzie zarówno badań prowadzonych przez prywatne agencje badawcze, jak i badań realizowanych przez krajowe urzędy statystyczne.

Zagadnieniu wykorzystania podejścia kalibracyjnego z wykorzystaniem regresji LASSO (ang. *Least Absolute Shrinkage and Selection Operator*) w kontekście prób nielosowych poświęcone są prace Chena (2016) oraz Chena i współautorów (2018a; 2018b). Kluczową sprawą w wypadku możliwości wykorzystania kalibracji w próbach nielosowych jest utworzenie wektora złożonego z wag d_k , który następnie podlegać będzie procesowi korygowania z uwzględnieniem zmiennych pomocniczych. W próbach nielosowych nie mają jednak zastosowania znane z metody reprezentacyjnej schematy losowania, w związku z czym pojawia się zasadniczy problem konstruowania tego typu wag. We wspomnianych powyżej pracach autorzy, wykorzystując koncepcję regresji LASSO, dokonują konstrukcji wag w estymatorze kalibracyjnym wartości globalnej (1.1). Na potrzeby ich zastosowania tworzą sztuczne próby, do których jednostki są wylosowane z różnymi prawdopodobieństwami inkluzji pierwszego rzędu, przy czym przypisują wyjściowe wagi podlegające kalibracji jako $d_k = N/n$. Zabieg taki ma symulować próby nielosowe.

Innym sposobem konstrukcji wag d_k w próbach nielosowych, które można kalibrować, jest utworzenie tzw. pseudowag. W podejściu tym zakłada się, że próba nielosowa podlega pewnemu mechanizmowi losowego doboru jednostek z prawdopodobieństwem, które należy oszacować przy spełnieniu określonych założeń. Celem jest zatem estymacja prawdopodobieństw wylosowania jednostki do takiej próby z wykorzystaniem na przykład danych spisowych, zawierających wspólne zmienne, które wyjaśnia nieznaną schemat losowania (Elliot i Valliant, 2017). Zagadnieniu modelowania pseudowag, które są przypisane wszystkim jednostkom znajdującym się w próbie nielosowej, a wykorzystywanym w procesie uogólniania wyników na całą populację, poświęcona jest również praca Elliota i współautorów (2010). Można również wszystkim jednostkom znajdującym się w próbie nielosowej przypisać sztuczne wagi tożsamościowo równe jeden i kalibrować je



Rysunek 1.1. Ewolucja teorii kalibracji

Źródło: na podstawie badań literaturowych.

z wykorzystaniem zmiennych pomocniczych pochodzących z alternatywnych źródeł informacji. Tego typu podejście było przedmiotem rozważań w pracy Beręsewicza i Szymkowiaka (2018), którzy zastosowali metodę kalibracji jako technikę redukcji obciążenia estymatorów na skutek występowania zjawiska autoselekcji. W tym celu wykorzystali dane z Badania Ofert Pracy, stanowiącego moduł badawczy w tzw. Bilansie Kapitału Ludzkiego. Jako zmienne pomocnicze uwzględnione zostały zmienne z Badania Popytu na Pracę realizowanego przez Główny Urząd Statystyczny.

Na rysunku 1.1 przedstawiono najważniejsze etapy w rozwoju podejścia kalibracyjnego, które uwzględniają kluczowe prace z tego zakresu⁶. Stanowi on podsumowanie rozważań zawartych w tym podrozdziale, jednocześnie wskazuje, że intensywny rozwój kalibracji przypada na lata 1992–2018. Warto zwrócić uwagę, że najnowsze prace z obszaru kalibracji odnoszą się do ostatnich 2–3 lat i będą wyznaczać najnowsze trendy w tym zakresie, uwzględniające przede wszystkim możliwości, jakie podejście kalibracyjne oferuje w wypadku prób nielosowych.

1.6. Wybrane zastosowania podejścia kalibracyjnego

Podejście kalibracyjne, które jest przedmiotem rozważań w książce, wykorzystywane jest na szeroką skalę przez krajowe urzędy statystyczne na całym świecie. Stosują je również agencje badania rynku w przeprowadzanych przez siebie badaniach ankietowych.

Niektóre z zastosowań estymatorów kalibracyjnych omówiono już w podrozdziale 1.4. Wskazano na praktyczne wykorzystanie estymatora GREG, będącego szczególnym przypadkiem estymatora kalibracyjnego, w wybranych badaniach statystycznych. W podrozdziale 1.5 poświęconym rozwojowi teorii podejścia kalibracyjnego wzmiankowano o wykorzystaniu kalibracji. Ze względu na różnorodne zastosowania omawianej metody w tym podrozdziale wskażemy na kilka przykładów użycia kalibracji w praktyce badań statystycznych. Przegląd ograniczono jedynie do najnowszych osiągnięć w tym zakresie.

Anastasiade i Tillé (2017) wykorzystali koncepcję kalibracji w badaniu różnic płacowych między kobietami a mężczyznami w Szwajcarii na przykładzie badania struktury wynagrodzeń (ang. *Structure of Earnings Survey*). W tym celu dokonali estymacji rozkładów kontrfaktycznych wynagrodzeń, w których rozpa-

⁶ Rysunek ten ma w pewnej mierze subiektywny charakter. Uwzględniono w nim wybrane, aczkolwiek kluczowe zdaniem autora, momenty z punktu widzenia rozwoju kalibracji.

trywano kobiety o cechach męskich oraz mężczyzn o cechach kobiet. Polegało to na takiej korekcie wag wynikających z planu losowania próby, a przypisanych w badaniu do zbiorowości kobiet, aby odtwarzać znane wartości globalne dla określonych charakterystyk w grupie mężczyzn. Dzięki temu możliwe było skonstruowanie estymatora średnich wynagrodzeń kontrfaktycznych w grupie kobiet (ang. *Women's Counterfactual Wage Mean Estimator*), który uwzględniał wagi kalibracyjne przypisane kobietom, ale odtwarzające wartości globalne dla mężczyzn. Podejście takie umożliwiło nie tylko oszacowanie wynagrodzeń w grupie kobiet i mężczyzn, ale również pogłębioną charakterystykę problemu dyskryminacji płacowej ze względu na płeć.

Klimanek i Szymkowiak (2017) przedstawili praktyczne zastosowanie kalibracji w badaniu pełnym na przykładzie Badania Losów Absolwentów Uniwersytetu Ekonomicznego w Poznaniu, w którym kluczową kwestią w procesie estymacji były odmowy wzięcia w nim udziału przez absolwentów uczelni. W swojej pracy wykorzystali podejście kalibracyjne, które na ogół jest stosowane w badaniach reprezentacyjnych do korygowania wag wynikających z planu losowania próby. Autorzy zaprezentowali, jak poprzez dobór właściwych zmiennych pomocniczych i odpowiednią konstrukcję wag wejściowych można zredukować negatywny wpływ braków odpowiedzi w pełnym badaniu losów absolwentów. W pracy zaprezentowali również zagadnienia dotyczące oceny jakości precyzji uzyskanych wyników estymacji.

Cohen, Ben-Hur i Burck (2017) przedstawili konstrukcję estymatora wariancji wartości globalnej w losowaniu wielofazowym, uwzględniając podejście kalibracyjne, w którym wagi są korygowane w każdej z wyszczególnionych faz. W tym celu wykorzystali dane pochodzące z ankietowego badania karier i mobilności doktorantów (ang. *Survey on Career and Mobility of Doctorate Holders*) i w badaniu symulacyjnym dokonali porównania własności zaproponowanego przez siebie estymatora wariancji z innymi estymatorami wariancji wartości globalnej.

Z kolei Wilak, Beręsewicz i Szymkowiak (2017a) opisali zastosowanie podejścia kalibracyjnego w Badaniu Aktywności Ekonomicznej Ludności w ujęciu miesięcznym. Przedstawione przez nich rozwiązanie stanowiło jeden z kluczowych elementów pracy badawczej zrealizowanej na zlecenie Eurostatu, a dotyczyło estymacji miesięcznej stopy bezrobocia (ang. *monthly unemployment rate – MUR*) w Polsce zgodnej z definicją Międzynarodowej Organizacji Pracy (ang. *International Labour Organization – ILO*) oraz wybranych estymatorów pośrednich. Ze względu na to, że BAEL jest badaniem kwartalnym, w którym występuje zjawisko braków danych, należało dokonać odpowiedniej kalibracji wag w ujęciu miesięcznym tak, aby zniwelować ujemny wpływ braków odpowiedzi. W tym celu

wykorzystano tzw. jednokrokowe i dwukrokowe estymatory kalibracyjne wartości globalnej, dzięki czemu możliwa była estymacja stopy bezrobocia w ujęciu miesięcznym dla dodatkowo uwzględnionych domen obejmujących płeć oraz wiek. Autorzy dokonali również oceny jakości precyzji dla wszystkich rozważanych w pracy estymatorów kalibracyjnych.

Ciekawą koncepcję wykorzystania regresji LASSO w sondażach wyborczych w kontekście omawianego w książce podejścia kalibracyjnego przedstawili Chen i współautorzy (2018b). Ze względu na rosnące koszty badań oraz malejące wskaźniki odpowiedzi w badaniu preferencji wyborczych w coraz większym stopniu wykorzystywane są próby nielosowe. Wykorzystanie tego typu danych jest jednak obarczone błędem selekcji. Wynika to z tego, że opinie w serwisach internetowych, na których zbierane są informacje na temat preferencji wyborczych, wyrażane są jedynie przez osoby mające dostęp do internetu, które na ogół mają inne charakterystyki (płeć, wykształcenie itp.) niż osoby, które z tego źródła informacji nie korzystają. Wykorzystując dane zgromadzone na portalu SurveyMonkey, Chen i współautorzy (2018b) podjęli próbę oszacowania preferencji wyborczych w wyborach do Senatu i na gubernatorów w Stanach Zjednoczonych w 2014 roku. W tym celu dokonali korekty odpowiednio utworzonych wag, aby odtwarzane były znane wartości globalne zmiennych pomocniczych (płeć, grupy wieku itd.) w populacji.

Podójście kalibracyjne wykorzystywane jest również w statystyce małych obszarów w nurcie badań określanym mianem mikrosymulacji przestrzennej (ang. *spatial microsimulation*). Mikrosymulacje przestrzenne są szeroko wykorzystywane w badaniach społecznych w celu generowania tzw. syntetycznych populacji, stanowiących podstawę do budowy wielu modeli mikrosymulacyjnych oraz estymacji dla małych obszarów. Roszka, Szymkowiak i Rahman (w druku), wykorzystując metodę GREGWT⁷ (ang. *Generalized Regression and Weighting*), należąca do technik kalibracyjnych i szeroko stosowaną w podejściu mikrosymulacyjnym, dostosowali wybrane brzegowe i łączne rozkłady cech z badania EU-SILC do znanych rozkładów tych cech w podregionach w Polsce, pochodzących z NSP 2011. Dzięki temu zabiegowi możliwe było oszacowanie średniego dochodu do dyspozycji gospodarstw domowych na niepublikowanym do tej pory przez Główny Urząd Statystyczny poziomie agregacji przestrzennej (podregiony).

Rahman (2017), również wykorzystując metodę GREGWT, dokonał oszacowania wskaźnika udziału kosztów lokalowych (ang. *housing stress*) w Australii

⁷ Metoda GREGWT de facto odpowiada estymacji kalibracyjnej dla tzw. uciętej liniowej funkcji kalibracyjnej, która została przedstawiona w rozdziale drugim, por. wzór (2.22).

w przekroju typu własności mieszkania. Opis wybranych aplikacji mikrosymulacji przestrzennej, wykorzystujących technikę GREGWT, z uwzględnieniem zastosowań w obszarze ubóstwa, omawia także w przeglądowym artykule Tanton (2018).

1.7. Wnioski

W niniejszym rozdziale wprowadzono najważniejsze pojęcia z zakresu metody reprezentacyjnej wykorzystywane w kontekście prezentowanej w książce teorii kalibracji i jej zastosowań w badaniach społeczno-ekonomicznych. Omówiono dwa kluczowe estymatory, tj. Horvitza-Thompsona i GREG, wraz z ich własnościami, które w badaniach prowadzonych przez krajowe urzędy statystyczne na całym świecie są wykorzystywane na szeroką skalę. Wskazano, że uogólniony liniowy estymator typu GREG jest szczególnym przypadkiem estymatora kalibracyjnego wartości globalnej dla odpowiednio dobranej funkcji odległości.

Ponadto w kompleksowy sposób przedstawiono rozwój podejścia kalibracyjnego. Uwzględniono przy tym różne aspekty odnoszące się do kalibracji w badaniach z brakami danych, kalibracji w badaniach z wartościami odstającymi, kalibracji zintegrowanej, podejścia funkcyjnego, metody powtarzanego ważenia czy też estymacji kalibracyjnej wspomaganej modelem. Wskazano także obszary, w których ta metoda jest stosowana w praktyce badań statystycznych (rynek pracy, niepełnosprawność, statystyka przedsiębiorstw, ubóstwo). Uwzględniono przy tym zarówno doświadczenia zagraniczne, jak i polskie.

Szerokie spektrum zastosowań podejścia kalibracyjnego, uwzględniające zarówno badania reprezentacyjne, badania pełne, jak i próby nielosowe, wielość idei o podłożu teoretycznym, a także to, że kalibracja jest wciąż młodą techniką korygowania wag, dają podstawy do stwierdzenia, że stanowi ona bardzo ważny obszar badawczy w teorii nowoczesnej estymacji. Rozkwitu tej metody w najbliższym czasie należy się spodziewać przede wszystkim w badaniach wykorzystujących próby nielosowe zasilanych przez alternatywne źródła informacji (big data czy dane z internetu).

Teoretyczne podstawy podejścia kalibracyjnego

2.1. Wprowadzenie

W badaniach reprezentacyjnych prowadzonych przez krajowe urzędy statystyczne przedmiotem estymacji są różnego rodzaju parametry: wartość globalna, średnia czy kwantyle. Spośród wymienionych parametrów populacji generalnej zdecydowanie najczęściej jest to wartość globalna. Dotyczy to również wartości globalnej w odpowiednio zdefiniowanych przekrojach o charakterze przestrzennym (województwo czy powiat), a także wynikających z uwzględnienia dodatkowych cech – najczęściej demograficznych (jak płeć, miejsce zamieszkania, grupy wieku czy wykształcenie). Wynika to z tego, że w publikacjach i różnego rodzaju raportach badawczych najczęściej przedstawia się odpowiednie tabele z liczebnościami. Liczebności te odnoszą się do oszacowanych wartości globalnych z wykorzystaniem odpowiednio skonstruowanych wag kalibracyjnych. Przykłady takich tabel można znaleźć w publikacjach odnoszących się do Badania Budżetów Gospodarstw Domowych (GUS, 2017a), Europejskiego Badania Warunków Życia Ludności (GUS, 2017b) czy Badania Aktywności Ekonomicznej Ludności (GUS, 2018a).

W poprzednim rozdziale wskazano, że kalibracja jest techniką estymacji powszechnie wykorzystywaną w badaniach statystycznych. Zwrócono uwagę na szerokie spektrum zastosowań tej metody w praktyce.

W tym rozdziale podejście kalibracyjne zostało sformalizowane od strony matematycznej. Uwaga została skoncentrowana na konstruowaniu estymatorów kalibracyjnych wartości globalnej, których budowa polega na znalezieniu minimum dla odpowiednio dobranej funkcji odległości, przy jednoczesnej konieczności spełnienia tzw. równań kalibracyjnych⁸. Rozważania dotyczyć będą nie tylko najczęściej wykorzystywanej w praktyce tzw. liniowej funkcji kalibracyjnej, która prowadzi do uzyskania estymatora typu GREG wartości globalnej, ale również innych rzadziej stosowanych funkcji. Omówiono wady i zalety poszczególnych funkcji odległości. Przedstawiono również estymatory wariancji rozważanych estymatorów kalibracyjnych zarówno w populacji generalnej, jak i w ujęciu domen, na które podzielona może być populacja.

W odpowiednio zaprojektowanym badaniu symulacyjnym, z wykorzystaniem rzeczywistych danych pochodzących z Badania Budżetów Gospodarstw Domyowych, dokonano oceny najważniejszych własności estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości. Dyskusji poddano również analizę wag kalibracyjnych, która w tego typu badaniach jest dość często pomijana⁹.

2.2. Estymatory kalibracyjne wartości globalnej – podejście minimalizujące funkcję odległości

Założmy, że celem badania jest oszacowanie wartości globalnej cechy statystycznej Y w populacji U danej wzorem (1.1). Z populacji tej pobieramy n -elementową próbę s zgodnie z określonym planem jej losowania. Zakładamy, że $d_k = \frac{1}{\pi_k}$, $k = 1, \dots, N$, jest odpowiednią wagą (por. wzór (1.21)), gdzie π_k oznacza prawdopodobieństwo inkluzji pierwszego rzędu (por. wzór (1.16)). W praktyce badań

⁸ W książce podejście to jest określane mianem podejścia minimalizującego funkcję odległości. W rozdziale trzecim pokazano, że nie jest to jedyna metoda konstruowania estymatorów kalibracyjnych wartości globalnej.

⁹ Najczęściej dokonuje się oceny oszacowań wartości globalnej, co zdaniem autora jest dużym uproszczeniem i może prowadzić do błędów w estymacji. Pożądane własności estymatora kalibracyjnego wartości globalnej, takie jak małe obciążenie czy wariancja dla jednej zmiennej Y , przy ustalonym wektorze wag, niekoniecznie musi się przełożyć na inne zmienne uwzględnione w badaniu. Z tego punktu widzenia konieczna jest ocena wag kalibracyjnych uwzględniająca występowanie wag ujemnych lub odstających. Jest to szczególnie istotne, gdy estymacji dokonujemy nie tylko na poziomie całej populacji, ale również w dodatkowych przekrojach.

krajowych urzędów statystycznych w procesie estymacji wartości globalnej (1.1) wykorzystuje się bardzo często wspomniany już estymator Horvitz-Thompsona, który wyraża się wzorem (1.35).

Zdarza się jednak, że wagi d_k wynikające z planu losowania próby nie odtwarzają znanych wartości globalnych w odniesieniu do niektórych kluczowych zmiennych. Przyjmijmy, że wektor:

$$\sum_{k \in U} \mathbf{x}_k = \left(\sum_{k \in U} x_{k1}, \dots, \sum_{k \in U} x_{kJ} \right)^T \quad (2.1)$$

jest wektorem wartości globalnych wszystkich zmiennych pomocniczych. Oznacza to, że co najmniej dla jednej zmiennej $j = 1, \dots, J$ nie jest spełniony poniższy warunek:

$$\sum_{k \in s} d_k x_{kj} = \sum_{k \in U} x_{kj}, \quad (2.2)$$

gdzie x_{kj} oznacza wartość j -tej zmiennej pomocniczej dla k -tej jednostki badania oraz $\sum_{k \in U} x_{kj}$ jest wartością globalną tej zmiennej. Wartości globalne takich zmiennych są znane zazwyczaj ze spisów powszechnych czy rejestrów administracyjnych. Ich przykładem mogą być informacje na temat liczby ludności w przekroju płci, klasy miejsca zamieszkania czy odpowiednich grup wieku. Wartości globalne zmiennych pomocniczych są wykorzystywane w procesie kalibracji wag d_k . Po jej zastosowaniu do wszystkich zmiennych pomocniczych nowe wagi – tzw. wagi kalibracyjne w_k – odtwarzają znane ich wartości globalne (2.1) dokładnie. Kalibracja wag jest zatem niezbędna w celu spełnienia często postulowanego w badaniach statystycznych wymogu zgodności. Oznacza to, że oszacowania tych samych zmiennych w różnych badaniach powinny dawać te same rezultaty. Co więcej, można oczekiwać, że jeśli zmienna Y jest skorelowana ze zmiennymi pomocniczymi, to również warunek ten w przybliżeniu powinien być spełniony i dla tej zmiennej.

Poniżej opiszemy proces kalibracji oraz sposób konstrukcji takich wag dla estymatora kalibracyjnego wartości globalnej zmiennej Y , aby dla każdej zmiennej pomocniczej odtwarzane były znane wartości globalne (2.1). Załóżmy w dalszym ciągu, że $\mathbf{d} = (d_1, \dots, d_n)^T$ jest wektorem wag wynikającym z planu losowania próby, a $\mathbf{w} = (w_1, \dots, w_n)^T$ poszukiwanym wektorem końcowych wag kalibracyjnych.

Zgodnie z ideą zaproponowaną w pracy Deville'a i Särndala (1992) estymator kalibracyjny wartości globalnej (1.1) jest postaci:

$$\hat{\tau}_{\text{CAL}} = \sum_{k \in s} w_k y_k, \quad (2.3)$$

gdzie wagi kalibracyjne w_k są rozwiązaniem następującego zadania optymalizacyjnego:

- (W1) – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s} \frac{d_k}{q_k} G\left(\frac{w_k}{d_k}\right) \rightarrow \min, \quad (2.4)$$

- (W2) – równania kalibracyjne:

$$\sum_{k \in s} w_k \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k. \quad (2.5)$$

Pierwszy warunek (W1) orzeka, że wagi kalibracyjne w_k powinny być w taki sposób wyznaczone, aby były możliwie bliskie – w sensie przyjętej funkcji odległości $D(\cdot)$ – wag d_k wynikających z planu losowania próby. Funkcja $G(\cdot)$ mierzy odległość między ilorazem wag w_k/d_k a 1 i jest wykorzystywana w konstruowaniu funkcji odległości $D(\cdot)$, dla której szuka się lokalnego minimum warunkowego. W tym miejscu należy podkreślić, że funkcja $D(\cdot)$ nie jest metryką. Na ogół nie jest bowiem spełniony warunek symetryczności. Co więcej, w wypadku tej funkcji odległości trudno w ogóle mówić o tzw. nierówności trójkąta. Nie bierze się bowiem pod uwagę dodatkowego – trzeciego wektora \mathbf{z} w celu sprawdzenia prawdziwości nierówności $D(\mathbf{d}, \mathbf{w}) \leq D(\mathbf{d}, \mathbf{z}) + D(\mathbf{z}, \mathbf{w})$. Czynniki q_k stanowią dodatkową wagę, która w zależności od jej postaci może prowadzić do uzyskania różnych estymatorów kalibracyjnych. Podobnie jak w wypadku uogólnionego estymatora regresyjnego typu GREG, często przyjmuje się, że $q_k = 1$, a w wypadku jednej zmiennej objaśniającej x_k można przykładowo założyć, że $q_k = 1/x_k$. Z kolei warunek (W2) stanowi istotę teorii kalibracji i orzeka, że wagi powinny być tak dobrane, aby po ich zastosowaniu dla wszystkich zmiennych pomocniczych można było odtworzyć ich znane wartości globalne. Jeśli ten warunek zostanie spełniony, to wykorzystanie wag kalibracyjnych w_k do innych zmiennych Y w badaniu powinno się przyczynić do lepszego oszacowania wartości globalnych tych zmiennych w populacji i do poprawienia precyzji.

Dla niektórych funkcji odległości $D(\cdot)$ może się zdarzyć, że wyznaczone wagi kalibracyjne w_k będą przyjmowały wartości ujemne. Jest to sprzeczne z definicją wagi, która stanowi odwrotność prawdopodobieństwa inkluzji pierwszego rzędu i powinna przyjmować wartość $w_k \geq 1$ dla każdego k . Może się również zdarzyć sytuacja, że wagi są dodatnie i większe od 1, ale przyjmują wartości ekstre-

malne, tj. znacznie odbiegają od wag d_k wynikających z planu losowania próby. Na ogół obserwuje się to, gdy dla danego przekroju w próbie istnieje niewielka liczba jednostek, podczas gdy w odpowiadającym mu przekroju w badaniu pełnym (na przykład spisie) liczba jednostek jest znacznie większa. Taki przypadek jest zatem szczególnie niebezpieczny, gdy estymatory kalibracyjne wykorzystywane są do szacowania wartości globalnych w domenach. W wypadku małej liczby jednostek w próbie reprezentujących daną domenę oszacowana wartość globalna ze względu na znaczne zniekształcenie wagi oryginalnej wpływa na istotne przeszacowanie nieznannej wartości prawdziwej. Z tego powodu w procesie wyznaczania wag kalibracyjnych w niektórych przypadkach wprowadza się dodatkowy warunek ograniczający na wagi, a dokładniej na iloraz wag w_k i d_k :

- (W3) – warunki ograniczające:

$$L \leq \frac{w_k}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, n. \quad (2.6)$$

Przy wyborze funkcji $G(\cdot)$ w procesie wyznaczania wag kalibracyjnych w_k istnieje pewna dowolność. Funkcja $G(\cdot)$ powinna jednak spełniać między innymi następujące własności matematyczne: $G(\cdot)$ jest ściśle wypukła i dwukrotnie różniczkowalna, $G(\cdot) \geq 0$, $G(1) = 0$, $G'(1) = 0$ oraz $G''(1) = 1$.

Do wyznaczania wag kalibracyjnych w_k szczególnie przydatna jest znajomość funkcji $F(\cdot)$ odwrotnej do pierwszej pochodnej funkcji $G(\cdot)$, tj.:

$$F(\cdot) = G'^{-1}(\cdot), \quad (2.7)$$

przy czym $F(0) = 1$. Jest to tzw. funkcja kalibracyjna (ang. *calibration function*). Wynika to ze sposobu poszukiwania rozwiązań zadania minimalizacji opisanego w warunku (W1), przy jednoczesnym spełnieniu równań kalibracyjnych (W2). Rozwiązania poszukać można bowiem, wykorzystując metodę czynników nieoznaczonych Lagrange'a służącą do znajdowania ekstremum warunkowego funkcji różniczkowalnej. Odpowiednia funkcja Lagrange'a wyraża się wzorem:

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}; \mathbf{d}) = \sum_{k \in S} \frac{d_k}{q_k} G\left(\frac{w_k}{d_k}\right) - \boldsymbol{\lambda}^T \left(\sum_{k \in S} w_k \mathbf{x}_k - \sum_{k \in U} \mathbf{x}_k \right), \quad (2.8)$$

gdzie $\boldsymbol{\lambda}^T = (\lambda_1, \dots, \lambda_J)^T$ jest tzw. wektorem mnożników (czynników nieoznaczonych) Lagrange'a. Równania Lagrange'a dla $k = 1, \dots, n$ wyrażają się zatem wzorem:

$$\frac{\partial \mathcal{L}}{\partial w_k} = \frac{d_k}{q_k} \cdot \frac{\partial}{\partial w_k} G\left(\frac{w_k}{d_k}\right) - \boldsymbol{\lambda}^T \mathbf{x}_k = 0. \quad (2.9)$$

Z powyższego równania i z twierdzenia o pochodnej funkcji złożonej wynika, że:

$$\frac{1}{q_k} G' \left(\frac{w_k}{d_k} \right) - \boldsymbol{\lambda}^T \mathbf{x}_k = 0. \quad (2.10)$$

Przyjmując, że $G'(\cdot) = \psi(\cdot)$ oraz że $F(\cdot) = \psi^{-1}(\cdot)$ jest funkcją odwrotną względem pierwszej pochodnej funkcji $G(\cdot)$, otrzymujemy, że:

$$\psi \left(\frac{w_k}{d_k} \right) = q_k \boldsymbol{\lambda}^T \mathbf{x}_k = q_k \mathbf{x}_k^T \boldsymbol{\lambda} \Rightarrow \frac{w_k}{d_k} = \psi^{-1} \left(q_k \mathbf{x}_k^T \boldsymbol{\lambda} \right) = F \left(q_k \mathbf{x}_k^T \boldsymbol{\lambda} \right). \quad (2.11)$$

Ostatecznie wagi kalibracyjne można przedstawić w następującej postaci:

$$w_k = d_k F \left(q_k \mathbf{x}_k^T \boldsymbol{\lambda} \right) = d_k g_k, \quad (2.12)$$

gdzie $g_k = F(q_k \mathbf{x}_k^T \boldsymbol{\lambda})$ to tzw. mnożniki wagowe (kalibracyjne). Do wyznaczenia wag kalibracyjnych (2.12) w dalszym ciągu niezbędna jest znajomość wektora mnożników Lagrange'a $\boldsymbol{\lambda}$, który można znaleźć, rozwiązując odpowiednie równanie kalibracyjne – por. równanie (2.5):

$$\sum_{k \in s} d_k F \left(q_k \mathbf{x}_k^T \boldsymbol{\lambda} \right) \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k. \quad (2.13)$$

Finalnie estymator kalibracyjny wartości globalnej można przedstawić w postaci:

$$\hat{\tau}_{\text{CAL}} = \sum_{k \in s} d_k F \left(q_k \mathbf{x}_k^T \boldsymbol{\lambda} \right) y_k = \sum_{k \in s} w_k y_k. \quad (2.14)$$

Wyznaczenie wektora $\boldsymbol{\lambda}$ czynników nieoznaczonych Lagrange'a, a w konsekwencji wag kalibracyjnych (2.12), zależy od postaci funkcji $G(\cdot)$. W wypadku niektórych funkcji istnieje możliwość jawnego przedstawienia wag kalibracyjnych w postaci odpowiedniego wzoru. W innych przypadkach wymagane jest z kolei zastosowanie podejścia iteracyjnego. Przegląd wybranych funkcji $G(\cdot)$ można znaleźć m.in. w pracach Deville'a i Särndala (1992) czy Szymkowiaka (2013)¹⁰. Szczególnie ważna jest funkcja $G(x) = \frac{1}{2}(x-1)^2$, dla której wektor wag kalibracyjnych można wyrazić za pomocą jawnego wzoru. Mówi o tym poniższe twierdzenie.

Twierdzenie 2.1 (Deville i Särndal, 1992). *Rozwiązaniem zadania minimalizacji (2.4) przy warunku (2.5) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag*

¹⁰ Przedstawiono je również wraz z odpowiadającymi im funkcjami kalibracyjnymi $F(\cdot)$ w tabeli 2.1.

kalibracyjnych $\mathbf{w} = (w_1, \dots, w_n)^T$, którego składowe, dla $k = 1, \dots, n$, wyrażają się wzorem:

$$w_k^{\text{lin}} = d_k + d_k q_k \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in s} q_k d_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \mathbf{x}_k. \quad (2.15)$$

Dowód. Dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jej pochodna $G'(x) = \psi(x) = x-1$, a funkcja $F(u) = u+1$. Korzystając z równania (2.13), otrzymujemy:

$$\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k = \sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \boldsymbol{\lambda}, \quad (2.16)$$

a stąd postać wektora $\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda} = \left(\sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right), \quad (2.17)$$

który po podstawieniu do wzoru (2.12) i prostym przekształceniu daje pożądaną postać wag kalibracyjnych (2.15). \square

W literaturze przedmiotu metoda ta jest określana jako liniowa (ang. *linear method*) z racji tego, że odpowiednia funkcja kalibracyjna $F(\cdot)$ ma postać funkcji liniowej¹¹. Warto podkreślić, że dla tak określonej funkcji $G(\cdot)$, odpowiadająca jej funkcja odległości $D(\cdot)$ wyraża się wzorem:

$$D(\mathbf{d}, \mathbf{w}) = \frac{1}{2} \sum_{k \in s} \frac{(w_k - d_k)^2}{d_k q_k} \quad (2.18)$$

i ze względu na podobieństwo do statystyki testowej w teście χ^2 zgodności czy niezależności określana jest mianem funkcji odległości χ^2 . Minimalizację funkcji podobnej postaci rozważali, jak wcześniej opisano, Deming i Stephan (1940) – por. wzór (1.69).

Wagi kalibracyjne (2.15) są bardzo często wykorzystywane w praktyce badań statystycznych. Zostały użyte m.in. przy okazji Narodowego Spisu Powszechnego Ludności i Mieszkań w 2011 roku (Szymkowiak, 2014)¹². Należy mieć jednak na uwadze, że dla niektórych jednostek wylosowanych do próby, mogą one

¹¹ Z tego powodu we wzorze (2.15) na wagi kalibracyjne użyto subskrypty „lin” dla podkreślenia tego faktu.

¹² Sposób wyznaczania wag kalibracyjnych w spisie został szczegółowo omówiony w dalszej części tej książki.

przyjmować wartości ujemne. Jest tak zazwyczaj, gdy w procesie kalibracji wykorzystywanych jest wiele zmiennych pomocniczych, przez co spełnionych musi być odpowiednio dużo równań kalibracyjnych. Mimo że wpływ wag ujemnych na proces estymacji może być niewielki, ich istnienie jest zwykle nieakceptowalne ze względu na trudną w takiej sytuacji ich interpretację. Warto również podkreślić, że estymator kalibracyjny, dla którego wagi są wyznaczone z wykorzystaniem wspomnianej powyżej funkcji odległości χ^2 , jest tożsamy z omówionym w podrozdziale (1.4) liniowym uogólnionym estymatorem regresyjnym GREG.

Inną często wykorzystywaną w praktyce funkcją $G(\cdot)$ służącą do wyznaczania wag kalibracyjnych jest:

$$G(x) = x \ln x - x + 1. \quad (2.19)$$

Odpowiadająca jej funkcja odległości $D(\cdot)$ jest wspomnianą wcześniej tzw. informacyjną funkcją odległości Kullbacka-Leiblera (dyskryminacyjną lub względnej entropii):

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in S} \frac{d_k}{q_k} \left[\frac{w_k}{d_k} \ln \frac{w_k}{d_k} - \frac{w_k}{d_k} + 1 \right]. \quad (2.20)$$

Dla tak zdefiniowanej funkcji odległości skonstruowanej na bazie funkcji (2.19) mamy $\psi(x) = \ln x$ oraz $F(u) = e^u$. Odpowiednie wagi kalibracyjne wyrażają się wzorem:

$$w_k = d_k e^{q_k \mathbf{x}_k^T \boldsymbol{\lambda}}. \quad (2.21)$$

Zaletą tak wyznaczonych wag kalibracyjnych jest to, że są one zawsze dodatnie. Wadą jest natomiast to, że dla niektórych jednostek wylosowanych do próby s wagi mogą przyjmować wartości ekstremalne, a odpowiadający im estymator kalibracyjny może dawać zawyżone oszacowania wartości globalnej (1.1), zwłaszcza w sytuacji, gdy zmienne pomocnicze są słabo skorelowane ze zmienną Y . Również końcowe rozwiązanie (2.13) może nie istnieć, a poszukiwanie wag kalibracyjnych ze względu na nieliniowy układ równań kalibracyjnych wymaga zastosowania metod numerycznych. Ze względu na postać funkcji $F(\cdot)$ metodę tę określa się mianem wykładniczej (ang. *exponential method*) lub rakingu (ang. *raking method*). Warto nadmienić, że to podejście stosowane jest w praktyce stosunkowo często. Na przykład, Shook-Sa i współautorzy (2017) wykorzystali metodę wykładniczą w badaniu nad przestępczością w celu wyznaczenia wag kalibracyjnych dla wybranych stanów oraz metropolitalnych obszarów statystycznych w Stanach Zjednoczonych.

Ponieważ w wypadku niektórych funkcji $G(\cdot)$ odpowiadające im funkcje odległości $D(\cdot)$ mogą generować wagi ujemne bądź ekstremalne, wprowadza się często pewne ograniczenia na wagi kalibracyjne w_k . Mówi o tym wspomniany wcześniej warunek (W3). Wynika z niego, że $d_k L \leq w_k \leq d_k U$. Przykładami metod, które uwzględniają ten dodatkowy warunek, mogą być: tzw. ucięta metoda liniowa (ang. *truncated linear method*), metoda logitowa (ang. *logit method*) bądź metoda mająca szczególne znaczenie w pracach Eurostatu i bazująca na funkcji odległości w wersji sinusa hiperbolicznego (metoda sinusa hiperbolicznego). W uciętej metodzie liniowej funkcja $G(\cdot)$ dla argumentu $x = w_k/d_k$ zdefiniowana jest w następujący sposób:

$$G\left(\frac{w_k}{d_k}\right) = \begin{cases} \frac{1}{2}\left(\frac{w_k}{d_k} - 1\right)^2, & \text{jeżeli } L < \frac{w_k}{d_k} < U, \\ \infty & \text{w przeciwnym razie,} \end{cases} \quad (2.22)$$

dla pewnych stałych L i U stanowiących odpowiednio dolne i górne ograniczenie ilorazu wag kalibracyjnych w_k i wag d_k podlegających procesowi kalibracji¹³. Odpowiednia funkcja kalibracyjna $F(\cdot)$ dla tak zdefiniowanej funkcji $G(\cdot)$ wyraża się wzorem:

$$F\left(q_k \mathbf{x}_k^T \boldsymbol{\lambda}\right) = \begin{cases} 1 + q_k \mathbf{x}_k^T \boldsymbol{\lambda}, & \text{jeżeli } L - 1 \leq q_k \mathbf{x}_k^T \boldsymbol{\lambda} \leq U - 1, \\ U, & \text{jeżeli } q_k \mathbf{x}_k^T \boldsymbol{\lambda} > U - 1, \\ L, & \text{jeżeli } q_k \mathbf{x}_k^T \boldsymbol{\lambda} < L - 1. \end{cases} \quad (2.23)$$

Z kolei w metodzie logitowej funkcja $G(\cdot)$ dla argumentu $x = w_k/d_k$ zdefiniowana jest w następujący sposób:

$$G\left(\frac{w_k}{d_k}\right) = \begin{cases} \left[\left(\frac{w_k}{d_k} - L \right) \ln \frac{\frac{w_k}{d_k} - L}{1-L} + \left(U - \frac{w_k}{d_k} \right) \ln \frac{U - \frac{w_k}{d_k}}{U-1} \right] A^{-1}, & \text{jeżeli } L < \frac{w_k}{d_k} < U, \\ \left[(U - L) \ln \frac{U-L}{U-1} \right] A^{-1}, & \text{jeżeli } \frac{w_k}{d_k} \leq L, \\ \left[(U - L) \ln \frac{U-L}{1-L} \right] A^{-1}, & \text{jeżeli } \frac{w_k}{d_k} \geq U, \end{cases} \quad (2.24)$$

gdzie $A = \frac{U-L}{(U-1)(1-L)}$. Odpowiednia funkcja kalibracyjna $F(\cdot)$ dla tak zdefiniowanej funkcji $G(\cdot)$ wyraża się wzorem:

$$F\left(q_k \mathbf{x}_k^T \boldsymbol{\lambda}\right) = \frac{L(U-1) + U(1-L)e^{Aq_k \mathbf{x}_k^T \boldsymbol{\lambda}}}{(U-1) + (1-L)e^{Aq_k \mathbf{x}_k^T \boldsymbol{\lambda}}}. \quad (2.25)$$

¹³ Jak wspomniano w rozdziale pierwszym, w mikrosymulacjach przestrzennych, stanowiących nowy przedmiot badań w statystyce małych obszarów metoda ta jest znana pod nazwą GREGWT.

Szczególnie ważna w praktyce krajowych urzędów statystycznych jest rekomendowana przez Eurostat w EU-SILC funkcja $G(\cdot)$ w wersji sinusa hiperbolicznego:

$$G(x) = \frac{1}{2\alpha} \int_1^x \sinh \left[\alpha \left(t - \frac{1}{t} \right) \right] dt, \quad (2.26)$$

gdzie α jest pewnym dodatnim parametrem, a $\sinh(x) = \frac{e^x - e^{-x}}{2}$ jest sinusem hiperbolicznym. Również w wypadku tej funkcji można nakładać na iloraz wag w_k/d_k ograniczenia w celu wyeliminowania odstających wartości wag kalibracyjnych.

Mimo że jest ona raczej rzadko wykorzystywaną funkcją w porównaniu z pozostałymi, w dalszej części zarówno w podejściu symulacyjnym, jak i w praktycznych zastosowaniach kalibracji w wybranych badaniach statystycznych (rozdział piąty i szósty) pokazano, że ma ona cenne zalety. W związku z powyższym przedstawimy teraz sposób konstruowania wag w_k dla funkcji (2.26).

Twierdzenie 2.2. *Wagi kalibracyjne w_k dla funkcji $G(\cdot)$ określonej wzorem (2.26) można wyrazić wzorem:*

$$w_k = d_k \frac{\operatorname{arsinh}(2\alpha q_k \mathbf{x}_k^T \boldsymbol{\lambda}) + \sqrt{\operatorname{arsinh}^2(2\alpha q_k \mathbf{x}_k^T \boldsymbol{\lambda}) + 4\alpha^2}}{2\alpha}. \quad (2.27)$$

Dowód. Ponieważ $\sinh \left[\alpha \left(1 - \frac{1}{1} \right) \right] = 0$, pochodna funkcji $G(x)$ jest postaci:

$$\psi(x) = \frac{1}{2\alpha} \sinh \left[\alpha \left(\frac{x^2 - 1}{x} \right) \right]. \quad (2.28)$$

W kolejnym kroku znajdziemy funkcję odwrotną $\psi^{-1}(x)$ do $\psi(x)$. Ponieważ funkcją odwrotną do funkcji $\sinh(x)$ jest $\operatorname{arsinh}(x) = \ln(x + \sqrt{x^2 + 1})$, należy rozwiązać poniższe równanie względem zmiennej x :

$$2\alpha u = \sinh \left[\alpha \left(\frac{x^2 - 1}{x} \right) \right]. \quad (2.29)$$

Uwzględniając znaną postać funkcji odwrotnej do $\sinh(x)$, powyższe równanie po prostych przekształceniach możemy zapisać w postaci równoważnego równania:

$$x^2 - \frac{\operatorname{arsinh}(2\alpha u)}{\alpha} x - 1 = 0. \quad (2.30)$$

Jest to równanie kwadratowe, dla którego wyróżnik Δ jest postaci:

$$\Delta = \frac{\operatorname{arsinh}^2(2\alpha u)}{\alpha^2} + 4. \quad (2.31)$$

Pierwiastki tego równania są zatem postaci:

$$x_{1,2} = \frac{\frac{\operatorname{arsinh}(2\alpha u)}{\alpha} \pm \sqrt{\frac{\operatorname{arsinh}^2(2\alpha u)}{\alpha^2} + 4}}{2}. \quad (2.32)$$

Biorąc pod uwagę pierwiastek x_1 i dokonując prostych przekształceń, otrzymujemy ostatecznie, że funkcja kalibracyjna $F(u)$ wyraża się wzorem:

$$F(u) = \frac{\operatorname{arsinh}(2\alpha u) + \sqrt{\operatorname{arsinh}^2(2\alpha u) + 4\alpha^2}}{2\alpha}. \quad (2.33)$$

Funkcję tę po odpowiednich przekształceniach można przedstawić jako:

$$F(u) = \frac{\frac{1}{\alpha} \ln(2\alpha u + \sqrt{4\alpha^2 u^2 + 1}) + \sqrt{\frac{1}{\alpha^2} \ln^2(2\alpha u + \sqrt{4\alpha^2 u^2 + 1}) + 4}}{2}. \quad (2.34)$$

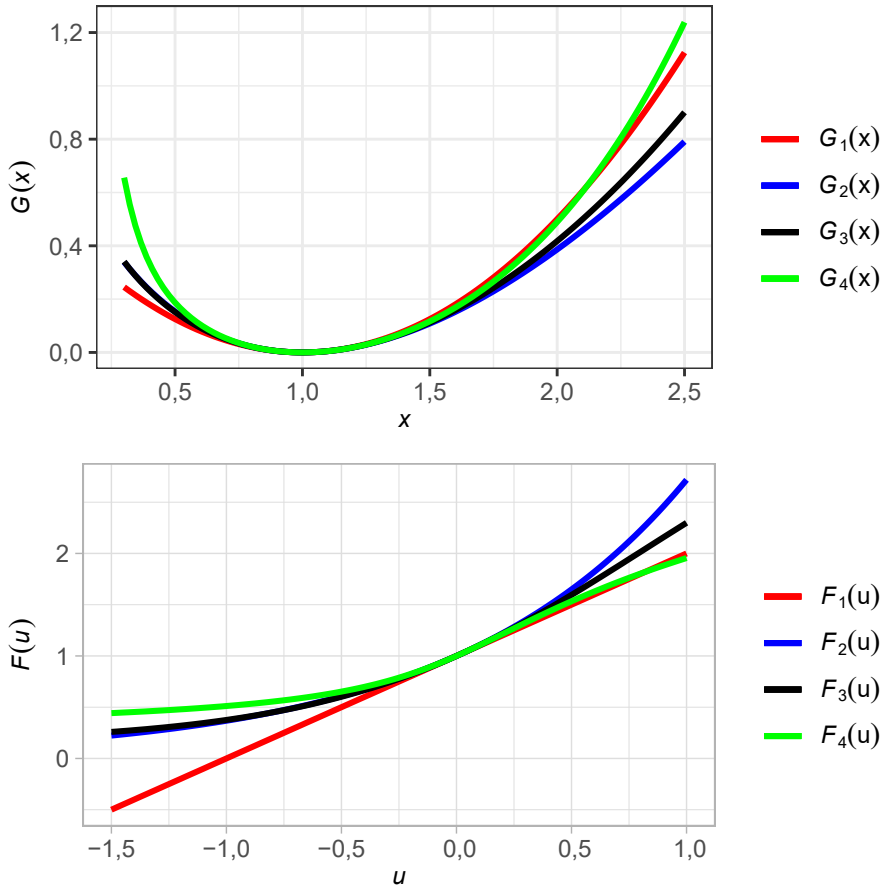
W ostatnim kroku podstawiamy do wzoru (2.12) funkcję (2.33), przyjmując $u = q_k \mathbf{x}_k^T \boldsymbol{\lambda}$. Otrzymujemy postać wag kalibracyjnych (2.27). \square

Na rysunku 2.1 przedstawiono wykresy najczęściej spotykanych w praktyce funkcji $G(\cdot)$ wraz z odpowiadającymi im funkcjami kalibracyjnymi $F(\cdot)$. Opisano je również w tabeli 2.1¹⁴.

Z wyjątkiem funkcji $G_1(\cdot)$, odnoszącej się do funkcji kwadratowej (por. twierdzenie 2.1), kluczową kwestią jest wyznaczenie wektora $\boldsymbol{\lambda}$ mnożników Lagrange'a. Zagadnieniu temu poświęcony będzie następny podrozdział.

Niewątpliwą zaletą niektórych z powyższych metod jest to, że wagi kalibracyjne w_k można wyspecyfikować tak, aby znajdowały się w odpowiednio dobranym przedziale. Jak podają Haziza i Beaumont (2017), w praktyce badań statystycznych przedziały $[L, U]$ mogą być postaci $[\frac{1}{2}, 2]$, $[\frac{1}{3}, 3]$ czy $[\frac{1}{4}, 4]$. Przeciwdziała to występowaniu wag ujemnych oraz ekstremalnych. Z drugiej strony narzucanie ograniczeń na wagi kalibracyjne powoduje, że maleje możliwość znalezienia rozwiązania odpowiednich równań kalibracyjnych. Im krótsze będą

¹⁴ Na rysunku 2.1 numeracja funkcji odpowiada ich kolejności w tabeli 2.1. Zrezygnowano z wykresu funkcji $G_5(x)$ oraz $F_5(u)$ (ucięta metoda liniowa) ze względu na ich podobieństwo do funkcji $G_1(x)$ oraz $F_1(u)$ (różnica dotyczy jedynie ograniczeń narzuczanych na iloraz wag).



Rysunek 2.1. Wykresy funkcji $G(x)$ i $F(u)$ w podejściu kalibracyjnym

Źródło: na podstawie przykładowych danych.

zdefiniowane przedziały, tym mniejsza jest szansa, że istnieje odpowiednie rozwiązanie. Ponadto, gdy przedziały będą zbyt krótkie, a odpowiednie rozwiązanie istnieje, wiele wag kalibracyjnych wykazywać będzie tendencję do skupiania się blisko wartości brzegowych wyznaczonych przedziałów. Warto jednak podkreślić, że gdy liczebność próby n jest stosunkowo duża, istnieje większa możliwość konstruowania bardziej restrykcyjnych przedziałów. Należy również każdorazowo dokonywać oceny rozkładu ilorazu wag w_k/d_k w celu zidentyfikowania właściwego przedziału $[L, U]$. Można również, jak proponują Folsom i Singh (2000), rozszerzyć konstruowanie przedziałów $[L, U]$ dla różnych podzbiorów wag.

Tabela 2.1. Zestawienie funkcji $G(x)$, $F(u)$ oraz ograniczeń na wagi w podejściu kalibracyjnym

Lp.	$G(x)$	$F(u)$	Ograniczenia
1	$\frac{(x-1)^2}{2}$, dla $x \in \mathbb{R}$	$1 + u$, dla $u \in \mathbb{R}$	brak
2	$x \ln(x) - x + 1$, dla $x \in \mathbb{R}^+$ $-x + 1$, dla $x = 0$	e^u , dla $u \in \mathbb{R}$	brak
3	$\left[\frac{(x-L) \ln \frac{x-L}{1-L} + (U-x) \ln \frac{U-x}{U-1} \right] A^{-1}$, dla $x \in (L, U)$ $\left[\frac{(U-L) \ln \frac{U-L}{U-1} \right] A^{-1}$, dla $x \leq L$ $\left[\frac{(U-L) \ln \frac{U-L}{1-L} \right] A^{-1}$, dla $x \geq U$	$\frac{L(U-1) + U(1-L)e^{Au}}{(U-1) + (1-L)e^{Au}}$, gdzie $u \in \mathbb{R}$ oraz $A = \frac{U-L}{(U-1)(1-L)}$	$0 \leq L < 1 < U$
4	$\frac{1}{2\alpha} \int_1^x \sinh \left[\alpha \left(t - \frac{1}{t} \right) \right] dt$, dla $\alpha > 0$	$\frac{\operatorname{arsinh}(2\alpha u) + \sqrt{\operatorname{arsinh}^2(2\alpha u) + 4\alpha^2}}{2\alpha}$, dla $u \in \mathbb{R}$	$0 \leq L < 1 < U$
5	$\frac{(x-1)^2}{2}$, dla $x \in \mathbb{R}$	$1 + u$, dla $u \in [L-1, U-1]$ L , dla $u = L-1$ U , dla $u = U-1$	$x \in [L, U]$ oraz $0 \leq L \leq 1 \leq U$

Źródło: na podstawie Deville i Särndal (1992) oraz Vanderhoef (2001).

W praktyce postuluje się w pierwszej kolejności, aby wagi kalibracyjne w_k wyznaczać z wykorzystaniem metody liniowej. Wynika to z tego, że w takim wypadku nie jest wymagane nakładanie na odpowiedni iloraz wag żadnych ograniczeń. Co więcej, wagi takie można wyznaczyć wprost ze wzoru, o czym mówi twierdzenie 2.1. W sytuacji gdy zaobserwuje się występowanie wag ujemnych, warto rozważyć z kolei podejście wykładnicze, które generuje tylko dodatnie wagi kalibracyjne. Jeżeli w takim wypadku stwierdzi się jednak istnienie dodatnich wag ekstremalnych, należy rozważyć nałożenie odpowiednich ograniczeń na iloraz wag w_k i d_k . Można w takiej sytuacji przyjąć dowolną funkcję $G(\cdot)$. Przy określaniu wartości L i U warto zacząć od szerszych przedziałów ograniczających i stopniowo je zawężać. Należy przy tym sprawdzać, czy odtwarzane są odpowiednie równania kalibracyjne.

Dla zadanego wektora \mathbf{x}_k zmiennych pomocniczych różne funkcje kalibracyjne $F(\cdot)$ prowadzą na ogół do nieco odmiennych oszacowań wartości globalnej (1.1) z wykorzystaniem estymatora kalibracyjnego (2.3). Mówi o tym twierdzenie 2.3. Twierdzenie to pokazuje, że oszacowania uzyskane dla estymatora kalibracyjnego uzyskanego z wykorzystaniem metody liniowej $\hat{\tau}_{\text{GREG}}^{\text{lin}}$ różnią się od oszacowań uzyskanych dla estymatorów kalibracyjnych $\hat{\tau}_{\text{CAL}}$ bazujących na innych funkcjach odległości o pewien czynnik.

Twierdzenie 2.3 (Estevao i Särndal, 2000). *Dla każdej próby s zachodzi poniższa relacja między $\hat{\tau}_{\text{GREG}}^{\text{lin}}$ oraz $\hat{\tau}_{\text{CAL}}$:*

$$\hat{\tau}_{\text{CAL}} = \hat{\tau}_{\text{GREG}}^{\text{lin}} + Z_s, \quad (2.35)$$

gdzie:

$$Z_s = \sum_{k \in s} (w_k - d_k) \hat{e}_k, \quad (2.36)$$

przy czym $\hat{e}_k = y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_s$ oraz $\hat{\boldsymbol{\beta}}_s$ wyraża się wzorem (1.42).

Można jednak pokazać, że dla niektórych – odpowiednio dobranych postaci wektora zmiennych pomocniczych \mathbf{x}_k , bez względu na przyjętą funkcję kalibracyjną $F(\cdot)$, uzyskujemy tę samą postać estymatora kalibracyjnego wartości globalnej. Poniżej rozważymy niektóre z takich przypadków.

Często w procesie wyznaczania wag kalibracyjnych wymaga się, aby sumowały się one do znanej liczebności N populacji U , tj. aby $\sum_{k \in s} w_k = N$. W tym celu wystarczy przyjąć, że $\mathbf{x}_k = q_k = 1$ dla wszystkich jednostek z populacji U .

W takiej sytuacji spełniony jest warunek $\sum_{k \in U} x_k = N$, a wagi kalibracyjne, o których mowa we wzorze (2.12), redukują się do:

$$w_k = d_k \frac{N}{\sum_{k \in s} d_k}. \quad (2.37)$$

Dla tak wyznaczonych wag estymator kalibracyjny (2.14) sprowadza się do tzw. estymatora Hájka (1971) postaci:

$$\hat{\tau}_{\text{HAJEK}} = N \frac{\sum_{k \in s} d_k y_k}{\sum_{k \in s} d_k}. \quad (2.38)$$

W wypadku jednej zmiennej pomocniczej x_k , dla której spełniony jest warunek $q_k = x_k$ oraz dla której znana jest jej wartość globalna $\sum_{k \in U} x_k$, wagi kalibracyjne wyrażają się wzorem:

$$w_k = d_k \frac{\sum_{k \in U} x_k}{\sum_{k \in s} d_k x_k}, \quad (2.39)$$

a estymator kalibracyjny wartości globalnej redukuje się do tzw. estymatora ilorazowego (ang. *ratio estimator* – RATIO):

$$\hat{\tau}_{\text{RATIO}} = \sum_{k \in U} x_k \frac{\sum_{k \in s} d_k y_k}{\sum_{k \in s} d_k x_k}. \quad (2.40)$$

Szczególnym przypadkiem estymatora kalibracyjnego jest stosowany w praktyce badań statystycznych tzw. estymator postratyfikacyjny. Załóżmy, że dysponujemy pewną zmienną jakościową (płeć, grupa wieku, klasa miejsca zamieszkania itd.), która ma P wariantów. Załóżmy ponadto, że $\mathbf{N} = (N_1, \dots, N_p, \dots, N_P)^T$ oznacza wektor zawierający liczebności populacyjne, tj. N_p oznacza liczbę jednostek populacji U , która należy do p -tej kategorii. Niech ponadto $\mathbf{x}_k^P = (x_{1k}, \dots, x_{Pk})^T$ będzie wektorem przynależącym do jednostki k , którego składowe spełniają warunek $x_{pk} = 1$, gdy jednostka k należy do p -tej kategorii rozpatrywanej zmiennej jakościowej, oraz $x_{pk} = 0$ w przeciwnym razie. Przyjmując za q_k we wzorze (2.12) q_p , jeśli jednostka k należy do p -tej kategorii, otrzymujemy tzw. wagi postratyfikacyjne:

$$w_k = d_k \frac{N_p}{\hat{N}_p}, \quad (2.41)$$

gdzie $\hat{N}_p = \sum_{k \in s \cap U_p} d_k$ oznacza oszacowaną liczbę jednostek mających p -ty wariant analizowanej zmiennej jakościowej, a U_p oznacza populację jednostek, które

mają p -ty wariant tej cechy. Ostatecznie estymator kalibracyjny (2.14) redukuje się do tzw. estymatora postratyfikacyjnego:

$$\hat{\tau}_{\text{PS}} = \sum_{p=1}^P \frac{N_p}{\hat{N}_p} \sum_{k \in s \cap U_p} d_k y_k. \quad (2.42)$$

W dalszej części uwaga skupiona będzie na wybranych miarach precyzji estymatorów kalibracyjnych. Dla dowolnej funkcji kalibracyjnej $F(\cdot)$, którą wykorzystuje się na potrzeby konstrukcji wag kalibracyjnych (2.12), wariancja estymatora kalibracyjnego (2.14) może być przedstawiona w postaci (Deville i Särndal, 1992, s. 379):

$$D^2(\hat{\tau}_{\text{CAL}}) = \sum_{k \in U} \sum_{l \in U} \Delta_{kl} (d_k e_k)(d_l e_l), \quad (2.43)$$

gdzie $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$, $e_k = y_k - \mathbf{x}_k^T \boldsymbol{\beta}_U$, przy czym $\boldsymbol{\beta}_U$ wyraża się wzorem (1.54).

Do oceny wariancji estymatora kalibracyjnego (2.43) można wykorzystać estymator postaci (Rao i Molina, 2015, s. 12):

$$\hat{D}_{(1)}^2(\hat{\tau}_{\text{CAL}}) = - \sum_{k < l} \sum_{k, l \in s} \frac{\Delta_{kl}}{\pi_{kl}} (d_k y_k - d_l y_l)^2, \quad (2.44)$$

który w takim wypadku redukuje się do estymatora wariancji wartości globalnej w populacji Sena-Yatesa-Grundy'ego (Cochran, 1977). Wartości y_k można również zastąpić przez reszty $\hat{e}_k = y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_s$, gdzie $\hat{\boldsymbol{\beta}}_s$ wyraża się wzorem (1.42), i wówczas do oceny wariancji (2.43) można wykorzystać estymator:

$$\hat{D}_{(2)}^2(\hat{\tau}_{\text{CAL}}) = - \sum_{k < l} \sum_{k, l \in s} \frac{\Delta_{kl}}{\pi_{kl}} (d_k \hat{e}_k - d_l \hat{e}_l)^2. \quad (2.45)$$

Ponieważ, jak pokazują w badaniach symulacyjnych Estevao i współautorzy (1995), estymator (2.45) może niedoszacowywać wariancji (2.43) estymatora kalibracyjnego, zaleca się do oceny wariancji estymator postaci (Rao i Molina, 2015, s. 15):

$$\hat{D}_{(3)}^2(\hat{\tau}_{\text{CAL}}) = - \sum_{k < l} \sum_{k, l \in s} \frac{\Delta_{kl}}{\pi_{kl}} (d_k g_k \hat{e}_k - d_l g_l \hat{e}_l)^2, \quad (2.46)$$

gdzie g_k wyrażają się wzorem (1.46).

Do oceny wariancji (2.43) można również wykorzystać estymator (Deville i Särndal, 1992, s. 380):

$$\hat{D}_{(4)}^2(\hat{\tau}_{\text{CAL}}) = \sum_{k \in s} \sum_{l \in s} (w_k \tilde{e}_k)(w_l \tilde{e}_l) \frac{\Delta_{kl}}{\pi_{kl}}, \quad (2.47)$$

przy czym $\tilde{e}_k = y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_{ws}$ oraz

$$\hat{\boldsymbol{\beta}}_{ws} = \left(\sum_{k \in s} q_k w_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in s} q_k w_k \mathbf{x}_k y_k. \quad (2.48)$$

Podobnie jak w wypadku uogólnionego liniowego estymatora typu GREG, który jest szczególnym przypadkiem estymatora kalibracyjnego wartości globalnej, również w odniesieniu do estymatorów kalibracyjnych można je wykorzystać do oszacowania wartości globalnej zmiennej Y w domenach. Zakładamy zatem, że populacja U składa się z D rozłącznych domen postaci U_1, \dots, U_D . Przyjmujemy ponadto, że w każdej z D wyróżnionych domen przedmiotem zainteresowania jest wartość globalna $\tau_{Y_1}, \dots, \tau_{Y_D}$, gdzie τ_{Y_d} , dla $d = 1, \dots, D$, wyraża się wzorem (1.58).

Funkcję przynależności δ_{dk} jednostki k do domeny d definiujemy wzorem (1.59), a wartość cechy Y dla jednostki k należącej do domeny d wzorem (1.60). Wartość globalną τ_{Y_d} w domenie d z kolei wyrazić można wzorem (1.61). Estymator kalibracyjny wartości globalnej (1.61) w domenie d może być zatem zapisany w postaci:

$$\hat{\tau}_{\text{CAL},d}^{(1)} = \sum_{k \in s} w_k y_{dk}, \quad (2.49)$$

gdzie wagi kalibracyjne w_k wyrażają się wzorem (2.12). Przy szacowaniu wartości globalnej τ_{Y_d} w domenie d uwzględnia się zatem wyłącznie wagi kalibracyjne odpowiadające jednostkom należącym do tej domeny, tj. $k \in s_d$, gdzie $s_d = s \cap U_d$. Estymator kalibracyjny wartości globalnej (1.61) w domenie d można również przedstawić w alternatywnej postaci jako $\hat{\tau}_{\text{CAL},d}^{(1)} = \sum_{k \in s_d} w_k y_k$. Estymator ten posiada następujące własności:

- 1) w wypadku gdy jakaś domena d nie jest reprezentowana w próbie przez żadną jednostkę, tj. gdy $s_d = \emptyset$, nie jest możliwe oszacowanie wartości globalnej τ_{Y_d} w tej domenie,
- 2) do wyznaczenia wag kalibracyjnych w_k nie jest wymagana znajomość wartości globalnych zmiennych pomocniczych na poziomie domeny, a jedynie na poziomie całej populacji,

- 3) spełniona jest zasada benchmarkingu rozumianego w ten sposób, że suma oszacowanych wartości globalnych w domenach jest równa oszacowanej wartości globalnej w populacji, tj.

$$\sum_{d=1}^D \hat{\tau}_{\text{CAL},d}^{(1)} = \hat{\tau}_{\text{CAL}}, \quad (2.50)$$

przy czym $\hat{\tau}_{\text{CAL}}$ określony jest wzorem (2.3).

Równość (2.50) wynika z tego, że:

$$\sum_{d=1}^D \hat{\tau}_{\text{CAL},d}^{(1)} = \sum_{d=1}^D \sum_{k \in s} w_k y_{dk} = \sum_{k \in s} w_k \sum_{d=1}^D y_{dk} = \sum_{k \in s} w_k y_k = \hat{\tau}_{\text{CAL}}.$$

Spełnienie własności benchmarkingu jest szczególnie cenne i pożądane w praktyce badań statystycznych. Nie wszystkie jednak estymatory, zwłaszcza z zakresu statystyki małych obszarów, ją mają. W konsekwencji rodzi to konieczność wprowadzania odpowiednich korekt, tak aby suma oszacowań wartości globalnych w domenach była spójna z jej szacunkiem na poziomie całej populacji (Pfeffermann, 2013).

Wariancję estymatora (2.49) można oszacować z wykorzystaniem wzoru (2.47), przy czym reszty \tilde{e}_k zastępowane są przez:

$$\tilde{e}_{dk}^{(1)} = \delta_{dk} y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_{dws}, \quad (2.51)$$

gdzie $\hat{\boldsymbol{\beta}}_{dws}$ dane jest wzorem (2.48), a w miejsce y_k należy wstawić $\delta_{dk} y_k$.

W wypadku gdy dysponujemy wektorem zmiennych pomocniczych na poziomie domeny d oraz znane są ich wartości globalne $\sum_{k \in U_d} \mathbf{x}_k$, estymator kalibracyjny wartości globalnej τ_{Y_d} w domenie d można zdefiniować w następujący sposób:

$$\hat{\tau}_{\text{CAL},d}^{(2)} = \sum_{k \in s_d} w_{dk} y_k, \quad (2.52)$$

przy czym wagi kalibracyjne w_{dk} są rozwiązaniem poniższego zadania optymalizacyjnego¹⁵:

- (W1)_d – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s_d} \frac{d_k}{q_k} G\left(\frac{w_{dk}}{d_k}\right) \rightarrow \min, \quad (2.53)$$

¹⁵ Subskrypt d w nazwie warunku odnosi się do tego, że wagi kalibrujemy w obrębie każdej z wyróżnionych domen oddzielnie.

- (W2)_d – równania kalibracyjne:

$$\sum_{k \in s_d} w_{dk} \mathbf{x}_k = \sum_{k \in U_d} \mathbf{x}_k, \quad (2.54)$$

- (W3)_d – warunki ograniczające:

$$L \leq \frac{w_{dk}}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, n_d, \quad (2.55)$$

n_d oznacza liczebność zbioru s_d .

W wypadku funkcji $G(x) = \frac{1}{2}(x-1)^2$ wektor wag kalibracyjnych w_{dk} można wyznaczyć wprost ze wzoru. Mówi o tym poniższe twierdzenie.

Twierdzenie 2.4. *Rozwiązaniem zadania minimalizacji (2.53) przy warunku (2.54) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag kalibracyjnych \mathbf{w} , którego składowe w domenie d , dla $d = 1, \dots, D$, wyrażają się wzorem:*

$$w_{dk}^{\text{lin}} = d_k + d_k q_k \left(\sum_{k \in U_d} \mathbf{x}_k - \sum_{k \in s_d} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in s_d} q_k d_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \mathbf{x}_k. \quad (2.56)$$

Dla tak zdefiniowanej funkcji $G(\cdot)$ estymator kalibracyjny (2.52) wartości globalnej τ_{Y_d} w domenie d sprowadza się do uogólnionego liniowego estymatora regresyjnego GREG w domenie d postaci:

$$\hat{\tau}_{\text{GREG},d}^{\text{lin}} = \sum_{k \in s_d} d_k y_k + \left(\sum_{k \in U_d} \mathbf{x}_k - \sum_{k \in s_d} d_k \mathbf{x}_k \right)^T \hat{\boldsymbol{\beta}}_d, \quad (2.57)$$

przy czym:

$$\hat{\boldsymbol{\beta}}_d = \left(\sum_{k \in s_d} d_k q_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in s_d} d_k q_k \mathbf{x}_k y_k. \quad (2.58)$$

Estymator kalibracyjny (2.52), którego szczególnym przypadkiem jest (2.57), ma podobne własności do estymatora (2.49). Spełnione są i tutaj własności 1) oraz 3) sformułowane na stronie 65. Do wyznaczenia wag kalibracyjnych w_{dk} w wypadku estymatora (2.52) niezbędna jest jednak znajomość zmiennych pomocniczych, jak i ich wartości globalnych $\sum_{k \in U_d} \mathbf{x}_k$, na poziomie domeny d .

Zasadnicza różnica między estymatorem kalibracyjnym $\hat{\tau}_{\text{CAL},d}^{(1)}$ a $\hat{\tau}_{\text{CAL},d}^{(2)}$ wartości globalnej τ_{Y_d} w domenie d polega więc na tym, że w wypadku pierwszego z estymatorów wagi kalibrujemy tak, aby odtworzone zostały wartości globalne

zmiennych pomocniczych na poziomie całej populacji, a dla drugiego wagi kalibracyjne mają odtwarzać znane wartości globalne zmiennych pomocniczych na poziomie domeny. Warto również podkreślić, że estymator $\hat{\tau}_{CAL,d}^{(1)}$ jest w przybliżeniu nieobciążony, nawet gdy wartość oczekiwana liczebności w domenie jest mała, podczas gdy estymator $\hat{\tau}_{CAL,d}^{(2)}$ jest w przybliżeniu nieobciążony, gdy wartość oczekiwana liczebności próby w domenie jest duża (Rao i Molina, 2015; Żądło, 2015).

Wariancję estymatora (2.52) można oszacować z wykorzystaniem wzoru (2.47), przy czym reszty \tilde{e}_k zastępowane są przez:

$$\tilde{e}_{dk}^{(2)} = \delta_{dk} \left(y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_d \right), \quad (2.59)$$

gdzie $\hat{\boldsymbol{\beta}}_d$ dane jest wzorem (2.58).

Na zakończenie tego podrozdziału warto wskazać, że w literaturze przedmiotu (Särndal, 2007; Haziza i Beaumont, 2017) wskazuje się na trzy zasadnicze zalety kalibracji wag:

- w wypadku gdy zmienne pomocnicze wzięte pod uwagę w procesie wyznaczania wag kalibracyjnych są silnie skorelowane ze zmienną Y , reszty \hat{e}_k są bliskie zera, a wariancja estymatora mała; w rezultacie należy oczekiwać zysku na precyzji;
- kalibracja umożliwia uzyskanie zgodności (spójności) rozumianej w ten sposób, że oszacowane na podstawie n -elementowej próby s wartości globalne dla wszystkich zmiennych pomocniczych odpowiadają znanym wartościom globalnym tych zmiennych z zewnętrznych źródeł danych (spisów czy rejestrów administracyjnych);
- podejście kalibracyjne umożliwia redukcję negatywnego wpływu błędów nielosowych, takich jak braki odpowiedzi czy błędy pokrycia, które są nieodłącznym elementem wielu badań statystycznych zarówno częściowych, jak i pełnych.

2.3. Algorytm wyznaczania mnożników Lagrange’a w podejściu kalibracyjnym

Wyznaczanie wag kalibracyjnych (2.12) wymaga znajomości funkcji kalibracyjnej $F(\cdot)$ oraz wektora mnożników Lagrange’a $\boldsymbol{\lambda}$. Dla zadanej funkcji $G(\cdot)$ stosunkowo łatwo zbudować odpowiadającą jej funkcję kalibracyjną $F(\cdot)$. Znalezienie wektora $\boldsymbol{\lambda}$ w wielu wypadkach wymaga już jednak zastosowania podejścia ite-

racyjnego. Poniżej opisano algorytm poszukiwania wektora λ zgodnie z propozycją Vanderhoefta (2001). Uwzględniono dwa przypadki wyznaczania wektora mnożników Lagrange’a: bez ograniczeń na iloraz wag kalibracyjnych w_k i wag d_k wynikających z planu losowania próby oraz z ograniczeniami. Szerszy przegląd algorytmów wykorzystywanych do wyznaczania wag kalibracyjnych uwzględniających ograniczenia można znaleźć w pracy Daviesa (2018). Warto przy tym zaznaczyć, że większość programów, w których jest zaimplementowana kalibracja (ReGenesees, Calmar 2, g-CALIB-S oraz pakiet **calibrate** programu R), w procesie poszukiwania wag kalibracyjnych wykorzystuje algorytm Newtona-Raphsona. W programie GES napisanym w języku 4GL (SAS) wykorzystywany jest algorytm PMA (ang. *Projection Method Algorithm*). Z kolei w programie Bascula zaimplementowano algorytm SMQ (ang. *Scale Modified Quadratic*). W książce uwaga została skupiona na wyznaczeniu wektora mnożników Lagrange’a λ , który stanowi kluczowy element wag kalibracyjnych (2.12), z wykorzystaniem metody Newtona-Raphsona, która jest iteracyjnym algorytmem znajdowania minimum zadanej funkcji celu.

Na potrzeby skonstruowania odpowiednich algorytmów wyznaczania wektora mnożników Lagrange’a λ , z ograniczeniami na iloraz wag w_k i d_k i bez ograniczeń, zapiszemy problem poszukiwania wag kalibracyjnych (W1)–(W3) w ujęciu macierzowym.

Niech \mathbf{X}_s będzie macierzą, w której element x_{kj} oznacza wartość j -tej zmiennej pomocniczej dla k -tej jednostki badania z próby s , dla $k = 1, \dots, n$ oraz $j = 1, \dots, J$:

$$\mathbf{X}_s = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1J} \\ x_{21} & x_{22} & \dots & x_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nJ} \end{pmatrix}. \quad (2.60)$$

W macierzy tej wektor $\mathbf{x}_k = (x_{k1}, \dots, x_{kJ})^T$ odpowiadający k -temu wierszowi, stanowi wektor złożony z wartości wszystkich J zmiennych pomocniczych dla k -tej jednostki z próby s . Niech ponadto $\mathbf{D}_s = \text{diag}(\mathbf{d}) = \text{diag}(d_1, \dots, d_n)$ oznacza macierz diagonalną zawierającą na głównej przekątnej wagi d_k wynikające z planu losowania próby, tj.

$$\mathbf{D}_s = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}. \quad (2.61)$$

Podobnie definiujemy macierz $\mathbf{W}_s = \text{diag}(\mathbf{w}) = \text{diag}(w_1, \dots, w_n)$ oraz macierz $\mathbf{G}_s = \text{diag}(\mathbf{g}) = \text{diag}(g_1, \dots, g_n)$. Macierz \mathbf{W}_s zawiera na głównej przekątnej wagi kalibracyjne, a macierz \mathbf{G}_s mnożniki kalibracyjne. Alternatywnie macierz \mathbf{W}_s można przedstawić w następującej postaci: $\mathbf{W}_s = \text{diag}(d_1 g_1, \dots, d_n g_n)$.

Przyjmując, że $\boldsymbol{\tau}_X = \sum_{k \in U} \mathbf{x}_k$ oznacza wektor złożony z wartości globalnych wszystkich zmiennych pomocniczych, równania kalibracyjne opisane w warunku (W2) można zapisać w ujęciu macierzowym w następującej postaci:

$$\mathbf{X}_s^T \mathbf{w} = \boldsymbol{\tau}_X. \quad (2.62)$$

Niech $\mathbf{g} = (g_1, \dots, g_n)^T$ oraz $\mathbf{1}_n$ oznaczają odpowiednio wektor złożony z mnożników kalibracyjnych i wektor kolumnowy złożony z n jedynek. Warunek ten można też zapisać alternatywnie jako $\mathbf{X}_s^T \mathbf{W}_s \mathbf{1}_n = \boldsymbol{\tau}_X$ lub $\mathbf{X}_s^T \mathbf{D}_s \mathbf{g} = \boldsymbol{\tau}_X$ czy $\check{\mathbf{X}}_s^T \mathbf{g} = \boldsymbol{\tau}_X$, gdzie $\check{\mathbf{X}}_s = \mathbf{D}_s \mathbf{X}_s$.

Również warunek (W1), który odnosi się do poszukiwania minimum odpowiedniej funkcji odległości $D(\cdot)$, można zapisać w postaci macierzowej. W tym celu zauważmy, że $G\left(\frac{w_k}{d_k}\right) = G(g_k)$, a funkcję odległości, dla której poszukujemy minimum warunkowego, można wyrazić w postaci $D(\mathbf{d}, \mathbf{w}) = \sum_{k \in S} \frac{d_k}{q_k} G(g_k)$. Przyjmując z kolei, że $G(\mathbf{g}) = (G(g_1), \dots, G(g_n))^T$ oraz $\mathbf{q} = (q_1, \dots, q_n)^T$, funkcję $D(\cdot)$ w warunku (W1) można przedstawić jako:

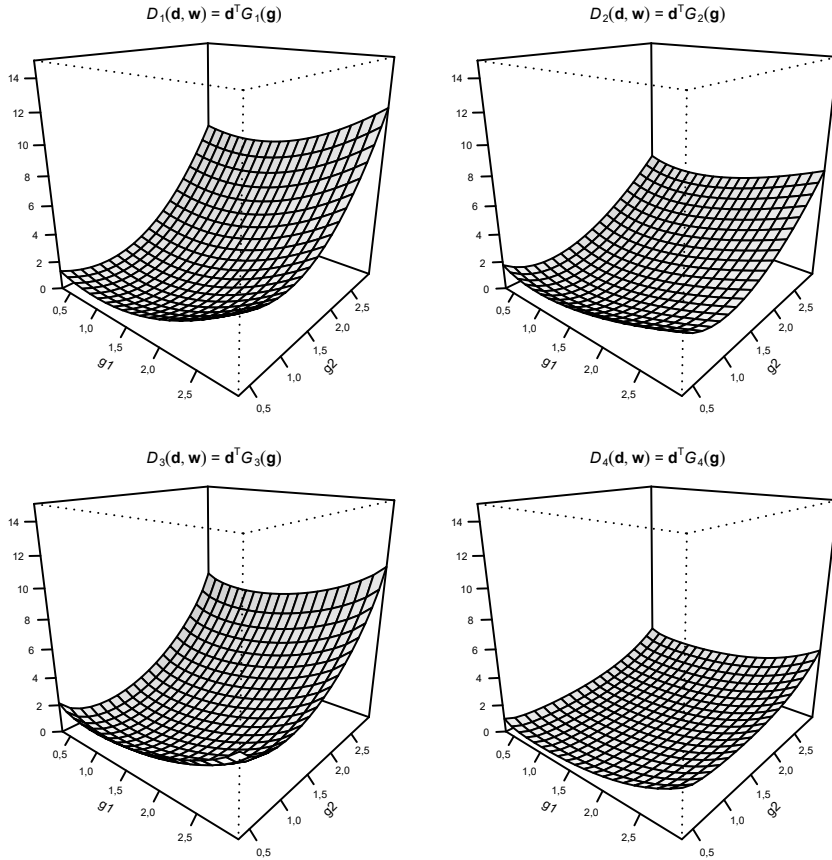
$$D(\mathbf{d}, \mathbf{w}) = \mathbf{d}^T \mathbf{Q}_s^{-1} G(\mathbf{g}), \quad (2.63)$$

gdzie $\mathbf{Q}_s = \text{diag}(\mathbf{q})$ oznacza macierz diagonalną zawierającą na głównej przekątnej wagi q_k . Przyjmując, że $q_k = 1$, powierzchnie odpowiadające funkcji celu, dla których poszukujemy minimum warunkowego, tj. $D(\mathbf{d}, \mathbf{w}) = \mathbf{d}^T G(\mathbf{g})$ dla różnych funkcji $G(\cdot)$, można przedstawić w przestrzeni \mathbb{R}^3 (rys. 2.2)¹⁶. Na przykład, jeśli weźmiemy pod uwagę funkcję (2.18), to odpowiadająca jej funkcja celu $D_1(\mathbf{d}, \mathbf{w})$ będzie postaci (metoda liniowa):

$$D_1(\mathbf{d}, \mathbf{w}) = (g_1 - 1)^2 + 2(g_2 - 1)^2.$$

Z kolei warunek (W3), w którym nakłada się ograniczenia na wagi kalibracyjne, a dokładniej na iloraz wag kalibracyjnych w_k i wag wynikających z planu losowania próby d_k , można zapisać w postaci $\mathbf{g} \in \Omega$, gdzie Ω jest ograniczonym

¹⁶ Na potrzeby wizualizacji przyjęto, że w wypadku funkcji $G_3(x)$ i $G_4(x)$: $d_1 = 2$, $d_2 = 4$ oraz $L = 0,33$ i $U = 3$. Numeracja funkcji celu $D(\cdot)$ jest zgodna z oznaczeniami funkcji $G(\cdot)$ w tabeli 2.1.



Rysunek 2.2. Wykresy powierzchni dla wybranych funkcji celu $D(\mathbf{d}, \mathbf{w})$

Źródło: na podstawie przykładowych danych.

podzbiorem $[L, U]^n$ przestrzeni euklidesowej \mathbb{R}^n . Ostatecznie problem poszukiwania wag kalibracyjnych w_k , gdzie $k = 1, \dots, n$, opisanych warunkami (W1)–(W3) można zapisać w postaci poniższego zadania optymalizacyjnego.

Model 1. Problem wyznaczania wag kalibracyjnych w_k

$$D(\mathbf{d}, \mathbf{w}) = \mathbf{d}^T \mathbf{Q}_s^{-1} G(\mathbf{g}) \rightarrow \min, \quad (2.64)$$

przy warunkach:

$$\mathbf{X}_s^T \mathbf{w} = \boldsymbol{\tau}_X, \quad (2.65)$$

$$\mathbf{g} \in \Omega. \quad (2.66)$$

Alternatywnie zagadnienie wyznaczania wag kalibracyjnych można zapisać jako problem poszukiwania mnożników kalibracyjnych g_k dla $k = 1, \dots, n$. Ich znajomość umożliwi wyznaczenie wag kalibracyjnych w_k ze wzoru (2.12)¹⁷.

Model 2. Problem wyznaczania mnożników kalibracyjnych g_k

$$D(\mathbf{d}, \mathbf{w}) = \mathbf{d}^T \mathbf{Q}_s^{-1} G(\mathbf{g}) \rightarrow \min, \quad (2.67)$$

przy warunkach:

$$\check{\mathbf{X}}_s^T \mathbf{g} = \boldsymbol{\tau}_X, \quad (2.68)$$

$$\mathbf{g} \in \Omega. \quad (2.69)$$

Punktem wyjścia do znalezienia wektora mnożników Lagrange'a jest równanie (2.13). Można je równoważnie zapisać w postaci:

$$\Phi(\boldsymbol{\lambda}) = \sum_{k \in s} d_k F(q_k \mathbf{x}_k^T \boldsymbol{\lambda}) \mathbf{x}_k - \boldsymbol{\tau}_X = \mathbf{X}_s^T \mathbf{w}(\boldsymbol{\lambda}) - \boldsymbol{\tau}_X = \mathbf{0}, \quad (2.70)$$

gdzie $\mathbf{w}(\boldsymbol{\lambda}) = \mathbf{D}_s F(\mathbf{X}_s \boldsymbol{\lambda})$. Wektor czynników nieoznaczonych Lagrange'a można znaleźć, rozwiązując nieliniowy układ J równań $\Phi(\boldsymbol{\lambda}) = \mathbf{0}$ z wykorzystaniem metody Newtona-Raphsona. Algorytm ten polega na wyborze punktu startowego $\boldsymbol{\lambda}$, którym jest zazwyczaj wektor złożony z samych zer, a następnie rekurencyjnie przekształca się ten wektor aż do momentu, gdy kolejne przybliżenia będą wystarczająco satysfakcjonujące. Wektory są przekształcane zgodnie z poniższym równaniem macierzowym:

$$\boldsymbol{\lambda}^{(l)} = \boldsymbol{\lambda}^{(l-1)} - [\Phi'(\boldsymbol{\lambda}^{(l-1)})]^{-1} \Phi(\boldsymbol{\lambda}^{(l-1)}), \quad (2.71)$$

które w sukcesywny sposób pozwala uaktualniać mnożniki Lagrange'a $\boldsymbol{\lambda}^{(l)}$ w l -tej iteracji, $l = 1, 2, \dots$, zaczynając od $\boldsymbol{\lambda}^{(0)} = \mathbf{0}$, gdzie Φ' jest pochodną (Frécheta), a macierz określona jako $[\Phi'(\cdot)]^{-1}$ oznacza tzw. uogólnioną macierz odwrotną do macierzy $\Phi'(\cdot)$. W każdej iteracji algorytmu w pierwszej kolejności obliczany jest wektor

$$\mathbf{w}(\boldsymbol{\lambda}^{(l-1)}) = \mathbf{D}_s F(\mathbf{X}_s \boldsymbol{\lambda}^{(l-1)}), \quad (2.72)$$

na bazie którego wyznaczana jest następująca macierz:

$$\Phi(\boldsymbol{\lambda}^{(l-1)}) = \mathbf{X}_s^T \mathbf{w}(\boldsymbol{\lambda}^{(l-1)}) - \boldsymbol{\tau}_X \quad (2.73)$$

¹⁷ Tak sformułowane problemy poszukiwania wag bądź mnożników kalibracyjnych można zatem rozpatrywać jako zagadnienia programowania matematycznego.

oraz macierz postaci:

$$\Phi'(\boldsymbol{\lambda}^{(l-1)}) = \mathbf{X}_s^T \mathbf{W}_s(\boldsymbol{\lambda}^{(l-1)}) \mathbf{X}_s, \quad (2.74)$$

gdzie $\mathbf{W}_s(\boldsymbol{\lambda}^{(l-1)})$ można przedstawić jako:

$$\mathbf{W}_s(\boldsymbol{\lambda}^{(l-1)}) = \text{diag}(\mathbf{w}(\boldsymbol{\lambda}^{(l-1)})). \quad (2.75)$$

W pierwszym kroku opiszemy algorytm wyznaczania mnożników Lagrange'a $\boldsymbol{\lambda}$ odpowiadający sformułowaniu zadania optymalizacyjnego poszukiwania wag kalibracyjnych w modelu 1 lub modelu 2. Pominiemy jednak założenie mówiące, że $\mathbf{g} \in \Omega$. Oznacza to, że nie wymagamy, aby wagi kalibracyjne należały do pewnego z góry określonego przedziału. Dopuszczamy zatem, w zależności od funkcji $G(\cdot)$, występowanie wag ujemnych czy ekstremalnych, tj. znacznie różniących się od wag wynikających z planu losowania próby.

Algorytm 1. Algorytm wyznaczania mnożników $\boldsymbol{\lambda}$

Krok 1: *Inicjalizacja*

Przyjmij wejściowy wektor mnożników Lagrange'a w postaci $\boldsymbol{\lambda}^{(0)} = \mathbf{0}$. Oblicz początkowe wartości wektora wag kalibracyjnych według następującego wzoru:

$$\mathbf{w}^{(0)} = \mathbf{D}_s F(\mathbf{X}_s \boldsymbol{\lambda}^{(0)}) = \mathbf{D}_s \mathbf{1}_n = \mathbf{d}. \quad (2.76)$$

Przejdź do kroku 2.

Krok 2: *Obliczenie mnożnika $\boldsymbol{\lambda}^{(1)}$*

Oblicz wartość mnożnika Lagrange'a $\boldsymbol{\lambda}^{(1)}$, korzystając z następującego wzoru:

$$\boldsymbol{\lambda}^{(1)} = -\left(\mathbf{X}_s^T \mathbf{D}_s \mathbf{X}_s\right)^{-1} \left(\mathbf{X}_s^T \mathbf{d} - \boldsymbol{\tau}_X\right) \quad (2.77)$$

oraz odpowiadające mu wagi kalibracyjne:

$$\mathbf{w}^{(1)} = \mathbf{D}_s F(\mathbf{X}_s \boldsymbol{\lambda}^{(1)}). \quad (2.78)$$

Ustal $l = 1$. Przejdź do kroku 3.

Krok 3: *Kryterium stopu* (por. wzory (2.81)–(2.83))

Jeżeli warunek stopu jest spełniony, to idź do kroku 6. W przeciwnym razie idź do kroku 4.

Krok 4: Obliczenie $\Phi^{(l)}$ oraz $\Phi'^{(l)}$

Przyjmując $\mathbf{W}_s^{(l)} = \text{diag}(\mathbf{w}^{(l)})$, oblicz $\Phi^{(l)} = \mathbf{X}_s^T \mathbf{w}^{(l)} - \tau_{\mathbf{X}}$ i $\Phi'^{(l)} = \mathbf{X}_s^T \mathbf{W}_s^{(l)} \mathbf{X}_s$, a następnie idź do kroku 5.

Krok 5: Obliczenie $\lambda^{(l)}$ oraz wektora wag kalibracyjnych $\mathbf{w}^{(l)}$

Ustaw l na $l+1$. Oblicz wektor mnożników Lagrange'a $\lambda^{(l)}$, korzystając ze wzoru (2.71) lub ze wzoru postaci:

$$\lambda^{(l)} = \lambda^{(l-1)} - \left(\mathbf{X}_s^T \mathbf{W}_s^{(l-1)} \mathbf{X}_s \right)^{-1} \left(\mathbf{X}_s^T \mathbf{w}^{(l-1)} - \tau_{\mathbf{X}} \right) \quad (2.79)$$

oraz wyznacz nowe wagi kalibracyjne:

$$\mathbf{w}^{(l)} = \mathbf{D}_s F \left(\mathbf{X}_s \lambda^{(l)} \right). \quad (2.80)$$

Wróć do kroku 3.

Krok 6: Końcowe rozwiązanie

Finalny wektor wag kalibracyjnych jest postaci: $\mathbf{w} = \mathbf{w}^{(l)} = \mathbf{D}_s F \left(\mathbf{X}_s \lambda^{(l)} \right)$.

W powyższym algorytmie kryterium stopu może być w różny sposób zdefiniowane. Przyjmując, że ε jest tzw. poziomem tolerancji, najczęściej zakłada się, że zbieżność jest osiągnięta, gdy spełniony jest poniższy warunek:

$$\max_{1 \leq k \leq n} \left| g_k^{(l)} - g_k^{(l-1)} \right| \leq \varepsilon. \quad (2.81)$$

Można jednak rozpatrywać inne kryteria, do których należą:

$$\max_{1 \leq k \leq n} \left| \frac{w_k^{(l)} - w_k^{(l-1)}}{w_k^{(l-1)}} \right| \leq \varepsilon \quad (2.82)$$

lub

$$\left\| \left(\mathbf{W}_s^{(l)} - \mathbf{W}_s^{(l-1)} \right) \left(\mathbf{W}_s^{(l)} \right)^{-1} \right\| \leq \varepsilon, \quad (2.83)$$

gdzie $\|\cdot\|$ oznacza normę euklidesową.

Powyżej sformułowany algorytm nie brał pod uwagę warunku $\mathbf{g} \in \Omega$. Poniżej zostaną podane kroki tzw. rozszerzonego algorytmu wyznaczania mnożników Lagrange'a λ , w którym zakładamy, że wagi kalibracyjne należą do pewnego z góry określonego przedziału. Z warunku (W3), w którym $L \leq \frac{w_k}{d_k} \leq U$, wynika, że $d_k L \leq w_k \leq d_k U$ dla pewnych stałych L i U takich, że $0 \leq L \leq 1 \leq U$. W algo-

rytmie tym, który jest modyfikacją opisanego wcześniej podstawowego algorytmu poszukiwania wektora mnożników Lagrange'a λ oraz końcowych wag kalibracyjnych, dokonuje się „przycięcia” wag kalibracyjnych w_k (ang. *truncated weights*) po ich uaktualnieniu, gdy nie leżą w założonym przedziale $d_k L \leq w_k \leq d_k U$. Oznacza to, że jeśli dla pewnego k zajdzie nierówność $w_k < d_k L$, przycięcie polega na tym, że przyjmujemy $w_k = d_k L$. Z drugiej strony, gdy $w_k > d_k U$, przyjmujemy, że po przycięciu $w_k = d_k U$. Zakładamy przy tym, że $L > 0$.

Ostatecznie, przyjmując kryterium stopu w postaci (2.81), rozszerzony algorytm wyznaczania mnożników Lagrange'a λ oraz wag kalibracyjnych w_k można sformułować w następujący sposób.

Algorytm 2. Rozszerzony algorytm wyznaczania mnożników λ

Krok 1: *Inicjalizacja*

Ustal wartości dla parametrów L i U . Następnie przyjmij wejściowy wektor mnożników Lagrange'a w postaci $\lambda^{(0)} = \mathbf{0}$. W dalszym ciągu oblicz początkowe wartości wektora wag kalibracyjnych ze wzoru (2.76). Przejdź do kroku 2.

Krok 2: *Obliczenie mnożnika $\lambda^{(1)}$*

Przyjmij, że $l = 1$. Oblicz wartość mnożnika Lagrange'a $\lambda^{(1)}$, korzystając ze wzoru (2.77), oraz odpowiadające mu wagi kalibracyjne ze wzoru (2.78). Sprawdź, czy wyznaczone wagi kalibracyjne wymagają operacji przycięcia. Jeśli tak, przejdź do kroku 3. Jeśli nie, przejdź do kroku 5.

Krok 3: *Przycięcie wag*

Przyjmij $w_k = d_k L$, gdy $w_k < d_k L$. W sytuacji gdy $w_k > d_k U$, przyjmij $w_k = d_k U$. Idź do kroku 4.

Krok 4: *Kryterium stopu*

Jeżeli warunek stopu jest spełniony, idź do kroku 7. W przeciwnym razie idź do kroku 5.

Krok 5: *Obliczenie $\Phi^{(l)}$ oraz $\Phi'^{(l)}$*

Oznaczając przez $\tilde{\mathbf{w}}^{(l)}$ wektor wag kalibracyjnych po operacji przycięcia (gdy była wymagana) i przyjmując $\tilde{\mathbf{W}}_s^{(l)} = \text{diag}(\tilde{\mathbf{w}}^{(l)})$, oblicz $\Phi^{(l)} = \mathbf{X}_s^T \tilde{\mathbf{w}}^{(l)} - \boldsymbol{\tau}_X$ oraz $\Phi'^{(l)} = \mathbf{X}_s^T \tilde{\mathbf{W}}_s^{(l)} \mathbf{X}_s$, a następnie idź do kroku 6.

Krok 6: *Obliczenie $\lambda^{(l)}$ oraz wektora wag kalibracyjnych $\mathbf{w}^{(l)}$*

Ustaw l na $l + 1$. Oblicz mnożnik Lagrange'a $\lambda^{(l)}$, korzystając ze wzoru (2.71) lub ze wzoru:

$$\lambda^{(l)} = \lambda^{(l-1)} - \left(\mathbf{X}_s^T \tilde{\mathbf{W}}_s^{(l-1)} \mathbf{X}_s \right)^{-1} \left(\mathbf{X}_s^T \tilde{\mathbf{w}}^{(l-1)} - \boldsymbol{\tau}_X \right). \quad (2.84)$$

Oblicz nowe wagi kalibracyjne:

$$\tilde{\mathbf{w}}^{(l)} = \mathbf{D}_s F(\mathbf{X}_s \boldsymbol{\lambda}^{(l)}). \quad (2.85)$$

Wróć do kroku 3.

Krok 7: Końcowe rozwiązanie

Finalny wektor wag kalibracyjnych jest postaci: $\mathbf{w} = \tilde{\mathbf{w}}^{(l)} = \mathbf{D}_s F(\mathbf{X}_s \boldsymbol{\lambda}^{(l)})$.

Opisany powyżej rozszerzony algorytm wyznaczania czynników nieoznaczonych Lagrange'a, a w konsekwencji wag kalibracyjnych w_k , umożliwia ich przycięcie przy jednoczesnym minimalizowaniu odpowiedniej funkcji odległości i spełnieniu równań kalibracyjnych. W konsekwencji zmniejsza się dyspersja wag kalibracyjnych. Operacja ta ma się przyczynić również do zmniejszenia błędu średniokwadratowego odpowiedniego estymatora kalibracyjnego wykorzystującego przycięte wagi.

Warto nadmienić, że istnieją również inne metody przycinania wag kalibracyjnych, które uznawane są za ekstremalne. Techniki te bazują na ogół na ustanowieniu pewnego punktu progowego w_0 , powyżej którego wagę kalibracyjną w_k uznaje się za odstającą. W takiej sytuacji wymusza się, aby po przycięciu waga kalibracyjna równa była wartości progowej w_0 . Wiąże się to jednocześnie z pewną modyfikacją wag kalibracyjnych, które nie przekraczają wartości progowej w_0 , po to, aby sumy wag kalibracyjnych przed operacją przycięcia i po niej były sobie równe. Formalnie, oznaczając odpowiednio przez w_k oraz $w_{k,t}$ wagę kalibracyjną przypisaną jednostce k przed i po operacji przycięcia, oznacza to, że:

$$w_{k,t} = \begin{cases} w_0, & \text{jeżeli } w_k \geq w_0, \\ \gamma w_k, & \text{jeżeli } w_k < w_0, \end{cases} \quad (2.86)$$

gdzie γ jest tzw. czynnikiem skalującym, który zapewnia spełnienie poniższego warunku:

$$\sum_{k \in s} w_{k,t} = \sum_{k \in s} w_k. \quad (2.87)$$

Odpowiednie ustalenie wartości progowej w_0 może wpływać na wartość błędu średniokwadratowego estymatora kalibracyjnego. Zła specyfikacja w_0 może prowadzić do wzrostu błędu średniokwadratowego estymatora kalibracyjnego z przyciętymi wagami w porównaniu z estymatorem kalibracyjnym, w którym wagi nie były w żaden sposób modyfikowane. Problem odpowiedniego doboru w_0 jest oma-

wiany w pracy Pottera (1990). Szczegółowo kwestia ta jest rozważana także w artykule Chena i współautorów (2017).

Warto w tym miejscu wspomnieć również o innej modyfikacji wag d_k wynikających z planu losowania próby, tzw. wygładzaniu wag (ang. *weight smoothing*), które zostało zaproponowane przez Beaumonta (2008). Idea tego podejścia sprowadza się do zbudowania odpowiedniego modelu regresji, w którym funkcję zmiennej objaśnianej będą pełniły wagi d_k , a zmienną objaśniającą będzie Y . Formalnie rozpatruje się zatem model postaci:

$$d_k = \mu(\mathbf{y}_k, \boldsymbol{\beta}) + \varepsilon_k, \quad (2.88)$$

gdzie wektor $\mathbf{y}_k = (y_{1k}, \dots, y_{Pk})^T$, a $\boldsymbol{\beta}$ jest P -wymiarowym wektorem nieznanych parametrów podlegających procesowi estymacji. Wagi wygładzone są wartościami teoretycznymi wyznaczonymi z wykorzystaniem powyższego modelu, tj. $\hat{d}_k = \mu(\mathbf{y}_k, \hat{\boldsymbol{\beta}})$, gdzie $\hat{\boldsymbol{\beta}}$ jest estymatorem wektora parametrów $\boldsymbol{\beta}$ uzyskanym uogólnioną metodą najmniejszych kwadratów. Ostatecznie wygładzony estymator wartości globalnej (1.1) wyraża się wzorem:

$$\hat{\tau}_{\text{SMOOTH}} = \sum_{k \in s} \hat{d}_k y_k. \quad (2.89)$$

Podejście to nie będzie jednak przedmiotem rozważań w niniejszej pracy. Szczegółowo zostało ono omówione we wspomnianym artykule Beaumonta (2008).

Na zakończenie tego podrozdziału warto podkreślić, że operacja przycinania wag w procesie kalibracyjnym może spowodować, że nie wszystkie wartości globalne zmiennych pomocniczych zostaną dokładnie odtworzone. Oznacza to, że oszacowane z wykorzystaniem wag kalibracyjnych, wartości globalne niektórych zmiennych pomocniczych mogą się nieznacznie różnić od ich wartości globalnych znanych ze spisu czy innych źródeł. Z sytuacją taką możemy mieć do czynienia, gdy dysponujemy znaczną liczbą zmiennych pomocniczych oraz ich znanymi wartościami globalnymi (Barranco-Chamorro, Jiménez-Gamero, Mayor-Gallego i Moreno-Rebollo, 2015; Guggemos i Tillé, 2010). W takim wypadku można wykorzystać tzw. kalibrację z karami (ang. *penalized calibration*), aby kontrolować różnicę pomiędzy wartościami globalnymi zmiennych pomocniczych oszacowanymi na podstawie wyznaczonych wag kalibracyjnych, a ich znanymi z innych źródeł wartościami prawdziwymi. Wartości globalne nie są zatem dokładnie odtworzone, a jedynie w pewnym przybliżeniu. Odbywa się to poprzez wprowadzenie do funkcji odległości $D(\mathbf{d}, \mathbf{w})$, dla której poszukuje się minimum warunkowego, dodatkowego składnika w postaci czynnika kary.

2.4. Badanie symulacyjne – podejście minimalizujące funkcję odległości

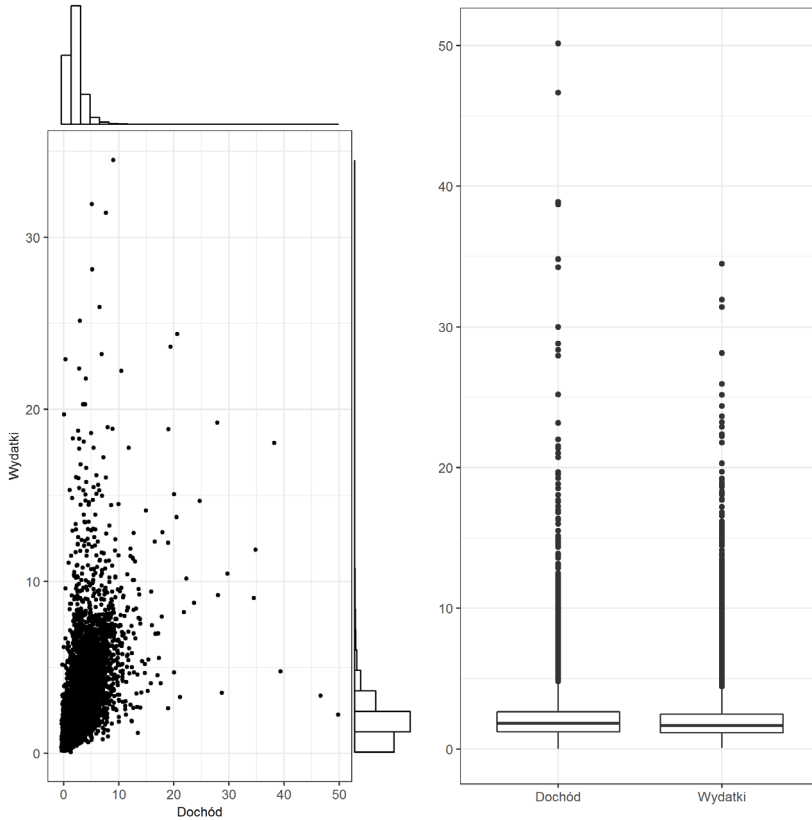
Celem przeprowadzonego badania symulacyjnego jest analiza najważniejszych własności estymatorów kalibracyjnych wartości globalnej (1.34) uzyskanych dla różnych postaci funkcji $G(\cdot)$ w podejściu minimalizującym funkcję odległości, opisanym szczegółowo w podrozdziale 2.2. Na potrzeby jego realizacji wykorzystano wspomniane już rzeczywiste dane pochodzące z Badania Budżetów Gospodarstw Domowych, które odgrywa ważną rolę w analizach poziomu życia ludności i stanowi podstawowe źródło informacji o rozchodach, spożyciu ilościowym żywności, przychodach oraz o innych aspektach warunków życia określonych grup ludności (GUS, 2017a). Za zmienną objaśnianą Y przyjęto wydatki ogółem (w tys. zł) gospodarstwa domowego, a jako zmienną objaśniającą w procesie kalibracji wzięto pod uwagę dochód rozporządzalny gospodarstwa domowego (w tys. zł). Innymi słowy, oszacowaniu podlegały wydatki ogółem wszystkich gospodarstw domowych.

Na potrzeby realizacji badania symulacyjnego utworzono tzw. pseudopopulację. Ze zbioru danych BBGD usunięto gospodarstwa, dla których nie było danych na temat dochodu rozporządzalnego lub wydatków ogółem. Usunięto również gospodarstwa, dla których obydwie rozpatrywane zmienne przyjmowały wartości niedodatnie. Tak utworzona pseudopopulacja liczyła ostatecznie $N = 32\,291$ gospodarstw domowych.

Na rysunku 2.3 zamieszczono korelacyjny diagram rozrzutu pomiędzy dochodem rozporządzalnym gospodarstw domowych a wydatkami ogółem oraz wykresy pudełkowe dla każdej zmiennej z osobna w tak utworzonej pseudopopulacji.

Można zauważyć, że istnieje dość silna dodatnia zależność między dochodem gospodarstw domowych a wydatkami ogółem. Potwierdzeniem tego jest oszacowany współczynnik korelacji liniowej Pearsona, który wynosi $r_{xy} = 0,65$. Zarówno dochód, jak i wydatki ogółem mają rozkład o asymetrii prawostronnej. Wykresy pudełkowe wskazują ponadto, że dla wielu gospodarstw domowych można zaobserwować odstające wartości dochodu rozporządzalnego i wydatków ogółem.

W kolejnym kroku badania z utworzonej pseudopopulacji losowano 2% próbę zgodnie z planem losowania prostego bez zwracania oraz dokonywano kalibracji wag d_k . Powyższy proces powtórzono $B = 200$ razy. W każdej replikacji, mając wyznaczone wagi kalibracyjne w_k , dokonano oszacowania wartości globalnej τ_Y , tj. łącznych wydatków ogółem gospodarstw domowych, korzystając ze wzoru (2.3). W badaniu symulacyjnym uwzględniono przy tym różne posta-



Rysunek 2.3. Korelacyjny diagram rozrzutu między wydatkami ogółem a dochodem rozporządzalnym gospodarstw domowych (lewy panel) oraz wykresy pudełkowe dochodu rozporządzalnego i wydatków ogółem (prawy panel)

Źródło: na podstawie BBGD.

cie funkcji $G(\cdot)$, które prowadziły do odmiennych zestawów wag kalibracyjnych w_k . Rozpatrzono wszystkie funkcje G_1 – G_5 , które w syntetyczny sposób przedstawiono w tabeli 2.1¹⁸. Odpowiednie estymatory kalibracyjne wartości globalnej dla różnych funkcji $G(\cdot)$ oznaczono jako $\hat{\pi}_{\text{lin}}$, $\hat{\pi}_{\text{rak}}$, $\hat{\pi}_{\text{log}}$, $\hat{\pi}_{\text{sinh}}$, $\hat{\pi}_{\text{lin_tr}}$, odpowiadające im zestawy wag kalibracyjnych przez w_{lin} , w_{rak} , w_{log} , w_{sinh} , $w_{\text{lin_tr}}$, a mnożniki kalibracyjne jako g_{lin} , g_{rak} , g_{log} , g_{sinh} , $g_{\text{lin_tr}}$ ¹⁹. W wypadku funkcji G_3 , G_4 oraz G_5 przyjęto ponadto ograniczenia na iloraz w_k/d_k : $L = \frac{2}{3}$ i $U = \frac{3}{2}$. W każdym z roz-

¹⁸ Numeracja funkcji jest zgodna z kolumną Lp. Na przykład G_4 odpowiada funkcja $G(\cdot)$ bazująca na sinusie hiperbolicznym.

¹⁹ Skrótowe oznaczenia pochodzą od angielskich nazw tych metod: lin – linear, rak – raking, log – logit, sinh – hyperbolic sine, lin_tr – linear truncated.

patrywanych przypadków założono, że $q_k = 1$. Na potrzeby estymacji wartości globalnej wydatków ogółem gospodarstw domowych wzięto również pod uwagę estymator Horvitz-Thompsona (1.35).

Jak wcześniej wspomniano, w procesie kalibracji wag jako zmienną pomocniczą wykorzystano dochód rozporządzalny gospodarstw domowych. Wzięto ponadto sztuczną zmienną przyjmującą dla każdego gospodarstwa wylosowanego do próby s wartość 1. Zapewniono sobie w ten sposób sumowalność wag kalibracyjnych w_k do liczebności pseudopopulacji N . W celu oceny jakości rozważanych estymatorów kalibracyjnych i estymatora Horvitz-Thompsona w badaniu symulacyjnym obliczono następujące miary²⁰:

- względne obciążenie (ang. *relative bias*):

$$\text{RB}_{\text{SIM}}(\hat{\tau}_{\text{CAL}}) = \frac{1}{B} \sum_{b=1}^B \frac{\hat{\tau}_{\text{CAL}}^{(b)} - \tau_Y}{\tau_Y} \cdot 100\%, \quad (2.90)$$

- błąd średniokwadratowy (ang. *mean squared error*):

$$\text{MSE}_{\text{SIM}}(\hat{\tau}_{\text{CAL}}) = \frac{1}{B} \sum_{b=1}^B \left(\hat{\tau}_{\text{CAL}}^{(b)} - \tau_Y \right)^2, \quad (2.91)$$

- względny błąd średniokwadratowy²¹ (ang. *relative root mean squared error*):

$$\text{RRMSE}_{\text{SIM}}(\hat{\tau}_{\text{CAL}}) = \frac{\sqrt{\text{MSE}_{\text{SIM}}(\hat{\tau}_{\text{CAL}})}}{\tau_Y} \cdot 100\%, \quad (2.92)$$

- wskaźnik pokrycia wartości globalnej odpowiadający 95% przedziałowi ufności bazującemu na ocenie odpowiedniego estymatora kalibracyjnego $\hat{\tau}_{\text{CAL}}$ i ocenie jego wariancji $\hat{D}^2(\hat{\tau}_{\text{CAL}})$ (ang. *coverage rate*):

$$\text{CR}_{\text{SIM}} \left[\hat{D}^2(\hat{\tau}_{\text{CAL}}) \right] = \frac{1}{B} \sum_{b=1}^B I^{(b)}, \quad (2.93)$$

- średnia długość przedziału ufności (ang. *average length of the confidence interval*):

$$\text{AL}_{\text{SIM}}(\hat{\tau}_{\text{CAL}}) = \frac{1}{B} \sum_{b=1}^B \left(a_2^{(b)} - a_1^{(b)} \right). \quad (2.94)$$

²⁰ W oznaczeniach wzorów dodano subskrypt dolny SIM, aby wyraźnie podkreślić, że ocena miar jest dokonywana w badaniu symulacyjnym metodą Monte Carlo.

²¹ Ze względu na znajomość „prawdziwej” wartości globalnej Y w pseudopopulacji w badaniu symulacyjnym, w mianowniku na wzór RRMSE, wykorzystano ją w miejsce jej oszacowania – por. wzór (1.33).

W powyższych wzorach $\hat{\tau}_{\text{CAL}}^{(b)}$ oznacza ocenę odpowiedniego estymatora kalibracyjnego wartości globalnej w b -tej iteracji badania symulacyjnego, a $I^{(b)}$ zdefiniowane jest jako:

$$I^{(b)} = \begin{cases} 1, & \text{jeżeli } [a_1^{(b)}, a_2^{(b)}] \text{ pokrywa } \tau_Y, \\ 0, & \text{w przeciwnym razie,} \end{cases} \quad (2.95)$$

gdzie

$$a_1^{(b)} = \hat{\tau}_{\text{CAL}}^{(b)} - z_{1-\frac{\alpha}{2}} \sqrt{\hat{D}^2(\hat{\tau}_{\text{CAL}}^{(b)})}, \quad a_2^{(b)} = \hat{\tau}_{\text{CAL}}^{(b)} + z_{1-\frac{\alpha}{2}} \sqrt{\hat{D}^2(\hat{\tau}_{\text{CAL}}^{(b)})},$$

przy czym $z_{1-\frac{\alpha}{2}}$ jest kwantylem rzędu $1 - \frac{\alpha}{2}$ w rozkładzie $N(0, 1)$, $\hat{D}^2(\hat{\tau}_{\text{CAL}}^{(b)})$ oznacza ocenę wariancji estymatora kalibracyjnego wyznaczoną zgodnie ze wzorem (2.47) w b -tej iteracji.

W analogiczny sposób obliczono wybrane miary jakości dla estymatora Horwitza-Thompsona. W związku z tym, że w badaniu symulacyjnym przyjęto losowanie proste bez zwracania, odpowiednie prawdopodobieństwa inkluzji pierwszego i drugiego rzędu można wyznaczyć ze wzorów (1.19) i (1.20). Wówczas:

$$\Delta_{kl} = \begin{cases} -\frac{n(N-n)}{N^2(N-1)}, & \text{jeżeli } k \neq l, \\ \frac{n(N-n)}{N^2}, & \text{jeżeli } k = l \end{cases} \quad (2.96)$$

i w takim wypadku $\hat{D}^2(\hat{\tau}_{\text{HT}})$ można wyznaczyć ze wzoru:

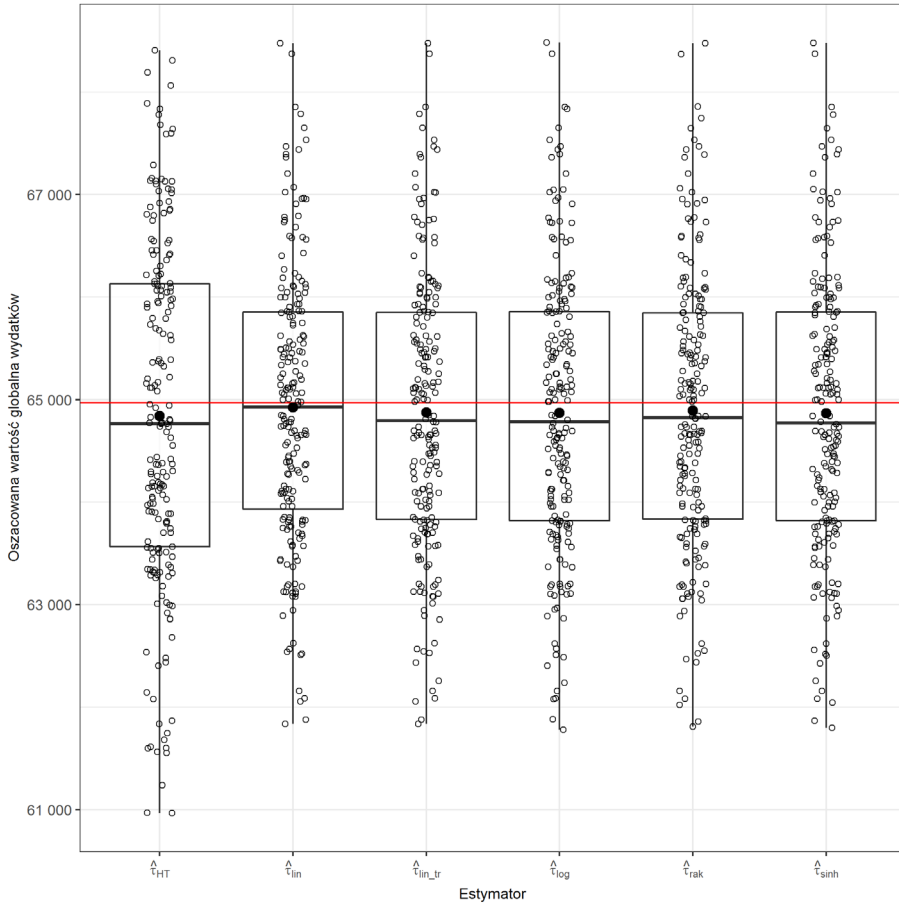
$$\hat{D}^2(\hat{\tau}_{\text{HT}}) = N^2 \left(1 - \frac{n}{N}\right) \frac{S_{ys}^2}{n}, \quad (2.97)$$

gdzie

$$S_{ys}^2 = \frac{1}{n-1} \sum_{k \in s} (y_k - \bar{y}_s)^2 \quad \text{oraz} \quad \bar{y}_s = \frac{1}{n} \sum_{k \in s} y_k.$$

Rysunek 2.4 pokazuje oszacowania punktowe wartości globalnej wydatków ogółem gospodarstw domowych wszystkich rozpatrywanych estymatorów kalibracyjnych oraz estymatora Horwitza-Thompsona.

Pozioma linia $\tau_Y = 64\,966,84$ tys. zł pokazuje prawdziwą wartość globalną τ_Y w utworzonej pseudopopulacji. Czarne punkty na każdym z wykresów ukazują z kolei realizacje konkretnego estymatora w każdej z $B = 200$ replikacji. Analiza tego wykresu wskazuje, że każdy z rozpatrywanych estymatorów charakteryzuje się niewielkim obciążeniem. W wypadku estymatora Horwitza-Thompsona



Rysunek 2.4. Oszacowana wartość globalna wydatków ogółem gospodarstw domowych

Źródło: na podstawie BBGD.

można jednak zauważyć większe zróżnicowanie uzyskanych wyników w poszczególnych replikacjach badania symulacyjnego w porównaniu z estymatorami kalibracyjnymi.

Potwierdzeniem powyższych spostrzeżeń są liczby przedstawione w tabeli 2.2. Względne obciążenie wszystkich estymatorów rozpatrywanych w badaniu symulacyjnym jest małe i oscyluje w granicach 0. Spośród wszystkich rozważanych estymatorów najmniejsze względne obciążenie zaobserwowano w wypadku estymatora kalibracyjnego, w którym wagi były wyznaczone zgodnie ze wzorem (2.15)²²,

²² Metoda liniowa.

a największe w odniesieniu do estymatora Horwitza-Thompsona. Do podobnych wniosków można dojść, analizując MSE_{SIM} oraz $RRMSE_{SIM}$ rozpatrywanych estymatorów. Estymatory kalibracyjne charakteryzują się większą dokładnością aniżeli estymator Horwitza-Thompsona, który w swojej konstrukcji nie wykorzystuje żadnej informacji o zmiennych dodatkowych. Oszacowane błędy średniokwadratowe w wypadku estymatorów kalibracyjnych są znacznie mniejsze w porównaniu z estymatorem Horwitza-Thompsona, przy czym najniższe MSE_{SIM} w grupie estymatorów kalibracyjnych zaobserwowano w wypadku estymatora $\hat{\tau}_{lin}$, a najwyższe dla estymatora $\hat{\tau}_{log}$. Względny błąd średniokwadratowy $RRMSE_{SIM}$ dla estymatorów kalibracyjnych był na podobnym poziomie. Większą jego wartość – 2,58% – odnotowano jedynie w wypadku estymatora Horwitza-Thompsona. Z kolei wskaźnik pokrycia wartości globalnej odpowiadający 95% przedziałowi ufności był dla wszystkich rozpatrywanych estymatorów na podobnym poziomie i oscylował w granicach 95–97%. Nieco wyższy CR_{SIM} odnotowano jedynie dla estymatorów $\hat{\tau}_{lin}$ oraz $\hat{\tau}_{rak}$. Zaobserwowane różnice nie są jednak znaczące. Widoczne są natomiast rozbieżności między średnimi długościami przedziałów ufności dla estymatora Horwitza-Thompsona i estymatorów kalibracyjnych. Wartość AL_{SIM} w wypadku estymatora Horwitza-Thompsona była około 35% wyższa aniżeli dla pozostałych estymatorów kalibracyjnych.

Tabela 2.2. Wartości miar jakości estymatorów kalibracyjnych

Estymator	RB_{SIM}	MSE_{SIM}	$RRMSE_{SIM}$	CR_{SIM}	AL_{SIM}
$\hat{\tau}_{HT}$	-0,20	2 809 107,99	2,58	0,950	7 417,04
$\hat{\tau}_{lin}$	-0,07	1 787 312,73	2,06	0,965	5 509,14
$\hat{\tau}_{lin_tr}$	-0,14	1 869 491,24	2,10	0,955	5 503,59
$\hat{\tau}_{log}$	-0,15	1 880 559,93	2,11	0,955	5 506,73
$\hat{\tau}_{rak}$	-0,12	1 814 227,08	2,07	0,965	5 494,40
$\hat{\tau}_{sinh}$	-0,15	1 877 382,08	2,11	0,955	5 508,16

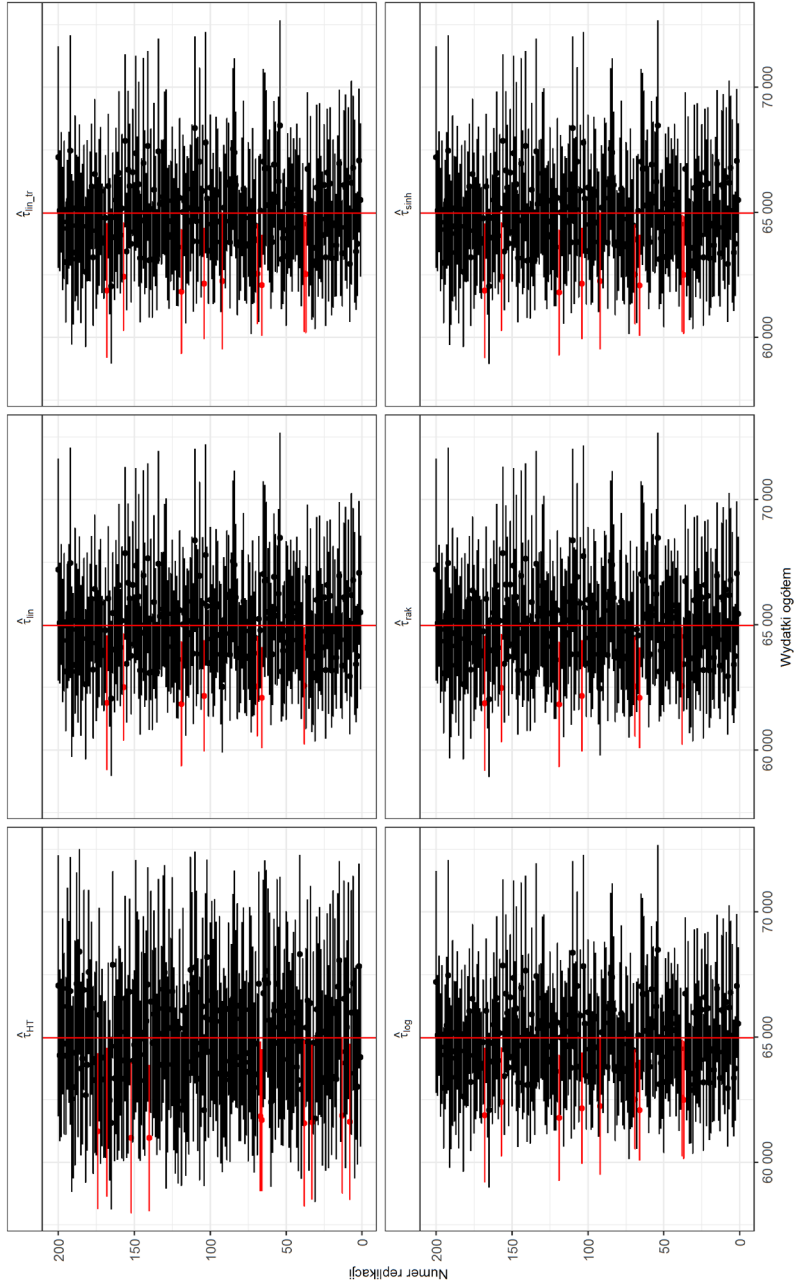
Objaśnienia: Wartości RB_{SIM} oraz $RRMSE_{SIM}$ wyrażone są w procentach.

Źródło: na podstawie BBGD.

Na rysunku 2.5 przedstawiono 95% przedziały ufności wydatków ogółem gospodarstw domowych uzyskanych w każdej z $B = 200$ replikacji badania symulacyjnego i wszystkich rozważanych estymatorów. Kolorem czerwonym zaznaczono te przedziały ufności, które nie pokrywają wartości globalnej wydatków ogółem (czerwona pionowa linia).

Analizując szczegółowo rysunek 2.5 można zauważyć, że w wypadku estymatora Horwitza-Thompsona otrzymujemy znacznie szersze przedziały ufności

Przedział ufności: ◆ Nie pokrywa ◆ Pokrywa



Rysunek 2.5. 95% przedziały ufności wartości globalnej wydatków ogółem gospodarstw domowych

Źródło: na podstawie BBGD.

aniżeli dla pozostałych estymatorów kalibracyjnych. Odpowiada to spostrzeżeniu, że wartość miary AL_{SIM} spośród wszystkich rozpatrywanych estymatorów jest najwyższa dla estymatora Horvitz-Thompsona. Warto również zauważyć, że przedziały ufności, które nie pokrywają wartości globalnej wydatków ogółem, występują z lewej strony „pionowej linii” oznaczającej wartość prawdziwą wydatków ogółem. Ma to związek z zauważoną prawostronną asymetrią analizowanych zmiennych i ich dość silną korelacją. Przedziały takie były konstruowane na podstawie prób, w których znalazły się głównie gospodarstwa domowe o niższych wartościach dochodów i wydatków, przez co koncentrowały się one wokół niewłaściwych wartości.

Odrębną kwestią wymagającą dyskusji jest analiza uzyskanych mnożników g_k i wag kalibracyjnych w_k . W tabeli 2.3 przedstawiono najważniejsze statystyki opisowe mnożników kalibracyjnych g_k oraz wag kalibracyjnych w_k dla wszystkich pięciu rozpatrywanych funkcji $G(\cdot)$.

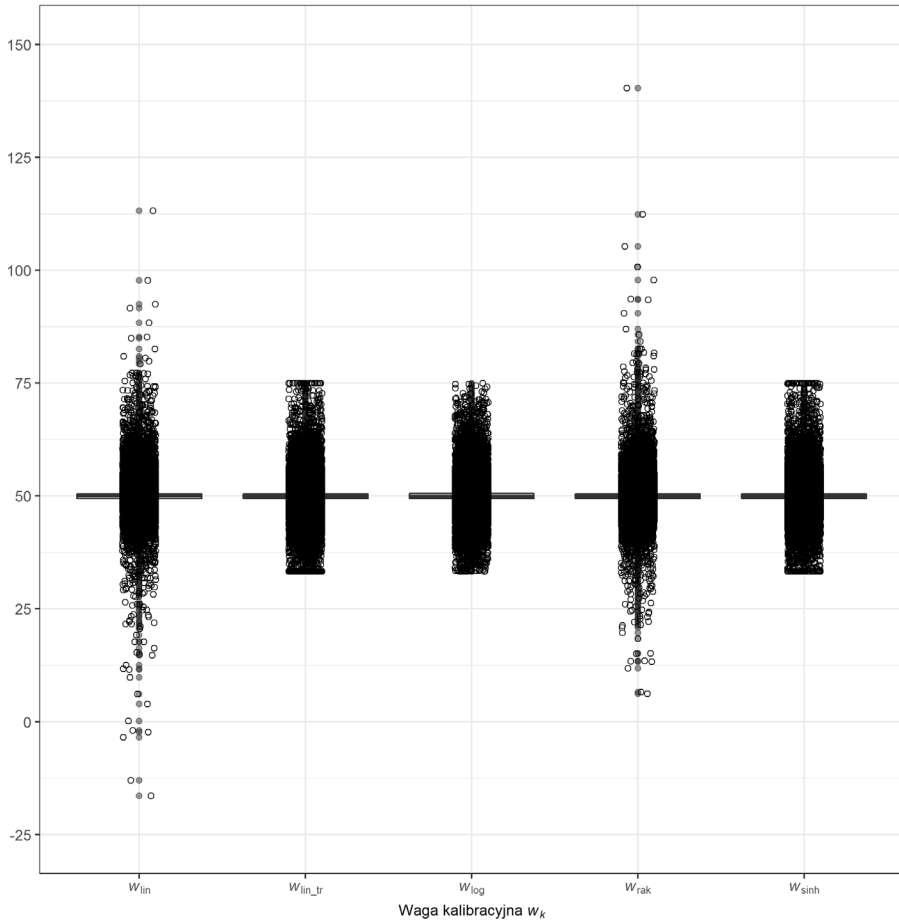
Tabela 2.3. Statystyki opisowe mnożników g_k i wag kalibracyjnych w_k dla różnych funkcji $G(\cdot)$

g	Min	Średnia	Std	Q ₂	Max
g_{lin}	-0,33	1,00	0,04	1,00	2,26
g_{lin_tr}	0,67	1,00	0,04	1,00	1,50
g_{log}	0,67	1,00	0,04	1,00	1,50
g_{rak}	0,12	1,00	0,04	1,00	2,81
g_{sinh}	0,67	1,00	0,04	1,00	1,50
w	Min	Średnia	Std	Q ₂	Max
w_{lin}	-16,47	49,99	1,99	49,99	113,17
w_{lin_tr}	33,32	49,99	2,06	49,99	74,98
w_{log}	33,32	49,99	2,08	49,98	74,96
w_{rak}	6,21	49,99	2,01	49,98	140,33
w_{sinh}	33,32	49,99	2,07	49,98	74,98

Objaśnienia: Std – odchylenie standardowe, Q₂ – kwartył drugi.

Źródło: na podstawie BBGD.

Należy zwrócić uwagę, że dla wszystkich funkcji $G(\cdot)$, w których narzucono ograniczenia na iloraz w_k/d_k , tj. G_3 , G_4 oraz G_5 , wagi w_k znajdują się w przedziale [33,32, 74,98], a odpowiadające im mnożniki kalibracyjne g_k w przedziale postaci [0,67, 1,50]. W wypadku wag w_{lin} w niektórych replikacjach zaobserwowano występowanie wag ujemnych. Jest to wyraźnie widoczne na rysunku 2.6. Stoi to w sprzeczności z definicją wagi, która jest odwrotnością prawdopodobieństwa inkluzji pierwszego rzędu.



Rysunek 2.6. Rozkład wag kalibracyjnych we wszystkich replikacjach

Źródło: na podstawie BBGD.

Mimo że estymatory kalibracyjne wartości globalnej bazujące na metodzie liniowej odznaczały się bardzo dobrymi własnościami (RB_{SIM} , MSE_{SIM} czy AL_{SIM}) w porównaniu z pozostałymi estymatorami kalibracyjnymi, należy mieć na względzie możliwość wystąpienia nieakceptowalnych wag, tj. ujemnych albo ekstremalnych. W takich przypadkach należy bądź narzucić odpowiednie ograniczenia na w_k/d_k , bądź zmienić postać funkcji $G(\cdot)$. Warto także zwrócić uwagę na funkcję G_2 , tj. na wagi w_{rak} , które w niektórych replikacjach znacznie odbiegały od wag d_k wynikających z przyjętego planu losowania próby. Minimalna wartość wagi wynosiła 6,21, podczas gdy wartość maksymalna 140,33 (przy wyjściowych wagach

$d_k = 49,98$). W wypadku funkcji G_2 mamy zatem do czynienia w niektórych replikacjach z występowaniem wag ekstremalnych. Również w takiej sytuacji, mimo że ten estymator kalibracyjny miał spośród innych estymatorów najmniejszą wartość miary AL_{SIM} oraz jedno z najniższych oszacowanych wartości względnego obciążenia i błędu średniokwadratowego, należy zachować szczególną ostrożność w procesie estymacji.

Na zakończenie rozważań nad własnościami wag kalibracyjnych należy podkreślić, że w każdej replikacji zaproponowanego badania symulacyjnego i dla każdej z rozpatrywanych funkcji odległości, wartość globalna zmiennej pomocniczej była dokładnie odtworzona. Sytuacja taka nie zawsze musi zachodzić, zwłaszcza wówczas, gdy dysponujemy dużą liczbą zmiennych pomocniczych oraz narzucamy na wagi kalibracyjne warunek przynależności do z góry określonego przedziału.

2.5. Wnioski

W rozdziale szczegółowo przedstawiono estymatory kalibracyjne wartości globalnej. Omówiono najważniejsze funkcje odległości, które prowadzą do odmiennych zestawów wag kalibracyjnych. Wskazano przy tym na zalety i wady tych funkcji. Rozważania prowadzono w kontekście estymacji wartości globalnej zarówno w populacji generalnej, jak i w odpowiednio zdefiniowanych domenach. Dyskusji poddano również sposób konstruowania wariancji rozważanych estymatorów.

W badaniu symulacyjnym z wykorzystaniem rzeczywistych danych pochodzących z Badania Budżetów Gospodarstw Domowych dokonano oceny najważniejszych miar jakości estymatorów kalibracyjnych wartości globalnej. Pokazano, że charakteryzują się one lepszymi własnościami w porównaniu z często wykorzystywanym w praktyce estymatorem Horvitz-Thompsona. Dotyczy to w zasadzie wszystkich estymatorów kalibracyjnych, bez względu na przyjętą funkcję odległości. Uzyskane wyniki są ponadto zgodne z rezultatami otrzymanymi w odpowiednio zaprojektowanych badaniach symulacyjnych przez innych badaczy (Deville i Särndal, 1992; Kim i Park, 2010; Nguyen, 2010). Uwzględnienie w podejściu kalibracyjnym właściwie dobranych zmiennych pomocniczych poprawia zatem proces estymacji z punktu widzenia najważniejszych własności wykorzystywanych estymatorów.

Odrębną kwestią, którą analizowano w badaniu symulacyjnym, była postać wag kalibracyjnych. Jak pokazano, dobór funkcji odległości ma wpływ na uzyskiwane wagi. W zależności od funkcji odległości mogą się bowiem pojawiać wagi

ujemne lub odstające. Mimo że oszacowanie wartości globalnej zmiennej analizowanej Y może być obciążone niewielkim błędem, należy się zastanowić nad zasadnością użycia takich wag w kontekście innych zmiennych celu. W praktyce bowiem dysponujemy zazwyczaj jednym systemem wag kalibracyjnych, który jest wykorzystywany do szacowania wartości globalnych różnych zmiennych. To, że dany zestaw wag kalibracyjnych przyczynia się do wzrostu precyzji i poprawnego oszacowania wartości globalnej jednej zmiennej, nie musi automatycznie oznaczać takiej sytuacji dla innych zmiennych. Jest to zgodne z najnowszymi wynikami, jakie uzyskali El Haj Tirari i Hdioud (2018).

Kwestia ta nabiera jeszcze większego znaczenia wtedy, gdy dokonujemy oszacowania wartości globalnej danej zmiennej w odpowiednio zdefiniowanych domenach. Występowanie wag ujemnych bądź ekstremalnych (najczęściej znacznie zawyżonych) może bowiem oznaczać, że w przekroju domen uzyska się ujemne oszacowania wartości globalnej analizowanej zmiennej lub jej znaczne przeszacowanie.

Z tego punktu widzenia kluczowa jest analiza końcowych wag kalibracyjnych. Powinna ona uwzględniać nie tylko występowanie wag ujemnych czy ekstremalnych, ale również stopień odtworzenia równań kalibracyjnych. W wypadku gdy nakładamy ograniczenia na wagi i zachodzi potrzeba ich przycięcia, można sobie zagwarantować, że wagi będą należały do z góry zdefiniowanego przedziału, ale niekoniecznie będą one jednocześnie dokładnie odtwarzać równania kalibracyjne dla wszystkich zmiennych pomocniczych.

Podjęcie kalibracyjne w ujęciu funkcyjnym

3.1. Wprowadzenie

W poprzednim rozdziale szczegółowo został opisany sposób wyznaczania wag kalibracyjnych $\mathbf{w} = (w_1, \dots, w_n)^T$, które minimalizują funkcję odległości $D(\cdot)$, przy jednoczesnym spełnieniu odpowiednich równań kalibracyjnych określonych wzorem (2.5). Wyznaczone wagi kalibracyjne mają następujące własności (Estevao i Särndal, 2000; Särndal i Lundström, 2005; Särndal, 2007):

- 1) **spójność systemu wag** – rozumianą w ten sposób, że wagi kalibracyjne spełniają odpowiednie równania kalibracyjne określone wzorem (2.5), tj. dla wszystkich zmiennych pomocniczych wziętych pod uwagę w procesie kalibracji odtwarzane są ich wartości globalne;
- 2) **bliskość** – wagi $\mathbf{d} = (d_1, \dots, d_n)^T$ wynikające z planu losowania próby mają pożądaną własność polegającą na tym, że oszacowania wartości globalnej (1.34) z wykorzystaniem estymatora Horvitz-Thompsona (1.35) są nieobciążone. Jakikolwiek modyfikacje tych wag powinny być niewielkie, aby zachować własność nieobciążoności: system wag $\mathbf{w} = (w_1, \dots, w_n)^T$, który otrzymujemy poprzez rozwiązanie odpowiedniego zadania optymalizacyjnego, ma tę własność;

3) **możliwość kontrolowania wartości globalnych wszystkich zmiennych pomocniczych** – która jest rozumiana w ten sposób, że uwzględnienie w procesie kalibracji większej liczby zmiennych pomocniczych i ich wartości globalnych prowadzi na ogół do lepszych oszacowań, tj. wariancja estymatorów kalibracyjnych wykazuje tendencję malejącą wraz ze wzrostem liczby zmiennych pomocniczych. Można zatem poprzez właściwy dobór zmiennych pomocniczych oraz odpowiadających im wartości globalnych w znacznej mierze kontrolować precyzję oszacowań zarówno na poziomie całej populacji, jak i w odpowiednio zdefiniowanych domenach.

Wagi kalibracyjne (2.12) uzyskane w podejściu minimalizującym odpowiednio dobraną funkcję odległości mają wszystkie trzy wspomniane powyżej własności. Nie jest to jednak jedyny sposób konstruowania różnych postaci estymatorów kalibracyjnych wartości globalnej i odpowiadających im wag w_k . Można również w tym celu wykorzystać tzw. podejście funkcyjne bazujące na koncepcji wektorów zmiennych instrumentalnych (Estevao i Särndal, 2000; 2004; Szymkowiak, 2010). W podejściu tym rezygnuje się z warunku, że wyznaczone wagi kalibracyjne minimalizują funkcję odległości. Zamiast tego proponuje się, aby wagi kalibracyjne spełniały jedynie równanie kalibracyjne (2.5) oraz aby były zapisane w odpowiedniej postaci. Podejście to opiszemy od strony formalnej w dalszej części tego rozdziału.

3.2. Estymatory kalibracyjne wartości globalnej – podejście funkcyjne

Podobnie jak wcześniej, niech $\mathbf{x}_k = (x_{k1}, \dots, x_{kJ})^T$ oznacza wektor złożony z wartości J zmiennych objaśniających dla k -tej jednostki w populacji, $k \in U$, przy czym J oznacza liczbę zmiennych pomocniczych. Zakładamy, że zmiennym tym odpowiada wektor wartości globalnych $\sum_{k \in U} \mathbf{x}_k$. Definiujemy ponadto wektor $\mathbf{z}_k = (z_{k1}, \dots, z_{kJ})^T$ dla $k \in s$ taki, że:

- (a) $\dim(\mathbf{z}_k) = \dim(\mathbf{x}_k) = J$,
- (b) macierz $\sum_{k \in s} c_k \mathbf{z}_k \mathbf{x}_k^T$ o wymiarach $J \times J$ jest nieosobliwa.

Dla tak ustalonego wektora \mathbf{z}_k oraz parametru c_k można zdefiniować wektor wag kalibracyjnych w podejściu funkcyjnym (Estevao i Särndal, 2000). Wykorzystuje on koncepcję tzw. zmiennych instrumentalnych.

Definicja 3.1. Wektorem wag kalibracyjnych w podejściu funkcyjnym nazywamy wektor $\mathbf{w} = (w_1, \dots, w_n)^\top$, którego składowe dla $k = 1, \dots, n$ są postaci:

$$w_k = d_k + c_k \boldsymbol{\lambda}^\top \mathbf{z}_k, \quad (3.1)$$

gdzie parametry c_k oraz \mathbf{z}_k są w taki sposób dobrane, aby spełnione były wspomniane powyżej własności (a) i (b). Wektor \mathbf{z}_k nazywamy wektorem zmiennych instrumentalnych.

Kluczową kwestię stanowi sposób znalezienia wektora $\boldsymbol{\lambda}^\top$. Zakładając, że postać wektora w_k w równaniu (3.1) jest znana, wektor $\boldsymbol{\lambda}^\top$ wyznaczany jest z równania (2.5). Rozwiązując to równanie, otrzymujemy:

$$\boldsymbol{\lambda}^\top = \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right)^\top \left(\sum_{k \in s} c_k \mathbf{z}_k \mathbf{x}_k^\top \right)^{-1}. \quad (3.2)$$

Prowadzi to do następującego wzoru na składową $w_k = w_{k\text{CALF}}$ w podejściu funkcyjnym:

$$w_{k\text{CALF}} = d_k + c_k \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right)^\top \left(\sum_{k \in s} c_k \mathbf{z}_k \mathbf{x}_k^\top \right)^{-1} \mathbf{z}_k. \quad (3.3)$$

Subskrypt CALF we wzorze (3.3) odnosi się do rodziny estymatorów kalibracyjnych uzyskanych z wykorzystaniem podejścia funkcyjnego (ang. *functional calibration*). Rozpatrywany wcześniej estymator typu GREG jest szczególnym przypadkiem estymatorów kalibracyjnych uzyskanych w podejściu funkcyjnym. Wystarczy bowiem we wzorze (3.3) przyjąć $\mathbf{z}_k = \mathbf{x}_k$ oraz podstawić za c_k wyrażenie $q_k d_k$, aby uzyskać postać wag kalibracyjnych w estymatorze typu GREG, który, jak pokazano wcześniej, odpowiada estymatorowi kalibracyjnemu dla liniowej funkcji kalibracyjnej $F(u) = 1 + u$. Nazwa „podejście funkcyjne” jest związana z tym, że wektor \mathbf{z}_k zmiennych instrumentalnych można zazwyczaj przedstawić jako pewną funkcję wektora \mathbf{x}_k zmiennych pomocniczych, tj. $\mathbf{z}_k = f(\mathbf{x}_k)$. Na przykład, biorąc pod uwagę funkcję $f(x) = x^{p-1}$, jednym ze sposobów ustalenia postaci wektora \mathbf{z}_k jest:

$$\mathbf{z}_k = \left(x_{1k}^{p-1}, \dots, x_{Jk}^{p-1} \right)^\top, \quad (3.4)$$

przy założeniu, że $x_{jk} > 0$ dla $j = 1, \dots, J$ oraz $p > 0$. W szczególnym przypadku, gdy $p = 2$, otrzymujemy $\mathbf{z}_k = \mathbf{x}_k$. Innym możliwym sposobem ustalenia wektora \mathbf{z}_k zmiennych instrumentalnych jest przyjęcie, że:

$$\mathbf{z}_k = \left((x_{k1} - \bar{x}_1)^{p-1}, \dots, (x_{kJ} - \bar{x}_J)^{p-1} \right)^T, \quad (3.5)$$

gdzie \bar{x}_j jest średnią ważoną j -tej zmiennej, $j = 1, \dots, J$, obliczoną na podstawie próby s , przy czym funkcję wag pełnią wagi d_k wynikające z planu losowania próby.

Nie każdy wektor \mathbf{z}_k zmiennych instrumentalnych może być jednak wzięty pod uwagę. Na przykład, wektor $\mathbf{z}_k = (1, \dots, 1)^T$ nie będzie spełniał warunku (b), gdyż odpowiednia macierz będzie osobliwa. Wyznaczenie wektora wag kalibracyjnych w podejściu funkcyjnym zgodnie ze wzorem (3.3) wymagać zatem będzie znajomości wektora $\sum_{k \in U} \mathbf{x}_k$ złożonego z wartości globalnych zmiennych pomocniczych, wartości parametru c_k oraz wektora \mathbf{z}_k zmiennych instrumentalnych. Przy odpowiednim ustaleniu wartości tych parametrów można uzyskać różne postacie wag kalibracyjnych w podejściu funkcyjnym, a w konsekwencji różne estymatory kalibracyjne wartości globalnej. Będzie o tym mowa w dalszej części podrozdziału. W pierwszej kolejności zdefiniowany zostanie estymator kalibracyjny wartości globalnej (1.34).

Definicja 3.2. *Estymatorem kalibracyjnym wartości globalnej (1.34) w podejściu funkcyjnym jest:*

$$\hat{\tau}_{\text{CALF}} = \sum_{k \in s} w_{k\text{CALF}} y_k, \quad (3.6)$$

gdzie wagi $w_{k\text{CALF}}$ wyrażają się wzorem (3.3). Estymator ten można przedstawić w następującej postaci:

$$\hat{\tau}_{\text{CALF}} = \sum_{k \in s} d_k y_k + \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right)^T \hat{\mathbf{Q}}, \quad (3.7)$$

przy czym:

$$\hat{\mathbf{Q}} = \left(\sum_{k \in s} c_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in s} c_k \mathbf{z}_k y_k. \quad (3.8)$$

Poniższe twierdzenie pokazuje, że obciążenie estymatora kalibracyjnego $\hat{\tau}_{\text{CALF}}$ wartości globalnej (1.34) jest rzędu $O(n^{-1})$. Mimo że estymatory kalibracyjne wartości globalnej otrzymane z wykorzystaniem podejścia funkcyjnego nie są nieobciążone, ich obciążenie jest zazwyczaj niewielkie, nawet dla prób o małych liczebnościach. Dowód tego twierdzenia można znaleźć w artykule Estevao i Särndala (2000).

Twierdzenie 3.1 (Estevao i Särndal, 2000). *Obciążenie estymatora kalibracyjnego $\hat{\tau}_{\text{CALF}}$ wartości globalnej (1.34) w podejściu funkcyjnym jest postaci:*

$$\begin{aligned} B(\hat{\tau}_{\text{CALF}}) &= E(\hat{\tau}_{\text{CALF}}) - \tau_Y = \\ &= -E \left[\left(\sum_{k \in s} d_k \mathbf{x}_k - \sum_{k \in U} \mathbf{x}_k \right)^T (\hat{\mathbf{Q}} - \mathbf{Q}) \right] = O(n^{-1}), \end{aligned} \quad (3.9)$$

gdzie:

$$\mathbf{Q} = \left(\sum_{k \in U} \pi_k c_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \sum_{k \in U} \pi_k c_k \mathbf{z}_k y_k. \quad (3.10)$$

Estymatory kalibracyjne wartości globalnej (1.34) uzyskane z wykorzystaniem podejścia funkcyjnego tworzą szeroką klasę estymatorów, których szczególnym przypadkiem, jak wspomniano, jest estymator typu GREG. Okazuje się, że przy odpowiednim doborze zmiennych pomocniczych w wektorze zmiennych instrumentalnych \mathbf{z}_k można jednak uzyskać inne powszechnie wykorzystywane i znane w praktyce badań statystycznych estymatory. W dalszej części książki rozważymy kilka takich ważnych przypadków (Estevao i Särndal, 2000; Särndal i Lundström, 2005; Szymkowiak, 2010).

- 1) Załóżmy, że dysponujemy jedną ciągłą zmienną pomocniczą x_k , tj. $\mathbf{x}_k = x_k$. Odpowiadająca jej wartość globalna jest postaci $\sum_{k \in U} x_k$. Przy założeniu, że dla wszystkich $k \in s$ wektor zmiennych instrumentalnych \mathbf{z}_k spełnia warunek $\mathbf{z}_k = 1$, tj. $\dim(\mathbf{x}_k) = \dim(\mathbf{z}_k) = 1$, oraz że $c_k = d_k$, estymator kalibracyjny wartości globalnej bazujący na podejściu funkcyjnym redukuje się do wspomnianego wcześniej estymatora ilorazowego $\hat{\tau}_{\text{RATIO}}$, który wyraża się wzorem (2.40).
- 2) Załóżmy, że podobnie jak wcześniej, dysponujemy jedną ciągłą zmienną pomocniczą x_k , której wartość globalna jest postaci $\sum_{k \in U} x_k$. Przyjmijmy jednak, że wektor \mathbf{x}_k zmiennych pomocniczych jest postaci $\mathbf{x}_k = (1, x_k)^T$. W takiej sytuacji wektor $\sum_{k \in U} \mathbf{x}_k$ złożony z wartości globalnych zmiennych pomocniczych opisany jest jako $\sum_{k \in U} \mathbf{x}_k = (N, \sum_{k \in U} x_k)^T$, gdzie N oznacza liczebność populacji. Przyjmując, że $\mathbf{z}_k = \mathbf{x}_k = (1, x_k)^T$, tj. $\dim(\mathbf{x}_k) = \dim(\mathbf{z}_k) = 2$, oraz $c_k = d_k$, otrzymujemy tzw. estymator regresyjny postaci:

$$\hat{\tau}_{\text{REG}} = N[\bar{y}_s + (\bar{x}_U - \bar{x}_s)\beta_s], \quad (3.11)$$

gdzie

$$\bar{x}_U = \frac{\sum_{k \in U} x_k}{N}, \quad \bar{y}_s = \frac{\sum_{k \in s} d_k y_k}{\sum_{k \in s} d_k}, \quad \bar{x}_s = \frac{\sum_{k \in s} d_k x_k}{\sum_{k \in s} d_k}$$

oraz

$$\beta_s = \frac{\sum_{k \in s} d_k (x_k - \bar{x}_s)(y_k - \bar{y}_s)}{\sum_{k \in s} d_k (x_k - \bar{x}_s)^2}. \quad (3.12)$$

- 3) Zakładamy, że dysponujemy zmienną jakościową, która przyjmuje P wariantów (na przykład płeć, wykształcenie, miejsce zamieszkania itd.). Niech ponadto x oznacza ciągłą zmienną, której wartości są znane dla k -tej jednostki, $k = 1, \dots, n$. Zakładamy ponadto, że wektor $\sum_{k \in U} \mathbf{x}_k$ składa się z P wartości globalnych $\sum_{k \in U_p} x_k$, $p = 1, \dots, P$, tzn. jest następującej postaci: $\sum_{k \in U} \mathbf{x}_k = \left(\sum_{k \in U_1} x_k, \dots, \sum_{k \in U_P} x_k \right)^T$. Definiujemy wektor $\boldsymbol{\gamma}_k$ przynależności jednostki k do odpowiedniej grupy jako $\boldsymbol{\gamma}_k = (\gamma_{1k}, \dots, \gamma_{Pk})^T$, przy czym dla $p = 1, \dots, P$, $\gamma_{pk} = 1$, jeśli jednostka k należy do grupy p , oraz $\gamma_{pk} = 0$ w przeciwnym razie. Wektor $\boldsymbol{\gamma}_k$ ma zatem $P - 1$ składowych, które są zerami, i jeden element równy 1. Wektor zmiennych pomocniczych \mathbf{x}_k określamy jako $\mathbf{x}_k = (\gamma_{1k} x_k, \dots, \gamma_{pk} x_k, \dots, \gamma_{Pk} x_k)^T$. Załóżmy, że $s = \bigcup_{p=1}^P s_p$, gdzie s_p jest częścią próby s o liczebności n_p , do której należą jednostki mające p -ty wariant rozpatrywanej cechy jakościowej. Przyjmując, że wektor zmiennych instrumentalnych $\mathbf{z}_k = \mathbf{x}_k$, a także $c_k = d_k$, otrzymujemy znany estymator ilorazowy (ang. *separate ratio estimator*) będący sumą estymatorów ilorazowych wyznaczonych dla każdej z P wyróżnionych grup:

$$\hat{\tau}_{\text{SEPR}} = \sum_{p=1}^P \left(\sum_{k \in U_p} x_k \right) \frac{\bar{y}_{s_p}}{\bar{x}_{s_p}}, \quad (3.13)$$

gdzie

$$\bar{y}_{s_p} = \frac{\sum_{k \in s_p} d_k y_k}{\sum_{k \in s_p} d_k} \quad \text{oraz} \quad \bar{x}_{s_p} = \frac{\sum_{k \in s_p} d_k x_k}{\sum_{k \in s_p} d_k}.$$

- 4) Podobnie jak w poprzednim punkcie, przyjmijmy, że dysponujemy zmienną jakościową, która przyjmuje P wariantów, oraz pewną zmienną ciągłą x . Załóżmy jednak, że oprócz wartości globalnych $\sum_{k \in U_p} x_k$ znane są również liczebności populacyjne N_p , tj. liczba jednostek w populacji U , które mają p -ty wariant cechy jakościowej, $p = 1, \dots, P$. Zakładamy zatem, że wektor wartości globalnych $\sum_{k \in U} \mathbf{x}_k$ można przedstawić następująco:

$$\sum_{k \in U} \mathbf{x}_k = \left(N_1, \dots, N_P, \sum_{k \in U_1} x_k, \dots, \sum_{k \in U_P} x_k \right)^T,$$

przy czym wektor \mathbf{x}_k można zapisać jako

$$\mathbf{x}_k = (\gamma_{1k}, \dots, \gamma_{pk}, \dots, \gamma_{Pk}, \gamma_{1k}x_k, \dots, \gamma_{pk}x_k, \dots, \gamma_{Pk}x_k)^T.$$

Przyjmując, że wektor zmiennych instrumentalnych $\mathbf{z}_k = \mathbf{x}_k$, a także $c_k = d_k$, otrzymujemy estymator regresyjny $\hat{\tau}_{\text{SEPREG}}$ (ang. *separate regression estimator*) będący sumą estymatorów regresyjnych wyznaczonych dla każdej z P wyróżnionych grup:

$$\hat{\tau}_{\text{SEPREG}} = \sum_{p=1}^P N_p [\bar{y}_{s_p} + (\bar{x}_{U_p} - \bar{x}_{s_p})\beta_{s_p}], \quad (3.14)$$

przy czym β_{s_p} jest zdefiniowane jako:

$$\beta_{s_p} = \frac{\sum_{k \in s_p} d_k (x_k - \bar{x}_{s_p})(y_k - \bar{y}_{s_p})}{\sum_{k \in s_p} d_k (x_k - \bar{x}_{s_p})^2}. \quad (3.15)$$

Poniższe twierdzenie pokazuje postać wariancji estymatora (3.6).

Twierdzenie 3.2 (Estevao i Särndal, 2000). *Wariancja estymatora kalibracyjnego (3.6) wartości globalnej (1.34) otrzymanego w podejściu funkcyjnym wyraża się wzorem:*

$$D^2(\hat{\tau}_{\text{CALF}}) = \sum_{k \in U} \sum_{l \in U} \left(\frac{d_k d_l}{d_{kl}} - 1 \right) e_k^* e_l^*, \quad (3.16)$$

gdzie $e_k^* = y_k - \mathbf{x}_k^T \mathbf{Q}$, przy czym \mathbf{Q} jest zdefiniowane wzorem (3.10).

W powyższych przykładach przyjmowaliśmy zazwyczaj, że $c_k = d_k$. Jest to najbardziej naturalny i najczęściej spotykany sposób ustalania wartości współczynnika c_k , gdyż prowadzi on do różnych znanych postaci estymatorów kalibracyjnych, przy odpowiednim doborze wektora zmiennych instrumentalnych \mathbf{z}_k . Jak pokazują Estevao i Särndal (2000), wybór współczynników c_k ma niewielki wpływ na wariancję estymatora kalibracyjnego $\hat{\tau}_{\text{CALF}}$ uzyskanego w podejściu funkcyjnym, nawet w sytuacji, gdy przypisze im się losowe wartości. W dalszej części będziemy zatem zakładać, że $c_k = d_k$. W takim przypadku wagi (3.1) można zapisać jako $w_k^* = d_k + d_k \boldsymbol{\lambda}^T \mathbf{z}_k = d_k v_k$, gdzie $v_k = 1 + \boldsymbol{\lambda}^T \mathbf{z}_k$.

Do oceny wariancji estymatora (3.16), przyjmując, że $c_k = d_k$, można wykozystać estymator postaci (Kott, 2006):

$$\hat{D}^2(\hat{\tau}_{\text{CALF}}) = \sum_{k \in s} \sum_{l \in s} \frac{\Delta_{kl}}{\pi_{kl}} (w_k^* \tilde{e}_k^*) (w_l^* \tilde{e}_l^*), \quad (3.17)$$

przy czym $\tilde{e}_k^* = y_k - \mathbf{x}_k^T \tilde{\boldsymbol{\beta}}_s^*$ oraz

$$\tilde{\boldsymbol{\beta}}_s^* = \left(\sum_{k \in s} d_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in s} d_k \mathbf{z}_k y_k \right). \quad (3.18)$$

Do oceny wariancji estymatora (3.16), przyjmując, że $c_k = d_k$, można również użyć estymatora zaproponowanego w pracy Särndala i Lundströma (2005) w postaci:

$$\begin{aligned} \hat{D}^2(\hat{\tau}_{\text{CALF}}) = & \sum_{k \in s} \sum_{l \in s} (d_k d_l - d_{kl}) (v_k \hat{e}_k^*) (v_l \hat{e}_l^*) + \\ & - \sum_{k \in s} d_k (d_k - 1) v_k (v_k - 1) (\hat{e}_k^*)^2, \end{aligned} \quad (3.19)$$

gdzie $\hat{e}_k^* = y_k - \mathbf{x}_k^T \hat{\boldsymbol{\beta}}_s^*$, przy czym

$$\hat{\boldsymbol{\beta}}_s^* = \left(\sum_{k \in s} w_k^* \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in s} w_k^* \mathbf{z}_k y_k \right). \quad (3.20)$$

W podejściu funkcyjnym wagi kalibracyjne, jak wcześniej wspomniano, najczęściej przedstawia się jako $w_k^* = d_k + d_k \boldsymbol{\lambda}_*^T \mathbf{z}_k$. Można jednak rozpatrywać pewne uogólnienia tych wag i zapisać je w postaci (Kott, 2006):

$$w_{k\text{GEN}}^* = d_k f(\boldsymbol{\lambda}_*^T \mathbf{z}_k) = d_k f(\mathbf{z}_k^T \boldsymbol{\lambda}_*), \quad (3.21)$$

gdzie f jest monotoniczną funkcją, dwukrotnie różniczkowalną i taką, że $f(0) = 1$ oraz $f'(0) = 1$, przy czym wektor $\boldsymbol{\lambda}_*^T$ jest w taki sposób dobrany, aby spełniony został układ równań kalibracyjnych (2.5). W szczególnym przypadku można założyć, że $f(u) = 1 + u$ i – przyjmując $u = \boldsymbol{\lambda}_*^T \mathbf{z}_k$ – otrzymać znaną już postać wag w_k^* , gdzie $\boldsymbol{\lambda}_*^T$ zdefiniowany jest wzorem (3.2). Przykładem nieliniowej funkcji f , którą można wykorzystać w procesie wyznaczania wag kalibracyjnych w podejściu funkcyjnym, jest $f(\boldsymbol{\lambda}_*^T \mathbf{z}_k) = e^{\boldsymbol{\lambda}_*^T \mathbf{z}_k}$. W tym jednak przypadku wyznaczenie wektora $\boldsymbol{\lambda}_*^T$ wymagać będzie zastosowania podejścia iteracyjnego zgodnie z opisanym poniżej algorytmem.

Algorytm 3. Algorytm wyznaczania mnożników λ_* w podejściu funkcyjnym

Krok 1: *Inicjalizacja*

Przyjmij wejsiowy wektor mnożników Lagrange'a w postaci $\lambda_*^{(0)} = \mathbf{0}$. Oblicz początkowe wartości wektora wag kalibracyjnych $w_{k\text{GEN}}^{*(0)} = d_k f(0) = d_k$. Przejdź do kroku 2.

Krok 2: *Obliczenie mnożnika $\lambda_*^{(1)}$*

Oblicz wartość mnożnika Lagrange'a $\lambda_*^{(1)}$, korzystając z następującego wzoru:

$$\lambda_*^{(1)} = \left(\sum_{k \in s} d_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} d_k \mathbf{x}_k \right), \quad (3.22)$$

oraz odpowiadające mu wagi kalibracyjne:

$$w_{k\text{GEN}}^{*(1)} = d_k f \left(\mathbf{z}_k^T \lambda_*^{(1)} \right). \quad (3.23)$$

Ustal $l = 1$. Przejdź do kroku 3.

Krok 3: *Kryterium stopu*

Jeżeli spełnione jest równanie kalibracyjne:

$$\sum_{k \in U} \mathbf{x}_k = \sum_{k \in s} w_{k\text{GEN}}^{*(l)} \mathbf{x}_k, \quad (3.24)$$

to zakończ algorytm poszukiwania wag kalibracyjnych. W przeciwnym razie idź do kroku 4.

Krok 4: *Obliczenie wektora $\lambda_*^{(l+1)}$ oraz wag kalibracyjnych $w_{k\text{GEN}}^{*(l+1)}$*

Ustaw l na $l + 1$. Oblicz wektor $\lambda_*^{(l+1)}$ mnożników Lagrange'a, korzystając ze wzoru:

$$\lambda_*^{(l+1)} = \lambda_*^{(l)} + \left(\sum_{k \in s} f' \left(\mathbf{z}_k^T \lambda_*^{(l)} \right) d_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} w_{k\text{GEN}}^{*(l)} \mathbf{x}_k \right), \quad (3.25)$$

a następnie wyznacz nowe wagi kalibracyjne:

$$w_{k\text{GEN}}^{*(l+1)} = d_k f \left(\mathbf{z}_k^T \lambda_*^{(l+1)} \right). \quad (3.26)$$

Wróć do kroku 3.

Powyżej opisany algorytm wyznaczania wag kalibracyjnych w podejściu funkcyjnym dla odpowiednio zdefiniowanej funkcji $f(\cdot)$ należy tak długo kontynuować, aż spełnione będzie równanie kalibracyjne (3.24). W praktyce algorytm ten można zakończyć, gdy spełniony będzie warunek postaci:

$$\max_{1 \leq j \leq J} \left| \sum_{k \in U} \mathbf{x}_k - \sum_{k \in s} w_{k\text{GEN}}^{*(l)} \mathbf{x}_k \right| \leq \varepsilon, \quad (3.27)$$

dla odpowiednio ustalonej wartości parametru ε .

W podejściu funkcyjnym również można dokonać oszacowania wartości globalnej zmiennej Y w domenach. Zakładamy zatem, podobnie jak wcześniej, że populacja U składa się z D rozłącznych domen postaci U_1, \dots, U_D . Przyjmujemy ponadto, że w każdej z D wyróżnionych domen przedmiotem zainteresowania jest wartość globalna $\tau_{Y_1}, \dots, \tau_{Y_D}$, gdzie τ_{Y_d} dla $d = 1, \dots, D$ wyraża się wzorem (1.58). Zakładamy także, że wektor wartości globalnych zmiennych pomocniczych $\sum_{k \in U_d} \mathbf{x}_d$ jest znany na poziomie domeny U_d .

W podejściu tym wagi kalibracyjne w domenie d , dla $d = 1, \dots, D$, można zatem wyrazić wzorem:

$$w_{dk} = d_k + c_k \boldsymbol{\lambda}_d^T \mathbf{z}_k, \quad (3.28)$$

gdzie parametry c_k oraz \mathbf{z}_k są w taki sposób dobrane, aby $\dim(\mathbf{z}_k) = \dim(\mathbf{x}_k) = J$ oraz macierz $\sum_{k \in s_d} c_k \mathbf{z}_k \mathbf{x}_k^T$ o wymiarach $J \times J$ była nieosobliwa.

Wektor $\boldsymbol{\lambda}_d$ występujący we wzorze (3.28) można przedstawić w następujący sposób:

$$\boldsymbol{\lambda}_d^T = \left(\sum_{k \in U_d} \mathbf{x}_k - \sum_{k \in s_d} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in s_d} c_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1}. \quad (3.29)$$

Prowadzi to do następującego wzoru na składową $w_{dk} = w_{dk\text{CALF}}$ w podejściu funkcyjnym:

$$w_{dk\text{CALF}} = d_k + c_k \left(\sum_{k \in U_d} \mathbf{x}_k - \sum_{k \in s_d} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in s_d} c_k \mathbf{z}_k \mathbf{x}_k^T \right)^{-1} \mathbf{z}_k. \quad (3.30)$$

Mając wyznaczone wagi kalibracyjne $w_{dk\text{CALF}}$ uzyskane w podejściu funkcyjnym, można zdefiniować estymator kalibracyjny wartości globalnej (1.58) w domenie d .

Definicja 3.3. Estymatorem kalibracyjnym wartości globalnej (1.58) w domenie d , dla $d = 1, \dots, D$, w podejściu funkcyjnym jest:

$$\hat{\tau}_{\text{CALF},d} = \sum_{k \in s_d} w_{dk_{\text{CALF}}} y_k, \quad (3.31)$$

gdzie wagi $w_{dk_{\text{CALF}}}$ wyrażają się wzorem (3.30). Estymator ten można przedstawić w następującej postaci:

$$\hat{\tau}_{\text{CALF},d} = \sum_{k \in s_d} d_k y_k + \left(\sum_{k \in U_d} \mathbf{x}_k - \sum_{k \in s_d} d_k \mathbf{x}_k \right)^{\text{T}} \hat{\mathbf{Q}}_d, \quad (3.32)$$

przy czym:

$$\hat{\mathbf{Q}}_d = \left(\sum_{k \in s_d} c_k \mathbf{z}_k \mathbf{x}_k^{\text{T}} \right)^{-1} \sum_{k \in s_d} c_k \mathbf{z}_k y_k. \quad (3.33)$$

Podobnie jak wcześniej, najczęściej zakłada się, że $c_k = d_k$, i w takiej sytuacji, wagi (3.28) można zapisać jako $w_{dk}^* = d_k + d_k \boldsymbol{\lambda}_d^{\text{T}} \mathbf{z}_k = d_k v_{dk}$, gdzie $v_{dk} = 1 + \boldsymbol{\lambda}_d^{\text{T}} \mathbf{z}_k$. Wówczas do oceny wariancji estymatora (3.31) można na przykład wykorzystać estymator postaci:

$$\hat{D}^2(\hat{\tau}_{\text{CALF},d}) = \sum_{k \in s_d} \sum_{l \in s_d} \frac{\Delta_{kl}}{\pi_{kl}} (w_{dk}^* \tilde{e}_{dk}^*) (w_{dl}^* \tilde{e}_{dl}^*), \quad (3.34)$$

przy czym $\tilde{e}_{dk}^* = y_{dk} - \mathbf{x}_k^{\text{T}} \tilde{\boldsymbol{\beta}}_d^*$ oraz

$$\tilde{\boldsymbol{\beta}}_d^* = \left(\sum_{k \in s_d} d_k \mathbf{z}_k \mathbf{x}_k^{\text{T}} \right)^{-1} \left(\sum_{k \in s_d} d_k \mathbf{z}_k y_k \right). \quad (3.35)$$

3.3. Badanie symulacyjne – podejście funkcyjne

Głównym celem przeprowadzonego badania symulacyjnego była analiza najważniejszych własności estymatorów kalibracyjnych wartości globalnej (1.34) uzyskanych w podejściu funkcyjnym, opisanym w podrozdziale 3.2. Analogicznie do przypadku badania symulacyjnego dotyczącego podejścia minimalizującego odpowiednią funkcję odległości, wykorzystano rzeczywiste dane pocho-

dzące z BBGD. Pseudopopulację utworzono zgodnie z jej opisem przedstawionym w podrozdziale 2.4. Jako zmienną objaśnianą Y w badaniu symulacyjnym wzięto ponownie pod uwagę wydatki ogółem (w tys. zł) gospodarstwa domowego. Rozszerzono jednak listę zmiennych pomocniczych w podejściu funkcyjnym do trzech:

- x_1 – dochód rozporządzalny gospodarstwa domowego (w tys. zł),
- x_2 – powierzchnia użytkowa mieszkania zajmowanego przez gospodarstwo domowe (w m²),
- x_3 – liczba osób w gospodarstwie domowym.

Ze względu na to, że zarówno zmienna x_2 , jak i x_3 były kompletne, tj. pozbawione braków danych, liczebność pseudopopulacji wynosiła, podobnie jak w wypadku badania symulacyjnego nad własnościami estymatorów kalibracyjnych minimalizujących funkcję odległości, $N = 32\,291$ gospodarstw domowych.

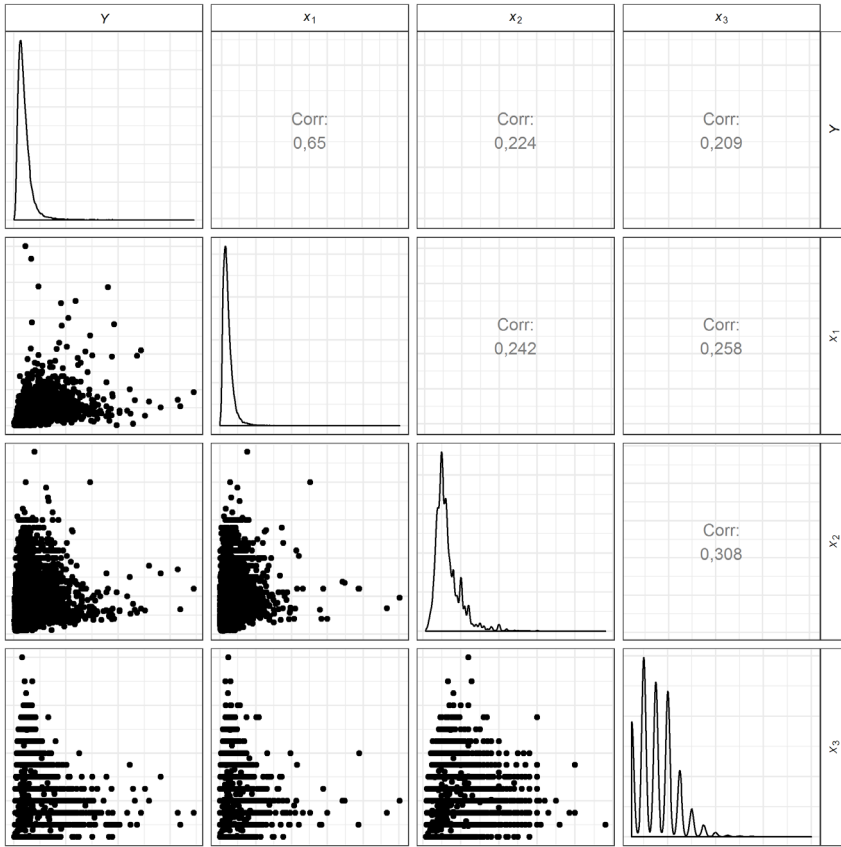
Na rysunku 3.1 zostały przedstawione korelacyjne diagramy rozrzutu pomiędzy wszystkimi zmiennymi wziętymi pod uwagę w badaniu symulacyjnym, oceny współczynników korelacji liniowej Pearsona oraz rozkłady rozważanych zmiennych.

Wszystkie zmienne charakteryzują się rozkładami o asymetrii prawostronnej. Zmienne są ze sobą skorelowane dodatnio, przy czym najsilniejszy związek obserwuje się w wypadku wydatków ogółem i dochodu rozporządzalnego gospodarstw domowych. Oszacowany współczynnik korelacji liniowej Pearsona między tymi zmiennymi wynosił 0,65.

Oszacowania wartości globalnej wydatków ogółem gospodarstw domowych dokonano z wykorzystaniem estymatora kalibracyjnego (3.7), przy czym wagi kalibracyjne w_{kCALF} obliczono, korzystając ze wzoru (3.3). Badanie symulacyjne zostało zrealizowane w dwóch wariantach, których założenia opisano poniżej.

Badanie symulacyjne 1 zostało zrealizowane przy założeniu, że $\dim(\mathbf{x}_k) = \dim(\mathbf{z}_k) = 1$, oraz przyjęciu następujących ustaleń co do zmiennych pomocniczych i wartości parametrów:

- (i) **Zmienne pomocnicze:** $\mathbf{x}_k = x_{1k}$ (jedna zmienna pomocnicza), przy czym za tę zmienną przyjęto dochód rozporządzalny gospodarstwa domowego (x_1). Oznacza to, że jedyną wartością globalną wziętą pod uwagę w procesie kalibracji była $\sum_{k \in U} x_{1k}$.
- (ii) **Wartości parametrów:** ustalono, że wektor \mathbf{z}_k zmiennych instrumentalnych jest postaci $\mathbf{z}_k = x_{1k}^{p-1}$ dla $p = 1, 2, \dots$. Przyjęto ponadto dwa warianty parametru c_k , tj. (1) $c_k = d_k$ oraz (2) $c_k \sim U(0, 1)$. W pierwszym przypadku założono zatem, że parametr c_k równy jest wadze d_k wynikającej z planu losowania próby, a w drugim, że został on wygenerowany z rozkładu jed-



Rysunek 3.1. Rozkłady zmiennych w badaniu symulacyjnym nad własnościami estymatorów kalibracyjnych wartości globalnej – podejście funkcyjne

Źródło: na podstawie BBGD.

nostajnego na przedziale $(0, 1)^{23}$. Przyjęcie wartości parametru $c_k = d_k$ jest często spotykane w praktyce. Z kolei założenie, że $c_k \sim U(0, 1)$, ma na celu pokazanie, że nawet losowe przypisanie wartości parametru c_k nie wpływa w istotny sposób na jakość estymacji (wariancję czy obciążenie estymatorów kalibracyjnych w podejściu funkcyjnym).

²³ Wybór parametrów przyjęto jak w pracy Estevao i Särndala (2000). W przeciwieństwie jednak do badania symulacyjnego nad własnościami estymatorów kalibracyjnych w podejściu funkcyjnym, zaproponowanego przez Estevao i Särndala (2000), wykorzystano rzeczywiste dane z BBGD. Estevao i Särndal (2000) w badaniu symulacyjnym dysponowali sztucznymi danymi wygenerowanymi z odpowiednio dobranego modelu. Prezentowane w książce wyniki badań nad podejściem funkcyjnym w kalibracji odpowiadać mają bardziej realnym problemom, jakie towarzyszą analizie danych w praktyce badań reprezentacyjnych.

Badanie symulacyjne 2 zostało z kolei zrealizowane przy założeniu, że $\dim(\mathbf{x}_k) = \dim(\mathbf{z}_k) = 4$, oraz przyjęciu następujących ustaleń co do zmiennych pomocniczych i wartości parametrów:

- (i) **Zmienne pomocnicze:** wektor $\mathbf{x}_k = (1, x_{1k}, x_{2k}, x_{3k})^T$, a jako zmienne pomocnicze uwzględniono w badaniu wszystkie trzy wspomniane wcześniej zmienne x_1, x_2, x_3 oraz 1. Przyjęcie zmiennej pomocniczej tożsamościowo równej 1 dla wszystkich jednostek miało zapewnić sumowalność wag kalibracyjnych do liczebności pseudopopulacji N . Oznacza to, że wektor wartości globalnych zmiennych pomocniczych w procesie kalibracji był postaci $(N, \sum_{k \in U} x_{1k}, \sum_{k \in U} x_{2k}, \sum_{k \in U} x_{3k})^T$.
- (ii) **Wartości parametrów:** ustalono, że wektor \mathbf{z}_k zmiennych instrumentalnych jest postaci $\mathbf{z}_k = (1, x_{1k}^{p-1}, x_{2k}^{p-1}, x_{3k}^{p-1})^T$, dla $p = 2, 3, \dots$. Pominięto przy tym przypadek $p = 1$ z racji tego, że wówczas macierz $\sum_{k \in s} c_k \mathbf{z}_k \mathbf{x}_k^T$ jest osobliwa. Przyjęto ponadto, jak w pierwszym badaniu symulacyjnym, dwa warianty dla parametru c_k , tj. (1) $c_k = d_k$ oraz (2) $c_k \sim U(0, 1)$.

W dalszej części, podobnie jak w wypadku badania dotyczącego własności estymatorów kalibracyjnych wartości globalnej w podejściu minimalizującym odpowiednią funkcję odległości, z pseudopopulacji losowano 2% próbę zgodnie z planem losowania prostego bez zwracania oraz dokonywano kalibracji wag d_k . Proces ten powtórzono $B = 200$ razy w obydwu opisanych powyżej badaniach symulacyjnych. W każdej replikacji, mając wyznaczone wagi kalibracyjne $w_{k\text{CALF}}$, dokonano oszacowania wartości globalnej τ_Y , tj. łącznych wydatków ogółem gospodarstw domowych, korzystając ze wzoru (3.7).

W celu oceny jakości rozważanych estymatorów kalibracyjnych w podejściu funkcyjnym w obydwu wariantach zaprojektowanego badania symulacyjnego obliczono następujące miary:

- względnego obciążenia:

$$\text{RB}_{\text{SIM}}(\hat{\tau}_{\text{CALF}}) = \frac{1}{B} \sum_{b=1}^B \frac{\hat{\tau}_{\text{CALF}}^{(b)} - \tau_Y}{\tau_Y} \cdot 100\%, \quad (3.36)$$

- błędu średniokwadratowego:

$$\text{MSE}_{\text{SIM}}(\hat{\tau}_{\text{CALF}}) = \frac{1}{B} \sum_{b=1}^B \left(\hat{\tau}_{\text{CALF}}^{(b)} - \tau_Y \right)^2, \quad (3.37)$$

- względnego błędu średniokwadratowego:

$$\text{RRMSE}_{\text{SIM}}(\hat{\tau}_{\text{CALF}}) = \frac{\sqrt{\text{MSE}_{\text{SIM}}(\hat{\tau}_{\text{CALF}})}}{\tau_Y} \cdot 100\%, \quad (3.38)$$

– wariancji empirycznej:

$$D_{\text{SIM}}^2(\hat{\tau}_{\text{CALF}}) = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\tau}_{\text{CALF}}^{(b)} - \bar{\tau}_{\text{CALF}} \right)^2, \quad (3.39)$$

– względnego średniego błędu szacunku:

$$\text{REE}_{\text{SIM}}(\hat{\tau}_{\text{CALF}}) = \frac{\sqrt{D_{\text{SIM}}^2(\hat{\tau}_{\text{CALF}})}}{\hat{\tau}_{\text{CALF}}} \cdot 100\%, \quad (3.40)$$

gdzie τ_Y oznacza prawdziwą wartość łącznych wydatków ogółem gospodarstw domowych w utworzonej pseudopopulacji, a

$$\bar{\tau}_{\text{CALF}} = \frac{1}{B} \sum_{b=1}^B \hat{\tau}_{\text{CALF}}^{(b)}.$$

W tabeli 3.1 przedstawiono najważniejsze miary oceny jakości estymatorów kalibracyjnych wartości globalnej w podejściu funkcyjnym. W tym celu skorzystano ze wzorów (3.36)–(3.40). W sytuacji zmiennych pomocniczych 1, x_1 , x_2 , x_3 nie wzięto pod uwagę przypadku $p = 1$, gdyż wówczas wektor \mathbf{z}_k jest postaci $\mathbf{z}_k = (1, 1, 1, 1)^T$, a macierz $\sum_{k \in s} c_k \mathbf{z}_k \mathbf{x}_k^T$ jest osobliwa. Zgodnie z opisem założeń badania symulacyjnego rozważono dwa sposoby konstrukcji parametru c_k . Warto podkreślić, że w wypadku $p = 1$, $c_k = d_k$ oraz jednej zmiennej pomocniczej x_1 estymator kalibracyjny wartości globalnej w podejściu funkcyjnym redukuje się do znanego estymatora ilorazowego $\hat{\tau}_{\text{RATIO}}$, który wyraża się wzorem (2.40). Z kolei w sytuacji gdy $p = 2$ oraz $c_k = d_k$, estymator ten sprowadza się do uogólnionego estymatora regresyjnego typu GREG (1.43), przy przyjęciu założenia, że $q_k = 1$.

W analizie uzyskanych wyników estymacji należy wziąć pod uwagę kilka elementów. Po pierwsze należy zauważyć, że nie ma większego znaczenia sposób ustalenia wartości parametru c_k . Bez względu na to, czy w badaniu symulacyjnym założono, że $c_k = d_k$, czy przyjęto, że $c_k \sim U(0,1)$, oszacowana w badaniu symulacyjnym wariancja empiryczna $D_{\text{SIM}}^2(\hat{\tau}_{\text{CALF}})$ lub błąd średniokwadratowy $\text{MSE}_{\text{SIM}}(\hat{\tau}_{\text{CALF}})$, przy ustalonym wektorze zmiennych pomocniczych, są na podobnym poziomie. W zdecydowanej większości wariancja oraz błąd średniokwadratowy są większe przy losowym doborze parametru c_k aniżeli przy jego ustaleniu jako $c_k = d_k$. Różnice są jednak na tyle niewielkie, że można przyjąć, że w ocenie jakości najważniejszych własności estymatorów kalibracyjnych wartości globalnej w podejściu funkcyjnym wpływ parametru c_k jest nieistotny.

Tabela 3.1. Ocena jakości estymatorów kalibracyjnych wartości globalnych w podejściu funkcyjnym

c_k	P	Badanie symulacyjne 1: Zmienna pomocnicza x_1					Badanie symulacyjne 2: Zmienne pomocnicze (x_1, x_2, x_3)				
		RB _{SIM}	D ² _{SIM}	REE _{SIM}	MSE _{SIM}	RRMSE _{SIM}	RB _{SIM}	D ² _{SIM}	REE _{SIM}	MSE _{SIM}	RRMSE _{SIM}
d_k	1	-0,18	2 435 974,89	2,41	2 449 800,99	2,41	—	—	—	—	—
	2	-0,11	2 065 426,53	2,21	2 070 223,24	2,21	-0,05	1 744 367,45	2,03	1 745 620,38	2,03
	3	0,00	1 860 360,18	2,10	1 860 369,63	2,10	-0,01	1 744 622,80	2,03	1 744 678,16	2,03
	4	0,01	1 852 589,93	2,09	1 852 645,84	2,10	-0,02	1 786 093,55	2,06	1 786 214,90	2,06
	5	-0,01	1 874 783,15	2,11	1 874 807,06	2,11	-0,04	1 834 538,73	2,09	1 835 176,61	2,09
	6	-0,03	1 895 595,24	2,12	1 895 909,94	2,12	-0,06	1 883 944,61	2,11	1 885 428,27	2,11
7	-0,04	1 911 521,62	2,13	1 912 258,81	2,13	-0,07	1 932 731,78	2,14	1 935 057,59	2,14	
20	-0,08	1 969 143,58	2,16	1 971 805,35	2,16	-0,25	2 585 696,68	2,48	2 612 196,39	2,49	
30	-0,08	1 978 702,74	2,17	1 981 345,69	2,17	-0,17	2 818 262,36	2,59	2 830 554,75	2,59	
40	-0,08	1 981 491,30	2,17	1 983 993,36	2,17	-0,19	2 785 877,32	2,57	2 801 189,78	2,58	
50	-0,08	1 980 543,73	2,17	1 982 918,82	2,17	-0,19	2 799 755,06	2,58	2 814 667,64	2,58	
60	-0,07	1 978 148,24	2,17	1 980 422,47	2,17	-0,18	2 807 561,90	2,58	2 821 583,55	2,59	
$U(0,1)$	1	-0,18	2 445 103,00	2,41	2 458 309,61	2,41	—	—	—	—	—
	2	-0,10	2 087 434,25	2,23	2 091 866,91	2,23	-0,05	1 773 226,41	2,05	1 774 481,19	2,05
	3	-0,02	1 892 192,25	2,12	1 892 299,46	2,12	-0,02	1 761 567,81	2,04	1 761 678,07	2,04
	4	0,01	1 861 311,75	2,10	1 861 329,25	2,10	-0,02	1 795 007,50	2,06	1 795 138,03	2,06
	5	-0,01	1 876 809,29	2,11	1 876 826,05	2,11	-0,04	1 840 284,28	2,09	1 840 867,25	2,09
	6	-0,02	1 893 639,53	2,12	1 893 868,75	2,12	-0,06	1 886 484,28	2,12	1 887 866,55	2,11
7	-0,04	1 905 086,59	2,13	1 905 624,15	2,12	-0,07	1 930 865,93	2,14	1 933 083,67	2,14	
20	-0,07	1 940 475,22	2,15	1 942 337,25	2,15	-0,24	2 986 831,40	2,67	3 011 147,77	2,67	
30	-0,07	1 945 177,76	2,15	1 947 334,45	2,15	-0,21	2 779 777,38	2,57	2 797 998,15	2,57	
40	-0,07	1 947 106,18	2,15	1 949 329,20	2,15	-0,20	2 790 440,49	2,58	2 806 940,39	2,58	
50	-0,07	1 947 357,28	2,15	1 949 566,46	2,15	-0,20	2 802 009,78	2,58	2 818 105,55	2,58	
60	-0,07	1 946 951,64	2,15	1 949 129,54	2,15	-0,19	2 811 420,42	2,59	2 826 475,74	2,59	

Objaśnienia: Wartości RB_{SIM}, REE_{SIM} oraz RRMSE_{SIM} wyrażone są w procentach.

Źródło: na podstawie BBGD.

Po drugie, analizując otrzymane wyniki, można zaobserwować, że bez względu na przyjętą wartość parametru p wyznaczone oszacowania charakteryzują się akceptowalną precyzją. Świadczą o tym niskie wartości względnego średniego błędu szacunku $REE_{SIM}(\hat{\tau}_{CALF})$. Odnosi się to w zasadzie także do względnego błędu średniokwadratowego $RRMSE_{SIM}(\hat{\tau}_{CALF})$.

Można przy tym zauważyć, że błędy te wzrastają stopniowo wraz ze wzrostem wartości parametru p . Sytuacja ta jest szczególnie widoczna w wypadku czterech zmiennych pomocniczych, tj. w drugim badaniu symulacyjnym. Dla jednej zmiennej pomocniczej wzrost błędów jest obserwowalny, począwszy od $p = 3$, przy czym zaczyna się on stabilizować dla $p \geq 30$.

Po trzecie można zauważyć, że przyjęcie w badaniu symulacyjnym większej liczby zmiennych pomocniczych, dla których spełnione mają być odpowiednie równania kalibracyjne, wpływa na redukcję wariancji oraz błędu średniokwadratowego estymatorów kalibracyjnych wartości globalnej w podejściu funkcyjnym dla niższych wartości parametru p , tj. dla $p \leq 6$. W wypadku większych wartości p uwzględnienie większej liczby zmiennych pomocniczych pogarsza nieco proces estymacji z punktu widzenia precyzji oszacowań.

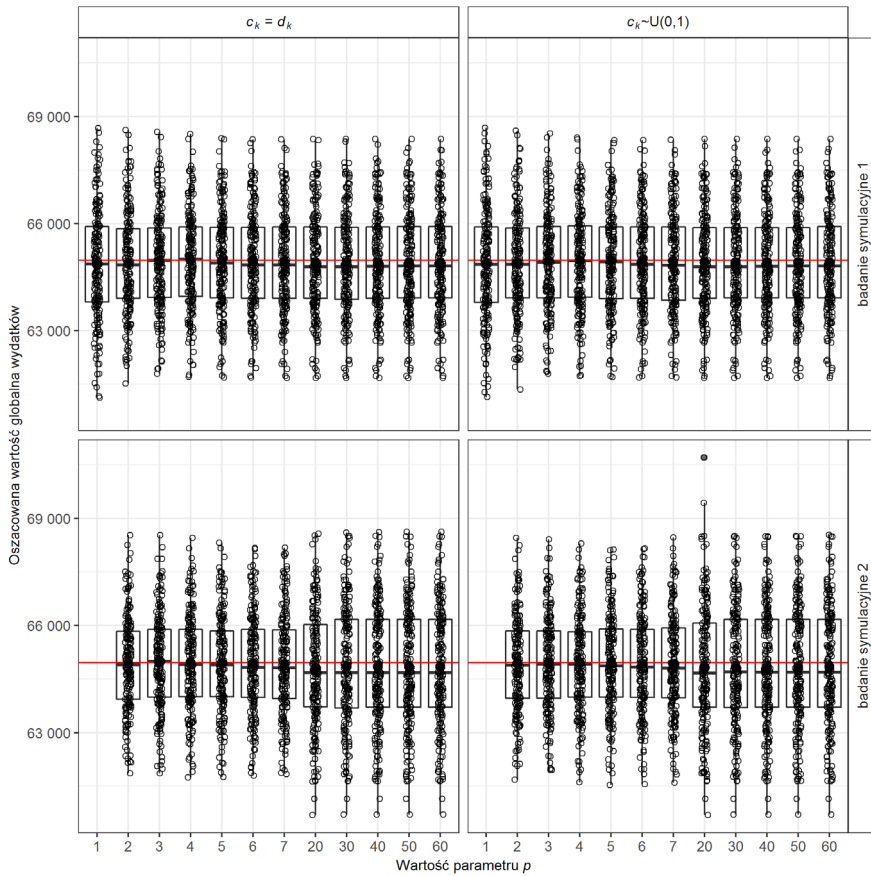
Warto również zwrócić uwagę, że bez względu na wartość parametrów p i c_k oraz liczbę zmiennych pomocniczych, względne obciążenie $RB_{SIM}(\hat{\tau}_{CALF})$ estymatorów kalibracyjnych w podejściu funkcyjnym oscyloowało wokół zera. Potwierdzeniem tego są wyniki zaprezentowane na rysunku 3.2.

Podobnie jak w wypadku podejścia minimalizującego odpowiednią funkcję odległości, pozioma linia $\tau_Y = 64\,966,84$ tys. zł pokazuje prawdziwą wartość globalną Y w pseudopopulacji utworzonej na potrzeby badania symulacyjnego. Punkty z kolei obrazują realizacje konkretnego estymatora w każdej z $B = 200$ replikacji (dla odpowiedniej wartości parametru p i współczynnika c_k).

Odrębną kwestią wymagającą wyjaśnienia jest postać wag kalibracyjnych uzyskanych w podejściu funkcyjnym. Na rysunku 3.3 przedstawiono ich rozkład w zależności od wartości parametru p oraz współczynnika c_k . Można zaobserwować ciekawą zależność, że wraz ze wzrostem wartości współczynnika p rośnie dyspersja wag kalibracyjnych w podejściu funkcyjnym²⁴.

Występują również wagi kalibracyjne ujemne oraz ekstremalne. Jest to widoczne zarówno przy przyjęciu współczynnika $c_k = d_k$, jak i przy losowym ustaleniu wartości tego parametru. Dotyczy to w zasadzie przypadku, w którym wzięto w procesie kalibracji jedną zmienną pomocniczą x_1 , a także, gdy wzięto cztery

²⁴ Wyjściowa waga wynikająca z planu losowania próby, podlegająca procesowi kalibracji, była postaci $d_k = N/n = 49,99$.



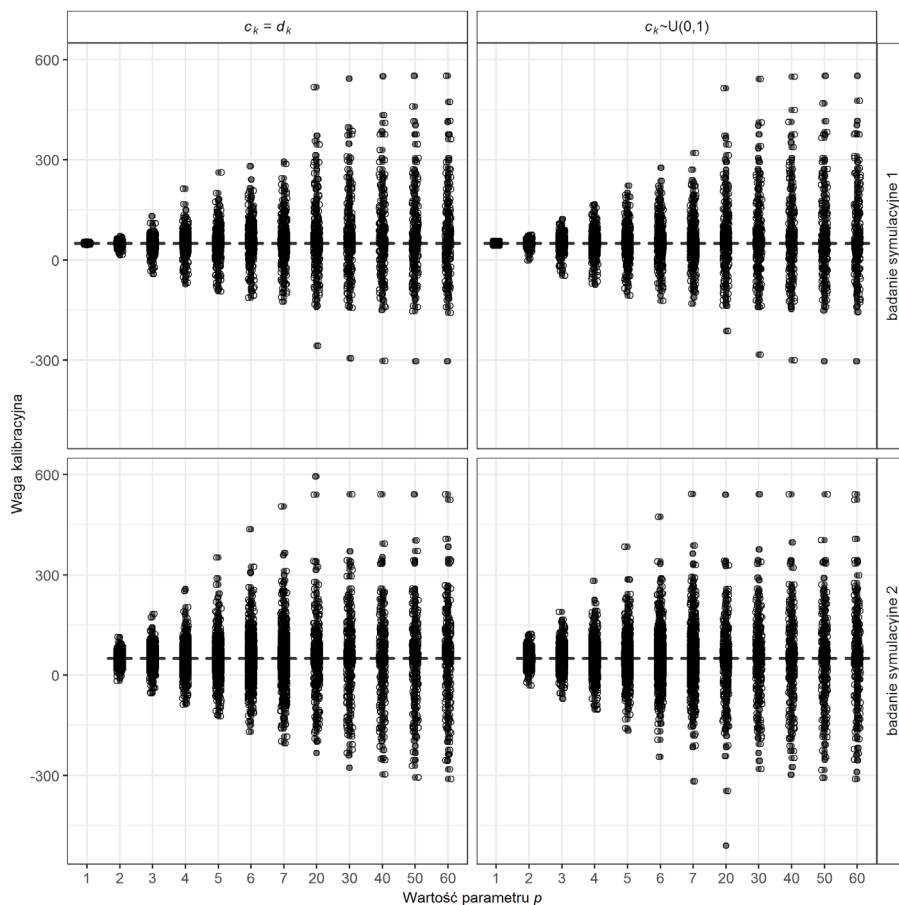
Rysunek 3.2. Oszacowana wartość globalna wydatków ogółem gospodarstw domowych – podejście funkcyjne

Źródło: na podstawie BBGD.

zmienne $1, x_1, x_2$ oraz x_3 . W badaniu symulacyjnym 1 wag ujemnych nie stwierdzono dla wartości parametru $p = 1$ oraz $p = 2$. W badaniu symulacyjnym 2 występowanie ujemnych wag kalibracyjnych stwierdzono w niektórych replikacjach nawet wówczas, gdy $p = 2$.

3.4. Wnioski

W rozdziale przedstawiono estymatory kalibracyjne wartości globalnej uzyskane w tzw. podejściu funkcyjnym. Rozważania prowadzono w kontekście estymacji



Rysunek 3.3. Porównanie rozkładów wag kalibracyjnych w rozpatrywanych wariantach – podejście funkcyjne

Źródło: na podstawie BBGD.

wartości globalnej zarówno w populacji generalnej, jak i w odpowiednio zdefiniowanych domenach. Dyskusji poddano również sposób konstruowania wariacji rozważanych estymatorów wartości globalnej otrzymanych w podejściu funkcyjnym. Pokazano, że wiele znanych w literaturze estymatorów wartości globalnej stanowi szczególny przypadek estymatorów kalibracyjnych uzyskanych w podejściu funkcyjnym.

W badaniu symulacyjnym z wykorzystaniem rzeczywistych danych pochodzących z Badania Budżetów Gospodarstw Domowych dokonano oceny najważniejszych miar jakości estymatorów kalibracyjnych wartości globalnej uzyskanych

w podejściu funkcyjnym. Należy podkreślić bardzo dobre własności rozpatrywanych estymatorów kalibracyjnych, bez względu na wartość parametru p i sposób ustalania wartości współczynnika c_k . Świadczą o tym niskie wartości obciążenia oraz względnych błędów szacunku. Uzyskane wyniki są ponadto w dużej mierze zgodne z rezultatami otrzymanymi przez Estevao i Särndala (2000). Warto podkreślić, że w swoim badaniu autorzy wykorzystali jednak sztuczne dane wygenerowane z odpowiednio zbudowanego modelu. Co więcej, nie przeprowadzili dyskusji nad własnościami wag kalibracyjnych uzyskanymi w podejściu funkcyjnym, co – jak pokazano w niniejszym rozdziale – jest niezwykle ważną kwestią.

Należy zatem wyraźnie mieć na uwadze, że zapewnienie odpowiednich własności tych estymatorów (obciążenia oraz precyzji) nie zawsze idzie w parze z pożądanymi własnościami wag kalibracyjnych (brak wag ujemnych oraz ekstremalnych). W badaniu symulacyjnym z wykorzystaniem rzeczywistych danych pokazano bowiem, że w sytuacji, gdy jesteśmy zainteresowani tylko estymacją interesującej nas charakterystyki w populacji generalnej, podejście funkcyjne stanowi ciekawą alternatywę w stosunku do podejścia minimalizującego odpowiednią funkcję odległości. Uzyskane estymatory również w tym wypadku charakteryzują się pożądanymi własnościami (na przykład niewielkim obciążeniem). Potwierdzeniem tego są zbliżone wartości wybranych miar jakości (RB_{SIM} oraz $RRMSE_{SIM}$) przedstawione w tabelach 2.2 i 3.1.

Warto jednak wyraźnie podkreślić, że gdy wyznaczone wagi będą wykorzystywane do szacowania wartości globalnych innych zmiennych aniżeli Y , wówczas należy dokonać kompleksowej ich oceny, zwłaszcza dla większych wartości parametrów p . W takiej sytuacji obserwuje się bowiem występowanie wag ujemnych oraz ekstremalnych, co generalnie w badaniach jest zjawiskiem niepożądanym.

W wypadku występowania ujemnych wag kalibracyjnych można również rozpatrzeć sposób ich konstruowania opisany w algorytmie 3. W ten sposób uniknie się wag ujemnych, nadal jednak mogą występować wagi ekstremalne.

Podjęcie kalibracyjne wspomagane modelem

4.1. Wprowadzenie

W badaniach reprezentacyjnych prowadzonych przez krajowe urzędy statystyczne wykorzystanie odpowiednio zbudowanych modeli w procesie konstruowania szacunków dla całej populacji lub wyodrębnionych domen jest w pewnym stopniu ograniczone. Wynika to z dwóch zasadniczych powodów. Po pierwsze sceptycyzm związany ze stosowaniem podejścia modelowego jest następstwem przywiązania w statystyce publicznej do tradycyjnych metod estymacji bazujących na odpowiednio zbudowanym systemie wag w ramach tzw. podejścia randomizacyjnego. Pominięcie w tym procesie roli modelu pozwala uniknąć konieczności spełnienia wielu założeń związanych z postacią modelu, metodą szacowania parametrów, zakładanymi rozkładami analizowanych zmiennych itd., które w praktycznych zastosowaniach są bardzo często trudne do zweryfikowania. Po drugie budowanie złożonych modeli jest na ogół czasochłonne, gdyż wymaga stosowania różnych ich postaci dla odmiennych zestawów parametrów czy odpowiedniego doboru zmiennych pomocniczych, co w pewnym stopniu stoi w sprzeczności z wymogiem bieżącego dostarczania wiarygodnych oszacowań (van den Brakel i Bethlehem, 2008; Särndal, 2010).

W tym miejscu warto podkreślić, że proces estymacji może obejmować wnioskowanie wykorzystujące wspomniane już podejście randomizacyjne (ang. *design-based*), wspomagane modelem (ang. *model-assisted approach*) oraz modelowe (ang. *model-based approach*). W wypadku podejścia wspomaganego modelem zakłada się, że dostępne są informacje z próby losowej, przy czym wnioskowanie statystyczne może mieć różny charakter. Z kolei w odniesieniu do podejścia modelowego wnioskowanie ma charakter warunkowy, tj. odbywa się na podstawie danej próby, przy czym nie musi być ona próbą losową. W tym podejściu wagi na ogół odgrywają mniej doniosłą rolę, a wnioskowanie opiera się na strukturze generowanej przez model (van den Brakel i Bethlehem, 2008; Żądło, 2015).

Podejście randomizacyjne oraz wspomagane modelem jest szeroko wykorzystywane w praktyce badań reprezentacyjnych realizowanych przez krajowe urzędy statystyczne na całym świecie. Warto wspomnieć o omówionych już estymatorach Horvitz-Thompsona czy GREG, które stanowią przykłady najważniejszych estymatorów w obrębie obydwu przytoczonych podejść odpowiednio. Również rozpatrywane w tym rozdziale estymatory kalibracyjne należy postrzegać w kontekście podejścia mieszanego, w którym proces wnioskowania jest wspomagany modelem.

Natomiast podejście modelowe jest w coraz większym zakresie wykorzystywane w estymacji pośredniej i znajduje liczne zastosowania w praktyce badań statystycznych. Ograniczając się tylko do statystyki publicznej w Polsce, warto wspomnieć o zastosowaniach modelu Faya-Herriota i M-kwantylogowego w kontekście badań nad ubóstwem (Marchetti, Beręsewicz, Salvati, Szymkowiak i Wawrowski, 2018; Szymkowiak, Młodak i Wawrowski, 2017) czy pracach nad zjawiskiem niepełnosprawności z wykorzystaniem estymacji typu SPREE (Szymkowiak, Józefowski i Klimanek, 2018) lub dotyczących estymacji miesięcznej stopy bezrobocia z użyciem strukturalnych modeli szeregów czasowych (Wilak, Beręsewicz i Szymkowiak, 2017b) prowadzonych w Ośrodku Statystyki Małych Obszarów Urzędu Statystycznego w Poznaniu.

W niniejszym rozdziale uwaga została skoncentrowana głównie na estymatorach kalibracyjnych wartości globalnej wspomaganych modelem²⁵. Ta klasa estymatorów, których koncepcję zaproponowali Wu i Sitter (2001), nie była do tej pory przedmiotem szerszego zainteresowania statystyki publicznej. Rozważania teoretyczne na temat podejścia mieszanego w kontekście estymatorów modelowo-kalibrowanych można znaleźć we wspomnianej już pracy Wu i Sittera (2001) oraz Wu (2003), a w polskiej literaturze w pracy Żądły (2015). W niniejszej książce

²⁵ Określane one będą zamiennie jako estymatory modelowo-kalibrowane – por. Żądło (2015).

skoncentrowano się także na najnowszych rozwiązaniach w tym zakresie obejmujących estymatory hybrydowo-kalibrowane (Lehtonen i Veijanen, 2015; 2017; 2018) oraz wykorzystujących regresję LASSO (Tibshirani, 1996) i grzbietową (Hoerl i Kennard, 1970) w kontekście podejścia kalibracyjnego (Chen i in., 2018a; McConville, Breidt, Lee i Moisen, 2017).

W badaniu symulacyjnym, wykorzystując rzeczywiste dane pochodzące ze wspomnianego już BBGD, analizie poddano wagi kalibracyjne oraz najważniejsze własności rozważanych estymatorów wspomaganym modelem. Dyskusja dotyczyć będzie także różnic pomiędzy estymatorami uzyskanymi w podejściu kalibracyjnym opisanym w rozdziale drugim a modelowo-kalibrowanymi.

4.2. Estymatory modelowo-kalibrowane wartości globalnej

Założmy w dalszym ciągu, że przedmiotem naszych rozważań jest estymacja wartości globalnej zmiennej Y w populacji, która wyraża się wzorem (1.34). W ujęciu zaproponowanym przez Deville'a i Särndala (1992) proces wyznaczania wag kalibracyjnych sprowadzał się do minimalizacji właściwie dobranej funkcji odległości $D(\cdot)$ przy założeniu, że spełnione będą odpowiednie równania kalibracyjne. W podejściu tym nie zakładało się istnienia żadnego modelu wiążącego zmienną objaśnianą ze zmiennymi pomocniczymi x_1, \dots, x_J (ang. *model-free calibration*). Istnieje jednak możliwość włączenia w procesie kalibracji odpowiedniego modelu, który uwzględniałby powiązania zmiennej objaśnianej ze zmiennymi pomocniczymi x_1, \dots, x_J . Tego typu koncepcja została sformułowana przez Wu i Sittera (2001), a następnie rozwinięta przez Wu (2003).

Z formalnego punktu widzenia Wu i Sitter (2001) rozważają problem wspomaganego modelem nadpopulacji, zakładając związek pomiędzy zmienną Y a zmiennymi pomocniczymi x_1, \dots, x_J :

$$\begin{cases} E_U(Y_k | \mathbf{x}_k) = \mu(\mathbf{x}_k, \boldsymbol{\beta}), \\ D_U^2(Y_k | \mathbf{x}_k) = v_k^2 \sigma^2, \end{cases} \quad (4.1)$$

gdzie $\mu(\mathbf{x}_k, \boldsymbol{\beta})$ jest pewną znaną funkcją \mathbf{x}_k i $\boldsymbol{\beta} = (\beta_1, \dots, \beta_J)^T$, $\boldsymbol{\beta}$ i σ^2 są nieznanymi parametrami modelu nadpopulacji, a v_k jest znaną funkcją wartości zmiennych pomocniczych dla $k = 1, \dots, N$.

Założmy w dalszym ciągu, że z populacji U losujemy zgodnie z określonym planem losowania n -elementową próbę s . Przez π_k oznaczamy, jak poprzed-

nio, prawdopodobieństwo inkluzji pierwszego rzędu, a przez d_k wagę wynikającą z planu losowania próby. Wu i Sitter (2001) proponują, aby za estymator parametru β przyjąć:

$$\hat{\beta} = (\mathbf{X}_s^T \mathbf{\Pi}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{X}_s^T \mathbf{\Pi}_s^{-1} \mathbf{y}_s, \quad (4.2)$$

gdzie macierz \mathbf{X}_s jest zdefiniowana wzorem (2.60), $\mathbf{y}_s = (y_1, \dots, y_n)^T$, a $\mathbf{\Pi}_s$ to macierz diagonalna o wymiarach $n \times n$ zawierająca na głównej przekątnej prawdopodobieństwa inkluzji pierwszego rzędu π_k :

$$\mathbf{\Pi}_s = \text{diag}(\pi_1, \dots, \pi_n) = \begin{pmatrix} \pi_1 & 0 & \dots & 0 \\ 0 & \pi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi_n \end{pmatrix}. \quad (4.3)$$

Alternatywnie estymator parametru β można zapisać w równoważnej postaci:

$$\hat{\beta} = (\mathbf{X}_s^T \mathbf{D}_s \mathbf{X}_s)^{-1} \mathbf{X}_s^T \mathbf{D}_s \mathbf{y}_s, \quad (4.4)$$

gdzie \mathbf{D}_s wyraża się wzorem (2.61).

Zakładając, że $\hat{\mu}_k = \hat{\mu}_k(\mathbf{x}_k, \hat{\beta})$ są wartościami teoretycznymi (modelowymi) wynikającymi z przyjętego modelu $\mu(\mathbf{x}_k, \beta)$, w którym oszacowano parametr β zgodnie ze wzorem (4.2), wagi kalibracyjne $w_{k\text{MC}}$ w estymatorze kalibracyjnym $\hat{\tau}_{\text{MC}}$ wartości globalnej (1.34) wspomaganym modelem znajduje się poprzez rozwiązanie następującego zadania warunkowej minimalizacji²⁶:

- (W1)_{MC} – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s} \frac{d_k}{q_k} G\left(\frac{w_{k\text{MC}}}{d_k}\right) \rightarrow \min, \quad (4.5)$$

- (W2)_{MC} – równania kalibracyjne:

$$\sum_{k \in s} w_{k\text{MC}} = N \quad \text{oraz} \quad \sum_{k \in s} w_{k\text{MC}} \hat{\mu}_k = \sum_{k \in U} \hat{\mu}_k, \quad (4.6)$$

- (W3)_{MC} – warunki ograniczające:

$$L \leq \frac{w_{k\text{MC}}}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, n. \quad (4.7)$$

²⁶ Subskrypt MC pochodzi od angielskiego wyrażenia *model-assisted calibration* – kalibracja wspomaganą modelem.

Podobnie jak w wypadku podejścia minimalizującego odpowiednią funkcję odległości opisanego w rozdziale drugim niniejszej książki, wagi w estymatorze modelowo-kalibrowanym wartości globalnej (1.34) poszukujemy, znajdując minimum warunkowe odpowiedniej funkcji celu – warunek (W1)_{MC}. Wyjaśnienia wymaga warunek (W2)_{MC}. Orzeka on, że wagi modelowo-kalibrowane powinny się sumować do liczebności populacji N . Jednocześnie wyznaczone są tak, aby ważona suma wartości teoretycznych wynikających z przyjętego modelu po wszystkich jednostkach wylosowanych do próby s równała się sumie wartości teoretycznych wyznaczonych dla wszystkich jednostek z populacji U . Warunek ten wymusza znajomość jednostkowych wartości wszystkich zmiennych pomocniczych na poziomie populacji. Odpowiedni model budowany jest bowiem na podstawie informacji zawartych w próbie s , przy czym wartości modelowe $\hat{\mu}_k$ należy wyznaczyć dla wszystkich jednostek z populacji U . Również w tym wypadku można nałożyć na iloraz wag modelowo-kalibrowanych w_{kMC} i wynikających z przyjętego planu losowania próby d_k pewne ograniczenia, które mają zapobiegać występowaniu wag ujemnych bądź ekstremalnych.

Postać wag modelowo-kalibrowanych w_{kMC} dla najczęściej wykorzystywanej w praktyce badań statystycznych funkcji $G(x) = \frac{1}{2}(x-1)^2$ oraz z uwzględnieniem tylko dwóch wspomnianych warunków (W1)_{MC} i (W2)_{MC} podaje poniższe twierdzenie. Oczywiście, nic nie stoi na przeszkodzie, aby wagi wyznaczać również dla innych funkcji odległości²⁷ i ograniczeń nakładanych w warunku (W3)_{MC}. W tym celu należy skorzystać z metody konstruowania wag opisanej w algorytmie 1 lub 2, w których kluczową rolę odgrywa sposób znajdowania mnożnika Lagrange’a λ .

Twierdzenie 4.1 (Wu i Sitter, 2001). *Rozwiązaniem zadania minimalizacji funkcji odległości (4.5) przy warunku (4.6) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag modelowo-kalibrowanych $\mathbf{w}_{MC} = (w_{1MC}, \dots, w_{nMC})^T$, którego składowe dla $k = 1, \dots, n$ wyrażają się wzorem:*

$$w_{kMC}^{\text{lin}} = d_k + d_k q_k \left(\sum_{k \in U} \mathbf{u}_{k1} - \sum_{k \in s} d_k \mathbf{u}_{k1} \right)^T \left(\sum_{k \in s} q_k d_k \mathbf{u}_{k1} \mathbf{u}_{k1}^T \right)^{-1} \mathbf{u}_{k1}, \quad (4.8)$$

gdzie $\mathbf{u}_{k1} = (1, \hat{\mu}_k)^T$.

Dowód powyższego twierdzenia jest analogiczny do dowodu twierdzenia 2.1. Zwróćmy przy tym uwagę, że wektor wartości globalnych $\sum_{k \in U} \mathbf{u}_{k1}$ występujący we wzorze (4.8) jest postaci $\boldsymbol{\tau}_{\mathbf{u}} = \sum_{k \in U} \mathbf{u}_{k1} = (N, \sum_{k \in U} \hat{\mu}_k)^T$.

²⁷ Funkcje te zostały syntetycznie przedstawione w tabeli 2.1.

W praktycznych zastosowaniach najczęściej przyjmuje się, że $q_k = 1$. W takiej sytuacji wektor wag modelowo-kalibrowanych można przedstawić w następującej postaci macierzowej:

$$\mathbf{w}_{\text{MC}}^{\text{lin}} = \mathbf{d} + \mathbf{D}_s \mathbf{M} \left(\mathbf{M}^T \mathbf{D}_s \mathbf{M} \right)^{-1} \left(\boldsymbol{\tau}_{\mathbf{u}}^T - \mathbf{d}^T \mathbf{M} \right)^T, \quad (4.9)$$

gdzie $\mathbf{M} = (\mathbf{d}, \hat{\boldsymbol{\mu}})$; \mathbf{M} jest zatem macierzą, której pierwszą kolumnę stanowi wektor wag wynikających z założonego planu losowania próby, a drugą kolumnę wektor złożony z wartości teoretycznych uzyskanych dla przyjętego modelu, a wyznaczony dla wszystkich jednostek należących do wylosowanej próby s .

Definicja 4.1. *Estymatorem kalibracyjnym wartości globalnej (1.34) wspomaganym modelem (estymatorem modelowo-kalibrowanym) jest:*

$$\hat{\tau}_{\text{MC}} = \sum_{k \in s} w_{k\text{MC}} y_k, \quad (4.10)$$

gdzie wagi modelowo-kalibrowane $w_{k\text{MC}}$ są rozwiązaniem zadania optymalizacyjnego (4.5)–(4.7). W szczególnym przypadku dla funkcji odległości zdefiniowanej jako (2.18) i uwzględnieniu jedynie warunków (4.5)–(4.6) wyrażają się one wzorem (4.8) lub (4.9)²⁸.

Powyżej zdefiniowany estymator jest w swojej konstrukcji podobny do estymatora (2.3), który został przedstawiony w rozdziale drugim. Różnica dotyczy jedynie sposobu budowania wag kalibracyjnych.

Przyjmując założenie, że $q_k = 1$, oraz biorąc pod uwagę warunki (4.5)–(4.6), estymator modelowo-kalibrowany wartości globalnej (1.34) dla funkcji odległości (2.18) można również zapisać w następującej postaci macierzowej:

$$\hat{\tau}_{\text{MC}} = \left(\mathbf{w}_{\text{MC}}^{\text{lin}} \right)^T \mathbf{y}_s = \mathbf{d}^T \mathbf{y}_s + \left(\boldsymbol{\tau}_{\mathbf{u}}^T - \mathbf{d}^T \mathbf{M} \right) \left(\mathbf{X}_s^T \mathbf{D}_s \mathbf{X}_s \right)^{-1} \mathbf{X}_s \mathbf{D}_s \mathbf{y}_s. \quad (4.11)$$

Alternatywnie, przy sformułowanych powyżej założeniach, estymator modelowo-kalibrowany $\hat{\tau}_{\text{MC}}$ wartości globalnej (1.34) można przedstawić jako (Wu i Sitter, 2001):

$$\hat{\tau}_{\text{MC}} = \sum_{k \in s} d_k y_k + \left(\sum_{k \in U} \hat{\mu}_k - \sum_{k \in s} d_k \hat{\mu}_k \right) \hat{\beta}_{\text{MC}}, \quad (4.12)$$

²⁸ W tym wypadku przy przyjęciu założenia, że $q_k = 1$.

gdzie:

$$\hat{\beta}_{\text{MC}} = \frac{\sum_{k \in s} d_k (\hat{\mu}_k - \hat{\mu})(y_k - \bar{y})}{\sum_{k \in s} d_k (\hat{\mu}_k - \hat{\mu})^2}, \quad (4.13)$$

przy czym

$$\hat{\mu} = \frac{\sum_{k \in s} d_k \hat{\mu}_k}{\sum_{k \in s} d_k} \quad \text{oraz} \quad \bar{y} = \frac{\sum_{k \in s} d_k y_k}{\sum_{k \in s} d_k}.$$

Warto przy tym zauważyć, że w wypadku gdy (4.1) będzie modelem liniowym, estymator modelowo-kalibrowany (4.10) wartości globalnej (1.34) dla funkcji odległości (2.18), przy uwzględnieniu warunków (4.5)–(4.6), jest równoważny uogólnionemu liniowemu estymatorowi regresyjnemu GREG (1.44), w którym za jedną ze zmiennych pomocniczych przyjmuje się wektor złożony z samych jedynek. Estymator modelowo-kalibrowany wspomagany modelem liniowym może być zatem traktowany jako szczególny przypadek estymatora kalibracyjnego wartości globalnej w podejściu minimalizującym funkcję odległości (2.18).

Dokonując analizy postaci wag kalibracyjnych $w_{k\text{MC}}$ w estymatorze modelowo-kalibrowanym wartości globalnej, można zauważyć, że wagi te zależą m.in. od wartości teoretycznych $\hat{\mu}_k$ poprzez odpowiednio zbudowany model. W podejściu tym przy szacowaniu wartości globalnej zmiennej Y nie są zatem odtwarzane wartości globalne wszystkich zmiennych pomocniczych x_1, \dots, x_J , a wartość globalna złożona z wartości teoretycznych $\hat{\mu}_k$ oraz liczebność populacji N . Odróżnia to w zasadniczy sposób estymatory modelowo-kalibrowane wartości globalnej od estymatorów uzyskanych poprzez minimalizację odpowiedniej funkcji odległości, opisanych w rozdziale drugim niniejszej książki. Oznacza to, że dla innej zmiennej celu aniżeli Y , dla której szacowana będzie wartość globalna, należy ponownie zbudować odpowiedni model i wyznaczyć wagi modelowo-kalibrowane $w_{k\text{MC}}$. W podejściu minimalizującym funkcję odległości wektor wag buduje się tylko raz, odtwarzając wartości globalne wszystkich zmiennych pomocniczych, a następnie wykorzystuje się go do szacowania wartości globalnych różnych zmiennych celu. Stanowi to zatem istotną różnicę, a ten sposób konstruowania wag (ang. *multipurpose weighting*) ma zdecydowanie bardziej uniwersalny charakter. W wypadku gdy interesuje nas oszacowanie wartości globalnej konkretnej zmiennej Y , użyć odpowiednio zbudowanego modelu może się przyczynić jednak do lepszych wyników w kontekście obciążenia i precyzji szacunków. Warto przy tym zaznaczyć, że nie są to jedyne rozbieżności pomiędzy obydwoimi rodzajami kalibracji. W tabeli 4.1 przedstawiono najważniejsze różnice między podejściem kalibracyj-

Tabela 4.1. Różnice między podejściem kalibracyjnym minimalizującym funkcję odległości a wspomaganym modelem

Cecha	Podejście kalibracyjne	
	klasyczne	wspomagane modelem
Cel	Zbudowanie wag kalibracyjnych tak, aby odtworzyć znane ze źródeł zewnętrznych (spisy oraz rejestry) wartości globalne zmiennych pomocniczych.	Zbudowanie wag kalibracyjnych tak, aby odtworzyć wartości globalne złożone z wartości teoretycznych dla założonego modelu oraz liczebności populacji N .
Popularność	Podejście powszechnie wykorzystywane w statystyce publicznej w wielu różnych badaniach.	Sporadyczne wykorzystanie w badaniach statystyki publicznej.
Dane	Wymagana znajomość zmiennych pomocniczych na poziomie próby, a także wartości globalnych tych zmiennych w populacji.	Wymagana znajomość jednostkowych danych dla zmiennych pomocniczych na poziomie populacji.
Model	W procesie wyznaczania wag kalibracyjnych nie jest wymagana znajomość żadnego modelu (ang. <i>model-free calibration</i>).	Niezbędne jest przyjęcie założenia o istnieniu modelu wiążącego zmienną Y ze wszystkimi zmiennymi pomocniczymi x_1, \dots, x_J .
Zgodność	System wag odtwarza znane wartości globalne wszystkich zmiennych pomocniczych.	System wag zazwyczaj nie odtwarza znanych wartości globalnych zmiennych pomocniczych.
Uniwersalność	System wag nie zależy od wartości zmiennej celu Y . Wagi można wykorzystywać w procesie estymacji parametrów dla wielu różnych zmiennych.	System wag zależy od oszacowanych wartości zmiennej Y . Nie jest zatem spełniona zasada uniwersalności. Wagi należy za każdym razem konstruować dla wielu różnych zmiennych Y .
Odporność	Odporne w sytuacji błędnie wyspecyfikowanego modelu.	Brak odporności w sytuacji błędnie wyspecyfikowanego modelu.
Precyzja i dokładność	Wymagana jest znajomość wysokiej jakości zmiennych pomocniczych.	W wypadku dobrze wyspecyfikowanego modelu obserwuje się poprawę na precyzji i dokładności oszacowań.
Czasowość	Podejście to jest stosowane przede wszystkim w badaniach statystyki publicznej, gdyż wymaga budowy tylko jednego systemu wag kalibracyjnych. Skraca to czas publikacji wyników.	Niezbędne jest wyznaczenie kilkunastu zestawów wag kalibracyjnych dla różnych zmiennych Y , co wydłuża proces estymacji oraz może ujemnie wpłynąć na proces produkcji danych statystycznych.

Źródło: na podstawie badań literaturowych.

nym minimalizującym odpowiednią funkcję odległości a estymatorami modelowo-kalibrowanymi, którym poświęcony jest niniejszy rozdział książki.

Kluczową kwestią jest również postać wariancji estymatorów kalibracyjnych wartości globalnej wspomaganym modelem. Wu i Sitter (2001) podają wzór na wariancję estymatora modelowo-kalibrowanego wartości globalnej (1.34) dla funkcji odległości (2.18), którą dla $q_k = 1$ można wyrazić wzorem:

$$D^2(\hat{\tau}_{MC}) = \sum_{k \in U} \sum_{\substack{l \in U \\ l > k}} (\pi_k \pi_l - \pi_{kl}) (d_k e_k^{MC} - d_l e_l^{MC})^2, \quad (4.14)$$

gdzie $e_k^{MC} = y_k - \mu(\mathbf{x}_k, \boldsymbol{\beta}_U) \beta_{MC}$, $\boldsymbol{\beta}_U = (\mathbf{X}_U^T \mathbf{X}_U)^{-1} \mathbf{X}_U^T \mathbf{y}$, a parametr β_{MC} jest postaci:

$$\beta_{MC} = \frac{\sum_{k \in U} (\mu(\mathbf{x}_k, \boldsymbol{\beta}_U) - \bar{\mu}_{MC}) (y_k - \bar{y}_{MC})}{\sum_{k \in U} (\mu(\mathbf{x}_k, \boldsymbol{\beta}_U) - \bar{\mu}_{MC})^2}, \quad (4.15)$$

przy czym

$$\bar{\mu}_{MC} = \frac{\sum_{k \in U} \mu(\mathbf{x}_k, \boldsymbol{\beta}_U)}{N} \quad \text{oraz} \quad \bar{y}_{MC} = \frac{\sum_{k \in U} y_k}{N},$$

a \mathbf{X}_U jest odpowiednikiem macierzy \mathbf{X}_s w populacji:

$$\mathbf{X}_U = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1J} \\ x_{21} & x_{22} & \dots & x_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NJ} \end{pmatrix}. \quad (4.16)$$

Zgodnie z propozycją Wu i Sittera (2001), estymator wariancji (4.14) może być przedstawiony w postaci:

$$\hat{D}^2(\hat{\tau}_{MC}) = - \sum_{k \in s} \sum_{\substack{l \in s \\ l > k}} \frac{\Delta_{kl}}{\pi_{kl}} (d_k \hat{e}_k^{MC} - d_l \hat{e}_l^{MC})^2, \quad (4.17)$$

gdzie $\hat{e}_k^{MC} = y_k - \hat{\mu}(\mathbf{x}_k, \hat{\boldsymbol{\beta}}) \hat{\beta}_{MC}$, $\hat{\boldsymbol{\beta}}$ wyraża się wzorem (4.2), a $\hat{\beta}_{MC}$ jest postaci (4.13).

Podobnie jak w wypadku estymatorów kalibracyjnych wartości globalnej w podejściu minimalizującym funkcję odległości, również dla estymatorów kalibracyjnych wartości globalnej wspomaganym modelem można rozważyć problem estymacji wartości globalnej w domenie.

Zakładamy zatem, że populacja U składa się z D rozłącznych domen postaci U_1, \dots, U_D . Zakładamy ponadto, że w każdej z D wyróżnionych domen przedmiotem zainteresowania jest wartość globalna $\tau_{Y_1}, \dots, \tau_{Y_D}$, gdzie τ_{Y_d} , dla $d = 1, \dots, D$, wyraża się wzorem (1.58). Problem poszukiwania wag modelowo-kalibrowanych $w_{dk_{MC}}$ estymatora kalibracyjnego wartości globalnej w domenie d dla jednostki k można zatem opisać w następujący sposób:

- (W1)_{d,MC} – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s_d} \frac{d_k}{q_k} G\left(\frac{w_{dk_{MC}}}{d_k}\right) \rightarrow \min, \quad (4.18)$$

- (W2)_{d,MC} – równania kalibracyjne:

$$\sum_{k \in s_d} w_{dk_{MC}} = N_d \quad \text{oraz} \quad \sum_{k \in s_d} w_{dk_{MC}} \hat{\mu}_k = \sum_{k \in U_d} \hat{\mu}_k, \quad (4.19)$$

- (W3)_{d,MC} – warunki ograniczające:

$$L \leq \frac{w_{dk_{MC}}}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, n. \quad (4.20)$$

W powyżej sformułowanym zadaniu optymalizacyjnym zakładamy, że odtwarzane są liczebności populacji w domenie oraz wartości globalne złożone z wartości teoretycznych otrzymanych z modelu zbudowanego na danych w ramach konkretnej domeny.

Przyjmując jak wcześniej funkcję $G(x) = \frac{1}{2}(x-1)^2$ oraz warunki (W1)_{d,MC} i (W2)_{d,MC}, można łatwo znaleźć postać wag modelowo-kalibrowanych w domenach, o czym mówi poniższe twierdzenie.

Twierdzenie 4.2. *Rozwiązaniem zadania minimalizacji funkcji odległości (4.18) przy warunku (4.19) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag modelowo-kalibrowanych $\mathbf{w}_{dk_{MC}}$, którego składowe dla jednostki k w domenie d , dla $d = 1, \dots, D$, wyrażają się wzorem:*

$$w_{dk_{MC}}^{\text{lin}} = d_k + d_k q_k \left(\sum_{k \in U_d} \mathbf{u}_{k_1} - \sum_{k \in s_d} d_k \mathbf{u}_{k_1} \right)^T \left(\sum_{k \in s_d} q_k d_k \mathbf{u}_{k_1} \mathbf{u}_{k_1}^T \right)^{-1} \mathbf{u}_{k_1}, \quad (4.21)$$

gdzie $\mathbf{u}_{k_1} = (1, \hat{\mu}_k)^T$.

Mając wyznaczone wagi modelowo-kalibrowane $w_{dk_{MC}}$ dla dowolnie określonej funkcji odległości i uwzględniając bądź nie warunki ograniczające (W3)_{d,MC},

można zdefiniować estymator kalibracyjny wartości globalnej τ_{Y_d} w domenie d wspomagany modelem.

Definicja 4.2. *Estymatorem kalibracyjnym wspomaganym modelem wartości globalnej (1.58) w domenie d , dla $d = 1, \dots, D$, jest:*

$$\hat{\tau}_{MC,d} = \sum_{k \in s_d} w_{dk_{MC}} y_k, \quad (4.22)$$

gdzie wagi modelowo-kalibrowane $w_{dk_{MC}}$ są rozwiązaniem zadania optymalizacyjnego (4.18)–(4.20). W szczególnym przypadku dla funkcji odległości określonej jako (2.18) i przy warunkach (4.18)–(4.19) wagi modelowo-kalibrowane $w_{dk_{MC}}$ można wyrazić wzorem (4.21).

Estymator modelowo-kalibrowany wartości globalnej w domenie d dla $d = 1, \dots, D$, który wyraża się wzorem (4.22), można zatem – przy $q_k = 1$ oraz uwzględniając wagi modelowo-kalibrowane (4.21) – zapisać w następującej postaci:

$$\hat{\tau}_{MC,d} = \sum_{k \in s_d} d_k y_k + \left(\sum_{k \in U_d} \hat{\mu}_k - \sum_{k \in s_d} d_k \hat{\mu}_k \right) \hat{\beta}_{MC,d}, \quad (4.23)$$

gdzie:

$$\hat{\beta}_{MC,d} = \frac{\sum_{k \in s_d} d_k (\hat{\mu}_k - \hat{\mu}_d) (y_k - \bar{y}_d)}{\sum_{k \in s_d} d_k (\hat{\mu}_k - \hat{\mu}_d)^2}, \quad (4.24)$$

przy czym

$$\hat{\mu}_d = \frac{\sum_{k \in s_d} d_k \hat{\mu}_k}{\sum_{k \in s_d} d_k} \quad \text{oraz} \quad \bar{y}_d = \frac{\sum_{k \in s_d} d_k y_k}{\sum_{k \in s_d} d_k}.$$

Warto zwrócić uwagę, że podobnie jak w wypadku estymatorów kalibracyjnych wartości globalnej w domenie opisanych w rozdziale drugim, również i tutaj nie jest możliwe oszacowanie wartości globalnej z wykorzystaniem estymatora modelowo-kalibrowanego (4.22), gdy domena nie jest reprezentowana w próbie przez żadną jednostkę.

Wariancję estymatora (4.22) można oszacować z wykorzystaniem wzoru (4.17), przy czym reszty \hat{e}_k^{MC} zastępowane są przez:

$$\hat{e}_{dk}^{MC} = \delta_{dk} (y_k - \hat{\mu}_k \hat{\beta}_{MC,d}), \quad (4.25)$$

gdzie $\hat{\beta}_{MC,d}$ dane jest wzorem (4.24).

4.3. Estymatory kalibracyjne wartości globalnej wspomagane modelem regresji LASSO i grzbietowej

Ciekawą koncepcję konstrukcji estymatorów kalibracyjnych wartości globalnej wspomaganych adaptacyjnym modelem regresji LASSO (ang. *adaptive LASSO*) zaproponował Chen (2016). Metoda ta znalazła praktyczne zastosowanie w ocenie stopnia pokrycia gruntów koronami drzew w stanie Utah (McConville i in., 2017) oraz we wspomnianych w podrozdziale 1.6 badaniach sondażowych w kontekście oszacowania preferencji wyborczych w Stanach Zjednoczonych (Chen i in., 2018a). Warto przy tym dodać, że podstawy modelu regresji LASSO zostały sformułowane przez Tibshiraniego (1996), a adaptacyjnego modelu regresji LASSO przez Zou (2006).

W podejściu kalibracyjnym wspomaganim adaptacyjnym modelem regresji LASSO nieznaną wektor parametrów β występujący w modelu (4.1) jest szacowany w następujący sposób (McConville i in., 2017; Chen i in., 2018a):

$$\hat{\beta}^{(AL)} = \underset{\beta}{\operatorname{argmin}} \left((\mathbf{y}_s - \mathbf{X}_s \beta)^T \mathbf{\Pi}_s^{-1} (\mathbf{y}_s - \mathbf{X}_s \beta) + \lambda_n \sum_{j=1}^J \alpha_j^\gamma |\beta_j| \right), \quad (4.26)$$

gdzie \mathbf{X}_s jest macierzą postaci (2.60), $\mathbf{\Pi}_s$ macierzą wyrażającą się wzorem (4.3), α_j^γ stanowią tzw. parametry regularyzacji zwane również współczynnikami tłumienia, λ_n jest parametrem kary²⁹, a $\gamma > 0$ jest odpowiednio dobraną stałą³⁰. Parametry regularyzacji α_j są proporcjonalne do odwrotności wartości bezwzględnych współczynników regresji β_j , tj. $\alpha_j \propto 1/|\beta_j|$. W praktycznych zastosowaniach najczęściej przyjmuje się, że:

$$\alpha_j = \frac{1}{|\hat{\beta}_j^{(ML)}|} \quad (4.27)$$

lub

$$\alpha_j = \frac{1}{|\hat{\beta}_j^{(RI)}|}, \quad (4.28)$$

gdzie $\hat{\beta}_j^{(ML)}$ oraz $\hat{\beta}_j^{(RI)}$ stanowią oszacowania parametru β_j uzyskane metodą największej wiarygodności (ang. *maximum likelihood* – ML) bądź występują-

²⁹ Sposób jego wyznaczania został opisany w artykule Chena i współautorów (2018a).

³⁰ Subskrypt *AL* pochodzi od angielskiego wyrażenia *adaptive LASSO*.

cych w modelu regresji grzbietowej³¹ (ang. *ridge regression* – RIDGE³²) odpowiednio.

Mając wyznaczony oszacowany wektor parametrów (4.26) uzyskany w adaptacyjnym modelu regresji LASSO, można skonstruować wagi kalibracyjne $w_{k_{AL}}$, rozwiązując zadanie optymalizacyjne (4.5)–(4.7). W szczególnym przypadku dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ oraz warunków (4.5)–(4.6) wagi modelowo-kalibrowane $w_{k_{AL}}$ można wyznaczyć ze wzoru (4.8). Estymator kalibracyjny wartości globalnej wspomagany adaptacyjnym modelem regresji LASSO będzie wówczas postaci (4.10), przy czym w miejsce wag $w_{k_{MC}}$ należy podstawić wagi $w_{k_{AL}}$.

W sposób analogiczny do przedstawionego w podrozdziale 4.2 można dokonać również oszacowania wariancji estymatora kalibracyjnego wartości globalnej wspomaganego adaptacyjnym modelem regresji LASSO czy konstrukcji wag w estymatorze kalibracyjnym wartości globalnej w domenie z uwzględnieniem jego wariancji. W tym celu należy skorzystać z odpowiednich wzorów przedstawionych w podrozdziale 4.2, mając na uwadze, że:

$$\hat{\mu}_k = \hat{\mu}_k(\mathbf{x}_k, \hat{\boldsymbol{\beta}}^{(AL)}). \quad (4.29)$$

Teoretyczne podstawy regresji grzbietowej zostały sformułowane przez Hoerla i Kennarda (1970). W wypadku estymatora kalibracyjnego wartości globalnej (4.10) wspomaganego modelem regresji grzbietowej nieznaną wektor parametrów $\boldsymbol{\beta}$ występujący w modelu (4.1) jest szacowany ze wzoru (McConville i in., 2017):

$$\hat{\boldsymbol{\beta}}^{(RI)} = \left(\mathbf{X}_s^T \boldsymbol{\Pi}_s^{-1} \mathbf{X}_s + \mu_{RI} \mathbf{Q}^- \right)^{-1} \mathbf{X}_s^T \boldsymbol{\Pi}_s^{-1} \mathbf{y}_s, \quad (4.30)$$

gdzie \mathbf{Q} jest macierzą diagonalną, która na głównej przekątnej ma elementy wektora $(|\hat{B}_1^{(L)}| + \eta, \dots, |\hat{B}_J^{(L)}| + \eta)$, przy czym η jest ustaloną małą liczbą dodatnią³³.

Z kolei parametr kary μ_{RI} jest w taki sposób dobrany, aby spełniony był następujący warunek:

$$\sum_{j=1}^J |\hat{\beta}_j^{(RI)}| = \sum_{j=1}^J |\hat{\beta}_j^{(L)}|, \quad (4.31)$$

³¹ Podejściu kalibracyjnemu wspomaganemu modelem regresji grzbietowej poświęcona będzie dalsza część tego podrozdziału.

³² W subskryptach używane będzie oznaczenie RI.

³³ \mathbf{Q}^- oznacza uogólnioną macierz odwrotną do macierzy \mathbf{Q} .

przy czym:

$$\hat{\beta}^{(L)} = \underset{\beta}{\operatorname{argmin}} \left((\mathbf{y}_s - \mathbf{X}_s \beta)^T \mathbf{\Pi}_s^{-1} (\mathbf{y}_s - \mathbf{X}_s \beta) + \lambda_n \sum_{j=1}^J |\beta_j| \right). \quad (4.32)$$

Również w tym wypadku, mając oszacowany wektor parametrów (4.30) uzyskany w modelu regresji grzbietowej, można skonstruować odpowiednie wagi kalibracyjne $w_{k_{\text{RI}}}$, rozwiązując zadanie optymalizacyjne (4.5)–(4.7). W szczególności, dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ oraz warunków (4.5)–(4.6), wagi modelowo-kalibrowane $w_{k_{\text{RI}}}$ można wyznaczyć ze wzoru (4.8). Estymator kalibracyjny wartości globalnej wspomagany modelem regresji grzbietowej będzie wówczas postaci (4.10), przy czym w miejsce wag $w_{k_{\text{MC}}}$ należy podstawić wagi $w_{k_{\text{RI}}}$.

W sposób analogiczny do przedstawionego w podrozdziale 4.2 można dokonać również oszacowania wariancji estymatora kalibracyjnego wartości globalnej wspomaganego modelem regresji grzbietowej czy konstrukcji wag w estymatorze kalibracyjnym wartości globalnej w domenie z uwzględnieniem jego wariancji. W tym celu należy skorzystać z odpowiednich wzorów przedstawionych w podrozdziale 4.2, mając tym razem na uwadze, że:

$$\hat{\mu}_k = \hat{\mu}_k \left(\mathbf{x}_k, \hat{\beta}^{(\text{RI})} \right). \quad (4.33)$$

4.4. Estymatory hybrydowo-kalibrowane wartości globalnej

Estymatory kalibracyjne wartości globalnej (1.34) zmiennej Y wspomagane modelem, omówione w poprzednim podrozdziale, mają tę własność, że wyznaczony system wag kalibracyjnych nie odtwarza na ogół znanych wartości globalnych zmiennych pomocniczych x_1, \dots, x_J . Odtwarzają one jedynie liczebność populacji N oraz wartość globalną złożoną z wartości teoretycznych wynikających z przyjętego modelu. Z tego punktu widzenia są one zdecydowanie rzadziej wykorzystywane w praktycznych zastosowaniach, gdzie preferuje się jeden system wag zapewniający odtwarzanie wartości globalnych kluczowych zmiennych pomocniczych. Tak zbudowany system wag jest wykorzystywany do estymacji wartości globalnych pozostałych zmiennych w badaniu. W wypadku podejścia wspomaganego modelem wymagana jest natomiast każdorazowo konstrukcja odpowiedniego modelu łączącego zmienną Y ze zmiennymi objaśniającymi, co z praktycznego punktu widzenia jest istotną niedogodnością.

Lehtonen i Veijanen (2015; 2017; 2018) zaproponowali podejście kalibracyjne, w którym łączy się korzyści wynikające ze stosowania klasycznej kalibracji (wolnej od modelu) opisaną w rozdziale drugim z podejściem wspomaganym modelem, któremu poświęcono podrozdział 4.2. Tego typu podejście w procesie wyznaczania wag kalibracyjnych nazwać będziemy hybrydowym (ang. *hybrid calibration*).

W kolejnej części pracy przedstawiono estymatory hybrydowe w podejściu kalibracyjnym. Załóżmy, że jak poprzednio, przedmiotem zainteresowania jest oszacowanie wartości globalnej zmiennej Y , która wyraża się wzorem (1.34). Zakładamy ponadto, że dysponujemy zmiennymi pomocniczymi x_1, \dots, x_J , które wykorzystywać będziemy w procesie wyznaczania wag kalibracyjnych. Podobnie jak wcześniej, $\mathbf{x}_k = (x_{k1}, \dots, x_{kJ})^T$ oznacza wektor złożony z wartości wszystkich zmiennych pomocniczych dla k -tej jednostki badania. Wektor ten dzielimy na dwa:

$$\mathbf{x}_k^{(1)} = (x_{k1}, \dots, x_{kj})^T \quad (4.34)$$

oraz

$$\mathbf{x}_k^{(2)} = (x_{k,j+1}, \dots, x_{kJ})^T. \quad (4.35)$$

Pierwszy wektor wykorzystany będzie w taki sposób, aby odtwarzane zostały w procesie kalibracji wartości globalne tych zmiennych. Z kolei drugi wektor zostanie wykorzystany jako zmienne pomocnicze w budowie odpowiedniego modelu, a w konsekwencji do wyznaczenia wartości modelowych $\hat{\mu}_k$. Warto zwrócić uwagę, że w podejściu tym zakładamy, że wektory (4.34) oraz (4.35) nie mają wspólnych zmiennych. Można jednak dopuścić sytuację, w której ta sama zmienna pomocnicza występować będzie w podwójnej roli: jako zmienna, dla której wyznaczone wagi kalibracyjne odtwarzać będą jej znaną wartość globalną, a z drugiej strony jako jedna ze zmiennych objaśniających wykorzystywanych na etapie budowy odpowiedniego modelu. W opisywanym podejściu zakładamy jednak, że konkretną zmienną pomocniczą wykorzystujemy tylko raz: do budowy wektora (4.34) albo wektora (4.35).

Zakładając, że $\hat{\mu}_k = \hat{\mu}_k(\mathbf{x}_k^{(2)}, \hat{\boldsymbol{\beta}})$ są wartościami teoretycznymi (modelowymi) wynikającymi z przyjętego modelu, w którym oszacowano parametr $\boldsymbol{\beta}$ zgodnie ze wzorem (4.2), (4.26) lub (4.30), wykorzystując w tym celu zmienne x_{j+1}, \dots, x_J , wagi hybrydowo-kalibrowane $w_{k\text{HC}}$ w estymatorze kalibracyjnym $\hat{\tau}_{\text{HC}}$ wartości

globalnej (1.34) w podejściu hybrydowym znajduje się poprzez rozwiązanie następującego zadania warunkowej minimalizacji³⁴:

- (W1)_{HC} – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s} \frac{d_k}{q_k} G\left(\frac{w_{k\text{HC}}}{d_k}\right) \rightarrow \min, \quad (4.36)$$

- (W2)_{HC} – równania kalibracyjne:

$$\sum_{k \in s} w_{k\text{HC}} \mathbf{u}_{k_2} = \sum_{k \in U} \mathbf{u}_{k_2}, \quad (4.37)$$

gdzie wektor \mathbf{u}_{k_2} jest postaci:

$$\mathbf{u}_{k_2} = (1, \hat{\mu}_k, x_{k_1}, \dots, x_{k_j})^T. \quad (4.38)$$

- (W3)_{HC} – warunki ograniczające:

$$L \leq \frac{w_{k\text{HC}}}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, n, \quad (4.39)$$

W opisanym powyżej zadaniu poszukiwania wag hybrydowo-kalibrowanych wektor wartości globalnych występujących w równaniu (4.37) jest zatem postaci:

$$\sum_{k \in U} \mathbf{u}_{k_2} = \left(N, \sum_{k \in U} \hat{\mu}_k, \sum_{k \in U} x_{k_1}, \dots, \sum_{k \in U} x_{k_j} \right)^T. \quad (4.40)$$

Wektor wag hybrydowo-kalibrowanych $w_{k\text{HC}}$ sumuje się nie tylko do znanej liczebności populacji U , ale także odtwarza wartość globalną złożoną z wartości teoretycznych dla założonego modelu oraz wartości globalne wszystkich zmiennych pomocniczych x_1, \dots, x_j , tj. wszystkich zmiennych wchodzących w skład wektora (4.34). Warto zwrócić uwagę, że w wypadku podejścia hybrydowego w kalibracji niezbędna jest znajomość wartości jednostkowych zmiennych x_{j+1}, \dots, x_J na poziomie całej populacji U , podczas gdy dla zmiennych x_1, \dots, x_j wystarczy ich znajomość na poziomie próby s oraz potrzebna jest informacja o wartościach globalnych tych zmiennych (na przykład ze spisu czy z odpowiedniego rejestru).

Podobnie jak wcześniej, wagi hybrydowo-kalibrowane $w_{k\text{HC}}$ można wyznaczyć wprost ze wzoru dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ oraz biorąc pod uwagę warunki (W1)_{HC} i (W2)_{HC}. Mówi o tym poniższe twierdzenie.

³⁴ Subskrypt HC pochodzi od angielskiego wyrażenia *hybrid calibration* – kalibracja hybrydowa.

Twierdzenie 4.3. Rozwiązaniem zadania minimalizacji funkcji odległości (4.36) przy warunku (4.37) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag hybrydowo-kalibrowanych $\mathbf{w}_{\text{HC}} = (w_{1\text{HC}}, \dots, w_{n\text{HC}})^T$, którego składowe, dla $k = 1, \dots, n$, wyrażają się wzorem:

$$w_{k\text{HC}}^{\text{lin}} = d_k + d_k q_k \left(\sum_{k \in U} \mathbf{u}_{k_2} - \sum_{k \in s} d_k \mathbf{u}_{k_2} \right)^T \left(\sum_{k \in s} q_k d_k \mathbf{u}_{k_2} \mathbf{u}_{k_2}^T \right)^{-1} \mathbf{u}_{k_2}, \quad (4.41)$$

gdzie wektor \mathbf{u}_{k_2} określony jest wzorem (4.38), a wektor wartości globalnych $\sum_{k \in U} \mathbf{u}_{k_2}$ wzorem (4.40).

Oczywiście, w podejściu hybrydowym nic nie stoi na przeszkodzie, aby wagi wyznaczać także dla innych funkcji odległości, uwzględniając dodatkowo ograniczenia nakładane na iloraz wag w warunku (W3)_{HC}. W tym celu należy skorzystać z metod konstruowania wag opisanych w algorytmach 1 lub 2, w których kluczową rolę odgrywa sposób znajdowania mnożnika Lagrange'a λ .

W związku z uzyskanym wektorem wag hybrydowo-kalibrowanych można sformułować następującą definicję estymatora kalibracyjnego wartości globalnej w podejściu hybrydowym.

Definicja 4.3. Estymatorem kalibracyjnym wartości globalnej (1.34) w podejściu hybrydowym (estymatorem hybrydowo-kalibrowanym) jest:

$$\hat{\tau}_{\text{HC}} = \sum_{k \in s} w_{k\text{HC}} y_k, \quad (4.42)$$

gdzie wagi hybrydowo-kalibrowane $w_{k\text{HC}}$ są rozwiązaniem zadania optymalizacyjnego (4.36)–(4.39). W szczególnym przypadku, dla funkcji odległości zdefiniowanej wzorem (2.18) z uwzględnieniem warunku (4.37), wyrażają się one wzorem (4.41).

Jako estymator wariancji estymatora hybrydowo-kalibrowanego, który wyraża się wzorem (4.42), można wykorzystać estymator postaci (2.43), gdzie w miejsce wag w_k i wektora \mathbf{x}_k należy podstawić $w_{k\text{HC}}$ oraz wektor \mathbf{u}_{k_2} odpowiednio, przy czym wektor \mathbf{u}_{k_2} wyraża się wzorem (4.38).

Również w wypadku podejścia hybrydowego można dokonać konstrukcji wektora wag hybrydowo-kalibrowanych dla odpowiednio zdefiniowanych domen. Problem poszukiwania wag hybrydowo-kalibrowanych można byłoby zapisać w sposób analogiczny do zadania (2.53)–(2.55). W miejsce wektora \mathbf{x}_k należałoby podstawić wektor \mathbf{u}_{k_2} . W takim ujęciu trzeba znać wartości globalne zmiennych x_1, \dots, x_j na poziomie domen.

4.5. Badanie symulacyjne – podejście kalibracyjne wspomagane modelem

Głównym celem przeprowadzonego badania symulacyjnego jest analiza najważniejszych własności estymatorów wartości globalnej (1.34) wspomaganych modelem, które zostały opisane w podrozdziałach 4.2–4.4. Dodatkowo ocenie poddano wagi modelowo-kalibrowane dla każdego estymatora wziętego pod uwagę w badaniu. Na potrzeby analizy wykorzystano rzeczywiste dane pochodzące z BBGD. Sposób utworzenia pseudopopulacji opisano w podrozdziale 2.4 oraz rozszerzono na pozostałe zmienne w podrozdziale 3.3. Jako zmienną objaśnianą Y w badaniu symulacyjnym wzięto wydatki ogółem gospodarstwa domowego (w tys. zł), a potencjalnymi zmiennymi pomocniczymi były, podobnie jak wcześniej, dochód rozporządzalny gospodarstwa domowego (x_1 – w tys. zł), powierzchnia użytkowa mieszkania zajmowanego przez gospodarstwo domowe (x_2 – w m^2) oraz liczba osób w gospodarstwie domowym (x_3).

Na potrzeby oszacowania wartości globalnej łącznych wydatków gospodarstw domowych uwzględniono 12 estymatorów³⁵:

- 1) estymator Horwitza-Thompsona (HT) dany wzorem (1.35),
- 2) estymator kalibracyjny (CAL) określony wzorem (2.3), w którym wzięto pod uwagę zmienne pomocnicze x_1 oraz x_2 ,
- 3) estymator kalibracyjny wspomagany modelem liniowym (MCAL11) postaci (4.10), w którym jedyną zmienną pomocniczą był dochód – x_1 ,
- 4) estymator kalibracyjny wspomagany modelem liniowym (MCAL12) postaci (4.10), w którym wzięto pod uwagę wszystkie zmienne pomocnicze x_1 , x_2 oraz x_3 ,
- 5) estymator kalibracyjny wspomagany modelem potęgowym (MCAL21) postaci (4.10), w którym jedyną zmienną pomocniczą był dochód – x_1 ,
- 6) estymator kalibracyjny wspomagany modelem potęgowym (MCAL22) postaci (4.10), w którym wzięto pod uwagę wszystkie zmienne pomocnicze x_1 , x_2 oraz x_3 ,
- 7) estymator kalibracyjny wspomagany adaptacyjnym modelem regresji LASSO (ALASSO), w którym wzięto pod uwagę wszystkie zmienne pomocnicze x_1 , x_2 , x_3 , przyjęto parametr regularyzacji α_j postaci (4.28) oraz $\gamma = 0, 1$,

³⁵ W nawiasach podano skrótove nazwy estymatorów, wykorzystywane w dalszej części podrozdziału.

- 8) estymator kalibracyjny wspomagany modelem regresji grzbietowej (RIDGE), w którym wzięto pod uwagę wszystkie zmienne pomocnicze x_1 , x_2 oraz x_3 ,
- 9) estymator hybrydowo-kalibrowany (HYBRID1) określony wzorem (4.42), w którym uwzględniono model liniowy, a także przyjęto, że $\mathbf{x}_k^{(1)} = x_2$ oraz $\mathbf{x}_k^{(2)} = x_1$ ³⁶,
- 10) estymator hybrydowo-kalibrowany (HYBRID2) określony wzorem (4.42), w którym uwzględniono model potęgowy, a także przyjęto, że $\mathbf{x}_k^{(1)} = x_2$ oraz $\mathbf{x}_k^{(2)} = x_1$,
- 11) estymator hybrydowo-kalibrowany (HYBALASSO), zdefiniowany wzorem (4.42), w którym uwzględniono adaptacyjny model regresji LASSO, a także przyjęto, że $\mathbf{x}_k^{(1)} = x_3$ oraz $\mathbf{x}_k^{(2)} = (x_1, x_2)^T$,
- 12) estymator hybrydowo-kalibrowany (HYBRIDGE), zdefiniowany wzorem (4.42), w którym uwzględniony został model regresji grzbietowej, a także przyjęto, że $\mathbf{x}_k^{(1)} = x_3$ oraz $\mathbf{x}_k^{(2)} = (x_1, x_2)^T$.

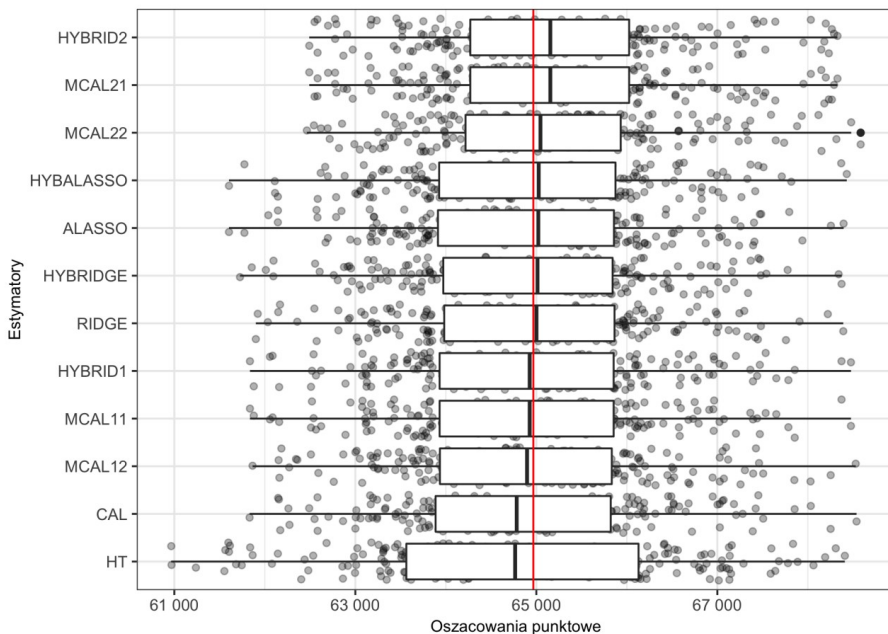
Spśród dwunastu wymienionych estymatorów jedynie dwa pierwsze nie wykorzystują w swojej konstrukcji modelu. Uwzględniono je w celu porównań z wynikami uzyskanymi dla pozostałych estymatorów. W tym miejscu warto podkreślić, że w procesie konstrukcji estymatorów modelowo-kalibrowanych rozpatrywano różne postacie modeli: liniowy, potęgowy, regresji grzbietowej i LASSO. Miało to na celu sprawdzenie, w jakim stopniu postać modelu może wpływać na otrzymane wyniki. W wypadku estymatorów 2–12, rozwiązując odpowiednie zadanie minimalizacji funkcji odległości, przyjęto funkcję w postaci sinusa hiperbolicznego, którą szczegółowo omówiono w podrozdziale 2.2. Nałożono przy tym ograniczenia na iloraz wag w postaci wartości parametrów $L = \frac{1}{2}$ oraz $U = \frac{3}{2}$. Wybór funkcji podyktowany był tym, że dla najczęściej wykorzystywanej w praktyce liniowej funkcji odległości dla niektórych estymatorów otrzymywano ujemne wagi kalibracyjne.

Na potrzeby oceny jakości rozważanych estymatorów z pseudopopulacji losowano 2% próbę zgodnie z planem losowania prostego bez zwracania oraz dokonywano kalibracji wag d_k . Proces ten powtórzono $B = 200$ razy, przy czym w każdej replikacji, mając wyznaczone odpowiednie wagi kalibracyjne (w wypadku estymatora CAL) lub modelowo-kalibrowane (dla pozostałych estymatorów), dokonano oszacowania wartości globalnej Y , tj. łącznych wydatków ogółem gospodarstw domowych. W dalszym kroku dla wszystkich estymatorów obli-

³⁶ Oznacza to, że model budowany był między wydatkami i dochodami, a dodatkowo wagi były w taki sposób konstruowane, aby odtworzyć wartość globalną powierzchni użytkowej mieszkania.

czono następujące miary: $RB_{SIM}(\hat{\tau})$, $MSE_{SIM}(\hat{\tau})$, $RRMSE_{SIM}(\hat{\tau})$, $D_{SIM}^2(\hat{\tau})$ oraz $REE_{SIM}(\hat{\tau})$, które zostały opisane we wzorach (3.36)–(3.40) w podrozdziale 3.3³⁷.

Rysunek 4.1 pokazuje oszacowania punktowe wartości globalnej wydatków ogółem gospodarstw domowych uzyskane dla dwunastu rozpatrywanych estymatorów we wszystkich 200 replikacjach. Czerwona pionowa linia oznacza prawdziwą wartość globalną τ_Y w utworzonej pseudopopulacji. Punkty na każdym z wykresów odpowiadają realizacjom konkretnego estymatora w każdej z 200 replikacji. Analiza tego wykresu wskazuje, że każdy z rozpatrywanych estymatorów charakteryzuje się niewielkim obciążeniem.



Rysunek 4.1. Oszacowana wartość globalna wydatków ogółem gospodarstw domowych w podejściu modelowo-kalibrowanym

Źródło: na podstawie BBGD.

Jak pokazuje tabela 4.2, względne obciążenie wszystkich dwunastu estymatorów jest małe i oscyluje w granicach zera. Analizując MSE_{SIM} oraz $RRMSE_{SIM}$ rozpatrywanych estymatorów, można jednak zauważyć, że wszystkie estymatory kalibracyjne mają niższe wartości tych miar jakości w porównaniu z estymatorem

³⁷ W miejsce estymatora $\hat{\tau}_{CALF}$ należy podstawić jeden z dwunastu rozważanych estymatorów.

Horvitz-Thompsona. Estymatory kalibracyjne charakteryzują się zatem większą dokładnością aniżeli estymator Horvitz-Thompsona, który w swojej konstrukcji nie wykorzystuje żadnej informacji o zmiennych dodatkowych.

Tabela 4.2. Wartości miar jakości estymatorów w podejściu modelowo-kalibrowanym dla funkcji odległości w postaci sinusa hiperbolicznego

Estymator	RB_{SIM}	MSE_{SIM}	$RRMSE_{SIM}$	D_{SIM}^2	REE_{SIM}
MCAL22	0,002	1 669 309,29	1,989	1 645 829,66	1,970
RIDGE	-0,000	1 723 018,89	2,020	1 722 765,66	2,021
MCAL12	-0,001	1 745 620,73	2,034	1 744 367,74	2,034
HYBRIDGE	-0,000	1 755 154,54	2,039	1 755 011,41	2,040
ALASSO	-0,000	1 767 071,40	2,046	1 766 859,50	2,046
HYBALASSO	-0,000	1 784 017,48	2,056	1 783 911,39	2,056
MCAL11	-0,001	1 787 312,17	2,058	1 785 287,70	2,058
HYBRID1	-0,001	1 787 312,79	2,058	1 785 288,31	2,058
HYBRID2	0,003	1 790 285,14	2,060	1 753 155,22	2,032
MCAL21	0,003	1 790 285,59	2,060	1 753 155,61	2,032
CAL	-0,001	1 792 193,68	2,061	1 787 095,83	2,060
HT	-0,002	2 809 108,35	2,580	2 792 851,82	2,577

Objaśnienia: Wartości RB_{SIM} , REE_{SIM} oraz $RRMSE_{SIM}$ wyrażone są w procentach.

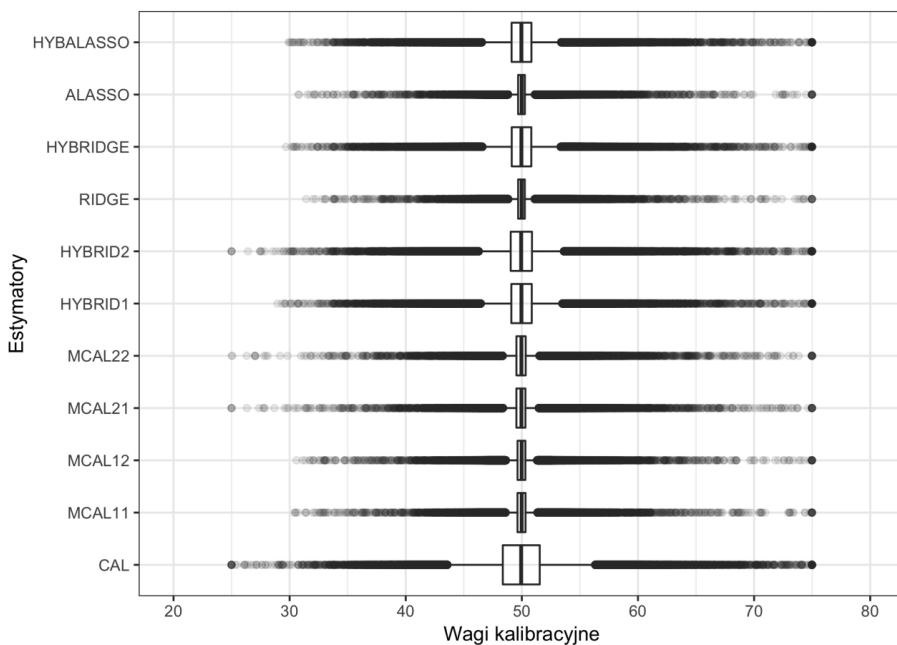
Źródło: na podstawie BBGD.

W grupie rozważanych estymatorów kalibracyjnych, estymatory modelowo-kalibrowane odznaczały się nieco lepszymi własnościami w porównaniu z estymatorem kalibracyjnym CAL, który nie był wspomagany modelem. Uwzględnienie w procesie estymacji wydatków ogółem gospodarstw domowych odpowiednio zbudowanego modelu poprawia zatem dokładność szacunku. Najniższe wartości MSE_{SIM} i $RRMSE_{SIM}$ w grupie estymatorów modelowo-kalibrowanych zaobserwowano wyraźnie w wypadku estymatora MCAL22, który wspomagany był modelem potęgowym uwzględniającym wszystkie zmienne pomocnicze. Najwyższe wartości MSE_{SIM} i $RRMSE_{SIM}$ odnotowano z kolei dla estymatorów MCAL21 oraz HYBRID2, które również wspomagane były modelem potęgowym, ale tylko z uwzględnieniem dochodu jako jedynej zmiennej objaśniającej w modelu. Warto podkreślić, że w wypadku estymatora HYBRID2, w porównaniu z MCAL21, wyznaczone wagi modelowo-kalibrowane dodatkowo odtwarzały wartość globalną powierzchni użytkowej mieszkania.

Dokonując szczegółowej analizy rysunku 4.1 oraz danych przedstawionych w tabeli 4.2, można dodatkowo zauważyć, że w wypadku estymatora Horvitz-Thompsona obserwowane jest większe zróżnicowanie uzyskanych oszacowań w poszczególnych replikacjach badania symulacyjnego w porównaniu z po-

zostałymi estymatorami. Względny średni błąd szacunku dla tego estymatora był na poziomie 2,6%, podczas gdy w wypadku estymatorów kalibracyjnych oscylował on w granicach 2%. Również tutaj można zaobserwować, że wspomaganie się w procesie estymacji odpowiednim modelem wpływa na poprawę precyzji oszacowań. Jest to szczególnie zauważalne w wypadku wspomnianego już estymatora MCAL22 oraz RIDGE wspomaganego modelem regresji grzbietowej.

Kluczową kwestią w ocenie własności rozważanych estymatorów jest również analiza uzyskanych wag kalibracyjnych. Ich rozkład przedstawiono na rysunku 4.2, a najważniejsze statystyki opisowe ujęte zostały w tabeli 4.3.



Rysunek 4.2. Rozkład wag kalibracyjnych w $B=200$ replikacjach w podejściu modelowym dla funkcji odległości w postaci sinusa hiperbolicznego

Źródło: na podstawie BBGD.

Można zauważyć, że najbardziej zróżnicowane wagi otrzymano w wypadku estymatora kalibracyjnego CAL, w którym nie wzięto pod uwagę żadnego modelu, oraz w dalszej kolejności dla estymatorów HYBRID2, MCAL21 oraz MCAL22. Z kolei dla estymatorów wspomaganých adaptacyjnym modelem regresji LASSO i modelem regresji grzbietowej wagi były w najmniejszym stopniu zróżnicowane.

Tabela 4.3. Statystyki opisowe wag w podejściu modelowo-kalibrowanym dla funkcji odległości w postaci sinusa hiperbolicznego

Estymator	Min	Q ₁	Q ₂	Średnia	Q ₃	Max
CAL	24,993	48,357	49,939	49,986	51,540	74,979
HYBRID2	24,993	49,037	49,940	49,986	50,872	74,979
MCAL21	24,993	49,530	49,963	49,986	50,306	74,979
MCAL22	24,993	49,533	49,962	49,986	50,328	74,979
HYBRID1	28,950	49,095	49,962	49,986	50,853	74,979
HYBRIDGE	29,683	49,127	49,960	49,986	50,813	74,979
HYBALASSO	29,913	49,115	49,959	49,986	50,824	74,979
MCAL11	30,408	49,623	49,983	49,986	50,315	74,979
MCAL12	30,590	49,626	49,981	49,986	50,302	74,979
ALASSO	30,756	49,677	49,986	49,986	50,260	74,979
RIDGE	31,418	49,678	49,981	49,986	50,254	74,979

Objaśnienia: Q₁ – kwartył pierwszy, Q₂ – kwartył drugi, Q₃ – kwartył trzeci.

Źródło: na podstawie BBGD.

Przyjęcie ograniczeń w postaci wspomnianych już wartości brzegowych $L = \frac{1}{2}$ oraz $U = \frac{3}{2}$ miało wpływ na zakres zmienności wag dla wszystkich estymatorów. Wagi znajdowały się w przedziale [24,993, 74,979]. W wypadku czterech estymatorów osiągnięte zostały obydwa końce danego przedziału. Nałożenie ograniczeń na wagi było jednak niezbędnym zabiegiem, gdyż w przeciwnym razie pojawiały się wagi ujemne.

4.6. Wnioski

W badaniach statystycznych prowadzonych przez krajowe urzędy statystyczne podejście randomizacyjne, w którym centralną rolę odgrywają wagi wynikające z przyjętego planu losowania próby, jest powszechnie stosowane w praktyce. Obserwuje się to od czasów Neymana, który wniósł fundamentalny wkład w rozwój metody reprezentacyjnej i zalecał jej wykorzystanie w procesie estymacji parametrów (Neyman, 1934). W podejściu randomizacyjnym kluczową rolę odgrywa kalibracja, której zastosowanie sprowadza się do odpowiedniej korekty wag wynikających z planu losowania próby tak, aby odtworzyć znane wartości globalne wszystkich istotnych zmiennych pomocniczych.

Oprócz wspomnianego podejścia randomizacyjnego, które ma ugruntowaną pozycję w praktyce badań próbkowych, funkcjonuje podejście wspomagane modelem i modelowe, którego wykorzystanie jest obecnie coraz bardziej powszechne i akceptowane. Dotyczy to wielu obszarów badawczych odnoszących się do esty-

macji pośredniej, zagadnień związanych z brakami danych czy metod estymacji wariancji (Kalton, 2002). Wykorzystanie modeli, które pełnią funkcję wspomagającą proces estymacji, jest także zauważalne w kontekście podejścia kalibracyjnego i zostało zapoczątkowane przez Wu i Sittera (2001). Jest ono obecnie dynamicznie rozwijającym się nurtem badań naukowych, wśród których główna uwaga jest skupiona na estymacji hybrydowej oraz wspomaganego modelem regresji grzbietowej czy LASSO. Warto przy tym zaznaczyć, że rozważania w tym obszarze dotyczą nie tylko prób losowych, ale także nielosowych, które odgrywają coraz większą rolę w statystyce publicznej.

Niniejszy rozdział stanowi podsumowanie aktualnego dorobku w kontekście podejścia kalibracyjnego wspomaganego modelem. Omówiono w nim estymatory modelowo-kalibrowane wartości globalnej zarówno w populacji generalnej, jak i w odpowiednio zdefiniowanych domenach. Dyskusji poddano również sposób konstruowania wariancji rozważanych estymatorów. Uwaga została skupiona również na estymatorach kalibracyjnych wspomaganym modelem regresji grzbietowej czy LASSO oraz podejściu hybrydowym, które stanowią najnowsze osiągnięcia w tym zakresie.

W badaniu symulacyjnym, wykorzystując rzeczywiste dane pochodzące z Badania Budżetów Gospodarstw Domowych, dokonano oceny najważniejszych własności rozważanych estymatorów modelowo-kalibrowanych wartości globalnej. Uwaga poświęcona była także analizie wag kalibracyjnych. W badaniu tym rozważono również dwie propozycje estymatorów hybrydowo-kalibrowanych, wspomaganym modelem regresji grzbietowej (HYBRIDGE) i adaptacyjnym LASSO (HYBALASSO).

Podsumowując wyniki uzyskane w przeprowadzonym badaniu symulacyjnym, należy zaznaczyć, że estymatory modelowo-kalibrowane charakteryzują się lepszymi własnościami w porównaniu z wykorzystywanym w praktyce estymatorem Horvitz-Thompsona oraz estymatorem kalibracyjnym, który w swojej konstrukcji nie uwzględniał żadnego modelu. Dla wszystkich rozpatrywanych estymatorów wspomaganym modelem, bez względu na przyjęty model, zaobserwowano bowiem niższe wartości najważniejszych miar jakości estymacji. Również finalne wagi wykorzystywane w estymatorach modelowo-kalibrowanych, w porównaniu z estymatorem kalibracyjnym, który nie był wspomagany modelem, na ogół charakteryzowały się mniejszym zróżnicowaniem. Jest to pożądana sytuacja, gdyż po kalibracji wagi wynikające z przyjętego planu losowania próby w mniejszym stopniu podlegały modyfikacjom. Z tego punktu widzenia kluczowym aspektem, który należy uwzględnić każdorazowo w procesie kalibracji wag, jest odpowiedni dobór funkcji odległości oraz ograniczeń nakładanych na iloraz wag.

Należy także wyraźnie podkreślić, że mimo zauważonych zalet estymatorów modelowo-kalibrowanych, odnoszących się do niższych względnych błędów średniokwadratowych oraz względnych średnich błędów szacunku, mają one pewne ograniczenia, których przewyciężenie będzie istotnym krokiem naprzód w ich bieżącym wykorzystaniu w praktyce badań statystycznych. Ograniczenia te są związane z tym, że w badaniach preferowany jest jeden system wag, który może być wykorzystany w procesie estymacji parametrów dla wielu zmiennych jednocześnie. W podejściu wspomaganym modelem taki system wag należałoby każdorazowo budować dla różnych zmiennych, co w połączeniu z dyskusją nad wyborem odpowiedniego modelu mogłoby się przełożyć na skomplikowanie stopnia procesu estymacji. Pewne remedium stanowią mogą estymatory hybrydowo-kalibrowane, w tym zaproponowane estymatory hybrydowe wspomagane adaptacyjnym modelem regresji LASSO i grzbietowej, w wypadku których wagi są tak skonstruowane, aby odtwarzane były wartości globalne zmiennych pomocniczych, których nie wzięto pod uwagę w procesie budowania modelu. Można także rozważyć w tym kontekście wspomnianą już w podrozdziale 1.5 propozycję sformułowaną przez Montanariego i Ranalli (2009), dotyczącą utworzenia jednego – uniwersalnego systemu wag kalibracyjnych.

Podjęcie kalibracyjne w badaniach z brakami odpowiedzi

5.1. Wprowadzenie

W badaniach statystycznych prowadzonych przez krajowe urzędy statystyczne na całym świecie do jednych z najpoważniejszych problemów należy zaliczyć brak możliwości uzyskania odpowiedzi od jednostek wylosowanych do próby (Särndal i Lundström, 2005; Szreder, 2015b). W literaturze przedmiotu przez brak odpowiedzi³⁸ należy rozumieć sytuację, w której pomimo próby kontaktu ze strony ankietera nie jest możliwe uzyskanie od respondenta odpowiedzi na pytania zawarte w ankiecie, niezależnie od przyczyny (AAPOR, 2016). Przedstawiona definicja odnosi się przede wszystkim do badań prowadzonych przez krajowe urzędy statystyczne realizowanych z wykorzystaniem ankietera. Można ją jednak także odnosić do badań, które wykorzystują inne kanały pozyskiwania informacji, bez udziału ankietera (na przykład badania realizowane przez internet bazujące na technice Web Surveys czy e-mailingu), w których nie jest możliwe uzyskanie odpowiedzi na zawarte pytania.

³⁸ Zamiennie używany będzie termin brak danych.

Konieczność redukcji kosztów, rosnące obciążenia respondentów oraz to, że znane z klasycznej metody reprezentacyjnej estymatory (na przykład wspomniany wcześniej estymator Horvitz-Thompsona) charakteryzują się dużą wariancją i obciążeniem w wypadku istnienia braków danych, wymusiła potrzebę rozwoju odpowiednich metod statystycznych, które mogą stanowić remedium na występujące w każdym badaniu braki odpowiedzi. W związku z powyższym w światowej literaturze przedmiotu zjawisko braków danych przeszło istotną ewolucję od podejścia polegającego na ignorowaniu występowania w badaniu braków odpowiedzi poprzez nowoczesne metody estymacji niwelujące ujemny wpływ tych błędów nielosowych.

W rozdziale tym uwagę skupiono na możliwości wykorzystania podejścia kalibracyjnego w badaniach statystycznych z brakami odpowiedzi. Występowanie braków danych dotyczy wszystkich typów badań: zarówno pełnych, które obejmują całą populację (spisy, niektóre badania ankietowe), jak i badań częściowych (Badanie Aktywności Ekonomicznej Ludności, Badanie Budżetów Gospodarstw Domowych, Europejskie Badanie Warunków Życia Ludności itp.).

Od strony teoretycznej opisano estymatory kalibracyjne wartości globalnej w bardziej realnej sytuacji, tj. gdy w badaniu występują braki danych, które mają wpływ na jakość finalnych oszacowań. Dotyczy to zarówno szacowania wartości globalnej pewnej zmiennej w populacji, jak i w odpowiednio określonych domkach. Omówiono też koncepcję tzw. jednokrokowego i dwukrokowego sposobu wyznaczania wag kalibracyjnych, gdy w badaniu występują braki danych.

Wykorzystując rzeczywiste dane pochodzące z Badania Aktywności Ekonomicznej Ludności, dyskusji poddano wyniki estymacji liczby osób pracujących, bezrobotnych oraz biernych zawodowo w przekroju podregionów, tj. na poziomie, dla którego brakuje oficjalnych szacunków publikowanych przez Główny Urząd Statystyczny. W tym celu zastosowano omówione w tym rozdziale estymatory kalibracyjne wartości globalnej dla różnych funkcji odległości. Analizie poddano zarówno uzyskane wagi, jak i precyzję oszacowań.

5.2. Braki odpowiedzi jako źródło błędów nielosowych

Problem braków odpowiedzi w badaniach statystycznych, jako najważniejszej kategorii błędów nielosowych, był dostrzegany od wczesnych lat 40. XX wieku i wraz z rozwojem metody reprezentacyjnej wzbudzał rosnące zainteresowanie środowiska naukowego oraz praktyków. Colley (1945), Deming (1953), Ferber

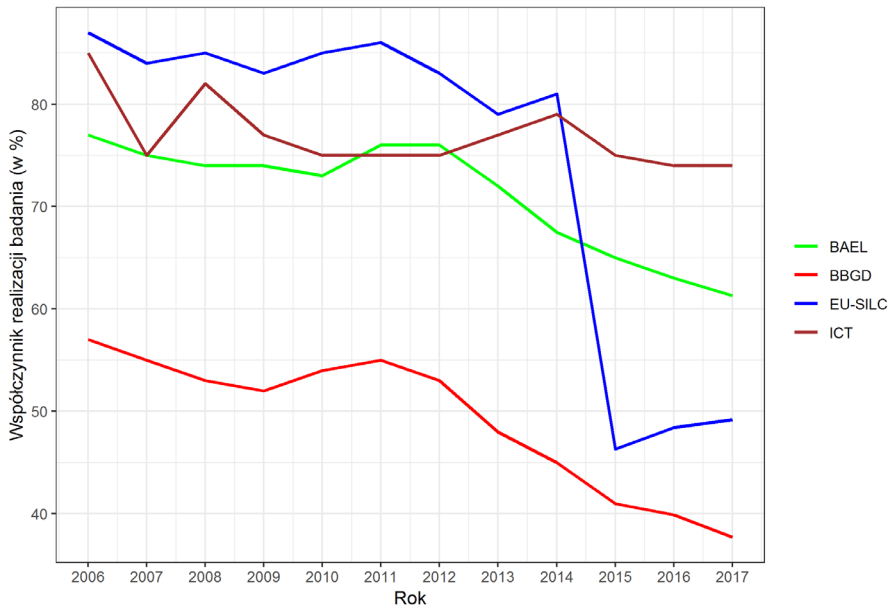
(1949), Hansen i Hurwitz (1946) czy Yates (1946) byli pierwszymi statystykami, którzy zwrócili uwagę na wpływ braków danych na jakość oszacowań. Jednocześnie wskazali na metody statystyczne odnoszące się do ważenia, które mogą stanowić remedium na to niekorzystne zjawisko występujące w badaniach statystycznych. Kompleksowy opis tych metod, uwzględniający m.in. podejście kalibracyjne, można znaleźć w przeglądowym artykule Kaltona i Floresa-Cervantes (2003). Zagadnienie braków danych w badaniach statystycznych opisywało wielu autorów, na przykład Bethlehem, Cobben i Schouten (2011), Brehm (1993), Goyder (1987), Groves i Couper (2012), Groves, Dillman, Eltinge i Little (2002), Kalton (1983), Kordos (1988), Laaksonen (2018), Longford (2006), Särndal i Lundström (2005), Stoop (2005) czy Stoop, Billiet, Koch i Fitzgerald (2010). Wielu autorów opublikowało na ten temat dziesiątki artykułów w specjalistycznych czasopismach statystycznych, w tym artykuły przeglądowe, np. Brick (2013), Chen i współautorzy (2017), Haziza i Beaumont (2017), Haziza i Lesage (2016), Särndal (2007), Szreder (2011) czy Wu i Lu (2016). Poświęcone one były różnym aspektom braków odpowiedzi: mechanizmom ich powstawania, technikom zbierania danych w celu zwiększenia wskaźników kompletności czy statystycznym metodom niwelowania ujemnego wpływu braków danych na jakość oszacowań w procesie estymacji.

Ograniczając się do badań statystyki publicznej w Polsce, można stwierdzić, że problem braków odpowiedzi dotyczy w zasadzie wszystkich badań częściowych, w szczególności Badania Budżetów Gospodarstw Domowych (GUS, 2017a), Badania Aktywności Ekonomicznej Ludności (GUS, 2018a), Europejskiego Badania Dochodów i Warunków Życia (GUS, 2017b) czy badania przedsiębiorstw należących do sektora ICT (ang. *information and communication technology*) (GUS, 2017c).

W badaniach tych obserwuje się niepokojący trend związany ze zmniejszającym się współczynnikiem realizacji, zwanym również wskaźnikiem kompletności³⁹. Sytuację tę ilustruje rysunek 5.1, który pokazuje kształtowanie się współczynników kompletności dla czterech wspomnianych badań prowadzonych przez Główny Urząd Statystyczny w latach 2006–2017⁴⁰.

³⁹ Wskaźnik realizacji (wskaźnik kompletności) oznacza udział jednostek, które zostały przebadane i ukończyły badanie, w grupie wszystkich wylosowanych jednostek kwalifikujących się do uczestniczenia w badaniu.

⁴⁰ Ze względu na to, że Badanie Aktywności Ekonomicznej Ludności jest badaniem kwartalnym, odpowiednie roczne współczynniki kompletności zostały przedstawione jako średnie wartości kwartalne z danego roku.



Rysunek 5.1. Wskaźniki realizacji dla wybranych badań reprezentacyjnych

Źródło: na podstawie danych GUS.

Szczególnie niepokojący jest przypadek Badania Budżetów Gospodarstw Domowych, dla którego wskaźnik realizacji jest poniżej progu 40%. Nieco korzystniejsza sytuacja wygląda w wypadku Badania Aktywności Ekonomicznej Ludności i przedsiębiorstw, jednak i tutaj obserwuje się trend spadkowy wskaźnika kompletności. Podobne trendy występują w badaniach realizowanych przez inne kraje należące do Unii Europejskiej (Eurostat, 2015). Różne są przyczyny niepodejmowania udziału w badaniach przez gospodarstwa domowe. Wyjaśnieniu determinant występowania braków danych w badaniach gospodarstw domowych realizowanych w Polsce poświęcona jest praca Rószkiewicz (2015). Z kolei Cobben (2009) omawia ten problem na przykładzie badań prowadzonych przez holenderski urząd statystyczny. Warto jednak zaznaczyć, że literatura przedmiotu poświęcona zagadnieniu braków odpowiedzi w badaniach gospodarstw jest bardzo obszerna. Dotyczy to zarówno przyczyn niepodejmowania udziału w badaniu, jak i metod korygowania finalnych wyników (Fricker i Tourangeau, 2010; Groves, 2006; Groves i Couper, 2012).

Paradoksalnie, problem braków odpowiedzi dotyczy nie tylko badań reprezentacyjnych mających charakter dobrowolny, ale również badań pełnych, takich jak spisy czy sprawozdawczość statystyczna, w których udział jednostek jest czę-

sto obligatoryjny (Gołata, 2018; Hora, 2009). Na przykład, w badaniach towarzyszących spisowi, takich jak badanie dzietności kobiet w Narodowym Spisie Powszechnym 1970 czy 1988, braki odpowiedzi były na poziomie 30%. Z kolei w Narodowym Spisie Powszechnym Ludności i Mieszkań 2011 około 1,5 mln osób odmówiło odpowiedzi na temat niepełnosprawności (GUS, 2012). Również w rejestrach administracyjnych, które na potrzeby statystyki publicznej przekształcane są w rejestry statystyczne, braki danych stanowią poważne źródło błędów (Groen, 2012; Laitila, A. Wallgren i B. Wallgren, 2011; 2014). Na przykład, w rejestrze PESEL, będącym głównym źródłem informacji na temat wieku czy płci mieszkańców Polski, stan cywilny jest zmienną, dla której obserwuje się spory odsetek braków danych.

Mówiąc o brakach danych jako najważniejszym źródle błędów nielosowych, należy wyróżnić całkowite (ang. *unit nonresponse*) oraz częściowe (ang. *item nonresponse*) braki odpowiedzi (Särndal i Lundström, 2005; Yan i Curtin, 2010). Omówione rodzaje braków ukazuje w syntetyczny sposób tabela 5.1.

Tabela 5.1. Rodzaje braków danych w hipotetycznym badaniu statystycznym

Jednostka	Zmienne identyfikacyjne			Zmienne w kwestionariuszu		
	1	2	3	1	2	3
1	x	x	x	x	x	x
2	x	x	x	x	x	x
3	x	x	x	.	.	x
4	x	x	x	x	x	.
5	x	x	x	.	x	x
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$n - 1$	x	x	x	.	.	.
n	x	x	x	.	.	.

Objaśnienia: x – istniejące informacje; . – brak danych.

Źródło: na podstawie Särndal i Lundström (2005).

Z pierwszą sytuacją mamy do czynienia, gdy znane są jedynie dane identyfikacyjne jednostki (najczęściej adresowe), która powinna wziąć, a nie bierze udziału w badaniu, na przykład na skutek odmowy, utrudnionego kontaktu, rozpadu gospodarstwa domowego czy nieobecności podczas badania. Z kolei z drugą sytuacją mamy do czynienia, gdy jednostka bierze udział w badaniu, ale nie udziela odpowiedzi na niektóre zadane pytania. Najczęściej związane jest to z drażliwością tych pytań lub obawą, że udzielone odpowiedzi zostaną wykorzystane prze-

ciwko respondentowi (na przykład pytanie o dochody, kwestie etniczne, obyczajowe itp.). Przyczyny niewzięcia udziału w badaniu bądź udzielenia odpowiedzi tylko na niektóre pytania kwestionariusza mogą mieć charakter obiektywny (choroba, podeszły wiek, nieobecność w mieszkaniu w trakcie przeprowadzenia badania ankietowego czy zmiana miejsca zamieszkania) bądź subiektywny (niechęć do badania czy brak czasu).

Należy podkreślić, że bez względu na rodzaj braków danych i przyczyny ich powstania, ich występowanie w badaniu jest zazwyczaj źródłem wielu problemów, zwłaszcza w procesie estymacji. Wynika to z tego, że respondenci różnią się na ogół od nierespondentów ze względu na pewne kluczowe cechy⁴¹. Powoduje to w konsekwencji powstanie błędu systematycznego. Braki danych, stanowiące główną kategorię błędów nielosowych, mają ponadto wpływ na (Lundström i Särndal, 1999; Manski, 2016; Peytchev, 2013; Toepoel i Schonlau, 2017):

- efektywną liczebność badanej próby bądź populacji, przez co zwiększa się wariancja wykorzystywanych estymatorów, tj. zmniejsza precyzja oszacowań,
- obciążenie uzyskanych wyników – oszacowane parametry znacznie odbiegają od ich „prawdziwych” wartości, a wyznaczone przedziały ufności koncentrują się wokół niewłaściwych wartości,
- zniekształcenie rozkładów analizowanych cech, przez co estymacja parametrów jest utrudniona,
- zmniejszenie zaufania do wyników badania przez końcowych odbiorców danych statystycznych.

W praktyce badań statystycznych stosuje się wiele różnego rodzaju metod, które mają zapobiegać występowaniu braków danych lub wpływać na zwiększenie wskaźnika realizacji badania oraz poprawę procesu estymacji (Brick, 2013). Mają one zastosowanie zarówno na etapie zbierania danych (stosowanie bodźców materialnych), jak i ich opracowywania (techniki korygujące wyniki). Metody te można podzielić na trzy grupy:

- **Techniki prewencyjne** – które polegają przede wszystkim na zapobieganiu występowaniu w badaniu braków odpowiedzi zarówno całkowitych, jak i częściowych. Mają one zniwelować sceptycyzm oraz niechęć respondenta do wzięcia udziału w badaniu, a także promować pozytywne do niego nastawienie. Działania prewencyjne mogą również obejmować odpowiednie przeszkolenie ankierów czy właściwe przygotowanie kwestionariu-

⁴¹ W literaturze anglojęzycznej funkcjonuje termin *nonrespondent* na określenie jednostki, która nie bierze udziału w badaniu. W polskiej terminologii w zasadzie brak odpowiedniego tłumaczenia. W pracy na określenie takiej jednostki przyjmować będziemy termin „nierespondent”.

- sza ankietowego oraz operatu losowania. Ze względu na to, że metody prewencyjne wymagają kontaktu ankietera z respondentem, niezbędna jest tutaj wiedza z nauk o zachowaniu się jednostek (socjologii oraz psychologii). Literatura poświęcona działaniom prewencyjnym mającym niwelować frakcję braków odpowiedzi już na etapie planowania i zbierania danych przez ankieterów jest niezwykle bogata (Campanelli, 1997; Dillman, Eltinge, Groves i Little, 2002; Hidiroglou, Douglas i Gerald, 1993; Leeuw, Hox i Huisman, 2003; Toepoel i Schonlau, 2017).
- **Techniki redukujące** – które polegają na stosowaniu bodźców finansowych i materialnych (czynnik zachęcający jednostki do wzięcia udziału w badaniu), wysyłaniu monitów, ponownym kontakcie telefonicznym czy mailowym, zastępowaniu jednostek, które nie biorą udziału w badaniu, innymi o podobnych charakterystykach (na przykład z wylosowanej próby rezerwowej). Literatura przedmiotu poświęcona metodom redukującym frakcję braków odpowiedzi jest również niezwykle bogata (Brown i Calderwood, 2014; Groves, Cialdini i Couper 1992; Leeuw, Callegaro, Hox, Korendijk i Lensvelt-Mulders, 2007; Lund i Gulløy, 2016; Singer i Ye, 2013; Toepoel i Schonlau, 2017).
 - **Techniki korygujące** – obejmujące metody statystyczne, których zadaniem jest zniwelowanie obciążenia będącego skutkiem występowania w badaniu braków odpowiedzi. Do tej grupy metod należą zarówno imputacja, tj. metoda uzupełniania brakujących danych, jak i różnego rodzaju techniki korygowania wag wynikających z planu losowania próby (kalibracja). Literatura w obszarze statystycznych metod niwelowania wpływu braków danych na jakość oszacowań jest obszerna. Dotyczy to opracowań polskich i zagranicznych, zarówno w odniesieniu do imputacji (Balicki, 2004; Little i Rubin, 2014; Rubin, 2004; Wesołowski i Tarczyński, 2016; van Buuren, 2018), jak i metod ważenia danych, w tym kalibracji (Beaumont, 2005; Beręsewicz i Szymkowiak, 2011; Brick, 2013; Chang i Kott, 2008; Haziza i Beaumont, 2017; Haziza i Lesage, 2016; Kott, 2006; Kott i Chang, 2010; Kozłowski, 2016; Särndal, 2007; Särndal i Lundström, 2005; Särndal, Traat i Lumiste, 2018; Szymkowiak, 2009a; 2009b; Thibaudeau i Slud, 2009). Warto przy tym wspomnieć, że podstawową metodą niwelowania skutków jednostkowych braków danych (ang. *unit nonresponse*) jest odpowiednia korekta wag uogólniających (kalibracja), a imputacja na ogół jest wykorzystywana w wypadku pozycyjnych (ang. *item nonresponse*) braków danych (Piasecki, 2014). Jak jednak zauważają Särndal i Lundström (2005), można w badaniach zastosować podejście złożone (ang. *combined approach*), które

polega na jednoczesnym wykorzystaniu kalibracji w wypadku całkowitych braków danych i imputacji dla braków pozycyjnych. W pracy uwagę skupiono na podejściu kalibracyjnym w sytuacji, gdy jednostki nie biorą całkowicie udziału w badaniu.

5.3. Kalibracja wag – podejście jednokrokowe

Założmy, że celem badania jest oszacowanie wartości globalnej zmiennej Y , która wyraża się wzorem (1.34). W dalszym ciągu zakładamy, że z N -elementowej populacji U pobieramy n -elementową próbę s zgodnie z określonym planem jej losowania. Do oszacowania wartości globalnej (1.34) w praktyce badań statystycznych bardzo często stosuje się wspomniany już estymator Horvitz-Thompsona, który wyraża się wzorem (1.35). W wypadku braków danych (na przykład na skutek odmów jednostek wylosowanych do próby) estymator Horvitz-Thompsona charakteryzuje się na ogół znacznym obciążeniem i wariancją. Wynika to z tego, że braki danych nie mają najczęściej charakteru losowego, a błędy powstałe w procesie estymacji są konsekwencją różnic między respondentami a nierespondentami.

W związku z tym założmy bardziej realną sytuację, tj. gdy w badaniu występują jednostkowe braki danych. Ich istnienie może być następstwem odmów czy niezastania w domu osób przez ankietera. Niech $r \subseteq s$ oznacza zbiór respondentów, tj. jednostek, które wzięły udział w badaniu. Założmy ponadto, że jest to zbiór m -elementowy, przy czym $m \leq n$. W szczególnym przypadku może się zdarzyć, że $r = s$, tj. gdy wszystkie jednostki wylosowane do próby wzięły udział w badaniu. Wówczas $m = n$. W praktyce badań statystycznych taka sytuacja zdarza się jednak bardzo rzadko i na ogół $m < n$. W takim wypadku estymator Horvitz-Thompsona wartości globalnej (1.34) można przedstawić w postaci:

$$\hat{\tau}_{\text{HT}} = \sum_{k \in r} d_k y_k = \sum_{k=1}^m d_k y_k. \quad (5.1)$$

Sumowanie w powyższym wzorze odbywa się zatem po zbiorze wszystkich respondentów. Oszacowana wartość globalna (5.1), ze względu na braki danych, jest jednak zazwyczaj niedoszacowana w stosunku do wartości globalnej (1.34). Zgodnie z ideą kalibracji, wagi d_k powinny zostać zmodyfikowane w taki sposób, aby zrekompensować utratę informacji związaną z brakami danych. Zagadnienie poszukiwania wag kalibracyjnych $\mathbf{w} = (w_1, \dots, w_m)^T$ można sformułować analogicznie do rozdziału drugiego:

- (W1') – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in r} \frac{d_k}{q_k} G\left(\frac{w_k}{d_k}\right) \rightarrow \min, \quad (5.2)$$

- (W2') – równania kalibracyjne:

$$\sum_{k \in r} w_k \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k, \quad (5.3)$$

- (W3') – warunki ograniczające:

$$L \leq \frac{w_k}{d_k} \leq U, \text{ gdzie: } 0 \leq L \leq 1 \leq U, k = 1, \dots, m. \quad (5.4)$$

Zasadnicza różnica polega więc na tym, że wagi kalibracyjne poszukiwane są w zbiorze respondentów. Kluczowy w takim wypadku jest dostęp do odpowiedniej jakości zmiennych pomocniczych. Kalibrację wag można przeprowadzić w różny sposób, w zależności od poziomu, na jakim wykorzystywane są zmienne pomocnicze. W dalszej części rozpatrzono trzy kluczowe przypadki (Haziza i Lesage, 2016; Lundström i Särndal, 1999; Särndal i Lundström, 2005).

Przypadek 1. Zakładamy, że informacja na temat zmiennych pomocniczych jest dostępna na poziomie populacji U . Oznaczmy przez \mathbf{x}_k^* wektor o wymiarze $J^* \geq 1$ taki, że:

- 1) znane są wartości wektora zmiennych pomocniczych \mathbf{x}_k^* dla wszystkich jednostek z populacji U oraz znany jest wektor wartości globalnych postaci $\tau_{\mathbf{X}^*} = \sum_{k \in U} \mathbf{x}_k^*$,
- 2) dla każdego $k \in r$ wektor wartości zmiennych pomocniczych \mathbf{x}_k^* jest znany.

W podejściu tym zakładamy, że znany jest wektor wartości globalnych dla wszystkich zmiennych pomocniczych wchodzących w skład wektora \mathbf{x}_k^* . Jest to często spotykany w praktyce przypadek, w którym wartości globalne zmiennych pomocniczych znane są z zewnętrznych źródeł, takich jak spisy czy rejestry administracyjne. W podejściu tym zakładamy również znajomość wartości jednostkowych zmiennych pomocniczych na poziomie populacji U . W konsekwencji znane są wartości jednostkowe zmiennych pomocniczych \mathbf{x}_k^* dla wszystkich jednostek wylosowanych do próby s , ale również w zbiorze respondentów, tj. dla $k \in r$. Przykładami takich zmiennych, których wartości znane są często zarówno na poziomie populacji U , próby s , jak i zbioru respondentów r , są płeć, grupa wieku czy miejsce zamieszkania. W tym wypadku wektor wartości zmiennych pomocniczych \mathbf{x}_k jest zatem postaci: $\mathbf{x}_k = \mathbf{x}_k^*$.

Przypadek 2. Zakładamy, że informacja na temat zmiennych pomocniczych jest dostępna na poziomie próby s , ale nie jest znana na poziomie populacji U . Oznaczmy przez \mathbf{x}_k° wektor o wymiarze $J^\circ \geq 1$ taki, że:

1) znane są wartości wektora złożonego ze zmiennych pomocniczych \mathbf{x}_k° dla wszystkich jednostek z próby s , ale nie jest znany wektor wartości globalnych $\boldsymbol{\tau}_{\mathbf{X}^\circ} = \sum_{k \in U} \mathbf{x}_k^\circ$,

2) dla każdego $k \in r$ wektor wartości zmiennych pomocniczych \mathbf{x}_k° jest znany. W podejściu tym nie zakłada się zatem, że znane są wartości zmiennych pomocniczych \mathbf{x}_k° na poziomie całej populacji U . Przyjmuje się jednak, że wartości te znane są na poziomie całej próby, a w konsekwencji na poziomie zbioru respondentów r . Przypadek taki można również spotkać w praktyce badań statystycznych. Może się bowiem zdarzyć, że jednostki, które odmawiają wzięcia udziału w badaniu, są skłonne do udzielenia odpowiedzi na jedno czy dwa zasadnicze pytania. Wiedza na temat informacji zawartych w tych pytaniach może odgrywać kluczową rolę w ocenie stopnia skłonności do wzięcia udziału w badaniu. Przykładem tego typu zmiennej, która może być dostępna na poziomie całej próby s , a w następnym dla wszystkich jednostek $k \in r$, może być płeć czy miejsce zamieszkania osoby. W podejściu tym przyjmuje się ponadto, że nie jest znany wektor wartości globalnych $\boldsymbol{\tau}_{\mathbf{X}^\circ} = \sum_{k \in U} \mathbf{x}_k^\circ$. Z racji tego, że znane są wartości zmiennych pomocniczych \mathbf{x}_k° na poziomie całej próby s , istnieje jednak możliwość oszacowania nieznaney wartości globalnej $\boldsymbol{\tau}_{\mathbf{X}^\circ} = \sum_{k \in U} \mathbf{x}_k^\circ$. W tym celu można skorzystać z estymatora Horwitza-Thompsona wartości globalnej, który wyraża się wzorem $\boldsymbol{\tau}_{\hat{\mathbf{X}}^\circ} = \sum_{k \in s} d_k \mathbf{x}_k^\circ$. W tym wypadku wektor wartości zmiennych pomocniczych \mathbf{x}_k jest zatem postaci: $\mathbf{x}_k = \mathbf{x}_k^\circ$.

Przypadek 3. W podejściu tym zakładamy, że:

1) znane są wartości wektora zmiennych pomocniczych \mathbf{x}_k^* dla wszystkich jednostek z populacji U , próby s oraz zbioru respondentów r , a także znany jest wektor wartości globalnych postaci $\boldsymbol{\tau}_{\mathbf{X}^*} = \sum_{k \in U} \mathbf{x}_k^*$,

2) znane są wartości wektora zmiennych pomocniczych \mathbf{x}_k° dla wszystkich jednostek z próby s oraz w zbiorze respondentów r , a także znane są oszacowania $\boldsymbol{\tau}_{\hat{\mathbf{X}}^\circ} = \sum_{k \in s} d_k \mathbf{x}_k^\circ$, podczas gdy same wartości globalne $\boldsymbol{\tau}_{\mathbf{X}^\circ} = \sum_{k \in U} \mathbf{x}_k^\circ$ zmiennych pomocniczych \mathbf{x}_k° nie są znane.

Powyższe podejście łączy zatem informacje dostępne zarówno na poziomie całej populacji U , jak i próby s . Oznacza to, że istnieją zmienne pomocnicze, dla których znane są jednostkowe wartości na poziomie całej populacji U , oraz zmienne pomocnicze, dla których wartości jednostkowe znane są jedynie na

poziomie próby s . W takim wypadku wektor zmiennych pomocniczych \mathbf{x}_k wymiarów $J^* + J^\circ$ w podejściu kalibracyjnym można zdefiniować jako:

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_k^* \\ \mathbf{x}_k^\circ \end{pmatrix}. \quad (5.5)$$

Opisane powyżej przypadki przedstawiono syntetycznie w tabeli 5.2. Można zauważyć, że przypadek 1 jest szczególną wersją przypadku 3, w sytuacji gdy nie dysponujemy wektorem zmiennych pomocniczych \mathbf{x}_k° . Wówczas wektor zmiennych pomocniczych \mathbf{x}_k redukuje się do wektora \mathbf{x}_k^* , a wektor wartości globalnych $\tau_{\mathbf{X}}^{\text{glob}}$ sprowadza się do wektora $\tau_{\mathbf{X}} = \sum_{k \in U} \mathbf{x}_k$ zdefiniowanego w podrozdziale 2.3.

Tabela 5.2. Wektor zmiennych pomocniczych oraz odpowiadający mu wektor wartości globalnych w podejściu kalibracyjnym z brakami danych

Przypadek	Wektor zmiennych pomocniczych \mathbf{x}_k	Wektor wartości globalnych $\tau_{\mathbf{X}}^{\text{glob}}$
1	\mathbf{x}_k^*	$\tau_{\mathbf{X}^*} = \sum_{k \in U} \mathbf{x}_k^*$
2	\mathbf{x}_k°	$\tau_{\hat{\mathbf{X}}^\circ} = \sum_{k \in s} d_k \mathbf{x}_k^\circ$
3	$\begin{pmatrix} \mathbf{x}_k^* \\ \mathbf{x}_k^\circ \end{pmatrix}$	$\begin{pmatrix} \tau_{\mathbf{X}^*} \\ \tau_{\hat{\mathbf{X}}^\circ} \end{pmatrix}$

Źródło: na podstawie Särndal i Lundström (2005).

Również tutaj przypadek 2 można rozpatrywać jako szczególną wersję przypadku 3, tj. w sytuacji gdy nie dysponuje się informacją na temat wektora zmiennych pomocniczych \mathbf{x}_k^* . Wówczas wektor zmiennych pomocniczych \mathbf{x}_k redukuje się do wektora \mathbf{x}_k° , a wektor wartości globalnych $\tau_{\mathbf{X}}^{\text{glob}}$ sprowadza się do wektora $\tau_{\hat{\mathbf{X}}^\circ}$. W praktycznych zastosowaniach taka sytuacja jest jednak zdecydowanie rzadziej spotykana, gdyż zazwyczaj w procesie kalibracji wag znane są wartości globalne pewnych zmiennych pomocniczych na poziomie całej populacji z zewnętrznych źródeł danych. Dostępne są bowiem na ogół dane pochodzące ze spisów czy rejestrów.

W dalszym ciągu rozpatrywać będziemy najbardziej ogólną sytuację, tj. gdy dysponujemy wektorem zmiennych pomocniczych \mathbf{x}_k^* na poziomie populacji U oraz wektorem zmiennych pomocniczych \mathbf{x}_k° na poziomie próby s (przypadek 3). Kluczową rolę w procesie znajdowania wag kalibracyjnych w_k , dla $k = 1, \dots, m$, odgrywa funkcja $G(\cdot)$, na podstawie której dokonuje się minimalizacji odpowied-

niej funkcji odległości $D(\cdot)$ – warunek (W1’), przy dodatkowych ograniczeniach w postaci równań kalibracyjnych – warunek (W2’) i warunków ograniczających na iloraz wag w_k i d_k – warunek (W3’). Własności funkcji $G(\cdot)$ omówione zostały w podrozdziale 2.2. Z kolei najważniejsze postacie funkcji $G(\cdot)$, odpowiadające im funkcje kalibracyjne $F(\cdot)$ i ograniczenia na wagi zaprezentowano w tabeli 2.1. Przy przyjętych oznaczeniach estymator kalibracyjny wartości globalnej (1.34), przy założeniu, że w badaniu występują jednostkowe braki danych, można zdefiniować w następujący sposób.

Definicja 5.1. *Estymatorem kalibracyjnym wartości globalnej (1.34), przy znanym wektorze zmiennych pomocniczych (5.5) i odpowiadającym mu wektorze wartości globalnych $\tau_{\mathbf{X}}^{\text{glob}}$, jest:*

$$\hat{\tau}_{\text{CAL}}^r = \sum_{k \in r} w_{k_r} y_k, \quad (5.6)$$

gdzie wagi w_{k_r} są rozwiązaniem zadania optymalizacyjnego (W1’)–(W3’), a wektor $\tau_{\mathbf{X}}^{\text{glob}}$ wyraża się wzorem⁴²:

$$\tau_{\mathbf{X}}^{\text{glob}} = \begin{pmatrix} \tau_{\mathbf{X}^*} \\ \tau_{\hat{\mathbf{X}}^{\circ}} \end{pmatrix}. \quad (5.7)$$

W szczególnym przypadku dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ oraz warunków (W1’)–(W2’) można znaleźć wektor wag kalibracyjnych wprost ze wzoru. Postać wag kalibracyjnych podaje poniższe twierdzenie⁴³.

Twierdzenie 5.1 (Särndal i Lundström, 2005). *Rozwiązaniem zadania minimalizacji (5.2) przy warunku (5.3) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag kalibracyjnych $\mathbf{w} = (w_{1_r}, \dots, w_{m_r})^T$, którego składowe, dla $k = 1, \dots, m$, wyrażają się wzorem:*

$$w_{k_r}^{\text{lin}} = d_k + d_k q_k \left(\tau_{\mathbf{X}}^{\text{glob}} - \sum_{k \in r} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in r} q_k d_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \mathbf{x}_k, \quad (5.8)$$

gdzie wektor $\tau_{\mathbf{X}}^{\text{glob}}$ wyraża się wzorem (5.7), a wektor \mathbf{x}_k jest postaci (5.5).

⁴² Dla podkreślenia, że proces kalibracji wag uwzględnia występowanie braków danych, tj. że bierze się pod uwagę korektę wag dla jednostek należących do zbioru respondentów r , oznaczono je przez w_{k_r} .

⁴³ Wyznaczone wagi kalibracyjne mają postać podobną do wag kalibracyjnych, o których mowa w twierdzeniu 2.1.

Kluczową sprawą w wypadku estymatora kalibracyjnego (5.6) wartości globalnej (1.34) jest wyznaczenie jego wariancji. Postać estymatora wariancji estymatora kalibracyjnego podają Särndal i Lundström (2005). Zgodnie z ich ideą, estymator wariancji estymatora kalibracyjnego wartości globalnej (5.6) może być przedstawiony jako suma dwóch komponentów, w których uwzględniane są dwa źródła losowości: będące konsekwencją planu losowania oraz rozkładu generowanego przez braki danych. Mówi o tym poniższe twierdzenie, którego uzasadnienie można znaleźć we wspomnianej monografii⁴⁴.

Twierdzenie 5.2 (Särndal i Lundström, 2005, s. 136). *Estymator wariancji estymatora kalibracyjnego (5.6) wartości globalnej (1.34) jest postaci:*

$$\hat{D}^2(\hat{\tau}_{\text{CAL}}^r) = \hat{D}_{\text{SAM}}^2 + \hat{D}_{\text{NR}}^2, \quad (5.9)$$

gdzie:

$$\begin{aligned} \hat{D}_{\text{SAM}}^2 = & \sum_{k \in r} \sum_{l \in r} (d_k d_l - d_{kl}) (g_k \hat{e}_{k_r}^*) (g_l \hat{e}_{l_r}^*) + \\ & - \sum_{k \in r} d_k (d_k - 1) g_k (g_k - 1) (\hat{e}_{k_r}^*)^2, \end{aligned} \quad (5.10)$$

$$\hat{D}_{\text{NR}}^2 = \sum_{k \in r} g_k (g_k - 1) (d_k \hat{e}_{k_r}^*)^2, \quad (5.11)$$

przy czym

$$\hat{e}_{k_r}^* = y_k - (\mathbf{x}_k^*)^T \mathbf{B}_r^*, \quad (5.12)$$

$$\hat{e}_{k_r} = y_k - \mathbf{x}_k^T \mathbf{B}_r = y_k - (\mathbf{x}_k^*)^T \mathbf{B}_r^* - (\mathbf{x}_k^\circ)^T \mathbf{B}_r^\circ \quad (5.13)$$

oraz

$$\mathbf{B}_r = \begin{pmatrix} \mathbf{B}_r^* \\ \mathbf{B}_r^\circ \end{pmatrix} = \left(\sum_{k \in r} d_k g_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \left(\sum_{k \in r} d_k g_k \mathbf{x}_k y_k \right). \quad (5.14)$$

Powyższe twierdzenie może być w oczywisty sposób zastosowane w odniesieniu do przypadków 1 i 2. W przypadku 1 wektor zmiennych pomocniczych jest postaci $\mathbf{x}_k = \mathbf{x}_k^*$, a reszty (5.12) oraz (5.13) można wyrazić wzorem:

⁴⁴ Podana jest bardziej ogólna postać estymatora wariancji estymatora kalibracyjnego wartości globalnej. Särndal i Lundström (2005) ograniczają się do podania postaci tego estymatora jedynie dla funkcji $G(x) = \frac{1}{2}(x-1)^2$.

$$\hat{e}_{k_r}^* = \hat{e}_{k_r} = y_k - (\mathbf{x}_k^*)^\top \left(\sum_{k \in r} d_k g_k \mathbf{x}_k^* (\mathbf{x}_k^*)^\top \right)^{-1} \left(\sum_{k \in r} d_k g_k \mathbf{x}_k^* y_k \right). \quad (5.15)$$

Z kolei w przypadku 2 wektor zmiennych pomocniczych jest postaci $\mathbf{x}_k = \mathbf{x}_k^\circ$, a reszty (5.12) oraz (5.13) można wyrazić wzorami:

$$\hat{e}_{k_r}^* = y_k, \quad \hat{e}_{k_r} = y_k - (\mathbf{x}_k^\circ)^\top \left(\sum_{k \in r} d_k g_k \mathbf{x}_k^\circ (\mathbf{x}_k^\circ)^\top \right)^{-1} \left(\sum_{k \in r} d_k g_k \mathbf{x}_k^\circ y_k \right). \quad (5.16)$$

Przedstawione powyżej rozważania dotyczące estymatora kalibracyjnego (5.6) wartości globalnej (1.34) można w dość łatwy sposób przenieść na przypadek estymacji wartości globalnej w domenach. Zakładamy zatem, że populacja U składa się z D rozłącznych domen postaci U_1, \dots, U_D . Przyjmujemy ponadto, że w każdej z D wyróżnionych domen przedmiotem zainteresowania jest wartość globalna $\tau_{Y_1}, \dots, \tau_{Y_D}$, gdzie τ_{Y_d} , dla $d = 1, \dots, D$, wyraża się wzorem (1.58). Wartość globalną τ_{Y_d} w domenie d można również wyrazić wzorem (1.61). Estymator kalibracyjny wartości globalnej (1.61) w domenie d w wypadku występowania braków danych może być zatem zapisany w postaci:

$$\hat{\tau}_{\text{CAL},d}^{(1r)} = \sum_{k \in r} w_{k_r} y_{dk}, \quad (5.17)$$

gdzie wagi kalibracyjne w_{k_r} są rozwiązaniem zadania optymalizacyjnego (W1')–(W3')⁴⁵. Przy szacowaniu wartości globalnej τ_{Y_d} w domenie d uwzględnia się zatem wyłącznie wagi kalibracyjne odpowiadające jednostkom należącym do tej domeny. Nie wymaga się w tym podejściu, aby znane były wartości globalne zmiennych pomocniczych na poziomie domeny d .

Estymator wariancji estymatora kalibracyjnego (5.6), w wypadku gdy dokonujemy oszacowania wartości globalnej w domenie d , może być przedstawiony podobnie jak w twierdzeniu 5.2, wzorami (5.9)–(5.14). Wystarczy jedynie dokonać zamiany y_k na y_{dk} . Wówczas wektor $\mathbf{B}_{r,d}$ można wyrazić następującym wzorem:

$$\mathbf{B}_{r,d} = \begin{pmatrix} \mathbf{B}_{r,d}^* \\ \mathbf{B}_{r,d}^\circ \end{pmatrix} = \begin{pmatrix} \sum_{k \in r} d_k g_k \mathbf{x}_k \mathbf{x}_k^\top \\ \sum_{k \in r} d_k g_k \mathbf{x}_k y_{dk} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{k \in r} d_k g_k \mathbf{x}_k y_{dk} \end{pmatrix}. \quad (5.18)$$

⁴⁵ W szczególnym przypadku dla liniowej funkcji kalibracyjnej $F(\cdot)$ wyrażają się one wzorem (5.8).

W wypadku gdy znane są wartości globalne $\sum_{k \in U_d} \mathbf{x}_k^*$ oraz oszacowane wartości globalne $\sum_{k \in r_d} d_k \mathbf{x}_k^\circ$, estymator kalibracyjny wartości globalnej τ_{Y_d} w domenie d można zdefiniować następująco:

$$\hat{\tau}_{\text{CAL},d}^{(2r)} = \sum_{k \in r_d} w_{dk_r} y_k, \quad (5.19)$$

gdzie r_d oznacza zbiór respondentów w domenie d , a wagi kalibracyjne w_{dk_r} są rozwiązaniem poniższego zadania optymalizacyjnego:

- (W1')_d – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in r_d} \frac{d_k}{q_k} G\left(\frac{w_{dk_r}}{d_k}\right) \rightarrow \min, \quad (5.20)$$

- (W2')_d – równania kalibracyjne:

$$\sum_{k \in r_d} w_{dk_r} \mathbf{x}_k = \tau_{\mathbf{X}_d}^{\text{glob}}. \quad (5.21)$$

- (W3')_d – warunki ograniczające:

$$L \leq \frac{w_{dk_r}}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, m_d, \quad (5.22)$$

m_d oznacza liczebność zbioru r_d , a $\tau_{\mathbf{X}_d}^{\text{glob}}$ wyraża się wzorem:

$$\tau_{\mathbf{X}_d}^{\text{glob}} = \left(\frac{\sum_{k \in U_d} \mathbf{x}_k^*}{\sum_{k \in s_d} d_k \mathbf{x}_k^\circ} \right). \quad (5.23)$$

W powyżej sformułowanym zadaniu poszukiwania wag kalibracyjnych wektor \mathbf{x}_k wyraża się wzorem (5.5). W wypadku funkcji $G(x) = \frac{1}{2}(x-1)^2$ wektor wag kalibracyjnych w_{dk_r} także można wyznaczyć wprost ze wzoru. Mówi o tym poniższe twierdzenie.

Twierdzenie 5.3. *Rozwiązaniem zadania minimalizacji (5.20) przy warunku (5.21) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor wag kalibracyjnych \mathbf{w} , którego składowe w domenie d , $d = 1, \dots, D$, wyrażają się wzorem:*

$$w_{dk_r}^{\text{lin}} = d_k + d_k q_k \left(\tau_{\mathbf{X}_d}^{\text{glob}} - \sum_{k \in r_d} d_k \mathbf{x}_k \right)^T \left(\sum_{k \in r_d} q_k d_k \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \mathbf{x}_k. \quad (5.24)$$

Oszacowanie wariancji estymatora kalibracyjnego wartości globalnej (5.19) w domenie d można znaleźć, korzystając z twierdzenia 5.2. W tym celu należy w zasadzie dokonać sumowania w odpowiednich wzorach tego twierdzenia po zbiorze respondentów r_d należących do domeny d .

Opisany powyżej sposób wyznaczania wag kalibracyjnych w_{kr} ⁴⁶ w badaniach z brakami danych polega na odpowiednim wykorzystaniu informacji o zmiennych pomocniczych, które mogą być dostępne zarówno na poziomie populacji U , jak i próby s . Taki sposób wykorzystania informacji o zmiennych pomocniczych jest szeroko omawiany w literaturze. Na przykład, mechanizm różnego wykorzystania zmiennych pomocniczych, w kontekście losowania dwufazowego, dostępnych na dwóch poziomach, tj. populacji i próby, omówili szczegółowo Estevao i Särndal (2002). Dyskusji poddany został również przez autorów wpływ odpowiedniego wykorzystania zmiennych pomocniczych z dwóch poziomów dostępności na wariancję estymatorów kalibracyjnych. Należy jednak podkreślić, że Estevao i Särndal (2002) dokonali analizy własności estymatorów kalibracyjnych wartości globalnej w badaniu symulacyjnym dla sztucznie wygenerowanych danych w sytuacji, gdy w badaniu nie ma zjawiska braków danych.

Z kolei Szymkowiak (2009a) w badaniu symulacyjnym z wykorzystaniem rzeczywistych danych pochodzących z Narodowego Spisu Powszechnego Ludności i Mieszkań (NSP 2002) dokonał oceny wariancji i obciążenia estymatora kalibracyjnego wartości globalnej w sytuacji występowania braków odpowiedzi. Rozważał przy tym dostępność informacji o zmiennych pomocniczych zarówno na poziomie populacji U , jak i próby s .

Warto nadmienić, że proces wyznaczania wag kalibracyjnych opisany w tym podrozdziale odbywa się w jednym kroku – bez wyznaczania wag pośrednich. Z tego powodu proces kalibracyjny określany jest mianem jednokrokowego (ang. *single-step procedure*). Równanie kalibracyjne (5.3) przyjmuje w takiej sytuacji postać:

$$\begin{pmatrix} \sum_{k \in r} w_k \mathbf{x}_k^* \\ \sum_{k \in r} w_k \mathbf{x}_k^o \end{pmatrix} = \begin{pmatrix} \boldsymbol{\tau} \mathbf{X}^* \\ \boldsymbol{\tau} \hat{\mathbf{X}}^o \end{pmatrix}. \quad (5.25)$$

Istnieje jednak możliwość konstrukcji wag kalibracyjnych w dwóch krokach poprzez utworzenie w pierwszym kroku tzw. wag pośrednich, które w kroku drugim podlegają dalszemu procesowi kalibracji. Sposób ten opisano w kolejnym podrozdziale książki.

⁴⁶ Dotyczy to również wyznaczania wag w domenach.

5.4. Kalibracja wag – podejście dwukrokowe

W dwukrokowej procedurze wyznaczania wag proces kalibracji przeprowadza się z poziomu zbioru respondentów r na poziom próby s , wykorzystując wartości wektora \mathbf{x}_k° w celu uzyskania wag pośrednich. Wagi te następnie wykorzystywane są jako wagi wejściowe w drugim kroku kalibracji z poziomu zbioru respondentów r na poziom populacji U z użyciem wektorów \mathbf{x}_k^* i \mathbf{x}_k° (Andersson i Särndal, 2016).

Zgodnie z propozycją, jaką podają Särndal i Lundström (2005), należy tutaj rozróżnić dwa przypadki określane mianem dwukrokowego podejścia kalibracyjnego typu A i B. W pierwszym kroku przedstawimy algorytm wyznaczania wag kalibracyjnych w podejściu dwukrokowym typu A. Zakładając będziemy dowolną postać funkcji $G(\cdot)$ ⁴⁷. Następnie przedstawiony zostanie algorytm wyznaczania wag kalibracyjnych w dwukrokowym podejściu typu B, również przy założeniu dowolnej postaci funkcji $G(\cdot)$ i odpowiadającej jej funkcji kalibracyjnej $F(\cdot)$.

Algorytm 4. Wagi kalibracyjne w podejściu dwukrokowym typu A

Krok 1: Wyznaczenie wag pośrednich w_k°

Przyjmując wagi $d_k = \frac{1}{\pi_k}$ wynikające z planu losowania próby oraz wektor zmiennych pomocniczych \mathbf{x}_k° , wyznacz pośrednie wagi w_k° , kalibrując je z poziomu zbioru respondentów r na poziom próby s , tak aby spełnione było następujące równanie kalibracyjne:

$$\sum_{k \in r} w_k^\circ \mathbf{x}_k^\circ = \sum_{k \in s} d_k \mathbf{x}_k^\circ. \quad (5.26)$$

Wagi w_k° oblicz ze wzoru:

$$w_k^\circ = d_k g_k^\circ, \quad (5.27)$$

gdzie

$$g_k^\circ = F\left(q_k(\mathbf{x}_k^\circ)^T \boldsymbol{\lambda}^\circ\right), \quad (5.28)$$

przy czym mnożnik Lagrange'a wyznacz zgodnie z algorytmem 1 lub 2. W kroku tym, ze względu na występujące braki danych, zbyt małe wagi d_k są zwiększane do odpowiedniego poziomu.

⁴⁷ Särndal i Lundström (2005) podają postać wag jedynie dla funkcji $G(x) = \frac{1}{2}(x-1)^2$.

Krok 2: Wyznaczenie końcowych wag kalibracyjnych w_{2Ak_r} .

Wykorzystując wagi pośrednie w_k° jako wagi wejściowe, wyznacz końcowe wagi kalibracyjne w_k , tak aby spełnione było następujące równanie kalibracyjne:

$$\sum_{k \in r} w_k \mathbf{x}_k = \boldsymbol{\tau}_{\mathbf{X}}^{\text{glob}}, \quad (5.29)$$

gdzie $\boldsymbol{\tau}_{\mathbf{X}}^{\text{glob}}$ wyraża się wzorem (5.7), a \mathbf{x}_k wzorem (5.5). Wagi kalibracyjne, które spełniają powyższe równanie kalibracyjne, wyrażają się wzorem:

$$w_{2Ak_r} = w_k^\circ g_k, \quad (5.30)$$

gdzie

$$g_k = F\left(q_k(\mathbf{x}_k)^T \boldsymbol{\lambda}\right), \quad (5.31)$$

przy czym mnożnik Lagrange'a wyznacz zgodnie z algorytmem 1 lub 2.

Krok 3: Estymator kalibracyjny wartości globalnej (1.34)

Wykorzystując wyznaczone wagi kalibracyjne w_{2Ak_r} , dokonaj oszacowania wartości globalnej (1.34) ze wzoru:

$$\hat{\tau}_{\text{CAL}}^{2A,r} = \sum_{k \in r} w_{2Ak_r} y_k. \quad (5.32)$$

Wyznaczone powyżej wagi kalibracyjne w_{2Ak_r} odtwarzają znane wartości globalne zmiennych pomocniczych na poziomie populacji U , tj. spełniają równanie kalibracyjne postaci $\sum_{k \in r} w_{2Ak_r} \mathbf{x}_k^* = \sum_{k \in U} \mathbf{x}_k^*$. Z racji tego, że wektor \mathbf{x}_k zawiera również jako składową wektor \mathbf{x}_k° , wagi te jednocześnie spełniają równanie kalibracyjne $\sum_{k \in r} w_{2Ak_r} \mathbf{x}_k^\circ = \sum_{k \in s} d_k \mathbf{x}_k^\circ$, podobnie jak w podejściu jednokrokowym. Oznacza to, że w podejściu dwukrokowym typu A odtwarzane są zarówno wartości globalne zmiennych pomocniczych \mathbf{x}_k^* na poziomie populacji U , jak i oszacowane wartości globalne zmiennych pomocniczych \mathbf{x}_k° na poziomie próby s .

Algorytm 5. Wagi kalibracyjne w podejściu dwukrokowym typu B

Krok 1: Wyznaczenie wag pośrednich w_k°

Wyznacz wagi pośrednie w_k° w sposób analogiczny do kroku 1 algorytmu 4.

Krok 2: Wyznaczenie końcowych wag kalibracyjnych w_{2Bk_r}

Wykorzystując wagi pośrednie w_k° jako wagi wejściowe, wyznacz końcowe wagi kalibracyjne w_k tak, aby spełnione było następujące równanie kalibracyjne:

$$\sum_{k \in r} w_k \mathbf{x}_k^* = \boldsymbol{\tau}_{\mathbf{x}^*}. \quad (5.33)$$

Wagi kalibracyjne, które spełniają powyższe równanie kalibracyjne, wyrażają się wzorem:

$$w_{2Bk_r} = w_k^\circ g_k^*, \quad (5.34)$$

gdzie

$$g_k^* = F\left(q_k(\mathbf{x}_k^*)^T \boldsymbol{\lambda}\right), \quad (5.35)$$

przy czym mnożnik Lagrange'a wyznaczone zgodnie z algorytmem 1 lub 2.

Krok 3: Estymator kalibracyjny wartości globalnej (1.34)

Wykorzystując wyznaczone wagi kalibracyjne w_{2Bk_r} , dokonaj oszacowania wartości globalnej (1.34) ze wzoru:

$$\hat{\tau}_{\text{CAL}}^{2B,r} = \sum_{k \in r} w_{2Bk_r} y_k. \quad (5.36)$$

Warto zauważyć, że zarówno w podejściu dwukrokowym typu A, jak i w B wagi kalibracyjne wyznaczone są z wykorzystaniem zmiennych pomocniczych dostępnych na obydwu poziomach, tj. próby s i populacji U . Istnieje jednak zasadnicza różnica między wagami w_{2Ak_r} a w_{2Bk_r} . Mianowicie w dwukrokowym podejściu typu B wagi kalibracyjne w_{2Bk_r} na ogół nie spełniają równania kalibracyjnego $\sum_{k \in r} w_{2Bk_r} \mathbf{x}_k^\circ = \sum_{k \in s} d_k \mathbf{x}_k^\circ$, co może mieć wpływ na obciążenie i wariancję estymatora (5.36). Warto również podkreślić, że wagi w_{k_r} uzyskane w podejściu jednokrokowym oraz wagi w_{2Ak_r} i w_{2Bk_r} uzyskane w podejściu dwukrokowym typu A i B odpowiednio, na ogół nie prowadzą do identycznych zestawów końcowych wag kalibracyjnych. Estymatory kalibracyjne (5.6), (5.32) oraz (5.36) wykorzystujące różne zestawy wag kalibracyjnych nie uwzględniają ponadto w swojej konstrukcji podejścia modelowego (ang. *model-free calibration*), o którym mowa była w rozdziale czwartym.

Kluczową kwestią jest również wyznaczenie estymatorów wariancji dla estymatorów kalibracyjnych wartości globalnej (1.34) w podejściu dwukrokowym typu A i B. W wypadku estymatora kalibracyjnego $\hat{\tau}_{\text{CAL}}^{2A,r}$, który wyraża się wzo-

rem (5.32), jego estymator wariancji można przedstawić równaniem (5.9), przy czym we wzorach (5.10)–(5.14) w miejsce g_k należy podstawić g_k° opisane w formule (5.28). W podobny sposób można wyrazić estymator wariancji dla estymatora kalibracyjnego $\hat{\tau}_{CAL}^{2B,r}$, który wyraża się wzorem (5.36). Mianowicie dla tego estymatora jego estymator wariancji można przedstawić jako (5.9), przy czym we wzorach (5.10)–(5.14) w miejsce g_k należy podstawić g_k° , a reszty wyrazić w następujący sposób:

$$\hat{e}_{k_r}^* = \hat{e}_{k_r} = y_k - (\mathbf{x}_k^*)^T \left(\sum_{k \in r} d_k g_k^\circ \mathbf{x}_k^* (\mathbf{x}_k^*)^T \right)^{-1} \sum_{k \in r} d_k g_k^\circ \mathbf{x}_k^* y_k. \quad (5.37)$$

Rozważane powyżej podejście dwukrokowe typu A i B można w dość łatwy sposób przenieść na problem szacowania wartości globalnej τ_{Y_d} w domenie d , dla $d = 1, \dots, D$. W tym celu zakłada się, że znane są wartości globalne $\sum_{k \in U_d} \mathbf{x}_k^*$. W algorytmie 4, w celu znalezienia końcowych wag kalibracyjnych w domenie d , należy dokonać zmiany zbiorów r i s na r_d i s_d odpowiednio, a w kroku 2 uwzględnić wektor wartości globalnych $\tau_{\mathbf{X}_d}^{\text{glob}}$, który wyraża się wzorem (5.23). Z kolei w algorytmie 5 należy dokonać zamiany zbioru r na r_d , a w miejsce wektora $\tau_{\mathbf{X}^*}$ uwzględnić wektor $\sum_{k \in U_d} \mathbf{x}_k^*$. Na potrzeby oszacowania wariancji tak skonstruowanych estymatorów kalibracyjnych wartości globalnej τ_{Y_d} w domenie d należy również skorzystać z twierdzenia 5.2 i dokonać stosownych zmian, zgodnie z opisem poczynionym powyżej.

Spośród omówionych trzech podejść w konstrukcji wag kalibracyjnych, tj. jednokrokowym, dwukrokowym typu A i dwukrokowym typu B, to ostatnie podejście w praktyce badań statystycznych jest powszechnie wykorzystywane. W podejściu tym kalibracja wag d_k wynikających z przyjętego planu losowania próby na pierwszym etapie polega na ich korekcie tak, aby uwzględniony został wpływ braków danych. W wypadku ich występowania wagi d_k są bowiem zbyt niskie i należy je odpowiednio zwiększyć. W tym celu najczęściej dokonuje się podziału próby s na pewne jednorodnie podgrupy i wyznacza w ich obrębie odwrotność wskaźnika frakcji odpowiedzi (ang. *inverse of the response fraction*), przez który, w celu skorygowania, należy pomnożyć wagi d_k . W drugim kroku tak skorygowane wagi kalibruje się, aby odtworzyć znane wartości globalne zmiennych pomocniczych na poziomie populacji U . W niektórych przypadkach proces korygowania wag obejmuje dodatkowy etap, podczas którego skalibrowane wagi podlegają dalszym modyfikacjom. Jest tak w sytuacji, gdy wagi kalibracyjne są bardzo zróżnicowane, przez co odpowiednie estymatory kalibracyjne wartości globalnej mogą wykazywać dużą wariancję. W obszarze tym wypracowano wiele podejść, w tym przy-

cinanie wag czy ich wygładzanie, w celu zwiększenia efektywności oszacowań (Chen i in., 2017; Haziza i Beaumont, 2017; Kott, 2011; Wu i Lu, 2016). Problem ten poruszono również w podrozdziale 2.3 niniejszej książki. Celem tak kompleksowo przeprowadzonego procesu kalibracji wag jest zmniejszenie obciążenia na skutek istnienia braków danych, redukcja wariancji (wzrost precyzji oszacowań) oraz uzyskanie spójnego systemu wag.

Dwukrokowe podejście typu B do kalibracji wag w badaniach statystycznych z brakami odpowiedzi jest stosowane przez wiele krajowych urzędów statystycznych na całym świecie. Jak wskazują Särndal i Lundström (2005), wykorzystywane jest ono przez Statistics Canada w Badaniu Aktywności Ekonomicznej Ludności (ang. *Labour Force Survey*) czy przez United States Census Bureau w Badaniu Dochodów i Uczestnictwa w Programie (ang. *Survey of Income and Program Participation*).

5.5. Podejście kalibracyjne w Badaniu Aktywności Ekonomicznej Ludności

Głównym celem badania było oszacowanie liczby osób pracujących, bezrobotnych oraz biernych zawodowo z wykorzystaniem estymatorów kalibracyjnych wartości globalnej z uwzględnieniem podejścia jednokrokowego oraz dwukrokowego typu A i B na poziomie podregionów w Polsce w 2015 roku w ujęciu kwartalnym. Na potrzeby estymacji, dla tak zdefiniowanych charakterystyk rynku pracy, wykorzystano dane pochodzące z Badania Aktywności Ekonomicznej Ludności. Należy podkreślić, że Główny Urząd Statystyczny publikuje wyniki oszacowań z tego badania na poziomie całego kraju z uwzględnieniem dodatkowych zmiennych (jak miejsce zamieszkania, wiek, stan cywilny, wykształcenie itp.), na poziomie makroregionów (grupujących województwa), regionów (grupujących województwa lub część województwa) oraz województw (z uwzględnieniem wykształcenia). Nie są natomiast publikowane wyniki z liczbą osób pracujących, biernych zawodowo czy bezrobotnych na poziomie podregionów (GUS, 2015). Wynika to ze sposobu konstrukcji wag, w których nie są brane pod uwagę dane ludnościowe na poziomie podregionów.

Badanie Aktywności Ekonomicznej Ludności jest jednym z najważniejszych badań prowadzonych przez Główny Urząd Statystyczny umożliwiającym bieżącą ocenę stopnia wykorzystania zasobów pracy, a zarazem pozwalającym na szerszą charakterystykę poszczególnych grup ludności wyróżnionych ze względu na ich status na rynku pracy (pracujący, bezrobotni) lub pozostających poza rynkiem

pracy (bierni zawodowo) (GUS, 2018b). W Badaniu Aktywności Ekonomicznej Ludności analizą objęte są gospodarstwa domowe oraz osoby w wieku 15 lat i więcej, a zatem populację generalną, do której odnosi się uogólnione wyniki, stanowi ludność Polski w wieku 15+. Badanie to ma charakter dobrowolny i jest przeprowadzane w cyklu kwartalnym metodą reprezentacyjną.

Dobór jednostek do próby w Badaniu Aktywności Ekonomicznej Ludności przebiega zgodnie z omówionym w podrozdziale 1.2 losowaniem dwustopniowym. Jednostkami losowania pierwszego stopnia (JPS) są w miastach rejony statystyczne, a na wsi obwody spisowe. Losowanie JPS-ów wewnątrz warstw przeprowadza się ze zróżnicowanymi prawdopodobieństwami wyboru, proporcjonalnymi do liczby mieszkań w JPS-ie. Podstawą podziału na warstwy jest podział Polski na województwa⁴⁸. Z kolei jednostkami losowania drugiego stopnia w ramach JPS-ów są mieszkania, a do ich selekcji stosowane jest losowanie proste bez zwracania.

W 2015 roku algorytm konstruowania wag w Badaniu Aktywności Ekonomicznej Ludności przebiegał według procedury obejmującej cztery zasadnicze etapy (GUS, 2015).

Algorytm 6. Konstrukcja wag w Badaniu Aktywności Ekonomicznej Ludności

Krok 1: *Wyznaczenie wag pierwotnych*

W pierwszym kroku wyznaczane są wagi pierwotne będące odwrotnościami prawdopodobieństw wyboru do próby poszczególnych mieszkań w warstwach:

$$d_k = \frac{M_k}{m_k}, \quad (5.38)$$

gdzie M_k jest liczbą wszystkich mieszkań w warstwie k -tej, a m_k jest liczbą mieszkań losowanych z warstwy k -tej, dla $k = 1, \dots, K$ ⁴⁹. Wszystkie osoby w wieku 15+ należące do wylosowanych mieszkań z danej warstwy otrzymują zatem taką samą wagę pierwotną.

Krok 2: *Obliczenie tzw. współczynników realizacji R*

Ze względu na to, że nie wszystkie mieszkania wylosowane do zakładanej próby i kwalifikujące się do badania zostają de facto zbadane, wyznacza się współczynnik realizacji R zgodnie ze wzorem:

⁴⁸ Obecnie jest to podział na 17 regionów NUTS2.

⁴⁹ Jak wcześniej nadmieniono, podstawą podziału na warstwy jest podział na województwa. W ramach województw tworzone są warstwy wewnątrzwojewódzkie, przy czym szczegółowy sposób ich tworzenia został opisany w publikacji GUS (2015).

$$R = \frac{K_M - N_M}{K_M}, \quad (5.39)$$

gdzie K_M to oszacowanie (obliczone za pomocą wag pierwotnych) liczby mieszkań kwalifikujących się do badania, z kolei N_M jest oszacowaniem liczby mieszkań kwalifikujących się do badania, lecz niedających się zbadać z różnych względów. Współczynniki realizacji wyznacza się w obrębie sześciu grup, które wyróżnia się ze względu na klasę miejscowości⁵⁰. Współczynniki te mierzą nasilenie zjawiska braków odpowiedzi ze względu na niemożliwość zbadania wszystkich kwalifikujących się do badania mieszkań.

Krok 3: Wyznaczenie wag wtórnych

W drodze dzielenia wag pierwotnych d_k przez współczynnik realizacji R , przy czym współczynnik R bierze się w zależności od klasy miejscowości, w której znajduje się dane mieszkanie, wyznacza się wagi wtórne:

$$w_{kh} = \frac{d_k}{R_h}, \quad (5.40)$$

gdzie $k = 1, \dots, K$ oraz $h = 1, \dots, 6$. Wagi wtórne są zarazem wagami finalnymi dla wyników dotyczących gospodarstw domowych.

Krok 4: Wyznaczenie wag finalnych dla respondentów

Wyznaczenie wag finalnych dla danych ludnościowych przeprowadza się tak, aby dostosować wyniki Badania Aktywności Ekonomicznej Ludności do bieżących szacunków demograficznych. Dokonuje się tego przez obliczenie tzw. modyfikatorów, które oblicza się osobno dla każdej z 48 kategorii zdefiniowanych przez: miejsce zamieszkania (miasto/wieś) \times płeć (mężczyzna/kobieta) \times 12 grup wieku (15–17, 18–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65+), w drodze dzielenia liczby osób w danej kategorii według skorygowanych szacunków demograficznych przez liczbę osób w tejże kategorii obliczoną z danych z Badania Aktywności Ekonomicznej Ludności przy zastosowaniu wag wtórnych uzyskanych w kroku 3. Wagi finalne otrzymuje się w wyniku mnożenia wag wtórnych przez odpowiednie modyfikatory.

Bardzo ważną kwestią w Badaniu Aktywności Ekonomicznej Ludności jest konstrukcja estymatorów wariancji w procesie oceny precyzji szacowanych cha-

⁵⁰ 1 – Warszawa, 2 – miasta od 500 000 do 1 miliona mieszkańców, 3 – miasta od 100 000 do 500 000 mieszkańców, 4 – miasta od 20 000 do 100 000 mieszkańców, 5 – inne miasta, 6 – wsie. Obecnie współczynniki realizacji wyznacza się w zależności od województwa oraz klasy miejscowości.

rakterystyk rynku pracy w określonych przekrojach. Począwszy od 2003 roku Główny Urząd Statystyczny wykorzystuje w tym celu metody przybliżone, oparte na podpróbkach replikacyjnych i zasadzie „bootstrap”. Wynika to z tego, że metody oparte na podpróbkach dają możliwość jednolitego traktowania bardzo różnorodnych parametrów, a także eliminują konieczność wyprowadzania skomplikowanych wzorów analitycznych. Zastosowanie zasady bootstrap do najczęściej występujących w praktyce metody reprezentacyjnej złożonych planów losowania wymaga odpowiednich modyfikacji. Szczegółowy opis tego typu zagadnień został przedstawiony w opracowaniu Shao i Tu (1995).

W wypadku złożonego, dwustopniowego schematu losowania próby, jaki występuje w Badaniu Aktywności Ekonomicznej Ludności, należy podkreślić, że szacowanie wariancji odbywa się na podstawie danych z prób kwartalnych, wylosowanych z uwzględnieniem podziału na warstwy na poziomie jednostek losowania pierwszego stopnia. Zastosowany wariant metody bootstrap realizowany jest oddzielnie w każdej warstwie. Oznacza to, że niezależnie w każdej warstwie ($k = 1, \dots, K$), zawierającej n_k wylosowanych JPS-ów, losowana jest spośród nich podpróbka o licznosci $n_k - 1$ metodą losowania prostego ze zwracaniem. Po wylosowaniu jednostek pierwszego stopnia do podpróbki bootstrapowej „przenosi się” wylosowane z nich do Badania Aktywności Ekonomicznej Ludności mieszkania wraz z zamieszkującymi je respondentami. Losowanie takie jest powtarzane B razy⁵¹, przy czym za każdym razem wyznacza się dla uzyskanej b -tej podpróbki ($b = 1, \dots, B$) zmodyfikowane i skalibrowane wagi w celu zapewnienia zgodności z danymi demograficznymi dla odpowiedniego kwartału. Następnie dla każdej z podpróbek i wyznaczonych wag dokonuje się oszacowania interesującego parametru $\hat{\tau}^{(b)}$, a po wykonaniu B iteracji – oszacowania wariancji estymatora $\hat{\tau}$ badanego parametru (w rozważanej warstwie) zgodnie ze wzorem:

$$D^2(\hat{\tau}) = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\tau}^{(b)} - \bar{\hat{\tau}} \right)^2, \quad (5.41)$$

gdzie

$$\bar{\hat{\tau}} = \frac{1}{B} \sum_{b=1}^B \hat{\tau}^{(b)}. \quad (5.42)$$

Następnie, w celu uzyskania oszacowania wariancji dla rozważanej podpopulacji, należy dokonać sumowania oszacowanych wariancji otrzymanych niezależnie we wszystkich jej warstwach. W ostatnim kroku można wyznaczyć względny średni błąd szacunku (REE) estymatora $\hat{\tau}$ interesującego parametru zgodnie ze wzorem

⁵¹ Przyjmuje się, że $B = 500$.

(1.32). Szczegółowy opis tej procedury można znaleźć w opracowaniu Popińskiego (2006) czy w zeszytach metodologicznych publikowanych przez Główny Urząd Statystyczny (GUS, 2018b).

Jak pokazuje rysunek 5.1, w 2015 roku wskaźnik realizacji w Badaniu Aktywności Ekonomicznej Ludności wynosił około 65%. Wynika to z tego, że w próbie znalazły się mieszkania, które zostały całkowicie niezbadane, a które powinny zostać zbadane. Dotyczyć to może na przykład mieszkań niezbadanych z takiego powodu jak: brak dostępu do mieszkania (niemożność dotarcia do mieszkania, ochrona itp.), brak kontaktu z mieszkańcami (nie zastano mieszkańców, choroba, podeszły wiek, bariera językowa), krótkotrwała nieobecność mieszkańców (do 12 miesięcy) czy odmowa wzięcia udziału w badaniu. Warto przy tym zwrócić uwagę, że oprócz jednostkowych braków danych, których istnienie wymusza odpowiednią korektę wag pierwotnych zgodnie z opisanym powyżej algorytmem kalibracji, w Badaniu Aktywności Ekonomicznej Ludności istnieją również pozycyjne braki odpowiedzi. Oznacza to, że respondenci, którzy zdecydowali się wziąć udział w badaniu, ze względu na drażliwy charakter pytania lub trudności z udzieleniem odpowiedzi mogących wystąpić w wypadku pytań dotyczących przeszłości, mają prawo odmowy. Szczególnie dotyczy to pytania o wielkość wynagrodzenia z pracy głównej, które z powodu drażliwego charakteru jest obciążone dużą frakcją braków odpowiedzi. Należy jednak podkreślić, że w odniesieniu do kilku pytań, w których występują pozycyjne braki odpowiedzi, nie są stosowane żadne techniki imputacji.

W wyżej opisanym sposobie konstruowania wag ujętym w algorytmie 6, stosowanym przez Główny Urząd Statystyczny w Badaniu Aktywności Ekonomicznej Ludności, nie uwzględnia się jako zmiennych pomocniczych informacji o liczbie ludności w przekroju podregionów. Ze względu na to, że głównym celem tego podrozdziału jest oszacowanie wybranych charakterystyk rynku pracy na poziomie podregionów, podjęto próbę kalibracji wag tak, aby odtwarzane były również wartości globalne odnoszące się do liczby ludności na tym poziomie agregacji przestrzennej. Oznacza to zatem, że poszerzono liczbę grup zdefiniowanych w kroku czwartym algorytmu 6 o 73 podregiony. Zabieg ten był podyktowany tym, że w kwartałach 2015 roku obserwowano duże zróżnicowanie wartości globalnych liczby ludności oszacowanych na podstawie wag finalnych w Badaniu Aktywności Ekonomicznej Ludności na poziomie podregionów. Zmienność ta odzwierciedlała się także dużym zróżnicowaniem szacunków wartości globalnych rozpatrywanych charakterystyk rynku pracy na poziomie podregionów z uwzględnieniem dodatkowych przekrojów (na przykład liczby pracujących czy biernych zawodowo). W związku z powyższym przyjęto założenie, że proces kalibracji

wag przeprowadzony zostanie tak, aby odtwarzać znane kwartalne struktury ludnościowe w 2015 roku dla podregionów i dodatkowych przekrojów.

Jako wartości globalne dla liczby ludności w podregionach wykorzystano dane ludnościowe opracowane na podstawie wyników Narodowego Spisu Powszechnego Ludności i Mieszkań z 2011 roku z uwzględnieniem zmian spowodowanych ruchem naturalnym (urodzeniami, zgonami), migracjami ludności (na pobyt stały i czasowy) oraz przemieszczeniami związanymi ze zmianami administracyjnymi.

W procesie budowania finalnych wag zastosowano cztery wybrane funkcje odległości $D(\cdot)$, które wykorzystują w swej konstrukcji funkcje kalibracyjne $F(\cdot)$, tj. linear, raking, logit oraz \sinh ⁵². Ponadto zastosowano trzy metody ich wyznaczania (jednokrokową, dwukrokową typu A oraz dwukrokową typu B), które zostały opisane w podrozdziale 5.3 i 5.4. Wykorzystano ponadto dwa odmienne zestawy zmiennych pomocniczych odpowiadających wektorowi \mathbf{x}_k^* :

- **zestaw 1:** (48 wariantów): płeć (2 warianty: mężczyzna, kobieta) \times miejsce zamieszkania (2 warianty: miasto, wieś) \times grupy wieku (12 wariantów: 15–17, 18–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65+),
- **zestaw 2:** (121 wariantów): zestaw 1 (48) + podregiony (73).

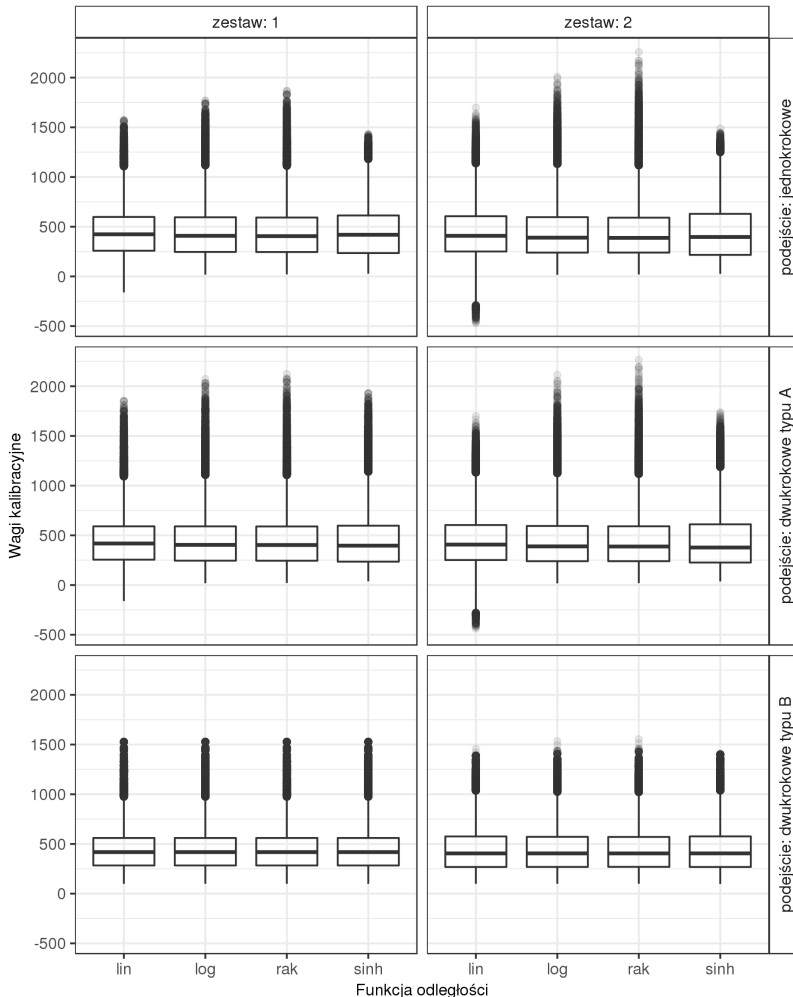
Powyżej określony zestaw 1 jest analogiczny do tego, który był stosowany w Badaniu Aktywności Ekonomicznej Ludności w 2015 roku (por. krok czwarty w algorytmie 6). Nie uwzględniono w nim zatem informacji o wartościach globalnych liczby ludności w przekroju podregionów. W przeciwieństwie jednak do zestawu 1, w zestawie 2 wykorzystano jako zmienną pomocniczą informację o przynależności danej osoby do odpowiedniego podregionu. W procesie kalibracji wag wykorzystano wartości globalne z danymi ludnościowymi na poziomie agregacji przestrzennej odpowiadającej podregionom. Zabieg taki miał z jednej strony służyć sprawdzeniu, w jakim stopniu uwzględnienie informacji o liczbie ludności w przekroju podregionów w postaci wartości globalnych wpłynie na postać wag kalibracyjnych, tj. na ich zróżnicowanie, występowanie wag ujemnych lub ekstremalnych. Z drugiej strony umożliwił estymację wybranych charakterystyk rynku pracy w przekroju podregionów z akceptowalną precyzją.

Z kolei za wektor zmiennych pomocniczych \mathbf{x}_k^o we wszystkich analizowanych podejściach przyjęto zmienne pomocnicze określające przynależność badanego mieszkania do jednej z sześciu wyróżnionych typów klasy miejscowości. Łącznie uzyskano zatem $2 \times 4 \times 3 = 24$ zestawy wag kalibracyjnych⁵³.

⁵² Funkcje kalibracyjne zostały przedstawione w tabeli 2.1.

⁵³ Dwa zestawy zmiennych pomocniczych \mathbf{x}_k^* , cztery funkcje kalibracyjne, trzy metody wyznaczania wag kalibracyjnych oraz jeden zestaw zmiennych pomocniczych \mathbf{x}_k^o .

Dla liniowej funkcji kalibracyjnej $F(\cdot)$, pierwszego zestawu zmiennych pomocniczych \mathbf{x}_k^* , zmiennych pomocniczych wchodzących w skład wektora \mathbf{x}_k° oraz podejścia dwukrokowego typu B uzyskuje się dokładnie takie same końcowe wagi kalibracyjne jak w Badaniu Aktywności Ekonomicznej Ludności w poszczególnych kwartałach 2015 roku. Jest to zgodne z wcześniejszym komentarzem, że w praktyce badań statystycznych właśnie to podejście jest powszechnie wykorzystywane przez krajowe urzędy statystyczne, przede wszystkim w BAEL.



Rysunek 5.2. Porównanie rozkładów wag kalibracyjnych w rozpatrywanych podejściach

Źródło: na podstawie BAEL.

W pierwszej kolejności w zastosowanym podejściu badawczym kompleksowej analizie poddano wagi kalibracyjne uzyskane z wykorzystaniem czterech wspomnianych już funkcji odległości, dwóch zestawów zmiennych pomocniczych x_k^* , wektora zmiennych pomocniczych x_k^o oraz trzech rozważanych metod konstruowania wag (jednokrokowej, dwukrokowej typu A i B). Ich rozkład ilustruje rysunek 5.2, przy czym wzięto pod uwagę wagi kalibracyjne ze wszystkich czterech kwartałów. Dodatkowo w tabeli 5.3 przedstawiono najważniejsze statystyki opisowe wag kalibracyjnych dla wszystkich 24 rozpatrywanych wariantów i czterech kwartałów 2015 roku.

Tabela 5.3. Charakterystyki wag kalibracyjnych

Podejście	F(u)	Min	Q ₁	Q ₂	Q ₃	Max	Średnia	Std
Zestaw 1								
Jednokrokowe	lin	-160,7	258,5	423,7	598,2	1573,0	443,7	253,2
Jednokrokowe	log	17,7	246,3	408,3	594,8	1767,8	443,7	257,6
Jednokrokowe	rak	20,6	245,5	405,3	592,3	1867,1	443,7	259,0
Jednokrokowe	sinh	27,4	234,8	419,0	613,2	1431,0	443,7	256,5
Dwukrokowe typu A	lin	-160,7	255,4	418,0	590,4	1851,9	443,7	257,6
Dwukrokowe typu A	log	18,5	245,0	404,0	590,2	2072,2	443,7	261,2
Dwukrokowe typu A	rak	20,6	244,5	402,7	589,4	2124,4	443,7	262,0
Dwukrokowe typu A	sinh	37,8	235,1	396,1	597,0	1930,8	443,7	264,9
Dwukrokowe typu B	lin	98,6	283,6	418,0	560,2	1526,8	443,7	215,6
Dwukrokowe typu B	log	98,6	283,6	418,0	560,2	1526,8	443,7	215,6
Dwukrokowe typu B	rak	98,6	283,6	418,0	560,2	1526,8	443,7	215,6
Dwukrokowe typu B	sinh	98,6	283,6	418,0	560,2	1526,8	443,7	215,6
Zestaw 2								
Jednokrokowe	lin	-468,9	250,8	408,3	605,6	1698,1	443,7	267,2
Jednokrokowe	log	16,3	239,5	389,9	596,3	2007,9	443,7	271,8
Jednokrokowe	rak	19,6	239,7	387,7	591,1	2256,4	443,7	273,8
Jednokrokowe	sinh	25,8	216,7	396,6	629,4	1488,5	443,7	273,1
Dwukrokowe typu A	lin	-437,9	251,6	406,9	603,6	1699,1	443,7	267,5
Dwukrokowe typu A	log	17,2	239,6	388,8	594,1	2113,8	443,7	272,6
Dwukrokowe typu A	rak	19,5	239,9	387,6	591,0	2264,0	443,7	273,9
Dwukrokowe typu A	sinh	36,1	226,2	377,4	610,8	1737,0	443,7	277,0
Dwukrokowe typu B	lin	99,0	268,4	405,0	575,6	1457,5	443,7	222,9
Dwukrokowe typu B	log	99,0	268,9	405,0	571,5	1533,8	443,7	223,1
Dwukrokowe typu B	rak	99,2	269,6	405,0	570,5	1553,7	443,7	223,2
Dwukrokowe typu B	sinh	96,8	268,4	405,1	576,3	1413,6	443,7	223,0

Źródło: na podstawie BAEL.

Jak nadmieniono powyżej, wagi otrzymane na podstawie podejścia dwukrokowego typu B z funkcją liniową i z pierwszym zestawem zmiennych pomocniczych x_k^* są równe wagom finalnym w Badaniu Aktywności Ekonomicznej Ludności wy-

korzystywanym do produkcji oficjalnych szacunków. To podejście będzie służyć jako punkt odniesienia do pozostałych rozważanych wariantów i określimy jako bazowe. Szczególną uwagę poświęcimy podejściom z drugim zestawem zmiennych pomocniczych. Wynika to z tego, że w tym wypadku odtworzone będą dodatkowo znane wartości populacyjne liczby ludności w przekroju podregionów.

Analiza wag kalibracyjnych w_k przedstawionych na rysunku 5.2 oraz w tabeli 5.3 pozwala zauważyć, że uwzględnienie zmiennej pomocniczej x_k^o w postaci informacji o podregionie w procesie kalibracji (zestaw 2) zwiększa na ogół ich zmienność. Dotyczy to w zasadzie każdej z czterech przyjętych funkcji odległości oraz trzech metod wyznaczania wag (jednokrokowej, dwukrokowej typu A i w mniejszym stopniu dwukrokowej typu B). Jest to szczególnie zauważalne w wypadku wykładniczej funkcji odległości (raking). W dwukrokowej procedurze typu B uwzględnienie dodatkowego równania kalibracyjnego odnoszącego się do informacji o przynależności osoby do podregionu (zestaw 2) powoduje, że zmniejsza się rozstęp wag kalibracyjnych w porównaniu z tym, jaki otrzymuje się dla zestawu 1 dla liniowej funkcji odległości i w wersji sinusa hiperbolicznego. Jest to widoczne przede wszystkim dla drugiej z wymienionych funkcji.

Warto również zwrócić uwagę, że z wyjątkiem dwukrokowej procedury typu B wyznaczania wag kalibracyjnych, liniowa funkcja kalibracyjna generuje dla niektórych respondentów ujemne wagi kalibracyjne, bez względu na przyjęty zestaw zmiennych pomocniczych x_k^* . Jest to sprzeczne z definicją wagi i tego typu funkcje kalibracyjne nie powinny być wykorzystywane w procesie estymacji. W podejściu bazowym, stosowanym przez Główny Urząd Statystyczny, funkcja liniowa pozwala uzyskać wagi dodatnie dla wszystkich respondentów, dlatego jej praktyczne zastosowanie jest w pełni uzasadnione.

W wypadku drugiego zestawu zmiennych pomocniczych najmniej zróżnicowane są wagi kalibracyjne w_k dla funkcji odległości w wersji sinusa hiperbolicznego. Dotyczy to każdej z trzech metod ich konstrukcji. W wypadku podejścia jednokrokowego i drugiego zestawu zmiennych pomocniczych najbardziej zróżnicowane wagi kalibracyjne daje liniowa funkcja kalibracyjna oraz raking. Podobna sytuacja jest zauważalna w wypadku dwukrokowego podejścia typu A.

Dokonując podsumowania wyników uzyskanych dla finalnych wag kalibracyjnych we wszystkich rozpatrywanych wariantach, można stwierdzić, że uwzględnienie jako wartości globalnych informacji o liczbie ludności w przekroju podregionów zwiększa nieco zmienność wag kalibracyjnych, przede wszystkim w wypadku procedury jednokrokowej i dwukrokowej typu A, w porównaniu z zestawem 1, w którym nie brano pod uwagę informacji o podregionach. Dla funkcji kalibracyjnej w postaci sinusa hiperbolicznego najbardziej widoczne jest ponadto

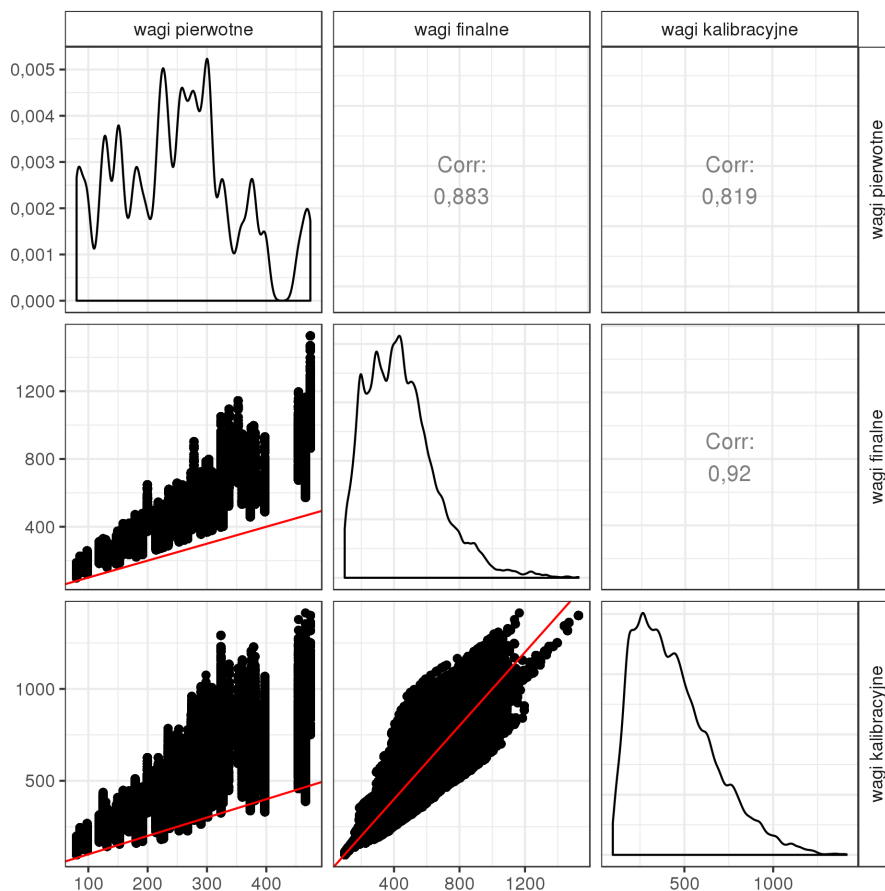
zmniejszenie rozstępu wag. Co więcej, bez względu na przyjętą funkcję kalibracyjną, uzyskuje się również w tym podejściu wagi dodatnie, a jednocześnie zachowana jest zgodność struktur ludnościowych w Badaniu Aktywności Ekonomicznej Ludności na poziomie podregionów z wartościami spisowymi skorygowanymi o ruch naturalny, migracje i przemieszczenia związane ze zmianami administracyjnymi.

Warto przy tym podkreślić, że w podejściu dwukrokowym typu B dla drugiego zestawu zmiennych pomocniczych wagi kalibracyjne uzyskane z wykorzystaniem funkcji odległości w postaci sinusa hiperbolicznego odznaczały się najlepszymi własnościami. Dotyczyło to również pozostałych podejść (jednokrokowego i dwukrokowego typu A). W związku z tym w procesie estymacji wybranych charakterystyk rynku pracy w przekroju podregionów, w pierwszej kolejności warto wziąć pod uwagę wagi kalibracyjne przypisane respondentom, uzyskane w podejściu dwukrokowym typu B dla funkcji kalibracyjnej w wersji sinusa hiperbolicznego.

Jak można zauważyć na rysunku 5.3, wagi te są silnie i dodatnio skorelowane zarówno z wagami pierwotnymi d_k , jak i z wagami finalnymi wykorzystywanymi przez Główny Urząd Statystyczny w Badaniu Aktywności Ekonomicznej Ludności do uogólniania wyników. Oszacowany współczynnik korelacji liniowej Pearsona pomiędzy wagami pierwotnymi a wagami kalibracyjnymi jest na poziomie 0,82. Z kolei oszacowany współczynnik korelacji liniowej Pearsona między wagami bazowymi (finalnymi) a wyznaczonymi z uwzględnieniem informacji o podregionach wagami kalibracyjnymi wynosi 0,92.

Na rysunku 5.4 przedstawiono oszacowaną, z wykorzystaniem estymatora kalibracyjnego wartości globalnej, liczbę osób pracujących, bezrobotnych i biernych zawodowo w poszczególnych kwartałach 2015 roku. Ograniczono się przy tym, w celu egzemplifikacji, do podregionów województwa wielkopolskiego. Wzięto ponadto pod uwagę wszystkie rozpatrywane podejścia uwzględniające trzy procedury kalibracyjne, dwa zestawy zmiennych pomocniczych oraz cztery różne funkcje odległości.

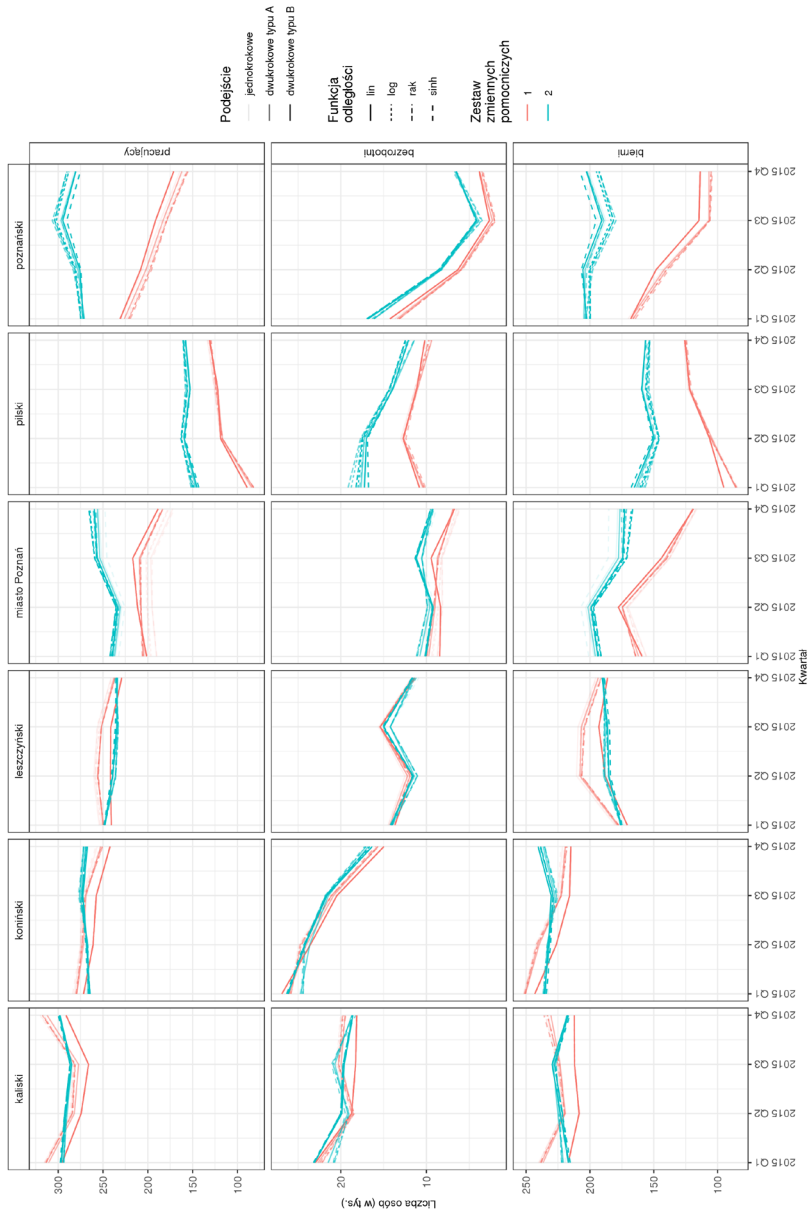
W analizie uzyskanych wyników należy zwrócić uwagę na dwa bardzo ważne aspekty. Po pierwsze przy ustalonym wspólnym wektorze zmiennych pomocniczych x_k^o , w zależności od wziętego zestawu zmiennych pomocniczych x_k^* , uzyskuje się często dość różne oszacowania liczby osób pracujących, bezrobotnych czy biernych zawodowo. Dotyczy to w dużej mierze oszacowanej liczby osób pracujących w podregionach pilskim, poznańskim oraz mieście Poznań, oszacowanej liczby osób bezrobotnych w podregionach pilskim i poznańskim oraz oszacowanej liczby osób biernych zawodowo, w zasadzie we wszystkich sześciu podregionach



Rysunek 5.3. Porównanie wag pierwotnych d_k , bazowych (finalnych w Badaniu Aktywności Ekonomicznej Ludności) oraz wag kalibracyjnych w_k otrzymanych dla drugiego zestawu zmiennych pomocniczych i dwukrokowego podejścia kalibracyjnego typu B wraz z funkcją kalibracyjną sinus hiperboliczny

Źródło: na podstawie BAEL.

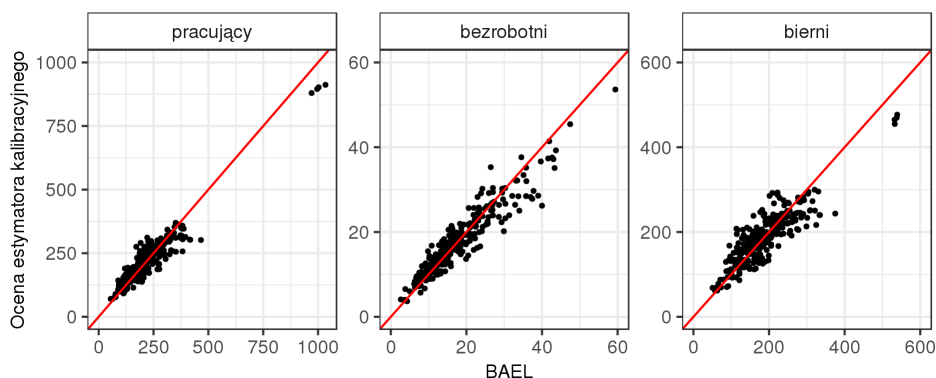
województwa wielkopolskiego. Zauważone duże różnice w oszacowaniach wynikają z tego, że wagi kalibracyjne uzyskane dla zestawu 1 nie brały pod uwagę informacji o przynależności danej osoby do odpowiedniego podregionu. Oznacza to, że tak uzyskane wagi nie odtwarzały znanych struktur ludnościowych odnoszących się do liczby osób w wieku 15+ w poszczególnych podregionach. W przeciwieństwie do tego, drugi zestaw zmiennych pomocniczych x_k^* uwzględniał informację o podregionie. W związku z tym wyznaczone wagi kalibracyjne, bez względu na zastosowaną procedurę (jednokrokowa, dwukrokowa typu A i B) oraz przyjętą



Rysunek 5.4. Oszacowana liczba osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów województwa wielkopolskiego

Źródło: na podstawie BAEŁ.

funkcję odległości, odtwarzały znaną liczbę osób w wieku 15+ w przekroju podregionów. Rysunek ten pokazuje zatem, że w procesie estymacji wartości globalnej dla domen zastosowanie wag, których nie uwzględniono na etapie kalibracji, może skutkować poważnymi błędami. Oszacowana liczba osób pracujących, bezrobotnych i biernych zawodowo we wspomnianych powyżej podregionach była bowiem dla pierwszego zestawu zmiennych pomocniczych znacznie poniżej oszacowań uzyskanych dla zestawu drugiego. Różnice te są bardziej widoczne na rysunku 5.5.



Rysunek 5.5. Porównanie oszacowań liczby osób pracujących, bezrobotnych i biernych zawodowo w 2015 roku (w tys.)

Źródło: na podstawie BAEL.

Na osi odciętych na rysunku 5.5 przedstawiono oszacowania liczby osób pracujących, bezrobotnych i biernych zawodowo z wykorzystaniem podejścia bazowego (dwukrokowego typu B dla liniowej funkcji kalibracyjnej i zestawu 1) dla wszystkich podregionów i kwartałów 2015 roku. Na osi rzędnych przedstawiono podobne oszacowania, ale dla procedury dwukrokowej typu B, funkcji odległości w wersji sinusa hiperbolicznego i drugiego zestawu zmiennych pomocniczych \mathbf{x}_k^* . Oprócz dostrzegalnej silnej i dodatniej zależności korelacyjnej między obydwoma typami oszacowań, można zauważyć dość istotne różnice między nimi w każdej z trzech rozważanych kategorii osób na rynku pracy. Jest to oczywiście konsekwencją odmiennych zestawów zmiennych pomocniczych \mathbf{x}_k^* , które mają wpływ na oszacowania uzyskiwane z wykorzystaniem estymatora kalibracyjnego wartości globalnej dla wybranych charakterystyk rynku pracy w przekroju podregionów.

Ponadto, przy ustalonym wektorze zmiennych pomocniczych \mathbf{x}_k^* , oszacowania liczby osób pracujących, bezrobotnych czy biernych zawodowo mogą się różnić w zależności od przyjętego podejścia (jednokrokowe, dwukrokowe typu A i B)

oraz w pewnym stopniu od przyjętej funkcji odległości⁵⁴. Jest to widoczne na przykład w trzecim kwartale 2015 roku w podregionie poznańskim dla drugiego zestawu zmiennych pomocniczych, gdzie występują pewne różnice w oszacowaniach liczby osób pracujących czy biernych zawodowo w zależności od przyjętego podejścia.

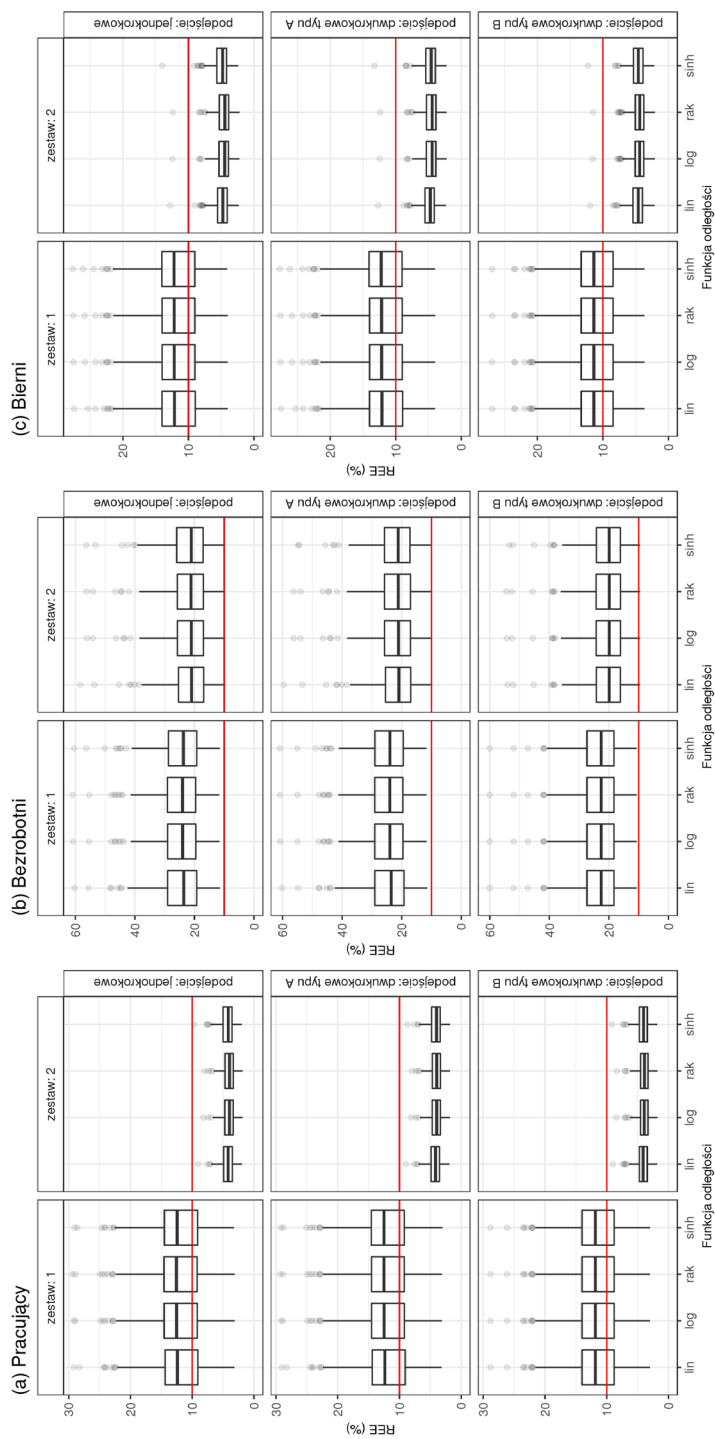
Oprócz szczegółowej analizy wag kalibracyjnych dyskusji zostały poddane również względne średnie błędy szacunku (REE) dla oszacowanej liczby osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów w Polsce we wszystkich kwartałach 2015 roku. Błędy te zostały przedstawione na rysunku 5.6.

Błędy te wyznaczono zgodnie z opisaną wyżej i stosowaną w BAEL metodą bootstrap. Jak można zauważyć, w wypadku osób pracujących oraz biernych zawodowo wzięcie drugiego zestawu zmiennych pomocniczych x_k^* powoduje znaczny wzrost precyzji oszacowań dla każdej z uwzględnionych w analizie funkcji odległości. Co więcej, kalibracja wag z uwzględnieniem informacji o podregionie powoduje, że oszacowania liczby osób pracujących i biernych zawodowo nie przekraczają przyjętego w tego typu analizach przez Główny Urząd Statystyczny 10% prognozy⁵⁵. Dotyczy to w zasadzie każdego podregionu dla czterech kwartałów 2015 roku. Można je więc uznać za akceptowalne.

Nieco odmienną sytuację można zaobserwować w wypadku osób bezrobotnych. Uwzględnienie drugiego zestawu zmiennych pomocniczych x_k^* wpływa na wzrost precyzji oszacowań liczby osób bezrobotnych w porównaniu z oszacowaniami uzyskanymi dla zestawu pierwszego. Świadczy o tym chociażby mediana oszacowanych względnych średnich błędów szacunku, która dla zestawu 1 była na poziomie 25% dla każdej funkcji odległości i przyjętej metody wyznaczania wag kalibracyjnych (jednokrokowej, dwukrokowej typu A i B). W wypadku zestawu 2 mediana oscylowała w okolicach 20%. Należy mieć jednak na uwadze, że w takiej sytuacji w dalszym ciągu w 50% przypadków możliwe jest publikowanie wyników, niemniej jednak należy do nich podchodzić bardziej sceptycznie. W pozostałych przypadkach, jak podkreślono na stronie 24, wyniki należałoby zagregować i przedstawić na wyższym poziomie agregacji przestrzennej (województwo). Warto przy tym podkreślić, że większe oszacowane względne średnie błędy sza-

⁵⁴ Jak pokazano w badaniu symulacyjnym w rozdziale drugim, dobór funkcji odległości tylko w pewnym stopniu wpływa na wielkość oszacowanej wartości globalnej. Należy jednak uwzględnić w procesie kalibracji na ogół różne funkcje, w celu wyboru optymalnej, gdyż dla niektórych funkcji mogą wystąpić wagi ujemne lub ekstremalne.

⁵⁵ Dyskusja na temat prognoz stosowanych przez Główny Urząd Statystyczny została podjęta na stronie 24.



Rysunek 5.6. Względne średnie błędy szacunku liczby osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów w Polsce we wszystkich kwartalach 2015 roku

Źródło: na podstawie BAEL.

cunków liczby osób bezrobotnych wynikają z tego, że w próbie będącej podstawą Badania Aktywności Ekonomicznej Ludności w poszczególnych kwartałach 2015 roku obserwowano w niektórych podregionach niewielką liczbę reprezentantów, tj. osób bezrobotnych. Mediana liczby osób bezrobotnych w próbie (z uwzględnieniem wszystkich czterech kwartałów) w podregionach wynosiła bowiem 31 (wartość minimalna 4, a maksymalna 161). Dla porównania, mediana liczby osób pracujących i biernych zawodowo w przekroju podregionów dla wszystkich kwartałów wynosiła 385⁵⁶ i 412⁵⁷ odpowiednio.

W wypadku estymacji liczby osób bezrobotnych w przekroju podregionów warto w pierwszej kolejności rozważyć w procesie kalibracji wag jako zmienne pomocnicze informacje z rejestru bezrobotnych w celu zwiększenia precyzji oszacowań. Można wziąć pod uwagę też inne metody estymacji, które oferuje statystyka małych obszarów (Rao i Molina, 2015). Pewne próby związane z estymacją liczby osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów w latach 2010–2015 podjęto już w projekcie „Rozszerzenie Badania Aktywności Ekonomicznej Ludności”, w którym rolę koordynatora merytorycznego pełnił autor niniejszej książki (GUS, 2018c). W tym celu wykorzystano m.in. wielomianowy mieszany model logistyczny i wielomianowy model mieszany z efektami losowymi skorelowanymi w czasie i przestrzeni (López-Vizcaíno, Lombardia i Morales, 2013; 2015). Przykładem tego typu metod, które z powodzeniem wykorzystano w Badaniu Aktywności Ekonomicznej Ludności dla bardziej szczegółowo zdefiniowanych domen, mogą być również strukturalne modele szeregów czasowych (Wilak, 2014).

5.6. Wnioski

Rozważania zaprezentowane w niniejszym rozdziale nawiązują do dynamicznie rozwijającej się na świecie metody kalibracji, poświęconej estymacji w badaniach statystycznych z brakami odpowiedzi. Przedstawiono konstrukcję estymatorów kalibracyjnych wartości globalnej uzyskanych w podejściu jednokrokowym oraz dwukrokowym typu A i B w sytuacji, gdy w badaniu występują jednostkowe braki odpowiedzi. Dyskusji poddano również sposób konstrukcji wariancji rozważanych estymatorów wartości globalnej otrzymanych w trzech rozpatrywanych podejściach. Rozważania prowadzono w kontekście estymacji wartości global-

⁵⁶ Wartość minimalna 123, a maksymalna 1606.

⁵⁷ Wartość minimalna 119, a maksymalna 1585.

nej i estymatora wariancji wartości globalnej zarówno w populacji generalnej, jak i w odpowiednio zdefiniowanych domenach.

W podejściu jednokrokowym zaproponowano sposób konstruowania wag estymatora kalibracyjnego wartości globalnej w domenie w sytuacji, gdy znane są wartości globalne zmiennych pomocniczych x_k^* w domenie oraz oszacowane wartości globalne zmiennych pomocniczych x_k° na zbiorze r_d dla dowolnie zdefiniowanej funkcji odległości. W algorytmach 4 i 5, będących uogólnieniem metody opisanej w pracy Särndala i Lundströma (2005) dla dowolnie zdefiniowanej funkcji odległości, zaproponowano ponadto sposób konstruowania wag kalibracyjnych w podejściu dwukrokowym typu A i B.

Z wykorzystaniem rzeczywistych danych pochodzących z Badania Aktywności Ekonomicznej Ludności, w którym występuje zjawisko braków odpowiedzi, dokonano estymacji liczby osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów, tj. na poziomie, dla którego Główny Urząd Statystyczny nie dostarcza oficjalnych szacunków. W tym celu wykorzystano estymatory kalibracyjne wartości globalnej w domenach dla różnie zdefiniowanej funkcji odległości oraz trzy metody wyznaczania wag kalibracyjnych (podejście jednokrokowe oraz dwukrokowe typu A i B). Rozważano przy tym dwa zestawy zmiennych pomocniczych x_k^* , z których drugi brał pod uwagę jako jedną ze zmiennych pomocniczych przynależność respondenta do odpowiedniego podregionu. Pokazano w ten sposób, że nie tylko występuje zysk na precyzji w stosunku do oszacowań, które można by uzyskać z wykorzystaniem zestawu wag kalibracyjnych stosowanych aktualnie w Badaniu Aktywności Ekonomicznej Ludności (podejście bazowe), ale także, że uzyskane oszacowania wybranych charakterystyk rynku pracy na poziomie podregionów są na ogół wiarygodne (akceptowalna precyzja).

Na zakończenie rozważań można się pokusić o opinię, że kalibracja w ujęciu prezentowanym w tym rozdziale, będzie odgrywać w polskiej praktyce badań statystycznych coraz większą rolę. Wynika to z tego, że w zasadzie wszystkie badania obarczone są błędami nielosowymi w postaci braków odpowiedzi.

Kalibracja zintegrowana w spisach mieszanych

6.1. Wprowadzenie

Przez pojęcie spisu ludności (spisu) należy rozumieć okresowe badanie, którego głównym celem jest pozyskanie pełnej informacji o stanie i strukturze ludności według podstawowych cech demograficznych, społecznych i ekonomicznych z uwzględnieniem najbardziej szczegółowych przekrojów terytorialnych (Gołata, 2018). Spisy powszechne, obok badań reprezentacyjnych, należą do najważniejszych źródeł informacji o warunkach życia ludności i są prowadzone w zasadzie przez wszystkie krajowe urzędy statystyczne na całym świecie (Krywult-Albańska, 2013; Paradysz, 2009). Są one jednocześnie najstarszymi badaniami statystycznymi, których historia sięga starożytnych Chin, Japonii, Grecji, Persji oraz Egiptu (Bethlehem, 2009). Jak podaje Gołata (2018), pierwszych wzmianek na temat spisów można się doszukiwać już w Babilonii, w której około 4 tys. lat p.n.e. przeprowadzano badanie populacji w celu określenia zapotrzebowania na żywność, czy w starożytnym Egipcie (2,5 tys. lat p.n.e.), gdzie spisywano ludność na potrzeby ustalenia zasobów siły roboczej w procesie budowy piramid.

Wzmianki na temat spisu powszechnego, który miał pomóc w ustaleniu stanu majątkowego mieszkańców rzymskiej prowincji Syrii oraz Judei, można również znaleźć w czasach narodzin Jezusa Chrystusa w Ewangelii według świętego Łukasza⁵⁸.

Historia nowożytnych spisów jest jednak znacznie krótsza i przypada na XVII i XVIII wiek. Dotyczy to m.in. spisów, które przeprowadzono w Szwecji, Francji, Rosji czy Prusach. Pierwszy spis ludności w Polsce również przypada na ten okres. Odbył się w 1789 roku, a jego inicjatorem był Fryderyk Józef Moszyński. Spis przeprowadzony został na podstawie uchwalonej przez Sejm Czteroletni (1788–1792) konstytucji z 22 czerwca 1789 roku „Lustracya dymów i podanie ludności”.

Współcześnie można wyodrębnić kilka różnych sposobów przeprowadzania spisu powszechnego jako najważniejszego źródła danych statystycznych o społeczeństwie (Beręsewicz i Gołata, 2018; Dygaszewicz, 2009; Gołata, 2018; Valente, 2010):

- **Spis tradycyjny** – który jest zazwyczaj przeprowadzany co 10 lat (rzadziej co 5 lat) i jest jedną z najczęściej spotykanych form spisu, w którym dane pozyskiwane są zarówno w drodze wypełnienia kwestionariusza spisowego przez respondentów, jak i w drodze wywiadu przeprowadzonego przez ankietatorów. Tego typu spisy przeprowadza się między innymi w takich państwach jak Kanada, Rosja, Wielka Brytania, Słowacja, Węgry czy Kolumbia. Również Narodowy Spis Powszechny Ludności i Mieszkań w 2002 roku przeprowadzony był w Polsce z wykorzystaniem metody tradycyjnej.
- **Spis tradycyjny z coroczną aktualizacją danych** – jest to pewna odmiana spisu tradycyjnego, w której dokonuje się corocznej modyfikacji szacunków liczby ludności i podstawowych cech demograficznych w roku spisowym. W odniesieniu do gospodarstw domowych każdego roku przez kolejne dziesięć lat do kolejnego spisu przeprowadzane są natomiast odpowiednie badania reprezentacyjne o charakterze społeczno-ekonomicznym. Tego typu spisy przeprowadzane są m.in. w Stanach Zjednoczonych.
- **Spis kroczący** – jest to rodzaj spisu, który ma charakter badania ciągłego. Informacje o osobach i gospodarstwach domowych są zbierane stopniowo poprzez kumulację wcześniejszych wyników obejmujących obszar całego państwa, w czasie zazwyczaj kilku lat, w przeciwieństwie do jednego dnia tzw. momentu krytycznego lub krótkiego okresu. Tego typu spis może być zaprojektowany jako pewien cykl corocznych badań reprezentacyjnych,

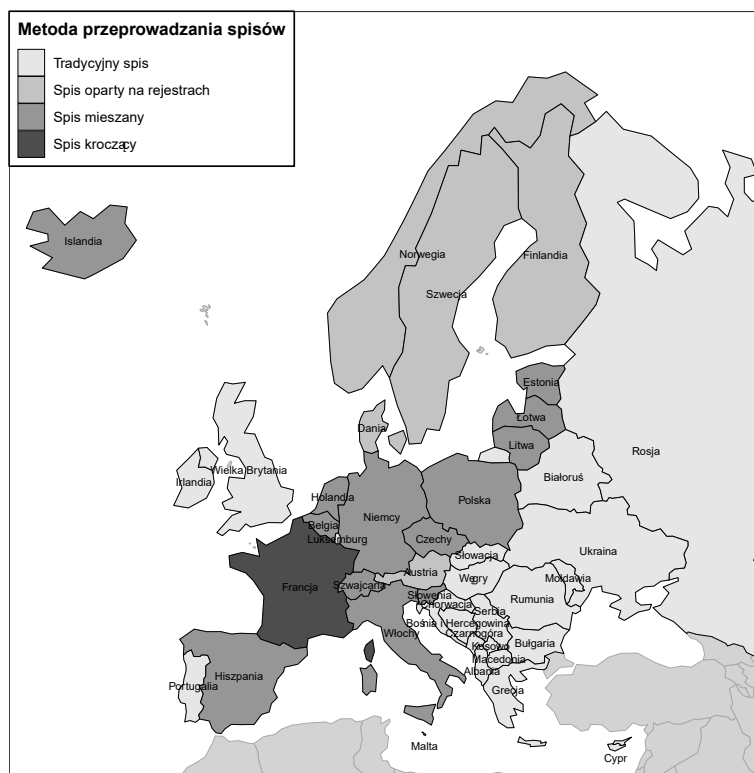
⁵⁸ W owym czasie wyszło rozporządzenie Cezara Augusta, żeby przeprowadzić spis ludności w całym państwie. Pierwszy ten spis odbył się wówczas, gdy wielkorządcą Syrii był Kwiryniusz.

które umożliwiają uogólnianie wyników na całą populację. W tym celu można dokonywać odpowiedniej kalibracji wag, aby dostosować uzyskane szacunki dla kluczowych zmiennych demograficznych do znanych struktur ludnościowych (Brown i Beaujouan, 2013). Spis kroczący przeprowadzony został m.in. we Francji w 2010 roku.

- **Spis oparty na rejestrach** – tego rodzaju spisy powszechne przeprowadzane są przede wszystkim w krajach skandynawskich (Danii, Szwecji czy Norwegii). Począwszy od 1970 roku, są one również realizowane w Finlandii. Jak nazwa spisu wskazuje, ich głównym celem jest jak najpełniejsze wykorzystanie istniejących źródeł informacji w postaci rejestrów administracyjnych, które na potrzeby spisu przekształcane są w rejestry statystyczne. Aby możliwe było przeprowadzenie tego typu spisu, w danym państwie muszą funkcjonować odpowiedniej jakości rejestry administracyjne, wśród których kluczową rolę odgrywa rejestr ludności obejmujący całą populację. Łączenie danych pochodzących z rejestrów odbywa się poprzez wykorzystanie klucza identyfikacyjnego bądź statystycznych metod integracji danych takich jak parowanie statystyczne czy probabilistyczne łączenie rekordów (D’Orazio, Di Zio i Scanu, 2006; Roszka, 2013). W spisach opartych na rejestrach można również wykorzystać w odpowiedni sposób kalibrację, zgodnie z propozycją sformułowaną przez Wallgrenów (2014). Można także skorzystać z tzw. podejścia Renssena (1998), które zostało wykorzystane w pracy Roszki i Szymkowiaka (2014) na potrzeby konstrukcji tabel dla zmiennych nieobserwowanych jednocześnie w dwóch źródłach informacji na przykładzie danych pochodzących z Europejskiego Badania Warunków Życia Ludności i Badania Budżetów Gospodarstw Domowych.
- **Spis wirtualny** – to szczególny rodzaj spisu, w którym wykorzystuje się dane z rejestrów administracyjnych oraz z badań reprezentacyjnych. Jest on stosowany z dużym powodzeniem w Holandii, a jego głównym celem jest budowa zintegrowanego systemu mikrobaz złożonego z danych o spójnych definicjach, które zawierają szczegółowe informacje demograficzne, ekonomiczne oraz społeczne dotyczące populacji osób, gospodarstw domowych czy mieszkań (Gerards, 2012). Również w tego typu spisach, będących pewną odmianą metody mieszanej, wykorzystywane jest podejście kalibracyjne w postaci wspomnianej już metody powtarzanego ważenia (de Waal, 2015; Knottnerus i van Duin, 2006; Lumiste, 2011).
- **Spis oparty na metodach mieszanych** – spis ten polega na wykorzystaniu wielu różnych źródeł danych statystycznych. W ujęciu klasycznym oznacza

to wykorzystanie rejestrów administracyjnych i informacji pochodzących ze spisu tradycyjnego. W praktyce tego typu łączenie danych nie znalazło jednak zastosowania. Zdecydowanie częściej stosuje się łączenie informacji z rejestrów administracyjnych, które uzupełniane są wynikami odpowiedniego badania reprezentacyjnego⁵⁹. Spis metodą mieszaną przeprowadzony był między innymi w 2011 roku w Niemczech oraz w Polsce.

Jak pokazuje powyższy opis, nie ma jednej uniwersalnej metody przeprowadzenia spisu. Różne państwa w latach 2010–2011 wdrożyły do praktyki odmienne modele spisu. Przedstawiono je na rysunku 6.1. Dwoma dominującymi typami spisów był spis tradycyjny oraz realizowany metodą mieszaną. Francja była w zasadzie jedynym państwem europejskim, które zdecydowało się na spis kroczący.



Rysunek 6.1. Metody przeprowadzania spisów w krajach europejskich w latach 2010–2011

Źródło: na podstawie Valente (2010).

⁵⁹ Więcej informacji na temat tego typu spisów można znaleźć w podrozdziale 6.2.

Z kolei spisy oparte w całości na rejestrach są domeną państw, w których istnieje bardzo dobra infrastruktura statystyczna, której podstawą są wysokiej jakości rejestry administracyjne. Dotyczy to w pierwszej kolejności Szwecji, Danii, Norwegii oraz Finlandii. Na tego typu spis zdecydowała się również po raz pierwszy w 2011 roku Austria.

Podejście kalibracyjne, które jest przedmiotem szczegółowych rozważań w niniejszej książce, jest szeroko wykorzystywane w spisach powszechnych. W zasadzie można je wykorzystać w spisach bazujących na dowolnej metodzie gromadzenia danych. W zależności od przyjętej metody spisu, w naturalny sposób zastosowanie kalibracji wymaga nieco innych rozwiązań metodologicznych. Z powodzeniem podejście kalibracyjne można wykorzystać w spisach opartych całkowicie na rejestrach administracyjnych (A. Wallgren i B. Wallgren, 2014), spisach kroczących (Brown i Beaujouan, 2013), spisie wirtualnym (Boonstra, 2004) czy w spisach opartych na metodzie mieszanej (Münnich, Burgard, Gabler, Ganninger i Kolb, 2014; Münnich, Wagner i Sachs, 2012). Również w Polsce, w której ostatni spis powszechny z 2011 roku był przeprowadzony metodą mieszaną, zastosowano podejście kalibracyjne. Zostało ono opisane w kontekście ludności rezydującej w artykule Szymkowiaka (2014).

W dalszej części książki szczegółowo przedstawiono opis podejścia kalibracyjnego w spisach mieszanych na przykładzie Narodowego Spisu Powszechnego Ludności i Mieszkań 2011 w kontekście ludności faktycznie zamieszkałej. Przedstawiono również propozycję zastosowania kalibracji zintegrowanej na danych pochodzących z tego spisu. Mimo że podejście zintegrowane nie zostało uwzględnione w ostatnim spisie, będzie je można z powodzeniem wykorzystać w zbliżającym się Narodowym Spisie Powszechnym Ludności i Mieszkań 2021, który zostanie również przeprowadzony metodą mieszaną.

6.2. Idea spisów mieszanych na przykładzie NSP 2011

Spisy mieszane, obok spisów realizowanych z wykorzystaniem podejścia tradycyjnego, należą do najczęściej stosowanych w praktyce metod gromadzenia danych. W rundzie spisowej 2010–2011, oprócz Polski, na tę formę spisu zdecydowały się również Niemcy, Hiszpania, Włochy, Słowacja, Słowenia, Belgia, Czechy, Islandia, Litwa, Łotwa i Estonia. Z jednej strony wynikało to z tego, że państwa te miały odpowiednią infrastrukturę statystyczną w postaci rejestrów. Z drugiej zaś było to konsekwencją konieczności pozyskania pewnych informacji w drodze badania

reprezentacyjnego, które nie były dostępne w rejestrach. Uzupełnienie danych rejestrowych wynikami badania reprezentacyjnego, na podstawie którego możliwe było uogólnianie wyników na całą populację, jest fundamentem podejścia mieszane- go. W tego typu rozwiązaniu spisowym łączy się zatem dane z rejestrów administracyjnych na poziomie jednostkowym z danymi z badania reprezentacyjnego. W rezultacie otrzymuje się tzw. złoty rekord, w którym dane pochodzą z różnego rodzaju źródeł statystycznych.

Podejście mieszane może być również zastosowane w innym ujęciu. Można dokonać połączenia danych pochodzących z rejestrów administracyjnych z tymi uzyskanymi w drodze spisu pełnego. Tego typu podejście nie znalazło jednak ostatecznie zastosowania w praktyce, mimo że pewne kraje, takie jak Hiszpania, wstępnie rozważały tę metodę gromadzenia danych. W związku z powyższym nie będzie ono dalej rozpatrywane w książce w kontekście spisów mieszanych.

Głównym uzasadnieniem wprowadzenia mieszanej metody spisu przez kilka państw europejskich, w której łączy się dane z rejestrów z informacjami z badania reprezentacyjnego, jest redukcja kosztów oraz zmniejszenie obciążeń respondentów w porównaniu z podejściem tradycyjnym. Szacuje się na przykład, że w Hiszpanii w 2011 roku spis przeprowadzony metodą mieszaną kosztował 85 milionów euro, co stanowiło około 20% kosztów, jakie należałoby ponieść w związku z zastosowaniem tradycyjnej formy jego realizacji (Argüeso i Vega, 2014). Ważnym aspektem jest również odpowiednia jakość spisów, w których w całości bądź w pewnym zakresie wykorzystywane są dane pochodzące ze źródeł administracyjnych. Temu zagadnieniu oraz dyskusji i ocenie jakości informacji pozyskiwanych w spisach ludności poświęcone są prace Gołaty (2012; 2014; 2018).

W państwach, które zdecydowały się na wykorzystanie w spisie mieszanym danych z rejestrów i systemów administracyjnych, kluczowy był rejestr ludności. Oprócz niego wykorzystywano na ogół wiele innych rejestrów – odnoszących się do ubezpieczeń społecznych, danych podatkowych, związanych z edukacją, zdrowiem itd. W kontekście badania reprezentacyjnego, którego głównym celem było pozyskanie informacji niedostępnych w rejestrach administracyjnych, zastosowano różne rozwiązania metodologiczne i liczebność próby. Na przykład w Niemczech badanie reprezentacyjne obejmowało próbę liczącą 10% (Bechtold, 2013; Federal Statistical Office, 2017), w Hiszpanii również 10% (Argüeso i Vega, 2014), a w Polsce objęto nim 20% mieszkań, co oznaczało, że w badaniu reprezentacyjnym udział wzięło około 20% populacji mieszkańców Polski (GUS, 2012).

W Polsce w 2011 roku przeprowadzono po raz pierwszy w historii spis powszechny z uwzględnieniem podejścia mieszane- go. Oznaczało to, że spis przeprowadzono w formie badania pełnego, które obejmowało całą populację, oraz bada-

nia reprezentacyjnego. W wypadku badania pełnego wykorzystano tzw. krótki formularz, który został wypełniony danymi pochodzącymi z rejestrów administracyjnych oraz zweryfikowanymi przez respondentów w samospisie internetowym⁶⁰. W kontekście badania reprezentacyjnego wykorzystano natomiast tzw. formularz długi, który obejmował szeroki zakres tematyczny z dużą liczbą pytań (ponad 120⁶¹), które po uogólnieniu na całą populację pozwoliły na szeroką charakterystykę demograficzno-społeczną i ekonomiczną ludności. W przeprowadzonym badaniu reprezentacyjnym poruszane były kwestie odnoszące się do sześciu grup tematycznych:

- ludności i jej charakterystyki demograficzno-społecznej,
- aktywności ekonomicznej,
- migracji wewnętrznych i zagranicznych ludności,
- narodowości i wyznania,
- gospodarstw domowych i rodzin,
- budynków i mieszkań.

Warto wspomnieć, że spis, który odbył się w 2011 roku, był pierwszym powszechnym badaniem zrealizowanym wyłącznie przy wykorzystaniu formularzy elektronicznych. Mianowicie, oprócz wywiadów przeprowadzanych przez rachmistrzów spisowych z wykorzystaniem urządzeń typu hand-held (metoda CAPI), w spisie zastosowano dwie nowe metody zbierania danych: wywiady telefoniczne wspomagane przez komputer (metoda CATI) oraz samospis internetowy (metoda CAII) (Dygaszewicz, 2012; Paradysz, 2012).

Odrębną kwestią wymagającą wyjaśnienia, niezbędną z punktu widzenia zastosowań kalibracji w Narodowym Spisie Powszechnym Ludności i Mieszkań 2011, opisanych w podrozdziałach 6.3 i 6.5, jest przedstawienie sposobu losowania próby w badaniu reprezentacyjnym.

Jak wcześniej wspomniano, badanie reprezentacyjne w NSP 2011 zaplanowano w taki sposób, aby obejmowało dużą dwudziestoprocentową próbę losową. Oznaczało to, że w próbie znalazło się około 8 mln osób mieszkających lub przebywających w około 2744 tys. mieszkań wylosowanych spośród blisko 13,5 mln mieszkań znajdujących się w operacie losowania⁶². Jednostką losowania było

⁶⁰ Szacuje się, że około 10% osób skorzystało z możliwości weryfikacji danych pochodzących z rejestrów administracyjnych. W wypadku respondentów, którzy nie skorzystali z możliwości samospisu, przyjęto zasadę, że będą odpowiadały im dane pobrane ze źródeł administracyjnych.

⁶¹ W trakcie badania reprezentacyjnego pytano m.in. o wykształcenie osób, zagadnienia związane z niepełnosprawnością czy dojazdami do pracy itp.

⁶² Opis sposobu losowania próby, obejmujący najważniejsze elementy, został przedstawiony na podstawie opracowania GUS (2012).

mieszkanie, a dokładniej jego adres. Operat losowania mieszkań został utworzony na podstawie dostępnych rejestrów i systemów informacyjnych, a następnie podlegał odpowiedniemu uporządkowaniu i podzieleniu na poszczególne warstwy. W operacie losowania próby pominięto obiekty zbiorowego zamieszkania niezawierające mieszkań, obiekty zamknięte oraz mieszkania na terenie obiektów zamkniętych, mieszkania w obiektach zbiorowego zakwaterowania stanowiących obiekty zamknięte, pomieszczenia prowizoryczne oraz mieszkania niezamieszkałe, zniszczone na skutek klęsk żywiołowych, w szczególności powodzi.

Ze względu na to, że głównym celem przeprowadzonego badania reprezentacyjnego było dostarczenie informacji o sytuacji społeczno-demograficznej na poziomie powiatów, niezbędne było dokonanie podziału założonej dwudziestoprocentowej próby mieszkań dla Polski na tym poziomie przestrzennej agregacji. W przeprowadzonym badaniu reprezentacyjnym przyjęto alokację pierwiastkową, w której liczba mieszkań losowanych w p -tym powiecie była proporcjonalna do pierwiastka kwadratowego z populacyjnej liczby mieszkań. Wyrażała się ona wzorem:

$$n_p = n \cdot \frac{\sqrt{N_p}}{\sum_p \sqrt{N_p}}, \quad (6.1)$$

gdzie n_p to liczebność próby w p -tym powiecie, n to założona liczebność próby dla Polski, a N_p to liczba mieszkań w p -tym powiecie.

Taki sposób alokacji próby pomiędzy powiaty wynikał z tego, że alokacja pierwiastkowa stanowiła kompromis pomiędzy alokacją proporcjonalną a alokacją zapewniającą jednakową precyzję dla subpopulacji. Gdyby zdecydowano się na proporcjonalne losowanie próby, oznaczałoby to, że w każdym powiecie próba stanowiłaby 20% populacji tego powiatu. Ponieważ precyzja oszacowań, tj. średnie błędy szacunku, zależą od liczebności próby, możliwe byłoby uzyskanie niedostatecznej precyzji dla mniejszych powiatów. W drugim przypadku, tj. gdyby chciano uzyskać jednakową precyzję oszacowań dla wszystkich powiatów, odbyłoby się to kosztem istotnego „spłaszczenia” próby⁶³.

W celu wylosowania w każdym z 379 powiatów⁶⁴ próby o ustalonej liczebności zgodnie ze wzorem (6.1) zastosowany został schemat losowania jednostopniowego warstwowego. Jednostki losowania, jak już zasygnalizowano, stanowiły mieszkania, przy czym zostały one przed losowaniem pogrupowane w warstwy w celu zwiększenia efektywności losowania. Warto przy tym nadmienić, że Główny

⁶³ Pewną modyfikację alokacji próby przeprowadzono dla Warszawy. Jej sposób przedstawiono w opracowaniu GUS (2012).

⁶⁴ Liczba powiatów w Polsce w 2011 roku.

ny Urząd Statystyczny zastosował różne podejście do warstwowania w zależności od typu powiatu i gminy. W miastach na prawach powiatów, ale również w większych miastach niebędących powiatami, na pierwszym etapie wyróżniono podział mieszkań na dwie kategorie: mieszkania w „blokowiskach” oraz pozostałe mieszkania. Następnie w każdej z dwóch wyodrębnionych grup dokonano warstwowania w zależności od liczby osób w mieszkaniu. W kolejnym kroku w obrębie tak utworzonych kategorii przeprowadzono dalsze warstwowanie na cztery grupy ze względu na:

- występowanie osoby pracującej w mieszkaniu,
- występowanie emeryta lub rencisty w wypadku braku pracującego,
- występowanie bezrobotnego w mieszkaniu w wypadku braku wyżej wymienionych osób,
- mieszkania z innymi osobami.

W wypadku pozostałych powiatów, w pierwszym kroku dokonano warstwowania według gmin, przy czym gminy miejsko-wiejskie traktowane były jako dwie oddzielne gminy. W mniejszych miastach pominięto etap podziału na mieszkania w „blokowiskach” i pozostałe mieszkania, a warstwowanie odbywało się jak w dużych miastach. Z kolei w gminach wiejskich w pierwszym kroku mieszkania zostały podzielone na dwie kategorie: mieszkania z użytkownikiem gospodarstwa rolnego oraz pozostałe mieszkania. W pierwszej kategorii dokonano warstwowania według liczby osób zamieszkałych w mieszkaniu, a następnie według powierzchni gospodarstwa rolnego. W odniesieniu do drugiej kategorii mieszkań wykonano warstwowanie analogicznie do małych miast.

Ostatecznie w wyniku tak przeprowadzonego warstwowania utworzono około 70,5 tys. warstw, w których wielkość próby wahała się od 6 do ponad 49%. Warto podkreślić, że po utworzeniu warstw ustalone zostały liczebności prób do wylosowania w poszczególnych warstwach. Określona w wyniku alokacji pierwiastkowej zgodnie ze wzorem (6.1) frakcja losowania w danym powiecie obowiązywała we wszystkich gminach danego powiatu oraz w warstwach utworzonych wewnątrz gmin.

Do oszacowania wartości globalnej liczby osób na poziomie agregacji typu: Polska, województwa i powiaty wykorzystywano estymator kalibracyjny wartości globalnej postaci:

$$\hat{\tau}_{\text{CAL}} = \sum_{h,i,j} w_{hij} y_{hij}, \quad (6.2)$$

gdzie h oznacza odpowiednią warstwę, i – kod mieszkania, czyli jednostkę losowania w próbie, a j – osobę zaliczoną do składu danego mieszkania. Występujące

we wzorze (6.2) wagi w_{hij} oznaczają z kolei wagi kalibracyjne, których sposób konstruowania przedstawiono w podrozdziale 6.3.

W procesie oszacowania wariancji estymatora kalibracyjnego wartości globalnej (6.2) skorzystano z metody linearyzacji Taylora. Estymator wariancji estymatora kalibracyjnego wartości globalnej (6.2) był wyznaczony z następującego wzoru:

$$\hat{D}^2(\hat{\tau}_{CAL}) = \sum_h \frac{n_h \left(1 - \frac{n_h}{N_h}\right)}{n_h - 1} \sum_i \left(\sum_j w_{hij} y_{hij} - \frac{\sum_i \sum_j w_{hij} y_{hij}}{n_h} \right)^2, \quad (6.3)$$

gdzie n_h oznacza liczbę mieszkań w danej warstwie w zrealizowanej próbie, a N_h jest liczbą mieszkań z danej warstwy w populacji. Ostatecznie do oceny jakości estymatora kalibracyjnego wartości globalnej wyznaczono wskaźniki precyzji estymacji w postaci względnego średniego błędu szacunku REE, który wyraża się wzorem (1.32).

W wypadku wyznaczania oszacowań punktowych wartości globalnej, ich wariancji oraz względnych średnich błędów szacunku na poziomach agregacji, które nie pokrywają się z warstwami stosowanymi w losowaniu próby, można dokonać estymacji wartości globalnej w domenach, w których konstruuje się zmienną pomocniczą równą analizowanej zmiennej dla rozważanej domeny (podpopulacji) oraz przyjmującą wartości zerowe poza rozważaną domeną⁶⁵.

Kluczową rolę w procesie estymacji odgrywają zatem wagi kalibracyjne, które umożliwiają uogólnianie wyników na poziom populacji oraz w odpowiednio zdefiniowanych domenach. Wykorzystanie danych z różnych źródeł wymagało jednak wcześniejszego skorygowania wag wynikających z planu losowania próby dla wszystkich osób biorących udział w badaniu reprezentacyjnym. Zabieg taki był niezbędny, aby w kontekście kluczowych zmiennych demograficznych zachować spójność pomiędzy wynikami badania reprezentacyjnego a danymi ze spisu pełnego bazującego na źródłach administracyjnych.

6.3. Kalibracja wag w NSP 2011

Wykorzystana w NSP 2011 procedura wyznaczania wag kalibracyjnych obejmowała trzy zasadnicze etapy. Zostały one w syntetyczny sposób opisane w algorytmie 7. Wagi kalibracyjne utworzono oddzielnie dla ludności rezydującej i fak-

⁶⁵ Podejście takie rozważane było we wcześniejszych rozdziałach – por. wzory (1.58)–(1.60).

tycznie zamieszkałej. W algorytmie 7 przedstawiono ogólne kroki, jakie podjęto przy wyznaczaniu wag dla obydwu kategorii ludności. Różnice w procesie kalibracji dotyczyły jedynie sposobu określania przekrojów, dla których wyznaczano wagi. W dalszej części tego podrozdziału opisano sposób wyznaczania wag kalibracyjnych dla ludności faktycznej. Dokładnie przedstawiono podejście wyznaczania wag kalibracyjnych stanowiących trzeci etap algorytmu 7. Z kolei w artykule Szymkowiaka (2014) przedstawiono sposób konstruowania wag kalibracyjnych dla ludności rezydującej⁶⁶.

Algorytm 7. Konstrukcja wag kalibracyjnych w NSP 2011

Krok 1: Wyznaczenie wag pierwotnych

W pierwszym kroku wyznaczono wagi pierwotne d_k wynikające z przyjętego sposobu losowania próby, który został szczegółowo opisany w podrozdziale 6.2. Wagi pierwotne zostały skonstruowane jako odwrotność frakcji losowania dla wspomnianych wcześniej 70,5 tys. warstw. W ramach każdej z wyznaczonych warstw wagi były identyczne. Oznacza to, że wszystkie osoby w gospodarstwach domowych z mieszkań przynależących do tej samej warstwy otrzymały tę samą wagę pierwotną.

Krok 2: Korekta wag ze względu na jednostkowe braki danych

Na drugim etapie wagi pierwotne musiały zostać skorygowane ze względu na to, że 13,7% ankiet mieszkaniowych w badaniu reprezentacyjnym z różnych powodów nie zostało wypełnionych. Tak wyznaczone wagi skorygowane, wyznaczone na drugim etapie (oznaczane przez v_k), zostały wykorzystane do uogólniania wyników spisu w zakresie mieszkań, gospodarstw domowych oraz rodzin.

Krok 3: Wyznaczenie wag kalibracyjnych dla osób

Ostatni (trzeci) etap obejmował kalibrację wag wyznaczonych w drugim kroku ich korygowania. Wynikał z tego, że nie była zachowana zgodność pomiędzy wynikami badania reprezentacyjnego a danymi ze spisu pełnego zasilanego rejestrami i systemami administracyjnymi. Brak spójności dostrzegalny był przede wszystkim w odniesieniu do podstawowych zmiennych demograficznych dotyczących płci, wieku (roczników lub grup wieku) oraz miejsca zamieszkania – poziom powiatu z wyodrębnieniem części miejskiej i wiejskiej. Wagi te zostały oznaczone jako w_k .

⁶⁶ Definicje ludności rezydującej i faktycznie zamieszkałej przyjęte na potrzeby spisu można znaleźć w opracowaniach GUS (2012) oraz Gołaty (2018).

Konieczność zastosowania podejścia kalibracyjnego opisanego szczegółowo w trzecim kroku algorytmu 7 wynikała z tego, że suma wag v_k ⁶⁷ na poziomie całego kraju, w odniesieniu do ludności faktycznie zamieszkałej, nie odtwarzała znanej liczby osób z badania pełnego (wyprowadzonej na podstawie rejestrów). Podobna sytuacja występowała w innych przekrojach wyznaczonych przez kombinację wariantów płci, roczników lub grup wieku czy miejsca zamieszkania. Również tutaj wagi v_k nie sumowały się do znanych wartości globalnych ze spisu pełnego w odpowiednich przekrojach.

Finalnie, jako zmienne pomocnicze wykorzystane w procesie kalibracji zdecydowano się wykorzystać wspomniane już zmienne demograficzne: płeć, wiek oraz miejsce zamieszkania. Wynikało to z dwóch zasadniczych powodów. Po pierwsze informacje na temat tych zmiennych demograficznych były kompletne zarówno w spisie pełnym bazującym na rejestrach, jak i w przeprowadzonym badaniu reprezentacyjnym, tj. nie zawierały braków danych. Po drugie w badaniach społecznych na potrzeby kalibracji wag, jak wskazuje literatura przedmiotu, najczęściej wykorzystuje się zmienne demograficzne (Särndal i Lundström, 2005). Można również rozpatrywać inne metody doboru zmiennych, na przykład wykorzystujące analizę składowych głównych (Cardot, Goga i Shehzad, 2017; Szymkowiak, 2017), regresję LASSO (McConville i in., 2017) czy estymację kalibracyjną z parametrem kary (Guggemos i Tillé, 2010). Tego typu podejścia nie były jednak rozpatrywane w procesie ustalania najważniejszych zmiennych pomocniczych, które należało uwzględnić w kalibracji wag w NSP 2011.

W odniesieniu do zmiennych pomocniczych kalibrację przeprowadzono w ramach każdego z 379 powiatów i 16 województw oraz w każdym przekroju wyznaczonym przez kategorię zmiennej płeć (mężczyzna, kobieta), miejsce zamieszkania (część miejska i wiejska) oraz wiek. W wypadku wieku przyjęto ponadto założenie, że na poziomie województwa kalibracji dokonywano w obrębie pojedynczych roczników wieku, tj. 0, 1, ..., 84, 85+, przy czym dla osób w wieku 85 lat i więcej utworzono jedną kategorię 85+. Oznaczało to zatem, że wagi kalibracyjne w_k wyznaczone były w taki sposób, aby odtwarzać znane struktury demograficzne w województwach ze spisu pełnego w każdym z przekrojów wyznaczonych przez kategorię płci, miejsca zamieszkania i rocznika wieku.

Ze względu na to, że nie wszystkie roczniki wieku były reprezentowane w badaniu reprezentacyjnym na poziomie powiatów, przyjęto zasadę, że na tym poziomie agregacji przestrzennej utworzone zostaną pięcioletnie grupy wieku: 0–4, 5–9, 10–12, 13–14, 15–17, 18–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49,

⁶⁷ Wagi te, jak już wspomniano, stanowiły pewną korektę wag d_k wynikających z przyjętego w badaniu reprezentacyjnym planu losowania próby.

50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84 oraz 85+. Wyjątek stanowiły osoby z przedziału wiekowego 10–19 lat, dla których utworzono dwu- oraz trzy-letnie grupy wieku. Było to związane z tym, że osoby takie były w wieku szkolnym i zachodziła potrzeba bardziej szczegółowych zestawień spisowych w kontekście edukacji młodzieży. Przyjęto ponadto zasadę, że dla pięciu największych ze względu na liczbę ludności w Polsce miast (Warszawy, Krakowa, Łodzi, Wrocławia i Poznania) wagi kalibracyjne powinny odtwarzać znaną liczbę ludności faktycznie zamieszkałej w rocznikach wieku 0, 1, ..., 99, 100+ (w wypadku Warszawy) oraz 0, 1, ..., 84, 85+ (dla pozostałych czterech miast). Nie uwzględniano podziału Warszawy na dzielnice oraz Krakowa, Łodzi, Wrocławia i Poznania na delegatury⁶⁸.

Tworzenie odpowiednich przekrojów na potrzeby kalibracji wag w NSP 2011, w kontekście ludności faktycznej, przedstawiono w tabeli 6.1. Zawarto w niej informacje na temat przekrojów, dla których możliwe było zastosowanie podejścia kalibracyjnego (1) oraz dla których odtworzenie znanych wartości globalnych w odpowiedniej podpopulacji nie było możliwe (0). Na przykład, w powiatach nie było możliwe na ogół zastosowanie kalibracji w przekroju wyznaczonym przez kombinację miejsca zamieszkania, płci oraz pojedynczych roczników wieku z tego względu, że w takim przekroju nie było w części reprezentacyjnej spisu żadnego reprezentanta. Z tego powodu dokonano agregacji wieku do odpowiednich grup. W niektórych przypadkach z oczywistych powodów (dotyczyło to miast na prawach powiatu) nie brano również pod uwagę miejsca zamieszkania (część miejska

Tabela 6.1. Sposób tworzenia przekrojów w podejściu kalibracyjnym w NSP 2011 – kategoria ludności faktycznie zamieszkałej

Poziom agregacji/ Zmienna pomocnicza	Część miejska/ wiejska – 1, 2	Płeć – 1, 2	Grupy wieku – 0–4, ... 80–84, 85+	Pojedyncze roczniki wieku	
				0, 1, ..., 83, 84, 85+	0, 1, ..., 98, 99, 100+
Polska	1	1	1	1	0
Województwa	1	1	1	1	0
Powiaty ^a	1	1	1	0	0
Warszawa	×	1	1	1	1
Pozostałe cztery miasta ^b	×	1	1	1	0

^a Bez pięciu największych miast.

^b Kraków, Łódź, Wrocław, Poznań.

Objaśnienia: 1 – kalibracja możliwa, 0 – kalibracja niemożliwa, × – przekrój nieadekwatny.

⁶⁸ Tego typu dodatkowe warunki zostały nałożone m.in. na wagi kalibracyjne w kontekście ludności rezydującej.

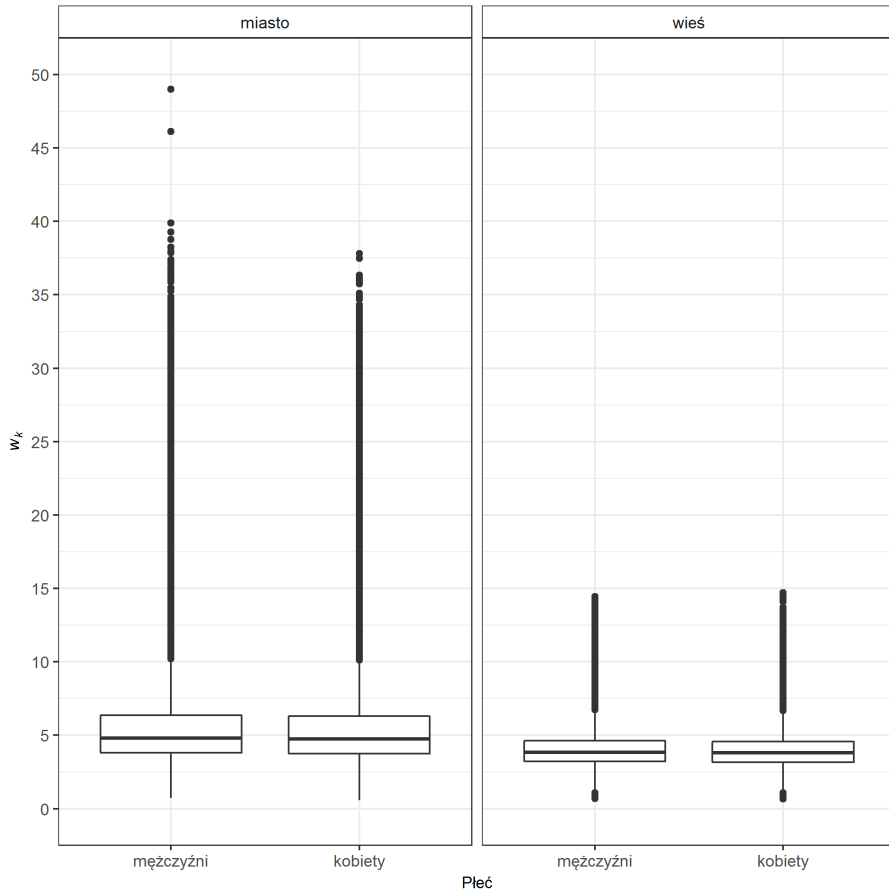
i wiejska), stąd w tabeli taki przekrój oznaczono jako nieadekwatny (\times). Ostatecznie kalibrację należało przeprowadzić w ponad 40 tys. domen wyznaczonych dla kombinacji przekrojów zdefiniowanych przez warianty trzech zmiennych pomocniczych: płci, miejsca zamieszkania i wieku.

Ze względu na to, że proces kalibracji przeprowadzono w domenach, wagi kalibracyjne zostały wyznaczone poprzez rozwiązanie odpowiedniego zadania optymalizacyjnego $(W1)_d$ – $(W3)_d$ określonego we wzorach (2.53)–(2.55). Jak wcześniej wspomniano, w pierwszej kolejności poszukiwanie wag kalibracyjnych rozpoczęto od metody liniowej. Było to konsekwencją tego, że w takim wypadku nie jest wymagane nakładanie na odpowiedni iloraz wag żadnych ograniczeń oraz że można je wyznaczyć wprost ze wzoru. Ze względu na to, że wagi kalibracyjne wykorzystujące podejście liniowe charakteryzowały się dobrymi własnościami, nie zaszła konieczność zastosowania innych aniżeli liniowa funkcji kalibracyjnych $F(\cdot)$. Ostatecznie wagi kalibracyjne wyznaczono ze wzoru (2.56), przy czym na potrzeby oceny ich własności uwaga została skupiona na czterech kluczowych aspektach odnoszących się do:

- 1) występowania wag ujemnych,
- 2) stopnia odtworzenia znanych ze spisu pełnego struktur demograficznych we wszystkich zdefiniowanych przekrojach,
- 3) stopnia skorelowania wag v_k , stanowiących korektę oryginalnych wag d_k wynikających z przyjętego planu losowania próby (ze względu na jednostkowe braki danych będące konsekwencją tego, że nie we wszystkich mieszkaniach można było przeprowadzić badanie), i wag kalibracyjnych oraz do analizy histogramów ilorazów tych wag,
- 4) występowania wag ekstremalnych.

W kontekście pierwszej z analizowanych własności wag kalibracyjnych nie stwierdzono występowania wag ujemnych w żadnym z rozpatrywanych przekrojów. Jest to szczególnie pożądana własność, gdyż nie stoi w sprzeczności z definicją wagi. Potwierdzeniem tego są rozkłady wag kalibracyjnych przedstawionych na rysunku 6.2, a wyznaczonych na poziomie całego kraju w zależności od płci i miejsca zamieszkania.

Również drugi postulat odnoszący się do stopnia odtworzenia ze spisu pełnego znanych struktur demograficznych w zdefiniowanych domenach został spełniony. Oznaczało to, że tablice spisowe wyznaczone z wykorzystaniem wag kalibracyjnych i danych pochodzących z badania reprezentacyjnego dla kombinacji wariantów zmiennych: płeć, miejsce zamieszkania oraz wiek (grupy wieku w wypadku powiatów lub pojedyncze roczniki wieku dla pięciu największych miast, województw czy całej Polski) odtwarzały dokładnie znane tablice ze spisu pełnego.

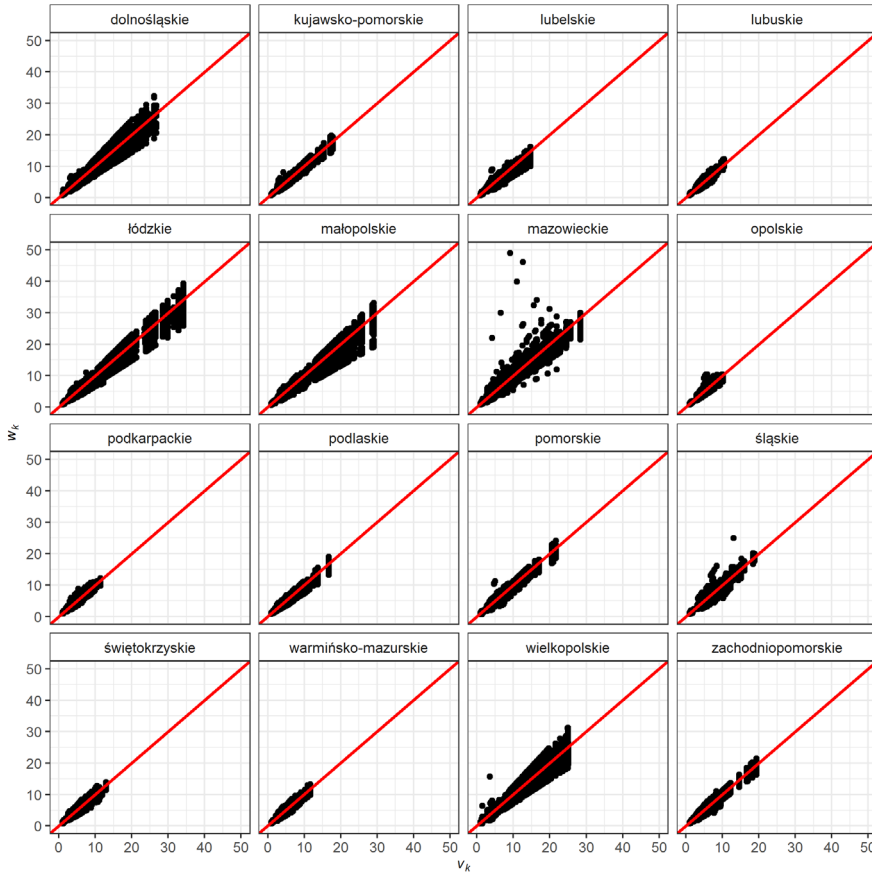


Rysunek 6.2. Rozkład wag kalibracyjnych w zależności od płci i miejsca zamieszkania

Źródło: na podstawie NSP 2011.

Własność ta dla niektórych funkcji kalibracyjnych nie zawsze musi być spełniona, o czym była mowa wcześniej.

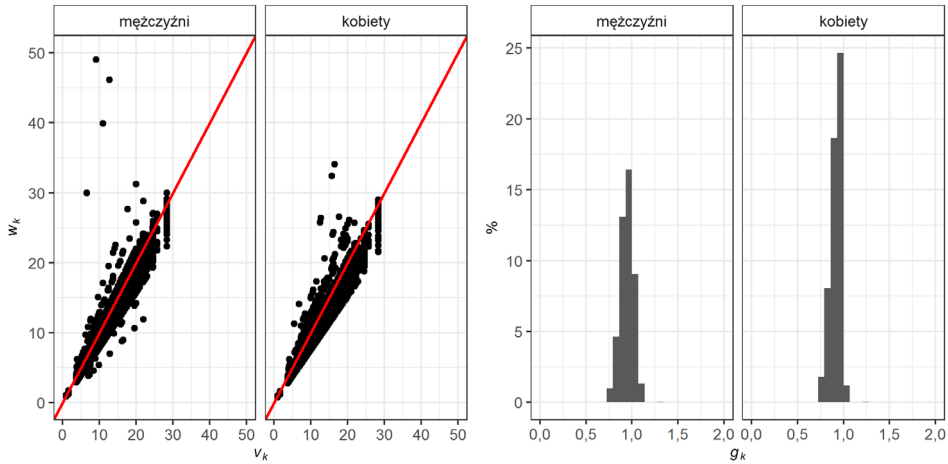
W analizie własności wyznaczonych wag kalibracyjnych we wszystkich założonych przekrojach dokonano również oceny stopnia skorelowania wag kalibracyjnych w_k i wag v_k . Na przykład, na rysunku 6.3 przedstawiono korelacyjne diagramy rozrzutu wag v_k i w_k w przekroju województw. Można zauważyć wyraźną silną dodatnią zależność między dwoma rozpatrywanymi rodzajami wag. Jest to pożądana własność, co oznacza, że wyższym wagom v_k odpowiadają wyższe wartości wag w_k po kalibracji.



Rysunek 6.3. Korelacyjny diagram rozrzutu wag v_k i wag kalibracyjnych w_k w przekroju województw

Źródło: na podstawie NSP 2011.

Warto również zwrócić uwagę, że przy analizie rozrzutu wag ważny jest nie tylko stopień skorelowania, ale także sposób układania się punktów. Pożądana jest przy tym sytuacja, w której punkty będą się koncentrować wzdłuż prostej $y = x$. Oznacza to bowiem, że wagi w_k nie będą odbiegać zbyt mocno od wag v_k . W zasadzie jedynie dla województwa mazowieckiego zaobserwowano występowanie kilku punktów, które świadczą o tym, że po kalibracji waga v_k znacznie zmieniła swoją wartość. W tym wypadku było to związane z wpływem Warszawy, dla której przyjęto założenie, że wagi kalibracyjne odtwarzać będą znane wartości globalne ze spisu pełnego w pojedynczych rocznikach wieku $0, 1, \dots, 99, 100+$ z uwzględnieniem zmiennej odnoszącej się do płci osoby.



Rysunek 6.4. Korelacyjny diagram rozrzutu wag v_k i w_k i histogram mnożników kalibracyjnych g_k dla Warszawy w przekroju płci

Źródło: na podstawie NSP 2011.

Potwierdzeniem tego jest rysunek 6.4, na którym przedstawiono korelacyjny diagram rozrzutu wag v_k i w_k i histogram mnożników kalibracyjnych $g_k = w_k/v_k$ dla Warszawy w przekroju płci. Wartości wag kalibracyjnych w_k , które znacznie różniły się od wag v_k , występują zarówno w grupie mężczyzn, jak i kobiet. W wypadku Warszawy można jednak zaobserwować, że wag ekstremalnych jest kilka. Co więcej, dotyczyły one starszych roczników wieku i było to związane z tym, że w danym roczniku w badaniu reprezentacyjnym mógł wystąpić jeden reprezentant w próbie, podczas gdy w spisie pełnym było ich więcej. Analiza histogramu mnożników kalibracyjnych g_k pozwala z kolei zauważyć, że koncentrowały się one w pobliżu jedynki, przy czym wartość maksymalna wynosiła $g_{\max} = 5,378^{69}$.

Warto również wyraźnie podkreślić, że w domenach, w których występują wagi ekstremalne, w procesie estymacji należy zachować szczególną uwagę. Byłoby to istotne zwłaszcza wówczas, gdyby omawiane zjawisko było obserwowalne na dużą skalę. W odniesieniu do ludności faktycznej, ze względu na odpowiednio dobrane zmienne pomocnicze oraz przyjęty w spisie sposób losowania próby, w większości domen nie stwierdzono występowania ekstremalnych wag kalibracyjnych. Ich wpływ w procesie uogólniania wyników jest również marginalny w wypadku wielu poziomów agregacji (jak województwo czy powiat). Ich rola

⁶⁹ Wykres mnożników kalibracyjnych g_k w celu czytelności ograniczono do przedziału $[0, 2]$. Wynikało to z tego, że mnożników kalibracyjnych g_k większych od 2 w wypadku Warszawy było jedynie 10.

może rosnąć w bardziej szczegółowo zdefiniowanych domenach uwzględniających starsze roczniki wieku. Ze względu na to, że wag ekstremalnych w takich przekrojach nie było zbyt wiele, a i tablice spisowe w większości wykorzystywały w swojej konstrukcji grupy wieku jako jedną ze zmiennych, nie miały one większego znaczenia w procesie estymacji.

Wagi v_k , które stanowiły korektę wag d_k ze względu na jednostkowe braki danych, miały tę własność, że przyjmowały taką samą wartość dla wszystkich osób w ramach warstwy. Oznacza to w konsekwencji, że wszystkie osoby wchodzące w skład danego gospodarstwa domowego miały również tę samą wagę. Waga ta została użyta do uogólniania wyników spisu w kontekście gospodarstw domowych. Przeprowadzona na trzecim etapie kalibracja wag, mimo że odtwarzała znane wartości globalne dla kluczowych zmiennych demograficznych oraz miała inne pożądane własności, zmieniała wagi przypisane osobom w obrębie tego samego gospodarstwa domowego. W konsekwencji oznaczało to, że inny zestaw wag musiał być wykorzystywany w odniesieniu do gospodarstw domowych, a inny dla osób w procesie uogólniania wyników spisowych.

Powyzszą niedogodność eliminuje tzw. kalibracja zintegrowana, która nie została wykorzystana w praktyce w NSP 2011. Zastosowanie podejścia zintegrowanego spowodowałoby, że nie tylko mogłyby zostać odtworzone jednocześnie odpowiednie wartości globalne zmiennych pomocniczych na poziomie osób i gospodarstw domowych, ale dodatkowo każda osoba w ramach danego gospodarstwa domowego otrzymałaby identyczną wagę. Ten sam zestaw wag kalibracyjnych mógłby zatem zostać wykorzystany w procesie uogólniania wyników dla gospodarstw domowych oraz osób.

W dalszej części książki, w podrozdziale 6.4, przedstawiono wybrane aspekty teoretyczne kalibracji zintegrowanej. Z kolei w podrozdziale 6.5 dokonano egzemplifikacji możliwości zastosowania podejścia zintegrowanego w spisach mieszanych na przykładzie NSP 2011.

6.4. Teoretyczne podstawy kalibracji zintegrowanej

W badaniach statystycznych dotyczących gospodarstw domowych bardzo często zachodzi potrzeba jednoczesnej estymacji parametrów dla gospodarstw domowych oraz osób. W związku z tym wymagane są dwa różne systemy wag: pierwszy do szacowania charakterystyk na poziomie gospodarstw domowych, a drugi – na poziomie osób. Takie podejście ma jednak dwie zasadnicze niedogodności. Po

pierwsze wymagana jest konstrukcja dwóch systemów wag. Po drugie wagi przypisane do osób powstałe w wyniku procesu kalibracji nie uwzględniają przynależności kilku osób do tego samego gospodarstwa domowego. W praktyce wagi przypisywane gospodarstwom domowym często są pochodną wag przypisanych do osób należących do danego gospodarstwa domowego. Na przykład, Alexander (1987) proponuje, aby waga gospodarstwa domowego była tożsama z wagą przypisaną do głowy gospodarstwa domowego. Z kolei Chowdhury (1994) sugeruje wyznaczenie średniej harmonicznej wag przypisanych osobom z danego gospodarstwa domowego w celu uzyskania wagi dla tego gospodarstwa.

Zaproponowane rozwiązania nie zyskały jednak w praktyce badań statystycznych szerszego uznania. Postulowano bowiem konieczność utworzenia innego systemu wag, który mógłby zostać wykorzystany zarówno w uogólnianiu wyników dla gospodarstw domowych, jak i osób. W podejściu takim zakłada się, że wszystkie osoby wchodzące w skład danego gospodarstwa domowego powinny mieć identyczne wagi. Waga przypisana do tych osób jest jednocześnie wagą odnoszącą się do danego gospodarstwa domowego. Tego typu rozwiązanie jest szczególnie zalecane i rekomendowane przez Eurostat we wszystkich badaniach dotyczących gospodarstw domowych i jest przedmiotem badań tzw. kalibracji zintegrowanej (Estevao i Särndal, 2006).

Początków utworzenia jednego systemu wag w badaniach statystycznych dla gospodarstw domowych i osób, w którym wszystkie osoby w ramach tego samego gospodarstwa domowego otrzymałyby identyczną wagę, można się doszukiwać już w opracowaniu Lemaître'a i Dufoura (1987). Zaproponowane przez autorów rozwiązanie sprowadzało się do utworzenia zmiennych kalibracyjnych wykorzystywanych w procesie kalibracji na poziomie osób, którym odpowiadają średnie wartości zmiennych pomocniczych wyznaczanych w obrębie każdego gospodarstwa domowego oddzielnie. Ideę zaproponowaną przez Lemaître'a i Dufoura (1987) rozwinęli następnie Renssen i Nieuwenbroek (1997), a także Neethling i Galpin (2006). W pracach tych autorów rozpatrywano problem wyznaczania wag z wykorzystaniem tylko zmiennych pomocniczych na poziomie gospodarstwa domowego bądź gospodarstwa domowego i osób jednocześnie.

W pierwszej kolejności w książce rozpatrzona zostanie idea Lemaître'a i Dufoura (1987) konstrukcji takich samych wag dla osób w ramach danego gospodarstwa domowego⁷⁰. W podejściu tym zakłada się jedynie, że zmienne pomocnicze

⁷⁰ W gruncie rzeczy autorzy nie użyli pojęcia „kalibracja wag” – artykuł Lemaître'a i Dufoura ukazał się pięć lat przed wspomnianą już pracą Deville'a i Särndala (1992), którą uznaje się formalnie za początek nowoczesnej teorii kalibracji. W niniejszej książce pokazano, jak ideę tę można przełożyć na „język” kalibracji.

są dostępne na poziomie osób. Na kolejnym etapie uwzględnione zostaną warunki, że zmienne pomocnicze mogą być dostępne także na poziomie gospodarstw domowych. Na końcu, zgodnie z propozycją Estevao i Särndala (2006), przedstawiona zostanie idea kalibracji zintegrowanej w kontekście losowania dwustopniowego, dla której obydwa wspomniane podejścia stanowią szczególne przypadki.

Rozważmy w dalszym ciągu populację złożoną z N osób, tj. $U = \{1, \dots, N\}$, które wchodzi w skład gospodarstw domowych. Niech U_M oznacza populację złożoną z N_M gospodarstw domowych, przy czym $U_M = \{1, \dots, N_M\}$. W podejściu rozpatrywanym w pracy Lemaître'a i Dufoura (1987) zakłada się, że próba s_M złożona z M gospodarstw domowych jest losowana zgodnie z określonym planem losowania z populacji U_M . Wszystkie osoby z wylosowanego gospodarstwa domowego biorą dalej udział w badaniu⁷¹. Przyjmujemy, że π_k oznacza prawdopodobieństwo, że dane gospodarstwo domowe zostanie wylosowane do próby, oraz że $d_k = 1/\pi_k$. Ponieważ wszystkie osoby z wylosowanego gospodarstwa domowego biorą udział w badaniu, waga ta jest jednocześnie wagą przypisaną osobom i jest jednakowa dla wszystkich osób z danego gospodarstwa domowego. Zakłada się przy tym, że w próbie s , do której należą wszystkie osoby z wylosowanych gospodarstw domowych, znalazło się n osób. Przyjmując, że gospodarstwo h składa się z m_h osób, prawdziwa jest następująca równość:

$$\sum_{h=1}^M m_h = n. \quad (6.4)$$

Punktem wyjścia w podejściu Lemaître'a i Dufoura (1987) jest macierz \mathbf{X}_s zdefiniowana wzorem (2.60). W tym podejściu zakładamy, że znane są jedynie zmienne pomocnicze na poziomie osób. Nie rozpatrujemy zmiennych pomocniczych na poziomie gospodarstw domowych⁷². Dla odróżnienia, że mamy tutaj do czynienia z sytuacją, w której osoby należą do odpowiednich gospodarstw domowych, oznaczaj ją będziemy przez \mathbf{X}_{pp} ⁷³. Macierz \mathbf{X}_{pp} można wyrazić wzorem

⁷¹ Można również rozpatrywać inne warianty polegające na tym, że tylko niektóre osoby z wylosowanych gospodarstw domowych biorą udział w badaniu. W niniejszym opracowaniu zakładamy jednak będziemy, że wszystkie osoby z wylosowanego gospodarstwa domowego włączane są do próby. Wynika to z tego, że w części reprezentacyjnej NSP 2011 badano wszystkie osoby z gospodarstw domowych zamieszkujących mieszkania wylosowane z danej warstwy.

⁷² W oryginalnym artykule Lemaître'a i Dufoura (1987) nie brano pod uwagę zmiennych pomocniczych na poziomie gospodarstw domowych, a jedynie na poziomie osób. Precyzując, autorzy rozpatrywali w badaniu tylko dwie takie zmienne: płeć oraz grupy wieku.

⁷³ Subskrypt pp odnosi się do tego, że informacje w postaci zmiennych pomocniczych mierzone są jedynie na poziomie osób (ang. *person-person*).

(6.5). W macierzy tej wektor $\mathbf{x}_k = (x_{k1}, \dots, x_{kJ_1})^T$ jest złożony z wartości wszystkich J_1 zmiennych pomocniczych dla k -tej osoby, $k = 1, \dots, n$ ⁷⁴. Wektorowi temu odpowiada k -ty wiersz macierzy \mathbf{X}_{pp} , tj. x_{kj} oznacza wartość j -tej zmiennej pomocniczej dla k -tej osoby.

$$\mathbf{X}_{pp} = \begin{matrix} & x_1 & x_2 & x_3 & \dots & x_{J_1} \\ h_1p_1 & \left(x_{11} & x_{12} & x_{13} & \dots & x_{1J_1} \right) \\ h_1p_2 & \left(x_{21} & x_{22} & x_{23} & \dots & x_{2J_1} \right) \\ h_1p_3 & \left(x_{31} & x_{32} & x_{33} & \dots & x_{3J_1} \right) \\ h_2p_1 & \left(x_{41} & x_{42} & x_{43} & \dots & x_{4J_1} \right) \\ h_2p_2 & \left(x_{51} & x_{52} & x_{53} & \dots & x_{5J_1} \right) \\ \vdots & \left(\vdots & \vdots & \vdots & \vdots & \vdots \right) \\ h_m p_{m_h} & \left(x_{n1} & x_{n2} & x_{n3} & \dots & x_{nJ_1} \right) \end{matrix}. \quad (6.5)$$

Na przykład, przy przyjętym sposobie etykietowania wierszy, w macierzy tej trzy pierwsze osoby należą do pierwszego gospodarstwa domowego, kolejne dwie osoby należą do gospodarstwa drugiego itd.⁷⁵

Zakładamy, że naszym celem jest kalibracja wag d_k ⁷⁶ przypisanych wszystkim osobom wylosowanym do próby tak, aby odtworzone zostały wartości globalne wszystkich zmiennych pomocniczych x_1, \dots, x_{J_1} oraz aby skalibrowane wagi w_k były jednakowe w obrębie danego gospodarstwa domowego. Znalezienie wag kalibracyjnych w_k wymaga rozwiązania zadania optymalizacyjnego (2.4)–(2.6). Niestety, tak wyznaczone wagi kalibracyjne w_k na ogół będą się różniły w obrębie danego gospodarstwa domowego, mimo że odtwarzać będą wartości globalne zmiennych pomocniczych.

Problem ten jednak można w dość łatwy sposób rozwiązać. Zgodnie z ideą Lemaître'a i Dufoura (1987), wystarczy w tym celu dokonać pewnego przekształcenia zmiennych pomocniczych x_1, \dots, x_{J_1} i utworzyć macierz \mathbf{Z}_{pp} o wymiarach $n \times J_1$, która zawiera średnie arytmetyczne odpowiednich zmiennych pomocni-

⁷⁴ Zakładamy, że liczba zmiennych pomocniczych na poziomie osób wynosi J_1 . W dalszej części będziemy również przyjmować, że dysponujemy J_2 zmiennymi na poziomie gospodarstw domowych, co da łącznie $J_1 + J_2 = J$ zmiennych pomocniczych, które można wykorzystać w procesie kalibracji.

⁷⁵ Symbol h odnosi się do gospodarstwa, a p do osoby. Na przykład, h_2p_1 oznacza pierwszą osobę w drugim gospodarstwie. Ogólnie, $h_i p_j$ oznacza j -tą osobę w i -tym gospodarstwie domowym.

⁷⁶ Wagi d_k przed procesem kalibracji mogą być wcześniej skorygowane ze względu na jednostkowe braki danych – gdy wylosowane do próby gospodarstwo domowe z różnych powodów nie bierze w nim udziału. W tym celu wykorzystywane są wskaźniki odpowiedzi.

czych. Z racji tego, że zmienne pomocnicze są zazwyczaj jakościowe (płeć, grupy wieku, miejsce zamieszkania itd.) oraz kodowane w macierzy \mathbf{X}_{pp} w postaci zero-jedynkowej, macierz \mathbf{Z}_{pp} będzie zawierała odpowiednie frakcje. W związku z tym, dla każdej osoby z gospodarstwa domowego h , którego liczebność wynosi m_h , wartość przekształconej j -tej zmiennej pomocniczej, dla $j = 1, \dots, J_1$, wynosić będzie:

$$z_{hj} = \frac{a_{hj}}{m_h}, \quad (6.6)$$

gdzie:

$$a_{hj} = \sum_{k \in h} x_{kj}. \quad (6.7)$$

Przyjmując, że $\mathbf{z}_{k_1} = (z_{h_1}, \dots, z_{h_{J_1}})^T$ oznacza wektor zmiennych pomocniczych dla k -tej osoby należącej do h -tego gospodarstwa domowego, proces poszukiwania zintegrowanych wag kalibracyjnych $w_{k_{IN}}$ w ujęciu Lemaître'a i Dufoura można przedstawić w postaci następującego zadania optymalizacyjnego⁷⁷:

- (W1)_{IN} – minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s} \frac{d_k}{q_k} G\left(\frac{w_{k_{IN}}}{d_k}\right) \rightarrow \min, \quad (6.8)$$

- (W2)_{IN} – równania kalibracyjne:

$$\sum_{k \in s} w_{k_{IN}} \mathbf{z}_{k_1} = \sum_{k \in U} \mathbf{z}_{k_1}, \quad (6.9)$$

- (W3)_{IN} – warunki ograniczające:

$$L \leq \frac{w_{k_{IN}}}{d_k} \leq U, \quad \text{gdzie: } 0 \leq L \leq 1 \leq U, \quad k = 1, \dots, n. \quad (6.10)$$

W równaniu (6.9) wektor wartości globalnych jest postaci:

$$\sum_{k \in U} \mathbf{z}_{k_1} = \sum_{k \in U} \mathbf{x}_k = \left(\sum_{k \in U} x_{k_1}, \dots, \sum_{k \in U} x_{k_{J_1}} \right)^T. \quad (6.11)$$

W wypadku funkcji $G(x) = \frac{1}{2}(x-1)^2$ oraz warunków (6.8)–(6.9) wektor zintegrowanych wag kalibracyjnych można wyznaczyć wprost ze wzoru. Mówi o tym poniższe twierdzenie.

⁷⁷ Subskrypt IN pochodzi od angielskiego wyrażenia *integrated calibration* – kalibracja zintegrowana.

Twierdzenie 6.1. Rozwiązaniem zadania minimalizacji (6.8) przy warunku (6.9) dla funkcji $G(x) = \frac{1}{2}(x-1)^2$ jest wektor zintegrowanych wag kalibracyjnych $w_{\text{IN}} = (w_{1\text{IN}}, \dots, w_{n\text{IN}})^T$, którego składowe, dla $k = 1, \dots, n$, wyrażają się wzorem:

$$w_{k\text{IN}}^{\text{lin}} = d_k + d_k q_k \left(\sum_{k \in U} \mathbf{z}_{k_1} - \sum_{k \in S} d_k \mathbf{z}_{k_1} \right)^T \left(\sum_{k \in S} q_k d_k \mathbf{z}_{k_1} \mathbf{z}_{k_1}^T \right)^{-1} \mathbf{z}_{k_1}. \quad (6.12)$$

W wypadku innych funkcji odległości oraz ograniczeń nakładanych na ilość wag, w celu znalezienia wag zintegrowanych należy skorzystać z algorytmu 1 lub 2. Wagi wyznaczone jako rozwiązanie zadania optymalizacyjnego (6.8)–(6.10) będą miały dwie zasadnicze własności. Po pierwsze będą odtwarzały znane wartości globalne zmiennych pomocniczych x_1, \dots, x_{J_1} ustalonych na poziomie osób. Po drugie w ramach danego gospodarstwa domowego wszyscy jego członkowie uzyskają jednakowe wagi. Tak sformułowane zagadnienie poszukiwania wag prowadzi zatem do znalezienia tzw. zintegrowanych wag kalibracyjnych.

Gdyby w procesie estymacji należało dokonać oszacowania wartości globalnej (1.34) pewnej zmiennej Y na poziomie osób, wówczas można byłoby skorzystać z estymatora kalibracyjnego wartości globalnej (2.14), w którym w miejsce wag w_k należałoby wstawić wagi zintegrowane $w_{k\text{IN}}$. Z kolei estymator wariancji tak skonstruowanego estymatora kalibracyjnego wartości globalnej mógłby być postaci (2.47), przy czym wagi w_k należałoby zastąpić przez $w_{k\text{IN}}$.

Podejście wyznaczania wag zintegrowanych $w_{k\text{IN}}$ wykorzystujące zmienne pomocnicze dostępne jedynie na poziomie osób można byłoby również zastosować w przekroju domen, na które podzielona może być populacja. W tym celu należałoby rozwiązać problem optymalizacyjny podobny do tego, który został sformułowany w zadaniu (2.53)–(2.55). W takim przypadku, podobnie jak wcześniej, w miejsce wektora \mathbf{x}_k należałoby podstawić wektor \mathbf{z}_{k_1} . Estymator wariancji estymatora wartości globalnej w domenie d można byłoby również wyznaczyć ze wzoru (2.47), przy czym reszty obliczane byłyby ze wzoru (2.59) i przy przyjęciu założenia, że w miejsce wag w_k podstawiane są wagi zintegrowane $w_{k\text{IN}}$. W podejściu tym przyjmujemy ponadto, że znane są wartości globalne zmiennych pomocniczych na poziomie domeny.

Opisana powyżej metoda wyznaczania wag zintegrowanych $w_{k\text{IN}}$, wykorzystująca koncepcję Lemaître'a i Dufoura (1987) przekształcania oryginalnego wektora zmiennych pomocniczych \mathbf{x}_k w wektor \mathbf{z}_{k_1} , nie brała pod uwagę tego, że zmienne pomocnicze mogą być również dostępne na poziomie gospodarstw domowych. Może się bowiem zdarzyć, że będziemy dysponować informacją na temat wielko-

ści gospodarstwa domowego (1-osobowe, 2-osobowe itd.), województwa zamieszkania czy typu miejsca zamieszkania (obszar wiejski i miejski) danego gospodarstwa domowego. W procesie kalibracji wagi zintegrowane można zatem wyznaczyć, wykorzystując zarówno zmienne pomocnicze na poziomie osób (płeć, grupa wieku itd.), jak i wspomniane zmienne pomocnicze mierzone na poziomie gospodarstw domowych.

Założmy w dalszym ciągu, że oprócz J_1 zmiennych pomocniczych na poziomie osób, dysponujemy również J_2 zmiennymi pomocniczymi na poziomie gospodarstw domowych. Zmienne te będą również najczęściej kodowane w postaci zero-jedynkowej. Na przykład, mając informację o typie miejsca zamieszkania (obszar wiejski oraz miejski) danego gospodarstwa domowego, można utworzyć dwie zmienne dychotomiczne. Jedna będzie odpowiadać gospodarstwom domowym zamieszkującym obszar wiejski, a druga gospodarstwom domowym zamieszkującym obszar miejski⁷⁸. Również w tym przypadku można utworzyć odpowiednią macierz \mathbf{X}_{ph} o wymiarach $n \times J$, gdzie $J = J_1 + J_2$ ⁷⁹. Macierz ta będzie przechowywała informacje o zmiennych pomocniczych na poziomie osoby i gospodarstwa domowego. W wypadku wspomnianych już dwóch zmiennych dychotomicznych odnoszących się do typu obszaru zamieszkiwanego przez dane gospodarstwo domowe, wszystkie osoby z gospodarstw domowych zamieszkujących na wsi otrzymałyby wartość 1 dla pierwszej zmiennej oraz 0 w przeciwnym razie. Podobnie dla drugiej zmiennej dychotomicznej wartość 1 otrzymałyby wszystkie osoby z gospodarstw domowych zamieszkujących obszar miejski i 0 w przeciwnym razie.

Macierz \mathbf{X}_{ph} można ogólnie przedstawić w następującej postaci:

$$\mathbf{X}_{ph} = \begin{matrix} & x_1 & \dots & x_{J_1} & \dots & x_{J_2} \\ h_1p_1 & \left(\begin{matrix} x_{11} & \dots & x_{1J_1} & \dots & x_{1J_2} \\ x_{21} & \dots & x_{2J_1} & \dots & x_{2J_2} \\ x_{31} & \dots & x_{3J_1} & \dots & x_{3J_2} \\ x_{41} & \dots & x_{4J_1} & \dots & x_{4J_2} \\ x_{51} & \dots & x_{5J_1} & \dots & x_{5J_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & \dots & x_{nJ_1} & \dots & x_{nJ_2} \end{matrix} \right) \end{matrix} \quad (6.13)$$

⁷⁸ W odniesieniu do pierwszej zmiennej 1 będzie oznaczało, że gospodarstwo domowe zamieszkuje obszar wiejski, a 0 – że zamieszkuje obszar miejski. W analogiczny sposób można dokonać kodowania drugiej zmiennej.

⁷⁹ Subskrypt *ph* odnosi się do tego, że informacje w postaci zmiennych pomocniczych mierzone są na poziomie osób oraz gospodarstw domowych (ang. *person-household*).

Również w tym przypadku możemy dokonać konstrukcji odpowiedniej macierzy \mathbf{Z}_{ph} o wymiarach $n \times J$, która będzie zawierać przekształcone wartości zawarte w macierzy \mathbf{X}_{ph} . Przekształcenia zmiennych pomocniczych zawartych w macierzy \mathbf{X}_{ph} mierzonych na poziomie osób (w pierwszych J_1 kolumnach) dokonujemy zgodnie z wcześniej poczynionym opisem, tj. wykorzystując wzór (6.6). Z kolei w odniesieniu do zmiennych pomocniczych zawartych w macierzy \mathbf{X}_{ph} mierzonych na poziomie gospodarstw domowych (w kolejnych J_2 kolumnach) ich przekształcenie polega na przyjęciu odwrotności liczebności gospodarstwa domowego m_h dla zmiennej dychotomicznej, do której przynależy dane gospodarstwo, oraz wartości 0 dla pozostałych zmiennych⁸⁰.

Przyjmując, że $\mathbf{z}_{k_2} = (z_{h1}, \dots, z_{hJ})^T$ jest wektorem zmiennych pomocniczych odnoszącym się do k -tej osoby wchodzącej w skład h -tego gospodarstwa domowego, w którym pierwszych J_1 składowych dotyczy przekształconych zmiennych na poziomie osoby, a kolejnych J_2 zmiennych na poziomie gospodarstwa domowego, proces poszukiwania zintegrowanych wag kalibracyjnych $w_{k_{IN}}$ można zapisać w postaci zadania optymalizacyjnego (6.8)–(6.10), przy czym w miejsce wektora \mathbf{z}_{k_1} należy podstawić wektor \mathbf{z}_{k_2} . Warto przy tym podkreślić, że wektor wartości globalnych $\sum_{k \in U} \mathbf{z}_{k_2}$ w takiej sytuacji składałby się z J_1 wartości globalnych odnoszących się do osób oraz J_2 wartości globalnych odnoszących się do gospodarstw domowych. Na przykład, dla typu obszaru zamieszkania danego gospodarstwa domowego niezbędna byłaby informacja o wartościach globalnych liczby gospodarstw domowych zamieszkujących obszar wiejski i miejski.

Wyznaczone w ten sposób wagi będą miały dwie zasadnicze własności. Po pierwsze będą odtwarzały nie tylko znane wartości globalne zmiennych pomocniczych ustalonych na poziomie osób, ale również wartości globalne zmiennych pomocniczych dotyczących gospodarstw domowych. Po drugie również w tym wypadku, w ramach danego gospodarstwa domowego wszyscy jego członkowie otrzymają jednakowe wagi. Prowadzi to zatem do uzyskania zintegrowanych wag kalibracyjnych, w których procesie konstrukcji wykorzystywano zmienne pomocnicze odnoszące się zarówno do osób, jak i gospodarstw domowych. Opisany proces kalibracji wag jest zatem rozszerzeniem podejścia kalibracyjnego wykorzystującego koncepcję Lemaître'a i Dufoura (1987). Uwzględniana jest bowiem

⁸⁰ Na przykład, gdyby w macierzy \mathbf{X}_{ph} określonej wzorem (6.13) pierwsze gospodarstwo zamieszkiwało wieś, wówczas w macierzy tej wszystkie osoby z tego gospodarstwa domowego otrzymałyby wartość 1 dla pierwszej zmiennej dychotomicznej (1 – obszar wiejski, 0 – obszar miejski) oraz wartość 0 dla drugiej zmiennej dychotomicznej (0 – obszar wiejski, 1 – obszar miejski). Z kolei w macierzy \mathbf{Z}_{ph} zawierającej transformowane wartości z macierzy \mathbf{X}_{ph} pierwsza z przekształconych zmiennych dla tego gospodarstwa domowego przyjmowałaby wartość $1/3$, a druga wartość 0.

dotatkowo informacja w postaci zmiennych pomocniczych mierzonych na poziomie gospodarstwa domowego⁸¹.

W praktyce badań statystycznych, o czym była mowa w podrozdziale 1.2, najczęściej wykorzystywane jest losowanie dwustopniowe. Dotyczy to m.in. wspomnianego już Badania Aktywności Ekonomicznej Ludności, Badania Budżetów Domowych czy Europejskiego Badania Warunków Życia Ludności. W związku z tym zostanie opisany sposób konstrukcji wag zintegrowanych w wypadku, gdy próba jest pobierana zgodnie ze schematem losowania dwustopniowego (Estevao i Särndal, 2006).

Przyjmijmy w dalszym ciągu, że populacja $U = \{1, \dots, N\}$ składa się z N jednostek, połączonych w grupy (zespoły). Z populacji grup $U_M = \{1, \dots, N_M\}$ na pierwszym stopniu losujemy próbę s_M złożoną z M grup, przy czym znane są prawdopodobieństwa inkluzji pierwszego rzędu π_{Mi} dla $i \in U_M$ oraz wagi $d_{Mi} = \frac{1}{\pi_{Mi}}$. Na drugim stopniu z wylosowanych grup losowane są próby s_i , dla $i \in s_M$, złożone z jednostek losowania drugiego stopnia, ze znanymi prawdopodobieństwami inkluzji drugiego rzędu (warunkowymi) $\pi_{k|i}$, dla $k \in s_i$, oraz wagami $d_{k|i} = \frac{1}{\pi_{k|i}}$. Wówczas $d_k = d_{Mi}d_{k|i}$ jest całkowitą wagą przypisaną jednostce k . Ponadto s jest n -elementową próbą złożoną z jednostek losowania drugiego stopnia, przy czym $s = \bigcup_{i \in s_M} s_i$ ⁸².

W badaniach statystycznych, w których próba jest pobierana zgodnie z opisanym powyżej losowaniem dwustopniowym, celem może być oszacowanie pewnych parametrów zarówno na poziomie populacji U złożonej z jednostek, które przynależą do grup, jak i populacji U_M złożonej z grup⁸³. Załóżmy w dalszym ciągu, że przedmiotem zainteresowania jest oszacowanie wartości globalnej pewnej zmiennej $Y_{(u)}$ na poziomie jednostek drugiego stopnia⁸⁴:

$$\tau_{Y_{(u)}} = \sum_{k \in U} y_{(u)k}, \quad (6.14)$$

⁸¹ Również w tym wypadku, dla tak wyznaczonych zintegrowanych wag kalibracyjnych $w_{k_{IN}}$, można dokonać estymacji wartości globalnej (1.34) pewnej zmiennej Y na poziomie osób. Sposób konstrukcji estymatora wariancji estymatora kalibracyjnego wartości globalnej, wyznaczania wag zintegrowanych dla domen oraz estymacji wartości globalnej w domenach, wraz z budową estymatora wariancji, jest analogiczny do opisanego wcześniej. Wystarczy jedynie w odpowiednich wzorach w miejsce wag w_k podstawić wagi zintegrowane $w_{k_{IN}}$.

⁸² Grupami mogą być na przykład rejony statystyczne, a jednostkami mieszkania. Grupami mogą być także gospodarstwa domowe, a jednostkami osoby w wieku 15+.

⁸³ Na przykład, gdy grupami są gospodarstwa domowe, a jednostkami osoby w wieku 15+, celem może być oszacowanie średniego dochodu gospodarstw domowych i jednocześnie liczby osób bezrobotnych.

⁸⁴ Subskrypt (u) odnosi się do jednostek losowania drugiego stopnia (ang. *unit*). W dalszej części wywodu będą one określane mianem jednostek.

gdzie $y_{(u)k}$ oznacza wartość zmiennej $Y_{(u)}$ dla jednostki $k \in U$ oraz wartości globalnej pewnej zmiennej $Y_{(g)}$ na poziomie jednostek pierwszego stopnia (grup)⁸⁵:

$$\tau_{Y_{(g)}} = \sum_{i \in U_M} y_{(g)i}, \quad (6.15)$$

gdzie $y_{(g)i}$ oznacza wartość zmiennej $Y_{(g)}$ dla każdej grupy $i \in U_M$.

Założmy, że estymacji wartości globalnej (6.14) oraz (6.15) będziemy dokonywać z wykorzystaniem zmiennych pomocniczych, które mogą być dostępne zarówno na poziomie jednostek, jak i grup. Przyjmujemy zatem następujące założenia:

- znany jest wektor wartości globalnych $\sum_{i \in U_M} \mathbf{x}_{(g)i}$ zmiennych pomocniczych na poziomie grup,
- dla każdej grupy $i \in s_M$ znany jest wektor $\mathbf{x}_{(g)i}$ wartości zmiennych pomocniczych,
- znany jest wektor wartości globalnych $\sum_{k \in U} \mathbf{x}_{(u)k}$ zmiennych pomocniczych na poziomie jednostek,
- dla każdej jednostki $k \in s$ znany jest wektor $\mathbf{x}_{(u)k}$ wartości zmiennych pomocniczych.

Przy tak przyjętych założeniach estymatorem kalibracyjnym wartości globalnej (6.14) jest:

$$\hat{\tau}_{(u)\text{CAL}} = \sum_{k \in s} w_k y_{(u)k}, \quad (6.16)$$

a estymatorem kalibracyjnym wartości globalnej (6.15) jest:

$$\hat{\tau}_{(g)\text{CAL}} = \sum_{i \in s_M} w_{Mi} y_{(g)i}, \quad (6.17)$$

przy czym wagi w_{Mi} spełniają równanie kalibracyjne:

$$\sum_{i \in s_M} w_{Mi} \mathbf{x}_{(g)i} = \sum_{i \in U_M} \mathbf{x}_{(g)i}, \quad (6.18)$$

a wagi w_k równanie kalibracyjne postaci⁸⁶:

$$\sum_{k \in s} w_k \mathbf{x}_{(u)k} = \sum_{k \in U} \mathbf{x}_{(u)k}. \quad (6.19)$$

⁸⁵ Subskrypt (g) odnosi się do jednostek losowania pierwszego stopnia, które stanowią grupy (ang. *groups*).

⁸⁶ Również w tym przypadku można nakładać ograniczenia na odpowiednie ilorazy wag. W dalszej części rozpatrzony zostanie przypadek liniowej funkcji kalibracyjnej bez ograniczeń nakładanych na iloraz wag.

Równanie (6.18) orzeka, że wagi kalibracyjne w_{Mi} na poziomie grupy powinny być w taki sposób ustalone, aby odtwarzały znane wartości globalne zmiennych pomocniczych $\mathbf{x}_{(g)}$. Z kolei wagi w_k przypisane jednostkom powinny odtwarzać znane wartości globalne zmiennych pomocniczych $\mathbf{x}_{(u)}$. Mówi o tym równanie kalibracyjne (6.19). Zakładamy przy tym, że zmiennych pomocniczych na poziomie jednostek jest J_1 , a zmiennych pomocniczych na poziomie grup jest J_2 . Stąd w procesie kalibracji wykorzystywanych jest $J = J_1 + J_2$ zmiennych pomocniczych. Przy przyjętych powyżej oznaczeniach wektor wartości globalnych wszystkich J zmiennych pomocniczych wziętych pod uwagę w procesie kalibracji jest postaci:

$$\boldsymbol{\tau}_{\mathbf{X}}^{(gu)} = \begin{pmatrix} \sum_{i \in U_M} \mathbf{x}_{(g)i} \\ \sum_{k \in U} \mathbf{x}_{(u)k} \end{pmatrix}. \quad (6.20)$$

Kluczową kwestią, którą należy wziąć pod uwagę w procesie kalibracji wag, uwzględniającą dwa rodzaje zmiennych pomocniczych, jest odpowiednie powiązanie wag w_{Mi} i w_k . Stanowi to podstawę tzw. zintegrowanego systemu ważenia (ang. *integrated weighting*). Jeden z takich sposobów opisuje równanie:

$$w_k = d_{k|i} w_{Mi} \quad (6.21)$$

dla wylosowanych k -tych jednostek z grupy $i \in s_M$. Warto przy tym zwrócić uwagę, że dla tak określonego równania (6.21) wynikający z niego iloraz wag $w_k/w_{Mi} = d_{k|i}$ „naśladuje” własność $d_k/d_{Mi} = d_{k|i}$ będącą następstwem konstrukcji oryginalnych wag wynikających z założonego dwustopniowego losowania próby.

W wypadku losowania jednostopniowego, które jest szczególnym przypadkiem losowania dwustopniowego, tj. gdy w badaniu biorą udział wszystkie jednostki k z wylosowanej grupy i , spełniony jest warunek $d_{k|i} = 1$ i wówczas równanie (6.21) upraszcza się do:

$$w_k = w_{Mi}. \quad (6.22)$$

Oznacza to, że wszystkie jednostki z wylosowanej grupy będą miały jednakową wagę równą wadze kalibracyjnej w_{Mi} . Wyznaczenie wag kalibracyjnych w_{Mi} oraz w_k , które będą spełniały równania kalibracyjne (6.18) i (6.19) odpowiednio, przy jednoczesnym spełnieniu warunku (6.21), wymaga zastosowania następującego algorytmu⁸⁷.

⁸⁷ Rozpatrzone przypadek liniowej funkcji kalibracyjnej bez ograniczeń, dla której wagi w_{Mi} , a w konsekwencji wagi w_k , można wyznaczyć wprost ze wzoru. W kolejnym podrozdziale w badaniu empirycznym uzyskano jednak zintegrowane wagi kalibracyjne dla różnych funkcji.

Algorytm 8. Konstrukcja wag w_{Mi} oraz w_k w losowaniu dwustopniowym

Krok 1: *Modyfikacja równań kalibracyjnych*

W równaniu kalibracyjnym (6.19) w miejsce wagi w_k podstaw $d_{k|i}w_{Mi}$ zgodnie ze wzorem (6.21). W ten sposób równanie (6.19) staje się równaniem zmiennej w_{Mi} .

Krok 2: *Konstrukcja wektora zmiennych pomocniczych $\mathbf{x}_{(gu)i}$*

Ustal wektor $\mathbf{x}_{(gu)i}$ postaci:

$$\mathbf{x}_{(gu)i} = \begin{pmatrix} \mathbf{x}_{(g)i} \\ \hat{\mathbf{x}}_{(u)i} \end{pmatrix}, \quad (6.23)$$

gdzie:

$$\hat{\mathbf{x}}_{(u)i} = \sum_{s_i} d_{k|i} \mathbf{x}_{(u)k}. \quad (6.24)$$

Krok 3: *Wyznaczenie wag kalibracyjnych w_{Mi}*

Zaczynając od wag d_{Mi} , wyznacz wagi kalibracyjne w_{Mi} dla $i \in s_M$ tak, aby spełnione było równanie kalibracyjne:

$$\sum_{i \in s_M} w_{Mi} \mathbf{x}_{(gu)i} = \boldsymbol{\tau}_{\mathbf{X}}^{(gu)}, \quad (6.25)$$

gdzie $\boldsymbol{\tau}_{\mathbf{X}}^{(gu)}$ wyraża się wzorem (6.20). Wagi w_{Mi} są postaci:

$$w_{Mi} = d_{Mi} \left(1 + \boldsymbol{\lambda}_{s_M}^T \mathbf{z}_i \right), \quad (6.26)$$

przy czym:

$$\boldsymbol{\lambda}_{s_M}^T = \left(\boldsymbol{\tau}_{\mathbf{X}}^{(gu)} - \sum_{i \in s_M} d_{Mi} \mathbf{x}_{(gu)i} \right)^T \left(\sum_{i \in s_M} d_{Mi} \mathbf{z}_i \mathbf{x}_{(gu)i}^T \right)^{-1}, \quad (6.27)$$

gdzie \mathbf{z}_i to wektor zmiennych instrumentalnych⁸⁸.

Krok 4: *Wyznaczenie wag kalibracyjnych w_k*

Wyznacz wagi kalibracyjne w_k , korzystając ze wzoru (6.21).

Mając wyznaczone wagi kalibracyjne w_{Mi} oraz w_k , można dokonać estymacji wartości globalnej (6.14) i (6.15), korzystając z estymatorów kalibracyjnych war-

⁸⁸ Koncepcja wektorów instrumentalnych została omówiona w rozdziale trzecim.

tości globalnej, które wyrażają się wzorami (6.16) oraz (6.17) odpowiednio. Warto jednocześnie podkreślić cztery kwestie. Po pierwsze w algorytmie 8 nie nakładano żadnych ograniczeń na wagi w_{Mi} , a dokładniej na iloraz wag w_{Mi} i d_{Mi} . Nic jednak nie stoi na przeszkodzie, aby zapewnić przynależność wag kalibracyjnych w_{Mi} do z góry zdefiniowanego przedziału. Wówczas dla odpowiednio dobranej funkcji $G(\cdot)$ oraz parametrów L i U można znaleźć wagi kalibracyjne w_{Mi} , korzystając z algorytmu 2 wyznaczania mnożników λ . Po drugie algorytm 8 można również wykorzystać w wypadku losowania jednostopniowego, w którym na drugim etapie brane są pod uwagę wszystkie jednostki z wylosowanych grup, tj. przy założeniu, że $d_{k|i} = 1$. W tym celu wystarczy w odpowiednich krokach i wzorach tego algorytmu dokonać podstawienia $d_{k|i} = 1$. W rezultacie uzyskamy zintegrowane wagi kalibracyjne spełniające warunek (6.22). Po trzecie ważną sprawą jest również dobór wektora zmiennych instrumentalnych \mathbf{z}_i we wzorze (6.26). W praktyce najczęściej przyjmuje się, że $\mathbf{z}_i = \mathbf{x}_{(gu)i}$. Po czwarte rozważane wcześniej podejście wykorzystujące ideę Lemaître'a i Dufoura (1987), w której zmienne pomocnicze mierzone są jedynie na poziomie osób (wektor \mathbf{z}_{k_1}) bądź na poziomie osób i gospodarstw domowych (wektor \mathbf{z}_{k_2}), są szczególnymi przypadkami idei konstrukcji wag zintegrowanych przedstawionych w algorytmie 8. Na przykład, zakładając, że mamy do czynienia z losowaniem jednostopniowym, w którym losowane są gospodarstwa domowe i badaniu podlegają wszystkie osoby z wylosowanych gospodarstw, oraz przyjmując, że $\mathbf{x}_{(gu)i} = \mathbf{x}_{(g)i}$ ⁸⁹, a także $\mathbf{z}_i = \mathbf{x}_{(g)i}/N_i$, gdzie N_i jest liczbą osób w danym gospodarstwie domowym, otrzymuje się dokładnie te same wagi kalibracyjne spełniające warunek $w_k = w_{Mi}$ co w podejściu wykorzystującym ideę Lemaître'a i Dufoura (1987), w której zmienne pomocnicze mierzone są jedynie na poziomie osób.

Kluczową kwestią jest również wyznaczenie wariancji estymatorów kalibracyjnych (6.16) i (6.17). Estevao i Särndal (2006) proponują, aby jako wariancję estymatora (6.16) wykorzystać:

$$D^2\left(\hat{\tau}_{(u)CAL}\right) = \sum_{i \in U_M} \sum_{j \in U_M} \Delta_{Mij} e_{Mi} e_{Mj} + \sum_{i \in U_M} d_{Mi} V_i, \quad (6.28)$$

gdzie

$$\Delta_{Mij} = \frac{d_{Mi} d_{Mj}}{d_{Mij}} - 1, \quad V_i = \sum_{U_i} \sum_{U_i} \Delta_{kl|i} e_k e_l, \quad \Delta_{kl|i} = \frac{d_{k|i} d_{l|i}}{d_{kl|i}} - 1,$$

⁸⁹ Wektor $\mathbf{x}_{(g)i}$ przechowuje w takim wypadku informacje o liczbie osób w każdym gospodarstwie z podziałem na rozważane charakterystyki (na przykład płeć, grupy wieku itd.).

a reszty e_{Mi} oraz e_k wyrażają się wzorami:

$$e_{Mi} = y_{(u)i} - \mathbf{x}_{(g)i}^T \mathbf{B}_{M1}^{(u)} - \mathbf{x}_{(u)i}^T \mathbf{B}_{M2}^{(u)}, \quad (6.29)$$

$$e_k = y_{(u)k} - \mathbf{x}_{(u)k}^T \mathbf{B}_{M2}^{(u)}, \quad (6.30)$$

przy czym:

$$\begin{pmatrix} \mathbf{B}_{M1}^{(u)} \\ \mathbf{B}_{M2}^{(u)} \end{pmatrix} = \left(\sum_{i \in U_M} \mathbf{z}_i \mathbf{x}_{(gu)i}^T \right)^{-1} \left(\sum_{i \in U_M} \mathbf{z}_i y_{(u)i} \right). \quad (6.31)$$

Z kolei w wypadku wariancji $D^2(\hat{\tau}_{(g)CAL})$ estymatora (6.17) Estevao i Särndal (2006) proponują, aby skorzystać ze wzoru (6.28), przy czym reszty e_{Mi} oraz e_k należy zastąpić przez:

$$e_{Mi} = y_{(g)i} - \mathbf{x}_{(g)i}^T \mathbf{B}_{M1} - \mathbf{x}_{(u)i}^T \mathbf{B}_{M2}, \quad (6.32)$$

$$e_k = \mathbf{x}_{(u)k}^T \mathbf{B}_{M2}, \quad (6.33)$$

gdzie:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{M1} \\ \mathbf{B}_{M2} \end{pmatrix} = \left(\sum_{i \in U_M} \mathbf{z}_i \mathbf{x}_{(gu)i}^T \right)^{-1} \left(\sum_{i \in U_M} \mathbf{z}_i y_{(g)i} \right). \quad (6.34)$$

Warto przy tym podkreślić, że w wypadku schematu losowania jednostopniowego, wzory na wariancję estymatorów (6.16) i (6.17) znacznie się upraszczają, gdyż $V_i = 0$. Jest to konsekwencją tego, że dla tego typu losowania $\Delta_{kl|i} = 0$. Co więcej, wymagane jest znalezienie estymatora $\hat{\mathbf{B}}$ parametru \mathbf{B} określonego wzorem (6.34) tak, aby możliwe było znalezienie estymatorów wariancji $\hat{D}^2(\hat{\tau}_{(u)CAL})$ oraz $\hat{D}^2(\hat{\tau}_{(g)CAL})$. Opis procedury konstrukcji obydwu estymatorów wariancji w wypadku losowania dwustopniowego podają Särndal i współautorzy (1992).

Opisana w niniejszym podrozdziale procedura wyznaczania zintegrowanych wag kalibracyjnych jest wykorzystywana w praktyce badań statystycznych zarówno przez krajowe urzędy statystyczne, jak i ośrodki badawczo-naukowe. Na przykład Główny Urząd Statystyczny wykorzystuje metodę kalibracji zintegrowanej w Europejskim Badaniu Warunków Życia Ludności (GUS, 2017b). Ma ona również zastosowanie w badaniu „Diagnoza społeczna” (Panek, Czapiński i Kottowska, 2011) czy realizowanym przez Departament Analiz Ekonomicznych i Departament Stabilności Finansowej Narodowego Banku Polskiego badaniu „Zasobność gospodarstw domowych w Polsce” (NBP, 2017).

6.5. Propozycja kalibracji zintegrowanej w NSP 2011

W niniejszym podrozdziale przedstawiono propozycję konstrukcji wag zintegrowanych i analizę ich własności w spisach prowadzonych metodą mieszaną na przykładzie NSP 2011⁹⁰. Opisano również proces estymacji liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy w przekroju powiatów, płci oraz miejsca zamieszkania z wykorzystaniem estymatorów kalibracyjnych wartości globalnej dla różnych zestawów wag⁹¹. Tego typu oszacowania uzyskane z wykorzystaniem wag kalibracyjnych w_k , których sposób konstruowania omówiono szczegółowo w podrozdziale 6.3, zostały zamieszczone przez Główny Urząd Statystyczny w najważniejszych tablicach spisowych⁹². Rozważania oraz zaprezentowane wyniki ograniczono tutaj do województwa wielkopolskiego.

Zgodnie z opisem kalibracji wag, przedstawionym w podrozdziale 6.3 w algorytmie 7, w NSP 2011 nie zastosowano podejścia zintegrowanego. Oznaczało to, że osoby w ramach danego gospodarstwa domowego na ogół miały przypisane różne końcowe wagi kalibracyjne. Jednocześnie inny system wag był wykorzystywany w kontekście szacowania wartości globalnych dla gospodarstw domowych, a inny dla osób. Utworzenie wag zintegrowanych umożliwiłoby utworzenie jednego systemu wag kalibracyjnych, który mógłby mieć zastosowanie w kontekście estymacji parametrów na poziomie gospodarstw domowych i osób.

Konstrukcję wag zintegrowanych przeprowadzono zgodnie z algorytmem 8. Ze względu na to, że w NSP 2011 zastosowano jednostopniowy schemat losowania próby, wszystkie osoby z wylosowanych mieszkań z danej warstwy otrzymały tę samą wagę pierwotną. Oznaczało to jednocześnie, że waga ta była przypisana wszystkim gospodarstwom domowym zamieszkującym wylosowane mieszkania. W algorytmie 8 przyjęto zatem założenie, że $d_{k|i} = 1$, tj. rozpatrywano jego szczególny przypadek. Konstrukcji wag zintegrowanych dokonywano po wcześniejszym skorygowaniu wag pierwotnych ze względu na wspomniane już jednostkowe braki danych. Korekta polegała na podzieleniu wag pierwotnych przez odpowiedni

⁹⁰ Uwaga odnosić się będzie ponownie do ludności faktycznie zamieszkałej.

⁹¹ W rozdziale tym uwagę skupiono na konstrukcji wag zintegrowanych. Oznacza to, że ten sam system wag kalibracyjnych można wykorzystać zarówno dla osób, jak i gospodarstw domowych. Zastosowanie wag w procesie estymacji, wraz z dyskusją na temat wskaźników precyzji, ograniczono w celu egzemplifikacji do konkretnej tablicy spisowej odnoszącej się do aktywności zawodowej osób.

⁹² Odpowiednią tablicę z oszacowaniami i wskaźnikiem precyzji można utworzyć, korzystając z informacji zawartych w Banku Danych Lokalnych w części poświęconej Narodowym Spisom Powszechnym.

wskaźnik odpowiedzi obliczany w ramach warstwy. Mówiąc precyzyjniej, kalibracji zintegrowanej podlegały wagi v_k , o których mowa w kroku 2 algorytmu 7. Wynikało to z tego, że wagi v_k nie odtwarzały znanych struktur demograficznych ze spisu pełnego bazującego na rejestrach i systemach administracyjnych.

W procesie konstruowania wag kalibracyjnych w_k przypisanych osobom wzięto pod uwagę pięć podejść⁹³:

- **podejście 0** – określane mianem bazowego. W tym podejściu wagi kalibracyjne były skonstruowane dokładnie w taki sam sposób jak w NSP 2011 (por. opis w podrozdziale 6.3). W procesie kalibracji wag uwzględniana była zatem informacja o płci (kobiety, mężczyźni), miejscu zamieszkania (obszar miejski i wiejski) i roczniku wieku (0, 1, ..., 84, 85+) na poziomie województwa oraz płci, miejscu zamieszkania i grup wieku dla osób na poziomie powiatów. W wypadku Poznania wzięto dodatkowo jako zmienną pomocniczą pojedyncze roczniki wieku, tak jak dla całego województwa. W tym podejściu nie brano pod uwagę zmiennych pomocniczych na poziomie gospodarstwa domowego. W wyniku przeprowadzonego procesu kalibracji wagi przypisane osobom różniły się na ogół w ramach danego gospodarstwa domowego (nie rozpatrywano podejścia zintegrowanego).
- **podejście 1** – w tym podejściu wzięto pod uwagę informację dla osób na poziomie całego województwa w odniesieniu do płci, miejsca zamieszkania oraz roczników wieku. Nie brano jako zmiennych pomocniczych płci, miejsca zamieszkania oraz grup wieku na poziomie powiatów. Ponadto na poziomie gospodarstw domowych uwzględniono typ obszaru zamieszkania (miejski i wiejski) na poziomie całego województwa wielkopolskiego. Użytkowano w ten sposób zintegrowane wagi kalibracyjne, tj. wszyscy członkowie gospodarstw domowych otrzymali te same wagi. Wagi te jednocześnie odtwarzały znaną liczbę gospodarstw domowych na obszarach miejskich i wiejskich oraz liczbę osób w województwie wielkopolskim w przekroju wyznaczonym przez płeć, miejsce zamieszkania oraz pojedyncze roczniki wieku. Wagi te jednak nie odtwarzały liczby osób w przekroju płci, miejsca zamieszkania i grup wieku w poszczególnych powiatach.
- **podejście 2** – w tym podejściu wzięto pod uwagę informację dla osób na poziomie całego województwa w odniesieniu do płci, miejsca zamieszkania oraz roczników wieku. Na poziomie powiatów uwzględniono ponadto informację o płci i grupach wieku, a nie wzięto pod uwagę miejsca zamieszkania.

⁹³ Przyjęto numerację 0, ..., 4, aby podkreślić, że podejściem bazowym jest to, które zostało zastosowane w NSP 2011.

Na poziomie gospodarstw domowych uwzględniono typ obszaru zamieszkania (miejski i wiejski) dla całego województwa wielkopolskiego. W tym podejściu także uzyskano zintegrowane wagi kalibracyjne. Wagi te odtwarzały znaną liczbę gospodarstw domowych na obszarach miejskich i wiejskich oraz liczbę osób w województwie wielkopolskim w przekroju wyznaczonym przez płeć, miejsce zamieszkania oraz pojedyncze roczniki wieku. Wagi te odtwarzały także liczbę osób w przekroju płci i grup wieku w poszczególnych powiatach.

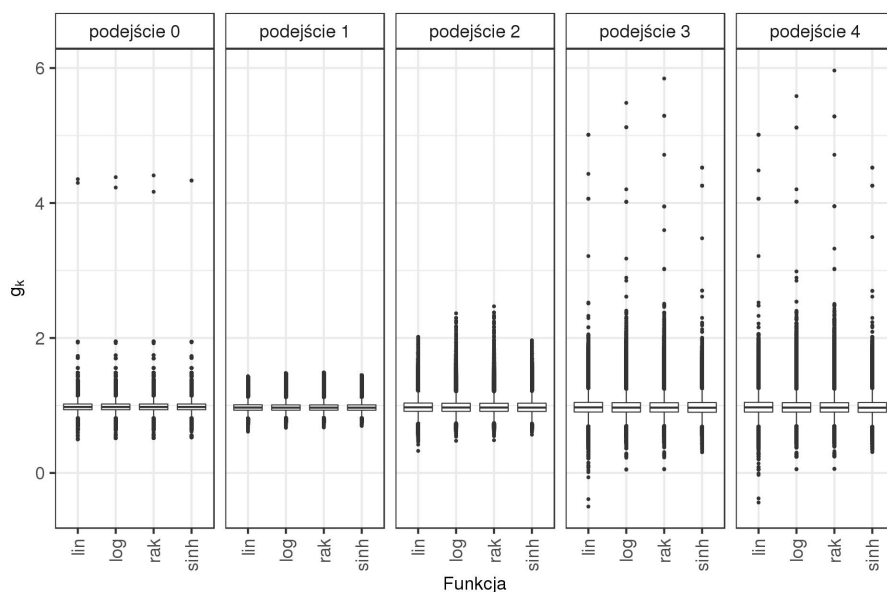
- **podejście 3** – tutaj wzięto pod uwagę taki sam układ zmiennych pomocniczych jak w podejściu 2. Dodatkowo dla osób na poziomie powiatów uwzględniono ich miejsce zamieszkania. Zintegrowane wagi kalibracyjne odtwarzały zatem dodatkowo znaną liczbę osób w przekroju płci, miejsca zamieszkania i grup wieku dla poszczególnych powiatów.
- **podejście 4** – rozpatrzony został taki sam układ zmiennych pomocniczych jak w podejściu 3. Dodatkowo dla osób zamieszkujących Poznań wzięto pod uwagę zmienną pomocniczą w postaci pojedynczych roczników wieku. W odróżnieniu od podejścia bazowego, uzyskano zintegrowane wagi kalibracyjne, które nie tylko odtwarzały struktury spisowe na poziomie osób, jak w NSP 2011, ale dodatkowo znane liczebności gospodarstw domowych na obszarach miejskich i wiejskich w województwie wielkopolskim.

Uwzględnienie kilku odmiennych podejść miało na celu sprawdzenie, w jakim stopniu wzięcie w procesie kalibracji różnych przekrojów i zmiennych pomocniczych mierzonych na poziomie osób i gospodarstw domowych wpływało na postać wag, uzyskiwane szacunki i ich precyzję. Z wyjątkiem podejścia bazowego, które było zastosowane w NSP 2011, w pozostałych otrzymano zintegrowane wagi kalibracyjne. W procesie wyznaczania wag kalibracyjnych wzięto pod uwagę cztery metody, które zostały syntetycznie przedstawione w tabeli 2.1: liniową (lin), raking (rak), logitową (log) oraz w wersji sinusa hiperbolicznego (sinh). Nie rozpatrzono jedynie liniowej funkcji z ograniczeniami. Dla funkcji liniowej wagi zintegrowane (podejścia 1–4) zostały wyznaczone ze wzoru (6.26). Dla pozostałych trzech metod, ze względu na to, że wagi nie można wyrazić wprost ze wzoru, uwzględniono dodatkowo algorytm 1 (dla metody raking) oraz algorytm 2 (dla metody logitowej i w wersji sinusa hiperbolicznego). Wektor zmiennych instrumentalnych ustalono jako $\mathbf{z}_i = \mathbf{x}_{(gu)i}$.

Na rysunkach 6.5 oraz 6.6 przedstawiono rozkłady mnożników kalibracyjnych $g_k = w_k/v_k$ oraz wag kalibracyjnych w_k przypisanych osobom w województwie wielkopolskim w zależności od podejścia i metody wyznaczania wag. W pierwszej kolejności należy podkreślić, że podejście bazowe (0) otrzymujemy jedynie

dla metody liniowej. Tę technikę kalibracji wag wykorzystano bowiem w NSP 2011 ze względu na brak wag ujemnych i odtwarzanie wszystkich wartości globalnych w założonych przekrojach. Jest to zgodne z opisem poczynionym w podrozdziale 6.3. Wagi kalibracyjne w_k uzyskane w podejściu bazowym nie są jednak zintegrowane. Oznacza to, że w ramach danego gospodarstwa domowego jego członkowie mieli przypisane różne wagi. W procesie estymacji należało zatem korzystać z innego zestawu wag dla osób i gospodarstw domowych.

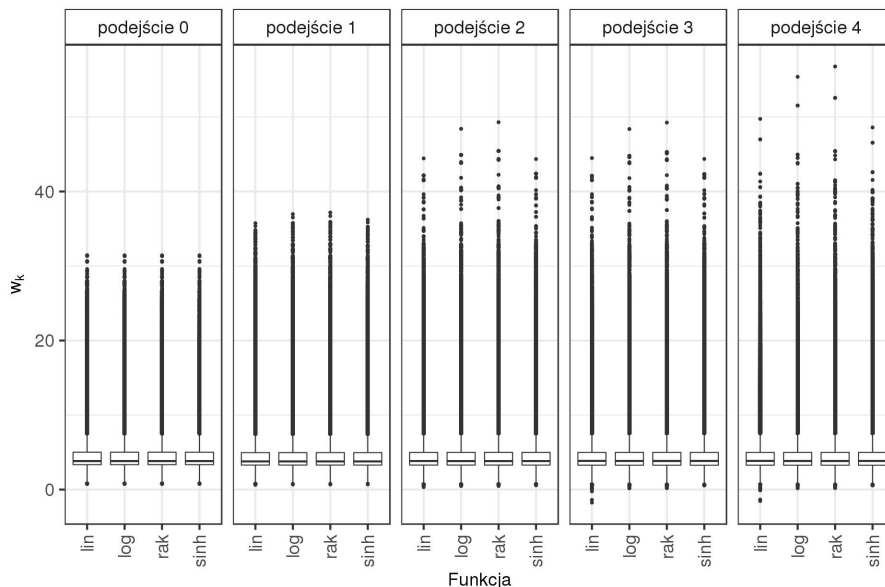
Począwszy od podejścia 1 mamy do czynienia ze zintegrowanym systemem wag kalibracyjnych. W najmniejszym stopniu wagi kalibracyjne są zróżnicowane w podejściu 1. W podejściu tym nie brano bowiem pod uwagę informacji na temat zmiennych pomocniczych na poziomie osób odnoszących się do płci, miejsca zamieszkania, grup wieku i powiatu.



Rysunek 6.5. Rozkład mnożników kalibracyjnych g_k w zależności od podejścia i wykorzystanej funkcji

Źródło: na podstawie NSP 2011.

Ponieważ w NSP 2011 przyjęto założenie, że estymacja parametrów dla osób będzie się odbywać co najmniej na poziomie powiatów i dodatkowo wyodrębnionych przekrojów, w podejściach 2–4 uwzględniono w różnym zakresie informacje o zmiennych pomocniczych w postaci płci, miejsca zamieszkania i grupy wieku w tym układzie regionalnym. Najbardziej zróżnicowane wagi kalibracyjne



Rysunek 6.6. Rozkład wag kalibracyjnych w_k w zależności od podejścia i wykorzystanej funkcji

Źródło: na podstawie NSP 2011.

otrzymano w podejściach 3 i 4, przede wszystkim dla metody wykładniczej (raking) oraz liniowej. Należy ponadto zwrócić uwagę, że dla liniowej funkcji kalibracyjnej i podejścia 3 oraz 4 uzyskuje się ujemne wagi kalibracyjne. Stoi to w sprzeczności z definicją wagi. W związku z tym wykorzystanie liniowej funkcji kalibracyjnej w podejściu zintegrowanym 3 oraz 4 nie byłoby możliwe. Mogłoby się bowiem zdarzyć, że wystąpiłyby ujemne oszacowania wartości globalnej pewnej zmiennej w szczegółowo zdefiniowanych domenach. Warto jednak podkreślić, że dla pozostałych funkcji uzyskujemy dodatnie zintegrowane wagi kalibracyjne. W najmniejszym stopniu są one przy tym zróżnicowane dla funkcji kalibracyjnej w wersji sinusa hiperbolicznego.

Analizując rysunki 6.5 oraz 6.6, można zatem dokonać podsumowania, że zwiększenie liczby zmiennych pomocniczych powoduje większe zróżnicowanie mnożników i wag kalibracyjnych. Wynika to z tego, że w takim wypadku musi być spełniona większa liczba równań kalibracyjnych.

Dla funkcji kalibracyjnych z ograniczeniami L i U można jednak zapewnić, że wartości wag kalibracyjnych będą się znajdować w odpowiednio dobranym przedziale. Z punktu widzenia jakości wag zintegrowanych najlepszymi własnościami dla wszystkich rozważanych podejść charakteryzowała się funkcja kalibracyjna

w postaci sinusa hiperbolicznego. Wynikało to przede wszystkim z tego, że używane dla niej wagi kalibracyjne w_k były najbliżej skupione wokół wag v_k , tj. wag pierwotnych skorygowanych o jednostkowe braki danych. Potwierdzeniem tego są najważniejsze statystyki opisowe mnożników kalibracyjnych g_k zawarte w tabeli 6.2. Bez względu na rozpatrywane podejście, rozstęp mnożników g_k jest każdorazowo najmniejszy dla funkcji kalibracyjnej w postaci sinusa hiperbolicznego.

Z tabeli 6.2 wynika jeszcze jedna ciekawa własność. Mianowicie, mediana mnożników kalibracyjnych g_k dla wszystkich rozpatrywanych podejść jest mniejsza od jedności. Oznacza to, że po kalibracji większość wag w_k musiała zmniejszyć swoje wartości w stosunku do odpowiednich wag v_k . Było to konsekwencją tego, że suma wag v_k w większości analizowanych w danym podejściu przekrojów była wyższa od rzeczywistej liczby ludności faktycznie zamieszkałej w tym

Tabela 6.2. Statystyki opisowe mnożników kalibracyjnych g_k dla rozważanych podejść i funkcji kalibracyjnych

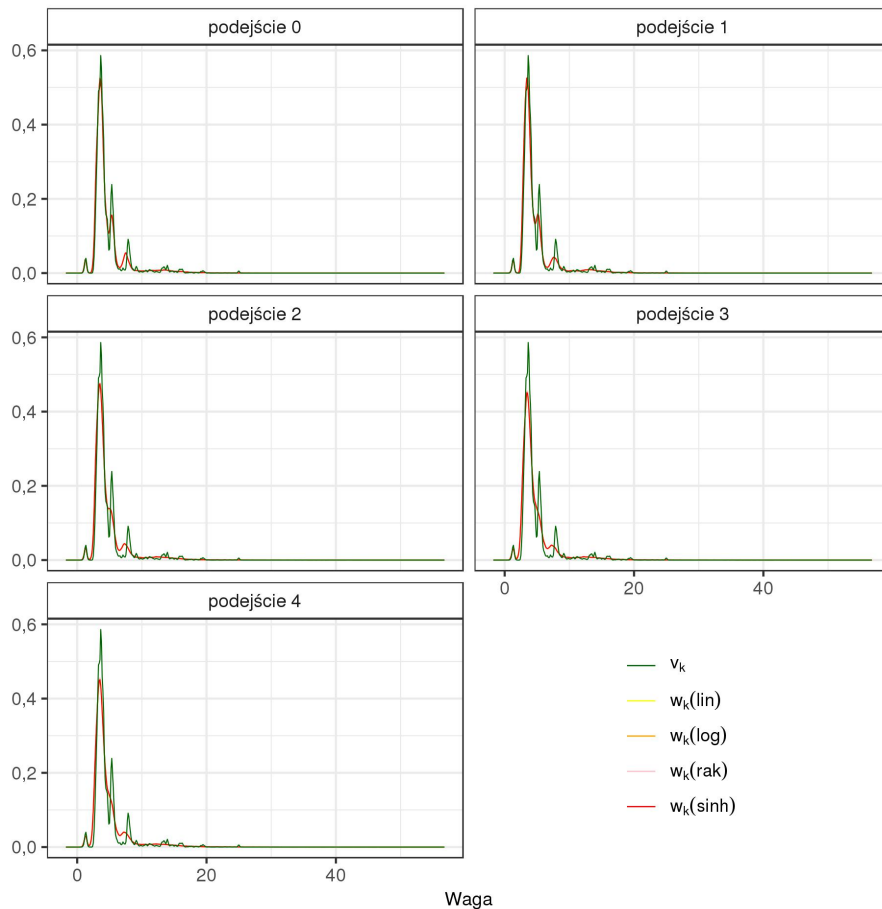
Funkcja	Min	Q ₁	Q ₂	Q ₃	Max	Średnia	Std
Podejście 0							
lin	0,50	0,94	0,98	1,02	4,35	0,98	0,07
log	0,52	0,94	0,98	1,02	4,38	0,98	0,07
rak	0,52	0,94	0,98	1,02	4,41	0,98	0,07
sinh	0,53	0,94	0,98	1,02	4,34	0,98	0,07
Podejście 1							
lin	0,61	0,93	0,97	1,01	1,43	0,97	0,07
log	0,67	0,93	0,97	1,01	1,48	0,97	0,07
rak	0,67	0,93	0,97	1,01	1,49	0,97	0,07
sinh	0,70	0,93	0,97	1,01	1,45	0,97	0,07
Podejście 2							
lin	0,33	0,91	0,97	1,04	2,02	0,98	0,10
log	0,48	0,91	0,97	1,03	2,37	0,98	0,10
rak	0,48	0,91	0,97	1,03	2,47	0,98	0,10
sinh	0,57	0,91	0,97	1,03	1,97	0,98	0,10
Podejście 3							
lin	-0,50	0,90	0,97	1,05	5,01	0,98	0,12
log	0,05	0,90	0,97	1,04	5,48	0,98	0,12
rak	0,05	0,90	0,97	1,04	5,84	0,98	0,12
sinh	0,31	0,90	0,97	1,04	4,52	0,98	0,12
Podejście 4							
lin	-0,44	0,90	0,97	1,05	5,01	0,98	0,12
log	0,05	0,90	0,97	1,04	5,58	0,98	0,12
rak	0,06	0,90	0,97	1,04	5,96	0,98	0,12
sinh	0,31	0,90	0,97	1,04	4,52	0,98	0,12

Źródło: na podstawie NSP 2011.

przekroju, a wynikającej ze spisu pełnego bazującego na rejestrach i systemach administracyjnych. Oznacza to zatem, że w wyniku kalibracji wagi musiały zmaleć, aby ich suma była równa znanym wartościom globalnym z badania pełnego.

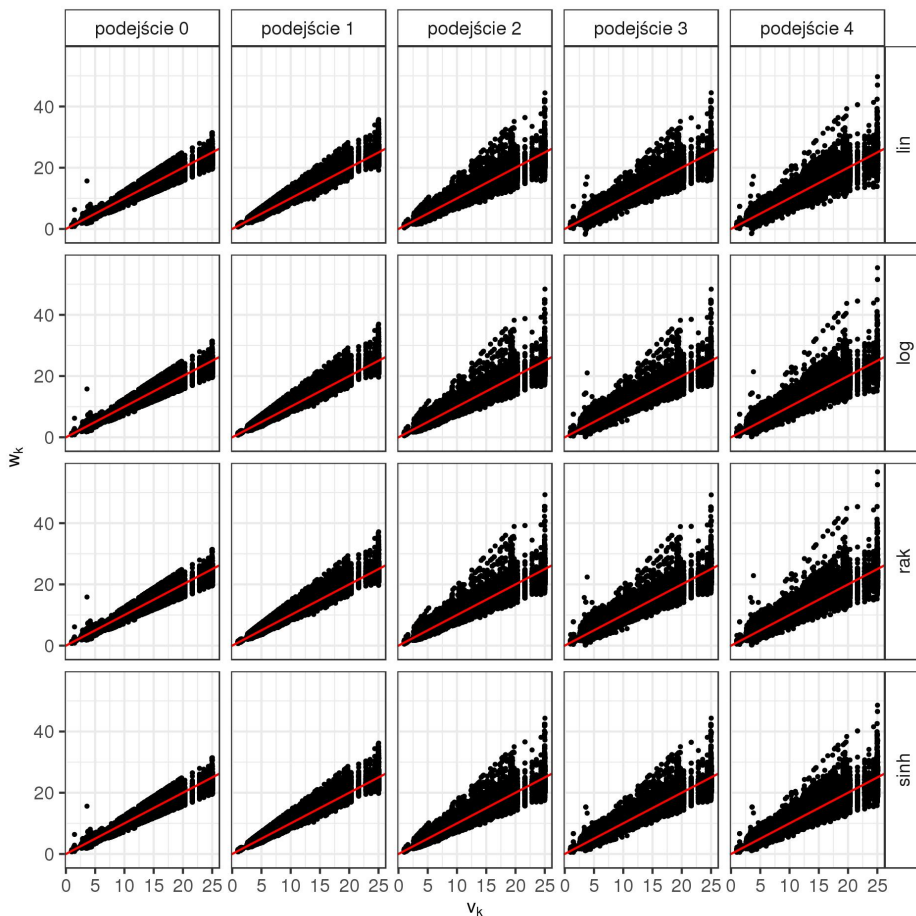
Potwierdzeniem tego są rozkłady wag v_k oraz kalibracyjnych w_k przedstawionych na rysunku 6.7 dla wszystkich analizowanych podejść i funkcji kalibracyjnych. Odpowiednie rozkłady wag kalibracyjnych „są mniej strome”.

Warto również zwrócić uwagę, że wagi v_k i w_k są dodatnio skorelowane w każdym z rozpatrywanych podejść i dla wszystkich uwzględnionych funkcji kalibracyjnych. Rysunek 6.8 jednocześnie potwierdza zauważone wcześniej prawidłowo-



Rysunek 6.7. Rozkłady wag v_k i kalibracyjnych w_k w zależności od podejścia i metody ich wyznaczania

Źródło: na podstawie NSP 2011.



Rysunek 6.8. Korelacyjny diagram rozrzutu wag v_k i kalibracyjnych w_k w zależności od podejścia i wykorzystanej funkcji

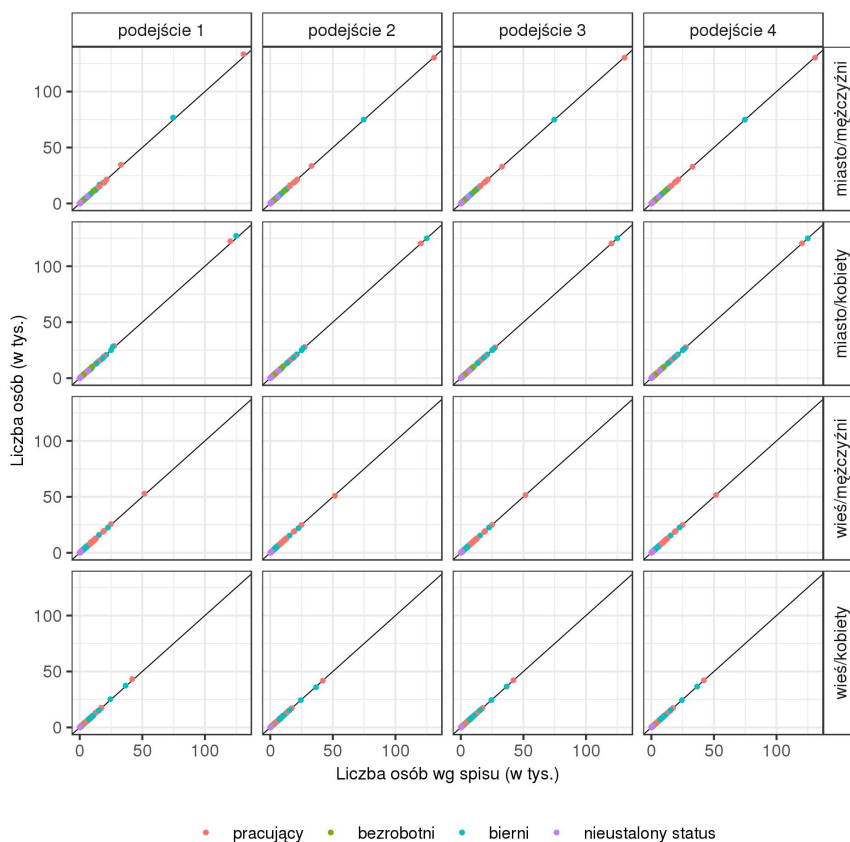
Źródło: na podstawie NSP 2011.

ści. Uwzględnienie w procesie kalibracji warunku zintegrowania wag zwiększa ich zmienność. Wzięcie pod uwagę większej liczby zmiennych pomocniczych dodatkowo nasila to zjawisko. Jest to szczególnie zauważalne w podejściu 3 i 4. Stanowi to swego rodzaju cenę, jaką należy zapłacić za możliwość skonstruowania jednego systemu wag kalibracyjnych, wspólnego dla osób i gospodarstw domowych.

Uzyskane w każdym podejściu wagi wykorzystano w dalszym kroku do oszacowania liczby osób pracujących, bezrobotnych, biernych zawodowo i o niestabilnym statusie na rynku pracy w przekroju powiatów województwa wielkopolskiego, płci oraz miejsca zamieszkania.

W tym miejscu warto podkreślić, że Główny Urząd Statystyczny nie dokonywał imputacji dla pozycyjnych braków danych. W kroku 2 algorytmu 7 wprowadzona korekta na wagi d_k wynikające z przyjętego planu losowania próby miała zniwelować ujemny wpływ jednostkowych braków danych. Pozycyjne braki danych mogły się jednak pojawić również wówczas, gdy osoba wzięła udział w badaniu, ale nie udzieliła odpowiedzi na wszystkie pytania. Z tego powodu w tablicach spisowych, m.in. w kontekście aktywności zawodowej, pojawiała się dodatkowa kategoria związana z nieustalonym statusem na rynku pracy.

Na rysunku 6.9 przedstawiono porównanie oszacowań liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy



Rysunek 6.9. Porównanie oszacowań spisowych liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy z uzyskanymi dla wag zintegrowanych w podejściach 1–4

Źródło: na podstawie NSP 2011.

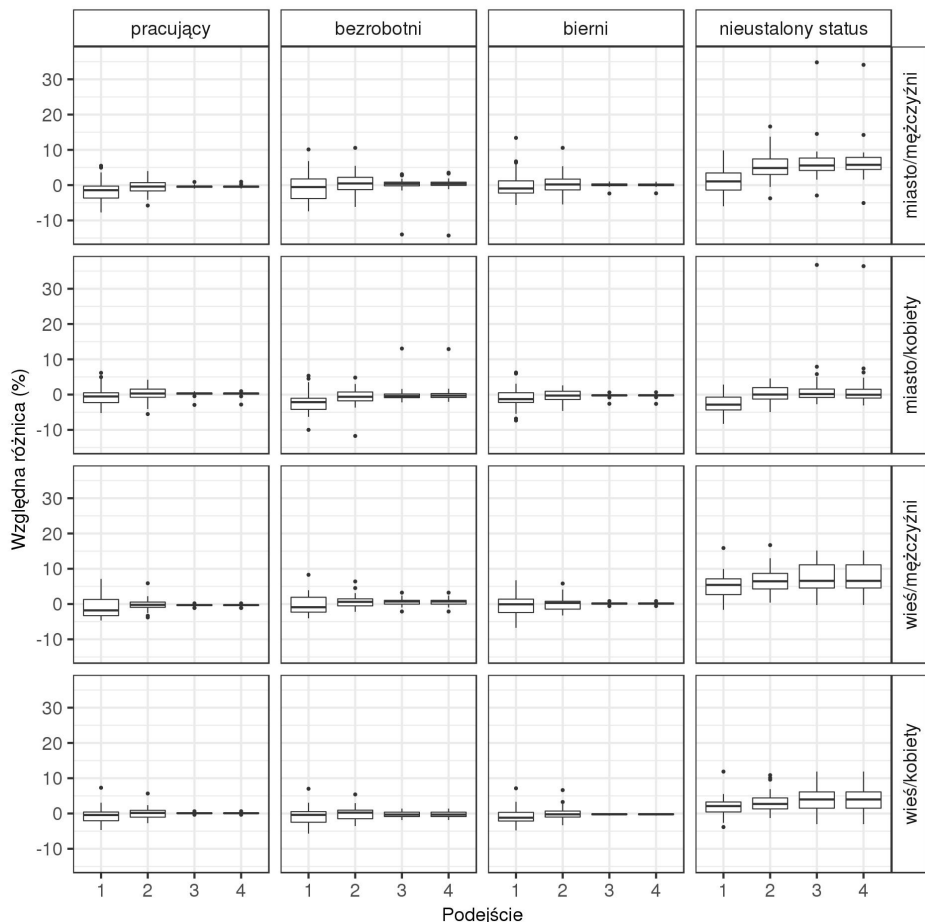
według wag kalibracyjnych, jakie zastosowano w NSP 2011 (metoda liniowa w podejściu bazowym), z oszacowaniami uzyskanymi dla wag zintegrowanych w podejściach 1–4 dla funkcji kalibracyjnej w postaci sinusa hiperbolicznego. Wybór tej funkcji podyktowany był tym, że uzyskiwano dzięki niej zintegrowane wagi kalibracyjne o najlepszych własnościach. Uwzględniono przy tym cztery domeny będące kombinacją wariantów płci i miejsca zamieszkania.

Każdy z punktów na rysunku 6.9, w zależności od użytego koloru, symbolizuje oszacowaną liczbę osób pracujących, bezrobotnych, biernych zawodowo czy o nieustalonym statusie na rynku pracy, oddzielnie dla każdego z 35 powiatów województwa wielkopolskiego, z uwzględnieniem płci i miejsca zamieszkania. Na osi odciętych zaznaczono oszacowania spisowe uzyskane dla wag w podejściu bazowym (metoda liniowa), a na osi rzędnych oszacowania otrzymane z wykorzystaniem wag zintegrowanych dla podejść 1–4 (dla metody w wersji sinusa hiperbolicznego).

Można zauważyć, że oszacowania, jakie uzyskałoby się z wykorzystaniem zintegrowanych wag kalibracyjnych, są silnie i dodatnio skorelowane z oszacowaniami, jakie uzyskano w NSP 2011, bazując na wagach, których sposób konstrukcji został opisany w podrozdziale 6.3. Mimo zauważonych silnych powiązań pomiędzy oszacowaniami liczby osób dla odpowiednich kategorii rynku pracy, widoczne są jednak punkty nieleżące na prostej $y = x$, które wskazują na różnice w oszacowaniach dla odmiennych zestawów wag kalibracyjnych.

Odpowiednie różnice w ujęciu względnym pomiędzy oszacowaniami spisowymi a oszacowaniami, jakie można byłoby uzyskać z wykorzystaniem zestawów wag zintegrowanych otrzymanych w podejściach 1–4, przedstawiono na rysunku 6.10. Można zauważyć, że w wypadku osób pracujących, bezrobotnych i biernych zawodowo największe różnice występują między oszacowaniami spisowymi a uzyskanymi dla wag zintegrowanych w podejściu 1. Jest to konsekwencją tego, że w podejściu 1 w procesie wyznaczania zintegrowanych wag kalibracyjnych nie wzięto pod uwagę zmiennych pomocniczych na poziomie powiatu.

Najmniejsze różnice w stosunku do oszacowań spisowych zaobserwowano dla podejść 3 i 4. W podejściach tych wykorzystywano bowiem zestawy zmiennych pomocniczych podobne do podejścia bazowego dla metody liniowej, które stanowi punkt odniesienia. Uwzględnienie dodatkowych zmiennych pomocniczych mierzonych na poziomie gospodarstw domowych nie tylko umożliwiło uzyskanie zintegrowanych wag kalibracyjnych, ale nie wpłynęło w istotnym stopniu na zmianę oszacowań liczby osób w odpowiednich kategoriach statusu na rynku pracy. Jedyne w odniesieniu do bezrobotnych mężczyzn oraz kobiet zamieszkujących w czę-



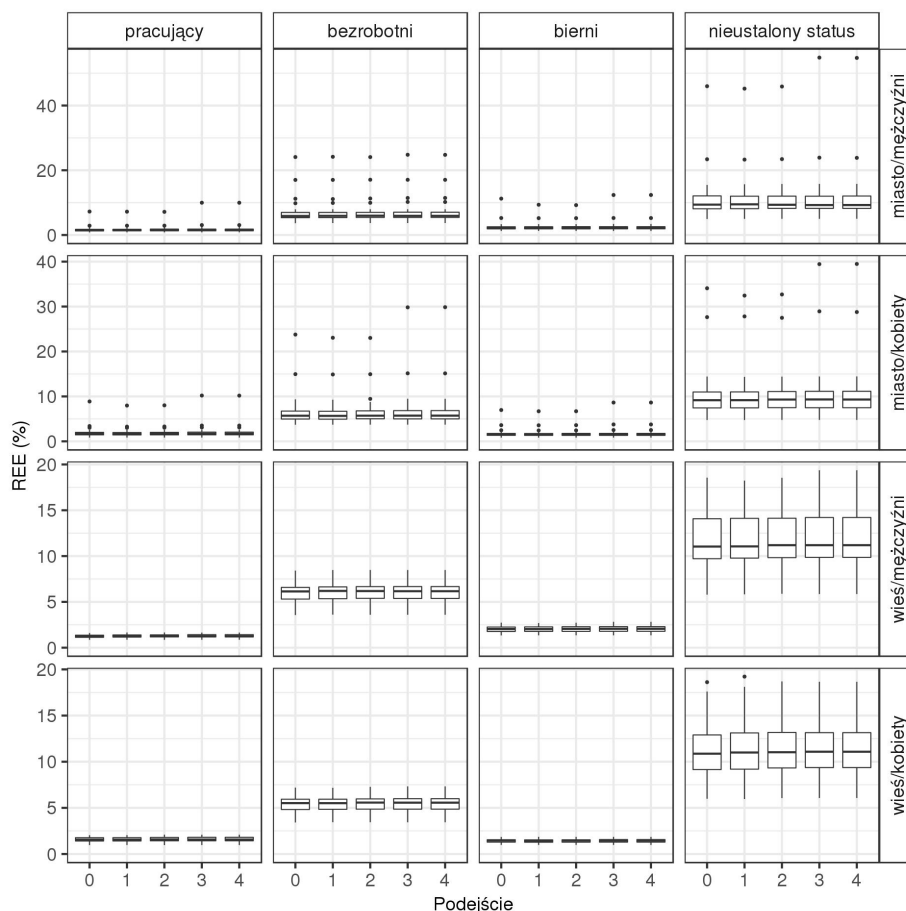
Rysunek 6.10. Względne różnice między oszacowaniami spisowymi liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy a otrzymanymi dla wag z podejść 1–4

Źródło: na podstawie NSP 2011.

ści miejskiej oraz wiejskiej powiatu w jednym przypadku stwierdzono większą różnicę niż w oszacowaniach spisowych (rzędu 15% co do wartości bezwzględnej). Większe względne różnice między oszacowaniami otrzymanymi dla wag zintegrowanych we wszystkich podejściach 1–4 a oszacowaniami spisowymi zauważono natomiast w kontekście osób o nieustalonym statusie na rynku pracy. Na ogół różnice te nie przekraczały jednak progu 10% i były związane z tym, że oszacowana liczba osób o nieustalonym statusie na rynku pracy w odpowiednich przekrojach była na znacznie niższym poziomie w porównaniu z szacunkami dla osób

pracujących, bezrobotnych czy biernych zawodowo. Nawet niewielkie rozbieżności w oszacowaniach uzyskanych w takich przypadkach mogły się przełożyć na większe względne różnice.

Kluczową kwestią, oprócz analizy najważniejszych własności wag kalibracyjnych i oszacowań punktowych, jest również ocena precyzji (rys. 6.11).



Rysunek 6.11. Względne średnie błędy szacunków liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy w przekroju płci i miejsca zamieszkania dla różnych podejść

Źródło: na podstawie NSP 2011.

W pierwszej kolejności, w związku z koniecznością wyznaczenia średnich błędów szacunku oszacowanej liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy w założonych przekrojach,

skorzystano ze wzoru (6.3). Następnie, w celu oceny jakości precyzji uzyskanych oszacowań, wyznaczono względne średnie błędy szacunku REE, zgodnie ze wzorem (1.32)⁹⁴. W wypadku podejścia bazowego uwzględniono metodę liniową. W wypadku pozostałych podejść konstrukcji zintegrowanych wag kalibracyjnych wzięto ponownie pod uwagę funkcję kalibracyjną w postaci sinusa hiperbolicznego.

W ocenie precyzji uzyskanych oszacowań należy zwrócić szczególną uwagę na to, że bez względu na rozważane podejście oraz przekrój (miasto/mężczyźni, miasto/kobiety, wieś/mężczyźni, wieś/kobiety), w odniesieniu do osób pracujących i biernych zawodowo uzyskujemy akceptowalne wyniki, tj. względne średnie błędy szacunku nie przekraczają zakładanego przez Główny Urząd Statystyczny progu 10%. Dotyczy to także w zasadzie kategorii osób bezrobotnych. Z wyjątkiem nielicznych powiatów oraz przekrojów obejmujących płęć, z uwzględnieniem części miejskiej, wyniki można uznać za wiarygodne. W sporadycznych przypadkach uzyskano miernik precyzji REE przekraczający próg 20%. Co ciekawe, akceptowalne błędy szacunku otrzymano również dla podejścia 1, w którym w procesie kalibracji wag nie wzięto pod uwagę zmiennych pomocniczych odnoszących się do płci, miejsca zamieszkania i grup wieku na poziomie powiatów.

Większe względne średnie błędy szacunku otrzymano dla kategorii osób o nieustalonym statusie na rynku pracy. Mediana tych błędów, bez względu na przyjęte podejście, oscylowała w granicach 10%. W wypadku osób zamieszkujących wieś, bez względu na płęć, względne średnie błędy szacunku nie przekraczały jednak progu 20%, dla których Główny Urząd Statystyczny w dalszym ciągu dopuszcza publikowanie wyników. W odniesieniu do mężczyzn i kobiet z miasta błędy te w nielicznych przypadkach przekraczały jednak ten poziom akceptowalności wyników.

6.6. Wnioski

Rozważania zaprezentowane w niniejszym rozdziale dotyczą metod konstruowania zintegrowanych wag kalibracyjnych. To podejście jest szczególnie preferowane w badaniach gospodarstw domowych, w których zachodzi potrzeba utworzenia jednego systemu wag kalibracyjnych. Wagi te można jednocześnie wyko-

⁹⁴ W ocenie precyzji szacunków dla danych z badania reprezentacyjnego NSP 2011 można byłoby również wykorzystać metodę bootstrap. Tego typu podejście zostało wykorzystane przez Młodaka (2014) w badaniu dojazdów do pracy. Ze względu na to, że Główny Urząd Statystyczny w ocenie precyzji oszacowań w NSP 2011 wykorzystał wspomnianą już metodę linearyzacji Taylora, w celu zachowania porównywalności wyników w analizach zostało wykorzystane to podejście.

rzystać w procesie estymacji wybranych charakterystyk odnoszących się zarówno do osób, jak i gospodarstw domowych. Jest to niewątpliwą zaletą tej metody konstruowania wag.

Zagadnienia poddane dyskusji prowadzone były przede wszystkim w kontekście spisów realizowanych metodą mieszaną. Mimo że podejście kalibracyjne można wykorzystać w spisach, w których dane są gromadzone w zasadzie w różny sposób, naturalne wydaje się jego zastosowanie właśnie w spisach mieszanych. Wynika to z tego, że oprócz najważniejszych danych pozyskiwanych z rejestrów i systemów administracyjnych, kluczowe są również informacje pochodzące z badania reprezentacyjnego towarzyszącego spisowi. Rola kalibracji sprowadza się w takim wypadku do takiego ustalenia systemu wag, który zapewnia zgodność najważniejszych struktur spisowych z części reprezentacyjnej i badania pełnego zasilanego rejestrami.

W rozdziale opisano metodę kalibracji zintegrowanej. Wskazano na sposób konstruowania wag zintegrowanych odwołujący się do koncepcji Lemaître'a i Dufoura (1987), polegającej na odpowiednim utworzeniu zmiennych pomocniczych na poziomie osób. Podejście to rozszerzono na przypadek, gdy zmienne pomocnicze dostępne są zarówno na poziomie osób, jak i gospodarstw domowych. Za Estevao i Särndalem (2006) przedstawiono także ideę kalibracji zintegrowanej dla losowania dwustopniowego, które jest wykorzystywane w wielu badaniach gospodarstw domowych. W wypadku gdy na drugim stopniu do badania brane są wszystkie jednostki z wylosowanych grup, podejście to może być zastosowane w losowaniu jednostopniowym. W takiej sytuacji dwa wcześniej rozważane warianty kalibracji zintegrowanej, w których wykorzystuje się zmienne pomocnicze mierzone jedynie na poziomie osób lub na poziomie osób i gospodarstw domowych, stanowią jej szczególne przypadki.

W rozdziale główny nacisk został położony na warstwę aplikacyjną omawianych metod w obszarze spisów mieszanych. Rozważania przeprowadzono w kontekście spisu mieszanego zrealizowanego w Polsce w 2011 roku. W pierwszej kolejności dokładnie opisano sposób konstruowania wag kalibracyjnych dla ludności faktycznie zamieszkałej zastosowany w NSP 2011⁹⁵. Kompleksowej ocenie poddano uzyskane wagi kalibracyjne.

Ze względu na to, że w NSP 2011 nie podjęto próby zbudowania wag zintegrowanych, w rozdziale 6.5 przedstawiono również propozycję ich konstrukcji w tym badaniu dla różnych funkcji odległości. Uwaga została skupiona nie tylko na analizie ich najważniejszych własności, ale także na ocenie procesu estymacji

⁹⁵ Wagi wyznaczono w Ośrodku Statystyki Małych Obszarów Urzędu Statystycznego w Poznaniu, a autor książki przewodniczył pracom grupy metodologów odpowiedzialnych za ich powstanie.

na przykładzie wybranych charakterystyk rynku pracy w szczegółowo zdefiniowanych przekrojach z wykorzystaniem tak utworzonych wag. Uzyskane wyniki wskazują, że możliwa jest konstrukcja wag zintegrowanych, dla których odpowiednie oszacowania można uznać na ogół za akceptowalne.

Otrzymane w rozdziale rezultaty przeprowadzonych prac nabiorą szczególnego znaczenia w NSP 2021, w którym na jeszcze większą skalę planuje się wykorzystanie danych pochodzących z rejestrów i systemów administracyjnych oraz z badania reprezentacyjnego towarzyszącego spisowi. Planuje się bowiem jego mieszany model realizacyjny, który będzie kontynuacją metod i rozwiązań wdrożonych w poprzedniej edycji. W tym kontekście zastosowanie zaawansowanych metod badawczych w postaci kalibracji zintegrowanej czy estymacji pośredniej może stanowić istotny krok w procesie uzyskiwania wiarygodnych szacunków dla niskich poziomów agregacji przestrzennej.

Zakończenie

Praca nawiązuje do dynamicznie rozwijającej się na całym świecie w badaniach reprezentacyjnych oraz spisach metody kalibracji, która od momentu jej formalnego opisu przez Deville'a i Särndala (1992) nieustannie pozostaje w centrum zainteresowania naukowców i krajowych urzędów statystycznych. Przedstawione podejście jest pierwszą w Polsce próbą opisaną estymatorów kalibracyjnych w tak szerokim ujęciu, obejmującym aktualny stan wiedzy zarówno z zakresu teorii, jak i praktycznych zastosowań. W polskiej literaturze przedmiotu brakuje bowiem kompleksowego opracowania na temat kalibracji. Podjęta w książce tematyka stanowi próbę wypełnienia dostrzegalnej luki w tym zakresie i jest pierwszym całościowym spojrzeniem na poruszaną w niej problematykę.

Prezentowana praca ma charakter metodologiczno-poznawczy z równym rozłożeniem akcentów w obszarze teorii estymacji oraz jej praktycznych zastosowań. W warstwie metodologicznej opisany został aktualny stan badań naukowych uwzględniający klasyczną kalibrację, podejście funkcyjne w procesie wyznaczania wag, kalibrację wspomaganą modelem z uwzględnieniem najnowszych rozwiązań dotyczących m.in. adaptacyjnej regresji LASSO, kalibrację hybrydową, kalibrację wag w spisach realizowanych metodą mieszaną, w tym kalibrację zintegrowaną, oraz kalibrację w badaniach z brakami danych. W pracy przedstawiono także własne rozwiązania związane z uogólnieniem podejścia jednokrokowego i dwukrokowego typu A i B prezentowanych przez Särndala i Lundströma (2005) na przykład dowolnie zdefiniowanej funkcji odległości, a także w kontekście estymatorów hybrydowych, w których funkcję modelu wspomagającego pełni regresja grzbie-

towa bądź adaptacyjny model regresji LASSO. W opisie aktualnego stanu wiedzy w ujęciu teoretycznym wartością dodaną jest też szczegółowy przegląd literatury, który może być punktem wyjścia do rozważań dla innych autorów chcących zająć się prezentowaną problematyką.

W warstwie empirycznej, w pierwszej kolejności, korzystając z rzeczywistych danych pochodzących z Badania Budżetów Gospodarstw Domowych, w odpowiednio zaprojektowanym badaniu symulacyjnym dokonano kompleksowej analizy wag i miar jakości estymatorów kalibracyjnych. Uwagę skupiono na różnych klasach estymatorów: otrzymanych w klasycznym podejściu polegającym na minimalizacji odpowiednio dobranej funkcji odległości, w ujęciu funkcyjnym oraz modelowym. W dalszym ciągu opisane zostały własne propozycje zastosowania kalibracji w dwóch kluczowych badaniach prowadzonych przez Główny Urząd Statystyczny: Badaniu Aktywności Ekonomicznej Ludności, w którym jednym z głównych problemów są niskie wskaźniki kompletności i brak dostępnych oszacowań na poziomie podregionów, oraz w spisach prowadzonych metodą mieszaną.

Proponowane rozwiązania mogłyby zostać wykorzystane w innych badaniach reprezentacyjnych realizowanych przez Główny Urząd Statystyczny, takich jak Badanie Budżetów Gospodarstw Domowych czy Europejskie Badanie Warunków Życia Ludności. W badaniach tych, ze względu na przyjęty sposób losowania próby, wyniki są zazwyczaj prezentowane na dość wysokim poziomie agregacji przestrzennej, a braki danych są jednym z głównych źródeł błędów nielosowych. W tym kontekście kalibracja może stanowić zadowalającą odpowiedź na braki odpowiedzi występujące we wszystkich badaniach społeczno-ekonomicznych.

Propozycja rozwiązań może również dotyczyć możliwości zastosowania zintegrowanego podejścia kalibracyjnego w zbliżającym się Narodowym Spisie Powszechnym Ludności i Mieszkań 2021, którego koncepcja zakłada mieszany model realizacyjny. Jak pokazały wyniki przeprowadzonych analiz, utworzenie jednego systemu wag dla gospodarstw domowych i osób umożliwia uzyskanie oszacowań z akceptowalnym poziomem błędu nawet dla dość szczegółowo zdefiniowanych domen. Można zatem przypuszczać, że omówione w pracy podejście kalibracyjne znajdzie w polskich warunkach szersze zastosowanie i umożliwi uzyskanie wiarygodnych informacji na niższych poziomach agregacji przestrzennej i/lub dla bardziej szczegółowo określonych przekrojów.

Należy jednak wyraźnie zaznaczyć, że aby kalibracja spełniła swoją ważną funkcję, niezbędne jest posiadanie wysokiej jakości zmiennych pomocniczych spoza próby. Mówiąc precyzyjnie, chodzi tutaj o dostępność rzetelnych wartości globalnych takich zmiennych, w różnych układach i przekrojach. Posiadanie odpowiedniej infrastruktury statystycznej w tym zakresie, na którą mogą się

składać źródła administracyjne czy statystyczna baza spisowa, są warunkiem *sine qua non* praktycznego stosowania podejścia kalibracyjnego.

Wśród kluczowych zagadnień, które będą wymagać w najbliższej przyszłości uwagi, można wskazać na kilka zasadniczych wątków, które prawdopodobnie wyznaczać będą kierunki dalszych badań w omawianym obszarze:

1) **Konieczność włączenia do badań statystycznych z brakami odpowiedzi podejścia modelowo-kalibrowanego oraz estymacji hybrydowej**

Rozważane w pracy estymatory kalibracyjne wartości globalnej wspomaganie odpowiednio zbudowanym modelem abstrahowały od występującej w rzeczywistości sytuacji związanej z istnieniem braków danych. Wydaje się zasadne przypuszczenie, że uwzględnienie w praktycznych zastosowaniach kalibracji wspomaganiej modelem może poprawić proces estymacji i prowadzić do uzyskania bardziej przekonujących wyników.

2) **Pogłębienie badań nad odpornymi estymatorami kalibracyjnymi**

Istniejąca literatura w tym zakresie jest wyjątkowo uboga i obejmuje w zasadzie jedną pozycję (Duchesne, 1999). Jest to jednak bardzo ważna tematyka, gdyż w wielu badaniach statystycznych (zwłaszcza dotyczących przedsiębiorstw) rozkłady cech charakteryzują się silną asymetrią prawostronną. Istniejące wartości odstające mogą mieć zatem wpływ na uogólniane wyniki, a odpowiednio skalibrowane wagi mogą się przyczynić w takiej sytuacji do poprawy procesu estymacji.

3) **Pogłębienie badań nad kalibracją zintegrowaną w spisach mieszanych**

W pracy zaproponowano sposób wyznaczania wag zintegrowanych, tak aby każda osoba z danego gospodarstwa domowego otrzymała tę samą wagę kalibracyjną. Warto jednak w procesie tworzenia wag uwzględnić przynależność danej osoby nie tylko do gospodarstwa domowego, ale również rodziny. Wynika to z tego, że opracowania spisowe odnoszą się do osób, gospodarstw domowych, rodzin i mieszkań.

4) **Pogłębienie badań nad metodą powtarzanego ważenia**

Wspomniana technika, wykorzystująca ideę kalibracji, jest stosowana w Holandii przez Statistics Netherlands w tzw. spisach wirtualnych, które stanowią pewną odmianę spisów realizowanych metodą mieszaną (Knottnerus i van Duin, 2006; Lumiste, 2011; de Waal, 2015). Celem tej metody jest każdorazowe tworzenie dla z góry zdefiniowanych tabel spisowych odpowiednich wag, aby zapewnić zgodność i spójność wszystkich tabel wynikowych, które powstają z wykorzystaniem różnych źródeł danych zasilających spis. Warto zatem podjąć prace nad wdrożeniem w kolejnym spisie w Polsce tej techniki konstrukcji wag kalibracyjnych.

5) **Rozwój podejścia kalibracyjnego w kontekście zintegrowanych źródeł danych**

Statystyczna integracja danych jest grupą metod umożliwiających łączną obserwację cech nieobserwowanych wspólnie w żadnym z dostępnych źródeł danych. Rezultatem zastosowania tych metod jest utworzenie jednostkowego zbioru danych zawierającego informacje o zmiennych ze wszystkich zintegrowanych źródeł. W konsekwencji możliwa jest analiza zmiennych pochodzących z różnych źródeł informacji. Kluczowymi kwestiami są zatem sposób tworzenia wag dla jednostek znajdujących się w tak utworzonym zbiorze danych oraz zagadnienia związane z estymacją bazującą na takich wagach.

6) **Rozwój podejścia kalibracyjnego dla prób nielosowych**

Fundamentalną sprawą dla możliwości wykorzystania kalibracji w próbach nielosowych jest utworzenie wektora złożonego z wag d_k , który następnie będzie podlegać procesowi korygowania z uwzględnieniem zmiennych pomocniczych. W rozdziale pierwszym zasygnalizowano, że w próbach nielosowych nie mają zastosowania znane z metody reprezentacyjnej schematy losowania, w związku z czym pojawia się zasadniczy problem konstrukcji tego typu wag. Pewne rozwiązania wskazanego problemu zaproponowano w opracowaniu Chena (2016) oraz Chena i współautorów (2018a; 2018b). Dalszych prac badawczych wymagać będą jednak kwestie jakości uzyskanych wyników z prób nielosowych, w sytuacji gdy w procesie uogólniania wykorzystywane będą utworzone wagi kalibracyjne, które stanowią korektę wspomnianych wag d_k .

Dokonując podsumowania prowadzonych w pracy rozważań, można stwierdzić, że kalibracja jako metoda estymacji będzie odgrywać w polskiej praktyce badań statystycznych coraz większą rolę. Wynika to z kilku zasadniczych powodów. Po pierwsze wiele badań prowadzonych przez Główny Urząd Statystyczny jest obarczonych błędami nielosowymi w postaci braków odpowiedzi. Po drugie w systemie statystyki publicznej w Polsce coraz większą wagę zaczyna się przywiązywać do pozyskiwania informacji z pozastatystycznych źródeł danych (internetu czy zasobów big data), które można postrzegać w kontekście prób nielosowych. Wreszcie rosnący popyt na informacje na niskich poziomach agregacji przestrzennej oraz dla bardziej szczegółowo zdefiniowanych domen będzie wymuszać na instytucjach przeprowadzających badanie stosowanie bardziej wyrafinowanych metod. Kalibracja jako technika estymacji w ujęciu zaprezentowanym w książce ma szansę sprostać temu zapotrzebowaniu.

Bibliografia

- AAPOR. (2016). Standard definitions. Final dispositions of case codes and outcome rates for surveys. The American Association for Public Opinion Research, 1–80.
- Alexander, C. H. (1987). A class of methods for using person controls in household weighting. *Survey Methodology*, 13(1), 183–198.
- Alfons, A. i Templ, M. (2013). Estimation of social exclusion indicators from complex surveys: The R package laeken. *Journal of Statistical Software*, 54(15), 1–25. <https://doi.org/10.2139/ssrn.2244876>
- Anastasiade, M. C. i Tillé, Y. (2017). Decomposition of gender wage inequalities through calibration: Application to the Swiss Structure of Earnings Survey. *Survey Methodology*, 43(2), 211–234.
- Andersson, P. G. i Särndal, C. E. (2016). Calibration for nonresponse treatment: In one or two steps? *Statistical Journal of the IAOS*, 32(3), 375–381. <https://doi.org/10.3233/sji-160996>
- Argüeso, A. i Vega, J. L. (2014). A population census based on registers and a “10% survey”. Methodological challenges and conclusions. *Statistical Journal of the IAOS*, 30(1), 35–39. doi: 10.3233/SJI-140797
- Balicki, A. (2004). Metody imputacji brakujących danych w badaniach statystycznych. *Wiadomości Statystyczne*, (9), 1–19.
- Bardsley, P. i Chambers, R. L. (1984). Multipurpose estimation from unbalanced samples. *Applied Statistics*, 33(3), 290–299. <https://doi.org/10.2307/2347706>
- Barranco-Chamorro, I., Jiménez-Gamero, M. D., Mayor-Gallego, J. A. i Moreno-Rebollo, J. L. (2015). A case-deletion diagnostic for penalized calibration estimators and BLUP under linear mixed models in survey sampling. *Computational Statistics & Data Analysis*, 87, 18–33. <https://doi.org/10.1016/j.csda.2015.01.004>
- Bates, D. i Maechler, M. (2018). Matrix: Sparse and dense matrix classes and methods. *R package version 1.2–14*.
- Beaumont, J. F. (2005). On the use of data collection process information for the treatment of unit nonresponse through weight adjustment. *Survey Methodology*, 31(2), 227–231.
- Beaumont, J. F. (2008). A new approach to weighting and inference in sample surveys. *Biometrika*, 95(3), 539–553. <https://doi.org/10.1093/biomet/asn028>

- Beaumont, J. F. i Alavi, A. (2004). Robust generalized regression estimation. *Survey Methodology*, 30(2), 195–208.
- Bechtold, S. (2013). The new register-based census of Germany – A multiple source mixed mode approach. Presentation on the 59th World Statistics Congress (WSC), Hong Kong.
- Beręsewicz, M. (2016). *Internet data sources for real estate market statistics* (Unpublished doctoral dissertation), Poznań: Uniwersytet Ekonomiczny w Poznaniu.
- Beręsewicz, M. i Gołata, E. (2018). *Przyszłość badań statystycznych wobec potrzeb społecznych i gospodarczych na przykładzie spisu ludności*. Pobrane z http://www.pte.pl/pliki/2/1/E_Golata_M_Beresewicz_referat_14_03_2018.pdf.
- Beręsewicz, M. i Szymkowiak, M. (2011). Kilka uwag na temat metod ważenia w badaniach marketingowych z brakami odpowiedzi. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Poznaniu*, (204), 116–124.
- Beręsewicz, M. i Szymkowiak, M. (2015). Big data w statystyce publicznej – nadzieje, osiągnięcia, wyzwania i zagrożenia. *Ekonometria*, 2(48), 9–22.
- Beręsewicz, M. i Szymkowiak, M. (2018). Metody redukcji obciążenia estymatorów w badaniach wykorzystujących big data. II Kongres Statystyki Polskiej, Warszawa.
- Bethlehem, J. (2009). The rise of survey sampling (Discussion Paper No. 09015), The Hague/Heerlen, Statistics Netherlands, 1–28.
- Bethlehem, J., Cobben, F. i Schouten, B. (2011). *Handbook in nonresponse in household surveys*. New York: John Wiley & Sons.
- Binder, D. A. i Théberge, A. (1988). Estimating the variance of raking-ratio estimators. *The Canadian Journal of Statistics*, 1651, 47–55. <https://doi.org/10.2307/3315215>
- Boonstra, H. J. (2004). A simulation study of repeated weighting estimation (Discussion Paper No. 04033), Statistics Netherlands, Voorburg/Heerlen.
- Bracha, C. (1996). *Teoretyczne podstawy metody reprezentacyjnej*. Warszawa: Wydawnictwo Naukowe PWN.
- Bracha, C. (1998). *Metoda reprezentacyjna w badaniu opinii publicznej i marketingu*. Warszawa: Efekt.
- Brackstone, G. J. i Rao, J. N. K. (1979). An investigation of raking ratio estimators. *Sankhya*, 41, 97–114.
- Brakel, J. van den i Bethlehem, J. (2008). Model-based estimation for official statistics (Discussion Paper No. 08002), Voorburg/Heerlen, Statistics Netherlands.
- Brehm, J. (1993). *The phantom respondents: Opinion surveys and political representation*. Ann Arbor: University of Michigan Press.
- Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A critical review. *Journal of Official Statistics*, 29(3), 329–353. <https://doi.org/10.2478/jos-2013-0026>
- Brown, J. J. i Beaujouan, E. (2013). Review of the rolling census approach: And other survey-based options (ESRC Centre for Population Change Working Paper No. 32).
- Brown, M. i Calderwood, L. (2014). Can encouraging respondents to contact interviewers to make appointments reduce fieldwork effort? Evidence from a randomized experiment in the UK. *Journal of Survey Statistics and Methodology*, 2(4), 484–498. <https://doi.org/10.1093/jssam/smu017>
- Buuren, S. van (2018). *Flexible imputation of missing data*. Chapman and Hall/CRC.

- Campanelli, P. (1997). Testing survey questions: New directions in cognitive interviewing. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 55(1), 5–17. <https://doi.org/10.1177/075910639705500103>
- Cardot, H., Goga, C. i Shehzad, M. A. (2017). Calibration and partial calibration on principal components when the number of auxiliary variables is large. *Statistica Sinica*, 243–260.
- Cassel, C. M., Särndal, C. E. i Wretman, J. H. (1977). *Foundations of inference in survey sampling*. New York-London-Sydney-Toronto: John Wiley & Sons.
- Chambers, R. L. i Dunstan, R. (1986). Estimating distribution functions from survey data. *Biometrika*, 73(3), 597–604. <https://doi.org/10.2307/2336524>
- Chambers, R. L., Dorfman, A. H. i Dunstan, R. (1992). Properties of estimators of finite population distribution function. *Biometrika*, 79(3), 577–582. <https://doi.org/10.2307/2336788>
- Chang, T. i Kott, P. S. (2008). Using calibration weighting to adjust for nonresponse under a plausible model. *Biometrika*, 95(3), 555–571. <https://doi.org/10.1093/biomet/asn022>
- Chen, J. K. T. (2016). *Using LASSO to calibrate non-probability samples using probability samples* (Unpublished doctoral dissertation), University of Michigan.
- Chen, J. K. T., Valliant, R. L. i Elliott, M. R. (2018a). Model-assisted calibration of non-probability sample survey data using adaptive LASSO. *Survey Methodology*, 44(1), 117–144.
- Chen, J. K. T., Valliant, R. L. i Elliott, M. R. (2018b). Calibrating non-probability surveys to estimated control totals using LASSO, with an application to political polling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 68(3), 1–25. <https://doi.org/10.1111/rssc.12327>
- Chen, Q., Elliott, M. R., Haziza, D., Yang, Y., Ghosh, M., Little, R. J. A., ... Thompson, M. (2017). Approaches to improving survey-weighted estimates. *Statistical Science*, 32(2), 227–248. <https://doi.org/10.1214/17-sts609>
- Chowdhury, S. (1994). Estimation in household surveys: An investigation on linear weighting. Internal Report, Australian Bureau of Statistics.
- Cobben, F. (2009). Nonresponse in sample surveys. Methods for analysis and adjustment. The Hague, Statistics Netherlands.
- Cohen, N., Ben-Hur, D. i Burck, L. (2017). Variance estimation in multi-phase calibration. *Survey Methodology*, 43(1), 125–140.
- Cochran, W. G. (1977). *Sampling techniques*. New York: John Wiley & Sons.
- Colley, R. H. (1945). Don't look down your nose at mail questionnaires. *Printers' Ink, March*, 16, 21–108.
- Daas, P. J. H. i Puts, M. J. (2014). Big data as a source of statistical information. *The Survey Statistician*, 69(1), 22–31.
- Daas, P. J. H., Puts, M. J., Buelens, B. i Hurk, P. A. M. van den (2015). Big data as a source for official statistics. *Journal of Official Statistics*, 31(2), 249–262. <https://doi.org/10.1515/jos-2015-0016>
- Davies, G. (2018). *Examination of approaches to calibration in survey sampling* (Unpublished doctoral dissertation). Cardiff University.

- Dehnel, G. (2010). *Rozwój mikroprzedsiębiorczości w Polsce w świetle estymacji dla małych domen*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Dehnel, G. (2012). Estymacja lokalna w szacowaniu informacji o działalności gospodarczej przedsiębiorstw. *Acta Universitatis Lodzianensis, Folia Oeconomica*, 271, 5–18.
- Deming, W. E. (1950). *Some theory of sampling*. New York: John Wiley & Sons.
- Deming, W. E. (1953). On a probability mechanism to attain an economic balance between resultant error of response and the bias of nonresponse. *Journal of the American Statistical Association*, 48(264), 743–772. <https://doi.org/10.1080/01621459.1953.10501197>
- Deming, W. E. i Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4), 427–444. <https://doi.org/10.1214/aoms/1177731829>
- Deville, J. C. i Särndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87(418), 376–382. <https://doi.org/10.1080/01621459.1992.10475217>
- Dillman, D. A., Eltinge, J. J., Groves, R. M. i Little, R. J. A. (2002). Survey nonresponse in design, data collection and analysis. W: R. M. Groves, D. A. Dillman, J. L. Eltinge, R. J. A. Little (Eds.), *Survey Nonresponse* (s. 3–26), John Wiley & Sons.
- D’Orazio, M., Di Zio, M. i Scanu, M. (2006). *Statistical matching: Theory and practice*, John Wiley & Sons.
- Duchesne, P. (1999). Robust calibration estimators. *Survey Methodology*, 25(1), 43–56.
- Dygazewicz, J. (2009). Spisy powszechnie XXI wieku. *Wiadomości Statystyczne*, (6), 13–27.
- Dygazewicz, J. (2012). Spisy powszechnie jako źródło danych do analiz geoprzestrzennych. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 23, 91–100.
- El Haj Tirari, M. i Hdioud, B. (2018). Criteria for choosing between calibration weighting and survey weighting. *Survey Methodology*, 44(2), 339–349.
- Elliot, M. R., Resler, A., Flannagan, C. i Rupp, J. (2010). Combining data from probability and non-probability samples using pseudo-weights. *Accident Analysis and Prevention*, 42, 530–539.
- Elliot, M. R. i Valliant, R. (2017). Inference for non-probability samples. *Statistical Science*, 32(2), 249–264. <https://doi.org/10.1214/16-sts598>
- Estevao, V. M., Hidiroglou, M. A. i Särndal, C. E. (1995). Methodological principles for a generalized estimation system at Statistics Canada. *Journal of Official Statistics*, 11, 181–204.
- Estevao, V. M. i Särndal, C. E. (2000). A functional form approach to calibration. *Journal of Official Statistics*, 16(4), 379–399.
- Estevao, V. M. i Särndal, C. E. (2002). The ten cases of auxiliary information for calibration in two-phase sampling. *Journal of Official Statistics*, 18(2), 233–255.
- Estevao, V. M. i Särndal, C. E. (2003). A new perspective on calibration estimators. Po-brane z <http://www.asasrms.org/Proceedings/y2003/Files/JSM2003-000462.pdf>
- Estevao, V. M. i Särndal, C. E. (2004). Borrowing strength is not the best technique within a wide class of design-consistent domain estimators. *Journal of Official Statistics*, 20(4), 645–669.

- Estevao, V. M. i Särndal, C. E. (2006). Survey estimates by calibration on complex auxiliary information. *International Statistical Review*, 74(2), 127–147. <https://doi.org/10.1111/j.1751-5823.2006.tb00165.x>
- Eurostat. (2015). Quality report of the European Union Labour Force Survey (Statistical Working Papers). Luxembourg: Publications Office of the European Union.
- Federal Statistical Office (2017). The combined census model in Germany – origins, lessons learned and future perspectives (Working Paper No. 25), Group of Experts on Population and Housing Censuses. Pobrane z https://www.zensus2021.de/Zensus2021/DE/Infothek/Publikationen/Workingpaper25.pdf?__blob=publicationFile&v=3
- Ferber, R. (1949). The problem of bias in mail returns: A solution. *Public Opinion Quarterly*, 12(4), 669–676. <https://doi.org/10.1086/266009>
- Folsom, R. E. i Singh, A. C. (2000). The generalized exponential model for sampling weight calibration for extreme values, nonresponse and poststratification. Pobrane z http://www.asasrms.org/Proceedings/papers/2000_099.pdf
- Friedman, J., Hastie, T. i Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22. <https://doi.org/10.18637/jss.v033.i01>
- Fricker, S. i Tourangeau, R. (2010). Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly*, 74(5), 934–955. <https://doi.org/10.1093/poq/nfq064>
- Gerards, E. (2012). *Dutch virtual census*. Presentation at the International Seminar on Population and Housing Censuses; Beyond the 2010 Round, Statistics Netherlands, Pobrane z <https://unstats.un.org/unsd/demographic/meetings/Conferences/Korea/2012/docs/s04-5-1-Netherlands.ppt>
- Goga, C. i Shehzad, M. A. (2014). A note on partially penalized calibration. *Pakistan Journal of Statistics*, 30(4), 429–438.
- Gołata, E. (2004). Estymacja pośrednia bezrobocia na lokalnym rynku pracy. Seria: Prace Habilitacyjne/Akademia Ekonomiczna w Poznaniu, 11(380).
- Gołata, E. (2012). Spis ludności i prawda. *Studia Demograficzne*, 1(161), 23–55. <https://doi.org/10.2478/v10274-012-0002-y>
- Gołata, E. (2014). Jakość spisów nowej generacji. *Wiadomości Statystyczne*, (5), 26–38.
- Gołata, E. (2018). *Koniec ery tradycyjnych spisów ludności*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Goyder, J. (1987). *The silent minority: Nonrespondents on sample surveys*. Westview Press.
- Graham, R. L., Knuth, D. E. i Patashnik, O. (2006). *Matematyka konkretna*. Warszawa: Wydawnictwo Naukowe PWN.
- Groen, J. A. (2012). Sources of error in survey and administrative data: The importance of reporting procedures. *Journal of Official Statistics*, 28(2), 173–198.
- Groves, R. M., Cialdini, R. B. i Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475–495. <https://doi.org/10.1086/269338>

- Groves, R. M., Dillman, D., Eltinge, J. i Little, R. (2002). *Survey nonresponse*. New York: John Wiley & Sons.
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675. <https://doi.org/10.1093/poq/nfl033>
- Groves, R. M. i Couper, M. P. (2012). *Nonresponse in household interview surveys*. New York: John Wiley & Sons.
- Guggemos, F. i Tillé, Y. (2010). Penalized calibration in survey sampling: Design-based estimation assisted by mixed models. *Journal of Statistical Planning and Inference*, 140(11), 3199–3212. <https://doi.org/10.1016/j.jspi.2010.04.010>
- GUS. (2012). *Raport z wyników. Narodowy Spis Powszechny Ludności i Mieszkań 2011*. Warszawa.
- GUS. (2013). *Ludność. Stan i struktura demograficzno-społeczna. Narodowy Spis Powszechny Ludności i Mieszkań 2011*. Warszawa.
- GUS. (2015). *Aktywność Ekonomiczna Ludności Polski – I kwartał 2015*. Warszawa.
- GUS. (2017a). *Budżety gospodarstw domowych w 2016 r.* Warszawa.
- GUS. (2017b). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2015)*. Warszawa.
- GUS. (2017c). *Społeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2013–2017*. Szczecin.
- GUS. (2018a). *Aktywność Ekonomiczna Ludności Polski*. Warszawa.
- GUS. (2018b). *Zeszyt metodologiczny. Badanie Aktywności Ekonomicznej Ludności*. Warszawa.
- GUS. (2018c). *Rozszerzenie Badania Aktywności Ekonomicznej Ludności*. Warszawa.
- Hájek, J. (1971). Comment on „An essay on the logical foundations of survey sampling D. Basu”. *The Foundations of Survey Sampling*, 236.
- Hansen, M. H. i Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, 41(236), 517–529. <https://doi.org/10.1080/01621459.1946.10501894>
- Hansen, M. H., Hurwitz, W. N. i Madow, W. G. (1953). *Sample survey methods and theory*, Vols. I and II, New York: John Wiley & Sons.
- Harms, T. i Duchesne, P. (2006). On calibration estimation for quantiles. *Survey Methodology*, 32(1), 37–52.
- Haziza, D. i Beaumont, J. F. (2017). Construction of weights in surveys: A review. *Statistical Science*, 32(2), 206–226. <https://doi.org/10.1214/16-sts608>
- Haziza, D. i Lesage, É. (2016). A discussion of weighting procedures for unit nonresponse. *Journal of Official Statistics*, 32(1), 129–145. <https://doi.org/10.1515/jos-2016-0006>
- Hidiroglou, M. A., Douglas, D. J. i Gerald, G. B. (1993). A framework for measuring and reducing nonresponse in surveys. *Survey Methodology*, 19(1), 81–94.
- Hidiroglou, M. A. i Patak, Z. (2006). Raking ratio estimation: An application to the Canadian Retail Trade Survey. *Journal of Official Statistics*, 22(1), 71–80.
- Hoerl, A. E. i Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>

- Hora, J. (2009). Problem of missing data in census – Who are the non-response respondents? Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic. Pobrane z <http://library.utia.cas.cz/separaty/2009/RO/hora-problem%20of%20missing%20data%20in%20census%20who%20are%20the%20non%20re-sponse%20respondents.pdf>
- Horvitz, D. G. i Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(26), 663–685. <https://doi.org/10.1080/01621459.1952.10483446>
- Ireland, C. T. i Kullback, S. (1968). Minimum discrimination information estimation. *Biometrics*, 24(3), 707–713. <https://doi.org/10.1002/9781118445112.stat01691>
- Józefowski, T. i Szymkowiak, M. (2013). Zastosowanie estymatora typu SPREE w szacowaniu liczby osób bezrobotnych w przekroju podregionów. *Studia Oeconomica Posnaniensia*, 10(1), 120–135.
- Kalton, G. (1983). Compensating for missing survey data. Ann Arbor: University of Michigan Press.
- Kalton, G. (2002). Models in the practice of survey sampling (revisited). *Journal of Official Statistics*, 18(2), 129–154.
- Kalton, G. i Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19, 81–97.
- Kennel, T. L. i Valliant, R. (2010). Logistic generalized regression (LGREG) estimator in cluster samples. Pobrane z http://www.asasrms.org/Proceedings/y2010/Files/308721_61470.pdf
- Kiaer, A. (1897). The representative method of statistical surveys (1976 English translation of the original Norwegian). Oslo, Central Bureau of Statistics of Norway.
- Kim, J. K. i Park, M. (2010). Calibration estimation in survey sampling. *International Statistical Review*, 78(1), 21–39.
- Kish, L. (1965). *Survey sampling*. New York: John Wiley & Sons.
- Klimanek, T. i Szymkowiak, M. (2017). Podejście kalibracyjne w badaniu losów absolwentów na przykładzie projektu „Kadry dla gospodarki”. *Wiadomości Statystyczne*, 9(676), 7–25.
- Klimanek, T., Szymkowiak, M. i Józefowski, T. (2018). Badanie zjawiska niepełnosprawności w przekroju powiatów województwa wielkopolskiego z wykorzystaniem metod statystyki małych obszarów. *Przegląd Statystyczny*, 65(4), 449–472.
- Knottnerus, P. i Duin, C. van (2006). Variances in repeated weighting with an application to the Dutch Labour Force Survey. *Journal of Official Statistics*, 22(3), 565–584.
- Kolenikov, S. i Hammer, H. (2015). Simultaneous raking of survey weights at multiple levels. Pobrane z <https://surveyinsights.org/?p=5099>
- Kordos, J. (1988). *Jakość danych statystycznych*. Warszawa: PWE.
- Koronacki, J. i Mielniczuk, J. (2018). *Statystyka dla studentów kierunków technicznych i przyrodniczych*. Warszawa: Wydawnictwo Naukowe PWN.
- Kott, P. S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, 32(2), 133–142.

- Kott, P. S. (2011). A nearly pseudo-optimal method for keeping calibration weights from falling below unity in the absence of nonresponse or frame errors. *Pakistan Journal of Statistics*, 27(4), 391–396.
- Kott, P. S. i Chang, T. (2010). Using calibration weighting to adjust for nonignorable unit nonresponse. *Journal of the American Statistical Association*, 105(491), 1265–1275. <https://doi.org/10.1198/jasa.2010.tm09016>
- Kovačević, M. S. (1997). Calibration estimation of cumulative distribution and quantile functions from survey data. W: *Proceedings of the Survey Methods Section* (s. 139–144). Statistical Society of Canada.
- Kozłowski, A. (2016). Losowanie zrównoważone i kalibracja. *Wiadomości Statystyczne*, (3), 38–60.
- Krapavickaitė, D. i Plikusas, A. (2005). Estimation of a ratio in the finite population. *Informatica*, 16(3), 347–364.
- Krywult-Albańska, M. (2013). Powszechne spisy ludności na przykładzie wybranych państw. Aspekty metodologiczne. *Politeja*, 4(26), 275–290. <https://doi.org/10.12797/politeja.10.2013.26.12>
- Krzyśko, M. (2004). *Statystyka matematyczna*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.
- Laaksonen, S. (2018). *Survey methodology and missing data. Tools and techniques for practitioners*. Springer.
- Laitila, T., Wallgren, A. i Wallgren, B. (2011). *Quality assessment of administrative data*. Statistiska Centralbyrån.
- Leeuw, E. D., Callegaro, M., Hox, J., Korendijk, E. i Lensvelt-Mulders, G. (2007). The influence of advance letters on response in telephone surveys: A meta-analysis. *Public Opinion Quarterly*, 71(3), 413–443. <https://doi.org/10.1093/poq/nfm014>
- Leeuw, E. D., Hox, J. i Huisman, M. (2003). Prevention and treatment of item nonresponse. *Journal of Official Statistics*, 19(2), 153–176.
- Lehtonen, R. (2009). A short note on extended GREG family estimators for domains. W: M. Carlson, H. Nyquist i M. Villani (Eds.), *Official Statistics – Methodology and Applications in Honour of Daniel Thorburn* (s. 107–116). Stockholm: Department of Statistics, Stockholm University.
- Lehtonen, R. i Veijanen, A. (1998). Logistic generalized regression estimators. *Survey Methodology*, 24(1), 51–55.
- Lehtonen, R. i Veijanen, A. (2012). Small area poverty estimation by model calibration. *Journal of the Indian Society of Agricultural Statistics*, 66(1), 125–133.
- Lehtonen, R. i Veijanen, A. (2015). Small area estimation by calibration methods. WSC 2015 of the ISL, Rio de Janeiro.
- Lehtonen, R. i Veijanen, A. (2017). A two-level hybrid calibration technique for small area estimation. SAE 2017 Conference, Paris.
- Lehtonen, R. i Veijanen, A. (2018). Small area estimation with calibration methods (Draft Working Paper), 1–11.
- Lemaître, G. i Dufour, J. (1987). An integrated method for weighting persons and families. *Survey Methodology*, 13(2), 199–207.

- Little, R. J. i Rubin, D. B. (2014). *Statistical analysis with missing data*. New York: John Wiley & Sons.
- Longford, N. T. (2006). *Missing data and small area estimation: Modern analytical equipment for the survey statistician*, Springer Science & Business Media.
- López-Vizcaíno, E., Lombardía, M. J. i Morales, D. (2013). Multinomial-based small area estimation of labour force indicators. *Statistical Modelling*, 13(2), 153–178. <https://doi.org/10.1177/1471082x13478873>
- López-Vizcaíno, E., Lombardía, M. J. i Morales, D. (2015). Small area estimation of labour force indicators under a multinomial model with correlated time and area effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3), 535–565. <https://doi.org/10.1111/rssa.12085>
- Lumiste, K. (2011). Consistent estimation of cross-classified domains. *Statistics in Transition new series*, 12(2), 253–264.
- Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9(8), 1–19. <https://doi.org/10.18637/jss.v009.i08>
- Lumley, T. (2017). Survey: Analysis of complex survey samples. *R package version 3.32*.
- Lund, K. A. i Gulløy, E. (2016). The use of incentives in household surveys: A future strategy to save response rates, or a relic of the past?. *Nordiskt Statistiker möte*, Stockholm.
- Lundström, S. i Särndal, C. E. (1999). Calibration as a standard method for treatment of nonresponse. *Journal of Official Statistics*, 15(2), 305–327.
- Manski, C. F. (2016). Credible interval estimates for official statistics with survey nonresponse. *Journal of Econometrics*, 19(2), 293–301. <https://doi.org/10.1016/j.jeconom.2015.12.002>
- Marchetti, S., Beręsewicz, M., Salvati, N., Szymkowiak, M. i Wawrowski, Ł. (2018). The use of a three-level M-quantile model to map poverty at local administrative unit 1 in Poland. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(4), 1077–1104. <https://doi.org/10.1111/rssa.12349>
- McConville, K. S., Breidt, F. J., Lee, T. C. M. i Moisen, G. G. (2017). Model-assisted survey regression estimation with the LASSO. *Journal of Survey Statistics and Methodology*, 5, 131–158. <https://doi.org/10.1093/jssam/smw041>
- Młodak, A. (2014). Zastosowanie metody bootstrapowej do oceny precyzji szacunków danych z badania reprezentacyjnego NSP 2011 na przykładzie dojazdów do pracy. *Wiadomości Statystyczne*, (12), 38–57.
- Montanari, G. E. i Ranalli, M. G. (2005). Nonparametric model calibration estimation in survey sampling. *Journal of the American Statistical Association*, 100(472), 1429–1442. <https://doi.org/10.1198/016214505000000141>
- Montanari, G. E. i Ranalli, M. G. (2009). Multiple and ridge model calibration for sample surveys. Pobrane z <http://www.stat.unipg.it/~giovanna/papers/ranalliWCES.pdf>
- Moser, C. A. i Stuart, A. (1953). An experimental study of quota sampling. *Journal of the Royal Statistical Society. Series A*, 116(4), 349–405.
- Münnich, R., Burgard, J. P., Gabler, S., Ganninger, M. i Kolb, J. P. (2016). Small area estimation in the German census 2011. *Statistics in Transition new series*, 17(1), 25–40. <https://doi.org/10.21307/stattrans-2016-004>

- Münnich, R., Wagner, M. i Sachs, E. W. (2012). Calibration benchmarking for small area estimates: An application to the German census 2011. *Symposium on the Analysis of Survey Data and Small Area Estimation in Honour of the 75th Birthday of J. N. K. Rao*.
- Myrskylä, M. (2007). Generalised regression estimation for domain class frequencies. *Research Reports*, (247), Statistics Finland.
- NBP. (2017). *Zasobność gospodarstw domowych w Polsce*. Aneks metodyczny do badania 2016 r. Warszawa: NBP, Departament Analiz Ekonomicznych i Departament Stabilności Finansowej.
- Neethling, A. i Galpin, J. S. (2006). Weighting of household survey data: A comparison of various calibration, integrated and cosmetic estimators. *South African Statistical Journal*, 40, 123–150.
- Neyman, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society (Series A)*, 97(4), 558–606. <https://doi.org/10.2307/2342192>
- Nguyen, J. H. (2010). An Introduction to calibration estimators. Pobrane z <http://mathreview.uwaterloo.ca/archive/voli/2/nguyen.pdf>
- Oh, H. L. i Scheuren, F. (1987). Modified raking ratio estimation. *Survey Methodology*, 13(2), 209–219.
- Panek, T., Czapiński, J. i Kotowska, I. E. (2011). Metodologia badań. Diagnoza społeczna 2011. Warunki i jakość życia Polaków. *Contemporary Economics*, 5(3), 35–44.
- Paradysz, J. (2009). Spisy jako źródło informacji o warunkach życia ludności w Polsce. *Wiadomości Statystyczne*, (7), 1–9.
- Paradysz, J. (2012). Statystyka regionalna: Stan, problemy i kierunki rozwoju. *Przegląd Statystyczny*, 2, 191–204.
- Pawłowski, R. (1972). *Wstęp do statystycznej metody reprezentacyjnej*. Warszawa: PWN.
- Peytchev, A. (2013). Consequences of survey nonresponse. *The Annals of the American Academy of Political and Social Science*, 645(1), 88–111.
- Pfeffermann, D. (2013). New important developments in small area estimation. *Statistical Science*, 28(1), 40–68. <https://doi.org/10.1214/12-sts395>
- Piasecki, T. (2014). Metody imputacji w badaniach gospodarstw domowych. *Wiadomości Statystyczne*, (9), 1–20.
- Plikusas, A. (2001). Calibrated estimators of the ratio. *Lithuanian Mathematical Journal*, 41, Special Issue, 457–462.
- Plikusas, A. (2003). Calibrated weights for the estimators of the ratio. *Lithuanian Mathematical Journal*, 43, 543–547.
- Plikusas, A. i Pumputis, D. (2007). Calibrated estimators of the population covariance. *Acta Applicandae Mathematicae*, 97(1–3), 177–187. <https://doi.org/10.1007/s10440-007-9129-7>
- Plikusas, A. i Pumputis, D. (2007). Estimation of the finite population covariance using calibration. *Nonlinear Analysis: Modelling and Control*, 15(3), 325–340.
- Popiński, W. (2006). Development of the Polish Labour Force Survey. *Statistics in Transition*, 7(5), 1009–1030.

- Potter, F. (1990). A study of procedures to identify and trim extreme sampling weights. W: *Proceedings of the Section on Survey Research Methods* (s. 225–230), American Statistical Association.
- Rahman, A. (2017). Small area housing stress estimation in Australia: Calculating confidence intervals for a spatial microsimulation model. *Communications in Statistics-Simulation and Computation*, 46(9), 7466–7484. <https://doi.org/10.1080/03610918.2016.1241406>
- Rao, J. N. K. i Fuller, W. A. (2017). Sample survey theory and methods: Past, present, and future directions. *Survey Methodology*, 43(2), 145–160.
- Rao, J. N. K., Kovar, J. G. i Mantel, H. J. (1990). On estimating distribution functions and quantiles from survey data using auxiliary information. *Biometrika*, 77, 365–375. <https://doi.org/10.1093/biomet/77.2.365>
- Rao, J. N. K. i Molina, I. (2015). *Small area estimation*. New York: John Wiley & Sons.
- R Core Team. (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria.
- Ren, R. (2002). Estimation de la fonction de répartition et des fractiles d'une population finie. *Actes des journées de méthodologie statistique*, INSEE Méthodes, 1(100), 263–289.
- Ren, R. i Chambers, R. (2002). Outlier robust methods: Outlier robust estimation and outlier robust imputation by reverse calibration. Report for Euredit, 1–170.
- Renssen, R. H. (1998). Use of statistical matching techniques in calibration estimation. *Survey Methodology*, 24, 171–184.
- Renssen, R. H. i Nieuwenbroek, N. J. (1997). Aligning estimates for common variables in two or more sample surveys. *Journal of the American Statistical Association*, 92(437), 368–364. <https://doi.org/10.1080/01621459.1997.10473635>
- Roszka, W. (2013). *Statystyczna integracja danych w badaniach społeczno-ekonomicznych* (Niepublikowana rozprawa doktorska). Poznań: Uniwersytet Ekonomiczny w Poznaniu.
- Roszka, W. i Szymkowiak, M. (2014). Podejście kalibracyjne w statystycznej integracji danych. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Taksonomia* 23(328), 308–315.
- Roszka, W., Szymkowiak, M. i Rahman, A. (w druku). Household income estimation using spatial microsimulation – the GREGWT approach. *Statistics in Transition New Series*.
- Rószkiewicz, M. (2015). Identyfikacja determinant braku odpowiedzi w badaniu polskich gospodarstw domowych. *Przegląd Statystyczny*, 62(4), 361–378.
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
- Rueda, M., Martínez-Puertas, S., Martínez-Puertas, H. i Arcos, A. (2006). Estimation of the distribution function with calibration methods. *Journal of Statistical Planning and Inference*, 137, 435–448.
- Rueda, M., Martínez-Puertas, S., Martínez-Puertas, H. i Arcos, A. (2007). Calibration methods for estimating quantiles. *Metrika*, 66(3), 355–371. <https://doi.org/10.1007/s00184-006-0116-1>

- Rueda, M., Sánchez-Borrego, I., Arcos, A. i Martínez, S. (2010). Model-calibration estimation of the distribution function using nonparametric regression. *Metrika*, 71(1), 33–44. <https://doi.org/10.1007/s00184-008-0199-y>
- Särndal, C. E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, 33(2), 99–119.
- Särndal, C. E. (2010). Models in survey sampling. *Statistics in Transition new series*, 11(3), 539–554.
- Särndal, C. E. i Lundström, S. (2005). *Estimation in surveys with nonresponse*. New York: John Wiley & Sons.
- Särndal, C. E., Swensson, B. i Wretman, J. (1992). *Model assisted survey sampling*. New York: Springer-Verlag.
- Särndal, C. E. i Traat, I. (2011). Domain estimators calibrated on information from another survey. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, 15(2), 43–60.
- Särndal, C. E., Traat, I. i Lumiste, K. (2018). Interaction between data collection and estimation phases in surveys with nonresponse. *Statistics in Transition New Series*, 19(2), 183–200. <https://doi.org/10.21307/stattrans-2018-011>
- Shao, J. i Tu, D. (1995). *The jackknife and bootstrap*. New York: Springer-Verlag.
- Shook-Sa, B. E., Kott, P. S., Berzofsky, A. E., Couzens, G. L., Moore, A., Lee, P., ... Planty, M. (2017). Assessing alternative precision measures when adjusting for conditional bias at the subnational level through calibration weighting. *Survey Research Methods*, 11(4), 405–414. <https://doi.org/10.18148/srm/2017.v11i4.6789>
- Singer, E. i Ye, C. (2013). The use and effects of incentives in surveys. *The Annals of the American Academy of Political and Social Science*, 645(1), 112–141. <https://doi.org/10.1177/0002716212458082>
- Singh, A. i Mohl, C. (1996). Understanding calibration estimators in survey sampling. *Survey Methodology*, 22(2), 107–115.
- Steczkowski, J. (1995). *Metoda reprezentacyjna w badaniach ekonomiczno-społecznych*. Warszawa: PWE.
- Stephan, F. i McCarthy, P. J. (1958). *Sampling opinions*. New York: John Wiley & Sons.
- Stoop, I. A. L. (2005). *The hunt for the last respondent: Nonresponse in sample surveys*, The Hague: Social and Cultural Planning Office.
- Stoop, I. A. L., Billiet, J., Koch, A. i Fitzgerald, R. (2010). *Improving survey response: Lessons learned from the European Social Survey*. New York: John Wiley & Sons.
- Sudman, S. (1966). Probability sampling with quotas. *Journal of the American Statistical Association*, 61(315), 749–791. <https://doi.org/10.1080/01621459.1966.10480903>
- Sudman, S. (1976). *Applied Sampling*, New York: Academic Press.
- Szreder, M. (2010a). *Metody i techniki sondażowych badań opinii*, Warszawa: PWE.
- Szreder, M. (2010b). Losowe i nielosowe próby w badaniach statystycznych. *Przegląd Statystyczny*, 4(57), 168–174.
- Szreder, M. (2011). O niektórych źródłach i konsekwencjach braków odpowiedzi w badaniach ankietowych. *Marketing i Rynek*, (5), 2–5.
- Szreder, M. (2015a). Big data wyzwaniem dla człowieka i statystyki. *Wiadomości Statystyczne*, (8), 1–11.

- Szreder, M. (2015b). Zmiany w strukturze całkowitego błędu badania próbkowego. *Wiadomości Statystyczne*, (1), 4–12.
- Szreder, M. (2016). *Rola badań statystycznych w naukach ekonomicznych w świetle nowych możliwości określanych mianem big data*. Pobrane z http://www.pte.pl/pliki/2/1/Konferencja_ENE_referat_M_Szreder.pdf
- Szreder, M. (2017). Nowe źródła informacji i ich wykorzystywanie w podejmowaniu decyzji. *Wiadomości Statystyczne*, 7(674), 5–17.
- Szymkowiak, M. (2009a). *Estymatory kalibracyjne w badaniu budżetów gospodarstw domowych* (Niepublikowana rozprawa doktorska). Poznań: Uniwersytet Ekonomiczny w Poznaniu.
- Szymkowiak, M. (2009b). Calibration estimators for quantiles in surveys with nonresponse. W: J. L., Wywiół i T. Żądło (Eds.), *Survey Sampling in Economic and Social Research* (s. 60–75). *Zeszyty Naukowe Akademii Ekonomicznej w Katowicach*, Katowice: Wydawnictwo Akademii Ekonomicznej w Katowicach.
- Szymkowiak, M. (2010). Konstrukcja estymatorów kalibracyjnych wartości globalnej z uwzględnieniem wektora zmiennych instrumentalnych. W: E. Gołata (red.), *Pomiar i informacja w gospodarce* (s. 122–144). Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Szymkowiak, M. (2013). Konstrukcja estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 279, 164–173.
- Szymkowiak, M. (2014). Estymatory kalibracyjne stosowane w Narodowym Spisie Powszechnym w 2011 r. *Wiadomości Statystyczne*, (11), 1–16.
- Szymkowiak, M. (2017). Podejście kalibracyjne wykorzystujące analizę składowych głównych w badaniach statystycznych z brakami odpowiedzi. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 468, 230–238.
- Szymkowiak, M., Józefowski, T. i Klimanek, T. (2018). SPREE estimation of the number of people with disability in terms of economic activity atatus. Improving the Quality of the EU Census (2021 and beyond), Eurostat project, Central Statistical Office, Warsaw.
- Szymkowiak, M., Józefowski, T. i Klimanek, T. (w druku). Modele log-liniowe w badaniu niepełnosprawności. *Wiadomości Statystyczne*.
- Szymkowiak, M., Młodak, A. i Wawrowski, Ł. (2017). Mapping poverty at the level of subregions in Poland using indirect estimation. *Statistics in Transition new series*, 18(4), 609–635. <https://doi.org/10.21307/stattrans-2017-003>
- Tanton, R. (2018). Spatial microsimulation: Developments and potential future directions. *International Journal of Microsimulation*, 11(1), 143–161.
- Thibaudeau, Y. i Slud, E. (2009). Simultaneous calibration and nonresponse adjustment. W: *Proceedings of the Survey Research Methods Section of the American Statistical Association* (s. 2263–2272).
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.

- Toepoel, V. i Schonlau, M. (2017). Dealing with nonresponse: Strategies to increase participation and methods for postsurvey adjustments. *Mathematical Population Studies*, 24(2), 79–83. <https://doi.org/10.1080/08898480.2017.1299988>
- Tzavidis, N., Zhang, L. C., Luna Hernandez, A., Schmid, T. i Rojas-Perilla, N. (2018). From start to finish: A framework for the production of small area official statistics. *Journal of the Royal Statistical Society A*, 181(4), 927–979. <https://doi.org/10.1111/rssa.12364>
- Urząd Statystyczny w Poznaniu. (2017). Ocena bieżącej działalności gospodarczej przedsiębiorstw – miesięczny meldunek DG–1. Poznań.
- Valente, S. (2010). Census taking in Europe: How are populations counted in 2010? *Population and Societies*, 467(1), 1–4.
- Vanderhoeft, C. (2001). Generalised calibration at Statistics Belgium. SPSS module g-CALIB-S and current practices (Working Paper No. 3), Statistics Belgium.
- Waal, T. de (2015). General approaches for consistent estimation based on administrative data and surveys (Discussion Paper). Statistics Netherlands, 1–38.
- Wallgren, A. i Wallgren, B. (2014). *Register-based statistics: Statistical methods for administrative data*. New York: John Wiley & Sons, Ltd.
- Welsh, A. H. i Ronchetti, E. (1998). Bias-calibrated estimation from sample surveys containing outliers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2), 413–428. <https://doi.org/10.1111/1467-9868.00133>
- Wesołowski, J. i Tarczyński, J. (2016). Podstawy matematyczne technik imputacyjnych. *Wiadomości Statystyczne*, (9), 7–54.
- Wilak, K. (2014). Strukturalne modele szeregów czasowych w estymacji stopy bezrobocia w dezagregacji na województwa, płeć i wiek. *Przegląd Statystyczny*, 61(4), 409–431.
- Wilak, K., Beręsewicz, M. i Szymkowiak, M. (2017a). Weight calibration for MUR estimation. W: A. Zgierska (Ed.). *Quality improvement of the monthly unemployment rate* (s. 50–61), Central Statistical Office, Warsaw.
- Wilak, K., Beręsewicz, M. i Szymkowiak, M. (2017b). Indirect estimation results. W: A. Zgierska (Ed.). *Quality improvement of the monthly unemployment rate* (s. 74–89), Central Statistical Office, Warsaw.
- Witkowski, J. (2010). Rola statystyki publicznej we współczesnym świecie. *Wiadomości Statystyczne*, 585(2), 1–18.
- Wright, T. i Farmer, J. (2000). A bibliography of selected statistical methods and development related to census 2000. Statistical Research Report Series, No. RR 2000/02, Washington, D.C.
- Wu, C. (2003). Optimal calibration estimators in survey sampling. *Biometrika*, 90(4), 937–951. <https://doi.org/10.1093/biomet/90.4.937>
- Wu, C. i Lu, W. W. (2016). Calibration weighting methods for complex surveys. *International Statistical Review*, 84(1), 79–98. <https://doi.org/10.1111/insr.12097>
- Wu, C. i Sitter, R. R. (2001). A model-calibration approach to using complete auxiliary information from survey data. *Journal of the American Statistical Association*, 96(453), 185–193. <https://doi.org/10.1198/016214501750333054>

- Wywił, J. L. (1999). *Elementy metody reprezentacyjnej z wykorzystaniem statystycznego pakietu SPSS*. Katowice: Wydawnictwo Akademii Ekonomicznej w Katowicach.
- Wywił, J. L. (2010). *Wprowadzenie do metody reprezentacyjnej*. Katowice: Wydawnictwo Akademii Ekonomicznej w Katowicach.
- Yan, T. i Curtin, R. (2010). The relation between unit nonresponse and item nonresponse: A response continuum perspective. *International Journal of Public Opinion Research*, 22(4), 535–551. <https://doi.org/10.1093/ijpor/edq037>
- Yates, F. (1946). A review of recent statistical developments in sampling and sample surveys. *Journal of the Royal Statistical Society*, 109(1), 12–43. <https://doi.org/10.2307/2981390>
- Zasępa, R. (1962). *Badania statystyczne metodą reprezentacyjną*. Warszawa: PWN.
- Zasępa, R. (1972). *Metoda reprezentacyjna*. Warszawa: PWE.
- Zeelenberg, K. (2016). Big data and methodological challenges in official statistics (Discussion Paper No. 8), Statistics Netherlands.
- Zou, H. (2006). The adaptive LASSO and its oracle properties. *Journal of the American Statistical Association*, 101(476), 1418–1429. <https://doi.org/10.1198/016214506000000735>
- Żądło, T. (2011). On some calibration estimators of subpopulation total for longitudinal data. *Acta Universitatis Lodzianensis. Folia Oeconomica*, (252), 191–204.
- Żądło, T. (2015). *Statystyka małych obszarów w badaniach ekonomicznych. Podejście modelowe i mieszane*, Katowice: Wydawnictwo Naukowe Uniwersytetu Ekonomicznego w Katowicach.

Spis tabel

1.1. Dwudzielcza tablica kontyngencji ($t \times w$)	35
2.1. Zestawienie funkcji $G(x)$, $F(u)$ oraz ograniczeń na wagi w podejściu kalibracyjnym	61
2.2. Wartości miar jakości estymatorów kalibracyjnych	83
2.3. Statystyki opisowe mnożników g_k i wag kalibracyjnych w_k dla różnych funkcji $G(\cdot)$	85
3.1. Ocena jakości estymatorów kalibracyjnych wartości globalnych w podejściu funkcyjnym	104
4.1. Różnice między podejściem kalibracyjnym minimalizującym funkcję odległości a wspomaganym modelem	116
4.2. Wartości miar jakości estymatorów w podejściu modelowo-kalibrowanym dla funkcji odległości w postaci sinusa hiperbolicznego	129
4.3. Statystyki opisowe wag w podejściu modelowo-kalibrowanym dla funkcji odległości w postaci sinusa hiperbolicznego	131
5.1. Rodzaje braków danych w hipotetycznym badaniu statystycznym	138
5.2. Wektor zmiennych pomocniczych oraz odpowiadający mu wektor wartości globalnych w podejściu kalibracyjnym z brakami danych	144
5.3. Charakterystyki wag kalibracyjnych	161
6.1. Sposób tworzenia przekrojów w podejściu kalibracyjnym w NSP 2011 – kategoria ludności faktycznie zamieszkałej	183
6.2. Statystyki opisowe mnożników kalibracyjnych g_k dla rozważanych podejść i funkcji kalibracyjnych	207
D.1. Skróty wykorzystane w pracy	240
D.2. Wykaz najważniejszych symboli wykorzystanych w pracy	242

Spis rysunków

1.1. Ewolucja teorii kalibracji	44
2.1. Wykresy funkcji $G(x)$ i $F(u)$ w podejściu kalibracyjnym	60
2.2. Wykresy powierzchni dla wybranych funkcji celu $D(\mathbf{d}, \mathbf{w})$	71
2.3. Korelacyjny diagram rozrzutu między wydatkami ogółem a dochodem rozporządzalnym gospodarstw domowych (lewy panel) oraz wykresy pudełkowe dochodu rozporządzalnego i wydatków ogółem (prawy panel)	79
2.4. Oszacowana wartość globalna wydatków ogółem gospodarstw domowych	82
2.5. 95% przedziały ufności wartości globalnej wydatków ogółem gospodarstw domowych	84
2.6. Rozkład wag kalibracyjnych we wszystkich replikacjach	86
3.1. Rozkłady zmiennych w badaniu symulacyjnym nad własnościami estymatorów kalibracyjnych wartości globalnej – podejście funkcyjne	101
3.2. Oszacowana wartość globalna wydatków ogółem gospodarstw domowych – podejście funkcyjne	106
3.3. Porównanie rozkładów wag kalibracyjnych w rozpatrywanych wariantach – podejście funkcyjne	107
4.1. Oszacowana wartość globalna wydatków ogółem gospodarstw domowych w podejściu modelowo-kalibrowanym	128
4.2. Rozkład wag kalibracyjnych w $B = 200$ replikacjach w podejściu modelowym dla funkcji odległości w postaci sinusa hiperbolicznego	130
5.1. Wskaźniki realizacji dla wybranych badań reprezentacyjnych	137
5.2. Porównanie rozkładów wag kalibracyjnych w rozpatrywanych podejściach	160
5.3. Porównanie wag pierwotnych d_k , bazowych (finalnych w Badaniu Aktywności Ekonomicznej Ludności) oraz wag kalibracyjnych w_k otrzymanych dla drugiego zestawu zmiennych pomocniczych i dwukrokowego podejścia kalibracyjnego typu B wraz z funkcją kalibracyjną sinus hiperboliczny	164

5.4. Oszacowana liczba osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów województwa wielkopolskiego	165
5.5. Porównanie oszacowań liczby osób pracujących, bezrobotnych i biernych zawodowo w 2015 roku (w tys.)	166
5.6. Względne średnie błędy szacunku liczby osób pracujących, bezrobotnych i biernych zawodowo w przekroju podregionów w Polsce we wszystkich kwartałach 2015 roku	168
6.1. Metody przeprowadzania spisów w krajach europejskich w latach 2010–2011	174
6.2. Rozkład wag kalibracyjnych w zależności od płci i miejsca zamieszkania . .	185
6.3. Korelacyjny diagram rozrzutu wag v_k i wag kalibracyjnych w_k w przekroju województw	186
6.4. Korelacyjny diagram rozrzutu wag v_k i w_k i histogram mnożników kalibracyjnych g_k dla Warszawy w przekroju płci	187
6.5. Rozkład mnożników kalibracyjnych g_k w zależności od podejścia i wykorzystanej funkcji	205
6.6. Rozkład wag kalibracyjnych w_k w zależności od podejścia i wykorzystanej funkcji	206
6.7. Rozkłady wag v_k i kalibracyjnych w_k w zależności od podejścia i metody ich wyznaczania	208
6.8. Korelacyjny diagram rozrzutu wag v_k i kalibracyjnych w_k w zależności od podejścia i wykorzystanej funkcji	209
6.9. Porównanie oszacowań spisowych liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy z uzyskanymi dla wag zintegrowanych w podejściach 1–4	210
6.10. Względne różnice między oszacowaniami spisowymi liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy a otrzymanymi dla wag z podejść 1–4	212
6.11. Względne średnie błędy szacunków liczby osób pracujących, bezrobotnych, biernych zawodowo i o nieustalonym statusie na rynku pracy w przekroju płci i miejsca zamieszkania dla różnych podejść	213

Spis algorytmów

1.	Algorytm wyznaczania mnożników λ	73
2.	Rozszerzony algorytm wyznaczania mnożników λ	75
3.	Algorytm wyznaczania mnożników λ_* w podejściu funkcyjnym	97
4.	Wagi kalibracyjne w podejściu dwukrokovym typu A	150
5.	Wagi kalibracyjne w podejściu dwukrokovym typu B	151
6.	Konstrukcja wag w Badaniu Aktywności Ekonomicznej Ludności	155
7.	Konstrukcja wag kalibracyjnych w NSP 2011	181
8.	Konstrukcja wag w_{Mi} oraz w_k w losowaniu dwustopniowym	199

Dodatek: Skróty i symbole

Tabela D.1. Skróty wykorzystane w pracy

SKRÓT	ZNACZENIE	TERMIN ANGIELSKI
AL	średnia długość przedziału ufności	average length of the confidence interval
BBGD	Badanie Budżetów Gospodarstw Domowych	Household Budget Survey
BAEL	Badanie Aktywności Ekonomicznej Ludności	Labour Force Survey
CAII	badanie internetowe wspomagane komputerowo	computer assisted internet interview
CAPI	badanie ankietowe wspomagane komputerowo	computer assisted personal interview
CATI	wspomagany komputerowo wywiad telefoniczny	computer assisted telephone interview
CR	wskaźnik pokrycia przez przedział ufności	coverage rate
GUS	Główny Urząd Statystyczny	Statistics Poland
GREG	uogólniony estymator regresyjny	generalized regression estimator
GREGWT	metoda uogólnionej regresji i ważenia	generalized regression and weighting
EU-SILC	Europejskie Badanie Warunków Życia Ludności	European Union Statistics on Income and Living Conditions
ICT	technologie informacyjne i komunikacyjne	information and communication technology
ILO	Międzynarodowa Organizacja Pracy	International Labour Organization
IPF	algorytm iteracyjnego proporcjonalnego dopasowania	iterative proportional fitting
JPS	jednostki losowania pierwszego stopnia	primary sampling units
LASSO	regresja typu LASSO	least absolute shrinkage and selection operator LASSO
MSE	błąd średniokwadratowy estymatora	mean squared error
MLE	metoda największej wiarygodności	maximum likelihood method
MUR	miesięczna stopa bezrobocia	monthly unemployment rate

Tabela D.1 – ciąg dalszy

SKRÓT	ZNACZENIE	TERMIN ANGIELSKI
NSP 2002	Narodowy Spis Powszechny Ludności i Mieszkań 2002	National Census of Population and Housing 2002
NSP 2011	Narodowy Spis Powszechny Ludności i Mieszkań 2011	National Census of Population and Housing 2011
NSP 2021	Narodowy Spis Powszechny Ludności i Mieszkań 2021	National Census of Population and Housing 2021
PESEL	Powszechny Elektroniczny System Ewidencji Ludności	Universal Electronic System for Registration of the Population
PKD	Polska Klasyfikacja Działalności	Polish Classification of Activities
RATIO	estymator ilorazowy	ratio estimator
RB	obciążenie względne	relative bias
REE	względny średni błąd szacunku	relative estimation error
REG	estymator regresyjny	regression estimator
RIDGE	regresja grzbietowa	ridge regression
RMSE	pierwiastek z błędu średniokwadratowego	root mean squared error
RRMSE	względny błąd średniokwadratowy	relative root mean squared error
RWM	metoda powtarzanego ważenia	repeated weighting method
SAS	program do analiz statystycznych SAS	Statistical Analysis Software
SEPPRA	suma estymatorów ilorazowych w grupach	separate ratio estimator
SEPPREG	suma estymatorów regresyjnych w grupach	separate regression estimator
SPREE	estymator zachowujący strukturę	structure preserving estimation
SP3	Badanie Działalności Gospodarczej Przedsiębiorstw	Business Activity Survey

Tabela D.2. Wykaz najważniejszych symboli wykorzystanych w pracy

SYMBOL	ZNACZENIE (jeżeli w tekście nie podano innego)
$\frac{\partial f}{\partial x_j}$	pochodna cząstkowa funkcji f względem zmiennej x_j
A^T	macierz transponowana macierzy A
A^{-1}	macierz odwrotna do A
A^-	uogólniona macierz odwrotna do A
d_k	waga dla k -tej jednostki i wynikająca z planu losowania
\mathbf{d}	wektor wag wynikających z planu losowania próby
$E(X)$	wartość oczekiwana zmiennej losowej X
\mathbf{I}_n	macierz jednostkowa o n wierszach i n kolumnach
$\mathbf{0}_{m,n}$	macierz zerowa o m wierszach i n kolumnach
$\mathbf{J}_{m,n}$	macierz jedynek o m wierszach i n kolumnach
$\boldsymbol{\lambda}$	wektor mnożników Lagrange'a
J	liczba zmiennych pomocniczych w kalibracji
n	liczebność próby
N	liczebność populacji
$P(s)$	rozkład prawdopodobieństwa określony na przestrzeni prób \mathbf{S}
π_k	prawdopodobieństwo inkluzji pierwszego rzędu
π_{kl}	prawdopodobieństwo inkluzji drugiego rzędu
r	zbiór respondentów
s	próba
\mathbf{S}	przestrzeń prób
σ_Y^2	wariancja cechy Y
$\text{Cov}(X, Y)$	kowariancja cech X i Y
$Q_{Y,\alpha}$	kwantyl rzędu α
U	populacja skończona
w_k	waga kalibracyjna dla k -tej jednostki
\mathbf{w}	wektor wag kalibracyjnych
τ_Y	wartość globalna cechy Y
Y_{\max}	maksymalna wartość cechy Y
Y_{\min}	minimalna wartość cechy Y
μ_Y	średnia cechy Y
\sum	suma
\prod	iloczyn
\cup	suma zbiorów

Summary: The calibration approach in socio-economic surveys

The book explores calibration methods in survey sampling and censuses, a field which is undergoing a dynamic development all over the world. In particular, the calibration approach is presented on the basis of the author's many years of experience involving the use of this technique for statistical research and its applications in official statistics in Poland. The book addresses a wide range of issues concerning the use of calibration in surveys with missing data or in mixed-mode censuses. Analysis also focuses on a functional form approach to the calculation of calibration weights, model-assisted calibration and the hybrid approach.

The author covers both methodological and empirical aspects, maintaining a sound balance between estimation theory and applications. The methodological part describes the state of the art of calibration research, including a detailed review of literature, which can serve as the starting point for other authors interested in exploring these issues. The empirical part draws on real data from surveys conducted by Poland's Central Statistical Office to provide a comprehensive assessment of weights and quality measures used for the calibration estimators of interest and their applications in the labour market.

The book is an attempt to fill the existing knowledge gap concerning the calibration approach and is the first comprehensive analysis of this field in the Polish literature. What makes the book even more relevant is the fact that calibration as a method of estimation is bound to play an increasingly important role in statistical surveys.

Keywords: calibration approach, calibration estimators, integrated calibration, nonresponse, survey sampling, censuses, labour market.

W recenzowanej książce został przedstawiony aktualny stan wiedzy na temat metod kalibracji stosowanych w badaniach reprezentacyjnych i spisach pełnych. [...] Obok szczegółowego przeglądu literatury zaproponowano wiele rozwiązań autorskich, m.in. konstrukcję estymatorów hybrydowo-kalibrowanych z uwzględnieniem regresji LASSO oraz grzbietowej, a także uogólniono sposób wyznaczania wag w podejściu dwukrokowym dla dowolnie zdefiniowanej funkcji odległości. Na uwagę zasługują również cenne badania symulacyjne przeprowadzone w celu zbadania podstawowych własności uzyskanych estymatorów kalibracyjnych. Część teoretyczna książki jest wsparta wieloma ciekawymi zastosowaniami praktycznymi, m.in. w badaniu budżetów gospodarstw domowych oraz badaniu aktywności ekonomicznej ludności. [...]

Nie było dotychczas na rynku wydawniczym tak wszechstronnego opracowania dotyczącego metod kalibracji.

(z recenzji wydawniczej prof. dra hab. Mirosława Krzyśki)

Prezentowana książka dotyczy ważnego dla teorii i praktyki statystycznych badań niewyczerpujących (próbkowych) zagadnienia kalibracji, czyli korygowania wag w próbie w celu poprawy jakości wnioskowania. [...] Cieszy się ono dużym zainteresowaniem wśród badaczy [...], w szczególności w naukach ekonomicznych i społecznych.

Monografia na temat wykorzystania kalibracji w badaniach społeczno-ekonomicznych jest – o ile mi wiadomo – pierwszą na ten temat na polskim rynku. [...] Autor potrafi wnieść do tych badań oryginalne i naukowo wartościowe propozycje wykorzystania różnych technik kalibracji i zaproponować rozwiązania użyteczne dla przyszłych badań oficjalnej statystyki. [...]

Wysoko oceniam metodykę badawczą wykorzystaną przez Autora. Obejmuje ona: matematyczną teorię wnioskowania statystycznego, metodykę badań reprezentacyjnych, eksperymenty symulacyjne i wnioskowanie na podstawie empirycznych zastosowań określonych rozwiązań z zakresu estymacji w badaniach statystyki publicznej.

(z recenzji wydawniczej prof. dra hab. Mirosława Szredera)

ISBN 978-83-66199-23-1

