

Piotr Jabkowski



# Sondażowe zwierciadło europeizacji

Wydawnictwo Naukowe UAM

# Sondażowe zwierciadło europeizacji



UNIwersytet IM. ADAMA MICKIEWICZA W POZNANIU  
SERIA SOCJOLOGIA NR 85

Piotr Jabkowski

## Sondażowe zwierciadło europeizacji

Monitorowanie trendów europejskiej opinii publicznej  
w międzykrajowych sondażach porównawczych  
jako wyzwanie metodologiczne



POZNAŃ 2021

Recenzenci: prof. IFiS PAN dr hab. Zbigniew Sawiński  
dr hab. Michał Nowosielski

© Piotr Jabkowski 2021

This edition © Uniwersytet im. Adama Mickiewicza w Poznaniu,  
Wydawnictwo Naukowe UAM, 2021



Open Access book, distributed under the terms of the CC licence (BY-NC-ND,  
<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Publikacja powstała w ramach realizacji projektu badawczego pt. *Zwierciadło europeizacji: skumulowane zbiory danych sondaży międzykrajowych jako narzędzie monitorowania trendów europejskiej opinii publicznej*, finansowanego przez Narodowe Centrum Nauki (grant nr 2018/31/B/HS6/00403).

Redakcja i korekty: Ewa Dobosz

Projekt okładki, opracowanie techniczne i łamanie: Marcin Tyma

ISBN 978-83-232-3915-4 (PDF)

ISSN 0554-8225

DOI: 10.14746/amup.9788323239154

WYDAWNICTWO NAUKOWE UNIWERSYTETU IM. ADAMA MICKIEWICZA W POZNANIU

61-701 POZNAŃ, UL. FREDRY 10

[www.press.amu.edu.pl](http://www.press.amu.edu.pl)

Sekretariat: tel. 61 829 46 46, faks 61 829 46 47, e-mail: [wyd nauk@amu.edu.pl](mailto:wyd nauk@amu.edu.pl)

Dział Promocji i Sprzedaży: tel. 61 829 46 40, e-mail: [press@amu.edu.pl](mailto:press@amu.edu.pl)

Wydanie I – uzupełnione. Ark. wyd. 10,00. Ark. druk. 9,125

DRUK I OPRAWA: VOLUMINA.PL DANIEL KRZANOWSKI, SZCZECIN, UL. KS. WITOLDA 7-9

# Spis treści

WPROWADZENIE	7
<b>ROZDZIAŁ 1. EUROPEJSKIE DZIEDZICTWO SONDAŻOWE</b>	11
1. Podstawowa charakterystyka głównych europejskich sondaży porównawczych	11
1.1. Eurobarometer	11
1.2. European Social Survey	16
1.3. European Values Study	22
1.4. International Social Survey Programme	27
2. Co poza „wielką czwórką”?	32
2.1. European Election Studies	32
2.2. European Quality of Life Survey	35
3. Podsumowanie	36
<b>ROZDZIAŁ 2. SONDAŻOWE ZWIERCIADŁO PROCESU INTEGRACJI EUROPEJSKIEJ W PROJEKTACH EB, ESS, EVS, ISSP</b>	39
1. Międzyprojektowe zróżnicowanie treści pytań kwestionariuszowych	40
2. Metodologiczne ograniczenia dla porównań międzyprojektowych na przykładzie harmonizacji pytania o zaufanie do Parlamentu Europejskiego w projektach EB oraz ESS	46
2.1. Monotoniczne przekształcenie harmonizacyjne jako nadzieja na międzyprojektową porównywalność	47
2.2. Wyzwania harmonizacyjne na przykładzie pytania o zaufanie do Parlamentu Europejskiego	49
3. Podsumowanie	56
<b>ROZDZIAŁ 3. METODOLOGICZNY TYGIEL – ZRÓŻNICOWANIE PRAKTYK SONDAŻOWYCH W PROJEKTACH EB, ESS, EVS, ISSP</b>	59
1. Zróżnicowanie procedur badawczych jako wyzwanie dla porównywalności danych – analiza literatury przedmiotu	60
1.1. Sposoby definiowania badanych populacji	60
1.2. Rodzaje operatorów doboru prób badawczych	61
1.3. Schematy losowania próby	62
1.4. Typy sondażowych prób badawczych	64
1.5. Procedury selekcji respondentów docelowych w adresowych próbach osób	65

1.6. Procedury terenowej realizacji prób sondażowych . . . . .	68
1.7. Wskaźniki terenowej realizacji próby . . . . .	69
2. Źródła danych, zakres inwentaryzacji praktyk sondażowych oraz ocena jakości dokumentacji metodologicznej projektów EB, ESS, EVS, ISSP . . . . .	70
2.1. Baza empiryczna . . . . .	70
2.2. Charakterystyka źródeł danych wykorzystanych w metodologicznej inwentaryzacji projektów EB, ESS, EVS oraz ISSP . . . . .	72
2.3. Zakres metodologicznej inwentaryzacji projektów . . . . .	73
2.4. Ocena jakości dokumentacji metodologicznej projektów . . . . .	78
3. Porównawcze analizy zróżnicowania praktyk sondażowych w projektach EB, ESS, EVS, ISSP . . . . .	82
3.1. Sposoby definiowania badanych populacji . . . . .	82
3.2. Rodzaje operatorów doboru prób badawczych oraz schematy losowania jednostek . . . . .	84
3.3. Typy sondażowych prób badawczych oraz procedury selekcji respondentów docelowych w adresowych próbach osób . . . . .	86
3.4. Procedury terenowej realizacji prób sondażowych . . . . .	88
3.5. Wskaźniki terenowej realizacji prób sondażowych . . . . .	92
4. Podsumowanie . . . . .	95
<b>ROZDZIAŁ 4. OD ZRÓŻNICOWANIA PRAKTYK SONDAŻOWYCH DO ZRÓŻNICOWANIA JAKOŚCI PRÓBY . . . . .</b>	<b>97</b>
1. Wyzwania dla oceny jakości prób sondażowych w międzykrajowych projektach porównawczych o charakterze sondażowym . . . . .	98
1.1. Procedury oceny jakości sondażowych prób badawczych oraz ich ograniczenia w ewaluacji sondaży prowadzonej na dużą skalę . . . . .	98
1.2. Rola wewnętrznych kryteriów i zewnętrznych kryteriów oceny jakości prób badawczych w sondażowych projektach porównawczych . . . . .	100
2. Porównawcze analizy jakości prób sondażowych w wybranych projektach . . . . .	102
2.1. Międzyprojektowe oraz wzdłużczasowe zróżnicowanie jakości prób sondażowych . . . . .	102
2.2. Zróżnicowanie jakości prób sondażowych z uwagi na procedury doboru próby . . . . .	108
2.3. Zależności pomiędzy jakością prób sondażowych oraz wskaźnikami jej terenowej realizacji . . . . .	115
3. Podsumowanie . . . . .	119
<b>ZAKOŃCZENIE . . . . .</b>	<b>121</b>
<b>ANEKS 1. Odniesienia literaturowe do zbiorów danych oraz dokumentacji projektów EB, ESS, EVS oraz ISSP . . . . .</b>	<b>125</b>
Eurobarometer . . . . .	125
European Social Survey . . . . .	126
European Values Study . . . . .	127
International Social Survey Programme . . . . .	127
<b>BIBLIOGRAFIA . . . . .</b>	<b>131</b>
<b>SUMMARY (Reflecting Europeanisation. Monitoring trends in European public opinion in cross-national comparative surveys as a methodological challenge) . . . . .</b>	<b>145</b>

## Wprowadzenie

Od długiego już czasu w Europie realizowane są dobrej jakości projekty sondażowe o charakterze porównawczym. Taki stan rzeczy umożliwia monitorowanie zmian społecznych w długiej perspektywie czasowej oraz daje sposobność analizowania międzykulturowych podobieństw oraz odmienności w postawach i opiniach obywateli różnych europejskich państw. Z drugiej jednak strony, znacząca liczba projektów oraz ich relatywnie wysoka jakość, w porównaniu z badaniami prowadzonymi na innych kontynentach (por. Kołczyńska, Schoene 2018), napotyka w Europie na znaczącą przeszkodę w postaci daleko posuniętej różnorodności metodologicznej. Europejskie sondaże wykazują przy tym znaczne zróżnicowanie zarówno pod względem: (a) celów badawczych i poruszanych kwestii, (b) pokrycia poszczególnych krajów sondażami oraz (c) podejść metodologicznych wykorzystywanych w terenowej realizacji prób badawczych (Jabkowski, Kołczyńska 2020a). Co więcej, ze względu na brak metodologicznej koordynacji pomiędzy projektami, problemem okazuje się również często ekwiwalentność pomiarowa, albowiem zadawane respondentom pytania kwestionariuszowe wykazują w najlepszym przypadku jedynie fasadowe podobieństwo oraz charakteryzują się zróżnicowaną jakością pomiaru.

Przedmiotem analiz zaprezentowanych w monografii uczyniono postawy Europejczyków wobec procesu integracji europejskiej mierzone w ramach czterech paneuropejskich wielofalowych sondażowych projektów porównawczych, w tym badań Eurobarometru/Eurobarometer (EB), Europejskiego Sondażu Społecznego / European Social Survey (ESS), Europejskiego Sondażu Wartości / European Values Study (EVS) oraz Międzynarodowego Programu Badań Społecznych / International Social Survey Programme (ISSP). Choć dane uwzględnione w *Sondażowym zwiercadle europeizacji* obejmują okres od początku lat 80. XX wieku, to należy pamiętać, że historia europejskich międzykrajowych badań sondażowych sięga połowy lat 70. ubiegłego wieku i związana jest z przeprowadzeniem pierwszej edycji (realizowanego nieprzerwanie od 1974 roku) projektu standardowego Eurobarometru. Dla rozwoju europejskich sondaży porównawczych przełomowa



okazała się jednak pierwsza dekada XXI wieku, głównie za sprawą uruchomienia (wzorcowego pod względem metodologicznym) Europejskiego Sondażu Społecznego oraz znaczącego rozszerzenia zasięgu geograficznego pozostałych dwóch projektów będących przedmiotem analiz, tj. realizowanego od połowy lat 80. XX wieku (corocznego) projektu ISSP, jak również – prowadzonego od 1981 roku w cyklach 10-letnich – Europejskiego Sondażu Wartości. Z drugiej strony, początek XXI wieku okazał się również przełomowy dla procesu integracji europejskiej, ze względu na największe w historii rozszerzenie Unii Europejskiej o kraje Europy Środkowo-Wschodniej, a także serię bezprecedensowych kryzysów politycznych, społecznych i gospodarczych, z jakimi przyszło się mierzyć od samego początku nowego milenium.

Przyjmując za punkt wyjścia rozróżnienie na integrację społeczną oraz systemową (zob. Lockwood 1964; Archer 1996; Mouzelis 1997), przeprowadzone w *Sondażowym zwierciadle europeizacji* analizy postaw obywateli krajów Europy wobec procesu integracji europejskiej ograniczone zostały do dwóch podstawowych grup zagadnień. W obszarze integracji społecznej zawężono je do analizy wskaźników subiektywnej identyfikacji respondentów z Europą oraz Unią Europejską. Natomiast w wymiarze integracji systemowej ograniczono je do studiów nad legitymizacją instytucji UE, w tym do analizy zaufania do Unii Europejskiej oraz Parlamentu Europejskiego, a także deklarowanego wsparcia członkostwa kraju w UE. Co oczywiste, skupienie uwagi na tak wąsko zdefiniowanym obszarze problemowym pomija wiele istotnych oraz interesujących kwestii związanych z integracją europejską, jednak ograniczenie pola analiz było konieczne ze względu na znaczący nadmiar dostępnych danych sondażowych oraz zróżnicowanie podejść badawczych i pomiarowych stosowanych w międzykrajowych sondażach porównawczych.

Warto przy tym zauważyć, iż niejednorodność metodologiczna, konceptualizacyjna oraz operacjonalizacyjna poszczególnych przedsięwzięć badawczych utrudnia w sposób znaczący możliwość przeprowadzenia bezpośrednich międzyprojektowych porównań wyników. Istotnie bowiem, poza metodologicznie zorientowanymi metaanalizami nad niepełną realizacją próby badawczej (np. Groves, Peytcheva 2008; Peytcheva, Groves 2009), wpływem procedur selekcji respondentów docelowych na jakość pomiaru w adresowych próbach osób (np. Gaziano 2005; Krenzkle, Li, Rust 2010), zjawiskiem fałszowania danych poprzez duplikowanie wyników w bazach danych (Słomczynski, Powalko, Krauze 2017), studiami nad ogólną oceną jakości różnych typów prób badawczych (Kohler 2007), czy też ekwiwalencją pomiarową konstruktów latentnych (Karpiński 2016), niewiele jest w literaturze światowej analiz zorientowanych na porównywanie substancywnych wyników między projektami (por. Ruiter, Van Tubergen 2009; Hess 2017; Kropp 2017; Zmerli, Van der Meer 2017; Kropp 2018). Normą jest raczej analizowanie danych z konkretnego projektu, w oderwaniu od rezultatów uzyskiwanych

w innych sondażach, które to ewentualnie pojawiają się we wstępie lub w dyskusji wyników jako opisowe punkty odniesienia (np. Heath, Martin, Spreckelsen 2009; Quaranta 2018).

Tymczasem w obszarze europeizacji istnieje przynajmniej kilka wspólnych pól problemowych pozwalających na międzyprojektowe zestawienia wyników. Dla przykładu można wskazać, że analizy poczucia związku z UE, krajem lub regionem oparte na wynikach Eurobarometru (zob. Antonsich 2008, 2009; Antonsich, Holland 2012), pozostają komplementarne względem analiz wykorzystujących wyniki z innych projektów, np. studiów opartych na danych EVS porównujących międzykrajowe różnice w sile związku z instytucjami europejskimi oraz krajowymi (np. Arts, Halman 2006), czy też prac wykorzystujących wyniki ESS w studiach nad tożsamością europejską (np. Fligstein, Polyakova, Sandholtz 2012). Warto jednak zauważyć, że istniejący potencjał dla prowadzenia międzyprojektowych analiz pozostaje znacząco niewykorzystany głównie z powodu dość uzasadnionych obaw badaczy o międzyprojektową równoważność pomiaru wynikającą z odmienności projektowych specyfik. Oczywiście wyzwania związane z porównywalnością danych pochodzących z różnych projektów stanowią powszechną przeszkodę w badaniach sondażowych (De Jonge, Veenhoven, Kalmijn 2017), okazują się jednak szczególnie zniechęcające przy badaniu opinii publicznej w Europie. W konsekwencji tendencja do skupiania się na danych z pojedynczego projektu, daje konglomerat fragmentarycznych szkiców o różnej jakości, zamiast panoramicznego obrazu postaw i opinii wyrażanych wobec procesu europeizacji.

Głównym celem tej książki jest zaprezentowanie podstawowych wyzwań metodologicznych, z jakimi uporać muszą się badacze dążący do analizy danych pochodzących z realizacji wielu odmiennych międzykrajowych projektów sondażowych. Wśród takich wyzwań pojawiają się zarówno trudności mające swoje źródło w różnorodności problematyki badawczej oraz odmiennych celach stawianych w ramach poszczególnych projektów, jak też w specyficznych sposobach formułowania pytań, odmiennej metodologii pomiaru oraz zróżnicowanej jakości sondaży. Rozdział pierwszy przedstawia bogactwo projektów sondażowych prowadzonych na kontynencie europejskim. W sposób szczegółowy opisano cztery renomowane przedsięwzięcia sondażowe, tj. badania Eurobarometru, Europejskiego Sondażu Społecznego, Europejskiego Sondażu Wartości oraz Międzynarodowego Programu Badań Społecznych. Rozdział drugi poświęcono ukazaniu projektowych specyfik operacjonalizacji poszczególnych wskaźników mierzących postawy wobec procesu europeizacji oraz studiom nad wpływem sposobów sformułowania pytań kwestionariuszowych na uzyskiwane wyniki. W rozdziale trzecim zaprezentowano bogactwo procedur badawczych służących losowaniu oraz terenowej realizacji krajowych sondaży. Analizie poddano ograniczenia dla międzyprojektowych porównań wynikające ze sposobów definiowania badanych populacji, jak również

będące efektem wykorzystania odmiennych rodzajów operatów doboru jednostek oraz różnych typów prób badawczych, a także stosowania zróżnicowanych procedur terenowych. W ostatnim rozdziale pracy zaproponowano narzędzia pozwalające na ewaluację jakości sondaży przy wykorzystaniu informacji standardowo gromadzonych w większości międzykrajowych projektów porównawczych. Na podstawie zaproponowanych narzędzi przeanalizowano jakość ponad 1,4 tys. prób sondażowych z czterech projektów, co pozwoliło na zidentyfikowanie procedur terenowych, posiadających istotny wpływ na jakość pomiaru sondażowego. Wyniki przeprowadzonych analiz pozwoliły na sformułowanie rekomendacji ukazujących konieczność uwzględniania projektowych specyfik terenowej realizacji prób sondażowych w międzyprojektowej analizie danych.

## Europejskie dziedzictwo sondażowe

Celem pierwszego rozdziału pracy jest opis oraz charakterystyka czterech wiodących sondażowych projektów porównawczych realizowanych w krajach europejskich, tj. badań Eurobarometru, Europejskiego Sondażu Społecznego (European Social Survey), Europejskiego Sondażu Wartości (European Values Study) oraz Międzynarodowego Programu Badań Społecznych (International Social Survey Programme). W rozdziale omówiono również inne projekty sondażowe prowadzone nie tylko w krajach europejskich.

### 1. Podstawowa charakterystyka głównych europejskich sondaży porównawczych

#### 1.1. Eurobarometer

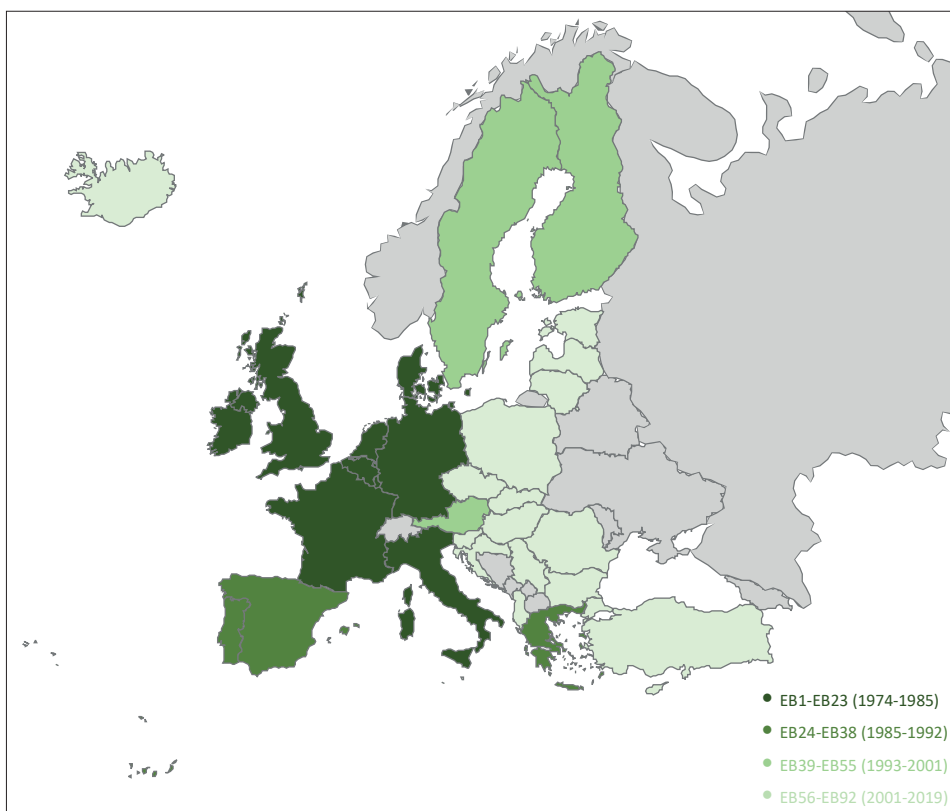
Pierwsze przymiarki do wprowadzenia cyklicznych pomiarów opinii publicznej w krajach członkowskich Wspólnoty Europejskiej zainicjowane zostały przez Komisję Wspólnot Europejskich (obecnie Komisję Europejską) już na początku lat 70. XX wieku. Parlament Europejski wyraził zainteresowanie pomysłem utworzenia systemu monitorującego europejską opinię publiczną, co zostało potwierdzone przeprowadzeniem dwóch pilotażowych sondaży wstępnych w 1971 oraz 1973 roku. Ojcowie założyciele EB, za których powszechnie uważani są Jacques-René Rabier oraz Wilhelmus Schuijt, określili główne zadania i cele stojące przed nowym wówczas projektem, podkreślając zarówno konieczność lepszego rozpoznania przez instytucje UE opinii publicznej w krajach członkowskich oraz społecznego odbioru prowadzonej polityki europejskiej, jak też konieczność wzajemnego poznania przez Europejczyków z różnych krajów wyrażanych przez siebie poglądów i opinii (Inglehart, Reif 1991). Pierwszy spośród corocznych wiosennych oraz jesiennych standardowych Eurobarometrów zrealizowano w roku 1974. Projekt trwa do dziś i jest najdłużej funkcjonującym europejskim międzykrajowym son-

dażem o charakterze porównawczym. Według stanu na koniec 2019 roku przeprowadzono łącznie 92 edycje EB w jego standardowej wersji.

Pomiary standardowego Eurobarometru uzupełniane są zarówno (a) Eurobarometrami specjalnymi, czyli pogłębionymi badaniami sondażowymi, dotyczącymi kwestii związanych z wybranymi obszarami działań inicjowanych przez Komisję Europejską lub przez inne instytucje UE, jak też (b) realizowanymi *ad hoc* (na podstawie wywiadów telefonicznych) sondażami w ramach *Flash Eurobarometer*, które to z kolei pozwalają na szybkie uzyskanie wyników dotyczących specyficznych tematów i grup docelowych (Höpner, Jurczyk 2015). Warto zauważyć, że przeprowadzono dotąd kilkaset pomiarów eurobarometrów „pogłębionych” oraz „migawkowych”. Niezwykłą zaletą wszystkich tych przedsięwzięć jest publiczny dostęp (poprzez strony Komisji Europejskiej) do ogromnego zasobu informacji zawierającego raporty ze wszystkich edycji Eurobarometrów standardowych, specjalnych oraz migawkowych (<https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm>), jak też pełny dostęp poprzez repozytorium GESIS (prowadzone przez Leibniz Institute for Social Science z siedzibą w Mannheim), do zbiorów danych wyników wszystkich Eurobarometrów standardowych oraz specjalnych (<https://zcat.gesis.org/webview/>).

Wydawać by się mogło, że publiczny charakter danych EB daje badaczom zainteresowanym europejską sferą publiczną nieograniczone wręcz możliwości prowadzenia studiów empirycznych pozwalających ‘cofnąć’ się w analizie społeczeństw krajów Unii Europejskiej o prawie pięć dekad wstecz. Nic bardziej mylnego, ponieważ z uwagi na dość swobodne podejście do wzdłużczasowej stabilności narzędzi pomiarowych, przejawiające się we włączaniu pewnych itemów oraz ich wyłączeniu z kwestionariuszy w kolejnych edycjach EB właściwie bez żadnego uzasadnienia (Signorelli 2012), a nawet zmienianiu treści poszczególnych pytań (Nissen 2014), Eurobarometry standardowe pozwalają – co najwyżej – na prowadzenie analiz obejmujących krótkookresową perspektywę kilku lub – w najlepszym przypadku – kilkunastu lat. Problem braku wzdłużnej ekwiwalentności pytań w EB podjęty zostanie w kolejnym rozdziale tej pracy.

Sondaże standardowego Eurobarometru prowadzone są obecnie we wszystkich krajach członkowskich Unii Europejskiej, krajach kandydujących, a także w Turcji i Tureckiej Republice Cypru Północnego (rys. 1). Warto przy tym zaznaczyć, że Wielka Brytania – reprezentowana we wszystkich edycjach projektu w latach 1974-2019 – po wyjściu tego kraju z UE nie będzie już więcej uwzględniana w pomiarze. Jak dotąd, we wszystkich 92 dotychczasowych rundach standardowego Eurobarometru, poza wspomnianą już Wielką Brytanią, udział wzięły także Belgia, Dania, Niemcy (przed zjednoczeniem w 1990 roku część zachodnia tego kraju), Francja, Holandia oraz Włochy.



Rys. 1. Zasięg geograficzny projektu EB (1974-2019)

Źródło: opracowanie własne.

Historycznie rzecz ujmując, włączanie kolejnych państw do sondażu EB związane było z procesem rozszerzenia Unii Europejskiej (Schmitt 2003). Dla przykładu, począwszy od 14. edycji EB przeprowadzonej jesienią 1980 roku, projekt zaczął być realizowany w Grecji, która dołączyła do Wspólnoty Europejskiej pierwszego dnia stycznia 1981 roku. Hiszpanię oraz Portugalię, które dołączyły do Wspólnoty z początkiem 1986 roku, zaczęto uwzględniać w badaniach standardowego EB od 24. edycji z jesieni 1985 roku. Począwszy od 34. jesiennej edycji EB z 1990 roku, pomiarem – po zjednoczeniu Niemiec – objęto obywateli Republiki Federalnej Niemiec zamieszkujących obszar dawnej Niemieckiej Republiki Demokratycznej. Austrię, Finlandię i Szwecję zaczęto uwzględniać w EB od 39. edycji zrealizowanej jesienią 1993 roku, dwa lata przed akcesją tych krajów do UE. Podobny mechanizm poszerzania zasięgu geograficznego EB zastosowano również w 2001 roku, w odniesieniu do krajów Europy Środkowo-Wschodniej, wstępujących do UE w efekcie kilku kolejnych rozszerzeń wspólnoty, odpowiednio w 2004, 2007 oraz

2013 roku (w 2001 roku do badań EB dołączyły także Turcja oraz Turecka Republika Cypru Północnego). Ostatnie zwiększenie zasięgu standardowego EB odbyło się w trakcie 67. edycji z 2007 oraz 76. edycji z 2011 roku, kiedy projekt zaczęto realizować w Islandii (od 2007) oraz Albanii i Serbii (od 2011 roku), tj. trzech krajach, które w 2009 roku złożyły wnioski akcesyjne do UE. Dokładne dane o udziale poszczególnych państw europejskich w kolejnych edycjach standardowego Eurobarometru przedstawione zostały w tabeli 1.

Tabela 1. Udział krajów europejskich w kolejnych edycjach standardowego Eurobarometru (1974-2019)

Symbol kraju <sup>(i)</sup>	EB1-EB23 (1974-1985)	EB24-EB38 (1985-1992)	EB39-EB55 (1993-2001)	EB56-EB66 (2001-2006)	EB67-EB92 (2007-2019)	Częstość udziału
AL					•	9
AT			•	•	•	51
BE	•	•	•	•	•	92
BG				•	•	37
CY				•	•	37
CYTCC <sup>(ii)</sup>				•	•	27
CZ				•	•	37
DE-E <sup>(iii)</sup>		•	•	•	•	59
DE-W <sup>(iv)</sup>	•	•	•	•	•	92
DK	•	•	•	•	•	92
EE				•	•	37
ES		•	•	•	•	69
FI			•	•	•	54
FR	•	•	•	•	•	92
GB-GBN <sup>(v)</sup>	•	•	•	•	•	92
GB-NIR <sup>(vi)</sup>	•	•	•	•	•	87
GR		•	•	•	•	79
HR				•	•	31
HU				•	•	37
IE	•	•	•	•	•	92
IS					•	9
IT	•	•	•	•	•	92
LT				•	•	37
LU	•	•	•	•	•	92
LV				•	•	37

Symbol kraju <sup>(i)</sup>	EB1-EB23 (1974-1985)	EB24-EB38 (1985-1992)	EB39-EB55 (1993-2001)	EB56-EB66 (2001-2006)	EB67-EB92 (2007-2019)	Częstość udziału
ME					•	16
MK					•	23
MT				•	•	37
NL	•	•	•	•	•	92
PL				•	•	37
PT		•	•	•	•	69
RO				•	•	37
RS					•	14
SE			•	•	•	51
SI				•	•	37
SK				•	•	37
TR				•	•	35

Uwagi: <sup>(i)</sup> Kody krajów zgodne ze standardem ISO 3166-1 alfa-2; <sup>(ii)</sup> Turecka Republika Cypru Północnego; <sup>(iii)</sup> RFN część wschodnia; <sup>(iv)</sup> RFN część zachodnia; <sup>(v)</sup> Wielka Brytania; <sup>(vi)</sup> Irlandia Północna

Źródło: opracowanie własne.

We wszystkich pomiarach standardowego Eurobarometru przeprowadzonych w okresie 1974-1992, aż do edycji oznaczonej numerem 40., sondażowe próby badawcze dobierane były z populacji osób w wieku 15 lat i więcej zamieszkujących na terenie danego kraju oraz posiadających jego obywatelstwo. Począwszy jednak od EB41 z 1993 roku, nastąpiła w tym względzie zmiana, gdyż w skład krajowych populacji zaliczono już nie tylko obywateli danego państwa, ale też wszystkie osoby w wieku 15 lat i więcej mieszkające w danym kraju posiadające obywatelstwo któregośkolwiek z krajów członkowskich. Fakt ten był zresztą podnoszony w literaturze przedmiotu jako jedno z ograniczeń dla prowadzenia wzdłużnych analiz na danych EB. Zmieniono bowiem zakres populacji objętej badaniami, ale nie uwzględniono tego w longitudinalnych zestawieniach danych (Nissen 2014).

Wielkość krajowych prób badawczych realizowanych w standardowych Eurobarometrach wynosi około 1000 respondentów, z wyjątkiem mniejszych krajów (takich jak Luksemburg oraz Malta), w których próba liczy mniej więcej 500 osób. Jeżeli chodzi o stosowane schematy doboru próby, to należy wskazać, że począwszy od jesiennej edycji EB z 1989 roku (oznaczonej numerem 32.) jednostki losowane są według wielostopniowego schematu warstwowego z zachowaniem probabilistycznej natury procesu doboru, podczas gdy we wcześniejszych edycjach standardowego Eurobarometru próby miały zarówno charakter losowy, jak i nielosowy. W części krajów stosowano nieprobabilistyczne schematy typu kwotowego



(z wagami nakładanymi na płeć, wiek i zawód), w innych natomiast adresowe próby osób (z komponentem losowej ścieżki), w pozostałych natomiast próby ściśle probabilistyczne (Nissen 2014). Co więcej, dla pierwszych siedmiu edycji Eurobarometru nie zachowały się żadne informacje odnośnie do stosowanych schematów doboru próby. Innymi słowy, we wczesnych edycjach EB panowała dość duża swoboda odnośnie do metodologicznych rygorów, jakie spełniać powinna próba sondażowa w badaniu porównawczym i dopiero po 1989 roku zdecydowano się na losowanie prób wyłącznie probabilistycznych.

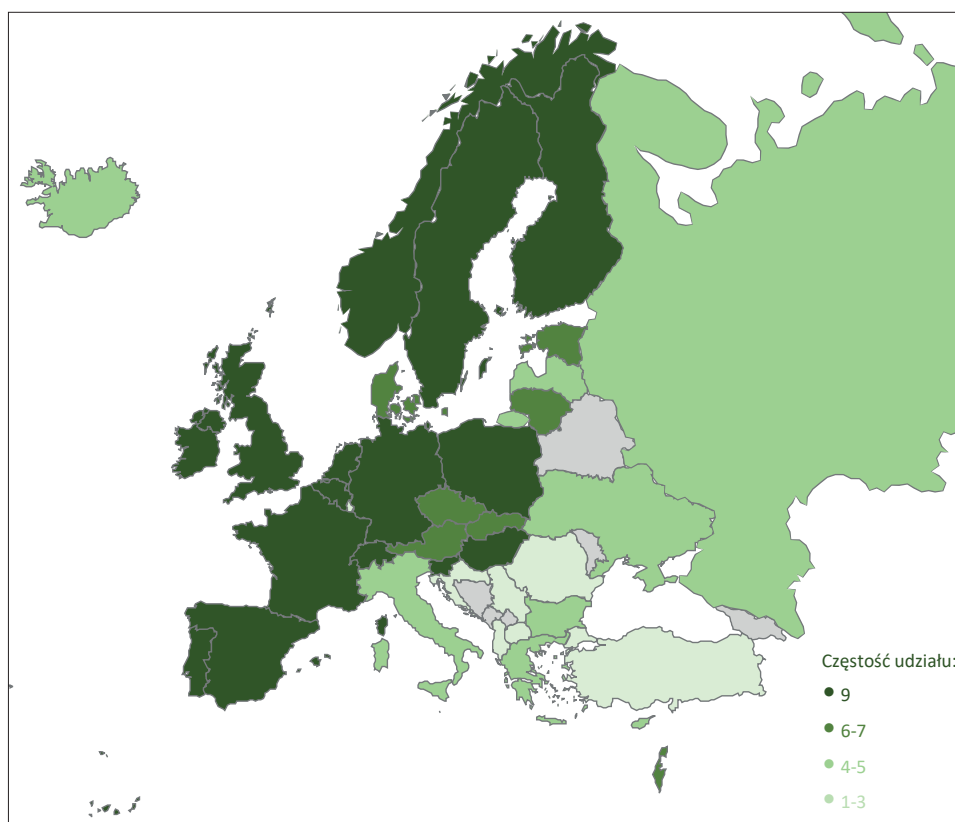
Warto również dodać, że realizacja wszystkich standardowych Eurobarometrów opierała się jak dotąd na wywiadach bezpośrednich prowadzonych przez ankierów w domach respondentów. Jedynym wyjątkiem od tej reguły był szwedzki pomiar EB42 z 1994 roku przeprowadzony przy pomocy wywiadu telefonicznego. Od 2004 roku zaczęto również (w krajach, w których było to możliwe) wdrażać procedury pomiaru wykorzystujące techniki wywiadów osobistych wspomaganých komputerowo (CAPI). Co istotne, kwestionariusze wywiadów przygotowywane były oryginalnie w dwóch wersjach językowych, tj. angielskiej oraz francuskiej (Harkness, Schoua-Glusberg 1998), a następnie tłumaczone na języki narodowe z wykorzystaniem uznanych metodologicznie procedur weryfikujących jakość translacji oraz międzykrajową ekwiwalentność treści pytań (Harkness 2003; Silber i in. 2018). W niektórych krajach, w tym w Belgii, Estonii, Finlandii, Hiszpanii, Luksemburgu, Łotwie, Malcie oraz w Turcji, kwestionariusze tłumaczone były nie tylko na języki urzędowe, ale też na inne języki mniejszości narodowych. Przykładowe pytania zadawane w ramach projektu EB służące pomiarowi społecznych postaw wobec procesu integracji europejskiej (w tym poczucia związku z UE i Europą, zaufania do Parlamentu Europejskiego oraz do UE, a także opinii o preferowanej głębokości integracji) omówione zostały w kolejnym rozdziale pracy. Do metodologicznych charakterystyk EB powrócimy jeszcze w rozdziale trzecim, omawiając międzyprojektowe zróżnicowanie wykorzystywanych praktyk badawczych.

## 1.2. European Social Survey

Geneza jednego z najbardziej uznanych metodologicznie międzykrajowych projektów porównawczych, jakim bez wątpienia jest Europejski Sondaż Społeczny, sięga lat 80. XX wieku, kiedy to Roger Jowell oraz Max Kaase podjęli w ramach Europejskiej Fundacji Nauki współpracę zmierzającą do wypracowania standardów realizacji porównawczych sondaży międzykrajowych (Breen, Healy 2016). Roger Jowell – pierwotnie związany z projektem International Social Survey Programme – poirytowany niską jakością metodologicznej dokumentacji badawczej, raportującej o krajowych realizacjach sondaży w ramach wczesnych fal projektu ISSP (Park, Jowell 1997), opublikował dokument zawierający dziesięć dyspozycji

oraz wymagań, jakie powinny zostać spełnione, by można było prowadzić metodologicznie uzasadnione analizy na zbiorach danych gromadzonych w sondażach realizowanych w wielu różnych krajach (Jowell 1998). Praca ta zaowocowała większą świadomością metodologów w zakresie jakości procedur wykorzystywanych w terenowej realizacji badań porównawczych oraz stała się podstawą dokumentu stawiającego podwaliny teoretyczne, metodologiczne oraz organizacyjne dla przyszłego konsorcjum realizującego projekt ESS (ESF 1999).

Zasięg geograficzny Europejskiego Sondażu Społecznego obejmuje większość krajów europejskich, w tym także Turcję oraz Izrael. Pomimo wysiłków zmierzających do objęcia przez ESS krajów całego kontynentu europejskiego, reprezentacja skrzywiona jest w kierunku państw należących do Unii Europejskiej.



Rys. 2. Zasięg geograficzny projektu ESS (2002-2018)

Źródło: opracowanie własne.

W przeprowadzonych obecnie dziewięciu edycjach ESS udział wzięło w sumie 38 różnych krajów, przy czym zbiór państw uczestniczących w kolejnych run-

dach projektu ulegał dość znaczącym zmianom. Najmniejsza liczba państw brała udział w pierwszej oraz siódmej rundzie ESS z 2002 oraz 2014 roku (o obu rundach uwzględniono 22 kraje), najliczniejsza była natomiast runda czwarta z 2010 roku, w której badaniami objętych zostało 31 różnych państw. Warto przy tym zaznaczyć, że jedynie 15 państw wzięło udział we wszystkich dotychczasowych edycjach ESS. Do grupy tej zalicza się również Polska (w której badania koordynowane są przez zespół socjologów afiliowanych w Instytucie Filozofii i Socjologii PAN w Warszawie) oraz Belgia, Finlandia, Francja, Hiszpania, Holandia, Irlandia, Niemcy, Norwegia, Portugalia, Słowenia, Szwajcaria, Szwecja, Węgry oraz Wielka Brytania. Dokładne zestawienie krajów uczestniczących w kolejnych dziewięciu rundach ESS zamieszczone zostały w tabeli 2 (obecnie trwają przygotowania do rundy 10., której terenowa realizacja rozpocznie się pod koniec 2020 roku).

Tabela 2. Udział krajów europejskich w kolejnych edycjach projektu ESS (2002-2018)

Symbol kraju <sup>(1)</sup>	ESS 1 2002	ESS 2 2004	ESS 3 2006	ESS 4 2008	ESS 5 2010	ESS 6 2012	ESS 7 2014	ESS 8 2016	ESS 9 2018	Częstość udziału
AL						•			•	2
AT	•	•	•	•	•		•	•	•	8
BE	•	•	•	•	•	•	•	•	•	9
BG			•	•	•	•			•	5
CH	•	•	•	•	•	•	•	•	•	9
CY			•	•	•	•			•	5
CZ	•	•		•	•	•	•	•	•	8
DE	•	•	•	•	•	•	•	•	•	9
DK	•	•	•	•	•	•	•		•	8
EE		•	•	•	•	•	•	•	•	8
ES	•	•	•	•	•	•	•	•	•	9
FI	•	•	•	•	•	•	•	•	•	9
FR	•	•	•	•	•	•	•	•	•	9
GR	•	•		•	•					4
HR				•	•				•	3
HU	•	•	•	•	•	•	•	•	•	9
IE	•	•	•	•	•	•	•	•	•	9
IL	•			•	•	•	•	•		6
IS		•				•		•	•	4
IT	•	•				•		•	•	5
LT				•	•	•	•	•	•	6

Symbol kraju <sup>(i)</sup>	ESS 1 2002	ESS 2 2004	ESS 3 2006	ESS 4 2008	ESS 5 2010	ESS 6 2012	ESS 7 2014	ESS 8 2016	ESS 9 2018	Częstość udziału
LU	•	•								2
LV			•	•			•		•	4
ME									•	1
NL	•	•	•	•	•	•	•	•	•	9
NO	•	•	•	•	•	•	•	•	•	9
PL	•	•	•	•	•	•	•	•	•	9
PT	•	•	•	•	•	•	•	•	•	9
RO			•	•						2
RS									•	1
RU			•	•	•	•		•		5
SE	•	•	•	•	•	•	•	•	•	9
SI	•	•	•	•	•	•	•	•	•	9
SK		•	•	•	•	•			•	6
TR		•		•						2
UA		•	•	•	•	•				5
UK	•	•	•	•	•	•	•	•	•	9
XK <sup>(ii)</sup>						•				1

Uwagi: <sup>(i)</sup> Kody krajów zgodne ze standardem ISO 3166-1 alfa-2; <sup>(ii)</sup> Kosowo

Źródło: opracowanie własne na podstawie [https://www.europeansocialsurvey.org/data/country\\_index.html](https://www.europeansocialsurvey.org/data/country_index.html).

Poza zebraniem kluczowych dla Europy danych społecznych w Europejskim Sondażu Społecznym realizuje się także ważne cele praktyczne, z których bodaj najważniejszym jest wypracowanie standardów przygotowania i realizacji naukowych badań. Po każdej rundzie publikowane są raporty dokumentujące metodologiczną warstwę projektu. Zawierają one bogaty zestaw informacji o operatach doboru próby, schematach losowania jednostek, typach prób badawczych, wskaźnikach terenowej realizacji próby, procedurach terenowych, a także o innych ważnych aspektach związanych z realizacją ESS. Dokumentacja wraz z bazami danych dostępna jest w domenie publicznej (<https://www.europeansocialsurvey.org/>).

W Europejskim Sondażu Społecznym wywiady prowadzone są w domach respondentów (przez doświadczonych ankieterów), przy wykorzystaniu standaryzowanych kwestionariuszy w wersji papierowej (PAPI) lub elektronicznej (CAPI). Badaną populację stanowią „osoby w wieku 15 lat i więcej zamieszkujące w prywatnych gospodarstwach domowych, niezależnie od narodowości, obywatelstwa, języka oraz

statusu prawnego” (ESS 2016: 6). Należy podkreślić, że ESS kładzie duży nacisk na standaryzację procedur badawczych pozwalających na uzyskanie danych o najwyższej jakości metodologicznej (Fitzgerald, Jowell 2010). Próby opierają się wyłącznie na schematach doboru zrandomizowanego (nie dopuszcza się możliwości realizacji prób kwotowych) oraz przyjmują jedną z pięciu możliwych form:

- (1) prostych lub warstwowych prób losowych dobieranych z operatów imiennych,
- (2) złożonych prób dobieranych z operatów imiennych,
- (3) wielostopniowych prób adresowych opartych na operatach gospodarstw domowych,
- (4) wielostopniowych prób adresowych opartych na operatach pozwalających wylosować adres budynku mieszkalnego,
- (5) wielostopniowych prób adresowych wykorzystujących dobór małych jednostek terytorialnych (Kohler 2007).

Każdy krajowy zespół koordynujący badania ESS odpowiedzialny jest za przedstawienie szczegółowego schematu doboru próby zgodnie z wytycznymi panelu ekspertów ds. próbkowania i ważenia danych oraz dostarczenie informacji umożliwiającej wyliczenie indywidualnych prawdopodobieństw losowania jednostek z populacji do próby badawczej (Kaminska, Lynn 2017). Minimalna efektywna wielkość próby ustalona została na poziomie 1500 osób (lub 800 w krajach o populacji liczącej do 2 milionów mieszkańców). W rzeczywistości próby badawcze są jednak liczniejsze, gdyż w ustalaniu ich wielkości uwzględnia się utratę precyzji wniosku, wynikającą z zastosowanego schematu doboru próby (Lynn i in. 2007). Dokładne metodologiczne charakterystyki projektu ESS przedstawione będą w trzecim rozdziale pracy.

Kwestionariusze ESS przyjmują charakter modułowy (tabela 3) i składają się z tematycznie powiązanych części, zawierających łącznie kilkaset pytań, tłumaczonych na języki narodowe z zachowaniem standardów umożliwiających zachowanie ekwiwalencji wielojęzycznych wersji narzędzia badawczego (Fitzgerald, Zavala-Rojas 2020). Mniej więcej połowa wszystkich itemów powtarzana jest w każdej rundzie projektu w ramach modułów podstawowych, pozostałe pytania stanowią składową modułów rotacyjnych, czyli tematycznie pogrupowanych sekcji, po jednej lub dwie na każdą rundę, niekiedy powtarzanych w kolejnych odsłonach projektu (Fitzgerald, Jowell 2010; Schnaudt i in. 2014). Warto zauważyć, że choć badanie ESS jest w swej naturze skoncentrowane na problematyce dotyczącej krajów europejskich, to zagadnienia społecznej oraz politycznej integracji w ramach UE nie były przedmiotem szczegółowych analiz w żadnym z dotychczasowych modułów rotacyjnych, a nieliczne pytania dedykowane bezpośrednio Unii Europejskiej „rozrzucone” były w kilku miejscach modułów podstawowych.

Tabela 3. Moduły tematyczne uwzględnione w kwestionariuszach kolejnych edycji ESS (2002-2018)

Nazwa modułu <sup>(i)</sup>	Edycje uwzględniające moduł
Media and social trust	wszystkie
Politics	wszystkie
Subjective well-being, social exclusion, religion, national and ethnic identity	wszystkie
Gender, year of birth and household grid	wszystkie
Socio demographics	wszystkie
Human values	wszystkie
Immigration	ESS1-2002 & ESS7-2014
Citizen involvement	ESS1-2002
Health and care seeking	ESS2-2004
Economic morality	ESS2-2004
Family work and well-being	ESS2-2004 & ESS5-2010
Timing of life	ESS3-2006 & ESS9-2018
Personal and social well-being	ESS3-2006 & ESS6-2012
Welfare attitudes	ESS4-2008 & ESS8-2016
Ageism	ESS4-2008
Justice	ESS5-2010
Democracy	ESS6-2012
Social inequalities in health	ESS7-2014
Public attitudes to climate change	ESS8-2016
Justice and fairness in Europe	ESS9-2018

Uwaga: <sup>(i)</sup> Zachowano oryginalne nazwy modułów.

Źródło: opracowanie własne na podstawie <https://www.europeansocialsurvey.org/data/module-index.html>.

Otwarty charakter baz danych, akademicki profil projektu oraz stosowanie najlepszych praktyk metodologicznych pozwala badaczom na prowadzenie licznych studiów zarówno nad wybranymi problemami zawartymi w podstawowych oraz tematycznych modułach kwestionariuszy wywiadów, jak też nad metodologią badań sondażowych w ogóle. Główne obszary analiz metodologicznych obejmują refleksję nad schematami doboru próby oraz typami prób badawczych (Häder, Lynn 2007; Menold 2014; Eckman, Koch 2019; Jabkowski, Cichocki 2019), procedurami umożliwiającymi zwiększanie wskaźników realizacji próby oraz minimalizowanie błędu niepełnej realizacji próby (Billiet i in. 2007; Kreuter, Kohler 2009; Matsuo i in. 2010; Sztabiński, Sztabiński, Przybysz 2012), efektem wynikającym z prowadzenia pomiaru w oparciu o różne techniki badawcze, *mixed-mode design* (Jäckle, Roberts, Lynn 2010; Revilla 2010; Vannieuwenhuyze, Loosveldt,

Molenberghs 2010), efektem ankierskim (Loosveldt, Beullens 2013; Beullens, Loosveldt 2016), źródłami błędów pomiaru oraz sposobami ich redukcji (Saris, Satorra, Coenders 2004; Coromina, Saris 2009; Saris, Revilla 2016), czy też międzykrajową inwariancją pomiarową różnych konstruktów latentnych (Reeskens, Hooghe 2008; Davidov 2010; Pokropek 2012).

Badanie ESS uznawane jest powszechnie za laboratorium najlepszych praktyk sondażowych (Słomczyński 2004), umożliwiających uzyskanie (w porównaniu z innymi projektami o charakterze międzykrajowym) wyników pomiaru o relatywnie wyższej jakości (Kohler 2007). To przekonanie o metodologicznej wyjątkowości ESS nie jest pozbawione podstaw, gdyż w 2005 roku ESS otrzymał nagrodę Kartezjusza za wybitne projekty i działania popularyzujące naukę. Było to pierwsze wyróżnienie kiedykolwiek przyznane projektowi o charakterze badań sondażowych (Fitzgerald, Jowell 2010), natomiast w 2013 roku uhonorowany został prestiżową nagrodą European Research Infrastructure Consortium (Ságvári, Messing 2019).

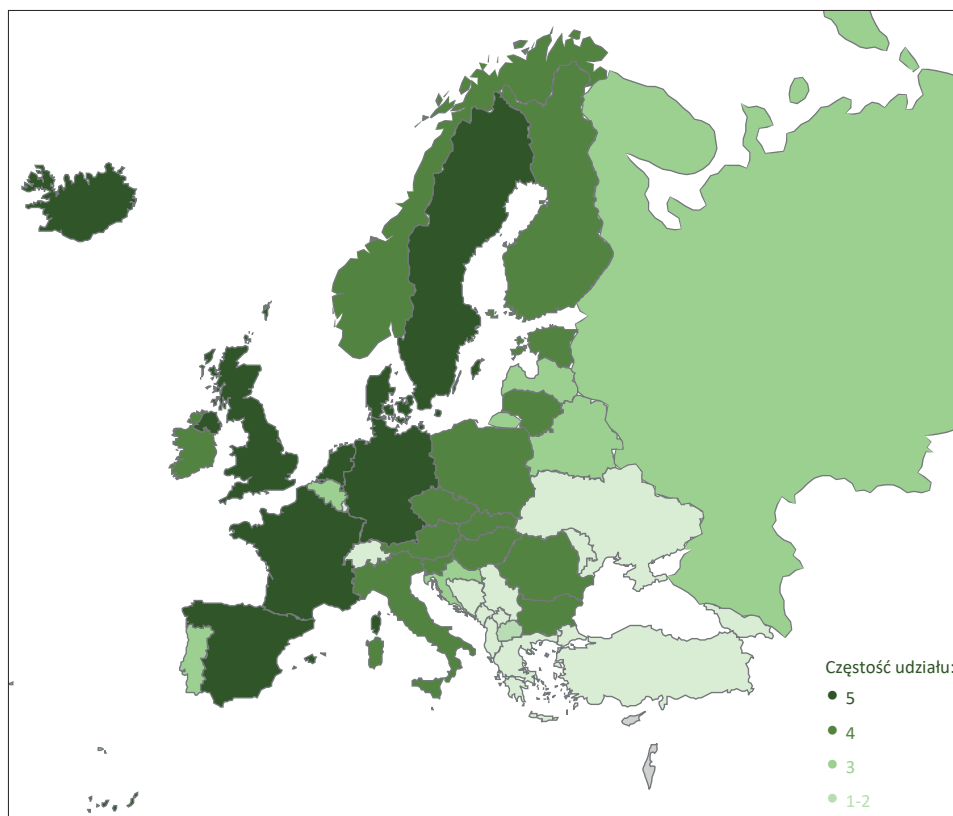
### 1.3. European Values Study

European Values Study należy do najstarszych (po Eurobarometrze) międzykrajowych sondaży porównawczych realizowanych na kontynencie europejskim. Jego podstawowym celem jest mapowanie wartości, przekonań oraz postaw wyrażanych przez obywateli różnych krajów europejskich w odniesieniu do ważnych dla nich sfer życia prywatnego i społecznego. Historia projektu sięga lat 70. XX wieku, kiedy to Jan Kerkhofs z Katolickiego Uniwersytetu w Leuven oraz Ruud de Moor z holenderskiego Uniwersytetu w Tilburgu podjęli współpracę zmierzającą do utworzenia nowego projektu badań międzykrajowych, który w sposób systematyczny oraz zgodny z uznanymi praktykami metodologicznymi monitorowałby międzykrajowe podobieństwa oraz odmienności we wzorach wartości uznawanych przez Europejczyków za ważne (Luijckx i in. 2016). W szczególności byli oni zainteresowani odpowiedzią na trzy główne pytania badawcze, a mianowicie:

- (1) jaki jest stopień międzykrajowego podobieństwa we wzorach wartości wyznawanych przez obywateli różnych państw,
- (2) jaka jest dynamika zmian tych wzorów i w jakim kierunku zmiany te podążają, a także
- (3) jakie mogą być konsekwencje owych międzykrajowych odmienności oraz przemian we wzorach wartości Europejczyków dla jedności kontynentu (Gelissen, Halman 2018).

Zamierzeniem Kerhofsza oraz Moora była realizacja sondaży zarówno w krajach Europy Zachodniej, jak też Środkowo-Wschodniej, jednakże z uwagi na uwarunkowania geopolityczne nie udało się – przygotowując pierwszą edycję EVS –

nawiązać współpracy z badaczami reprezentującymi kraje znajdujące się wówczas za żelazną kurtyną. Realizację pierwszej fali projektu poprzedziły intensywne prace koordynowane przez zespół zrzeszony w ramach European Values System Study Group, który we współpracy z agencjami badawczymi z trzech krajów, tj. Francji (Faits et Opinions, Paris), Republiki Federalnej Niemiec (Institut für Demoskopie Allensbach) oraz Wielkiej Brytanii (Gallup UK, London) przeprowadził serię jakościowych oraz ilościowych studiów pilotażowych, pozwalających na opracowanie wzorcowego kwestionariusza wywiadu oraz wytycznych dla terenowej realizacji pierwszej edycji EVS.



Rys. 3. Zasięg geograficzny projektu EVS (1981-2017)

Źródło: opracowanie własne.

Na rysunku 3 przedstawiono informacje na temat pokrycia krajów europejskich sondażami realizowanymi w ramach EVS (łącznie we wszystkich jego rundach). W tabeli 4 zawarto z kolei szczegółowe informacje o udziale poszczególnych państw w każdej z kolejnych edycji projektu. Warto zauważyć, że choć zbiór



krajów biorących udział w pięciu odsłonach EVS ulegał (z rundy na rundę) dość znaczącym modyfikacjom, to każdy z krajów Europy (poza księstwami Andory, Lichtensteinu, Monako oraz Republiką San Marino) uczestniczył przynajmniej w jednej edycji EVS. Co ciekawe, badania oparte na kwestionariuszach EVS prowadzone były też okazjonalnie poza Europą (np. w dwóch pierwszych rundach zrealizowano je w Kanadzie oraz USA). Zaangażowany w to przedsięwzięcie Ronald Inglehart doprowadził następnie do ustanowienia Światowego Sondażu Wartości (*World Values Survey*), jako projektu niezależnego od EVS, choć wywodzącego się z niego w sposób bezpośredni (Halman, Vloet 1994).

Tabela 4. Udział krajów europejskich w kolejnych edycjach projektu EVS (1981-2017)

Symbol kraju <sup>(1)</sup>	EVS 1 1981	EVS 2 1990	EVS 3 1999	EVS 4 2008	EVS 5 2017	Częstość udziału
AL				•	•	2
AM				•	•	2
AT		•	•	•	•	4
AZ					•	1
BA				•		1
BE	•	•	•			3
BG		•	•	•	•	4
BY			•	•	•	3
CH				•	•	2
CZ		•	•	•	•	4
CY				•		1
DE	•	•	•	•	•	5
DK	•	•	•	•	•	5
EE		•	•	•	•	4
ES	•	•	•	•	•	5
FI		•	•	•	•	4
FR	•	•	•	•	•	5
GE				•	•	2
GR			•	•		2
HR			•	•	•	3
HU		•	•	•	•	4
IE	•	•	•	•		4
IS	•	•	•	•	•	5

Symbol kraju <sup>(i)</sup>	EVS 1 1981	EVS 2 1990	EVS 3 1999	EVS 4 2008	EVS 5 2017	Częstość udziału
IT	•		•	•	•	4
LT		•	•	•	•	4
LU			•	•		2
LV		•	•	•		3
MD				•		1
ME				•		1
MK				•		1
MT	•	•	•	•		4
NL	•	•	•	•	•	5
NO	•	•		•	•	4
PL		•	•	•	•	4
PT		•	•	•		3
RO		•	•	•	•	4
RS				•	•	2
RU			•	•	•	3
SE	•	•	•	•	•	5
SI		•	•	•	•	4
SK		•	•	•	•	4
TR			•	•		2
UA			•	•		2
UK	•	•	•	•	•	5
XK <sup>(ii)</sup>				•		1

Uwagi: <sup>(i)</sup> Kody krajów zgodne ze standardem ISO 3166-1 alfa-2; <sup>(ii)</sup> Kosowo

Źródło: opracowanie własne, dla edycji 1-4 na podstawie Luijkx i in. 2016.

Biorąc pod uwagę cele projektu zorientowane na analizę przemian zachodzących w systemie wartości, które – co do swej istoty – nie podlegają tak dynamicznym zmianom w czasie jak opinie, czy też nawet postawy, kolejne edycje EVS planowane były z częstotliwością mniej więcej raz na dekadę. Posiada to swoje oczywiste konsekwencje dla analiz wzdluznych, gdyż – inaczej niż w projektach Eurobarometru oraz Europejskiego Sondażu Społecznego, prowadzonych ze zdecydowanie większą częstotliwością – dane uzyskane w badaniu EVS pozwalają na analizę zmian zachodzących jedynie w długiej perspektywie czasowej. Co więcej, problematyka skoncentrowana na społecznej oraz systemowej integracji w ramach

UE nie stanowiła nigdy trzonu zainteresowania badaczy realizujących EVS. Koncentrowano się za to na wartościach odnoszonych do życia respondenta, rodziny, pracy, religii, polityki i społeczeństwa, choć – co warto podkreślić – w kwestionariuszach kolejnych edycji zawarto również kilka pytań dedykowanych postawom wobec UE. Dla przykładu, pytanie o zaufanie wyrażane wobec Unii Europejskiej po raz pierwszy pojawiło się drugiej edycji EVS z 1990 roku, z kolei pytanie o postawy wobec pogłębiania procesu integracji europejskiej zadane było w dwóch edycjach projektu z 2008 oraz 2017 roku.

Pierwsza edycja EVS, datowana na rok 1981, objęła swym zasięgiem wszystkie kraje należące do Wspólnoty Europejskiej (z wyjątkiem Grecji), uwzględniono jednak Hiszpanię, która wówczas nie należała do UE. Badania terenowe koordynowane były w całości przez Gallup UK z siedzibą w Londynie oraz przeprowadzone w okresie od marca do maja 1981 roku. W każdym z krajów biorących udział w pierwszej edycji EVS dobrano reprezentatywne próby liczące około 1000 osób, dodatkowo wylosowano również 200 jednostek do tzw. *próby boosterowej* (Listhaug 1984), celem nieproporcjonalnego (w stosunku do rozkładu populacyjnego) zwiększenia udziału osób w wieku od 18 do 24 roku życia (zbiory danych wynikowych zwiernają wagi dostosowujące to celowe skrzywienie do danych populacyjnych).

Realizacja drugiej edycji EVS zbiegła się z okresem rewolucyjnych przemian zachodzących w Europie Środkowo-Wschodniej po 1989 roku. Przygotowania do tej rundy rozpoczęto już jednak w 1986 roku serią pilotażowych studiów we Francji, Hiszpanii, Republice Federalnej Niemiec, Wielkiej Brytanii, a poza Europą również w USA. Rezultatem tych badań było opracowanie wzorcowej wersji kwestionariusza w języku angielskim, który następnie poddano tłumaczeniu na języki urzędowe poszczególnych państw. Zasadnicza część badań zrealizowana została na przełomie 1990 i 1991 roku we wszystkich krajach Wspólnoty Europejskiej (z wyjątkiem Grecji), w krajach skandynawskich, Austrii, w dziesięciu krajach Europy Środkowo-Wschodniej, w tym również w Polsce, a także w byłej Niemieckiej Republice Demokratycznej. Badania opierały się na próbach dobieranych (w każdym kraju z osobna) z populacji osób w wieku 18 lat i więcej. Warto przy tym zaznaczyć, że w niektórych państwach próba przyjmowała charakter losowy, w innych natomiast zdecydowano się na nielosowe próby kwotowe z wagami nakładanymi na płeć i wiek.

Trzecia odsłona EVS zrealizowana została na przełomie 1999 oraz 2000 roku. Badania koordynowane były przez Uniwersytet w Tilburgu, natomiast za terenową realizację badań odpowiedzialne były krajowe instytuty badawcze. Wywiady przeprowadzono z wykorzystaniem technik ankierskich na próbach dobieranych w sposób losowy z populacji osób w wieku 18 lat i więcej (nie stosowano górnych ograniczeń wiekowych). W badaniu uczestniczyły wszystkie kraje, które

wzięły udział we wcześniejszych edycjach, do projektu dołączyło przy tym siedem nowych państw, w tym Białoruś, Chorwacja, Luksemburg, Grecja, Rosja, Turcja i Ukraina.

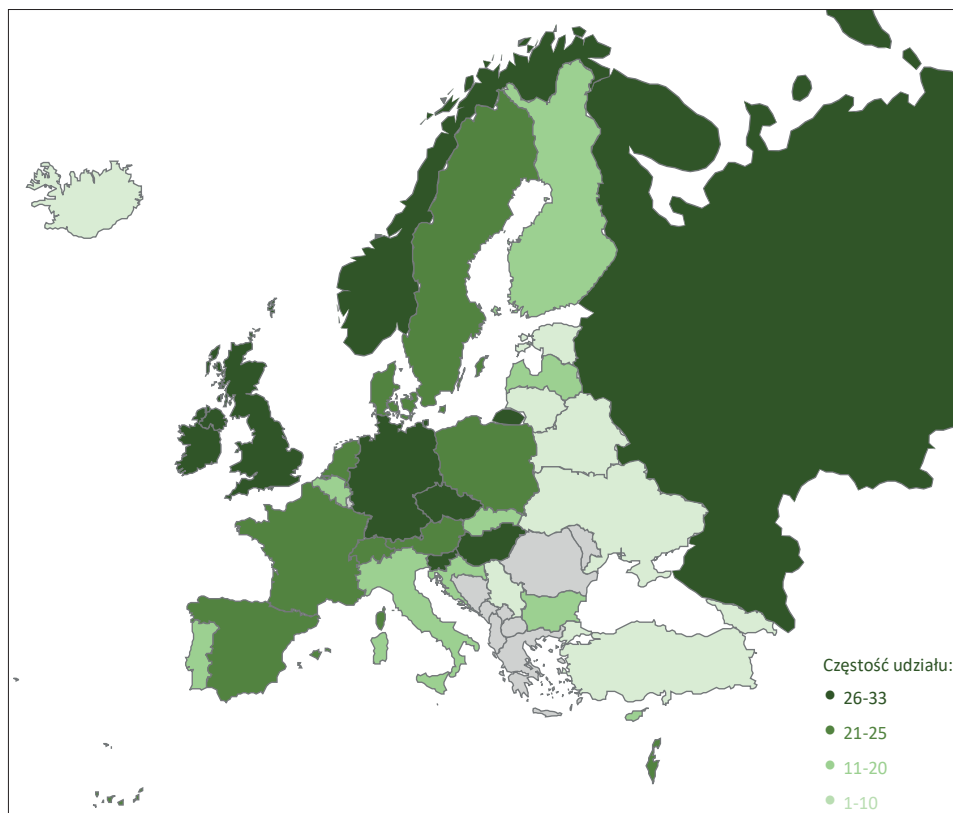
Badaniami czwartej rundy EVS z 2008 roku objęto już praktycznie całą Europę. Zdecydowano się również na rewizję stosowanych dotąd podejść metodologicznych, czego efektem było wprowadzenie najlepszych uznanych standardów realizacji międzykrajowych sondaży porównawczych, jakie wówczas były możliwe (próby miały posiadać już wyłącznie charakter losowy, w ustalaniu minimalnych wielkości prób badawczych zaczęto uwzględniać efekt schematu losowania, ulepszo procedury wyliczania wag, wprowadzono procedury kontrolujące pracę ankieterów). Interesującą inwentaryzację najważniejszych usprawnień metodologicznych oraz organizacyjnych przyjętych w edycji EVS z 2008 roku odnalazć można w opracowaniu (Luijckx i in. 2016). Najnowsza, piąta odsłona EVS z 2017 roku podążyła w pełni za najlepszymi praktykami realizacji międzykrajowych badań porównawczych, choć – co warto zaznaczyć – wzięło w niej udział o wiele mniej państw niż w rundzie z 2008 roku.

Wyczerpujących informacji na temat badań EVS wraz z bogatą bazą literaturową zawierającą odnośniki do pozycji wykorzystujących dane projektu w analizach zorientowanych zarówno na problemy metodologiczne oraz substancyjne odnaleźć można na stronie dedykowanej EVS (<https://europeanvaluesstudy.eu>). Zbiory danych wynikowych wraz z dokumentacją zdeponowane zostały w repozytorium GESIS (<https://zacam.gesis.org/webview/index/en/ZACAT/ZACAT.c.ZACAT/European-Values-Study-EVS-/fCatalog/Catalog5>).

#### **1.4. International Social Survey Programme**

International Social Survey Programme (ISSP) jest znanym międzykrajowym projektem sondażowym realizowanym corocznie, począwszy od połowy lat 80. XX wieku. Pomysł na ustanowienie cyklicznych pomiarów integrujących różne krajowe sondaże zrodził się w wyniku współpracy podjętej pomiędzy dwoma znanymi ośrodkami badawczymi, tj. niemieckim Zentrum für Umfragen, Methoden, und Analysen oraz amerykańskim National Opinion Research Centre Univeristy of Chicago (Skjåk 2010). Oba ośrodki realizowały wówczas krajowe generalne sondaże społeczne (ALLBUS w Niemczech oraz GSS w USA). Pod koniec 1983 roku do współpracy włączyli się również przedstawiciele brytyjskiego National Centre for Social Research oraz australijskiego Research School of Social Science Australian National University, którzy podjęli wysiłki zmierzające do ustanowienia i przeprowadzenia w 1985 roku pierwszej edycji ISSP (Smith 2008). Badaniami objęto łącznie sześć krajów, tj. Australię, Austrię, Niemcy (część zachodnia), USA, Wielką Brytanię (bez Irlandii Północnej) oraz Włochy. W 1989 roku James W. Davis oraz

Roger Jowell – uznawani za ojców założycieli ISSP – nieco zaskoczeni wzrastającą popularnością projektu postanowili podjąć dalsze wysiłki, zmierzające do opracowania trwałych procedur, które miały za zadanie sprostać rosnącym wyzwaniom metodologicznym i organizacyjnym, wynikającym z rosnącej różnorodności i złożoności tego przedsięwzięcia. Z czasem popularność ISSP wzrosła jeszcze bardziej, a obecnie (stan na wrzesień 2020) projekt obejmuje 42 kraje członkowskie z Afryki, Australii i Oceanii, Azji, Europy, Ameryki Północnej i Południowej (<http://w.issp.org/members/member-states/>).



Rys. 4. Zasięg geograficzny projektu ISSP w Europie (1985-2017)

Źródło: opracowanie własne.

Zbiór krajów europejskich objętych badaniami ISSP zaprezentowany został na rys. 4 oraz w tabeli 5. Warto jednak pamiętać, że w różnych edycjach tego projektu udział wzięło w sumie kilkadziesiąt krajów położonych na wszystkich kontynentach. A zatem, choć kraje europejskie dominują, to liczna jest również reprezentacja

Tabela 5. Udział krajów europejskich w kolejnych edycjach projektu ISSP (1985-2017)

Symbol kraju <sup>(0)</sup>	ISSP 1985-1990	ISSP 1991-2000	ISSP 2001-2010	ISSP 2011-2017	Częstość udziału
AT	.	.	.	.	25
BA	.				1
BE			.	.	13
BG		.	.	.	18
CH	.	.	.	.	24
CY		.	.		11
CZ		.	.	.	26
DE	.	.	.	.	33
DK		.	.	.	21
EE			.	.	3
ES		.	.	.	25
FI		.	.	.	18
FR		.	.	.	21
GE				.	4
HR	.		.	.	13
HU	.	.	.	.	29
IE	.	.	.	.	24
IL	.	.	.	.	27
IS			.	.	7
IT	.	.	.	.	16
LT			.	.	9
LV		.	.	.	19
NL	.	.	.	.	25
NO	.	.	.	.	28
PL	.	.	.	.	22
PT		.	.	.	15
RS	.				1
RU		.	.	.	27
SE		.	.	.	24
SK		.	.	.	18
SI	.	.	.	.	28
TR			.	.	8
UA			.		2
UK	.	.	.	.	33

Uwagi: <sup>(0)</sup> Kody krajów zgodne ze standardem ISO 3166-1 alfa-2

Źródło: opracowanie własne.

państw azjatyckich. Co więcej, do krajów biorących systematyczny udział w ISSP zaliczają się także Stany Zjednoczone, Australia i Nowa Zelandia. Jeżeli chodzi o pokrycie Europy, to we wszystkich dotychczasowych 33 edycjach ISSP z lat 1985-2017 (dla których we wrześniu 2020 roku udostępniono bazy danych oraz dokumentację projektową) udział wzięły wyłącznie Niemcy oraz Wielka Brytania, przy czym do krajów o najczęstszym udziale zaliczyć można także Czechy, Irlandię, Norwegię, Rosję oraz Słowenię. Polska uwzględniona była w pomiarze ISSP już w 1987 roku (kierownikiem badań był wówczas Bogdan W. Mach), chociaż w sposób systematyczny pojawiała się w projekcie w latach 1991-2015. Do 2010 roku polska część badań ISSP realizowana była pod kierunkiem Bogdana Cichomskiego (Sawiński, Zieliński 2015). W latach 2011-2015 polskie badania ISSP afiliowane były przy Instytucie Studiów Społecznych Uniwersytetu Warszawskiego oraz kierowane przez Marcina W. Zielińskiego i Tomasza Jerzyńskiego. Po 2016 roku projekt ISSP nie był w Polsce kontynuowany. Polska oficjalnie nie opuściła jednak konsorcjum projektowego i następane badania pod kierunkiem M. W. Zielińskiego i T. Jerzyńskiego zaplanowane są na rok 2022.

Przechodząc do krótkiej metodologicznej charakterystyki badań terenowych w ramach ISSP, należy powiedzieć, że nie podlegają one tak ścisłej centralizacji, jak w wypadku omówionych wcześniej projektach EB, ESS oraz EVS. Skutkuje to znaczącym międzykrajowym zróżnicowaniem wykorzystywanych praktyk sondażowych. Dla przykładu, chociaż w każdym z państw uczestniczących w kolejnych edycjach ISSP badaną populację tworzyły osoby w wieku 18 lat i więcej (ISSP 2017), to jednocześnie w wielu krajowych sondażach wprowadzane były specyficzne dolne (w zakresie od 14 do 21 lat) oraz górne punkty odcięcia dla wieku respondentów (w zakresie od 65 do 94 lat). Wywiady realizowane były w większości przy pomocy technik ankierskich opartych na wywiadach bezpośrednich, prowadzonych w domach badanych osób (techniki PAPI lub CATI), jednakże w części krajów wykorzystywano wywiady telefoniczne (CATI), ankiety mailowe i internetowe (CAWI), a nawet ankiety roznoszone z prośbą o samodzielne ich wypełnienie przez respondentów. Niejednokrotnie stosowano również rozwiązania mieszane, tj. część pytań zadawana była w trakcie wywiadu bezpośredniego, natomiast zmienne społeczno-demograficzne inwentaryzowane były z wykorzystaniem wybranej techniki ankietowej (Jabkowski, Kołczyńska 2020a). Wielkość krajowych prób badawczych pozostawała przy tym dość znacząco zróżnicowana i wynosiła od kilkuset do kilku tysięcy respondentów, przy czym w większości krajów stosowano wielostopniowe próby adresowe, rzadziej próby imienne, choć niejednokrotnie zdarzały się również próby nielosowe. Dodatkowo, w celu ograniczenia kosztów terenowej realizacji badań, kwestionariusze z ISSP dołączane były do innych krajowych badań lub też dwie edycje projektu łączone były ze sobą i realizowane są w ramach jednej rundy. Doskonałym tego przykładem jest realizacja ISSP w Niemczech, gdzie dwie różne edycje ISSP

włączane były raz na dwa lata do jednej rundy niemieckiego Generalnego Sondażu Społecznego ALLBUS (ISSP 2013: 38-39). Krajowe zespoły koordynujące realizację badań posiadały również dużą swobodę w decydowaniu o włączaniu lub wyłączeniu pewnych pytań z kwestionariuszy wywiadów, o ile tylko poinformowały o tym badaczy z innych państw (ISSP 2019). Z drugiej strony należy mieć na względzie, że w połowie lat 90. XX wielu, m.in. za sprawą zaangażowania się w projekt Rogera Jowella (późniejszego inicjatora badań Europejskiego Sondażu Społecznego), ISSP wyznaczył nowe standardy w zakresie raportowania o krajowych procedurach doboru próby oraz terenowej realizacji sondaży, stając się w tym względzie pionierem innowacyjności metodologicznej (Park, Jowell 1997). Do charakterystyki procedur terenowych wykorzystywanych w projekcie ISSP powrócimy jeszcze w trzecim rozdziale tej pracy.

Każda kolejna edycja ISSP poświęcona była specyficznym modułom tematycznym (tabela 6), niekiedy powtarzanym w kolejnych rundach, z częstotliwością mniej więcej raz na 10 lat. Poza społeczno-demograficznymi charakterystykami badanych osób, pytania kwestionariuszowe zadawane respondentom w ramach konkretnych modułów tematycznych nie pokrywały się z pytaniami zadawanymi w innych modułach. A zatem, chociaż pomiar ISSP realizowany był corocznie (nieprzerwanie od 1985 roku), to uzyskiwane dane pozwalają raczej na prowadzenie analiz wzdłużnych w długiej perspektywie czasowej, nie nadają się jednak do monitorowania trendów z roku na rok. W tym sensie, pomimo

Tabela 6. Moduły tematyczne uwzględnione w ISSP (1985-2017&2018-2020)

Nazwa modułu <sup>(i)</sup>	Liczba wystąpień	Edycje uwzględniające moduł
Role of government	5	1985, 1990, 1996, 2006, 2016
Social inequality	5	1987, 1992, 1999, 2009, 2019 <sup>(ii)</sup>
Family and changing gender roles	4	1988, 1994, 2002, 2012
Work orientations	4	1989, 1997, 2005, 2015
Religion	4	1991, 1998, 2008, 2018 <sup>(ii)</sup>
Environment	4	1993, 2000, 2010, 2020 <sup>(ii)</sup>
Social networks	3	1986, 2001, 2017
National identity	3	1995, 2003, 2013
Citizenship	2	2004, 2014
Leisure time and sports	1	2007
Health and health care	1	2001

Uwaga: <sup>(i)</sup> Zachowano oryginalne nazwy modułów, <sup>(ii)</sup> Dane przygotowywane do upublicznienia lub moduły w trakcie realizacji

Źródło: opracowanie własne na podstawie <http://w.issp.org/data-download/by-topic/>.



corocznej realizacji, projekt ISSP podobny jest bardziej do prowadzonego raz na dekadę badania EVS aniżeli do pomiarów EB oraz ESS. Warto zaznaczyć, że problematyka społecznej oraz systemowej integracji w ramach Unii Europejskiej nie była przedmiotem szczegółowych i dogłębnych analiz w żadnym z dotychczasowych modułów ISSP, co nie jest szczególnym zaskoczeniem, zważając na globalny charakter projektu wykraczający poza Europę. Pytania odnoszące się wprost do UE odnaleźć można jedynie w module pt. *National identity*, powtórzonym trzykrotnie w 1995, 2003 oraz 2013 roku, w którym pojawia się pytanie o preferowany przez respondentów model władzy ulokowany bądź to na poziomie krajowym (rządów poszczególnych państw), bądź też ponadnarodowym, tj. Unii Europejskiej. Dokładną charakterystykę tego pytania odnaleźć można się w kolejnym rozdziale pracy.

## 2. Co poza „wielką czwórką”?

### 2.1. European Election Studies

Niezwykle interesującym, choć zdecydowanie mniej rozpoznawalnym na mapie europejskich sondaży o charakterze porównawczym, jest projekt European Election Studies (EES). Posiada on bogatą tradycję sięgającą końca lat 70. XX wieku i związany jest z 5-letnimi cyklami wyborczymi do Parlamentu Europejskiego. Najnowsza edycja EES z 2019 roku, była 9. z kolei odsłoną tego projektu (wcześniejsze rundy przeprowadzono w latach 1979, 1984, 1989, 1994, 1999, 2004, 2009 oraz 2014). Badanie EES posiada swoją stronę internetową (<http://europeanelectionstudies.net/>), zawierającą podstawowe informacje o projekcie oraz dostęp do baz danych wynikowych. Metodologiczne charakterystyki tego badania przedstawione zostały w kilku ciekawych opracowaniach dedykowanych wybranym jego edycjom (Schmitt i in. 1997; Van der Kolk i in. 1997; Van der Eijk i in. 2002; Schmitt i in. 2009).

Projekt EES skoncentrowany jest na problematyce partycypacji obywatelskiej w państwach członkowskich UE, decyzjach wyborczych podejmowanych w wyborach do Parlamentu Europejskiego, diagnozie społecznych postaw wobec polityki UE, ocenie efektywności polityki w wybranych obszarach, europejskiej sferze publicznej, jak również na opiniach dotyczących dalszego pogłębiania procesu integracji europejskiej (De Vries 2017). Na projekt EES składają się zarówno reprezentatywne sondaże powyborcze (*voter study*), jak też inne komponenty, np. analizy kontekstowe programów wyborczych partii uczestniczących w wyborach do PE (*euromanifesto study*), badania realizowane z kandydatami w wyborach do PE, europośłami, deputowanymi krajowych parlamentów oraz członkami partii

politycznych (*elite study*), a także analizy treści przekazów medialnych głównych nadawców w krajach członkowskich UE (*media study*).

Sondaże Voter Study prowadzone są na populacji wyborców i stanowią główny komponent projektu EES obecny prawie we wszystkich jego rundach (z wyjątkiem 1984 roku). Pomiar pełni funkcję studiów powyborczych realizowanych możliwie najszybciej po wyborach do Parlamentu Europejskiego. Metodologia projektu ulegała jednak znacznym modyfikacjom w czasie, co utrudnia wzdlużne porównania wyników. Poniżej zaprezentowano główne charakterystyki metodologiczne kolejnych edycji Voter Study zrealizowanych w ramach EES. Materiały przygotowano na podstawie następujących publikacji: van der Eijk, Oppenhuis i Schmitt (1993); Schmitt i in. (1997); van der Eijk i in. (1999); Schmitt i in. (2009); Egmond i in. (2013); Schmitt i in. (2016); Schmitt i in. (2019).

## **Podstawowe charakterystyki kolejnych edycji Voter Study zrealizowanych w ramach EES**

### **EES Voter Study 1979, 1989, 1994**

Trzy pierwsze edycje Voter Study dołączane były do różnych fal Eurobarometru. W konsekwencji, pomimo iż projekt w swym założeniu miał diagnozować postawy wyborców bezpośrednio po wyborach do Parlamentu Europejskiego, to część pytań zadawana była jeszcze przez wyborami. Dla przykładu, w 1994 roku pytania z projektu EES stanowiły zarówno komponent specjalnego powyborczego Eurobarometru (EB41.1), jak też rozrzucone były po różnych kwestionariuszach z wielu standardowych Eurobarometrów przeprowadzonych zarówno przed wyborami (EB40-1993, EB41-1994), jak i po nich (EB42-1994). Podobną strategię badawczą przyjęto również w 1979 oraz 1989 roku. Warto zauważyć, że w 1984 roku nie zrealizowano Voter Study, a jedynie przeanalizowano programy partyjne (Euromanifesto).

### **EES Voter Study 1999**

Projekt zrealizowano w 15 krajach członkowskich, przy czym głównymi fundatorami badań EES z 1999 roku była Holenderska Rada ds. Badań Naukowych, Uniwersytet w Amsterdamie, Uniwersytet w Mannheim oraz Trinity College w Hartford. W przeciwieństwie do wcześniejszych edycji EES, badań zrealizowanych w 1999 nie włączono w cykl Eurobarometru, ale przeprowadzono jako niezależny projekt badawczy oparty na wywiadach telefonicznych. Liczebności prób badawczych ustalono proporcjonalnie do wielkości populacji poszczególnych państw.

### **EES Voter Study 2004**

Pomiar przeprowadzono w 24 z 25 państw członkowskich (nie uwzględniono Malty). Z uwagi na brak centralnego finansowania projektu nie udało się we wszyst-

kich krajach przeprowadzić wywiadów bezpośrednio po wyborach do Parlamentu Europejskiego. Wielkość krajowych prób badawczych podlegała znaczącym różnicowaniom. Co więcej, w dziewięciu krajach terenowa realizacja opierała się na wywiadach telefonicznych, w czterech na ankietach rozsyłanych mailem, natomiast w pozostałych jedenastu na osobistych wywiadach kwestionariuszowych prowadzonych przez ankieterów. Znaczące różnicowanie technik zbierania danych wynika z faktu realizacji badań przez 25 różnych krajowych instytucji badawczych.

### **EES Voter Study 2009**

Badania EES z 2009 roku uzyskały finansowanie z projektu PIRDEU (Providing an Infrastructure for Research on Electoral Democracy in the European Union), koordynowanego przez przedstawicieli Robert Schuman Centre for Advanced Studies. Badania objęły 27 państw członkowskich, natomiast wielkość krajowych prób badawczych wynosiła około 1000 osób. Dane zbierane były z wykorzystaniem technik wywiadu telefonicznego (CATI), z wyjątkiem Bułgarii, Czech, Estonii, Litwy, Łotwy, Polski, Rumunii, Słowacji i Węgier, gdzie 70% wywiadów zrealizowano w oparciu o techniki wywiadu bezpośredniego, a jedynie 30% telefonicznie. Kwestionariusz wywiadu zawierał te same pytania we wszystkich krajach. Dla zachowania porównywalności wyników z wcześniejszymi rundami EES powtórzono znaczną większość itemów z wcześniejszych edycji.

### **EES Voter Study 2014**

Pomiar EES z 2014 roku przeprowadzony został we wszystkich 28 krajach członkowskich w kooperacji z Parlamentem Europejskim (badania sfinansowano ze źródeł Volkswagen Foundation). Krajowe próby badawcze liczyły około 1100 respondentów, z wyjątkiem Malty i Luksemburga (500 osób) oraz Wielkiej Brytanii (1300 respondentów, z czego 300 osób stanowiła próba w Irlandii Północnej). Badania terenowe przeprowadzone zostały przy pomocy wywiadów bezpośrednich wspomaganých komputerowo (CAPI) oraz zrealizowane przez TNS Opinion. Kwestionariusz zawierał mniej więcej 250 pytań, w znaczniej części tych samych, które zadano respondentom w 2009 roku.

### **EES Voter Study 2019**

Badanie zrealizowane zostało we wszystkich 28 krajach członkowskich UE przez konsorcjum Gallup International. Była to pierwsza edycja EES, w której dane zebrane zostały z wykorzystaniem technik *online*, poza Maltą i Cyprem, gdzie zdecydowano się na wykorzystanie wywiadów telefonicznych. We wszystkich krajach próba badawcza podlegała stratyfikacji z uwagi na płeć, wiek, region zamieszkiwania oraz typ miejscowości. Wielkości krajowych prób badawczych wynosiły mniej

więcej 1000 osób, z wyjątkiem Cypru, Malty i Luksemburga, w których przebadano 500 osób. Kwestionariusz zawierał ponad 100 itemów, w większości zbieżnych z poprzednimi edycjami projektu.

Studia powyborcze prowadzone w ramach EES stanowią bez wątpienia przykład niezwykle interesujących porównawczych badań sondażowych, dających możliwość opisania krajowych systemów politycznych w państwach Unii Europejskiej. Niemniej jednak znaczącym ograniczeniem dla międzykrajowej oraz wzdłużczasowej porównywalności wyników okazuje się zróżnicowanie podejść metodologicznych pomiędzy kolejnymi rundami projektu, a w ramach tej samej rundy projektu pomiędzy poszczególnymi krajami. Pomimo tej różnorodności udało się wypracować zestaw trzonowych pytań powtarzanych w kolejnych edycjach EES dających możliwość prowadzenia analiz wzdłużnych.

## 2.2. European Quality of Life Survey

Europejskie Badanie Jakości Życia (EQLS) realizowane jest w cyklach czteroletnich pod auspicjami Europejskiej Fundacji na rzecz Poprawy Warunków Życia i Pracy (Eurofound), czyli agencji Unii Europejskiej mającej za zadanie realizację – poza EQLS – również dwóch innych projektów sondażowych, w tym badań European Working Condition Survey (EWCS) oraz European Company Survey (ECS). Wszystkie te trzy projekty obejmują swym zasięgiem terytorialnym kraje członkowskie UE, kraje kandydujące oraz potencjalnych kandydatów do UE, a także niektóre państwa stowarzyszone w ramach Europejskiego Stowarzyszenia Wolnego Handlu (EFTA). Głównym celem sondaży nadzorowanych przez Eurofund jest dostarczenie danych o wysokiej jakości metodologicznej, pozwalających m.in. na diagnozowanie warunków życia ludności (w projekcie EQLS), warunków pracy (w badaniu EWCS) oraz sytuacji przedsiębiorców (w sondażu ECS). Każde z trzech badań oparte jest na reprezentatywnych próbach losowanych z populacji obejmujących odpowiednio: dorosłych obywateli, osoby pracujące oraz przedsiębiorców, przy czym badanie EQLS – posiadające charakter diagnozy jakości życia ludności – prowadzone jest na reprezentatywnych próbach losowanych z populacji „wszystkich osób w wieku 18 lat i więcej, zamieszkujących na terytorium państw biorących udział w badaniach” (EQLS 2013).

Warto zaznaczyć, że EQLS nie zawiera pytań dedykowanych wprost procesowi integracji europejskiej, skupia się bowiem na inwentaryzacji obiektywnych warunków życia obywateli krajów europejskich, w odniesieniu do ich pracy, dochodów, edukacji, miejsca zamieszkania, rodziny oraz zdrowia, jak również na subiektywnej samocenie badanych odnośnie do ich indywidualnych warunków życia oraz dokonywanej przez respondentów ogólnej ocenie zadowolenia z życia. Metodo-

logia badań EQLS podlega ścisłej kontroli oraz procesowi standaryzacji, na wzór Europejskiego Sondażu Społecznego. Dla przykładu, próby dobierane są wyłącznie w sposób losowy pod nadzorem panelu ekspertów mających za zadanie wybór optymalnego dla każdego kraju schematu losowania. Terenową realizacją zajmuje się Kantar Public, przy czym wywiady prowadzone są z wykorzystaniem technik ankierskich PAPI lub CAPI. Zdecydowana większość krajowych prób badawczych ma charakter wielostopniowy, z komponentem losowej ścieżki (Jabkowski 2019). Protokoły realizacyjne nie przewidują możliwości substytucji jednostek niedostępnych (EQLS 2013). Wielkość krajowych prób badawczych podlega przy tym znacznym wahaniom. W krajach o mniej licznych populacjach, pomiarem objętych zostaje zazwyczaj około 600-700 respondentów, standardowa wielkość krajowej próby wynosi około 2000 osób, a w Niemczech dochodzi nawet do 3000 osób.

Dotąd zrealizowano cztery rundy tego projektu (w latach 2003, 2007/2008, 2011/2012 oraz 2016). W pierwszej odsłonie EQLS wzięło udział 27 krajów członkowskich UE (uwzględniając w tym kraje włączone do UE podczas rozszerzenia z 2004 oraz 2007 roku) oraz Turcja. W drugiej rundzie uwzględniono 31 krajów (27 członkowskich UE oraz Chorwację, Macedonię Północną, Turcję oraz Norwegię), w trzeciej rundzie 34 kraje (poza 27 krajami UE również Chorwację, Czarnogórę, Islandię, Kosowo, Macedonię Północną, Serbię i Turcję), natomiast w najnowszej czwartej rundzie badaniami objęto 33 kraje (28 krajów członkowskich oraz Albanie, Czarnogórę, Kosowo, Macedonię Północną, Serbię i Turcję). Dokumentacja oraz wyniki projektu EQLS dostępne są za pośrednictwem repozytorium UK Data Archive (<https://beta.ukdataservice.ac.uk>).

### 3. Podsumowanie

Niezwykłe bogactwo europejskich międzykrajowych sondaży o charakterze porównawczym sprawia, że przedstawione wyliczenie projektów nie jest wyczerpujące. Do innych ciekawych przedsięwzięć metodologicznych prowadzonych w Europie należy zaliczyć chociażby projekty porównawcze mające na celu agregację danych z krajowych sondaży wyborczych, np. Comparative National Election Project (Gunther i in. 2015) oraz Comparative Study of Electoral Systems (Klingemann 2009). Cechuje je już jednak zdecydowanie mniejsza międzykrajowa koordynacją metodologiczną niż omówione wcześniej sondaże porównawcze. Co więcej, skoncentrowane są one w ogromnej mierze na specyficznych politycznych kwestiach krajowych i przypisują znacznie mniejszą wagę międzykrajowej ekwiwalencji wyników.

Czytelnika zainteresowanego projektami porównawczymi prowadzonymi nie tylko w Europie, ale również na innych kontynentach, odesłać można do dwóch ciekawych opracowań poświęconych tej tematyce (Słomczyński 2004; Słomczyński, Tomescu-Dubrow, Jenkins 2016). Znaleźć tam można opis kilkunastu interesujących przedsięwzięć o charakterze porównawczym, w tym takich projektów sondażowych jak: Afrobarometer (<https://www.afrobarometer.org/>), Americas Barometer (<https://www.vanderbilt.edu/lapop/>), *Arab Barometer* (<https://www.arabbarometer.org/>), Balkan Barometer (<https://www.rcc.int/balkanbarometer/>), Caucasus Barometer (<https://caucasusbarometer.org/en/>), Latinobarometro (<https://www.latinobarometro.org/lat.jsp>), Global Barometer (<https://www.globalbarometer.net/>), Pew Global Attitudes and Trends (<https://www.pewresearch.org/global/>), czy też wreszcie badania najbardziej znanego w tej grupie World Values Survey (<http://www.worldvaluessurvey.org/wvs.jsp>), którego europejskim odpowiednikiem jest omówiony wcześniej projekt Europejskiego Sondażu Wartości (EVS).



## Sondażowe zwierciadło procesu integracji europejskiej w projektach EB, ESS, EVS, ISSP

Międzynarodowe porównawcze badania sondażowe były przez lata jednym z głównych narzędzi rozpoznawania społecznych postaw i opinii wyrażanych przez obywateli państw członkowskich wobec Unii Europejskiej (Hobolt, Vries 2016). O ile jednak na poziomie państw narodowych badania opinii społecznej stanowiły tylko jeden z wielu punktów odniesienia dla zorientowanej krajowo debaty publicznej, o tyle rola przekraczających krajowe granice badań sondażowych w kształtowaniu dyskursu na poziomie ponadnarodowym była już zdecydowanie bardziej znacząca. Wyniki badań Eurobarometru należały przy tym, przez wiele lat, do jedynych wiarygodnych źródeł wiedzy o opiniach i postawach obywateli poszczególnych państw członkowskich wyrażanych w stosunku do UE.

Eurobarometr pełnił przy tym nie tylko funkcje użytkowe dla instytucji UE, mając za zadanie monitorowanie zmian postrzegania UE i jej polityki, ale stanowił również pionierski wysiłek metodologiczny w śledzeniu długo- i krótkoterminowej dynamiki zmian opinii publicznej w perspektywie porównawczej. Należy bowiem pamiętać, że pierwsze przymiarki do cyklicznych badań sondażowych w krajach UE podjęto już na początku lat 70. XX wieku, a projekt EB w swojej pełnej odsłonie z pomiarem przeprowadzonym wiosną oraz jesienią każdego roku ruszył w 1974 roku. Praktycznie do końca lat 80. XX wieku EB nie miał konkurencji na polu międzynarodowych badań porównawczych i pomimo formułowanych obecnie zastrzeżeń co do warsztatu badawczego i standardów metodologicznych (Nissen 2014) oraz niskiej jakości dokumentacji projektowej (Jabkowski, Kołczyńska 2020a), stanowił bez wątpienia źródło inspiracji dla wielu innych międzynarodowych badań porównawczych (Moschner 2014). Dopiero pojawienie się akademicko zorientowanego Europejskiego Sondażu Społecznego, stawiającego cele nie tylko o charakterze substancywnym, ale mającego na względzie również wyzwania metodologiczne w dążeniu do uzyskania międzynarodowej porównywalności danych sondażowych (Lynn i in. 2007; Fitzgerald, Jowell 2010), ukazało kruchość metodologiczną międzynarodowych projektów



porównawczych z lat 70. i 80., a w niektórych przypadkach nawet tych prowadzonych jeszcze w latach 90. XX wieku. Obecne standardy metodologiczne i wynikające z tego zasady raportowania o procedurach terenowej realizacji próby sprawiają, że projekt EB wydaje się już nieco anachroniczny, o niskim standardzie metodologicznym (Jabkowski 2019). Z drugiej jednak strony wyniki badań EB wykorzystywane są nadal w studiach nad dynamiką procesu integracji europejskiej, gdyż większość europejskich projektów sondażowych o charakterze porównawczym nie ma tak długiej tradycji i nie służy bezpośrednio potrzebom diagnozowania integracji politycznej oraz systemowej w ramach UE. Mimo wszystko, projekty takie jak ESS, EVS, czy też nawet ISSP stanowią ważne źródło informacji na temat integracji europejskiej oraz opinii publicznej w krajach członkowskich, uzupełniając poznawcze i metodologiczne słabości Eurobarometru.

Rozdział poświęcony jest analizie międzyprojektowego zróżnicowania sposobów formułowania konkretnych pytań kwestionariuszowych, odnoszących się do procesu integracji europejskiej oraz analizie wpływu tych zróżnicowań na uzyskiwane wyniki pomiaru. Przedstawione w rozdziale analizy nie wyczerpują podjętej problematyki w całości, niemniej jednak pozwalają na wielostronne spojrzenie na problemy pojawiające się w studiach nad procesem europeizacji prowadzonych na podstawie danych z różnych sondażowych badań międzykrajowych. Przyjmując przy tym rozróżnienie na społeczną oraz systemową integrację na poziomie europejskim (Archer 1996), przedmiotem zaprezentowanych analiz uczyniono pytania dotyczące trzech podstawowych wymiarów europeizacji, tj. odczuwanego związku z Europą oraz z Unią Europejską, zaufania instytucjonalnego do Parlamentu Europejskiego oraz do UE, a także postaw wobec dalszego pogłębiania procesu integracji europejskiej.

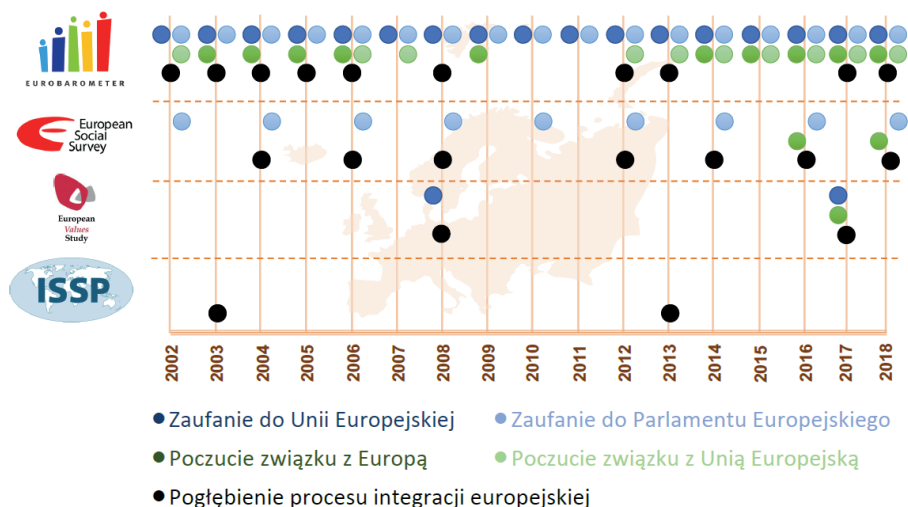
## **1. Międzyprojektowe zróżnicowanie treści pytań kwestionariuszowych**

Próba odnalezienia takich samych lub podobnych pytań, które w różnych projektach sondażowych mierzyłyby społeczne opinie i postawy wobec procesu integracji europejskiej, nie jest zadaniem łatwym (dla przykładu pytanie o poczucie związku z Unią Europejską zadawane jest wyłącznie w Eurobarometrze). O ile bowiem EB w swojej naturze skoncentrowany jest – prawie wyłącznie – na problematyce Unii Europejskiej, o tyle w pozostałych projektach zagadnienia z tym związane stanowią, w najlepszym przypadku, wątek poboczny. Punkty wspólne należą do rzadkości i dotyczą raczej pojedynczych pytań niż baterii itemów, czy też całych bloków tematycznych. Przeglądając dokumentację projektową oraz bazy danych projektów EB, ESS, EVS oraz ISSP, zidentyfikować można zaledwie kilka podobnych pytań dotyczących Unii Europejskiej. Dotyczą one:

- (1) zaufania do Unii Europejskiej (EB, EVS),
- (2) zaufania do Parlamentu Europejskiego (EB, ESS),
- (3) poczucia związku z Europą (EB, ESS, EVS), oraz
- (4) postaw wobec pogłębiania procesu europejskiej integracji (EB, ESS, EVS, ISSP).

Wyliczenie to nie jest oczywiście zupełne – w wielu projektach odnaleźć można podobne lub nawet takie same pytania dotyczące zaufania uogólnionego, zaufania do instytucji krajowych i międzynarodowych, skale ludzkich wartości oraz inne pytania – jeżeli chodzi jednak o kwestie europejskości, to wiele więcej wspólnych pytań nie udało się odnaleźć.

Występowanie w kolejnych edycjach czterech porównywanych projektów tych samych lub podobnych pytań dotyczących kwestii związanych z UE pokazane zostało na rysunku 5. Wybrany zakres czasowy (2002-2018) obejmował 34 wiosenne i jesienne edycyjne Eurobarometru (do EB57 do EB90), dziewięć pierwszych rund Europejskiego Sondażu Społecznego (ESS1-2002 – ESS9-2018), dwie z pięciu zrealizowanych edycji Europejskiego Sondażu Wartości (EVS4-2008 oraz EVS5-2017), a także siedemnaście fal International Social Survey Programme, przy czym wyłącznie w dwóch rundach ISSP poświęconych tożsamości narodowej (z 2003 oraz 2013 roku) znajdowały się pytania dotyczące postaw respondentów wobec procesu integracji Europejskiej.



Rys. 5. Występowanie wybranych pytań w kolejnych edycjach projektów (2002-2018)

Źródło: opracowanie własne na podstawie Jabkowski i Cichocki (2020).

Warto zaznaczyć, że problemem dla międzyprojektowej porównywalności danych nie jest wyłącznie brak lub obecność wspólnych pytań w różnych projektach, ale także występowanie w obrębie jednego projektu znaczących fluktuacji w odniesieniu do kontekstu, w jakim określone pytania są zadawane, jak rów-

niez arbitralne decyzje badaczy o włączaniu pewnych pytań do kwestionariuszy oraz ich wyłączeniu z wywiadów w kolejnych rundach tych samych projektów. Dla przykładu w badaniu ESS pytanie o poczucie związku z Europą pojawiło się dopiero w ósmej rundzie z 2016 roku, a w badaniu EVS odpowiednie pytanie o związek z Europą włączone zostało do kwestionariusza wywiadu jedynie w ostatniej piątej fali projektu datowanej na 2017 rok. Jeszcze większą labilnością cechuje się – wydawać by się mogło najbardziej stabilny z uwagi na swoją długą historię – projekt Eurobarometru. Pytania pojawiają się i znikają w kolejnych edycjach w zasadzie bez żadnego uzasadnienia, czasami zmieniana jest forma, a nawet treść pytania, nie wspominając choćby o zmianie umiejscowienia poszczególnych pytań w kwestionariuszach wywiadów. Dobitym tego przykładem jest pytanie o poczucie związku z Europą oraz z Unią Europejską. Te dwa na pozór proste pytania pojawiają się w różnych konfiguracjach: często zadawane są razem wszystkim respondentom, czasami następuje podział próby i połowa respondentów odpowiada na pytania o związek z Europą, druga zaś o związek z UE, pojawiają się nawet oddzielnie lub nie są zadawane wcale. Z punktu widzenia ekwiwalentności pomiaru sondażowego zobrazowana powyżej różnorodność ogranicza nie tylko możliwość prowadzenia międzyprojektowych porównań we wspólnych czasowo pomiarach, ale nawet wzdluzczasowych zestawień wyników w ramach jednego projektu (Nissen 2014).

Porównania międzyprojektowe utrudnione są jednak nie tylko z powodu braku ciągłości w zadawaniu pewnych pytań w kolejnych odsłonach danego projektu, czy też z powodu zróżnicowanego umiejscowienia poszczególnych pytań w kwestionariuszach wywiadów, ale przede wszystkim z odmiennych sposobów formułowania konkretnych pytań oraz definiowania kategorii odpowiedzi. Nie jest wielkim zaskoczeniem, że w ramach każdego projektu badawczego istnieje tendencja do zadawania własnych pytań oraz definiowania formatu odpowiedzi, służącego zaspokojeniu konkretnych potrzeb badaczy. Przykłady tej różnorodności, w odniesieniu do wybranych pytań mierzących społeczne postawy wobec procesu integracji europejskiej w projektach EB, ESS, EVS oraz ISSP, w tym pytań o zaufanie do UE oraz Parlamentu Europejskiego, poczucie związku z Europą oraz UE, a także pytań o opinie dotyczące pogłębiania procesu integracji europejskiej, przedstawiono w tabeli 7.

Najbardziej jaskrawym przykładem różnic w sposobach zadawania pytań wydaje się z pozoru proste pytanie, mające na celu zdiagnozowanie społecznych postaw wobec procesu pogłębiania integracji europejskiej. Pomiędzy projektami zaobserwować można zarówno znaczące niepodobieństwo w treści pytań, jak również w układzie dostępnych respondentom kafeterii odpowiedzi. Dla przykładu w Eurobarometrze pytanie dotyczące postaw wobec procesu pogłębiania integracji mierzy w swoim założeniu opinie wobec aktualnej „prędkości” tego procesu, czy też obecnego „tempa” budowania zjednoczonej Europy (w EB zadaje się również

Tabela 7. Zróżnicowanie treści pytań dotyczących wybranych aspektów procesu integracji europejskiej w projektach EB, ESS, EVS, ISSP

Treść pytania	Format odpowiedzi
<b>Zaufanie do Unii Europejskiej</b>	
EB: <i>Chciał(a)by Pan(i) zadać Panu(i) pytanie o to, na ile ufa Pan(i) niektórym mediom i instytucjom. Jeżeli chodzi o następujące media i instytucje, proszę powiedzieć, czy raczej im Pan(i) ufa, czy raczej im nie ufa?</i> – Unia Europejska	1: raczej ufam 2: raczej nie ufam 3: trudno powiedzieć
ESS: brak pytania	–
EVS: <i>Teraz wymienię różne instytucje i organizacje. Proszę powiedzieć, w jakim stopniu darzy je Pan/i/ zaufaniem: w bardzo dużym stopniu, raczej w dużym stopniu, w niezbyt dużym stopniu, czy w ogóle nie darzy zaufaniem?</i> – Unia Europejska	1: w bardzo dużym stopniu 2: raczej w dużym stopniu 3: w niezbyt dużym stopniu 4: w ogóle nie darzę zaufaniem 8: (trudno powiedzieć) 9: (brak odpowiedzi)
ISSP: brak pytania	–
<b>Zaufanie do Parlamentu Europejskiego</b>	
EB: <i>Proszę powiedzieć, czy raczej Pan(i) ufa, czy raczej nie ufa Pan(i) następującym europejskim instytucjom?</i> – Parlament Europejski	1: raczej ufam 2: raczej nie ufam 3: trudno powiedzieć
ESS: <i>Posługując się tą kartą, proszę wskazać na skali od 0 do 10, na ile osobiście ma P. zaufanie do każdej z instytucji, którą wymienię. '0' oznacza całkowity brak zaufania wobec instytucji, a '10' oznacza, że ma P. do niej całkowite zaufanie. Pozostałe liczby służą do wyrażenia opinii pośrednich. Na ile ma P. zaufanie do:</i> – Parlamentu Europejskiego	00: całkowity brak zaufania 10: całkowite zaufanie 77: (odmowa odpowiedzi) 88: (trudno powiedzieć)
EVS: brak pytania	–
ISSP: brak pytania	–
<b>Poczucie związku z Europą</b>	
EB: <i>Proszę powiedzieć, na ile Pan(i) czuje się związany(a):</i> – z Europą	1: bardzo związany 2: raczej związany 3: raczej niezwiązany 4: w ogóle niezwiązany 5: trudno powiedzieć
ESS: <i>A na ile czuje się P. emocjonalnie związany(a) z Europą?</i>	00: w ogóle nie czuję się emocjonalnie związany 10: czuję się bardzo emocjonalnie związany 77: (odmowa odpowiedzi) 88: (trudno powiedzieć)
EVS: <i>Jak silnie czuje się Pan(i) związany(a) z tymi miejscami?</i> – Europa	1: bardzo silnie 2: silnie 3: niezbyt silnie 4: w ogóle nie czuję się związany 8: (trudno powiedzieć) 9: (brak odpowiedzi)

Treść pytania	Format odpowiedzi
ISSP: brak pytania	-
<b>Poczucie związku z Unią Europejską</b>	
EB: Proszę powiedzieć, na ile Pan(i) czuje się związany(a): - z Unią Europejską	1: bardzo związany 2: raczej związany 3: raczej niezwiązany 4: w ogóle niezwiązany 5: trudno powiedzieć
ESS: brak pytania	-
EVS: brak pytania	-
ISSP: brak pytania	-
<b>Pogłębienie procesu integracji europejskiej</b>	
EB: Jakie jest Pana(i) zdaniem obecne tempo budowania Europy? Proszę spojrzeć na te cyfry. Numer 1 oznacza, że stoi w miejscu, numer 7 oznacza, że jest tak szybkie, jak to możliwe. Proszę wybrać jedną cyfrę, która najlepiej odzwierciedla Pana(i) opinię na temat obecnego tempa budowania Europy.	1: stoi w miejscu 7: jest tak szybkie, jak to możliwe 8: trudno powiedzieć
ESS: A teraz proszę pomyśleć o Unii Europejskiej. Niektórzy twierdzą, że integracja europejska, proces jednoczenia się krajów członkowskich powinien postępować coraz dalej. Inni twierdzą, że proces jednoczenia się Europy i tak zaszedł już zbyt daleko. Posługując się tą kartą, proszę powiedzieć, która liczba na skali najbardziej odpowiada P. opinii w tej sprawie.	00: proces jednoczenia się krajów członkowskich i tak zaszedł już zbyt daleko 10: proces jednoczenia się krajów członkowskich powinien postępować coraz dalej 77: (odmowa odpowiedzi) 88: (trudno powiedzieć)
EVS: Niektórzy uważają, że Unia Europejska powinna nadal się rozszerzać. Inni mówią, że już obecnie rozszerzenie zaszło za daleko. Proszę wybrać z tej karty liczbę, która najlepiej charakteryzuje Pana(i) pogląd w tej sprawie. „1” oznacza „powinna nadal się rozszerzać”, „10” oznacza „rozszerzenie zaszło za daleko”	1: powinien dalej się rozszerzać 10: rozszerzenie zaszło za daleko 88: (trudno powiedzieć) 99: (brak odpowiedzi)
ISSP: Ogólnie mówiąc, czy uważa Pan(i), że Unia Europejska powinna mieć znacznie więcej, więcej, tyle samo, mniej, czy też znacznie mniej władzy niż narodowe rządy krajów członkowskich?	1: znacznie więcej 2: więcej 3: tyle samo 4: mniej 5: znacznie mniej 8: trudno powiedzieć

Źródło: opracowanie własne na podstawie polskich wersji kwestionariuszy wywiadów; dostęp do kwestionariuszy EB, EVS oraz ISSP za pośrednictwem <https://zacat.gesis.org/webview/>; kwestionariusz ESS dostępny na stronie <https://www.europeansocialsurvey.org/>.

podobne pytanie o preferowaną szybkość integracji w przyszłości). Z kolei w Europejskim Sondażu Społecznym pada pytanie o to, czy proces zjednoczenia się Europy nie zaszedł bynajmniej za daleko, czy też powinien postępować. Kwestionariusz EVS zawiera natomiast sformułowane wprost pytanie o rozszerzanie Unii Europejskiej o inne państwa, z kolei w drugiej i trzeciej fali ISSP poświęconej te-

matycznie tożsamości narodowej (badania z 2003 o 2013 roku), pytanie o zakres i stopień zaawansowania integracji europejskiej dotyczyło opinii obywateli wobec ograniczania władzy rządów krajów członkowskich na rzecz władzy na poziomie UE. Wszystkie te pytania, w swojej istocie, dotyczą problematyki pogłębiania procesu integracji, jednak każdy z nich mierzy w zasadzie co innego.

Co więcej, różnice występują również w kafeterii odpowiedzi oraz w sposobach kodowania braków danych. Po pierwsze, nawet jeśli we wszystkich projektach w odniesieniu do pytań o pogłębianie procesu integracji wykorzystano porządkowy format odpowiedzi (co nie jest normą), to w EB i ESS owo uporządkowanie dla akceptacji pogłębiania procesu integracji jest rosnące, natomiast w EVS i ISSP już malejące. Po drugie, w ESS oraz EVS opcja „trudno powiedzieć” została przed respondentami ukryta i musi być spontanicznie przez nich wywołana, by ankieterzy ją zakodowali, podczas gdy w EB i ISSP odpowiedź „trudno powiedzieć” jest już widoczna dla osoby uczestniczącej w wywiadzie kwestionariuszowym, nawet jeśli jej udzielenie oznacza przypadek braku danych.

Pozostałe pytania uwzględnione w zestawieniach zaprezentowanych w tabeli 7 nie różnią się co prawda już tak znacząco pod względem treści zadanego pytania, jednak nadal zauważyć można wyraźne międzyprojektowe zróżnicowania w wykorzystywanych przez badaczy kafeteriach odpowiedzi. W zdecydowanej większości pytań o opinie w projekcie ESS stosuje się jedenastopunktową *quasi*-ciągłą skalę, z kafeterią uporządkowaną rosnąco na kontinuum rozciągającym się pomiędzy 0 a 11 punktów (etykietowane są wyłącznie skrajne punkty skali). Z kolei w projekcie EB w pytaniach o zaufanie do Unii Europejskiej oraz do Parlamentu Europejskiego pojawiają się dychotomiczne opcje odpowiedzi wymuszające na respondencie opowiedzenie się w sposób jednoznaczny po jednej lub drugiej stronie, natomiast w pytaniach o poczucie związku (z Europą lub Unią Europejską) stosowana jest w EB porządkowa czteropunktowa kafeteria uporządkowana malejąco. W badaniach EVS popularne są czteropunktowe kafeterie z kategoriami uporządkowanymi w kolejności malejącej.

Warto również zauważyć, że dodatkowe komplikacje o charakterze metodologicznym dla międzyprojektowej porównywalności danych pojawiają się również wtedy, gdy w grę wchodzi analiza oparte na wielowymiarowych modelach pomiarowych. Dobrym tego przykładem jest wykorzystanie przez badaczy baterii pytań mierzących zaufanie do instytucji krajowych oraz ponadnarodowych. Nawet pobieżna analiza literatury dotyczącej badań empirycznych nad zaufaniem do instytucji politycznych ukazuje, że zestaw pytań wykorzystywany jest przez poszczególnych badaczy na wiele różnych, często osobliwych sposobów. Dla przykładu Allum, Read oraz Sturgis (2011) wprowadzają konstrukt łączący zaufanie polityczne (mierzone zaufaniem do parlamentu oraz polityków) z tzw. zaufaniem uogólnionym (mierzonym znanymi pytaniami o zaufanie w kontaktach z nieznanymi). Zavala-Rojas (2012) definiuje konstrukt odnoszący się do zaufania po-

litycznego na podstawie dwóch niezależnych, choć skorelowanych ze sobą konstruktów, takich jak: zaufanie do władz (oparte na wskaźnikach zaufania do partii politycznych oraz polityków), a także zaufanie do instytucji politycznych (oparte na wskaźnikach zaufania do systemu prawnego oraz parlamentu). Listhaug i Ringdal (2008) wyróżniają dwa całkowicie odmienne wymiary politycznego zaufania, tj. zaufanie do systemu wyborczego (krajowego parlamentu, polityków, partii politycznych) oraz zaufanie do instytucji prawa (systemu prawnego, policji). Z kolei Marien i Hooghe (2011), Hooghe i Marien (2013) oraz McLaren (2012) konstruują skalę zaufania politycznego jednowymiarowo, wykorzystując różne konfiguracje pojedynczych itemów. Innymi słowy, nawet jeśli w danym projekcie zadawane są te same pytania, to badacze wykorzystują je w zupełnie odmiennych konfiguracjach prowadzących do odmiennych konstruktów. Nie powinno zatem dziwić, że w połączeniu z pluralizmem metodologicznym oraz mnogością podejść pomiarowych charakterystycznych dla każdego projektu z osobna, uzyskanie choćby fasadowej międzyprojektowej porównywalności konstruktów wykraczających poza jedno pytanie w kwestionariuszu wydaje się wyzwaniem prawie beznadziejnym.

## **2. Metodologiczne ograniczenia dla porównań międzyprojektowych na przykładzie pytania o zaufanie do Parlamentu Europejskiego w projektach EB oraz ESS**

Obecne osiągnięcia w dziedzinie gromadzenia danych uzyskiwanych w różnych projektach sondażowych pozwalają na podjęcie próby integracji odmiennych podejść metodologicznych ustalonych w odniesieniu do losowania próby i jej terenowej realizacji, a także procesu pomiaru (Japiec i in. 2015; Biemer i in. 2017). Co prawda, skumulowane zbiory danych z pojedynczych projektów badawczych udostępniane były publicznie już od dawna, a obecnie się to upowszechniło, to jednak porównawcze analizy na zbiorach danych z różnych projektów znalazły szerokie zastosowanie dopiero stosunkowo niedawno (Tomescu-Dubrow, Słomczyński 2014; Ganzeboom, Treiman 2020). W szczególności należy tutaj zwrócić uwagę na coraz większą popularność projektów harmonizacji danych, w których stosuje się szeroki wachlarz procedur umożliwiających integrację różnorodnych zbiorów w jedną zbiorczą bazę, zawierającą ujednolicony zestaw zmiennych (Granda, Wolf, Hadorn 2010). Tego rodzaju harmonizacja napotyka co prawda na pewne istotne przeszkody, wynikające wprost z różnorodności w sposobach zadawania pytań kwestionariuszowych oraz standardów gromadzenia danych (Tomescu-Dubrow, Słomczyński 2016), jednak staje się podejściem wykorzystywanym coraz częściej w integracji wyników pochodzących z różnych jakościowo baz danych. Międzyprojektowe harmonizacje ograniczone są jednak zwykle do wąskich i specyficznych kwestii oraz wykonywane na niewielkiej liczbie wskaźników (Dubrow, Tomescu-Dubrow 2016), a pełna harmo-



nizacja okazuje się możliwa wyłącznie w odniesieniu do kilku zmiennych społeczno-demograficznych (Hoffmeyer-Zlotnik, Wolf 2003).

## 2.1. Monotoniczne przekształcenie harmonizacyjne jako nadzieja na międzyprojektową porównywalność danych

Podstawowym celem harmonizacji danych jest przekształcenie określonego zbioru zmiennych/itemów, pochodzących z różnych badań sondażowych (a więc często odmiennych co do treści zadawanych pytań, kafeterii odpowiedzi oraz wartości służących kodowaniu danych) do zharmonizowanego zbioru z oryginalnymi wartościami zmiennych przekształconymi – zazwyczaj monotonicznie – do wspólnego przedziału, np.  $[0;10]$  lub  $[0;1]$ . Dla procesu harmonizacji nie ma oczywiście żadnego znaczenia to, czy przekształcenie źródłowych wartości przeprowadzane jest do przedziału z wartościami skrajnymi rozciągającymi się na kontinuum 0-10, czy też 0-1 lub do zbioru o dowolnych innych wartościach skrajnych, jednakże w projektach harmonizacyjnych najczęściej spotkać można transformacje zmiennych do przedziału  $[0;10]$  lub  $[0;1]$  (zob. Słomczyński, Tomescu-Dubrow, Jenkins 2016).

Proces harmonizacji zmiennych pochodzących z różnych sondaży można w sposób formalny zdefiniować jako przekształcenie, które dowolną zmienną  $X^s$  zamienia na zharmonizowaną zmienną  $H^s$ , przy czym  $X_j^s: j \in [1, 2, \dots, k_s]$ , gdzie  $s \in [1, 2, \dots, S]$  indeksuje sondaże uwzględnione w harmonizacji, natomiast  $k_s$  oznacza liczbę odpowiedzi, które respondenci mogli wybrać odpowiadając na pytanie  $X^s$  w sondażu oznaczonym indeksem  $s$ . Jeżeli ponadto przyjmie się założenie, że:

$$\forall_{X^s: s \in [1, 2, \dots, S]; i, j \in [1, 2, \dots, k_s]} k_s > 2 \wedge i < j \implies X_j^s - X_{j-1}^s = X_i^s - X_{i-1}^s \quad (1)$$

stanowiące o tym, iż różnica pomiędzy dwoma dowolnymi parami następujących po sobie uporządkowanych wartości zmiennej (z wyłączeniem zmiennej dychotomicznej) jest taka sama, to harmonizacja jest funkcją  $h: X^s \rightarrow H^s$ , która każdemu elementowi  $X_j^s$  ze zbioru  $X^s$  przyporządkowuje element ze zbioru w taki sposób, że:

- jeżeli wartości zmiennej  $X_j^s: j \in [1, 2, \dots, k_s]$  uporządkowane są rosnąco, tzn.  $X_1^s < X_2^s < \dots < X_{k_s}^s$ , to zharmonizowaną wartością  $X^s$  jest wyrażenie postaci:

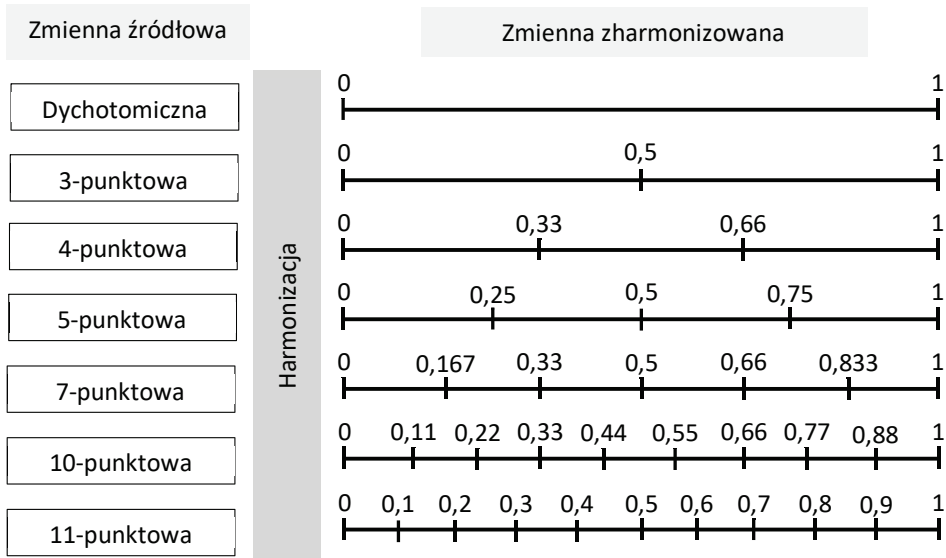
$$H_j^s = \frac{j-1}{k_s-1}; \quad (2)$$

- jeżeli wartość zmiennej  $X_j^s: j \in [1, 2, \dots, k_s]$  uporządkowana jest malejąco, tzn.  $X_1^s > X_2^s > \dots > X_{k_s}^s$ , to zharmonizowaną wartością  $X^s$  jest wyrażenie postaci:

$$H_j^s = \frac{k_s-j}{k_s-1}. \quad (3)$$



Graficzna prezentacją przekształcenia harmonizacyjnego jest poniższy schemat przedstawiony na rysunku 6.



Rys. 6. Liniowa harmonizacja zmiennych źródłowych

Źródło: opracowanie własne.

Zaprezentowane przekształcenie harmonizacyjne nie jest oczywiście jedyną możliwą – a tym bardziej najlepszą – procedurą sprowadzania wartości zmiennych pochodzących z różnych źródeł do wspólnego mianownika (De Jonge, Veenhoven, Kalmijn 2017), jednakże z uwagi na swą prostotę okazuje się niezwykle popularna (Słomczyński, Tomescu-Dubrow, Jenkins 2016). Jest ona przy tym typowym przykładem liniowego przekształcenia wartości zmiennych polegającego na ich zawężeniu do wspólnego przedziału  $[0;1]$ , z zachowaniem proporcji oryginalnych odległości pomiędzy kategoriami odpowiedzi. Harmonizacyjne przekształcenia liniowe posiadają jednak szereg ograniczeń o charakterze metodologicznym, wynikających przede wszystkim z konieczności przyjęcia dwóch kłopotliwych założeń mówiących o tym, że: a) w zbiorze zmiennych poddawanych harmonizacji odległości pomiędzy poszczególnymi kategoriami odpowiedzi są takie same, b) etykietowanie kolejnych kategorii odpowiedzi nie ma wpływu na proces harmonizacji. Monotoniczne liniowe harmonizacje abstrahują w zasadzie od stylu odpowiadania respondentów, dostępnych opcji odpowiedzi oraz ich etykietowania (Batz, Parrigon, Tay 2016). Dla przykładu jeżeli międzyprojektowej harmonizacji poddane byłoby poczucie związku z Europą mierzone w EB oraz EVS odmiennymi

pytaniami zawierającymi, co prawda, w obu projektach tę samą liczbę możliwych odpowiedzi, ale różnie etykietowanych, to owo odmienne etykietowanie w EB (1: bardzo związany, 2: raczej związany, 3: raczej niezwiązany, 4: w ogóle niezwiązany) oraz EVS (1: bardzo silnie, 2: silnie, 3: niezbyt silnie, 4: w ogóle nie czuję się związany) nie miałyby wpływu na liniowe przekształcenie harmonizacyjne i każda para odpowiadających sobie odpowiedzi, np. „raczej związany” w EB oraz „silnie” w EVS, byłaby przekształcona na identyczną zharmonizowaną wartość wynikową.

## 2.2. Wyzwania harmonizacyjne na przykładzie pytania o zaufanie do Parlamentu Europejskiego

Wyzwania metodologiczne oraz trudności praktyczne, wynikające z dążenia do międzyprojektowego zestawiania wyników pomiaru pewnych zmiennych najlepiej zobrazować, odwołując się do przykładów empirycznych. Dobrą egzemplifikacją kłopotliwej natury projektów harmonizacyjnych będzie pytanie dotyczące zaufania do Parlamentu Europejskiego, zadawane respondentom we wszystkich edycjach Eurobarometru oraz obecne we wszystkich odsłonach Europejskiego Sondażu Społecznego. Oczywiście, zbiór krajów objętych badaniami EB nie był tożsamy ze zbiorem krajów uwzględnionych w kolejnych rundach ESS, istnieje jednak część wspólna obejmująca łącznie 23 kraje (rys. 8), którą można włączyć do zestawień porównawczych. Ponieważ oba pomiary sondażowe przeprowadzane zostały w tych samych latach, na tych samych populacjach, przy wykorzystaniu tej samej – co do charakteru – metody reprezentatywnej, to należałoby oczekiwać (przynajmniej na wynikach zagregowanych do poziomu kraju), że oba pomiary wskażą na podobny poziom krajowego zaufania (De Jonge, Veenhoven, Kalmijn 2017). Oczywiście czynników różnicujących wartości pomiaru przeprowadzonego w dwóch różnych projektach może być wiele (niekoniecznie najważniejszym z nich będzie sposób sformułowania pytania), jednakże w projektach harmonizacyjnych, nawet jeśli badacze są świadomi potencjalnego wpływu wszystkich projektowych specyfik na uzyskiwane wyniki, to i tak dąży się do ujednoczenia danych w zasadzie w oderwaniu od kontekstu prowadzonego pomiaru.

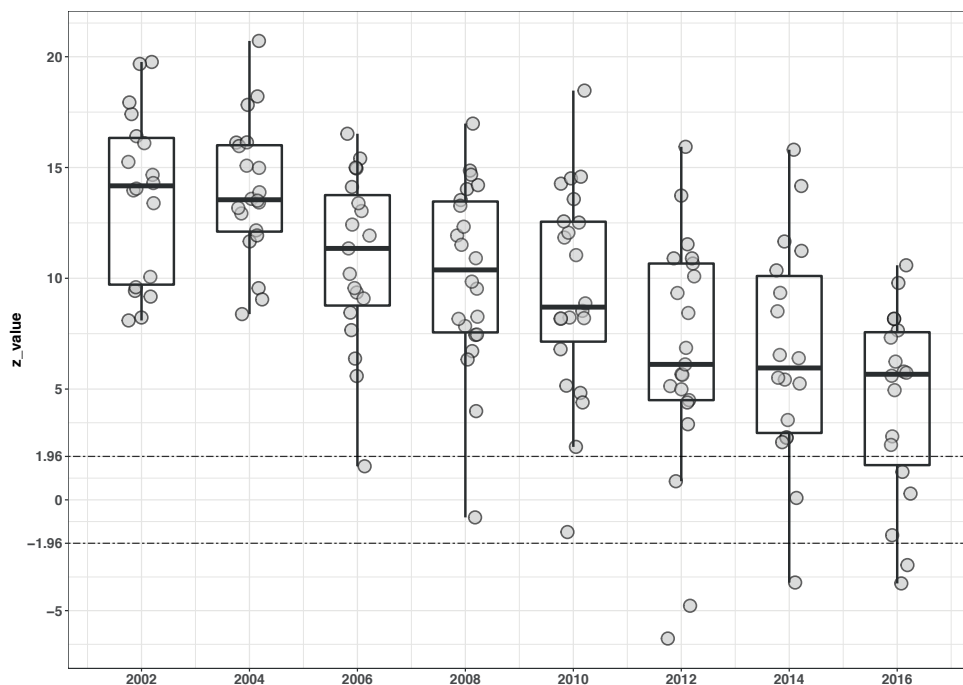
Pytanie o zaufanie do Parlamentu Europejskiego, co do swojej istoty, jest w obu projektach treściowo bardzo podobne. W ESS respondenci proszeni są o wskazanie, na ile osobiście wyrażają oni zaufanie do wymienionych instytucji, jedną z nich jest Parlament Europejski, natomiast w EB pada pytanie o to, na ile respondent ufa europejskim instytucjom, wśród nich wymieniony jest Parlament Europejski. Na tym kończą się jednak podobieństwa. Wprawdzie w obu pytaniach pojawia się wprost sformułowanie o zaufanie do Parlamentu Europejskiego, jednakże w baterii pytań o zaufanie do tej konkretnej instytucji respondenci oceniają również inne. W Eurobarometrze są to instytucje UE, takie jak Komisja Europej-

ska, Europejski Bank Centralny, Rada Europejska oraz Rada Unii Europejskiej, natomiast w ESS zarówno instytucje krajowe (parlament, system prawny, policja, politycy, partie polityczne), jak też Organizacja Narodów Zjednoczonych, o zaufanie do której respondenci są pytani po ocenie zaufania do Parlamentu Europejskiego. Na porównywalności danych z obu projektów ważyć może już sam efekt tła. Wystarczy zauważyć, że wyniki wielu badań empirycznych dotyczących zaufania do instytucji europejskich ukazują, iż ich funkcjonowanie oceniane jest przez pryzmat sprawności funkcjonowania instytucji krajowych (Martinotti, Stefanizzi 1995; Armingeon, Ceka 2014; Ruelens, Nicaise 2020). Jeżeli te ostatnie nie działają w opinii respondentów sprawnie, to instytucje europejskie, w kontrze do krajowych, odbierane są jako efektywne. Z drugiej strony, instytucje krajowe oceniane jako działające sprawnie mogą prowadzić do oceny instytucji europejskich jako zbiurokratyzowanych oraz nieefektywnych. A zatem zadawanie pytania o zaufanie do Parlamentu Europejskiego z asystą lub bez asysty pytań o zaufanie do instytucji krajowych lub innych instytucji ponadnarodowych może mieć znaczący wpływ na udzielane odpowiedzi.

Zasadnicze odmienności w sposobie zadawania pytania o zaufanie do Parlamentu Europejskiego w EB oraz ESS ujawniają się jednak dopiero wtedy, gdy pod uwagę weźmie się stosowane kafeterie odpowiedzi. Dychotomiczny podział, dający w EB respondentom jedynie możliwość opowiedzenia się po stronie osób wyrażających zaufanie lub brak zaufania do Parlamentu Europejskiego (odpowiedź „trudno powiedzieć” jest widoczna, ale traktowana jako brak danych), kontrastuje z *quasi*-ciągłą 11-punktową skalą w ESS, rozciągającą się na kontinuum od 0 oznaczającym „całkowity brak zaufania” do 10 oznaczającym „całkowite zaufanie”, przy czym zarówno „odmowa odpowiedzi”, jak i „trudno powiedzieć” traktowane są jako braki danych i nie są widoczne w kafeterii możliwych do wyboru opcji. Innymi słowy, w ESS respondenci posiadają możliwość stopniowania swojego zaufania do Parlamentu Europejskiego, natomiast w EB muszą się zdecydowanie opowiedzieć po jednej ze stron. Taki odmienny sposób zadawania pytania powinien mieć jednak przełożenie na zróżnicowanie błędów pomiarowych spowodowane odmiennym sposobem reakcji respondentów na skale o różnej długości (Saris, Gallhofer 2014).

Całkowicie odmiennie podejście do formułowania pytań w projektach międzykrajowych ma przede wszystkim swoje przełożenie na sposób interpretacji uzyskiwanych wyników. Można bowiem zauważyć, że nawet gdyby zastosowano podejście harmonizacyjne polegające na liniowym przekształceniu mierzonego poziomu zaufania do wspólnego przedziału  $[0;1]$  oraz obliczono średnie krajowe wartości, to interpretacja średniej pozostałaby i tak pozostać całkowicie odmienna. W projekcie EB średnia wartość zmiennej zharmonizowanej (tożsama z średnią zmiennej oryginalnej) pokazywałaby frakcję respondentów, którzy opowiedzieli

się po stronie osób wyrażających zaufanie do Parlamentowi Europejskiemu na dychotomicznej skali, podczas gdy w badaniu ESS – pomimo harmonizacji do przedziału [0;1] – ukazywałyby przeciętne natężenie zaufania na skali 11-punktowej, tyle że już o innych wartościach skrajnych. Co więcej, trudno oczekiwać uzyskania choć podobnych średnich krajowych wartości zaufania w obu projektach. Nieporównywalność harmonizowanych średnich zaufania do Parlamentu Europejskiego w oparciu o pomiary przeprowadzone (w latach 2002-2016) w projektach EB oraz ESS przedstawia rysunek 7.



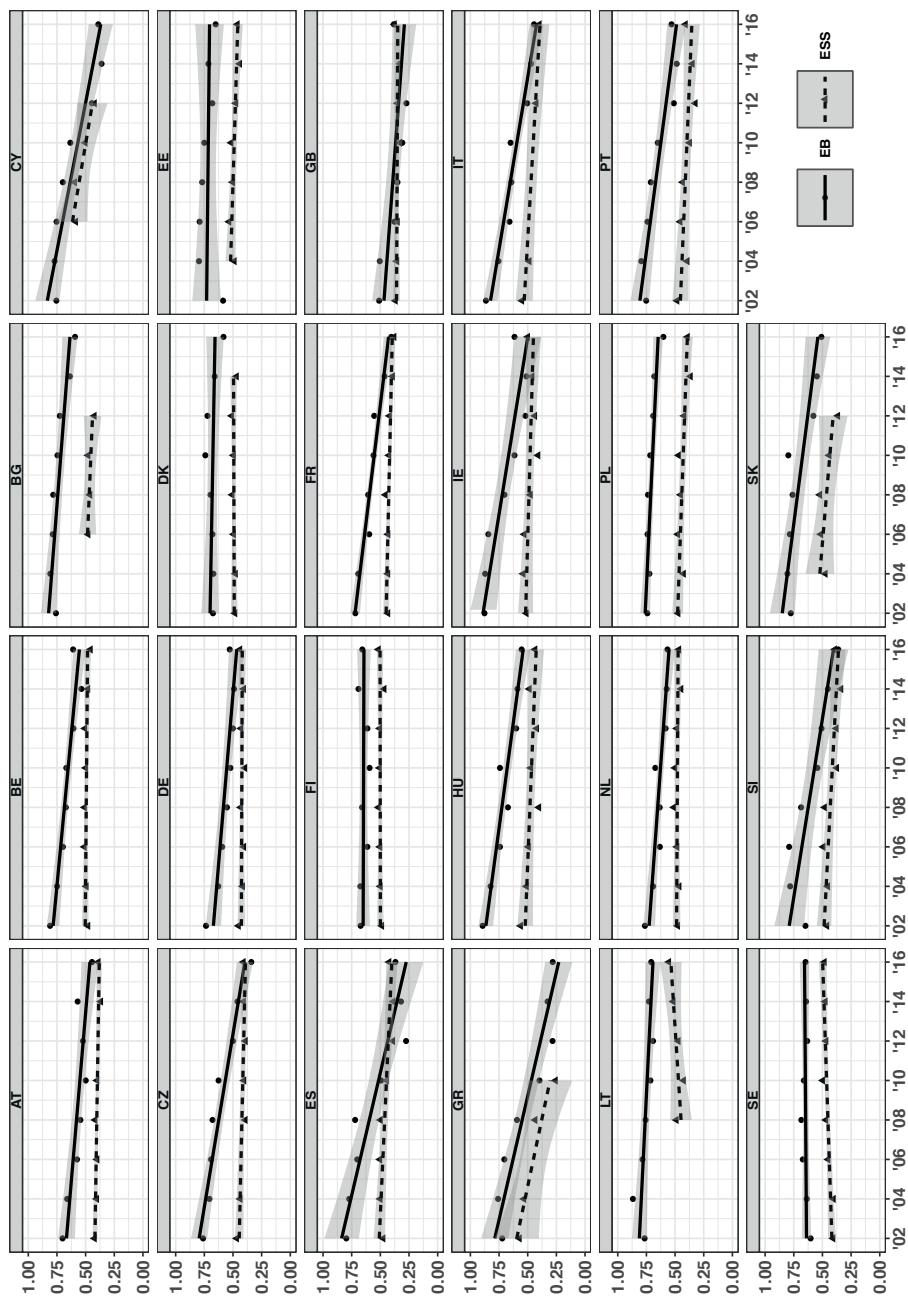
Rys. 7. Różnica pomiędzy zharmonizowaną średnią wartością zaufania do Parlamentu Europejskiego w pomiarze EB oraz ESS w latach 2002-2016

Źródło: opracowanie wykresu wspólnie z P. Cichockim.

W ramach każdej z odpowiadających sobie edycji EB oraz ESS porównane zostały średnie wartości zharmonizowanego zaufania do Parlamentu Europejskiego w tych krajach, które w danym roku brały udział zarówno w pomiarze EB, jak też w ESS. Różnica pomiędzy średnią wartością uzyskaną w pomiarze EB oraz ESS poddana została standaryzacji i zaprezentowana na wykresie w postaci punktów na osi rzędnych (z-value). Dzięki procesowi standaryzacji wszystkie różnice niemieszczące się w przedziale  $\pm 1,96$  wskazują (z 95% przedziałem ufności) na statystyczną istotność różnicy. Dodatkowo przedstawiono również

wykresy skrzynkowe zawierające informacje o medianie, kwartyłu dolnym i górnym oraz wartościach skrajnych analizowanych różnic występujących pomiędzy zharmonizowaną średnią w EB oraz w ESS.

Dane przedstawione na rysunku 7 ukazują jednoznaczną prawidłowość, a mianowicie że dla zdecydowanej większości krajów biorących udział zarówno w badaniu EB, jak i ESS, po zharmonizowaniu wartości zmiennej mierzącej zaufanie do Parlamentu Europejskiego, średnie krajowe wartości wykazują znaczące międzyprojektowe różnice. Wystarczy zauważyć, że jedynie w 8 przypadkach na 167 wszystkich porównywanych w okresie 2002-2016 krajowych par wyników EB w porównaniu z ESS, różnica średnich zharmonizowanych wartości zaufania do PE mieściła się w obszarze krytycznym. A zatem w pozostałych 159 parach różnice były znaczące, co więcej, prawie zawsze (tj. poza 8 przypadkami) wartość różnicy była dodatnia. Innymi słowy, pomiar EB wskazywał na wyższy poziom zaufania do Parlamentu Europejskiego w analizowanych krajach niż pomiar przeprowadzony w ramach ESS. Może to oczywiście po części wynikać z faktu, iż w badaniu Eurobarometru większą skłonnością do wzięcia udziału w sondażu charakteryzują się osoby bardziej przychylnie nastawione do Unii Europejskiej (Nissen 2014), jednakże nie bez znaczenia jest również konstrukcja pytania oraz format odpowiedzi. Otóż jeżeli żąda się od respondentów jednoznacznego opowiedzenia po jednej lub drugiej stronie (zadeklarowania zaufania lub braku zaufania do Parlamentu Europejskiego), to narzuca się im dychotomiczny schemat myślenia, nawet jeżeli stopień ich zaufania lub nieufności nie jest jednoznaczny. A zatem osoby, które posiadają pewne wątpliwości (a ich poziom zaufania nie jest możliwy do uchwycenia zero-jedynkowo), będą i tak musiały się zdecydowanie określić, do której ze stron im bliżej (lub też uciec do kategorii trudno powiedzieć<sup>2</sup>). Inaczej jest w badaniu ESS, gdzie respondenci mogą stopniować poziom swojego zaufania i nie są 'przymuszani' do jednoznacznych deklaracji. Co więcej, taka *quasi*-ciągła konstrukcja kafeterii odpowiedzi okazuje się niezwykle pożądana w analizach ukierunkowanych na ocenę międzykrajowej inwariancji pomiarowej prowadzonej przy wykorzystaniu modelowania strukturalnego (Jabkowski 2018), gdyż jednym z założeń przyjmowanych w tych modelach jest pomiar wskaźników na skali ciągłej (Konarski 2010). Oznacza to, że zastosowanie 11-punktowej skali odpowiedzi spełnia założenie o ciągłości skali w znacznie większym stopniu niż pomiar prowadzony na „krótszej” skali. Otwarte pozostaje przy tym pytanie o to, jaki jest ów próg zaufania, powyżej którego – przy dychotomicznym podziale – respondenci będą mieli tendencje do deklarowania raczej zaufania niż jego braku. Niemniej jednak nawet proste zestawienia wyników krajowych średnich z projektów EB oraz ESS zaprezentowane na rysunku 7 pokazały, iż schematy pomiaru wykorzystujące dwie całkowicie odmienne kafeterie odpowiedzi nie mogą być traktowane jako tożsame.



Rys. 8. Porównanie średniej zharmonizowanych wartości zaufania do PE w wybranych krajach uczestniczących w badaniach EB oraz ESS w latach 2002-2016

Źródło: opracowanie wykresu wspólnie z P. Cichoćkim.

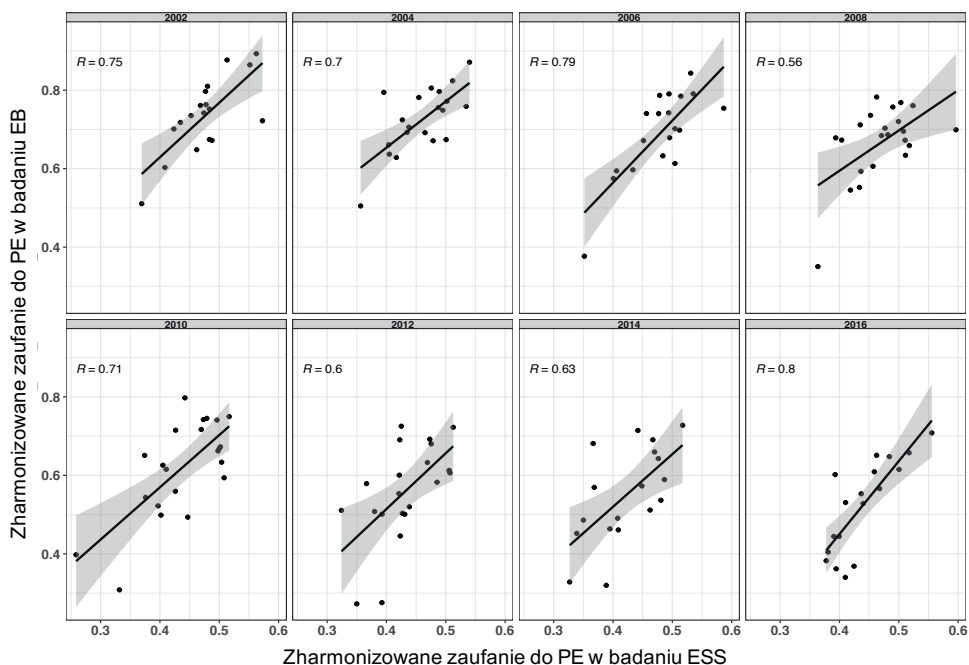
Zjawisku braku zbieżności wniosków formułowanych na podstawie wyników uzyskiwanych w różnych międzykrajowych projektach sondażowych warto przyrzeć się nieco dokładniej, by ustalić, na ile zastosowanie różnych kafeterii odpowiedzi prowadzi do formułowania sprzecznych konkluzji nie tylko w odniesieniu do krajowego agregatu w postaci średniej, ale w analizach innego typu. Dla przykładu można postawić całkiem zasadne pytanie o to, na ile wyniki badań EB oraz ESS będą prowadziły do podobnych wniosków odnośnie do longitudinalnych zmian poziomu zaufania do Parlamentu Europejskiego lub też na ile zgodne będzie uporządkowanie krajów na podstawie uśrednionych wartości poziomu zaufania do Parlamentu Europejskiego w obu projektach.

Niespójność programów EB oraz ESS w odniesieniu do uzyskanych wartości pomiaru zaufania do PE przedstawiono na rysunku 8 (dane obejmują okres od roku 2002 do 2016). Punkty na wykresach wskazują na uśrednioną wartość pomiaru w danym roku, natomiast linie są najlepiej dopasowanymi do danych empirycznych liniami trendu (szare pola wokół linii wyznaczają obszar 95% przedziału ufności dla linii trendu). Wyniki przedstawionych analiz ukazują kilka ciekawych właściwości zarówno w odniesieniu do problemów związanych z porównywaniem wyników pomiaru z różnych projektów badawczych, jak też w zakresie ograniczonej przydatności projektów harmonizacji danych. Po pierwsze, zaprezentowane dane potwierdzają zaobserwowaną już wcześniej prawidłowość wskazującą na to, iż średnie krajowe wartości wyliczone na podstawie zharmonizowanych wyników dychotomicznej skali odpowiedzi w projekcie EB są w przeważającej większości znacząco wyższe od średnich zharmonizowanych wartości mierzonych oryginalnie na 11-punktowej skali stosowanej w ESS. Po drugie, analizując wyniki w wymiarze wzdłużnym, można zauważyć, iż poziom zaufania do Parlamentu Europejskiego mierzony pytaniem ESS charakteryzuje się znacząco większą stabilnością w czasie niż w przypadku EB. Wystarczy spojrzeć na wyniki uzyskane w Austrii, Belgii, Czechach, Hiszpanii, Irlandii, Finlandii, Francji, Niemczech, Węgrach oraz Zjednoczonym Królestwie Wielkiej Brytanii i Irlandii Północnej, by zauważyć praktycznie brak zmian średnich krajowych wyników zaufania do Parlamentu Europejskiego w kolejnych rundach ESS, przy jednocześnie wyraźnym spadku przeciętnego zaufania do PE, notowanego w tych krajach w badaniach Eurobarometru. Rzadko kiedy zdarza się, by linie trendu w EB i ESS były równoległe i wskazywały na taki sam kierunek zmian (Bułgaria, Cypr, Dania, Estonia, Grecja, Polska, Szwecja). Najbardziej kłopotliwe dla międzyprojektowej porównywalności danych są przy tym takie sytuacje, w których krajowe linie trendu wyznaczone z pomiarów przeprowadzonych w kolejnych latach posiadają nachylenie o współczynnikach kierunkowych wyrażonych różnymi znakami. Najlepszym tego przykładem są kolejne pomiary zaufania do Parlamentu Europejskiego przeprowadzone na Litwie. Wyniki z projektu EB ukazują wyraźny trend spadkowy w poziomie



zaufania, podczas gdy wnioski formułowane na podstawie wyników pomiaru przeprowadzonego od 4. do 8. rundy ESS wskazują na wzrost zaufania wyrażanego w stosunku do europejskiego parlamentu.

Dane zaprezentowane na kolejnym rysunku 9 pozwalają spojrzeć na problem porównywalności wyników pochodzących z różnych projektów jeszcze z nieco innej perspektywy. Dla każdej odpowiadającej sobie edycji EB oraz ESS, na kolejnych diagramach korelacyjnych zaprezentowano umiejscowienie każdego kraju z uwagi na średnią wartość zaufania do Parlamentu Europejskiego wyliczoną na podstawie pomiaru EB (oś rzędnych) oraz pomiaru ESS (oś odciętych). Współczynniki korelacji liniowej traktować można przy tym jako miarę zbieżności rankingów, tj. miarę zgodności uszeregowania krajów pod względem średniej wartości zaufania do europejskiego parlamentu w obu projektach (prosta obrazuje zależność liniową pomiędzy pomiarami EB oraz ESS, natomiast szary obszar odpowiada 95% przedziałowi ufności dla linii trendu).



Rys. 9. Diagramy korelacyjne dla średnich zharmonizowanych wartości zaufania do PE w wybranych krajach uczestniczących w badaniach EB oraz ESS w latach 2002-2016

Źródło: opracowanie wykresu wspólnie z P. Cichockim.

Analizując dane przedstawione na powyższym rysunku, można wyraźnie stwierdzić, że zgodność rankingów z uwagi na średni poziom zaufania do Par-



lamentu Europejskiego nie jest szczególnie znaczna. Największą zgodność uszeregowania krajów na podstawie pomiarów EB oraz ESS uzyskano w 2018 roku (współczynnik korelacji liniowej Pearsona wynosi 0,8; oznacza to, że mniej więcej 64% zmienności średnich krajowych wartości zaufania do europejskiego parlamentu w jednym projekcie można wyjaśnić średnią krajową wartością zaufania w drugim projekcie), najmniejszą natomiast w 2008 roku (współczynnik korelacji liniowej wynoszący 0,56 oznacza, że wartości średnich krajowych w jednym projekcie udaje się wyjaśnić jedynie w mniej więcej 32% średnimi krajowymi wartościami w drugim projekcie). Dane te uwidaczniają zatem kolejny poważny problem z międzyprojektową porównywalnością wyników danych sondażowych, a mianowicie w zależności od projektu badawczego, z którego pochodzą zestawienia, niektóre kraje jawić się będą jako bardziej lub mniej pozytywnie nastawione od innych pod względem wyrażanego zaufania do Parlamentu Europejskiego.

### 3. Podsumowanie

Systematycznie realizowane sondażowe badania o charakterze porównawczym stanowią podstawę analiz próbujących odnaleźć podobieństwa i odmienności w postawach i opiniach wyrażanych przez obywateli różnych krajów Europy, jak również źródło danych dla naukowo zorientowanego dyskursu, prowadzonego nad społeczną i systemową integracją na poziomie ponadnarodowym. Już na początku lat 70. XX wieku ukonstytuował się cykliczny program Eurobarometru, będący wówczas pionierskim przedsięwzięciem dającym obraz długo- oraz krótkoterminowej dynamiki zmian opinii publicznej w krajach UE. Z Eurobarometrem wiązano wielkie nadzieje, tak dosadnie wyrażone we wprowadzeniu do raportu z pierwszego pomiaru EB w 1974 roku, w którym napisano: „tak jak barometr może służyć do pomiaru ciśnienia atmosferycznego, a tym samym do przedstawiania krótkookresowej prognozy pogody, tak Eurobarometr może służyć do obserwacji i w pewnym stopniu prognozowania postaw społecznych wobec najważniejszych aktualnych wydarzeń, związanych bezpośrednio lub pośrednio z rozwojem Wspólnoty Europejskiej i zjednoczeniem Europy” (Hewstone 1991).

Znacząca inflacja międzykrajowych projektów sondażowych nadeszła wraz z pojawieniem się – w latach 80. i 90. XX wieku – projektów EVS, ISSP oraz WVS (Smith 2010), jednak dopiero Europejski Sondaż Społeczny przyniósł (na początku XXI wieku) jakościową zmianę w sposobach organizacji badań porównawczych skutkującą większą wrażliwością badaczy na problemy międzykrajowej porównywalności wyników pomiaru sondażowego (Jowell i in. 2007). Z drugiej strony, okres ten był również przełomowy dla integracji europejskiej ze względu na kolejne rozszerzenia UE o kraje Europy Środkowo-Wschodniej, a także serię bez-

precedensowych kryzysów politycznych, społecznych i gospodarczych, jakie dotknęły Europę. O ile większość europejskich projektów porównawczych nie służyła (tak jak EB) zaspokojeniu bezpośrednich potrzeb instytucji unijnych, to mimo wszystko stanowiły one niezależne od instytucji europejskich źródło informacji o opinii publicznej i to nie tylko w krajach członkowskich, ale też w innych europejskich państwach. Należy przy tym pamiętać, że projekty porównawcze służyć miały przede wszystkim porównaniom międzykrajowym, nie zaś międzyprojektowym. W konsekwencji, choć od dawna podejmowano znaczne wysiłki zmierzające do osiągnięcia międzykrajowej ekwiwalencji wewnątrzprojektowej, to dopiero niedawne pojawienie się projektów harmonizacji danych pochodzących z wielu różnych sondaży rozbudziło nadzieję na możliwość przeprowadzenia metodologicznie umocowanych analiz integrujących odmienne źródła danych. Znaczącym ograniczeniem dla takich przedsięwzięć okazało się jednak przede wszystkim

- (1) zróżnicowanie sposobów formułowania pytań w poszczególnych projektach (o ile oczywiście w różnych sondażach udaje się odnaleźć treściowo oraz znaczeniowo podobne pytania kwestionariuszowe),
- (2) jak też zróżnicowanie charakteru procedur wykorzystywanych w losowaniu oraz terenowej realizacji sondażowych prób badawczych.

Na problemy wynikające z odmienności pytań kwestionariuszowych w różnych projektach zwrócono już uwagę w tym rozdziale, natomiast międzyprojektowemu zróżnicowaniu procedur stosowanych w doborze próby oraz jej terenowej realizacji poświęcony będzie trzeci rozdział tej pracy.



## Metodologiczny tygiel – zróżnicowanie praktyk sondażowych w projektach EB, ESS, EVS, ISSP

Zasadniczym celem realizacji międzykrajowych sondaży porównawczych jest uzyskanie danych pozwalających na formułowanie uprawnionych metodologicznie wniosków dotyczących międzykrajowych różnic w pomiarze postaw i opinii społecznych. Cel ten musi zostać osiągnięty pomimo odmiennych procedur terenowych stosowanych w poszczególnych krajowych próbach badawczych. Problematyka porównywalności procedur badawczych pozostaje przy tym jednym z najbardziej intensywnie eksplorowanych obszarów metodologicznej refleksji nad jakością badań porównawczych w ogóle (Jowell 1998; Harkness i in. 2010). Wysiłki skupiają się przy tym wokół dwóch podstawowych wyzwań:

- (1) zapewnienia ekwiwalentności procesu losowania oraz terenowej realizacji prób badawczych – pomimo występowania międzykrajowych różnic w operatach doboru próby, schematach losowania jednostek, czy też procedurach terenowej realizacji badań (Stoop i in. 2010), a także
- (2) zapewnienia ekwiwalentności pomiaru pewnych konstruktów latentnych – pomimo odmiennych kontekstów kulturowych, w których prowadzone są badania (Jowell i in. 2007).

Nie są to problemy posiadające proste rozwiązania, albowiem uzyskanie porównywalności w badaniach międzykrajowych wymaga uważnego namysłu nad wszystkimi etapami procesu badawczego wdrażanymi jednocześnie w wielu różnych krajowych próbach sondażowych (Lynn, Japac, Lyberg 2006).

Głównym celem tego rozdziału jest prezentacja zróżnicowania praktyk terenowych w czterech, rozpatrywanych w tej pracy, międzykrajowych badaniach porównawczych. Analizy empiryczne przeprowadzone zostały na podstawie kwerendy procedur badawczych, udokumentowanych w raportach metodologicznych kolejnych edycji EB, ESS, EVS oraz ISSP. Scharakteryzowano bazę źródłową oraz opisano zakres przeprowadzonej metodologicznej inwentaryzacji krajowych sondaży. Biorąc pod uwagę kryterium kompletności informacji zawartych w dokumenta-

cji badawczej czterech projektów, ocenie poddano również jakość dokumentacji. Główna część rozdziału poświęcona została omówieniu międzyprojektowego, międzykrajowego oraz w dłuższym czasie zróżnicowania praktyk sondażowych, w tym ocenie różnic w sposobach definiowania badanych populacji, rodzajach wykorzystanych operatów doboru próby, typach prób sondażowych, wdrażanych procedurach wewnątrzspołecznej selekcji respondentów docelowych w adresowych próbach osób, procedurach terenowych oraz standardach raportowania o efektach realizacji prób sondażowych. Rozdział rozpoczyna jednak wyczerpujący przegląd literatury przedmiotu ukazujący obecny stan wiedzy w zakresie wpływu różnych procedur badawczych na jakość pomiaru sondażowego.

## **1. Zróżnicowanie procedur badawczych jako wyzwanie dla porównywalności danych – analiza literatury przedmiotu**

### **1.1. Sposoby definiowania badanych populacji**

Jednym z kluczowych wyzwań dla międzykrajowej ekwiwalencji sondaży w projektach porównawczych pozostaje konieczność zdefiniowania populacji w jednolity sposób we wszystkich krajach uwzględnionych w pomiarze. A zatem jeżeli krajowe pomiary realizowane w ramach danego projektu mają posiadać charakter porównawczy, to w każdym z nich populacja powinna uwzględniać jednostki o dokładnie takich samych cechach. Co oczywiste, z faktu prowadzenia sondaży w różnych krajach wynika już wprost, iż w badaniach o charakterze porównawczym ma się do czynienia z wieloma odmiennymi krajowymi populacjami, jednak zapewnienie ekwiwalencji w tym względzie wymaga, by owe krajowe populacje określić w ten sam sposób, co wymaga podania cech jednostek oraz ram terytorialnego i czasowego włączania pewnych osób do populacji.

Warto przy tym zaznaczyć, że w notach metodologicznych większości międzykrajowych projektów sondażowych populacje definiowane są w sposób deskryptywny, tj. poprzez podanie cech, jakie powinny posiadać jednostki, by uznać je za elementy populacji (Lissowski, Haman, Jasiński 2008). Należy jednak mieć na względzie, iż w wielu krajowych sondażach wprowadza się dodatkowe ograniczenia wyłączające pewne kategorie jednostek poza zakres badanej populacji; mowa wówczas o populacji docelowej (Kish 1979; Kruskal, Mosteller 1979; Groves 2004; Cox 2008). W konsekwencji, nawet jeśli definicje populacji generalnej w krajowych sondażach realizowanych w ramach określonego projektu porównawczego będą jednakowe, to kraje różnicować może i tak sposób definiowania populacji docelowych. Decyzja taka jest w gruncie rzeczy arbitralna, choć bywa motywowana względami metodologicznymi lub praktycznymi, ważne jednak jest to, że ogra-

nicza międzykrajową porównywalność danych. Problem ten będzie przedmiotem szczegółowych analiz w kolejnej części tego rozdziału, w której przyjrzymy się zróżnicowaniu sposobów definiowania populacji w czterech porównywanych projektach międzykrajowych.

## 1.2. Rodzaje operatów doboru prób badawczych

Ważną cechą różnicującą krajowe sondaże pozostaje rodzaj wykorzystanego operatu doboru próby. Można przy tym wyróżnić kilka głównych typów operatów (rys. 10), w tym

- (1) rejestry imienne,
- (2) operaty zawierające wykazy budynków mieszkalnych oraz
- (3) operaty zawierające wykazy gospodarstw domowych (Stoop i in. 2010).

Co więcej, w wielu krajach europejskich nie istnieją wiarygodne wykazy jednostek, mieszkań lub choćby budynków, które można by wykorzystać w losowaniu próby (Lynn i in. 2007). To ostatnie oznacza konieczność doboru prób przestrzennych (zob. Marker, Stevens 2009).

Typ operatu	Imienny	Gospodarstw domowych	Adresowy budynków	Brak wykazu
Operat losowania	<ul style="list-style-type: none"> <li>wykaz jednostek</li> </ul>	<ul style="list-style-type: none"> <li>wykaz mieszkań</li> </ul>	<ul style="list-style-type: none"> <li>wykaz budynków</li> </ul>	<ul style="list-style-type: none"> <li>brak wiarygodnych rejestrów</li> </ul>
Etap 1: Wybór budynków mieszkalnych	<ul style="list-style-type: none"> <li>nie stosuje się</li> </ul>	<ul style="list-style-type: none"> <li>nie stosuje się</li> </ul>	<ul style="list-style-type: none"> <li>losowy dobór budynków mieszkalnych</li> </ul>	<ul style="list-style-type: none"> <li>losowy lub systematyczny dobór miejscowości lub ulic</li> <li>spis budynków mieszkalnych oraz systematyczny lub losowy dobór budynku</li> </ul>
Etap 2: Wybór gospodarstw domowych	<ul style="list-style-type: none"> <li>nie stosuje się</li> </ul>	<ul style="list-style-type: none"> <li>losowy dobór gospodarstw domowych</li> </ul>	<ul style="list-style-type: none"> <li>spis gospodarstw domowych oraz systematyczny lub losowy dobór mieszkań</li> </ul>	<ul style="list-style-type: none"> <li>spis gospodarstw domowych oraz systematyczny lub losowy dobór mieszkań</li> </ul>
Etap 3: Wybór jednostek do próby	<ul style="list-style-type: none"> <li>losowa selekcja z operatu</li> </ul>	<ul style="list-style-type: none"> <li>losowa, quasi-losowa lub nielosowa selekcja respondenta docelowego w obrębie wybranego mieszkania</li> </ul>		

Rys. 10. Typy operatów a działania zmierzające do losowania jednostek populacji

Źródło: opracowanie własne.

Oddzielną kategorię operatorów stanowią wykazy numerów telefonii stacjonarnej oraz mobilnej, choć wiążą się one z koniecznością realizacji wywiadów przy pomocy technik wspomaganych telefonicznie. Warto mieć również na względzie fakt, że o ile populację docelową zdefiniować można w sposób deskryptywny (podając cechy posiadane przez jednostki), to operat musi zawierać już wykaz jakichś konkretnych elementów pozwalających na dobór jednostek do próby w sposób bezpośredni (z operatorów imiennych) lub pośredni (z operatorów innych niż imienne).

Dla międzykrajowej porównywalności wyników najbardziej znaczącym ograniczeniem okazuje się przy tym dostęp do operatorów różnej jakości. Warto również pamiętać, że spośród wielu różnych czynników posiadających wpływ na jakość operatu, zdecydowanie największe znaczenie przypisuje się błędom niepełnego pokrycia populacji (Lyberg, Japac, Tongur 2018). Jeżeli bowiem badacz nie posiada kompletnej listy jednostek tworzących populację, to – z uwagi na niepełne pokrycie – część osób nie ma żadnych szans wylosowania (Weisberg 2005). Zjawisko to jest przedmiotem szczegółowych analiz w odniesieniu do badań realizowanych technikami wywiadów telefonicznych (Curtin, Presser, Singer 2005; Vincente, Reis 2009) lub ankiet internetowych (Best, Krueger 2002; Couper i in. 2007; Dever, Rafferty, Valliant 2008; Heerwegh, Loosveldt 2008), niemniej jednak staje się szczególnie ważne w sondażach międzykrajowych, wymagających zastosowania operatorów różnego typu, ale o porównywalnej jakości.

### 1.3. Schematy losowania próby

Punktem odniesienia w ocenie schematu doboru próby (nie tylko w międzykrajowych projektach porównawczych, ale w ogóle) pozostaje prosta próba losowa wraz ze swoimi charakterystykami definicyjnymi, tj.:

- (1) losowaniem jednostopniowym przeprowadzanym wprost z rejestru populacyjnego (bez konieczności pośredniej selekcji jednostek z operatorów adresowych),
- (2) losowaniem pojedynczych osób (a nie ich wiązek lub zespołów),
- (3) losowaniem prowadzonym z całej populacji (bez podziału na jej poszczególne warstwy), oraz
- (4) losowaniem z zachowaniem równości szans selekcji jednostek (Jabkowski 2015).

Każdy odbiegający od schematu prostego dobór próby wiąże się już z koniecznością uwzględnienia efektu schematu losowania, albowiem nietrywialną konsekwencją doboru innego niż prosty jest zmniejszenie precyzji estymacji. Innymi słowy, aby próby dobierane w różnych krajach według odmiennych schematów losowania pozwalały na szacowanie wartości parametrów populacyjnych z podobną precyzją, należy (w sposób odpowiedni) zróżnicować liczebności prób

w krajowych sondażach (Lynn i in. 2007; Lee 2012). Zresztą nie o schemat doboru próby rozgrywa się największa batalia w sondażach porównawczych, lecz o liczbowe skwantyfikowanie schematu i uwzględnienie jego właściwości w ustaleniu minimalnej liczebności próby badawczej potrzebnej do osiągnięcia zakładanego maksymalnego błędu wnioskowania (przy danym poziomie istotności).

Biorąc pod uwagę różnorodność możliwych podejść badawczych, stworzenie wyczerpującej klasyfikacji schematów losowania nie ma w ogóle sensu. Należy skupiać się raczej na motywacjach skłaniających badaczy do zastosowania określonych schematów doboru próby oraz na konsekwencjach podejmowanych decyzji. Dla przykładu można wskazać, że w celu uniknięcia znacznego terytorialnego rozproszenia jednostek dokonuje się często wyboru małych (kilkuosobowych) wiązek osób zamieszkujących te same miejscowości (uchyla się tym samym jedno z głównych założeń próby prostej, mówiące o losowaniu indywidualnym). Przyjęcie takiej zasady skutkuje homogenizacją cech respondentów wewnątrz danej wiązki osób (Kish 1965), obniża efektywność próby i wymaga zwiększenia jej liczebności (Gabler, Hader, Lynn 2008). Pozwala jednak znacząco ograniczyć koszty dotarcia do respondentów (Groves 2004; Weisberg 2005). Podobnie, w celu zwiększenia kontroli nad populacyjnymi rozkładami pewnych społeczno-demograficznych cech jednostek, dokonuje się często losowania warstwowego (niejednokrotnie nieproporcjonalnego do populacyjnych wielkości warstw), co skutkuje zróżnicowaniem szans selekcji jednostek (Lynn i in. 2007). Na zróżnicowanie szans selekcji jednostek znaczący wpływ w próbach adresowych ma także dobór wielostopniowy; prawdopodobieństwo losowania konkretnej osoby jest odwrotnie proporcjonalne do liczby członków gospodarstwa domowego (Jabkowski 2015). Warto jednocześnie wskazać, że dla poprawności wnioskowania statystycznego oraz dla zachowania porównawczego charakteru krajowych badań kluczowe jest to, by owe szanse doboru były znane (Dorofeev, Grant 2006). Nierówne prawdopodobieństwa selekcji można bowiem równoważyć waznieniem przypisującym każdej jednostce wartość będącą odwrotnością szans jej selekcji (Horvitz, Thompson 1952), co skutkuje przyrostem wariancji estymatorów i wymaga losowania prób o większej liczebności (Kish 1992; Lee 2012). Gdyby jednak prawdopodobieństwa selekcji nie były znane, to proces estymacji z całą pewnością pozostałby obciążony (Berger, Tillé 2009).

Wyprzedzając analizy zaprezentowane w drugiej części tego rozdziału, można już teraz wspomnieć, że w zdecydowanej większości międzykrajowych projektów sondażowych o charakterze porównawczym schematy losowania okazały się jednym ze słabiej udokumentowanych etapów procesu badawczego (z wyjątkiem Europejskiego Sondażu Społecznego). Co więcej, ponieważ w niektórych projektach przywiązywano duże znaczenie do metodologicznej poprawności wnioskowania i wyznaczano wagi korygujące nierówne prawdopodobieństwa selekcji, w innych



natomiast tego nie robiono, to w konsekwencji utrudnione okazują się nie tylko uprawomocnione metodologicznie analizy międzyprojektowe, ale też międzykrajowe analizy porównawcze prowadzone na wynikach tych konkretnych projektów, w których wag nie wyznaczono. Zestawienie najbardziej popularnych schematów losowania w czterech analizowanych w tej pracy projektach porównawczych przedstawione zostanie w kolejnych częściach tego rozdziału.

#### 1.4. Typy sondażowych prób badawczych

Sondaże realizowane w projektach o charakterze porównawczym wykazują znaczące międzykrajowe oraz międzyprojektowe zróżnicowania z uwagi na typy prób badawczych. Decyzja o zastosowaniu określonego typu zależy oczywiście od wielu czynników, spośród których największe znaczenie ma jakość dostępnych operatorów doboru próby, tradycja realizacji sondaży w danym kraju, zasoby infrastrukturalne i finansowe oraz wymagania metodologiczne stawiane przed krajowymi koordynatorami projektów (Heeringa, O’Muirheartaigh 2010). Podstawowym rozróżnieniem typów prób sondażowych – pojawiającym się w wielu opracowaniach metodologicznych – jest ich podział na próby losowe (w których każdy respondent ma znane i niezerowe prawdopodobieństwa selekcji) oraz próby nielosowe (Baker i in. 2013). Te ostatnie dobierane są najczęściej z zachowaniem ustalonego z góry udziału określonych cech społeczno-demograficznych jednostek oraz posiadają komponent losowej ścieżki (Lohr 2008). Warto przy tym pamiętać, że ów komponent – nazywany w metodologii sondażowej mianem *random route* – pojawić może się również w próbach losowych. Innymi słowy, od kontekstu oraz sposobu wdrożenia tej procedury zależy probabilistyczny lub nieprobabilistyczny charakter próby (Bauer 2014, 2016). Co istotne, losowość próby jest warunkiem koniecznym dla prowadzenia uprawomocnionych metodologicznie międzykrajowych zestawień wyników (Häder, Gabler 2003), co oczywiście nie oznacza, że we wszystkich rozważanych w tej pracy projektach porównawczych krajowe próby miały wyłącznie taki charakter. Będzie to przedmiotem pogłębionych analiz w dalszej części tego rozdziału.

Niezwykle przydatną klasyfikację typów prób badawczych odnaleźć można w pracy Ulrich Kohler (2007). W analizach prowadzonych na danych pochodzących z kilku projektów porównawczych autor ten wprowadził sześć kategorii prób sondażowych:

- (1) próby kwotowe typu nieprobabilistycznego oraz pięć typów prób probabilistycznych, w tym:
- (2) imienne proste (lub warstwowe) próby losowe,
- (3) wielostopniowe próby imienne,
- (4) wielostopniowe adresowe próby osób zawierające wykazy budynków lub gospodarstw domowych,

- (5) wielostopniowe próby adresowe z elementami *random route* służące ustaleniu adresu konkretnego gospodarstwa domowego oraz
- (6) próby adresowe bez sprecyzowanych informacji o etapach losowania. Do tej kategorii Kohler (2007) zaliczył wszystkie te sondaże realizowane na próbach losowych dobieranych z operatów adresowych, które nie zostały dostatecznie dokładnie opisane w raportach metodologicznych analizowanych przez niego projektów.

Warto przy tym pamiętać, że decyzja o zastosowaniu określonego typu próby badawczej ma bezpośrednie przełożenie na jakość pomiaru (Eckman, Koch 2019). O ile bowiem próby imienne pozwalają na losowanie jednostek wprost z rejestru, który zawiera wykazy osób należących do badanej populacji, to próby adresowe wiążą się z koniecznością pośredniej selekcji jednostek. Będzie o tym mowa w kolejnej sekcji tego rozdziału. W tym momencie warto wspomnieć, że w próbach innych niż imienne badacze mają ograniczoną możliwość weryfikacji tego, czy do próby wybrany został rzeczywiście ten respondent, który powinien się w niej znaleźć (Jabkowski 2015). Próby imienne stawiają bowiem badacza w uprzywilejowanej sytuacji względem pozycji ankietowanych z uwagi na fakt, iż potencjalny respondent znany jest z imienia i nazwiska (dostępne są również inne cechy jednostek), a zatem każda podmiana wylosowanej osoby na inną dla ankietera wiąże się z wysokim ryzykiem demaskacji jego fałszerstw dokonanych na etapie selekcji (Jabkowski, Cichocki 2019). Próby imienne posiadają w konsekwencji wyższą jakość od prób adresowych (Kohler 2007; Menold 2014).

### 1.5. Procedury selekcji respondentów docelowych w adresowych próbach osób

Z koniecznością selekcji wewnątrzspołowej badacze spotykają się w sytuacji realizacji sondaży opartych na próbach adresowych (Sawiński 2005), tj. wtedy, gdy losowaniu podlegać powinny pojedyncze osoby, natomiast operat pozwala na dobór budynku mieszkalnego lub gospodarstwa domowego. W takich sytuacjach zachodzi konieczność wielostopniowego doboru jednostek, przy czym ostatnim etapem losowania jest wybór jednej osoby spośród domowników tworzących dane gospodarstwo. Warto już na wstępie wskazać, że w literaturze badań sondażowych odnaleźć można przynajmniej kilkanaście różnych procedur służących doborowi respondentów docelowych w próbach adresowych (Gaziano 2005; Jabkowski 2017). Niektóre z nich posiadają charakter ściśle probabilistyczny, tzn. opierają się na zrandomizowanym doborze respondentów, podczas gdy inne mają charakter *quasi*-losowy lub nawet nielosowy.

Do najbardziej znanych procedur probabilistycznych zalicza się przy tym siatka Kisha (Kish 1949), procedura szeregowania respondentów „wyłącznie-według-

-wieku” (Denk, Hall 2000) oraz metoda Rizzo, która stanowi łatwiejszą do implementacji wersję siatki Kisha (Rizzo, Brick, Park 2004; Beebe i in. 2007). Z kolei typowymi przykładami procedur *quasi*-losowych są metody wykorzystujące daty urodzin respondentów (Salmon, Nichols 1983), natomiast do procedur nielosowych zalicza się metody kwotowe, w tym m.in. technikę zaproponowaną w pracy Troldahl i Carter (1964), wraz z jej różnymi modyfikacjami wprowadzonymi m.in. przez: Paisley i Parker (1965), Sabin i Godley (1987), Groves i Kahn (1979), Czaja, Blair i Sebestik (1982), Hagan i Collier (1983), Keeter i Fisher (1997), czy też Le i in. (2013).

Procedury selekcji respondentów docelowych w próbach adresowych różnią się nie tylko założeniami teoretyczno-metodologicznymi, ale też specyficznymi wyzwaniem z terenową realizacją próby oraz rodzajami błędów, jakie niesie za sobą wdrożenie poszczególnych technik losowania. Dla przykładu można wskazać, iż procedury probabilistyczne wydłużają czas, jaki ankieter poświęca na realizację wywiadu oraz zwiększają ryzyko nieprzeprowadzenia wywiadu z uwagi na odmowy spowodowane niechęcią do udzielania informacji o wszystkich domownikach (O'Rourke, Blair 1983). *Quasi*-probabilistyczne procedury nie wymagają już wyliczania wszystkich członków gospodarstwa domowego, jednakże ceną za to jest rezygnacja z losowego charakteru procesu selekcji. Z drugiej jednak strony nie są one już tak uciążliwe i czasochłonne, pozostają łatwe do terenowej implementacji, co skutkuje niższymi wskaźnikami odmów i wyższą kooperacją (Jabkowski 2017). Z kolei nieprobabilistyczne procedury selekcji respondentów docelowych opierają się na idei prób kwotowych, a ich podstawowym celem jest uzyskanie odpowiednio sprofilowanej próby według pewnych cech społeczno-demograficznych jednostek, takich choćby jak płeć i wiek. Kosztem, jaki płaci się za implementację procedur nielosowych, jest jednak brak informacji o prawdopodobieństwach selekcji jednostek oraz (w konsekwencji) brak możliwości zastosowania klasycznych metod wnioskowania statystycznego. Jest to zresztą jeden z głównych powodów, dla którego unika się wykorzystywania tego typu procedur w badaniach o charakterze sondażowym, w tym zwłaszcza w badaniach porównawczych (Gaziano 2005).

Do dwóch najbardziej popularnych technik umożliwiających wewnątrzspółową selekcję respondentów docelowych należy siatka Kisha oraz procedura daty urodzin. Siatka Kisha polega przy tym na spisaniu oraz uszeregowaniu według płci i wieku wszystkich dorosłych członków wylosowanych gospodarstw domowych, a następnie na zastosowaniu – z ustaloną częstotliwością – jednej z sześciu (lub jednej z ośmiu) tablic zawierających numer osoby, którą należy wybrać do próby (Kish 1949). Co do istoty obecnie stosowane modyfikacje tej procedury niewiele różnią się od pierwowzoru. Z drugiej strony, metody wykorzystujące daty urodzin polegają na losowaniu takiej osoby, która

- (1) obchodziła urodziny jako ostatnia,
- (2) będzie następną obchodzić urodziny lub też
- (3) której data urodzin przypada najbliżej dnia wizyty ankietera (Jabkowski 2015).

Ponieważ procedury te okazują się niezwykle często wykorzystywane w międzykrajowych projektach sondażowych o charakterze porównawczym, o czym będzie zresztą mowa w dalszej części pracy, to warto przyrzeć się ustaleniom literaturowym traktującym o ich wpływie na jakość procesu selekcji oraz na wskaźniki terenowej realizacji próby.

W pierwszej kolejności można wskazać, że obie procedury dają próby charakteryzujące się znaczącą nadreprezentacją kobiet (Denk, Guterbock, Gold 1996; Yan 2009; Jabkowski 2017). Większość analiz ukazywała przy tym, że siatka Kisha prowadzi do mniejszej nadreprezentacji kobiet, niż ma to miejsce przy wyborze respondentów w oparciu o daty urodzin (Denk, Hall 2000; Yan 2009; Jabkowski 2017). Zarówno siatka Kisha, jak i procedury oparte na dacie urodzin wykazują jednak tendencje do znaczącego niedoreprezentowania osób młodszych oraz starszych. Należy przy tym zaznaczyć, że młodszy respondenci pozostają niedoreprezentowani z uwagi na niższą dostępność oraz wyższe wskaźniki odmów, podczas gdy starsi głównie z powodu niskiej gotowości do udziału w badaniu (Stoop 2004). Co istotne, próby dobierane z wykorzystaniem siatki Kisha zawierają relatywnie mniej osób młodszych oraz mniej osób starszych niż w próbach wykorzystujących daty urodzin respondentów (O'Rourke, Blair 1983). Procedura wykorzystująca daty urodzin pociąga również za sobą większe ryzyko nieintencjonalnego lub intencjonalnego zniekształcenia przebiegu procesu selekcji, ponieważ respondenci oraz ankieterzy mają większą możliwość wskazania osoby przebywającej w domu jako tej, która spełnia warunki selekcji (Jabkowski 2015). Innymi słowy, konieczność pełnej enumeracji wszystkich członków określonego gospodarstwa domowego w siatce Kisha ogranicza możliwość wpływu ankietera na proces selekcji i skutkuje wyższą jakością prób dobieranych z wykorzystaniem tej procedury (Yan 2009). Z drugiej strony należy jednak pamiętać o wpływie siatki Kisha na wskaźniki terenowej realizacji próby. W literaturze metodologicznej panuje w zasadzie zgoda co do tego, że konieczność sporządzania przez ankieterów pełnego wykazu członków gospodarstwa domowego ma niekorzystny wpływ na wskaźniki terenowej realizacji próby (Yan 2009), prowadząc do niższych wskaźników kooperacji (Binson, Catania 2000) oraz wyższych wskaźników odmów (O'Rourke, Blair 1983; Lavrakas, Bauman, Merkle 1993).

Podsumowując, należy wskazać, że każda procedura selekcji respondentów docelowych posiada swoje zalety oraz ograniczenia. Niektóre okazują się stosunkowo łatwe do implementacji, podczas gdy inne wymagają więcej wysiłku ze strony ankieterów, więcej czasu oraz wiążą się z większymi kosztami. Wybór optymalnej

procedury selekcji w porównawczych sondażach międzykrajowych stanowi duże wyzwanie, gdyż od terenowej realizacji krajowych sondaży oczekuje się daleko posuniętej ekwiwalentności uzyskiwanych wyników (Lynn i in. 2007).

## 1.6. Procedury terenowej realizacji prób sondażowych

Projekty porównawcze o charakterze sondażowym wykazują znaczące zróżnicowanie z uwagi na stosowane procedury terenowe (Jabkowski, Kołczyńska 2020a). Można przy tym zauważyć nie tylko istotne różnice międzyprojektowe w tym względzie (posiadające swoje źródło w specyficznych dla każdego projektu celach badawczych, sposobach organizacji i zarządzania projektem, czy też w swoistej kulturze metodologicznej każdego przedsięwzięcia), ale również znaczące międzykrajowe zróżnicowania wewnątrzprojektowe. Różnice te dotyczą przede wszystkim:

- (1) czasu przeznaczanego na terenową fazę badań,
- (2) stosowanych technik badawczych,
- (3) możliwości dokonywania substytucji, tj. podmiany jednostek niedostępnych na respondentów zastępczych o podobnych cechach społeczno-demograficznych,
- (4) stosowania gratyfikacji pieniężnych lub innych nagród warunkowanych (lub nie) zgodą na udział w badaniu,
- (5) wysyłania listów zapowiednich informujących o celu badania oraz przewidywanej dacie podjęcia próby nawiązania przez ankietera kontaktu z wylosowaną osobą,
- (6) wykorzystywania procedur konwersji odmów mających na celu nakłonienie osób niechętnych na udział w badaniu, jak również
- (7) stosowania procedur kontrolnych weryfikujących poprawność realizacji wywiadów oraz jakość pracy ankieterów.

Co istotne, wymienione działania pozostają pod całkowitą kontrolą badacza (De Heer 1999), a ich wdrożenie posiada przełożenie na wskaźniki realizacji próby oraz jakość pomiaru (De Heer, De Leeuw 2002; Groves 2004; Weisberg 2005; Brick, Montaquila 2009; Groves, Lyberg 2010; Stoop i in. 2010). Dla przykładu można wskazać opracowania metodologiczne stwierdzające istotny wpływ techniki badawczej na wskaźniki realizacji próby (Bethlehem, Cobben, Schouten 2011; Krzewińska, Grzeszkiewicz-Radulska 2013), czy też prace potwierdzające dodatnią korelację pomiędzy długością terenowej fazy badania oraz poziomem realizacji próby (Sztabinski, Sztabinski, Przybysz 2009; Jabkowski 2011; Fuchs, Bossert, Stukowski 2013). W literaturze podkreśla się także pozytywny wpływ wysyłania listów zapowiednich (Biemer, Lyberg 2003; Vonder Lippe, Schmich, Lange 2011), stosowania gratyfikacji (Grauenhorst, Blohm, Koch 2016), konwersji odmów (Sto-

op i in. 2010), czy też procedur kontrolnych (Kohler 2007) oraz negatywne oddziaływanie dokonywania substytucji niedostępnych respondentów osobami o podobnych cechach (Elliot 1993; Lynn 2004; Menold 2014; Muñoz-Conde i in. 2018; Jabkowski, Cichocki 2019; Baldissera, Ferrante, Campostrini 2020).

### 1.7. Wskaźniki terenowej realizacji próby

Raportowanie o efektach terenowej realizacji prób sondażowych sprowadza się w praktyce do ustalenia liczby respondentów (tj. zrealizowanych wywiadów) oraz liczby jednostek niedostępnych (tj. osób, z którymi wywiadów nie udało się przeprowadzić), a także do obliczenia kilku standardowych mierników realizacji próby, w tym ogólnego wskaźnika poziomu jej realizacji (*response rate*), wskaźnika kooperacji (*cooperation rate*), wskaźnika odmów udziału w badaniu (*refusal rate*) oraz wskaźnika kontaktów (*contact rate*). W realizowanych obecnie projektach sondażowych miary te wykorzystywane są w zasadzie powszechnie. Co więcej, coraz częściej ustala się je zgodnie z uznanymi w światowej metodologii standardami zaproponowanymi na początku XXI wieku przez American Association for Public Opinion Research (AAPOR 2016).

Warto jednocześnie zauważyć, że pomimo licznych dowodów na niski stopień korelacji pomiędzy wskaźnikami terenowej realizacji próby oraz jakością pomiaru (Keeter i in. 2000; Curtin, Presser, Singer 2005; Groves 2006; Cornesse, Bosnjak 2018), wskaźniki realizacji próby pozostają nadal powszechnie stosowanymi miarami oceny jej jakości (Biemer, Lyberg 2003). Coraz częściej uznaje się je jednak za miary wtórne wobec innych wskaźników pozwalających na bezpośrednią ocenę jakości próby (Couper, De Leeuw 2003; Schouten, Cobben, Bethlehem 2009; Mohler 2019). Te ostatnie wymagają jednak dostępu do niestandardowych danych o wylosowanych jednostkach, rzadko kiedy raportowanych w dokumentacji metodologicznej projektów badawczych. Źródłem takich dodatkowych informacji o jednostkach, pozwalających na wyliczenie bardziej miarodajnych mierników jakości sondażu, mogą być jednak

- (1) dane zawarte w operatach doboru próby oraz w repozytoriach statystyki publicznej (Johnson i in. 2006; Kreuter, Müller, Trappmann 2010; Sakshaug, Kreuter 2012),
- (2) badania *follow-up* przeprowadzane po badaniach właściwych na grupie wszystkich lub losowo dobranych jednostek niedostępnych (Hansen, Hurwitz 1946; Stoop 2004; Olson 2006; Peytchev, Baxter, Carley-Baxter 2009),
- (3) paradane charakteryzujące fazę kontaktu ankietera z respondentem (Beaumont 2005; Matsuo i in. 2010; West 2011; Lyberg 2012; Kreuter, Olson 2013),



- (4) charakterystyki jednostek niedostępnych pozyskiwane metodą Basic Question Procedure (Kersten, Bethlehem 1984; Bethlehem, Kersten 1985) lub jej zmieniona wersja w postaci metodologii PEDAKSI (Lynn 2003).

Uzyskiwanie takich dodatkowych informacji wymaga jednak włożenia wielu wysiłku i dodatkowych środków finansowych, trudno zatem wyobrazić sobie podejmowanie takich działań na dużą skalę w międzykrajowych projektach porównawczych. Wyjątkiem jest Europejski Sondaż Społeczny, w którym od niedawna przeprowadza się analizę jakości próby na podstawie charakterystyk miejsca zamieszkania respondentów zbieranych przez ankierów w trakcie realizacji wywiadów oraz dodatkowych danych pochodzących z rejestrów populacyjnych dotyczących wieku i płci (Beullens, Loosveldt 2016; Wuyts, Loosveldt 2019). Biorąc pod uwagę perspektywę użytkownika danych wtórnych, wskaźniki jakości próby opierać muszą się przy tym na danych standardowo gromadzonych w trakcie badań, udostępnianych publicznie oraz możliwych do zastosowania na dużą skalę. Innymi słowy, wymóg gromadzenia dodatkowych informacji o jednostkach wylosowanych do próby nie jest żadnym realnym rozwiązaniem w prowadzonej *ex post* analizie jakości sondaży.

## 2. Źródła danych, zakres inwentaryzacji praktyk sondażowych oraz ocena jakości dokumentacji metodologicznej projektów EB, ESS, EVS, ISSP

### 2.1. Baza empiryczna

Zaprezentowane w kolejnych częściach tego rozdziału analizy empiryczne opierają się na metodologicznej inwentaryzacji procedur badawczych wykorzystanych w czterech – opisanych w rozdziale pierwszym – europejskich projektach sondażowych o charakterze porównawczym<sup>1</sup>. Przedmiotem szczegółowych analiz uczyniono

- (1) siedemnaście jesiennych fal EB z lat 2001-2017, w tym 3 rundy tego projektu z lat 2001-2003 zrealizowane w krajach kandydujących oraz potencjalnych kandydatach do Unii Europejskiej (Candidate Countries Eurobarometer),
- (2) osiem rund ESS realizowanych w cyklach dwuletnich od 2002 roku,
- (3) cztery odsłony EVS z 1981, 1990, 1999 oraz 2008 roku<sup>2</sup>, a także
- (4) trzydzieści jeden edycji ISSP z okresu 1985-2015.

---

<sup>1</sup> Przedstawione w tym rozdziale porównania zaprezentowane zostały w dwóch wcześniejszych pracach autora (Jabkowski 2009; Jabkowski, Kołczyńska 2020). Zostały one jednak znacząco uzupełnione o dodatkowe wątki problemowe.

<sup>2</sup> Dokumentacja metodologiczna piątej rundy EVS z 2017, jak również wyników EB z 2018 roku oraz EVS z 2016, 2017 oraz 2018 roku nie była dostępna w momencie składania monografii do wydawnictwa.

Przeprowadzona inwentaryzacja objęła w sumie 60 różnych edycji czterech projektów oraz 1412 krajowych sondaży, przy czym zakres inwentaryzacji ograniczony został wyłącznie do sondaży pochodzących z krajów europejskich, nawet jeśli projekt badawczy (ISSP) przekraczał granice Europy (tabela 8).

Tabela 8. Podstawowe charakterystyki projektów EB, ESS, EVS, ISSP

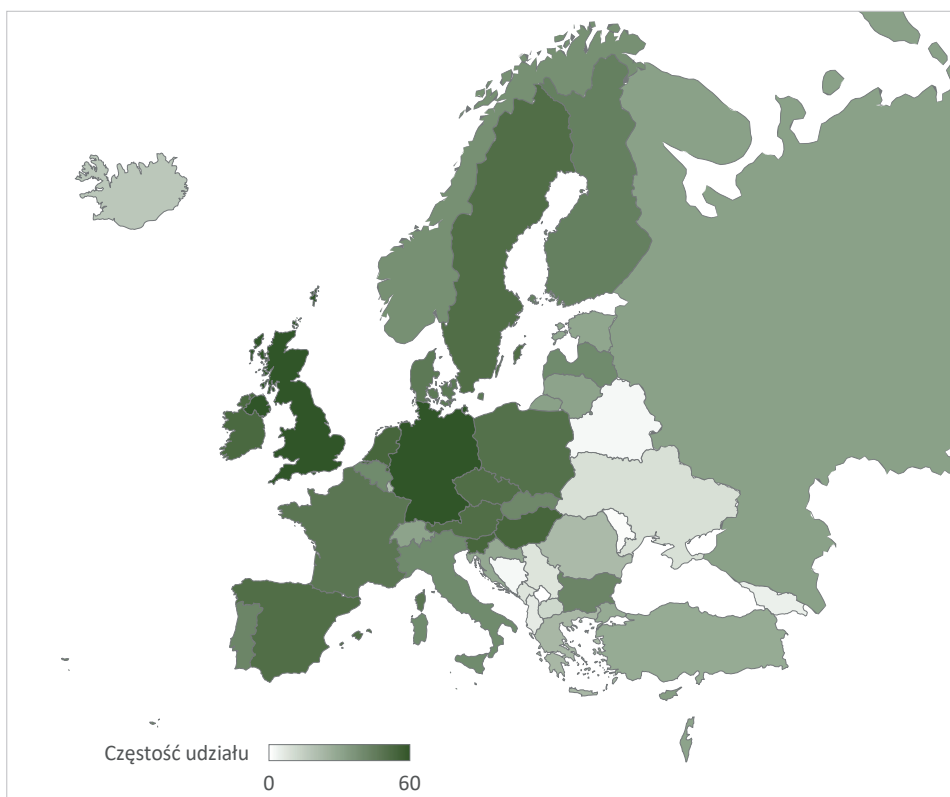
Symbol projektu	Nazwa projektu	Zakres czasowy	Liczba edycji	Liczba sondaży
EB	Eurobarometer	2001-2017	17	523
ESS	European Social Survey	2002-2016	8	199
EVS	European Values Study	1981-2008	4	112
ISSP	International Social Survey Programme	1985-2015	31	578
Razem			60	1412

Uwaga: <sup>①</sup> Jesienna fala standardowych Eurobarometrów

Źródło: opracowanie własne.

Rysunek 11 przedstawia częstość udziału krajów europejskich we wszystkich 60 różnych edycjach czterech analizowanych w tej pracy projektów o charakterze porównawczym. Przyglądając się zaprezentowanemu danym, zauważyć można znacznie częstszy udział państw Europy Zachodniej (włączając w to kraje południa, takie jak Hiszpania, Portugalia oraz Włochy), a także państw Europy Północnej (w tym przede wszystkim Finlandii i Szwecji) oraz Europy Środkowej (szczególnie krajów Grupy Wyszehradzkiej). Z drugiej strony, reprezentacja krajów Europy Wschodniej oraz Południowo-Wschodniej (w szczególności państw Półwyspu Bałkańskiego) była już zdecydowanie mniejsza. Występowanie znaczącego skrzywienia przejawiającego się w częstszym udziale w paneuropejskich projektach międzykrajowych państw zachodniej części kontynentu zauważone zostało w wielu wcześniejszych studiach empirycznych i metodologicznych (por. Słomczyński, Tomescu-Dubrow 2006; Wymułek 2018). Istotnie, nawet jeżeli z prezentowanych zestawień wyłączy się Eurobarometr (specyficzny z uwagi na zasięg terytorialny ograniczony do obszaru Unii Europejskiej), to i tak Niemcy oraz Wielka Brytania obecne były we wszystkich – bez wyjątku – edycjach ESS, EVS oraz ISSP, a pozostałe kraje Europy Zachodniej, Północnej i Środkowej w zdecydowanej większości z nich. Z drugiej strony, Mołdawia wzięła udział tylko w jednej rundzie EVS, a Białoruś jedynie w dwóch edycjach EVS, Bośnia i Hercegowina partycypowały zaledwie w jednej rundzie ESS oraz jednej edycji ISSP, natomiast Gruzja w jednej rundzie EVS oraz trzech ISSP. Cztery kraje europejskie – Andora, Liechtenstein, Monako oraz San Marino – nie brały w ogóle udziału w analizowanych projektach. Pamiętać należy jednak o tym, że poza „wielką czwórką” sondaży porównawczych,





Rys. 11. Częstość udziału krajów europejskich w kolejnych edycjach porównywanych projektów  
 Źródło: opracowanie własne.

w niektórych regionach Europy prowadzone są inne międzykrajowe projekty (por. Słomczyński, Tomescu-Dubrow, Jenkins 2016), czego najlepszym przykładem są Bałkany, będące polem intensywnej eksploracji sondażowej w projektach znajdujących się poza spectrum analiz zaprezentowanych w tej książce (zob. Kołczyńska 2014; Baranowski, Jabkowski 2019).

## 2.2. Charakterystyka źródeł danych wykorzystanych w metodologicznej inwentaryzacji projektów EB, ESS, EVS oraz ISSP

Dokumentami pozwalającymi na przeprowadzenie metodologicznej inwentaryzacji procedur wykorzystanych na etapie doboru sondażowych prób badawczych oraz ich terenowej realizacji w surveyach prowadzonych w kolejnych edycjach EB, ESS, EVS oraz ISSP były raporty oraz sprawozdania metodologiczne towarzyszące zbiorom danych wynikowych. Należy przy tym zaznaczyć, że raporty metodologiczne oraz sprawozdania techniczne z realizacji badań przyjmowały całkowicie

odmienne formy oraz charakteryzowały się różnym stopniem międzyprojektowej, międzykrajowej oraz wzdłużczasowej standaryzacji. Dla przykładu na metodologiczną dokumentację Eurobarometru składały się wyłącznie ogólne sprawozdania udostępniane na stronach internetowych projektu (niezawierające szczegółowych specyfikacji realizacji krajowych prób badawczych), podczas gdy w projekcie ESS, jak również w realizowanych od drugiej połowy lat 90. XX wieku edycjach EVS oraz ISSP, dokumentacja projektowa przyjmowała postać dość rozbudowanych opracowań zawierających opis schematów losowania oraz procedur terenowej realizacji krajowych sondaży. W projekcie ESS metodologiczne charakterystyki krajowych prób badawczych odnaleźć można w dokumentach noszących nazwę ESS Data Documentation Reports, podczas gdy źródłem informacji o realizacji EVS są Methods Reports oraz Variable Reports. Z kolei w ISSP dokumenty metodologiczne każdej z krajowych prób sondażowych publikowane są niezależnie w opracowaniach noszących nazwę ISSP Study Description<sup>3</sup>.

### 2.3. Zakres metodologicznej inwentaryzacji projektów

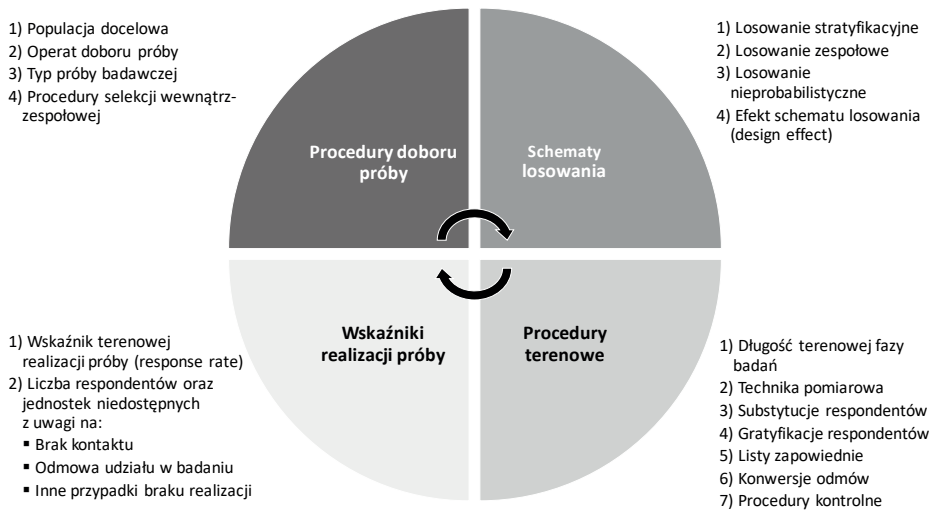
Dokumentacja oraz raporty metodologiczne każdego z czterech rozpatrywanych w tej pracy międzykrajowych projektów porównawczych przeszukiwane były ze szczególnym uwzględnieniem informacji o czterech aspektach procesu realizacji krajowych sondaży, w tym

- (1) procedurach doboru próby (tj. sposobach definiowania badanych populacji, typach operatów losowania jednostek, typach prób badawczych oraz procedurach selekcji respondentów docelowych w próbach adresowych),
- (2) schematach losowania próby (tj. zastosowaniu losowania warstwowego, zespołowego, nieprobabilistycznego oraz wyznaczaniu miar efektu losowania próby),
- (3) procedurach terenowych (tj. długości terenowej fazy badań, zastosowanych technik badawczych, substytucjach jednostek niedostępnych, stosowaniu listów zapowiednich, gratyfikacji za udział w badaniu, konwersji odmów oraz stosowaniu procedur kontrolnych), a także
- (4) na wskaźnikach terenowej realizacji próby, w tym liczbie respondentów oraz jednostek niedostępnych.

Rysunek 12 przedstawia schemat przeprowadzonej inwentaryzacji krajowych sondaży w projektach EB, ESS, EVS oraz ISSP.

---

<sup>3</sup> Odniesienia bibliograficzne do wszystkich zbiorów danych oraz do raportów metodologicznych wykorzystanych w analizach przedstawianych w kolejnych częściach tego rozdziału zawarte zostały w Aneksie.



Rys. 12. Zakres metodologicznej inwentaryzacji projektów

Źródło: opracowanie własne.

Dokładną specyfikację wszystkich zmiennych wykorzystanych w procesie metodologicznej inwentaryzacji czterech projektów zawiera tabela 9. Czytelnika zainteresowanego dostępem do zbioru, zawierającego dokładne informacje o procedurach doboru próby, schematach losowania, procedurach terenowych oraz wskaźnikach terenowej realizacji prób w każdym z 1412 krajowych sondaży odsłać można do elektronicznych materiałów uzupełniających stanowiących załącznik do tej publikacji (zob. Jabkowski i Kołczyńska 2020b).

Tabela 9. Charakterystyki projektów EB, ESS, EVS oraz ISSP poddane inwentaryzacji metodologicznej – klucz kodowy

Charakterystyki sondaży	Sposób kodowania wartości zmiennych
	1. Procedury doboru próby:
<i>Populacja docelowa</i>	[-9999] brak danych; [1] dolny limit wieku (w latach); [2] górny limit wieku (w latach)
<i>Operat doboru próby</i>	[-9999] brak danych; [1] imienny; [2] adresowy; [3] telefoniczny; [4] brak operatu (losowanie przestrzenne); [5] dobór nielosowy
<i>Typ próby</i>	[-9999] brak danych; [1] nieprobabilistyczne kwotowe; [2] losowe z komponentem random route; [3] wielostopniowe adresowe; [4] wielostopniowe imienne; [5] prosta lub warstwowa próba imienna; [6] wielostopniowa próba niesprecyzowana;
<i>Procedury selekcji respondenta wewnątrz gospodarstwa domowego</i>	[-9999] brak danych; [0] nie ma zastosowania; [1] siatka Kisha; [2] inne losowe; [3] ostatnie urodziny; [4] następne urodziny; [5] najbliższe urodziny; [6] urodziny bez specyfikacji; [7] nielosowe

Charakterystyki sondaży	Sposób kodowania wartości zmiennych
2. Schematy losowania:	
<i>Schemat losowania warstwowego</i>	[-9999] brak danych; [0] nie zastosowano; [1] zastosowano
<i>Schemat losowania zespołowego</i>	[-9999] brak danych; [0] nie zastosowano; [1] zastosowano
<i>Schemat nieprobabilistyczny</i>	[-9999] brak danych; [0] schemat losowy; [1] schemat nielosowy
<i>Efekt schematu losowania</i>	[-9999] brak danych; [0] nie wyliczono efektu schematu doboru próby; [1] podano wartość DEFF (design effect)
3. Procedury terenowej realizacji próby:	
<i>Okres realizacji badania</i>	[-9999] brak danych; [1] data rozpoczęcia fazy terenowej; [2] data zakończenia fazy terenowej
<i>Technika pomiarowa</i>	[-9999] brak danych; [1] PAPI; [2] CAPI; [3] F2P (PAPI lub CAPI); [4] CATI; [5] CAWI; [6] Ankieta rozsyłana; [7] Ankieta roznoszona
<i>Substytucje</i>	[-9999] brak danych; [0] brak możliwości; [1] dopuszczono
<i>Gratyfikacje respondentów</i>	[-9999] brak danych; [0] nie zastosowano; [1] zastosowano
<i>Listy zapowiednie</i>	[-9999] brak danych; [0] nie zastosowano; [1] zastosowano
<i>Konwersje odmów</i>	[-9999] brak danych; [0] nie zastosowano; [1] zastosowano
<i>Procedury kontrolne</i>	[-9999] brak danych; [0] nie zastosowano; [1] zastosowano
4. Wskaźniki terenowej realizacji próby:	
<i>Wskaźnik realizacji próby (response rate)</i>	[-9999] brak danych; [1] nie podano wartości response rate; [1] podano wartość response rate
<i>Struktura próby zrealizowanej</i>	[-9999] brak danych; [1] liczba zrealizowanych wywiadów; [2] liczba odmów, [3] brak kontaktu; [4] inne przypadki niezrealizowania wywiadów; [5] wylosowano jednostki spoza populacji

Źródło: opracowanie własne.

Inwentaryzując procedury wykorzystane w doborze prób badawczych, zwrócono w pierwszej kolejności uwagę na sposoby definiowania badanych populacji, co w praktyce sprowadzało się do kodowania (dla każdego krajowego sondażu) dolnego oraz górnego limitu wieku osób objętych badaniem. Działanie takie miało na celu sprawdzenie, czy we wszystkich krajach uczestniczących w danej rundzie określonego projektu obowiązywała ujednolicona definicja populacji, jak również ocenę tego, czy w kolejnych edycjach projektów zachodziły w tym względzie zmiany. W drugiej kolejności zwrócono również uwagę na typ operatu wykorzystanego w losowaniu jednostek, przyporządkowując krajowe sondaże do jednej z pięciu rozłącznych kategorii, w tym do

- (1) operatów imiennych,
- (2) operatów adresowych (wykazów budynków lub gospodarstw domowych),
- (3) operatów abonentów telefonicznych oraz do dwóch kategorii opisujących przypadki,

- (4) braku wiarygodnych operatów prowadzących do konieczności losowania probabilistycznych prób przestrzennych oraz
- (5) braku wiarygodnych operatów prowadzących do nieprobabilistycznych prób kwotowych.

Co więcej, jeżeli w dokumentacji projektowej brakowało informacji o typie operatu, to fakt ten traktowano jako brak danych obniżających jakość dokumentacji. W trzeciej kolejności zwrócono również uwagę na typ próby badawczej, przypisując krajowe sondáže do jednej z sześciu kategorii, w tym

- (1) do prób kwotowych typu nieprobabilistycznego oraz prób probabilistycznych obejmujących
- (2) imienne próby proste (lub warstwowe),
- (3) wielostopniowe próby oparte na operatach imiennych,
- (4) wielostopniowe adresowe próby osób oparte na wykazach budynków lub gospodarstw domowych,
- (5) wielostopniowe próby adresowe z elementami *random route* (wykorzystywane w sytuacji braku dostępu do wiarygodnych operatów imiennych lub adresowych) oraz
- (6) wielostopniowe próby adresowe bez dokładnej specyfikacji poszczególnych etapów losowania. Co więcej, ponieważ dokumentacja metodologiczna krajowych sondaży mogła zawierać informacje niepozwalające na szczegółowe określenie typu próby, mogła także nie zawierać żadnej informacji w tym zakresie, to obie sytuacje traktowano jako uchybienia dokumentacji obniżające jej jakość.

Po czwarte, dokonano także identyfikacji procedur selekcji wewnętrzzespołowej respondentów docelowych w próbach adresowych, w tym zastosowanie

- (a) siatki Kisha,
- (b) innych procedur losowych, a także quasi-losowych procedur opartych na danych
  - ostatnich urodzin,
  - następnym urodzin,
  - urodzin przypadających najbliższej dacie wizyty ankietera,
  - urodzin (bez dokładnej specyfikacji kryteriów), jak również
- (c) procedur nielosowych (dowolnych).

Jeżeli w dokumentacji pojawiała się informacja o zastosowaniu losowania wewnętrzzespołowego, jednak bez dokładnej specyfikacji, o jaką procedurę może chodzić, to sytuację taką traktowano jako brak danych obniżający jakość dokumentacji.

Drugim obszarem metodologicznej inwentaryzacji projektów były schematy doboru próby, przy czym kodowano informację o zastosowaniu (1) schematu lo-

sowania warstwowego oraz (2) schematu losowania zespołowego. Ponadto, ponieważ konsekwencją doboru próby według schematu innego niż losowanie proste (lub warstwowe proporcjonalne losowanie proste) jest utrata precyzji estymacji (o czym była już mowa w poprzednim podrozdziale), to sprawdzano również, czy w dokumentacji projektowej podana została wartość miernika efektu schematu losowania (*design effect*).

Trzecim obszarem inwentaryzacji metodologicznej krajowych sondaży były procedury terenowe. Poszukiwano przy tym informacji dotyczących

- (1) czasu trwania terenowej fazy badań (liczba dni liczona na podstawie daty rozpoczęcia oraz zakończenia prac terenowych),
- (2) techniki lub technik badawczych wykorzystanych do gromadzenia danych (wywiady osobiste z kwestionariuszem papierowym (PAPI), wywiady osobiste wspomagane komputerowo (CAPI), wywiady osobiste bez specyfikacji narzędzia (PAPI lub CAPI), wywiady telefoniczne (CATI), ankiety webowe lub ankiety rozsyłane drogą elektroniczną (CAWI), ankiety pocztowe, ankiety roznoszone przez ankietera do samodzielnego wypełnienia),
- (3) stosowania substytucji jednostek niedostępnych respondentami zastępczymi,
- (4) występowania gratyfikacji za udział w badaniu,
- (5) stosowania listów zapowiednich informujących o celu badania oraz przewidywanej dacie wizyty ankietera,
- (6) stosowania procedur konwersji odmów, a także
- (7) stosowania procedur kontrolnych weryfikujących poprawność realizacji wywiadów oraz pracę ankietarów.

Czwartym i ostatnim obszarem inwentaryzacji krajowych sondaży były efekty terenowej realizacji próby, które kodowano zgodnie ze standardami wypracowanymi przez *American Association for Public Opinion Research* (AAPOR 2016). Sprowadzało się to do ustalenia

- (1) liczby respondentów, tj. liczby wywiadów zrealizowanych w całości lub w części,
- (2) liczby osób, z którymi nie udało się nawiązać kontaktu,
- (3) liczby osób odmawiających udziału w badaniu,
- (4) liczby osób niedostępnych z innych powodów,
- (5) liczby osób o nieustalonym statusie przynależności do populacji oraz
- (6) liczby osób, które do populacji nie należą.

Odnutowywano również podaną w dokumentacji projektowej wartość wskaźnika terenowej realizacji próby (o ile informacje takie były dostępne), co w połączeniu z gromadzonymi i informacjami o terenowej realizacji próby (tj. o jednostkach dostępnych i niedostępnych) miało pozwolić na weryfikację tego, czy wskaźnik *response rate* wyznaczono zgodnie ze standardem AAPOR.

## 2.4. Ocena jakości dokumentacji metodologicznej projektów

Zanim zaprezentowane oraz omówione zostaną międzyprojektowe oraz międzykrajowe zróżnicowania procedur doboru oraz terenowej realizacji prób badawczych w krajowych sondażach prowadzonych w ramach kolejnych edycji czterech rozpatrywanych w tej pracy projektów porównawczych, analizie oraz szczegółowej ocenie poddana zostanie jakość dokumentacji projektowej, która przez użytkowników danych wtórnych traktowana jest często jako wyznacznik przydatności danych do analiz substancywnych oraz (choć niesłusznie) jako miernik jakości pomiaru (Biemer 2010b). Warto przy tym zaznaczyć, że przeprowadzona analiza jakości dokumentacji opierała się na ocenie kompletności raportów metodologicznych, tj. na ustaleniu tego, czy i w jakim stopniu projektowe sprawozdania zawierały kluczowe dane pozwalające na opisanie procedur badawczych w krajowych sondażach. Ponieważ przedstawiona w poprzedniej części tego rozdziału metodologiczna inwentaryzacja procedur badawczych skupiała się na czterech głównych obszarach praktyki sondażowej, w tym na (1) procedurach losowania, (2) schematach doboru jednostek, (3) procedurach terenowych oraz (4) raportowaniu o wskaźnikach realizacji próby, to miary jakości dokumentacji skonstruowane zostały osobno dla każdego z tych obszarów. Dodatkowo wyznaczono ogólny wskaźnik jakości dokumentacji projektowej uśredniając cztery miary składowe (tabela 10).

Tabela 10. Wskaźniki jakości dokumentacji metodologicznej projektów międzykrajowych

Wskaźniki	Sposób kodowania
1. Jakość dokumentacji w zakresie procedur doboru próby <sup>01</sup> :	
<i>Populacja docelowa</i>	0: brak informacji o zakresie wieku populacji docelowej 1: informacje dostępne w dokumentacji projektowej
<i>Operat doboru próby</i>	0: brak informacji o wykorzystanym operacie 1: informacje dostępne w dokumentacji projektowej
<i>Typ próby</i>	0: brak informacji o typie próby badawczej 1: informacje dostępne w dokumentacji projektowej
<i>Procedury selekcji respondenta wewnątrz gospodarstwa domowego</i>	0: brak informacji o procedurach selekcji wewnątrz-zespołowej 1: informacje dostępne w dokumentacji projektowej
2. Jakość dokumentacji w zakresie schematów losowania <sup>01</sup> :	
<i>Schemat losowania próba warstwowa</i>	0: brak informacji o losowaniu warstwowym 1: informacje dostępne w dokumentacji projektowej
<i>Schemat losowania: próba zespołowa</i>	0: brak informacji o losowaniu zespołowym 1: informacje dostępne w dokumentacji projektowej
<i>Schemat losowania: nieprobabilistyczny</i>	0: brak informacji o doborze nieprobabilistycznym 1: informacje dostępne w dokumentacji projektowej
<i>Efekt schematu losowania</i>	0: nie podano wartości design effect 1: podano wartość design effect

Wskaźniki	Sposób kodowania
3. Jakość dokumentacji w zakresie procedur terenowej realizacji próby <sup>(1)</sup> :	
<i>Okres realizacji badania</i>	0: brak informacji o dacie rozpoczęcia oraz zakończenia badań 1: informacje dostępne w dokumentacji projektowej
<i>Technika pomiarowa</i>	0: brak informacji o technice pomiarowej 1: informacje dostępne w dokumentacji projektowej
<i>Substytucje respondentów docelowych</i>	0: brak informacji o występowaniu substytucji 1: informacje dostępne w dokumentacji projektowej
<i>Gratyfikacje respondentów</i>	0: brak informacji o gratyfikacjach respondentów 1: informacje dostępne w dokumentacji projektowej
<i>Listy zapowiednie</i>	0: brak informacji o listach zapowiednich 1: informacje dostępne w dokumentacji projektowej
<i>Konwersje odmów</i>	0: brak informacji o stosowaniu konwersji odmów 1: informacje dostępne w dokumentacji projektowej
<i>Procedury kontrolne realizacji próby</i>	0: brak informacji o zastosowanych procedurach kontrolnych 1: informacje dostępne w dokumentacji projektowej
4. Jakość dokumentacji w zakresie wskaźników terenowej realizacji próby <sup>(1)</sup> :	
<i>Wskaźnik realizacji próby</i>	0: brak informacji o wartościach response rate 1: informacje dostępne w dokumentacji projektowej
<i>Liczba respondentów oraz jednostek niedostępnych</i>	0: brak informacji w dokumentacji projektowej 1: informacje dostępne w dokumentacji projektowej

Uwaga: <sup>(1)</sup> Dla każdego pojedynczego sondażu (*projekt\*edycja\*sondaż*) wartość wskaźnika jest średnią arytmetyczną z kodów przypisanych określonoemu typowi informacji zawartej w dokumentacji projektowej. Średnia przyjmuje wartości z przedziału [0;1].

Źródło: opracowanie własne.

Dla każdego z czterech obszarów wybrano zestaw od dwóch do sześciu zmiennych (zero-jedynkowych), opisujących zawartość informacyjną raportów oraz sprawozdań metodologicznych, przy czym jeżeli w dokumentacji projektowej nie występował poszukiwany typ informacji, to zmienna przyjmowała wartość równą zero. Co więcej, każdy z czterech szczegółowych wskaźników (oraz wskaźnik ogólny) wyliczony został niezależnie dla każdego z 1412 krajowych sondaży (*projekt\*edycja\*sondaż*) jako średnia arytmetyczna z kodów 0-1, przy czym owe wartości zostały uśrednione dla kolejnych edycji projektów (*projekt\*edycja*) oraz każdego projektu z osobna. Taki sposób oceny jakości dokumentacji badawczej stosowany był we wcześniejszych opracowaniach metodologicznych, m.in. w pracach Jabkowski i Kołczyńska (2020a), Jabkowski (2019), Kołczyńska i Schoene (2018) oraz Kołczyńska (2014).

Tabela 11 zawiera uśrednione wartości wskaźników jakości dokumentacji projektów EB, ESS, EVS oraz ISSP. Podano również wartość odchylenia standardowego jako



miarę wewnątrzprojektowego zróżnicowania jakości dokumentacji. Należy przy tym pamiętać, że ponieważ raporty metodologiczne EB przyjmowały postać ogólnych sprawozdań (bez wyróżnienia procedur badawczych wdrażanych w krajowych sondażach), to wartości odchylenia standardowego dla EB osiągają wartość równą zero.

Tabela 11. Porównanie jakości dokumentacji metodologicznej projektów EB, ESS, EVS, ISSP

Oznaczenie projektu	Liczba sondaży	Wskaźniki jakości dokumentacji metodologicznej <sup>(i)</sup>				Ogólny wskaźnik jakości <sup>(ii)</sup>
		Procedury losowania	Schematy doboru próby	Procedury terenowe	Realizacja próby	
EB <sup>(iii)</sup> : 2001-2003	84	0,75 (0,00)	0,80 (0,00)	0,33 (0,00)	0,00 (0,00)	0,47 (0,00)
EB <sup>(iii)</sup> : 2004-2017	439	1,00 (0,00)	0,80 (0,00)	0,33 (0,00)	0,00 (0,00)	0,53 (0,00)
ESS	199	0,99 (0,02)	0,92 (0,12)	0,99 (0,03)	1,00 (0,00)	0,98 (0,03)
EVS	112	0,84 (0,21)	0,34 (0,29)	0,56 (0,36)	0,59 (0,46)	0,58 (0,28)
ISSP	578	0,84 (0,31)	0,30 (0,20)	0,71 (0,33)	0,57 (0,36)	0,61 (0,25)

Uwagi: <sup>(i)</sup> Średnia wartość krajowych wskaźników (*projekt\*edycja\*sondaż*) wraz z odchyleniem standardowym; <sup>(ii)</sup> Dla EB dostępna jest wyłącznie ogólna dokumentacja projektowa (*projekt\*edycja*) bez podania szczegółowych informacji o metodologii pomiaru w każdym kraju (*projekt\*edycja\*sondaż*).

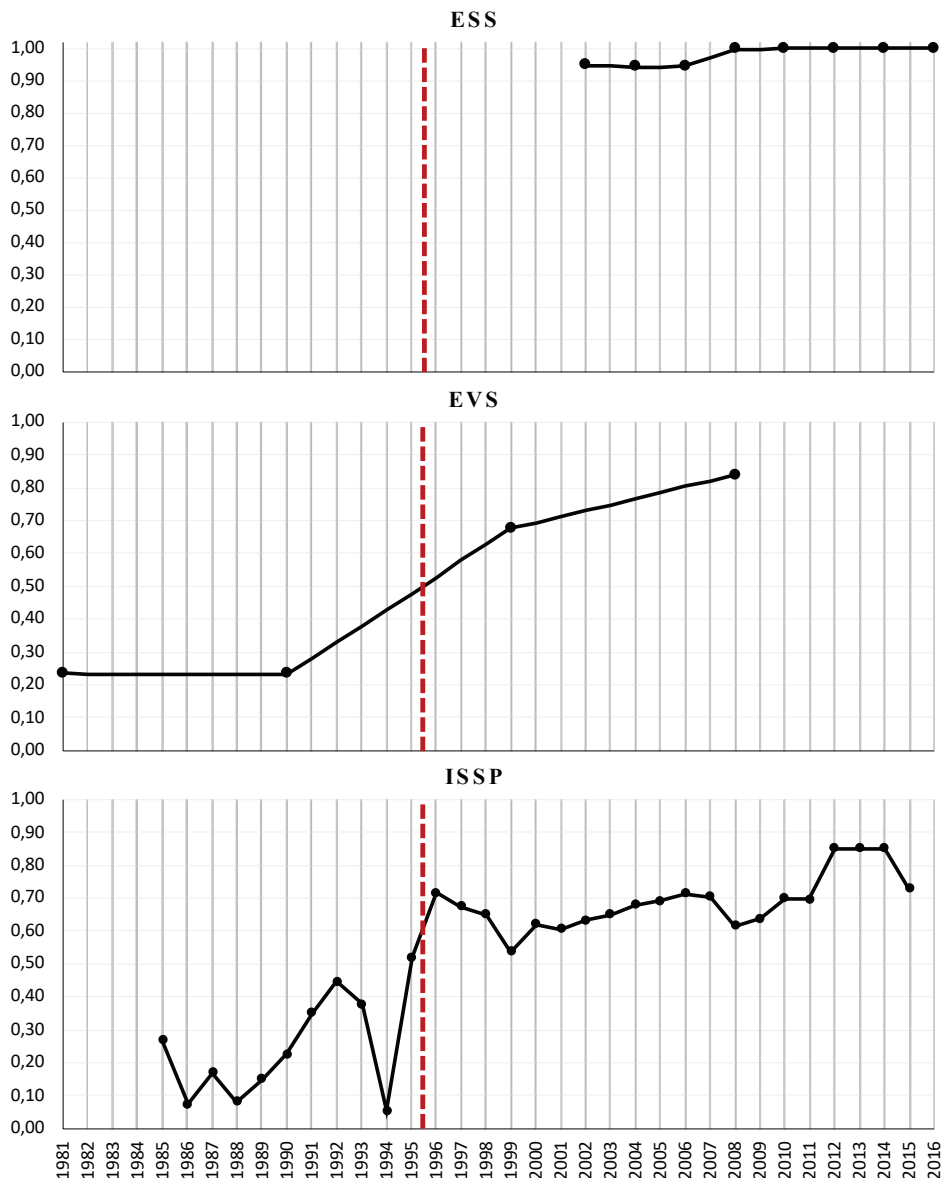
Źródło: opracowanie własne.

Przyglądając się danym przedstawionym w powyższej tabeli, można zauważyć, iż EB stanowi najslabiej udokumentowany metodologicznie projekt badawczy spośród czterech analizowanych w tej pracy. Wystarczy wskazać, że w EB nie informuje się w ogóle o wskaźnikach terenowej realizacji próby oraz bardzo słabo dokumentuje procedury terenowe. Oznacza to, że Eurobarometr stanowiący projekt zupełnie fundamentalny dla badaczy podejmujących analizy nad społeczną oraz systemową integracją dokonującą się na poziomie Unii Europejskiej, cechuje się jednocześnie największą niepewnością, co do stosowanych standardów metodologicznych. Fakt ten był już zresztą przedmiotem ostrej krytyki formułowanej w odniesieniu do użyteczności EB dla zorientowanych europejsko analiz substancyjnych (Nissen 2014).

Najlepiej udokumentowanym projektem pozostaje przy tym ESS, który charakteryzuje się wysoką standaryzacją oraz znaczącą międzykrajową spójnością raportowania o procedurach doboru oraz terenowej realizacji próby (w porównaniu do innych projektów uzyskano najwyższe średnie wartości każdego ze wskaźników oraz najniższe wartości odchylenia standardowych). Jakość dokumentacji EVS oraz ISSP okazała się przy tym już istotnie niższa oraz znacząco bardziej zróżnicowana (stosunkowo dobrze udokumentowano procedury losowania, w tym przyjęte definicje populacji, rodzaje operatów, typy prób badawczych oraz procedury selekcji wewnątrzspołowej, słabiej natomiast procedury terenowe oraz wskaź-

niki realizacji próby, najslabiej natomiast schematy losowania). Co więcej, jakość dokumentacji EVS oraz ISSP charakteryzowała się znaczącymi zmianami w czasie.

Wykresy przedstawione na rysunku 13 ukazują zmiany jakości dokumentacji kolejnych edycji ESS, EVS oraz ISSP. Z zestawień wyłączono EB, dla którego



Rys. 13. Zmiany jakości dokumentacji metodologicznej projektów w czasie

Źródło: opracowanie własne.

jedyną zmianą było włączenie do dokumentacji projektowej w 2004 roku informacji o stosowaniu dat urodzin w wewnątrzgrupowej selekcji respondentów docelowych; we wszystkich edycjach poprzedzających rok 2004 oraz w kolejnych edycjach po tym roku, dokumenty metodologiczne EB zawierały niezmiennie taki sam zestaw informacji.

Analizując zmiany jakości dokumentacji projektowej (w okresie obejmującym ponad 35 lat), wyraźnie uwidacznia się stosowanie najwyższych standardów raportowania o metodologii krajowych sondaży w (stosunkowo młodym) projekcie ESS. Istotnie, dla najslabiej opisanej drugiej rundy tego projektu z 2004 roku, jakość dokumentacji okazała się i tak znacząco wyższa od najlepiej opisanej edycji któregośkolwiek z pozostałych projektów. Fakt ten nie powinien być jednak żadnym zaskoczeniem, gdyż ESS projektowano od samego początku z zamierzeniem implementacji nie tylko najlepszych możliwych praktyk sondażowych, ale również wysokich standardów raportowania o procedurach badawczych. W tym sensie należy docenić, że po przełomowym dla ISSP raporcie krytykującym niską jakość dokumentacji krajowych sondaży prowadzonych w ramach wczesnych edycji tego projektu (Park, Jowell 1997), który to odbił się szerokim echem w środowisku metodologów badań sondażowych na całym świecie, znacznie poprawiło się raportowanie o krajowych praktykach sondażowych nie tylko w ISSP, ale też w EVS. Oba programy charakteryzuje bardzo niska jakość dokumentacji wczesnych edycji oraz znacząco wyższa jakość dokumentacji edycji po 1995 roku. Najnowsze odsłony ISSP oraz EVS nie odbiegają już tak bardzo pod względem jakości raportowania o operatach losowania, typach prób badawczych, procedurach selekcji wewnątrzgrupowej, czy też miernikach terenowej realizacji próby, choć nadal nie dokumentują tak dokładnie (jak ESS) procedur terenowych oraz schematów losowania jednostek.

### **3. Porównawcze analizy zróżnicowania praktyk sondażowych w projektach EB, ESS, EVS, ISSP**

#### **3.1. Sposoby definiowania badanych populacji**

Definicje badanych populacji zamieszczone zostały w dokumentacji metodologicznej wszystkich czterech analizowanych projektów porównawczych. Odpowiednio można wskazać, że w EB na populację składała się zbiorowość wszystkich obywateli państw członkowskich (lub kandydujących do UE), w wieku 15 lat i więcej (por. European Commission 2017). W projekcie ESS w skład populacji wchodziły wszystkie osoby w wieku 15 lat i więcej zamieszkujące w prywatnych gospodarstwach domowych niezależnie od ich narodowości, obywatelstwa lub statusu prawnego (por. ESS 2018); w EVS przyjęto praktycznie taką samą definicję

populacji, jedyna różnica w porównaniu do ESS dotyczyła dolnej granicy wieku ustalonej na poziomie 18 lat (por. EVS 2011). Z kolei w skład populacji w badaniach ISSP wchodziły wszystkie osoby w wieku 18 lat i więcej (por. ISSP 2017). Z dokumentacji projektowych wynikało również, że definicje populacji pozostawały niezmiennie w czasie.

Czynnikiem różnicującym krajowe sondaże (w ramach projektów) okazał się jednak sposób określania populacji docelowych, czyli takich zbiorowości, które w praktyce (a nie tylko w założeniu) objęte zostały pomiarem (tabela 12). Na wstępie warto jednak zauważyć, że we wszystkich krajowych próbach zrealizowanych w ramach EB oraz ESS przedziały wieku osób wchodzących w skład populacji docelowych były jednakowe i nie występowały różnice pomiędzy krajowymi sondażami. Znaczące zróżnicowania pojawiły się już jednak w EVS (w 7 sondażach na 103 wprowadzono górny limit wieku) oraz ISSP (we wszystkich edycjach można zauważyć liczne międzysondażowe różnice w określeniu dolnej oraz górnej granicy wieku osób należących do populacji).

Tabela 12. Zróżnicowanie dolnych oraz górnych ograniczeń wieku jednostek włączonych w skład populacji docelowych w projektach EB, ESS, EVS, ISSP

Symbol projektu	Brak danych <sup>(i)</sup>	Zbiorowość generalna	Zakres wieku populacji docelowej	
			Dolna granica	Górna granica
EB	0/484	15+	15 lat [484; 100%]	b.o. [484; 100%]
ESS	0/199	15+	15 lat [199; 100%]	b.o. [199; 100%]
EVS	2/112	18+	18 lat [110; 100%]	74 lata [1; 0,9%]
				75 lat [3; 2,7%]
				80 lat [3; 2,7%]
				b.o. [103; 93,7%]
				65 lat [3; 0,6%]
				69 lat [1; 0,2%]
				70 lat [4; 0,8%]
ISSP	72/578	18+	14 lat [3; 0,6%]	74 lata [31; 6,1%]
			15 lat [28; 5,5%]	75 lat [7; 1,4%]
			16 lat [45; 8,9%]	76 lat [2; 0,4%]
			17 lat [7; 1,4%]	78 lat [2; 0,4%]
			18 lat [415; 82,0%]	79 lat [36; 0,7%]
			19 lat [6; 1,2%]	80 lat [7; 1,4%]
			21 lat [2; 0,4%]	85 lat [8; 1,6%]
				94 lata [1; 0,2%]
				b.o. [403; 79,6%]

Uwagi: <sup>(i)</sup> Liczba krajowych sondaży, których dokumentacja nie zawierała informacji o górnym i dolnym limicie wieku populacji docelowej / ogólna liczba sondaży. <sup>(ii)</sup> W nawiasie kwadratowym podano liczbę krajowych sondaży oraz procent całości. <sup>(iii)</sup> Symbol: b.o. oznacza, iż nie wprowadzono górnego ograniczenia wieku.

Źródło: opracowanie własne.

Z wielu krajowych populacji wyłączano również takie kategorie osób, z którymi realizacja badań była problematyczna (Jabkowski 2015). Doskonałą egzemplifikacją takich praktyk okazują się operacyjne definicje populacji przyjmowane w projekcie ESS. Wystarczy przywołać badania realizowane w Izraelu, by wskazać, iż poza zakresem populacji znajduje się mniejszość palestyńska we Wschodniej Jerozolimie oraz na Zachodnim Brzegu Jordanu (ESS 2018), jak również pomiar przeprowadzony w Grecji, wyłączający jednostki populacji zamieszkujące słabo zaludnione wyspy archipelagu Cyklad oraz Dodekanez, czy też hiszpańską odśloną pierwszej edycji ESS (2002), w której z populacji badawczej wyłączono mieszkańców dwóch afrykańskich miast, tzn. Ceuta oraz Melilla. Ponieważ próba typologizacji wszystkich motywów stojących za decyzjami badaczy o wyłączeniu poza zakres badanej populacji pewnych kategorii jednostek, pozbawiona byłaby większego sensu, to dane zawarte w tabeli 12 podają jedynie informacje o wewnątrzprojektowych zróżnicowaniach w sposobach określania górnych oraz dolnych granic wieku osób objętych badaniami w poszczególnych projektach. Trzeba jednak pamiętać, że nawet zróżnicowania w zakresie takiej zmiennej jak wiek osób objętych badaniem (a więc w odniesieniu do jednej tylko charakterystyki populacji) okazuje się dla porównywalności danych niezwykle problematyczne. Obniżają bowiem nie tylko międzyprojektową porównywalność wyników, ale również możliwość prowadzenia międzykrajowych porównań w obrębie konkretnych projektów oraz wiążą się z koniecznością uwzględnienia owych zróżnicowań w analizach danych.

### **3.2. Rodzaje operatów doboru prób badawczych oraz schematy losowania jednostek**

Tabela 13 zawiera informacje o wykorzystaniu (w ramach poszczególnych sondażowych projektów porównawczych) określonych rodzajów operatów doboru prób oraz schematów losowania jednostek. Nawet pobieżna analiza wyników uwiadczenia znaczące międzyprojektowe zróżnicowania w tym zakresie. Dla przykładu w EB oraz ESS próby miały wyłącznie charakter probabilistyczny; w EB oparte były na próbach przestrzennych, natomiast w ESS głównie na operatach imiennych lub adresowych. W EVS znaczący był udział prób nielosowych (dla których typ operatu nie ma większego znaczenia) oraz prób probabilistycznych opartych na operatach adresowych, z kolei operaty imienne oraz próby przestrzenne wykorzystywane były z mniejszą częstotliwością. W ISSP dominowały operaty adresowe oraz rejestry imienne, natomiast pozostałe typy operatów wykorzystywane były już ze zdecydowanie mniejszą częstotliwością.

W większości krajowych sondaży zrealizowanych w czterech analizowanych projektach schemat losowania jednostek zawierał komponent doboru warstwowego oraz doboru zespołowego, przy czym warto mieć na względzie, że dokumen-

tację EVS oraz ISSP charakteryzowały poważne braki w tym zakresie. Co więcej, wskaźniki efektywności schematów losowania (DEFF) zawarto wyłącznie w dokumentacji ESS.

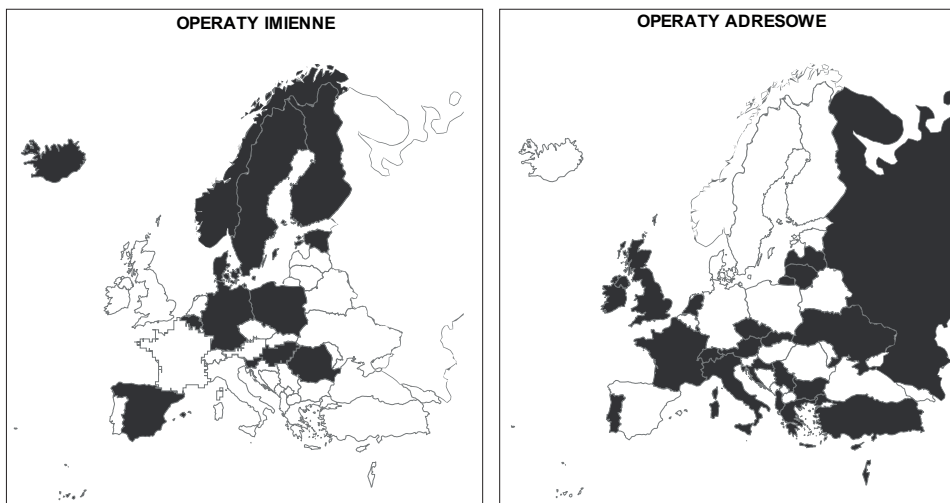
Tabela 13. Zróżnicowanie rodzajów operatów doboru prób badawczych oraz schematów losowania w projektach EB, ESS, EVS, ISSP

Aspekt porównań	EB	ESS	EVS	ISSP
Rodzaj operatu:				
[1] imienny	-	46,2%	17,3%	38,2%
[2] adresowy	-	45,2%	26,0%	47,5%
[3] telefoniczny	-	-	-	0,4%
[4] brak operatu / próby przestrzenne	100%	8,6%	18,3%	11,9%
[5] próba nielosowa	-	-	38,4%	2,0%
Brak danych / liczba sondaży	0/523	0/199	8/112	81/578
Schemat doboru warstwowego:				
[1] zastosowano warstwowanie	100%	81,2%	78,5%	70,2%
Brak danych / liczba sondaży	0/523	1/199	47/112	175/578
Schemat doboru zespołowego:				
[1] zastosowano wiązkanie	100%	64,1%	62,1%	63,3%
Brak danych / liczba sondaży	0/523	1/199	54/112	529/578
Efektywność schematu losowania:				
[1] podano wartość wskaźnika DEFF	-	63,8%	-	-
Brak danych / liczba sondaży	523/523	72/199	112/112	578/578

Źródło: opracowanie własne.

Na rysunku 14 zestawiono kraje europejskie z uwagi na dominujący typ operatów wykorzystywany w losowaniu sondażowych prób badawczych. Z uwagi na brak informacji o międzykrajowym zróżnicowaniu w stosowanych operatach doboru próby, z prezentowanych zestawień wyłączono dane dla EB.

Przyglądając się obu grafikom, można zauważyć, iż w niewielu europejskich krajach badacze mieli sposobność losowania jednostek na podstawie rejestrów imiennych. Rejestry tego typu dostępne były w krajach skandynawskich (Danii, Finlandii, Islandii, Norwegii, Szwecji), a także w Austrii, Belgii, Estonii, Hiszpanii, Polsce, Niemczech i na Węgrzech. Rejestry imienne, jeśli istniały, zasilane były najczęściej repozytoriami administracji publicznej. Charakteryzowały się przy tym bardzo wysoką jakością oraz kompletnością danych, tzn. zawierały aktualne informacje o prawie wszystkich jednostkach należących do populacji docelowych, charakteryzując się bardzo wysokimi wskaźnikami pokrycia populacji osiągnię-



Rys. 14. Kraje europejskie z uwagi na dominujący typ operatu

Źródło: opracowanie własne.

mi wartość powyżej 99 p.p. Z drugiej strony, operaty adresowe zasilane były rejestrami o zdecydowanie bardziej zróżnicowanym charakterze. Biorąc za przykład projekt ESS, można wskazać, że wykazy budynków mieszkalnych konstruowane były na podstawie repozytoriów zawierających kody pocztowe, natomiast operaty gospodarstw domowych budowane na wykazach abonentów telefonii stacjonarnej, komercyjnych bazach klientów usług komunalnych lub na urzędowych rejestrach gospodarstw domowych. W kontekście realizacji badań porównawczych pamiętać należy jednak o tym, że przydatność operatów adresowych w doborze prób do badań międzykrajowych zależy w głównej mierze od stopnia pokrycia jednostek populacji, nie zaś od pokrycia gospodarstw domowych, czy też pokrycia budynków mieszkalnych. Opierając się na danych ESS, można przykładowo wskazać, że choć stopień pokrycia badanych populacji operatami adresowymi był dość znaczny, osiągając poziom 95 p.p., to i tak okazywał się niższy niż w przypadku operatów imiennych. Fakt ten stanowi potencjalne źródło problemów w badaniach porównawczych, albowiem wpływa na zróżnicowanie jakości sondaży oraz wymaga każdorazowego rozpoznania przed włączeniem danych z różnych krajów do analizy.

### 3.3. Typy sondażowych prób badawczych oraz procedury selekcji respondentów docelowych w adresowych próbach osób

Typ próby badawczej oraz stosowanie określonej procedury służącej pośredniej selekcji respondentów docelowych zależy w zasadzie od tego, jaki rodzaj rejestru

jednostek populacji może zostać przez badaczy wykorzystany (lub inaczej, do jakiego operatu badacze mają dostęp) w procesie losowania próby sondażowej. Jak już zostało wcześniej powiedziane, operaty imienne dają możliwość bezpośredniego doboru jednostek populacji do próby, podczas gdy operaty adresowe, rejestry gospodarstw domowych oraz próby przestrzenne wiążą się z koniecznością losowania respondentów w sposób pośredni. W tabeli 14 zestawiono międzyprojekto- we zróżnicowanie wykorzystania określonych typów prób badawczych oraz procedur selekcji respondentów docelowych w próbach adresowych.

Tabela 14. Zróżnicowanie typów prób badawczych oraz procedur selekcji respondentów docelowych w adresowych próbach osób w projektach EB, ESS, EVS, ISSP

Aspekt porównań	EB	ESS	EVS	ISSP
Typ próby:				
[1] nieprobabilistyczna: kwotowa	-	-	48,0%	2,4%
[2] imienna prosta próba losowa	-	20,2%	6,9%	17,2%
[3] wielostopniowa imienna	-	26,3%	10,8%	19,6%
[4] wielostopniowa adresowa	-	44,9%	14,7%	49,4%
[5] wielostopniowa: random route	100%	8,6%	15,7%	8,6%
[6] wielostopniowa nieokreślona	-	-	3,9%	2,6%
Brak danych / liczba sondaży	0/523	1/199	6/112	102/578
Procedura selekcji wewnątrzspołecznej:				
[1] losowe: siatka Kisha	-	35,5%	40,5%	47,3%
[2] losowe: inne	-	-	-	1,4%
[3] quasi-losowe: daty urodzin	100%	64,5%	38,1%	44,9%
[4] nielosowe: dowolne	-	-	21,4%	6,4%
Brak danych / liczba sondaży	84/523	0/107	52/94	112/402

Źródło: opracowanie własne.

Przyglądając się przedstawionym zróżnicowaniom, można zauważyć, że w projekcie EB próby miały wyłącznie charakter wielostopniowy (gospodarstwa domowe losowane były z wykorzystaniem procedury *random route*), przy czym respondenci docelowi wybierani byli na podstawie dat urodzin. Specyfiką badań EVS jest znaczący udział prób nielosowych z kwotowym doбором jednostek. Należy przy tym podkreślić, że próby nielosowe występowały w tym projekcie wyłącznie w badaniu z 1981 oraz 1990 roku. Próby nielosowe pojawiały się także w ISSP, choć ich udział był w zasadzie marginalny. Z kolei w ESS występowały wyłącznie próby probabilistyczne, w tym wielostopniowe próby adresowe, wielostopniowe próby imienne oraz proste (warstwowe) próby losowe.

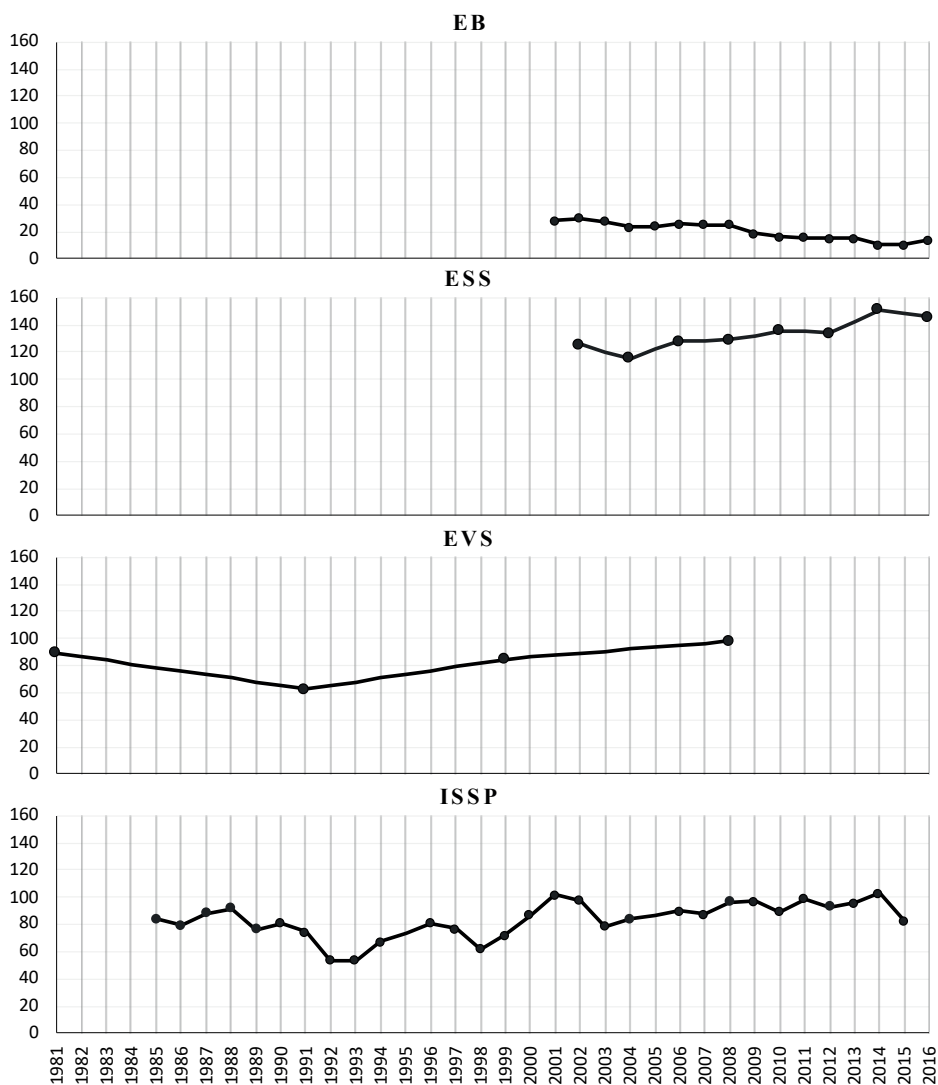


Do dwóch najczęściej wykorzystywanych procedur selekcji wewnątrzspółkowej należy zachowująca całkowicie losowy charakter siatka Kisha (dominująca w EVS oraz ISSP) oraz *quasi*-losowa procedura (grupa procedur) wykorzystująca daty urodzin domowników (stosowana najczęściej w ESS). W EVS znaczący okazał się udział procedur nielosowych mających jednak zastosowanie wyłącznie w dziewięciu krajowych sondażach w ramach edycji z 1999 roku (por. Halman 2001). Nielosowe procedury selekcji respondentów docelowych wykorzystywane były również w kilkunastu krajowych sondażach ISSP (ISSP 1996; Gendall 2011).

### 3.4. Procedury terenowej realizacji prób sondażowych

Sondażowe projekty o charakterze porównawczym wykazują znaczące różnicowanie z uwagi na wykorzystywane procedury terenowe oraz czas przeznaczony na terenową fazę badań. Należy przy tym wspomnieć, że raportowanie o dacie rozpoczęcia oraz zakończenia prac w terenie okazało się standardem wdrażanym we wszystkich analizowanych projektach badawczych. Dla zdecydowanej większości z analizowanych w tej pracy 1412 krajowych sondaży w dokumentacji podawane były dokładne daty pierwszego oraz ostatniego dnia fazy terenowej, jednakże dla stosunkowo niewielkiej części (dotyczy to zwłaszcza pierwszych edycji EVS oraz ISSP) informacje takie podawane były z dokładnością miesięczną. W takich sytuacjach podjęto w tej pracy decyzję o tym, że okres terenowej fazy badań liczony będzie od pierwszego do ostatniego dnia miesiąca wskazanego w dokumentacji, co oczywiście było znaczącym uproszczeniem przeszacowującym rzeczywistą liczbę dni trwania badań terenowych, jednak alternatywą dla takiego uproszczenia było wyłączenie poza zakres prezentowanych analiz krajowych sondaży oraz traktowanie ich jako braków uniemożliwiających wnioskowanie. Na rysunku 15 zawarto informację o średniej liczbie dni trwania terenowej fazy badań.

Zaprezentowane dane uwidaczniają, iż projekt EB charakteryzował się zdecydowanie najkrótszym okresem przeznaczonym na zbieranie danych w terenie. We wszystkich analizowanych edycjach EB średni czas nie przekraczał bowiem 30 dni, przy czym w badaniach realizowanych po 2009 roku wynosił mniej więcej 20 dni (tyle czasu upływało od momentu rozpoczęcia prac w terenie do uzyskania ostatniego wywiadu). Należy uznać okres ten za niezwykle krótki. Badania EB prowadzone były bowiem wyłącznie z wykorzystaniem techniki standaryzowanego wywiadu kwestionariuszowego w domach respondentów, na próbie przekraczającej 1000 jednostek, wymagały zatem dotarcia do całkiem znacznej liczby osób, nakłonienia tych osób do udziału w badaniu oraz uzyskania wywiadów. Nie sposób przy tym zweryfikować, na ile ten krótki czas przekładał się na osiągnięte wskaźniki terenowej realizacji próby, gdyż dokumentacja EB nie zawierała potrzebnych ku temu informacji.



Rys. 15. Średnia liczba dni przeznaczona na terenową realizację próby

Źródło: opracowanie własne.

W pozostałych projektach okres trwania terenowej fazy badań odbiegał już od standardów EB. Jest to szczególnie widoczne we wszystkich rundach ESS (średnio terenowa faza badań trwała od 120 do 160 dni, czyli od 17 do 22 tygodni). Warto jednak zauważyć, że w ESS wymaga się od krajowych koordynatorów, by prace terenowe nie trwały krócej niż cztery tygodnie (co ma zminimalizować wskaźniki braku kontaktów), podczas gdy maksymalna zalecana długość prac terenowych wynosić miała cztery miesiące (Koch, Blohm 2006). Z kolei w EVS oraz ISSP ba-

dania terenowe trwały znacznie krócej niż w ESS (nadal jednak dłużej niż w EB), zamykając się średnio w przedziale od 50 do 100 dni (czyli od 4 do 12 tygodni).

Analizując zaprezentowane dane, należy pamiętać, że czas trwania terenowej fazy badań nie pozostaje bez wpływu na terenową realizację próby badawczej. Wystarczy wskazać, że wydłużenie okresu przewidzianego na zbieranie danych jest jednym z prostszych zabiegów pozwalających na zwiększenie odsetka realizacji próby (Brick, Montaquila 2009; Stoop i in. 2010), służący ograniczeniu frakcji jednostek niedostępnych z powodu braku kontaktu (Jabkowski 2011), choć warto zarazem pamiętać, iż działania takie mogą mieć negatywny wpływ na trafność oraz rzetelność pomiaru (Domański 2006). Wydłużenie okresu przeznaczonego na realizację badań w terenie ma jednak przede wszystkim przełożenie na możliwość zastosowania szerokiego wachlarza procedur zwiększających szansę na realizację badań z osobami wylosowanymi do próby.

Tabela 15 zawiera dokładne dane o zróżnicowaniu procedur terenowych wykorzystanych w czterech międzykrajowych sondażowych projektach porównawczych. Podane informacje pozwalają wyjaśnić niezwykle krótki czas trwania terenowej fazy krajowych sondaży w ramach projektu EB. Wystarczy zauważyć, że we wszystkich krajowych sondażach EB dopuszczana była możliwość substytucji wylosowanych jednostek (Kohler 2007). Oznacza to, że ankieterzy nie musieli podejmować wielokrotnych prób dotarcia i nawiązania kontaktu z konkretnymi osobami wylosowanymi do próby (i tylko z tymi), lecz po nieudanej realizacji jakiegось wywiadu mieli możliwość dobrania respondenta zastępczego. Co więcej, w projekcie EB terenowa faza badań trwała do momentu osiągnięcia założonej liczby jednostek (wielkości próby), inaczej niż w pozostałych projektach (zwłaszcza w ESS), w których dużo wysiłku wkładane było w wielokrotne próby dotarcia do wylosowanych osób oraz przeprowadzenie wywiadów z jak największą liczbą jednostek. Opisane praktyki EB nie poprawiały w żadnym względzie jakości próby, skracały jednak czas potrzebny na realizację wywiadów (Vehovar 1999; Chapman 2003). Dla kontrastu, w projekcie ESS substytucje jednostek niedostępnych były zabronione we wszystkich krajowych sondażach oraz traktowane jako fałszerstwo wywiadu przez ankietera (Stoop i in. 2016), dopuszczane były jednak w części krajowych sondaży zrealizowanych w ramach EVS oraz ISSP.

Przyglądając się pozostałym danym zaprezentowanym w tabeli, można również wskazać, iż we wszystkich analizowanych projektach dominującą techniką badawczą były standaryzowane wywiady kwestionariuszowe. Dokumentacja EB nie określa przy tym, czy pomiar prowadzony był na podstawie techniki PAPI, czy też CAPI. W projekcie ESS przeważały wywiady kwestionariuszowe wspomagane komputerowo (CAPI), natomiast w EVS oraz ISSP wywiady oparte na tradycyjnych kwestionariuszach papierowych (PAPI). Dodatkowo, w niewielkiej części krajowych sondaży EVS dane zbierane były poprzez wywiady telefoniczne, ankiety

Tabela 15. Zróżnicowanie procedur terenowej realizacji próby w projektach EB, ESS, EVS, ISSP

Aspekt porównań	EB	ESS	EVS	ISSP
<b>Technika badawcza:</b>				
[1] PAPI	-	42,2%	85,6%	52,9%
[2] CAPI	-	57,8%	12,6%	46,7%
[3] PAPI lub CAPI	100%	100%	98,2%	99,6%
[3] CATI	-	-	0,9%	4,4%
[4] CAWI	-	-	0,9%	4,3%
[5] ankieta pocztowa	-	-	0,9%	21,7%
[6] ankieta roznoszona	-	-	-	18,0%
Brak danych / liczba sondaży	0/523	0/199	1/112	38/540
<b>Substytucje jednostek niedostępnych:</b>				
[1] Dopuszczono w trakcie badań	100%	0%	17,9%	15,4%
Brak danych / liczba sondaży	0/523	0/199	42/112	123/578
<b>Listy zapowiednie:</b>				
[1] Stosowano	-	83,4%	25,9%	26,3%
Brak danych / liczba sondaży	523/523	0/199	71/112	255/578
<b>Gratyfikacje respondentów za udział w badaniu:</b>				
[1] Stosowano	-	63,8%	17,0%	13,5%
Brak danych / liczba sondaży	523/523	1/199	21/112	252/578
<b>Procedury konwersji odmów:</b>				
[1] Stosowano	-	74,9%	20,5%	24,4%
Brak danych / liczba sondaży	523/523	4/199	71/112	192/578
<b>Procedury kontroli pracy ankierów:</b>				
[1] Stosowano	-	96,5%	55,4%	58,0%
Brak danych / liczba sondaży	523/523	1/199	42/112	147/578

Źródło: opracowanie własne.

webowe oraz pocztowe. Warto także zauważyć, że specyfiką projektu ISSP jest równoczesne stosowanie wielu technik (stąd wartości procentowe dla tego projektu nie sumują się do całości). W prawie wszystkich krajowych sondażach ISSP dane zbierane były przy tym z wykorzystywaniem wywiadów kwestionariuszowych PAPI lub CAPI (zdecydowanie rzadziej CAWI lub CATI) oraz uzupełniane ankietami pocztowymi lub ankietami roznoszonymi przez ankierów. Zasadniczą część badań ISSP opierała się często na standaryzowanym wywiadzie kwestionariuszowym wspomaganym technikami niewymagającymi obecności ankierera, co miało na celu zebranie społeczno-demograficznych charakterystyk gospodarstw domowych. Co więcej,

w wielu krajach biorących udział w ISSP kwestionariusze kolejnych edycji (moduły tematyczne) dołączane były do innych krajowych sondaży lub też moduły z dwóch różnych edycji realizowano w jednym roku (ISSP 2013).

Przyglądając się różnicowaniom częstości wykorzystywania pozostałych procedur terenowych, można również zauważyć, iż listy zapowiednie informujące respondentów o tematyce projektu oraz spodziewanym terminie wizyty ankietera stosowane były w większości sondaży prowadzonych w ESS, z kolei w EVS oraz ISSP nie były już tak powszechne. Co więcej, respondentów nagradzano pieniędzmi lub rzeczowo przede wszystkim w ESS, natomiast w pozostałych projektach znacznie rzadziej. Jeśli pod uwagę weźmie się stosowanie konwersji odmów, czyli procedur służących nakłanianiu osób odmawiających udziału w badaniu, to wykorzystywano je w  $\frac{3}{4}$  wszystkich sondaży prowadzonych w ramach ESS, natomiast w EVS oraz ISSP już znacznie rzadziej (tj. jedynie w 1 na 5 krajowych sondaży). Można wreszcie wskazać, iż procedury kontrolujące pracę ankieterów stosowane były w prawie wszystkich krajowych sondażach realizowanych w ESS, nieco rzadziej (choć nadal w większości surveyów) w badaniach EVS oraz ISSP. Dokumentacja EB nie zawierała przy tym informacji pozwalających na ustalenie stosowania w krajowych sondażach listów zapowiednich, gratyfikacji respondentów, konwersji odmów oraz procedur kontrolnych. Mając jednak w pamięci zgłaszane wątpliwości co do niskiej metodologicznej jakości tego projektu (Nissen 2014), można z dużym prawdopodobieństwem przyjąć, iż wymienionych procedur wspomagających terenowe realizacje prób badawczych nie stosowano w ogóle.

### 3.5. Wskaźniki terenowej realizacji prób sondażowych

Wszystkie działania terenowe podejmowane przez badaczy w trakcie prowadzenia pomiaru posiadają swoje konkretne przełożenie na osiągnięte efekty, tj. na jakość próby oraz wskaźniki jej terenowej realizacji. Problematyka międzyprojektowego, międzykrajowego oraz wzdłużczasowego zróżnicowania jakości sondażowej próby badawczej w czterech porównywanych projektach, a także wpływu różnych procedur terenowych na jakość pomiaru, będzie przedmiotem szczegółowych analiz w kolejnym rozdziale tej pracy. W tym momencie warto przyjrzeć się wzdłużczasowym zmianom wartości wskaźników terenowej realizacji próby oraz przeanalizować międzyprojektowe różnice w tym względzie.

Jak już wspomniano, zestawienia takie możliwe są wyłącznie wtedy, gdy do obliczania wartości wskaźników stosuje się jednolite standardy postterenowej klasyfikacji jednostek wylosowanych do próby. Wymaga to jednak kodowania

- (1) liczby wywiadów (przeprowadzonych w całości lub w części),
- (2) liczby jednostek, z którymi nie zrealizowano wywiadów z powodu:
  - (a) braku kontaktu, (b) odmowy, (c) innej przyczyny,

- (3) liczby jednostek o nieustalonym statusie przynależności do populacji docelowej,
- (4) liczby jednostek, co do których ustalono, iż nie przynależą do badanej populacji.

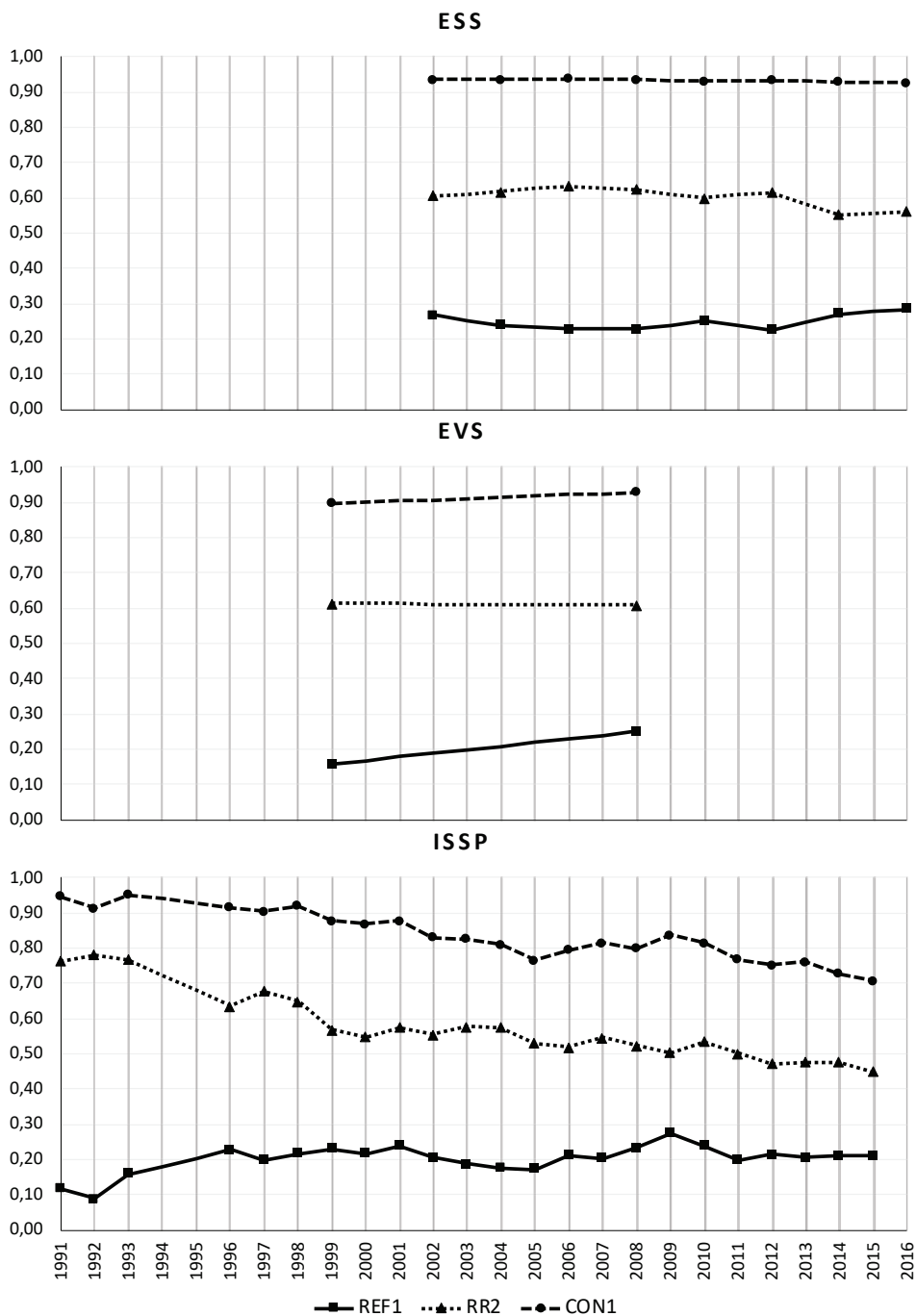
Szczegółowe wytyczne wskazujące na reguły klasyfikacji jednostek do którejs z tych czterech rozłącznych klas odnaleźć można w dokumentach AAPOR (2016).

Na rysunku 16 przedstawiono zestawienie wartości trzech głównych wskaźników terenowej realizacji próby w ESS, EVS oraz ISSP. Wykresy przedstawiają wartości wskaźnika ogólnego wyznaczonego zgodnie ze wzorem RR2 (por. AAPOR 2016: 61), wskaźnika kontaktów ustalonego w wersji CON1 (por. AAPOR 2016: 65), a także wskaźnika odmów w wariancie REF1 (por. AAPOR 2016: 64)<sup>4</sup>. Należy przy tym zaznaczyć, że wskaźnik RR2 jest frakcją jednostek, z którymi udało się zrealizować wywiady w całości lub w części (w zbiorze wszystkich jednostek o ustalonym lub nieustalonym statusie przynależności do populacji), CON1 jest frakcją jednostek, z którymi udało się nawiązać kontakt (ale tylko w zbiorze osób o potwierdzonym statusie przynależności do populacji), natomiast REF1 podaje proporcję osób odmawiających udziału w badaniu (w zbiorze wszystkich osób, z którymi udało się nawiązać kontakt). Informacje konieczne do ustalenia wartości powyższych wskaźników nie były dostępne w dokumentacji EB, jak również w niektórych edycjach EVS (z roku 1981 oraz 1990 roku) oraz ISSP (z lat 1985–1990 oraz 1994–1995).

Dane zaprezentowane na rysunku 16 potwierdzają obserwacje innych badaczy wskazujących na rosnące trudności w realizacji sondażowych prób badawczych (De Leeuw, Hox, Luiten 2018, Grzeszkiewicz-Radulska 2009). Istotnie, w kolejnych edycjach ESS, EVS oraz ISSP zaobserwowano znaczący wzrost wskaźników odmów, a także spadek wartości wskaźników kontaktów oraz wskaźników realizacji próby. Jest to w szczególności widoczne w projekcie ISSP, w którym jeszcze na początku lat 90. XX wieku wskaźniki terenowej realizacji próby wynosiły mniej więcej 80 p.p., natomiast w ostatniej zarchiwizowanej edycji z 2015 roku osiągały już tylko 45 p.p. Spadek ten był wynikiem znaczącej redukcji wskaźników kontaktu, w mniejszym stopniu natomiast konsekwencją wzrostu wskaźników odmów. Co więcej, porównując projekty we wspólnych punktach czasowych można zauważyć, iż najwyższe wartości RR2 i CON1 osiągnęte były zazwyczaj w ESS, najniższe natomiast w ISSP. Co ciekawe, ISSP był również projektem o najniższych notowanych wskaźnikach odmów.

---

<sup>4</sup> W zestawieniach pominięto wskaźnik kooperacji, gdyż pozostaje on pochodną wskaźnika odmów.



Rys. 16. Zmiany wybranych wskaźników terenowej realizacji próby w czasie

Źródło: opracowanie własne.

## 4. Podsumowanie

Dokumentacja metodologiczna czterech wybranych sondażowych projektów porównawczych ukazała znaczące międzyprojektowe i międzykrajowe różnicowania stosowanych schematów doboru jednostek oraz procedur terenowej realizacji prób badawczych. Przedstawiony materiał uwidocznił jednak przede wszystkim skalę problemów, z jakimi uporać muszą się zarówno badacze realizujący projekty międzykrajowe, jak też użytkownicy danych wtórnych wykorzystujący wyniki sondażowe do studiów nad problemami o charakterze substancywnym. Podstawowym wyzwaniem tych pierwszych jest „wyprodukowanie” danych pozwalających na uprawnione metodologicznie wnioskowanie o różnicach pomiędzy krajami. Wyzwaniem, z jakim mierzyć muszą się ci drudzy, jest konieczność uwzględnienia metodologicznego zróżnicowania projektów w analizach prowadzonych na zbiorach danych wynikowych.

Niezależnie od przyjętej perspektywy – wytwórcy danych lub też ich użytkownika – kryterium porównywalności sprowadza się w zasadzie do tego, by odmienne procedury badawcze skutkowały w krajowych sondażach podobnymi wielkościami błędów losowych oraz nielosowych (Słomczyński 2004). Wysokiej jakości próby sondażowej nie można bowiem nigdy przyjmować za pewnik, nawet jeśli dany projekt uznawany jest powszechnie za przedsięwzięcie o wysokich standardach metodologicznych. Najbardziej naturalnym sposobem zwiększania jakości pomiaru sondażowego w projektach międzykrajowych wydaje się przy tym minimalizowanie wszystkich źródeł błędów, które mogą pojawić się w toku realizacji badania (Smith 2018). Choć strategia taka wydaje się optymalna, to nie jest możliwa do wdrożenia przez użytkowników danych wtórnych. Operują oni na materiale już istniejącym i nic dla poprawy jakości pomiaru zrobić nie mogą. Tym samym szczególnego znaczenia nabiera postulowane w tej pracy możliwie najbardziej wiarygodne dokumentowanie wszystkich procedur badawczych, gdyż informacje o podejmowanych działaniach terenowych mogą zostać włączone do analiz substancywnych w postaci odpowiednio zdefiniowanych zmiennych kontrolnych. Otwarte pozostaje przy tym pytanie, na które w tej pracy dotąd nie odpowiedziano, czy i w jakim stopniu terenowe realizacje krajowych sondaży prowadzą do istotnego zróżnicowania jakości próby. Wymaga to jednak wcześniejszego zdefiniowania mierników jakości w sposób pozwalający na prowadzenie ewaluacji sondaży na dużą skalę, z pominięciem konieczności odwoływania się do zewnętrznych danych wymagających żmudnego gromadzenia materiału empirycznego, charakteryzującego populację lub wylosowane jednostki. Miarom pozwalającym użytkownikom danych wtórnych na przeprowadzenie oceny jakości próby w projektach sondażowych o charakterze porównawczym przyjrzymy się w ostatnim rozdziale pracy.





## Od zróżnicowania praktyk sondażowych do zróżnicowania jakości próby

Zróżnicowanie procedur badawczych w krajowych sondażach prowadzonych w ramach czterech prezentowanych projektów porównawczych nie powinno stanowić żadnego zaskoczenia. Trudno byłoby nawet oczekiwać, aby w prowadzonych na dużą skalę projektach międzykrajowych możliwe było zastosowanie jednakowych schematów losowania, identycznych operatów doboru próby, czy też takich samych procedur terenowych. Biorąc pod uwagę cele stawiane przed projektami porównawczymi, należy jednak pamiętać o tym, że odmienność procedur terenowych nie powinna przekładać się w znaczący sposób na zróżnicowanie jakości sondaży. Jeżeli bowiem badania mają mieć rzeczywiście charakter porównawczy, to terenowe realizacje krajowych pomiarów powinny prowadzić do identycznych (lub podobnych) błędów całkowitych, dając sondaże o porównywalnej jakości.

Celem tego rozdziału jest porównanie jakości prób badawczych w czterech międzykrajowych projektach sondażowych – EB, ESS, EVS oraz ISSP – a także analiza zależności zachodzących pomiędzy procedurami terenowymi oraz jakością próby. Rozdział podzielony został na dwie części. W pierwszej omówiono wyzwania wynikające z praktycznej implementacji zaproponowanych w literaturze miar służących ocenie jakości prób sondażowych do prowadzonej na dużą skalę ewaluacji jakości pomiarów z wielu różnych krajów. W części drugiej zaprezentowano wyniki analiz empirycznych, których głównym celem była ocena jakości krajowych prób w badaniach międzykrajowych.

Pierwszą część rozdziału otwiera omówienie założeń teoretycznych i metodologicznych stojących u podstaw paradygmatu „całkowitego błędu pomiaru” (Biemer i in. 2017), co stanowi punkt wyjścia do rozważań nad przydatnością miar jakości krajowych sondaży. Główna uwaga skupia się na ograniczeniach wynikających z implementacji poszczególnych mierników, a także na przydatności miar odwołujących się do tzw. zewnętrznych oraz wewnętrznych kryteriów oceny reprezentatywności próby. Analizy uwidoczniają, iż z perspektywy użytkowni-

ka danych wtórnych najbardziej optymalnym rozwiązaniem, jakie może zostać wdrożone w ewaluacji jakości sondaży pochodzących z międzykrajowych badań porównawczych, jest implementacja miar opartych na analizie rozkładu płci w dwuosobowych gospodarstwach domowych zamieszkiwanych przez heteroseksualnych partnerów (Kohler 2007).

Drugą część rozdziału trzeciego rozpoczną międzyprojektowe oraz wzdłużczasowe analizy zróżnicowania jakości krajowych sondaży, które pozwolą na identyfikację pomiarów o szczególnie niskiej jakości. Główna uwaga skupi się na konsekwencjach wynikających z zastosowania odmiennych procedur terenowych, a także na analizie współzależności pomiędzy jakością próby oraz wskaźnikami jej terenowej realizacji. W podsumowaniu przedstawione zostaną praktyczne rekomendacje dla prowadzenia międzyprojektowych oraz międzykrajowych zestawień jakości próby.

## **1. Wyzwania dla oceny jakości prób sondażowych w międzykrajowych projektach porównawczych o charakterze sondażowym**

### **1.1. Procedury oceny jakości sondażowych prób badawczych oraz ich ograniczenia w ewaluacji sondaży prowadzonej na dużą skalę**

Na jakość pomiaru patrzeć można w różny sposób, przyjmując za wyznacznik wysokiej jakości zestaw całkowicie odmiennych kryteriów (por. Groves, Lyberg 2010). Ta różnorodność podejść i możliwych sposobów definiowania jakości pomiaru nie zmienia faktu, że w metodologii badań sondażowych za (właściwie jedyny) uprawomocniony sposób określania jakości pomiaru przyjmuje się działania prowadzone w paradygmacie całkowitego błędu pomiaru (*Total Survey Error*, dalej TSE). Paradygmat ten pozwala nie tylko zdefiniować jakość sondażu poprzez różnicę między otrzymaną w badaniu wartością estymatora oraz prawdziwą (rzeczywistą) wartością parametru populacyjnego, ale również wskazuje na najpoważniejsze źródła uchybień składających się na całkowitą wielkość błędu (por. Biemer, Lyberg 2003; Groves 2004; Alwin 2007; Biemer 2010a). Można zatem wskazać, że poza błędami próbkowania, wynikającymi z samej natury badań prowadzonych na próbach losowych, do źródeł błędów posiadających największy wpływ na wielkość TSE zaliczyć należy przede wszystkim błąd niepełnego pokrycia populacji operatem losowania, błąd niepełnej realizacji wylosowanej próby, błąd pomiaru oraz błąd wynikający z niewłaściwego przetwarzania danych wynikowych. Wyliczenie możliwych źródeł błędów nie jest oczywiście wyczerpujące, choć zawiera zestaw wskazywany przez zdecydowaną większość badaczy odwołujących się

w swoich studiach metodologicznych do paradygmatu TSE (por. np. Weisberg 2005; Jabkowski 2015). Znaczące jest również to, że każda z wyróżnionych klas błędów posiada własne (specyficzne) źródła zanieczyszczeń pomiaru, może też w zróżnicowanym stopniu wpływać na przyrost wariancji (obniżenie precyzji estymacji) oraz systematyczne wypaczenie wyników (obciążenie estymatorów) (por. Stoop i in. 2010).

Sposób postrzegania jakości pomiaru sondażowego okazuje się znacząco zróżnicowany zwłaszcza wtedy, gdy pod uwagę weźmie się dwie odmienne perspektywy patrzenia na dane wynikowe, tj. badacza oraz użytkownika danych wtórnych (Biemer 2010a, b). O ile bowiem badacze kładą duży nacisk na jakość pomiaru, czyli na całkowitą wielkość błędu (analizując przy tym wielkości próby badawczej, wskaźniki terenowej realizacji próby, pokrycie populacji operatem losowania, schematy losowania, trafność i rzetelność pomiaru oraz procedury terenowe, przeznaczając na to znaczną część wysiłków i środków finansowych), o tyle dla użytkowników danych wtórnych najważniejszym kryterium nie jest wcale jakość pomiaru (której nie znają i przyjmują za coś pewnego), ale aktualność danych oraz ich dostępność i użyteczność dla prowadzonych analiz substancyjnych. Istnieje oczywiście wyjątek od tej prawidłowości. Jeżeli bowiem użytkownikiem danych jest osoba wykorzystująca zgromadzone w sondażach wyniki do prowadzonych przez siebie analiz o charakterze naukowym, to jakość pomiaru nie może stanowić kryterium drugorzędne, nawet jeśli wiąże się to z koniecznością podjęcia jakichś dodatkowych działań zmierzających do oceny jakości próby na podstawie dostępnych lub koniecznych do zgromadzenia dodatkowych danych.

Przeglądając pod tym kątem literaturę metodologiczną, należy podkreślić, że ocena jakości prób sondażowych dokonywana była dotychczas raczej w ramach jednego projektu (Beullens i in. 2014; Koch i in. 2014), nie zaś pomiędzy odmiennymi badaniami. Głównym celem, jaki temu przyświecał, było ustalenie relacji pomiędzy konkretnymi działaniami badawczymi oraz uzyskiwaną jakością próby (por. Lynn i in. 2007; Smith 2007). Innymi słowy, porównania międzyprojektowe nie były dotąd rozpowszechnione, choć – co warto podkreślić – i w tym zakresie odnaleźć można przykłady kilku ciekawych analiz empirycznych (Kohler 2007; Ortmanns, Schneider 2016). Tymczasem kompleksowe studia międzyprojektowe umożliwiłyby nie tylko porównanie międzykrajowych projektów sondażowych z uwagi na uzyskiwaną jakość poszczególnych pomiarów, ale przede wszystkim dostarczyłyby danych pozwalających na ustalenie wpływu projektowych specyfik terenowej realizacji próby na osiąganą jakość przeprowadzonego sondażu.

Robert Groves (2006) omówił podstawy kilku praktycznych podejść służących ocenie jakości sondażowych prób badawczych. Większość z nich okazuje się jednak niepraktyczna w prowadzonej na dużą skalę ewaluacji jakości pomiaru, gdyż wymaga dostępu do zewnętrznych danych charakteryzujących każdą osobę

wylosowaną do próby. Dla porządku można jednak wskazać, że pierwsza z zaproponowanych przez Grovesa klas metod polega na porównaniu wskaźników terenowej realizacji próby (*response rate*) w podpróbach wyróżnionych z uwagi na cechy społeczno-demograficzne jednostek, takie jak płeć, wiek czy też poziom wykształcenia. Zakłada się przy tym, że jeżeli wartości wskaźników realizacji próby są takie same (lub porównywalne) w poszczególnych podpróbach, to cała próba charakteryzuje się dobrą jakością. Problem z implementacją tego podejścia polega na tym, że w praktyce dostępne są co najwyżej wskaźniki realizacji całej próby (i to nie zawsze), rzadko kiedy udostępniane są dane o udziale w badaniu konkretnych wylosowanych osób.

Druga z grup procedur opisanych przez Grovesa polega na wykorzystaniu (w ocenie jakości próby) informacji pochodzących z operatów losowania lub też danych gromadzonych przez ankietatorów oraz informacji zbieranych w ramach *follow-up studies*, czyli badań uzupełniających prowadzonych z osobami niedostępnymi lub niechętnymi do udziału w sondażu. Wreszcie ostatnia z propozycji Grovesa polega na kontrastowaniu szacunków przed i po zastosowaniu metod ważenia danych (por. np. Billiet i in. 2009; Peytchev, Presser, Zhang 2018; Sakshaug, Antoni 2018). Nowsze podejścia do oceny jakości prób sondażowych wywodzące się wprost z pomysłów przedstawionych przez Grovesa obejmują analizę wskaźników reprezentatywności próby, opierających się na ocenie zróżnicowania w zbiorze jednostkowych skłonności do udziału w badaniu (por. Schouten i in. 2012). Inne metody odwołujące się do tego pomysłu polegają na analizie zbalansowanego zestawu odpowiedzi (Särndal 2011) i porównują skłonność do udziału w badaniu w oparciu o rozkłady cech respondentów oraz wszystkich osób wylosowanych do próby (Lundquist, Särndal 2013).

## **1.2. Rola wewnętrznych kryteriów i zewnętrznych kryteriów oceny jakości prób badawczych w sondażowych projektach porównawczych**

Omówione w poprzedniej sekcji tego rozdziału podejścia metodologiczne służące ocenie jakości sondażowych prób badawczych znajdują niewielkie zastosowanie w zakrojonej na dużą skalę ewaluacji sondaży pochodzących z wielu różnych projektów, wymagają bowiem dostępu do informacji niedostępnych w upublicznianych zbiorach danych sondażowych. Z perspektywy użytkownika danych wtórnych powinno poszukiwać się raczej takich miar jakości sondażu, które nie będą wymagały dostępu do jednostkowych informacji o wylosowanych osobach, a także będą bezpośrednio powiązane z koncepcją TSE. Co więcej, z uwagi na to, iż użytkownicy danych wtórnych mają ograniczony dostęp do dodatkowych informacji gromadzonych przez badaczy realizujących krajowe sondaże, zastosowane miary

jakości – aby miały szerokie zastosowanie – powinny odwoływać się do informacji rutynowo udostępnianych publicznie.

Dwa potencjalnie możliwe do wdrożenia na dużą skalę sposoby oceny jakości prób sondażowych w międzykrajowych badaniach porównawczych, które nie wymagają dostępu do żadnych dodatkowych informacji o wylosowanych do próby jednostkach, opierają się na analizie rozkładu płci w całej próbie lub w specyficznym zdefiniowanej podpróbie respondentów. Pierwsza propozycja polega na porównaniu otrzymanego w próbie rozkładu płci ze znanym rozkładem w populacji, co wymaga dostępu do zewnętrznych danych charakteryzujących całą populację, nie wymaga jednak informacji o poszczególnych jednostkach. Implementacja zewnętrznych kryteriów oceny jakości próby polega w praktyce na porównaniu odsetka kobiet w próbie ze znanym odsetkiem w całej populacji (Eckman, Koch 2019). Druga metoda, nosząca nazwę kryteriów wewnętrznych, polega z kolei na porównaniu frakcji kobiet w podpróbie respondentów zamieszkujących w dwuosobowych gospodarstwach domowych heteroseksualnych partnerów, z teoretyczną frakcją kobiet wynoszącą w takich gospodarstwach dokładnie 50% (Kohler 2007; Menold 2014; Eckman, Koch 2019; Jabkowski, Cichocki 2019). Kryterium wewnętrzne nie wymaga już dostępu do żadnych zewnętrznych źródeł danych i opiera się wyłącznie na informacjach gromadzonych w badaniu

Obie metody odwołują się wprost do paradygmatu TSE, porównują bowiem wartość pewnego estymatora (w obu przypadkach jest to frakcja kobiet) z prawdziwą wartością parametru w populacji lub subpopulacji. Dla kryteriów zewnętrznych wartością prawdziwą jest odsetek kobiet składających się na populację, natomiast dla kryteriów wewnętrznych aprioryczny udział kobiet w subpopulacji osób zamieszkujących dwuosobowe gospodarstwa domowe heteroseksualnych partnerów wynoszący 50%. Ponieważ jednak kryteria zewnętrzne wymagają zgromadzenia (dla konkretnego kraju i roku, w którym badanie zostało przeprowadzone) danych populacyjnych, opisujących rozkład płci (źródłem takich informacji dla krajów UE może być Eurostat, natomiast dla krajów pozaeuropejskich dane United Nations Population Division), to stosunkowo najmniej problematyczne wydaje się – mimo wszystko – zastosowanie kryteriów wewnętrznych niewymagających żadnych dodatkowych informacji, poza tymi gromadzonymi standardowo w kwestionariuszach.

Warto jednak pamiętać, że możliwość zastosowania kryteriów wewnętrznych warunkowana jest posiadaniem informacji nie tylko o płci respondenta, ale też o wielkości gospodarstwa domowego (by można było wyselekcjonować gospodarstwa zamieszkiwane wyłącznie przez dwie osoby) oraz o związku respondenta z drugim członkiem gospodarstwa (by można było wyselekcjonować pary pozostające w heteroseksualnym związku). Miarę jakości próby opartą na

kryterium wewnętrznym definiuje się wtedy jako wielkość oznaczoną wzorem  $Abs\_bias = \frac{|\hat{p}_i - 0.5|}{\sqrt{0.25/n_i}}$ , gdzie  $\hat{p}_i$  jest proporcją kobiet w podpróbie oraz  $n_i$  jest wielkością próby respondentów zamieszkujących w dwuosobowych gospodarstwach domowych heteroseksualnych partnerów. Odnosząc zdefiniowaną miarę do paradygmatu TSE należy zauważyć jej bezpośrednie powiązanie z błędami reprezentatywności próby (Groves i in. 2011). Innymi słowy, na wielkość miary jakości sondażu wyrażonej kryterium wewnętrznym wpływ mają przede wszystkim błędy próbkowania, niepełnego pokrycia populacji oraz niepełnej realizacji próby (Weisberg 2005).

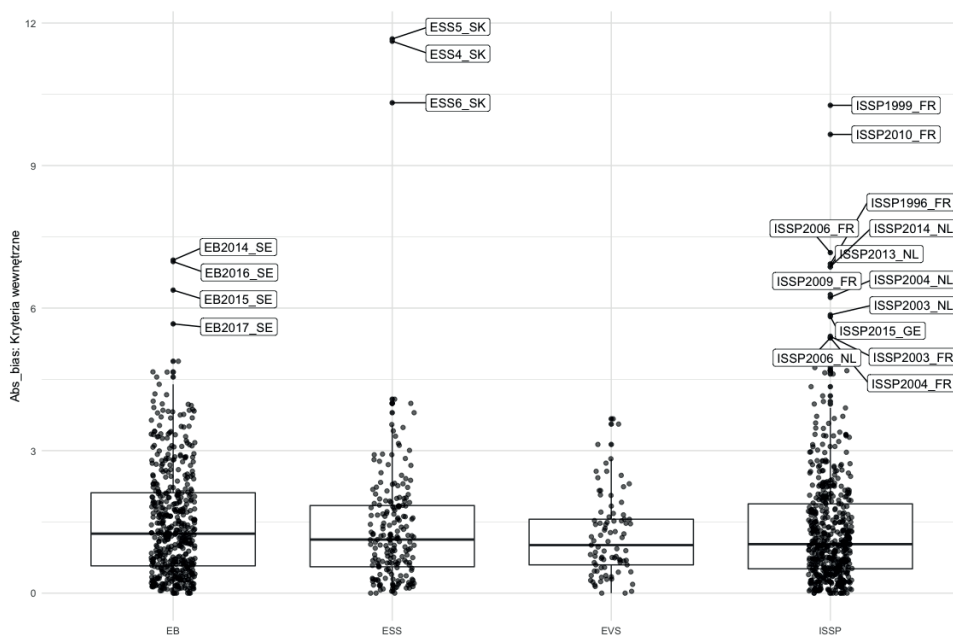
## 2. Porównawcze analizy jakości prób sondażowych w wybranych projektach

Druga część rozdziału trzeciego poświęcona jest analizie jakości sondażowych prób badawczych zrealizowanych w ramach EB, ESS, EVS oraz ISSP. W pierwszej kolejności zaprezentowano międzyprojektowe oraz wzdłużczasowe zróżnicowania jakości prób, a następnie przeprowadzono analizy weryfikujące wpływ wykorzystanych rodzajów operatów doboru jednostek, typów prób badawczych oraz procedur selekcji respondentów docelowych na uzyskiwaną jakość próby sondażowej. Przedmiotem analizy uczyniono także korelacje pomiędzy miernikami jakości próby oraz wskaźnikami jej terenowej realizacji. Bazę empiryczną stanowiły zbiory danych wyników wszystkich 1412 krajowych sondaży będących przedmiotem pogłębionych studiów metodologicznych zaprezentowanych w poprzednim rozdziale tej pracy. Ocena jakości próby opierała się na tzw. kryteriach wewnętrznych, czyli na analizie rozkładu płci w podpróbie osób zamieszkujących dwuosobowe gospodarstwa domowe heteroseksualnych partnerów. Zmienne pozwalające na zastosowanie wewnętrznych kryteriów oceny jakości próby (tj. płeć respondenta, wielkość gospodarstwa domowego, związek respondenta z drugą osobą zamieszkujejącą we wspólnym gospodarstwie) dostępne były dla 1326 krajowych sondaży, w tym dla 493 pomiarów EB, 199 sondaży ESS, 85 sondaży EVS oraz 549 krajowych prób ISSP.

### 2.1. Międzyprojektowe oraz wzdłużczasowe zróżnicowanie jakości prób sondażowych

Porównawcze zestawienia jakości prób sondażowych w EB, ESS, EVS oraz ISSP warto rozpocząć od analiz międzyprojektowych. Rysunek 17 przedstawia wartości miar jakości każdej krajowej próby sondażowej zrealizowanej w ramach wszystkich po-

równypanych edycji czterech projektów (punkty na wykresie). Na rysunku zamieszczono również wykresy typu ramka-wąs. Ramki ukazują wartość kwartyła dolnego, mediany oraz kwartyła górnego rozkładu jakości próby, „wąsy” wskazują natomiast na zakres wartości nieodstających ustalony na podstawie rozstępu ćwiartkowego z mnożnikiem 1,5. Na wykresie oznaczono także krajowe pomiary odstające ekstremalnie od mediany rozkładu, tj. sondaże o wartości miary jakości wykraczającej poza granicę określoną rozstępem ćwiartkowym z mnożnikiem 3. Warto przypomnieć, że miara jakości skonstruowana została w ten sposób, by mniejsza wartość świadczyła o lepszej jakości pomiaru.



Rys. 17. Porównanie jakości sondażowych prób badawczych w projektach EB, ESS, EVS, ISSP  
 Uwagi: test mediany ( $\chi^2 = 9,99$ ;  $df = 3$ ;  $p = ,019$ ); test K-W ( $\chi^2 = 7,77$ ;  $df = 3$ ;  $p = ,059$ );  $N = 1326$ .  
 Oznaczenia krajów: Francja (FR), Gruzja (GE), Holandia (NL), Słowacja (SK), Szwecja (SE).

Źródło: opracowanie własne.

Ponieważ wartością krytyczną miary jakości próby mierzonej kryterium wewnętrznym jest wielkość 1,96 (przy 5% poziomie istotności wnioskowania), to należy stwierdzić, że w zdecydowanej większości krajowych pomiarów (mniej więcej dla 75%) wartość miernika jakości próby mieściła się w tolerowanych granicach. Z drugiej strony oznacza to jednak, że co czwarty krajowy pomiar charakteryzował się niezadowalającą jakością. Szczególnie interesujące wydają się przy tym sondaże o ekstremalnie dużym obciążeniu, gdyż ich rozpoznanie może pozwolić na iden-



tyfikację czynników posiadających najbardziej istotne przełożenie na niską jakość próby sondażowej. Co więcej, w analizach prowadzonych przez użytkowników danych wtórnych, którzy nie posiadają dostępu do innych danych projektowych poza zbiorami wyników, zastosowanie miar oceny jakości próby oraz identyfikacja sondaży o szczególnie wątpliwej jakości może uwidocznić konieczność zachowania większej ostrożności w formułowaniu wniosków opartych na danych empirycznych o niskiej jakości.

Zdecydowanie najbardziej jaskrawymi przykładami ekstremalnego odchylenia jakości próby sondażowej od mediany rozkładu (wyliczonej dla każdego projektu z osobna) okazują się pomiary pochodzące z czwartej, piątej oraz szóstej rundy ESS na Słowacji. Są one interesujące o tyle, że charakteryzują się nie tylko najwyższymi (spośród wszystkich 1,4 tys. krajowych sondaży uwzględnionych w prezentowanych analizach) wartościami błędów (w czwartej, piątej oraz szóstej rundzie ESS wynoszącymi odpowiednio 11,62, 11,66 oraz 10,32), ale również znaczącym przyrostem obciążenia w porównaniu z wcześniejszymi pomiarami przeprowadzonymi w tym kraju w drugiej oraz trzeciej rundzie ESS (wynosiły one wówczas 0,31 oraz 1,01). Szczegółowe raporty metodologiczne projektu ESS, bogato informujące o terenowych specyfikach każdego pojedynczego sondażu, ujawniają przy tym fakt znaczącej modyfikacji procedur terenowych wdrażanych na Słowacji pomiędzy trzecią oraz czwartą rundą tego badania. Jeśli w drugiej oraz trzeciej rundzie zastosowano imienny operat losowania, to od czwartej edycji ESS opierał się on już na rejestrach adresowych. Wiązało się to nie tylko z techniczną zmianą typu operatu, ale pociągało za sobą szereg konsekwencji praktycznych. Dla przykładu w drugiej (por. ESS 2004) oraz trzeciej rundzie ESS (por. ESS 2006) próbę na Słowacji losowano wprost z operatu imiennego, natomiast od rundy czwartej przeprowadzano wstępną selekcję gmin, by następnie (w każdej gminie) wylosować pewne adresy, dokonać spisu wszystkich gospodarstw domowych na danej ulicy lub w bloku, systematycznie dobrać 10 gospodarstw domowych oraz przeprowadzić losowanie respondentów docelowych (por. ESS 2012).

Zmiana typu wykorzystanego operatu losowania oraz wprowadzenie pośredniej selekcji respondentów docelowych zmniejszyło możliwość skutecznego nadzoru badaczy nad procesem doboru jednostek (Eckman, Koch 2019). Co więcej, przyglądając się „surowym” wartościom miar obciążenia próby (dodatnia oznacza nadreprezentację kobiet), można zauważyć, że sondáže pochodzące z drugiej i trzeciej rundy ESS na Słowacji charakteryzowały się nieistotną nadreprezentacją mężczyzn, podczas gdy w rundach 4-6 można już zauważyć istotną statystycznie nadreprezentację kobiet, których udział w podpróbie respondentów zamieszkujących dwuosobowe gospodarstwa domowe heteroseksualnych partnerów (przy referencyjnej wartości prawdziwej równej 50%) wynosił ponad 70 p.p. Wnioski te pozostają zgodne z wcześniejszymi ustaleniami badaczy wskazującymi na to, iż

jedną z zasadniczych konsekwencji wykorzystania prób innych niż imienne okazuje się znacząca nadreprezentacja kobiet (zob. Jabkowski, Cichocki 2019).

Analizując dane ISSP, warto zwrócić również uwagę na systematycznie częstszą obecność sondaży prowadzonych we Francji oraz Holandii, w grupie badań o ekstremalnie gorszej jakości pomiaru. Wyjaśnienie tego fenomenu odnaleźć można w raportach metodologicznych projektu ISSP. Okazuje się bowiem, że w obu krajach sondaże realizowane były na podstawie ankiet pocztowych/e-mailowych, przy czym respondent docelowy dobierany był bez asysty ankietera, tj. samodzielnie wskazywany przez któregoś z członków gospodarstwa domowego. Trudno uznać taki przebieg selekcji jednostek za losowy lub nawet *quasi*-losowy, gdyż niemożliwa okazuje się w takich przypadkach skuteczna kontrola nad samoselekcją respondentów, co prowadziło w sposób zupełnie oczywisty do znacznie gorszej jakości próby poprzez istotną nadreprezentację osób o wysokiej skłonności do udziału w badaniu lub posiadających przekonanie, iż to właśnie one mają, z grona domowników, zdecydowanie największe kompetencje oraz wiedzę potrzebną do udzielenia odpowiedzi (por. Lavrakas 2008).

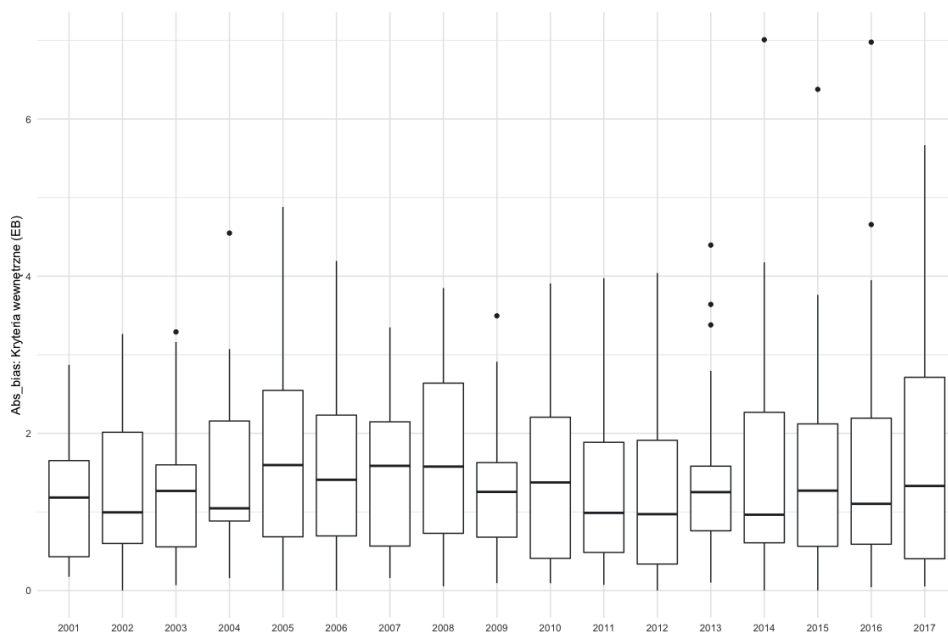
Powracając do zestawień międzyprojektowych, należy również wskazać, że cztery porównywane projekty różniły się w sposób statystycznie istotny z uwagi na jakość uzyskiwanych prób sondażowych. Wskazuje na to zarówno test mediany, jak też test Kruskala-Wallisa (odpowiednie wartości statystyk testowych oraz *p*-wartości podano w uwagach pod rysunkiem). Najwyższą medianę rozkładu wartości miar jakości sondaży, tj. największe przeciętne obciążenie próby uzyskano w projekcie EB<sup>1</sup>, najniższe natomiast w EVS oraz ISSP. Należy przy tym dodać, że zaobserwowanie międzyprojektowych różnic w jakości prób badawczych nie powinno zaskakiwać. Wszak projekty posiadają swoją specyfikę, różnią się wdrażanymi procedurami, stosuje się w nich odmienne techniki badawcze, podejmuje różne działania terenowe oraz wykorzystuje własne reguły raportowania o metodologii badawczej i osiągniętych rezultatach terenowej realizacji próby. Wszystko to ma określone przełożenie na jakość pomiaru i skutkuje zróżnicowaniem porównywanych projektów.

Odejdziemy w tym momencie jednak od zestawień międzyprojektowych, aby skupić uwagę na zmianach jakości sondażowych prób badawczych zachodzących w kolejnych edycjach czterech projektów. Do porównań międzyprojektowych powrócimy jednak w kolejnej sekcji tego podrozdziału. Na rysunkach oznaczonych numerami od 18 do 21 zaprezentowano rozkłady miar jakości prób

---

<sup>1</sup> Warto zwrócić uwagę na cztery pomiary zrealizowane w Szwecji (edycje 2014-2017). Niestety niepełna dokumentacja projektu EB uniemożliwia ustalenie powodów znaczącego wypaczenia próby we wskazanych pomiarach.

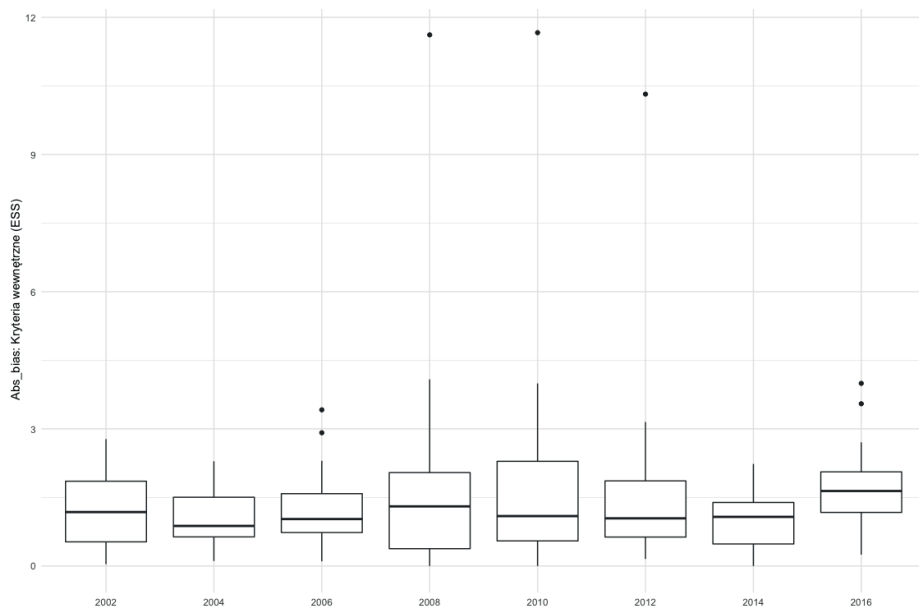
sondażowych oddzielnie dla każdej edycji EB, ESS, EVS oraz ISSP<sup>2</sup>. Pod każdym rysunkiem zawarto również informacje o wynikach testów weryfikujących hipotezy o wzdluzczasowej równości median oraz równości rozkładów miar jakości sondaży. Punkty na wykresie identyfikują krajowe pomiary o jakości próby odchylającej się znacząco od mediany rozkładu ustalonej osobno dla każdej edycji analizowanych projektów.



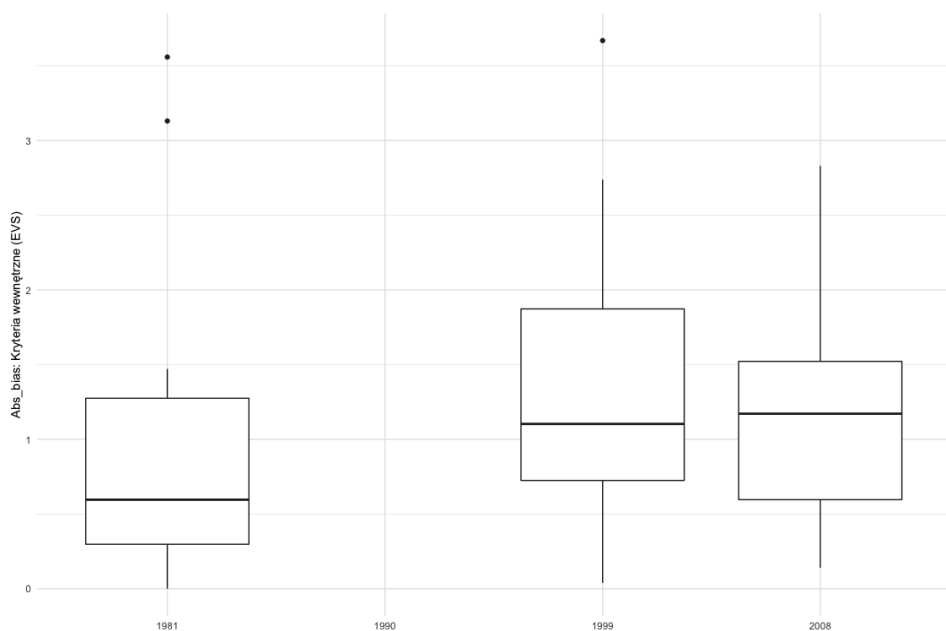
Rys. 18. Porównanie jakości sondażowych prób badawczych w kolejnych edycjach EB  
 Uwagi: test mediany ( $\chi^2 = 11,74; df = 16; p = ,762$ ); test K-W ( $\chi^2 = 10,35; df = 16; p = ,848$ );  $N = 493$ .  
 Źródło: opracowanie własne.

Zaobserwowane różnice nie dają wystarczających podstaw do odrzucenia hipotez o równości median oraz równości rozkładów miar jakości w kolejnych rundach EB, ESS, EVS oraz ISSP. Wynika to w głównej mierze z niewielkiej liczby (od kilkunastu do około trzydziestu) sondaży realizowanych w każdej rundzie porównywanych projektów, co oznacza, iż obserwowane różnice pomiędzy porównywanymi edycjami każdego projektu musiałyby w praktyce być dużo większe, by uznać je za statystycznie istotne.

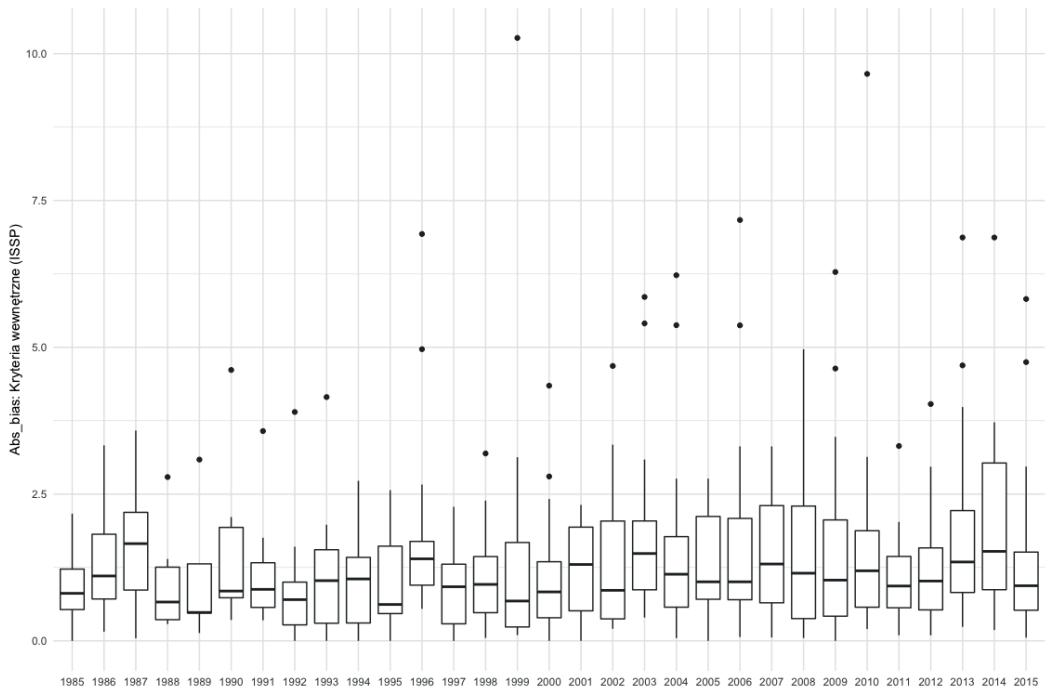
<sup>2</sup> W przypadku EVS nie było możliwości wyliczenia miar jakości próby dla sondaży zrealizowanych w 1990 roku.



Rys. 19. Porównanie jakości sondażowych prób badawczych w kolejnych edycjach ESS  
 Uwagi: test mediany ( $\chi^2 = 7,90$ ;  $df = 7$ ;  $p = ,341$ ); test K-W ( $\chi^2 = 6,89$ ;  $df = 7$ ;  $p = ,440$ );  $N = 199$ .  
 Źródło: opracowanie własne.



Rys. 20. Porównanie jakości sondażowych prób badawczych w kolejnych edycjach EVS  
 Uwagi: test mediany ( $\chi^2 = 2,16$ ;  $df = 2$ ;  $p = ,339$ ); test K-W ( $\chi^2 = 3,26$ ;  $df = 3$ ;  $p = ,196$ );  $N = 85$ .  
 Źródło: opracowanie własne.



Rys. 21. Porównanie jakości sondażowych prób badawczych w kolejnych edycjach ISSP

Uwagi: test mediany ( $\chi^2 = 20,9$ ;  $df = 30$ ;  $p = ,892$ ); test K-W ( $\chi^2 = 24,6$ ;  $df = 30$ ;  $p = ,744$ );  $N = 549$ .

Źródło: opracowanie własne.

Przyglądając się wyłącznie statystykom opisowym (np. medianie), można jednak wskazać na brak uniwersalnego wzoru opisującego zmiany jakości prób badawczych w czasie. Jest to szczególnie widoczne w (realizowanych co roku) badaniach EB oraz ISSP, dla których mediana miar jakości próby ulegała (z roku na rok) znacznym wahaniom. Brak prawidłowości w tym względzie jest w sumie dość oczywisty. Projektowa specyfika pomiaru oraz niewielka zmienność wykorzystywanych procedur badawczych skutkuje próbami o różnej, choć w każdej rundzie przeciętnie podobnej jakości.

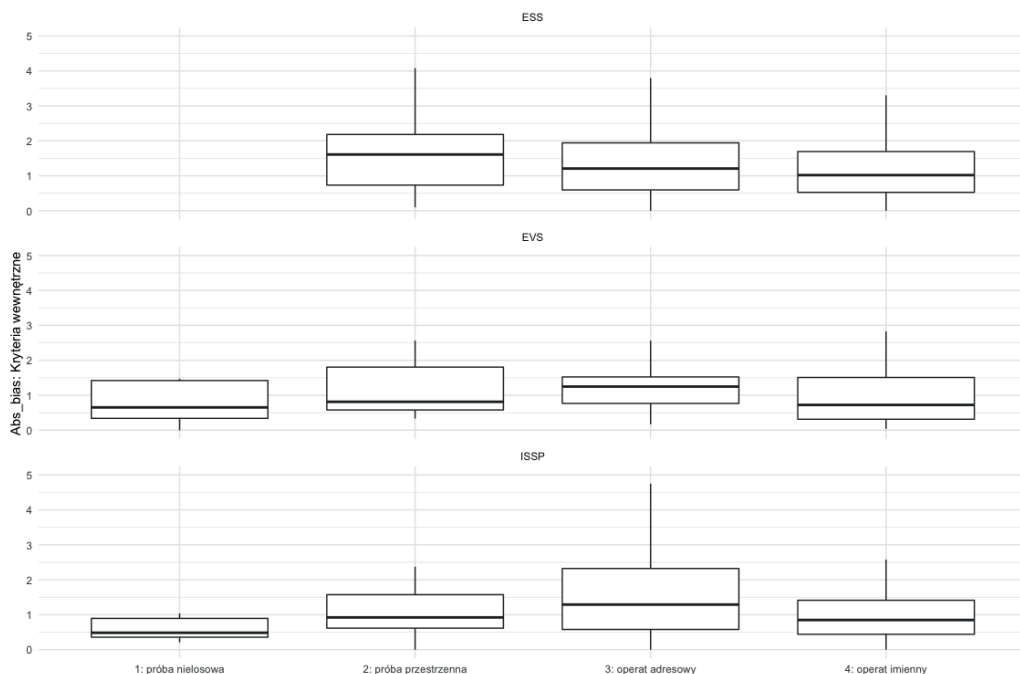
## 2.2. Zróźnicowanie jakości prób sondażowych z uwagi na procedury doboru próby

Jak już wskazywano wcześniej, analizy empiryczne przedstawione w poprzednim rozdziale tej pracy ukazały znaczące międzyprojektowe oraz międzykrajowe zróźnicowanie procedur losowania oraz terenowej realizacji sondażowych prób badawczych. Otwarte pozostaje jednak pytanie, czy i w jakim stopniu owo zróźni-

cowanie przekłada się w jakiś znaczący sposób na jakość prób badawczych. Warto przyjrzeć się temu bliżej głównie z tego powodu, że podstawową ideą przyświecającą prowadzeniu międzykrajowych sondaży porównawczych jest uzyskanie prób o podobnej jakości, pomimo odmiennych procedur służących ich terenowej realizacji. Zgromadzony w tej pracy materiał pozwala na weryfikację hipotez o wpływie, jaki na jakość próby mają czynniki takie jak

- (1) typ operatu losowania jednostek,
- (2) rodzaj próby badawczej oraz
- (3) stosowane procedury selekcji respondentów docelowych w próbach adresowych.

Na rysunku 22 przedstawiono porównanie jakości sondażowych prób badawczych z uwagi na rodzaj operatu (nielosowy, przestrzenny, adresowy, imienny). Z zestawień wyłączono projekt EB, którego dokumentacja nie zawierała informacji potrzebnych do przeprowadzenia porównań.



Rys. 22. Porównanie jakości sondażowych prób badawczych w projektach ESS, EVS oraz ISSP z uwagi na rodzaj operatu

Uwagi: ESS: test mediany ( $\chi^2 = 2,73$ ;  $df = 2$ ;  $p = ,255$ ); test K-W ( $\chi^2 = 2,33$ ;  $df = 2$ ;  $p = ,312$ );  $N = 199$ ; EVS: test mediany ( $\chi^2 = 2,33$ ;  $df = 3$ ;  $p = ,357$ ); test K-W ( $\chi^2 = 3,35$ ;  $df = 3$ ;  $p = ,340$ );  $N = 78$ ; ISSP: test mediany ( $\chi^2 = 17,26$ ;  $df = 3$ ;  $p = ,001$ ); test K-W ( $\chi^2 = 25,98$ ;  $df = 3$ ;  $p < ,001$ );  $N = 483$ .

Źródło: opracowanie własne.

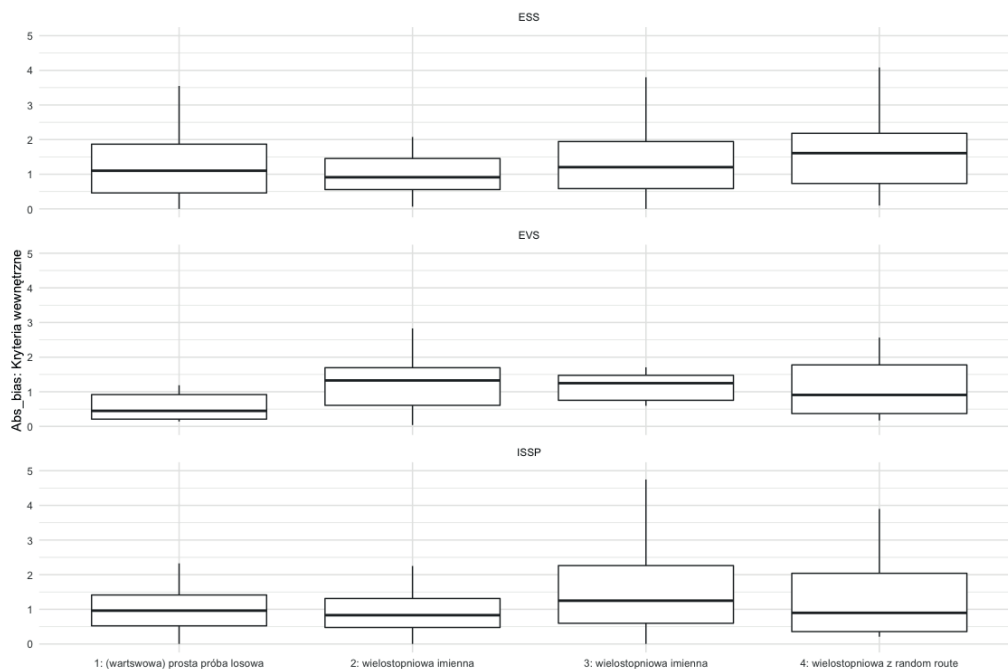
Zaprezentowane na rysunku dane prowadzą do dość nieoczekiwanej konstatacji, a mianowicie, że najmniejszym poziomem obciążenia procesu selekcji jednostek (mierzonym wskaźnikami jakości próby opartymi na kryteriach wewnętrznych) charakteryzowały się próby nielosowe. Nie oznacza to jednak, że próby tego typu są pożądane w badaniach sondażowych, gdyż nie biorą one pod uwagę prawdopodobieństw selekcji jednostek. Innymi słowy, nie spełniają podstawowego kryterium poprawności wnioskowania statystycznego w oparciu o wyniki reprezentatywnego pomiaru. Co więcej, wytłumaczenie niewielkiego obciążenia prób nielosowych jest bardzo proste. Otóż w próbach tego typu stosuje się najczęściej niezrandomizowany dobór kwotowy, który oznacza dowolność selekcji uczestników wywiadów, przy czym jedynym kryterium włączania do próby jest posiadanie przez jednostkę pewnych z góry ustalonych cech. Ponieważ kwoty nakładane są najczęściej na zmienną płęć, czyli charakterystykę wykorzystywaną w wewnętrznych miarach oceny jakości sondażu, to nie jest zaskoczeniem, iż odgórnie ustalony rozkład płci, odpowiadający mniej więcej rozkładowi populacyjnemu, znajdzie odzwierciedlenie w niskich wartościach obciążenia wyliczanego kryteriami wewnętrznymi, nawet jeżeli pod uwagę biorą one jedynie podpróbę respondentów zamieszkujących w dwuosobowych gospodarstwach domowych heteroseksualnych partnerów, a nie całą próbę, w której kontrolowano rozkład płci. Innymi słowy, próby nielosowe posiadają swoją specyfikę na tyle odmienną od prób probabilistycznych, że porównywanie tych pierwszych z drugimi pozbawione jest większego sensu. Zestawienie obu typów próby zaprezentowane zostało na rysunku 22 głównie po to, by ukazać brak podstaw do prowadzenia analizy jakości prób nielosowych przy pomocy narzędzi przeznaczonych do oceny sondaży opartych na próbach losowych. W dalszej części prezentacji próby nieprobabilistyczne zostaną wyłączone poza zakres prowadzonej analizy.

Biorąc zatem pod uwagę wyłącznie próby probabilistyczne, można przypomnieć poczynione w rozdziale drugim ustalenia, a mianowicie sondaże oparte na operatach imiennych powinny dawać próby lepszej jakości (zarówno z uwagi na możliwość bezpośredniego losowania jednostek wprost z operatu, jak też sposobność kontrolowania procesu terenowej realizacji próby), podczas gdy sondaże oparte na operatach adresowych (wymagające pośredniej selekcji jednostek) prowadzić powinny do prób o gorszej jakości (z uwagi na mniejszą możliwość skutecznej kontroli ankieterów, w tym weryfikacji poprawności implementacji procedur losowania wewnątrzzespołowego). Najniższą jakością charakteryzować powinny się przy tym sondaże oparte na próbach przestrzennych, gdyż stosuje się je w sytuacji braku możliwości wykorzystania wiarygodnych operatów imiennych lub adresowych.

Przyróżczenia te znajdują tylko częściowe potwierdzenie w danych empirycznych. Ujmując to bardziej szczegółowo, można wskazać, że operaty imienne

dawały co prawda próby o (przeciętnie) lepszej jakości od tych opartych na operatach adresowych, jednakże w EVS oraz ISSP próby przestrzenne charakteryzowały się już przeciętnie lepszą jakością od prób adresowych, zbliżoną do jakości prób imiennych. Trudno jednoznacznie wytłumaczyć taki stan rzeczy, gdyż na uzyskaną jakość próby wpływ musiały mieć inne czynniki, których znaczenie trudno określić bez dostępu do informacji o wszystkich kluczowych decyzjach podejmowanych w trakcie terenowej realizacji próby.

Na rysunku 23 zaprezentowano porównanie jakości sondażowych prób badawczych z uwagi na typ próby (z zestawień ponownie wyłączono projekt EB) oraz próby nielosowe. Jak już zresztą wspomniano, typ próby pozostaje w ścisłym związku z operatem losowania, pozwala jednak wprowadzić dodatkowe rozróżnienie w grupie sondaży opartych na operatach imiennych, tj. na imienne proste (lub warstwowe) próby losowe oraz wielostopniowe próby imienne, a także rozróżnienie w grupie sondaży opartych na operatach adresowych, tj. na adresowe próby wielostopniowe bez oraz z komponentem losowej ścieżki.



Rys. 23. Porównanie jakości sondażowych prób badawczych w projektach ESS, EVS oraz ISSP z uwagi na typ próby

Uwagi: ESS: test mediany ( $\chi^2 = 3,08$ ;  $df = 3$ ;  $p = ,379$ ); test K-W ( $\chi^2 = 3,31$ ;  $df = 3$ ;  $p = ,346$ );  $N = 198$ ; EVS: test mediany ( $\chi^2 = 2,85$ ;  $df = 3$ ;  $p = ,415$ ); test K-W ( $\chi^2 = 5,15$ ;  $df = 3$ ;  $p = ,161$ );  $N = 44$ ; ISSP: test mediany ( $\chi^2 = 13,08$ ;  $df = 4$ ;  $p = ,004$ ); test K-W ( $\chi^2 = 21,61$ ;  $df = 3$ ;  $p < ,001$ );  $N = 390$ .

Źródło: opracowanie własne.



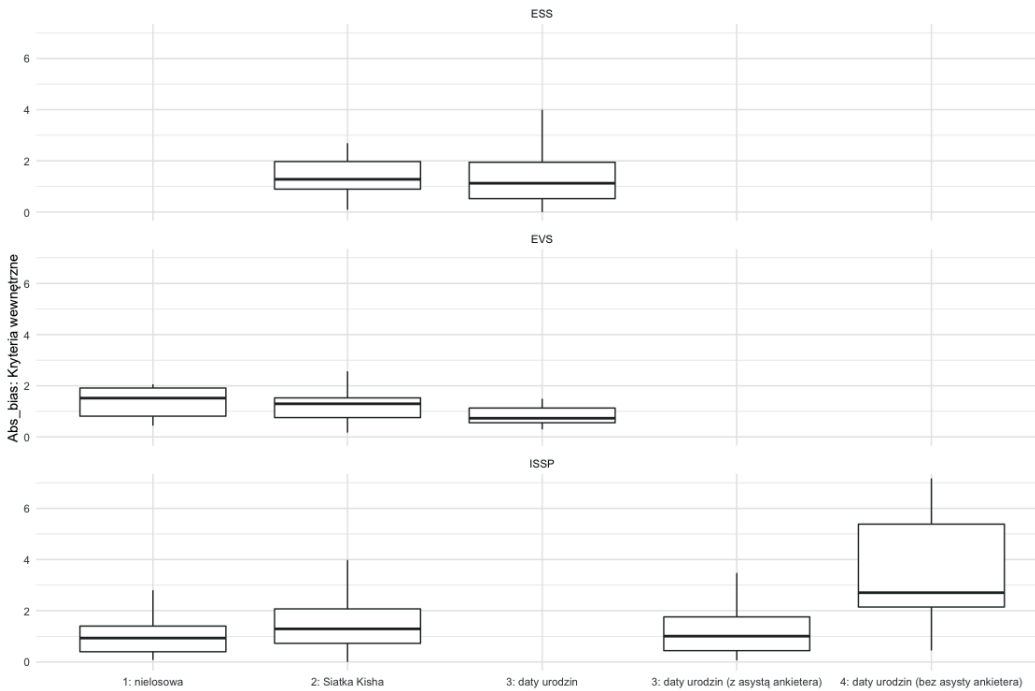
Ponieważ – tak jak już wcześniej wspomniano – zastosowanie określonego typu próby pozostaje w ściślejszej relacji z rodzajem wykorzystanego operatu losowania jednostek, to (warstwowe) proste próby imienne oraz wielostopniowe próby imienne powinny charakteryzować się lepszą jakością od wielostopniowych prób adresowych oraz prób adresowych z komponentem losowej ścieżki. Podobnie jak w odniesieniu do omówionych wcześniej relacji pomiędzy typem operatu a jakością próby, tak i tutaj dane empiryczne nie dają podstaw do jednoznacznych rozstrzygnięć. Przyglądając się zróżnicowaniu jakości próby w grupach sondaży wyróżnionych z uwagi na typ próby, trudno znaleźć jakiś uniwersalny wzór zależności. Dla przykładu w ESS oraz ISSP zdecydowanie najlepszą jakością charakteryzowały się wielostopniowe próby imienne, podczas gdy w EVS już najgorszą. Z kolei w EVS najwyższą jakość osiągały proste lub warstwowe próby imienne. Nie sposób jednak znaleźć wyjaśnienie powodu, dla którego wielostopniowe próby imienne charakteryzowały się w EVS przeciętnie najmniejszą jakością, natomiast próby proste lub warstwowe imienne – najwyższą, albowiem (co do istoty) próby te niewiele się od siebie różnią. Dla porządku należy również dodać, że w ESS najgorszą jakością cechowały się wielostopniowe próby adresowe z komponentem losowej ścieżki. Przykłady te potwierdzają brak ogólnego wzorca pozwalającego na opisanie zależności pomiędzy jakością próby oraz jej typem.

Rysunek 24 zawiera informacje o parametrach rozkładu miar jakości prób w grupach sondaży wyróżnionych z uwagi na zastosowaną procedurę wewnątrzspołowej selekcji respondentów docelowych. Warto przypomnieć, że procedury te nie mają zastosowania w sondażach opartych na próbach imiennych, które dają możliwość doboru jednostek wprost z istniejącego wykazu konkretnych osób.

W przypadku projektu ESS przedstawione na rysunku porównania zawierają wyłącznie zestawienia grup sondaży wykorzystujących siatkę Kisha oraz daty urodzin (innych procedur w ESS nie stosowano), z kolei dla EVS obejmują one dodatkowo trzecią kategorię grupującą sondaże z doborem wewnątrzspołowym prowadzonym według reguł nielosowych. Z kolei dla ISSP zestawiono sondaże wykorzystujące zarówno procedury nielosowe, siatkę Kisha, jak też daty urodzin, przy czym te ostatnie podzielono, biorąc pod uwagę to, czy w procesie doboru respondentów docelowych uczestniczył ankieter, czy też domownik sam wskazywał osobę biorącą udział w badaniu.

Wyniki przedstawionych analiz ukazują kilka ciekawych prawidłowości wynikających ze stosowania odmiennych procedur selekcji wewnątrzspołowej w próbach adresowych. Po pierwsze, zestawiając miary jakości prób badawczych otrzymane w grupie krajowych sondaży wykorzystujących siatkę Kisha z miarami jakości otrzymanymi w grupie sondaży wykorzystujących procedury dat urodzin można zauważyć, że niższa mediana w tej drugiej grupie prób badawczych wskazuje na fakt wyższej jakości prób adresowych w sondażach, w któ-

rych selekcji respondentów docelowych dokonywano na podstawie daty urodzin domowników. Wynik ten pozostaje w zgodzie z ustaleniami poczynionymi w poprzednim rozdziale. Wskazywano już bowiem, że siatka Kisha (pomimo ściśle losowego charakteru) jest procedurą znacznie bardziej ingerującą w prywatność respondentów, gdyż wymaga sporządzenia pełnego wykazu domowników zamieszkujących gospodarstwa domowe, co w efekcie prowadzi częściej do odmowy udziału w badaniu już na etapie wstępnym, zanim jeszcze wylosowany zostanie respondent docelowy. Niezwykle interesujące okazują się również zestawienia otrzymane w projekcie ISSP. Specyfika tego badania, opisana dokładnie w rozdziale drugim, polega bowiem na stosowaniu w pewnych krajach technik pomiarowych wyłączających obecność ankietera (ankiety roznoszone, mailowe, pocztowe). Konsekwencją zastosowania takiej praktyki badawczej okazały się próby o zdecydowanie gorszej jakości.

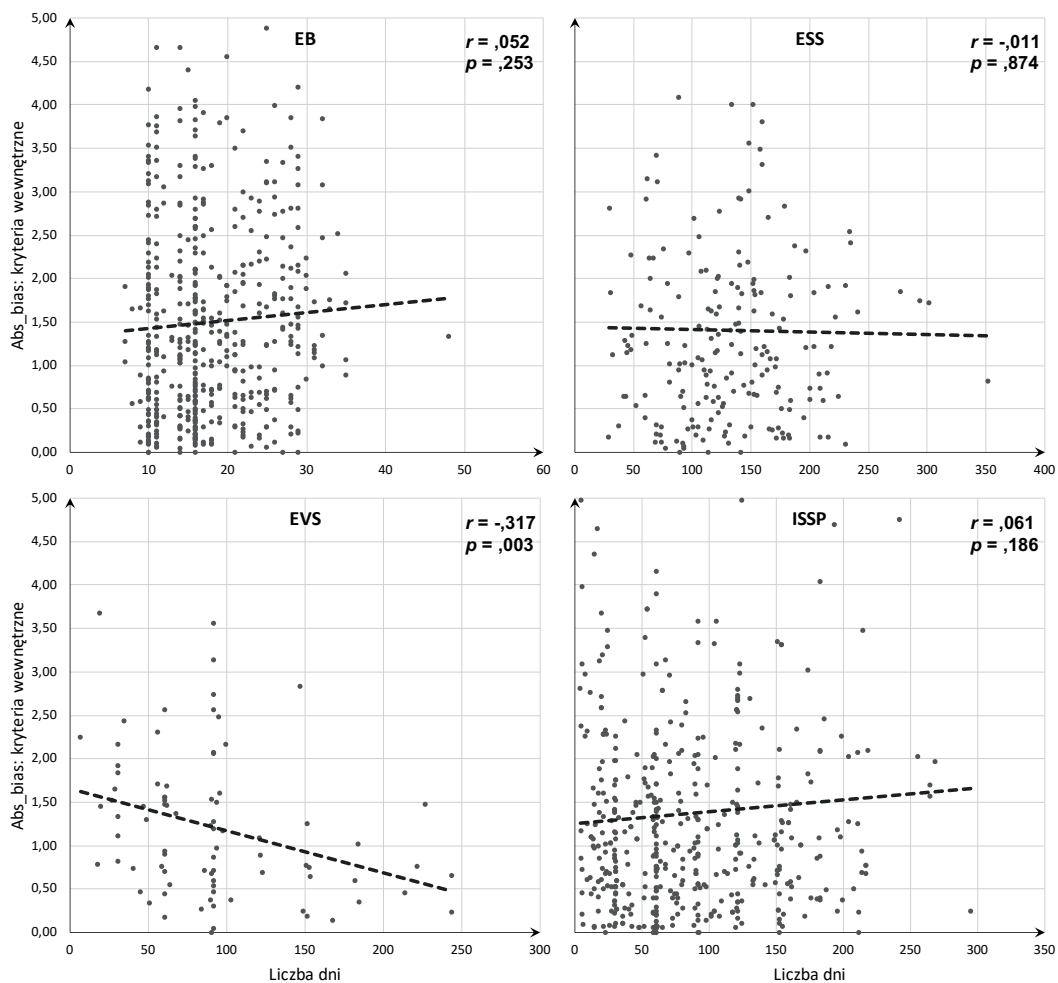


Rys. 24. Porównanie jakości sondażowych prób badawczych w projektach ESS, EVS oraz ISSP z uwagi na procedury selekcji respondentów docelowych w próbach adresowych

Uwagi: ESS: test mediany ( $\chi^2 = 1,17$ ;  $df = 1$ ;  $p = ,279$ ); test K-W ( $\chi^2 = ,97$ ;  $df = 1$ ;  $p = ,326$ );  $N = 107$ ; EVS: test mediany ( $\chi^2 = 1,02$ ;  $df = 2$ ;  $p = ,319$ ); test K-W ( $\chi^2 = 4,32$ ;  $df = 2$ ;  $p = ,115$ );  $N = 41$ ; ISSP: test mediany ( $\chi^2 = 29,09$ ;  $df = 3$ ;  $p < ,001$ ); test K-W ( $\chi^2 = 42,52$ ;  $df = 3$ ;  $p < ,001$ );  $N = 277$ .

Źródło: opracowanie własne.

Ostatni z serii wykresów prezentujących wpływ na jakość próby określonych procedur badawczych wykorzystywanych w losowaniu oraz terenowej realizacji krajowych prób sondażowych zawiera diagramy ukazujące korelacje zachodzące pomiędzy jakością sondażu oraz czasem trwania terenowej fazy badań (rys. 25). Dla każdego projektu podano informację o wartościach współczynników korelacji liniowej Pearsona wraz z  $p$ -wartościami pozwalającymi wnioskować o istotności liniowej zależności pomiędzy analizowanymi zmiennymi.



Rys. 25. Korelacja pomiędzy jakością próby oraz czasem trwania badań projektach EB, ESS, EVS oraz ISSP

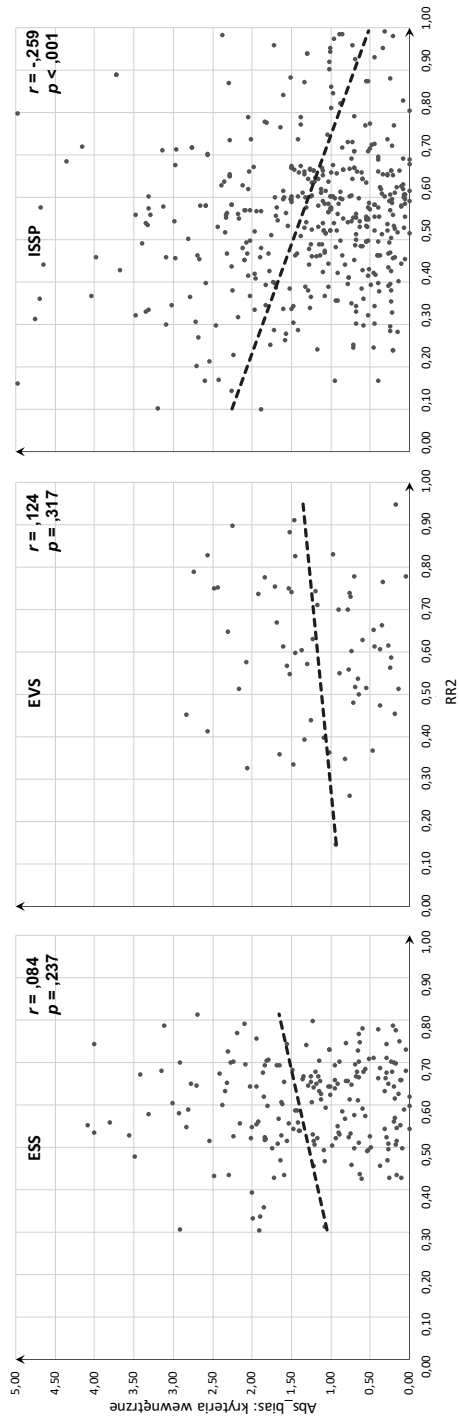
Źródło: opracowanie własne.

Przyglądając się zaprezentowanym zestawieniom, można zauważyć, że jedynie w projekcie EVS opisywana korelacja jest statystycznie istotna. Ma ona przy tym charakter odwrotnie proporcjonalny, co oznacza, że sondaże o dłuższym czasie przeznaczonym na terenową fazę badań (liczonym w dniach) charakteryzowały się lepszą jakością. Wniosek ten – choć pozostaje zgodny z intuicją – przyjąć należy z dużą ostrożnością, gdyż prezentowany układ zależności nie spełnia warunków homoskedastyczności reszt w modelu regresji ( $p$ -wartość w teście Breuscha-Pagana jest mniejsza niż 0,001). Jest to o tyle istotne, że w odniesieniu do pozostałych trzech porównywanych projektów zależności nie były już statystycznie istotne. Innymi słowy, przeprowadzone analizy nie dały wyraźnych podstaw do stwierdzenia istotnej zależności pomiędzy czasem trwania badań terenowych oraz uzyskiwaną jakością pomiaru.

### 2.3. Zależności pomiędzy jakością prób sondażowych oraz wskaźnikami jej terenowej realizacji

Na zakończenie warto przyjrzeć się jeszcze zależnościom pomiędzy jakością sondażowych prób badawczych oraz poziomem jej terenowej realizacji. Co prawda, problem ten podejmowany był w literaturze wielokrotnie, jednak perspektywa realizowanych na dużą skalę międzykrajowych sondaży porównawczych może rzucić – mimo wszystko – nowe światło na rozpoznane już zjawisko. Co więcej, wskaźniki realizacji próby (w szczególności doskonale znany miernik *response rate*) uznawane są nadal dość powszechnie za wyznaczniki jakości pomiaru. Innymi słowy, często bezrefleksyjnie przyjmuje się, że im wyższą wartość osiągnie terenowa realizacja próby, tym jakość sondażu będzie lepsza. Nie wchodząc w szczególności toczącej się debaty akademickiej, można przywołać konstatację Grovesa (2006) poczynioną w podsumowaniu przeprowadzonej przez niego metaanalizy kilkuset sondaży, z której wynika, iż „ślepa pogoń za wysokim wskaźnikiem realizacji próby jest nierozsądna; świadome dążenie do wysokiego odsetka realizacji próby jest mądre”. Oznacza to, że jakość pomiaru oraz poziom jej terenowej realizacji nie muszą być ze sobą istotnie skorelowane.

Rysunek 26 zawiera trzy diagramy obrazujące zależności pomiędzy poziomem realizacji próby w projektach ESS, EVS oraz ISSP (wyrażonej miernikiem *response rate* w wersji RR2) oraz jej jakością mierzoną kryterium wewnętrznym (z uwagi na niedostępność informacji o osiągniętych wartościach RR2 z analiz wyłączono EB). Dodatkowo podano również wartości współczynników korelacji liniowej Pearsona wraz z odpowiednimi  $p$ -wartościami pozwalającymi określić istotność współczynnika korelacji.

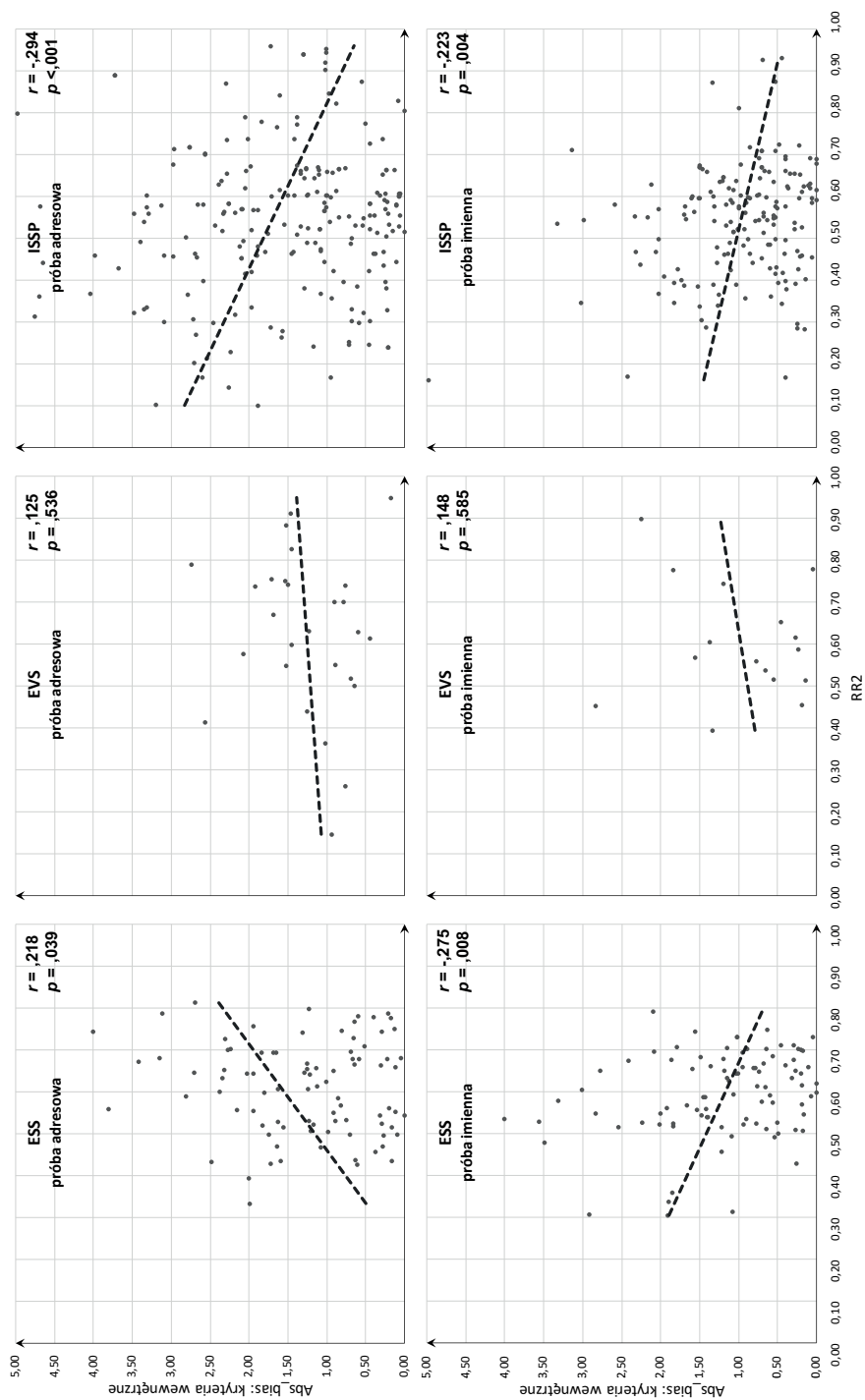


Rys. 26. Korelacja pomiędzy jakością próby oraz wskaźnikiem jej terenowej realizacji w projektach ESS, EVS oraz ISSP  
 Źródło: opracowanie własne.

Przyglądając się diagramom korelacyjnym, można zauważyć, że jedynie w projekcie ISSP jakość krajowych sondaży okazała się w sposób statystycznie istotny powiązana ze wskaźnikami terenowej realizacji próby. Ponieważ miary jakości sondaży pokazują wielkość obciążenia procesu selekcji (im mniejsza wartość, tym lepsza jakość pomiaru), to uzyskana odwrotnie proporcjonalna zależność ( $r = -0,259$ ;  $p < 0,01$ ) oznacza, iż krajowe sondaże o wyższych wartościach RR2 charakteryzowały się wyższą jakością pomiaru. Z drugiej strony można jednak zauważyć, że korelacja pomiędzy jakością krajowych prób badawczych oraz poziomem ich realizacji w projektach ESS oraz EVS okazała się już statystycznie nieistotna. Uzyskane wyniki nie dają zatem jednoznacznych konkluzji w kwestii opisu zależności pomiędzy wskaźnikami RR2 oraz jakością pomiaru.

Do niezwykle ciekawych wniosków wiedzie jednak dodatkowa analiza przeprowadzona w grupach sondaży wyróżnionych z uwagi na typ próby badawczej (rys. 27).

Analiza współzależności w podziale na próby imienne oraz adresowe wskazuje, co prawda, że w EVS korelacje są nadal nieistotne, natomiast w ISSP – niezależnie od typu próby – zachodzi umiarkowana (choć istotna statystycznie) zależność odwrotnie proporcjonalna. Do niezwykle ciekawych wniosków prowadzą jednak dane z projektu ESS. Okazuje się bowiem, że w grupie sondaży opartych na próbach imiennych wraz ze wzrostem wartości wskaźnika terenowej realizacji próby badawczej wzrastała jej jakość (istotnie mniejsze były bezwzględne wartości obciążeń wyliczone na podstawie kryteriów wewnętrznych), podczas gdy w próbach adresowych zależność okazała się odwrotna. Do podobnych ustaleń prowadzą zresztą analizy autorstwa Eckman i Koch (2019) oraz Jabkowski i Cichocki (2019). W obu pracach wskazano, że w zasadzie jedynym wyjaśnieniem takich różnic pomiędzy sondażami opartymi na próbach imiennych oraz adresowych jest większa swoboda, jaką w próbach adresowych posiadają ankieterzy w decydowaniu o wyborze konkretnych osób do badania. A zatem, ponieważ w próbach adresowych nadzór nad selekcją wewnątrzspołową jest ograniczony, to ankieterzy mają skłonność do wskazywania osób przebywających w domu jako spełniających kryteria selekcji (w tej kategorii przeważają kobiety). Działanie takie posiada w oczywisty sposób negatywny wpływ na wskaźniki jakości próby, prowadząc do nadreprezentacji kobiet, ale jednocześnie pozytywny wpływ na wskaźniki terenowej realizacji próby poprzez zmniejszenie odsetka odmów, który w efekcie prowadzi do wyższych wskaźników RR2.



Rys. 27. Korelacja pomiędzy jakością próby oraz wskaźnikami jej terenowej realizacji z uwagi na typ próby w projektach ESS, EVS oraz ISSP  
 Źródło: opracowanie własne.

### 3. Podsumowanie

Użytkownicy danych wtórnych podchodzą niejednokrotnie z nadmiernym zaufaniem do publicznie udostępnianych zbiorów danych z realizacji międzykrajowych sondaży porównawczych. Można to zaobserwować zwłaszcza wtedy, gdy projekty posiadają ustaloną renomę, a ich wyniki wykorzystywane są w opracowaniach naukowych. Z tego powodu – w zasadzie bezrefleksyjnie – użytkownicy danych traktują często rezultaty badań tak, jakby wysoka jakość realizacji krajowych pomiarów była czymś oczywistym. Taką postawę wymuszają praktyczne ograniczenia w dostępie do dokumentacji projektowej, albowiem znaczna część międzykrajowych badań porównawczych nie dostarcza nawet najbardziej podstawowych informacji o wykorzystanych procedurach doboru próby oraz jej terenowej realizacji. Oczywiście pewien stopień zaufania do standardów metodologicznych wdrażanych w poszczególnych projektach wydaje się konieczny, by wykorzystywać zbiory danych wynikowych w prowadzonych studiach o charakterze substancywnym, niemniej jednak bezkrytyczne zaufanie do wysokiej jakości pomiaru hamuje możliwość ulepszenia procedur terenowych oraz osłabia moc wnioskowania na podstawie wyników empirycznych uzyskanych na podstawie danych o niezwerifikowanej jakości.

Chociaż w literaturze badań sondażowych refleksja nad jakością pomiaru towarzyszyła od zawsze metodologicznie zorientowanym studiom nad metodą reprezentatywną, to jednak większość z zaproponowanych procedur służących ocenie jakości sondaży (osadzonych w dominującym paradygmacie całkowitego błędu pomiaru) nie znajduje zastosowania w zakrojonej na dużą skalę ewaluacji pomiarów pochodzących z wielu różnych międzykrajowych badań porównawczych. Podstawowym ograniczeniem stosowania większości takich procedur okazuje się przede wszystkim konieczność gromadzenia dodatkowych danych o wszystkich (bez wyjątku) jednostkach wylosowanych do próby, tj. także o osobach niedostępnych. Ta kłopotliwa konieczność wyklucza możliwość wykorzystania większości z dostępnych procedur w prowadzonej na dużą skalę ewaluacji sondaży.

Zaprezentowane analizy ukazały jednak, że użytkownicy danych wtórnych mogą wykorzystać miary oparte na kryteriach wewnętrznych, jako metodę przesiewania krajowych sondaży ułatwiającą podejmowanie decyzji o ich włączeniu do prowadzonych analiz. Kryteria wewnętrzne nie wymagają bowiem dostępu do żadnych dodatkowych informacji o jednostkach, poza tymi zbieranymi standardowo w zdecydowanej większości sondaży. Choć główną zaletą miar jakości pomiaru opartych na kryteriach wewnętrznych jest możliwość ich zastosowania na szeroką skalę, to jednak nawet jeżeli uda się zidentyfikować pomiary o podejrzanej jakości, to nadal niewiele wiadomo o ewentualnym wpływie jakości sondaży na analizy danych wynikowych. A zatem, choć do użytkownika danych wtórnych



należy ostateczna decyzja o wykluczeniu z analizy sondaży o podejrzonej jakości, to jednak już sama świadomość ograniczeń wynikających z wnioskowania na podstawie danych o dyskusyjnej jakości może okazać się niezwykle przydatna w interpretowaniu niespodziewanych wyników. Takie podejście wydaje się szczególnie istotne w badaniach porównawczych, których celem jest analiza oraz wyjaśnienie międzykrajowego zróżnicowania danych.

## Zakończenie

W okresie ostatnich dwudziestu lat zauważyć można znaczący przyrost liczby oraz jakości realizowanych w Europie międzykrajowych projektów porównawczych o charakterze sondażowym. Dotyczy to zarówno przedsięwzięć posiadających charakter paneuropejski, jak również projektów o zasięgu regionalnym, prowadzonych najczęściej w kilku sąsiadujących krajach. Jednak duża liczba i zwykle akceptowalna jakość europejskich badań międzykrajowych napotyka na prawdziwą przeszkodę w postaci niejasnej różnorodności. A zatem, chociaż znaczące europejskie projekty sondażowe posiadają wspólne cechy, jeśli chodzi o rozwiązania metodologiczne i podejmowane wątki problemowe, to jednocześnie okazują się zasadniczo zróżnicowane pod względem konkretnych celów badawczych i poruszanych kwestii, zakresu czasowego i przestrzennego prowadzonego pomiaru, a także stosowania określonych procedur terenowych oraz sposobów operacjonalizacji problematyki badawczej. Koordynacja między różnymi projektami jest niewielka, jeśli w ogóle występuje, zwłaszcza w odniesieniu do stosowania konkretnych rozwiązań metodologicznych, ustalenia wspólnego zbioru wskaźników, czy też formułowania pytań kwestionariuszowych. Ta niejednorodność utrudnia bezpośrednio międzyprojektowe porównania wyników oraz ogranicza możliwość prowadzenia analiz o charakterze substancyjnym.

Europejskie badania opinii społecznej mają jednocześnie kluczowe znaczenie dla debaty publicznej oraz akademickiej skoncentrowanej na problematyce integracji europejskiej. Jeśli jednak badania sondażowe stanowią jeden z wielu punktów odniesienia dla dyskursu publicznego na poziomie poszczególnych państw narodowych, to ich rola w kształtowaniu dyskusji na poziomie między- i ponadnarodowym okazuje się już dużo bardziej znacząca. W przypadku Unii Europejskiej agregaty sondażowe pozostają nadal jednym z niewielu uznanych oraz wiarygodnych źródeł wiedzy o opiniach i postawach obywateli różnych europejskich państw, stanowiąc tym samym wsparcie dla dyskursu na temat europeizacji. Z drugiej jednak strony, pomimo pełnienia funkcji użytkowej, zdecydowana większość europejskich sondaży nie służy potrzebom instytucji unijnych i dotyczy kwestii

jedynie luźno powiązanych z integracją europejską. Mimo wszystko projekty takie stanowią ważne źródło wiedzy na temat europejskiej opinii publicznej, wypełniając lukę informacyjną będącą efektem słabej integracji społecznej na poziomie ponadnarodowym. Polityczna rzeczywistość integracji europejskiej stworzyła trwały popyt na badania międzykrajowe już w latach 70. XX wieku, a zatem bogata i długa tradycja ich prowadzenia nie powinna dziwić. Co więcej, Europa pozostała by prawdopodobnie atrakcyjnym obiektem badań międzynarodowych nawet bez systemowej integracji europejskiej. Pomimo wielu cech wspólnych, historycznie ugruntowane państwa narodowe i ich społeczeństwa charakteryzują się w dalszym ciągu ogromną różnorodnością. Na tak pluralistycznym kontynencie, jakim jest Europa, analizy społecznych postaw obywateli opartej na danych pochodzących z pomiaru sondażowego pojedynczych państw cieszą się znacznie mniejszym zainteresowaniem, ponieważ kraje europejskie są liczne, względnie homogeniczne i w zdecydowanej większości małe, zwłaszcza jeżeli pod uwagę weźmie się obszar lub liczebność populacji.

Głównym celem tej publikacji było ukazanie zasadniczych wyzwań metodologicznych, jakim sprostać muszą badacze wykorzystujący wyniki badań z międzykrajowych projektów sondażowych. O ile wyzwania związane z niejednorodnością metodologiczną oraz brakiem międzykrajowej ekwiwalencji pomiarowej w obrębie wyników uzyskanych w ramach jednego projektu sondażowego są stosunkowo dobrze rozpoznane w literaturze światowej, o tyle namysł nad ograniczeniami wynikającymi z wykorzystania danych nie tyle z różnych krajów, co uzyskiwanych w ramach realizacji wielu – często całkowicie odmiennych – projektów sondażowych, nie były jak dotąd przedmiotem szczegółowych oraz systematycznych analiz metodologicznych. Podejmowano je co prawda (po części) w związku z realizacją modnych ostatnio projektów harmonizacyjnych, jednak dużo większą wagę przywiązywano w takich przedsięwzięciach do wypracowania procedur analitycznych przekształcających odmienne zakresy wartości zmiennych źródłowych do wspólnego (ujednoliconego) zbioru wartości, aniżeli do refleksji nad wpływem projektowych specyfik próbkowania, terenowej realizacji próby oraz procesu pomiaru na uzyskiwane wyniki.

Zaprezentowane dane oparte na analizie ponad 1,4 tys. krajowych sondaży pochodzących z realizacji czterech wiodących projektów sondażowych o charakterze porównawczym, tj. badań Eurobarometru, Europejskiego Sondażu Społecznego, Europejskiego Sondażu Wartości oraz Międzynarodowego Programu Badań Społecznych, ukazały skalę różnorodności metodologicznej nie tylko w wymiarze międzyprojektowym, ale również wewnątrzprojektowym oraz wzdłużczasowym. Owa różnorodność dotyczy przede wszystkim zasięgu geograficznego poszczególnych przedsięwzięć sondażowych, podejmowanej problematyki badawczej, konstruowania pytań kwestionariuszowych, procedur próbkowania oraz procedur

terenowych, standardów przyjmowanych w odniesieniu do raportowania o metodologii oraz rezultatach terenowej realizacji próby, czy też wreszcie uzyskanej jakości pomiaru sondażowego. Biorąc pod uwagę znaczące zróżnicowanie projektów oraz potencjalny wpływ specyfik projektowych na wyniki pomiaru należałoby rozważyć włączenie meta-charakterystyk krajowych sondaży, cech pomiaru, sposobów formułowania pytań, jak również miar jakości próby sondażowej, do międzyprojektowej oraz wewnątrzprojektowej analizy danych. Wprawdzie wymaga to każdorazowego podjęcia decyzji dotyczącej konstrukcji konkretnych modeli analitycznych, stanowi jednak nadzieję na uporanie się z różnorodnością projektową lub przynajmniej na kontrolowanie wpływu specyfik konkretnego pomiaru sondażowego w międzyprojektowo zorientowanej analizie danych.



## Odniesienia literaturowe do zbiorów danych oraz dokumentacji projektów EB, ESS, EVS oraz ISSP

### Eurobarometer

- European Commission (2001): Eurobarometer 56.2 (2001). European Opinion Research Group, Brussels [Producer]. GESIS Data Archive, Cologne. ZA3627 data file version 1.0.1, doi:10.4232/1.10946.
- European Commission (2002): Eurobarometer 58.1 (2002). European Opinion Research Group, Brussels [Producer]. GESIS Data Archive, Cologne. ZA3693 data file version 1.0.1, doi:10.4232/1.10953.
- European Commission (2003): Eurobarometer 60.1 (2003). European Opinion Research Group, Brussels [Producer]. GESIS Data Archive, Cologne. ZA3938 data file version 1.0.1, doi:10.4232/1.10958.
- European Commission (2004): Candidate Countries Eurobarometer 2001.1. The GALLUP Organization Hungary, Budapest. GESIS Data Archive, Cologne. ZA3978 Data file Version 1.0.0, doi:10.4232/1.3978.
- European Commission (2004): Candidate Countries Eurobarometer 2002.2. The GALLUP Organization Hungary, Budapest. GESIS Data Archive, Cologne. ZA3979 Data file Version 1.0.0, doi:10.4232/1.3979.
- European Commission (2004): Candidate Countries Eurobarometer 2003.4. The GALLUP Organization Hungary, Budapest. GESIS Data Archive, Cologne. ZA3986 Data file Version 1.0.0, doi:10.4232/1.3986.
- European Commission (2004): Eurobarometer 62.0 (2004). TNS OPINION & SOCIAL, Brussels [Producer]. GESIS Data Archive, Cologne. ZA4229 data file version 1.1.0, doi:10.4232/1.10962.
- European Commission (2005): Eurobarometer 64.2 (2005). TNS OPINION & SOCIAL, Brussels [Producer]. GESIS Data Archive, Cologne. ZA4414 data file version 1.1.0, doi:10.4232/1.10970.
- European Commission (2006): Eurobarometer 66.1 (2006). TNS OPINION & SOCIAL, Brussels [Producer]. GESIS Data Archive, Cologne. ZA4526 data file version 1.0.1, doi:10.4232/1.10980.
- European Commission (2007): Eurobarometer 68.1 (2007). TNS Opinion & Social, Brussels [Producer]. GESIS Data Archive, Cologne. ZA4565 data file version 4.0.1, doi:10.4232/1.10988.
- European Commission (2008): Eurobarometer 70.1 (2008). TNS Opinion & Social, Brussels [Producer]. GESIS Data Archive, Cologne. ZA4819 data file version 3.0.2, doi:10.4232/1.10989.
- European Commission (2009): Eurobarometer 72.4 (2009). TNS Opinion & Social, Brussels [Producer]. GESIS Data Archive, Cologne. ZA4994 data file version 3.0.0, doi:10.4232/1.11141.

- European Commission (2010): Eurobarometer 74.2 (2010). TNS Opinion & Social, Brussels [Producer]. GESIS Data Archive, Cologne. ZA5449 data file version 2.2.0, doi:10.4232/1.11626.
- European Commission (2014): Eurobarometer 76.3 (2011). TNS Opinion & Social. GESIS Data Archive, Cologne. ZA5567 Data file Version 2.0.1, doi:10.4232/1.12007.
- European Commission (2015): Eurobarometer 78.1 (2012). TNS Opinion. GESIS Data Archive, Cologne. ZA5685 Data file Version 2.0.0, doi:10.4232/1.12061.
- European Commission (2017): Eurobarometer 80.1 (2013). TNS Opinion. GESIS Data Archive, Cologne. ZA5876 Data file Version 2.0.0, doi:10.4232/1.12768.
- European Commission (2017): Eurobarometer 82.3 (2014). TNS Opinion. GESIS Data Archive, Cologne. ZA5932 Data file Version 2.1.0, doi:10.4232/1.12854.
- European Commission (2017): Eurobarometer 84.3 (2015). TNS Opinion. GESIS Data Archive, Cologne. ZA6643 Data file Version 3.1.0, doi:10.4232/1.12799.
- European Commission (2017): Eurobarometer 86.2 (2016). TNS Opinion. GESIS Data Archive, Cologne. ZA6788 Data file Version 1.3.0, doi:10.4232/1.12853.
- European Commission (2017): Eurobarometer 87.3 (2017). TNS Opinion. GESIS Data Archive, Cologne. ZA6863 Data file Version 1.0.0, doi:10.4232/1.12847.

## European Social Survey

- ESS Round 1: European Social Survey (2016): ESS-1 2002 Documentation Report. Edition 6.5. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 1: European Social Survey Round 1 Data (2002). Data file edition 6.5. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 2: European Social Survey (2016): ESS-2 2004 Documentation Report. Edition 3.6. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 2: European Social Survey Round 2 Data (2004). Data file edition 3.5. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 3: European Social Survey (2016): ESS-3 2006 Documentation Report. Edition 3.6. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 3: European Social Survey Round 3 Data (2006). Data file edition 3.6. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 4: European Social Survey (2016): ESS-4 2008 Documentation Report. Edition 5.4. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 4: European Social Survey Round 4 Data (2008). Data file edition 4.4. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 5: European Social Survey (2016): ESS-5 2010 Documentation Report. Edition 4.1. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 5: European Social Survey Round 5 Data (2010). Data file edition 3.3. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 6: European Social Survey (2016): ESS-6 2012 Documentation Report. Edition 2.3. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.

- ESS Round 6: European Social Survey Round 6 Data (2012). Data file edition 2.3. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 7: European Social Survey (2016): ESS-7 2014 Documentation Report. Edition 3.1. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 7: European Social Survey Round 7 Data (2014). Data file edition 2.1. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 8: European Social Survey (2018): ESS-8 2016 Documentation Report. Edition 2.0. Bergen, European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- ESS Round 8: European Social Survey Round 8 Data (2016). Data file edition 2.0. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

## European Values Study

- Brislinger, E., Nijs Bik, E. d., Harzenetter, K., Hauser, K., Kampmann, J., Kurti, D., Zenk-Möltgen, W. (2011). European Values Study 2008: project and data management (GESIS-Technical Reports, 2011/14). Köln: GESIS – Leibniz-Institut für Sozialwissenschaften. <http://nbn-resolving.de/urn:nbn:de:0168-ssaoar-288600>.
- EVS, GESIS (2010): EVS 2008 Method Report. GESIS-Technical Reports 2010/17. Retrieved from <http://www.europeanvaluesstudy.eu/>.
- EVS, GESIS (2011): EVS – European Values Study 1981 – Integrated Dataset. GESIS Data Archive, Cologne. ZA4438 Data file Version 3.0.0, doi:10.4232/1.10791.
- EVS, GESIS (2011): EVS – European Values Study 1990 – Integrated Dataset. GESIS Data Archive, Cologne. ZA4460 Data file Version 3.0.0, doi:10.4232/1.10790.
- EVS, GESIS (2011): EVS – European Values Study 1999 – Integrated Dataset. GESIS Data Archive, Cologne. ZA3811 Data file Version 3.0.0, doi:10.4232/1.10789.
- EVS, GESIS (2013): European Values Study 1981 – Variable Report. GESIS-Variable Reports 2013/24.
- EVS, GESIS (2015): European Values Study Longitudinal Data File 19812008 (EVS 1981-2008). GESIS Data Archive, Cologne. ZA4804 Data file Version 3.0.0, doi:10.4232/1.12253.
- EVS, GESIS (2016): European Values Study 2008: Integrated Dataset. GESIS Data Archive, Cologne. ZA4800 Data file Version 4.0.0, doi:10.4232/1.12458.
- EVS, GESIS (2016): EVS 2008 Method Report. GESIS Papers 2016|18. Retrieved from <http://www.europeanvaluesstudy.eu/>.
- Halman, L. C. J. M. (2001). The European values study: A third wave. Sourcebook of the 1999/2000 European values study surveys.

## International Social Survey Programme

- Andersen, Jørgen Goul; Tonboe, Jens Christian; Borre, Ole; Togeby, Lise; Nielsen, Hans Jørgen; Andersen, Bjarne Hjorth; Kjær, Ulrik; Fridberg, Torben (2014): International Social Survey Programme 2007 (Denmark). GESIS Data Archive, Cologne. ZA4831 Data file Version 1.0.0, doi:10.4232/1.4831.
- Andersen, Jørgen Goul; Tonboe, Jens Christian; Borre, Ole; Togeby, Lise; Nielsen, Hans Jørgen; Andersen, Bjarne Hjorth; Kjær, Ulrik (2014): International Social Survey Programme 1999 (Denmark). GESIS Data Archive, Cologne. ZA3562 Data file Version 1.0.0, doi:10.4232/1.11835.



- Becker, Jos; Niggebrugge, Dono E. (2014): International Social Survey Programme 2001 (Netherlands). GESIS Data Archive, Cologne. ZA3916 Data file Version 1.0.0, doi:10.4232/1.11832.
- Becker, Jos; Niggebrugge, Dono E. (2014): International Social Survey Programme 1999 (Netherlands). GESIS Data Archive, Cologne. ZA3293 Data file Version 1.0.0, doi:10.4232/1.11837.
- Diekmann, Andreas; Franzen, Axel; Meyer Schweizer, Ruth (2014): International Social Survey Programme 1993 (Switzerland). GESIS Data Archive, Cologne. ZA3024 Data file Version 1.0.0, doi:10.4232/1.11845.
- Diez-Nicolás, Juan (2014): International Social Survey Programme 1992 (Spain). GESIS Data Archive, Cologne. ZA2496 Data file Version 1.0.0, doi:10.4232/1.11843.
- Ganzeboom, Harry (2014): International Social Survey Programme: Environment III – ISSP 2010 (Netherlands). GESIS Data Archive, Cologne. ZA5513 Data file Version 1.0.0, doi:10.4232/1.11792.
- Ganzeboom, Harry (2017): International Social Survey Programme: National Identity III – ISSP 2013 (Netherlands). GESIS Data Archive, Cologne. ZA5517 Data file Version 1.0.0, doi:10.4232/1.12921.
- Ganzeboom, Harry B.G. (2014): International Social Survey Programme 2007 (Netherlands). GESIS Data Archive, Cologne. ZA4861 Data file Version 1.0.0, doi:10.4232/1.4861.
- Ganzeboom, Harry B.G. (2015): International Social Survey Programme: Social Inequality IV – ISSP 2009 (Netherlands). GESIS Data Archive, Cologne. ZA5995 Data file Version 1.0.0, doi:10.4232/1.12338.
- Höllinger, Franz; Haller, Max (2014): International Social Survey Programme 1993 (Austria). GESIS Data Archive, Cologne. ZA2793 Data file Version 1.0.0, doi:10.4232/1.11844.
- Iglic, Hajdeja (2014): International Social Survey Programme 1986 (Slovenia). GESIS Data Archive, Cologne. ZA1784 Data file Version 1.0.0, doi:10.4232/1.11841.
- ISSP Research Group (1986): International Social Survey Programme: Role of Government I – ISSP 1985. GESIS Data Archive, Cologne. ZA1490 Data file Version 1.0.0, doi:10.4232/1.1490.
- ISSP Research Group (1988): International Social Survey Programme: Social Networks and Support Systems – ISSP 1986. GESIS Data Archive, Cologne. ZA1620 Data file Version 1.0.0, doi:10.4232/1.1620.
- ISSP Research Group (1989): International Social Survey Programme: Social Inequality I – ISSP 1987. GESIS Data Archive, Cologne. ZA1680 Data file Version 1.0.0, doi:10.4232/1.1680.
- ISSP Research Group (1990): International Social Survey Programme: Family and Changing Gender Roles I – ISSP 1988. GESIS Data Archive, Cologne. ZA1700 Data file Version 1.0.0, doi:10.4232/1.1700.
- ISSP Research Group (1991): International Social Survey Programme: Work Orientations I – ISSP 1989. GESIS Data Archive, Cologne. ZA1840 Data file Version 1.0.0, doi:10.4232/1.1840.
- ISSP Research Group (1992): International Social Survey Programme: Role of Government II – ISSP 1990. GESIS Data Archive, Cologne. ZA1950 Data file Version 1.0.0, doi:10.4232/1.1950.
- ISSP Research Group (1993): International Social Survey Programme: Religion I – ISSP 1991. GESIS Data Archive, Cologne. ZA2150 Data file Version 1.0.0, doi:10.4232/1.2150.
- ISSP Research Group (1994): International Social Survey Programme: Social Inequality II – ISSP 1992. GESIS Data Archive, Cologne. ZA2310 Data file Version 1.0.0, doi:10.4232/1.2310.
- ISSP Research Group (1995): International Social Survey Programme: Environment I – ISSP 1993. GESIS Data Archive, Cologne. ZA2450 Data file Version 1.0.0, doi:10.4232/1.2450.

- ISSP Research Group (1997): International Social Survey Programme: Family and Changing Gender Roles II – ISSP 1994. GESIS Data Archive, Cologne. ZA2620 Data file Version 1.0.0, doi:10.4232/1.2620.
- ISSP Research Group (1998): International Social Survey Programme: National Identity I – ISSP 1995. GESIS Data Archive, Cologne. ZA2880 Data file Version 1.0.0, doi:10.4232/1.2880.
- ISSP Research Group (1999): International Social Survey Programme: Role of Government III – ISSP 1996. GESIS Data Archive, Cologne. ZA2900 Data file Version 1.0.0, doi:10.4232/1.2900.
- ISSP Research Group (1999): International Social Survey Programme: Work Orientations II – ISSP 1997. GESIS Data Archive, Cologne. ZA3090 Data file Version 1.0.0, doi:10.4232/1.3090.
- ISSP Research Group (2000): International Social Survey Programme: Religion II – ISSP 1998. GESIS Data Archive, Cologne. ZA3190 Data file Version 1.0.0, doi:10.4232/1.3190.
- ISSP Research Group (2002): International Social Survey Programme: Social Inequality III – ISSP 1999. GESIS Data Archive, Cologne. ZA3430 Data file Version 1.0.0, doi:10.4232/1.3430.
- ISSP Research Group (2003): International Social Survey Programme: Environment II – ISSP 2000. GESIS Data Archive, Cologne. ZA3440 Data file Version 1.0.0, doi:10.4232/1.3440.
- ISSP Research Group (2003): International Social Survey Programme: Social Relations and Support Systems / Social Networks II – ISSP 2001. GESIS Data Archive, Cologne. ZA3680 Data file Version 1.0.0, doi:10.4232/1.3680.
- ISSP Research Group (2008): International Social Survey Programme: Role of Government IV – ISSP 2006. GESIS Data Archive, Cologne. ZA4700 Data file Version 1.0.0, doi:10.4232/1.4700.
- ISSP Research Group (2009): International Social Survey Programme: Leisure Time and Sports – ISSP 2007. GESIS Data Archive, Cologne. ZA4850 Data file Version 2.0.0, doi:10.4232/1.10079.
- ISSP Research Group (2012): International Social Survey Programme: National Identity II – ISSP 2003. GESIS Data Archive, Cologne. ZA3910 Data file Version 2.1.0, doi:10.4232/1.11449.
- ISSP Research Group (2012): International Social Survey Programme: Citizenship – ISSP 2004. GESIS Data Archive, Cologne. ZA3950 Data file Version 1.3.0, doi:10.4232/1.11372.
- ISSP Research Group (2012): International Social Survey Programme: Environment III – ISSP 2010. GESIS Data Archive, Cologne. ZA5500 Data file Version 2.0.0, doi:10.4232/1.11418.
- ISSP Research Group (2013): International Social Survey Programme: Family and Changing Gender Roles III – ISSP 2002. GESIS Data Archive, Cologne. ZA3880 Data file Version 1.1.0, doi:10.4232/1.11564.
- ISSP Research Group (2013): International Social Survey Programme: Work Orientation III – ISSP 2005. GESIS Data Archive, Cologne. ZA4350 Data file Version 2.0.0, doi:10.4232/1.11648.
- ISSP Research Group (2013): Religion Around the World Study of the 2008 International Social Survey Programme (ISSP). GESIS Data Archive, Cologne. ZA5690 Data file Version 1.0.1, doi:10.4232/1.11762.
- ISSP Research Group (2015): International Social Survey Programme: Health and Health Care – ISSP 2011. GESIS Data Archive, Cologne. ZA5800 Data file Version 3.0.0, doi:10.4232/1.12252.
- ISSP Research Group (2015): International Social Survey Programme: National Identity III – ISSP 2013. GESIS Data Archive, Cologne. ZA5950 Data file Version 2.0.0, doi:10.4232/1.12312.
- ISSP Research Group (2016): International Social Survey Programme: Family and Changing Gender Roles IV – ISSP 2012. GESIS Data Archive, Cologne. ZA5900 Data file Version 4.0.0, doi:10.4232/1.12661.
- ISSP Research Group (2016): International Social Survey Programme: Citizenship II – ISSP 2014. GESIS Data Archive, Cologne. ZA6670 Data file Version 2.0.0, doi:10.4232/1.12590.

- ISSP Research Group (2017): International Social Survey Programme: Social Inequality IV – ISSP 2009. GESIS Data Archive, Cologne. ZA5400 Data file Version 4.0.0, doi:10.4232/1.12777.
- ISSP Research Group (2017): International Social Survey Programme: Work Orientations IV – ISSP 2015. GESIS Data Archive, Cologne. ZA6770 Data file Version 2.1.0, doi:10.4232/1.12848.
- Nic Ghiolla Phádraig, Máire (2014): International Social Survey Programme 1999 (Ireland). GESIS Data Archive, Cologne. ZA3613 Data file Version 1.0.0, doi:10.4232/1.11836.
- Nic Ghiolla Phádraig, Máire (2014): International Social Survey Programme 1997 (Ireland). GESIS Data Archive, Cologne. ZA3612 Data file Version 1.0.0, doi:10.4232/1.11839.
- Piscová, Magdalena (2014): International Social Survey Programme 2006 (Slovakia). GESIS Data Archive, Cologne. ZA4710 Data file Version 1.0.0, doi:10.4232/1.4710.
- Stamm, Hanspeter; Nef, Rolf; Suter, Christian; Joye, Dominique (2014): International Social Survey Programme 1999 (Switzerland). GESIS Data Archive, Cologne. ZA3297 Data file Version 1.0.0, doi:10.4232/1.11838.
- Vala, Jorge; Ramos, Alice (2014): International Social Survey Programme: Environment III – ISSP 2010 (Portugal). GESIS Data Archive, Cologne. ZA5516 Data file Version 1.0.0, doi:10.4232/1.11793.
- Whelan, Brendan; Ward, Conor (2014): International Social Survey Programme 1987 (Ireland). GESIS Data Archive, Cologne. ZA1306 Data file Version 1.0.0, doi:10.4232/1.1306.
- Whelan, Brendan; Ward, Conor (2014): International Social Survey Programme 1986 (Ireland). GESIS Data Archive, Cologne. ZA1303 Data file Version 1.0.0, doi:10.4232/1.1303.

## Bibliografia

- AAPOR (2016), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, American Association for Public Opinion Research.
- Allum, N., Read, S., Sturgis, P. (2011), *Evaluating Change in Social and Political Trust in Europe*, [w:] *Cross-Cultural Analysis: Methods and Applications*, red. E. Davidov, P. Schmidt, J. Billiet, New York: Routledge, s. 35-53.
- Alwin, D. F. (2007), *Margins of Error: A Study of Reliability in Survey Measurement*, New York: Wiley.
- Antonsich, M. (2008), *EUropean Attachment and Meanings of EUrope. A Qualitative Study in the EU-15*, *Political Geography*, 27(6), s. 691-710.
- Antonsich, M. (2009), *National Identities in the Age of Globalisation: The Case of Western Europe*, *National Identities*, 11(3), s. 281-299.
- Antonsich, M., Holland, E. C. (2012), *Territorial Attachment in the Age of Globalization: The Case of Western Europe*, *European Urban and Regional Studies*, 21(2), s. 206-221.
- Archer, M. (1996), *Social Integration and System Integration: Developing the Distinction*, *Sociology*, 30(4), s. 679-699.
- Armingeon, K., Ceka, B. (2014), *The Loss of Trust in the European Union during the Great Recession Since 2007: The Role of Heuristics from the National Political System*, *European Union Politics*, 15(1), s. 82-107.
- Arts, W., Halman, L. (2006), *Identity: The Case of the European Union*, *Journal of Civil Society*, 2(3), s. 179-198.
- Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., Tourangeau, R. (2013), *Summary Report of the AAPOR Task Force on Non-Probability Sampling*, *Journal of Survey Statistics and Methodology*, 1(2), s. 90-143.
- Baldissera, S., Ferrante, G., Campostrini, S. (2020), *Field Substitution in Surveys*, [w:] *Wiley StatsRef: Statistics Reference Online*, red. N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri, J. L. Teugels, New York: Wiley, s. 1-7.
- Baranowski, M., Jabkowski, P. (2019), *Perception of Social Welfare and Psychological Well-Being in The Western Balkan Countries*, [w:] *XII. Balkan Tarihi Kongresi „Bitişinin 100. Yılında I. Dünya Savaşı, Balkanlar ve Türkiye”*, red. B. Kocaoğlu, Samsun: Baski - Cilt, s. 17-34.
- Batz, C., Parrigon, S., Tay, L. (2016), *The impact of Scale Transformations on National Subjective Well-Being Scores*, *Social Indicators Research*, 129(1), s. 13-27.
- Bauer, J. J. (2014), *Selection Errors of Random Route Samples*, *Sociological Methods & Research*, 43(3), s. 519-544.

- Bauer, J. J. (2016), *Biases in Random Route Surveys*, *Journal of Survey Statistics and Methodology*, 4(2), s. 263-287.
- Beaumont, J. (2005), *On the Use of Data Collection Process Information for the Treatment of Unit Non-response through Weight Adjustment*, *Survey Methodology*, 31(2), s. 227.
- Beebe, T. J., Davern, M. E., McAlpine, D. D., Ziegenfuss, J. K. (2007), *Comparison of Two Within-Household Selection Methods in a Telephone Survey of Substance Abuse and Dependence*, *Annals of Epidemiology*, 17(6), s. 458-463.
- Berger, Y. G., Tillé, Y. (2009), *Sampling with Unequal Probabilities*, [w:] *Handbook of Statistics*, red. C. R. Rao, Amsterdam: Elsevier, s. 39-54.
- Best, S. J., Krueger, B. (2002), *New Approaches to Assessing Opinion: The Prospects for Electronic Mail Surveys*, *International Journal of Public Opinion Research*, 14(1), s. 73-92.
- Bethlehem, J., Cobben, F., Schouten, B. (2011), *Handbook of Nonresponse in Household Surveys*, New York: Wiley.
- Bethlehem, J. G., Kersten, M. (1985), *On the Treatment of Nonresponse in Sample Surveys*, *Journal of Official Statistics*, 1(3), s. 287.
- Beullens, K., Loosveldt, G. (2016), *Interviewer Effects in the European Social Survey*, *Survey Research Methods*, 10(2), s. 103-118.
- Beullens, K., Matsuo, H., Loosveldt, G., Vandenplas, C. (2014), *Quality Report for the European Social Survey, round 6*. [https://www.europeansocialsurvey.org/docs/round6/methods/ESS6\\_quality\\_report.pdf](https://www.europeansocialsurvey.org/docs/round6/methods/ESS6_quality_report.pdf) [dostęp: 26.05.2021].
- Biemer, P. P. (2010a), *Overview of Design Issues: Total Survey Error*, [w:] *Handbook of Survey Research*, red. P. V. Marsden, J. D. Wright, Bingley, UK: Emerald, s. 27-57.
- Biemer, P. P. (2010b), *Total Survey Error: Design, Implementation, Evaluation*, *Public Opinion Quarterly*, 74(5), s. 817-848.
- Biemer, P. P., De Leeuw, E. D., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L. E., Tucker, N. C., West, B. T. (2017), *Total Survey Error in Practice*, New York: John Wiley & Sons.
- Biemer, P. P., Lyberg, L. E. (2003), *Introduction to Survey Quality*, New York: Wiley.
- Billiet, J., Matsuo, H., Beullens, K., Vehovar, V. (2009), *Non-Response Bias in Cross-National Surveys: Designs for Detection and Adjustment in the ESS*, *ASK. Research & Methods*, 18(1), s. 3-43.
- Billiet, J., Philippens, M., Fitzgerald, R., Stoop, I. (2007), *Estimation of Nonresponse Bias in the European Social Survey: Using Information from Reluctant Respondents*, *Journal of Official Statistics*, 23(2), s. 135.
- Binson, D., Catania, J. A. (2000), *Random Selection in a National Telephone Survey: A Comparison of the Kish, Next-Birthday, Last-Birthday Methods*, *Journal of Official Statistics*, 16(1), s. 53.
- Breen, M. J., Healy, A. E. (2016), *Changing Values, Attitudes and Behaviours in Ireland: An Analysis of European Social Survey Data in Ireland, 2002-2012*, Cambridge: Cambridge Scholars Publishing.
- Brick, J. M., Montaquila, J. M. (2009), *Nonresponse and Weighting*, [w:] *Handbook of Statistics*, red. C. R. Rao, Amsterdam: Elsevier, s. 163-185.
- Chapman, D. W. (2003), *To Substitute or not to Substitute that is the Question*, *The Survey Statistician*, 48, s. 32-34.
- Commission, E. (2017), *Eurobarometer 87.3, 2017 (Study description in DDI format DDI-lifecycle 3.2) [Data sets, questionnaires and other documents]*. <https://doi.org/10.4232/1.12847> [dostęp: 26.05.2021].
- Cornesse, C., Bosnjak, M. (2018), *Is there an Association between Survey Characteristics and Representativeness? A Meta-Analysis*, *Survey Research Methods*, 12(1), s. 1-13.

- Coromina, L., Saris, W. E. (2009), *Quality of Media Use Measurement*, *International Journal of Public Opinion Research*, 21(4), s. 424-450.
- Couper, M. P., De Leeuw, E. D. (2003), *Nonresponse in Cross-Cultural and Cross-National Surveys*, [w:] *Cross-Cultural Survey Methods*, red. J. Harkness, F. van de Vijver, P. P. Mohler, New York: Wiley, s. 157-177.
- Couper, M. P., Kapteyn, A., Schonlau, M., Winter, J. (2007), *Noncoverage and Nonresponse in an Internet Survey*, *Social Science Research*, 36(1), s. 131-148.
- Cox, B. G. (2008), *Target Population*, [w:] *Encyclopedia of Survey Research Methods*, red. P. J. Lavrakas, New York: SAGE, s. 876-877.
- Curtin, R., Presser, S., Singer, E. (2005), *Changes in Telephone Survey Nonresponse over the Past Quarter Century*, *Public Opinion Quarterly*, 69(1), s. 87-98.
- Czajka, R., Blair, J., Sebestik, J. P. (1982), *Respondent Selection in a Telephone Survey: A Comparison of three Techniques*, *Journal of Marketing Research*, 19(3), s. 381-385.
- Davidov, E. (2010), *Testing for Comparability of Human Values Across Countries and Time with the Third Round of the European Social Survey*, *International Journal of Comparative Sociology*, 51(3), s. 171-191.
- Davis, J. A., Jowell, R. (1989), *Measuring National Differences: An Introduction to the International Social Survey Programme (ISSP)*, [w:] *British Social Attitudes: Special International Report*, red. J. Roger, S. Witherspoon, L. Broo, Aldershot: Gower, s. 1-13.
- De Heer, W. (1999), *International Response Trends: Results of an International Survey*, *Journal of Official Statistics*, 15(2), s. 129.
- De Heer, W., De Leeuw, E. (2002), *Trends in Household Survey Nonresponse: A Longitudinal and International Comparison*, *Survey Nonresponse*, 41, s. 41-54.
- De Jonge, T., Veenhoven, R., Kalmijn, W. (2017), *Diversity in Survey Questions on the Same Topic: Techniques for Improving Comparability* (Vol. 68), Amsterdam: Springer.
- De Leeuw, E., Hox, J., Luiten, A. (2018), *International Nonresponse Trends Across Countries and Years: An Analysis of 36 Years of Labour Force Survey Data*, *Survey Methods: Insights From the Field*, s. 1-11.
- De Vries, C. (2017), *Cross-National Data Sources: Opportunities & Challenges*, [w:] *Sage Handbook of Electoral Behavior*, red. K. Arzheimer, J. Evans, M.S. Lewis-Beck, London: Sage, s. 952-971.
- Denk, C., Guterbock, T., Gold, D. (1996), *Modelling Selection of Respondents Within Household in Telephone Surveys*, Annual Meeting of the American Association for Public Opinion Research, Salt Lake City, UT.
- Denk, C., Hall, J. (2000), *Respondent Selection in RDD Surveys: A Randomized Trial of Selection Performance*, Annual Meeting of the American Association for Public Opinion Research, Portland, OR.
- Dever, J. A., Rafferty, A., Valliant, R. (2008), *Internet Surveys: Can Statistical Adjustments Eliminate Coverage Bias?*, *Survey Research Methods*, 2(2), s. 47-60.
- Domański, H. (2006), *Liczba wizyt i czas trwania badań*, *ASK. Research & Methods*, 15, s. 29-49.
- Dorofeev, S., Grant, P. (2006), *Statistics for Real-Life Sample Surveys: Non-Simple-Random Samples and Weighted Data*, Cambridge: Cambridge University Press.
- Dubrow, J. K., Tomescu-Dubrow, I. (2016), *The Rise of Cross-National Survey Data Harmonization in the Social Sciences: Emergence of an Interdisciplinary Methodological Field*, *Quality & Quantity*, 50(4), s. 1449-1467.
- Eckman, S., Koch, A. (2019), *Interviewer Involvement in Sample Selection Shapes the Relationship between Response Rates and Data Quality*, *Public Opinion Quarterly*, 83(2), s. 313-337.



- Egmond, M., Van der Brug, W., Hobolt, S., Franklin, M., Sapir, E. V. (2013), *European Parliament Election Study 2009, Voter Study*, Mannheim: GESIS Data Archive.
- Elliot, D. (1993), *The Use of Substitution in Sampling*, Survey Methodology Bulletin, 33, s. 8-11.
- EQLS (2013), *3rd European Quality of Life Survey. Sampling Report EU27 and non-EU countries*. [http://doc.ukdataservice.ac.uk/doc/7316/mrdoc/pdf/7316\\_3rd\\_eqls\\_sampling\\_report.pdf](http://doc.ukdataservice.ac.uk/doc/7316/mrdoc/pdf/7316_3rd_eqls_sampling_report.pdf) [dostęp: 26.05.2021].
- ESS (1999), *The European Social Survey, ESS: A Research Instrument for the Social Sciences in Europe*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- ESS (2002), *ESS1 – 2002 Documentation Report. Edition 6.2.*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- ESS (2004), *ESS2 – 2004 Documentation Report, Edition 3.7.*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- ESS (2006), *ESS3 – 2006 Documentation Report, Edition 3.7.*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- ESS (2018), *ESS4 – 2008 Documentation Report. Edition 5.5.*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- ESS (2012), *ESS6 – 2012 Documentation Report, Edition 2.4.*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- ESS (2016), *ESS8 – 2016 Documentation Report (Edition 2.1)*, European Social Survey Data Archive/Norwegian Social Science Data Services.
- EVS (2011), *European Values Study 2008: Integrated dataset (EVS 2008)*. [https://dbk.gesis.org/dbksearch/file.asp?file=ZA4800\\_cdb.pdf](https://dbk.gesis.org/dbksearch/file.asp?file=ZA4800_cdb.pdf) [dostęp: 26.05.2021].
- Fitzgerald, R., Jowell, R. (2010), *Measurement Equivalence in Comparative Surveys: The European Social Survey (ESS) – From Design to Implementation and Beyond*, [w:] *Wiley Series In Survey Methodology. Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, red. J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B.-E. Pennell, T. W. Smith, New York: John Wiley & Sons, s. 485-495.
- Fitzgerald, R., Zavala-Rojas, D. (2020), *A Model for Cross-National Questionnaire Design and Pre-testing*, [w:] *Advances in Questionnaire Design, Development, Evaluation and Testing*, red. P. Beatty, D. Collins, L. Kaye, J. L. Padilla, G. Willis, A. Wilmot, New York: Wiley, s. 493-520.
- Fligstein, N., Polyakova, A., Sandholtz, W. (2012), *European Integration, Nationalism and European Identity*, JCMS: Journal of Common Market Studies, 50(1), s. 106-122.
- Fuchs, M., Bossert, D., Stukowski, S. (2013), *Response Rate and Nonresponse bias-Impact of the Number of Contact Attempts on Data Quality in the European Social Survey*, Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 117(1), s. 26-45.
- Gabler, S., Hader, S., Lynn, P. (2008), *Design Effects for Multiple Design Samples*, Survey Methodology, 32(1), s. 115-120.
- Ganzeboom, H. B., Treiman, D. J. (2020), *International Stratification and Mobility File: Conversion Tools*. <http://www.harryganzeboom.nl/ismf/> [dostęp: 26.05.2021].
- Gaziano, C. (2005), *Comparative Analysis of within-household Respondent Selection Techniques*, Public Opinion Quarterly, 69(1), s. 124-157.
- Gelissen, J., Halman, L. (2018), *European Values Survey*, [w:] *The Blackwell Encyclopedia of Sociology*, red. G. Ritzer, C. Rojek, New York: Wiley, s. 1-4.
- Gendall, P. (2011), *International Social Survey Programme Study Monitoring 2009*. [https://dbk.gesis.org/dbksearch/file.asp?file=ZA5400\\_mr.pdf](https://dbk.gesis.org/dbksearch/file.asp?file=ZA5400_mr.pdf) [dostęp: 26.05.2021].

- Granda, P., Wolf, C., Hadorn, R. (2010), *Harmonizing Survey Data*, [w:] *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, red. J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B. E. Pennell, T. W. Smith, New York: Wiley, s. 315-332.
- Grauenhorst, T., Blohm, M., Koch, A. (2016), *Respondent Incentives in a National Face-To-Face Survey: Do they Affect Response Quality?*, *Field Methods*, 28(3), s. 266-283.
- Groves, R. M. (2006), *Nonresponse Rates and Nonresponse Bias in Household Surveys*, *Public Opinion Quarterly*, 70(5), s. 646-675.
- Groves, R. M. (2004), *Survey Errors and Survey Costs*, New York: Wiley.
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., Tourangeau, R. (2011), *Survey Methodology*, New York: Wiley.
- Groves, R., Kahn, R. (1979), *Surveys by Telephone: A National Comparison with Personal Interviews*, New York: Academic Press.
- Groves, R. M., Lyberg, L. (2010), *Total Survey Error: Past, Present, and Future*, *Public Opinion Quarterly*, 74(5), s. 849-879.
- Groves, R. M., Peytcheva, E. (2008), *The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis*, *Public Opinion Quarterly*, 72(2), s. 167-189.
- Gunther, R., Beck, P. A., Magalhães, P. C., Moreno, A. (2015), *Voting in Old and New Democracies*, London: Routledge.
- Häder, S., Gabler, S. (2003), *Sampling and Estimation*, [w:] *Cross-Cultural Survey Methods*, red. J. A. Harkness, F. J. R. van de Vijver, P. P. Mohler, New York: Wiley, s. 117-136.
- Häder, S., Lynn, P. (2007), *How Representative can a Multi-Nation Survey be*, [w:] *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, red. R. Jowell, C. Roberts, R. Fitzgerald, G. Eva, London: Sage, s. 33-53.
- Hagan, D. E., Collier, C. M. (1983), *Must Respondent Selection Procedures for Telephone Surveys be Invasive?*, *Public Opinion Quarterly*, 47(4), s. 547-556.
- Halman, L., Vloet, A. (1994), *Measuring and Comparing Values in 16 Countries of the Western World: Documentation of the European Values Study 1981-1990 in Europe and North America*, Tilburg: Work and Organization Research Centre, Tilburg University.
- Halman, L. C. J. M. (2001), *The European Values Study: A Third Wave. Sourcebook of the 1999/2000 European Values Study Surveys*, Tilburg: Work and Organization Research Centre, Tilburg University.
- Hansen, M. H., Hurwitz, W. N. (1946), *The Problem of Non-Response in Sample Surveys*, *Journal of the American Statistical Association*, 41(236), s. 517-529.
- Harkness, J. (2003), *Questionnaire Translation*, [w:] *Cross-Cultural Survey Methods*, red. J. A. Harkness, F. J. van de Vijver, P. P. Mohler, New York: Wiley, s. 35-55.
- Harkness, J., Schoua-Glusberg, A. (1998), *Questionnaires in Translation*, [w:] *Cross-Cultural Survey Equivalence*, red. J. Harkness, Mannheim: Zentrum für Umfragen, Methoden und Analysen – ZUMA, s. 87-126.
- Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L. E., Mohler, P. P., Pennell, B.-E., Smith, T. W. (2010), *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, New York: Wiley.
- Heath, A., Martin, J., Spreckelsen, T. (2009), *Cross-National Comparability of Survey Attitude Measures*, *International Journal of Public Opinion Research*, 21(3), s. 293-315.
- Hearinga, S. G., O'Muircheartaigh, C. (2010), *Sample Design for Cross-Cultural and Cross-National Survey Programs*, [w:] *Survey Methods In Multinational, Multiregional, and Multicultural Con-*



- texts, red. J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B. E. Pennell, T. W. Smith, New York: Wiley, s. 251-267.
- Heerwegh, D., Loosveldt, G. (2008), *Face-To-Face versus Web Surveying in a High-Internet-Coverage Population: Differences in Response Quality*, *Public Opinion Quarterly*, 72(5), s. 836-846.
- Hess, M. (2017), *Rising Preferred Retirement Age in Europe: Are Europe's Future Pensioners Adapting to Pension System Reforms?*, *Journal of Aging & Social Policy*, 29(3), s. 245-261.
- Hewstone, M. (1991), *Public Opinion and Public Information Campaigns: the Value of the Eurobarometer*, [w:] *Eurobarometer: The Dynamics of European Public Opinion Essays in Honour of Jacques-René Rabier*, red. K. Reif, R. Inglehart, London: Palgrave Macmillan UK, s. 81-100.
- Hobolt, S. B., Vries, C. E. d. (2016), *Public Support for European Integration*, *Annual Review of Political Science*, 19(1), s. 413-432.
- Hoffmeyer-Zlotnik, J. H., Wolf, C. (2003), *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, Amsterdam: Springer.
- Hooghe, M., Marien, S. (2013), *A Comparative Analysis of the Relation Between Political Trust and Forms of Political Participation in Europe*, *European Societies*, 15(1), s. 131-152.
- Höpner, M., Jurczyk, B. (2015), *How the Eurobarometer Blurs the Line between Research and Propaganda*, MPIfG Discussion Paper, 16(6), s. 1-20.
- Horvitz, D. G., Thompson, D. J. (1952), *A Generalization of Sampling without Replacement from a Finite Universe*, *Journal of the American Statistical Association*, 47(260), s. 663-685.
- Inglehart, R., Reif, K. (1991), *Analyzing Trends in West European Opinion: The Role of the Eurobarometer Surveys*, [w:] *Eurobarometer*, red. R. Inglehart, K. Reif, Amsterdam: Springer, s. 1-26.
- ISSP (1996), *Study Description. ISSP 1996 – Role of Government III*, Mannheim: GESIS Data Archive.
- ISSP (2013), *International Social Survey Programme Study Monitoring 2011 Health. Report to the ISSP General Assembly on monitoring work undertaken for the ISSP by the Methodology Committee*, Mannheim: GESIS Data Archive.
- ISSP (2017), *International Social Survey Programme: Workorientations IV – ISSP 2015 [Data sets, questionnaires, code book and other documents]*, Mannheim: GESIS Data Archive.
- ISSP (2019), *International Social Survey Programme (ISSP). Working Principles (Amended May 2019)*, Mannheim: GESIS Data Archive.
- Jabkowski, P. (2011), *Do more Contact-Attempts Reduce Non-Response Bias in Representative Face-To-Face Interviews? Findings from a PAPI Survey with a Low Response Rate*, *ASK. Research & Methods*, 20, s. 27-58.
- Jabkowski, P. (2015), *Reprezentatywność badań reprezentatywnych. Analiza wybranych problemów metodologicznych oraz praktycznych w paradygmacie całkowitego błędu pomiaru*, Poznań: Wydawnictwo Naukowe UAM.
- Jabkowski, P. (2017), *A Meta-analysis of Within-Household Selection Impact on Survey Outcome Rates, Demographic Representation and Sample Quality in the European Social Survey*, *ASK. Research & Methods*, 26, s. 31-60.
- Jabkowski, P. (2018), *Ekwiwalentność pomiaru skali zaufania politycznego w Europejskim Sondażu Społecznym*, *Studia Socjologiczne*, 229(2), s. 65-92.
- Jabkowski, P. (2019), *Zróżnicowanie praktyk sondażowych w międzynarodowych badaniach porównawczych na podstawie archiwizacji 1537 surveyów zrealizowanych w krajach europejskich w latach 1981–2017*, *Przegląd Socjologiczny*, 68(1), s. 9-38.

- Jabkowski, P., Cichocki, P. (2019), *Within-household Selection of Target-Respondents Impairs Demographic Representativeness of Probabilistic Samples: Evidence From Seven Rounds of the European Social Survey*, *Survey Research Methods*, 13(2), s. 167-180.
- Jabkowski, P., Cichocki, P. (2020), *Reflecting Europeanisation: Cumulative Meta-data of Cross-Country Surveys as a Tool for Monitoring European Public Opinion Trends*, <https://osf.io/5xfcg/> [dostęp: 26.05.2021].
- Jabkowski, P., Kołczyńska, M. (2020a), *Sampling and Fieldwork Practices in Europe: Analysis of Methodological Documentation From 1,537 Surveys in Five Cross-National Projects, 1981-2017*, *Methodology. European Journal of Research Methods for the Behavioral Science*, 16(3), s. 186-207.
- Jabkowski, P., Kołczyńska, M. (2020b), *Supplementary Materials to: Sampling and Fieldwork Practices in Europe: Analysis of Methodological Documentation from 1,537 Surveys in Five Cross-National Projects, 1981-2017*. <https://doi.org/10.23668/psycharchives.3460> [dostęp: 26.05.2021].
- Jäckle, A., Roberts, C., Lynn, P. (2010), *Assessing the Effect of Data Collection Mode on Measurement*, *International Statistical Review*, 78(1), s. 3-20.
- Japac, L., Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C., Lane, J., O'Neil, C., Usher, A. (2015), *Big Data in Survey Research: AAPOR Task Force Report*, *Public Opinion Quarterly*, 79(4), s. 839-880.
- Johnson, T. P., Cho, Y. I., Campbell, R. T., Holbrook, A. L. (2006), *Using Community-Level Correlates to Evaluate Nonresponse Effects in a Telephone Survey*, *International Journal of Public Opinion Quarterly*, 70(5), s. 704-719.
- Jowell, R. (1998), *How Comparative is Comparative Research?*, *American Behavioral Scientist*, 42(2), s. 168-177.
- Jowell, R., Kaase, M., Fitzgerald, R., Eva, G. (2007), *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, London: SAGE.
- Kaminska, O., Lynn, P. (2017), *Survey-Based Cross-Country Comparisons where Countries Vary in Sample Design: Issues and Solutions*, *Journal of Official Statistics*, 33(1), s. 123-136.
- Karpiński, Z. (2016), *Zaufanie uogólnione a ostrożność w kontaktach z nieznanymi: Porównanie wyników dwóch sondaży*, *Studia Socjologiczne*, 22(3), s. 199-228.
- Keeter, S., Fisher, K. (1997), *A Comparison of „Last Birthday” and „Youngest Male/Oldest Female” Respondent Selection Procedures*, *Annual Meeting of the American Association for Public Opinion Research*, Norfolk, VA.
- Keeter, S., Miller, C., Kohut, A., Groves, R. M., Presser, S. (2000), *Consequences of Reducing Nonresponse in a National Telephone Survey*, *Public Opinion Quarterly*, 64(2), s. 125-148.
- Kersten, H. M. P., Bethlehem, J. G. (1984), *Exploring and Reducing the Nonresponse Bias by Asking the Basic Question*, *Statistical Journal of the United Nations Economic Commission for Europe*, 2(4), s. 369-380.
- Kish, L. (1949), *A Procedure for Objective Respondent Selection within the Household*, *Journal of the American Statistical Association*, 44 (247), s. 380-387.
- Kish, L. (1965), *Survey Sampling*, New York: Wiley.
- Kish, L. (1979), *Samples and Censuses*, *International Statistical Review*, 47 (2), s. 99-109.
- Kish, L. (1992), *Weighting for Unequal Pi*, *Journal of Official Statistics*, 8(2), s. 183-200.
- Klingemann, H.-D. (2009), *The Comparative Study of Electoral Systems*, Oxford: Oxford University Press.
- Koch, A., Blohm, M. (2006), *Fieldwork Details in the European Social Survey 2002/2003*, *International Workshop on Comparative Survey Design and Implementation (CSDI): DEU*.

- Koch, A., Halbherr, V., Stoop, I. A., Kappelhof, J. W. (2014), *Assessing ESS Sample Quality by Using External and Internal Criteria*, Mannheim: European Social Survey, GESIS.
- Kohler, U. (2007), *Surveys from Inside: An Assessment of Unit Nonresponse Bias with Internal Criteria*, *Survey Research Methods*, 1(2), s. 55-67.
- Kołczyńska, M. (2014), *Representation of Southeast European Countries in International Survey Projects: Assessing Data Quality*, *ASK. Research & Methods*, 23, s. 57-78.
- Kołczyńska, M., Schoene, M. (2018), *Survey Data Harmonization and the Quality of Data Documentation in Cross-National Surveys*, [w:] *Advances in Comparative Survey Methods: Multicultural, Multinational and Multiregional Contexts (3MC)*, red. T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, B. Dorer, New York: Wiley, s. 963-984.
- Konarski, R. (2010), *Modele równań strukturalnych: teoria i praktyka*, Warszawa: Wydawnictwo Naukowe PWN.
- Krenzkle, T., Li, L., Rust, K. (2010), *Evaluating Within Household Selection Rules under a Multi-Stage Design*, *Survey Methodology*, 36(1), s. 111-119.
- Kreuter, F., Kohler, U. (2009), *Analyzing Contact Sequences in Call Record Data: Potential and Limitations of Sequence Indicators for Nonresponse Adjustments in the European Social Survey*, *Journal of Official Statistics*, 25(2), s. 203-226.
- Kreuter, F., Müller, G., Trappmann, M. (2010), *Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data*, *Public Opinion Quarterly*, 74(5), s. 880-906.
- Kreuter, F., Olson, K. (2013), *Paradata for Nonresponse Error Investigation*, [w:] *Improving Surveys with Paradata: Analytic uses of Process Information*, red. F. Kreuter, New York: Wiley, s. 13-42.
- Kropp, K. (2017), *The cases of the European Values Study and the European Social Survey-European Constellations of Social Science Knowledge Production*, *Serendipities: Journal for the Sociology and History of the Social Sciences*, 2(1), s. 50-68.
- Kropp, K. (2018), *The European Social Survey and European Research Policy: Homological Structures and Conjunctural Alliances*, *European Journal of Cultural and Political Sociology*, 5(3), s. 295-319.
- Kruskal, W., Mosteller, F. (1979), *Representative Sampling. III: The Current Statistical Literature*, *International Statistical Review*, 47(3), s. 245-265.
- Krzewińska, A., Grzeszkiewicz-Radulska, K. (2013), *Klasyfikacja sondażowych technik otrzymania materiałów*, *Przegląd Socjologiczny*, 62(1), s. 9-31.
- Lavrakas, P. J. (2008), *Encyclopedia of Survey Research Methods*, London: Sage.
- Lavrakas, P. J., Bauman, S. L., Merkle, D. M. (1993), *The Last-Birthday Selection Method and within-unit COVERAGE PROBLEMS*, *Proceedings of the Survey Section, American Statistical Association*, [http://www.asasrms.org/Proceedings/papers/1993\\_190.pdf](http://www.asasrms.org/Proceedings/papers/1993_190.pdf) [dostęp: 26.05.2021].
- Le, K. T., Brick, J. M., Diop, A., Alemadi, D. (2013), *Within-household Sampling Conditioning on Household Size*, *International Journal of Public Opinion Research*, 25(1), s. 108-118.
- Lee, H. (2012), *How should one find out the Contributions to the Design Effect (Variance) made by each of the Design Components (Stratification, Clustering, Weighting) of a Complex Sample Design*, *Survey Statistician*, 66, s. 16-20.
- Lissowski, G., Haman, J., Jasiński, M. (2008), *Podstawy statystyki dla socjologów*, Warszawa: Scholar.
- Listhaug, O. (1984), *Confidence in Institutions: Findings from the Norwegian Values Study*, *Acta Sociologica*, 27(2), s. 111-122.
- Listhaug, O., Ringdal, K. (2008), *Trust in Political Institutions: The Nordic Countries Compared with Europe*, [w:] *Nordic Social Attitudes in a European Perspective*, red. E. Heikki, T. Fridberg, M. Hjerm, K. Ringdal, Northampton: Edward Elgar, s. 131-151.

- Lockwood, D. (1964), *Social Integration and System Integration*, [w:] *Explorations in Social Change*, red. G. Zollschan, W. Hirsch, London: Routledge, s. 244-257.
- Lohr, S. L. (2008), *Coverage and sMpling*, [w:] *International Handbook of Survey Methodology*, red. E. D. de Leeuw, J. J. Hox, D. A. Dillman, London: Routledge, s. 97-112.
- Loosveldt, G., Beullens, K. (2013), *The Impact of Respondents and Interviewers on Interview Speed in Face-To-Face Interviews*, *Social Science Research*, 42(6), s. 1422-1430.
- Luijckx, R., Halman, L., Sieben, I., Brislinger, E., Quandt, M. (2016), *European Values in Numbers: Trends and Traditions at the Turn of the Century*, Leiden: Brill.
- Lundquist, P. Särndal, C.-E. (2013), *Aspects of Responsive Design with Applications to the Swedish Living Conditions Survey*, *Journal of Official Statistics*, 29(4), s. 557-582.
- Lyberg, L. (2012), *Survey Quality*, *Survey Methodology*, 38(2), s. 107-130.
- Lyberg, L., Japac, L., Tongur, C. (2018), *Prevailing Issues and the Future of Comparative Surveys*, [w:] *Advances in Comparative Survey Methods: Multicultural, Multinational and Multiregional Contexts (3MC)*, red. T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, B. Dorer, New York: Wiley, s. 1053-1082.
- Lynn, P. (2003), *PEDAKSI: Methodology for Collecting Data about Survey Non-Respondents*, *Quality and Quantity*, 37(3), s. 239-261.
- Lynn, P. (2004), *The Use of Substitution in Surveys*, *The Survey Statistician*, 49, s. 14-16.
- Lynn, P., Häder, S., Gabler, S., Laaksonen, S. (2007), *Methods For Achieving Equivalence of Samples in Cross-National Surveys: The European Social Survey Experience*, *Journal of Official Statistics*, 23(1), s. 107-124.
- Lynn, P., Japac, L., Lyberg, L. (2006), *What's so Special about Cross-National Surveys?*, *International Workshop on Comparative Survey Design and Implementation (CSDI)*: DEU.
- Marien, S., Hooghe, M. (2011), *Does Political Trust Matter? An Empirical Investigation into the Relation between Political Trust and Support for Law Compliance*, *European Journal of Political Research*, 50(2), s. 267-291.
- Marker, D. A., Stevens, D. L. (2009), *Sampling and Inference in Environmental Surveys*, [w:] *Handbook of Statistics*, red. C. R. Rao, Amsterdam: Elsevier, s. 487-512.
- Martinotti, G., Stefanizzi, S. (1995), *Europeans and the Nation State*, *Public Opinion and Internationalized Governance*, 2, s. 163-189.
- Matsuo, H., Billiet, J., Loosveldt, G., Berglund, F., Kleven, Ø. (2010), *Measurement and Adjustment of Non-Response Bias Based on Non-Response Surveys: The Case of Belgium and Norway in the European Social Survey Round 3*, *Survey Research Methods*, 4(3), s. 165-178.
- McLaren, L. M. (2012), *The Cultural Divide in Europe: Migration, Multiculturalism, and Political Trust*, *World Politics*, 64(2), s. 199-241.
- Menold, N. (2014), *The Influence of Sampling Method and Interviewers on Sample Realization in the European Social Survey*, *Survey Methodology*, 40(1), s. 105-123.
- Mohler, P. P. (2019), *Why Do We Prefer Second Best Quality Indicators?*, *Comparative Survey Design and Implementation Workshop*.
- Moschner, M. (2014), *Eurobarometer*, [w:] *Encyclopedia of Quality of Life and Well-Being Research*, red. A. C. Michalos, Amsterdam: Springer, s. 2006-2009.
- Mouzelis, N. (1997), *Social and System Integration: Lockwood, Habermas, Giddens*, *Sociology*, 31(1), s. 111-119.
- Muñoz-Conde, M. M., Muñoz-García, J., Pascual-Acosta, A., Pino-Mejías, R. (2018), *Field Substitution and Sequential Sampling Method*, [w:] *The Mathematics of the Uncertain: A Tribute to Pedro Gil*, red. E. Gil, J. Gil, M. Á. Gil, Amsterdam: Springer, s. 329-336.

- Nissen, S. (2014), *The Eurobarometer and the Process of European Integration*, *Quality & Quantity*, 48(2), s. 713-727.
- O'Rourke, D., Blair, J. (1983), *Improving Random Respondent Selection in Telephone Surveys*, *Journal of Marketing Research*, 20(4), s. 428-432.
- Olson, K. (2006), *Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias*, *International Journal of Public Opinion Quarterly*, 70(5), s. 737-758.
- Ortmanns, V., Schneider, S. L. (2016), *Can we Assess Representativeness of Cross-National Surveys Using the Education Variable?*, *Survey Research Methods*, 10(3), s. 189-210.
- Paisley, W. J., Parker, E. B. (1965), *A Computer-Generated Sampling Table for Selecting Respondents within Households*, *The Public Opinion Quarterly*, 29(3), s. 431-436.
- Park, A., Jowell, R. (1997), *Consistencies and Differences in a Cross-National Survey: The International Social Survey Programme (1995)*, London: Social and Community Planning Research.
- Peytchev, A., Baxter, R. K., Carley-Baxter, L. R. (2009), *Not all Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error*, *Public Opinion Quarterly*, 73(4), s. 785-806.
- Peytchev, A., Presser, S., Zhang, M. (2018), *Improving Traditional Nonresponse Bias Adjustments: Combining Statistical Properties with Social Theory*, *Journal of Survey Statistics and Methodology*, 6(4), s. 491-515.
- Peytcheva, E., Groves, R. M. (2009), *Using Variation in Response Rates of Demographic Subgroups as Evidence of Nonresponse Bias in Survey Estimates*, *Journal of Official Statistics*, 25(2), s. 193.
- Pokropek, A. (2012), *Porównania międzynarodowe [w:] Metodologia badań nad stratyfikacją społeczną*, red. H. Domański, Warszawa: Wydawnictwo IFiS PAN, s. 261-283.
- Quaranta, M. (2018), *How Citizens Evaluate Democracy: An Assessment Using the European Social Survey*, *European Political Science Review*, 10(2), s. 191-217.
- Reeskens, T., Hooghe, M. (2008), *Cross-Cultural Measurement Equivalence of Generalized Trust. Evidence from the European Social Survey (2002 and 2004)*, *Social Indicators Research*, 85(3), s. 515-532.
- Revilla, M. A. (2010), *Quality in Unimode and Mixed-Mode Designs: A Multitrait-Multimethod Approach*, *Survey Research Methods*, 4 (3), s. 151-64.
- Rizzo, L., Brick, J. M., Park, I. (2004), *A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys*, *Public Opinion Quarterly*, 68(2), s. 267-274.
- Ruelens, A., Nicaise, I. (2020), *Investigating a Typology of Trust Orientations towards National and European Institutions: A Person-Centered Approach*, *Social Science Research*, 87, s. 1-14.
- Ruiter, S., Van Tubergen, F. (2009), *Religious Attendance in Cross-National Perspective: A Multilevel Analysis of 60 Countries*, *American Journal of Sociology*, 115(3), s. 863-895.
- Sabin, M. C., Godley, S. H. (1987), *Mental Health Citizen Surveys: A Comparison of Two within Household Telephone Sampling Techniques*, *Evaluation and Program Planning*, 10(2), s. 137-141.
- Ságvári, B., Messing, V. (2019), *Central European Societies on the Map of Europe*, *Intersections*, 5(1), s. 4-7.
- Sakshaug, J. W., Antoni, M. (2018), *Evaluating the Utility of Indirectly Linked Federal Administrative Records for Nonresponse Bias Adjustment*, *Journal of Survey Statistics and Methodology*, 7(2), s. 227-249.
- Sakshaug, J. W., Kreuter, F. (2012), *Assessing the Magnitude of Non-Consent Biases in Linked Survey and Administrative Data*, *Survey Research Methods*, 6(2), s. 113-122.
- Salmon, C. T., Nichols, J. S. (1983), *The Next-Birthday Method of Respondent Selection*, *Public Opinion Quarterly*, 47(2), s. 270-276.



- Saris, W. E., Gallhofer, I. N. (2014), *Design, Evaluation, and Analysis of Questionnaires for Survey Research*, New York: Wiley.
- Saris, W. E., Revilla, M. (2016), *Correction for Measurement Errors in Survey Research: Necessary and Possible*, *Social Indicators Research*, 127(3), s. 1005-1020.
- Saris, W. E., Satorra, A., Coenders, G. (2004), *A new Approach to Evaluating the Quality of Measurement Instruments: The Split-Ballot MTMM Design*, *Sociological methodology*, 34(1), s. 311-347.
- Särndal, C.-E. (2011), *Dealing with Survey Nonresponse in Data Collection, in Estimation*, *Journal of Official Statistics*, 27, s. 1-21.
- Sawiński, Z. (2005), *Metody doboru respondentów, [w:] Fieldwork jest sztuką. Jak dobrać respondentów, skłonić do udziału w wywiadzie, rzetelnie i sprawnie zrealizować badanie*, red. P. Sztabiński, Z. Sawiński, F. Sztabiński, Warszawa: Wydawnictwo IFIS PAN.
- Sawiński, Z., Zieliński, M. W. (2015), *Bogdan Cichomski (5 lutego 1941–2 maja 2015)*, *Studia Socjologiczne*, 2(217), s. 5-8.
- Schmitt, H. (2003), *The Eurobarometers: Their Evolution, Obvious Merits, and Ways to add Value to Them*, *European Union Politics*, 4(2), s. 243-251.
- Schmitt, H., Bartolini, S., Van der Brug, W., Van der Eijk, C., Franklin, M., Fuchs, D., Toka, G., Marsh, M., Thomassen, J. (2009), *European Election Study 2004, Design, Data Description and Documentation*. Mannheim: GESIS Data Archive.
- Schmitt, H., Hobolt, S. B., Popa, S. A., Teperoglou, E. (2016), *European Parliament Election Study 2014, Voter Study, First Post-Election Survey*. Mannheim: GESIS Data Archive.
- Schmitt, H., Hobolt, S. B., Van der Brug, W., Popa, S. A. (2019), *European Parliament Election Study 2019, Voter Study*. Mannheim: GESIS Data Archive.
- Schmitt, H., Van der Eijk, C., Pappi, F. U., Scholz, E. (1997), *European Elections Study 1994: Design, Implementation and Results*, Mannheim: GESIS Data Archive.
- Schnaudt, C., Weinhardt, M., Fitzgerald, R., Liebig, S. (2014), *The European Social Survey: Contents, Design, and Research Potential*, *Journal of Contextual Economics: Schmollers Jahrbuch*, 134(4), s. 487-506.
- Schouten, B., Bethlehem, J., Beullens, K., Kleven, Ø., Loosveldt, G., Luiten, A., Rutar, K., Shlomo, N., Skinner, C. (2012), *Evaluating, Comparing, Monitoring, and Improving Representativeness of Survey Response Through R-Indicators and Partial R-Indicators*, *International Statistical Review*, 80(3), s. 382-399.
- Schouten, B., Cobben, F., Bethlehem, J. (2009), *Indicators for the Representativeness of Survey Response*, *Survey Methodology*, 35(1), s. 101-113.
- Signorelli, S. (2012), *The EU and Public Opinions: A Love-Hate Relationship?*, Milan: Notre Europe. Jacques Delors Institute.
- Silber, H., Stark, T. H., Blom, A. G., Krosnick, J. A. (2018), *Implementing a Multinational Study of Questionnaire Design, [w:] Advances in Comparative Survey Methods*, red. T. P. Johnson, B. E. Pennell, I. Stoop, B. Dorer, New York: Wiley, s. 161-179.
- Skjåk, K. K. (2010), *The International Social Survey Programme: Annual Cross-National Social Surveys Since 1985, [w:] Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, red. J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B.-E. Pennell, T. W. Smith, New York: Wiley, s. 497-506.
- Słomczyński, K. M. (2004), *Europejski sondaż społeczny a inne międzynarodowe badania surveyowe: Czego można nauczyć się z analiz porównawczych?*, *ASK. Research & Methods*, 13, s. 9-25.
- Słomczyński, K. M., Powalko, P., Krauze, T. (2017), *Non-unique Records in International Survey Projects: The Need for Extending Data Quality Control*, *Survey Research Methods*, 11(1), s. 1-16.

- Ślomończynski, K. M., Tomescu-Dubrow, I. (2006), *Representation of Post-Communist European Countries in Cross-National Public Opinion Surveys*, *Problems of Post-Communism*, 53(4), s. 42-52.
- Ślomończyński, K. M., Tomescu-Dubrow, I., Jenkins, J. C. (2016), *Democratic Values and Protest Behavior: Harmonization of Data from International Survey Projects*, Warszawa: Wydawnictwo Naukowe IFiS PAN.
- Smith, T. W. (2007), *Survey Non-Response Procedures in Cross-National Perspective: The 2005 ISSP Non-Response Survey*, *Survey Research Methods*, 1(1), s. 45-54.
- Smith, T. W. (2008), *International Social Survey Programme (ISSP)*, [w:] *Encyclopedia of Survey Research Methods*, red. P.J. Lavrakas, London: Sage, s. 353-355.
- Smith, T. W. (2010), *The globalization of survey research*, [w:] *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, red. J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B.-E. Pennell, T. W. Smith, New York: Wiley, s. 475-484.
- Smith, T. W. (2018), *Improving Multinational, Multiregional, and Multicultural (3MC) Comparability Using the Total Survey Error (TSE) Paradigm*, [w:] *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC)*, red. T. P. Johnson, B.-E. Pennell, I. Stoop, B. Dorner, New York: Wiley, s. 13-43.
- Stoop, I. (2004), *Surveying nonrespondents*, *Field Methods*, 16(1), s. 23-54.
- Stoop, I., Billiet, J., Koch, A., Fitzgerald, R. (2010), *Improving Survey Response. Lessons Learned from the European Social Survey*, New York: Wiley.
- Stoop, I., Koch, A., Halbherr, V., Loosveldt, G., Fitzgerald, R. (2016), *Field Procedures in the European Social Survey Round 8: Guidelines for Enhancing Response Rates and Minimising Nonresponse Bias, Technical Report*, London: ESS ERIC Headquarters.
- Sztabiński, P. B., Sztabiński, F., Przybysz, D. (2009), *How Does Length of Fieldwork Period Influence Non-Response? Findings From ESS 2 in Poland*, *ASK. Research and Methods*, 18, s. 67-95.
- Sztabiński, P. B., Sztabiński, F., Przybysz, D. (2012), *What do Respondents and Non-Respondents Think of Incentives and how do They React to Them? The ESS Experience in Poland*, *ASK. Research and Methods*, 21, s. 87-122.
- Tomescu-Dubrow, I., Ślomończynski, K. M. (2014), *Democratic Values and Protest Behavior: Data Harmonization, Measurement Comparability, and Multi-Level Modeling in Cross-National Perspective*, *ASK. Research and Methods*, 23, s. 103-114.
- Tomescu-Dubrow, I., Ślomończynski, K. M. (2016), *Harmonization of Cross-National Survey Projects on Political Behavior: Developing The Analytic Framework of Survey Data Recycling*, *International Journal of Sociology*, 46(1), s. 58-72.
- Troldahl, V. C. Carter Jr, R. E. (1964), *Random Selection of Respondents within Households in Phone Surveys*, *Journal of Marketing Research*, 1(2), s. 71-76.
- Van der Eijk, C., Franklin, M., Schoenbach, K., Schmitt, H., Semetko, H. (1999), *European Election Study – 1999*, Mannheim: GESIS Data Archive.
- Van der Eijk, C., Franklin, M., Schönbach, K., Schmitt, H., Semetko, H. (2002), *European Election Study 1999: Design, Data Description and Documentation*, Mannheim: GESIS Data Archive.
- Van der Eijk, C., Oppenhuis, E., Schmitt, H. (1993), *European Election Study 1989*, Mannheim: GESIS Data Archive.
- Van der Kolk, H., Schmitt, H., Scholz, E., Thomassen, J. (1997), *The European Elections Study 1994*, *European Journal of Political Research*, 32(2), s. 283-289.
- Vannieuwenhuyze, J., Loosveldt, G., Molenberghs, G. (2010), *A Method for Evaluating Mode Effects in Mixed-Mode Surveys*, *Public Opinion Quarterly*, 74(5), s. 1027-1045.

- Vehovar, V. (1999), *Field Substitution and Unit Nonresponse*, Journal of Official Statistics, 15(2), s. 335.
- Vincente, P., Reis, E. (2009), *The Mobile-only Population in Portugal and Its Impact in a Dual Frame Telephone Survey*, Survey Research Methods, 3(2), s. 105-111.
- Von der Lippe, E., Schmich, P., Lange, C. (2011), *Advance Letters as a Way of Reducing Non-Response in a National Health Telephone Survey: Differences between Listed and Unlisted Numbers*, Survey Research Methods, 5(3), s. 103-116.
- Weisberg, H. F. (2005), *The Total Survey Error Approach: A Guide to the New Science Of Survey Research*, Chicago: University of Chicago Press.
- West, B. T. (2011), *Paradata in Survey Research*, Survey Practice, 4(4), s. 1-8.
- Wuyts, C., Loosveldt, G. (2019), *Interview Speed Patterns versus Deviating Behaviors Observed in Audio-Recorded Interviews for (early) Identification Of Field Interviewers' Harmful Interviewing Practice: A Test Case*, ESRA Conference, Zagreb.
- Wysmulek, I. (2018), *Europe of Uneven Data: Country Representation in International Surveys on Corruption, 1989-2017*, ASK. Research & Methods, 27, s. 87-104.
- Yan, T. (2009), *A Meta-Analysis of within-household Respondent Selection Methods*, AAPOR, 17, s. 6134-6146.
- Zavala-Rojas, D. (2012), *Evaluation of the Concepts Trust in Institutions and Trust in authorities*, RECSM Working Paper Number 29, Barcelona: Universitat Pompeu Fabra
- Zmerli, S., Van der Meer, T. W. (2017), *Handbook on Political Trust*, Northampton: Edward Elgar.





## Summary

# Reflecting Europeanisation

## Monitoring trends in European public opinion in cross-national comparative surveys as a methodological challenge

This book considers the methodological challenges of studying Europeans' attitudes towards the European Union and European integration process, measured over 20 years within four prominent pan-European multi-wave transnational comparative surveys, i.e., Eurobarometer, the European Social Survey, the European Values Studies and the International Social Survey Program. The primary goal is to present the main challenges stemming from the diversity of research issues and the different objectives of individual projects, as well as those stemming from the specific ways in which questions are formulated, the different measurement methodologies and survey procedures, and the variable quality of the surveys. The first chapter presents the abundance of survey projects carried out on the European continent since the mid-1970s. The second chapter analyses the variation in the wording of questions about European integration and the impact of this wording on the survey results. The third chapter presents the variation in sampling and fieldwork procedures implemented in four projects. The final chapter analyses the quality of over 1,400 national surveys, which allowed the impact of different fieldwork procedures on survey process quality to be assessed. The conclusion recommends taking into account project-specific field sample implementation in the analysis of data from transnational comparative surveys.



# WYDAWNICTWO NAUKOWE

---

## WYDAWNICTWO NAUKOWE

UNIwersytetu im. Adama Mickiewicza

W POZNANIU

ul. A. Fredry 10, 61-701 Poznań

tel. 61 829 46 46

e-mail: [wydnauk@amu.edu.pl](mailto:wydnauk@amu.edu.pl)

[www.press.amu.edu.pl](http://www.press.amu.edu.pl)