

Ilu jest w Polsce cudzoziemców? Ujęcie statystyczne i ekonomiczne

Dr Maciej Beręsewicz, prof. UEP

1 Wprowadzenie

Jednym z kluczowych zagadnień w statystyce publicznej stanowiących istotne wyzwanie metodologiczne jest estymacja wielkości populacji. Informacja o liczbie osób mieszkających w danym kraju lub regionie, czy też stanowiących populację określoną w inny sposób aniżeli przestrzennie, ma bowiem istotny wpływ na wiele aspektów życia, takich jak podział funduszy europejskich, środków na ochronę zdrowia czy kalkulację PKB. Jednocześnie strategia i wizja rozwoju spisów powszechnych formułowana przez Eurostat kładzie główny nacisk na ideę odejścia od tradycyjnego spisu organizowanego w odstępach dziesięcioletnich na rzecz kroczącego spisu wirtualnego bazującego na źródłach administracyjnych uzupełnionych o badania reprezentacyjne. Powyższe działania są wynikiem funkcjonowania w statystyce publicznej nowego paradygmatu, w którym kładzie się nacisk na ponowne wykorzystanie wszystkich dostępnych źródeł danych zamiast tworzenia nowych.

Wykorzystanie „niestatystycznych” źródeł danych prowadzi jednak do powstania problemów natury metodologicznej}. Jako przykład można podać zrealizowaną przez GUS pracę eksperymentalną „Populacja cudzoziemców w Polsce w czasie COVID-19”, w której w oparciu o dziewięć rejestrów administracyjnych oszacowano, że w dniu 31.12.2019 w Polsce przebywało 2,1 mln. obcokrajowców. Można zadać pytanie, jak bliska prawdy jest ta wartość? Jaki wpływ na ten szacunek mają błędy nielosowe i w jaki sposób możemy go ewentualnie skorygować? Ważna jest również kwestia precyzji uzyskanych szacunków i sposobu jej wyznaczania. Z ekonomicznego punktu widzenia kluczowa jest bowiem wiarygodność szacunków wskaźników, takich jak stopa bezrobocia czy PKP, służy bowiem ona do oceny tego, w jakim stopniu poszerzają one wiedzę z zakresu sytuacji gospodarczej. Z tej perspektywy istotne jest zatem określenie, na ile informacje zawarte w rejestrach administracyjnych są wystarczające jako źródło wiedzy. Uzyskanie odpowiedzi na wyżej postawione pytania z jednej strony wypełni lukę metodologiczną, z drugiej zaś strony służyć będzie wypracowaniu metodyki estymacji charakterystyk populacji cudzoziemców w Polsce, która pozwoli poszerzyć wiedzę na temat tej populacji.

Kolejnym wyzwaniem jest określenie wielkości populacji cudzoziemców uchodźców w przypadku konfliktów zbrojnych czy katastrof naturalnych. W wyniku napaści Rosji na Ukrainę Polską granicę przekroczyło ponad ... z czego ponad ... skorzystało ze specjalnego statusu UKR, który m.in. umożliwia uzyskanie numeru PESEL, legalny pobyt, ułatwienia w kontaktach z urzędami oraz instytucjami państwowymi czy wykonywanie pracy. Status UKR zostaje odebrany w momencie przebywania poza Polską ponad miesiąc.

Celem artykułu jest przedstawienie problemów związanych z określeniem liczby cudzoziemców w Polsce w świetle istniejących źródeł administracyjnych oraz wskazanie możliwych rozwiązań

2 Problematyka pomiaru wielkości populacji

2.1 Populacja trudna do zbadania

Bereświcz, Gudaszewski, i Szymkowiak (2019) rozróżniają następujące populacje trudne do zbadania:

- których jednostki trudno wylosować do próby (ang. *hard-to-sample*) – w przypadku populacji trudnych do zbadania bardzo rzadko zdarza się, aby istniał właściwy operat losowania, z którego jednostki można byłoby wylosować do próby wykorzystując odpowiedni schemat jej pobierania. Przykładem tego typu populacji mogą być osoby bezdomne lub pracujący cudzoziemcy;
- których jednostki trudno jest zidentyfikować (ang. *hard-to-identify*) – w niektórych przypadkach, szczególnie w odniesieniu do stygmatyzowanych grup społecznych, członkowie populacji mogą nie chcieć udostępnić swoich cech, co wiązać się może z lękiem przed ujawnieniem nielegalnego lub wstydliwego statusu społecznego. Przykład tego typu populacji stanowią mogą narkomani, alkoholicy czy różne mniejszości (na przykład osoby LGBT, wyznawcy określonych religii czy ideologii);
- których jednostki trudno znaleźć i nawiązać z nimi kontakt (ang. *hard-to-find-and-contact*) – trudność w nawiązaniu kontaktu związana jest przede wszystkim z mobilnością tego typu populacji. Przykładem tego typu populacji mogą być niezameldowani cudzoziemcy, członkowie koczowniczych kultur (Beduini z południowo-zachodniej Azji czy Tuaregowie z Afryki Północnej), mniejszości wędrownie (Romowie w Europie), osoby bezdomne;
- której jednostki trudno namówić do wzięcia udziału w badaniu (ang. *hard-to-persuade*) – niechęć do wzięcia udziału w badaniu związana może być z drażliwością poruszanej tematyki bądź z brakiem czasu. Przykładem tego typu populacji mogą być aktywni zawodowo, pracujący w szarej czy czarnej strefie, cudzoziemcy;
- której jednostki można zachęcić do wzięcia udziału w badaniu, ale ciężko przeprowadzić wywiad (ang. *hard-to-interview*) – trudność w przeprowadzeniu wywiadu może wynikać

z tego, że należy uzyskać na udział w badaniu danej jednostki zgody przełożonego, opiekuna prawnego czy rodzica. Przykład tego typu populacji stanowić mogą więźniowie, osoby niepełnosprawne psychicznie czy cudzoziemcy nie znający języka danego kraju.

Tak jest w przypadku populacji cudzoziemców w Polsce, której jednostki trudno wylosować do próby (brak pełnego operatu losowania oraz kompleksowych źródeł danych statystycznych, z których można czerpać wiedzę na temat cudzoziemców), z którą trudno nawiązać kontakt (mobilność cudzoziemców na rynku pracy oraz brak stałego miejsca zamieszkania) czy też przeszkodą może być bariera językowa.

2.2 Źródła danych

2.3 Estymacja wielkości populacji na podstawie źródeł administracyjnych

Jednakże, aby całkowicie zastąpić spis rejestrami administracyjnymi oraz metodą CR należy wziąć pod uwagę następujące założenia:

- *populacja jest zamknięta* (brak zgonów/urodzeń) – rejestr wg stanu na określony dzień,
- *źródła muszą być niezależne* – uwzględnienie rejestrów policji, uwzględnienie co najmniej 3 źródeł i/lub estymatorów opracowanych dla źródeł zależnych,
- *homogeniczność przynależności do rejestrów* – wykorzystujemy zmienne społeczno-demograficzne i zakładamy homogeniczność w tych grupach,
- *brak błędów łączenia* – łączymy rejestry po kluczu, ewentualnie uwzględniamy błąd wynikający z probabilistycznego łączenia rekordów,
- *brak błędów pokrycia* – wykorzystujemy rejestry aktualizowane w ujęciu miesięcznym/kwartalnym.

Aby w sposób nieobciążony oszacować wielkość populacji, przy powyższych założeniach, należy skorzystać z następującego estymatora Lincolna-Petersena:

$$\hat{N} = \frac{n_1 n_2}{m},$$

gdzie n_1 będzie wielkością próby wylosowaną za pierwszym razem, n_2 będzie wielkością próby wylosowaną za drugim razem, a m będzie częścią wspólną obydwu prób. Wariancja tego estymatora dana jest wzorem

$$\hat{V}(\hat{N}) \approx \frac{n_1^2 n_2 (n_2 - m)}{m^3}.$$

Przykładowo, Gerritse (2016) zastosowała metodę *capture-recapture* do oszacowania liczby obywateli Polski przebywających w Holandii w 2009 roku. W tym celu wykorzystwała dwa źródła

danych: Rejestr Populacji (Gemeentelijke BasisAdministratie; GBA) oraz rejestr Policji (Her-Kenningsdienst Systeem; HKS) uzyskując następującą tablicę:

Tablica 1: Liczba obywateli Polski w dwóch rejestrach w Holandii w 2009r.

		HKS	
		Tak	Nie
GBA	TAK	374	39 488
	NIE	1 445	?

Źródło: Opracowanie własne na podstawie...

Na podstawie tablicy 1 Gerritse (2016) wielkość populacji obywateli Polski Holandii na blisko 193 tys. Według NSP 2011 emigrantów przebywających czasowo w Holandii było 106 tys.

Niestety, aktualnie w Polsce niemożliwe jest dokonanie podobnego szacunku na podstawie danych pochodzących z Policji (dokładnie Krajowego Systemu Inforacyjnego Policji; KSIP) ponieważ regulacje prawne uniemożliwiają przekazanie identyfikowalnych danych z KSIP statystyce publicznej. W związku z tym jedyną możliwością zastosowania metod capture-recapture jest wykorzystanie wszelkich dostępnych źródeł administracyjnych.

... ZHANG I Dunne zaproponowali inne podejście oparte na zintegrowaniu rejestrów ... a jako źródło niezależne wykorzystali dane z ...

3 Populacja cudzoziemców w Polsce w świetle dostępnych danych

3.1 Narodowy Spis Powszechny Ludności i Mieszkań 2021

Tablica 2 zawiera wyniki Narodowego Spisu Powszechnego Ludności i Mieszkań 2021 (NSP 2021) w zakresie liczby cudzoziemców według obywatelstwa oraz długości pobytu. Kolumna *stali* odnosi się do tzw. ludności wg definicji krajowej¹, natomiast kolumna *stali i czasowi* odnosi się do ludności bez względu na okres przebywania. Należy zaznaczyć, że dane zawarte w tablicy 2 są oznaczone przez GUS jako wstępne i mogące ulec zmianie. Zgodnie z tymi szacunkami liczba cudzoziemców w Polsce wynosiła blisko 1,5 mln, z czego większość stanowili obywatele krajów Europy (bez Unii Europejskiej). Niestety, w momencie przygotowywania artykułu GUS nie udostępnił szczegółowych informacji o obywatelstwie. Jednak można domniemywać, że większość z grupy 1,1 mln cudzoziemców stanowią obywatele i obywatelki Ukrainy. We wszystkich przypadkach wymienionych w tablicy 2 relacja liczby cudzoziemców przebywających czasowo do stałych mieszkańców była wyższa niż 5, a w przypadku *pozostałych krajów Europy* wynosiła ponad 12.

¹Ludność według definicji krajowej – stali mieszkańcy Polski - w tym osoby, które przebywają czasowo za granicą (bez względu na okres przebywania), ale zachowały stałe zameldowanie w Polsce. Do ludności nie są natomiast zaliczani imigranci przebywający w Polsce czasowo.

Tablica 2: Cudzoziemcy wg grup krajów obywatelstwa oraz długości pobytu według NSP 2021 (w tys., stan na 31.03.2021)

obywatelstwo	stali	stali i czasowi	relacja
ogółem	111,8	1 459,2	12,1
Pozostałe kraje Europy	80,7	1 132,2	13,0
Azja	10,6	145,6	12,7
Kraje Unii Europejskiej	15,5	140,2	8,0
Afryka	2,8	18,5	5,6
Ameryka Północna i Środkowa	1,5	14,7	8,8
Ameryka Południowa	0,6	5,7	8,5
Australia i Oceania	0,1	1,2	11,0
Bezpaństwowcy	0,1	1,0	9,0
Obywatelstwo obce - kraj nieustalony	0,0	0,0	0,0

Źródło: Opracowanie własne na podstawie danych NSP 2021 (dane pobrane zgodnie ze stanem na 23.02.2023).

Dane z NSP 2021 mogą nam ukazać pewne zależności, które występowały w marcu 2021 ale po rozpoczęciu agresji Rosji na Ukrainę 24 lutego 2022 i fali migracji uchodźców, liczby te zdecydowanie różnią się od bieżącej sytuacji. Czy w takim razie można wykorzystać rejestry administracyjne do aktualizacji szacunków z NSP 2021?

3.2 Rejestry administracyjne

W pracy eksperymentalnej pt. zespół GUS ... podjął problem wyznaczenia populacji cudzoziemców na podstawie zintegrowanych rejestrów administracyjnych. Integracji danych dokonano z wykorzystaniem identyfikatorów tj. numerów PESEL.

Tablica 3: Liczba cudzoziemców według występowania w źródłowych rejestrach administracyjnych według stanu na 31.12.2019 r.

Rejestr źródłowy	Liczba cudzoziemców
PESEL	263 976
UdSC	426 266
ZUS	696 434
MF	919 491
MRPiPS (MRiPS)	767 033
MNiSW (MEiN)	66 824
MEN (MEiN)	18 032
KRUS	60 198

Rejestr źródłowy	Liczba cudzoziemców
ZUS	1 230 717

Źródło: Opracowanie własne na podstawie...

Występowanie cudzoziemców w rozważanych rejestrach przedstawiono w tablicy 4.

Tablica 4: Liczba cudzoziemców według liczby rejestrów

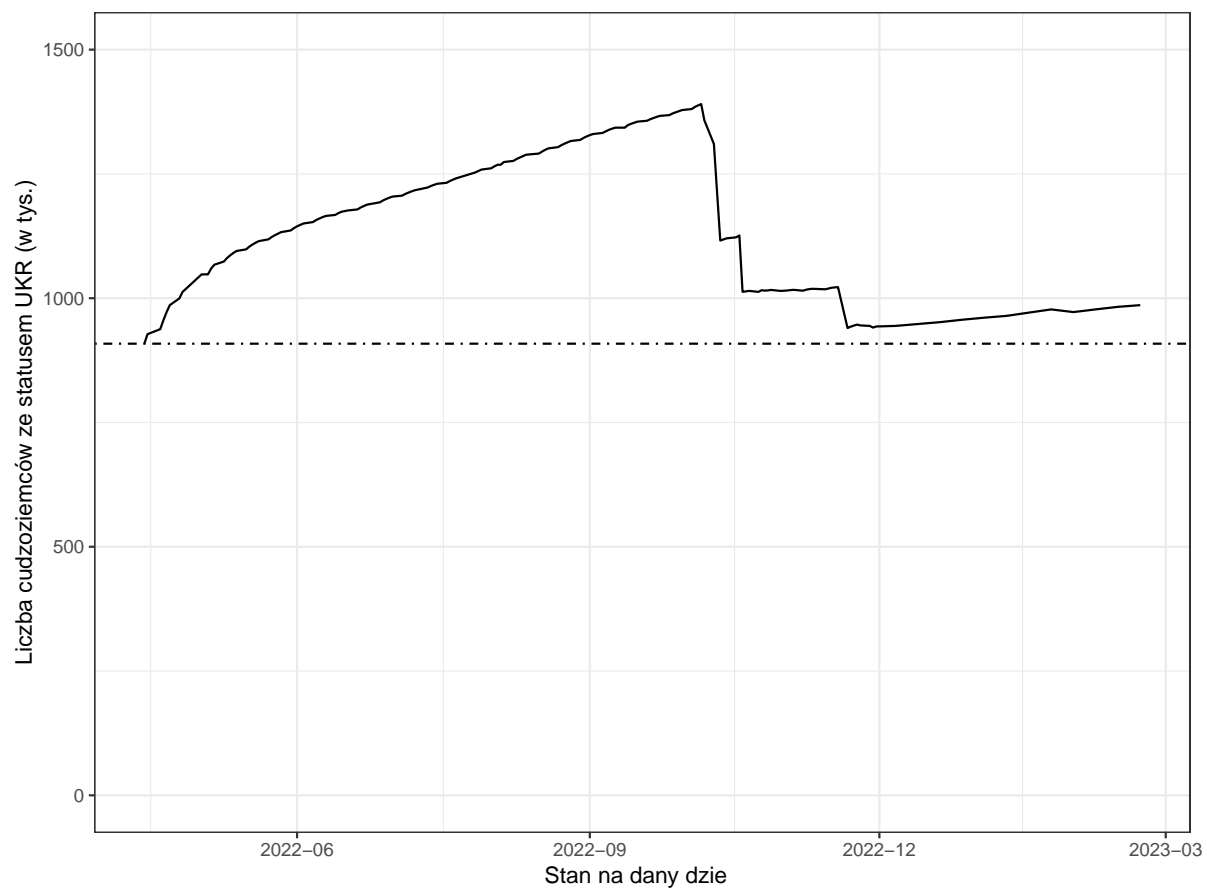
Współwystępowanie	Liczba cudzoziemców
W 1 zbiorze	988 235
W 2 zbiorach	461 233
W 3 zbiorach	303 767
W 4 zbiorach	222 911
W 5 zbiorach	111 896
W 6 zbiorach	17 921
W 7 zbiorach	138
W 8 zbiorach	0
W 9 zbiorach	0

Źródło: Opracowanie własne na podstawie...

Dane z NSP mogłyby zostać zaktualizowane informacją o liczbie cudzoziemców, którzy uzyskali status UKR. Rysunek 1 przedstawia liczbę cudzoziemców posiadających status UKR (w tys.) w okresie od 14 kwietnia 2022 do 21 lutego 2023. Początkowa liczba wynosiła ponad 908,5 tys. a końcowa 986 tys. W szczytowym momencie (6 października 2020) liczba osób ze statusem UKR wynosiła blisko 1,4 mln. Gwałtowny spadek związany jest m.in. wyjazdami z rozpoczęciem roku szkolnego oraz innymi przyczynami, a opóźnienie w raportowaniu wskazuje na znaczną zmianę w październiku oraz listopadzie. Widać także lekki wzrost liczby osób ze statusem UKR ale dostępne dane nie pozwalają określić czy są to osoby nowe czy uzyskujące status UKR poraz kolejny.

Źródło: Opracowanie własne na podstawie danych o statusie UKR udostępnionych w serwisie dane.gov.pl (<https://dane.gov.pl/pl/dataset/2715,zarejestrowane-wnioski-o-nadanie-statusu-ukr>). Linia przerywana oznacza początkową liczbę cudzoziemców wynoszącą 908,5 tys.

Jeżeli rejestry administracyjne



Rysunek 1: Liczba obywateli Ukrainy ze statusem UKR (w tys.) w okresie od 14 kwietnia 2022 do 21 lutego 2023

3.3 Źródła big data

4 Podsumowanie

Akualnie

5 Podziękowania

Autor dziękuje prof. dr hab. Markowi Świtońskiemu za zaproszenie do wygłoszenia referatu pt. „Ilu jest w Polsce cudzoziemców? Ujęcie statystyczne i ekonomiczne” w ramach Dwugłosu Nauki „Migracje – historia i teraźniejszość” 21.11.2022.

Badania autora są realizowane dzięki wsparciu Narodowego Centrum Nauki (grant NCN OPUS 20 nr 2020/39/B/HS4/00931 pt. „Statystyka cudzoziemców bez spisu powszechnego – jakość, integracja danych i estymacja”).

6 Literatura

Beręsewicz, Maciej, Grzegorz Gudaszewski, i Marcin Szymkowiak. 2019. „Estymacja liczby cudzoziemców w Polsce z wykorzystaniem metody capture-recapture”. *Wiadomości Statystyczne. The Polish Statistician* 64 (10): 7–35. <https://doi.org/10.5604/01.3001.0013.7589>.