



**CEFMR Working Paper
2/2007**

***STUDIUM METODOLOGICZNE OSZACOWANIA
RZECZYWISTEJ LICZBY LUDNOŚCI WARSZAWY***

Jakub Bijak, Anna Kicingier i Marek Kupiszewski

Współpraca: Przemysław Śleszyński

ul. Twarda 51/55, 00-818 Warsaw, Poland
tel. +48 22 697 88 34, fax +48 22 697 88 43
e-mail: cefmr@cefmr.pan.pl
Internet: www.cefmr.pan.pl

Central European Forum for Migration and Population Research (CEFMR) is a research partnership of the Foundation for Population, Migration and Environment, Institute of Geography and Spatial Organization of the Polish Academy of Sciences and IOM - International Organization for Migration



International Organization
For Migration



Foundation for Population,
Migration and Environment



Institute of Geography and Spatial Organisation,
Polish Academy of Sciences

STUDIUM METODOLOGICZNE OSZACOWANIA RZECZYWISTEJ LICZBY LUDNOŚCI WARSZAWY

Jakub Bijak, Anna Kicinger* i Marek Kupiszewski**

Współpraca: Przemysław Śleszyński[†]

* Środkowoeuropejskie Forum Badań Migracyjnych i Ludnościowych w Warszawie
Central European Forum for Migration and Population Research in Warsaw

[†] Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa
Institute of Geography and Spatial Organization, Polish Academy of Sciences, Warsaw

Streszczenie: Niniejszy raport, będący metodologicznym przyczynkiem do zagadnienia szacowania rzeczywistej liczby ludności Warszawy, przedstawia analizę możliwych do wykorzystania metod badawczych oraz standardowych i niestandardowych źródeł danych. Wstępne informacje o rzędzie wielkości spodziewanych korekt oficjalnych szacunków pozwalają oczekiwać dodatkowo około ćwierci miliona osób zamieszkałych w Warszawie oraz, oprócz tego, około pół miliona osób przyjeżdżających codziennie do miasta.

Słowa kluczowe: Liczba ludności, Szacunki demograficzne, m. st. Warszawa

FEASIBILITY STUDY FOR THE ESTIMATION OF THE ACTUAL POPULATION SIZE OF THE CITY OF WARSAW

Abstract: The report, being a methodological feasibility study for the task of estimation of the actual population size of the City of Warsaw, presents an overview of the possibly applicable research methods, as well as standard and non-standard data sources. Preliminary assessments of the order of magnitude of the expected corrections of the official figures indicate around a quarter of a million residents more than reported in the statistics, and on the top of that, about half a million daily commuters.

Keywords: Population size, Demographic estimates, City of Warsaw

Opinie przedstawione w opracowaniu stanowią prywatne przekonania autorów i nie muszą odzwierciedlać poglądów Urzędu m. st. Warszawy, IOM – Międzynarodowej Organizacji do spraw Migracji, Instytutu Geografii i Przestrzennego Zagospodarowania Polskiej Akademii Nauk, ani Fundacji na rzecz Ludności, Migracji i Środowiska.

The views presented in this paper are those of the authors and do not necessarily reflect those of the Warsaw Municipal Government Office, IOM – International Organization for Migration, the Institute of Geography and Spatial Organization of the Polish Academy of Sciences, or the Foundation for Population, Migration and Environment.

Redakcja / Editor

ul. Twarda 51/55, 00-818 Warsaw, Poland
tel. +48 22 697 88 34, fax +48 22 697 88 43
e-mail: cefmr@cefmr.pan.pl
Internet: www.cefmr.pan.pl

© Copyright by the Warsaw Municipal Government Office. Reprinted by a kind permission.
Warsaw, December 2007

ISSN 1732-0631
ISBN 978-83-60462-07-2

Spis treści

Podziękowania / Acknowledgements	2
Streszczenie.....	3
Executive Summary	5
1. Wprowadzenie.....	7
2. Rodzaje informacji statystycznej o ludności Warszawy	10
2.1. Ewidencja ludności a dostępność i kompletność informacji statystycznej	10
2.2. Definicje stanów ludności i ich użyteczność dla szacowania liczby mieszkańców stolicy.....	13
2.2.1. Ludność stała (stałe zamieszkała).....	13
2.2.2. Ludność faktycznie zamieszkała	13
2.2.3. Rezydenci (ludność rezydująca)	15
2.3. Ludność nocna a ludność dzienna.....	17
2.4. Dezagregacja liczby ludności według różnych cech	18
3. Propozycje wykorzystujące standardowe źródła danych i metody badawcze.....	23
3.1. Sugestie metodyczne dla szacowania liczebności wybranych grup.....	23
3.1.1. Uwagi ogólne.....	23
3.1.2. Studenci	24
3.1.3. Ludność aktywna zawodowo.....	25
3.1.4. Bezdomni.....	28
3.1.5. Podsumowanie i dyskusja.....	29
3.2. Korekty oparte na wynikach reprezentacyjnych badań społecznych.....	33
3.2.1. Badanie ankietowe <i>face-to-face</i> w terenie – zarys metody	33
3.2.2. Badanie CATI – zarys metody.	36
4. Możliwości szacunków opartych na niestandardowej informacji statystycznej.....	38
4.1. Dostępność danych statystycznych w wybranych instytucjach.....	38
4.1.1. Ministerstwo Finansów – dane z rejestrów podatkowych	38
4.1.2. Statystyka ubezpieczonych w Narodowym Funduszu Zdrowia	39
4.1.3. Rejestry Państwowej Komisji Wyborczej (PKW)	41
4.1.4. Korekty oparte na szacunkach Biura Administracji i Spraw Obywatelskich m. st. Warszawy	43
4.1.5. Statystyka kościelna	44
4.2. Obliczenia na podstawie danych o gospodarce komunalnej w Warszawie.....	46
4.2.1. Miejskie Przedsiębiorstwo Wodociągów i Kanalizacji – zużycie wody i produkcja ścieków.....	46
4.2.2. Zarząd Oczyszczania Miasta – gospodarka odpadami	47
4.3. Szacunki oparte na informacji ze zdjęć satelitarnych i lotniczych	48
5. Metody szacowania liczby ludności dziennej.....	54
5.1. Informacje o codziennych dojazdach do Warszawy.....	54
5.1.1. Badania natężenia ruchu drogowego	54
5.1.2. Dojazdy do Warszawy z wykorzystaniem transportu publicznego	61
5.2. Metoda stosowana przez US Bureau of Census	64
6. Podsumowanie i wnioski.....	66
6.1. Podsumowanie wyników studium	66
6.2. Wnioski i rekomendacje dla Urzędu m. st. Warszawy.....	69
Bibliografia	72
Źródła informacji bezpośredniej	75

Podziękowania / Acknowledgements

Niniejsze opracowanie sfinansowane zostało przez Urząd Miasta Stołecznego Warszawy (kontrakt nr AM/001/1.1/B/275/06/2162). Jakub Bijak był również beneficjentem krajowego stypendium dla młodych naukowców Fundacji na rzecz Nauki Polskiej za rok 2006. Powstanie tego raportu nie byłoby możliwe bez nieocenionego wsparcia ze strony pracowników Wydziału Strategii Biura Naczelnego Architekta m. st. Warszawy, kierowanego przez Panią Naczelnik Jadwigę Mijkowską. Autorzy kierują także wyrazy wdzięczności do wszystkich przedstawicieli instytucji, z którymi kontaktowali się w trakcie prac nad niniejszym raportem. Podziękowania należą się również Izabeli Koryś, Weronice Kloc-Nowak oraz Dorocie Kupiszewskiej ze Środkowoeuropejskiego Forum Badań Migracyjnych i Ludnościowych za pomoc przy wybranych fragmentach pracy. Jesteśmy również wdzięczni Prof. Markowi Okólskiemu, Prof. Janowi Paradyszowi i Prof. Zbigniewowi Strzeleckiemu za ich wnikliwe recenzje i cenne uwagi. Wszystkie pozostałe błędy obciążają autorów.

* * *

The study has been financed by the Warsaw Municipal Government Office (contract no. AM/001/1.1/B/275/06/2162). Jakub Bijak was also a beneficiary of a stipend for young scientists of the Foundation for Polish Science (FNP) for 2006. Completion of this report would not be possible without the invaluable support of the staff of the Strategy Department of the Bureau of the National Architect of the Capital City of Warsaw, led by Ms. Jadwiga Mijkowska. The authors wish to thank the representatives of all institutions contacted during the work on the current study. We also owe credits to Izabela Koryś, Weronika Kloc-Nowak and Dorota Kupiszewska from the Central European Forum for Migration and Population Research for their assistance in the selected fragments of the study. The authors are also grateful to Prof. Marek Okólski, Prof. Jan Paradysz and Prof. Zbigniew Strzelecki for their in-depth reviews and helpful remarks. All the remaining errors are ours.

Streszczenie

Od pewnego czasu poddaje się w wątpliwość oficjalną liczbę mieszkańców Warszawy szacowaną na około 1.7 miliona – przyjmując że jest ona zaniżona i nie odpowiada stanowi faktycznemu. Z drugiej strony, wiarygodna informacja o liczbie i strukturze ludności Warszawy stanowi kluczowy element konieczny dla prowadzenia racjonalnej polityki gospodarczej, społecznej oraz planowania przestrzennego w stolicy. Oficjalna informacja statystyczna, oparta na ewidencji ludności oraz zameldowaniach na pobyt stały i czasowy, jest jednak niekompletna – nie uwzględnia na przykład osób mieszkających w Warszawie bez meldunku. Ponadto, wiele osób dojeżdża każdego dnia do stolicy w związku z pracą lub nauką, znacząco zwiększając tak zwaną ludność dzienną miasta. Z powyższych względów, przed przystąpieniem do oszacowania rzeczywistej liczby mieszkańców stolicy dla różnych kategorii ludności, Biuro Naczelnego Architekta Urzędu m. st. Warszawy zleciło dokonanie inwentaryzacji zasobów danych mogących posłużyć takiemu oszacowaniu oraz opracowanie jego metodologii.

Niniejszy raport przedstawia analizę możliwych do wykorzystania metod badawczych oraz standardowych i niestandardowych źródeł informacji statystycznej. Opracowanie oferuje przegląd: danych gromadzonych w instytucjach publicznych; możliwości pozyskania danych z innych źródeł; oraz związanej z poruszaną problematyką literatury naukowej, ilustrującej przykłady wykorzystania różnych metod badawczych (na przykład zastosowania zdjęć lotniczych i satelitarnych).

Na podstawie przeanalizowanych informacji, udało się wstępnie zidentyfikować kilka obiecujących źródeł wiedzy o faktycznej liczbie ludności Warszawy. Po pierwsze, istnieją dane i oszacowania dotyczące liczby ludności w najbardziej mobilnych grupach wieku – studentów i osób pracujących. Porównując te wielkości z danymi GUS o studiujących i pracujących mieszkańcach stolicy oraz uwzględniając wiedzę o części wspólnej obu tych kategorii, można pokusić się o skorygowanie szacunków ludności dziennej.

Alternatywnie, drugą możliwością jest przeprowadzenie reprezentacyjnych badań sondażowych mających na celu zgromadzenie wiedzy o skali zjawiska jakim jest zamieszkiwanie w Warszawie bez formalnego zameldowania. Dodatkowo, oszacowania te mogłyby zostać podparte danymi o zasobach mieszkaniowych, pochodzącymi z oficjalnych rejestrów lub ze zdjęć lotniczych i satelitarnych. Otrzymane wielkości oszacowań można następnie zweryfikować w oparciu o niestandardowe źródła informacji, jak choćby dane

Narodowego Funduszu Zdrowia o osobach ubezpieczonych, lub dane Ministerstwa Finansów dotyczące podatników zamieszkałych w stolicy.

Do wyodrębnienia kategorii ludności dziennej posłużyć mogą informacje z Warszawskiego Badania Ruchu (WBR), przeprowadzonego w 2005 roku przez Biuro Planowania Rozwoju Warszawy SA, oraz z opracowanego na ich podstawie modelu ruchu, będącego w dyspozycji Biura Drogownictwa i Komunikacji Urzędu m. st. Warszawy.

Wszystkie proponowane metody szacunku w nieunikniony sposób obarczone będą pewnym marginesem błędu, charakterystycznym dla poszczególnych metod i źródeł danych. Mimo to, zgromadzone informacje, choć często fragmentaryczne, powinny umożliwić, przynajmniej w sposób przybliżony oszacowanie rzeczywistej liczby mieszkańców stolicy, jak również liczby osób przebywających dodatkowo w Warszawie każdego dnia. Bieżący raport dostarcza pewnych wstępnych informacji o rzędzie wielkości spodziewanych korekt oficjalnych szacunków: oprócz osób uwidocznionych w statystykach, można spodziewać się dodatkowo około ćwierć miliona osób zamieszkałych w Warszawie oraz, oprócz tego, około pół miliona osób przyjeżdżających codziennie do miasta. Dokładniejsze wartości mogą zostać uzyskane jedynie na podstawie wyników osobnego badania, poświęconego szacowaniu liczby ludności m. st. Warszawy.

Executive Summary

The official number of inhabitants of Warsaw, amounting to around 1.7 million people, has been recently questioned as underestimated and not corresponding to reality. On the other hand, credible information about the size and structure of the population of Warsaw constitutes a crucial element necessary for a rational policy of the city with respect to all areas, such as economy, social life, and spatial organisation alike. However, official statistical information, based on the number of persons officially registered for permanent or temporary residence in Warsaw is incomplete, as it does not include persons living in the city without registration. Moreover, many people commute to Warsaw every day, in order to work, to study, or for many other reasons, considerably increasing the so-called daytime population of the city. Therefore, before making an attempt to assess the number of different categories of population, i.e. people actually living or present in Warsaw, the Municipal Office of the City of Warsaw commissioned the preparation of an inventory of data potentially useful for the estimation, and proposing the methodology of such a study.

The current study analyses the possibly useful research methods, as well as both standard and non-standard sources of statistical information. The final report offers an overview of data gathered by public institutions, indicate possibilities of obtaining statistical information from other sources, and present the relevant scientific literature, illustrating examples of application of various research methods in practice (e.g., the use of aerial and satellite photography).

On the basis of the reviewed information, several promising sources of information on the actual size of the population of Warsaw have been identified. Firstly, data and estimates concerning population in the most mobile age groups – students and the working population – have been reviewed. Comparing these figures with the data of the Polish Central Statistical Office about the studying and working inhabitants of the capital, and including additional knowledge about the intersection of these categories, one can assess the size of a correction factor for the day-time population estimates.

Alternatively, the second possibility is to conduct a representative survey research with the aim to gather knowledge about the scale of the phenomenon of dwelling in Warsaw without formal registration. Additionally, such estimates might be supported by the data on the housing stock, available either from the official cadastral registers, or from aerial or satellite photography. The obtained figures can be subsequently verified on the basis of non-

standard sources of information, such as the data of the National Health Fund on the persons covered by the public medical insurance, or the data of the Ministry of Finance concerning the tax-payers living in Warsaw.

In order to distinguish between the day-time and the night-time (resident) population, data from the Warsaw Traffic Study (WBR) can be used, alongside with the traffic model for Warsaw constructed on their basis, which is at the disposal of the Municipal Office. The WBR is a survey-based study, conducted in 2005 by the Warsaw Development Planning Office (BPRW SA).

All the proposed methodologies of estimation are characterised by an inevitable margin of error, specific for particular methods and data sources. Nevertheless, information gathered within the framework of the study, although often partial, should enable at least an approximate calculation of the actual size of the population of Warsaw, as well as of the number of daily commuters to the capital of Poland. The current report provides some preliminary insights as to the order of magnitude of the expected corrections of the official figures: around a quarter of a million residents more than reported in the statistics, and on the top of that, about half a million daily commuters. More precise figures, however, need to be established on the basis of the results of a separate study, particularly devoted to the estimation of the population size of the City of Warsaw.

1. Wprowadzenie

Wiarygodna informacja o liczbie i strukturze ludności zamieszkującej dany obszar geograficzny jest elementem niezbędnym i kluczowym dla racjonalnego prowadzenia polityki gospodarczej, społecznej, jak i dla planowania przestrzennego. Niniejsze opracowanie ma na celu dokonanie inwentaryzacji zasobów danych mogących posłużyć oszacowaniu rzeczywistej liczby ludności m. st. Warszawy, jak również opracowanie metodologii takiego oszacowania. Jako, że „ludność rzeczywista” nie należy do kategorii pojęć jednoznacznie rozumianych w sensie statystyki demograficznej, w opracowaniu przeprowadzone zostały rozważania osobno dla ludności „diennej” i „nocnej”, zdefiniowanych w części 2.3 na tyle precyzyjnie, na ile pozwalały na to dostępne dane. Trzeba również podkreślić, że niniejsze *Studium...* rozważa Warszawę jako całość, nie zajmowaliśmy się problemem oszacowania ludności w poszczególnych dzielnicach, ani też w całym obszarze metropolitalnym, co wymagałoby modyfikacji części rozważanych metod. W obu przypadkach – definicji ludności oraz zakresu przestrzennego pracy – autorzy mają pełną świadomość niedoskonałości zaproponowanego podejścia, co jednak wynika w znacznej mierze z problemów z dostępnością i porównywalnością różnych źródeł informacji dla różnych jednostek przestrzennych.

Aby osiągnąć przedstawione zamierzenie, w studium zaprezentowano przede wszystkim informacje o różnych źródłach danych statystycznych, zarówno tradycyjnie wykorzystywanych w tego typu zagadnieniach badawczych, jak również niestandardowych. Dla różnych źródeł danych oraz metod szacunku przeprowadzona została dyskusja ich użyteczności z punktu widzenia ewentualnego projektu badawczego mającego na celu szacowanie liczby mieszkańców Warszawy. Dla tych spośród nich, które wydają się we wspomnianym kontekście obiecujące, w sposób szczegółowy przeanalizowano potencjalne możliwości i ograniczenia związane z pozyskaniem informacji i realizacją badania. Natomiast te źródła i metody, dla których wstępna analiza przydatności wykazała, że ich wykorzystanie do przeprowadzenia oszacowań byłoby niecelowe, lub pochłonięłoby nieproporcjonalnie dużo kosztów oraz czasu pracy w stosunku do spodziewanych korzyści, zostały w raporcie omówione jedynie skrótowo, z uzasadnieniem przyjętego stanowiska.

Studium oparte jest na trzech podstawowych źródłach informacji. Pierwszym z nich są oficjalne statystyki udostępniane przez instytucje, które zostały zidentyfikowane jako potencjalnie obiecujące źródła danych z punktu widzenia celu badania. Drugim źródłem są wywiady przeprowadzone bezpośrednio z przedstawicielami wybranych instytucji

publicznych, mające na celu sprawdzenie, czy możliwe jest pozyskanie innych informacji, niż publicznie dostępne. Wreszcie, trzecim komponentem wiedzy na temat możliwości szacowania rzeczywistej liczby ludności Warszawy są rozmaite raporty, ekspertyzy oraz publikacje naukowo-badawcze, ilustrujące przykłady wykorzystania różnych metod szacunku dla innych miast na świecie. W ten sposób informacja dotycząca danych i metod bezpośrednio związanych z szacowaniem ludności Warszawy jest pokazana na szerszym tle rozwoju metodologii badań geograficznych i demograficznych, z uwzględnieniem ograniczeń związanych ze specyfiką problematyki warszawskiej.

Raport składa się z sześciu rozdziałów. Po niniejszym wprowadzeniu, prezentującym cel, źródła i strukturę badania, w rozdziale drugim omówione zostały rodzaje informacji statystycznej, która może być wykorzystywana do szacowania liczby ludności Warszawy. Po przedstawieniu problemów związanych z dostępnością i kompletnością informacji statystycznej, wynikających z przepisów o ewidencji ludności, omówione zostały różne definicje stanów ludności (stała, faktycznie zamieszkała i rezydująca) oraz ich użyteczność dla oszacowania liczby mieszkańców miasta. Następnie przedyskutowane zostało rozróżnienie między ludnością nocną a dzienną i wynikające z niego implikacje dla niniejszego studium oraz metody rozszacowania liczby ludności według różnych kategorii.

Trzeci rozdział raportu zawiera propozycje badań wykorzystujących dostępne dane Głównego Urzędu Statystycznego, jak również standardowe metody badań społecznych. W szczególności, przedyskutowane zostały sugestie odnośnie metod szacowania liczebności wybranych grup ludności: studentów, ludności aktywnej zawodowo, oraz osób bezdomnych. Ponadto, zaoferowane zostały możliwości korekt stanów ludności w oparciu o wyniki sondażowych badań reprezentacyjnych, które mogłyby zostać przeprowadzone w oparciu o rejestry terytorialne GUS, lub o generowane komputerowo numery telefoniczne.

Kolejny, czwarty rozdział, poświęcony jest możliwościom przeprowadzenia szacunków w oparciu o niestandardową informację statystyczną oraz inne, potencjalnie interesujące źródła wiedzy o procesach urbanistycznych. Przedyskutowana została dostępność i możliwość wykorzystania danych statystycznych z wybranych instytucji, takich jak: Ministerstwo Finansów (Urzędy Skarbowe), Mazowiecki Oddział Narodowego Funduszu Zdrowia, Państwowa Komisja Wyborcza, Biuro Administracji i Spraw Obywatelskich m. st. Warszawy, jak również Instytut Statystyki Kościoła Katolickiego. Omówione zostały też możliwości i ograniczenia związane z przeprowadzeniem obliczeń bazujących na danych o wybranych segmentach gospodarki komunalnej w Warszawie: zużyciu wody, produkcji ścieków oraz gospodarki odpadami. Na koniec, na podstawie literatury przedmiotu

zarysowana została metodologia badań wykorzystujących do szacowania liczby ludności informację ze zdjęć satelitarnych i lotniczych.

Oddzielnie, w rozdziale piątym, przedstawiony został wybór metod szacowania liczby ludności dziennej Warszawy w oparciu o różne statystyki. W niniejszym studium rozważane były dwa źródła danych o codziennych dojazdach do miasta (*commuting*): pomiary natężenia ruchu drogowego oraz przewozów świadczonych przez przedsiębiorstwa komunikacji zbiorowej, jak również różne rodzaje danych o ludności i zatrudnieniu, pozwalające na szacowanie salda ludności dziennej według metody stosowanej w Stanach Zjednoczonych przez US Bureau of Census.

Na zakończenie, w rozdziale szóstym raportu, przedstawione zostało podsumowanie wyników całego studium oraz wynikających z nich wniosków wraz z praktycznymi rekomendacjami dla Urzędu m. st. Warszawy, dotyczącymi projektu badawczego mającego na celu szacowanie liczby mieszkańców miasta. Raport zamyka lista wykorzystanych pozycji bibliograficznych, wraz z publikacjami zawierającymi oficjalne dane upowszechniane przez poszczególne instytucje, oraz zestawienie źródeł wykorzystanej informacji bezpośredniej – wywiadów bezpośrednich lub telefonicznych z cytowanymi w opracowaniu przedstawicielami urzędów i przedsiębiorstw.

2. Rodzaje informacji statystycznej o ludności Warszawy

Rozpoczynając dyskusję dotyczącą oszacowania rzeczywistej liczby mieszkańców Warszawy, niezbędne jest przede wszystkim oszacowanie dostępności i kompletności dostępnej informacji statystycznej (część 2.1) oraz zdefiniowanie przedmiotu badania – różnych kategorii ludności, które mogą być potencjalnie interesujące z punktu widzenia Urzędu Miasta. W literaturze przedmiotu spotyka się różne określenia dotyczące ludności przebywającej na danym obszarze geograficznym. Z punktu widzenia niniejszego opracowania, najważniejszymi kategoriami ludności wydają się być: osoby stale i faktycznie zamieszkałe, rezydenci, oraz ludność dzienna i nocna, definicje których omówione zostały w częściach 2.2 oraz 2.3. Część 2.4 przedstawia pokrótce zarys metody dezagregacji łącznej liczby ludności według wybranych cech, takich jak wiek oraz dzielnica zamieszkania.

2.1. Ewidencja ludności a dostępność i kompletność informacji statystycznej

Przy szacowaniu liczby ludności według jakiegokolwiek definicji trzeba brać pod uwagę, że obliczenia zawsze są obciążone różnymi błędami, spośród których najważniejsze to błąd pomiaru, oraz błąd szacunku wynikający z konstrukcji metody szacowania. W niniejszej części przedstawione zostały problemy związane błędami pomiaru, podczas gdy błędy szacunku są omawiane w kolejnych rozdziałach raportu, przy okazji dyskusji nad poszczególnymi metodami badania.

W przypadku zjawisk ludnościowych błędy pomiaru najczęściej są związane z niekompletną rejestracją statystyczną, nieprawidłowymi wartościami zmiennych opisujących daną osobę lub zdarzenie demograficzne (np. podaną niewłaściwą datą urodzenia), oraz, zdecydowanie rzadziej, z obecnością zduplikowanej informacji na temat danej osoby lub zdarzenia. Dla szacowania rzeczywistej liczby ludności Warszawy kluczowy problem stanowi pierwsze spośród wymienionych źródeł błędów – braki w rejestracji. Możliwościom korekty tych braków poświęcone jest całe niniejsze opracowanie.

W Polsce kwestia rejestracji ludności regulowana jest przez *Ustawę o ewidencji ludności i dowodach osobistych* z dnia 10 kwietnia 1974 r., z późniejszymi zmianami¹. W odniesieniu do rejestracji ludności, przepisy wspomnianej ustawy stanowią, że:

¹ Tekst jednolity – Dz. U. z 2006 r., nr 139, poz. 993.

„Art. 4.

Obowiązek meldunkowy polega na:

- 1) zameldowaniu się w miejscu pobytu stałego lub czasowego;
- 2) wymeldowaniu się z miejsca pobytu stałego lub czasowego; [...]

Art. 5.

1. Osoba posiadająca obywatelstwo polskie i przebywająca stale na terytorium Rzeczypospolitej Polskiej jest obowiązana zameldować się w miejscu pobytu stałego.
2. W tym samym czasie można mieć tylko jedno miejsce pobytu stałego.

Art. 6.

1. Pobytem stałym jest zamieszkanie w określonej miejscowości pod oznaczonym adresem z zamiarem stałego przebywania. [...]

Art. 7.

1. Pobytem czasowym jest przebywanie bez zamiaru zmiany miejsca pobytu stałego w innej miejscowości pod oznaczonym adresem lub w tej samej miejscowości, lecz pod innym adresem. [...]

Art. 8.

1. Osoba zameldowana na pobyt czasowy i przebywająca w tej samej miejscowości nieprzerwanie dłużej niż 2 miesiące jest obowiązana zameldować się na pobyt stały [...]. Za okoliczności uzasadniające zameldowanie na pobyt czasowy trwający ponad 2 miesiące uważa się w szczególności:
 - 1) wykonywanie pracy poza miejscem pobytu stałego;
 - 2) pobyt związany z kształceniem się, leczeniem, wypoczynkiem lub ze względów rodzinnych;
 - 3) odbywanie czynnej służby wojskowej;
 - 4) pobyt w zakładach karnych i poprawczych, aresztach śledczych, schroniskach dla nieletnich i zakładach wychowawczych. [...]
3. W tym samym czasie można mieć tylko jedno miejsce pobytu czasowego trwającego ponad 2 miesiące. [...]

Art. 23.

Cudzoziemiec przebywający na terytorium Rzeczypospolitej Polskiej jest obowiązany wykonywać obowiązek meldunkowy na zasadach określonych w ustawie, jeżeli przepisy niniejszego rozdziału nie stanowią inaczej.”

Za niedopełnienie obowiązku meldunkowego grożą sankcje przewidziane w kodeksie wykroczeń²:

„Art. 147.

§ 1. Kto nie dopełnia ciążącego na nim obowiązku meldunkowego, podlega karze ograniczenia wolności, grzywny [w wysokości od 20 do 5.000 złotych – przyp. aut.] albo karze nagany.”

Pomijając kwestie sankcji oraz skuteczności egzekwowania wspomnianych przepisów, istnieje wiele potencjalnych przyczyn niedopełniania obowiązku meldunkowego przez obywateli. Na przykład, do niedawna w ustawie funkcjonował zapis, który wymagał od osoby meldującej się potwierdzenia uprawnienia do przebywania w lokalu mieszkalnym. Przepis ten został uchylony dopiero wyrokiem Trybunału Konstytucyjnego z dnia 27 maja 2002 r. jako niezgodny z zapisaną w ustawie zasadniczej zasadą wolności wyboru miejsca zamieszkania i pobytu³.

² Ustawa z dnia 20 maja 1971 r. *Kodeks wykroczeń*, Dz. U. z 1971 r., nr 12, poz. 114, z późniejszymi zmianami.

³ Sygn. Akt K 20/01, Dz. U. z 2002 r., nr 78, poz. 716.

Pomimo ułatwienia zameldowania na skutek wejścia w życie orzeczenia TK, można przypuszczać, że część osób wciąż nie dopełnia obowiązku meldunkowego. Potencjalnych przyczyn może być wiele: po pierwsze, kwestia posiadania (bądź braku) meldunku wciąż może być bardzo delikatna, zwłaszcza po kilkudziesięciu latach silnych restrykcji meldunkowych w okresie PRL. Część osób wynajmujących mieszkania może chcieć ukrywać fakt posiadania lokatorów ze względów fiskalnych. Ponadto, w grę wchodzić może obecność innych profitów, np. związanych z niższymi stawkami za ubezpieczenie pojazdów zarejestrowanych w innych miejscowościach, itp.

Wydaje się, że dopiero w sytuacji, gdy fakt zameldowania się przekłada się na dostęp do pewnych istotnych usług, jak np. w przypadku wyboru lekarza pierwszego kontaktu dla rodzin z dziećmi, miejsc w przedszkolach publicznych itp., może to znaleźć przełożenie na zwiększenie kompletności ewidencji ruchu ludności. To samo dotyczy migrantów zagranicznych, których do dopełnienia obowiązków meldunkowych w obu krajach może skłonić chęć uniknięcia podwójnego opodatkowania.

Ewidencja ludności jest jednym ze źródeł informacji statystycznej wykorzystywanym przez GUS. O ile dla wybranych lat dostępne są oszacowania liczby ludności przeprowadzone na podstawie narodowych spisów powszechnych (NSP), ostatnio w latach 1988 oraz 2002, o tyle dla okresów międzyspisowych wyjściowa liczba ludności ze spisu jest korygowana o saldo ruchu naturalnego (urodzeń i zgonów) oraz zameldowań i wymeldowań⁴. Wynika stąd, że błędy w liczbie ludności publikowanej przez GUS dla poszczególnych obszarów mogą wynikać zarówno z błędów oszacowań ludności spisowej, oraz z niedociągnięć periodycznych korekt opartych o informacje pochodzące z ewidencji ludności.

Dla przykładu, znaczna część emigrantów zagranicznych z Polski jest wciąż uwzględniona w krajowych rejestrach – na podstawie danych krajów docelowych szacuje się, że w 2002 r. z ewidencji ludności wymeldował się co szósty emigrant (Bijak i Koryś, 2006: 28). W przypadku Warszawy, o niekompletności zameldowań świadczy z kolei fakt, że spis powszechny z 2002 roku wykazał o 62 tysiące (3.7%) mieszkańców więcej, niż wynikałoby to z wcześniejszej oficjalnej rejestracji opartej na ewidencji ruchu naturalnego i migracyjnego (zameldowań i wymeldowań) ludności (Śleszyński, 2004a: 97).

Ponadto, należy pamiętać, że w oficjalnych statystykach GUS, ludność może być zliczana według kilku alternatywnych definicji, mających potencjalnie różną użyteczność dla

⁴ Szczegóły dotyczące ewidencji ludności, rejestracji statystycznej oraz danych o liczbie, strukturze, ruchu naturalnym i migracyjnym ludności w Polsce – zob. np. Kupiszewska, Nowok i Kupiszewski (2006: 577–589).

celów planowania przestrzennego. W NSP z 2002 roku GUS przyjął trzy różne definicje stanu ludności: ludność stałą (zameldowaną na pobyt stały), ludność faktycznie zamieszkałą oraz tzw. ludność rezydującą (rezydentów), przedstawione szczegółowo w kolejnej części. Warto przy tym zwrócić uwagę, za J. Paradyszem (2005a), że kategorie spisowe są kategoriami deklaracyjnymi, w przeciwieństwie do ewidencji zameldowania na pobyt stały lub czasowy, które jest ściśle określoną kategorią administracyjno-prawną.

2.2. Definicje stanów ludności i ich użyteczność dla szacowania liczby mieszkańców stolicy

2.2.1. Ludność stała (stałe zamieszkała)

Punktem wyjścia do definiowania wszystkich kategorii ludności jest ludność stała (stałe zamieszkała), będąca kategorią prawną, powiązaną z miejscem stałego pobytu według ewidencji ludności (ludność *de jure*), niekoniecznie odpowiadającą stanowi rzeczywistemu (ludności *de facto*). Główny Urząd Statystyczny w Narodowym Spisie Powszechnym zdefiniował ludność stałą w następujący sposób⁵:

„Kategoria obejmuje osoby mieszkające stale (z reguły zameldowane na pobyt stały) oraz:

- obecne w czasie spisu, a dokładnie w momencie krytycznym spisu,
- nieobecne w momencie spisu, bez względu na miejsce przebywania i czas trwania tej nieobecności.”

Ze względu na drugi element definicji („osoby [...] nieobecne w momencie spisu”), liczba ludności stałej jako taka, bez dokonania niezbędnych korekt przybliżających stan faktyczny, nie jest użyteczna z punktu widzenia racjonalizacji polityki gospodarczej i społecznej oraz planowania przestrzennego. Kategoria ludności stałej jest jednak istotna z innego powodu: podstawowe zdarzenia demograficzne – urodzenia, zgony i migracje – przypisywane są do miejsca stałego zameldowania odpowiednio matki noworodka, osoby zmarłej, lub migranta, co również nie musi odzwierciedlać faktycznego miejsca zamieszkania⁶.

2.2.2. Ludność faktycznie zamieszkała

Podstawową kategorią opisu stanów ludności *de facto* stosowaną w polskiej statystyce publicznej jest ludność faktycznie zamieszkała, o następującej definicji według GUS⁷:

⁵ Za: GUS (2003: 15).

⁶ Por. np.: Paradysz (2005a: 1), GUS (2006a: 15).

⁷ Za: GUS (2003: 15–16).

„Kategoria obejmuje następujące grupy osób:

1. Mieszkające stale, które

- a) były obecne w momencie spisu;
- b) były nieobecne w momencie spisu, ale ich nieobecność trwała mniej niż 2 miesiące;
- c) były nieobecne w momencie spisu przez okres dłuższy niż 2 miesiące, ale ich nieobecność wynikała z następujących powodów:
 - odbywania zasadniczej służby wojskowej lub innej w systemie skoszarowanym oraz uczestniczenia w misjach wojskowych;
 - przebywania w zakładzie karnym lub śledczym;
 - pobytu za granicą.

2. Przebywające czasowo przez okres powyżej 2 miesięcy. Dotyczy to osób, które mieszkają na stałe w innym miejscu (są tam zameldowane na pobyt stały), natomiast w miejscu spisania przebywają czasowo z następujących powodów: nauka, praca, warunki rodzinne lub mieszkaniowe, leczenie lub rehabilitacja, przebywanie w domu opieki.

Jako czas okresowej nieobecności lub przebywania przyjęty został czas faktyczny, a nie zamierzony. Kategoria ludności faktycznie zamieszkałej nie uwzględnia osób przybyłych z zagranicy na pobyt czasowy, tj. tych, które nie posiadają stałego pobytu w Polsce (pozwolenia na osiedlenie się).”

Według *Programu badań statystycznych statystyki publicznej*⁸, dane o ludności faktycznie zamieszkałej oraz ludności zameldowanej na stałe, przygotowane na poziomie podregionów i powiatów (tj. odpowiednim dla Warszawy), są udostępniane przez GUS co kwartał. Dostęp do informacji gromadzonej przez GUS dla organów administracji publicznej, takich, jak np. Urząd m. st. Warszawy, jest bezpłatny.

Stosowana przez GUS powyższa definicja ludności faktycznie zamieszkałej nie bierze pod uwagę zdecydowanej większości migracji zagranicznych, mianowicie tych na czas określony, co powoduje istotne zafałszowanie stanów ludności prezentowanych przez statystykę państwową. W efekcie, od ludności faktycznie zamieszkałej odliczani są jedynie długookresowi emigranci zagraniczni (przebywający za granicą ponad 12 miesięcy), a doliczani jedynie długookresowi imigranci posiadający prawo stałego pobytu.

Ludność faktycznie zamieszkała jest wykorzystywana do międzyspisowych korekt stanów ludności przygotowywanych przez GUS. Jest ona uzyskiwana poprzez modyfikację liczby ludności stałej dla kolejnych okresów o dodanie salda ludności czasowo zameldowanej. Taka metoda obliczeń jest wymuszana przez system rejestracji zdarzeń demograficznych w Polsce, który odnosi urodzenia, zgony, czy migracje do miejsca stałego, a nie faktycznego zamieszkania. Dokonane w ten sposób korekty nie są więc spójne metodycznie, biorąc pod uwagę, że zdarzenia przypisywane są do ludności stałej, a korekty dokonywane są w oparciu o ludność faktycznie zamieszkałą – następuje tu konfuzja obu kategorii ludności. Co ważniejsze, rozbieżności między ludnością stałą a faktycznie

⁸ Por. np.: *Program badań statystycznych statystyki publicznej na rok 2007*, Załącznik do rozporządzenia Rady Ministrów z dnia 5 września 2006 r., Dz. U. z 2006 r., nr 170, poz. 1219.

zamieszkałą, zwłaszcza dla niewielkich obszarów, mogą być znaczne: J. Paradysz (2005a) pokazał, że według danych NSP z 2002 roku, w przypadku niektórych jednostek urbanistycznych miasta Poznania różnice sięgały nawet kilkudziesięciu procent.

W przypadku zarówno ludności stałej, jak i faktycznie zamieszkałej, do podstawowych źródeł błędów oszacowań należy: (a) ignorowanie większości emigracji zagranicznych, powodujące traktowanie emigrantów jako zamieszkałych w Polsce do momentu wymeldowania się na stałe, (b) ignorowanie większości imigracji zagranicznych poprzez włączanie imigrantów do stanów ludności dopiero z chwilą uzyskaniu przez nich stałego prawa pobytu w Polsce, oraz (c) zbędny podział na ludność zameldowaną na stałe i tymczasowo, będący dziedzictwem po systemie kontroli migracji z okresu PRL, obecnie zaś niczemu nie służący (Paradysz, 2005a; Bijak i Kupiszewski, 2006).

2.2.3. Rezydenci (ludność rezydująca)

W świetle rekomendacji dotyczących rejestracji statystycznej, przyjętych przez Organizację Narodów Zjednoczonych w 1998 roku, za długookresowego migranta powinno uważać się osobę, która „przenosi się z kraju zwyczajowego zamieszkania (*usual residence*) do innego kraju, na okres co najmniej roku (12 miesięcy), tak, że kraj docelowy efektywnie staje się dla tej osoby nowym krajem zwyczajowego zamieszkania”⁹. Warto podkreślić, że taka definicja migranta może stosować się również do przepływów ludności między różnymi jednostkami administracyjnymi jednego kraju, oraz, że stanowi ona dopełnienie liczby ludności „zazwyczaj zamieszkałej” na danym obszarze. Kategoria ludności rezydującej (rezydentów), odpowiadająca definicji ludności zazwyczaj zamieszkałej przyjętej przez ONZ, została w polskiej statystyce publicznej uwzględniona dopiero w czasie NSP w 2002 r., i określona następująco¹⁰:

„Do rezydentów zalicza się:

- a) stałych mieszkańców, z wyjątkiem osób przebywających poza miejscem zamieszkania przez okres co najmniej 12 miesięcy – bez względu na ich miejsce przebywania (w kraju czy za granicą);
- b) osoby przebywające czasowo przez okres co najmniej 12 miesięcy, przybyłe z innego miejsca w kraju lub z zagranicy (cudzoziemcy).

Jako kryterium przemieszczania się ludności przy wyodrębnianiu tej kategorii przyjmuje się: naukę, pracę, warunki rodzinne i mieszkaniowe, leczenie i rehabilitację, pobyt w domu opieki. Oznacza to, że osoby odbywające zasadniczą służbę wojskową (w systemie skoszarowanym) lub przebywające w zakładach karnych czy aresztach – bez względu na czas nieobecności – są zaliczane do rezydentów miejscowości, w których zamieszkiwali przed jej «przymusowym» opuszczeniem.”

⁹ Za: ONZ (1998: 18), w tłumaczeniu autorów.

¹⁰ Za: GUS (2003: 16).

Ludność rezydująca daje potencjalnie najlepsze możliwe przybliżenie rzeczywistego stanu ludności, zgodne z rekomendacjami ONZ z 1998 roku. Kategoria ta powinna obejmować wszystkie osoby, które przebywały w Warszawie przez ostatnie 12 miesięcy, niezależnie od swojej sytuacji prawnej (czyli tzw. ludność *de facto*), oraz nie obejmować osób, które wyemigrowały z miasta na okres dłuższy niż rok. Problemem pozostaje, że w świetle *Programu badań statystycznych statystyki publicznej*, coroczne dane o rezydentach dla okresów międzypisowych nie są dostępne na niższym poziomie, niż wojewódzki, czyli nie są odpowiednie dla badania liczby ludności Warszawy.

Zastosowanie definicji ludności rezydującej w czasie spisu powszechnego w 2002 roku dało w efekcie oszacowanie liczby ludności Polski mniejsze o 610 tysięcy osób (1.6%) od populacji faktycznie zamieszkałej. Dla m. st. Warszawy oraz dzielnic według obecnego podziału administracyjnego, zestawienie liczb ludności faktycznie zamieszkałej i rezydującej, oszacowanych na podstawie wyników NSP z 2002 r. przedstawione jest w tablicy 1. Wynika z niego, że ludność rezydująca Warszawy, kategoria uznawana z przyczyn metodologicznych za bardziej wiarygodną i potencjalnie mogąca stanowić punkt wyjścia do dalszych analiz, jest o ok. 2.7% (45 tysięcy osób) mniej liczna od ludności faktycznie zamieszkałej. Podobne różnice występują na poziomie dzielnic, gdzie przeszacowanie liczby ludności faktycznie zamieszkałej w stosunku do rezydującej waha się od 1.6% w Wilanowie do 4.9% w Białołęce.

Tablica 1. Ludność rezydująca a ludność faktycznie zamieszkała w dzielnicach Warszawy wg NSP 2002

Dzielnice	Liczba ludności		Rezydenci jako % ludności faktycznej
	faktycznie zamieszkałej	rezydującej	
M. st. Warszawa	1 689 201	1 644 155	97.3%
- Bemowo	104 663	102 149	97.6%
- Białołęka	57 765	54 944	95.1%
- Bielany	138 356	134 387	97.1%
- Mokotów	235 381	228 109	96.9%
- Ochota	94 178	91 072	96.7%
- Praga Południe	189 251	185 347	97.9%
- Praga Północ	75 348	73 845	98.0%
- Rembertów	21 601	20 967	97.1%
- Śródmieście	138 299	134 465	97.2%
- Targówek	124 316	122 250	98.3%
- Ursus	42 547	41 591	97.8%
- Ursynów	134 440	129 861	96.6%
- Wawer	61 817	60 357	97.6%
- Wesoła*	17 531	17 063	97.3%
- Wilanów	13 731	13 507	98.4%
- Włochy	39 176	38 404	98.0%
- Wola	148 869	145 651	97.8%
- Żoliborz	51 932	50 186	96.6%

* W czasie NSP 2002 Wesoła była oddzielnym miastem i nie wchodziła w skład m. st. Warszawy.

Źródło: GUS - Narodowy Spis Powszechny Ludności i Mieszkań w 2002 r. GUS (2003)

W rzeczywistości, jak zauważył P. Śleszyński (2004b), biorąc pod uwagę braki odpowiedzi w formularzach spisowych, wskazujące na to, że dana osoba nie została spisana „z natury” w miejscu swojego zamieszkania, tylko na podstawie rejestrów ewidencji ludności, różnica między ludnością rezydującą a faktycznie zamieszkałą w Warszawie może być jeszcze większa. Na przykład, dla zmiennej opisującej źródło utrzymania danej osoby, brakowało 82.1 tysiąca odpowiedzi spośród osób wykazanych w NSP 2002 jako faktyczni mieszkańcy Warszawy (tamże: 106). Jednak, jak zauważa autor cytowanego opracowania, akurat w przypadku stolicy może wynikać z wielu przyczyn, jak na przykład przemieszczanie się ludności w obrębie miasta lub aglomeracji, a niekoniecznie emigracji.

2.3. Ludność nocna a ludność dzienna

Zaprezentowane powyżej oficjalne definicje stanów ludności według GUS, stosowane w polskiej statystyce publicznej, oparte są o kategorię miejsca zamieszkania – pojęcie odnoszące się do tzw. ludności nocnej, przebywającej (zamieszkującej) na danym obszarze w ciągu nocy. Ludność nocna jest więc terminem odpowiednim dla zastosowań w analizach demograficznych, dotyczących przemian ludnościowych na danym terytorium.

W przypadku miast, a w szczególności dużych aglomeracji, takich, jak warszawska, ludność nocna nie jest jednak odpowiednia do planowania rozwoju pewnych działań infrastruktury (obciążenia transportu miejskiego, sieci wodociągowych, zaopatrzenia detalicznego, itp.). Dla celów polityki społeczno-ekonomicznej oraz planowania przestrzennego pożądane jest więc posiadanie dodatkowo informacji o grupie ludności dziennej (spędzającej dzień na danym obszarze), która w dużej mierze ma wpływ na obciążenie infrastruktury miejskiej. Od liczebności obu kategorii ludności – nocnej i dziennej – zależy zapotrzebowanie zarówno na infrastrukturę miejską o charakterze sieciowym (transport, wodociągi, kanalizacja, elektryczność, gaz, łączność), jak i na infrastrukturę o charakterze społecznym (przedszkola, szkoły, przychodnie, urzędy).

Jako, że w polskiej statystyce publicznej informacja statystyczna o ludności jest kreowana głównie z punktu widzenia demograficznego, a nie urbanistycznego¹¹, dane o ludności dziennej nie są zbierane przez GUS. Ponadto, w przypadku ludności dziennej występuje dodatkowy problem z dokładnym określeniem tej kategorii. Liczba osób

¹¹ Pomijamy tu najważniejsze cele gromadzenia danych ludnościowych, związane z zagadnieniami *stricte* politycznymi, np. wyborczymi, oraz ekonomicznymi, np. podatkowymi (podziękowania pod adresem Prof. Marka Okólskiego za zwrócenie nam uwagi na te kwestie).

spędzających dzień na danym obszarze zależy od zdefiniowania zakresu czasu przebywania w danym miejscu – czy ludność dzienna obejmować będzie tylko osoby dojeżdżające do pracy lub szkoły w danej miejscowości, czy również osoby przybywające po zakupy, do urzędu, ośrodków kultury, itp.

Dodatkowo, czynnikiem praktycznie uniemożliwiającym jednoznaczne zdefiniowanie ludności dziennej, oraz zaburzającym pomiar jej liczby, jest wybór momentu odniesienia – dnia tygodnia oraz pory dnia. W przypadku dużego miasta takiego, jak Warszawa, przyciągającego codziennie w różnych celach wielu mieszkańców okolicznych miejscowości, można spodziewać się, że liczba ludności dziennej w sobotę lub niedzielę będzie niższa od tej w pozostałe dni tygodnia. Również liczba ludności przebywająca w mieście w dni powszednie około godziny 14:00 będzie wyższa od liczby ludności o godzinie 7:00, czy 20:00. Dlatego też dokładniejsza dyskusja możliwości szacowania ludności dziennej m. st. Warszawy w oparciu o różne dane, przedstawiona w rozdziale 5 niniejszego opracowania, nie jest oparta na jednej zunifikowanej definicji, lecz na szeregu kategorii pojęciowych wynikających bezpośrednio z charakteru dostępnych źródeł. W tym aspekcie niniejsze studium odwołuje się więc do metodyki zaproponowanej na przykład przez O. Bojanowską oraz G. Węclawowicza (1984), którzy oparli szacunki ludności dziennej Radomia *stricto* o dane na temat dojazdów do pracy i szkoły, dostępnych w NSP z 1978 roku.

2.4. Dezagregacja liczby ludności według różnych cech

Oprócz oszacowania samej liczby ludności, dla celów planistycznych potencjalnie bardzo przydatna jest także informacja o strukturze ludności według podstawowych charakterystyk demograficznych (płci i wieku) i geograficznych (np. dzielnicy zamieszkania). Z reguły taka informacja nie jest dostępna, zwłaszcza w odniesieniu do kategorii nieznanych i podlegających oszacowaniu (takich, jak wielkość korekty liczby ludności). Pomimo to, wspomniane struktury można w dużym przybliżeniu szacować nawet na podstawie fragmentarycznych danych, wykorzystując standardowe narzędzia analizy demograficznej.

Zakładając, że niedoszacowanie liczby ludności Warszawy jest wyłącznie wynikiem nierejestrowanych migracji, do rozszacowania wielkości korekty tej liczby według wieku można zastosować na przykład podwójny model wykładniczy natężenia migracji autorstwa A. Rogersa oraz L. J. Castro (np.: Rogers i Castro, 1981; Castro i Rogers, 1983)¹². W

¹² Założenie to wydaje się być dość realistyczne. Z pracy P. Śleszyńskiego (2005a), w której autor pokazuje wielkość korekty liczby ludności faktycznie zamieszkałej w oparciu o wyniki NSP 2002 w różnych

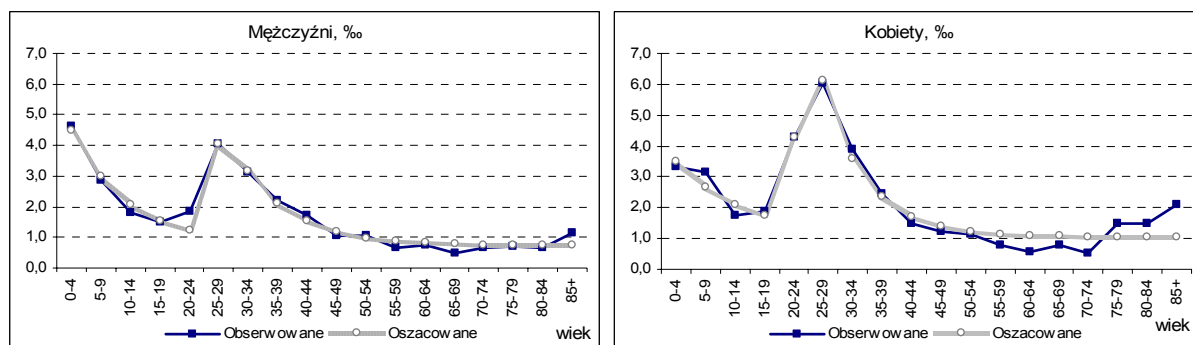
najprostszej postaci wspomniany model dla współczynnika natężenia migracji (mobilności) $\phi(x)$ dla ludności w wieku x , zdefiniowanego jako liczba migrantów w wieku x odniesiona w stosunku do 1,000 mieszkańców regionu pochodzenia, lub regionu docelowego migrantów będących w wieku x , ma postać daną równaniem (tamże):

$$(1) \quad \phi(x) = c + a_1 \cdot \exp(-\alpha_1 \cdot x) + a_2 \cdot \exp\{-\alpha_2 \cdot (x - \mu_2) - \exp[-\lambda_2 \cdot (x - \mu_2)]\}.$$

Wielkości c , a_1 , α_1 , a_2 , α_2 , μ_2 oraz λ_2 stanowią parametry modelu (1), które mogą zostać oszacowane na podstawie danych empirycznych, na przykład dotyczących migracji wewnętrznych w Polsce. Model Rogersa i Castro można szacować dla każdej płci z osobna.

Przykładowe dane dotyczące natężenia $\phi(x)$ zarejestrowanych migracji do i z Warszawy dla 1995 roku, pochodzące z oficjalnych, choć niepublikowanych danych GUS, zilustrowane zostały w pracy A. Potrykowskiej i P. Śleszyńskiego (1999), w przekrojach według płci i pięcioletnich grup wieku. Na podstawie tych danych, dotyczących na przykład migracji do Warszawy z obszaru ówczesnego województwa stołecznego warszawskiego (bez przepływów w obrębie miasta)¹³, można oszacować współczynniki natężenia napływu ludności w przeliczeniu na 1,000 faktycznych mieszkańców stolicy. Przeprowadzona na podstawie powyższych danych estymacja parametrów¹⁴ modelu (1) daje w efekcie krzywe typu Rogersa i Castro przedstawione na rysunku 1.

Rysunek 1. Natężenie migracji do Warszawy (na 1,000 mieszkańców) z obszaru woj. warszawskiego, 1995



Wartości współczynnika determinacji (miary dopasowania): $R^2 = 0.963$ dla kobiet i $R^2 = 0.924$ dla mężczyzn.

Oszacowania parametrów – kobiety: $c = 1.05$, $a_1 = 3.01$, $\alpha_1 = 0.08$, $a_2 = 9.69$, $\alpha_2 = 0.14$, $\mu_2 = 22.53$, $\lambda_2 = 8.42$;
– mężczyźni: $c = 0.73$, $a_1 = 4.85$, $\alpha_1 = 0.10$, $a_2 = 4.08$, $\alpha_2 = 0.11$, $\mu_2 = 27.37$, $\lambda_2 = 9.52$.

Źródło: obliczenia własne na podstawie danych z pracy P. Śleszyńskiego i A. Potrykowskiej (1999)

przekrojach, wynika, że np. dla Warszawy struktura tej korekty według wieku dla kobiet odpowiada wzorcowi natężenia migracji według modelu Rogersa i Castro (Śleszyński, 2005a: 201, Rysunek 3).

¹³ Autorzy są wdzięczni P. Śleszyńskiemu za udostępnienie kompletu danych z cytowanego opracowania.

¹⁴ Estymacja przy użyciu metody najmniejszych kwadratów, obliczenia wykonane zostały w pakiecie MS Excel 2003 z wykorzystaniem narzędzia Solver.

W celu otrzymania analogicznych struktur dla korekty liczby ludności mieszkańców miasta, należałoby następnie otrzymane krzywe przeskalować tak, by dla wszystkich grup wieku łącznie odpowiadały całkowitej wielkości korekty odniesionej w stosunku do 1,000 ludności faktycznie zamieszkanej w Warszawie.

Podstawowym problemem zastosowanej powyżej metody jest kwestia źródła danych wykorzystanych do oszacowania parametrów modelu. Założenie o tym, że struktury według wieku dla (a) natężenia imigracji wewnątrzwojewódzkich oraz (b) wielkości korekty liczby ludności, w przeliczeniu na 1,000 ludności Warszawy, są identyczne (z dokładnością do proporcjonalności), jest bardzo silne i może nie odpowiadać rzeczywistości¹⁵. Na przykład, w świetle argumentacji przedstawionej w części 2.1, udział rodzin z dziećmi wśród populacji osób niezameldowanych jest najprawdopodobniej dużo niższy, niż wynikałoby to z rysunku 1, ze względu na kwestie dostępu do opieki zdrowotnej, przedszkoli, szkół, itp. Dla odmiany, współczynniki wyższe niż pokazane charakteryzować będą dwie najbardziej mobilne grupy ludności – studentów oraz osoby pracujące. Kwestia ta została dokładniej omówiona w części 3.2 niniejszego opracowania.

Opcją alternatywną w stosunku do danych oficjalnych może być przyjęcie za punkt wyjścia na przykład struktur uzyskanych z reprezentacyjnych badań społecznych, takich jak zaprezentowanych w pracy B. Cichomskiego (2004), czy też szerzej, całego badania *Warsaw Area Study*, zrealizowanego w 2001 roku przez Instytut Studiów Społecznych Uniwersytetu Warszawskiego na zlecenie Urzędu m. st. Warszawy¹⁶. Można w tym celu także wykorzystać informacje z badań przeprowadzonych specjalnie w celu oszacowania rzeczywistej liczby ludności Warszawy (część 3.2). W obu wypadkach struktury mogą być następnie wygładzone za pomocą modelu Rogersa i Castro. Ze względu na problem wielkości próby w tego typu badaniach, oszacowania powinny być oparte na grupach wieku nie węższych, niż pięcioletnie.

Posiadając informację o przekrojach liczby ludności według kilku cech z osobna (np. wieku i dzielnicy zamieszkania), ale nie o łącznym rozkładzie według wszystkich tych zmiennych, można się pokusić o odtworzenie macierzy takiego łącznego rozkładu na podstawie danych o wartościach brzegowych. W demografii najbardziej rozpowszechnioną metodą takiego rozszacowania jest tzw. metoda iteracyjnego dopasowania proporcjonalnego

¹⁵ W niniejszym raporcie pojęcia ‘mocne / silne’ są używane w kontekście takich założeń, od których w znacznym stopniu zależy tok dalszego rozumowania i w konsekwencji wyniki obliczeń.

¹⁶ Poza strukturą ludności według wieku, badanie *Warsaw Area Study* może dostarczyć również wielu cennych informacji na temat innych charakterystyk socjo-demograficznych ludności miasta (podziękowania pod adresem Prof. Marka Okólskiego za zwrócenie naszej uwagi na wieloaspektowość wspomnianych prac).

(*iterative proportional fitting*, IPF), omówiona dla analiz demograficzno-urbanistycznych na przykład w pracy P. H. Rees'a (1994)¹⁷.

Rozważmy przypadek dwóch zmiennych A i B (np. wiek i dzielnica zamieszkania), z których pierwsza przyjmuje m wartości R_i ($i = 1, \dots, m$), druga zaś – n wartości C_j ($j = 1, \dots, n$). Szacowana jest wówczas macierz łącznego rozkładu liczby ludności według obu cech, $\mathbf{P} = [p_{i,j}]_{m \times n}$, o wymiarach m wierszy na n kolumn, przedstawiona w tablicy 2. Warto zauważyć, że wartości R_i oraz C_j są odpowiednimi wartościami brzegowymi macierzy $\mathbf{P}$, i są w procedurze IPF dane, w przeciwieństwie do wartości $p_{i,j}$, które są szacowane w sposób iteracyjny, opisany poniżej.

Tablica 2. Procedura iteracyjnego dopasowania proporcjonalnego – szacowana macierz $\mathbf{P}$

Zmienna B \ Zmienna A	Kategoria 1	Kategoria 2	...	Kategoria n	Suma
Kategoria 1	$p_{1,1}$	$p_{1,2}$	...	$p_{1,n}$	$R_1 = \sum_j p_{1,j}$
Kategoria 2	$p_{2,1}$	$p_{2,2}$	...	$p_{2,n}$	$R_2 = \sum_j p_{2,j}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
Kategoria m	$p_{m,1}$	$p_{m,2}$	...	$p_{m,n}$	$R_m = \sum_j p_{m,j}$
Suma	$C_1 = \sum_i p_{i,1}$	$C_2 = \sum_i p_{i,2}$	...	$C_n = \sum_i p_{i,n}$	$\sum_i R_i = \sum_j C_j$

Źródło: opracowanie własne

Za pracą P. Normana (1999), w pierwszym kroku procedury IPF ustalane są pewne wartości wyjściowe $p_{i,j}^{(1)}$, na przykład jednakowe dla każdej komórki tabeli (elementu macierzy). Dla k -tej iteracji, $k = 1, 2, \dots$, wykonywane są kolejno następujące dwa kroki algorytmu szacowania wartości $p_{i,j}^{(k)}$, przy czym indeks górny w nawiasie oznacza kolejny numer kroku:

$$(2) \quad \begin{cases} \hat{p}_{i,j}^{(2k)} = \frac{\hat{p}_{i,j}^{(2k-1)}}{\sum_j \hat{p}_{i,j}^{(2k-1)}} \cdot R_i \\ \hat{p}_{i,j}^{(2k+1)} = \frac{\hat{p}_{i,j}^{(2k)}}{\sum_i \hat{p}_{i,j}^{(2k)}} \cdot C_j \end{cases}.$$

Oba kroki w (2) wykonywane są dopóty, dopóki nie zachodzi (z pewną założoną wcześniej dokładnością, np. do 0.1 osoby), że $\sum_j \hat{p}_{i,j} = R_i$ oraz $\sum_i \hat{p}_{i,j} = C_j$, dla wszystkich i, j .

Mimo, że metoda IPF jest czysto mechaniczna, jej zaletą jest, że nie wymaga ona dodatkowej informacji, ani dużych nakładów pracy, zaś uzyskane w jej wyniku łączne

¹⁷ Więcej o metodzie piszą Y. M. M. Bishop, E. F. Fienberg i P. W. Holland (1975).

rozkłady według wszystkich badanych cech są spójne z rozkładami brzegowymi dla poszczególnych zmiennych. Ponadto, przy pewnych niezbyt silnych założeniach, wartości uzyskane w wyniku działania IPF mogą być interpretowane z punktu widzenia statystycznego jako łączne rozkłady prawdopodobieństwa otrzymane za pomocą metody największej wiarygodności (Bishop, Fienberg i Holland, 1975; cyt. za pracą P. Normana, 1999: 2).

Z tych wszystkich względów, metoda ta wydaje się być odpowiednia dla zastosowania w celu otrzymania dokładniejszej informacji o ludności Warszawy i jej dzielnic w różnych przekrojach. Również w tym przypadku, informację wsadową mogą stanowić wyniki reprezentacyjnych badań społecznych, których przykład stanowi cytowana uprzednio praca B. Cichomskiego (2004). Sama metodyka przeprowadzania badań sondażowych została dokładniej przedyskutowana w części 3.2.

3. Propozycje wykorzystujące standardowe źródła danych i metody badawcze

Do pierwszej grupy potencjalnych możliwości szacowania rzeczywistej liczby mieszkańców Warszawy należą proste metody szacunku oparte o ogólnodostępne dane GUS, dotyczące liczebności wybranych grup mieszkańców miasta, omówione w części 3.1. Ponadto zaproponowane zostało przeprowadzenie analizy z wykorzystaniem standardowej metodologii z dziedziny nauk społecznych, jaką stanowią reprezentacyjne badania sondażowe, których kilka wariantów zostało zaprezentowane w części 3.2.

3.1. Sugestie metodyczne dla szacowania liczebności wybranych grup

3.1.1. Uwagi ogólne

Do korekty stanów ludności Warszawy można wykorzystać dostępne informacje o dwóch najbardziej mobilnych grupach ludności – studentach oraz osobach aktywnych zawodowo, bez konieczności uciekania się do kosztownych i długotrwałych badań. Takie uproszczone oszacowania nie będą brały od uwagę mobilności dzieci, młodzieży, jak również osób nieaktywnych ekonomicznie i osób w wieku starszym, ale pomimo tego mogą oferować dość dobre przybliżenia przynajmniej rzędu wielkości korekt ogólnej liczby ludności.

Za przyjęciem takiego podejścia przemawiają dwa argumenty. Po pierwsze, badania empiryczne wskazują, że największe natężenie migracji można zaobserwować wśród ludności będącej w wieku odpowiadającym okresowi studiów, wstępowania w związek małżeński oraz aktywności zawodowej, przy czym w tym ostatnim przypadku natężenie to maleje wraz z wiekiem (zob. np. Rogers i Castro, 1981 oraz przykładowe oszacowania w części 2.4, z zastrzeżeniem dotyczącym prawdopodobnie niższego natężenia nierejestrowanych migracji wśród rodzin z dziećmi). Po drugie, podejście takie może być zastosowane przy relatywnie niewielkim nakładzie kosztów i pracy. Z drugiej strony, trzeba brać pod uwagę, że uzyskane oszacowania będą cechować się niską dokładnością i brakiem możliwości dokładniejszego wyznaczenia błędów szacunku.

3.1.2. Studenci

Według danych GUS, w roku akademickim 2005/2006 w Warszawie studiowało około 277.8 tysiąca studentów, z czego 125.5 tysiąca (45.2%) na studiach dziennych¹⁸. Według tego samego źródła, *współczynnik skolaryzacji brutto*, definiowany jako „(wyrażony procentowo) stosunek wszystkich osób uczących się na danym poziomie do całej populacji (według stanu w dniu 31 grudnia) osób będących w wieku nominalnie przypisanym temu poziomowi kształcenia”¹⁹, wynosił dla szkolnictwa wyższego w całej Polsce 48.9%. Przy założeniu, że:

- powyższa liczba odnosi się do *osób*, a nie *zajętych miejsc na różnych uczelniach*, czyli uwzględnia fakt, że część osób studiuje na kilku kierunkach;
- szkolnictwu wyższemu przypisany jest nominalnie przedział wieku 19–24 lat, oraz że:
- współczynnik skolaryzacji brutto dla Warszawy jest taki sam, jak dla całej Polski²⁰,

oraz znając oficjalną liczbę ludności w wieku 19–24 lat, faktycznie zamieszkałej w Warszawie, szacowaną przez GUS na koniec 2005 roku na 152,487 osób²¹, można pokusić się o wyznaczenie przybliżonej korekty dla liczby studentów w Warszawie. Z definicji współczynnika skolaryzacji brutto, liczba faktycznie zamieszkałych studentów może być w przybliżeniu oszacowana jako $152,487 \times 0.489 \approx 74.6$ tysiąca osób. Z porównania tej wielkości oraz liczby studentów cytowanej na początku niniejszego punktu wynika, że ta grupa ludności może być niedoszacowana o przynajmniej $277.8 - 74.6 = 203.2$ tysiąca osób, nie włączając szkół resortowych podległych MON oraz MSWiA.

Powyższy, przybliżony, szacunek można następnie doprecyzować poprzez uwzględnienie dokładniejszej informacji, pod warunkiem dysponowania odpowiednimi danymi dotyczącymi współczynnika skolaryzacji brutto obliczonego dla Warszawy, oraz liczby studentów z wliczeniem uczelni MON i MSWiA. Pierwszą z tych informacji można uzyskać na przykład przeprowadzając reprezentacyjny sondaż na próbie osób w wieku 19–24 lat, zaś pozostałe – z GUS, lub ze wspomnianych resortowych szkół wyższych. Z kolei informacja niezbędna do zweryfikowania założenia dotyczącego studentów studiujących na

¹⁸ GUS (2006b: 23–24). Dane nie obejmują szkół podległych resortowi obrony narodowej (Wojskowej Akademii Technicznej i Akademii Obrony Narodowej) oraz resortowi spraw wewnętrznych i administracji (Szkoly Głównej Służby Pożarniczej).

¹⁹ Tamże.

²⁰ W dwa miesiące po ukończeniu prac nad *Studium...* otrzymaliśmy z GUS informację, że współczynnik skolaryzacji brutto dla Warszawy w 2005 roku wynosił 53.2% (wliczając w to studentów uczelni ‘mundurowych’), a więc niewiele więcej, niż przyjęta ogólnokrajowa średnia (48.9%) – przyp. aut.

²¹ Źródło: GUS, Bank Danych Regionalnych. Dane dla Warszawy dostępne pod adresem: www.stat.gov.pl/bdr/bdrap.dane_cechter.czaskatpoag?P_TERY_ID=3890, zweryfikowane dn. 14 listopada 2006 r.

więcej niż jednym kierunku znajduje się w dyspozycji dziekanatów poszczególnych uczelni, do których w tej sprawie trzeba by zwracać się indywidualnie.

Ponadto, dane o części osób studiujących znajdują się również w posiadaniu wojskowych komend uzupełnień (WKU). Niemniej, jako że są one fragmentaryczne, dotyczą bowiem jedynie młodych mężczyzn zdolnych do służby wojskowej, oraz są zbierane w WKU odpowiednich dla miejsca stałego lub czasowego zameldowania poborowych²², nie są odpowiednie do oszacowania globalnej liczby osób studiujących w Warszawie. Jako, że w większości przypadków urzędy administracji lokalnej (wójtowie, burmistrzowie, prezydenci miast) zobowiązane są do przekazania informacji o stałych lub czasowych meldunkach do WKU²³, dane z ewidencji wojskowej charakteryzują się w ogólności potencjalnie tymi samymi niedostatkami, co dane o zameldowaniach na pobyt stały lub czasowy.

Niezależnie od zastrzeżeń wymienionych powyżej, dodatkowym problemem pozostaje jednak, w jakim stopniu oszacowaną korektę należy przypisać tylko do ludności dziennej, a w jakim również do nocnej. Kwestia ta dyskutowana jest szerzej w punkcie 3.1.5.

3.1.3. Ludność aktywna zawodowo

Kompleksowy przegląd różnych rodzajów danych dotyczących zatrudnienia w Warszawie został dokonany w opracowaniu P. Śleszyńskiego (2005b), przygotowanym na zlecenie Urzędu m. st. Warszawy²⁴. W świetle tego raportu, liczba osób zatrudnionych w Warszawie podawana przez różne źródła, przy przyjęciu różnych definicji, waha się w sposób znaczny. Dolną granicę wyznacza tu liczba 681.3 tysiąca osób pracujących spośród ludności faktycznie zamieszkałej w Warszawie, podawana przez GUS na podstawie wyników spisu powszechnego z 2002 roku, bez uwzględnienia dojazdów do pracy²⁵. Górną granicę – liczba 1,660.5 tysiąca osób, która według Urzędu Statystycznego w Warszawie odzwierciedla zatrudnionych w przedsiębiorstwach, które mają główną siedzibę w Warszawie, na wiosnę 2005 roku, niezależnie od faktycznego miejsca pracy. Druga z powyższych liczb jest też, jak wskazuje autor raportu, przeszacowana co najmniej z jeszcze dwóch powodów: wliczania podmiotów gospodarczych, które zawiesiły, lub zaprzestały działalności i nie zgłosiły tego

²² Art. 53, ust. 3 ustawy z dn. 21 listopada 1967 r. *O powszechnym obowiązku obrony Rzeczypospolitej Polskiej*, Dz. U. z 1967 r., nr 44, poz. 220, tekst jednolity – Dz. U. z 2004 r., nr 241, poz. 2416, z późniejszymi zmianami, oraz rozporządzenie Rady Ministrów z dn. 1 czerwca 2004 r. *W sprawie wojskowego obowiązku meldunkowego, obowiązku powiadamiania wojskowego komendanta uzupełnień oraz uzyskiwania zezwolenia na wyjazd i pobyt za granicą*, Dz. U. z 2004 r., nr 140, poz. 1480, z późniejszymi zmianami.

²³ Tamże.

²⁴ Raport P. Śleszyńskiego zawiera między innymi dokładną analizę danych Zakładu Ubezpieczeń Społecznych, która nie została w niniejszej pracy podjęta bezpośrednio, a jedynie pośrednio, poprzez cytowane opracowanie (podziękowania pod adresem Prof. Zbigniewa Strzeleckiego za wskazanie nam tego uchybienia).

²⁵ Wszystkie informacje w tym paragrafie cytowane za pracą P. Śleszyńskiego (2005b: 4–6).

faktu do rejestrów GUS, oraz faktu pracy na więcej, niż jeden etat, zwłaszcza w niektórych sektorach, takich jak choćby edukacja, szkolnictwo wyższe i część wolnych zawodów²⁶.

Warto wspomnieć, że dane dotyczące zatrudnienia najczęściej pojawiające się w oficjalnych publikacjach GUS, takie jak 851.9 tysiąca osób zatrudnionych w sektorze przedsiębiorstw w 2005 roku²⁷, również nie do końca odpowiadają rzeczywistości. Problem stanowi na przykład nieuwzględnianie osób zatrudnionych w przedsiębiorstwach liczących poniżej dziesięciorga pracowników, wliczając w to sporą liczbę samozatrudnionych²⁸. W oparciu o dane GUS przyjęte jako punkt wyjścia, informacje z rejestrów przedsiębiorstw, oraz korekty mające na celu weryfikację liczby zatrudnionych w największych zakładach, P. Śleszyński (2005b) oszacował łączną liczbę pracujących w Warszawie w 2005 roku na 1,049,652 osób. Autor cytowanego opracowania zastrzegł jednak, że w przyszłości wielkości zastosowanych korekt powinny zostać jeszcze dokładniej zweryfikowane w oparciu o indywidualne badania dla 2.7 tysiąca przedsiębiorstw zatrudniających powyżej 50 osób, oraz reprezentacyjne badania ankietowe dla wszystkich pozostałych.

Niezależnie od cytowanego raportu, sprawdzono również możliwość użycia do oszacowania ludności rezydentnej i dziennej Warszawy danych z prowadzonego co kwartał przez GUS Badania Aktywności Ekonomicznej Ludności (BAEL)²⁹. Stosownych danych nie można uzyskać bezpośrednio z badania, gdyż nie jest możliwe uzyskanie informacji o przyczynach „wypadania” z będącego przedmiotem analizy panelu gospodarstw domowych poszczególnych elementów (gospodarstw). Teoretycznie możliwe byłoby zlecenie Głównemu Urzędowi Statystycznemu dołączenia do kwestionariuszy dystrybuowanych w Warszawie dodatkowych pytań dotyczących formalnego miejsca zamieszkania członków badanych gospodarstw domowych, lub zlecenie porównania miejsca faktycznego i rzeczywistego zameldowania, lecz w praktyce nie jest to celowe ze względu na bardzo małą i niereprezentatywną próbę gospodarstw domowych badanych w Warszawie. BAEL jest badaniem ogólnopolskim i próba dobierana jest tak, aby uzyskać reprezentatywność na poziomie kraju i województwa, a nie na poziomach niższych od wojewódzkiego. Co za tym idzie, w ramach BAEL zbierane są informacje o województwie, w którym zatrudnieni są członkowie ankietowanego gospodarstwa domowego. W związku z brakiem informacji dla niższych jednostek przestrzennych (np. powiatów) nie ma możliwości użycia tych danych do

²⁶ Tamże: 8.

²⁷ GUS (2006a: 40).

²⁸ Śleszyński (2005b: 5).

²⁹ Źródło: [1a] (→ „Źródła informacji bezpośredniej” na końcu niniejszego opracowania).

oszacowania dojazdów do pracy wewnątrz województwa, a zatem i wykorzystania tego typu informacji do przeprowadzenia szacunków dziennej lub nocnej ludności m. st. Warszawy.

Przyjmując, za obliczeniami P. Śleszyńskiego (2005b), rzeczywistą liczbę zatrudnionych w Warszawie jako 1,049,652 osób, można pokusić się o oszacowanie liczby osób pracujących spośród ludności faktycznie zamieszkałej w Warszawie dla 2005 roku, która odpowiadałaby kategoriom zdefiniowanym w NSP 2002. Bez dodatkowych informacji oszacowanie takie może mieć wyłącznie charakter orientacyjny, będąc z konieczności opartym na dość mocnych założeniach. Przykładowo, można przyjąć, że dynamika liczby osób zatrudnionych w Warszawie w okresie 2002–2005 była taka sama w przypadku trzech różnych definicji zatrudnienia: (a) spisowej, odniesionej do faktycznych mieszkańców miasta, (b) faktycznie zatrudnionych w Warszawie w przedsiębiorstwach liczących powyżej 9 pracowników, oraz (c) zatrudnionych w przedsiębiorstwach mających główną siedzibę w Warszawie i liczących powyżej 9 pracowników. Dane, oparte na zestawieniu dokonanym w cytowanym opracowaniu (Śleszyński, 2005b: 6) zaprezentowane są w tablicy 3, wraz z oszacowaniami dla 2005 roku dokonanymi przy przyjęciu założeń o proporcjonalności trzech kategorii osób zatrudnionych (a) – (c).

Tablica 3. Wstępne oszacowanie liczby zatrudnionych w Warszawie w 2005 r., wg różnych definicji (tys. osób)

Kategoria osób zatrudnionych w Warszawie	2002	2003	2004	2005
Osoby faktycznie zamieszkałe w Warszawie	681.3			648.1
Faktyczne miejsca pracy w Warszawie, > 9 zatrudnionych*	740.6	698.7		704.5
Główna siedziba w Warszawie, > 9 zatrudnionych*		836.4	834.4	843.3
Oszacowanie P. Śleszyńskiego (2005b: 22), po korektach				1,049.7

* Dokładne definicje oraz wyjątki (wyłączenia) podane są w raporcie P. Śleszyńskiego (2005b: 6)

Źródła: Śleszyński (2005b: 6, 22), GUS (2006a: 40), szary kolor oznacza obliczenia własne.

Dokładniejsze szacunki wymagałyby albo zlecenia odpowiednich prac Głównemu Urzędowi Statystycznemu (choćby poprzez uwzględnienie dodatkowych pytań przy zbieraniu informacji o zatrudnieniu, dotyczących na przykład miejsca faktycznego zamieszkania osób zatrudnionych), albo przeprowadzenia reprezentacyjnych badań społecznych wśród ludności faktycznie zamieszkałej w Warszawie z pytaniem o miejsce pracy. Pierwszy przypadek wymagałby prawdopodobnie wprowadzenia zmian w rozporządzeniach wykonawczych do ustawy o statystyce publicznej³⁰. W tym drugim przypadku trzeba wziąć pod uwagę różne zastrzeżenia metodologiczne dotyczące przeprowadzenia badań ankietowych, które przedstawione zostały w części 3.2.

³⁰ Por. np. wspomniany uprzednio w punkcie 2.2.2 Program badań statystycznych statystyki publicznej na rok 2007, Załącznik do rozporządzenia Rady Ministrów z dn. 5 września 2006 r., Dz. U. z 2006 r., nr 170, poz. 1219.

Przy przyjęciu powyższych założeń, korektę rzeczywistej liczby osób zatrudnionych w stosunku do zatrudnionych faktycznie zamieszkałych w Warszawie można wyznaczyć dla 2005 r. na podstawie danych zawartych w tablicy 3, jako różnicę: $1,049.7 - 648.1 = 401.6$ tysiąca osób. Podobnie jak w przypadku studentów, korekta ta obejmuje jednak ludność dzienną, z której pewna, bliżej nieznana część stanowi również ludność nocną. Dodatkowy problem stanowi to, że wielkości obu korekt (studentów oraz zatrudnionych) nie można po prostu zsumować, gdyż obie te kategorie mają znaczącą część wspólną, obejmującą choćby większość studentów zaocznych. Obie te trudności oraz próby ich przewyżczenia prezentowane są bliżej w punkcie 3.1.5.

3.1.4. Bezdomni

Bardzo trudną do oszacowania kategorią osób nie ujętych w oficjalnych rejestrach ludności są osoby bezdomne. Wszelkie liczby podawane w tym kontekście mają charakter czysto orientacyjny i obarczone są dużą niepewnością, przy czym można domyślać się, że z różnych względów (np. związanych z redystrybucją środków w ramach pomocy społecznej) są one z reguły zawyżane. Przykładowo, w *Raporcie o stanie Warszawy* z 2004 roku stwierdza się, że³¹:

„Na terenie Miasta st. Warszawy przebywa największa w kraju liczba bezdomnych, według różnych badań szacunkowo od 5,000 do 10,000 osób. Liczba ta jest jednak płynna i ulega zmianom w różnych porach roku. W opinii Urzędu m. st. Warszawy podawana w badaniach liczba jest jednak nieco zawyżona. [...] W 2004 roku w badanych placówkach stacjonarnych dla osób bezdomnych, które współpracują z Urzędem Miasta Stołecznego Warszawy przebywało średnio [2,452] osób. [...] W 2004 r. dofinansowanych zostało 10 jadalni, świadczących pomoc żywnościową osobom bezdomnym i ubogim. Łącznie placówki te wydawały [3,380] posiłków dziennie.” [Podkreślenie autorów. Wartości w nawiasach kwadratowych zostały dodane na podstawie tablic z dalszej części cytowanego raportu.]

Przy interpretacji powyższych liczb trzeba pamiętać, że liczba posiłków wydawanych w jadalniach jest wyższa od liczby osób bezdomnych korzystających z tej formy pomocy, gdyż z posiłków takich mogą korzystać również najubożsi mieszkańcy Warszawy.

Przypuszczenia o zawyżeniu większości szacunków liczby osób bezdomnych w Warszawie znajdują potwierdzenie również w opinii praktyka, Siostry M. Chmielewskiej (1994: 109), wyrażonej dziesięć lat wcześniej w wywiadzie dla miesięcznika *Znak*:

„Czy znana jest liczba ludzi bezdomnych w Warszawie? Jaka jest skala tego problemu? Ilu ludzi mieszka w kanałach i na dworcu?”

[...] Trudno jest określić liczbę bezdomnych. Do Warszawy ściągają bezdomni z całej Polski; atrakcyjne są wielkie dworce, a ponadto istnieje tu większa możliwość przeżycia, istnieją różne ośrodki,

³¹ *Raport o stanie m. st. Warszawy* (2004), dokument elektroniczny. Część 12: Polityka społeczna, Sekcja 12.2.5: Pomoc osobom bezdomnym. Url: um.warszawa.pl/politykaspoleczna/filez/raportpolitykaspoleczna2004.zip, stan na dzień 16 listopada 2006 r.

kuchnie darmowe itd. Wielu ludzi zrozpaczonych przyjeżdża do Warszawy, poszukując mieszkania i pracy. Podawane są bardzo różne liczby, w większości chyba zawyżone. Moim zdaniem, w tej chwili nie ma ich więcej niż – wliczając mieszkańców różnych ośrodków – około trzech tysięcy. Trzeba jednak pamiętać, że jest to już efekt kilkuletniej, rzetelnej pracy. Dzięki niej wielu ludzi już wyszło ze swojej bezdomności, ale napływają ciągle nowi. [...] Bezdomny jest wyjęty spod prawa. Nie ma praw do uzyskania meldunku czasowego, chociaż niektóre urzędy rejonowe idą nam na rękę i meldują czasowo, po to, żeby ludzie bez stałego miejsca zamieszkania mogli podjąć pracę. Przecież jeśli pracodawca chce kogoś zatrudnić, to nie zatrudni człowieka bez meldunku, bo nie ma gdzie wpłacać za tego człowieka podatku i ZUS-u. [...]” [Podkreślenia autorów]

Istotną kwestią wydają się być również wspomniane problemy z meldunkiem, które pomimo zmian prawnych w omawianym okresie³², najprawdopodobniej nie przyczyniły się znacznie do poprawy sytuacji osób bezdomnych w ciągu dekady 1994–2004.

W świetle powyższej argumentacji można oczekiwać, że liczba bezdomnych w Warszawie może wahać się sezonowo około trzech-pięciu tysięcy osób. W porównaniu z korektami liczby studentów i osób zatrudnionych, przedstawionymi w punktach 3.1.2 oraz 3.1.3, jest to więc liczba o dwa rzędy wielkości mniejsza, która nie będzie miała dużego wpływu na ostateczną wielkość oszacowania. W przypadku osób bezdomnych dość realistycznym wydaje się założenie, że większość z nich nie będzie ujęta ani w grupie studentów, ani wśród osób zatrudnionych. Choć nie można zupełnie wykluczyć osób bezdomnych jednocześnie studiujących lub pracujących, ich ogólna liczba będzie najprawdopodobniej niewielka, co pozwala na proste dodanie liczby rzędu trzech-pięciu tysięcy osób do ostatecznej korekty liczby ludności. Można też w miarę realistycznie założyć, że cała wielkość tej korekty dotyczy ludności nocnej i jednocześnie dziennej.

3.1.5. Podsumowanie i dyskusja

Z zaprezentowaną powyżej analizą wiąże się kilka problemów. Pomijając błędy szacunku dla poszczególnych grup ludności jako takich, kłopotów może nastroczać też kwestia oszacowania ich części wspólnej. Wydaje się, że dobrym źródłem informacji o tym, ilu studentów jednocześnie pracuje, mogą być na przykład wyniki badań ankietowych – albo istniejących, albo przeprowadzonych na uczelniach warszawskich specjalnie dla celu tego typu oszacowań, co z kolei wiązałoby się z kosztami i problemami natury metodologicznej (zob. część 3.2). Z drugiej strony, dodatkową zaletą badań ankietowych byłaby tu możliwość oszacowania liczby studentów studiujących na więcej, niż jednym kierunku.

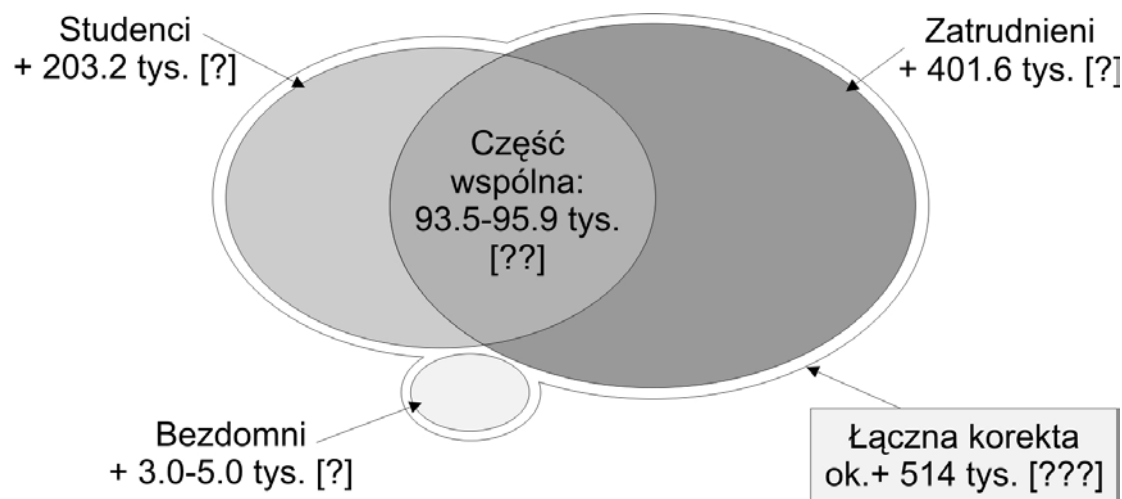
Istnieje kilka opracowań prezentujących wyniki ogólnopolskich badań ankietowych dotyczących powyższego zagadnienia. Przykładowe badania na próbach zarówno dziennych,

³² Por. dyskusję w części 2.1, dotyczącą uchylecia przez Trybunał Konstytucyjny ustawowego zapisu, który wymagał od osoby meldującej się potwierdzenia uprawnienia do przebywania w lokalu mieszkalnym.

jak i zaocznych studentów różnych uczelni w Polsce, przeprowadzone na zlecenie *Business Center Club*, jak również czasopisma *Education World*³³, wykazały zbieżne wyniki. W pierwszym z tych badań 46.0% studentów zadeklarowało się jako pracujący zarobkowo, w drugim zaś – 47.2%. Pozostaje pytanie, na ile próby ogólnopolskie odzwierciedlają zachowania studentów w Warszawie, ale odpowiedzi na nie dostarczyłoby jedynie osobne badanie ankietowe poświęcone studentom stołecznych szkół wyższych.

Dla zilustrowania przykładowego dalszego przebiegu tego typu analizy można założyć podobieństwo zachowań studentów z Warszawy i z całej Polski, oraz dodatkowo podobieństwo zachowań studentów faktycznie zamieszkałych w Warszawie i pozostałych. Otrzymujemy wówczas, że część wspólna korekty dla liczby studentów oraz osób zatrudnionych w Warszawie może wahać się w granicach około 93.5 – 95.9 tysiąca osób, co odpowiada 46.0 – 47.2% wielkości korekty dla wszystkich studentów. Przy takich (dość mocnych) założeniach, obliczona wielkość korekty liczby ludności dziennej uwzględniająca trzy omówione grupy ludności: studentów, osoby pracujące i bezdomnych, waha się między 511.9 a 516.3 tysięcy osób, czyli średnio wynosi około 514 tysięcy, jak przedstawiono na rysunku 2. Znaki zapytania przy poszczególnych liczbach przedstawionych na rysunku zwracają uwagę na bardzo niepewny charakter wymienionych oszacowań.

Rysunek 2. Korekta liczby ludności dziennej uwzględniająca studentów, osoby pracujące i bezdomnych



Źródło: opracowanie własne na podstawie danych omówionych w punktach 3.1.2 oraz 3.1.3.

Warto nadmienić, że przedstawiona powyżej metoda zawiera dużo silnych założeń, niezbędnych do przeprowadzenia obliczeń, z których każde nie musi odpowiadać

³³ Źródła: „Studenci a rynek pracy”, *Education World*, 3/2004, www.educationworld.pl/archiwum/nr3/studenci_a_rynek_pracy.htm, oraz „Studenci polscy 2002 – szczegóły stanowiska BCC”, Business Center Club, www.bcc.org.pl/stanowiska_szczegoly_archiwum.php?objectID=2602, zweryfikowane dn. 16 listopada 2006 r.

rzeczywistości. W związku z tym, że wyliczenia dla każdego z komponentów szacowanej korekty ludności z osobna charakteryzują się dużą niepewnością, tym większą niepewnością obarczona jest ich część wspólna, czy ostateczna suma. Z tego względu zaprezentowane wielkości stanowią raczej ilustrację proponowanej metodyki, a nie faktyczne wyliczenia wielkości korekt liczby ludności dziennej Warszawy. Przy praktycznym zastosowaniu tej metody, weryfikacja poszczególnych założeń może nastąpić na przykład w drodze badań ankietowych, bardziej szczegółowo dyskutowanych w części 3.2.

Osobny problem stanowi rozszacowanie wspomnianych korekt na ludność nocną i dzienną. Dysponując wynikami powyższych obliczeń dla ludności dziennej, można otrzymać ludność nocną przez odjęcie wielkości codziennych napływów studentów oraz osób pracujących do Warszawy, oszacowanych na podstawie badań natężenia ruchu, przedstawionych w części 5.1, lub zastosowania metody bilansowej, opisaną w części 5.2.

Potencjalnym mankamentem powyższej korekty łącznego stanu ludności Warszawy jest jej niekompletność. Pomimo, że rozważane są najbardziej mobilne grupy ludności (studenci i osoby zatrudnione), nie można wykluczyć, że kwestia braku meldunku w Warszawie lub codziennych dojazdów dotyczy również innych grup społecznych: osób w wieku poprodukcyjnym, dzieci, uczniów, itp. Problem polega na tym, że tego typu grup może być potencjalnie dużo, zaś informacja o wielkości potencjalnych korekt nie musi być dostępna na podstawie oficjalnych danych GUS.

Jedną ze statystycznych metod szacowania wielkości nieznanej populacji w oparciu o informację o pochodzących z niej dwóch próbach oraz ich części wspólnej, jest tak zwana metoda analizy dualnych systemów danych (*dual data systems analysis*), zwana też popularnie metodą *capture-recapture* („złap i złap ponownie”)³⁴. Technika ta, opisywana tu za pracą J. M. Alho (1990), w swej najprostszej postaci zakłada, że dysponujemy dwiema niezależnymi próbami losowymi (wielkości n_1 i n_2 obserwacji) z interesującej nas zamkniętej populacji o nieznanej liczebności N , i że znamy wielkość ich części wspólnej, m . Zakładając ponadto, że poszczególne osoby znalazły się w obu próbach niezależnie od siebie, ale z jednakowym prawdopodobieństwem, estymator oszacowania wielkości N , N_{est} , oraz estymator jego wariancji, $\text{Var}(N_{\text{est}})$ otrzymujemy z następujących wzorów (Alho, 1990: 624):

$$(3) \quad N_{\text{est}} = n_1 \cdot n_2 / m; \quad \text{Var}(N_{\text{est}}) = n_1 \cdot n_2 \cdot (n_1 - m) \cdot (n_2 - m) / m^3.$$

³⁴ Metoda stosowana jest głównie do szacowania populacji żyjących na danym obszarze dzikich zwierząt, spośród których losowo łapie się pewną grupę osobników, znakuje i puszcza wolno. Później gromadzi się drugą próbę zwierząt, po czym zlicza się, ile spośród nich jest już oznakowanych, czyli zostało złapane w obu próbach – przyp. aut.

Gdyby potraktować studentów i osoby zatrudnione jako dwie niezależne próby z populacji, zdefiniowanej jako dzienna ludność Warszawy nie ujęta w statystykach GUS, można na podstawie informacji przedstawionej na rysunku 2 pokusić się o oszacowanie wielkości nieznanego, brakującego elementu korekty. Przy założeniu, że $n_1 = 203.2$ tysiąca, $n_2 = 401.6$ tysiąca oraz $m =$ średnio 94.7 tysiąca osób, otrzymujemy ze wzoru (3) oszacowanie N_{est} w wysokości 861.7 tysiąca osób, z odchyleniem standardowym wynoszącym $\sqrt{\text{Var}(N_{est})} = 1.8$ tysiąca osób. Wynikałoby stąd, że oprócz wyliczonej wyżej korekty ludności dziennej o ok. 510 tysięcy studentów i osób pracujących (bez uwzględniania bezdomnych), należy liczyć się z dodatkową korektą w przybliżonej wysokości $861.7 - 510.0 \approx 351.7$ tysiąca osób.

W świetle wiedzy o tym, że dane użyte do szacunku dotyczą najbardziej mobilnych grup ludności, oszacowanie korekty dla grup mniej mobilnych wynoszące ponad 350 tysięcy osób wydaje się jednak bardzo zawyżone. Wynika to z faktu, że głównym mankamentem zastosowania powyższej metody w kontekście szacowania liczby ludności Warszawy są jej bardzo silne założenia, które niekoniecznie muszą być prawdziwe. Po pierwsze, „próby” (czyli zbiorowości studentów i osób pracujących) nie zostały dobrane w sposób losowy, ale celowy. Po drugie, fakt bycia studentem nie musi być niezależny od faktu wykonywania pracy zarobkowej. Po trzecie, grupa nieznaną i podlegającą szacowaniu najprawdopodobniej jest specyficzna i może pozostawać w rozmaitych zależnościach zwłaszcza z osobami pracującymi (problem dotyczy dzieci, młodzieży, oraz osób nieaktywnych ekonomicznie i osób w wieku starszym). Na pewno nie można więc bezpiecznie przyjąć założenia o wewnętrznej jednorodności całej szacowanej populacji.

W związku z powyższymi zastrzeżeniami, należy więc zachować ostrożność przy stosowaniu narzędzi typu *capture-recapture* do takich celów, jak oszacowanie rzeczywistej liczby ludności Warszawy. Dodatkowo, aby osłabić zaprezentowane wyżej założenia, trzeba by zastosować różne modyfikacje metody, jak choćby zaproponowaną przez J. M. Alho (1990), dotyczącą uwzględnienia wewnętrznego zróżnicowania (*heterogeneity*) badanych populacji. Proponowane metody opierają się jednak na znacznie bardziej skomplikowanych modelach matematycznych, niż zaprezentowane powyżej. Z drugiej strony, metodę *capture-recapture* w czystej postaci można by z powodzeniem zastosować do wyników podwójnych reprezentacyjnych badań społecznych, omawianych w kolejnej części (3.2). Niemniej jednak, ze względu na koszty takiego badania i problemy z kompletnością odpowiedzi dla obu prób, najprawdopodobniej analiza tego typu byłaby nieopłacalna, a potencjalne zyski informacyjne z jej przeprowadzenia – niewielkie.

3.2. Korekty oparte na wynikach reprezentacyjnych badań społecznych

Odrębnym sposobem wyznaczania nieznanej wielkości korekty liczby mieszkańców m. st. Warszawy mogą być szacunki oparte o wyniki reprezentacyjnych badań społecznych. W niniejszej części przedstawiono zarys metodyczny dwóch podstawowych typów takich badań, jakimi są bezpośrednie badania ankietowe „*face-to-face*” (punkt 3.2.1) oraz badania telefoniczne typu CATI (punkt 3.2.2), wraz z omówieniem problemów i ograniczeń związanych z przeprowadzeniem badania reprezentacyjnego.

Istotą reprezentacyjnych badań społecznych jest taki dobór próby losowej i takie przeprowadzenie badania, by uzyskane wyniki były maksymalnie zbliżone do wyników, które otrzymalibyśmy, gdyby przebadać całą populację. W interesującym nas przypadku, badaną populacją jest cała rzeczywista populacja Warszawy, w której pewną frakcję (q) stanowią zameldowani mieszkańcy (uchwyceni w statystykach GUS), a pewien, nieznany odsetek ($p = 1 - q$) – osoby niezameldowane w stolicy ani czasowo, ani na stałe. Celem badania byłoby ustalenie udziału procentowego p tej drugiej grupy w populacji stolicy, co pozwoli na wnioskowanie o rzeczywistej populacji Warszawy.

3.2.1. Badanie ankietowe *face-to-face* w terenie – zarys metody

Pierwszym ze sposobów uzyskania odpowiedzi na powyższe pytanie, mogą być terenowe badania ankietowe typu „*face-to-face*” (bezpośrednie), zakładające przeprowadzenie sondażu i zadanie pytań przez ankietera odwiedzającego lokal mieszkalny. Spośród kwestii związanych z wyłonieniem próby reprezentatywnej dla takich badań, na czoło wysuwa się kwestia doboru próby i wyboru właściwego operatu losowania. Nie można tu zastosować losowania z bazy PESEL (taką metodę zastosowano m.in. w badaniu kwestionariuszowym gospodarstw domowych w ramach Warszawskiego Badania Ruchu), gdyż dobór ten zakładałby losowanie jedynie spośród osób zameldowanych, a więc ujętych już przez statystyki GUS w kategoriach ludności stałej oraz faktycznie zamieszkałej³⁵.

Alternatywną, wobec PESEL, bazą do losowania próby, mogłaby być baza adresów w ramach Krajowego Rejestru Urzędowego Podziału Terytorialnego Kraju (TERYT), prowadzona przez Główny Urząd Statystyczny. Rejestr TERYT obejmuje m.in. adresy budynków i mieszkań w układzie rejonów statystycznych i obwodów spisowych, które mogą służyć do losowania prób i ustalenia adresów mieszkań celem badań reprezentacyjnych³⁶. W przypadku wyboru rejestru TERYT do losowania próby, konieczne byłyby dodatkowe

³⁵ Por. np.: Kupiszewska, Nowok i Kupiszewski (2006: 577–578).

³⁶ Url: www.stat.gov.pl/bip/rejestry/teryt/zakres_teryt.htm.

założenia, dotyczące tego, która z osób mieszkających w lokalu mieszkalnym będzie poproszona o udzielenie informacji. Przy takiej konstrukcji badania, z wykorzystaniem bazy adresowej jako operatu doboru próby, niemożliwy jest podział próby na tzw. warstwy na przykład ze względu na wiek, płeć bądź inne cechy społeczno-demograficzne, a jedynie ze względu na miejsce zamieszkania (dzielnice Warszawy).

Do oszacowania nieznanej wartości frakcji osób niezameldowanych p , wykorzystany zostałby estymator tej wielkości, $\hat{p}$, równy wielkości tego odsetka otrzymanej w próbie, czyli ilorazowi liczby zidentyfikowanych osób niezameldowanych y przez liczebność próby n ($\hat{p} = y / n$). Przy założeniu najprostszego schematu losowania prostego nieograniczonego, błąd oszacowania estymatora $\hat{p}$, $\text{Err}(\hat{p})$, definiowany jest jako połowa długości 95-procentowego przedziału ufności dla $\hat{p}$ (np.: J. Jóźwiak oraz J. Podgórski, 1997: 283):

$$(4a) \quad \text{Err}(\hat{p}) = u_{0.05} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n - 1}} \cdot \sqrt{\frac{N - n}{N}},$$

gdzie N oznacza wielkość populacji (tu: łączną liczbę ludności Warszawy, obejmującą zarówno osoby zameldowane, jak i nie), n – wielkość próby, zaś $u_{0.05}$ – kwantyl rzędu 0.975 z rozkładu normalnego. Warto zauważyć, że jeśli cała populacja jest o kilka rzędów wielkości większa od próby, wówczas czynnik $\sqrt{(N - n)/N}$ nie odgrywa większego znaczenia, jest bowiem bliski jedności.

Ze wzoru (4a) wynika, że próba losowa dla nieznanego rzędu wielkości badanej frakcji p powinna obejmować powyżej tysiąca respondentów, by błąd $\text{Err}(\hat{p})$ wynosił nie więcej, niż 3%. Jeśli istnieją dokładniejsze informacje przynajmniej co do rzędu wielkości p , przy czym wiadomo, że p nie ma wartości bliskiej 0.5, wówczas liczebność próby może być mniejsza. Przykładowo, zakładając w sposób bardzo uproszczony, że dla Warszawy $N \approx 1.93$ miliona osób, zaś $p \approx 0.135$ (podstawy dla przyjęcia tych założeń przedstawiono w części 6.1, zakłada się też homogeniczność całej Warszawy ze względu na p), wówczas trzyprocentowy błąd szacunku jest zagwarantowany już w przypadku próby 499-elementowej. Jest to o tyle ważne, że dobranie tysiącosobowej próby reprezentatywnej i przeprowadzenie na jej podstawie badania sondażowego wiąże się niewątpliwie z wysokimi kosztami.

W przypadku rejestrów administracyjnych, takich jak PESEL i TERYT, losowanie proste nie jest jednak rozwiązaniem optymalnym. Znacznie efektywniejsze estymatory (obarczone mniejszymi błędami szacunku) można uzyskać stosując losowanie zespołowe, gdzie zespoły np. mogłyby być złożone z mieszkań, lub też losowanie warstwowe, gdzie

warstwami byłyby dzielnice lub inne jednostki urbanistyczne. Ze względu na pożądane własności estymatorów, preferowanym rozwiązaniem byłoby losowanie warstwowe³⁷, w którym błąd szacunku dany jest wzorem (np.: J. Jóźwiak oraz J. Podgórski, 1997):

$$(4b) \quad \text{Err}(\hat{p}) = u_{0.05} \cdot \sqrt{\sum_{k=1}^K \frac{N_k^2}{N^2} \cdot \frac{\hat{p}_k \cdot (1 - \hat{p}_k)}{(n_k - 1)} \cdot \frac{N_k - n_k}{N_k}},$$

gdzie indeks k odnosi się odpowiednio do wielkości populacji N , liczebności próby n oraz estymatora frakcji p dla poszczególnych warstw spośród K możliwych. Warto nadmienić, że zmiana schematu losowania na losowanie warstwowe zapewniłaby reprezentatywność próby na poziomie poszczególnych dzielnic oraz pozwoliła na przyjęcie bardziej realistycznych założeń co do wielkości próby. W przypadku uwzględnienia dokładniejszego podziału administracyjnego, należałoby, w miarę możliwości, przyjąć zróżnicowane oszacowania wstępne wartości p dla poszczególnych części miasta.

Oprócz kwestii związanych z doбором próby, konieczne jest też rozwiązanie problemów związanych z samym przeprowadzeniem badania, dotyczących konstrukcji pytań kwestionariusza ankietowego i ich potencjalnej drażliwości społecznej oraz liczby (stopy) odmów odpowiedzi. W związku z tym, że chcemy otrzymać informację, ile osób faktycznie zamieszkuje dany lokal, a ile jest w nim zameldowanych, respondenci mogą obawiać się udzielenia prawdziwej odpowiedzi w przypadku zbyt bezpośredniego lub nieodpowiedniego sformułowania pytań ankiety.

Kolejnym potencjalnym problemem może być niska stopa zwrotu, czyli odsetek liczb wypełnionych ankiet w stosunku do liczby odmów udzielenia odpowiedzi, przerwanych wywiadów, lub przypadków niemożności zastania mieszkańców pod wylosowanym adresem. Ogólnie rzecz biorąc stopa zwrotu w badaniach społecznych sondażowych w Polsce wynosi średnio około 60% (Kołakowska, 2006), ale jest niższa w wielkich miastach ze względu na takie czynniki jak upowszechnienie domofonów, dłuższe godziny pracy i większa niechęć do poświęcenia czasu na ankietę. Z kolei takie problemy, jak strach przed wpuszczeniem obcej osoby do domu, dyskomfort odpowiadania na pytanie o zameldowanie w lokalu, w którym się przebywa i w którym trwa ankietę, mogą przyczynić się do dodatkowego obniżenia stopy zwrotu w przypadku badania ankietowego *face-to-face*, a w efekcie pod znakiem zapytania postawić kwestie właściwej interpretacji wyników.

³⁷ Serdeczne podziękowania dla Prof. Jana Paradysza za zwrócenie nam uwagi na tę kwestię oraz za sugestie rozwiązań. Szczegółowe informacje dotyczące teorii metody reprezentacyjnej oraz różnych schematów losowania zawierają monografie Cz. Brachy (1996), oraz J. Jóźwiak i J. Podgórskiego (1997: 277–302).

Problemy z uzyskaniem właściwej stopy zwrotu uwidoczniły się na przykład w trakcie badania ankietowego gospodarstw domowych w ramach Warszawskiego Badania Ruchu, kiedy to stopa zwrotu w próbie losowanej z rejestru PESEL wyniosła 55%, a pozostałe 45% odpowiedzi trzeba było wygenerować metodą dobierania respondentów przez ankieterów na zasadzie *random route*, czyli losowego doboru w terenie (BPRW, 2005: 9). Alternatywą może być założenie wyjściowej wielkości próby większej, niż wynikałoby to ze wzoru (4), o czynnik wynikający z uwzględnienia spodziewanej stopy zwrotu. Żaden ze wspomnianych sposobów radzenia sobie z niekompletnymi odpowiedziami nie gwarantuje jednak pełnej reprezentatywności ostatecznej próby, co stanowi nierozwiązywalny problem cechujący dowolne badania ankietowe.

Potencjalnym problemem związanym z badaniami ankietowymi jest również ich wysoki koszt, związany przede wszystkim z kosztami pracy ankieterów. Przeprowadzenie takich badań na tysiącosobowej próbie, jedynie w celu zbadania procentowego udziału osób niezameldowanych oraz udziału osób zameldowanych a nie zamieszkałych w Warszawie mogłoby być zbyt drogie, porównując koszty do uzyskanej wiedzy.

Bardziej racjonalnym rozwiązaniem, z punktu widzenia stosunku nakładów finansowych do wyników badań, mogłoby być przyłączenia badania stanu ludności do innych badań socjologicznych prowadzonych na populacji Warszawy. Wówczas pytania o kwestię zameldowania i zamieszkania zostałyby zadane dodatkowo obok innych pytań prowadzonego badania. Zaletą takiego rozwiązania, oprócz obniżenia kosztów, byłoby „ukrycie” potencjalnie drażliwych pytań o zameldowanie wśród szeregu innych, przypuszczalnie mniej drażliwych zagadnień poruszanych przez ankietera. Założenia odnośnie operatu doboru próby omówione powyżej musiałyby być spełnione również w przypadku dołączenia pytań o zameldowanie / zamieszkanie do szerszych badań socjologicznych w Warszawie.

3.2.2. Badanie CATI – zarys metody.

Alternatywną techniką przeprowadzenia reprezentacyjnych badań społecznych, są tzw. badania typu CATI (*Computer Assisted Telephone Interview*). Badania takie realizuje się przez ankietę telefoniczną wśród abonentów losowo wybranych numerów telefonów. Operat doboru próby może tu stanowić spis telefonów gospodarstw domowych w Warszawie (obejmujący wszystkich numery wszystkich dostawców usług telefonicznych), spośród którego losuje się określoną ilość numerów celem wykonania sondażu. Alternatywnym sposobem doboru próby jest w przypadku takich badań komputerowe generowanie numerów przy wykorzystaniu algorytmów pseudolosowych. Zakłada się wówczas, że losowanie

przeprowadzane jest do skutku, czyli do osiągnięcia kontaktu z na przykład tysiąc osobową próbą lokali mieszkalnych, przy zignorowaniu wygenerowanych numerów telefonicznych do instytucji, firm oraz numerów nieistniejących.

Dużą zaletą badania telefonicznego jest jego relatywnie niski koszt w porównaniu do badań ankietowych prowadzonych w terenie przez ankieterów oraz możliwość szybkiego przeprowadzenia sondażu. Według informacji uzyskanych w jednej z wiodących na rynku firm badawczych, prowadzącej między innymi sondaże typu CATI, zlecenie przeprowadzenia badania na tysiąc osobowej próbie respondentów w Warszawie, przy założeniu, że ankieta zawierałaby do dziesięciu pytań, byłoby możliwe do zrealizowania w ciągu dwóch dni.

Drugą zaletą badań CATI, oprócz niższych kosztów, jest większa anonimowość, którą zapewnia badanie metodą telefoniczną. Mogłaby ona przyczynić się do wzrostu stopy zwrotu, czyli wzrostu liczby osób, które zgodzą się na udzielenie odpowiedzi, oraz do wzrostu prawdopodobności, czyli wzrostu liczby odpowiedzi odpowiadających rzeczywistości.

Metoda CATI ma oczywiście również ograniczenia. Wyłączona z badania jest grupa lokali mieszkalnych, w których nie jest zainstalowana żadna stacjonarna linia telefoniczna. Trudno natomiast o jakiegokolwiek domniemania, czy osoby niezameldowane lub zameldowane a nie zamieszkałe, są w równomierny sposób rozmieszczone w lokalach mieszkalnych z linią telefoniczną oraz w lokalach bez takiej linii.

Podsumowując omówienie wad i zalet obu przedstawionych metod wykorzystania reprezentacyjnych badań społecznych do szacowania rzeczywistej liczby ludności Warszawy, celowe wydaje się zarekomendowanie przeprowadzenia takich badań jako uzupełnienia bądź potwierdzenia szacunków dokonanych na podstawie danych ze standardowych i niestandardowych źródeł informacji statystycznej. Na bazie dokonanego porównania techniki przeprowadzenia sondażu w terenie metodą „*face-to-face*” oraz metody CATI, biorąc również pod uwagę kwestię kosztów takiego badania, racjonalniejszą dla osiągnięcia postawionych celów badania wydaje się w omawianym przypadku metoda CATI.

4. Możliwości szacunków opartych na niestandardowej informacji statystycznej

W niniejszym rozdziale przedstawiony został przegląd niestandardowych źródeł danych statystycznych pochodzących z wybranych instytucji. W części 4.1 analizie pod kątem zawartości, dostępności, a wreszcie użyteczności dla oszacowań liczby ludności Warszawy poddane zostały dane: Ministerstwa Finansów (Urzędów Skarbowych), Mazowieckiego Oddziału Wojewódzkiego Narodowego Funduszu Zdrowia, rejestry Państwowej Komisji Wyborczej, oraz statystyki kościelne. Ponadto, zbadano możliwości ewentualnych szacunków w oparciu o dane Biura Administracji i Spraw Obywatelskich m. st. Warszawy dotyczące dodatkowych zameldowań w tzw. *strefie I* w czasie Europejskiego Szczytu Gospodarczego w kwietniu 2004 r. W części 4.2 dokonano przeglądu danych o gospodarce komunalnej w Warszawie pod kątem możliwości ich wykorzystania. Rozdział kończy część 4.3, w której przedstawione zostały i poddane analizie techniki szacowania liczby ludności w oparciu o informacje uzyskane ze zdjęć satelitarnych i lotniczych.

4.1. Dostępność danych statystycznych w wybranych instytucjach

4.1.1. Ministerstwo Finansów – dane z rejestrów podatkowych

Za potencjalne źródło danych o osobach zamieszkałych, a nie zameldowanych w Warszawie mogłaby posłużyć informacja gromadzona przez Ministerstwo Finansów, a wynikająca z danych dostarczanych Ministerstwu przez system rozliczeń podatkowych obywateli. Osoby fizyczne składając rozliczenie roczne (PIT-36 lub PIT-37) dostarczają do Urzędów Skarbowych podstawowe dane identyfikacyjne, w tym adres zamieszkania. Broszury informacyjne udostępniane przez Ministerstwo Finansów, nie zawierają żadnych sugestii odnośnie tego, jak definiować adres zamieszkania na w/w formularzach.

Z drugiej strony, formularze NIP-1 i NIP-3, składane w celu wprowadzenia bądź aktualizacji danych osobowych podatnika, jeśli te ulegają zmianie w przeciągu poprzedniego roku podatkowego, zawierają oddzielne rubryki dla adresu miejsca zamieszkania oraz adresu miejsca zameldowania (stałego lub czasowego, jeśli brak stałego), w przypadku, gdy adres zamieszkania nie pokrywa się z adresem zameldowania. Pozwala to domniemywać, że termin „adres zamieszkania” występujący w formularzach PIT-36 i PIT-37 rozumiany jest przez

Urząd Skarbowy jako miejsce faktycznego zamieszkiwania, niekoniecznie pokrywające się z miejscem zameldowania.

Analiza danych zgromadzonych przez Ministerstwo Finansów dla całego kraju mogłaby więc pozwolić na ustalenie liczby osób, które jako miejsce zamieszkania wskazują Warszawę, niezależnie od miejsca zameldowania. Mogłoby to pozwolić na ustalenie łącznej liczby podatników mieszkających w Warszawie. Informacja ta pozwoliłaby z kolei na dokonanie korekty oficjalnych statystyk dostarczanych przez GUS, np. dotyczących liczby osób pracujących wśród ludności faktycznie zamieszkałej, o podatników zamieszkałych w Warszawie, a zameldowanych w innych miejscowościach. Korekta ta pozwoliłaby na otrzymanie oszacowania liczby pracujących niezależnej od statystyk zatrudnienia zbieranych przez Główny Urząd Statystyczny oraz weryfikację szacunków P. Śleszyńskiego (2005b).

Wydaje się celowe przeprowadzenie tego typu porównania i analizy danych, niemniej jednak ustalenia wymaga jeszcze dokładna dostępność danych gromadzonych przez Ministerstwo Finansów na potrzeby badania. Wprawdzie z odpowiedzi Ministerstwa wynika, że nie sporządza ono statystyk uwzględniających miejsce zamieszkania podatnika³⁸, niemniej jednak z informacji uzyskanych w Departamencie Eksploatacji Systemów Informatycznych wynika³⁹, że wprawdzie dane z formularzy PIT-36 i PIT-37 są digitalizowane i przetwarzane jedynie w odniesieniu do kwot (dochodów, podatków) a nie innych danych o podatniku, natomiast digitalizowane i możliwe do przeszukania są dane z formularzy NIP-3 i NIP-1, zawierające również informacje zarówno o miejscu zamieszkania, jak i miejscu zameldowania. Istnieje więc potencjalnie możliwość wykorzystania danych Ministerstwa Finansów dla celu szacowania rzeczywistej liczby ludności Warszawy.

4.1.2. Statystyka ubezpieczonych w Narodowym Funduszu Zdrowia

Z informacji uzyskanych w Mazowieckim Oddziale Wojewódzkim Narodowego Funduszu Zdrowia wynika, że NFZ na bieżąco gromadzi i aktualizuje dane o pacjentach przypisanych do lekarzy pierwszego kontaktu według numerów PESEL świadczeniobiorców (pacjentów)⁴⁰. Deklaracje składane przez świadczeniobiorców celem wyboru lekarza podstawowej opieki zdrowotnej (POZ), czyli tzw. lekarza pierwszego kontaktu, zawierają obok numeru PESEL, także dane o adresie zameldowania i adresie zamieszkania osoby

³⁸ Źródło: [3b] (→ „Źródła informacji bezpośredniej”), pismo DD1-016-495/GK/06/1171 z 23 listopada 2006 r.

³⁹ Źródło: [3a] (→ „Źródła informacji bezpośredniej”).

⁴⁰ Źródło: [4] (→ „Źródła informacji bezpośredniej”).

ubezpieczonej. Ponadto, Centralny Wykaz Ubezpieczonych zawiera również dane o adresie do korespondencji pacjentów.

Według informacji uzyskanych w Mazowieckim Oddziale Wojewódzkim NFZ, tworzeniem bazy danych na podstawie informacji zawartych w deklaracjach pacjentów i przetwarzaniem informacji zawartych w bazie zajmuje się firma informatyczna CAMSOFT. Uzyskanie informacji statystycznych opartych o dane zawarte w deklaracjach pacjentów jest możliwe tylko po zleceniu dostawcy usług informatycznych (CAMSOFT-owi) przetwarzania danych pod kątem uzyskania odpowiedzi na poszczególne pytania. Mazowiecki Oddział Wojewódzki NFZ przetwarza dane tylko w zakresie ewidencji numerów PESEL osób zgłaszających się do lekarzy podstawowej opieki zdrowotnej i w zakresie wychwycenia osób zameldowanych poza województwem mazowieckim ze względu na inny system rozliczeń za usługi medyczne w takich przypadkach.

Jedną z wartych rozważenia możliwości szacowania rzeczywistej liczby ludności Warszawy, zważywszy na zakres danych gromadzonych przez NFZ, jest przeprowadzenie analizy danych osób zarejestrowanych do lekarza podstawowej opieki zdrowotnej (POZ) w Warszawie, a zameldowanych poza Warszawą. Na podstawie takiej analizy możliwe do uzyskania są dwie istotne informacje. Po pierwsze, można dowiedzieć się jaka jest liczba osób korzystających z POZ w Warszawie, zameldowanych poza Warszawą, a podających jako adres zamieszkania adres warszawski. Dane takie stanowiłyby jedno z możliwych źródeł dla szacunku ludności nocnej Warszawy. Po drugie, możliwa do uzyskania byłaby też informacja o osobach zameldowanych i zamieszkałych poza Warszawą, a korzystających z POZ w Warszawie z powodu zatrudnienia bądź nauki w stolicy. Te informacje są niewątpliwie mniej wartościowe z punktu widzenia szacunków ludnościowych, gdyż dotyczyłyby nieznannej części osób dojeżdżających do Warszawy (ludności dziennej stolicy), które zdecydowały się na zapisanie do lekarza pierwszego kontaktu w miejscu pracy, a nie w miejscu zamieszkania.

Obydwie powyższe informacje nie są możliwe do uzyskania na obecnym etapie studium nawet w przybliżeniu ze względu na brak gotowych przetworzonych danych będących w dyspozycji NFZ. Wydaje się jednak, że w badaniu szacującym rzeczywistą ludność stolicy warto będzie sięgnąć do danych NFZ w zakresie porównania liczby osób deklarujących miejsce zamieszkania w Warszawie, a miejsce zameldowania poza nią, dla uzupełnienia i porównania z innymi szacunkami ludności nocnej stolicy.

4.1.3. Rejestry Państwowej Komisji Wyborczej (PKW)⁴¹

Informacja o osobach mogących brać udział w głosowaniach, z definicji ograniczona do osób posiadających czynne prawo wyborcze, czyli będących w wieku powyżej 18 lat i nie pozbawionych tego prawa na mocy wyroku sądu, zawarta jest w rejestrze i spisie wyborców. Analiza przepisów ustaw w zakresie dotyczącym rejestru i spisu wyborców wskazuje na dwie możliwości wpisania do spisu osób, które nie są na stałe zameldowane w danej gminie, a w rzeczywistości w niej zamieszkują i chcą wziąć udział w głosowaniu⁴²:

- 1) wpisanie się do spisu wyborców na podstawie oświadczenia, a następnie wykreślenie w miejscu stałego zameldowania (o ile posiada się stałe zameldowanie, co stanowi poważny problem w przypadku osób bezdomnych), dokonywane administracyjnie bez udziału zainteresowanych;
- 2) wykreślenie ze spisu wyborców przez zainteresowanego i pobranie zaświadczenia o prawie głosowaniu w miejscu pobytu poza miejscem stałego zameldowania.

Pierwsza z metod wymaga uregulowania sprawy na 10 dni przed terminem wyborów, a druga w praktyce umożliwia pobranie zaświadczenia na dzień przed wyborami (termin nie jest określony w ustawie). Z drugiej strony, pierwszy sposób jest dużo łatwiejszy – nie wymaga udania się do urzędu gminy w stałym miejscu zamieszkania, często odległego od miejsca pobytu. Możliwe jest uzyskanie zaświadczenia o prawie głosowania w miejscu pobytu na przykład drogą korespondencyjną, lub poprzez upoważnienie do odebrania przez członka rodziny. Drugi przypadek przewidziany jest na wypadki losowe – krótkotrwałe wyjazdy służbowe, rodzinne, turystyczne, itp. W praktyce osoby zamieszkałe poza miejscem stałego zameldowania mogą korzystać z obu sposobów.

Na podstawie danych Krajowego Biura Wyborczego możliwe jest ustalenie liczby osób, które: (a) same zgłosiły się do urzędu, mając zamiar wzięcia udziału w głosowaniu w miejscu innym niż stałego zameldowania, (b) zostały wykreślane wskutek wpisania w innym spisie wyborców lub utraty praw publicznych (np. w rezultacie wyroku sądowego), oraz (c) pobrały karty do głosowania na podstawie zaświadczenia o prawie głosowania w miejscu pobytu. Wprawdzie daty sporządzania bilansów ludnościowych (koniec czerwca i grudnia) i terminy wyborów z reguły się nie pokrywają, ale w przypadku Warszawy nie stanowi to dużego problemu, gdyż liczba osób zameldowanych w ciągu roku zmienia się w niewielkim stopniu w stosunku do ogólnej liczby mieszkańców (zmiany rzędu kilku promili).

⁴¹ Część autorstwa dra Przemysława Śleszyńskiego z Instytutu Geografii i Przestrzennego Zagospodarowania PAN, uzupełniona przez pozostałych autorów opracowania.

⁴² Zob. np.: ustawa z dnia 12 kwietnia 2001 r. Ordynacja wyborcza do Sejmu Rzeczypospolitej Polskiej i do Senatu Rzeczypospolitej Polskiej, Dz. U. z 2001 r., nr 46, poz. 499, z późniejszymi zmianami.

Spisy wyborców w Warszawie prowadzone są przez dzielnicowe Wydziały Obsługi Mieszkańców (WOM). Zestawienie liczbowe podstawowych kategorii ludności można uzyskać w Krajowym Biurze Wyborczym, z możliwością dezagregacji według 18 dzielnic. Teoretycznie, nawet przy podstawowym zakresie danych, można więc próbować oszacować liczbę osób rzeczywiście przebywających w Warszawie, potwierdzoną przez uczestnictwo w wyborach. Liczba dodatkowych osób ujawniona w spisach wyborców podzielona przez wskaźnik frekwencji dałaby w wyniku szacowaną liczbę ludności (przykładowo, 20 tysięcy dodatkowych wyborców przy 62% frekwencji odpowiadałoby $20,000 / 0,62 = 32,000$ tysiącom osób w grupie wieku powyżej 18 lat, obejmującej osoby najbardziej mobilne⁴³). Uzyskanie bardziej szczegółowych informacji też byłoby możliwe, wymagałoby jednak dodatkowej pracy Komisji Wyborczych przy kolejnych wyborach⁴⁴.

Praktyczne stosowanie powyższej metody uniemożliwia jednak kwestia ustalenia faktycznej frekwencji wśród osób dopisujących się do spisów wyborców. Udział ludności w wyborach powszechnych jest zróżnicowany i zależy od szeregu czynników. Największym zainteresowaniem cieszą się z reguły wybory prezydenckie, a następnie parlamentarne. Frekwencja charakteryzuje się również silnym zróżnicowaniem w zależności od cech demograficznych (wiek, płeć) i społecznych (wykształcenie, zawód, miejsce zamieszkania). Wydaje się, że najprawdopodobniej frekwencja wśród osób dopisanych do spisu wyborców lub pobierających zaświadczenia może być niższa, niż średnia dla Warszawy (choć nie wiadomo, w jakim stopniu), gdyż fakt dopisania się do spisu lub uzyskania zaświadczenia wymaga dodatkowego wysiłku i nakładu czasu. Niestety, nie widać w tym przypadku żadnej możliwości oszacowania frekwencji wśród osób dopisanych do spisu, w porównaniu do ludności stale zameldowanej. Mimo braku przeszkód prawnych do zorganizowania np. badań ankietowych w wybranych obwodach wyborczych⁴⁵, trudno sobie wyobrazić jakąkolwiek racjonalną metodologię tego typu analizy.

Kolejną, bardzo poważną przeszkodą w stosowaniu danych Państwowej Komisji Wyborczej jest niekompletność rejestrów, na podstawie których tworzone są spisy wyborców. Przykładowo, w wyborach parlamentarnych z 25 września 2005 roku w rejestrach ludności wykorzystywanych przez PKW znajdowało się 1,575,958 mieszkańców Warszawy, z czego 1,343,728 osób mających prawo wyborcze⁴⁶. Ogólna liczba ludności jest więc według tego

⁴³ Por. dyskusja w części 3.1.

⁴⁴ Źródło: [5] (→ „Źródła informacji bezpośredniej”).

⁴⁵ Jak wyżej.

⁴⁶ Źródło: strona internetowa PKW (www.wybory2005.pkw.gov.pl/SJM/PL/KOMISJE/140000/146501.html), dane zweryfikowane dn. 14 listopada 2006 r.

źródła o ponad 121 tysięcy osób (7%) mniejsza, niż liczba ludności faktycznie zamieszkałej wykazywanej przez GUS dla analogicznego okresu (1,697,214)⁴⁷. Oba wymienione problemy: niekompletność rejestrów oraz trudności w oszacowaniu frekwencji dla różnych grup wyborców, powodują więc, że dane PKW nie wydają się odpowiednim źródłem informacji dla szacowania stanów ludności Warszawy.

4.1.4. Korekty oparte na szacunkach Biura Administracji i Spraw Obywatelskich m. st. Warszawy

Jako jedna z propozycji metodycznych szacowania liczby osób niezameldowanych, a zamieszkałych w Warszawie, rozważana była analiza oparta na informacji o dodatkowych meldunkach w dzielnicy Śródmieście w czasie Europejskiego Szczytu Gospodarczego, który odbywał się w Warszawie w dniach 28–30 kwietnia 2004 r. Ze względów bezpieczeństwa, w tych dniach wstęp do kilku kwartałów Śródmieścia, bezpośrednio otaczających miejsca związane ze szczytem (tzw. *Strefy I*, pokazanej na rysunku 3), był ograniczony wyłącznie do osób tam zameldowanych⁴⁸. Na podstawie informacji o liczbie dodatkowych zameldowań dokonanych w kwietniu 2004 r. w związku z wymienionymi powyżej ograniczeniami, potencjalnie można byłoby następnie dokonać ekstrapolacji odsetka osób uprzednio niezameldowanych na inne części miasta.

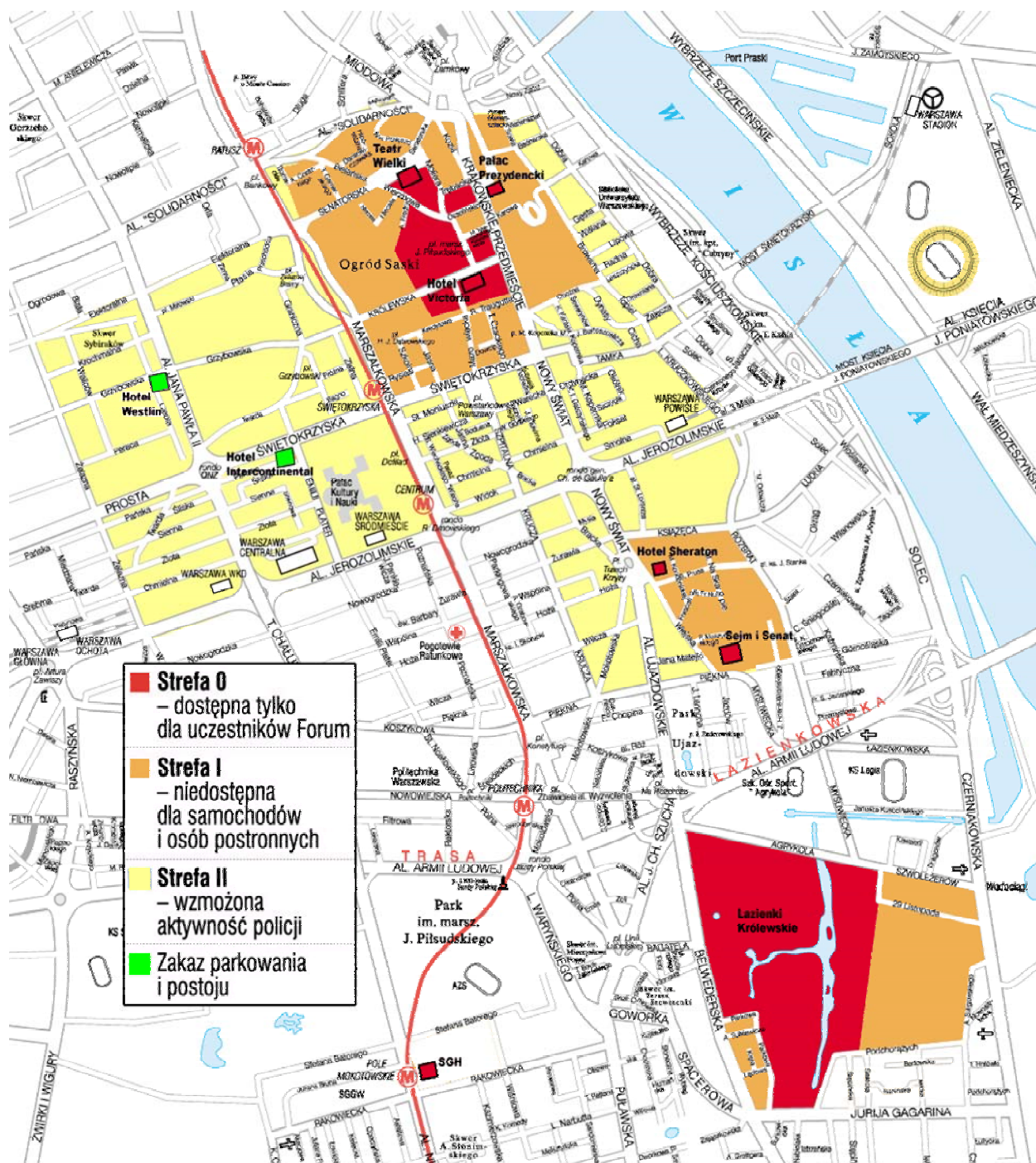
Pomimo potencjalnej atrakcyjności przedstawionego podejścia, przed praktycznym zastosowaniem powyższej informacji do szacowania liczby ludności zamieszkałej w Warszawie bez zameldowania powstrzymuje kilka bardzo istotnych problemów. Po pierwsze, część osób zgłaszała zameldowanie nie ze względu na fakt zamieszkiwania w Śródmieściu, ale na przykład z powodu stałej opieki nad zamieszkałymi tam osobami starszymi, lub zwierzętami domowymi. Po drugie, co dodatkowo potwierdza poprzednią informację, większość zameldowań była dokonywana na bardzo krótki okres (np. dwóch miesięcy), zaś dane ich dotyczące nie były komputeryzowane. Po trzecie, część meldunków wynika z naturalnego biegu rzeczy – niektóre osoby i tak zgłaszałyby się we wspomnianym okresie do wydziałów ewidencji ludności niezależnie od Europejskiego Szczytu Gospodarczego. Wreszcie, rząd wielkości liczby tych dodatkowych zameldowań nie jest zbyt duży – w grę wchodzi informacja o tysiącu kilkuset osobach, włączając w to wszystkie przypadki wymienione powyżej⁴⁹.

⁴⁷ GUS (2006a: 34), szacunek na dzień 30 września 2005 r.

⁴⁸ Informacje ze strony: www.um.warszawa.pl/v_syrenka/esg/index.php, stan na dzień 24 listopada 2006 r.

⁴⁹ Źródło: [6c] (→ „Źródła informacji bezpośredniej”).

Rysunek 3. Mapy stref ograniczonego dostępu w czasie Europejskiego Szczytu Gospodarczego w 2004 r.



Źródło: Urząd m. st. Warszawy, www.um.warszawa.pl/v_syrenka/esg/mapa-strefy.gif, stan na 20 listopada 2006 r.
 Przedruk za zgodą Urzędu m. st. Warszawy.

Oprócz powyższych zastrzeżeń natury administracyjnej, założenie metodyczne, że odsetki osób zamieszkałych bez meldunku są identyczne w różnych częściach Warszawy, wydaje się również nie być prawdziwe, gdyż powyższa informacja dotyczy dosyć specyficznego obszaru ścisłego Śródmieścia. W związku z tym, informacja o dodatkowych meldunkach w okresie Europejskiego Szczytu Gospodarczego nie może być w sposób

wiarygodny uogólniona, i, co za tym idzie, nie może posłużyć do oszacowania liczby osób w rzeczywistości zamieszkałych w Warszawie.

4.1.5. Statystyka kościelna

Kolejnym źródłem danych, mogącym potencjalnie stanowić podstawę do dalszych szacunków i korekt stanu ludności Warszawy, są statystyki kościelne. Można bowiem hipotetycznie wyobrazić sobie ekstrapolację danych o osobach wierzących zebranych przez Kościół Katolicki na pozostałych członków społeczeństwa celem oszacowania rzeczywistej liczby mieszkańców stolicy. Oszacowanie takie wymagałoby jednak przyjęcia dosyć silnego założenia o tym, że struktury społeczno-demograficzne wśród katolików oraz pozostałych grup ludności są takie same, co niekoniecznie musi znajdować odzwierciedlenie w rzeczywistości. Ponadto, potencjalnymi przeszkodami w bezpośrednim wykorzystaniu wspomnianych danych są ich niekompletność oraz szereg ograniczeń w zbieraniu statystyk przez Kościół Katolicki.

Zgodnie z danymi dostępnymi w Instytucie Statystyki Kościoła Katolickiego, zbierane dane obejmują liczbę osób uczęszczających na Mszę św. w niedziele i przystępujących do komunii. Badanie przeprowadzane jest we wszystkich kościołach katolickich w Polsce corocznie w niedzielę zwykłą, w październiku lub listopadzie. Jednak nawet w wypadku agregacji takich danych na poziomie Warszawy, nie można byłoby z nich korzystać celem szacunków ludności z wielu względów, m.in. z uwagi na brak korelacji faktu uczęszczania do kościoła w Warszawie z zamieszkiwaniem w stolicy, oraz z uwagi na danych o frekwencji dla populacji osób niezameldowanych⁵⁰.

Dodatkowym potencjalnym źródłem informacji mogłyby być dane gromadzone przez parafie w kartotekach parafialnych w czasie tzw. wizyty duszpasterskiej, popularnie zwanej „kolędą”. Dane obejmują liczbę, wiek i płeć osób zamieszkałych w danym mieszkaniu, bez względu na zameldowanie. Dane takie aktualizowane są w zależności od częstotliwości wizyt duszpasterskich. W mniejszych parafiach wizyty takie odbywają się corocznie, jednak w dużych parafiach warszawskich parafie dzielone są na sektory i wizyta duszpasterska, a z nią aktualizacja danych o mieszkańcach, przypada nawet raz na 4 lata. Należy zauważyć jednak, że dane te są niedostępne w postaci zagregowanej na szczeblu wyższym niż parafialny, oraz że są one niekompletne, nie obejmują bowiem mieszkań, w których nie zastano mieszkańców, bądź których mieszkańcy nie zdecydowali się na przyjęcie wizyty duszpasterskiej.

⁵⁰ Źródło: [2] (→ „Źródła informacji bezpośredniej”).

Podsumowując, wydaje się, że nie można w prosty sposób wykorzystać statystyk zbieranych i przetwarzanych przez Kościół Katolicki dla ekstrapolacji celem oszacowania rzeczywistej liczby mieszkańców stolicy, bez uciekania się do dodatkowych badań społecznych. Badania takie musiałyby z konieczności być bardziej złożone i brać pod uwagę więcej czynników, w stosunku do badań reprezentacyjnych omówionych w części 3.2, co stawia pod znakiem zapytania celowość ich przeprowadzania w konkretnym celu, jakim jest oszacowanie liczby ludności Warszawy.

4.2. Obliczenia na podstawie danych o gospodarce komunalnej w Warszawie

4.2.1. Miejskie Przedsiębiorstwo Wodociągów i Kanalizacji – zużycie wody i produkcja ścieków

Jednym z potencjalnych źródeł informacji statystycznej możliwej do wykorzystania w szacunkach mogłyby być dane Miejskiego Przedsiębiorstwa Wodociągów i Kanalizacji (MPWiK), dotyczące zużycia wody oraz produkcji ścieków w przeliczeniu na jednego mieszkańca, pod warunkiem jednak, że dane o liczbie mieszkańców zbierane byłyby przez MPWiK niezależnie od danych GUS. Powszechność zastosowania liczników w mieszkaniach sprawia jednak, że podstawowy sposób rozliczania zużycia wody i produkcji ścieków jest oparty o dane wynikające z odczytów pomiarów dla poszczególnych lokali mieszkalnych.

Powszechność wprowadzenia liczników w mieszkaniach oraz, co za tym idzie, upowszechnienie się ekonomicznych, ale też ekologicznych motywacji by oszczędzać wodę, sprawiła, że mimo wzrostu ludności stolicy zużycie wody i produkcja ścieków przez gospodarstwa domowe maleje w stolicy, i to zarówno w wartościach globalnych, jak i przeliczeniu na jednego mieszkańca, według oficjalnych statystyk GUS (GUS, 2006c: 81). Zmienna wartość zużycia wody na osobę uniemożliwia więc jakiegokolwiek szacunki liczby ludności na podstawie zużycia wody lub produkcji ścieków.

Alternatywną, choć bardzo kosztowną do przeprowadzenia, metodą badawczą mogłaby być również analiza zawartości chemicznej ścieków, przeprowadzana przy okazji ich oczyszczania, przy założeniu, że istniałyby ku temu odpowiednie możliwości technologiczne. Przykładowo, prace opublikowane w Japonii wskazują na to, że przeciętna zawartość azotu (N) przedostającego się do środowiska przez ludzkie fekalia wynosi około 9.0 gram na dobę na 1 osobę, zaś dla pozostałego typu ścieków (brudnej wody, tzw. *grey water*) wskaźnik ten wynosi około 1.7 grama (S. Itahashi, M. Komada i M. Takeuchi, 2006:

135). Oczywiście, problemem pozostaje, na ile wymienione wartości można w prosty sposób przełożyć na warunki europejskie, na przykład ze względu na różnice w diecie. Ponadto, oszacowania tego typu prawdopodobnie cechować się będą niską precyzją, może być też trudno oddzielić ścieki domowe od przemysłowych, ścieki produkowane przez ludność dzienną od wytwarzanych przez ludność nocną, itp. Wszystkie te problemy sprawiają, że praktyczne zastosowanie tego typu metody w przypadku Warszawy wydaje się niecelowe, zwłaszcza w kontekście bardzo wysokich kosztów, jakie związane byłyby z aparaturą pomiarową oraz z monitorowaniem jej wskazań przez pewien czas.

4.2.2. Zarząd Oczyszczania Miasta – gospodarka odpadami

Kolejnym potencjalnym źródłem danych mogących pomóc w oszacowaniu rzeczywistej liczby mieszkańców Warszawy są informacje dotyczące produkcji śmieci w przeliczeniu na mieszkańca. Dane takie byłyby możliwe do wykorzystania jedynie przy założeniu, że produkcja śmieci na osobę byłaby stała, bądź wiadoma, szacowana niezależnie od statystyk GUS, wówczas wzrost produkcji śmieci z gospodarstw domowych można byłoby wiązać z proporcjonalnym wzrostem liczby mieszkańców. Dane o produkcji śmieci przez gospodarstwa domowe nie są jednak łatwo dostępne. Odbiorem śmieci z gospodarstw domowych w Warszawie zajmują się liczne prywatne firmy, których jest około stu (MPO, SITA, Remondis, AG Complex, Bestja, Trans-formers, Byś, MPRO, Universitas, ...), choć sześć największych z nich ma ok. 80% udziału w rynku⁵¹.

Kwestia produkcji śmieci na osobę jest dodatkowo skomplikowana ze względu na fakt, że na fluktuacje i zmiany produkcji śmieci przez gospodarstwa domowe mają wpływ dwa przeciwstawne trendy. Z jednej strony wzrost konsumpcji powoduje wzrost ilości śmieci wynikający ze wzrostu liczby opakowań jednorazowych, z drugiej jednak strony upowszechnienie wiedzy i wzrost świadomości ekologicznej wśród mieszkańców, wpływa na świadome ograniczanie ilości produkowanych odpadów przez część gospodarstw domowych. Biorąc pod uwagę fluktuacje w produkcji śmieci przez gospodarstwa domowe wynikające z obydwu trendów, wydaje się, że dane takie byłyby trudne do uzyskania i mało wiarygodne, a co za tym idzie nieprzydatne do ewentualnych szacunków liczby ludności miasta.

⁵¹ Źródło: [6a] (→ „Źródła informacji bezpośredniej”).

4.3. Szacunki oparte na informacji ze zdjęć satelitarnych i lotniczych

W literaturze istnieje wiele przykładów oszacowań liczby ludności na danym obszarze w oparciu o informację pochodzącą ze zdjęć satelitarnych lub lotniczych (telemetryczną). Artykuły na ten temat można znaleźć w wielu czasopismach fachowych, a zwłaszcza w *IEEE Transactions on Geoscience and Remote Sensing* oraz *International Journal of Remote Sensing*, obu dostępnych w Bibliotece Głównej Politechniki Warszawskiej. Zaletą metod telemetrycznych jest, że mogą być one stosowane nawet dla obszarów, dla których trudno jest o jakąkolwiek wiarygodną informację statystyczną poza spisami ludności⁵².

Jeden z nurtów badań telemetrycznych oparty jest na poszukiwaniu w miarę uniwersalnych prawideł uzależniających liczbę ludności na danym obszarze od powierzchni tego obszaru. Do podstawowych propozycji należy tu tzw. prawo allometrycznego (zróżnicowanego) wzrostu (*the law of allometric growth*), które w wersji podanej przez W. R. Toblera (1969) ma postać:

$$(5) \quad r = a \cdot P^b, b < 1,$$

gdzie P oznacza liczbę ludności, r zaś promień koła opisanego na dwuwymiarowym odwzorowaniu danego obszaru. Dla Stanów Zjednoczonych końca lat sześćdziesiątych uzyskano oszacowania parametrów $a = 0.035$ oraz $b = 0.44$, w przypadku promienia wyrażonego w kilometrach.

W ogólności można przyjąć, że istnieje analogiczny związek między liczbą ludności, a powierzchnią danego obszaru, A , bez stosowania przybliżeń przez opisywanie kół. Odpowiednik zależności (5) przyjmuje wówczas postać $A = a \cdot P^b, b < 1$, jak proponują np. D. R. Vining i S. J. H. Louw (1978). Ci sami autorzy zwracają jednak uwagę, że funkcja (5) w ogólnym przypadku nie musi być stabilna w czasie oraz porównywalna pomiędzy krajami, i dlatego jej oszacowania nie są polecane do przeprowadzania, czy nawet wspomagania analiz bazujących na teledziurze. Jako uzasadnienie autorzy przytaczają różne badania przeprowadzone w latach sześćdziesiątych i siedemdziesiątych w Japonii, Norwegii, Szwecji oraz USA, które pokazały, że oszacowania parametrów a i b różnią się nie tylko między krajami, ale również między różnymi okresami dla tych samych krajów.

⁵² Por. np. studium J. F. Olorunfemi'ego (1982) dla Nigerii, lub S. Amarala i in. (2006) dla brazylijskiej Amazonii. Przykładem zastosowania teledziury dla szacowania liczby ludności w kraju rozwiniętym jest praca J. T. Harvey'a (2002) poświęcona Australii. Przegląd urbanistycznych zastosowań metod telemetrycznych opartych na zdjęciach satelitarnych oraz systemach informacji geograficznej (GIS), zarówno dla krajów rozwiniętych, jak i rozwijających się, prezentują D. Maktav, F. S. Erbek i C. Jürgens (2005).

Pewnym wyjściem z sytuacji mogłoby być tu zastosowanie funkcji oszacowanej w miarę niedawno dla danego obszaru (kraju). Przykładów tego typu informacji dostarcza praca P. Suttona i in. (2001), którzy dokonali estymacji zależności (5), osobno dla poszczególnych krajów świata, na podstawie nocnych zdjęć satelitarnych, identyfikujących prawie 23 tysiące ‘podświetlonych’ obszarów miejskich na Ziemi. Przeszkodą w bezpośrednim zastosowaniu ich pracy do szacowania liczby ludności Warszawy jest jednak to, że postać funkcyjna (5) została wyznaczona w oparciu o oficjalne dane o liczbie ludności dla podzbioru ok. 1.6 tysiąca miast, spośród których z dużym prawdopodobieństwem znalazła się Warszawa, jako potencjalnie „najjaśniejsze” miasto w Polsce. W takiej sytuacji uzyskane w oparciu o funkcję allometryczną (5) oszacowanie rzeczywistej ludności miasta będzie siłą rzeczy silnie obciążone w stronę statystyk oficjalnych.

Alternatywnie, można podjąć indywidualną próbę oszacowania liczby ludności dla danego obszaru. Na przykład, jedna spośród metod szacunku zaproponowanych przez J. F. Olorunfemiego (1982) jest niezależna od oficjalnej informacji statystycznej i dlatego mogłaby być wykorzystana w przypadku Warszawy. Metoda ta polega na zsumowaniu powierzchni zabudowanej na danym obszarze i przemnożenie jej przez średnią liczbę mieszkańców przypadającą na 1 m² budynku, uzyskaną na podstawie reprezentacyjnego badania ankietowego w terenie. Problemem może być tu otrzymanie reprezentacyjnej próby, uwzględniającej stratyfikację gęstości zaludnienia według różnych rodzajów zabudowy (jednorodzinna, bloki kilkupiętrowe, bloki wyższe, itp.). Warto jednak nadmienić, że mimo tej przeszkody metoda ta może być użyteczna w kontekście szacowania liczby ludności Warszawy – w cytowanych badaniach Olorunfemiego dała ona wyniki zbieżne z innymi, w szczególności opartymi na danych ze spisu powszechnego.

Jako źródła danych do wymienionych wyżej analiz najczęściej wykorzystywane są zdjęcia satelitarne, dostarczające informacji bardziej ogólnej, o mniejszej dokładności, oraz zdjęcia lotnicze w dużo dokładniejszej skali, stosowane dla mniejszych obszarów. Oprócz obu rodzajów zdjęć, do szacowania liczby ludności wykorzystywane mogą być także informacje z radarów, zwłaszcza z tzw. radarów z syntetyczną aperturą (*Synthetic Aperture Radar*, SAR)⁵³. Przegląd zastosowań informacji dostarczonych przez SAR do detekcji obszarów zamieszkałych oraz do estymacji wielkości zaludnienia danego obszaru dokonany został w pracy F. M. Hendersona oraz X. Zong-Guo (1997). W przypadku różnych wymienionych źródeł danych (zdjęcia satelitarne, lotnicze oraz informacje dostarczane przez

⁵³ Syntetyczny opis działania takiego radaru można znaleźć np. w artykule A. Kośnego (2006).

radar), metoda szacowania ludności byłaby podobna, różnice sprowadzałyby się jedynie do dokładności źródła oraz technicznych metod obróbki materiału wyjściowego.

W przypadku informacji satelitarnej (ogólniejszej), podstawowym źródłem danych do oszacowań bazujących na teledetekcji mogą być obrazy satelitarne oraz radarowe udostępniane w ramach różnych programów badawczych NASA. Przykład zdjęcia fragmentu Warszawy wykonanego w ramach programu Landsat 7, w technologii ETM+ (*Enhanced Thematic Mapper Plus*), pokazany jest na rysunku 4. Poniższe zdjęcie zostało wykonane 7 maja 2000 r. przez satelitę Landsat 7. Stanowi ono fragment większej całości, obejmującej znacznie rozleglejszy obszar (ok. 182 x 185 km), do którego należą również Płock, Bełchatów, czy Kielce⁵⁴. Pomimo zmniejszenia zdjęcia na rysunku 4 w stosunku do oryginału, widać z niego, że nawet obraz we względnie niskiej rozdzielczości (1 piksel = obszar wielkości 14.25 x 14.25 m w terenie), pozwala w przybliżeniu wyodrębnić obszary zabudowane, szlaki komunikacyjne, rzeki, tereny zielone, tereny rolnicze, itp., niezbędne do przeprowadzenia wstępnych oszacowań liczby ludności w oparciu o teledetekcję.

Rysunek 4. Przykładowe zdjęcie fragmentu Warszawy z satelity Landsat (ETM+, dokładność do 14.25 m)



Źródło: NASA Landsat program (landsat.gsfc.nasa.gov) oraz US Geological Survey (edcsns17.cr.usgs.gov/EarthExplorer). Fragment zdjęcia Landsat Orthorectified ETM+ Pan-sharpened No. epp188r024_7f20000507.

⁵⁴ Obszar zdjęcia określony jest przez współrzędne jego czterech wierzchołków: pn.-zach. (52°40'31"N, 19°31'15"E), pn.-wsch. (52°17'14"N, 22°09'54"E), pd.-zach. (51°04'42"N, 18°51'07"E) oraz pd.-wsch. (50°42'21"N, 21°24'36"E).

Przykładowa metodyka szacowania gęstości zaludnienia w oparciu o informacje ze zdjęć Landsat 7/ETM+ zaproponowana została w pracy C. Wu i A. T. Murray'a (2005). Procedura polega na wyodrębnieniu poszczególnych elementów obrazu, odpowiadających obszarom zabudowanym. Technicznie, na zdjęciach identyfikowana jest tzw. „powierzchnia nieprzepuszczalna” (*impervious surface*), definiowana jako „dowolny materiał zapobiegający przedostawaniu się wody do gleby”, której udział w ogólnej powierzchni, według autorów cytowanego opracowania, „dostarcza więcej informacji o terenach zabudowanych, niż jakiegokolwiek klasyfikacje oparte na wykorzystaniu terenu”⁵⁵. Do identyfikacji powierzchni nieprzepuszczalnej wykorzystywane są tu klasyczne statystyczne metody klasyfikacji, oparte na metodzie największej wiarygodności. Następnie, po „wyczyszczeniu” danych, dysponując wyrywkowymi informacjami o liczbie ludności na kilku wybranych obszarach miasta (np. otrzymanymi ze spisu powszechnego, badań ankietowych, itp.), dokonuje się estymacji modelu regresji uzależniającego liczbę ludności od wielkości zamieszkiwanej przez tę ludność powierzchni nieprzepuszczalnej⁵⁶. Zależność taką można następnie interpolować na łączną powierzchnię nieprzepuszczalną dla całego miasta (tu: Warszawy), w celu otrzymania całkowitej liczby zamieszkującej w nim ludności.

Pewien problem przy analizie zdjęć satelitarnych mogą stanowić brakujące dane (z uwagi na obecność chmur, złą jakość zdjęcia, niejawność pewnych informacji, itp.). W takich przypadkach niezbędne jest zastosowanie odpowiednich procedur korygujących, co dodatkowo może zwiększyć nakład pracy. Statystyczną metodologię przeprowadzania tego typu korekt zaproponowali D. A. Griffith, R. J. Bennett oraz R. P. Haining (1989), sugerując szacowanie brakujących informacji na zdjęciach satelitarnych w oparciu o estymację za pomocą metod statystycznych (tak jak poprzednio, metody największej wiarygodności), która mogłaby mieć zastosowanie w przypadku danych o niskiej jakości.

Jeśli chodzi o oszacowanie kosztów tego typu analizy, to z jednej strony zdjęcia takie, jak pokazane powyżej dostępne są za darmo ze stron internetowych programu Landsat (całość w formacie TIFF zajmuje ok. 0.8 GB pamięci), inne zaś za opłatą wahającą się od kilkudziesięciu do kilkuset dolarów USA. Ponadto, podstawowe oprogramowanie do analizy i obróbki danych dostępne jest w Internecie za darmo (np. *MultiSpec* stworzony na Uniwersytecie Purdue⁵⁷), zaś materiały i podręczniki dotyczące wykorzystywania w praktyce

⁵⁵ Cytaty za: C. Wu i A. T. Murray (2005: 561–562), w tłumaczeniu autorów.

⁵⁶ Szczegóły w cytowanym artykule (Wu i Murray, 2005: 562–575).

⁵⁷ Url: www.ece.purdue.edu/~biehl/MultiSpec.

danych z programu Landsat rozpowszechniane są przez NASA⁵⁸. Z drugiej strony, opracowanie danych przy wykorzystaniu systemów informacji geograficznej (GIS) dla całego miasta będzie prawdopodobnie zadaniem praco- i czasochłonnym. Przy dodatkowym stosowaniu badań reprezentacyjnych, pozostają w mocy uwagi przedstawione w części 3.2.

Alternatywnie, w przypadku badań telemetrycznych przeprowadzanych dla Warszawy można wykorzystać znacznie dokładniejsze zdjęcia lotnicze. Przykładem jest tu tzw. ortofotomapa stolicy, przygotowana na zlecenie Biura Geodezji i Katastru Urzędu m. st. Warszawy⁵⁹. W odróżnieniu od zdjęć satelitarnych, barwne fotografie lotnicze stolicy, na których oparta jest ortofotomapa, zostały wykonane w skali 1:10,870 (1 piksel odpowiada obszarowi o wielkości 20 x 20 cm w terenie). Mapa została wykonana w maju 2005 r. i obejmuje cały obszar administracyjny stolicy. Według informacji dostępnych na cytowanej stronie internetowej Biura Geodezji i Katastru,

„[o]rtofotomapa posiada właściwości zarówno mapy jak i zdjęcia lotniczego i może spełniać rolę obydwu tych materiałów. Zawarte są na niej takie same informacje jak na zdjęciu lotniczym, a dodatkowo istnieje możliwość kartometrycznego pomiaru współrzędnych płaskich odfotografowanych szczegółów, ponieważ w trakcie przetworzenia usunięte zostają zniekształcenia występujące na zdjęciach, które spowodowane zostały geometrią zdjęć i deniwelacją terenu.”

Ze względu na wymienione cechy mapy, dostępnej w formacie TIFF z georeferencjami, może ona podlegać obróbce w większości dostępnego oprogramowania systemów GIS. Należy tu jednak brać pod uwagę ograniczenia związane z wielkością plików: ortofotomapa Warszawy w mniej dokładnej skali (1:25,000) zajmuje ponad 1.3 GB pamięci. Przykładowe zdjęcie stanowiące fragment ortofotomapy, obejmujące obszar okolic Pałacu Kultury i Nauki, pokazane jest (w kilkukrotnym pomniejszeniu) na rysunku 5.

Oprócz zastosowań wymienionych przy okazji omawiania zdjęć satelitarnych, dokładniejsza mapa oparta na zdjęciach lotniczych mogłaby mieć kilka dodatkowych zastosowań w projekcie poświęconym szacowaniu rzeczywistej liczby ludności Warszawy. Przede wszystkim, mogłaby posłużyć jako narzędzie weryfikacji informacji o zasobach mieszkaniowych, wykorzystywanej na przykład do konstrukcji operatu losowania w badaniach reprezentacyjnych, omówionych w części 3.2. Warto zaznaczyć, że na mapie takiej uwidocznione są dokładnie poszczególne budynki, dla których można określić (przynajmniej w przybliżeniu) powierzchnię i liczbę kondygnacji (rysunek 5).

⁵⁸ Zob. np.: *Remote Sensing Tutorial* (rst.gsfc.nasa.gov), *Create A Landsat Dataset* (change.gsfc.nasa.gov/create.html), oraz *Science Data Users Handbook* (landsathandbook.gsfc.nasa.gov/handbook.html).

⁵⁹ Url: mapa.um.warszawa.pl, zweryfikowano dn. 27 listopada 2006 r.

Rysunek 5. Przykładowe zdjęcie centrum Warszawy z ortofotomapy Urzędu Miasta (dokładność do 20 cm)



Źródło: Przykładowe zdjęcie ze strony www.um.warszawa.pl/bgik/files/Image/przyklad1.jpg

Przedruk za zgodą Urzędu m. st. Warszawy.

Ponadto, informację z mapy lotniczej można wykorzystać do wyprowadzenia pewnej zależności funkcyjnej analogicznej do (5), uzależniającej liczbę mieszkańców nie tyle od powierzchni, co na przykład od kubatury uwidoczniionych na mapie budynków mieszkalnych. Empiryczna postać takiej zależności mogłaby wtedy zostać oszacowana na podstawie fragmentarycznych badań reprezentacyjnych przeprowadzonych w kilku dzielnicach dla kilku typów zabudowy (wysoka, niska) oraz przy wszelkich zastrzeżeniach przedstawionych w części 3.2. Po dokonaniu takiego oszacowania, jego wyniki można by ekstrapolować na wszystkie budynki mieszkalne uwidocznione na mapie (czyli z wyłączeniem obiektów biurowych, przemysłowych, handlowych, itp.), aby uzyskać ostateczne oszacowanie liczby mieszkańców Warszawy.

5. Metody szacowania liczby ludności dziennej

Podstawowymi danymi, które należy uwzględnić przy szacowaniu liczby ludności dziennej Warszawy są dane o codziennych dojazdach do miasta osób pracujących, uczących się lub przybywających do stolicy z innych powodów, a zamieszkałych poza granicami administracyjnymi miasta. Dotyczy to nie tylko dojazdów z miejscowościach okalających Warszawę i tworzących tzw. aglomerację warszawską ale również z innych miejscowości województwa mazowieckiego, a nawet położonych poza nim. Informacji o codziennych dojazdach, omówionych w części 5.1 na podstawie wyników badania natężenia ruchu w Warszawie, dostarcza zsumowanie danych o indywidualnych dojazdach samochodami prywatnymi, oraz o dojazdach środkami komunikacji zbiorowej (autobusy przewoźników miejskich i prywatnych, pociągi WKD, Kolei Mazowieckich oraz PKP – Przewozów Regionalnych). W części 5.2 przedstawione zostało pokrótce wykorzystanie metody szacowania liczby ludności dziennej opracowanej przez US Bureau of Census.

5.1. Informacje o codziennych dojazdach do Warszawy

5.1.1. Badania natężenia ruchu drogowego

Oszacowanie codziennych dojazdów do Warszawy w komunikacji indywidualnej (KI), czyli z wykorzystaniem samochodów osobowych (SO), wymaga:

- a) określenia natężenia ruchu drogowego na granicach Warszawy;
- b) obliczenia liczby osób korzystających z KI poprzez pomnożenie liczby samochodów przekraczających granice Warszawy przez wskaźnik średniego napelnienia SO;
- c) oszacowania, jaki procent samochodów oraz osób przekraczających granice Warszawy stanowi ludność dzienną stolicy poprzez badanie motywacji podróży. Możemy na przykład przyjąć, że w skład ludności dziennej wchodzić będą osoby pracujące, uczące się bądź studiujące w stolicy, choć lepszym rozwiązaniem wydaje się rozszerzenie tej definicji o osoby przybywające do Warszawy z innych przyczyn. Warto tu nadmienić, że z punktu widzenia planowania przestrzennego nie jest istotne, jak często pojedyncze osoby przyjeżdżają do stolicy, natomiast liczy się to, ile osób przeciętnie przybywa codziennie do Warszawy.

Podstawowym źródłem informacji o dojazdach indywidualnych do Warszawy są badania natężenia ruchu drogowego na granicach stolicy. Najpełniejszym obecnie źródłem

informacji o natężeniu ruchu drogowego w stolicy i na jej granicach są wyniki *Warszawskiego Badania Ruchu 2005*, przeprowadzonego przez zespół Biura Planowania Rozwoju Warszawy S.A. (BPRW, 2005) pod kierownictwem mgr inż. Sławomira Monkiewicza na zlecenie Miasta Stołecznego Warszawy. Warszawskie Badanie Ruchu (WBR) oparte jest na reprezentacyjnym badaniu ankietowym 2,245 gospodarstw domowych, przeprowadzonym przez GfK Polonia Sp. z o.o. w kwietniu, maju i czerwcu 2005 r. na obszarze Warszawy (1,119 przebadanych gospodarstw) i tzw. „strefy”, czyli obszaru obejmującego gminy dawnego województwa warszawskiego i powiatu mińskiego (1,126 przebadanych gospodarstw). Głównymi celami badania było (BPRW, 2005: 9):

- „- uzyskanie podstawowych informacji o zachowaniach komunikacyjnych mieszkańców Warszawy i okolic, wchodzących w skład aglomeracji (obszaru metropolitalnego);
- zebranie danych o potokach ruchu w komunikacji indywidualnej i zbiorowej w Warszawie;
- poznanie ocen i preferencji mieszkańców Warszawy i aglomeracji dotyczących systemu transportowego;
- zbudowanie komputerowego modelu ruchu dla komunikacji indywidualnej i zbiorowej dla stanu istniejącego w obszarze Warszawy”.

Jako uzupełnienie danych otrzymanych z badania ankietowego gospodarstw domowych, w ramach WBR 2005 przeprowadzono także badanie ankietowe kierowców na granicy Warszawy. Badanie objęło ruch dojazdowy do Warszawy na 22 wlotach drogowych w dniach 12 kwietnia do 16 czerwca 2005 (łącznie ponad 20.6 tysiąca zebranych ankiet). Dodatkowym źródłem danych, które posłużyło m.in. do weryfikacji budowanego modelu ruchu, są badania pomiarowe natężenia ruchu pojazdów i pomiary liczby osób w pojazdach komunikacji zbiorowej między innymi na tzw. kordonie Warszawy, czyli na granicy administracyjnej miasta.

Na podstawie badania ankietowego gospodarstw domowych uzupełnionego o badanie ankietowe kierowców i pomiary natężenia ruchu stworzono warszawski model ruchu, obejmujący następujące elementy (BPRW, 2005: 77):

- generowanie i absorpcję ruchu;
- rozkład przestrzenny ruchu dla dnia powszedniego, dla dwóch okresów doby: godzin szczytu porannego i popołudniowego;
- podział zadań przewozowych, obejmujący rozkład ruchu na sieć transportu zbiorowego i sieć drogowo-uliczną.

Korzystając m.in. z wyników WBR 2005, w Biurze Drogownictwa i Komunikacji Urzędu m. st. Warszawy stworzono w 2006 r. projekt *Strategii Zrównoważonego Rozwoju Systemu Transportowego Warszawy na lata 2007–2013* (Strategia..., 2006). Według informacji cytowanych przez jej autorów, w ciągu doby generowanych jest około 550 tysięcy podróży

wjazdowych oraz 550 tysięcy podróży wjazdowych ze stolicy (1 podróż = 1 osoba). Większość (63%) tych podróży wykonywanych jest samochodem osobowym. Dobowe potoki osób w samochodach osobowych wynoszą więc około 350 tysięcy osób w każdym kierunku, oraz dodatkowo 200 tysięcy osób w środkach komunikacji zbiorowej (*Strategia...*, 2006: 53). Po uwzględnieniu średniego wskaźnika napęlenia samochodu (1.5 osoby na samochód), około 230 tysięcy samochodów osobowych w ciągu doby wjeżdża i tyle samo wyjeżdża z Warszawy. Z kolei dobowy tranzyt przez Warszawę wynosi około 23 tysięcy samochodów osobowych (tamże: 54).

Badania pomiaru natężenia ruchu na kordonie (granicy) Warszawy pozwalają na jednoznaczne stwierdzenie, że ruch ten przyjmuje formę porannych dojazdów do Warszawy i popołudniowych powrotów do innych miejscowości wchodzących w skład aglomeracji, położonych poza granicami stolicy. W sposób szczegółowy liczba podróży na granicy Warszawy, badana na 22 wlotach komunikacyjnych do miasta, z uwzględnieniem godzin szczytu komunikacyjnego w dzień powszedni, przedstawiona jest w tabelicy 4. Wskazuje ona także na duże znaczenie komunikacji indywidualnej w dojazdach do Warszawy.

Tablica 4. Liczba podróży na granicy (kordonie) Warszawy, badana na 22 wlotach komunikacyjnych do miasta

Kierunek	Liczba SO w godz. 0.00-24.00	Liczba SO w godz. 7.00-9.00 i 15.30-17.30	Liczba pasażerów KI w godz. 7.00- 9.00 i 15.30-17.30	Łączna liczba pasażerów w godz. 7.00-9.00 i 15.30-17.30	Udział podróży KI w godz. 7.00-9.00 i 15.30-17.30.
Do Centrum	268 581	78 691	102 298	175 456	58%
Od Centrum	277 200	75 119	97 655	166 226	59%
Razem	545 781	153 810	199 953	341 682	59%

Źródło: Opracowanie własne na podstawie (Strategia..., 2006: 60, Tabela 17).

Jak wynika z tablicy 4, istnieją pewne rozbieżności w ramach danych przytaczanych przez autorów *Strategii...* (np. w różnych miejscach cytowanego raportu przytaczane są szacunki liczby samochodów osobowych dobowo wjeżdżających do Warszawy wahające się między 230 a 270 tysięcy), które na etapie właściwego projektu badawczego mającego na celu szacunek liczby ludności Warszawy powinny zostać wyjaśnione. Niezależnie od powyższych zastrzeżeń, liczbę osób dojeżdżających do Warszawy w ramach komunikacji indywidualnej można uzyskać poprzez przemnożenie liczby samochodów osobowych przekraczających granicę przez współczynnik średniego napęlenia samochodu osobowego na wlotach do Warszawy, a następnie odjęcie ruchu tranzytowego.

Nieco skomplikowana jest kwestia wyliczenia średniego napęlenia samochodów osobowych. Badanie ankietowe gospodarstw domowych (dane z WBR 2005) wskazało na różne średnie napęlenia samochodów mieszkańców aglomeracji (poza Warszawą)

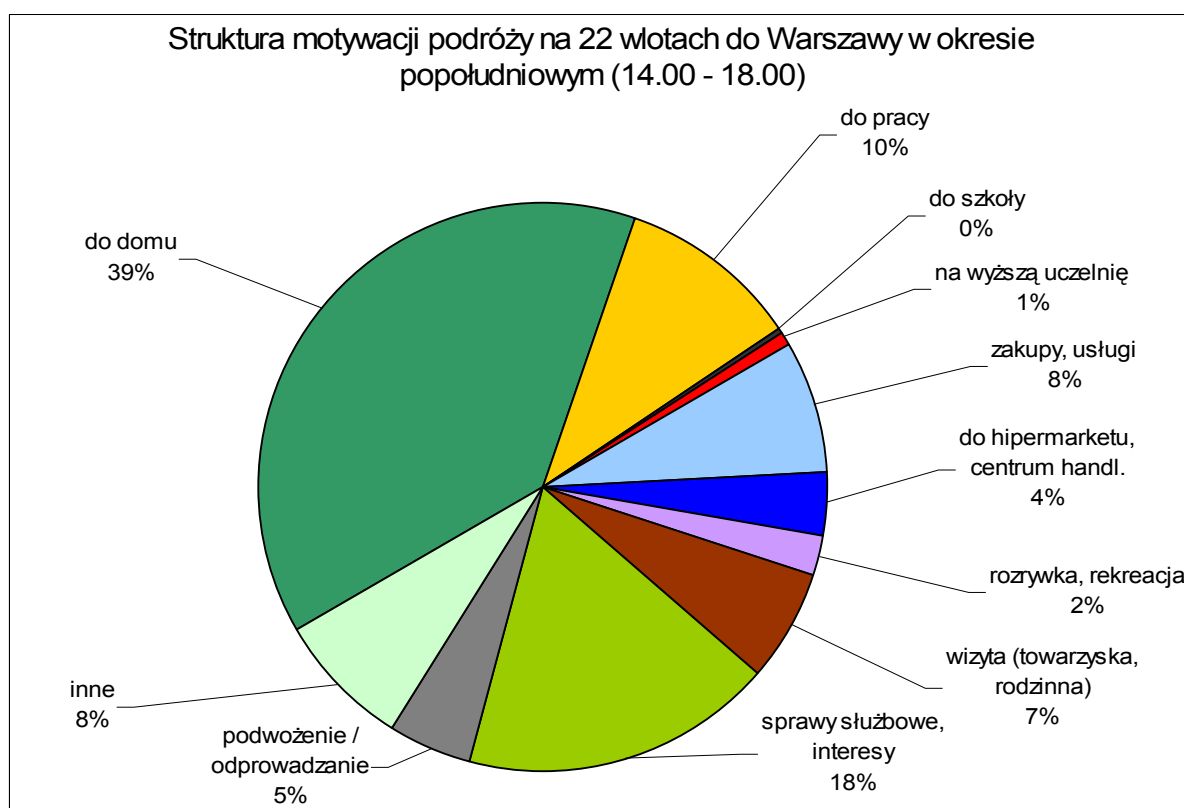
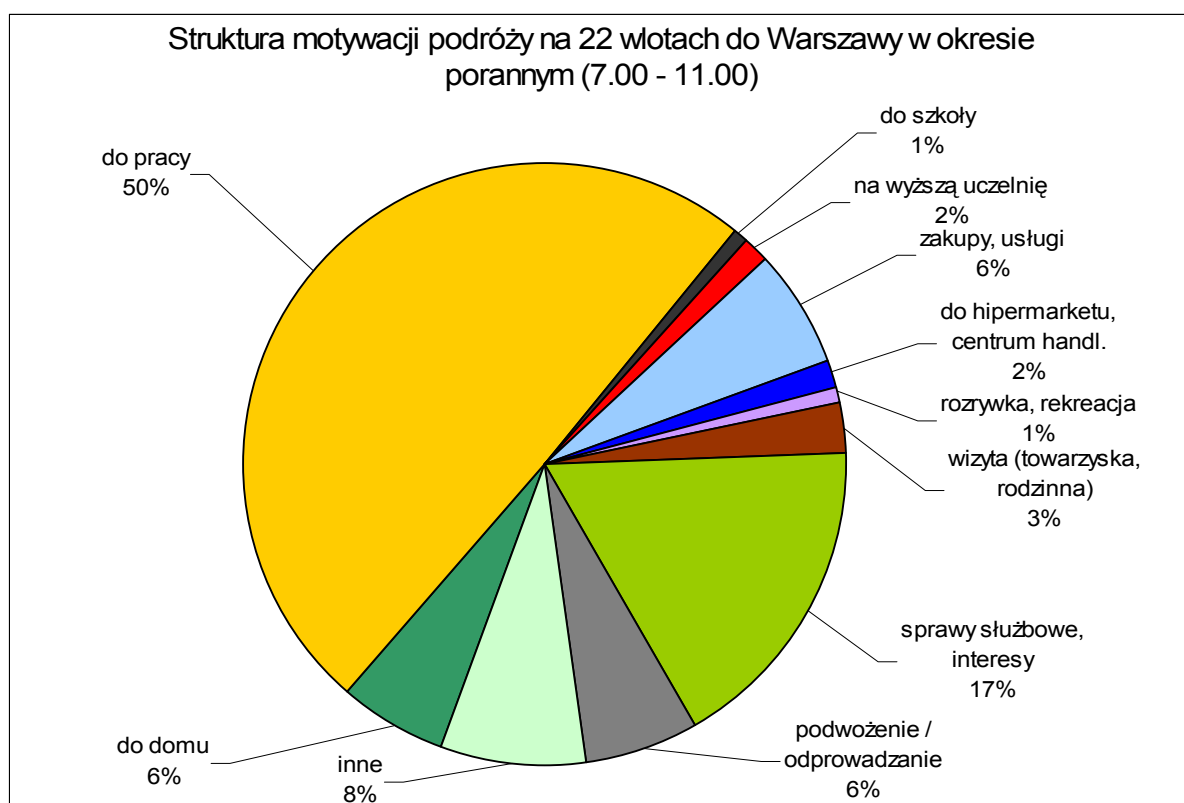
podróżujących do Warszawy i z Warszawy, w zależności od przyczyny (motywacji) podróży. I tak, średnie napelnienie samochodu osobowego mieszkańców aglomeracji w dzień powszedni w drodze do Warszawy wyniosło 1.3 osoby na samochód dla dojazdu do pracy, tyle samo w przypadku dojazdu do szkoły (1.3), a już 1.4 osoby na samochód w przypadku dojazdu na uczelnię wyższą. Z kolei napelnienie samochodu osobowego mieszkańca „strefy” przy wyjeździe z Warszawy poza jej granicę w dzień powszedni wyniosło 1.4 osoby na samochód osobowy dla podróży do domu. Trzeba tu nadmienić, że dane dotyczące średniego napelnienia samochodów mieszkańców „strefy” w podróżach do i z Warszawy w zależności od motywacji mogłyby zostać użyte w celu wyliczenia przeciętnego napelnienia samochodów osobowych na granicy Warszawy tylko przy zastosowaniu średniej ważonej. Do jej wyznaczenia konieczne byłoby wzięcie pod uwagę nie tylko średniej wartości napelnienia samochodu w zależności od motywacji podróży, ale również udział danego potoku pojazdów (również według motywacji) w całym natężeniu ruchu samochodów osobowych.

Alternatywnie, informacji o średnim wypełnieniu samochodów osobowych przekraczających granicę Warszawy można również szukać w wynikach badania ankietowego kierowców (BAK), drugim komponencie WBR 2005. Zgodnie z nim średnie napelnienie samochodów osobowych na 22 trasach wjazdowych do Warszawy wyniosło około 1.5 osoby na samochód. Trzeba zauważyć jednak, że badanie objęło wszystkie samochody wjeżdżające do stolicy, bez względu na cel podróży.

Badanie ankietowe kierowców dostarczyło dodatkowo wielu cennych informacji o częstotliwości podróży do Warszawy – niemal 64% wszystkich ankietowanych zadeklarowało fakt podróżowania do Warszawy codziennie lub kilka razy w tygodniu. Ponad 85% kierowców samochodów osobowych podawało Warszawę jako miejsce docelowe podróży, przy zaledwie niespełna dziesięcioprocentowym udziale tranzytu (niewielkie ilości respondentów deklarowały również podróż po Warszawie lub wyjazd z Warszawy).

Badanie struktury motywacji podróży wśród kierowców samochodów osobowych na wjeździe do Warszawy, również dostępnej na podstawie wyników BAK, wykazało w godzinach 7.00-11.00 dominację dojazdów do pracy, na uczelnię i do szkoły. Średnio dla badanych 22 wlotów, udział kierowców samochodów osobowych deklarujących jeden z trzech wymienionych celów podróży wyniósł w godzinach porannych i przedpołudniowych 51.7%, natomiast w godzinach popołudniowych (między 14.00 a 18.00) – zaledwie 11.5%. Uwzględnienie struktury motywacji podróży dla okresu porannego i popołudniowego, pokazana na rysunku 6, pozwala na modyfikowanie definicji ludności dziennej, stosownie do potrzeb informacyjnych Urzędu m. st. Warszawy.

Rysunek 6. Struktura motywacji podróży na 22 włotach do Warszawy, rano i po południu



Źródło: Badanie Ankieta Kierowców, WBR 2005.

Wszystkie opisane powyżej informacje uzyskane z badań ankietowych, zweryfikowane poprzez wyniki pomiarów natężenia ruchu, stanowią dane wsadowe do wspomnianego uprzednio modelu ruchu dla Warszawy. Ostateczna wersja modelu (BPRW, 2006), oparta na wynikach WBR 2005, w chwili obecnej oczekuje na wdrożenie do użytkowania⁶⁰. Z chwilą wdrożenia, wszystkie przedstawione powyżej informacje o kierunkach i natężeniu przepływów, itp., można będzie również uzyskać z modelu. W tej chwili dostępne są już pewne wstępne, wrywkowe oszacowania, jak na przykład dane o przepływach pojazdów na godzinę w okresie szczytu porannego, pokazane na rysunku 7, dostępnym na stronie internetowej Zarządu Dróg Miejskich. Przy interpretacji danych z rysunku 7 należy wziąć pod uwagę, że pokazane liczby dotyczą tzw. pojazdów umownych, będących ekwiwalentami samochodów osobowych, przy czym na przykład samochód dostawczy odpowiada dwóm pojazdom umownym, TIR – trzem, itp.⁶¹

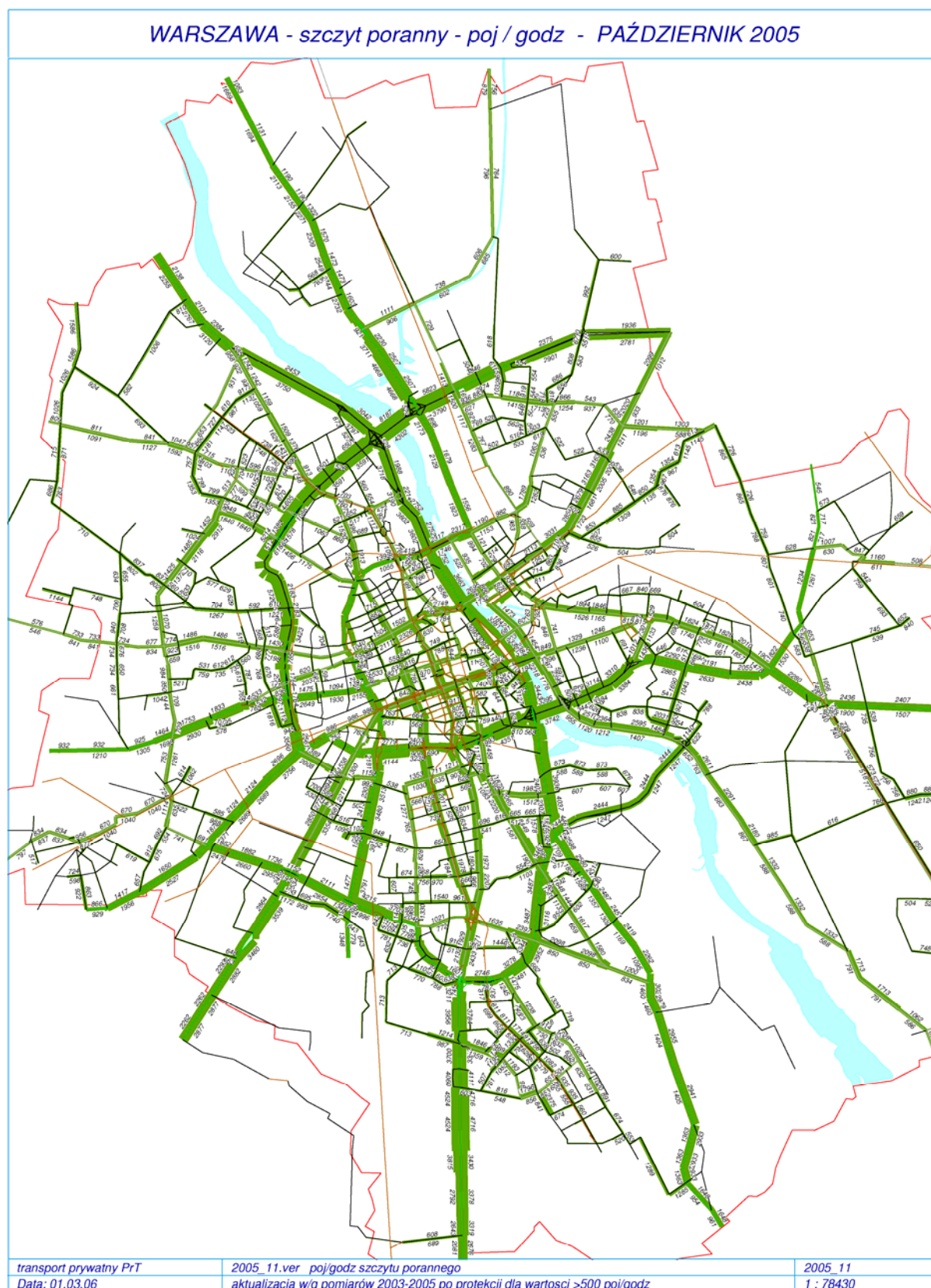
Trzeba tu jednak wspomnieć, że z modelu mogą zostać wygenerowane wartości teoretyczne opisujące poszczególne przepływy, które zostały uzyskane na podstawie danych empirycznych, w oparciu o pewną strukturę powiązań między modelowanymi zjawiskami opisaną za pomocą zależności matematycznych, oraz przy przyjęciu szeregu dodatkowych założeń. W związku z tym, wyniki otrzymane na podstawie modelu ruchu, oprócz błędów informacji wsadowej (tu: danych z badań ankietowych i pomiarów), obarczone są jeszcze błędami związanymi z taką, a nie inną formalną specyfikacją równań modelu, oraz z niepewnością dotyczącą poszczególnych założeń. Z drugiej strony, model jest pewną generalizacją mającą na celu opis fragmentu rzeczywistości w interesującym badacza aspekcie, więc jego wyniki mają charakter dosyć ogólny, syntetyzujący informację pochodzącą z różnych źródeł. Wydaje się więc, że dla celów szacowania liczby ludności dziennej Warszawy, model ruchu może potencjalnie zostać (po wdrożeniu) wykorzystany do uzyskania informacji dotyczących codziennych napływów i odpływów ludności. Alternatywą pozostają wyniki badań ankietowych, takich jak cytowane WBR 2005.

Podsumowując, można stwierdzić, że wyniki Warszawskiego Badania Ruchu, zarówno badania ankietowego gospodarstw domowych, badania ankietowego kierowców, jak i dane pochodzące z tworzonego modelu ruchu, stanowią obiecujące pole do dalszej eksploracji w celu szacunku liczby dojazdów do Warszawy komunikacją indywidualną.

⁶⁰ Informacja na dzień 20 listopada 2006 r. Źródła: [6b], [7] (→ „Źródła informacji bezpośredniej”).

⁶¹ Źródło: [7] (→ „Źródła informacji bezpośredniej”).

Rysunek 7. Przykład mapy natężenia ruchu w Warszawie dla szczytu porannego, otrzymana z modelu ruchu



Źródło: Zarząd Dróg Miejskich: www.zdm.waw.pl/docs/doc1192.pdf, stan na dzień 21 listopada 2006 r.

Przedruk za zgodą Urzędu m. st. Warszawy.

Szczególnie cenna wydaje się dezagregacja danych w zależności od celu (motywacji) podróży. Pozwala to bowiem na wydzielenie zarówno populacji osób stale dojeżdżających (dojazd do pracy, szkoły i na uczelnię) oraz osób okazjonalnie dojeżdżających do Warszawy w innych celach. Wobec dostępności tego typu danych, liczbę ludności dziennej Warszawy można oszacować w dwóch wariantach: dodając do ludności nocnej tylko osoby dojeżdżające stale (praca, szkoła, uczelnia) lub uwzględniając częstotliwość innych, mniej regularnych, ale stanowiących przecież prawie połowę dojazdów, podróży w pozostałych celach (wyjazd na zakupy, wizyta towarzyska, sprawy służbowe, rozrywka).

W celu dalszych szacunków niniejszej wstępnej ekspertyzy, jako liczbę osób dojeżdżających komunikacją indywidualną do Warszawy, tworzącą pulę ludności dziennej, można przyjąć za autorami *Strategii...* (2006) liczbę 320 tysięcy osób, z zaznaczeniem jednak, że liczba ta będzie wymagała dodatkowej weryfikacji w świetle poszczególnych wyników WBR w ramach właściwego badania liczby ludności dziennej Warszawy.

5.1.2. Dojazdy do Warszawy z wykorzystaniem transportu publicznego

Drugim komponentem składowym koniecznym do dodania do ludności nocnej celem oszacowania ludności dziennej, obok dojazdów komunikacją indywidualną, są dojazdy do stolicy środkami komunikacji zbiorowej: autobusowej i kolejowej. Podobnie jak w przypadku dojazdów komunikacją indywidualną, również w przypadku transportu zbiorowego (TZ) podstawowym obecnie źródłem informacji są dane z Warszawskiego Badania Ruchu (WBR) z 2005 r., zarówno z badania ankietowego gospodarstw domowych, jak i z pomiarów natężenia ruchu pojazdów na granicy Warszawy, oraz pomiaru liczby podróżnych w komunikacji zbiorowej.

Badanie ankietowe gospodarstw domowych przeprowadzone w ramach WBR 2005 nie dostarcza wielu istotnych danych odnośnie transportu zbiorowego. Analiza zachowań komunikacyjnych mieszkańców „strefy” nie wydaje się szczególnie przydatna w oszacowaniu liczby dojazdów do stolicy komunikacją zbiorową, zważywszy na fakt, że dokładniejsze dane uzyskujemy np. dzięki badaniom pomiaru natężenia ruchu i pomiarom liczebności pasażerów TZ na granicach Warszawy, które również zostały przeprowadzone w ramach WBR.

Dane WBR dotyczące pomiarów pasażerów i pojazdów komunikacji zbiorowej (KZ) dostępne są również w dezagregacji na wyniki badania w 22 punktach wlotowych do stolicy, w których badanie było przeprowadzane. Przy analizie powyższych danych trzeba jednak zaznaczyć, że pomiary liczby pasażerów na granicy Warszawy dokonywane były w dwóch dwugodzinnych cyklach w ciągu doby (7.00–9.00 i 15.30–17.30), natomiast zaledwie w

trzech punktach na kordonie Warszawy pomiar pasażerów komunikacji zbiorowej dokonywany był w godzinach 6.00–22.00 (ul. Radzymińska, Trakt Brzeski i Al. Krakowska).

Zbiorecze zestawienie danych WBR odnośnie dojazdów komunikacją zbiorową do Warszawy znajdujemy w projekcie *Strategii...* (2006). Według obliczeń autorów projektu *Strategii...*, dobowe potoki pasażerów przemieszczających się komunikacją zbiorową przez granice Warszawy wynoszą w obu kierunkach około 200 tysięcy osób (*Strategia...*2006: 53). Największe natężenie ruchu w kierunku wjazdowym do stolicy występuje w godzinach szczytu porannego (7.00-9.00), największe natężenie ruchu wyjazdowego w godzinach szczytu popołudniowego (15.30-17.30). Udział przewozu komunikacją zbiorową w godzinach szczytów (rannego i popołudniowego) na granicach Warszawy wynosi jednak zaledwie 41%, co wskazuje na fakt, że większość dojazdów do Warszawy na granicy miasta odbywa się za pomocą indywidualnych środków transportu, jak to zostało pokazane w tablicach 5 i 6.

Tablica 5. Liczba pasażerów TZ na granicy Warszawy w godzinach szczytu porannego i popołudniowego

Kierunek	Liczba pasażerów w TZ	Łączna liczba pasażerów	Udział podróży TZ
Do centrum	73 158	175 456	42%
Od centrum	68 571	166 226	41%
Razem	141 729	341 682	41%

Uwaga: Szczyt poranny zdefiniowany jest jako godziny 7.00–9.00, a popołudniowy jako 15.30–17.30

Źródło: opracowanie własne na podstawie danych ze *Strategii...* (2006: 60, Tabela 17)

Tablica 6. Kordon Warszawy – pasażerowie komunikacji zbiorowej

Potok pasażerski [pas/h]	7.00-8.00	16.00-17.00	Cała doba
Ruch do Warszawy	28 345	7 837	200 913
<i>Z czego:</i>			
- autobus	14 240	6 063	131 713
- kolej podmiejska	14 105	1 774	69 200
Ruch z Warszawy	8 049	20 748	200 268
<i>Z czego:</i>			
- autobus	6 434	11 781	131 068
- kolej podmiejska	1 615	8 967	69 200

Źródło: BPRW (2005: 63)

Wyniki badania WBR 2005 wykazały na rosnące znaczenie przewoźników prywatnych w komunikacji autobusowej na granicy stolicy. Urząd Marszałkowski Województwa Mazowieckiego wydał do tej pory zezwolenia na wykonywanie przewozów regularnych około 150 przewoźnikom (w tym 19 przedsiębiorstwom PKS i około 120 małym przewoźnikom prywatnym obsługującym ruch podmiejski, związany z dojazdami do

Warszawy) (za opracowaniem Faber Mausell/AECOM, 2005: 19). Udział procentowy przewoźników prywatnych w ruchu autobusowym na granicach Warszawy różni się znacznie w zależności od trasy wjazdowej do stolicy, sięgając nawet wartości 100% w przypadku porannej godziny szczytu na ulicy Okuniewskiej (*Strategia...*2006: 61).

Transport kolejowy na granicy Warszawy jest realizowany obecnie przez 4 podmioty:

- a) spółkę „Koleje Mazowieckie – KM” Sp. z o.o., utworzona przez Samorząd Województwa Mazowieckiego (51% udziałów) i PKP Przewozy Regionalne (49% udziałów);
- b) spółkę „Szybka Kolej Miejska” Sp. z o.o., której większościowym akcjonariuszem (99,7% udziałów) jest miasto stołeczne Warszawa. Spółka ta została zawiązana w celu zorganizowania, uruchomienia i zarządzania systemem przewozów pasażerów w formule „Parkuj i jedź” w oparciu o istniejące w aglomeracji warszawskiej linie kolejowe, obecnie obsługuje dojazdy do Warszawy z Piastowa i Pruszkowa.
- c) spółkę „PKP Warszawska Kolej Dojazdowa”, która zarządza, eksploatuje i prowadzi przewozy na linii podmiejskiej Warszawa Śródmieście WKD – Grodzisk Mazowiecki, z odnogą: Podkowa Leśna – Milanówek (Faber Mausell/AECOM, 2006: 19–20);
- d) PKP Przewozy Regionalne, na pozostałych trasach dojazdowych do Warszawy (np. Radom – Warszawa, Łódź – Warszawa);

Podsumowując omówienie danych możliwych do uzyskania z WBR, warto podkreślić, że w przeciwieństwie do komunikacji indywidualnej, w przypadku komunikacji zbiorowej nie przeprowadzono na granicy miasta badania motywacji podróży do lub z Warszawy. Dane takie można więc uzyskać jedynie z badania ankietowego gospodarstw domowych. Nałożenie na siebie powyższych informacji – o liczbie pojazdów KZ i liczbie pasażerów w KZ uzyskanych w badaniu pomiarowym, oraz o motywacji podróży do lub z Warszawy uzyskanych w badaniu ankietowym innej próby (próby mieszkańców „strefy”) – niesłoby ze sobą konieczność dokonania wielu założeń i uogólnień.

Alternatywnym (i jednocześnie komplementarnym) wobec badania WBR źródłem danych o dojazdach KZ do Warszawy mogłyby być dane o biletach miesięcznych kolejowych i autobusowych. Można domniemać, że 100% osób posiadających bilety miesięczne stanowi regularnie ludność dzienną Warszawy poprzez pracę bądź naukę w stolicy. Podporą takiego założenia mogą być wyniki badania ankietowego mieszkańców strefy w ramach WBR 2005, według których 89% posiadaczy biletów miesięcznych dojeżdżających do Warszawy podróżuje do pracy, szkoły lub na uczelnię wyższą (choć warto zauważyć, że 65% osób nie posiadających biletu okresowego również deklarowało te 3 cele). Uzyskanie dokładniejszych

danych byłoby jednak czasochłonne⁶², wymagałoby zwrócenia się z prośbą o udostępnienie statystyk o wydanych biletach okresowych do ZTM oraz czworga wymienionych wyżej przewoźników kolejowych, obsługujących ruch dojazdowy do stolicy. Ponadto, dla uzyskania pełnego obrazu dojazdów KZ do Warszawy, informacje o biletach okresowych należałoby jeszcze uzupełnić danymi dojazdach autobusami prywatnymi uzyskanymi z WBR. W związku z powyższymi zastrzeżeniami, wydaje się więc, że przy szacowaniu liczby ludności dziennej Warszawy celowe jest ograniczenie się do jednego źródła danych obejmującego wszystkie kategorie dojazdów, jakim jest warszawskie badanie ruchu.

5.2. Metoda stosowana przez US Bureau of Census

L. F. Schnore (1960) zauważył, że informacje o ludności dziennej i dojazdach mogą pochodzić z trzech różnych źródeł: (a) badań ruchu, (b) administracyjnych źródeł o zatrudnieniu zbieranych przez poszczególne zakłady pracy, oraz (c) spisów powszechnych, w których respondenci podają informację o miejscu zatrudnienia. Do celów badawczych w skali makro, autor cytowanego opracowania zaleca stosowanie źródeł (a) oraz (c), jako dających więcej możliwości przeprowadzania różnego typu analiz przestrzennych. Jako najdokładniejsze źródło, niosące potencjalnie najwięcej informacji i nie będące z drugiej strony ograniczone wymogami publikacyjnymi lub problemami ochrony informacji, wskazywane są badania ruchu (tamże: 20).

Powyższe uwagi znajdują odzwierciedlenie w metodzie US Bureau of Census, według której liczba ludności dziennej może być szacowana na dwa równoważne sposoby⁶³. W pierwszym z nich (definitywnym), liczba ludności dziennej jest równa liczbie mieszkańców (tj. ludności nocnej), powiększonej o liczbę osób codziennie dojeżdżających, oraz pomniejszonej o liczbę osób codziennie wyjeżdżających. Informację o dojazdach i wyjazdach otrzymuje się tu z badań ruchu – źródła (a) w cytowanym powyżej zestawieniu L. F. Schnore’a (1960). Dla uproszczenia, zakłada się tu, że kategoria osób dojeżdżających (*commuters*) równoważna jest z osobami dojeżdżającymi do pracy, co w kontekście wyników badania ruchu zaprezentowanych w części 5.1 zupełnie nie odpowiada *status quo* dla Warszawy, gdzie bardziej kompletne dane można uzyskać z WBR 2005.

Według drugiej metody (bilansowej), według autorów będącej prostszej obliczeniowo i nie wymagającej tak szczegółowych danych, jak pierwsza, liczbę ludności dziennej

⁶² Źródło: [8] (→ „Źródła informacji bezpośredniej”).

⁶³ Url: www.census.gov/population/www/socdemo/daytime/daytimepopcalc.html, stan na 28 listopada 2006 r.

otrzymuje się jako liczbę mieszkańców, powiększoną o liczbę pracowników pracujących w danej miejscowości, a następnie pomniejszoną o liczbę pracowników zamieszkałych w danej miejscowości. Informacje te otrzymuje się ze spisów powszechnych, czyli alternatywnie ze źródła (c) według systematyki L. F. Schnore'a (1960), ewentualnie uzupełnionego przez dane o zatrudnieniu (b). Autorzy metody w bardzo prosty algebraicznie sposób wyprowadzają, dlaczego obie metody dają w efekcie jednakowe wyniki, przy założeniu, że wspomniana korekta dotyczy jedynie osób pracujących, a nie na przykład dojazdów związanych z nauką, rekreacją, korzystaniem z handlu i usług, itp., co wynika z charakteru dostępnych w Stanach Zjednoczonych danych ze spisów powszechnych⁶⁴.

Przy zastosowaniu powyższych metod do oszacowania liczby ludności dziennej Warszawy, należy wziąć pod uwagę nie tylko, jakie dane są dostępne jeśli chodzi o same dojazdy, ale także, co wiadomo o wielkości i kompozycji korekty całkowitej liczby ludności. Jeśli istnieje pewność, że korekta ta dotyczy wyłącznie ludności nocnej, jak w przypadku badań ankietowych zaproponowanych w części 3.2, można wówczas wykorzystać zarówno wyniki badań ruchu, jak i metodę bilansową. Przy stosowaniu metody bilansowej trzeba jednak zwrócić szczególną uwagę na to, jakie kategorie ludności dziennej podlegają zliczaniu. W przypadku Warszawy nie można ograniczyć się do osób pracujących, gdyż według WBR 2005 stanowią oni około połowy osób dojeżdżających codziennie do stolicy (rysunek 6). W efekcie, najpewniejszym źródłem danych pozostają i tak badania ruchu. Oszacowane na ich podstawie liczby codziennych dojazdów do oraz wyjazdów z Warszawy można nałożyć na liczbę ludności nocnej skorygowaną o wielkość uzyskaną na podstawie reprezentacyjnych badań ankietowych, aby otrzymać w efekcie oszacowanie liczby ludności dziennej miasta.

Jeżeli natomiast korekta całkowitej liczby ludności obejmuje zarówno ludność nocną, jak i dzienną, tak, jak w przypadku oszacowań dla najbardziej mobilnych grup wieku (jak w przypadku opisanym w części 3.1), to wtedy i tak nie można skorzystać z metody bilansowej i trzeba użyć danych z badania ruchu (np. WBR 2005). W takiej sytuacji szacowanie odbywa się *de facto* w drugą stronę: od ogólnej korekty dla ludności dziennej trzeba odjąć bilans codziennych napływów i odpływów, żeby otrzymać saldo korekty wyłącznie dla ludności nocnej. Trzeba tu jednak wziąć pod uwagę, że odejmowane saldo napływów i odpływów musi dotyczyć tych samych kategorii ludności, jakie są brane pod uwagę w korekcie ludności.

⁶⁴ Tamże.

6. Podsumowanie i wnioski

6.1. Podsumowanie wyników studium

Dla informacji zaprezentowanych w poprzednich częściach opracowania, można pokusić się o syntezę i wyznaczenie przynajmniej przybliżonych rzędów wielkości korekt liczby ludności nocnej i dziennej. Ze względu na fragmentaryczny charakter jakichkolwiek dostępnych danych, metoda szacunku będzie z konieczności musiała mieć charakter hybrydowy, łącząc elementy szacowania definicyjnego i bilansowego, według definicji US Bureau of Census cytowanych w części 5.2.

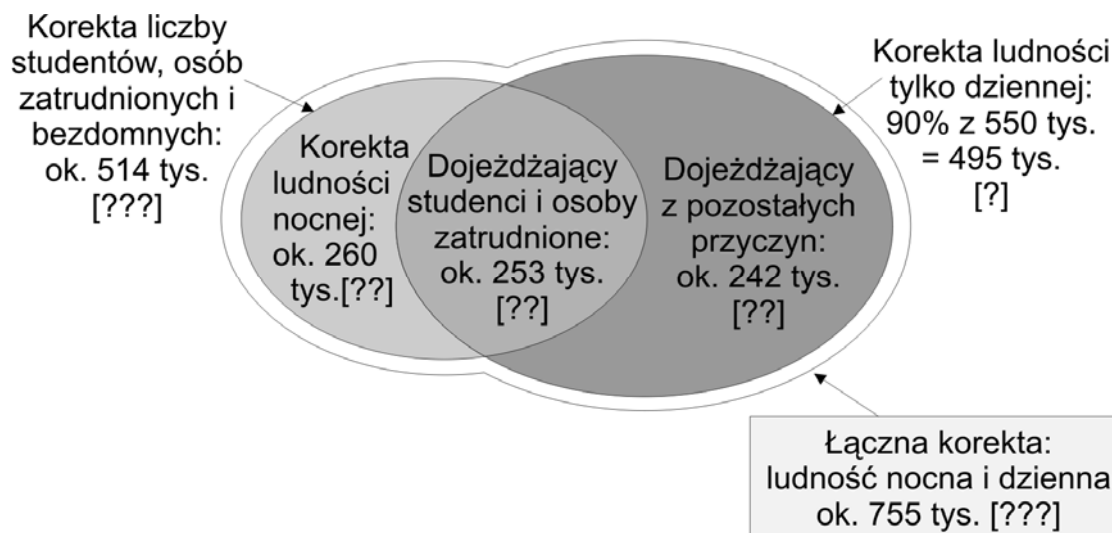
Według wyników przykładowego oszacowania pokazanego w części 3.1, korekta liczby ludności dziennej dla studentów i osób zatrudnionych (oraz bezdomnych) wynosi łącznie około 514 tysięcy osób (zob. rysunek 2). Z drugiej strony, z badań ruchu wynika, że w ciągu doby do Warszawy przyjeżdża i wyjeżdża z niej około 550 tysięcy osób (WBR 2005, część 5.1), z czego po odjęciu mniej więcej dziesięciu procent ruchu tranzytowego (w dużym uproszczeniu), otrzymujemy 495 tysięcy osób dojeżdżających do Warszawy. Ponadto, dla komunikacji indywidualnej wiadomo, że z pracą lub studiami związane jest 51.1% codziennych podróży (por. rysunek 6). Założywszy, że odsetek ten jest taki sam również dla komunikacji zbiorowej, otrzymalibyśmy, że do pracy lub na wyższe uczelnie dojeżdża do stolicy codziennie około 253 tysięcy osób.

Uwzględniając powyższe informacje, można wyznaczyć przybliżoną wielkość korekty dla ludności nocnej jako około 260 tysięcy osób (514 tysięcy – 253 tysięcy, po zaokrągleniu), oprócz których każdego dnia w Warszawie przebywa dodatkowo około 495 tysięcy osób dojeżdżających z różnych powodów do stolicy (bez ruchu tranzytowego). Łączna korekta dla liczby ludności dziennej wynosi wtedy w przybliżeniu $260 \text{ tysięcy} + 495 \text{ tysięcy} \approx 755 \text{ tysięcy}$ osób, jak pokazano na rysunku 8. W zależności od potrzeb, wielkości powyższych korekt można odnieść albo do liczby ludności faktycznie zamieszkałej, albo liczby ludności rezydującej publikowanych przez GUS, równych odpowiednio około 1.70 miliona i 1.65 miliona osób⁶⁵. Liczba ludności nocnej Warszawy wynosiłaby wtedy w przybliżeniu 1.91 – 1.96 miliona osób, zaś ludności dziennej – między 2.41 a 2.46 miliona. Co za tym idzie, udział wielkości korekty w skorygowanej liczbie ludności nocnej, istotny na przykład dla

⁶⁵ Dla ludności faktycznie zamieszkałej: GUS (2006a: 34), dla ludności rezydującej proste oszacowanie na podstawie wyników NSP 2002, zakładające jej proporcjonalność do liczby ludności faktycznie zamieszkałej. Dane o ludności rezydującej dla okresów międzypisowych nie są zbierane przez GUS poniżej poziomu województw, por. punkt 2.2.3. Źródło: [1b] (→ „Źródła informacji bezpośredniej” na końcu niniejszego opracowania).

wyznaczenia wielkości próby w reprezentacyjnym badaniu ankietowym, proponowanym w części 3.1, wynosi w tym przykładzie około 13.5%.

Rysunek 8. Łączna korekta liczby ludności dziennej uwzględniająca wszystkie grupy osób dojeżdżających



Źródło: opracowanie własne na podstawie danych omówionych w częściach 3.1 oraz 5.1.

Na powyższym schemacie, tak samo, jak na rysunku 2, znaki zapytania przy poszczególnych wielkościach wskazują na dużą niepewność oszacowań. W ogólności, kwestia niepewności i błędów szacunku stanowi bardzo poważny problem każdej z omawianych w raporcie metod wyznaczania liczby dziennej i nocnej ludności Warszawy. Dla niektórych metod (np. sondażowych badań reprezentacyjnych) można wyznaczyć wielkości błędów oszacowań z odpowiednich wzorów, np. (4). W przypadku korekt mechanicznych, takich, jak zaproponowane w części 3.1, wyznaczenie wielkości błędu nie jest jednak w ogólnym wypadku możliwe. Z tego względu, żeby odpowiednio uwzględnić niepewność co do ostatecznego wyniku, należałoby przeprowadzić kilka niezależnych oszacowań przy zastosowaniu różnych metod, które pozwoliłyby przynajmniej na ocenę zakresu zmienności szacowanych wielkości korekt nocnej i dziennej ludności m. st. Warszawy.

Pewnych wskazówek co do precyzji oszacowania opartego na przedstawionych wyżej zestawieniach i bilansach dla różnych grup ludności może dostarczyć uproszczona analiza wrażliwości wyników na zmiany niektórych spośród przyjętych założeń. Przykładowo, można w ten sposób poddać weryfikacji na trzy silne założenia przyjęte w częściach 3.1, 5.1 oraz 6.1, dotyczące: (a) wysokości współczynnika skolaryzacji brutto dla szkolnictwa wyższego w Warszawie, (b) odsetka osób pracujących wśród warszawskich studentów, oraz (c) udziału ruchu tranzytowego w codziennych dojazdach do stolicy.

Oprócz „wariantu podstawowego”, opartego na założeniach przyjętych uprzednio w poszczególnych częściach raportu, którego rezultaty pokazane są na rysunku 8, dla każdego z założeń (a) – (c) przyjęto dwa dodatkowe warianty – „niski” i „wysoki”. Dla współczynnika skolaryzacji oraz odsetka pracujących studentów odpowiednie założenia wynoszą w obu tych wariantach 33.3% oraz 66.7%, natomiast dla udziału ruchu tranzytowego – 5% i 15%. Wydaje się, że wartości te można przyjąć za potencjalnie najszersze granice zmienności dla wymienionych parametrów. Wyniki korekt wyznaczonych dla różnych kombinacji powyższych założeń przedstawione są w tablicy 7.

Tablica 7. Analiza wrażliwości korekt liczby ludności Warszawy na zmiany wybranych założeń (tys. osób)

Założony procent pracujących wśród studentów w Warszawie	Współczynnik skolaryzacji brutto (WSB) dla szkolnictwa wyższego w Warszawie		
	WSB = 33.3%	WSB = 48.9%*	WSB = 66.7%
Założony tranzyt przez Warszawę = 5% łącznego ruchu			
33.3% studentów pracuje	811.6 (289.1)	795.7 (273.2)	777.7 (255.2)
46.6% studentów pracuje*	781.5 (259.0)	768.8 (246.3)	754.3 (231.8)
66.7% studentów pracuje	735.9 (213.4)	728.0 (205.5)	719.0 (196.5)
Założony tranzyt przez Warszawę = 10% łącznego ruchu*			
33.3% studentów pracuje	798.1 (303.1)	782.3 (287.3)	764.2 (269.2)
46.6% studentów pracuje*	768.0 (273.0)	*755.3 (260.3)	740.9 (245.9)
66.7% studentów pracuje	722.5 (227.5)	714.5 (219.5)	705.5 (210.5)
Założony tranzyt przez Warszawę = 15% łącznego ruchu			
33.3% studentów pracuje	784.7 (317.2)	768.8 (301.3)	750.8 (283.3)
46.6% studentów pracuje*	754.6 (287.1)	741.9 (274.4)	727.4 (259.9)
66.7% studentów pracuje	709.0 (241.5)	701.1 (233.6)	692.1 (224.6)

Liczby w komórkach tabeli dotyczą łącznej korekty dla ludności nocnej i dziennej oraz tylko nocnej (w nawiasach)

* Wartości przeciętne przyjęte w podstawowym wariantcie korekty przedstawionym w niniejszym opracowaniu

Źródło: obliczenia własne, dane i metodyka obliczeń jak dla rysunków 2 i 8.

Z tablicy 7 wynika, że zakres zmienności wyliczonych korekt nie jest duży: łączna korekta dla liczby ludności nocnej i dziennej waha się między 692.0 a 811.6 tysiąca osób, natomiast dla samej ludności nocnej – od 196.5 do 317.2 tysiąca osób. W obu przypadkach mamy więc do czynienia z modyfikacją wyniku o około ± 60.0 tysięcy osób w stosunku do wariantu podstawowego, co biorąc pod uwagę rząd wielkości obliczonych korekt, wskazuje, że powyższe oszacowania nie są bardzo wrażliwe na zmiany trzech wymienionych założeń.

Oprócz innych wspomnianych możliwości, badanie liczby ludności Warszawy mogłoby dodatkowo zawierać element formalnej analizy wrażliwości, opartej na metodologii statystycznej, zwłaszcza w ujęciu bayesowskim, gdzie sprawdzanie odporności wyników na zmiany założeń stanowi nieodzowny element analizy.

Wyliczone wielkości korekt należy traktować jako wartości wstępne, przybliżone i oparte na szeregu bardzo silnych założeń, wymienionych w poszczególnych częściach niniejszego opracowania. Ich weryfikacja wymaga przeprowadzenia osobnego badania, w całości poświęconego szacowaniu liczby ludności m. st. Warszawy. Przedstawione oszacowania mogą co najwyżej dostarczyć pewnych wstępnych informacji o rzędzie wielkości spodziewanych korekt oficjalnych szacunków: oprócz osób uwidoczniionych w statystykach, można spodziewać się dodatkowo około ćwierć miliona osób zamieszkałych w Warszawie oraz, oprócz tego, pół miliona osób dojeżdżających codziennie do miasta.

6.2. Wnioski i rekomendacje dla Urzędu m. st. Warszawy

W niniejszej ekspertyzie dokonano przeglądu potencjalnych źródeł danych i metod badawczych możliwych do zastosowania w celu dokonania korekty rzeczywistej liczby ludności Warszawy w stosunku do podważanych jako zaniżone oficjalnych statystyk GUS. Na podstawie dokonanego przeglądu danych i metod, w części 6.1 podjęto również próbę ogólnego oszacowania rzędu wielkości korekty liczby ludności stolicy, wykazując, że dla 2005 roku powinna ona wynieść około ćwierć miliona osób dla ludności nocnej i dodatkowo około pół miliona osób dla ludności dziennej, ponad wielkości podawane w oficjalnych statystykach. Okazało się ponadto, że wspomniane wyniki są dość mało wrażliwe na przyjęte roboczo dość silne założenia. Wstępne szacunki są jednak, jak zaznaczono we wstępie, bardzo istotne dla operacjonalizacji przyszłych badań, zwłaszcza dla rozważania kwestii doboru próby w przypadku podjęcia reprezentacyjnych badań społecznych.

Wobec wniosków płynących z niniejszego opracowania, postawić można pytanie o celowość lub racjonalność dalszych badań w tym zakresie, zważywszy na określone powyżej rzędy wielkości korekt zarówno dla dziennej, jak i nocnej ludności Warszawy. Niewątpliwie, jak wykazano w niniejszym studium, z metodologicznego punktu widzenia, istnieje cały szereg możliwych do wykorzystania danych i metod badawczych, dzięki którym można dokonać właściwej korekty ludności stolicy. Poszczególne metody wzajemnie się uzupełniają i weryfikują, prowadząc z jednej strony do otrzymania dokładniejszej informacji o rzeczywistej liczbie ludności stolicy, z drugiej zaś do możliwości dokładniejszego zbadania cech społeczno-demograficznych populacji nieujętej w statystykach GUS dotyczących ludności stałej bądź ludności faktycznie zamieszkałej w Warszawie. Dokładniejsza wiedza nie tylko o wysokości korekty ludności, ale również o jej rozkładzie według takich charakterystyk, jak wiek, płeć, dzielnica zamieszkania, powód przebywania w stolicy, oraz,

ewentualnie, powód braku zameldowania, pozwoliłaby niewątpliwie na lepsze planowanie rozwoju miasta i wyznaczanie różnych aspektów polityki miasta w zakresie życia społecznego, ekonomii, zagospodarowania przestrzennego, itp. Dostarczana przez dokładniejsze badania wiedza uwzględniałaby nie tylko korekty liczby ludności, ale też strukturę, potrzeby i charakter pobytu w stolicy ludności nieujętej w oficjalnych statystykach.

W świetle argumentacji przedstawionej w raporcie, do zgromadzenia dokładniejszej informacji na temat wysokości i struktury korekty dla liczby ludności Warszawy, można zarekomendować następujące metody:

- sondaż na reprezentatywnej próbie mieszkańców stolicy, ze względu na koszty mający raczej postać badania telefonicznego typu CATI (część 3.2);
- ewentualne powiązanie sondażu z analizą telemetryczną, bazującą na informacji ze zdjęć satelitarnych lub lotniczych (część 4.3);
- alternatywnie, oparcie wyliczeń na analizie wielkości korekt dla wybranych, najbardziej mobilnych grup ludności (część 3.1), w oparciu o hybrydową metodę definicyjno-bilansową według nomenklatury US Bureau of Census (część 5.2);
- opcjonalne uzupełnienie powyższych wyliczeń o formalną statystyczną analizę wrażliwości uzyskanych wyników na zmiany poszczególnych założeń (część 6.1);
- wykorzystanie do wyznaczenia liczby i struktury codziennych dojazdów do Warszawy (czyli liczby ludności dziennej) statystycznej analizy wyników Warszawskiego Badania Ruchu (WBR 2005) oraz ewentualnie modelu ruchu dla Warszawy, po jego wdrożeniu do użytkowania w Urzędzie Miasta (część 5.1);
- zastosowanie standardowych narzędzi statystycznych, takich, jak model Rogersa i Castro, czy metoda iteracyjnego dopasowania proporcjonalnego, do uzupełnienia brakującej informacji o strukturach ludności według wieku, dzielnicy zamieszkania, itp. (część 2.4).

Dla weryfikacji otrzymanych wyników, potencjalnie warto przeprowadzić równolegle kilka spośród wymienionych analiz.

Oprócz źródeł danych wymienionych w opracowaniu, obiecujące wydają się być również informacje pochodzące z Głównego Urzędu Geodezji i Kartografii, gdzie obecnie prowadzone są prace nad integracją różnych baz danych przestrzennych, jak również z portalu www.geoportal.gov.pl, które na przykład mogą służyć w przyszłości jako źródła zdjęć ortofotograficznych w przypadku stosowania metody opisanej w części 4.3. Warto skorzystać również z doświadczeń programu tzw. audytu miejskiego, gdzie zbierane są informacje o

warunkach życia w 258 miastach Unii Europejskiej (*Urban Audit*, www.urbanaudit.org, por. też prace D. Felczak i D. Stachowiak, 2005, oraz J. Paradysza, 2005b)⁶⁶. Nie rozważaliśmy natomiast wykorzystania na przykład źródeł policyjnych, które, pomijając kwestię ograniczonej dostępności, wydają się nam być bardzo selektywne, jak również systemów związanych z udzielaniem pomocy społecznej (PULS / POMOST), które są bardzo słabo udokumentowane. Te ostatnie mogą być użyteczne w ograniczonym zakresie (np. do oszacowania liczby bezdomnych), jednakże należy do nich podchodzić z pewnym sceptycyzmem, wychodząc z założenia, że znaczne obszary biedy pozostają poza jakąkolwiek rejestracją instytucji państwowych, podczas gdy rejestry bezrobocia mogą być w części zafałszowane ze względu na nieprawdziwe oświadczenia osób rejestrujących się.

Niezależnie od wyników i rekomendacji przedstawionych w opracowaniu, wydaje się również, że niewątpliwą bolączką z punktu widzenia planowania racjonalnej polityki miasta, jest oprócz nieadekwatności oficjalnych danych GUS, kwestia braku przepływu zebranych informacji między różnymi instytucjami publicznymi, zarówno centralnego, jak i niższego szczebla. Niniejsze studium wykazało istnienie dość zróżnicowanych źródeł potencjalnych danych, które mogą w sposób bezpośredni bądź pośredni posłużyć do szacunków liczby ludności stolicy (Ministerstwo Finansów, Mazowiecki Oddział Wojewódzki NFZ, dane o natężeniu i rozkładzie ruchu drogowego będące w posiadaniu Biura Drogownictwa i Komunikacji Urzędu m. st. Warszawy), a które nie znajdowały się początkowo w posiadaniu zleceńodawcy (Biura Naczelnego Architekta Urzędu Miasta). Wydaje się więc, że jedną z kwestii, która wymaga usprawnienia, niezależnie od przyszłych badań mających na celu szacunek korekty liczby ludności stolicy, jest lepszy przepływ posiadanych zasobów danych pomiędzy instytucjami publicznymi, tak by dane gromadzone przez jedną instytucję mogły służyć także innym instytucjom publicznym w realizacji ich celów ustawowych.

⁶⁶ Bardzo dziękujemy Prof. Janowi Paradyszowi za zwrócenie nam uwagi na wszystkie źródła danych wymienione w niniejszym akapicie.

Bibliografia

- Alho, J. M. (1990). Logistic Regression in Capture-Recapture Models. *Biometrics*, 46 (3): 623–635.
- Amaral, S., A. M. V. Monteiro, G. Camara, J. A. Quintanilha (2006). DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon. *International Journal of Remote Sensing*, 27 (5): 855–870.
- Bijak, J., I. Koryś (2006). *Statistics or reality? International migration in Poland*. Working Paper 3/2006. Central European Forum for Migration Research, Warszawa.
- Bishop, Y. M. M., E. F. Fienberg, P. W. Holland (1975). *Discrete multivariate analysis*. MIT Press, Cambridge, MA.
- BPRW [Biuro Planowania Rozwoju Warszawy] (2005). *Warszawskie Badanie Ruchu 2005, wraz z opracowaniem modelu ruchu*. Opracowanie wykonane na zlecenie Urzędu m. st. Warszawy. Biuro Planowania Rozwoju Warszawy SA, Warszawa.
- BPRW (2006). *Warszawskie Badanie Ruchu 2005 wraz z opracowaniem modelu ruchu. Raport V. Model ruchu – aktualizacja*. Opracowanie wykonane na zlecenie Urzędu m. st. Warszawy. Biuro Planowania Rozwoju Warszawy SA, Warszawa.
- Bracha, Cz. (1996). *Teoretyczne podstawy metody reprezentacyjnej*. PWN, Warszawa.
- Budzynowska, O., G. Węclawowicz (1984). Zmiany gęstości zaludnienia w mieście w ciągu doby na przykładzie Radomia. *Przegląd Geograficzny*, 56 (1–2): 141–153.
- Castro, L. J., A. Rogers (1983). Patterns of Family Migration: Two Methodological Approaches. *Environment and Planning A*, 15 (2): 237–254.
- Chmielewska, M. (1994). Wybrałam bycie motłochem... Z Siostrą Małgorzatą Chmielewską rozmawia Grzegorz Chrzanowski OP. *Znak*, 46 (5): 99–111, www.opoka.org.pl/biblioteka/T/TA/TAC/bycie_motloch.html
- Cichomski, B. (2004). Zróżnicowanie społeczne mieszkańców Warszawy w perspektywie porównawczej. [w:] J. Grzelak, T. Zarycki (red.), *Spoleczna mapa Warszawy. Interdyscyplinarne studium metropolii warszawskiej*, Scholar, Warszawa: 11–42.
- Faber Maunsell/AECOM (2005). *Analiza i ocena funkcjonowania komunikacji zbiorowej w Warszawie i aglomeracji warszawskiej. Raport końcowy*. Opracowanie wykonane na zlecenie Biura Komunikacji Urzędu m. st. Warszawy, Warszawa.
- Felczak, D., D. Stachowiak (2005). Porównanie warunków życia w miastach europejskich (program Urban Audit II). *Wiadomości Statystyczne*, L (11): 102–118.
- Griffith, D. A., R. J. Bennett, R. P. Haining (1989). Statistical analysis of spatial data in the presence of missing observations: a methodological guide and an application to urban census data. *Environment and Planning A*, 21 (11): 1511–1523.
- GUS [Główny Urząd Statystyczny] (2003). *Raport z wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2002*. Główny Urząd Statystyczny, Warszawa.

- GUS [Główny Urząd Statystyczny] (2006a). *Przegląd Statystyczny. Warszawa. rocznik XIV*, nr 4. Urząd Statystyczny w Warszawie, Warszawa.
- GUS [Główny Urząd Statystyczny] (2006b). *Szkoły wyższe i ich finanse w 2005 r.* Główny Urząd Statystyczny, Warszawa.
- GUS [Główny Urząd Statystyczny] (2006c). *Panorama dzielnic Warszawy w 2004 r.* Główny Urząd Statystyczny, Warszawa.
- Harvey, J. T. (2002). Estimating census district populations from satellite imagery: some approaches and limitations. *International Journal of Remote Sensing*, 23 (10): 2071–2095.
- Henderson, F. M., X. Zong-Guo (1997). SAR applications in human settlement detection, population estimation and urban land use pattern analysis: a status report. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (1): 79–85.
- Itahashi, S., M. Komada, M. Takeuchi (2006). Methods for Estimating the Nitrogen Load on a Catchment Scale. *Japan Agricultural Research Quarterly*, 40 (2): 131–142.
- Józwiak, J., J. Podgórski (1997). *Statystyka od podstaw*. PWE, Warszawa.
- Kołąkowska, Z. (2006). *O osobliwościach rozwoju badań sondażowych w Polsce*. Materiały Zakładu Statystyki, Demografii i Socjologii Matematycznej, Instytut Socjologii Uniwersytetu Warszawskiego, Warszawa. Url: stat1.is.uw.edu.pl.
- Kośny, A. (2006). Radar z syntetyczną aperturą. *Przegląd Sił Powietrznych*, 4 (4): 43–49.
- Kupiszewska, D., B. Nowok, M. Kupiszewski (2006). Country report: Poland. [w:] M. Poulain, N. Perrin, A. Singleton (red.), *THESIM – Towards Harmonised European Statistics on International Migration*. Presses Universitaires de Louvain, Louvain-la-Neuve: 577–589.
- Kupiszewski, M., J. Bijak (2006). *Ocena prognozy ludności GUS 2003 z punktu widzenia aglomeracji warszawskiej*. Working Paper 1/2006. Central European Forum for Migration Research, Warszawa.
- Maktav, D., F. S. Erbek, C. Jürgens (2005). Remote sensing of urban areas. *International Journal of Remote Sensing*, 26 (4): 655–659.
- Norman, P. (1999). *Putting Iterative Proportional Fitting on the Researcher's Desk*. Working Paper 99/03. School of Geography, University of Leeds, Leeds.
- Olorunfemi, J. F. (1982). Applications of aerial photography to population estimation in Nigeria. *GeoJournal*, 6 (3): 225–230.
- ONZ [Organizacja Narodów Zjednoczonych] (1998). *Recommendation on statistics of international migration*. United Nations Population Division, New York.
- Paradysz, J. (2005a). *Prognozowanie demograficzne dla małych obszarów na przykładzie miasta Poznania w latach 2005–2030*. Referat przedstawiony na VI Konferencji Naukowej nt. „Statystyka regionalna. Wielowymiarowa analiza statystyczna. Metoda reprezentacyjna w badaniach ekonomiczno-społecznych.” Poznań-Kiekrz, 13–15 czerwca 2005 r.

- Paradysz, J. (2005b). Audyt miejski. [w:] A. Zeliaś (red.), *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*. Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków: 205–214.
- Potrykowska, A., P. Śleszyński (1999). *Migracje wewnętrzne w Warszawie i województwie warszawskim*. Atlas Warszawy, zeszyt 7. Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa.
- Raport o stanie m. st. Warszawy* (2004). Urząd m. st. Warszawy, Warszawa.
- Rees, P. (1994). Estimating and projecting the populations of urban communities. *Environment and Planning A*, 26 (11): 1671–1697.
- Rogers, A., L. J. Castro (1981). *Model Migration Schedules*. IIASA Report RR 81 30. International Institute for Applied Systems Analysis, Laxenburg.
- Schnore, L. F. (1960). Three Sources of Data on Commuting: Problems and Possibilities. *Journal of the American Statistical Association*, 55 (289): 8–22.
- Sutton, P., D. Roberts, C. Elvidge, K. Baugh (2001). Census from Heaven: an estimate of the global human population using night-time satellite imagery. *International Journal of Remote Sensing*, 22 (16): 3061–3076.
- Strategia zrównoważonego rozwoju systemu transportowego Warszawy na lata 2007-2013 i dalsze (projekt, wersja 30 wrzesień)* (2006). Urząd m. st. Warszawy, nie publikowane.
- Śleszyński, P. (2004a). Regionalne różnice pomiędzy liczbą ludności według narodowego spisu powszechnego w 2002 r. i rejestrowaną na podstawie ewidencji bieżącej. *Studia Demograficzne*, 145 (1): 93–103.
- Śleszyński, P. (2004b). Różnice liczby ludności wykazane w NSP 2002 – Suplement. *Studia Demograficzne*, 146 (2): 104–109.
- Śleszyński, P. (2005a). Różnice w spisie ludności ujawnione w Narodowym Spisie Powszechnym 2002. *Przegląd Geograficzny*, 77 (2): 193–212.
- Śleszyński, P. (2005b). *Szacunek liczby pracujących w Warszawie w 2005 roku według sekcji PKD w układzie rejonów komunikacyjnych*. Opracowanie wykonane na zlecenie Urzędu m. st. Warszawy. Instytut Geografii i Zagospodarowania Przestrzennego PAN, Warszawa.
- Tobler, W. R. (1969). Satellite Confirmation of Settlement Size Coefficients. *Area*, 1 (3): 30–34.
- Vining Jr., D. R., S. J. H. Louw (1978). A cautionary note on the use of the allometric function to estimate urban populations. *The Professional Geographer*, 30 (4): 365–370.
- Wu, C., A. T. Murray (2005). A cokriging method for estimating population density in urban areas. *Computers, Environment and Urban Systems*, 29 (5): 558–579.

Źródła informacji bezpośredniej

[1] *Główny Urząd Statystyczny*

[1a] Departament Statystyki Społecznej, Al. Niepodległości 208, 00-925 Warszawa

[1b] Centralne Informatorium Statystyczne, Departament Udostępniania Informacji

e-mail: dane@stat.gov.pl. Url: www.stat.gov.pl

[2] *Instytut Statystyki Kościoła Katolickiego*

Księża Pallotyni, ul. Skaryszewska 12, 03-802 Warszawa

e-mail: iskk@pallotti.ecclesia.org.pl. Url: www.iskk.ecclesia.org.pl

[3] *Ministerstwo Finansów*

[3a] Departament Eksploatacji Systemów Informatycznych, ul. Świętokrzyska 12, 00-916 Warszawa

[3b] Departament Podatków Dochodowych

Url: www.mf.gov.pl

[4] *Narodowy Fundusz Zdrowia*

Mazowiecki Oddział Wojewódzki NFZ, Chałubińskiego 8, 00-613 Warszawa

e-mail: oddzial@nfz-warszawa.pl. Url: www.nfz-warszawa.pl

[5] *Państwowa Komisja Wyborcza*

Krajowe Biuro Wyborcze – Zespół Prawny i Organizacji Wyborów, Pl. Bankowy 3/5, 00-950 Warszawa

e-mail: ziw-mli@poczta.kbw.gov.pl. Url: www.pkw.gov.pl

[6] *Urząd m. st. Warszawy*

[6a] Biuro Naczelnego Architekta Miasta, PKiN, Pl. Defilad 1, 00-901 Warszawa

e-mail: bnam@warszawa.um.gov.pl

[6b] Biuro Drogownictwa i Komunikacji, ul. Górskiego 7, 00-033 Warszawa

e-mail: drogownictwo@warszawa.um.gov.pl

[6c] Biuro Administracji i Spraw Obywatelskich, ul. Nowogrodzka 43, 00-691 Warszawa

e-mail: ao-elido@srodmiescie.warszawa.pl

[7] *Zarząd Dróg Miejskich*

Wydział Badań Ruchu i Rozwoju Dróg, ul. Chmielna 120, 00-801 Warszawa

e-mail: zdm@zdm.waw.pl. Url: www.zdm.waw.pl

[8] *Koleje Mazowieckie Sp. z o.o*

Biuro Rzecznika Prasowego, ul. Lubelska 1, 03-802 Warszawa

e-mail: sekretariat@mazowieckie.com.pl. Url: www.mazowieckie.com.pl



CEFMR Working Papers Series

- 1/2003: M. Kupiszewski, *Consequences of EU enlargement for freedom of movement between Council of Europe Member States*, ISBN 83-920313-0-X.
- 2/2003: M. Kupiszewski, D. Kupiszewska, *Internal migration component in subnational population projections in Member States of the European Union*, ISBN 83-920313-1-8.
- 3/2003: M. Kupiszewski, J. Bijak, K. Saczuk, R. Serek, *Komentarz do założeń prognozy ludności na lata 2003–2030 przygotowywanej przez GUS*, ISBN 83-920313-2-6.
- 4/2003: K. Saczuk, *A development and critique of the concept of replacement migration*, ISBN 83-920313-3-4.
- 5/2003: I. Koryś, *Migration Trends in Selected EU Applicant Countries: Poland*, ISBN 83-920313-4-2.
- 1/2004: A. Kicingier, K. Saczuk, *Migration policy in the European perspective – development and future trends*, ISBN 83-920313-5-0.
- 2/2004: A. Kicingier, *International migration as a non-traditional security threat and the EU responses to this phenomenon*, ISBN 83-920313-6-9.
- 3/2004: J. Bijak, *Fertility and mortality scenarios for 27 European countries, 2002–2052*, ISBN 83-920313-7-7.
- 4/2004: J. Bijak, M. Kupiszewski, A. Kicingier, *International migration scenarios for 27 European countries, 2002–2052*, ISBN 83-920313-8-5.
- 5/2004: K. Saczuk, *Labour force participation scenarios for 27 European countries, 2002–2052*, with Annex, ISBN 83-920313-9-3.
- 1/2005: J. Bijak, D. Kupiszewska, M. Kupiszewski, K. Saczuk, *Impact of international migration on population dynamics and labour force resources in Europe*, with Annex, ISBN 83-921915-0-1.
- 2/2005: A. Kicingier, *Unia Europejska wobec zagadnienia integracji imigrantów*, ISBN 83-921915-1-X.
- 3/2005: I. Koryś, *Dimensions of Integration: Migrant Youth in Poland*, ISBN 83-921915-2-8.
- 4/2005: A. Kicingier, *Polityka emigracyjna II Rzeczypospolitej*, ISBN 83-921915-3-6.
- 5/2005: B. Nowok, D. Kupiszewska, *Official European statistics on international migration flows: availability, sources and coverage*, ISBN 83-921915-4-4.
- 6/2005: J. Bijak, *Bayesian methods in international migration forecasting*, ISBN 83-921915-5-2.
- 7/2005: D. Kupiszewska, B. Nowok, *Comparability of statistics on international migration flows in the European Union*, ISBN 83-921915-6-0.
- 8/2005: B. Nowok, *Evolution of international migration statistics in selected Central European countries*, ISBN 83-921915-7-9.
- 9/2005: A. Kicingier, *Between Polish interests and the EU influence – Polish migration policy development 1989–2004*, ISBN 83-921915-8-7.
- 10/2005: D. Kupiszewska, M. Kupiszewski, *A revision of a traditional multiregional model to better capture international migration: The MULTIPOLES model and its applications*, ISBN 83-921915-9-5.
- 1/2006: M. Kupiszewski, J. Bijak, *Ocena prognozy ludności GUS 2003 z punktu widzenia aglomeracji warszawskiej*, ISBN 83-60462-00-3.
- 2/2006: M. Mazur-Rafał, *Zmiana paradygmatu w niemieckiej polityce imigracyjnej w latach 1998–2004? Wnioski dla Polski*, ISBN 83-60462-01-1.
- 3/2006: J. Bijak, I. Koryś, *Statistics or reality? International migration in Poland*, ISBN 83-60462-02-X.

- 4/2006: J. Bijak, *Forecasting international migration: Selected theories, models, and methods*, ISBN 83-60462-03-8
- 5/2006: I. Koryś, *Między instrumentalnym a symbolicznym wymiarem posiadania drugiego obywatelstwa*, ISBN 83-60462-04-6
- 6/2006: M. Kupiszewski, J. Bijak, B. Nowok, *Impact of future demographic trends in Europe*, ISBN 83-60462-05-4
- 1/2007: A. Kicingier, A. Weinar (eds.), *State of the Art of the Migration Research in Poland*, ISBN 978-83-60462-06-5
- 2/2007: J. Bijak, A. Kicingier, M. Kupiszewski, współpraca P. Śleszyński, *Studium metodologiczne oszacowania rzeczywistej liczby ludności Warszawy*, ISBN 978-83-60462-07-2**

All CEFMR Working Papers are available at: «www.cefmr.pan.pl»

ISSN 1732-0631
ISBN 978-83-60462-07-2

Printed in Poland