

Zygmunt Gostkowski

**OSOBY NIEDOSTĘPNE DLA WYWIADÓW
I MIARY STOPNIA REPREZENTATYWNOŚCI PRÓB BADAWCZYCH
CBOS-u JAKO WSKAŹNIKI STOSUNKU SPOŁECZEŃSTWA
DO SONDAŻY OPINII PUBLICZNEJ**

Analiza danych dotyczących liczebności wypadków niedochodzenia do skutku wywiadu oraz zapisanych przez ankierów okoliczności i przyczyn tego zjawiska pozwala podjąć następujące problemy i dokonać próby ich rozwiązania.

1. Stopień reprezentatywności próby zrealizowanej, który zależy od liczebności respondentów niedostępnych w danym badaniu (tj. tych, z którymi nie udało się przeprowadzić wywiadów) oraz od rodzaju tej niedostępności (na przykład inne konsekwencje dla stopnia reprezentatywności będzie miało nieprzeprowadzenie wywiadu pod wskazanym adresem, gdyż okazało się, że pod danym numerem nie ma domu – a inne te wypadki, gdy respondent odmówił wywiadu itp.).

2. Ogólne nastawienie (postawy) respondentów wobec występowania w roli respondenta w sondażu CBOS-u; podjęcie tego problemu od strony niedosłysz do skutku wywiadów pozwala rozpatrywać go z punktu widzenia behawioralnych i psychospołecznych barier, stojących na drodze do uzyskania maksymalnej kompletności, a więc reprezentatywności próby zrealizowanej – reprezentatywności wyników badań dla określonego universum. W szczególności istotne dla **tego problemu** jest uwzględnienie różnego rodzaju barier (na przykład z jednej strony takich jak niesprawność psychofizyczna wylosowanego respondenta czy jego niepełnoletniość, a z drugiej – odmowa udzielenia wywiadu z powodu niechętnego nastawienia do badań). Kluczowe znaczenie będzie tu miało liczbowe zróżnicowanie występowania wypadków ingerencji poszczególnych barier o różnym stopniu negatywnego wpływu na reprezentatywność uzyskanej próby (na przykład wypadki odmowy wywiadu z powodu obawy wpuszczenia ankierę do mieszkania – a odmowy z powodu niechęci do wypowiadania się na dany temat czy obawy przed konsekwencjami swoich wypowiedzi).

3. Struktura (skład) pod względem głównych cech demograficznych i społeczno-zawodowych (oraz najprawdopodobniej skorelowanych z nimi postaw i opinii) zbiorowości wylosowanych do badania osób, z którymi nie udało się przeprowadzić wywiadu. Czy występuje stałość tej struktury w kolejnych sondażach? Czy wśród tej zbiorowości występują jakieś kategorie osób „nadreprezentowane” liczebnie w stosunku do proporcji ich występowania w uniwersum, dla którego chcemy uzyskać reprezentatywność badania? Inaczej mówiąc – czy istnieją takie kategorie osób, dla których szanse znalezienia się w próbie **zrealizowanej** są mniejsze niż dla innych kategorii? (Można by wysunąć hipotezę, że owymi kategoriami o mniejszych szansach znalezienia się w zrealizowanej próbie będą osoby o niższym poziomie wykształcenia, kobiety i osoby starsze).

Powtarzalność – stałość – struktury zbiorowości osób, z którymi nie udało się uzyskać wywiadu w kolejnych badaniach, pozwoli na potraktowanie cech takiej zbiorowości jako charakterystycznych dla stałych „odmawiaczy” uczestnictwa w sondażach CBOS-u, a więc zbiorowości niedostępnej sondażom. Wskazywać to będzie, pod względem jakich cech próby **zrealizowane** są tendencyjne, tj. w jakimś stopniu umniejszają systematycznie liczebność pewnych, określonych kategorii osób.

4. Ingerencja badacza poprzez listy zapowiadające wizytę ankietera – w celu zmniejszenia liczby odmów wywiadu.

A. REPREZENTATYWNOŚĆ PRÓB ZREALIZOWANYCH

W zasadzie stopień reprezentatywności **uzyskanej** ostatecznie do badania próby losowej zależy nie tylko od poprawności operatu losowania (który powinien spełniać podstawowy warunek reprezentatywności, mówiący, że każdy element uniwersum mającego być reprezentowanym w próbie powinien mieć taką samą szansę znalezienia się w niej). Nawet, jeśli warunek ten jest spełniony, to zawsze pozostaje problem praktyczny faktycznego dotarcia do wylosowanych elementów; liczebność elementów niedostępnych badaniu z różnych względów określa w istocie stopień reprezentatywności **efektywnej** próby – tj. próby poddanej badaniu.

W praktyce CBOS-u losuje się najpierw rejony statystyczne, a w każdym z nich sześć mieszkań (tj. adresów)¹. Następnie w wylosowanym adresie

¹ Rejony statystyczne różnią się znacznie liczbą adresów – a więc mieszkańców będących potencjalnymi respondentami. Najliczniej zaludnione rejony są w miastach, a najrzadziej na wsiach. Tak więc, mieszkaniec rejonu miejskiego ma mniejszą szansę znalezienia się w wylosowanej próbie niż wiejskiego; mieszkańcy wsi są więc w próbach nadreprezentowani, a miast niedoreprezentowani. Te wypaczenia w „surowych” próbach stara się CBOS korygować przez stosowne

losuje się do wywiadu, spośród pełnoletnich osób na stałe tam mieszkających, jedną osobę – jako respondenta. Nie przewiduje się dobierania zastępczych respondentów w miejsce tych, z którymi nie udało się przeprowadzić wywiadu. Uzyskiwane w ten sposób zaprojektowane próby liczą najczęściej 1500 adresów.

Zbiorowość próby **zaprojektowanej** ulega zmniejszeniu w trakcie **realizacji** badań – w wyniku barier różnego rodzaju. Charakter tych barier decyduje, czy „odsiew” elementów z próby wylosowanej miał charakter losowy, czy – przypuszczalnie – systematyczny, tj. tendencyjny.

Tak więc, zbiór wylosowanych adresów/osób dzieli się w efekcie końcowym na następujące podzbiory:

1) zbiór wylosowanych adresów, które okazały się puste (brak domu w oczekiwanym miejscu, dom niemieszkalny, mieszkanie niezamieszkałe itp.), oraz zbiór adresów, pod którymi nikogo nie zastano pomimo ponawiania wizyty; we wszystkich tych wypadkach nie dochodzi do kontaktu z jakąkolwiek osobą pod wylosowanym adresem;

2) zbiór adresów realnych, a więc zbiór wypadków nawiązania kontaktu z osobą czy osobami tam zamieszkałymi; zbiór ten dzieli się na dwa podzbiory:

2.1 – zbiór wypadków niedojścia do skutku wywiadu, z różnych przyczyn;

2.2 – zbiór wypadków dojścia do skutku wywiadu; liczebność tego zbioru to liczebność **próby zrealizowanej**.

Zobaczmy, jak kształtowały się liczebności owych zbiorów w każdym z 12 kolejnych sondaży poddanych analizie (tab. 1).

Procentowy „odsiew” adresów „pustych” i takich, pod którymi nikogo nie zastano, a więc odsiew, o którym można założyć, że ma charakter przypadkowego „odlosowania” (a więc nie wpływa na stopień reprezentatywności **umniejszonej** w ten sposób próby wyjściowej) – zawiera się w granicach 8,8 – 14,7%, przy czym w 10 wypadkach na 12 przekracza nieco 10% (mieści się w zakresie 10,4–14,7%).

Patrząc diachronicznie na trend – można stwierdzić, że ten rodzaj ubytku prób wyjściowych zwiększał się: podczas gdy dla pierwszych 6 sondaży wynosił przeciętnie 10,3%, to dla końcowych 6 sondaży aż 13,0%; niepodobna stwierdzić, jakie czynniki w praktyce doboru prób wyjściowych i pracy sieci ankierskiej mogły mieć na to wpływ. Należy podkreślić, że ów zwiększający się odsiew nie ma związku z postawami występującymi wśród kandydatów na respondentów wobec badań czy CBOS-u; wynika on z postępującej dezaktualizacji dokumentacji adresowej i – być może – wskazuje na pogarszającą się pracę ankierów (na przykład niedochodzenie do skutku wywiadów

przeliczenia, aby uzyskać w nich proporcje „ważone”. Nie udało się uzyskać bliższych informacji o metodzie przeliczeń.

Tabela 1

Udziały w wylosowanych próbach CBOS-u trzech głównych kategorii elementów próby

		III	III	III-IV	IV	IV-V	V	VI	VII	IX	X	XI	XII	Razem
		N %	N %	N %	N %	N %	N %	N %	N %	N %	N %	N %	N %	N %
Zbiór adresów wylosowanych		1500 100%	1500 100%	2250 100%	2250 100%	1500 100%	1500 100%	1500 100%	1500 100%	1500 100%	1500 100%	1500 100%	1500 100%	19 500 100%
Zbiór adresów „pustych” (adresy nieistniejące i adresy, gdzie nikogo nie zastano)		161 10,7	138 9,2	197 8,8	235 10,4	189 12,6	152 10,1	165 11,0	209 13,9	180 12,0	187 12,5	207 13,8	221 14,7	2241 11,5
Zbiór adresów realnych, pod którymi nawiązano kontakt z jakąś osobą, wśród których:	liczba wypad- ków niedojś- cia do skut- ku wywiadu	174 11,6	164 10,9	230 10,2	208 9,2	141 9,4	172 11,5	112 7,5	166 11,1	168 11,2	101 6,7	161 10,7	91 6,1	1888 9,7
	liczba wypad- ków dojścia do skutku wywiadu	1165 77,7	1198 79,9	1823 81,0	1807 80,3	1170 78,0	1176 78,4	1223 81,5	1125 75,0	1152 76,8	1212 80,8	1132 75,5	1188 79,5	15 371 78,8

Liczbę rzymskie w główkach oznaczają miesiąc(e), kiedy przeprowadzano sondaż.

usprawiedliwione rzekomym brakiem posesji lub niemożnością odnalezienia adresu, zaniechanie powtórnych wizyt itp.). Zagadnienie to wymagałoby wyjaśnienia przez organizatorów sieci i kontrolę pracy ankierów².

Drugi rodzaj ubytków z próby wyjściowej – to wywiady niedoszące do skutku pomimo nawiązania kontaktu z osobą czy osobami zamieszkałymi pod wylosowanym adresem; ubytków tych nie można traktować jako „odlosowania”, gdyż wynikają one z przyczyn leżących po stronie osób z próby adresowej, a zatem kształtują się w sposób tendencyjny (wiążą się w większości z różnymi negatywnymi reakcjami na prośbę o wywiad u niedoszących respondentów). Procentowy ich udział w próbie wyjściowej zawiera się w zakresie zmienności od 6,1 do 11,6%, przy czym ma on tendencję do zmniejszania się: podczas gdy dla pierwszych 6 sondaży przeciętny procent tego typu „odsiewu” wynosi 10,5%, to dla kolejnych 6 sondaży już tylko 8,9% i począwszy od 4 sondaży w 7 przypadkach na 8 jest mniejszy niż procenty ubytków raczej losowych (spowodowanych „pustymi” adresami).

Biorąc pod uwagę to, co powiedziano już o losowości ubytków z próby pierwotnej, spowodowanych „pustymi” adresami – ocena stopnia reprezentatywności próby zrealizowanej powinna brać za punkt odniesienia zbiór obejmujący tylko wypadki adresów realnych jako stuprocentową zbiorowość wyjściową; pozytywną miarą reprezentatywności zrealizowanej próby będzie więc procent przeprowadzonych wywiadów obliczony od liczby adresów realnych, tj. takich, pod którymi nawiązano kontakt z jakąś zamieszkałą tam osobą.

Dla kolejnych 12 sondaży procenty te wynoszą: 87; 88; 88,8; 89,7; 89,2; 87,2; 91,6; 87,1; 87,3; 92,3; 87,5; 92,9.

Widać więc, że stopień reprezentatywności próby **osób** namówionych do wywiadów jest wysoki: poza jej zasięgiem pozostaje zaledwie od 13 do 7,1% próby, którą można uważać za losową reprezentację uniwersum dorosłych mieszkańców kraju. Przeciętna wielkość tego residuum dla 12 sondaży wynosi 11%.

Intuicyjnie wydaje się, że wynikające z takiego „ubytku reprezentatywności” błędy rozkładów procentowych, obliczonych dla próby zrealizowanej, muszą być minimalne. Spróbujmy dokonać szacunku takiego błędu, używając do

² Założenie o losowości ubytków spowodowanych „pustymi” adresami trzeba traktować jako hipotezę; wydaje się, że jest ona mniej trafna, jeśli chodzi o podział w próbie adresów na miasto i wieś, gdyż na wsi prawdopodobnie trafia się mniej adresów pustych. Zagadnienie to trzeba by zbadać, dokonując porównania procentu występowania adresów pustych dla wsi i dla miasta, aby stwierdzić, czy szansa ich pojawienia się jest taka sama w obydwu tych środowiskach. Można tu dodać, że wśród adresów pustych, w pierwszych 8 sondażach wypadki niezastania nikogo stanowiły zawsze większość (od 56% do 84%), przy czym procent ten wzrastał w kolejnych sondażach. Dla pozostałych 4 sondaży brak tej informacji.

tego celu przeciętnej liczebności osób, z którymi nawiązano kontakt pod realnymi adresami w 12 analizowanych sondażach. Taka przeciętna próba liczyłaby 1440 osób; przeciętny zaś ubytek, a więc obarczony tendencyjnością odsiew osób, które w jakiś sposób uchylły się od uczestnictwa w sondażu – 160 osób (obydwie liczby zaokrąglono do 10). Residuum 160 utraconych dla sondażu osób stanowi 11%, a więc przeciętną dla residuów z 12 sondaży.

Wyobraźmy sobie, że w próbie zrealizowanej, złożonej z 1280 (tj. 1440 – 160) osób – zdeklarowanych zwolenników agencji sondażowych, jako środka wspierającego „demokrację poprzez opinię publiczną”, było na przykład 60%, tj. 768 osób. Wyobraźmy sobie dalej, że wśród owych 160 osób, które – tak czy inaczej – uniknęły czy „wykręciły się” od udzielenia wywiadu w sondażu, **wszystkie** były przeciwnikami agencji sondażowych jako środka wspierania „demokracji poprzez opinię publiczną”. Gdyby jakimś sposobem udało się je namówić do udzielenia wywiadu, to po dodaniu ich do próby 1280 osób – procent zwolenników sondaży zmalałby

do $53\% = \frac{768}{1280 + 160} \times 100$. Różnica 7 punktów procentowych wydaje

się dość znacząca w porównaniu z 60% z próby tendencyjnej, bo obarczonej ubytkiem przeciwników sondaży na skutek selektywnego „odrzutu” 160 niedoszłych respondentów.

Jest to jednak wypadek skrajny, idealny; zakładając hipotetycznie i bardziej realistycznie, że wśród owych 160 niedoszłych respondentów tylko na przykład połowa odmówiła wywiadu motywowana niechęcią wobec agencji sondażowych, a z drugą połową nie przeprowadzono wywiadu z powodu czynników bardziej losowych, takich jak rzeczywisty brak czasu, nieobecność, chwilowa niedyspozycja (na przykład nietrzeźwość)³ itp. – uzyskujemy znacznie słabsze umniejszenie proporcji respondentów będących zwolennikami sondaży opinii publicznej: zamiast 60% w próbie zrealizowanej – w nowej hipotetycznej próbie, obejmującej niedoszłych respondentów (przy założeniu, że pierwsza połowa z nich byłaby przeciwnikami sondaży, a druga połowa – bądź zwolennikami, bądź nie miała zdania, przy czym było wśród nich 60% zwolenników sondaży, tj. 48 takich osób) – uzyskujemy

$$\frac{768 + 48}{1280 + 80 + 32} \times 100 = 58,6\%.$$

Przy tych, chyba najbardziej realistycznych założeniach, różnica między procentem zwolenników sondaży w próbie zrealizowanej (tendencyjnej)

³ Ten podział na dwie części niedoszłych respondentów może być oparty na empirycznych danych, gdyż ankietrzy w protokole niedojścia do skutku wywiadu rejestrują przyczyny tego faktu (o czym dalej).

a hipotetycznej próbie w pełni reprezentatywnej wynosiłaby tylko 1,4 punkta procentowego, czyli 5 razy mniej niż różnica poprzednia, uzyskana przy możliwie najskrajniejszym i najbardziej pesymistycznym założeniu co do wypaczającego efektu pominięcia 160 osób w próbie zrealizowanej.

Oszacowanie błędu tendencyjności próby umniejszonej o respondentów odmawiających czy unikających wywiadu może opierać się na mniej dowolnych i arbitralnych założeniach, jeśli zarejestrowane zostały w protokołach niedojścia do skutku wywiadu pewne istotne cechy niedoszłych respondentów (kategoria wieku, płeć, miejsce zamieszkania, na przykład miasto – wieś).

Jeśli z przeprowadzonych analiz statystycznych na zrealizowanej próbie (czy też na innych wcześniejszych próbach) wynika, że istnieje określony i stały związek między tymi cechami a daną opinią – to przypuszczalny jej rozkład w zbiorze niedoszłych respondentów może być oszacowany bardziej realistycznie i zbiór ten dołączony do próby zrealizowanej pozwoli na uzyskanie rozkładu bliższego rzeczywistości. Tego rodzaju korektury bywają dokonywane w agencjach sondażowych na „surowych” wynikach, przed ich ogłoszeniem – przy czym założenia, na których się opierają, i procedury przeliczeniowe nie są podawane do publicznej wiadomości, gdyż stanowią tajemnicę zawodową. W szczególności stosuje się takie szacunkowe korekty w sondażach przedwyborczych, których wyniki są weryfikowane *ex post* przez rzeczywiste rezultaty wyborów (frekwencję i rozkłady głosów na poszczególne partie czy/i kandydatów).

B. WPŁYW NEGATYWNEGO NASTAWIENIA NIEDOSZŁYCH RESPONDENTÓW WOBEĆ PROŚBY O UDZIELENIE WYWIADU NA STOPIEŃ REALIZACJI PRÓBY LOSOWEJ I JEJ REPREZENTATYWNOŚĆ

Różnorodność okoliczności i przyczyn niedochodzenia do skutku wywiadu w kolejnych 12 surveyach CBOS-u ukazuje tab. 2, zawierająca rozkłady procentowe 14 rodzajów tych okoliczności i przyczyn. Zgodnie z założeniami przyjętymi w poprzednim podrozdziale odnosi się ona tylko do zbiorowości adresów „realnych”, tj. do tych wypadków, w których pod odnalezionym adresem, w wylosowanym mieszkaniu nawiązano kontakt z jakąś zamieszkałą tam na stałe osobą, czy osobami, w celu wylosowania spośród nich respondenta do wywiadu. Osobę wylosowaną proszono o udzielenie wywiadu, jeśli była na miejscu; jeśli była poza mieszkaniem w momencie losowania, dowiadywano się, kiedy wróci, i zapowiadano na ten czas ponowną wizytę w celu przeprowadzenia wywiadu. Jeśli wylosowana osoba była w mieszkaniu, od razu proszono ją o wywiad, jeśli zaś na przykład spieszyła się do wyjścia w pilnej sprawie – umawiano się z nią na najbliższy termin dla niej dogodny.

Jak widać, sytuacje „startowe” do wywiadu miały najczęściej charakter grupowy, przede wszystkim dlatego że „kandydatami” do wywiadu byli na początku wszyscy zastani stali mieszkańcy wylosowanego lokalu mieszkalnego.

Zarejestrowane przez ankiera na podstawie obserwacji i wstępnej rozmowy z napotkanymi w wylosowanym mieszkaniu osobami, okoliczności i sytuacje, wiążące się z niedojściem do skutku wywiadu, mają różny stopień trafności jako wskaźniki ogólniejszego negatywnego nastawienia ludzi wobec przeprowadzania sondaży, występowania w roli respondenta czy też wobec agencji sondażowych jako instytucji (państwowych, rynkowych itp. – w zależności od sposobu ich postrzegania). Te okoliczności i sytuacje różnią się też znacznie pod względem miejsca, które zajmują na kontinuum przypadkowość–tendencyjność w generowaniu efektu ostatecznego, jakim jest niedojście do skutku wywiadu.

Rozpatrzmy z tego punktu widzenia każdy z 14 rodzajów okoliczności i sytuacji, aby dokonać ich stosownego pogrupowania i następnie skomasywania procentów występowania zbliżonych pod względem trafności oraz przypadkowości/tendencyjności rodzajów tych okoliczności i sytuacji.

Rzut oka na tab. 2 pozwala stwierdzić, duże rozdrobnienie 14 kategorii okoliczności i sytuacji, a więc mały stopień koncentracji frekwencji ich występowania. Stosunkowo najczęściej spotykamy dwa rodzaje tych okoliczności: „nieobecność wylosowanej osoby przez cały czas realizacji badań” oraz „odmowa wpuszczenia do mieszkania”. Przeciętnie, dla całej zbiorowości 17 259 kontaktów nawiązanych w „realnych” adresach, pierwszy z tych dwu rodzajów okoliczności wystąpił z częstotliwością 3,1%, a drugi 2,7%. Zajmują one w **każdym z 12 sondaży zawsze** pierwsze lub drugie miejsce pod względem częstotliwości występowania; zakres zmienności, w zależności od sondażu, zawiera się dla pierwszego rodzaju w granicach 1,8–5,3%, a dla drugiego 1,3–4,4%.

Patrząc diachronicznie na procenty wyrażające częstotliwość występowania omawianych dwu rodzajów okoliczności i sytuacji, trudno zauważyć wyraźne trendy od marca do grudnia. Jeśli jednak wziąć pod uwagę dla każdego ich rodzaju dwa przeciętne procenty – dla pierwszych sześciu sondaży (od marca do maja) i kolejnych sześciu sondaży (od czerwca do grudnia), to ujawnią się dwa przeciwne trendy. Częstość stałej nieobecności wylosowanej do wywiadu osoby wzrosła przeciętnie od 3 do 3,3%, a odmowy wpuszczenia do mieszkania – spadła z 3 do 2,5%.

Zwróćmy uwagę na trzy kategorie wypadków nieobecności wylosowanej do wywiadu osoby (wiersze 4, 5 i 6). Można domniemywać, że za częścią owych „nieobecności” kryją się „uniki” przed spotkaniem z ankierem, a więc ukryte odmowy udzielenia wywiadu; z drugiej strony – za wypadkami takimi mogła kryć się również rzeczywista nieobecność danej osoby w momencie wizyty ankiera. Jeśli przeważały liczbowo takie właśnie sytuacje, to

Tabela 2

Rodzaje okoliczności i przyczyn niedochodzenia do skutku wywiadów w 12 sondażach, w okresie od marca do grudnia 1992 r.

Rodzaje okoliczności i przyczyn	III 05		III 06		III-IV 07		IV 11		IV-V 12		V 13		VI ^a 15		VII 17		IX 19		X ^a 24		XI 27		XII ^a 33		Razem		Ranga częst. wy- stępowania
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
1. Odmowa wpuszczenia do miesz- kania	31	2,3	58	4,3	59	2,9	54	2,7	38	2,9	41	3,0	21	1,6	25	1,9	58	4,4	23	1,8	48	3,7	17	1,3	473	2,7	2
2. Odmowa informacji o osobach stałe zamieszkujących	6	0,4	11	0,8	16	0,8	8	0,4	8	0,6	2	0,1	8	0,6	4	0,3	3	0,2	1	0,1	3	0,2	3	0,2	73	0,4	8
3. Brak osoby odpowiadającej kryte- rium wyboru (np. niepełnoletnia lub zamieszkała czasowo)	—	—	—	—	1	0,04	—	—	1	0,1	4	0,3	3	0,2	—	—	1	0,08	1	0,1	1	0,1	—	—	12	0,07	14
4. Chwilowa nieobecność wylosowa- nej osoby	—	—	1	0,1	1	0,04	—	—	—	—	3	0,2	—	—	2	0,2	1	0,08	—	—	3	0,2	4	0,3	15	0,09	12,5
5. Nieobecność wylosowanej osoby przez cały okres realizacji badań	51	3,8	34	2,5	37	1,8	57	2,8	40	3,1	54	4,0	34	2,5	68	5,3	40	3,0	42	3,2	47	3,6	30	2,3	534	3,1	1
6. Nieobecność wylosowanej osoby w uzgodnionym wcześniej terminie	3	0,2	2	0,1	11	0,5	7	0,3	—	—	5	0,4	4	0,3	12	0,9	5	0,4	2	0,2	7	0,5	3	0,2	61	0,4	8
7. Niesprawność wylosowanej osoby (chora, głucha itp.)	16	1,2	10	0,7	22	1,1	15	0,7	9	0,7	14	1,0	7	0,5	13	1,0	10	0,8	6	0,5	10	0,8	11	0,9	143	0,8	4,5
8. Odmowa z powodu „braku czasu”	11	0,8	11	0,8	5	0,2	8	0,4	7	0,5	6	0,4	2	0,1	5	0,4	3	0,2	1	0,1	10	0,8	2	0,2	71	0,4	8
9. Odmowa z powodu niechęci do wypowiadania się na dany temat	9	0,7	7	0,5	19	0,9	16	0,8	9	0,7	12	0,9	4	0,3	7	0,5	9	0,7	6	0,5	12	0,9	2	0,2	112	0,6	6
10. Odmowa z powodu nie widzenia sensu (potrzeby) ankietowania	21	1,6	13	1,0	23	1,0	20	1,0	15	1,1	16	1,2	13	1,0	14	1,1	15	1,1	10	0,8	8	0,6	10	0,8	178	1,0	3
11. Odmowa uzasadniana obawą przed konsekwencjami wypowiedzi	3	0,2	—	—	2	0,1	1	0,05	4	0,3	—	—	1	0,07	1	0,08	—	—	—	—	2	0,2	2	0,2	16	0,09	12,5
12. Inne podawane powody odmowy	6	0,4	3	0,2	5	0,2	2	0,1	3	0,2	2	0,1	4	0,3	2	0,2	7	0,5	—	—	4	0,3	3	0,2	41	0,2	10,5
13. Odmowa bez podania przyczyn	17	1,3	12	0,9	24	1,2	19	0,9	5	0,4	11	0,8	6	0,4	12	0,9	14	1,1	4	0,3	5	0,4	3	0,2	132	0,8	4,5
14. Okresowa niedyspozycja (np. nie- trzeźwość)	—	—	2	0,1	5	0,2	1	0,05	2	0,2	2	0,1	5	0,4	1	0,08	2	0,2	5	0,4	1	0,1	1	0,08	27	0,2	10,5
15. Liczba zrealizowanych wywiadów	1165	87,0	1198	88,0	1823	88,8	1807	89,7	1170	89,2	1176	87,2	1223	91,6	1125	87,1	1152	87,3	1212	92,3	1132	87,5	1188	92,9	15 371	89,1	
16. Razem (liczebność kontaktów w realnych adresach)	1339	100	1362	100	2053	100	2015	100	1311	100	1348	100	1335	100	1291	100	1320	100	1313	100	1293	100	1279	100	17 259	100	

U w a g a: Procenty zaokrąglono do pierwszego miejsca dziesiętnego, aby uprościć ogląd danych w tabeli; z tego powodu zsumowania procentowe w rzeczywistości w większości wypadków dawałyby nie 100%, lecz liczby bardzo zbliżone, niące się do 100% kilkoma miejscami dziesiętnymi procenta; z powodu zaokrągleń występują – w przypadku małych liczb bezwzględnych – pewne anomalie w relacjach między procentami – na przykład dwie różne małe liczby względne wyrażone są takim samym procentem. Liczby rzymskie w główkach kolumn oznaczają miesiąc realizacji danego sondażu.

^a W danym sondażu zastosowano listy zapowiednie.

najczęściej powinny występować w typowo wakacyjnych miesiącach, gdy mieszkańcy miast wyjeżdżają na urlop⁴. Miesiącem takim jest w tab. 2 lipiec; otóż okazuje się, że właśnie w lipcu wystąpił najwyższy łączny procent (6,4%) trzech kategorii nieobecności, a zwłaszcza kategorii 5 (5,3%) oraz kategorii 6 (0,9%).

Tak więc, nieobecności wylosowanej do wywiadu osoby nie mogą stanowić w pełni trafnych wskaźników negatywnego nastawienia do udzielenia wywiadu, a są raczej efektem okoliczności obiektywnych – w dużym stopniu przypadkowych. Podobnie należy potraktować niktę liczebnie wypadki niedojścia do skutku wywiadu, z takich powodów, jak: „brak osoby odpowiadającej kryterium wyboru” (kategoria 3), „niesprawność wylosowanej osoby” (kategoria 7), czy jej chwilowa niedyspozycja (kategoria 14).

Drugi pod względem częstotliwości występowania wśród ogółu 17 259 kontaktów rodzaj niedochodzenia do skutku wywiadów to „odmowa wpuszczenia do mieszkania” (2,7%). Sprawia on poważne kłopoty interpretacyjne jako wskaźnik stosunku do wywiadu jako takiego. Są to sytuacje rozmowy z kimś w wylosowanym mieszkaniu przez domofon lub zamknięte drzwi. Należy przypuszczać, że za takimi wypadkami kryje się obawa przed wpuszczeniem do domu przestępcy; zastanawia jednak wspomniany znaczny przeciętny spadek częstotliwości (z 3 do 2,5%) takich sytuacji; przypisać go należy wpływowi listów zapowiednich, które zastosowano w trzech sondażach ich drugiej szóstki.

Ostatecznie zdecydowano się wziąć jako wskaźniki negatywnego nastawienia wobec próby ankietera o wywiad tylko najmniej kontrowersyjne pod tym względem reakcje osób czy osoby w wylosowanym mieszkaniu; dokonano więc skomasowania procentów występowania 7 kategorii: 2, 8, 9, 10, 11, 12, 13. Wśród nich „ostra” wskaźnikowo jest kategoria 10, tj. „odmowa z powodu niewiedzenia sensu, potrzeby ankietowania”. Występuje też ona względnie najczęściej (1,0%, tj. trzecia ranga częstotliwości), a więc wyraża takie stanowisko co setna osoba proszona o wywiad. Bliska znaczeniowo jest też kategoria 9 – „odmowa z powodu niechęci do wypowiadania się na dany temat (0,6% – szósta ranga). Kategorię 13 („odmowa bez podania przyczyn” – 0,8%, ranga 4,5) oraz kategorię 8 („odmowa z powodu braku czasu” – 0,4%, ranga 8) można też potraktować jako zdradzające ogólne negatywne nastawienie do badań. Oczywiście – najbardziej drastyczną wskaźnikowo jest „odmowa uzasadniona obawą przed konsekwencjami wypowiedzi”; jest ona pozostałością po epoce dyktatury komunistycznej, ale występuje w całkowicie znikomej liczbie wypadków (ogółem 0,09% i ostatnia ranga, tj. 16 osób reagujących na prośbę o wywiad w tak idiosynkratyczny sposób).

⁴ Wśród mieszkańców wsi częstsze nieobecności w lipcu mogą być spowodowane całodziwnym zaangażowaniem w prace żniwne.

Tabela 3 ukazuje poziom oraz ewolucję w czasie łącznego procentu występowania wymienionych siedmiu kategorii niedochodzenia do skutku wywiadu. Szereg czasowy ukazanych w zestawieniu procentów zdaje się świadczyć o polepszaniu się stopnia reprezentatywności prób CBOS-u w badanym okresie dziewięciu miesięcy: w grudniu odsiew z próby kontaktów nawiązanych w wylosowanych adresach „realnych” był niemal trzykrotnie mniejszy niż w marcu, a więc zaledwie co 50. osoby nie udało się skłonić do wywiadu z powodu jakichś negatywnych nastawień do roli respondenta, podczas gdy w marcu – około co 18. Wobec znacznych wahań w kolejnych sondażach omawianych procentowych wskaźników rozmiaru tendencyjnego odsiewu warto posłużyć się – jak poprzednio – dwoma przeciętnymi procentami dla dwu równolicznych grup sondaży. Otóż dla pierwszych sześciu sondaży procent ten wynosi 4,3%, a dla kolejnych sześciu spada do 2,9%, a więc aż o 1/3!

Ten znaczący spadek procentu odmów udzielenia wywiadu jest spowodowany tym, że w trzech sondażach drugiej szóstki zastosowano listy zapowiednie, których skutkiem jest znaczne zmniejszenie liczby odmów, będących wyrazem negatywnego nastawienia do próby o wywiad; przeciętny procent takich odmów w tych trzech sondażach wynosi zaledwie 2,17%, a w trzech pozostałych sondażach (bez listu zapowiedniego) aż 3,6% (bardziej szczegółowe omówienie skutków listu zapowiedniego znajdzie czytelnik w ostatnim paragrafie niniejszego studium).

Można z pewnością stwierdzić, że nawet maksymalnie skrajnie tendencyjny odsiew z próby reprezentatywnej, wynoszący zaledwie 2–3%, może spowodować (przy podanych niżej założeniach) wypaczenie rozkładów procentowych w hipotetycznie zrealizowanej próbie, rzędu zaledwie 1,5 punkta procentowego. Niech zilustruje to następujący przykład.

Próba reprezentatywna kontaktów z wylosowanymi realnymi mieszkaniami (gdzie nawiązano kontakt z kandydatami na respondentów) wynosiła 1000. Odmówiło wywiadu 3% – tj. 30 niedoszłych respondentów. Chodzi o rozkład opinii na pewne pytanie; założmy maksymalną tendencyjność tego odsiewu, tj. iż **wszystkich** 30 „odmawiaczy” wyraziłoby opinię negatywną. W próbie zrealizowanej, 970 osób, rozkład opinii pozytywnych i negatywnych wyniósł 50% i 50%. Gdyby udało się jednak uzyskać opinie od wszystkich zaplanowanych do wywiadu 1000 osób, to uzyskano by wśród nich następujący rozkład: $485 + (485 + 30) = 1000$ – czyli 485 opinii pozytywnych plus $485 + 30 = 515$ negatywnych. A zatem opinie negatywne w próbie w pełni zrealizowanej (tj. w próbie nietendencyjnej) byłoby 51,5%, czyli zaledwie o 1,5% więcej niż w uzyskanej próbie tendencyjnej. Jak widać, ów błąd zawyżenia negatywnych opinii na skutek maksymalnej tendencyjności odsiewu z powodu odmów stanowi 1,5 punkta procentowego. Oczywiście – wielkość ta pozostaje taka sama dla wszystkich absolutnych liczebności reprezentatywnych prób wylosowanych kontaktów.

Tabela 3

Wskaźnik reakcji na prośbę o wywiad wskazujący na najbardziej wyrażnie negatywne nastawienie do roli respondenta

	III N %	III N %	III-IV N %	IV N %	IV-V N %	V N %	VI ^a N %	VII N %	IX N %	X ^a N %	XI N %	XII ^a N %	Razem N %
Łączna liczba niedoszłych wywiadów z powodów wskazujących na negatywne nastawienie respondenta	73 5,5	57 4,2	94 4,6	74 3,7	51 3,9	49 3,6	38 2,8	45 3,5	51 3,9	22 1,7	44 3,4	25 2,0	623 3,6
Razem (liczebność kontaktów w realnych adresach)	1339 100	1362 100	2053 100	2015 100	1311 100	1348 100	1335 100	1291 100	1320 100	1313 100	1293 100	1279 100	17 259 100

U w a g a: Procenty obliczono z liczb bezwzględnych z tab. 2 – aby uniknąć błędów wynikających z zaokrąglenia procentów do pierwszego miejsca po przecinku.

^a List zapowiedni.

Wielkość błędu spowodowanego tendencyjnością odsiewu w wyniku odmów **zależy** jednak również od rozkładu opinii w próbie zrealizowanej: im bardziej rozkład ten odbiega od przyjętego w przykładzie 50 na 50%, tj. im bardziej zbliża się do 100% występowania tylko jednej z alternatywnych opinii, tym błąd wynikający z tendencyjności odsiewu w wyniku odmów jest większy. Na przykład gdyby w przykładzie opinie pozytywne i negatywne stanowiły nie 50 i 50%, lecz 70 i 30% – to analogicznie wyliczony błąd tendencyjności wynosiłby 2,1 punkta procentowego (tj. w faktycznie nieosiągalnej próbie w pełni reprezentatywnej byłoby 67,9% respondentów mających opinię pozytywną i 32,1% negatywną).

I ten procent błędu można uważać za minimalny – zbliżony do zwykłego błędu standardowego próby. W rzeczywistości jednak musi on być w realnych sytuacjach badawczych znacznie mniejszy, bo chyba nigdy nie będzie się zdarzać, że wśród osób utraconych na skutek odmowy wywiadu dokładnie **wszystkie** miałyby jeden tylko rodzaj opinii (na przykład wielu nie miało by żadnej opinii, tj. odpowiadałoby „nie wiem” czy „trudno powiedzieć”).

Sposoby szacowania wyrażonego procentowo błędu tendencyjności próby B_t , wynikłego z utracenia części zaplanowanych do wywiadu osób, można przedstawić za pomocą następującego wzoru matematycznego:

$$B_t = P_t - \left(\frac{\frac{N_t \times P_t}{100} + \frac{(N_r - N_t) \times (P_{Nr - Nt})}{100}}{N_r} \times 100 \right).$$

Po uproszczeniu, tj. zastąpieniu procentów przez proporcje, wzór przybiera postać:

$$B_t = P_t - \frac{N_t \times P_t + (N_r - N_t) \times (P_{Nr - Nt})}{N_r},$$

gdzie:

N_r – liczebność całej próby reprezentatywnej (tj. zaplanowanej próby kontaktów z kandydatami na respondentów);

N_t – liczebność uzyskanych efektywnie wywiadów (tj. liczebność próby zrealizowanej, a więc umniejszonej o osoby, które odmówiły wywiadu, czyli prawdopodobnie tendencyjnej);

$N_r - N_t$ – liczebność osób, które odmówiły wywiadu;

P_t – proporcja osób wyrażających daną opinię w próbie tendencyjnej;

$P_{Nr - Nt}$ – proporcja osób wyrażających daną opinię wśród osób, które odmówiły wywiadu, tj. w zbiorowości $N_r - N_t$.

Jeśli prawa część tego wzoru jest większa niż lewa – oznacza to, że w próbie zrealizowanej P_t (tendencyjnej) wystąpił niższy procent wyrażania danej opinii niż w hipotetycznej próbie w pełni reprezentatywnej; znak minus przy wartości uzyskanej ze wzoru oznacza więc, że w próbie tendencyjnej nastąpiło zaniżenie danego procentu (a odwrotnie – jeśli prawa część we wzorze jest mniejsza niż lewa).

Jak widać ze struktury podanego wzoru, błąd w procencie występowania danej opinii w próbie zrealizowanej zależy od trzech czynników: wielkości tego procentu w próbie zrealizowanej oraz liczebności (procentu) niezrealizowanych wywiadów (na skutek wypaczającego, tj. nielosowego czynnika, jakim są różnego rodzaju uniknięcia wywiadu). Podczas, gdy te dwa czynniki są badaczowi w CBOS-ie znane, trzeci czynnik, będący dla badacza zawsze niewiadomą, stanowi procent występowania danej opinii (ewentualnie jej brak) wśród niedoszłych respondentów. Tutaj badacz może tylko dokonywać hipotetycznych oszacowań tego procentu; na przykład może przyjąć hipotetycznie wersję maksymalnie pesymistyczną i optymistyczną takiego szacunku i do wyliczenia przyjąć wartość środkową.

W wersji maksymalnie optymistycznej badacz przyjmowałby, że procent, o który chodzi wśród niedoszłych respondentów, byłby taki sam – lub bardzo podobny – jak stwierdzony w próbie zrealizowanej. Pewne uzasadnienie takiego założenia istniałoby wówczas, gdyby sprawa, jakiej dotyczy opinia, o którą chodzi, miała całkowicie neutralny charakter dla faktu odmowy wywiadu.

Inaczej mówiąc – wedle zdroworozsądkowego przypuszczenia nie istniałaby korelacja między opiniami dotyczącymi tej sprawy a negatywną postawą uzewnętrznioną odmową wywiadu. Przykładem takiej sprawy byłoby na przykład zagadnienie upodobań do koloru opakowania jakiegoś produktu (w badaniach efektywności reklamy i preferencji konsumenckich). Przykładem przeciwnym byłaby sprawa deklarowanego zamiaru uczestnictwa (czy nieuczestnictwa) w najbliższych wyborach; tutaj najbardziej rozsądnym przypuszczeniem byłoby założenie, że istnieje silny związek między odmową (jako uzewnętrznieniem zgeneralizowanej negatywnej postawy wycofania się ze spraw publicznych, braku zainteresowania nimi itp.) a deklarowaną absencją wyborczą (jako przejawem tego negatywizmu w sferze politycznej). Wymagałoby to raczej hipotezy maksymalnie negatywnej co do zbiorowości niedoszłych respondentów: należałoby przypuszczać, że ogromna większość wśród nich (jeśli nie wszyscy) nie ma zamiaru iść do wyborów⁵. Możliwie trafny

⁵ Wobec stwierdzanego dotąd w Polsce faktu, że w odpowiedziach sondażowych respondenci silnie zawyżają swą przyszłą partycypację wyborczą (rzeczywista frekwencja w dotychczasowych wyborach była zawsze znacznie niższa niż frekwencja deklarowana), jest prawdopodobne, że wskaźnik negatywizmu wobec udziału w sondażu – jakim jest odmowa wywiadu – może pośrednio trafnie zwiastować absencję wyborczą niż deklaracje słowne wobec ankietiera.

szacunek błędu tendencyjności występującego w zrealizowanej próbie miałby tutaj szczególnie duże znaczenie, gdyż 1 punkt procentowy we frekwencji wyborczej (czy w liczbie głosów oddanych na kandydata) może przesądzić o wynikach wyborów (w zależności od ordynacji – jeśli na przykład wymaga, aby dla ważności wyników wyborów było 50% + 1 głosujących wśród wszystkich uprawnionych).

Podawane dotąd przykładowo sposoby szacowania błędu w próbie zrealizowanej, powstałego w wyniku niedochodzenia do skutku wywiadów z powodu odmowy czy innych sposobów ich uniknięcia, mają charakter raczej zdroworozsądkowych założeń i spekulacji. Wykorzystanie pewnych empirycznych wskazówek, mogących bardziej uprawdopodobnić takie założenia i szacunki, jest możliwe na podstawie szczegółowszej analizy cech zbiorowości osób, które odmawiają wywiadów w kolejnych sondażach. Zagadnieniu temu poświęcona jest kolejna część.

C. PŁEĆ, WIEK, WYKSZTAŁCENIE I MIEJSCE ZAMIESZKANIA JAKO CECHY ZWIĄZANE Z TENDENCYJNOŚCIĄ UBYTKU NIEDOSZŁYCH RESPONDENTÓW Z PRÓBY REALNYCH ADRESÓW

Zajmiemy się teraz zagadnieniem sformułowanym w pkt 3 wprowadzenia do niniejszego studium: chodziło będzie o to, jakie z wymienionych czterech cech predestynują ludzi proszonych o wywiad do uchylenia się od niego w taki czy inny sposób – a więc o to, czy i w jakim stopniu owe cztery rodzaje kategorii społecznych są nadreprezentowane – czy niedoreprezentowane – wśród „odmawiaczy” wywiadu? Częstsze odmawianie wywiadu charakteryzujące daną kategorię społeczną będzie dla nas behawioralnym wskaźnikiem jej bardziej negatywnego nastawienia do uczestnictwa w sondażach opinii – do przyjmowania roli respondenta. W konsekwencji – przedstawiciele takiej kategorii społecznej będą nadreprezentowani wśród „odmawiaczy”, a niedoreprezentowani w efektywnie uzyskanej do badań próbie. Analogicznie rzecz się ma z częstszym wyrażaniem zgody na wywiad: osoby takie będą niedoreprezentowane wśród „odmawiaczy”, a nadreprezentowane w próbach, co będzie wskazywało na ich bardziej pozytywny stosunek do sondaży i przyjmowania roli respondenta.

Tak określone przejawy nad- i niedoreprezentacji można traktować jako wskaźniki wyrażone w liczbie punktów procentowych, stanowiących różnicę w stosunku do stosownie dobranych procentów odniesienia. Jednakże możliwość skonstruowania takich wskaźników zależy od dostępności danych charakteryzujących zbiorowość osób, które – choć proszone – uchyliły się od wywiadu. Struktura płci, wieku i wykształcenia takich osób nie dała się odtworzyć z dwu powodów.

Po pierwsze – w tych wszystkich wypadkach, gdy z różnych przyczyn nie doszło do wylosowania respondenta spośród dorosłych osób stale zamieszkujących pod wylosowanym adresem; wchodzi tu w grę takie wypadki, jak odmowa wpuszczenia do mieszkania (kontakt przez domofon lub zamknięte drzwi), odmowa informacji o osobach stale tam zamieszkujących, brak osoby odpowiadającej kryteriom wyboru (bardzo rzadko). W istocie są to sytuacje, które sprawiły, że nie mogło wchodzić w grę skierowanie prośby o wywiad do konkretnej, wylosowanej osoby. W pozostałych wypadkach, gdy jednak wylosowano jakąś osobę, ale nastąpiła odmowa przez nią wywiadu, w przeważającej liczbie wypadków ankieterzy nie nanosili na „kartę realizacji badania” płci, wieku i poziomu wykształcenia takiej osoby. Zresztą informacji takich, poza wizualnie stwierdzoną płcią „odmawiacza”, w wielu wypadkach nie starano się uzyskać.

Jedyną informacją nanoszoną na „kartę realizacji badania”, w wypadku niedoszedłego do skutku wywiadu, była informacja o regionie kraju, województwie oraz miejscu zamieszkania (wieś oraz miasto wedle klas liczby mieszkańców).

Owe różne informacyjnie sytuacje nawiązywania kontaktu z niedoszłymi respondentami determinują sposób obliczania ich nad- i niedoreprezentacji pod względem wymienionych cech wśród ogółu „odmawiaczy” i respondentów. W wypadku kategorii płci, wieku i poziomu wykształcenia jedyna dostępna metoda ustalenia wskaźników nad- i niedoreprezentacji opierać się musi na procentach odniesienia ustalonych przez GUS dla ogółu dorosłej ludności w kraju; jednocześnie trzeba przyjąć założenie, że udziały procentowe tych kategorii w próbie reprezentatywnej osób proszonych o wywiad w realnych adresach były dostatecznie zbliżone (a właściwie takie same) do udziałów GUS-owskich, tj. występujących w uniwersum. Wówczas – na zasadzie, że układ nad- i niedoreprezentacji wśród zbiorowości efektywnych respondentów stanowi „lustrzane odbicie” takiegoż układu wśród niedoszłych respondentów („odmawiaczy”) – można wnioskować o hipotetycznej strukturze tych ostatnich.

Analiza oparta na tych założeniach pozwala odpowiedzieć na dwa następujące pytania badawcze:

a) u których kategorii osób w badanych zbiorowościach (w kolejnych 12 sondażach) występowało mniejsze, a u których większe prawdopodobieństwo zgody na wywiad, tj. przyjęcie roli respondenta CBOS-u (z czym wiązało się mniej lub bardziej pozytywne nastawienie do prośby o wywiad i – przypuszczalnie – do sondaży opinii publicznej);

b) w jakim stopniu to zróżnicowane prawdopodobieństwo (i przypuszczalne nastawienie do sondaży) wpływało wypaczająco na reprezentatywność uzyskiwanych prób badawczych pod względem płci, wieku i poziomu wykształcenia.

Tabele (4–6) ukazują układy wskaźników nad- i niedoreprezentacji pod względem płci, wieku i wykształcenia wśród respondentów w 12 kolejnych sondażach CBOS-u. Wskaźniki te obliczono odejmując procenty odniesienia, tj. udziały procentowe tych kategorii wśród ogółu dorosłej ludności (według GUS) od procentów znalezionych w próbach respondentów (nieduże różnice między ogólnymi liczebnościami respondentów w tych tabelach a liczebnościami w tab. 1 i 2 wynikają z niewliczenia do nich pewnych wypadków „braku danych”).

Tabele (4–6) ujawniają fakt systematycznie występujących wypaczeń reprezentatywności we wszystkich 12 sondażach i to zawsze w tym samym kierunku. Tak więc, występuje stała nadreprezentacja kobiet i – stanowiąca jej dopełnienie – taka sama niedoreprezentacja mężczyzn (zawierająca się w granicach od 0,6 do 5,1 punktów procentowych); przeciętna nad- i niedoreprezentacji dla 12 sondaży wynosi 2,79 punkta procentowego. Wartości te można przyjąć jako stałą w czasie miarę większej szansy kobiet na występowanie w roli respondenta (nie widać żadnego trendu w badanym okresie). Pozostaje jednak sprawą otwartą – czy owa „większa szansa” jest efektem ich bardziej pozytywnego nastawienia do próśby o wywiad, czy większej „uległości społecznej”⁶, czy też wynika z faktu, iż kobiety częściej przebywają w domu. Najprawdopodobniej wszystkie te czynniki występują łącznie – ich skumulowany wpływ daje ową stałą nadreprezentację – 2,79 punkta procentowego kobiet wśród respondentów.

Silniejsze wypaczenia reprezentatywności uzyskiwanych prób występują pod względem struktury wieku. Jak widać z tab. 5, w 11 kolejnych sondażach mamy stały układ nad- i niedoreprezentacji tych samych grup wiekowych: stale niedoreprezentowane (poza 1 wyjątkiem) są osoby w dwóch grupach wiekowych: 18–24 i 25–34 lat. Zakres niedoreprezentacji występującej w tych grupach zawiera się w granicach od –0,07 do –5,78 punktów procentowych, ze średnimi za cały okres wynoszącymi odpowiednio: –4,7 i –2,26 punktów procentowych. Widać stąd, że wraz ze spadkiem wieku aż ponad dwukrotnie wzrasta średnia niedoreprezentacji, co oznacza, że im młodsze osoby proszone są o wywiad, tym częściej odmawiają lub inaczej unikają przyjęcia roli respondenta (co hipotetycznie można traktować jak behawioralny wskaźnik ich bardziej negatywnego nastawienia do sondaży czy agencji sondażowych).

W trendach wskaźników niedoreprezentacji widać zarazem, że z czasem następował w omawianych grupach wzrost wskaźników „negatywizmu respondenckiego”: dla osób w wieku 18–24 lat, w okresie od marca do

⁶ Większą „uległość społeczną” (*social dependency*) kobiet wykryto w różnych badaniach psychospołecznych. Por. np. M. Oetzel, *Selected Bibliography on Sex Differences*, „Social Science Research Council”, Committee on Socialization and Social Structures, Stanford 1962.

Tabela 4

Nadreprezentacja i niedoreprezentacja w zrealizowanych próbach respondentów pod względem płci w 12 sondażach CBOS-u, w roku 1992 (w punktach procentowych)

Płeć	Miesiąc i kod sondażu oraz liczba respondentów w próbie zrealizowanej											
	III 05 N = 1168	III 06 N = 1207	III-IV 07 N = 1778	IV 11 N = 1812	IV-V 12 N = 1190	V 13 N = 1176	VI 15 N = 1222	VII 17 N = 1124	IX 19 N = 1153	X 24 N = 1211	XI 27 N = 1139	XII 33 N = 1188
Mężczyźni: 47,7%	-1,91	-1,97	-2,48	-3,5	-3,5	-2,89	-1,7	-0,6	-4,9	-5,1	-2,92	-1,99
Kobiety: 52,3%	+1,91	+1,97	+2,48	+3,5	+3,5	+2,89	+1,7	+0,6	+4,9	+5,1	+2,92	+1,99

U w a g a: Rozkład procentowy w próbie zaplanowanej wedle rozkładu podawanego przez GUS dla uniwersum w roku 1992.

Tabela 5

Nadreprezentacja i niedoreprezentacja w zrealizowanych próbach respondentów pod względem wieku w 12 sondażach CBOS-u, w roku 1992 (w punktach procentowych)

Wiek w przedziałach i procentowy udział klasy wieku wśród ogółu dorosłej ludności według GUS		Miesiąc i kod sondażu oraz liczba respondentów w próbie zrealizowanej											
		III 05 N = 1164	III 06 N = 1196	III-IV 07 N = 1798	IV 11 N = 1184	IV-V 12 N = 1164	V 13 N = 1209	VI 15 N = 1116	VII 17 N = 1146	IX 19 N = 1209	X 24 N = 1137	XI 27 N = 1187	XII 33 N = 1187
18-24:	13,7%	-4,76	-4,92	-	-5,64	-5,34	-2,36	-4,44	-5,19	-4,19	-3,61	-5,52	-5,78
25-34:	20,6%	-1,27	-1,1	-	-0,63	+0,43	-0,07	-2,57	-3,75	-4,28	-4,14	-4,42	-3,08
35-49:	30,5%	+1,08	+2,36	-	+0,87	+0,24	-0,09	+0,02	+1,67	+1,00	+0,93	+2,92	+2,36
50-59(64):	17,1%	-0,43	+1,88	-	+1,31	+0,89	-0,52	+3,5	+2,43	+1,31	+2,09	+2,78	-0,34
Powyżej 59(64):	18,2%	+4,57	+2,03	-	+3,99	+3,67	+2,93	+3,39	+4,74	+6,06	+4,63	+4,14	+6,74

U w a g a: Nad- i niedoreprezentacje (w punktach procentowych) obliczone od udziałów procentowych wyodrębnionych 5 klas wieku w całej populacji, według danych GUS na koniec roku 1992. Dwie ostatnie klasy wieku są zróżnicowane dla kobiet i mężczyzn: kobiety 50-59 i ponad 60 lat; mężczyźni 50-64 i ponad 65 lata. Dla sondażu 07 brak danych.

Tabela 6

Nadreprezentacja i niedoreprezentacja w zrealizowanych próbach respondentów pod względem poziomu wykształcenia w 12 sondażach CBOS-u, w roku 1992 (w punktach procentowych)

Poziom wykształcenia i jego procentowy udział wśród ogółu dorosłej ludności według GUS	Miesiąc i kod sondażu oraz liczba respondentów w próbie zrealizowanej											
	III 05 N = 1172	III 06 N = 1207	III-IV 07 N = 1775	IV 11 N = 1813	IV-V 12 N = 1190	V 13 N = 1164	VI 15 N = 1209	VII 17 N = 1121	IX 19 N = 1146	X 24 N = 1209	XI 27 N = 1140	XII 33 N = 1187
Wyższe: 7,3%	- 1,16	- 0,51	- 0,65	- 0,24	+ 0,52	- 1,11	- 1,51	+ 0,19	- 1,11	+ 1,00	+ 0,07	- 0,9
Średnie: 29,1%	- 2,39	- 0,52	- 2,4	- 3,89	- 2,8	+ 1,48	+ 1,34	- 2,87	- 2,5	- 0,8	- 1,03	- 3,4
Zasadnicze zawodowe: 27,1%	- 0,74	- 2,49	- 2,76	+ 1,91	- 2,39	- 2,96	- 0,47	- 1,05	- 1,8	- 3,4	- 2,1	- 3,26
Podstawowe: 36,5%	+ 4,28	+ 3,52	+ 5,81	+ 2,22	+ 4,68	+ 2,59	+ 0,64	+ 3,73	+ 5,4	+ 3,3	+ 3,06	+ 7,56

U w a g a: Definicje czterech kategorii wykształcenia: 1) wyższe – absolwenci szkół wyższych, 2) średnie – ukończone średnie, łącznie z pomaturalnym i niepełnym wyższym, 3) zasadnicze zawodowe – po zasadniczych szkołach zawodowych, 4) podstawowe – ukończone podstawowe łącznie z niepełnym średnim i niepełnym podstawowym oraz osoby bez wykształcenia szkolnego.

czerwca (6 sondaży) – przeciętny wskaźnik niedoreprezentacji wynosił $-4,58$ punkta procentowego, a między lipcem a grudniem (5 sondaży) $-4,86$; odpowiednie przeciętne dla osób starszych, w wieku od 25 do 34 lat, wynosiły: $-0,87$ i $-3,93$.

Jak widać, tylko w starszej grupie wiekowej, odznaczającej się ogólnie niższym wskaźnikiem, nastąpił wyraźny, bo ponad czterokrotny jego wzrost między lipcem a grudniem, pomimo tego, że w okresie tym w dwóch sondażach zastosowano listy zapowiednie. (Brak pozytywnego wpływu listów zapowiednich widać też w wypadku najmłodszej grupy wiekowej.)

W tab. 5 zauważyć można punkt zwrotny kierunku wypaczeń zależnych od wieku respondentów: od wieku 35 lat i więcej występuje narastająca wyraźnie dominacja wypadków nadreprezentacji, podczas gdy wśród osób w wieku od 35 do 49 lat jest tylko jeden wypadek bardzo słabej niedoreprezentacji, a wśród osób w grupie wiekowej od 50 do 59 lat występują trzy takie wypadki; w ostatniej, najstarszej grupie widzimy bez wyjątku wypadki stosunkowo najsilniejszej nadreprezentacji. Obserwacje te można uogólnić w następujący sposób: w kolejnych wiekowo trzech grupach respondentów od 35 lat wzwyż występuje wyraźna prawidłowość: im są oni starsi – tym silniej są nadreprezentowani w uzyskanych próbach. Ilustruje to tab. 7.

Tabela 7

Zakresy i średnie nadreprezentacji 3 grup wiekowych w 11 kolejnych próbach CBOS-u (w punktach procentowych)

Grupy wiekowe	Zakres	Średnia
35-49	od $-0,09$ do $+2,92$	$+1,21$
50-59(64)	od $-0,50$ do $+3,5$	$+1,35$
Powyżej 59(64)	od $+2,03$ do $+6,74$	$+4,25$

U w a g a: Por. uwagę w tab. 5.

Z tab. 7 widać wyraźnie, że do wieku przedemerytalnego nastąpił słaby wzrost wskaźnika nadreprezentacji – a więc wzrost prawdopodobieństwa przyjęcia roli respondenta; natomiast od wieku emerytalnego wzwyż – wystąpił gwałtowny skok tego wskaźnika: jego wartość potroiła się! Ów skok należałoby chyba przypisać nie tylko – i nie tyle – większej skłonności emerytów do zgody na udzielenie wywiadu, ale raczej faktowi, że ludzie ci znacznie częściej przebywają w domu, a więc są dostępniejsi w momencie wizyty ankietera.

I tym razem uważne prześledzenie diachronicznych zmian wartości analizowanych wskaźników w omówionych trzech grupach wiekowych pozwala stwierdzić występowanie podobnego trendu jak w dwóch poprzednich młodszych grupach: w badanym okresie dziewięciu miesięcy nastąpiło

zaostrenie wypaczeń procentowych udziałów tych grup wiekowych w próbach. Wskazuje to tab. 8.

Należy zauważyć, że i tym razem nie widać wpływu listów zapowiednich na polepszenie „adekwatności reprezentacyjnej” uzyskanych prób pod względem struktury wieku respondentów.

Mówiąc inaczej – w okresie 9 miesięcy objętych analizą poddane badaniu 11 prób stawało się coraz mniej reprezentatywne: ubywało w nich coraz więcej ludzi młodych (do 34 roku życia), a przybywało starszych (zwłaszcza w wieku ponad 59 (64) lat). Najsilniej zaznaczyło się to w odniesieniu do dwóch skrajnych grup wiekowych.

Tabela 8

Przeciętne nadreprezentacji 3 grup wiekowych w próbach w pierwszych 6 i dalszych 5 sondażach

Grupa wiekowa	Pierwsze 6 sondaży	Dalsze 5 sondaży
35–49	+0,75	+1,78
50–59(64)	+1,11	+1,65
Powyżej 59(64)	+3,43	+5,62

Tabela 9

Zakresy i średnie nad- i niedoreprezentacji 4 kategorii wykształcenia w 12 kolejnych próbach CBOS-u (punktach procentowych)

Kategorie wykształcenia	Zakres	Średnia
Podstawowe	od +0,64 do +7,56	+3,9
Zasadnicze zawodowe	od – 3,4 do +1,91	– 1,79
Średnie	od – 3,89 do +1,48	– 1,65
Wyższe	od – 1,51 do +1,00	– 0,45

Mniej jednoznaczny jest kierunek nad- i niedoreprezentacji w zrealizowanych próbach czterech kategorii poziomu wykształcenia. Na podstawie danych w tab. 6 oraz 9 można jednak generalnie stwierdzić, że w 12 kolejnych próbach występuje bez wyjątku tylko nadreprezentacja osób z najniższym poziomem wykształcenia: zawiera się ona w granicach od +0,64 do +7,56 punktów procentowych, przy średniej wynoszącej +3,9. Wraz ze wzrostem poziomu wykształcenia pojawia się w większości sondaży najczęściej niedoreprezentacja osób z zasadniczym wykształceniem zawodowym oraz wykształceniem średnim i wyższym.

Układ wskaźników nad- i niedoreprezentacji czterech kategorii wykształcenia w tab. 9 można zinterpretować w sposób następujący. Stosunkowo

najwyższy wskaźnik nadreprezentacji osób z najniższym poziomem wykształcenia można tłumaczyć łącznym oddziaływaniem dwóch czynników: a) większą podatnością takich osób na prośbę ankietera o udzielenie wywiadu – z czym może się wiązać jego wyższy prestiż, jako kogoś reprezentującego „ważną instytucję” publiczną; b) osoby te są w większości wypadków mieszkańcami wsi, a – jak zobaczymy dalej w tab. 12 – mieszkańcy wsi są systematycznie nadreprezentowani w uzyskiwanych próbach (przeciętna dla 12 sondaży wynosi +2,39 punktu procentowego). Należy również pamiętać, że ludzie z najniższym poziomem wykształcenia występują częściej wśród emerytów, którzy są silnie nadreprezentowani w próbach (przeciętna +4,25).

Wahania wskaźników nad- i niedoreprezentacji w pozostałych trzech grupach wykształcenia mogą mieć charakter losowy – jakkolwiek w zdecydowanej większej liczbie wypadków występuje niedoreprezentacja. Zaznacza się to najsilniej dla osób z wykształceniem średnim (przeciętna –1,65) i zasadniczym zawodowym (przeciętna –1,79); niedoreprezentacja osób z wykształceniem wyższym jest tak nikła (–0,45), że można ją traktować jako nieistotną: oznacza to, że udział tych osób w próbach jest adekwatny do ich udziału w uniwersum, gdyż ich nastawienie do roli respondenta i prośby o wywiad nie cechuje ani szczególna uległość, ani negatywizm powodujący częstsze niedochodzenie do skutku wywiadu (co jednak zaznacza się wśród osób z wykształceniem średnim i zasadniczym zawodowym).

Należy też zwrócić uwagę na tę okoliczność, że niedoreprezentacja owych trzech kategorii wykształcenia w próbach może być odbiciem stwierdzonej dalej niedoreprezentacji ludności miejskiej (wśród której występuje znacznie wyższy niż wśród wiejskiej udział osób z wykształceniem ponadpodstawowym).

Tak jak i w odniesieniu do zmiennych płci i wieku – zobaczmy teraz, czy wypaczenie prób pod względem wykształcenia zmieniało się w okresie badanych dziewięciu miesięcy: czy wystąpiły różnice średniego wskaźnika wypaczenia w pierwszych sześciu i następnych sześciu sondażach. Odpowiedzi na to pytanie przynosi tab. 10.

Tabela 10

Przeciętne nad- i niedoreprezentacji kategorii wykształcenia w próbach pierwszych 6 i dalszych 6 sondażach

Wykształcenie	Pierwsze 6 sondaży	Dalsze 6 sondaży
Wyższe	– 0,53	– 0,38
Średnie	– 1,75	– 1,54
Zasadnicze zawodowe	– 1,57	– 2,01
Podstawowe	+ 3,85	+ 3,95

W tab. 10 widać diachroniczną prawidłowość, zaznaczającą się w ewolucji średnich wskaźnika nad- i niedoreprezentacji: niedoreprezentacja osób z wyższym i średnim poziomem wykształcenia uległa złagodzeniu, a tym samym wzrosła niedoreprezentacja osób z wykształceniem zasadniczym zawodowym (i to najsilniej) oraz bardzo słabo nadreprezentacja osób z wykształceniem podstawowym.

Faktu, że wystąpiło przeciętnie zmniejszenie niedoreprezentacji wśród respondentów z wykształceniem średnim i wyższym, a najsilniej wzrosła niedoreprezentacja kategorii osób z wykształceniem zasadniczym zawodowym, przy niemal braku zmian u osób z wykształceniem podstawowym – nie można wiązać ze zróżnicowanym wpływem listów zapowiednich. (Brak stosownych różnic między wskaźnikami nad- i niedoreprezentacji między trzema próbami, przy których zastosowano takie listy, a próbami sąsiednimi.)

Wykryte w tab. 10 różnice między średnimi dla dwu szóstek prób – należy więc raczej przypisać jakimś innym czynnikom.

Użyte w tej analizie wskaźniki procentowej nad- i niedoreprezentacji trzech rodzajów kategorii osób w uzyskanych próbach respondentów, jako wskaźniki mniej lub bardziej przychylnego ogólnego nastawienia wobec uczestnictwa w sondażach CBOS-u, mają ograniczoną trafność. Oprócz bowiem nastawienia do sondaży odzwierciedlają one w pewnym stopniu również działanie innego rodzaju czynników. Liczbowe wartości tych wskaźników są jakąś wypadkową łącznego oddziaływania owego nastawienia oraz różnych czynników na proces selekcji dokonujący się w trakcie poszukiwania przez ankierów kontaktu z kandydatami na respondentów i kształtowanie się tego kontaktu oraz reakcji na prośbę o wywiad. Owe inne czynniki, nie mające związku z ogólnym nastawieniem do sondaży i podjęciem decyzji udzielenia wywiadu, to m. in.:

a) większa szansa zastania kandydata na respondenta w domu (skorelowana dodatnio z wiekiem emerytalnym i płcią żeńską);

b) mniejsze lub większe nasilenie obawy przed przestępczością, występującą w środowisku respondenta (uwarunkowane stopniem urbanizacji – o czym będzie mowa szczegółowiej w dalszej części studium).

W grę wchodzi również zupełnie inne czynniki, związane z procedurą losowania przez CBOS wyjściowych prób adresów, pod którymi będzie się nawiązywać kontakt z kandydatami na respondentów. Otóż, sposób losowania sześciu adresów w każdym wylosowanym rejonie statystycznym (por. przyp. 1) powoduje nadreprezentację wylosowanych adresów w rejonach wiejskich, a niedoreprezentację – w miejskich. Z drugiej zaś strony, przeciętnie większa liczebność osób w jednym adresie (w gospodarstwie domowym) na wsi niż w mieście sprawia, że – przy zasadzie losowania zawsze tylko jednego respondenta spośród osób w adresie – osoby te mają nieco mniejszą szansę bycia wylosowanymi do próby, a więc występuje tendencja do ich niedorep-

reprezentacji. Jest to tendencja przeciwna kierunkowo wobec tendencji wynikającej ze sposobu losowania adresów w rejonach statystycznych, a zatem występuje tu w jakimś stopniu znoszenie się obydwu kierunków tendencyjności. Jak można jednak sądzić, ów korygujący wpływ tendencji do niedoreprezentacji wylosowanych kandydatów na respondentów na wsi jest znacznie słabszy niż wpływ wynikający ze znacznie mniejszej liczebności adresów w rejonach wiejskich, co powoduje nadreprezentację mieszkańców wsi⁷.

Przedstawiona procedura trójstopniowego losowania kandydatów na respondentów (najpierw rejonów statystycznych, następnie adresów w rejonie i wreszcie osób w adresach) wpływa najbardziej wypaczająco na „nieważoną” strukturę prób wylosowanych do badania osób pod względem ich miejsca zamieszkania (podziału na wieś i miasto). Natomiast w mniejszym stopniu musi dotyczyć struktury płci oraz wieku, a w nieco większym struktury wykształcenia (gdyż ta ostatnia jest silniej skorelowana z podziałem na miasto i wieś, tj. wśród ludności miejskiej jest znacznie wyższa proporcja osób z wykształceniem średnim i wyższym).

Tak więc omówione wskaźniki nad- i niedoreprezentacji – jako wskaźniki stosunku do przyjmowania roli respondenta – są stosunkowo najmniej „zanieczyszczone” informacyjnie wpływem tendencyjności procedury losowania tam, gdzie odnoszą się do płci i wieku respondentów, a najbardziej tam, gdzie chodzi o strukturę wykształcenia. W najwyższym zaś stopniu trafność takich wskaźników musi być osłabiona wtedy, gdy w grę wchodzi „nieważona” struktura uzyskanych prób pod względem miejsca zamieszkania. Na szczęście, akurat w tym wypadku mamy do dyspozycji zarówno dane dotyczące prób uzyskanych, jak i dane charakteryzujące niedoszłych respondentów. Pozwala to na skonstruowanie wskaźników nad- i niedoreprezentacji najbardziej trafnych, bo niezależnych od adekwatności GUS-owskiego (tj. dotyczącego uniwersum dorosłej ludności) procentowego podziału wedle kategorii miejsca zamieszkania, traktowanego jako tożsamy z podziałem występującym w zbiorowości osób proszonych o wywiad.

Wskaźniki nad- i niedoreprezentacji osób z czterech wyodrębnionych w badaniach kategorii miejsca zamieszkania wśród zbiorowości niedoszłych respondentów można więc traktować jako behawioralne wskaźniki ich stosunku do prośby o wywiad, tj. negatywnego lub pozytywnego nastawienia do uczestnictwa w sondażach CBOS-u, czyli do roli respondenta. Nastawienia

⁷ Wedle *Rocznika statystycznego 1994* – w roku 1988 (a więc cztery lata przed rokiem 1992) w przeciętnym gospodarstwie domowym na wsi było 3,55 osób, a w miastach 2,86; oznacza to, że w miastach w gospodarstwie domowym było o 19,4% mniej osób niż na wsi. W roku 1984 procent ten wynosił 21,5 – a więc występuje tendencja do zmniejszania się tej różnicy. (Brak w *Roczniku* owej informacji w odniesieniu do roku 1992). Opieramy się na założeniu, że 1 adres = 1 gospodarstwo domowe.

te można poddać pomiarowi z zastosowaniem ilościowych wskaźników na skali interwałowej.

Powstaje jednak problem metodologiczny: w stosunku do jakiego „procentu odniesienia” obliczać owe procentowe odchylenia, mające stanowić wspomniane wskaźniki?

Istnieją tutaj trzy możliwości.

1. Bierzemy do porównania dwie zbiorowości (każdą jako 100%): zbiorowość niedoszłych respondentów i zbiorowość respondentów efektywnych. Gdyby w trakcie procesu kontaktowania się ankietowanych z osobami w realnych adresach odmowy wywiadu **nie** dawały w efekcie jakiegos selektywnego (tendencyjnego) odsiewu osób proszonych o wywiad (inaczej mówiąc, gdyby ów odsiew był przypadkowy), to w obydwu zbiorowościach uzyskiwano by analogiczne (w ramach błędu standardowego) proporcje (czy procenty) kategorii zamieszkania. Miara stopnia owej losowości (a więc braku tendencyjności) byłyby **różnice** między procentami występującymi wśród zbiorowości niedoszłych respondentów a procentami występującymi w zbiorowości respondentów efektywnych. Brak takich różnic oznaczałby, że środowiska terytorialne (miaso, wieś) nie różnią się pod względem nastawienia do roli respondenta.

2. Rekonstruujemy całą zbiorowość pierwotnych „kontaktów ankietarskich” w realnych adresach przez dodanie do siebie dwu wspomnianych zbiorowości (w punkcie 1). W obrębie tej całej zbiorowości obliczamy rozkład procentowy osób z czterech kategorii miejsca zamieszkania; uzyskane 4 liczby procentowe będą stanowiły „procenty odniesienia” dla obliczania różnic punktów procentowych będących wskaźnikami nastawienia do roli respondenta (tj. do prośby o udzielenie wywiadu). Inaczej mówiąc – od każdego z czterech procentowych udziałów wśród **niedoszłych** respondentów danej kategorii zamieszkania odejmujemy odpowiadający mu „procent odniesienia” i uzyskaną różnicę traktujemy jako wskaźnik nad- lub niedoreprezentacji tej kategorii wśród niedoszłych respondentów („odmawiaczy”); różnica ujemna oznacza, że dana kategoria jest procentowo niedoreprezentowana wśród „odmawiaczy”, a więc odznacza się bardziej pozytywnym stosunkiem do prośby o wywiad niż kategoria, dla której uzyskano różnicę dodatnią (a więc jest ona nadreprezentowana wśród „odmawiaczy”), a zatem ma bardziej negatywny stosunek do prośby o wywiad.

I tutaj – różnice równe lub bliskie zeru oznaczałyby, że odsiew jest losowy, czyli nietendencyjny pod względem miejsca zamieszkania, a więc że środowiska różniące się terytorialnie nie różnią się nastawieniem do roli respondenta.

3. Porównujemy dwie zbiorowości: całą zbiorowość „zrekonstruowaną” ze zbiorowości uzyskanych do wywiadu respondentów. Podobnie jak poprzednio, odejmujemy „procenty odniesienia” (tj. występujące w całej zbiorowości „zrekonstruowanej”) od procentów, jakie wystąpiły w zbiorowości efektywnych respondentów (tj. w zbiorowości umniejszonej o osoby, które

uchyliły się od wywiadu). Różnice dodatnie wystąpią przy procentach tych kategorii zamieszkania, które są w zbiorowości efektywnych respondentów (tj. w próbie zrealizowanej) wyższe, czyli nadreprezentowane z powodu mniej częstego uchylania się od wywiadu; odwrotnie – różnice ze znakiem ujemnym wystąpią dla tych kategorii zamieszkania, które są w próbie zrealizowanej niższe, czyli niedoreprezentowane (z powodu częstszego uchylania się od wywiadu osób w tej kategorii zamieszkania). W pierwszym wypadku dana kategoria zamieszkania zawiera mniejszą liczbę osób uchylających się od udzielenia wywiadu niż druga, a zatem – odznacza się (na szczeblu zbiorowym) bardziej pozytywnym stosunkiem do akceptacji roli respondenta.

Można teraz w sposób syntetyczny zestawić sens wskaźnikowy uzyskanych na drodze tej analizy sześciu relacyjnych układów nierówności między udziałami procentowymi osób danej kategorii zamieszkania wśród trzech wyróżnionych zbiorowości: 1) tylko wśród tych, którzy uchylili się od wystąpienia w roli respondenta; 2) wśród ogółu osób proszonych o wywiad (tj. respondentów łącznie z tymi, którzy uchylili się od wystąpienia w roli respondenta); 3) tylko wśród respondentów.

Oznaczenia:

u – procent osób należących do danej kategorii miejsca zamieszkania wśród tych, którzy na prośbę o wywiad zareagowali uchYLENIEM się od wywiadu (odmówili wystąpienia w roli respondenta);

r – procent osób należących do danej kategorii zamieszkania wśród respondentów (tj. reagujących na prośbę o wywiad przyjęciem roli respondenta);

o – procent takichże osób wśród ogółu proszonych o wywiad;

1) $r < u$ – dana kategoria ma większy udział procentowy wśród uchylających się od wywiadu niż wśród respondentów, a więc jest nadreprezentowana wśród „odmawiaczy” – czym wykazuje negatywne nastawienie do roli respondenta;

2) $r > u$ – dana kategoria ma mniejszy udział procentowy wśród uchylających się od wywiadu niż wśród respondentów, a więc jest niedoreprezentowana wśród „odmawiaczy” – czym wykazuje pozytywne nastawienie do roli respondenta;

3) $o < u$ – dana kategoria ma większy udział procentowy wśród uchylających się od wywiadu niż wśród ogółu proszonych o wywiad, a więc jest nadreprezentowana wśród „odmawiaczy” – czym wykazuje negatywne nastawienie do roli respondenta;

4) $o > u$ – dana kategoria ma mniejszy udział procentowy wśród uchylających się od wywiadu niż wśród ogółu proszonych o wywiad, a więc jest niedoreprezentowana wśród „odmawiaczy” – czym wykazuje pozytywne nastawienie do roli respondenta;

5) $o < r$ – dana kategoria ma większy udział procentowy wśród respondentów niż wśród ogółu proszonych o wywiad, a więc jest nadreprezentowana wśród respondentów – czym wykazuje pozytywne nastawienie do roli respondenta;

6) $o > r$ – dana kategoria ma mniejszy udział procentowy wśród respondentów niż wśród ogółu proszonych o wywiad, a więc jest niedoreprezentowana wśród respondentów – czym wykazuje negatywne nastawienie do roli respondenta.

Należy zauważyć, że trzy spośród sześciu układów nierówności – a mianowicie $r < u$, $o < u$ i $o > r$ – wskazują na bardziej negatywne niż pozytywne nastawienie danej kategorii zamieszkania do prośby o wywiad, zaś trzy pozostałe, tj. $r > u$, $o > u$ i $o < r$, na nastawienie pozytywne.

Można obliczyć trzy alternatywne miary (tj. trzy wersje wskaźnika) nastawienia danej kategorii zamieszkania do roli respondenta; będą to różnice: $u - r$, $u - o$, $r - o$.

Powstaje zagadnienie, czym się te miary (w swym sensie) różnią między sobą, jakie zachodzą między nimi rozbieżności i która z miar jest najbardziej trafna jako miara nastawienia osób należących do danej kategorii miejsca zamieszkania wobec roli respondenta. Postarajmy się sens uzyskiwanych w ten sposób wskaźników rozpatrzeć w sposób konkretno-empiryczny, tj. poprzez obliczenie trzech wersji miar dla wybranego sondażu i ich analizę.

Zgodnie z ogólnie przyjmowaną w statystyce konwencją jako procenty odniesienia przyjęto procenty **odejmowane** (podobnie jak przy obliczaniu odchylen od średniej, którą odejmuje się od danego procentu, aby wyrazić go jako odchylenie od niej). Tak więc, różnica $u - r$ mówi o tym, czy dana kategoria osób jest nadreprezentowana (ma wtedy znak dodatni, bo $u > r$), czy niedoreprezentowana (ma wtedy znak ujemny) wśród **uchylających się** od wywiadu w stosunku do godzących się na wywiad. W pierwszym wypadku kategoria ta ma bardziej negatywne nastawienie do prośby o wywiad, a w drugim – bardziej pozytywne (moduł tej różnicy mówi o sile tego nastawienia, a znak algebraiczny o jego kierunku). Różnica $u - o$ mówi również o nad- czy niedoreprezentacji danej kategorii osób wśród **uchylających się** od wywiadu, ale tym razem w stosunku do ogółu osób proszonych o wywiad: sens znaków algebraicznych mówiących o kierunku nastawienia do prośby o wywiad, jest taki sam jak w wypadku różnicy $u - r$; natomiast mniejsza musi być wartość modalna tej różnicy (a więc większa wartość algebraiczna), gdyż mniejsza jest wartość o niż r . Ostatnia różnica, $r - o$, ujawnia nad- czy niedoreprezentację danej kategorii osób wśród respondentów w stosunku do ogółu proszonych o wywiad; nadreprezentacja oznacza, że kategoria ta ma bardziej pozytywne nastawienie do takiej prośby, a niedoreprezentacja – że bardziej negatywne.

Tabela 11

Trzy wersje wskaźnika (miar) nastawienia czterech kategorii zamieszkania do roli respondenta w sondażu CBOS-u z marca 1992 r. (kod: 05)

Kategorie miejsca zamieszkania	Różnice między procentowymi udziałami osób z poszczególnych kategorii zamieszkania w zbiorowościach:		
	wersja $u - r$	wersja $u - o$	wersja $r - o$
Wieś	$25,4 - 42,53 = -17,13$	$25,4 - 40,3 = -14,9$	$42,53 - 40,30 = +2,23$
Miasta do 20 000 mieszkańców	$10,9 - 10,24 = +0,66$	$10,9 - 10,41 = +0,49$	$10,24 - 10,41 = -0,17$
Miasta od 21 000 do 101 000 mieszkańców	$24,1 - 19,04 = +5,06$	$24,1 - 19,7 = +4,4$	$19,04 - 19,70 = -0,66$
Miasta powyżej 101 000 mieszkańców	$39,6 - 28,18 = +11,42$	$39,6 - 29,59 = +10,01$	$28,18 - 29,59 = -1,41$

U w a g a: znaki „-” przy wartościach różnic $u - r$ i $u - o$ oraz znak „+” przy wartości różnicy $r - o$ dla ludności wsi wskazują na jej bardziej pozytywne nastawienie do próby o wywiad niż ludności miast, gdzie znaki „+” dla wersji $u - r$ i $u - o$, a znaki „-” dla wersji $r - o$ wskazują na jej negatywne nastawienie. Liczebności trzech zbiorowości, od których obliczono procentowe udziały kategorii zamieszkania: respondenci: 1171, „odmawiacze”: 174, ogół proszonych o wywiad: 1345.

Układ znaków algebraicznych przy wszystkich trzech wersjach wskaźnika dowodzi, że mieszkańcy wsi mają bardziej pozytywne nastawienie do próby o wywiad niż mieszkańcy każdej z pozostałych trzech kategorii miast: wiejskie różnice (tj. wskaźniki stosownej nad- i niedoreprezentacji) dowodzą, że mieszkańcy wsi są procentowo nadreprezentowani wśród respondentów w stosunku do ogółu osób proszonych o wywiad, a zarazem są oni niedoreprezentowani wśród uchylających się od wywiadu – zarówno w stosunku do respondentów, jak i ogółu proszonych o wywiad. Jest to wspomniany układ nierówności: $r > u$, $o > u$ i $o < r$ – i taki właśnie układ tworzą trzy pary liczb procentowych dla wsi w tab. 11.

Układ znaków algebraicznych, wskazujący na ogólnie negatywne nastawienie mieszkańców miast do próby o wywiad, to wspomniany układ $r < u$, $o < u$ i $o > r$; jak widzimy, występuje on dla wszystkich trzech par udziałów procentowych dla każdej z trzech kategorii zamieszkania ludności miejskiej.

Liczby bezwzględne punktów procentowych, towarzyszące omówionym różnicom dodatnim i ujemnym, świadczą o stopniu nad- i niedoreprezentacji. Widzimy tutaj wyraźną prawidłowość: wraz ze wzrostem wielkomiejskości rośnie nasilenie bezwzględnych wartości wskaźników negatywnego nastawienia do próby o wywiad – i to przy zastosowaniu każdej z trzech wersji różnicy procentowych udziałów mieszkańców trzech rodzajów miast. Dla wersji $u - r$ oraz $u - o$ wskaźniki te regularnie rosną w kierunku coraz to większych **wartości dodatnich**, w miarę jak od wsi (z wartością ujemną wskaźnika) przechodzimy do coraz większych miast – odpowiednio: $-17,13$; $+0,66$; $+5,06$ i $+11,42$ oraz $-14,9$; $+0,49$; $+4,4$ i $+10,01$. Taki sam kierunek (z odwrotnym układem znaków algebraicznych) ma ten sam szereg wskaźników dla wersji $r - o$: od $+2,23$ dla wsi do $-0,17$; $-0,66$ i $-1,41$ dla kolejno coraz większych miast.

Zwróćmy teraz uwagę na różnice między wartościami wskaźników w zależności od przyjętej wersji sposobu ich obliczania: $u - r$, $u - o$ i $r - o$. Zastanówmy się zarazem, która z wersji sposobu obliczania nad- i niedoreprezentacji daje najbardziej stabilny i syntetyczny wskaźnik.

Wersja $u - r$ abstrahuje od tego, jaką liczebność bezwzględną (jak i w stosunku do całej zbiorowości proszonych o wywiad) ma zbiorowość „odmawiaczy”, lub – inaczej mówiąc – jaki jest stopień realizacji próby (proporcja i liczebność uzyskanych do wywiadów osób w stosunku do zbiorowości zaplanowanej). Na wartość różnicy $u - r$ nie wpływa bowiem to, ilu było respondentów i odmawiających; gdyby na przykład „odmawiaczy” była **znikoma** liczba – to uzyskany wśród nich wysoki procent osób danej kategorii zamieszkania mógłby być **przypadkowy** (nieistotny statystycznie), a mimo to wpłynąć decydująco na kierunek i wielkość bezwzględną wskaźnika ich nadreprezentacji w relacji do próby zrealizowanej. Taki wskaźnik byłby nieporównywalny ze wskaźnikiem obliczonym dla sondażu, w którym procent u wśród „odmawiaczy” byłby znacznie większy, a ich zbiorowość znacznie liczniejsza. Dlatego też z wersji $u - r$ należy zrezygnować wtedy, gdy w grę wchodzi do porównywania sondaże różniące się znacznie stopniem realizacji próby (co – jak widzieliśmy – ma miejsce w niniejszym studium 12 sondaży CBOS-u).

Pozostają do rozważenia dwie pozostałe wersje obliczania wskaźnika: $u - o$ i $r - o$ – a więc różniące się tylko tym, od czego odejmujemy ten sam procent danej kategorii osób występujący wśród **ogółu** osób proszonych o wywiad: raz odejmujemy go od procentu występującego wśród „odmawiaczy”, a raz wśród respondentów. Biorąc pod uwagę fakt, że zbiorowości respondentów są z reguły znacznie liczniejsze niż „odmawiaczy” – uzyskany wśród nich procent danej kategorii zamieszkania jest bardziej stabilny statystycznie (mniej przypadkowy). Sens znaków algebraicznych oznacza

Tabela 12

Nad- i niedoreprezentacja mieszkańców czterech kategorii miejscowości wśród respondentów – obliczona w stosunku do ogółu osób proszonych o wywiad
($r - o$) w punktach procentowych

Kategorie miejsca zamieszkania	Miesiąc i kod sondażu oraz wskaźniki nastawienia do roli respondenta											
	III 05	III 06	III-IV 07	IV 11	IV-V 12	V 13	VI 15 ^a	VII 17	IX 19	X 24 ^a	XI 27	XII 33 ^a
Wieś	+2,23	+2,66	+2,52	+2,32	+2,33	+2,48	+1,99	+2,89	+2,96	+2,12	+2,92	+1,23
Miasta do 20 000 miesz- kańców	-0,17	-0,66	-0,33	-0,05	-0,40	+0,21	-0,55	+0,02	-0,33	-0,62	-0,33	+0,07
Miasta od 21 000 do 101 000 mieszkańców	-0,66	-0,86	-0,13	+0,39	-0,27	+0,23	-0,16	+0,07	+0,53	-0,84	+0,15	+0,05
Miasta powyżej 101 000 mieszkańców	-1,41	-1,12	-2,06	-2,65	-1,68	-2,92	-1,28	-2,99	-3,16	-0,67	-2,75	-1,31

^a W sondażach zastosowano listy zapowiednie.

tutaj jednak coś odwrotnego niż w wersji $u - o$; dla różnicy $r - o$ znak plus sygnalizuje nam, że dana kategoria zamieszkania jest nadreprezentowana **wśród respondentów** (bo $r > o$), a znak minus – że jest niedoreprezentowana (bo $r < o$); w pierwszym wypadku oznacza to bardziej pozytywne nastawienie do prośby o wywiad, a w drugim – bardziej negatywne. Poza tym, tak obliczone wskaźniki nad- czy niedoreprezentacji danej kategorii w uzyskanej próbie sygnalizują badaczowi, czy w opracowaniu ma do czynienia z danymi wypaczonymi i w jakim stopniu.

Pamiętając o tej konwencji, spójrzmy teraz na powyższą tab. 12, ukazującą wskaźniki nastawienia do prośby o wywiad w 12 kolejnych sondażach CBOS-u w roku 1992 – obliczone jako różnica $r - o$.

Uważny ogląd przedstawionych w tab. 12 wskaźników nad- i niedoreprezentacji – jako tym razem pełnych wskaźników skłonności osób proszonych o wywiad do przyjmowania roli respondenta CBOS-u – prowadzi do następujących stwierdzeń:

1. Występuje bez wyjątku prawidłowość w nastawieniu do wywiadu ludności wsi i miast powyżej 101 000 mieszkańców: ci pierwsi wykazują zawsze bardziej pozytywne nastawienie do prośby o wywiad (tj. rzadziej uchylają się od niego), w rezultacie czego są zawsze nadreprezentowani w próbach, w granicach od +1,23 do +2,92 punktów procentowych, przy średniej wynoszącej +2,39; ci drudzy zaś – odwrotnie: zdradzają zawsze bardziej negatywne nastawienie, co wyraża się w bezwyjątkowym niedoreprezentowaniu w próbach (tj. częstszym uchylaniu się od wywiadu) w granicach -0,67 do -3,16 punktów procentowych, przy średniej wynoszącej -2,04. Tak więc, ludność wiejska generalnie w silniejszym stopniu przejawia pozytywne nastawienie do prośby o wywiad dla CBOS-u, i to w sposób bardziej stabilny niż swe negatywne nastawienie przejawia ludność wielkomiejska, która uzewnętrznia znacznie silniejszą zmienność pod tym względem: odchylenie przeciętne od średniej dla ludności wiejskiej w 12 sondażach wynosi 0,35 punkta procentowego, a dla ludności największych miast aż 0,75 punkta.

2. Występują interesujące trendy diachroniczne dla wskaźników ludności wsi i największych miast – co ukazuje tab. 13.

Zaznacza się wzrost wskaźnika negatywnego nastawienia do prośby o wywiad wśród mieszkańców największych miast o 0,06 punkta procentowego (z -1,97 do -2,03), pomimo tego że w drugiej szóstce sondaży, w trzech z nich zastosowano listy zapowiednie! Natomiast w odniesieniu do wsi widzimy **spadek** wskaźnika pozytywnego nastawienia, będący jak gdyby „lustrzanym odbiciem” stwierdzonego wzrostu wskaźnika negatywnego dla największych miast – i to o niemal taką samą różnicę! (o 0,07 punkta

Tabela 13

Przeciętne nad- i niedoreprezentacji 4 kategorii miejsca zamieszkania w próbach w pierwszych 6 i dalszych 6 sondażach

Miejsce zamieszkania	Pierwsze 6 sondaży	Dalsze 6 sondaży
Wieś	+2,42	+2,35
Miasta do 20 000 mieszkańców	-0,23	-0,29
Miasta od 21 000 do 101 000 mieszkańców	-0,22	-0,03
Miasta powyżej 101 000 mieszkańców	-1,97	-2,03

– z +2,42 do +2,35). Jest więc w tym układzie zmian pewna logika: w obydwu skrajnych kategoriach zamieszkania nastąpiło pewne pogorszenie „atmosfery psychologicznej” wokół rekrutacji respondentów do wywiadu – i to pomimo listów zapowiednich (okoliczność ta nadaje owej nikłej statystycznie różnicy większe znaczenie).

3. Jeśli zaś chodzi o dwie kategorie ludności między wsią a największymi miastami – tj. ludność małych miast i średnich miast – to wskaźniki nad- i niedoreprezentacji są tak nikłe, iż można je uważać za zerowe (nigdy nie przekraczają 1 punkta procentowego – a więc mieszczą się w granicach nie tylko błędu statystycznego, ale i błędów zaokrąglania). Oznacza to, że w odniesieniu do tych kategorii ludności w trakcie realizacji badań w terenie (tj. w momentach kontaktu ankierów z osobami proszonymi o wywiad) nie następował tendencyjny ich odsiew i ich procentowy udział w całej próbie jest stosunkowo najbardziej adekwatny.

Wspomniany wcześniej w tab. 11 układ wskaźników, mający świadczyć o prawidłowości typu „im bardziej wielkomiejskie środowiska – tym bardziej negatywne nastawienie do prośby o wywiad” – wystąpił tylko w dwóch sondażach (kody 05, 06). Fakt ten można więc potraktować jako wyjątek od reguły, że wartości algebraiczne wskaźników i ich wartości modalne (w małych i średnich miastach) skaczą nieregularnie, co jest zresztą rzeczą oczekiwaną przy tak małych ich wartościach i związanych z tym wahaniach losowych.

D. WPLYW LISTÓW ZAPOWIEDNICH NA STOPIEŃ REALIZACJI ADRESOWEJ PRÓBY W SONDAŻACH CBOS-II

W poprzednich analizach już kilkakrotnie odwoływano się do wpływu listów zapowiednich w interpretacji diachronicznych zmian wartości niektórych wskaźników nastawienia do prośby o wywiad. Obecnie postaramy się spojrzeć na wpływ listów zapowiednich w sposób całościowy.

Oprócz czynników związanych z sezonem, w którym przeprowadza się sondaż, wpływ na rozmiary niecosowego ubytku z próby osób przewidzianych do wywiadu (tj. w realnych adresach) może mieć czynnik aranżacyjny, jakim jest wysłanie na wylosowany adres listu zapowiadającego wizytę ankietera, wyjaśniającego cel i temat wywiadu i ujawniającego sponsora badań. Stosowanie takiej procedury aranżacji kontaktu badawczego jest realizacją postulatu „humanizacji moralnej” – nakazującego pełne informowanie badanych osób o tym, kto przeprowadza badanie i w jakim celu. W ten sposób szanuje się podmiotowość badanych, unika zaskakiwania ich czy niepokojenia nieoczekiwanym najściem obcej osoby⁸. Gdyby okazało się, że list zapowiedni równocześnie zwiększa stopień realizacji próby (tj. zmniejsza liczbę odmów różnego rodzaju), to wykazano by, że rzetelność i przyzwoitość w sposobie traktowania ludzi wylosowanych do badań daje zarazem pozytywne efekty metodologiczne, zwiększa poziom reprezentatywności próby i wiarygodności wyników badawczych.

List zapowiedni wysłano w roku 1992 tylko w trzech sondażach CBOS-u, tj. przeprowadzonych w czerwcu, październiku i grudniu. Analizując dane z tab. 2, zobaczymy, jak procenty wypadków niedojścia do skutku wywiadów, z powodu odmów, kształtowały się w tych trzech sondażach na tle pozostałych dziewięciu – bez listu zapowiedniego.

Patrząc od „pozytywnej” strony na efekt listu zapowiedniego, tj. obserwując poziom realizacji próby w realnych adresach (po odrzuceniu adresów nie istniejących czy też takich, gdzie nikogo nie zastano pomimo ponawiania wizyty), widzimy natychmiast, że owe trzy sondaże mają trzy najwyższe pozycje rangowe pod tym względem: procenty zrealizowanych wywiadów wynoszą w nich 91,6%, 92,3% i 92,9%; zaś **najwyższy** procent zrealizowanych wywiadów wśród pozostałych 9 sondaży bez listu zapowiedniego wynosi 89,7% (sondaż z kwietnia na próbie 2015 realnych adresów).

Porównując procenty wyrażające stopień realizacji próby w owych trzech sondażach z procentami realizacji próby w czterech sondażach czasowo sąsiednich, tj. przeprowadzonych o miesiąc wcześniej lub później, stwierdzamy, że w trzech sondażach z listem zapowiednim występuje zawsze poziom realizacji próby wyższy w granicach od 4,4 do 5,4 punktów procentowych różnicy. Jest to miarą ostatecznej efektywności procedury zwiększania stopnia reprezentatywności próby poprzez stosowanie listu zapowiedniego; gdyby owe pięć różnic przedstawić jako średnią, to wynosiłaby ona 4,8 punkta procentowego.

⁸ Postulat taki sformułował autor niniejszego opracowania w t. 9 „*Analiz i Prób Technik Badawczych w Socjologii*”, w artykule *Respondent jako pełny człowiek w badaniach surveyowych*. Inne artykuły tam zawarte ukazują, jak partnerskie podejście do badanych i uszanowanie ich podmiotowości sprzyja uzyskiwaniu ich współdziałania i tym samym przyczynia się do polepszenia rezultatów badawczych.

Spójrzmy teraz na efektywność listu zapowiedniego od strony „negatywnej” – i to w poszczególnych kategoriach odmów. Okazuje się, że list zapowiedni – zgodnie ze zdroworozsądkowym oczekiwaniem – najbardziej przyczynia się do usunięcia obawy przed wpuszczeniem do mieszkania nieznanej czy niebezpiecznej osoby. Wszystkie trzy procentowe wskaźniki udziału wypadków odmowy wpuszczenia do mieszkania – występujące w trzech sondażach z listem zapowiednim – są **najniższe** (1,6%; 1,8%; 1,3%) w porównaniu ze wskaźnikami w pozostałych dziewięciu sondażach, wśród których najniższy procent wynosi 1,9. Ponieważ trzy sondaże z listem zapowiednim należą do drugiej szóstki analizowanych wcześniej pod tym względem sondaży, znajdujemy tym samym wyjaśnienie stwierdzonego faktu, że w szóstce tej nastąpił spadek przeciętnego wskaźnika obawy wpuszczenia do domu ankietera z 3 do 2,5%.

Podobny – choć mniej wyraźny – wpływ listu zapowiedniego widzimy w wypadku kategorii „odmów bez podania przyczyn” (są to reakcje niedoszłych respondentów zdradzające lekceważący, wrogi czy niechętny stosunek do prośby o wywiad). Otóż okazuje się, że w trzech sondażach, w których wylosowane osoby miały jakąś szansę⁹ zetknięcia się z listem zapowiednim (bądź wiedzieć o nim z relacji członków rodziny), procenty tego typu odmów były niższe niż w pozostałych sondażach (tylko w dwu z pozostałych sondaży procent ten był tak samo niski – 0,4% – jak w czerwcowym sondażu z listem zapowiednim). Porównajmy tym razem dwa przeciętne procenty występowania „odmów bez podania przyczyn”: jeden dla trzech sondaży z listem zapowiednim i drugi – dla dziewięciu sondaży bez takiego listu; pierwszy z nich wynosi 0,3%, a drugi trzykrotnie więcej – bo 0,9%.

W najliczniejszej kategorii wypadków niedochodzenia do skutku wywiadu ankieterzy zarejestrowali jako przyczynę „nieobecność wylosowanej osoby przez cały okres realizacji badania”. Można podejrzewać, że w wielu wypadkach kryło się za tą „nieobecnością” unikanie kontaktu z ankieterem – „chowanie się” przed nim. Jeśliby występowały tutaj rzeczywiście takie formy uniknięcia wywiadu – to również w tej kategorii powinno się zaobserwować podobny układ różnic procentowych jak w poprzednio omawianych kategoriach odmów. Analiza danych w tab. 2 częściowo potwierdza tę hipotezę: dla dziewięciu sondaży bez listu zapowiedniego przeciętny procent wypadków wyjaśniania niedojścia do skutku wywiadu dłuższą nieobecnością wynosi 3,3%, a dla sondaży z listem zapowiednim – tylko 2,7%. A więc wygląda to tak, jakby list zapowiadający zmniejszał procent „**dłuższych nieobecności**” osób w mieszkaniach, gdzie zapukał

⁹ Mowa tylko o szansie – bo w rzeczywistości list zapowiedni nie zawsze dotarł akurat do rąk osoby, która została wylosowana do badań.

ankieter z prośbą o wywiad! Narzuca się nieodparcie przypuszczenie, że w jakiejś liczbie wypadków „dłuższa nieobecność” musiała być sposobem „wykręcenia się” od wywiadu, przy czym zachowanie takie występowało nieco częściej wtedy, gdy przed danym sondażem nie były wysyłane listy zapowiednie. Warto zauważyć, że w wypadku „nieobecności wylosowanej osoby w uzgodnionym wcześniej terminie” kierunek tego rodzaju różnic procentowych kształtuje się podobnie, ale tylko w siedmiu ostatnich sondażach (zakres pięciu różnic „międzysondażowych” zawiera się między 0,1% i 0,6% – z przeciętną różnicą wynoszącą 0,3 punktu procentowego).

Wpływ listu zapowiedniego na zmniejszenie liczby odmów widać wreszcie w kategorii odmów z powodu „niechęci do wypowiedzania się na dany temat”: trzy wskaźniki procentowe dla trzech sondaży z listem zapowiednim wynoszą 0,3; 0,5 i 0,2; wśród pozostałych dziewięciu sondaży w siedmiu z nich wskaźniki te są wyższe niż 0,5, a tylko w dwu sondażach ów wskaźnik procentowy wynosi 0,5. Przeciętna wartość wskaźnika dla owych dziewięciu sondaży wynosi 0,7 – a dla trzech sondaży z listem zapowiednim zaledwie 0,3, a więc ponad dwukrotnie mniej. Pięć różnic „międzysondażowych” wynosi kolejno: 0,6; 0,2; 0,2; 0,4; 0,7 – ze średnią równą 0,4. Tak więc – zgodnie ze zdroworozsądkowym oczekiwaniem – wyjaśnienia dostarczone w liście zapowiednim (m. in. o tematyce sondażu) przyczyniają się w pewnym stopniu do zmniejszenia ilości odmów.

Jest rzeczą szczególnie interesującą, że występują dwie kategorie odmów, gdzie nie znajdujemy układu i kierunku różnic procentowych wskazujących na stwierdzony dotąd tak wyraźny wpływ listu zapowiedniego: w trzeciej pod względem ogólnej frekwencji występowania kategorii „odmowa z powodu niewidzenia sensu (potrzeby) ankietowania” oraz w kategorii frekwencyjnie ósmej – „odmowa informacji o osobach stale zamieszkujących”.

Hipotetycznie wyjaśnić to można ingerencją czynników nie poddających się perswazyjnemu wpływowi treści zawartych w liście zapowiednim. Tak więc w wypadku, gdy nie widzi się sensu (potrzeby) ankietowania, mamy do czynienia ze zgeneralizowaną i trwałą postawą negatywną wobec sondaży, której nie może zmienić list zapowiedni; jak widzimy w tab. 2 – postawę taką zdaje się żywić, na szczęście, zaledwie co 100. osoba proszona o wywiad. W mniej licznej (tylko 0,4%) kategorii „odmów udzielenia informacji o osobach stale zamieszkujących” (co uniemożliwia wylosowanie kandydata do wywiadu) w grę musi wchodzić całkowicie inny czynnik – najprawdopodobniej jakaś obawa przed ujawnieniem stanu zamieszkania lokalu (być może wchodzi w grę obecność tam osób oficjalnie nie zameldowanych itp.). Są to – jak widzieliśmy – wypadki bardzo rzadkie (w mniej niż co 200. adresie).

Na zakończenie chciałbym dołączyć uwagę krytyczno-metodologiczną. Gdyby trzymać się ściśle reguł statystycznych, należałoby poddać stwierdzone

rozkłady procentowe testowi χ^2 , aby zobaczyć, na jakim poziomie istotności można dokonywać konstatacji dotyczących wpływu listu zapowiedniego na procenty odmów (tj. istotności różnic procentowych). Wobec jednak silnego dystrybucyjnego rozdrobnienia badanych kategorii odmów – absolutne teoretyczne (oczekiwane) liczebności wypadków 14 wyodrębnionych rodzajów odmów byłyby w poszczególnych okienkach tabeli zbyt małe (nawet poniżej 5 wypadków), co – wedle opinii statystyków – czyni test χ^2 bezprzedmiotowym.

Niemniej – niezależnie od rygorystycznie traktowanego testu χ^2 – można spojrzeć na uzyskane wyniki analizy różnic między procentami, jak i różnic „międzysondażowych”, w sposób całościowy: pomimo każdorazowo małych liczebności (a więc i nikłych procentów) **kierunek** tych różnic jest we wszystkich analizowanych wypadkach porównań ten sam, tj. wskazuje na określony, korzystny metodologicznie wpływ listu zapowiedniego. Na tej zasadzie można przyjąć, że stwierdzone tak liczne różnice procentowe w tym samym kierunku nie są przypadkowe.

Zygmunt Gostkowski

PERSONS INACCESSIBLE TO INTERVIEWS AND THE MEASURES OF DEGREE OF
REPRESENTATIVENESS OF SURVEY SAMPLES AS
INDICATORS OF A SOCIETY'S ATTITUDE TO PUBLIC OPINION SURVEYING

In 1992 12 consecutive random samples of the Public Opinion Research Center (CBOS) have been analysed together with all the 1888 cases of avoidance and refusals to be interviewed. 14 various categories of refusals or other kinds of avoidance were distinguished as indicative of a negative attitude toward the acceptance of respondent's role. Percentage points measuring over- and under-representation of sex, age, education level and place-of-residence categories in the samples of respondents were used as indicators of these categories' willingness/unwillingness to be interviewed. In all samples over-represented were women, retirement age and least educated persons and inhabitants of the countryside – which fact testifies their more positive attitude toward surveys and acceptance of respondent's role. Influence of letters announcing the interviewer's visit upon the degree of realization of planned address samples was ascertained by comparing refusal rates in 3 surveys using such letters with the remaining 9 ones without them: it turned out that in the former consistently lower percentages of refusals were found which shows that a researcher can create in potential respondents a more cooperative attitude.