

Henryk Domański
Polska Akademia Nauk

LICZBA WIZYT I CZAS TRWANIA BADAŃ

Odwołując się do danych badania Europejski Sondaż Społeczny próbuję odpowiedzieć na dwa pytania. Po pierwsze, od jakich czynników zależy dostępność respondentów – definiowana przez liczbę wizyt, podejmowanych przez ankietera w celu nawiązania kontaktu. Po drugie, w jakim stopniu odpowiedzi respondentów zależą od punktu czasowego, w jakim realizowany jest wywiad. Przeanalizuję tu rolę tych zjawisk na danych z Polski. Występowanie tych zależności wskazywałoby, że pomiary obciążone są systematycznym błędem polegającym na zmianie kontekstu badania – wpływ czasu był dowodem na to, że w trakcie badania zmieniają się warunki realizacji wywiadów. Z moich analiz wynika, że czynnikami utrudniającymi dostępność były: wyższy poziom wykształcenia, mieszkanie w większej miejscowości i młody wiek, natomiast stosunkowo łatwo dociera się do osób bezrobotnych, mieszkańców wsi i niższych rangą pracowników umysłowych. Jeżeli chodzi o stwierdzenie, czy data realizacji wywiadu wywiera niezależny wpływ na odpowiedzi badanych, to okazało się, że – w istocie – zależność taka istnieje.

Główne pojęcia: dostępność respondenta, długość trwania badania, liczba wizyt, realizacja próby, błąd systematyczny.

Posługując się metodą wywiadu kwestionariuszowego musimy brać pod uwagę czas trwania badań. Rozciągnięcie badania w czasie zmniejsza prawdopodobieństwo, że nie zmieniło się *indicatum*, czyli badane obiekty. Ewentualna zmiana grozi zmniejszeniem porównywalności wyników, obniża trafność zmiennych i rzetelność narzędzi pomiaru. Jednym z czynników wydłużających czas badań są trudności w docieraniu do respondentów. Ankieterzy nie zastają ich w domu, spotykają się z odmowami, muszą powtarzać wizyty. Problem ten

doczekał się wielu analiz, których celem jest ustalenie wpływu różnych zjawisk na dostępność definiowaną, na ogół, w postaci liczby wizyt ankieterów. Analizowany jest wpływ wzorów spędzania wolnego czasu, liczby godzin pracy, przebywania w domu innych osób, które ułatwiają nawiązanie kontaktu, telefonicznego umawiania się na wywiad, jak również takich cech respondentów, jak narodowość czy posługiwanie się innym językiem, które utrudniają przeprowadzenie wywiadu (Smith 1983; Ward i in. 1985; Groves 1989; Smeets 1995).

Analizy te prowadzone są głównie w Stanach Zjednoczonych i w krajach zachodnich. Przeanalizuję tu rolę tych zjawisk na danych z Polski, gdzie nie były one przedmiotem systematycznych analiz. Należy podkreślić, że problematyka dostępności respondenta i czasu badań wymaga specjalnych danych, niewystępujących na ogół w standardowym badaniu. Podjęcie tego problemu umożliwił Europejski Sondaż Społeczny, w którego kwestionariuszu zamieszczono zestaw pytań metodologicznych, pozwalających na różne eksperymenty, dotyczące doskonalenia narzędzi badawczych.

Odwołując się do polskiej edycji ESS z 2002/2003 roku spróbuję odpowiedzieć na dwa pytania. Po pierwsze, od jakich czynników zależy dostępność respondentów – definiowana przez liczbę wizyt podejmowanych przez ankietera w celu nawiązania kontaktu. Po drugie, w jakim stopniu odpowiedzi respondentów zależą od punktu czasowego, kiedy realizowany jest wywiad. Chodzi o ustalenie, czy data realizacji wywiadu jest zmienną różnicującą odpowiedzi, niezależnie od wpływu innych cech. Występowanie takiej zależności wskazywałoby, że pomiary obciążone są systematycznym błędem, polegającym na zmianie kontekstu badania – niezależny wpływ czasu byłby dowodem na to, że w trakcie badania zmieniają się *indicatum* i warunki realizacji wywiadów.

Dane, zmienne, schemat analizy

Europejski Sondaż Społeczny jest realizowany na próbach dobieranych metodą losową, reprezentujących populację w wieku od 15 lat i więcej. Podstawą moich ustaleń będą dane pochodzące z pierwszej rundy ESS, zrealizowanej w 2002 i 2003 roku, która objęła w sumie 24 kraje. Badanie polskie przeprowadzono na przełomie 2002 i 2003 roku, a zrealizowana próba liczyła 2110 przypadków.

Szczegółową charakterystykę ESS od strony metodologii i celów badawczych można znaleźć w artykule Pawła Sztabińskiego (2004). W tym miejscu istotne jest to, że Europejski Sondaż Społeczny jest stosunkowo najnowszym przedsię-

wzięciem, jeżeli chodzi o międzynarodowe programy badawcze, podjętym przez Europejską Fundację Nauki w celu podniesienia standardów porównywalności międzykrajowych badań surveyowych. Z tego też względu, realizacji merytorycznego celu ESS, którym jest monitorowanie zmian w ramach integrującej się Europy, towarzyszą eksperymenty metodologiczne, realizowane za pomocą modułów dotyczących doboru próby, pomiaru zmiennych, zasad prowadzenia wywiadu i zapewnienia wysokiej realizowalności.

(a) Dostępność respondenta

Podstawowymi zmiennymi w analizach, które przedstawiam poniżej, są dostępność respondenta i data przeprowadzenia wywiadu. W większości dotychczasowych analiz, wskaźnikiem dostępności respondenta była liczba wizyt ankietera pod wylosowanym adresem zakończonych przeprowadzeniem (lub nie) wywiadu. Większą liczbę wizyt zwykle się traktować jako świadectwo mniejszej dostępności.

Wskaźnik ten nie jest bez wad, na co wielokrotnie zwracano uwagę. Wynikają one z niejednoznacznego charakteru tej informacji, a mianowicie, liczba wizyt nie odróżnia powtarzających się odmów respondenta od jego długotrwałej nieobecności w domu – spowodowanej np. urlopem – sytuacji choroby i wielu innych, „bardziej obiektywnych”, powodów (Stoop 2005). Za sytuacjami tymi kryją się różne mechanizmy, których nie można oddzielić posługując się liczbą wizyt – argumentują krytycy. Z tego też względu lepszym wskaźnikiem dostępności jest liczba wizyt do pierwszego kontaktu z respondentem, ponieważ eliminuje on sytuacje powtarzania się odmów, identyfikujących bardziej niechęć do „kooperacji” niż trudność „kontaktu” (Groves i Couper 1998).

Podstawowym warunkiem trafności każdej zmiennej jest jednowymiarowość, stąd też trudno się z tymi argumentami nie zgodzić. Z kolei zaletą liczby wizyt jest sumaryczny charakter tej zmiennej. W naszych analizach posłużymy się zarówno liczbą wizyt definiowaną ogółem, jak i liczbą wizyt ograniczonych do pierwszego kontaktu. W badaniu z 2002/2003 roku ankieterzy wypełniali dodatkowy kwestionariusz (będący rozszerzoną wersją „ankiety dla ankietera”), gdzie odnotowywali m.in. liczbę wizyt, ich daty (miesiąc, dzień, godzina, minuta) i rezultat wizyty: (i) wywiad zrealizowany w całości lub częściowo, (ii) kontakt z respondentem niezakończony wywiadem, (iii) kontakt pod wylosowanym adresem z kimś innym, (iv) brak jakiegokolwiek kontaktu, (v) respondent niedostępny (wyjechał, umarł). Informacje te zakodowano dla trzech ankieterów, ponieważ – zgodnie z praktyką – ankieterzy mogli się zmieniać. Należy

stwierdzić, że występowało to rzadko - na 3128 wszystkich wizyt, tylko w 145 przypadkach wysyłano do tych samych respondentów drugiego ankietera i tylko w 5 sytuacjach trzeciego.

Informacje o liczbie wizyt zakodowano w postaci zmiennych o wartościach od 1 do 7. Konstruując zmienną „dostępność respondenta” połączyłem te 3 zmienne składowe w sumaryczny wskaźnik o wartościach od 1 do 7. Analogicznie zoperacjonalizowałem „dostępność respondenta” w wersji rekomendowanej przez cytowanych wyżej krytyków tradycyjnego podejścia - gdzie wskazuje się na większą trafność liczby wizyt ograniczonej do pierwszego kontaktu. Różnica między dwiema wersjami polega na tym, że w tej drugiej, rekomendowanej - „liczba wizyt” została ograniczona do sytuacji przeprowadzenia wywiadu lub kontaktu niezakończonych wywiadem. Ujmując problem technicznie, w przypadku pierwszej wersji, zmienna „dostępność respondenta” przybiera wartość 1, 2, ..., 7, dla wszystkich kolejnych wizyt, tzn. występujących również po pierwszym kontakcie z respondentem, gdy respondent przy kolejnym spotkaniu odmawia, nie ma go w domu lub do wywiadu nie dochodzi z jakichś innych powodów. W drugiej wersji - zmienna ta przybiera wartość 1, gdy rezultatem pierwszej wizyty ankietera był wywiad lub jakiś kontakt z respondentem niezakończony wywiadem, wartość 2, gdy do wywiadu lub kontaktu doszło przy drugiej wizycie itd.

Tabela 1. Rozkład liczby wizyt ankieterów w badaniu Europejski Sondaż Społeczny. Polska 2002/2003

	Liczba wizyt	
	do przeprowadzenia wywiadu	do pierwszego kontaktu
1	54,4	81,2
2	26,4	15,2
3	10,9	2,6
4	5,3	0,6
5	1,6	0,2
6	0,7	0,1
7	0,7	0,1
	100,0	100,0

W tabeli 1 przedstawiono rozkłady wartości dla tej zmiennej w obu wersjach. Z danych tych wynika - co stanowi prawidłowość występującą w przypadku każdego badania - że większość kontaktów między ankieterem a respondentem

ogranicza się do jednej wizyty (przy szerokim rozumieniu kontaktu, obejmującym również dotarcie przez ankietera do wylosowanego adresu). W pierwszej kolumnie tabeli przedstawiono rozkłady wartości dla liczby wszystkich wizyt. Odsetek kontaktów ograniczających się do jednej wizyty (zakończonych lub niezrealizowania wywiadu) stanowił 54,4%. W trakcie drugiej wizyty, do zakończenia sytuacji badania dochodziło w 26,4% przypadków. Obydwie sytuacje objęły w sumie cztery piąte (80,8%) wszystkich prób podejmowanych przez ankieterów. Trzy wizyty objęły aż 91,7% zrealizowanych wywiadów.

W drugiej kolumnie tabeli 1 przedstawiamy rozkład liczby wizyt do pierwszego kontaktu. Jest on jeszcze bardziej prawostronnie skośny niż liczba wizyt ogółem. Podczas pierwszej wizyty ankietera miało miejsce aż 81,2% przypadków pierwszego kontaktu. Podczas drugiej wizyty było ich 15,2%, natomiast tylko 1% sytuacji pierwszego kontaktu występowało w trakcie czwartej wizyty i następnych.

Przejdźmy do schematu analizy tych danych. Nasze główne pytanie badawcze – czynniki, od których zależy dostępność respondenta – dotyczy zjawiska usytuowanego w przedziale czasowym. Jest to sytuacja, którą na gruncie analiz statystycznych zwykło się rozpatrywać w ramach podejścia nazywanego „analizami historii przypadku” (*event history analysis*). Dla danego przypadku (u nas: wizyty ankietera) istnieje określony przedział czasu, który dzieli się na okres „ryzyka” (*the risk period*) – zakończonego wystąpieniem zjawiska – i okres po jego wystąpieniu (*non-risk*). W przedziale tym określony jest rozkład szans wystąpienia danego zjawiska, liczonych od punktu początkowego – u nas jest nim rozpoczęcie badania. W miarę upływu czasu szanse te – określane mianem *hazard rate* (lub *transition rate*) – ulegają zmianie (Yamaguchi 1992; Box-Steffensmeier i Jones 2004). Głównym celem analizy jest ustalenie siły wpływu różnych czynników na szanse przejścia ze stanu ryzyka do nieryzyka, którym to przejściem może być przeprowadzenie wywiadu lub zakończenie kontaktu w jakiś inny sposób. Modele historii przypadku pozwalają na dokładniejszą estymację parametrów tych zależności w porównaniu z klasycznymi modelami regresji, z tym, że wymagają one danych zdefiniowanych na trajektorii czasu, na ogół rzadko występującymi w ramach standardowego badania.

Wizyty ankietera można więc traktować jako przypadek (*event*), mający określoną historię, ze zmieniającą się szansą realizacji. Chcemy ustalić, od czego ta szansa zależy. W celu rozstrzygnięcia tej kwestii posłużyłem się modelem występującym w literaturze statystycznej pod nazwą *continuous-time proportional hazard model* – model proporcjonalnego ryzyka. Jest on jednym z najczęściej stosowanych modeli nieparametrycznych, wprowadzonych przez Davida Coxa –

zaletą modeli nieparametrycznych jest to, że opierają się one na założeniu o zależności rozpatrywanych zjawisk od czasu, ale bez konieczności wyspecyfikowania tej zależności w postaci funkcyjnej (Cox i Oakes 1984). W modelu proporcjonalnego ryzyka zakłada się, że szansa wystąpienia rozpatrywanego zjawiska jest funkcją parametrów ustalonych dla zmiennych niezależnych (*covariates*). Konkretna wartość tej funkcji, czyli szansy wystąpienia analizowanego zjawiska dla przypadku i w czasie t wynosi $h_i(t)$ i określona jest wzorem:

$$h_i(t) = h_0(t)\exp(\beta'x),$$

gdzie $h_0(t)$ identyfikuje funkcję ryzyka w modelu wyjściowym (*baseline*), a $\beta'x$ są parametrami regresji i zmiennymi niezależnymi – tzn. wyznacznikami ryzyka.

Formułując te zależności w języku naszego problemu można powiedzieć, że analizie poddamy dostępność respondentów, definiowaną przez liczbę wizyt zakończonych „sukcesem”. Sukcesem jest tu zakończenie kontaktu z respondentem (adresem) w postaci realizacji wywiadu lub stwierdzenia definitywnej niedostępności respondenta. Sukces zdefiniowałem w postaci dwóch zmiennych, identyfikujących łatwiejszą i trudniejszą dostępność. Dla łatwiejszej dostępności respondenta sukcesem będzie zakończenie kontaktu podczas pierwszej wizyty; dla dostępności trudniejszej – sytuacja zakończenia kontaktu w czasie od wizyty trzeciej do siódmej. Estymacja wartości $\beta'x$ dostarczy parametrów, charakteryzujących zależność szans tak definiowanych sukcesów od wybranych czynników.

Wyniki dotychczasowych analiz wskazują, że o dostępności respondentów decyduje wiele czynników. Można je podzielić na kilka grup. Pierwsza z nich związana jest z cechami położenia społecznego respondentów, takimi jak wiek, miejsce zamieszkania, wykształcenie czy status zawodowy. W świetle badań, prowadzonych głównie w Stanach Zjednoczonych i krajach Europy Zachodniej, jednostki niedostępne rekrutują się bardziej z ludzi młodych niż starych, mieszkańców dużych miast i kategorii o wyższym statusie, bardziej zajętych pracą zawodową i mających stosunkowo mniej czasu na wywiad. Druga grupa wyznaczników dostępności obejmuje zaufanie do innych, skłonność do kooperacji, poczucie obowiązku obywatelskiego i inne postawy. Okazuje się, że – niezależnie od pozycji zawodowej czy wieku – większej dostępności sprzyjają wyższe zaufanie do innych osób, zainteresowanie polityką, mniejszy indywidualizm (w sensie polegania na sobie) i aktywny udział w organizacjach samorządowych. Przedmiotem osobnych analiz były przekonania o wartości badań surveyowych, gdzie – zgodnie z oczekiwaniami – respondenci, bardziej przekonani o ich użyteczności i mający pozytywny stosunek do badań, rzadziej odmawiają

wywiadu i łatwiej się z nimi umówić. Warto podkreślić, że czynniki te występują w różnych kombinacjach i bardziej pogłębione analizy dowodzą, że np. wyższe wykształcenie i pozycja zawodowa jednostek ułatwiają rozmowę z ankietowanym, sprzyjają traktowaniu wywiadu w kategoriach powinności obywatelskich, a przez to mogą zmniejszać prawdopodobieństwo odmowy. Stosunkowo najnowszego przeglądu tych analiz dokonała Ineke Stoop (2005) w książce *The Hunt for the Last Respondent*.

W kwestionariuszu ESS zamieszczono pytania, które pozwoliły na skonstruowanie wskaźników dla niektórych z tych zjawisk. Z podstawowych cech położenia społecznego uwzględniłem we wstępnych analizach płeć (mężczyźni = 1, kobiety = 0), poziom wykształcenia (liczba efektywnych lat nauki), dochody rodziny na osobę, stan cywilny, fakt bezrobocia (bezrobotni = 1, pozostali = 0) i tygodniowy czas pracy. W kontekście niedostępności ważną zmienną jest wiek, którą wprowadziłem do analizowanego modelu w postaci kategoryjnej (15-25, 26-40, 41-55, 56-65, 66 i powyżej), miejsce zamieszkania (zmienna kategoryjna w podziale na: duże miasto, przedmieścia dużego miasta, małe miasto, wieś) oraz pozycja społeczno-zawodowa. W tym ostatnim przypadku posłużyłem się podziałem na pięć „klas społecznych”, korzystając ze schematu operacjonalizacji występującego pod nazwą EGP. Jest to standardowy i - w analizach porównawczych - najczęściej stosowany wskaźnik miejsca w strukturze społecznej, odwołujący się do teorii Webera. W polskiej literaturze socjologicznej szczegółowy opis jego konstrukcji przedstawili Domański i Przybysz (2003). EGP wprowadziłem do modelu w podziale na: inteligencję, pracowników umysłowych niższego szczebla, właścicieli firm, robotników i rolników.

W analizie czynników dostępności wstępnej selekcji poddałem również: udział w ostatnich wyborach parlamentarnych (dychotomiczna zmienna „tak-nie”), częstość spotykania się ze znajomymi i krewnymi (skala od „nigdy” do „codziennie”), a także zmienne świadomościowe - stopień zaufania do innych, samoocenę poglądów politycznych na dziesięciopunktowej skali lewica-prawica, ocenę polityków w Polsce (skonstruowaną z odpowiedzi na dwa pytania: czy polityków interesują sprawy zwykłych ludzi, i czy dbają oni raczej o opinie czy głosy wyborców), samoocenę poziomu szczęścia z życia (na dziesięciopunktowej skali) i religijność (zmienna utworzona z odpowiedzi na pytania dotyczące tego, jako często respondent modli się i chodzi na msze). Miernikiem konserwatyzmu obyczajowego jest opinia na temat tego, czy „gejowie i lesbijki powinni mieć swobodę do układania sobie życia według własnych przekonań?”, a tradycjonalizmu - „czy dla kraju jest lepiej, gdy istnieje więcej wyznań religijnych?” (w obu przypadkach pięciopunktowe skale, od „zdecydowa-

nie się zgadzam” do „zdecydowanie nie zgadzam”). Ostatnią pozycją na liście zmiennych niezależnych było poczucie zagrożenia, zdefiniowane w postaci dychotomicznej, na podstawie odpowiedzi na pytanie: „czy P. lub ktoś z P. gospodarstwa domowego był w ciągu ostatnich 5 lat ofiarą napadu, włamania” (tak=1, nie=0).¹

(b) Data przeprowadzenia wywiadu

Do skonstruowania tej zmiennej posłużyły informacje dotyczące miesiąca i dnia przeprowadzenia wywiadu. Realizacja ESS zajęła cztery miesiące. Pierwszy wywiad przeprowadzono 30 października 2002 roku, ostatni – 19 stycznia 2003 roku. Łącząc dane, dotyczące miesięcy i dni otrzymujemy zmienną ilościową zakodowaną w postaci trzech cyfr, gdzie pierwsza identyfikuje miesiąc, a dwie następne – kolejne dni miesiąca. Tak więc wartości tej zmiennej zawierają się w przedziale od 130 do 419. Potraktujemy ją jako zmienną wyjaśniającą, próbując ustalić, na ile moment realizacji wywiadu różnicuje w znaczącym stopniu odpowiedzi respondentów. Zmiennejmi wyjaśnianymi będą niektóre ze zmiennych zdefiniowanych powyżej. Poza tym analizuję wpływ daty realizacji badania na stosunek do imigrantów, definiowany przez odpowiedzi na dwa pytania: „czy zgodził(a)by się P., żeby P. bliski(a) krewny(a) zawarł(a) związek małżeński z osobą innej rasy, lub przynależności etnicznej?” i „czy zgodził(a)by się P., żeby osoba innej rasy, lub przynależności etnicznej była P. szefem w pracy?”. Odpowiadając, respondenci posługiwali się skalami od 0 do 10 – skonstruowałem na tej podstawie wskaźnik będący sumą tych odpowiedzi na skali malejącej akceptacji od 0 do 20. Analizuję też stosunek do prawa („czy zgadza się P. ze stwierdzeniem, że należy zawsze przestrzegać prawa?” – mierzony na pięciopunktowej skali, od „zdecydowanie się zgadzam” do „zdecydowanie nie zgadzam”), samoocenę satysfakcji z życia (na skali od 0 do 10), poczucie bezpieczeństwa (odpowieź na pytanie, „czy czuje się P. bezpiecznie, spacerując samemu tam gdzie P. mieszka, po zmierzchu?” – czteropunktowa skala od „bardzo bezpiecznie” do „bardzo niebezpiecznie”), oraz opinię na temat tego, czy przyjaciele są ważną wartością (skala od 0 do 10). W tym ostatnim przypadku respondentom przedstawiono listę kilku wartości prosząc o ocenę, na ile są one ważne. Uwzględniam tu tylko oceny dotyczące wartości przyjaciół, ponieważ tylko one znacząco korelują z datą realizacji wywiadu.

¹ Nie podaję szczegółowego opisu operacjonalizacji niektórych z tych zmiennych i brzmienia pytań, ponieważ okazały się one słabo związane z liczbą wizyt i pominąłem je w dalszych analizach.

Wyniki analiz

Liczba wizyt

Zacznijmy od przyjrzenia się zależnościom między liczbą wizyt, traktowaną jako wskaźnik dostępności respondenta, a możliwymi korelatami dostępności. W tabeli 2 przedstawiłem wartości stosunków korelacyjnych między liczbą wizyt zdefiniowaną w postaci zmiennej ilościowej, na skali od 1 do 7, a zmiennymi „obiektywnymi” i wskaźnikami postaw. Posługujemy się stosunkiem korelacyjnym, ponieważ rozkład liczby wizyt jest skośny, co oznacza duże prawdopodobieństwo występowania zależności nieliniowych.²

Tabela 2. Zależności między liczbą wizyt ankietera a wybranymi zmiennymi

Stosunki korelacyjne między liczbą wizyt a:	Eta
Płcią	0,08
Wiekem	0,11**
Miejscem zamieszkania	0,28**
Liczbą lat nauki	0,23**
Liczbą godzin pracy w tygodniu	0,06
Bezrobociem	0,09*
Wysokością dochodów rodziny na osobę	0,17**
Stopniem religijności respondenta	0,16**
Oceną polityków	0,07
Zaufaniem do innych	0,06
Zainteresowaniem polityką	0,11**
Stosunkiem do pluralizmu wyznań religijnych w Polsce	0,15**
Głosowaniem w ostatnich wyborach	0,04
Lokowaniem się na skali lewicowość–prawicowość	0,07
Satysfakcją z życia	0,07
Stosunkiem do przyznawania praw gejom i lesbijkom	0,13**
Przekonaniem, że prawa należy zawsze przestrzegać	0,05
Samooceną szczęścia	0,07
Częstością spotkań z przyjaciółmi i rodziną	0,05
Poczuciem bezpieczeństwa	0,07
Byciem ofiarą napadu lub kradzieży – respondenta lub członka jego rodziny	0,15**

* $p < 0,05$; ** $p < 0,01$

² Stosunek korelacyjny jest miernikiem uwzględniającym występowanie zależności nieliniowych, którego wartości zawierają się w przedziale od 0 do 1. Przy obliczaniu wartości przedstawionych w tabeli 2, potraktowałem liczbę wizyt jako zmienną niezależną.

Mimo że uwzględniliśmy zmienne, które powinny być czynnikami różnicującymi dostępność respondentów, zależności są niskie. Stosunkowo najsilniejszymi korelatami liczby wizyt okazują się: poziom wykształcenia i wielkość miejsca zamieszkania. Zgodnie z oczekiwaniami, dostępność respondentów maleje ze wzrostem poziomu wykształcenia (0,23) i liczby mieszkańców (0,28). Przejawem tej pierwszej tendencji jest wzrost średniej liczby lat nauki w kategoriach respondentów wyróżnionych ze względu na numer kolejnej wizyty. W świetle danych, których nie przedstawiam w tabelach, liczba lat nauki zwiększa się z 10,8 lat dla respondentów, z którymi wywiad przeprowadzono w czasie pierwszej wizyty, do 13,1 lat dla osób, z którymi zrealizowano go w czasie wizyty siódmej.

Znaczącymi korelatami są również: poziom dochodów (0,17) i to, czy było się ofiarą rabunku (0,15). Obydwa te czynniki zmniejszają dostępność, przy czym zależność z dochodami jest krzywoliniowa. Najniższe dochody na osobę (1045 zł) mają respondenci udzielający wywiadu w czasie pierwszej wizyty, następnie średnia ta wzrasta, do wielkości 2063 zł dla czwartej wizyty, po czym trochę się zmniejsza, osiągając wielkość 1143 zł dla kategorii respondentów, z którym zrealizowano wywiad w czasie siódmej wizyty.

W świetle wcześniejszych analiz (które cytowałem powyżej) mniej dostępnymi kategoriami respondentów powinni być bardziej mężczyźni – ponieważ kobiety częściej są w domu – i osoby, które dłużej pracują. W istocie, ze wzrostem liczby wizyt odsetek respondentek-kobiet systematycznie maleje. Było ich 54% wśród respondentów, z którymi zrealizowano wywiad w czasie pierwszej wizyty i 37% w czasie ostatniej, natomiast dotarcie do mężczyzn jest stosunkowo trudniejsze. Trudniej dociera się również do respondentów poświęcających więcej czasu na pracę, chociaż nie jest to zależność liniowa. Tygodniowy czas pracy w kategorii dostępnej już w czasie pierwszej wizyty wynosi średnio 46,5 godzin, dla drugiej i trzeciej wizyty kształtuje się on na poziomie 45,7–45,8 godzin, wzrasta do 50,6 dla piątej wizyty, następnie (w przypadku szóstej wizyty) zmniejsza się do 45 godzin, aby dla ostatniej wizyty wzrosnąć do 57,4 godzin. Występowanie tych zależności nie zmienia faktu, że nie są one statystycznie znaczące, co nie pozwala ich uogólniać na populację respondentów. Natomiast statystycznie istotny, chociaż stosunkowo nieduży, okazuje się wpływ sytuacji bezrobocia. Z jednej strony, pozostawanie bez pracy sprzyja poczuciu izolacji i wyłączenia, zniechęcając do współpracy z ankierem, z drugiej – osoby bezrobotne mają mniej ograniczeń czasowych. Czynnik czasu jest prawdopodobnie ważniejszy, czego świadectwem jest relatywnie największy udział osób bezrobotnych wśród respondentów, z którymi zrealizowano wywiad w czasie pierwszego kontaktu

(11% w stosunku do 9,3% bezrobotnych wśród ogółu badanych). Odsetek ten dla kolejnych wizyt maleje – w czasie drugiej wizyty wyniósł 9%, a w kolejnych odpowiednio: 4,5; 4,3; 4,6 i zero.

Badania realizowane w Holandii wskazują, że większą dostępnością charakteryzują się respondenci mający wyższy poziom zaufania i liczby kontaktów, których wskaźnikiem może być częstość spotykania się ze znajomymi i krewnymi. Ważne było również zainteresowanie polityką, którego wzrost był czynnikiem ułatwiającym realizację wywiadu (Stoop 2005). Posługując się podobnymi zmiennymi, próbowałem potwierdzić występowanie tych zależności dla Polski, uwzględniając również ocenę działalności polityków i samoocenę poglądów na skali „lewica–prawica”. W przypadku Polski wielkości stosunków korelacyjnych dla wszystkich tych zależności okazują się statystycznie nieznaczące i niskie.

Realistyczne wydaje się założenie, że czynnikami ułatwiającymi dostępność są poczucie szczęścia, satysfakcji i pozytywne nastawienie do życia, które powinny osłabiać bariery nawiązywania kontaktów. Wskaźnikiem tych orientacji jest samoocena szczęścia, która jednak – jak widać – nie wchodzi w istotne korelacje z liczbą wizyt. Nikt tego dotąd nie badał, ale czynnikiem zwiększającym dostępność może być konserwatyzm obyczajowy, zakładając, że osoby konserwatywne, o bardziej restryktywnych poglądach, charakteryzują się większym poczuciem obowiązku, co dotyczy też obowiązku udzielania ankieterowi wywiadu. Konserwatyzm obyczajowy mierzony jest tu przez stopień akceptacji dla praw lesbijek i gejów i dla występowania w kraju wielu różnych religii. Przypomnijmy, że postawy te mierzone są na skalach od 1 do 5, gdzie odpowiednio wyższe wartości oznaczają mniejszą aprobatę dla przyznawania praw homoseksualistom i dla pluralizmu religii.

Porównanie średnich wartości obu tych skal (których tu nie przedstawiam) wskazuje, że – rzeczywiście – najwyższymi wartościami charakteryzują się respondenci, z którymi wywiad zrealizowano w czasie pierwszej wizyty. Sugeruje to, że najbardziej dostępnymi respondentami są ludzie o małej tolerancji obyczajowej, których cechuje konserwatyzm. Innym elementem tego syndromu jest religijność, czego świadectwem jest znacząca wielkość stosunku korelacyjnego (0,16) między liczbą wizyt a stopniem religijności mierzonym na sumarycznej skali obejmującej częstość modlenia się i chodzenia na msze. Osoby o największej religijności okazują się nadreprezentowane wśród respondentów uchwyconych już w czasie pierwszej wizyty.

Dostępność respondenta

Należy sądzić, że niektóre z tych zależności mają charakter pozorny. Dotyczy to zwłaszcza religijności, konserwatyizmu obyczajowego i wieku. Religijność kojarzy się z tradycjonalizmem i niską mobilnością społeczną ludzi starszych i gospodyń domowych przebywających w domu. Chcąc ustalić czysty wpływ tych czynników na stopień dostępności respondentów musimy je wyspecyfikować w ramach jednego modelu. Modelem, którym się tu posłużyłem, a który pozwala zidentyfikować niezależny wpływ wielu zmiennych na zjawisko zmieniające się w czasie jest model proporcjonalnego ryzyka. Zastosujemy go dla dwóch sytuacji – w celu ustalenia czynników różnicujących prawdopodobieństwo realizacji wywiadu w czasie pierwszej wizyty i realizacji wywiadu po kilkakrotnym niepowodzeniu, gdy ankieter próbuje dotrzeć do respondenta po raz trzeci, czwarty, piąty, szósty i siódmy.³

W tabeli 3 przedstawiono wartości parametrów dla finalnej wersji tego modelu, po przetestowaniu go dla różnych zestawów zmiennych wyjaśniających. Jest to model najbardziej adekwatny w tym sensie, że zmienne uwzględnione w tabeli 3 okazały się w sumie najsilniejszymi wyznacznikami dostępności. Wartości parametrów zamieszczamy w postaci multiplikatywnej (wykładniczej), gdzie 1 identyfikuje brak zależności, wartości powyżej 1 informują o zależnościach dodatnich, a poniżej 1 – o zależnościach ujemnych.

Okazuje się, że zależności są niskie. Szanse przeprowadzenia wywiadu w czasie pierwszej wizyty zależą tylko od poziomu wykształcenia i przynależności społeczno-zawodowej respondentów. Zgodnie z oczekiwaniami, zależność z wykształceniem jest negatywna, tzn. im więcej lat nauki, tym mniejsze są szanse realizacji wywiadu, czy ujmując problem ogólniej – zakończenia kontaktu. Wiadomo, że wyższym wykształceniem charakteryzują się kadry kierownicze i reprezentanci zawodów, których wykonywanie angażuje czasowo, podczas gdy reprezentanci niższych szczebli wykształcenia są mniej zajęci pracą i częściej są w domu. W przypadku zawodu najłatwiej dostępną kategorią są niżsi rangą pracownicy umysłowi, czyli technicy, pielęgniarki, księgowi, referenci, szeregowi pracownicy handlu i usług. Z przedstawionych tu analiz wynika, że ankieterzy 1,16 razy częściej realizują z nimi wywiad – w stosunku do hipotetycznej

³ W obydwu przypadkach zmienna wyjaśniana jest zmienną dychotomiczną. Za pierwszym razem respondenci, z którymi zrealizowano wywiad w czasie pierwszej wizyty, otrzymują kod 1, a wszyscy pozostali 0; za drugim razem wszyscy respondenci, z którymi zrealizowano wywiad w czasie od trzeciej wizyty do siódmej otrzymują kod 1, a pozostali 0).

Tabela 3. Zależności między dostępnością respondenta a wybranymi cechami respondenta. Współczynniki regresji w modelu proporcjonalnego ryzyka

Zmienne niezależne	Wywiad zrealizowany w czasie pierwszej wizyty	Wywiad zrealizowany dla wizyt 3-7
Kategorie społeczno-zawodowe EGP ^a		
Inteligencja	0,96	1,20
Pracownicy administracyjni średniego szczebla	1,16*	0,94
Właściciele	0,88	1,25
Robotnicy	0,94	0,87
Rolnicy	1,07	0,77
Wiek ^a		
15-25	0,89	1,41**
26-40	0,93	1,43**
41-55	0,99	1,03
56-65	1,11	1,00
66 i powyżej	1,15	0,53**
Bezrobocie	1,08	0,59**
Liczba lat nauki	0,96*	1,02
Płeć (1=mężczyźni)	0,96**	1,17
Ofiara napadu	1,10	0,79**
Miejsce zamieszkania ^a		
Duże miasto	0,92	1,60**
Przedmieście	0,98	1,25
Małe miasto	1,11	1,08
Wieś	0,96	0,43**

*p<0,05; ** p<0,01

^a Charakterystyką statystycznej istotności współczynników regresji w modelu proporcjonalnego ryzyka są statystyki Wald. W przypadku zależności dla wszystkich 5 kategorii wieku wartość Wald (w drugim modelu w tabeli 3) wyniosła 19,53, przy p<0,001. Z kolei, dla zależności określającej istotność wpływu wszystkich 4 kategorii miejsca zamieszkania wartość Wald (również w modelu drugim) wyniosła 48,92, przy p<0,000. Statystycznie nieistotny jest całkowity wpływ przynależności zawodowej (5 kategorii) na realizację wywiadu w czasie pierwszej wizyty i wizyt 3-7 (pierwszy i drugi model) i całkowity wpływ wieku i miejsca zamieszkania na realizację wywiadu w czasie pierwszej wizyty (pierwszy model).

średniej - w czasie pierwszej wizyty. Natomiast faktem bez istotnego znaczenia okazuje się przynależność respondentów do pozostałych kategorii - inteligencji, właścicieli, robotników i rolników.⁴

Dla dostępności definiowanej przez realizację wywiadu w czasie wizyt od 3 do 7, szanse te kształtowały się trochę inaczej. Z danych zamieszczonych w drugiej kolumnie tabeli 3 wynika, że na tym późniejszym etapie przestają mieć istotne znaczenie poziom wykształcenia i zawód, natomiast dochodzą do głosu wiek respondenta, miejsce zamieszkania i fakt bezrobocia. Jeżeli chodzi o wiek, to najtrudniej dostępnymi respondentami są ludzie młodzi. Respondenci z dwóch najmłodszych kategorii (15-40 lat), byli znacznie częściej osiągalni w czasie ostatnich wizyt (w granicach 1,41-1,43) w stosunku do średniej. Odwrotnie z bezrobociem, które jest czynnikiem zwiększającym dostępność. Osoby bezrobotne mają - w porównaniu z osobami pracującymi - 1,69 (1/0,59) mniej szans na znalezienie się wśród respondentów najtrudniej dostępnych. Wśród niedostępnych respondentów wyraźnie nadreprezentowani są również mieszkańcy największych miast (1,6 razy częściej w stosunku do średniej). Zgodnie z oczekiwaniami, „szanse” niedostępności maleją dla respondentów z małych miejscowości, a wśród mieszkańców wsi są one znacząco najmniejsze (0,43).

Liczba wizyt jest sumarycznym wskaźnikiem dostępności respondentów. Jednak - jak odnotowaliśmy we wstępie - jego krytycy nie bez słuszności wskazują, że identyfikuje on bardzo różne aspekty, które należy wyłączać, co w szczególności dotyczy powtarzającej się odmowy na wywiad. Rozwiązaniem stosowanym przez wielu badaczy jest posługiwanie się liczbą wizyt do pierwszego kontaktu. W tabeli A1 (zamieszczonej w aneksie na końcu) przedstawiamy wyniki analiz uzyskanych dla modelu proporcjonalnego ryzyka, po zastąpieniu liczby wizyt ogółem przez liczbę wizyt do pierwszego kontaktu z respondentem. Zależności są nieistotne statystycznie, trudno więc na tej podstawie dokonywać uogólnień, poza stwierdzeniem braku systematycznych tendencji. W kwestii niedostępności wyniki te nie wnoszą nic ponad to, co powiedzieliśmy

⁴ Wpływ kategorii zawodowej, wieku i miejsca zamieszkania zdefiniowany jest w postaci serii zmiennych zero-jedynkowych metodą „kodowania efektów”. Pozwala to interpretować wartości współczynników regresji dla poszczególnych kategorii jako różnicę między szansami dostępności respondentów z tej kategorii a hipotetyczną średnią dostępnością w próbie. Wartości dla kategorii, które należało pominąć - co jest warunkiem estymacji parametrów dla zmiennych kodowanych w postaci zestawu kategorii - zostały ustalone w ramach modeli, w których pominąłem odpowiednio inne kategorie, a uwzględniłem kategorie pominięte w modelu wyjściowym (byli to odpowiednio: rolnicy, osoby najstarsze i mieszkańcy wsi).

powyżej. Niedostępność ograniczona do sytuacji pierwszego kontaktu okazuje się mniej czytelna, jeżeli chodzi o wyznaczniki tych zjawisk.

Czas

Wpływ cech respondentów na dostępność sygnalizuje jedno ze źródeł błędów obniżających trafność pomiaru. Błędów tych nie można wyeliminować do końca, natomiast istnieją techniki statystyczne, pozwalające je kontrolować na etapie analiz (Domański 2002). Innym źródłem błędów może być wpływ momentu realizacji wywiadu. Chodzi o porównywalność wyników uzyskiwanych, powiedzmy, na początku, w środku i na końcu badania. Empirycznym świadectwem występowania tej zależności byłby wpływ daty przeprowadzenia wywiadu na zjawiska będące przedmiotem analiz.

Aby to stwierdzić zacznijmy od przyjrzenia się sile zależności między datą przeprowadzenia wywiadu, a różnymi zmiennymi identyfikującymi zjawiska, które w analizach socjologicznych są przedmiotem zainteresowania badaczy. Data przeprowadzenia wywiadu jest typową zmienną ilościową, o wartościach zawierających się w przedziale wyznaczonym przez dzień rozpoczęcia i zakończenia badania. Europejski Sondaż Społeczny realizowany był od 30 października 2002 do 19 stycznia 2003 roku. Tak więc zmienna ta liczy kilkadziesiąt wartości. Aby nie zamykać sobie możliwości uchwycenia związków nieliniowych, zdefiniowałem ją w postaci kategorialnej. Jest ona podziałem na 10 równolicznych kategorii – przedziałów czasowych – wyznaczonych na podstawie rozkładu decylogowego zmiennej ilościowej.

Pierwszym krokiem było ustalenie siły różnicującej tego podziału dla kilkadziesiątu zmiennych skonstruowanych na podstawie odpowiedzi respondentów. Były to wybrane aspekty położenia społecznego i świadomości jednostek. Należy podkreślić, że dobór tych zmiennych był przypadkowy, tzn. nie przyświecał nam cel rozstrzygnięcia jakiegoś merytorycznego problemu i potraktowaliśmy je wyłącznie jako kryteria pozwalające ocenić różnicujący wpływ czasu realizacji badania pod kątem trafności i rzetelności pomiaru.

Miernikiem siły różnicującej podziału na 10 okresów realizacji były wartości stosunków korelacyjnych. Po przeprowadzeniu wstępnych analiz okazało się, że data realizacji jest znacząco skorelowana z 7 zmiennymi: poziomem dochodów rodziny na osobę (0,12), faktem bezrobocia (0,09), samooceną satysfakcji z życia (0,12), poczuciem bezpieczeństwa (0,14) – identyfikowanym przez pytanie, czy respondent czuje się bezpiecznie, spacerując samemu po zmierzchu, konserwatyzmem obyczajowym, mierzonym przez stosunek do praw lesbijek i ge-

jów (0,09), stosunkiem do imigrantów (wskaźnik akceptacji dla małżeństwa i podlegania w pracy osobie innej narodowości lub rasy) – (0,11), oraz opinią na temat tego, czy przyjaciele są w życiu ważną wartością (0,10).

Występowanie znaczących korelacji może wskazywać, że data przeprowadzenia wywiadu wywiera pewien wpływ na odpowiedzi badanych. Z punktu widzenia trafności i rzetelności pomiaru byłby to fakt, który skłania do obaw. Zależność od punktu czasowego badania zmniejsza trafność, ponieważ świadczy o tym, że pytania mierzą coś wykraczającego poza operacyjną definicję wskaźnika. Na przykład, wartości na skali satysfakcji z życia byłyby nie tylko odzwierciedleniem trwałych postaw respondentów, ale również przejściowych nastrojów i jakichś innych okoliczności związanych z czasem realizacji wywiadu. Występowanie tych zależności obniża również rzetelność, ponieważ dowodzi braku stabilności narzędzia pomiaru.

Aby to sprawdzić, należało wyspecyfikować w miarę „czyste” zależności tych zjawisk od daty realizacji wywiadu. Dla zmiennych świadomościowych ustaliłem je, kontrolując podstawowe cechy położenia społecznego respondentów: płeć, wiek, wielkość miejsca zamieszkania, przynależność społeczno-zawodową, fakt bezrobocia, dochody rodziny na osobę, poziom wykształcenia i tygodniowy czas pracy. Te same zmienne kontrolowałem analizując wpływ daty realizacji wywiadu na deklarowane dochody i fakt bezrobocia – oczywiście po wyłączeniu każdej z tych zmiennych z modelu, w którym występują jako zmienna zależna.

Tabela 4. Zależności między datą realizacji wywiadu a wybranymi zmiennymi

Siły korelacji między liczbą wizyt a:	Eta	Współczynnik korelacji częściowej	Statystyka F	Prawdopodobieństwo
Wysokością dochodów rodziny na osobę	0,12	0,002	0,948	0,390
Stopniem akceptacji imigrantów	0,11	0,018	3,189	0,001
Satysfakcją z życia	0,12	0,014	2,434	0,010
Stosunkiem do przyznawania praw gejom i lesbijkom	0,09	0,006	0,996	0,441
Uznanie, że przyjaciele są w życiu ważną wartością	0,10	0,010	1,757	0,072
Poczuciem bezpieczeństwa	0,14	0,009	1,706	0,083

W tabeli 4 przedstawiono kolejno: wielkości stosunków korelacyjnych, współczynników korelacji częściowej oraz statystyki F i poziomów istotności

dla zależności między sześcioma zmiennymi a datą realizacji wywiadu.⁵ Współczynniki korelacji częściowej – które ustaliłem posługując się analizą wariancji – można traktować jako charakterystyki zróżnicowania między odpowiedziami respondentów, wynikającego z daty realizacji wywiadu. Z kolei analizując wpływ tego czynnika na bezrobocie (zmienna zero-jedynkowa) posłużyłem się regresją logistyczną (danych tych nie przedstawiam) W obu przypadkach data realizacji badania została zdefiniowana w postaci 9 zmiennych zero-jedynkowych, gdzie ostatni przedział czasu potraktowano jako tzw. „zmienną referencyjną”, która się w danym modelu pomija.

Data realizacji wywiadu okazuje się czynnikiem różnicującym tylko stosunek do imigrantów i samoocenę satysfakcji z życia. Wprawdzie są to tylko dwie zmienne na siedem, ale nie powinna ona ich różnicować w ogóle. Występowanie tych zależności nie pozwala przyjąć bezpiecznego założenia, że moment realizacji nie wywiera żadnego wpływu na odpowiedzi badanych. Jest to znaczący czynnik, z którym trzeba się liczyć.

Tabela 5. Wpływ daty realizacji wywiadu na stopień akceptacji imigrantów i satysfakcję z życia. Metryczne współczynniki regresji^a

Przedziały czasowe daty realizacji wywiadu	Stopień akceptacji imigrantów	Satysfakcja z życia
1 – od 30 października 2002	1,30*	0,68**
2	0,58	-0,02
3	-0,21	-0,23
4	0,86	0,61*
5	1,42	0,29
6	-0,19	0,52
7	0,41	0,17
8	-1,61*	0,45
9	-0,87	0,19
10 – do 19 stycznia 2003	-0,16	-2,56**

* $p < 0,05$; ** $p < 0,01$

^a Ustalane przy kontroli płci, wieku, wielkości miejsca zamieszkania, przynależności społeczno-zawodowej, faktu bezrobocia, dochodów rodziny na osobę, poziomu wykształcenia i tygodniowego czasu pracy.

⁵ Wielkości F i prawdopodobieństw p informują o statystycznej istotności tych zależności w modelach wielozmiennych.

Tak więc wyniki te sygnalizują obecność systematycznego błędu, który można zmniejszyć – przede wszystkim skracając czas badań. Dokładniejszą charakterystyką tego błędu są wielkości współczynników regresji zamieszczone w tabeli 5. Wartości te – ustalone w ramach analizy wariancji (w modelach, do których odwołujemy się w tabeli 4) – informują o tym, jak kształtował się stosunek do imigrantów i poziom satysfakcji z życia na poszczególnych etapach realizacji badania, po wyłączeniu płci, wieku, wykształcenia i pozostałych czynników. Pamiętając, że niższe wartości na skali stosunku do imigrantów identyfikują postawy pozytywne, stwierdzamy, że wystąpiły one częściej w końcowych etapach badania. Z trudnych do odgadnięcia powodów, respondenci, z którymi przeprowadzono wywiad w przedziałach czasu 8 i 9, stosunkowo częściej niż inni deklarowali brak sprzeciwu wobec bycia podwładnymi i małżeństwa bliskiego krewnego z imigrantami innej rasy/narodu. Jeżeli chodzi o zróżnicowanie satysfakcji z życia, to oceny te kształtują się przypadkowo, tzn. nie występuje tu żaden wzór, a najwyższymi samoocenami na skali satysfakcji charakteryzowali się respondenci na samym początku i bliżej środkowego punktu badania – w przedziale 1 (0,68) i 4 (0,61). Z kolei zdecydowanie najniższą samoocenę mieli respondenci udzielający wywiadu najpóźniej (-2,56).⁶

Wnioski

Przedstawione tu analizy dotyczyły kwestii dostępności respondentów i związku między odpowiedziami a datą realizacji wywiadu. Rozpatrywaliśmy je w kontekście błędów w badaniach surveyowych, odwołując się do danych Europejskiego Sondażu Społecznego z 2002/3 roku. Wymiernym świadectwem błędu, związanego z niedostępnością respondentów, jest występowanie zależności między cechami respondentów a szansą realizacji wywiadu. Wiadomo, że respondenci różnią się pod względem łatwości nawiązania kontaktu z ankietarem i wyrażania zgody na wywiad. Może to wywierać wpływ na udzielane odpowiedzi, nie mówiąc o błędach wynikających z całkowitej niedostępności – w postaci wypadania z próby – co potwierdzają wyniki analiz prowadzonych

⁶ Ścisłe mówiąc, wartości współczynników regresji informują o wielkości różnicy między ocenami dokonywanymi przez respondentów badanych w danym przedziale czasu w stosunku do średniej w próbie. Wartość dla kategorii referencyjnej (ostatni przedział czasowy) ustalono w innym modelu, w którym pominiętą kategorią referencyjną był pierwszy przedział czasu.

w krajach zachodnich. Próby rozpoznania wyznaczników niedostępności można zatem traktować jako krok w kierunku doskonalenia narzędzi pomiaru.

Wskaźnikiem niedostępności respondentów była liczba wizyt ponawianych przez ankietera w celu przeprowadzenia wywiadu. Okazuje się, że z rozpatrywanych tu cech respondentów, czynnikami utrudniającymi dostępność były wyższy poziom wykształcenia, mieszkanie w większej miejscowości i młody wiek, natomiast stosunkowo łatwo dociera się do osób bezrobotnych, mieszkańców wsi i niższych rangą pracowników umysłowych. To ostatnie można interpretować jako efekt lepszego kontaktu pracowników umysłowych z ankieterem i większych kompetencji w zakresie rozumienia pytań – w porównaniu z kategoriami o niższym statusie – a równocześnie mniejszych ograniczeń czasowych, w porównaniu z respondentami w zawodach inteligenckich.

Drugim celem analizy było stwierdzenie, czy data realizacji wywiadu wywiera niezależny wpływ na odpowiedzi badanych. Można założyć, że odpowiedzi uzyskiwane w ciągu trzech miesięcy są mniej wiarygodnym materiałem niż odpowiedzi na to samo pytanie uzyskiwane w ciągu trzech dni. Wydłużenie badania w czasie zwiększa niebezpieczeństwo nieporównywalności wyników. Okazuje się, że – w istocie rzeczy – zależność taka istnieje. Świadczy o tym pozytywny wynik testu dla dwóch zmiennych skonstruowanych na podstawie odpowiedzi badanych.

Literatura

- Box-Steffensmeier, Janet M. i Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.
- Cox, David i D. Oakes. 1984. *Analysis of Survival Data*. London: Chapman and Hall.
- Domański, Henryk. 2002. *Identyfikacja błędów pomiaru w modelach zależności strukturalnych*. „Ask. Społeczeństwo, Badania, Metody” 11: 27–44.
- Domański, Henryk i Dariusz Przybysz. 2003. *Analiza przydatności EGP jako wskaźnika pozycji społecznej*. „Ask. Społeczeństwo, Badania, Metody” 12: 85–116.
- Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley & Sons, Inc.
- Groves, Robert M. i Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Smeets, Ingrid. 1995. *Facing another Gap: An Exploration of the Discrepancies between Voting Turnout in Survey Research and Official Statistics*. „Acta Politica” 30: 307–334.
- Smith, Tom W. 1983. *The Hidden 25 Percent: An Analysis of Nonresponse of the 1980 General Social Survey*. „Public Opinion Quarterly” Vol. 47: 386–404.

- Stoop, Ineke. 2005. *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. Hague: Social and Cultural Planning Office,
- Sztabiński, Paweł. 2004. *Metodologia badania Europejski Sondaż Społeczny*. „Ask. Społeczeństwo. Badania. Metody” 13: 27-37.
- Ward, James C., Bertram Russick i William Rudelius. 1985. *A Test of Reducing Callbacks and Not-At-Home Bias in Personal Interviews by Weighting At-Home Respondents*. „Journal of Marketing Research” Vol. XXII (February 1985): 66-73.
- Yamaguchi, Kazuo. 1992. *Accelerated Failure-Time Regression Models with a Regression Model of Surviving Fraction: An Application to the Analysis of ‘Permanent Employment’ in Japan*. „Journal of the American Statistical Association” Vol. 87: 284-292.
- Yamaguchi, Kazuo. 1991. *Event History Analysis*. Sage Publications.

NUMBER OF VISITS AND DURATION OF SURVEYS

Using data from European Social Survey 2002 I attempt to determine two questions. First, what are basic determinants of availability of respondents as defined in terms of number of visits to make interview. Second, to what extent responses are affected by date of interview in time trajectory of the research. One may assume that significant relationships between date of interview and responses would indicate systematic error resulting from effect of the context of interview related to time. My analysis is based on the Polish data. According to main findings, basic impediments of availability of respondents included: higher education, residence in bigger city, and younger category of age. Conversely, relatively most accesible are: unemployed, countryside dwellers, and lower non-manual workers. This analysis show also that date of interview exerts significant effect on responses.

Key words: duration of interview, number of visits, availability of respondents, systematic error, non-responses.

Aneks

Tabela A1. Zależności między dostępnością respondenta do pierwszego kontaktu a wybranymi cechami respondenta. Współczynniki regresji w modelu proporcjonalnego ryzyka^a

Zmienne niezależne	Wywiad zrealizowany w czasie pierwszej wizyty	Wywiad zrealizowany dla wizyt 3-7
Kategorie społeczno-zawodowe EGP		
Inteligencja	1,04	0,75
Pracownicy administracyjni średniego szczebla	1,08	1,46
Właściciele	0,88	1,95
Robotnicy	0,95	0,79
Rolnicy	Kategoria pominięta	
Wiek		
15-25	1,02	1,03
26-40	0,99	1,98
41-55	0,97	1,52
56-65	1,02	0,81
66 i powyżej	Kategoria pominięta	
Bezrobocie	0,98	0,35
Liczba lat nauki	0,97	0,97
Płeć (1=mężczyźni)	1,02	1,48
Ofiara napadu	1,11	1,38
Miejsce zamieszkania		
Duże miasto	1,01	1,73
Przedmieście	0,93	1,26
Małe miasto	1,11	0,86
Wieś	Kategoria pominięta	

^a Ze względu na brak istotnych zależności nie przeprowadziłem tu dodatkowych estymacji parametrów dla „pominiętych” kategorii w przypadku zawodu, wieku i miejsca zamieszkania.

