

WOJCIECH HARDY\*

## **Długość kwestionariusza a precyzja odpowiedzi na pytania ilościowe w badaniach ankietowych<sup>1</sup>**

### **Wstęp**

Nierówności społeczne, często mierzone poprzez rozkład dochodów w społeczeństwie, to temat poruszany w wielu pracach, zarówno naukowych, jak i dziennikarskich. Teoretycy ekonomii wielokrotnie przedstawiali ten problem jako zjawisko naturalne i pożądane lub odwrotnie – negatywne w skutkach i charakterze. Niezależnie jednak od punktu widzenia, nierówności dochodowe można zaobserwować jedynie poprzez konstruowane specjalnie w tym celu miary. Większość z nich oparta jest na danych z reprezentatywnych badań ankietowych, które stanowią jedno z podstawowych źródeł wiedzy o procesach społecznych i ekonomicznych. Wyzwaniem metodologicznym w badaniach ankietowych jest zarówno procedura zbierania danych, jak i zapewnienie dobrej jakości zebranej informacji. Obydwa te czynniki mogą zależeć m.in. od struktury kwestionariusza.

W przypadku systematycznych, a nieobserwowalnych zaburzeń w jakości danych, pomiary nierówności dochodów w społeczeństwie mogą odzwierciedlać problemy zawarte w samym zbiorze danych, nie zaś rzeczywiste nierówności społeczne. Ponadto wyniki przeprowadzanych procedur statystycznych mogą podlegać obciążeniom – badania oparte na danych o wariancji zawyżonej lub zaniżonej względem wariancji danych doskonale precyzyjnych dostarczą błędne wnioski. Mimo że oszacowania parametrów w modelach mogą pozostać prawidłowe, zniekształcone zostaną oszacowania błędów losowych, co podważa wiarygodność statystyk świadczących m.in. o istotności wykorzystanych zmiennych.

W tym miejscu chcemy zbadać wpływ wydłużenia ankiety na precyzję otrzymywanych w niej odpowiedzi. W tym celu wykorzystano dane BAEL o dochodach gospodarstw domowych w Polsce w latach 2005–2006. Struktura kwestionariusza zmieniła się wówczas istotnie z powodu rozszerzenia go o ponad 20 dodatkowych pytań. Wskutek tego pytanie o średnie miesięczne wynagrodzenie respondenta przesunęło się w ankiecie o około 17 pozycji dalej<sup>2</sup>. Ten naturalny *quasi*-eksperyment można wykorzystać, aby przeanalizo-

---

\* Mgr Wojciech Hardy – Wydział Nauk Ekonomicznych Uniwersytetu Warszawskiego i Instytut Badań Strukturalnych w Warszawie; e-mail: hardywojciech@gmail.com

<sup>1</sup> Dziękuję: dr hab. Joannie Tyrowicz za pomoc w przeprowadzeniu niniejszego badania, uczestnikom konferencji „Wzrost gospodarczy – Rynek pracy – Innowacyjność gospodarki” w Łodzi za konstruktywne komentarze, dr. Leszkowi Morawskiemu za uwagi do wcześniejszej wersji artykułu.

<sup>2</sup> Faktyczne przesunięcie mogło wahać się w zależności od udzielanych wcześniej odpowiedzi.

wać wpływ długości ankiet BAEL na wariancję odpowiedzi na pytanie o dochód. Testy statystyczne wskazują na istotne pogorszenie precyzji udzielanych odpowiedzi. W ramach kontroli odporności na inne czynniki testy przeprowadzono także dla innych, wybranych z kwestionariusza pytań. Pierwotne wnioski zostały potwierdzone.

W pierwszej części artykułu przedstawiono ogólną problematykę badań ankietowych, uwzględniając przegląd dotychczasowej literatury na ten temat. W drugiej części przedstawiono przeprowadzone badanie, wykorzystując dane BAEL. Wyniki zostały omówione w części trzeciej. Tekst kończy się interpretacją uzyskanych wyników i podsumowaniem.

## 1. Przegląd literatury

Literaturę na temat konstruowania ankiet można podzielić na trzy główne grupy. Pierwsza z nich podejmuje temat z punktu widzenia psychologii respondenta i jego motywacji podczas wypełniania kwestionariusza. Wiedza ta stanowi podstawę do pełnego zrozumienia związku między niektórymi elementami kwestionariusza a tym, jak z ankietą poradzi sobie respondent. Druga grupa dotyczy metod zbierania danych i samego skłaniania ankietowanego do wzięcia udziału w badaniu i odpowiedzenia na wszystkie zawarte w nim pytania. Trzecia grupa porusza problematykę jakości odpowiedzi, czyli prawdziwości i precyzji informacji otrzymanych od respondenta.

Poniżej omówiono wybrane badania na ten temat według zaproponowanej wyżej klasyfikacji.

### 1.1. Psychologia udzielania odpowiedzi

Bardzo istotną rolę w kształtowaniu się jakości odpowiedzi, jak wskazują Krosnick i Presser (2009, s. 265–266), odgrywa struktura procesu myślowego odpowiadającego za ich udzielanie. Proces ten składa się z czterech złożonych etapów, z których każdy może wymagać istotnego wkładu myślowego. Te etapy to:

- 1) interpretacja pytania i dedukcja jego właściwego znaczenia;
- 2) przeszukanie pamięci w celu odnalezienia odpowiedniej informacji;
- 3) zintegrowanie odnalezionej informacji w pojedynczy osąd;
- 4) przetłumaczenie osądu na odpowiedź według podanych w ankiecie możliwości.

Według licznych prac psychologicznych wkład potrzebny do wykonania wszystkich części powyższego procesu może być motywowany m.in. czynnikami takimi jak: możliwość autoekspresji, wyzwanie intelektualne czy po prostu altruizm. Jeśli respondent podejmie trud przejścia przez każdy z wyżej wymienionych etapów jego postawę możemy nazwać postawą „optymalizującą” (*optimizing*). Oczywistym celem każdego badacza jest zmaksymalizowanie dokładnie takiego zachowania wśród respondentów. Jest to jednak utrudnione przez wiele zjawisk, od wypełniania ankiety z przymusu po znużenie czy dekoncentrację wiele samego procesu.

Na drugim końcu skali zachowań respondentów kładzie się postawę satysfakcjonującą/wystarczającą (*satisficing*). Zjawisko to jest związane z szeregiem niedoskonałości podczas wspomnianego wyżej procesu myślowego związanego z odpowiadaniem na pytania. Respondenci mogą mieć problem ze zrozumieniem znaczenia pytania lub poświęcić mu zbyt mało uwagi. Przeszukanie pamięci może odbyć się mało dokładnie, a integracja odnalezionych informacji zostać przeprowadzona nieuważnie. Wreszcie, dobór odpowiedzi może ucierpieć na precyzji. Powodem tego nie musi być zła wola, lecz nawet znużenie czy

dekoncentracja. Niezależnie jednak od powodu, proces decyzyjny zbiega się do procesu wystarczającego z punktu widzenia respondenta, lecz dającego odpowiedzi rozmiągające się z prawdą. W skrajnych przypadkach ankietowani mogą pominąć część etapów i od razu skupić się na wypełnianiu kwestionariusza według domniemanych oczekiwań badacza, tym samym także zaburzając wyniki badań.

Krosnick (1991, s. 213–236) podaje trzy główne czynniki odpowiedzialne za zachowanie respondenta:

- 1) trudność zadania – związana ze skomplikowaniem pytania i procesu pozyskania informacji oraz czynników zewnętrznych (np. prędkość zadawania pytań lub rozpraszające okoliczności);
- 2) zdolność respondenta – związana z ogólnymi możliwościami umysłowymi, jego wiedzą o temacie i wcześniejszymi przemyśleniami o badanej kwestii;
- 3) motywacja respondenta – związana z potrzebą pracy umysłowej, indywidualnym odbiorem tematyki pytań, wiarą w istotność badania, zmęczeniem oraz sposobem aplikacji kwestionariusza.

Należy zauważyć, że w przypadku pytania o wynagrodzenie respondenta może występować wysokie ryzyko zachowania typu *satisficing*. Po pierwsze, respondent może mieć problemy w przypomnieniu sobie dokładnej wartości miesięcznego dochodu. Po drugie, zadanie to może być dodatkowo utrudnione, jeśli wypełniany przez respondenta kwestionariusz jest zbyt długi. Doprowadzić to może do zmęczenia i dekoncentracji ankietowanego, który tym samym straci zdolność i motywację do znalezienia precyzyjnej odpowiedzi. Jeśli przyjąć istnienie jednego rozkładu dochodu w społeczeństwie polskim, konstrukcja ankiety sprzyjająca zjawisku *satisficing* może doprowadzić do zebrania odpowiedzi znajdujących się jedynie „w pobliżu” odpowiedzi prawdziwych. Tym samym badana zmienna wykaże sztucznie zawyżoną wariancję. Zachodzenie tego zjawiska jest przedmiotem niniejszego artykułu.

## 1.2. Problemy z pozyskiwaniem odpowiedzi

Istnieje wiele badań dotyczących problemu niskiej frekwencji ankietowanych, czyli rzadkiej zgody na wypełnienie kwestionariusza. Część owych badań dotyka jeszcze ankiet zwrotnych przesyłanych drogą pocztową. Wielu autorów (Dillman, Sinclair, Clark 1993; Heberlein, Baumgartner 1978) rozważało ten problem w swoich pracach, dochodząc nie raz do przeciwstawnych wniosków na temat stopnia wpływu długości ankiety na stopę odpowiedzi (*response rate*). Dillman, Sinclair i Clark (1993) wskazywali na istotny, negatywny wpływ długości kwestionariusza, lecz zauważyli także znaczącą rolę „przyjazności” formy kwestionariusza oraz obecności trudnych lub kontrowersyjnych pytań, które w sposób oczywisty zniechęcają do uczestnictwa w badaniu.

Kolejnym problemem związanym z przeprowadzaniem ankiet jest wysoki odsetek rezygnacji w trakcie wypełniania ankiety, który podobnie jak niski udział w badaniu może zaburzyć reprezentatywność próby i w sposób znaczny wpłynąć na jej wielkość. Problem ten pojawia się, gdy dany respondent w trakcie wypełniania ankiety utraci motywację do kontynuacji odpowiedzi. W szczególności może to wynikać z długiego czasu trwania badania. W eksperymencie przeprowadzonym przez Galesica i Bosnjaka (2009) zależność ta została potwierdzona – dłuższy czas trwania ankiety zwiększa ryzyko rezygnacji respondenta z dalszego udziału w badaniu.

Pokrewny dwóm powyższym problemom, i istotniejszy w kontekście pytań o wynagrodzenie, jest problem porzucania niektórych pytań, tzw. stopnia niewypełnialności.

Problem ten jest poruszony w badaniu spółki Lightspeed Research (2008). Ich wyniki wyraźnie wskazują na wzrost niewypełnialności wraz z trwaniem badania, choć możliwe jest przeciwdziałanie temu zjawisku poprzez dodatkową motywację (np. pieniężną) w trakcie badania.

Wykorzystywanie dodatkowej zachęty pieniężnej wiązałoby się jednak ze znacznie większymi kosztami badań ankietowych. Ważne jest więc skonstruowanie kwestionariuszy tak, aby minimalizować to niekorzystne zjawisko. Respondenci, w szczególności, omijają pytania, które uważają za zbyt prywatne lub kontrowersyjne. Pytanie o wynagrodzenie zalicza się do grupy pytań o szczególnie wysokim stopniu niewypełnialności, co sprawia, że wymaga ono szczególnej uwagi (np. Riphahn, Serfling 2005).

### 1.3. Problemy z jakością odpowiedzi

Istnieje wiele ujęć problemu jakości odpowiedzi. Herzog i Bachman (1981) wyróżniają problemy takie jak odpowiadanie w prostej linii<sup>3</sup> na pytania zamknięte. Subar i in. (2001) pokazują, że jakość odpowiedzi może być silniej związana z łatwością interpretacji ankiety niż z samą jej długością<sup>4</sup>. Autorzy definiują jednak jakość jako odsetek wspomnianego już stopnia niewypełnialności (*incompletion rate*) z włączeniem odpowiedzi błędnych (niemożliwych do zinterpretowania).

Wiele efektów zwiększenia długości ankiety zbadali Galesic i Bosnjak (2009). Skonstruowali oni kilka miar jakości otrzymywanych danych: stopień niewypełnialności, długość odpowiedzi udzielanych na pytania otwarte, czas odpowiadania na pytania oraz wariancję odpowiedzi na pytania z siatkami odpowiedzi. Według tych badaczy długość udzielanych odpowiedzi malała wraz z czasem trwania ankiety – respondenci udzielali więc odpowiedzi mniej wyczerpujących. Podobna zależność zaszła w przypadku czasu poświęcanego na pojedyncze pytanie – respondenci krócej więc zastanawiali się nad pojedynczą odpowiedzią, co mogło wpłynąć na precyzję udzielanych odpowiedzi. Czas trwania ankiety miał ujemny wpływ na wariancję odpowiedzi udzielanych w pytaniach z siatką odpowiedzi.

## 2. Badanie

### 2.1. Przedmiot badania

Celem tego badania jest analiza wpływu długości ankiety na precyzję otrzymywanych w niej odpowiedzi. Na istnienie takiej zależności wyraźnie wskazuje przytoczona powyżej literatura. Aby przetestować hipotezę o negatywnym wpływie czasu trwania ankiety na jakość zbieranych w niej danych, wykorzystane zostaną dane ankietowe Badania Aktywności Ekonomicznej Ludności<sup>5</sup> dla lat 2005/2006, kiedy to nastąpiła znacząca zmiana

<sup>3</sup> Herzog i Bachman przyglądają się zachowaniu polegającemu na zaznaczaniu „tych samych” odpowiedzi w wielu pytaniach z rzędu (np: środkowa albo C). Zwracają także uwagę na tendencję do zaznaczania odpowiedzi, które wydają się stereotypowe i schematyczne.

<sup>4</sup> W swoim badaniu Subar i in. porównują wyniki dwóch ankiet różniących się między sobą długością i skomplikowaniem (ankieta dłuższa jest równocześnie prostsza do interpretacji).

<sup>5</sup> Badanie Aktywności Ekonomicznej Ludności (BAEL) prowadzone jest od roku 1992, co kwartał, przez Główny Urząd Statystyczny, zgodnie z metodyką Międzynarodowej Organizacji Pracy. BAEL ma na celu analizę charakterystyki (za stroną internetową GUS): „ludności w wieku 15 lat i więcej aktywnej ekonomicznie, bezrobotnej i biernej zawodowo“ według cech demograficznych i społecznych. W celu zebrania

w formie wykorzystywanych kwestionariuszy. Dotychczasowy zestaw pytań został uzupełniony o około 20 dodatkowych pytań, co znacznie rozszerzało kwestionariusz, przesuwając wiele ze znajdujących się w nim pytań na dalsze pozycje. Wśród przesuniętych w ten sposób pytań znajdowało się pytanie o średnie miesięczne wynagrodzenie respondenta. Numery wybranych pytań dla 2005 r. i 2006 r. podano w tabeli 1.

Zgodnie z dotychczasową literaturą, zmiana ta mogła mieć negatywny wpływ na otrzymywane za pomocą ankiety dane. W szczególności można tu wyróżnić dwa potencjalne efekty:

- 1) znużenie (dekoncentracja) respondenta obniżające precyzję odpowiedzi, w tym informacji na temat średniego miesięcznego dochodu (wyższa wariancja);
- 2) demotywacja i zniecierpliwienie respondenta, wywołująca upraszczanie odpowiedzi do odpowiedzi „wystarczających” z punktu widzenia respondenta (np. częstsze zaokrąglenia odpowiedzi).

**Tabela 1**  
**Pytania ankiety BAEL wykorzystane w badaniu**

Rok	Pozycja (numer*)	Treść
2005	20	Proszę podać wysokość swoich zarobków netto w poprzednim miesiącu w głównym miejscu pracy.
2006	37	
2005	29	Ile godzin łącznie (w pracy głównej i dodatkowej) chciał(a)by Pan(i) pracować w tygodniu?
2006	48	
2005	34	Ile godzin pracował(a) Pan(i) w badanym tygodniu w pracy dodatkowej? (rzeczywista liczba godzin)
2006	54	
2005	51a	Przy jakim minimalnym miesięcznym wynagrodzeniu brutto był(a)by Pan(i) gotów (gotowa) podjąć pracę?
2006	72	
2005	72	Ile godzin łącznie poświęcił(a) Pani na doksztalcanie w ciągu ostatnich czterech tygodni?
2006	98	

\* Dokładna pozycja pytań w obu latach mogła różnić się w zależności od wcześniejszych odpowiedzi indywidualnego respondenta.

Nasze badanie weryfikuje hipotezę o zachodzeniu powyższych dwóch efektów. Pierwszy z nich zmierzony zostanie poprzez zmianę w wariancji deklarowanych średnich dochodów społeczeństwa w latach 2005 i 2006. Jeśli pierwszy efekt ma miejsce, to w 2006 r. zauważalny będzie wyraźny wzrost wariancji w deklarowanych dochodach. Drugi z efektów zostanie przeanalizowany za pomocą porównania odsetka zaokrągleń w odpowiedziach na pytanie o miesięczne wynagrodzenie w latach 2005 i 2006. Jeśli efekt ma miejsce, to w 2006 r. zauważalny będzie wzrost częstotliwości zaokrąglenia przez respondentów podawanych liczb do liczb „okrągłych” (wielokrotności 50, 100 lub 500).

wszystkich danych respondentom przedstawiane są kwestionariusze złożone z pytań zamkniętych oraz mniej licznych otwartych. Wypełniane są one wraz z zadającym pytania ankieterem. Jednostką badaną jest pojedynczy respondent.

## 2.2. Metoda badawcza

Aby określić statystyczny wpływ zmiany kwestionariusza, zostaną wykorzystane trzy testy na równość wariancji wynagrodzenia w latach 2005 i 2006.

Pierwszym jest test *F* Fishera. Test ten sprawdza hipotezę zerową o równości wariancji w dwóch populacjach przeciwko hipotezom alternatywnym o większej wartości którejś z wariancji. Test *F* opiera się na założeniu normalności badanego rozkładu, tym samym jego wynik może być obciążony, jeśli rozkład dochodu nie spełnia tego warunku<sup>6</sup>.

Drugim jest test *U* Manna-Whitneya-Wilcoxon, znany też jako test rang. Test ten sprawdza hipotezę zerową o symetrii prawdopodobieństwa przyjmowania większej wartości przez zmienną w jednej z porównywanych grup. Innymi słowy, test rang bada, czy deklarowane wynagrodzenie w 2005 r. jest większe od wynagrodzenia w 2006 r. z takim samym prawdopodobieństwem, z jakim wynagrodzenie w 2006 r. jest większe od tego w 2005 r. W uproszczeniu hipoteza zerowa oznacza równość rozkładów w dwóch grupach. Test ten nie ma żadnych założeń odnośnie do rozkładu badanych zmiennych, tym samym brak rozkładu normalnego nie jest dla niego problemem. Jego wadą jest stosunkowo niska moc statystyczna, przez co hipoteza zerowa może nie zostać odrzucona w przypadku, w którym średnie są sobie równe, choć to samo nie może być powiedziane o wariancjach.

Ostatnim testem jest test Levene'a na równość wariancji. Razem z nim policzone zostaną jego dwie alternatywne statystyki – testy Browna-Forsythe'a zastępujące wartość średniej odpowiednio medianą lub 10% średnią obciążoną.

Wyniki wszystkich testów są przedstawione w następnej części artykułu. W celu analizy odporności wyników na nieidentyfikowalne wpływy testy te zostały przeprowadzone także dla czterech innych pytań przytoczonych w tabeli 1. Ponadto dokonane zostanie omówienie i objaśnienie otrzymanych wyników oraz oceniona zostanie ich odporność na inne problemy.

W celu oszacowania zmiany w częstotliwości zaokrągleń deklarowanego dochodu wykonany zostanie prosty test *t*-Studenta, porównujący między latami 2005 i 2006 średnie trzech zmiennych zerojedynkowych oznaczających wynagrodzenia będące wielokrotnością liczb 50, 100 i 500.

## 3. Wyniki

### 3.1. Prezentacja wyników

Tabela 2 zawiera wyniki testów statystycznych dla równości wariancji deklarowanego dochodu w dwóch próbach, czyli w latach 2005 i 2006. Wszystkie testy wskazują na istotną zmianę w odchyleniach i w rozkładzie badanej zmiennej. Warto zauważyć, że każdy z wyników otrzymany jest na poziomie 1%, co dodatkowo świadczy o ich wysokiej istotności. Tym samym potwierdzona zostaje teza stawiana w tym badaniu – w 2006 r. nastąpił znaczny wzrost w rozrzucie deklarowanych wynagrodzeń.

Wynik ten może być podyktowany czymś więcej niż zmiana długości kwestionariusza. W szczególności możliwe jest, że w społeczeństwie polskim w ciągu roku wzrósł poziom nierówności społecznych, owocując zaobserwowaną zmianą. Aby wykluczyć wpływ czynników nieobserwowalnych, analogiczna procedura testowa została przeprowadzona dla czterech innych pytań z ankiet BAEL, przytoczonych już wcześniej.

<sup>6</sup> Test na skośność i kurtozę rozkładu dochodu potwierdza istotną różnicę od rozkładu normalnego.

**Tabela 2**  
**Zestawienie wniosków z testów dla pytań wykorzystanych w badaniu**

Pytanie	Dochód miesięczny		Ile chciałby pracować?		Czas pracy dodatkowej		Minimalny zarobek		Czas doksztalcania	
	2005	2006	2005	2006	2005	2006	2005	2006	2005	2006
Odchylenie	~673,17	~691,76	10,23	9,9	8,72	8,7	398,7	423,31	16,74	21,08
Test <i>F</i>	wzrost odchylenia*		spadek odchylenia*		brak zmian		wzrost odchylenia*		wzrost odchylenia*	
Test Levene'a	zmiana odchylenia*		zmiana odchylenia*		brak zmian		zmiana odchylenia*		zmiana odchylenia*	
Test Browna-Forsythe'a W50	zmiana odchylenia*		zmiana odchylenia*		brak zmian		zmiana odchylenia**		zmiana odchylenia*	
Test Browna-Forsythe'a W10	zmiana odchylenia*		zmiana odchylenia*		brak zmian		zmiana odchylenia*		zmiana odchylenia*	
Test MW	zmiana rozkładu*		zmiana rozkładu*		brak zmian		zmiana rozkładu*		brak zmian	

\* Na poziomie istotności 1%.

\*\* Na poziomie istotności 5%.

Rezultaty dodatkowych testów zawarte są w tabeli 2. Otrzymane wyniki wskazują analogiczną zależność w przypadku dwóch pozostałych pytań, brak zależności w przypadku jednego z nich i odwrotną zależność w przypadku ostatniego.

**Tabela 3**  
**Wyniki testów *t*-Studenta na równość częstotliwości zaokrągleń**

Częstotliwość zaokrągleń do:	50	100	500
2005	0,4085	0,3567	0,1119
2006	0,4262	0,3743	0,1179
H1: $\bar{sr}(2005) < \bar{sr}(2006)$	0,0002	0,0001	0,0307
H2: $\bar{sr}(2005) \neq \bar{sr}(2006)$	0,0004	0,0003	0,0614
H3: $\bar{sr}(2005) > \bar{sr}(2006)$	0,9998	0,9999	0,9693

Tabela 3 przedstawia wyniki testu na równość częstotliwości zaokrągleń w obydwu latach wysokości podawanego wynagrodzenia do wielokrotności 50, 100 i 500. Niezależnie od przyjętego stopnia zaokrąglenia, testy jednoznacznie wskazują na istotny wzrost częstotliwości zaokrągleń deklarowanego dochodu.

### 3.2. Omówienie wyników

Rezultaty przeprowadzonego badania wskazują na wzrost wariancji deklarowanych wynagrodzeń w 2006 r., w którym nastąpiło rozszerzenie kwestionariusza BAEL o około 20 nowych pytań. Wynik ten wydaje się potwierdzać tezę o zmniejszonej precyzji udzielanych przez respondentów odpowiedzi.

Głównym zastrzeżeniem dla otrzymanych wyników są zmiany w rozrzucie dochodów na przestrzeni lat. Rzeczywiście, w debacie publicznej często wspomina się o wzrastających nierównościach w zakresie wynagrodzeń. Jeśli więc w 2006 r. nastąpiła znacząca zmiana w zróżnicowaniu dochodów, można się spodziewać, że owa zmiana sytuacji mogła znacząco wpłynąć na wyniki testów, a tym samym uzyskane wyniki mogą jedynie wskazywać na wzrost nierówności społecznych w 2006 r., w stosunku do 2005 r.

Pewną informację na temat wielkości tego problemu mogą dostarczyć miary nierówności dochodów zastosowane dla społeczeństwa polskiego, w szczególności współczynnik Giniego. W tabeli 4 przedstawione są wartości współczynnika Giniego dla Polski w latach 2005–2006, według danych UNU-WIDER<sup>7</sup>, EU-SILC<sup>8</sup>, Banku Światowego oraz na podstawie danych BAEL<sup>9</sup>.

**Tabela 4**  
**Wartości współczynnika Giniego dla Polski w latach 2005–2006**

Rok	Współczynnik Giniego	Źródło
2005	36,6	UNU-WIDER
2006	34,0	UNU-WIDER
2005	35,6	EU-SILC
2006	33,3	EU-SILC
2005	34,92	Bank Światowy
2006	34,13	Bank Światowy
2005	26,1	BAEL
2006	27,0	BAEL

Jak widać, współczynnik Giniego wskazuje na zmianę nierówności w dochodach w 2006 r. Co jednak ciekawe i nieco zaskakujące na tle trendu z innych lat, według większości źródeł wskazuje on na spadek nierówności. Jedynie dane BAEL wskazują na wzrost współczynnika Giniego. Jeżeli przyjąć, że stopień nierówności dochodowych w 2006 r. obniżył się, to fakt ten potwierdza interpretowanie wyników przeprowadzonych testów jako świadectwa negatywnego wpływu zmiany konstrukcji kwestionariuszy BAEL na precyzję uzyskiwanych za ich pomocą odpowiedzi.

Sprzeczne z hipotezą postawioną w badaniu pozostają wyniki testów dla pytania o wymarzony czas pracy tygodniowo. Zgodnie wskazują one na zmianę w rozkładzie zmiennej poprzez spadek jej odchylenia standardowego. Jest to więc efekt dokładnie przeciwny do szukanego. By lepiej przyjrzeć się temu zjawisku należy przyjrzeć się formie kwestionariuszy. Potencjalnym wytłumaczeniem tego wyniku jest fakt, że pytanie to pojawia się w ankiecie po pytaniu o stosunkowo podobnym charakterze – o liczbę godzin zwykle pracowanych przez respondenta w pojedynczym tygodniu. Obydwie wartości prawdopodobnie są ze sobą powiązane, w związku z czym znalezienie odpowiedzi na pierwsze z tych pytań może wywierać wpływ na odpowiedź udzielaną na drugie – pojawiające się trochę później. Zgodnie z etapami procesu myślowego udzielania odpowiedzi przedstawionymi

<sup>7</sup> World Institute for Development Economics Research of the United Nations University.

<sup>8</sup> The European Union Statistics on Income and Living Conditions.

<sup>9</sup> Wartości wyliczone na podstawie danych BAEL zostały zaczerpnięte z raportu *Badanie struktury i zmian rozkładu wynagrodzeń w Polsce w latach 2000–2006* (Marcinkowska i in. 2008).



przez Krosnicka (2009) respondent dokona więc przeszukania pamięci w poszukiwaniu podobnej informacji przed analizowanym pytaniem. Tym samym bardziej znużony respondent, zapytany o wymarzony tygodniowy czas pracy, może ulec zachowaniu *satisficing* i zamiast wyszukać nową informację, udzielić podobnej, znalezionej już przedtem – na temat liczby godzin zwykle przepracowywanych tygodniowo.

Jeśli przedstawione rozumowanie jest poprawne, udzielona odpowiedź powinna zbiegać się ze wcześniejszą w tym większym stopniu, im dłuższy jest kwestionariusz. Zależność tę można zweryfikować, ponieważ pozycja pytania o pożądaną liczbę godzin pracy zmienia się w 2006 r. o 19 miejsc, podczas gdy pytanie o liczbę zwykle przepracowanych tygodniowo godzin nie tylko nie przesuwają się w przód, ale wręcz cofa o jedną pozycję bliżej początku kwestionariusza. Tym samym zmiana konstrukcji kwestionariusza nie powinna istotnie wpłynąć na wyniki pierwszego z pytań, podczas gdy może wpłynąć na zachowanie respondenta podczas odpowiedzi na drugie. W 2006 r. średnie odpowiedzi z obu pytań powinny więc być sobie bliższe niż w 2005 r. z uwagi na częstsze odnoszenie się do wcześniej wyszukanej informacji.

**Tabela 5**  
**Wyniki testu *t* na równość średnich w dwóch próbach**

	2005	2006
Obserwacje	19795	19682
Średnia	4,11	3,59
Odchylenie standardowe	7,86	7,49
H1: diff<0	Pr( $T < t$ )	= 1,0000
H2: diff≠0	Pr( $ T  >  t $ )	= 0,0000
H3: diff>0	Pr( $T > t$ )	= 0,0000

Istnienie tego efektu sprawdzamy, porównując średnie absolutne różnice obu deklarowanych przez respondentów wartości w latach 2005 i 2006. Jeśli omawiany efekt zachodzi, test na równość średnich powinien wykazać spadek średniej w 2006 r., czyli zjawisko zbiegania odpowiedzi na pytanie o pożądaną liczbę godzin pracy do wcześniej deklarowanych wartości. Wyniki testu *t*-Studenta na równość średnich przedstawione w tabeli 5 wskazują na spadek wartości różnic w 2006 r., w stosunku do roku poprzedniego, tym samym wskazując na prawdopodobne zachowanie *satisficing* wśród respondentów. Wynik ten może częściowo tłumaczyć wyniki testów na równość odchyłeń dokonanych wcześniej. Trudno jednoznacznie ustalić, czy jest to przyczyna nieuznania hipotezy alternatywnej podczas testowania odpowiedzi na pytanie o wymarzoną liczbę godzin pracy. Jest to jednak potencjalne wytłumaczenie, a wskazana różnica w kwestionariuszu dotyczy jedynie tej zmiennej.

Dodatkowych wniosków dostarczyła analiza częstotliwości zaokrągleń dla obu badanych lat. Wskazuje ona przypuszczalnie na jeszcze jeden efekt zachodzący po podniesieniu liczby pytań kwestionariusza BAEL. W 2006 r. większa część respondentów deklarowała „okrągłe” liczby w odpowiedzi na pytanie o miesięczne wynagrodzenie. Wynik ten jest istotny zarówno dla wielokrotności 50, jak i 100, czy nawet 500. Świadczy to o dodatkowym obniżeniu precyzji pozyskiwanych danych.

#### 4. Interpretacja i wnioski

W artykule został przetestowany wpływ wydłużenia kwestionariusza BAEL na precyzję odpowiedzi otrzymywanych na pięć z pytań zamieszczonych w ankiecie. Uwzględniono pytania o dochód i inne wartości wyrażane liczbowo. W wyniku zmiany formy kwestionariusza w 2006 r. badane pytania przesunęły się o prawie 20 pozycji dalej. Właściwie testowano więc wpływ pozycji pytania (czyli długości ankiety do danego momentu) na odchylenie standardowe udzielonych odpowiedzi. Zgodnie z wynikami testów trzy z pięciu pytań (w tym pytania o dochód) cechowały się mniej precyzyjnymi odpowiedziami w 2006 r. w porównaniu z 2005 r.; przy jednym pytaniu nie zauważono różnicy; przy jednym zaś wykryto zjawisko odwrotne.

Analiza wartości współczynników Giniego dla lat 2005 i 2006, na podstawie innych zbiorów danych, pozwalają ocenić odporność wyników na inne nieobserwowane zjawiska (jak możliwość rzeczywistego wzrostu nierówności dochodów). Ponadto dla pytania wskazującego na wpływ odmienny od zakładanego zaobserwowano działający oddzielny efekt typu *satisficing*, mogący tłumaczyć odwrotne wyniki testów. Pytanie o charakterze „neutralnym” wydaje się także prostsze od pozostałych, a tym samym w mniejszym stopniu podatne na zmęczenie respondentów. Wyniki, poparte dodatkową analizą zmiennych, potwierdzają hipotezę stawianą w tym badaniu.

Zjawisko umniejszonej precyzji odpowiedzi zostało zbadane dla pytań otwartych o charakterze liczbowym. Nie powinno jednak postrzegać się go tylko jako problemu tego typu zagadnień. Zawężenie to wynikało z potrzeby ograniczenia innych zjawisk oddziałujących na wariancję odpowiedzi, która w tym badaniu stanowi główny wskaźnik badanego problemu. Głównym z tych efektów jest skłonność do udzielania odpowiedzi łatwych, szybkich, neutralnych, stereotypowych czy schematycznych, która obniża różnorodność udzielanych odpowiedzi, a więc i wariancję do stopnia, w którym zjawisko obniżonej precyzji nie daje się zaobserwować (zjawisko to znajduje pewne odzwierciedlenie w zaokrągleniach odpowiedzi liczbowych). Prawdopodobnym wzorem działania zniecierpliwionych respondentów może być także wybór odpowiedzi na chybił trafił lub też wybór jedynie przybliżeń prawdziwych odpowiedzi.

W przypadku wielu typów rozkładów zmiennych okazać się może, że brak precyzji odpowiedzi oraz większy ich rozrzut nie wpływają na wartość średniej. Tym samym wnioskowanie np. o średnich dochodach w społeczeństwie nie powinno nieść ze sobą żadnego obciążenia. Także estymacja parametrów w modelach statystycznych wykorzystujących daną zmienną nie powinna przynieść błędnych oszacowań współczynników. Problemem pozostaje jednak interpretacja statystyk testowych badanych zmiennych, które ściśle zależą od ich odchyłeń standardowych. Brak precyzyjnych odpowiedzi może zaburzyć wnioskowanie statystyczne na podstawie badanego zbioru danych, prowadząc np. do mylnego uznania zmiennych za nieistotne.

Często najbardziej interesującym elementem zbiorów danych są właśnie różnice w badanych odpowiedziach. Doskonałym tego przykładem jest deklarowany dochód, którego różnice w społeczeństwie stanowią temat wielu prac, artykułów i debat. Mimo rzeczywistego wzrostu nierówności społecznych na przestrzeni lat, mogą być one częściowo zawyżane poprzez zjawisko nieprecyzyjnych odpowiedzi. Problem pozycji pytania w ankiecie jest więc istotny i powinien być brany pod uwagę podczas każdej analizy opartej na danych z kwestionariuszy ankietowych. Wyniki dowolnej ankiety mogą być nierównomierne i zawyżać odchylenia standardowe z powodu nawet małego znużenia respondentów. Ponadto

utrudnione może być ich porównywanie z wynikami pozyskanymi na podstawie innych kwestionariuszy, w których te same pytania umieszczone na innych pozycjach mogą przynieść odmienne rezultaty.

Potencjalnym rozwiązaniem mogłaby być metoda wykorzystująca informacje na temat trendu wzrostu odchylenia dla poszczególnych pytań w zależności od ich pozycji. Odpowiedni algorytm mógłby wprowadzać poprawkę do uzyskiwanych wyników, standaryzując je do prawidłowych wielkości. Algorytm ten mógłby ponadto wykorzystywać mieszanie kolejności pytań dla każdego respondenta – umożliwiłoby to skontrolowanie poszczególnych odchylen standardowych w zależności od pozycji pytania i pozbycie się efektów znalezionego trendu. Nie jest to jednak przedmiotem tego artykułu, a metoda ta wymagałaby przemyślanej techniki zbierania odpowiedzi, a tym samym poniesienia większych kosztów badania.

Tekst wpłynął: 29 lipca 2013 r.

## Bibliografia

- Badanie Aktywności Ekonomicznej Ludności*, dane: BAEL II kwartał 2005, BAEL II kwartał 2006.
- Dillman D.A., Sinclair M.D., Clark J.R., *Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys*, „Public Opinion Quarterly” 1993, nr 57(3).
- Galesic M., Bosnjak M., *Effect of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey*, „Public Opinion Quarterly” 2009, nr 73(2).
- Heberlein T.A., Baumgartner R., *Factors Affecting Response Rates to Mailed Questionnaires – A Quantitative Analysis of the Published Literature*, „American Sociological Review” 1978, nr 43(4).
- Herzog A.R., Bachman J.G., *Effects of Questionnaire Length on Data Quality*, „Public Opinion Quarterly” 1981, nr 45(4).
- Krosnick J.A., Alwin D.F., *An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement*, „Public Opinion Quarterly” 1987, nr 51(2).
- Krosnick J.A., *Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys*, „Applied Cognitive Psychology” 1991, nr 5(3).
- Krosnick J.A., Presser S., *Question and Questionnaire Design*, w: *Handbook of Survey Research (2ed)*, red. J.D. Wright, P.V. Marsden, Elsevier, San Diego 2009.
- Krosnick J.A., *Survey Research*, „Annual Review Psychology” 1999, nr 50.
- Lightspeed Research 2008, *Questionnaire Length*, [http://www.lightspeedresearch.com/pdfs/Questionnaire\\_Length.pdf](http://www.lightspeedresearch.com/pdfs/Questionnaire_Length.pdf) [14.08.2011].
- Marcinkowska I., Ruzik A., Strawiński P., Walewski M., *Badanie struktury i zmian rozkładu wynagrodzeń w Polsce w latach 2000–2006*, Ministerstwo Pracy i Polityki Społecznej, Warszawa 2008.
- Polski Generalny Sondaż Społeczny*, dane: PGSS 1992–1995, PGSS 1997/8.
- Riphahn R.T., Serfling O., *Item Non-response on Income and Wealth Questions*, „Empirical Economics” 2005, nr 30(2).
- Subar A.F., et al., *Is Shorter Always Better – Relative Importance of Questionnaire Length and Cognitive Ease on Response Rates and Data Quality for Two Dietary Questionnaires*, „American Journal of Epidemiology” 2001, nr 153(4).
- World Institute for Development Economics Research of the United Nations University 2008, UNU-WIDER World Income Inequality Database, Version 2.0c, [http://www.wider.unu.edu/research/Database/en\\_GB/database/](http://www.wider.unu.edu/research/Database/en_GB/database/) [15.08.2011].

## **DLUGOŚĆ KWESTIONARIUSZA A PRECYZJA ODPOWIEDZI NA PYTANIA ILOŚCIOWE W BADANIACH ANKIETOWYCH**

### **Streszczenie**

Cechy ankiety, takie jak jej długość, trudność czy jej subiektywnie postrzegana ciekawość, mają bezpośredni wpływ na jakość otrzymanych danych. Autor analizuje związek pomiędzy długością kwestionariusza a precyzją pozyskanych z jego pomocą danych ilościowych, wykorzystując dane ankietowe GUS na temat rozkładu wynagrodzeń w społeczeństwie polskim. Badanie opiera się na naturalnym quasi-eksperymentie, jakim było zwiększenie liczby pytań w kwestionariuszu między 2005 r. i 2006 r. Analizie poddano wpływ tej zmiany na rozrzut deklarowanych przez respondentów dochodów. Wyniki wskazują na istotny negatywny związek pomiędzy długością ankiety i precyzją udzielanych odpowiedzi (pytania przesunięte na dalsze miejsca w ankiecie dają mniej precyzyjne odpowiedzi). Uzyskane wyniki są względnie odporne na oddziaływanie innych czynników.

**Słowa kluczowe:** dane ankietowe, rozkład dochodów, kwestionariusz

## **LENGTH OF THE QUESTIONNAIRE AND THE PRECISION OF RESPONSE TO QUANTITATIVE QUESTIONS IN SURVEYS**

### **Summary**

The features of the questionnaire, such as its length, difficulty, or subjectively perceived curiosity, have a direct effect on the quality of survey data received. The author analyses the relationship between the length of the questionnaire and the precision of quantitative data obtained in a survey, using survey data on wage income distribution in the Polish households, collected by the Central Statistical Office. The research is based on a natural quasi-experiment provided by raising the number of questions asked between 2005 and 2006. The author analyses the impact of this change on the dispersion of incomes declared by the respondents. The results show a negative relationship between the length of the questionnaire and the precision of the responses obtained (questions that have been shifted to further places in the questionnaire yield less precise answers). The results are relatively robust against the impact of other factors.

**Key words:** survey data, income distribution, questionnaire

## **ДЛИНА АНКЕТЫ И ТОЧНОСТЬ ОТВЕТОВ НА КОЛИЧЕСТВЕННЫЕ ВОПРОСЫ В АНКЕТНЫХ ИССЛЕДОВАНИЯХ**

### **Резюме**

Такие характеристики анкеты как длина, трудность или субъективно ощущаемая привлекательность, имеют прямое влияние на качество получаемых данных. Автор анализирует связь между длиной анкеты и точностью полученных с ее помощью количественных данных, используя анкетные данные ГСУ на тему расклада заработной платы в польском обществе. Исследование опирается на естественный квазиэксперимент, каковым было увеличение количества вопросов в анкете между 2005 и 2006 гг. Анализировалось влияние изменения анкеты на диапазон декларируемых респондентами доходов. Результаты указывают на существенную отрицательную связь между длиной анкеты и точностью даваемых ответов (вопросы, переставленные на более удаленные места в анкете, дают менее точные ответы). Полученные результаты относительно устойчивы к воздействию других факторов.

**Ключевые слова:** анкетные данные, расклад доходов, анкета