

**Wojciech Hardy\***

## **IM WCZEŚNIEJ TYM LEPIEJ. O PRECYZJI ODPOWIEDZI NA PYTANIA ILOŚCIOWE W BADANIACH ANKIETOWYCH**

### **THE SOONER THE BETTER. ON THE PRECISION OF RESPONSES TO QUANTITATIVE QUESTIONS IN SURVEYS**

#### **Abstract**

*Questionnaire's design, length, difficulty and general appeal may directly influence the quality of the collected data. I analyse the relationship between a survey's length and the precision of acquired responses using Polish Labour Force Survey's data on income distribution. I utilize a natural experiment of elongation of the LFS questionnaire that took place between the years 2005 and 2006 and measure the impact of this change on the distribution of the respondents' declared earnings. The results of the performed tests indicate a significant, negative relationship between the length of the survey and the precision of the acquired data. The robustness of the results is discussed and conclusions are drawn.*

**Key words:** quantitative research, survey, questionnaire, income, BAEL

**JEL classification:** C42, D0

#### **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 negatywne w skutkach i charakterze. Niezależnie jednak od punktu widzenia, nierówności dochodowe przybliżane są przez wiele konstruowanych specjalnie w tym celu miar. 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 przy wykorzystaniu kwestionariuszy jest zarówno

---

\* Doktorant, Zakład Teorii Rozwoju Gospodarczego, Wydział Nauk Ekonomicznych, Uniwersytet Warszawski.

procedura zbierania danych, jak i zapewnienie najlepszej jakości zebranej informacji. Obydwa te czynniki zależą m.in. od struktury ankiety.

W przypadku systematycznych a nieobserwowalnych zaburzeń w jakości danych, pomiary nierówności dochodów w społeczeństwie mogą odzwierciedlać problemy zawarte w samym zbiorze informacji, nie zaś rzeczywiste nierówności społeczne. Ponadto, wyniki przeprowadzanych procedur statystycznych mogą podlegać obciążeniom – badania oparte na materiale 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 pracy tej zbadano wpływ wydłużenia ankiety na precyzję otrzymywanych w niej odpowiedzi. W tym celu wykorzystano dane BAEL dla lat 2005–2006. Struktura kwestionariusza zmieniła się wówczas istotnie z powodu rozszerzenia go o ponad dwadzieścia dodatkowych zagadnień. Wskutek tego działania pytanie o średnie miesięczne wynagrodzenie respondenta przesunęło się w ankiecie o około 17 pozycji dalej<sup>1</sup>. Ten naturalny quasi-eksperyment wykorzystano by przeanalizować 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. Wnioski zostają potwierdzone.

W pierwszej części pracy przedstawiono ogólną problematykę badań ankietowych, uwzględniając przegląd dotychczasowej literatury na ten temat. W sekcji drugiej przeprowadzono analizę wykorzystując, dane Badania Aktywności Ekonomicznej Ludności. Potencjalne ograniczenia i problemy z wynikami zostają omówione. Praca kończy się interpretacją uzyskanych danych i podsumowaniem.

## Przegląd literatury

Literaturę na temat konstruowania ankiet można podzielić na trzy główne grupy. Pierwsza z nich przedstawia temat z punktu widzenia psychologii respondenta i jego motywacji podczas wypełniania kwestionariusza. Druga dotyczy metod zbierania danych i samego skłaniania ankietowanego do brania udziału w badaniu i odpowiadania na wszystkie zawarte w nim pytania. Trzecia grupa porusza problematykę jakości odpowiedzi, czyli prawdziwości i precyzji informacji otrzymywanych od respondenta.

Poniżej przedstawiono najważniejsze badania na ten temat, według zaproponowanej wyżej klasyfikacji.

---

<sup>1</sup> Faktyczne przesunięcie mogło wahać się w zależności od udzielanych wcześniej odpowiedzi.

## Psychologia udzielania odpowiedzi

Bardzo istotną rolę w kształtowaniu się jakości odpowiedzi, jak wskazuje Krosnick i Presser (2009), 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 stadia to<sup>2</sup>:

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

Według licznych prac psychologicznych wkład potrzebny do wykonania wszystkich części powyższego procesu może być motywowany między innymi czynnikami, takimi jak: możliwość autoekspresji, wyzwanie intelektualne czy po prostu altruizm. Jeśli respondent podejmie trud przejścia przez każdy z wyżej wymienionych stadiów, jego postawę możemy nazwać postawą „optymalizującą”<sup>3</sup>. Oczywistym celem każdego badacza jest zmaksymalizowanie dokładnie takiego zachowania wśród respondentów, co jednak może być utrudnione przez szereg zjawisk, począwszy od wypełniania ankiety z przymusu, aż po znużenie czy dekoncentrację podczas samego procesu.

Na drugim końcu skali zachowań respondentów kładzie się postawę satysfakcjonującą/wystarczającą<sup>4</sup>. Zjawisko to jest związane z szeregiem niedoskonałości podczas wymienionego powyżej procesu myślowego. Działanie to dotyczy przede wszystkim udzielania odpowiedzi na pytania. U respondentów może wystąpić problem ze zrozumieniem treści pytania. Ponadto mogą poświęcić mu zbyt mało uwagi. Przeszukanie pamięci może odbyć się mniej dokładnie, zaś integracja odnalezionych informacji nieuważnie. Wreszcie dobór odpowiedzi może ucierpieć na precyzji. Powodem tego nie musi być zła wola, ale nawet znużenie czy dekoncentracja (tzw. *weak satisficing*). Niezależnie jednak od powodu, proces decyzyjny zbiega do procesu wystarczającego z punktu widzenia respondenta, lecz dającego odpowiedzi rozmiijające się z prawdą. W skrajnych przypadkach ankietowani mogą pominąć część etapów i od razu skupić się na wypełnianiu kwestionariusza według domniemych oczekiwań badaczy (tzw. *strong satisficing*), tym samym podważając wiarygodność wyników badań.

---

<sup>2</sup> On A. Krosnick, Stanley Presser 2009, Question and Questionnaire Design, *Handbook of Survey Research* (2ed), James D. Wright and Peter V. Marsden (Eds). San Diego, CA: Elsevier, s. 265–266.

<sup>3</sup> (Ang. *optimizing*).

<sup>4</sup> (Ang. *satisficing*).

Krosnick (1991) podaje trzy główne czynniki odpowiedzialne za zachowanie respondenta<sup>5</sup>:

- trudność zadania – związana ze stopniem trudności pytania i procesu pozyskania informacji oraz czynnikami zewnętrznymi, takimi jak np.: prędkość otrzymywania pytań, lub rozpraszające okoliczności,
- zdolność respondenta – dotyczy to ogólnych możliwości umysłowych respondenta, jego wiedzy o zagadnieniu i poczynieniu wcześniejszych przemyśleń na temat poruszanych kwestii,
- motywacja respondenta – wiąże się z potrzebą pracy umysłowej, indywidualną wagą tematu pytań, wiarą w istotność badania, zmęczenie oraz sposób aplikacji kwestionariusza.

Należy zauważyć, że w przypadku pytania o wynagrodzenie respondenta może występować wysokie ryzyko zachowania typu *satisficing*. Po pierwsze, respondent nie zawsze potrafi przypomnieć sobie dokładną wartość miesięcznego dochodu. Po drugie, zadanie to może być dodatkowo utrudnione, jeśli wypełniany przez ankietowanego kwestionariusz jest zbyt długi. Dość często prowadzi to do zmęczenia i dekoncentracji ankietowanego, który jednocześnie straci zdolność i motywację znalezienia precyzyjnej odpowiedzi. Jeśli przyjąć istnienie jednego rozkładu dochodu w społeczeństwie polskim, konstrukcja ankiety, sprzyjająca zjawisku *satisficing*, może doprowadzić do zebrania przykładów, znajdujących się jedynie „w pobliżu” odpowiedzi prawdziwych. Tym samym badana zmienna wykaże sztucznie zawyżoną wariancję. Efekt ten zostaje pokazany w niniejszej pracy.

## Problemy z pozyskiwaniem odpowiedzi

Istnieje szereg badań dotyczących problemu niskiej frekwencji udziału ankietowanych, czyli rzadkiej zgody na wypełnienie kwestionariusza. Część owych analiz dotyka jeszcze ankiet zwrotnych przesyłanych drogą pocztową. Wielu autorów (Dillman, Sinclair, Clark 1993; Heberlein, Baumgartner 1978) wskazywało na ten problem w swoich pracach, dochodząc nieraz do przeciwstawnych wniosków na temat stopnia wpływu długości ankiety na *response rate*. Dillman, Sinclair oraz Clark (1993) zwrócili uwagę na istotne, negatywne oddziaływanie długości kwestionariusza, zauważyli także znaczącą rolę „przyjazności” formy arkusza oraz obecności trudnych lub kontrowersyjnych pytań, które w sposób oczywisty zniechęcają do uczestnictwa w badaniu.

---

<sup>5</sup> Jon A. Krosnick, Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, *Applied Cognitive Psychology*, Vol. 5, John Wiley & Sons, Ltd. 1991, s. 213–236.

Kolejnym problemem związanym z przeprowadzaniem ankiet jest wysoki odsetek rezygnacji podczas uczestniczenia w badaniu. Ten powód, a także niski udział w ankietowaniu mogą zaburzyć reprezentatywność próby i w znaczący sposób wpłynąć na jej wielkość. Problem ten pojawia się, gdy dany respondent w trakcie wypełniania ankiety utraci motywację do kontynuacji. W szczególności może to wynikać z długiego czasu trwania badania. W eksperymencie przeprowadzonym przez Galesic i Bosnjak (2009) zależność ta została potwierdzona – im dłużej wypełnianio kwestionariusze, tym zwiększało się ryzyko rezygnacji respondenta z dalszego udziału w ankietowaniu.

Pokrewna dwóm powyższym problemom i istotniejsza w kontekście tego badania jest analiza porzucania niektórych pytań, tzw. stopnia niewypełnialności. Problem ten poruszony został w badaniu spółki Lightspeed Research (2008). Ich wyniki wyraźnie wskazują na wzrost niewypełnialności wraz z trwaniem ankietowania, choć możliwe jest przeciwdziałanie temu zjawisku poprzez dodatkową motywację (np.: pieniężną), udzielaną w trakcie przeprowadzania ankiety.

Wykorzystywanie dodatkowej zachęty pieniężnej wiązałoby się jednak ze znacznie większymi kosztami badań ankietowych. Ważnym staje skonstruowanie kwestionariuszy tak, by minimalizować to niekorzystne zjawisko. Respondenci najczęściej omijają pytania, które uważają za zbyt prywatne lub kontrowersyjne. Kwestia wynagrodzenia zalicza się do grupy pytań o szczególnie wysokim stopniu niewypełnialności, co sprawia, że wymaga wyjątkowej uwagi (np.: Riphahn i Serfling, 2005). Analiza zależności między wielkością tego problemu a długością ankiety może opierać się na przetestowaniu prostej hipotezy dotyczącej relacji między zmianą konstrukcji kwestionariusza a liczbą przypadków, w których respondent zdecydował się ominąć to pytanie.

## Problemy z jakością odpowiedzi

Istnieje wiele ujęć kwestii jakości odpowiedzi. Herzog i Bachman (1981)<sup>6</sup> wyróżniają problemy, takie jak odpowiadanie w prostej linii na pytania zamknięte. Subar i inni (2001) pokazują, że jakość odpowiedzi może być silniej związana z łatwością interpretacji ankiety niż z samą jej długością<sup>7</sup>. Autorzy definiują jednak

<sup>6</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 ‘*strong satisficing*’, czyli w tym wypadku: zaznaczania odpowiedzi, które wydają się być stereotypowymi i schematycznymi.

<sup>7</sup> W swoim badaniu Subar et al. porównują wyniki dwóch ankiet różniących się między sobą długością, jak i skomplikowaniem (ankieta dłuższa jest równocześnie prostsza do interpretacji).

jakość jako odsetek wspomnianego już *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: *incompletion rate*, długość odpowiedzi udzielanych na pytania otwarte, czas odpowiadania na pytania, oraz wariancję odpowiedzi na pytania z siatkami odpowiedzi. Według badaczy, długość udzielanych informacji malała wraz z czasem trwania ankiety. Respondenci udzielali odpowiedzi mniej wyczerpujących. Podobna zależność zaszła w przypadku czasu poświęcanego na pojedyncze pytanie, kiedy to respondenci krócej zastanawiali się nad pojedynczą kwestią, co mogło wpłynąć na precyzję udzielanych informacji. Czas trwania miał ujemny wpływ na wariancję danych udzielanych w pytaniach z siatką odpowiedzi, co można tłumaczyć zjawiskiem *strong satisficing*. Galesic i Bosnjak nie zauważyli żadnego oddziaływania na obserwowany stopień niewypełnialności, choć, jak zaznaczają, może to wynikać z formy przeprowadzenia ankiety.

## Przedmiot badania

Celem tego badania jest analiza wpływu długości ankiety na precyzję otrzymywanych w niej odpowiedzi. Na istnienie takiej zależności wyraźnie wskazuje literatura przytoczona we wcześniejszej części tej pracy. By przetestować teorię o negatywnym wpływie czasu trwania ankiety na jakość zbieranych w niej danych, wykorzystane zostaną dane ankietowe – Badania Aktywności Ekonomicznej Ludności<sup>8</sup> dla lat 2005–2006, kiedy to nastąpiła znacząca zmiana w formie wykorzystywanych kwestionariuszy. Dotychczasowy zestaw pytań został uzupełniony o około dwadzieścia dodatkowych, tym samym znacznie rozszerzając kwestionariusz i przesuując wiele ze znajdujących się w nim pytań na dalsze pozycje. Wśród przesuniętych w ten sposób zagadnień znajdowało się pytanie o średnie miesięczne wynagrodzenie respondenta. Numery wybranych pytań dla roku 2005 i 2006 zebrane są w tabeli nr 1.

---

<sup>8</sup> Badanie Aktywności Ekonomicznej Ludności (w skrócie: 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 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. Źródło: <http://www.stat.gov.pl/gus/>.

Tabela 1. Pytania ankiet BAEL, wykorzystane w badaniu

| rok  | pozycja (numer <sup>9</sup> ) | treść  |
|------|-------------------------------|--|
| 2005 | 20                            | Proszę podać wysokość swoich zarobków netto w poprzednim miesiącu w głównym miejscu pracy (w zł)           |
| 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? (w zł) |
| 2006 | 72                            |  |
| 2005 | 72                            | Ile godzin łącznie poświęcił(a) Pani na doksztalcanie w ciągu ostatnich czterech tygodni?                  |
| 2006 | 98                            |  |

*Źródło: opracowanie własne na podstawie ankiet BAEL.*

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) i zniecierpliwienie respondenta obniżające precyzję udzielanych przez niego odpowiedzi, w tym informacji na temat średniego miesięcznego dochodu,

2. Demotywacja respondenta, wpływająca negatywnie na prawdopodobieństwo udzielenia przez niego odpowiedzi na osobiste pytanie o średni miesięczny dochód.

Praca ta weryfikuje hipotezę o zachodzeniu powyższych dwóch efektów. Pierwszy z nich zmierzony zostanie poprzez zmianę w wariancji deklarowanych średnich dochodów społeczeństwa w latach 2005–2006. Jeśli efekt ma miejsce, w 2006 roku zauważalny będzie wyraźny wzrost wariancji w deklarowanych dochodach. Drugi z rezultatów zostanie przeanalizowany za pomocą porównania

<sup>9</sup> Dokładna pozycja pytań w obu latach mogła różnić się w zależności od wcześniejszych odpowiedzi indywidualnego respondenta.

odsetka niewypełnionych odpowiedzi na pytanie o miesięczne wynagrodzenie w latach 2005 i 2006. Jeśli otrzymuje się ten rodzaj efektów, wówczas w 2006 roku będzie można zaobserwować spadek otrzymywanych odpowiedzi na pytanie o dochód respondenta.

## Metoda badawcza

By określić statystyczny wpływ zmiany konstrukcji scenariusza wykorzystane zostają trzy testy na równość wariancji wynagrodzenia w latach 2005 i 2006.

Pierwszym z wykorzystanych testów jest test F (Fisher – lata 20. XX wieku). 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>10</sup>.

Drugim testem jest test U Manna-Whitneya-Wilcoxon, znany też jako test rang (Wilcoxon – 1945; Mann, Whitney – 1947). 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 roku 2005 jest większe od wynagrodzenia w roku 2006 z takim samym prawdopodobieństwem, z jakim wynagrodzenie w roku 2006 jest większe od tego w 2005. w uproszczeniu, hipoteza zerowa oznacza równość rozkładów w dwóch grupach. Test ten nie posiada żadnych założeń odnośnie 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 zostają jego dwie, pokrewne, alternatywy – testy Browna-Forsythe'a zastępujące w formie testu wartość średniej odpowiednio medianą lub 10% średnią obcięta.

Wyniki wszystkich testów są zaraportowane w następnej sekcji pracy. w celu analizy odporności wyników na nieidentyfikowalne wpływy, testy te zostają przeprowadzone także dla czterech innych pytań, przytoczonych w tabeli 1. Ponadto dokonane zostaje omówienie otrzymanych rezultatów, ich możliwe powody, oraz odporność na inne problemy.

---

<sup>10</sup> Test na skośność i kurtozę rozkładu dochodu potwierdza istotną różnicę od rozkładu normalnego.



## Wyniki

Tabela 2 zawiera wyniki testów statystycznych dla równości wariancji deklarowanego dochodu w dwóch podpróbach, 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 je wzmacnia. Tym samym potwierdzona zostaje teza stawiana w tym badaniu – w roku 2006 nastąpił znaczny wzrost w rozrzucie deklarowanych wynagrodzeń.

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

| Pytanie   | ‘Jaki dochód miesięczny’ |          | ‘Ile chciałby pracować’ |      | ‘Godziny pracy dod.’ |      | ‘Minimalny zarobek’ |        | ‘Ile godz. doksztalcenia’ |       |
|---|--------------------------|----------|-------------------------|------|----------------------|------|---------------------|--------|---------------------------|-------|
|   | 2005                     | 2006     | 2005                    | 2006 | 2005                 | 2006 | 2005                | 2006   | 2005                      | 2006  |
| Rok   | 2005                     | 2006     | 2005                    | 2006 | 2005                 | 2006 | 2005                | 2006   | 2005                      | 2006  |
| Odchylenie  | 673,1709                 | 691,7568 | 10,23                   | 9,9  | 8,72                 | 8,7  | 398,7               | 423,31 | 16,74                     | 21,08 |
| Test F  | wzrost odchylenia*       |          | spadek odch.*           |      | brak zmian           |      | wzrost odch.*       |        | wzrost odch.*             |       |
| Test Levene’a   | zmiana odchylenia*       |          | zmiana odch.*           |      | brak zmian           |      | zmiana odch.*       |        | zmiana odch.*             |       |
| Test Browna-Forsythe’a W50                                    | zmiana odchylenia*       |          | zmiana odch.*           |      | brak zmian           |      | zmiana odch.**      |        | zmiana odch.*             |       |
| Test Browna-Forsythe’a W10                                    | zmiana odchylenia*       |          | zmiana odch.*           |      | brak zmian           |      | zmiana odch.*       |        | zmiana odch.*             |       |
| Test MW   | zmiana rozkładu*         |          | zmiana rozkładu*        |      | brak zmian           |      | zmiana rozkładu*    |        | brak zmian                |       |
| * – na poziomie istotności 1%; ** – na poziomie istotności 5% |                          |          |                         |      |                      |      |                     |        |                           |       |

*Źródło: opracowanie własne.*

Wynik ten może być podyktowany czymś więcej niż zmianą 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 zostaje przeprowadzona dla czterech innych pytań z ankiet BAEL, przytoczonych już wcześniej.

Rezultaty dodatkowych testów zawarte są w tabeli 2. Otrzymane wyniki wskazują analogiczną zależność w przypadku dwóch z pozostałych pytań, brak zależności w przypadku jednego z nich i odwrotną zależność w przypadku ostatniego. Wyniki te zostają szerzej omówione w następnym rozdziale.

## Omówienie wyników

Rezultaty przeprowadzonego badania wskazują na wzrost wariancji deklarowanych wynagrodzeń w roku 2006, 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 dochodu na przestrzeni lat. Rzeczywiście, w debacie publicznej często wspomina się o wzrastających nierównościach w zakresie wynagrodzeń. Jeśli więc w roku 2006 nastąpiła znacząca zmiana w zróżnicowaniu dochodów, można się spodziewać, że owa zmiana sytuacji będzie znacząco wpływać na wyniki testów i tym samym uzyskane wyniki będą jedynie wskazywać na wzrost nierówności społecznych w roku 2006, w stosunku do roku 2005.

**Tabela 3. Wartości współczynnika Giniego dla Polski, w latach 2005–2006**

| rok  | wsp. giniego | źródło  |
|------|--------------|---|
| 2005 | 36,6         | Household Budget Survey (Transmonee 2007)         |
| 2006 | 34           | Survey of Earnings Distribution (Transmonee 2007) |
| 2005 | 35,6         | EU-SILC <sup>11</sup>                             |
| 2006 | 33,3         | EU-SILC   |
| 2005 | 34,92        | Bank Światowy                                     |
| 2006 | 34,13        | Bank Światowy                                     |
| 2005 | 27,13        | BAEL  |
| 2006 | 27,48        | BAEL  |

**Źródło:** opracowanie własne na podstawie danych według danych i obliczeń UNU-WIDER, Banku Światowego, oraz według danych BAEL.

<sup>11</sup> The European Union Statistics on Income and Living Conditions.

Pewną informację na temat wielkości tego problemu mogą dostarczyć inne miary nierówności zastosowane dla społeczeństwa polskiego, w szczególności tzw. współczynnik Giniego. Wskaźnik ten służy do mierzenia równości w dochodach w zakresie całych krajów. Przyjmuje on wartości z przedziału  $<0,1>$ , gdzie 0 oznacza całkowitą równość, zaś 1 całkowite zróżnicowanie. Dane dotyczące wartości tego wskaźnika, dla wielu krajów, zbierane są przez liczne niezależne Instytuty, w tym przez UNU-WIDER<sup>12</sup>. W tabeli 3 przedstawione są wartości współczynnika Giniego dla Polski, w latach 2005–2006, według danych zgromadzonych przez UNU-WIDER.

Jak widać, współczynnik Giniego wskazuje na zmianę nierówności w dochodach w roku 2006. Co jednak ciekawe i nieco zaskakujące na tle trendu z sąsiadujących lat, wskazuje on na spadek nierówności. Rzeczywiście, wskaźnik Giniego w poprzedzających latach ma tendencję wzrostową aż do roku 2005, po czym ‘gwałtownie’ spada jak przedstawiono w tabeli 3. Oznacza to, że sam fakt upływu czasu między latami 2005 i 2006 powinien mieć efekt przeciwny do tego zaobserwowanego na podstawie danych BAEL. Wniosek ten potwierdza interpretowanie wyników przeprowadzonych testów jako świadectwo 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 przepracowanych 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 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 do wcześniejszej w tym większym stopniu, im dłuższy jest kwestionariusz. Zależność tę można zweryfikować, ponieważ pozycja pytania o pożądaną liczbę godzin pracy zmienia się w roku 2006 o 19 miejsc,

---

<sup>12</sup> World Institute for Development Economics Research of the United Nations University.

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 roku 2006, średnie odpowiedzi z obu pytań powinny więc być sobie bliższe niż w roku 2005 z uwagi na częstsze odnoszenie się do wcześniej wyszukanej informacji.

Tabela 4. Wyniki testu  $t$  na równość średnich w dwóch próbach.

|                          | Grupa                  |          |
|--------------------------|------------------------|----------|
|                          | 2005                   | 2006     |
| Obserwacje               | 19 795                 | 19 682   |
| Średnia                  | 4,11                   | 3,59     |
| Odch. Stand.             | 7,86                   | 7,49     |
| H1: $\text{diff} < 0$    | $\text{Pr}(T < t)$     | = 1.0000 |
| H2: $\text{diff} \neq 0$ | $\text{Pr}( T  >  t )$ | = 0.0000 |
| H3: $\text{diff} > 0$    | $\text{Pr}(T > t)$     | = 0.0000 |

**Źródło:** opracowanie własne.

Istnienie tego efektu sprawdzamy porównując średnie absolutne różnice obu deklarowanych przez respondentów wartości w roku 2005 i 2006. Jeśli efekt zachodzi, test na równość średnich powinien wykazać spadek średniej w roku 2006, czyli zjawisko zbiegania odpowiedzi na pytanie o pożądaną liczbę godzin pracy do wcześniej deklarowanych wartości. Wyniki testu  $t$  na równość średnich przedstawione są w tabeli 4 i wskazują na spadek wartości różnic w roku 2006, 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 dla pytania o wymarzoną liczbę godzin. Jest to jednak potencjalne wytłumaczenie, a wskazana różnica w kwestionariuszu dotyczy jedynie tej zmiennej.

## Podsumowanie i wnioski

W pracy przetestowany został 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 latach 2005–2006, badane pytania przesunęły się o około 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 roku 2006; przy jednym pytaniu nie zauważono różnicy; przy jednym zaś wykryto zjawisko odwrotne.

Analiza wartości wskaźników 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 być także prostsze od pozostałych, tym samym w mniejszym stopniu podatne na zmęczenie respondentów. Wyniki, poparte dodatkową analizą zmiennych, potwierdzają hipotezę stawianą w badaniu.

Przebadane 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 oddziaływujących na wariancję odpowiedzi, która w niniejszym badaniu stanowi główny wskaźnik badanego problemu. Głównym z tych efektów jest skłonność do udzielania odpowiedzi łatwych, szybkich, neutralnych, stereotypowych, czy schematycznych, która to obniża różnorodność udzielanych odpowiedzi, a więc i wariancję, do stopnia w którym zjawisko obniżonej precyzji nie daje się zaobserwować. Trudno jednak nie zgodzić się, że 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 o np.: średnich dochodach w społeczeństwie nie powinno nieść ze sobą żadnego obciążenia. Rzeczywiście także estymacja parametrów w modelach statystycznych wykorzystujących daną zmienną nie powinna przynieść błędnych współczynników. Problemem pozostaje jednak interpretacja statystyk testowych badanych zmiennych, które ściśle zależą od ich odchyień

standardowych. Brak precyzyjnych odpowiedzi może zaburzyć wnioskowanie statystyczne na podstawie badanego zbioru danych prowadząc do np.: mylnego uznania zmiennych za nieistotne.

Często najbardziej interesującym elementem zbiorów danych są właśnie różnice w badanych odpowiedziach. Doskonałym tego przykładem jest najczęściej poruszana w tej pracy kwestia – 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 jest więc istotny i powinien być brany pod uwagę podczas każdej analizy na podstawie danych z kwestionariuszy. 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 odchyleń standardowych w zależności od pozycji pytania i pozbycie się efektów znalezionego trendu. Nie jest to jednak przedmiotem tej pracy, a metoda ta wymagałaby przemyślanej techniki zbierania odpowiedzi, a tym samym poniesienia większych kosztów badania.

## Bibliografia

- Badanie Aktywności Ekonomicznej Ludności, dane: BAEL II kwartał 2005, BAEL II kwartał 2006.
- Dillman Don A., Michael D. Sinclair, Jon R., Clark Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys, *Public Opinion Quarterly*, Vol. 57, No. 3, Oxford University Press on behalf of the American Association for Public Opinion Research 1993.
- Krosnick Jon A., Duane F., Alwin An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement, *Public Opinion Quarterly*, Vol. 51, No. 2, The University of Chicago Press on behalf of the American Association for Public Opinion Research 1987.
- Krosnick Jon A., Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, *Applied Cognitive Psychology*, Vol. 5, John Wiley & Sons, Ltd. 1991.
- Krosnick Jon A., Stanley Presser, Question and Questionnaire Design, *Handbook of Survey Research (2ed)*, James D. Wright and Peter V. Marsden (Eds). San Diego, CA: Elsevier 2009.
- Krosnick Jon A., Survey Research, *Annual Review Psychology*, Vol. 50, Annual Reviews 1999.
- Lightspeed Research, Questionnaire Length [online], Lightspeed Research 2008, Dostępne: [http://www.lightspeedresearch.com/pdfs/Questionnaire\\_Length.pdf](http://www.lightspeedresearch.com/pdfs/Questionnaire_Length.pdf), [14 sierpień 2011]

- Mirta Galesic, Michael Bosnjak, Effect of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey, *Public Opinion Quarterly*, Vol. 73, No. 2, Oxford University Press 2009.
- Polski Generalny Sondaż Społeczny, dane: PGSS 1992–1995, PGSS 1997/8
- Regula Herzog A., Jerald G., Bachman Effects of Questionnaire Length on Data Quality, *Public Opinion Quarterly*, Vol. 45, by the Trustees of Columbia University, Elsevier North-Holland, Inc. 1981.
- Riphahn, Regina T., Oliver Serfling, Item Non-response on Income and Wealth Questions, *Empirical Economics* 30, 2005, No. 2.
- Subar Amy F., Ziegler Regina G., Frances E. Thompson, Christine Cole Johnson, Joel L. Weissfeld, Douglas Reding, Katherine H. Kavounis, and Richard B. Hayes for the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Investigators, 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*, Vol. 153, No. 4, The Johns Hopkins University School of Hygiene and Public Health 2001.
- Thomas A. Heberlein, Robert Baumgartner Factors Affecting Response Rates to Mailed Questionnaires – A Quantitative Analysis of the Published Literature, *American Sociological Review*, Vol. 43, No. 4, American Sociological Association, 1978.
- World Institute for Development Economics Research of the United Nations University 2008, UNU-WIDER World Income Inequality Database [online], Version 2.0c, Dostępne: [http://www.wider.unu.edu/research/Database/en\\_GB/database/](http://www.wider.unu.edu/research/Database/en_GB/database/), [15 sierpień 2011]

### Streszczenie

Cechy ankiety, takie jak jej długość, trudność, czy jej subiektywnie postrzegana ciekawość mają bezpośrednie przełożenie na jakość otrzymywanych danych. Praca analizuje związek pomiędzy długością kwestionariusza, a precyzją pozyskanych z jego pomocą danych, wykorzystując dane BAEL na temat rozkładu wynagrodzeń w społeczeństwie polskim. Badanie opiera się na naturalnym quasi-eksperymentcie, jakim było zwiększenie liczby pytań w kwestionariuszu BAEL między rokiem 2005, a 2006. Analizie poddano wpływ tej zmiany na rozrzut deklarowanych przez respondentów dochodów. Wyniki wskazują na istotny, negatywny związek pomiędzy długością ankiety i precyzją udzielanych odpowiedzi. Uzyskane wyniki są względnie odporne.

**Słowa kluczowe:** badania ilościowe, ankieta, kwestionariusz, dochód, BAEL

**Numer klasyfikacji JEL:** C42, D0