

Metodologiczne Inspiracje 2018

Badania Ilościowe w Naukach Społecznych — Wyzwania i Problemy

24–26 września, Jabłonna

24 września

Sesja 1. Prowadzący: Dariusz Przybysz

TOMASZ ŻÓŁTAK, INSTYTUT BADAŃ EDUKACYJNYCH

Rozłączność przedziałów ufności jako kryterium weryfikacji hipotez statystycznych

Wyznaczanie przedziałów ufności oraz weryfikacja hipotez statystycznych z wykorzystaniem procedury Neymana-Pearsona to dwie wyraźnie różne koncepcje prowadzenia wnioskowania statystycznego. Przy realizacji obu procedur odwołuje się jednak do rozkładów statystyk, które są ze sobą ściśle powiązane formalnie. W związku z tym w wielu przypadkach znając wynik realizacji jednej procedury, można łatwo wyciągnąć pewne wnioski na temat rezultatu przeprowadzenia drugiej z nich. Przykładowo dosyć powszechnie znany jest fakt, że przy wnioskowaniu na temat średniej z próby z wykorzystaniem aproksymacji normalnej, gdy przedział ufności dla wartości populacyjnej wyznaczony na poziomie ufności γ nie obejmuje wartości c , to hipoteza zerowa głosząca, że wartość średniej w populacji jest równa c , zostanie w wyniku przeprowadzenia procedury Neymana-Pearsona, przy zastosowaniu testu dwustronnego, odrzucona, jeśli przyjmie się poziom istotności α większy lub równy $1 - \gamma$.

Chęć stosowania podobnych reguł wnioskowania pojawia się zwykle w sytuacji, gdy wyniki analiz prezentowane są w formie graficznej z wykorzystaniem przedziałów ufności, podczas gdy badacz chciałby wypowiedzieć się na temat wyniku procedury weryfikacji hipotezy statystycznej. Często dotyczy jednak bardziej skomplikowanych sytuacji, gdy związki pomiędzy oboma podejściami są znacznie bardziej złożone niż te wspomniane powyżej i stosowanie takich prostych reguł jest nieuprawnione.

Sytuacją, która opisana zostanie w tym wystąpieniu, jest wykorzystanie przedziałów ufności wyznaczonych dla dwóch różnych parametrów do wnioskowania o rezultacie weryfikacji hipotezy głoszącej równość wartości obu tych parametrów (przeciwko hipotezie, że wartości te się od siebie różnią). Problem taki występuje np. przy analizie efektów brzegowych (*marginal effects*) przy interpretacji wyników estymacji modeli regresji. Często popełnianym w takiej sytuacji błędem jest wnioskowanie o istotności różnic wartości parametrów na podstawie rozłączności przedziałów ufności wyznaczonych oddzielnie dla każdego z tych parametrów, przy założeniu, że poziom ufności i poziom istotności pozostają ze sobą w takim związku, jak w opisanym wcześniej wnioskowaniu nt. średniej.

W wystąpieniu zaprezentowane zostanie, jakie związki łączą w takiej sytuacji poziom ufności, dla jakiego wyznaczone zostały przedziały ufności i poziom istotności, na jakim można by odrzucić hipotezę zerową głoszącą równość parametrów w wyniku realizacji procedury Neymana-Pearsona. Omówione zostanie, na jakie błędy wnioskowania naraża badaczy stosowanie w takiej sytuacji wspomnianej powyżej

uproszczonej i niepoprawnej reguły oraz od jakich własności analizowanych danych zależy wielkość takich błędów. Przedstawione zostaną też rekomendacje odnośnie do sposobu graficznej prezentacji wyników analizy regresji w sposób bardziej adekwatny do wykorzystania we wnioskowaniu z wykorzystaniem procedury Neymana-Pearsona.

TADEUSZ KRAUZE

Częstość i zgodność odpowiedzi w pytaniu kwestionariuszowym

W pracy przedstawione są wyniki badania związków między częstością i zgodnością odpowiedzi na jedno pytanie kwestionariuszowe zawierające s kategorii (czyli opcji) odpowiedzi. Przez częstość rozumiemy $p_i = \frac{m_i}{m}$, gdzie m jest liczbą respondentów, zaś m_i jest liczbą respondentów, którzy wybrali i -tą opcję. Oczywiście

$$p_1 + p_2 + p_3 + \dots + p_s = 1.$$

Zgodność odpowiedzi definiujemy jako sumę kwadratów częstości

$$p_1^2 + p_2^2 + \dots + p_s^2 = a.$$

Parametr a jest proporcją par respondentów, którzy wybrali tę samą opcję. Pokazujemy, że minimalna wartość a wynosi $\frac{1}{s}$, zaś maksymalna wynosi 1, oraz że dla $\frac{1}{s} \leq a \leq 1$ zachodzi nierówność

$$a \leq \max(p_1, p_2, \dots, p_s).$$

Podajemy geometryczną interpretację tej nierówności. Wskazujemy również nietrywialne warunki, przy których zachodzi równość $a = \max(p_1, p_2, \dots, p_s)$.

Dla ustalonego $b = \max(p_1, p_2, \dots, p_s)$, gdzie $\frac{1}{s} < b < 1$, istnieje wiele możliwych zgodności $a = p_1^2 + p_2^2 + \dots + p_s^2$. Stawiamy pytania typu: jeżeli dana jest maksymalna częstość, to jaka jest maksymalna i minimalna zgodność? Innymi słowy, staramy się obliczyć górne ograniczenie $g(b) = \max(a)$ oraz dolne ograniczenie $d(b) = \min(a)$. Odwrotnie, jeżeli jest dane a (czyli zgodność), to czy można oszacować najmniejszą i największą maksymalną częstość?

Zadania matematyczne tego rodzaju należą do teorii optymalizacji, gdzie ważną rolę odgrywa zasada dwoistości. Przedstawiamy ideę tej zasady i podajemy postać analityczną funkcji, które rozwiązują cztery problemy postawione w pytaniach. Pokazujemy geometryczną postać zbioru punktów (b, a) spełniających wymienione ograniczenia (w zależności od s).

Na zakończenie szkicujemy zastosowania przedstawionych rozwiązań.

MACIEJ KONIEWSKI, PAULINA SKÓRSKA, INSTYTUT SOCJOLOGII UNIwersytet Jagielloński

Aplikacja techniki dopasowania statystycznego do doboru próby w badaniach jakościowych jako sposób lepszego rozumienia wyników analiz kontrfaktycznych

Analizy kontrfaktyczne na świecie stosowane są od kilku dekad, choć w ewaluacjach *ex-post* w Polsce pierwsze zastosowania miały miejsce stosunkowo niedawno. Dyskusja teoretyków i praktyków ogniskuje się obecnie na ograniczeniach analiz kontrfaktycznych wiążących się głównie z tym, że umożliwiają jedynie oszacowanie uśrednionych efektów jakiejś interwencji/oddziaływania. Nie pozwalają natomiast na prześledzenie mechanizmów, które do nich doprowadziły (nie odpowiadają na pytania o przyczyny). Stąd coraz częściej pojawia się postulat uzupełnienia tych analiz o perspektywę jakościową — jednym z takich paradygmatów jest oparta na teorii ewaluacja wpływu (*theory-based impact evaluation*). Niestety w praktyce to podejście nie było do tej pory implementowane na szeroką skalę.

W prezentacji zostanie omówione zaproponowane przez autorów podejście łączące w praktyce perspektywę kontrfaktyczną i jakościową. Przybliżona zostanie technika doboru beneficjentów do badań jakościowych, oparta na statystycznym dopasowaniu (*statistical matching*) i wskazane korzyści wynikające z zastosowania podejścia. Technika zostanie omówiona na przykładzie jej wykorzystania w ramach interpretacji efektów działania 4.4 Programu Operacyjnego Innowacyjna Gospodarka — „Nowe inwestycje o wysokim potencjale innowacyjnym”. Statystyczne dopasowanie posłużyło w badaniu identyfikacji „firm sukcesu” (przedsiębiorstw, którym udało się przekuć otrzymane z PO IG dofinansowanie na wzrost wskaźników konkurencyjności) oraz firmy, z historią porażki (brak wpływu wsparcia na konkurencyjność). W zaprezentowanym badaniu zastosowano dopasowanie statystyczne oparte na odległości Mahalanobisa (*Mahalanobis distance matching*, MDM) między firmami wg wartości wskaźnika konkurencyjności i kwoty uzyskanego dofinansowania oraz jednocześnie dopasowanie dokładne wg: roku rozpoczęcia i zakończenia inwestycji, wielkości firmy i branży głównej działalności. Następnie w tak dobranych kontrastowych parach firm (sukces vs porażka) przeprowadzono badania jakościowe. Pozwoliły one zidentyfikować czynniki mające potencjalny wpływ na przełożenie dofinansowania na konkurencyjność przedsiębiorstw.

ARTUR POKROPEK, INSTYTUT FILOZOFII I SOCJOLOGII PAN

(Nie)porównywalność pomiarowa w typie porównawczych badań sondażowych. Koncepcje, modele i walidacja.

W badaniach porównawczych do analizy wykorzystuje się zazwyczaj konstrukty mierzone przez kilka wskaźników uzyskanych zazwyczaj jako odpowiedzi na pytania kwestionariuszowe. Odpowiedzi te za pomocą różnych modeli skalowania wykorzystywane są do budowania jednej skali, która ma odzwierciedlać intensywność badanego zjawiska. Na przykład, w badaniu ESS do pomiaru poziomu uogólnionego zaufania politycznego wykorzystuje się odpowiedzi respondentów mówiące, na ile ufają oni parlamentowi, na ile politykom, a na ile partiom politycznym. Mimo iż badania porównawcze mogą poszczycić się długą tradycją, to do tej pory problemy konstrukcji międzynarodowo porównywalnych wskaźników stanowią poważną przeszkodę stojącą na drodze do trafności dociekań substancyjnych (van de Vijver, 2011).

Standardowym podejściem zarówno do testowania hipotez odnoszących się do porównywalności wskaźników, jak i konstrukcji skal w badaniach społecznych jest wielogrupowa confirmacyjna analiza czynnikowa (*Multi-Group Confirmatory Factor Analysis*, MGCFA). Podejście, które umożliwia budowę skal, nawet gdy część wskaźników nie jest porównywalnych, nazywane jest częściową ekwiwalencją (*partial equivalence*) (Byrne i wsp. 1989). Mimo iż procedury oparte na MGCFA nie są nowe, to w kontekście badań międzynarodowych metoda ta wykazuje problemy, które wciąż czekają na rozwiązanie (van de Vijver, 2011)

W ostatnich latach rozwinięto również inne metody, które można wspólnie określić jako metody przybliżonej ekwiwalencji (*approximate equivalence*). Należą do niej: wielopoziomowa analiza czynnikowa (*Multilevel Confirmatory Factor Analysis*) (Fox, 2010; Hox, de Leeuw i Brinkhuis, 2010), Bayesowskie modele strukturalne (*Bayesian Structural Equating Modeling*) (De Jong, Steenkamp i Fox, 2007; Muthén i Asparouhov, 2012) oraz metody optymalizacji wyrównującej (*Alignment optimization methods*) (Asparouhov i Muthén, 2014; Muthén i Asparouhov, 2014).

Prezentacja obejmowała będzie przedstawienie metod badania porównywalności oraz wyniki badań symulacyjnych które objęły zarówno nowe jak i standardowe podejścia. Główne pytanie badawcze tego projektu odwołuje się do trzech kwestii: (i) na ile wykorzystywane w sondażowych badaniach porównawczych wskaźniki są porównywalne, (ii) na ile złamanie warunków porównywalności może prowadzić do błędów wnioskowania oraz (iii) na ile złamanie warunków porównywalności można kontrolować i kompensować dostępnymi metodami statystycznymi.

MAREK STYCZEŃ, INSTYTUT SOCJOLOGII UW

O potrzebie i możliwości łącznego analizowania dynamiki jednostek i zbiorowości.

Jest jakaś zmienna X nominalna. Dla uproszczenia może być dwuwartościowa: pracuje — tak/nie, głosuje — tak/nie. Wartość tej zmiennej przysługuje jednostkom tworzącym zbiorowość. Obserwujemy wartości tej zmiennej w czasie dyskretnym, codziennie, co miesiąc lub co pokolenie.

Na poziomie jednostki ciąg wartości tej zmiennej dla kolejnych momentów obserwacji możemy traktować jako realizacje pewnego procesu stochastycznego i opisać ten proces szacując jego parametry lub poprzestać na sumarycznym określeniu, w ilu momentach jednostka znajdowała się w każdym ze stanów.

Na poziomie zbiorowości możemy dla wybranego momentu stwierdzić, ile (lub jaka część) jednostek znajduje się w którym ze stanów, a dla kolejnych momentów wyznaczyć funkcje przedstawiające wartości tych parametrów w tych momentach.

Obserwując zbiorowość w dwu kolejnych momentach moglibyśmy badać przepływy jednostek między stanami. Rezultatem tych przepływów jest różnica między rozkładami zmiennej X w obu momentach.

Gdybyśmy wcześniej oszacowali parametry indywidualnych procesów, których realizacjami są indywidualne wartości zmiennej X , to znając indywidualne wartości zmiennej X w zbiorowości w momencie t moglibyśmy oszacować hipotetyczną tablicę prawdopodobieństw przepływów między momentami t i $t + 1$.

Jeżeli zaobserwowane przepływy, różnią się (w nie dający się przez przypadek wyjaśnić stopniu) od wyliczonych na podstawie oszacowanych wcześniej danych, to przyczyny tego będziemy szukać poza stabilnymi cechami jednostek w otoczeniu: np. zmieniła się liczba stanowisk pracy, lub wielkość zbiorowości potencjalnych pracowników.

Pełny model powinien obejmować parametry indywidualnych procesów, być może wspólne dla grup jednostek oraz parametry „otoczenia” i pozwalać na dekompozycję tablicy przepływów na ruchliwość wynikającą z natury indywidualnych procesów stochastycznych i ruchliwość stymulowaną (lub ograniczaną) przez „otoczenie”.

Projekt powiązania analizy jednostkowych procesów z przepływami między stanami w zbiorowości wydaje mi się wart zastanowienia.

Trójka moich studentów ucząc się korzystania z informacji zapisanej w relacyjnej bazie danych oszacowała dla każdego badanego (2000 absolwentów UW obserwowanych przez 5 lat) metodą najwyższej wiarygodności stałe i specyficzne dla niego prawdopodobieństwo tego, że będąc niezatrudnionym na etacie uzyska on etat w następnym miesiącu. Był to najprostszy wariant procesu ze stałym prawdopodobieństwem przejścia. Ze względu na brak czasu i uzasadnienia w „procesie dydaktycznym” nie szacowali prawdopodobieństwa przejścia w kierunku przeciwnym. Technicznie byłoby to zadanie identyczne.

Dysponując danymi o wszystkich absolwentach polskich uczelni z 2014 roku chciałbym zilustrować zaproponowane terminy porównując tablice przepływów i przeciętne indywidualnych prawdopodobieństwa odzyskania w następnym miesiącu pracy wcześniej utraconej i analogiczne prawdopodobieństwo utraty pracy posiadanej.

25 września

Sesja 2. Prowadzący: Henryk Domański

PAWEŁ B. SZTABIŃSKI, FRANEK SZTABIŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Czy zastosowanie wywiadu osobistego (face-to-face) w mixed mode design poprawia strukturę zrealizowanej próby?

Zgodnie z założeniami, zastosowanie *mixed mode design* ma na celu poprawę współczynnika realizacji i zmniejszenie błędu związanego z brakiem odpowiedzi w porównaniu z użyciem jednej tylko techniki, a jednocześnie redukcję kosztów surveyu (de Leeuw, 2005; de Leeuw, Hox, and Dillman 2008; Dillman, Smyth, and Christian 2014). Badania metodologiczne, zmierzające do doskonalenia tej procedury, dotyczą jednak na ogół łącznego użycia wyłącznie najtańszych technik badawczych, tzn. ankiety pocztowej i internetowej. Niewiele natomiast wiadomo o rezultatach zastosowania jako jednej z technik wywiadu osobistego (F2F). Tymczasem zdecydowana większość liczących się badań międzykrajowych, np. WVS, GSS, ESS, EVS, Eurobarometr, realizowana jest obecnie z użyciem F2F. Zastosowanie w nich, obok F2F, także innych, tańszych technik, mogłoby więc poprawić zrealizowaną próbę, a także obniżyć koszty surveyu.

W prezentacji podejmujemy dwa zagadnienia:

- (i) Czy użycie w *mixed mode* F2F, obok ankiet wypełnianych samodzielnie przez respondentów, umożliwi poprawę struktury próby zrealizowanej w porównaniu z F2F jako jedynej techniki?
- (ii) Czy bardziej efektywne jest użycie F2F w pierwszej fazie badania *mixed mode*, czy też w drugiej?

Zrealizowane w Polsce eksperymenty w związku z ESS 2012 i 2015 oraz POLPAN 2013 wskazują, że użycie ankiet samodzielnie wypełnianych drugiej fazie (z F2F w pierwszej) bardzo nieznacznie zwiększa współczynnik realizacji. W roku 2015 przeprowadzono eksperyment, w którym zastosowano kolejność odwrotną: w pierwszej fazie użyto ankiet (pocztowej i internetowej), a w drugiej F2F. W eksperymencie użyto kwestionariusza ESS 7. W rezultacie zastosowania F2F współczynnik realizacji znacznie zwiększył się, a struktura zrealizowanej próby znacznie uległa znaczącej poprawie w porównaniu z użyciem ankiet samodzielnie wypełnianych.

ADAM RYBAK, INSTYTUT SOCJOLOGII UAM

Analiza behawioralnych i deklaracyjnych preferencji respondentów: próba rozwiązania problemu doboru do techniki sondażu w schemacie mixed-mode

Trwający od dekad w Polsce i na świecie spadek wartości wskaźników realizacji badań sondażowych powoduje obniżenie jakości dokonywanego za ich pomocą pomiaru. Jedną z koncepcji, której wyjściowym celem była minimalizacja problemów wynikających z braków odpowiedzi są sondaże w schemacie mieszania technik.

Jednak najbardziej intuicyjny schemat realizacji sondażu *mixed-mode*, a więc taki w którym respondent ma możliwość samodzielnego wyboru techniki, w której będzie uczestniczył ma zasadniczą wadę – dotychczasowe metaanalizy metodologiczne sugerują, że takie podejście skutkuje zmniejszeniem odsetka zrealizowanych wywiadów w stosunku do badań narzucających technikę. Potencjalnym rozwiązaniem tego problemu wydaje się próba wcześniejszego ustalenia prawdopodobieństw pozytywnych odpowiedzi respondentów na prośby o udział w sondażu za pomocą różnych technik.

Punktem wyjścia wystąpienia chciałbym uczynić aktualny stan badań nad skutkami, jakie preferencje respondentów wywierają na wskaźnikach realizacji i na jakości pomiaru w badaniach sondażowych, a następnie przedstawić własne modele regresji logistycznej oparte na hipotezie istnienia względnie stałej preferencji techniki. Źródłem danych do ich opracowania jest eksperyment sondażowy *mixed-mode* przeprowadzony przez prof. Pawła Sztabińskiego w 2015 roku. Modele te służą ustaleniu prawdopodobieństw preferowania danej techniki dla kategorii respondentów wyróżnionych na podstawie danych dostępnych w operacie losowania do sondażu.

ILONA WYSMULEK, NIKA PALAGUTA, WERONIKA BORUC, MARTA KOŁCZYŃSKA, PRZEMEK POWAŁKO I DENYS LAVRYK, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Wyzwania harmonizacji danych w ponad 3500 badaniach sondażowych z całego świata: SDR Project 2.0

Referat dotyczy wyzwań związanych w przetwarzaniem i zarządzaniem wyjątkowo dużym i złożonym zbiorem danych sondażowych w projekcie Survey Data Recycling prowadzonym przez Kazimierza M. Słomczyńskiego, Irinę Tomescu-Dubrow i Craiga Jenkinsa, będącym kontynuacją wcześniejszego projektu harmonizacji danych sondażowych. Celem projektu SDR jest harmonizacja wybranych zmiennych z ponad 3500 badań sondażowych z 24 projektów realizowanych na przestrzeni lat 1966-2017, co stanowi czterokrotny wzrost ilości przetwarzania danych w porównaniu z pierwszym projektem harmonizacyjnym zespołu. SDR proponuje nowe podejście analityczne łączące harmonizację istniejących danych sondażowych, analizę jakości sondaży, oraz modelowanie wielopoziomowe z uwzględnieniem różnic metodologicznych między sondażami. W ten sposób SDR przyczynia się do wzmocnienia infrastruktury badawczej nauk społecznych nie tylko poprzez stworzenie zbioru zharmonizowanych danych sondażowych, ale też narzędzi do harmonizacji sondaży *ex-post*, oraz - ogólniej - poprzez wypracowanie odpowiedzi na wyzwania związane z pracą z dużymi zbiorami danych. Te wyzwania dotyczą zwłaszcza gromadzenia danych, zarządzania nimi i ich przetwarzania.

Doświadczenia pierwszego projektu harmonizacyjnego oraz zwiększona ilość danych w nowym projekcie SDR doprowadziły do wypracowania standardów, procedur, oraz narzędzi do automatyzacji przetwarzania danych i dokumentacji, jak również ułatwiających współpracę, ulepszających dokumentację projektu, oraz zapewniających replikowalność. Zestaw narzędzi wykorzystywanych w SDR obejmuje zarówno najbardziej popularne programy (np. szablony w MS Excel z makrami do tworzenia kodu do transformacji zmiennych, Onenote do optycznego rozpoznawania znaków), jak i otwarte oprogramowanie do wykonywania poszczególnych zadań (np. Cygwin jako platforma do automatyzacji powtarzalnych czynności, języki skryptowe do przetwarzania tekstu, MariaDB do zarządzania danymi). Ograniczone możliwości automatyzacji złożonych czynności pozostawiają pole do dalszych usprawnień.

Niniejszy referat skupia się na wyzwaniach pracy z danymi w projekcie SDR i proponowanych rozwiązaniach wypracowanych przez zespół projektowy. Wyzwania

obejmują: (i) konieczność zapewnienia przejrzystości procesu gromadzenia danych i dokumentacji, (ii) automatyzację złożonych czynności, (iii) zapewnianie replikowalności ręcznego przetwarzania danych i dokumentacji, (iv) równoległe wykonywanie wybranych ręcznych czynności. Tym referatem chcielibyśmy otworzyć dyskusję o narzędziach i metodach pracy z dużymi danymi, jak również na temat wyzwań piętrzących się wraz z ilością danych a związanych z transparentnością, dokumentacją, i replikowalnością.

PIOTR JABKOWSKI, INSTYTUT SOCJOLOGII UAM

O zróżnicowaniu praktyk sondażowych w międzykrajowych badaniach porównawczych na podstawie metodologicznej archiwizacji 1537 surveyów zrealizowanych w latach 1981-2017 w krajach europejskich

W ramach prezentacji omówione zostaną wybrane wyniki grantu Narodowego Centrum Nauki (2017/01/X/HS6/01304) pt. Porównawcza analiza jakości prób sondażowych w badaniach międzykrajowych z wykorzystaniem wewnętrznych i zewnętrznych kryteriów oceny: inwentaryzacja projektów oraz meta-baza danych wynikowych. Głównym celem grantu była metodologiczna archiwizacja ponad 1,5 tys. sondaży (zrealizowanych w latach 1981-2017 na reprezentatywnych próbach krajowych populacji w państwach europejskich) pochodzących z pięciu międzykrajowych projektów porównawczych o uznanej renomie naukowej (takich jak: Eurobarometer, European Quality of Life Survey, European Social Survey, European Values Study, International Social Survey Programme), jak również utworzenie zharmonizowanej meta-bazy danych wynikowych pozwalającej na ocenę wpływu zastosowania różnych procedur badawczych na jakość pomiaru sondażowego. Autor przedstawi między-projektowe, między-krajowe oraz wzdłuż-czasowe różnice w: (1) jakości uzyskiwanych wyników, jak również (2) stosowanych procedurach badawczych, w tym: (a) sposobach definiowania populacji docelowych; (b) typach operatów doboru próby, (c) typach prób, (d) schematach losowania jednostek, (e) procedurach doboru respondentów, (f) wykorzystanych technikach pozyskiwania danych, (g) procedurach terenowych, (h) procedurach kontrolnych.

Sesja 3. Prowadzący: Piotr Jabkowski

MAREK BOŻYKOWSKI^{1,2}, AGNIESZKA CHŁOŃ-DOMIŃCZAK^{3,2}, MIKOŁAJ JASIŃSKI^{1,2}, TOMASZ ZAJĄC^{1,2}

¹ Pracownia Ewaluacji Jakości Kształcenia (Uniwersytet Warszawski)

² OPI-BIP

³ Instytut Statystyki i Demografii SGH

Badania populacyjne PEJK UW. Najnowsze wyniki. Nowe obszary badań

Problematyka związków pomiędzy edukacją a rynkiem pracy jest wciąż bardzo aktualna zarówno w Polsce, jak i na świecie. W Polsce badaniu karier zawodowych absolwentów uczelni służy Ogólnopolski system monitorowania ekonomicznych losów absolwentów szkół wyższych (ELA), którego trzecia edycja została zrealizowana w tym roku.

W referacie przedstawimy najnowsze wyniki badań dotyczących zatrudnialności po studiach przeprowadzonych z wykorzystaniem danych systemu ELA. Omówimy zestawy modeli służących analizie czynników wpływających na zatrudnialność oraz wynagrodzenia po uzyskaniu dyplomu, przygotowane osobno dla poszczególnych

roczników studiów. Modele te wskazują na wyjątkową stabilność zjawisk mających miejsce w trakcie wchodzenia absolwentów na rynek pracy.

W trakcie wystąpienia chcielibyśmy także przedstawić kierunki rozwoju systemu ELA. W kolejnej edycji system zostanie znacznie rozszerzony. Poza absolwentami badaniem zostaną objęci studenci oraz doktoranci. Pozwoli to m.in. na zbadanie ścieżek edukacyjnych studentów, w tym np. badanie zjawiska odpadania i odchodzenia ze studiów, a także związków między studiami, pracą zawodową w trakcie nauki oraz późniejszymi karierami.

Dodatkowo system ELA przyczynił się do rozwoju badań wykorzystujących dane administracyjne. Rozwiązania metodologiczne, prawne i organizacyjne wykorzystane w ELA posłużyły jako wzór przy przygotowywaniu Zintegrowanej Platformy Analitycznej (ZPA), która ma służyć łączeniu informacji ze znacznie większej liczby źródeł. Jednym z zadań naszego zespołu w projekcie ZPA jest otwarcie dostępu do badań populacyjnych opartych na informacjach z rejestrów administracyjnych. W ostatniej części wystąpienia przedstawimy pokrótce nową platformę oraz planowane mechanizmy udostępniania zasobów i pokazać potencjał tego przedsięwzięcia dla badań społecznych.

MICHAŁ SITEK, INSTYTUT FILOZOFII I SOCJOLOGII PAN ORAZ INSTYTUT BADAŃ EDUKACYJNYCH

Wykorzystanie analizy sekwencji do opisu zmian w przebiegu życia. Zalety, wady i nowe techniki analizy i opisu sekwencji.

W wystąpieniu podsumuję najnowsze kierunki rozwoju analizy sekwencji — najczęściej wykorzystywanej techniki do opisu różnic i zmian karier edukacyjno-zawodowych i rodzinnych, ściśle związanej perspektywą przebiegu życia (*lifecourse*). Empiryczne ilustracje będą się opierać na danych dotyczących Polski, różnicach między płciami i zmianami w procesie przebiegu życia między wybranymi kohortami wieku. W wystąpieniu skupię się na konsekwencjach wyborów metodologicznych w analizie skupień oraz nowych rozwiązaniach metodologicznych wykorzystywanych w ostatnich latach w tego typu analizach w grupowaniu porównywaniu i wizualizacji różnic w sekwencjach.

KONRAD TUREK, NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE,
UNIwersytet Jagielloński

MAGDALENA JELONEK, UNIwersytet Ekonomiczny w Krakowie, CEIAPP UJ

BARBARA WOREK, UNIwersytet Jagielloński

Jak pytać o uczenie się dorosłych? Ocena wpływu konstrukcji pytań na wyniki na przykładzie danych z Bilansu Kapitału Ludzkiego w Polsce

Aktywność edukacyjna osób dorosłych jest jednym z ważniejszych obszarów zainteresowania polityk publicznych. Skuteczne projektowanie interwencji publicznych i ich późniejsza ewaluacja wymagają trafnych i rzetelnych wskaźników pozwalających na monitorowanie zmian w tym zakresie. Uczenie się przez całe życie (and. *lifelong learning*) stanowi jednak wymagający przedmiot badań ze względu na trudne do wyznaczenia granice i różne definicje kluczowych pojęć. Odmienne podejście w zakresie konceptualizacji i operacjonalizacji, np. pomiędzy krajami, lub nawet niewielkie zmiany w kolejnych falach badań mogą skutkować znaczącymi różnicami w wartościach

wskaźników. Może to prowadzić do nadinterpretacji, w których różnice wynikające z odmiennej metodologii pomiaru, są oceniane przez decydentów w kategoriach mniejszej lub większej skuteczności podejmowanych interwencji.

W naszym wystąpieniu, prezentujemy ocenę wpływu modyfikacji pytań kwestionariuszowych dotyczących uczestnictwa w uczeniu się przez całe życie na odpowiedzi osób badanych, a w efekcie — na wartość wyliczonych wskaźników uczestnictwa. Analizy opierają się na danych z Badania Ludności — Bilans Kapitału Ludzkiego (BKL I i BKL II). Wyniki pokazują jak nawet stosunkowo niewielkie zmiany w konstrukcji pytań mogą prowadzić do znacznie odmiennych reakcji respondentów, zróżnicowanych dodatkowo ze względu na ich wykształcenie lub wiek.

ANNA KIERSZTYN, INSTYTUT SOCJOLOGII UW

Jak uchwycić prekaryjne zatrudnienie w badaniach ilościowych? Znaczenie danych o przebiegu karier i przykład Polskiego Badania Panelowego POLPAN.

W ostatnich latach widoczny jest wzrost zainteresowania socjologów i ekonomistów problematyką niepewności na rynku pracy, co ma związek m.in. z rosnącą popularnością tzw. niestandardowych czy elastycznych form zatrudnienia. Powyższa zmiana niejednokrotnie była interpretowana w kategoriach przemian wzorów kariery zawodowej, polegających na zastąpieniu modelu stabilnego zatrudnienia na pełnym etacie, zapewniającego bezpieczeństwo ekonomiczne pracownikom i ich rodzinom, modelem powtarzających się epizodów tymczasowego zatrudnienia przeplatanych okresami bezrobocia. Zwracano również uwagę, iż przemiany w dziedzinie zatrudnienia mogą w różnym stopniu dotyczyć różne kategorie pracowników, prowadząc do dualizacji rynku pracy. Powyższe obawy znalazły swój wyraz w rozważaniach Guya Standinga dotyczących przemian struktury klasowej i wyłaniającego się w niej prekariatu. Do tej kategorii Standing zalicza osoby zatrudnione na tymczasowych kontraktach, o niskich dochodach i bez dostępu do systemu świadczeń socjalnych, dotknięte niepewnością ekonomiczną i poczuciem tymczasowości, a także pozbawione tożsamości związanej z wykonywaną pracą.

Mimo dość powszechnych wskazań na wagę problemu prekaryjnego zatrudnienia i jego potencjalnie dalekosiężne skutki (społeczne, demograficzne, czy polityczne), badacze wciąż napotykają trudności związane z definicją i operacjonalizacją tego zjawiska. Celem prezentacji jest analiza tych trudności oraz przedstawienie rozwiązań przyjętych w najnowszej fali Polskiego Badania Panelowego POLPAN 2018.

W dotychczasowej praktyce badawczej najczęściej stosowane są dwa podejścia do analiz niepewności na rynku pracy: pierwsze za kryterium prekaryjnego zatrudnienia przyjmuje rodzaj umowy, na podstawie której świadczona jest praca, drugie zaś subiektywnie odczuwaną niepewność związaną z możliwością utraty zatrudnienia i trudnościami w znalezieniu nowego. Oba podejścia są niezadowolające. W przypadku pierwszego problematyczny jest fakt, iż pomija ono kwestię zróżnicowań wewnątrz kategorii niestandardowo zatrudnionych pracowników: obok prekariuszy, dużą część grupy stanowią mogą pracujący na kontraktach wysoko wykwalifikowani i dobrze zarabiający specjaliści, a także osoby zatrudnione na okres próbny poprzedzający stabilne zatrudnienie. Na adekwatność drugiego podejścia rzutuje z kolei fakt, iż subiektywnie odczuwane poczucie bezpieczeństwa nie zawsze trafnie odzwierciedla rzeczywiste perspektywy danej jednostki na rynku pracy. Wpływ na to mają m.in. takie czynniki, jak: porównania z grupą odniesienia, czy też psychologiczne mechanizmy adaptacji do

trudności życiowych, skłaniające jednostki do „normalizacji” prekaryjnej sytuacji zawodowej.

W świetle powyższych wątpliwości, informacje o sytuacji pracy bądź odczuwanej niepewności w jednym punkcie czasowym trudności nie pozwalają ocenić, czy dana jednostka rzeczywiście doświadcza prekaryjnego zatrudnienia. Jest to możliwe dopiero poprzez całościowe badanie losów zawodowych osób, które pracowały na tymczasowych umowach bądź doświadczały lęku przed bezrobociem. Dlatego też we współczesnych analizach rynku pracy coraz większą wagę przykładana się do danych panelowych, pozwalających na analizę trajektorii jednostek w celu identyfikacji wzorów karier nacechowanych przedłużającą się niepewnością i brakiem stabilizacji. Szczególnie cenne są w tym przypadku badania, w ramach których rejestruje się przebieg historii zawodowej na przestrzeni wielu lat, począwszy od pierwszej pracy. Zarazem konieczność wiernego i szczegółowego odtworzenia sekwencji prac zawodowych w warunkach coraz większego różnicowania się form zatrudnienia, czasu i miejsca wykonywania pracy, a także wzorów karier zawodowych, rodzi nowe wyzwania dla twórców narzędzi badawczych mających umożliwić analizę niepewności na rynku pracy na dużych próbach respondentów.

Sesja 4. Prowadzący: Zbigniew Sawiński

PIOTR KOCYBA, UNIWERSYTET TECHNICZNY W CHEMNITZ

Metodologia ankietowania skrajnie prawicowych protestów - doświadczenia z drezdeńskiej Pegidy

PEGIDA (Patriotyczni Europejczycy Przeciwko Islamizacji Zachodu) to najbardziej dziś znany przykład nowej formy ruchów skrajnej prawicy. Jego program to niechęć do imigracji (szczególnie z krajów Bliskiego Wschodu i Afryki), reprezentatywnej demokracji oraz jej kultury politycznej. Ogromnym sukcesem PEGIDY była mobilizacja uliczna. W latach 2015 i 2016 nawet 25 tysięcy osób co tydzień demonstrowało w starówce Drezna. Regularne protesty w stolicy Saksonii trwają zresztą do dziś. Zachowanie oraz liczba demonstrantów skutkowało zainteresowaniem medialnym, ale i naukowym: demonstracje w sumie 11 razy badało pięć różnych zespołów badawczych. Celem referatu jest wprowadzenie w *research design* poszczególnych zespołów, opisanie specyfiki ankietowania demonstracji skrajnej prawicy oraz możliwe konsekwencje metodologiczne.

KATARZYNA MALINOWSKA, PAULINA GÓRSKA, CENTRUM BADAŃ NAD UPRZEDZENIAMI UW

KATARZYNA LIPOWSKA, WYDZIAŁ PSYCHOLOGII UW

Honor, Ojczyzna! — czyli o badaniu polskich demonstracji

Nasz zespół od kilku lat prowadzi badania demonstracji w Polsce, zaczynając od Marszu Niepodległości 2015r. aż po Paradę Równości 2018r. Próbujemy dotrzeć nie tylko do uczestników i uczestniczek, ale też do osób, które z różnych powodów nie biorą udziału w zbiorowym działaniu.

O ile wyjście poza klasyczne paradygmaty i badania laboratoryjne umożliwia poznanie czegoś więcej niż tendencji behawioralnych, badania terenowe stanowią duże

wyzwanie metodologiczne i wiążą się z brakiem możliwości kontrolowania wszystkich pożądaných czynników.

W trakcie naszego wystąpienia chcielibyśmy podzielić się zarówno dobrymi praktykami jak i trudnościami, które napotkaliśmy. Zamierzamy omówić poszczególne procedury przeprowadzanych badań i skutki podejmowanych przez nas decyzji. Chcielibyśmy zaprosić uczestników konferencji do dyskusji dotyczącej m.in. przeszkolenia i doboru ankietów, sposobu rekrutacji uczestników, konstrukcji kwestionariuszy, czy zasadności nawiązywania kontaktu i podejmowania współpracy z organizatorami demonstracji.

TOMASZ ŻÓŁTAK, INSTYTUT BADAŃ EDUKACYJNYCH

ADAM GENDŹWIŁŁ, WYDZIAŁ GEOGRAFII I STUDIÓW REGIONALNYCH UW

Podejście kontrfaktyczne w ocenie skutków wprowadzenia parytetów płci na listach wyborczych w Polsce

W ostatnich dekadach w wielu krajach na świecie, w tym również w Polsce, wprowadzone zostały regulacje prawne nakazujące stosowanie parytetów płci na listach wyborczych. W proporcjonalnych systemach wyborczych jest to jedno z najczęściej stosowanych rozwiązań mających na celu zmniejszenie dysproporcji między udziałem mężczyzn i kobiet w ciałach przedstawicielskich. Wzrost reprezentacji kobiet wśród wybranych reprezentantów w wyniku wprowadzenia parytetów co do zasady jest jednak znacząco mniejszy niż wzrost ich udziału wśród kandydatów (tak też stało się w Polsce po wprowadzeniu parytetów w 2011 r.). Otwiera to pole do dyskusji nad tym, czy parytety są efektywnym narzędziem zwiększania udziału kobiet w polityce i jakie czynniki mają decydujące znaczenie dla ograniczania skuteczności ich oddziaływania.

Wcześniejsze badania i analizy pozwalają wnioskować, że w polskich wyborach do Sejmu, przeprowadzanych w systemie proporcjonalnym z otwartą listą, kluczowe znaczenie miały trzy zjawiska. Po pierwsze – trudności z zapewnieniem odpowiedniej liczby kobiet-kandydatów posiadających podobne do mężczyzn wcześniejsze doświadczenie polityczne (będące jedną z najważniejszych cech wpływających na popularność kandydata). Po drugie – zmiana strategii komitetów wyborczych w zakresie ustalania kolejności kandydatów na listach (przejście na schematy mniej korzystne w stosunku do kandydatek). Po trzecie – zmiana preferencji wyborców dla kandydatów określonej płci. O ile w literaturze identyfikowano potencjalne źródła ograniczonej skuteczności parytetów, o tyle zasadnicze wyzwanie stanowiło oszacowanie, jak duże jest znaczenie każdego ze wspomnianych zjawisk.

W metodach stosowanych dotychczas do analizy efektu wprowadzenia parytetów płci nie było to możliwe. Badania odwołujące się do modelowania przeciętnych szans kandydatów na uzyskanie mandatu (czy obsadzenie na wysokiej pozycji na liście wyborczej) pozwalały co prawda stwierdzić, czy sposoby głosowania (a także sposoby ustalania kolejności kandydatów na listach przez komitety wyborcze) zmieniły się w wyniku wprowadzania parytetów przeciętnie na korzyść, czy też na niekorzyść kobiet, przy kontroli zmiany rozkładu innych cech kandydatów, decydujących o ich popularności wśród wyborców. Nie pozwalały jednak wyrazić, w jakim stopniu takie zmiany przekładały się na udział kobiet w gronie wybranych reprezentantów. Nie dawały też szansy na porównanie ich znaczenia z innymi czynnikami – przede wszystkim zmianą rozkładu cech kandydatów.

Aby przewyciężyć to analityczne ograniczenie, proponujemy w naszym wystąpieniu nową metodę analizy skutków wprowadzenia parytetów, bazującą na podejściu kontrfaktycznym. Modele opisujące indywidualne szanse uzyskania mandatu przed wprowadzeniem parytetu płci wykorzystywane są w niej do obliczenia przewidywanego prawdopodobieństwa bycia wybranym przez kandydatów startujących w wyborach, w których obowiązywały parytety, a następnie – przewidywanego udziału kobiet wśród reprezentantów wybranych w tych wyborach. Odpowiedni dobór zestawu uwzględnionych w modelu zmiennych kontrolnych pozwala na oszacowanie hipotetycznych udziałów kobiet w różnych sytuacjach, a w konsekwencji na ocenę znaczenia poszczególnych wymienionych wcześniej czynników na reprezentację płci.

Działanie metody zaprezentujemy w odniesieniu do wyborów do polskiego Sejmu z lat 2005-2015. Uzyskane przez nas wyniki wskazują, że decydujące znaczenie w ograniczeniu oddziaływania parytetów na zwiększenie reprezentacji kobiet, zwłaszcza bezpośrednio po ich wprowadzeniu, miała niemożliwość zapewnienia odpowiedniej liczby kobiet-kandydatów posiadających istotne wcześniejsze doświadczenie polityczne. Znaczenie tego czynnika zmalało jednak w kolejnych wyborach, podczas gdy wzmocnieniu uległy w nich niekorzystne względem kobiet wzorce kolejności kandydatów na listach wyborczych.

JACEK HAMAN, INSTYTUT SOCJOLOGII UW

Badanie społeczne – badanie obywatelskie: czy badania mogą przywrócić wiarę w wybory?

Kryzys związany w wyborami samorządowymi 2014 – załamanie się systemu informatycznego i opóźnienia w ogłoszeniu oficjalnego wyniku, wysoka liczba głosów nieważnych, zaskakujący w świetle wyników exit-pollu wynik PSL – wywołał reakcję w postaci formułowania przez niektóre środowiska oskarżeń o masowe fałszerstwa wyborcze nie tylko w tej, ale również w innych polskich elekcjach. Oskarżenia te formułowane były głównie przez działaczy związanych z Ruchem (ruchami) Kontroli Wyborów, ale powtarzane są także przez polityków związanych z obecnym obozem rządzącym. Z kolei kryzys związany z przeforsowanymi przez PiS zmianami w Kodeksie Wyborczym, wprowadzanymi na niespełna rok przed kolejnymi wyborami na przełomie 2017/2018 roku, wywołał zarzuty zamiaru fałszowania wyborów ze strony aktualnej opozycji – tym samym fala nieufności względem procesu wyborczego rozlała się już na wszystkie strony sceny politycznej.

Wiarygodność wyborów jest czynnikiem kluczowym dla funkcjonowania demokracji. Ważne jest przy tym zarówno to, czy wybory obiektywnie są wiarygodne, jak i to, czy są jako wiarygodne postrzegane. Jeśli ich wiarygodność jest kwestionowana (czy post, czy ante factum), należy te zarzuty sprawdzić. Jeśli zarzuty się potwierdzą, należy działać w celu przywrócenia praworządności, ale jeśli się one nie potwierdzą, istotne jest, aby przekonać o ich niezasadności opinię publiczną.

Ocena wiarygodności wyborów nie jest oczywiście zadaniem, które może wykonać całościowo socjolog czy statystyk, jednakże metody ilościowych badań społecznych i statystycznej analizy danych mogą się okazać bardzo przydatne przy projektowaniu systematycznej obserwacji przebiegu głosowania oraz ocenie i interpretacji zebranych w niej danych. Jednocześnie świadomość konsekwencji faktu, że zarówno proces przygotowania i przeprowadzenia wyborów, jak i sam proces ich obserwacji i badania, same w sobie są „zjawiskami społecznymi” (a nie tylko „prawnymi” czy „politycznymi”) pozwala na lepsze zaprojektowanie działań badawczych.

W referacie przedstawię możliwe schematy badania reprezentacyjnego, mogącego pozwolić ocenić prawidłowość przynajmniej pewnych aspektów procesu wyborczego w 2018 roku, a jednocześnie – poprzez zaangażowanie w proces badawczy przedstawicieli różnorodnych środowisk – mającego szansę na większą „moc przekonywania” niż badanie wyłącznie „akademickie”. Czy i w jakim wariantcie badanie takie w czasie wyborów samorządowych w 2018 roku zostanie przeprowadzone – w chwili zgłaszania tego referatu (przełom czerwca i lipca) trudno jeszcze ocenić; w czasie jego wygłoszenia (koniec września) zapewne będzie już to przynajmniej wstępnie wiadome; obecnie w ramach programu „Masz Głos - Masz Wybór” Fundacji Batorego trwają prace koncepcyjne oraz rozmowy z potencjalnymi partnerami społecznymi (co, skądinąd, wymusza rozpatrywanie różnych podejść i rozwiązań metodologicznych, tak ze względu na konieczność „mierzenia zamiarów na siły”, jak i kompromisu z oczekiwaniami partnerów).

Drugą kwestią poruszoną w referacie będą statystyczne metody oceny wiarygodności wyników wyborów.

SŁAWOMIR BARTNICKI, MACIEJ ALIMOWSKI, INSTYTUT SOCJOLOGII I KOGNITYWISTYKI UWB

Możliwości wykorzystania regresji nieciągłej (Regression Discontinuity Design) w analizach zjawiska incumbency advantage w wyborach bezpośrednich do gminnej egzekutywy w Polsce.

Obserwacje ($N = 33680$) z dotychczasowych startów w wyborach bezpośrednich do stanowisk gminnej egzekutywy w Polsce zostaną wykorzystane do analizy efektu inkumbenta (*incumbency advantage*), przy pomocy RDD. W tym celu charakterystyce podlegać będzie specyfika rozkładu obserwacji (brak koncentracji wokół punktu cięcia), przy przyjętym schemacie kodowania danych, (konieczność uwzględnienia II tury). W analizie dyskusji poddane zostaną wyniki testów kwalifikujących stosowanie RDD uwzględniające randomizację cech obserwacji i ich liczebności wokół punktu podziału, (np.: testy podobieństwa cech obserwacji czy test gęstości McCrary (McCrary 2008)). Randomizacja cech wśród ledwie przegrywających i ledwie wygrywających wybory proponowana przez Lee (Lee 2008), została podważona w zakresie analiz tzw. bliskich wyborów (Snyder 2005; Grimmer i in. 2011; Caughey, Sekhon 2011) i podlega dyskusji (Eggers i in. 2015). Nadal oczekuje się jednak zbliżonej liczby obserwacji wokół punktu cięcia (Cuesta, Imai 2016). W analizie ukazane będą kwestie potencjalnie problematyczne związane z dysproporcjonalnością liczebności i weryfikacja wpływu kwantyfikowalnych cech o teoretycznym wpływie na sortowanie danych. Następnie, przedstawione zostaną propozycje modeli nieparametrycznych RDD, które wykażą zwykle określane przewagi inkumbentów, jak siłę (*vote share*) czy przewagę nad najsilniejszym rywalem (*vote margin*). Poruszona będzie też kwestia prawidłowego wnioskowania modeli nieparametrycznych i przypadków, w których wnioskowanie to może być wadliwe (Hyytinen i in. 2018). Ostatecznie, zaprezentowana zostanie propozycja szacowania prawdopodobieństw kandydowania i wygrywania w kolejnych wyborach.

Podsumowując, w wystąpieniu można wyodrębnić trzy kwestie do dyskusji: 1) kodowanie danych w schemacie na kandydata; 2) wyniki testów walidacyjnych; i 3) schemat(y) estymowania szans kandydowania i wygrywania w kolejnych wyborach.

Referat w ramach projektu NCN (2014/13/D/HS5/02010), pt.: *Uwarunkowania reelekcji egzekutywy gminnej po 2002 roku w Polsce.*

26 września

Sesja 5. Prowadzący: Zbigniew Karpiński

ZBIGNIEW SAWIŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Spójność poglądów na temat zasad społecznej stratyfikacji: empiryczne argumenty a artefakty badawcze

Analizując długofalowe zmiany w postawach społeczeństwa nie sposób abstrahować od równoczesnych zmian, jakie dokonują się w podejściu respondentów do badań sondażowych. Wzrastająca niechęć do udziału w badaniach prowadzi nie tylko do malejących wskaźników realizacji próby (*response rate*), lecz również objawia się rosnącą tendencją do udzielania przez respondentów zdawkowych i nieprzemyślanych odpowiedzi. Owa tendencja wypaczać może obraz badanych zjawisk. W Polsce w latach 1958-2008 zrealizowano serię badań nad prestiżem zawodów. Kolejne badania potwierdzały, że respondenci w coraz mniejszym stopniu są zgodni co do prestiżu przynależnego wykonawcom różnych ról zawodowych. Na tej podstawie sformułowano hipotezę o malejącym znaczeniu norm uwarstwienia społecznego, które to normy w teoriach funkcjonalnych uznaje się za podstawę legitymizacji systemu (Domański, Sawiński i Słomczyński, 2010). Zwrócono przy tym uwagę, że uzyskane rezultaty stanowią argument na rzecz proponowanej hipotezy jedynie pod warunkiem, że nie są efektem wzrastającej skłonności do udzielania powierzchownych odpowiedzi, co samo w sobie prowadziłoby do pozornego osłabienia spójności poglądów badanych osób.

W wystąpieniu sformułowano tezę, że nawet jeśli w dłuższych okresach spada zaangażowanie respondentów i rośnie odsetek nieprzemyślanych odpowiedzi, to owa tendencja nie jest na tyle silna, aby prowadzić do pozornego spadku spójności poglądów społeczeństwa. Do analiz wybrano dwie sfery poglądów. Pierwsza dotyczyła głównych osi konfliktów w społeczeństwie, zaś druga czynników, od których zależy sukces w życiu. Do pomiaru spójności poglądów zaproponowano dwa wskaźniki: (a) odsetek wariancji ocen respondentów wyjaśniony przez zestaw wymiarów podlegających ocenie; (b) przeciętną korelację ocen każdej osoby z ocenami pozostałych osób. W analizach wykorzystano dane z badania panelowego POLPAN, porównując wyniki pierwszej fali badania (1988), z wynikami uzyskanymi 25 lat później (2013). W przypadku sfery konfliktów obie miary spójności okazały się znacząco wyższe w 2013 roku, natomiast spójność poglądów na temat czynników decydujących o sukcesie kształtowała się podobnie w obu okresach. Nie stwierdzono więc podstaw do odrzucenia hipotezy, w myśl której wzrastająca tendencja do udzielania zdawkowych i nieprzemyślanych odpowiedzi nie musi prowadzić do artefaktów badawczych, polegających na pozornym obniżaniu się społecznego konsensusu.

Z przeprowadzonej analizy płyną implikacje zarówno dla metodologii badań sondażowych, jak również dla teoretycznych rozważań nad strukturą społeczną. W sferze metodologii stwierdzono, że w badaniu przeprowadzonym z odpowiednią starannością oraz dbałością o jak najwyższy odsetek realizacji próby, zniekształcenia związane z niskim zaangażowaniem respondentów i udzielaniem nieprzemyślanych odpowiedzi nie muszą być na tyle silne, aby wypaczyć obraz badanych zjawisk. Natomiast dla ustaleń teoretycznych największe znaczenie ma wniosek dotyczący rosnącej roli konfliktów jako płaszczyzny postrzegania zróżnicowania społecznego. Stwarza to nowe wyzwania w dyskusjach między zwolennikami funkcjonalnych i konfliktowych teorii społeczeństwa.

SŁAWOMIR PASIKOWSKI, WYDZIAŁ NAUK O WYCHOWANIU UŁ

Skale szacunkowe i algebraiczne formuły integracji danych jako czynniki trafności pomiaru cech modelowanych w strukturze opozycyjnych par

Pomiar psychologiczny oparty na operacjonalizacji przewidującej opozycyjną strukturę mierzonej cechy wymaga zastosowania rozwiązań wrażliwych na stosunki pomiędzy przeciwnymi jej składnikami, a których wypadkową jest określona gotowość lub kierunek zachowania podmiotu. Wśród najczęściej stosowanych znajdują się odpowiednio konstruowane skale szacunkowe oraz formuły algebraiczne. Te ostatnie służą do integracji danych pochodzących z tych skal i odtwarzania naprężenia zachodzącego pomiędzy składnikami cechy.

O ile dyskusja i badania nad optymalizowaniem skal szacunkowych nie ustają w środowisku badaczy (problem skal bipolarnych vs unipolarnych, skale typu GRID, itp.), o tyle kwestia formuł, które odwzorowują przewidywane teorią naprężenia między składnikami cech, wydaje się nie budzić już większych wątpliwości. Upowszechnione rozwiązania algebraiczne posiadają jednak rzadko dostrzegane wady obniżające trafność odwzorowań.

W związku z tym w toku wystąpienia dokonany zostanie skrócony przegląd formuł przywoływanych w literaturze przedmiotu wraz ze wskazaniem ich zalet oraz grupy kluczowych zakłóceń, na które wrażliwe są przeliczenia realizowane za ich pomocą. Następnie przedstawione zostanie autorskie rozwiązanie wychodzące naprzeciw założeniom stawianym tego rodzaju formułom. W prezentacji wykorzystane zostaną wyniki badań symulacyjnych polegających na poszukiwaniu funkcji matematycznych optymalnie opisujących generowane dane i pomagających charakteryzować słabe i mocne strony zaprezentowanych formuł algebraicznych. Niektóre z własności tych formuł ujawniają się dopiero po zastosowaniu zaproponowanego postępowania badawczego.

MAREK MUSZYŃSKI, INSTYTUT SOCJOLOGII UJ, EVIDENCE INSTITUTE

PAWEŁ GRYGIEL, INSTYTUT PEDAGOGIKI UJ

GRZEGORZ HUMENNY, INSTYTUT SOCJOLOGII UJ

KAROLINA ŚWIST, AISECLAB

Odpowiadanie nieuważne w badaniach samoopisowych - porównanie i walidacja metod identyfikacji.

Odpowiadanie nieuważne najczęściej definiuje się jako odpowiadanie w sposób losowy lub bez pełnego zaangażowania (Curran, 2015). Prowadzi ono do zwiększenia losowego błędu pomiaru, co obniża jakość danych i może skutkować błędami wnioskowania (I i II rodzaju) (Maniaci i Rogge, 2014). Może także powodować błędy w oszacowaniu miar dopasowania modelu do danych (Mansolf & Reise, 2018). Szacuje się, że w badaniach niskiej stawki odsetek odpowiadających losowo może sięgać nawet 50%, choć zwykle oscyluje wokół 10% (Maniaci i Rogge, 2014; Oppenheimer, Meyvis i Davidenko, 2009).

Problemowi obciążenia odpowiadaniem losowym wypada poświęcić szczególną uwagę nie tylko ze względów teoretyczno-metodologicznych, lecz także praktycznych. Wyniki badań wykorzystujących narzędzia samoopisowe niskiej stawki (szczególnie

zagrożone odpowiedziami nieuważnymi/losowymi; Curran, 2015), np. Bilans Kapitału Ludzkiego, PIAAC, PISA, Diagnoza Społeczna, wpływają wszak na kształtowanie polityk publicznych.

Metody identyfikacji badanych odpowiadających nieuważnie/losowo generalnie podzielić można na dwa podstawowe rodzaje: (1) wymagające zaplanowania przed badaniem (metody a priori); (2) możliwe do użycia już po zebraniu danych (metody post hoc). Pierwsza grupa metod obejmuje m.in. sprawdziany uważności badanych za pomocą itemów-pułapek (*instructional manipulation checks*; Oppenheimer i in., 2009). Do drugiej z kolei zaliczyć można m.in. miary: (a) badające spójność indywidualnych danych (np. rzetelność połówkowa skali, korelacje między pytaniami „synonimami” i „antonimami”); (b) szukające przypadków odstających (np. miary dopasowania osób — *person fit statistics*, bazujące na odchyleniu się od „wzorca odpowiedzi”), miary wielowymiarowych odległości); (c) dzielące badanych na grupy (np. analiza klas latentnych, czynnikowe modele mieszane (factor mixture models), modele mieszanek IRT (mixture IRT models) (Meade i Craig, 2012).

W ramach prezentacji (oprócz przedstawienia założeń teoretycznych i metodologiczno-statystycznych) przedstawione zostanie porównanie skuteczności wybranych metod identyfikacji odpowiedzi nieuważnych/losowych (a priori i post hoc) na jednym zbiorze danych, pochodzącym z ogólnopolskiego badania kompetencji uczniów ($N > 25$ tys.). Analizy koncentrować się będą na skali Lęku przed Matematyką (AMAS) w polskiej adaptacji Krzysztofa Cipory i współpracowników (Cipora i in., 2015). Narzędzie posiada sprawdzone (i dobre) właściwości psychometryczne, potwierdzone także dla wersji komputerowej (Cipora i in., 2018). W analizie wykorzystane zostaną także pozycje-pułapki oraz wyniki kognitywnego testu kompetencji matematycznych. Niezbędne estymacje przeprowadzone zostaną z wykorzystaniem pakietu Mplus i R.

ANDRZEJ JAN KUTYŁOWSKI, UNIWERSYTET W OSLO

Ludzkie wartości według Schwartza: przegląd, krytyka metodyczna i alternatywne podejście do modelowania danych

Szereg autorów skupiło swoją uwagę na wyszczególnieniu i pomiarze stabilnych cech osobowości ludzkiej, często określanych jako „wartości”. W proponowanym wystąpieniu dokonam przeglądu tego rodzaju podejścia autorstwa S. Schwartza, które zyskało wielką popularność wśród badaczy w socjologii i psychologii wyrażającą się, m. inn. umieszczeniem pewnej wersji kwestionariusza tego autora w Europejskim Sondażu Społecznym.

Z metodycznego punktu widzenia podkreślona zostanie cecha tego podejścia opierająca się na konsekwentnym, chociaż *explicite* nie podkreślanym i nie interpretowanym założeniu, że kategorie odpowiedzi na pytania tego kwestionariusza są traktowane jako wartości zmiennych na poziomie interwałowym, nie zaś porządkowym. Założenie to bezpodstawnie przyjęte jest także przez szereg innych badaczy tej problematyki i stanowi podstawową wadę opublikowanych dotychczas analiz bogatego materiału empirycznego zebranego z inspiracji rozważań tego autora.

Pierwszą częścią alternatywnych propozycji stanowić będzie specyfikacja prawidłowo rozumianej porządkowości zmiennych sformułowanych przez Schwartza, ze szczególnym uwzględnieniem wymiarowości i subiektywności danych.

W drugiej, głównej części przedstawione zostaną niektóre alternatywy, odnoszące się do teorii i praktyki modelowania graficznego (*graphical models*), w szczególności sieci Markova i ich zmodyfikowanej semi-parametrycznej wersji rozwiniętej przez S. Kreinera,

oraz ich rozszerzeń. Zademonstrowany zostanie także sposób testowania podstawowego aspektu teorii Schwartza odnoszącego się do kołowej konfiguracji podstawowych „wartości”, z wykorzystaniem modelu z ukryta reakcja sformułowanego przez J. Ramseya oraz jego rozszerzeniami zaproponowanymi przez autora. Przedstawione zostaną przykłady modelowania danych w ramach tych podejść z wykorzystaniem narzędzi badawczych Digram (S. Kreiner) i MIM (D. Edwards).

Część końcowa dotyczyć będzie zmian perspektywy, tak aby zapobiec w przyszłości kontynuacji analiz danych opartych na nieuprawnionej interpretacji zmiennych porządkowych jako zmiennych przedziałowych.

MARTA BOŻEWICZ, INSTYTUT SOCJOLOGII UW

Sondaże religijności Polaków – elaboracja nowej metodologii

Problematyka mierzenia tego, co w pewien sposób niemierzalne i uchwycenia wiary jako jednorazowego aktu objawiającego się w przekonaniach lub zachowaniach obecna jest w polskiej socjologii od lat 80. (Grabowska, Szawiel 2005). Badania socjologiczne nad religijnością w Polsce wymagają – więcej niż dotychczas – wsparcia metodologicznego, które wniosłoby większą precyzję i rzetelność w stosowanych narzędziach badawczych. O przemianach w obszarze religijności społeczeństwa polskiego świadczą liczne badania o charakterze jakościowym, studia przypadków i sondaże o mniejszym zasięgu (Baniak 2001, Boguszewski 2008, Grabowska, Szawiel 2005, Jasińska-Kania, Marody 2002, Marody 1996, Zdaniewicz, Zembrzuski 2000, Zaręba 2003 i in.). Na ich podstawie w literaturze przedmiotu formułuje się wnioski o selektywności wiary, jej prywatyzacji czy indywidualizacji. Socjologowie zwracają jednak uwagę na brak całościowych analiz i kompleksowego badania nad treścią wiary współczesnych Polaków (CBOS 2012; Grabowska, Szawiel 2005; Mariański, Zaręba 2002). Metodologia w zakresie socjologii religii jest w Polsce bardzo rzadko podejmowana jako przedmiot badań, a jednocześnie religijność jako zmienna wyjaśniająca jest chętnie używana do analiz bardzo różnorodnych zjawisk. Wskaźniki, takie jak autodeklaracja wiary lub częstość uczestniczenia w nabożeństwach stają się jednak coraz bardziej niewystarczające.

Badania (Kamińska 1993, Grotowska 1999, Libiszowska-Żółtkowska 2010), prowadzone m.in. przez zespół CBOS w ramach projektu Metoda i znaczenie (2012), czy przeze mnie (2014) ukazują, że do rozumienia zawartości autodeklaracji religijnych niezbędne jest poznanie języka samych wierzących i niewierzących. W tym celu pytania dotyczące religijności z międzynarodowych i ogólnopolskich sondaży (ISKK, CBOS, ESS, EVS, WVS) poddałam badaniom kognitywnym, a więc wywiadam z przedstawicielami różnych grup społecznych, aby przyjrzeć się ich sposobom rozumienia pytań i dochodzenia do odpowiedzi. Celem mojego badania było dogłębne rozpoznanie i szczegółowe opisanie języka mówienia o religijności i obszarów jej ujawniania się.

Zgodnie z uzyskaną wiedzą zaproponowałam nowe wersje pytań, zbudowanych w oparciu o głęboko eksplorowany w poprzednim etapie język doświadczenia religijnego i poddawane kolejnym wywiadam kognitywnym. Dostosowywanie kwestionariusza odbywało się także w konsultacjach z socjologami, specjalistami w zakresie metodologii, oraz teologami. „Nie przypadkiem układanie pytań nazywano sztuką albo rzemiosłem – nigdy teorią (Sułek 2011: 172)”. Tak przygotowane narzędzie zostało zademonstrowane w ogólnopolskim badaniu ilościowym (realizowanym przez CBOS), które umożliwiło również porównanie wyników dotychczas wykorzystywanego kwestionariusza badania religijności i nowej, eksperymentalnej metodologii. Wyniki ogólnopolskiego sondażu

stanowią demonstrację możliwości formułowania bardziej trafnych i rzetelnych pytań o religijność.

Nowatorskość przyjętego przeze mnie podejścia polega na tym, że wykorzystując dotychczasową wiedzę (a szczególnie metodologię i wyniki reprezentatywnych badań) projekt „wychodzi od dołu”. To znaczy, poczynając od dogłębnego rozpoznania sposobów rozumienia własnej religijności i wiary, eksploruje pojęcia i interpretacje badanych. Centralnym jego punktem jest aktywność w testowaniu języka. Następnie definiuje kluczowe kategorie i dąży do zbadania tego, co badani rozumieją pod pojęciami, pytaniami i udzielanymi odpowiedziami. W efekcie tego podejścia zaproponowana została nowa metodologia badania religijności, której język pytań do respondenta jest maksymalnie zbliżony do naturalnego. Ponieważ do tej pory tak dokładnie nie przyglądano się kwestionariuszowemu językowi dotyczącemu religijności i nie podejmowano badań kognitywnych tego obszaru, moja praca stanowi demonstrację nowych możliwości formułowania bardziej trafnych i rzetelnych pytań sondażowych. W referacie przedstawię najważniejsze wnioski rozprawy doktorskiej (planowany termin obrony: wrzesień 2018).