Jolanta Perek-Białas

Uniwersytet Jagielloński
Szkoła Główna Handlowa[1]

Tradycyjne i uzupełniające podejścia do ewaluacji efektów polityki senioralnej w Polsce[2]

Streszczenie

Celem artykułu jest pokazanie, w jaki sposób można oceniać efekty działań, które składają się na politykę senioralną w Polsce. W artykule pokazano sposób przedstawiania polityki senioralnej w dokumentach rządowych oraz zaprezentowano jej ocenę za pomocą wskaźników wykorzystywanych w międzynarodowych porównaniach. Wskazanie tradycyjnego podejścia do takiej oceny zostało zilustrowane na przykładzie Polski, zarówno na poziomie krajowym, jak i w ujęciu regionalnym. Tradycyjne sposoby oceny zostały ponadto rozszerzone o nowe ujęcie — znane w analizie polityk publicznych od dawna, ale niestosowane w analizach polityki senioralnej w Polsce — czyli tzw. analizę stanów kontrfaktycznych. Celem artykułu jest wskazanie możliwości i zapoczątkowanie systematycznych — tego typu — analiz dla całościowej oceny efektów działań dotyczących osób starszych. Podjęto próbę analizy kontrfaktycznej na podstawie danych Diagnozy Społecznej (2015), której

[1] Instytut Socjologii Uniwersytet Jagielloński, Grodzka 52, 31-004 Kraków oraz Instytut Statystyki i Demografii, Szkoła Główna Handlowa, ul. Madalińskiego 6/8, 02-554 Warszawa; adresy elektroniczne autorki: jperek@sgh.waw.pl oraz jolanta.perek-bialas@uj.edu.pl

[2] Wstępna wersja opracowania powstała w ramach badań statutowych Instytutu Statystyki i Demografii Kolegium Analiz Ekonomicznych Szkoły Głównej Handlowej w Warszawie (KAE/S15/23/15), pt. Monitorowanie polityki senioralnej (recenzent prof. dr hab. P. Błędowski). Autorka pragnie podziękować także Redakcji jak i anonimowym recenzentom za uwagi do wcześniejszej wersji artykułu. Także serdeczne podziękowania za inspirujące sugestie dotyczące wykorzystania metody *exact matching* dla małych prób dla Maćka Koniewskiego oraz prof. dr hab. Jarosław Górniaka.

wyniki zostały przedstawione i skomentowane w artykule. W Polsce konieczny jest stały monitoring efektów oddziaływania polityki senioralnej nie tylko za pomocą wskaźników (znanych i modyfikowanych czy nowych w zależności od powiązania z celami polityk). Potrzebna jest także ewaluacja działań za pomocą zaawansowanych metod analitycznych (m.in. z wykorzystaniem analizy stanów kontrfaktycznych).

Słowa kluczowe: polityka senioralna, ludzie starzy, wskaźnik aktywnego starzenia się, analiza stanów kontrfaktycznych

Wprowadzenie

W przyjętych przez Radę Ministrów w roku 2013 Założeniach Długofalowej Polityki Senioralnej (ZDPS) w Polsce na lata 2014–2020 (Uchwała nr 238) zdefiniowano politykę senioralną jako „ogół celowych działań organów administracji publicznej wszystkich szczebli oraz innych organizacji i instytucji, które realizują zadania i inicjatywy kształtujące warunki godnego i zdrowego starzenia się" (s. 4). W tym dokumencie rządowym polityka senioralna jest „rozumiana szeroko jako ogół działań, w okresie całego życia człowieka, prowadzących do zapewnienia warunków wydłużenia aktywności, zarówno zawodowej, jak i społecznej, oraz samodzielnego, zdrowego, bezpiecznego i niezależnego życia osób starszych. […] Polityka senioralna powinna być zaprojektowana horyzontalnie, tzn. powinna być spójna z polityką państwa w obszarach zabezpieczenia społecznego, rynku pracy, zdrowia, edukacji (uczenia się przez całe życie), infrastruktury, mieszkalnictwa i innych" (tamże; zob. też Szatur-Jaworska 2015).

ZDPS miały wyznaczać kierunki działań w ramach polityki senioralnej w Polsce i określać działania skierowane do osób starszych (w wieku od 60 lat). Wykorzystanie potencjału osób starszych ma się przekładać głównie na aktywizację zasobów ich możliwości i zdolności, które są podstawowym elementem w rozwoju polityki zdrowego i aktywnego starzenia się (WHO 2002; Walker 2002; ActivAge Consortium 2008; Walker, Maltby 2012). Różnego rodzaju działania z założenia mają przyczynić się do skutecznej realizacji głównego celu polityki senioralnej, czyli „**poprawy sytuacji osób starszych dla godnego starzenia się w dobrym zdrowiu**".

Uchwalenie ZDPS jest częścią ważnego procesu kształtowania polityki senioralnej, która rozwijana jest szczególnie od roku 2012, ogłoszonego Europejskim Rokiem Aktywności Osób Starszych i Solidarności Międzypokoleniowej (European Year for Active Ageing and Solidarity between Generations — w skrócie: EY) (Council of the European Union 2012; Szatur-Jaworska 2015). Przygotowanie dokumentów koordynował Departament Polityki Senioralnej Ministerstwa Pracy i Polityki Społecznej, który powstał w 2012 r. W tworzeniu ZDPS istotną rolę odegrali członkowie Rady ds. Polityki Senioralnej działającej przy Ministerstwie Pracy i Polityki Społecznej w Warszawie.

Polityka senioralna (polityka „wobec" starości i „wobec" osób starszych, zob. Szatur-Jaworska 2000) obejmuje wiele kwestii, które są także w centrum uwagi innych polityk publicznych, takich jak polityka rynku pracy, edukacyjna czy zdrowotna. Nie jest łatwo

analizować łączny efekt wszystkich niezależnych działań, które powinny wspólnie przyczyniać się do pozytywnych zmian w polityce senioralnej w Polsce. To powiązanie z innymi politykami publicznymi utrudnia jej ocenę za pomocą prostych wskaźników (np. tylko wskaźników zatrudnienia w ramach polityki rynku pracy). Stwierdzenie, w jakim stopniu polityka senioralna jest skuteczna, wymaga bardziej kompleksowego podejścia i uwzględnienia różnych efektów wielu polityk łącznie, gdyż wprowadzanie zmian w programach dotyczących osób starszych ma się opierać na różnych analizach, a nie tylko na możliwościach i/lub chęciach decydentów.

Głównym celem artykułu jest próba wskazania, w jaki sposób można oceniać efekty polityki senioralnej. Zostanie to przedstawione — z jednej strony — w ujęciu tradycyjnym, czyli za pomocą wskaźników, które dają możliwość określenia pozytywnych lub negatywnych trendów w analizowanym zjawisku. Z drugiej strony zostanie podjęta także próba rozważenia następującego zagadnienia metodologicznego: jeśli w ewaluacji polityk publicznych można wykorzystywać analizę stanów kontrfaktycznych, to czy jest to możliwe w analizie efektów polityki senioralnej w Polsce? Celem jest sprawdzenie, czy — i w jakim stopniu — zaawansowane analizy (nie tylko wskaźniki) mogą pomóc w ocenie efektów określonych interwencji powiązanych z wybranymi politykami składającymi się na politykę senioralną.

W kontekście skuteczności polityki senioralnej można postawić pytanie, jak dalece udział seniorów w określonych działaniach aktywizujących (np. edukacyjnych) wpływa pozytywnie na ich ocenę jakości życia. Tak postawione pytanie nie tylko wpisuje się w próbę oceny działań polityki senioralnej, ale też uwzględnia koncepcję aktywnego i zdrowego starzenia się. Jest to tylko jedno z możliwych i przykładowych pytań, które by można było zadać w tego typu ewaluacji efektów polityki senioralnej. Będzie też ono miało swoje uzasadnienie w kontekście dostępnych danych, które posłużą do analiz w dalszej części tekstu. Analizy te ilustrują możliwości zastosowania zaawansowanych metod statystycznych w ewaluacji. Oczywiście najlepszą metodą sprawdzenia, w jakim stopniu działania podejmowane na rzecz seniorów przynoszą korzyści, byłby eksperyment polegający na przydzieleniu badanych w sposób losowy do grupy osób poddawanych oddziaływaniu bodźca (interwencji) lub niepoddawanych temu oddziaływaniu (interwencji) (Kurowska 2010). Jednak zorganizowanie tego typu badań jest w praktyce rzadko możliwe a nawet jeśli mamy próby randomizowane w tego typu badaniach, pojawiają się dodatkowe trudności. Przykładowo takie jak utrata osób czy zmiany przypisania do określonych grup eksperymentalnej lub kontrolnej. Mimo różnych strategii i sposobów radzenia sobie z tymi trudnościami nadal — nawet w eksperymencie — nie dysponujemy idealnym podejściem badawczym (Martini 2013).

Tradycyjne podejście — ujęcie wskaźnikowe

Szansę na ocenę zmian będących rezultatem działań składających się na politykę aktywnego starzenia się stwarza wykorzystanie różnego rodzaju wskaźników, które odpowiednio mają wskazywać zmiany i być sygnałem, ale też motywacją do modyfikacji programów określonych polityk. Kwestia wykorzystania wskaźników w polityce społecznej (Kurowska

2011), a szczególnie tych dotyczących osób starszych, w polskiej literaturze przedmiotu była często poruszana. Warto wspomnieć szczególnie publikację Marka Bednarskiego i Barbary Szatur-Jaworskiej (1999), gdzie kompleksowo ujęto możliwości stosowania wskaźników społecznych w ocenie efektów określonych polityk (także tych kierowanych do osób starszych). Kolejne prace próbowały wskazywać zarówno na możliwości, jak i na ograniczenia stosowania wskaźników w ocenie sytuacji starszego pokolenia (m.in. Perek-Białas, Małodzińska 2006; Perek-Białas 2008; Perek-Białas 2010). Jednak trzeba podkreślić, że wskaźniki nie mogą być oderwane od celów programów i określonych polityk i mają pełnić określoną rolę. Trzeba mieć na uwadze zasadę SMART uwzględniającą następujące cechy wskaźników: konkretne (ang. *specific*) — chodzi o to, co chcemy osiągnąć; mierzalne (ang. *measurable*) — trzeba umieć zmierzyć to, co się chce osiągnąć; osiągalne (ang. *achievable*); realistyczne (ang. *realistic*) — przy istniejących zasobach; określone w czasie (ang. *time-bounded*) (Górniak, Keler 2007).

W tym kontekście monitorowanie elementów, które składają się na politykę senioralną, umożliwia wskaźnik aktywnego starzenia się (ang. *active ageing index* — AAI) zaproponowany w Europejskim Roku Aktywności Osób Starszych i Solidarności Międzypokoleniowej (Zaidi i in. 2013). Sam wskaźnik i jego idea, metodologia oraz ocena to materiał na oddzielną publikację. W niniejszym tekście odniesiono się jednak do kilku jego aspektów, aby wykazać, że to „tradycyjne" podejście może być — nie tylko dla Polski ogółem, ale i dla poszczególnych województw — narzędziem ułatwiającym ocenę polityki senioralnej.

Warto jednak na początku tych rozważań przypomnieć, że ujęcie wskaźnikowe stało się w ostatnich latach popularne, służąc nie tylko do oceny efektów polityki senioralnej, ale również, przykładowo, do oceny jakości życia osób starszych. Jednym ze znanych wskaźników wykorzystywanych do porównania sytuacji osób starszych na świecie jest globalny wskaźnik oceny jakości życia seniorów (ang. *global age watch index*). Indeks prezentowany jest cyklicznie od 2013 r. przez sieć organizacji pozarządowych HelpAge International[3] działających na rzecz ludzi starszych. Uwzględnia się w nim wskaźniki, które dotyczą: bezpieczeństwa socjalnego (materialnego), zdrowia, edukacji, zatrudnienia, sprzyjających warunków życia. W pierwszej edycji (2013) prezentacji tego wskaźnika Polska zajęła 62. miejsce, głównie w wyniku niskiej oceny samopoczucia seniorów i małej aktywności zawodowej osób starszych[4]. W drugiej edycji (2014) Polska znalazła się na 32. miejscu, dlatego że zmieniła się wartość wskaźnika wykorzystywanego do oceny samopoczucia starszych osób (tj. osoby starsze — w porównaniu z młodszymi grupami — lepiej niż wcześniej oceniły swoją chęć do życia) oraz przez większą aktywność zawodową osób w wieku 55–64 lat (nastąpił wzrost wskaźnika zatrudnienia w tej grupie).

[3] http://www.helpage.org
[4] Tamże.

Wskaźnik aktywnego starzenia się

W ramach projektu Komisji Europejskiej i Europejskiej Komisji Gospodarczej Narodów Zjednoczonych (United Nations Economic Commission for Europe — UNECE) opracowano — wspomniany już wyżej — syntetyczny **wskaźnik aktywnego starzenia się** dla krajów europejskich (Zaidi i in. 2013; UNECE/European Commission 2015[5]). Wskaźnik ma pokazywać, w jakim stopniu wykorzystywany jest potencjał osób starszych w pracy zarobkowej i niezarobkowej (ang. *employment*). Uwzględnia się też ich aktywność społeczną, w tym wolontariat (ang. *social participation*), a także ocenia się, w jakim stopniu osoby starsze mogą funkcjonować niezależnie, bezpiecznie i zgodnie ze zdrowym stylem życia (ang. *independent, healthy and secure living*). Ponadto bierze się pod uwagę to, czy w konkretnym kraju w Europie istnieją możliwości realizacji polityki aktywnego starzenia się (ang. *capacity and enabling environment for active ageing*). Wskaźnik może być narzędziem do oceny efektów polityki senioralnej na poziomie krajowym, regionalnym i lokalnym, uwzględniającym różny zakres odpowiedzialności decydentów za określone działania w ramach tej polityki (Karpinska, Dykstra 2014, 2015).

W obrębie wskaźnika zostały wyróżnione cztery następujące obszary aktywności: 1) zatrudnienie i/lub aktywność zawodowa; 2) aktywność społeczna; 3) zdrowe, bezpieczne i niezależne od innych życie; 4) zdolność i/lub przygotowanie otoczenia do wykorzystania potencjału osób starszych. Szczegółowe wskaźniki AAI wykorzystane w opisie tych obszarów prezentuje tabela 1[6]. W pierwszej edycji prezentacji wskaźnika, tj. w 2012 r., Polska zajęła ostatnie miejsce w UE. Wyniki naszego kraju dla trzech pomiarów pokazują, że przede wszystkim nastąpiła zmiana w obszarze aktywności zawodowej, dzięki reformom w polityce rynku pracy (m.in. ograniczenie możliwości przechodzenia na wcześniejszą emeryturę). Obszar aktywności społecznej (zarówno w zakresie wolontariatu, opieki, jak i partycypacji obywatelskiej) nie wykazał zmian, co w 2012 r. było też powodem uruchomienia Rządowego Programu na rzecz Aktywności Społecznej Osób Starszych (ASOS, Uchwała nr 137, Uchwała nr 237). Potencjalne efekty tych działań mogą być widoczne dopiero po 2016 r. (dla danych z 2014 — koniec I edycji programu ASOS). Sytuacja materialna osób starszych wydaje się nie zmieniać od 2008 r. (nawet uległa poprawie, jeśli patrzeć nie tylko na sytuację dochodową osób starszych, ale również na deprywację materialną, której od edycji 2008 w kolejnych latach doświadczało mniej osób starszych w Polsce). Natomiast to, że osoby starsze nie uczestniczą w żadnej formie edukacji, nie zmieniło się w analizowanych okresach, co nie tylko znajduje potwierdzenie w danych wykorzystanych w tych wskaźnikach, ale jest także spójne z wynikami analiz na podstawie danych badania SHARE (Portret generacji… 2015).

---

[5] http://www1.unece.org/stat...

[6] W tabeli 1 zaprezentowano skrócone nazwy wskaźników, rozszerzone opisy wskaźników w języku angielskim znajdują się w opracowaniach: Zaidi i in. 2013; UNECE/European Commission 2015; Perek-Białas, Mysińska 2013; Perek-Białas, Zwierzchowski 2014.

Tabela 1. Wyniki wskaźnika aktywnego starzenia się dla Polski z trzech okresów: 2008/2010, 2010/2012 i 2012/2014* (w %)

Obszar	Wskaźniki	2008/2010		2010/2012		2012/2014	
		M	K	M	K	M	K
Zatrudnienie/ Aktywność na rynku pracy	Wskaźnik zatrudnienia 55–59	54,7	27,1	59,4	33,6	63	43,1
	Wskaźnik zatrudnienia 60–64	27,8	11,4	26,7	12,7	32,9	13,7
	Wskaźnik zatrudnienia 65–69	13	6,5	13,4	6,4	13,4	6,7
	Wskaźnik zatrudnienia 70–74	7,3	3,2	7,5	3,3	7,3	3
Aktywność społeczna	Wolontariat	3,7	4,4	4	1,7	4	1,7
	Opieka nad dziećmi	22,4	21,6	17,4	26,2	17,4	26,2
	Opieka nad osobami dorosłymi	16	12,6	11	14,8	11	14,8
	Zaangażowanie polityczne	16,4	10,2	10,7	8,4	10,7	8,4
Niezależność, zdrowie i bezpieczeństwo	Aktywność fizyczna	8,1	6,2	8,1	6,2	8,1	6,2
	Dostęp do opieki medycznej	81,4	81,8	78,1	76,2	78,5	75,9
	Niezależne mieszkanie	71,2	73,6	65,5	70,8	68,2	69,1
	Relacja mediany dochodów	100	91,8	100	87,7	100	89,5
	Brak zagrożenia ubóstwem	96,2	94	95,4	91,9	96	91,9
	Brak deprywacji materialnej	82,4	77,2	87,1	81,4	89,5	82,7
	Bezpieczeństwo w miejscu zamieszkania	77,8	66,9	86,1	75,2	79,8	74,3
	Uczenie się przez całe życie	0,7	0,7	0,8	0,8	0,6	0,6

Zdolność i przygotowanie otoczenia do wykorzystania potencjału starzejącego się społeczeństwa	Pozostałe średnie dalsze oczekiwanie życia w wieku 55++	42,6	54,8	43,5	55,8	44,2	56,4
	Relacja lat w zdrowiu do średniej oczekiwanej długości życia w wieku 55	54,9	49,6	53,2	47,4	56,1	48,9
	Zdrowie psychiczne	54	54,5	52,2	51,6	52,2	51,6
	Korzystanie z Internetu	17	11	23	15	26	22
	Więzi społeczne	31,8	25,3	30,8	30,7	25,1	28,6
	Wykształcenie (co najmniej średnie)	73	62,7	76,2	67,7	78,4	71

* Różne lata, gdyż dane wskazane jako odnoszące się do roku 2010 dotyczą de facto roku 2008.

++ W polskim podejściu wskaźnik jest inaczej liczony, w oryginale: *Remaing life expectancy achievement of 50 years at age 55.*

Źródło: opracowanie na podstawie: http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=76287845 [dostęp: 23.08.2015].

Prace UNECE i Komisji Europejskiej, które były prowadzone pod kierunkiem prof. Asghara Zaidiego, stały się inspiracją do przygotowania indeksu dla województw w Polsce (Perek-Białas, Mysińska 2013; Perek-Białas, Zwierzchowski 2014). Celem tych prac było określenie, czy można wykorzystać wskaźnik aktywnego starzenia się nie tylko do planowania działań na rzecz osób starszych w skali całego kraju, lecz także jako narzędzie (tj. indeks aktywnego starzenia się w ujęciu regionalnym w Polsce — na poziomie województw) pomocne w planowaniu polityk regionalnych dotyczących omawianych zagadnień. Wyniki i szczegółowe omówienie oryginalnej koncepcji wskaźnika oraz zasad jego przygotowania wraz z analizami wskaźnika w ujęciu regionalnym dla Polski można znaleźć w opracowaniach: Perek-Białas, Mysińska (2013), Perek-Białas, Zwierzchowski (2014) (www.senior.gov.pl). Wskaźnik ten mógł wesprzeć przygotowanie, rewizję i zmiany w zakresie określonych polityk regionalnych w Polsce poprzez wskazanie, co można zmienić/poprawić w danym obszarze polityki aktywnego starzenia się/polityki senioralnej w konkretnym województwie/regionie kraju. Tak więc, odpowiednio zaplanowane działania w strategiach województw[7], podporządkowane zwłaszcza celowi jak najlepszego i dostosowanego do potrzeb wykorzystania środków europejskich w kolejnych latach, są powiązane z wynikami analiz, które wskazują, że określone działania muszą ulec zmianie. W związku z tymi pracami, w październiku 2014 r. w Krakowie odbyło się spotkanie w ramach programu wzajemnej oceny w zakresie ochrony socjalnej i włączenia społecznego, poświęcone wskaźnikowi aktywnego starzenia się i rozszerzeniu go na szczebel regionalny (Peer Review in Poland: The Active Ageing Index and its extension to the regional level[8]) (Breza, Perek-Białas 2014) oraz możliwości wykorzystania podejścia regionalnego w innych krajach europejskich. Celem spotkania była wymiana doświadczeń i dyskusja, w jakim stopniu na poziomie regionalnym/lokalnym można wykorzystywać wskaźnik aktywnego starzenia się w planowaniu i prowadzeniu krajowych polityk wpisujących się w koncepcję aktywnego starzenia się.

Zestawienie wyników z obu edycji wskaźnika aktywnego starzenia się dla Polski w ujęciu regionalnym zostało zaprezentowane w tabeli 2. Wyniki wyliczonego wskaźnika nie odbiegają od wartości ogółem liczonej dla Polski, mimo konieczności zmiany lub rezygnacji z niektórych wskaźników, gdyż nie było odpowiednich danych, z których by można było skorzystać. Warto zauważyć, że wyraźnie widoczne są zmiany w obszarze zatrudnienia dla wszystkich województw, co wpływa na ogólny wzrost wartości wskaźnika, gdyż ten obszar ma największy udział w ogólnej wartości wskaźnika aktywnego starzenia się (por. wagi obszarów, tj. obszar 1 oraz obszar 2 mają w sumie 70% udziału). Prezentacja tego podejścia — mimo niewątpliwie dalszej potrzeby rozwoju i rewizji określonych rozwiązań

[7] Zapisy dotyczące polityki aktywnego starzenia się nie są wprost zawarte w dokumentach strategicznych województw, ale przykładowo mieszczą się pod hasłem „srebrna gospodarka" — jak to miało miejsce w przypadku Małopolski. Można m.in. wskazać prace zespołu ekspertów pod kierunkiem prof. dr hab. S. Golinowskiej dotyczące przygotowania tego regionu do procesu starzenia się populacji, przeprowadzone na zlecenie Urzędu Marszałkowskiego województwa małopolskiego (Golinowska 2010).

[8] http://ec.europa.eu/social/main...

metodologicznych — pokazuje możliwości wykorzystania wskaźników, które z osobna, ale i całościowo (łącznie) pozwalają na ocenę zmian w polityce senioralnej, w zakresie aktywnego starzenia się nie tylko na poziomie kraju, ale też w ujęciu wojewódzkim. Obecnie podejmuje się próby sprawdzenia, w jakim stopniu budowa wskaźnika aktywnego starzenia się mogłaby być zastosowana na niższym poziomie zarządzania (poziom lokalny, aczkolwiek w różny sposób rozumiany w różnych krajach). Przykładowo NUTS-3 (według nomenklatury statystycznej) w przypadku Polski dotyczy powiatów, a w związku z tym, że wiele się dzieje na poziomie lokalnym, ważne jest też monitorowanie polityki senioralnej na poziomie LAU-2 (gminy) — w tym z uwzględnieniem gmin miejskich, np. miast powyżej 100 tys. mieszkańców[9].

Tabela 2. Wskaźnik aktywnego starzenia się w Polsce w ujęciu regionalnym (system wag jak w ujęciu regionalnym)

Województwa		2013 (I edycja)					2014 (II edycja)				
		Wartości składowych wskaźnika wg obszarów				Ogółem	Wartości składowych wskaźnika wg obszarów				Ogółem
		1	2	3	4	Wartość	1	2	3	4	Wartość
1	dolnośląskie	29,7	13,3	49,7	25,6	25,1	35,9	19,85	61,01	28,3	31,3
2	kujawsko--pomorskie	24,3	11,0	49,3	26,7	22,6	35,6	15,75	59,07	27,7	29,4
3	lubelskie	**38,3**	**17,3**	**47,3**	**23,7**	**28,9**	40,0	19,90	59,11	24,7	31,8
4	lubuskie	30,3	16,5	48,1	24,1	26,0	34,4	17,85	59,05	26,5	29,5
5	łódzkie	36,2	11,2	46,8	24,7	26,2	38,3	13,76	58,95	27,7	29,7
6	małopolskie	32,5	13,9	50,3	25,5	26,4	36,8	17,79	59,41	25,2	30,1
7	mazowieckie	**39,5**	**13,5**	**51,7**	**27,5**	**29,2**	**43,6**	**19,61**	**59,43**	**27,8**	**33,6**
8	opolskie	29,3	12,3	49,9	26,4	24,8	36,3	17,01	64,86	27,5	30,6
9	podkarpackie	**39,8**	**13,3**	**46,5**	**24,8**	**28,2**	**39,5**	**22,58**	**57,19**	**25,5**	**32,6**
10	podlaskie	32,9	11,0	47,5	20,5	24,2	38,0	13,98	56,72	22,3	28,3
11	pomorskie	28,9	12,1	50,7	30,2	25,5	**37,9**	**18,26**	**61,64**	**31,7**	**32,2**
12	śląskie	27,9	12,0	53,9	28,4	25,0	32,5	16,82	64,95	30,4	29,8

[9] Obecnie w Niemczech są prowadzone prace analityczne w ramach inicjatywy UNECE i Komisji Europejskiej mające na celu ocenę możliwości wyliczenia wskaźnika aktywnego starzenia się na poziomie lokalnym (spotkanie poświęcone tej kwestii — *Peer Review on Active Ageing Index at local level* — odbyło się w kwietniu 2016 r.), ale dla NUTS-3, czyli polskiego odpowiednika powiatów, które z punktu widzenia kształtowania polityki senioralnej nie będą miały w Polsce takiego znaczenia jak np. polityka senioralna gmin miejskich (w tym miast na prawach powiatu).

Tabela 2. cd.

	Województwa	2013 (I edycja)					2014 (II edycja)				
		Wartości składowych wskaźnika wg obszarów				Ogółem	Wartości składowych wskaźnika wg obszarów				Ogółem
		1	2	3	4	Wartość	1	2	3	4	Wartość
13	świętokrzyskie	36,7	11,6	44,4	21,5	25,6	37,4	16,27	57,35	24,2	29,4
14	warmińsko--mazurskie	27,8	9,6	45,6	23,2	22,3	34,0	14,14	54,93	26,3	27,6
15	wielkopolskie	30,0	13,0	52,8	24,0	25,1	37,4	18,38	62,22	26,8	31,1
16	zachodniopomorskie	25,2	13,7	47,5	25,4	23,5	37,8	16,49	63,76	27,2	30,8
	Wagi	**35**	**35**	**10**	**20**		**35**	**35**	**10**	**20**	

Obszary: 1. Zatrudnienie; 2. Aktywność społeczna; 3. Niezależne, zdrowe i bezpieczne życie; 4. Zdolność i przygotowanie otoczenia do wykorzystania potencjału starzejącego się społeczeństwa.

Źródło: opracowanie na podstawie Perek-Białas, Mysińska (2013) oraz Perek-Białas, Zwierzchowski (2014).

Uzupełniające podejście — analiza stanów kontrfaktycznych w ocenie skuteczności działań polityki senioralnej

Jednym z celów niniejszego artykułu jest próba wykazania, że szczególnie w ewaluacji polityk publicznych można wykorzystywać alternatywne podejścia (nie tylko wskaźniki), które pozwalają ocenić, w jakim stopniu określone interwencje dają efekty netto[10] dla określonej grupy.

Przedstawione zostaną możliwości oceny, przykładowo, czy udział seniorów w określonych działaniach aktywizujących (edukacyjnych, które wpisują się w realizację polityki senioralnej w tym obszarze aktywności społecznej) mają wpływ na jakość ich życia. Chodzi na przykład o to, czy osoby biorące udział w zajęciach uniwersytetów trzeciego wieku (UTW) uzyskują w związku z tym korzyści, których nie doświadczają osoby niebiorące udziału w tego typu aktywnościach? I jeśli tak, to jak duże są te efekty? Powyżej we wstępie wyjaśniono krótko, że najlepszą metodą byłoby zastosowanie eksperymentu klasycznego, który pozwalałby ocenić, czy faktycznie dana przyczyna (np. udział w zajęciach UTW czy innych aktywnościach, m.in. w projektach finansowanych w ramach programu ASOS) powoduje określony skutek (np. zadowolenie z życia, brak depresji itp.). Jednak wspomniano wyżej, że takie podejścia badawcze nie są możliwe do zastosowania nie tylko

[10] W ewaluacji ustala się związek przyczynowy między podjętą interwencją a obserwowanymi wynikami, ale wtedy wymagane jest: zmierzenie efektu całkowitego interwencji jako ogólnej zmiany w zakresie zdefiniowanym przez cel (zmiany na poziomie przyjętych wskaźników) lub/oraz oddzielenie zmiany niezależnej od podjętej interwencji od zmiany, którą można przypisać wpływowi interwencji (efekt netto).

z przyczyn organizacyjnych. Dlatego alternatywą może być zastosowanie metody analizy stanów kontrfaktycznych, dzięki której można próbować uchwycić efekty trudne do wykazania na poziomie indywidualnym.

Metoda analizy stanów kontrfaktycznych

Generalnie idea tego podejścia sprowadza się do wykazania różnic między grupą osób poddanych pewnemu oddziaływaniu a grupą osób, które takiemu oddziaływaniu nie są poddawane. Sytuacja jest modelowa, jeśli w obu grupach znajdują się porównywalne osoby, gdyż wtedy różnice powinny być — na mocy przyjętego założenia — oceniane jako wynik działania bodźca. Problemem z jednej strony jest znalezienie osób porównywalnych, z drugiej zaś powiązanie ich (*matching*) z osobami z grupy obserwowalnej (eksperymentalnej) i kontrolnej. Sama metoda nie będzie omawiana szczegółowo w niniejszym artykule, gdyż jest dość dobrze opisana w literaturze przedmiotu, w tym w polskich pracach dotyczących zagadnień ewaluacji (m.in. Konarski, Kotnarowski 2007; Strawiński 2008; Trzciński 2009). Główne zasady i założenia metody zostaną pokrótce przedstawione poniżej, gdyż były one wykorzystywane w analizie prezentowanej w tym artykule, m.in. z zastosowaniem odpowiednich pakietów w programie R, ale też z użyciem programu PS IMAGO oraz STATA.

Rolą *matchingu* (Ho i in. 2007) jest przekształcenie danych pochodzących z badań nieeksperymentalnych w taki sposób, aby ostatecznie grupy różniące się faktem ekspozycji na bodziec były do siebie jak najbardziej podobne (eksperyment *ex post facto*). *Matching* przeprowadza się na macierzy wcześniej wyliczonych dystansów Mahalanobisa, Euklidesa lub najczęściej na wartościach przewidywanych z jakiegoś modelu dwumianowego (logit lub probit). Takie wartości przewidywane nazywa się *propensity scores* (PS) (Rosenbaum, Rubin 1983). *Propensity scores* (PS) są szacowane z wykorzystaniem regresji dwumianowej. Mogą być wyrażone w jednostkach prawdopodobieństwa (probit) lub logitowych (logit), w zależności od zastosowanego modelu.

W celu bilansowania rozkładu cech w porównywanych grupach można zastosować PS do *matchingu*, subklasyfikacji lub ważenia danych (Stuart 2010). W omawianym przykładzie zastosowany został *matching*. Istnieją różne metody wykorzystania PS do *matchingu*, przykładowo: dokładne (*exact matching*) — pary tworzą obserwacje identyczne na wszystkich kowariantach; pełne (*full matching*) — wykorzystane są wszystkie obserwacje w próbie, dopasowane jednostki tworzą konfiguracje 1:wielu lub wiele:1; najbliższego sąsiedztwa (*nearest neighbor*) — dla każdej jednostki w grupie eksperymentalnej szuka się jednostki z grupy kontrolnej, która ma do niej najbardziej zbliżoną odległość; optymalne (*optimal matching*) — wszystkie pary są rozpatrywane symultanicznie, tak aby ostateczne rozwiązanie skutkowało najmniejszą sumą odległości we wszystkich dopasowanych parach.

W przypadku metody najbliższego sąsiedztwa możliwe są dodatkowe opcje. Niedopasowane jednostki mogą zostać usunięte lub zostać w puli do wykorzystania jako kolejne jednostki do połączenia kolejnych obserwacji (tym samym procedura może być prowadzona na dwa sposoby: bez zwracania lub ze zwracaniem). Możliwe jest też zastosowanie różnego stosunku (*ratio*) obserwacji z grupy kontrolnej (C) do eksperymentalnej (T) (C:T).

Np. ratio = 1 będzie skutkowało parami 1:1; ratio = 2 będzie skutkowało dopasowaniem 2:1 (czyli 2 jednostki z C do jednej z T); ratio = 0,5 będzie skutkowało dopasowaniem 1:2 (czyli 1 jednostka z C do dwóch z T). Dodatkowo możliwe jest zadanie punktu odcięcia (*caliper*), definiującego, z jaką precyzją chcemy dopasowywać pary (pewna wartość wyrażona w jednostkach odchylenia standardowego PS, np. powyżej 2 odchyleń standardowych wartości PS, powyżej której nie dopuszczamy łączenia w pary).

Na końcowym etapie wylicza się różnicę między średnimi zmiennej wynikowej (Y) w grupie objętej interwencją (Y1) i grupie nieobjętej interwencją (Y0), która jest trafnym i nieobciążonym szacunkiem średniego efektu przyczynowego (*average treatment effect* — ATE), czyli ATE = E[Y1–Y0]. ATE to oczekiwany przyrost poziomu zmiennej wynikowej dla losowo wybranych jednostek z populacji.

Przykład zastosowania analizy stanów kontrfaktycznych w ocenie działań aktywizujących osoby starsze

W prezentacji możliwości zastosowania metod opisanych powyżej wykorzystano dane z badania Diagnoza Społeczna, które jest prowadzone w Polsce od 2000 r., ale dopiero w edycji w 2015 r. uwzględniono pytanie, które pozwala uchwycić jednostki biorące udział w aktywnościach edukacyjnych przeznaczonych specjalnie dla seniorów oraz te, które nie brały w nich udziału. Dlaczego to może mieć znaczenie? Chociażby dlatego, że wspomniany wcześniej ASOS — program Aktywności Społecznej Osób Starszych na lata 2013, 2014–2015 oraz od 2016 — powinien przynosić zwiększenie aktywności starszych osób w Polsce w dziedzinie społecznej/edukacyjnej. Można też chcieć sprawdzić, czy takie działania przekładają się na lepszą ocenę jakości życia tych, którzy biorą w nich udział.

Niestety tych efektów oddziaływania programu ASOS nie da się w łatwy sposób sprawdzić, gdyż wyliczenia we wspomnianym wyżej wskaźniku aktywnego starzenia się dotyczą wcześniejszych okresów z opóźnieniem o dwa lata (np. edycja wskaźnika dla 2014 r. dotyczy obliczeń dla wcześniejszych lat, tj. 2012, a więc przed EY 2012 oraz przed wprowadzeniem ASOS). Zasadne jest stałe i systematyczne monitorowanie jak osoby starsze przejawiają swoją aktywność zarówno na rynku pracy, jak i w innych obszarach (np. opieka nad wnukami, innymi starszymi osobami), oraz czy mają one możliwość prowadzenia niezależnego, bezpiecznego, zdrowego życia w środowisku, które wspiera ideę aktywnego starzenia się. Jak napisano powyżej, ujęcie wskaźnikowe jest możliwe, ale nie pozwala na bezpośrednie wykazanie, że udział w określonych aktywnościach ma bezpośredni wpływ na lepszą jakość życia osób starszych wyrażoną obiektywną lub subiektywną oceną tych osób.

Analizę stanów kontrfaktycznych należy rozpocząć od zdefiniowania kluczowej zmiennej (najczęściej dychotomicznej) grupującej obserwacje do stanów eksperymentalnych (stan eksperymentalny, czyli grupa eksperymentalna *versus* kontrolna). W omawianym przypadku jest to zmienna *UTW*, która ma wartość 1, jeśli osoba w wieku 54+[11] brała

[11] Najmłodsza osoba deklarująca udział w zajęciach UTW miała 54 lata, dlatego też analizy dotyczą osób w wieku 54+.

udział w edukacji w ramach uniwersytetu trzeciego wieku, a 0 — jeśli nie była uczestnikiem tej formy edukacji/partycypacji społecznej. Oczywiście wykorzystanie tej zmiennej nie jest tożsame z założeniem, że istnienie i działalność UTW jest tylko i wyłącznie wynikiem funkcjonującej polityki senioralnej oraz że tylko UTW mogą przyczyniać się do zmiany oceny jakości życia seniorów, która jest elementem polityki senioralnej. Zdecydowano się na taki wybór zmiennej do analizy, gdyż w badaniu w ramach Diagnozy Społecznej dopiero w 2015 r. pojawiła się możliwość udzielania odpowiedzi na pytanie o udział w tej formie edukacji/kształcenia[12]. Tym samym dopiero teraz pojawiła się szansa na próbę analiz z użyciem informacji, która zawiera pytanie o UTW, ale niestety, bez uwzględnienia innych form edukacyjnych (które też mogły być działaniem w ramach np. ASOS).

W grupie kontrolnej powinno być więcej obserwacji niż w grupie eksperymentalnej, najlepiej minimum dwa razy więcej. Jest to konieczne, gdyż celem jest znalezienie dla obserwacji w grupie eksperymentalnej ich „statystycznych bliźniaków", czyli najbardziej podobnych obserwacji w grupie kontrolnej. Im więcej obserwacji jest dostępnych w grupie kontrolnej, tym większa jest szansa na znalezienie lepszych par. Osoby w dwóch grupach mają być jak najbardziej do siebie podobne. Należy więc sprawdzić, czy różnią się ze względu na zmienne, które (a) wiążą się z udziałem w zajęciach UTW i/lub (b) zmienną wynikową (w tym przypadku: ocena swojego dotychczasowego życia[13], tzw. *outcome* — przekształcenie zmiennej oryginalnej w taki sposób, żeby w analizie przyjmowała wartości maksymalnie do 70; im wyższy wynik, tym lepsza ocena, przykładowo wartość 70 oznacza umownie „wspaniałe" życie, a wartość 10 — „okropne", zmienna *ZYCIE2*). Niestety, na tym etapie analiz nie zostały wykorzystane możliwości zastosowania innych zmiennych wynikowych, które by można było dodatkowo sprawdzić.

Okazało się, że osób, które wskazały UTW jako formę kształcenia, było w całej bazie Diagnozy Społecznej 2015 tylko szesnaście[14]. Tak mała liczebność stwarza pewne problemy z porównaniem z grupą osób, które nie wskazały, że uczą się w UTW. Osobami, które wskazały na udział w zajęciach UTW, były w większości kobiety (10 na 16 osób), osoby z województw: lubelskiego, małopolskiego, pomorskiego, śląskiego, mazowieckiego, wielkopolskiego i warmińsko-mazurskiego. Średnia ich wieku wynosiła 71 lat, a średni

[12] Jest to tylko jedna z form edukacji, uczenia się, a nie jedyna forma, która odpowiada za politykę senioralną.

[13] W przypadku badań edukacyjnych przykładowo wykorzystuje się ocenę wyników z egzaminu z matematyki w grupie uczniów biorących udział w zajęciach szachowych i w grupie tych, którzy w nich nie uczestniczyli. Zdaję sobie sprawę z tego, że ta zmienna nie oddaje najlepiej obiektywnej możliwości określenia wpływu udziału w zajęciach na UTW na ocenę jakości własnego życia, szczególnie z uwzględnieniem wcześniejszych etapów życia, które mogły być decydujące dla tej oceny (*life-course perspective*). Przyjęłam założenie, że nawet jeśli wcześniejsze etapy życia mogłyby być uznane za „okropne", to właśnie fakt udziału w zajęciach, spotykanie się z różnymi osobami, może zmienić taką opinię i nie będzie to dotyczyło tylko teraźniejszej oceny swojego zadowolenia z życia.

[14] Niestety, ten fakt niesie ze sobą zagrożenie i stwarza podstawę do możliwej krytyki tych analiz z powodu małych liczebności grupy quasi-eksperymentalnej. Jednak jest to przykład i próba pokazania możliwości takich analiz, dlatego zaprezentowane wyniki należy traktować z ostrożnością, a nie jako „statystyczne" dowody na korzyści/niekorzyści płynące z udziału w zajęciach UTW.

miesięczny dochód — 2239 zł; w większości były to osoby z wykształceniem wyższym i policealnym (7 na 16 osób), pełnosprawne (12 osób), niewykonujące pracy zarobkowej (tylko jedna osoba oświadczyła, że oprócz zajęć na UTW także pracuje), aktywnie działające społecznie (8 osób); zadowolenie ze stanu zdrowia zadeklarowała jedna osoba, z opieki zdrowotnej — również jedna, zadowolonych z warunków mieszkaniowych było 9 osób, korzystających z Internetu — również 9. Tylko jedna osoba przyznała, że pali papierosy. Mała liczebność próby ogranicza prezentację rozkładów. Gdyby porównano obie grupy ze względu na ich ocenę dotychczasowego życia (zmienna: ocena życia), to (na tzw. grupach niedopasowanych bez uwzględnienia wag) osoby biorące udział w zajęciach UTW w porównaniu z osobami nieuczestniczącymi w tej formie zajęć lepiej oceniają życie, efekt wyniósł: 0,458[15] (tabela 3). Jeśli uwzględni się odpowiednie wagi to okazuje się, że osoby uczęszczające na zajęcia UTW gorzej oceniają swoje życie niż osoby z grupy porównywalnej (efekt wyliczony z uwzględnieniem wag wyniósł: -0,026, tabela 4). Wynik ten jednak nie może być tak wykorzystany, gdyż istnieją znaczne różnice między tymi grupami, co potwierdzono analizą z zastosowaniem procedury *imbalance* (tabela 5) z użyciem następujących zmiennych: płeć, wykształcenie, niepełnosprawność, wykonywanie pracy przynoszącej zarobek, palenie papierosów, korzystanie z Internetu, aktywność fizyczna (im wartość odległości L1 jest bliższa 1, tym gorzej, bo to oznacza większe rozbieżności między grupami).

W związku z tym, że w grupie kontrolnej było tak mało obserwacji, zdecydowano się zrezygnować z często stosowanego — w tego typu ocenach — podejścia, a mianowicie szacowania indywidualnych wyników skłonności do udziału w zajęciach UTW np. za pomocą modelu regresji logistycznej, który jest wykorzystany do wyliczenia tzw. *propensity scores* (Rubin, Thomas 1996; Konarski, Kotnarowski 2007). Zastosowano metodę *exact matching* zgodnie z procedurą *coarsened exact matching* (CEM) (Blackwell i in. 2009). Procedura zastosowanego algorytmu jest zaprezentowana w tabeli 6 wraz z wynikami, które wskazują, że udało się zmniejszyć dystanse między obiema grupami[16] (odległość L1 bliska 0 — oznacza idealne dopasowanie, a im bliżej zera, tym lepiej). W kolejnym kroku sprawdzono raz jeszcze, czy przynależność do grupy osób uczestniczących w zajęciach UTW ma wpływ na ocenę życia w porównaniu z osobami z tzw. grupy kontrolnej i jaki jest ten wpływ. Wyniki wskazują, że osoby uczęszczające na zajęcia UTW w porównaniu do osób, które nie brały udziału w zajęciach UTW, a na wybranych zmiennych były w takiej samej sytuacji, oceniały gorzej jakość swojego życia (znak ujemny przy współczynniku regresji), ale wynik ten nie jest istotny statystycznie (tabela 7).

[15] Bez oceny istotności statystycznej tych wyników.

[16] Do tej metody przy tak małej liczebności osób w interesującej nas grupie nie powinno się używać zbyt dużej liczby zmiennych.

Tabela 3. Wyniki regresji dla zmiennej oceniającej życie z uwzględnieniem zmiennej UTW (dane nieważone)

```
. reg ZYCIE2 UTW

      Source |       SS       df       MS              liczba obserwacji =   10233
-------------+------------------------------           F(  1, 10231) =    0.03
       Model |  3.14743051     1  3.14743051           Prob > F      =  0.8559
    Residual |   976751.49 10231  95.4697967           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared = -0.0001
       Total |  976754.637 10232  95.4607737           Root MSE      =  9.7709

------------------------------------------------------------------------------
      ZYCIE2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         UTW |   .4584067   2.524677     0.18   0.856    -4.490456    5.407269
       _cons |   50.87493   .0966607   526.32   0.000     50.68545     51.0644
------------------------------------------------------------------------------
```

Źródło: opracowanie własne.

Tabela 4. Wyniki regresji dla zmiennej oceniającej życie z uwzględnieniem zmiennej UTW (dane ważone)

```
. regress ZYCIE2 UTW [iweight = WAGA_2015_IND]

      Source |       SS       df       MS              liczba obserwacji =    8411
-------------+------------------------------           F(  1,  8409) =    0.00
       Model |  .009866672     1  .009866672           Prob > F      =  0.9920
    Residual |  828836.296  8409  98.5653818           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared = -0.0000
       Total |  828836.305  8410   98.553663           Root MSE      =  9.9275

------------------------------------------------------------------------------
      ZYCIE2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         UTW |  -.0269553   2.694004    -0.01   0.992    -5.307867    5.253956
       _cons |   50.93266   .1083292   470.17   0.000     50.72031    51.14501
------------------------------------------------------------------------------
```

Źródło: opracowanie własne.

Tabela 5. Wyniki analizy pokazującej różnice między grupami osób uczestniczących oraz nieuczestniczących w zajęciach UTW, z zastosowaniem procedury *imbalance*

```
. imb PLEC_ALL eduk4_2015 HC30 HD2 HP43 HP100 HP106_01 , treatment (UTW)
(using the scott break method for L1 distance)

Multivariate L1 distance: .91043575

Univariate imbalance:

                 L1      mean       min       25%       50%       75%       max
   PLEC_ALL   .10933    .10933         0         0         0         0         0
 eduk4_2015   .54659    1.2767         0         2         2         1         0
       HC30   .13783    .85894         1         0         0         0         0
        HD2   .17285    .17285         0         0         0         0         0
       HP43   .21475    .21475         1         0         0         0         0
      HP100   .24632   -.24632         0         0        -1         0         0
   HP106_01   .47554    .47554         0         0         1         1         0
```

Źródło: opracowanie własne.

Tabela 6. Wyniki zastosowanego *matchingu* metodą *coarsened exact matching* (CEM)

```
. cem PLEC_ALL eduk4_2015 HC30 HD2 HP43 HP100 HP106_01 , tr(UTW)

Matching Summary:
-----------------
Number of strata: 664
Number of matched strata: 12

               0      1
      All  12926     16
  Matched    965     14
Unmatched  11961      2


Multivariate L1 distance: 2.907e-15

Univariate imbalance:

                  L1      mean       min       25%       50%       75%       max
   PLEC_ALL   1.3e-15  -6.2e-15         0         0         0         0         0
 eduk4_2015   3.0e-15   1.5e-14         0         0         0         0         0
       HC30   2.3e-16  -8.9e-16         0         0         0         0         0
        HD2   6.1e-16  -5.8e-15         0         0         0         0         0
       HP43   1.0e-15   3.1e-15         0         0         0         0         .
      HP100   4.1e-15   7.5e-15         0         0        -1         0         .
   HP106_01   3.7e-15   1.9e-14         0         0         0         0         .
```

Źródło: opracowanie własne.

Tabela 7. Wyniki regresji dla zmiennej oceniającej życie po zastosowanej metodzie *matchingu*

```
. reg ZYCIE2 UTW [iweight=cem_weight]

      Source |       SS       df       MS              liczba obserwacji =     909
-------------+------------------------------           F(  1,   907) =    1.13
       Model |  88.9760147     1  88.9760147           Prob > F      =  0.2870
    Residual |  71124.1578   907  78.4169325           R-squared     =  0.0012
-------------+------------------------------           Adj R-squared =  0.0002
       Total |  71213.1338   908  78.4285615           Root MSE      =   8.855

------------------------------------------------------------------------------
      ZYCIE2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         UTW |  -2.635072   2.473682    -1.07   0.287    -7.489878    2.219734
       _cons |    53.4043   .2958125   180.53   0.000     52.82375    53.98486
------------------------------------------------------------------------------
```

Źródło: opracowanie własne.

Głównym problemem, jaki wiąże się z próbą zastosowania analizy stanów kontrfaktycznych w ewaluacji skuteczności wybranego działania/aktywności wpisującej się w politykę senioralną, jest fakt, że w tym przykładzie bardzo mało osób badanych wskazało na UTW jako formę edukacji, a to było podstawą analizy. To znacznie utrudniło przeprowadzenie analizy z użyciem proponowanej metody (niezależnie od sposobu *matchingu*). Kwestia braków danych na wybranych do analizy zmiennych też nie jest w tekście szerzej omówiona, ale trzeba pamiętać, że może mieć wpływ zarówno na zbilansowanie obu grup, jak i na ocenę efektów.

Podsumowanie

Zaplanowane i realizowane działania w ramach polityki senioralnej powinny być odpowiednio monitorowane i ewaluowane, aby można było stwierdzić, czy są skuteczne, czy jednak wymagają dalszych zmian. Nie jest łatwo całościowo wykazać sukces lub porażkę działań w ramach określonych polityk, które składają się na politykę senioralną, tutaj powiązaną z polityką aktywnego starzenia się (Błędowski 2012; Theiss 2012). Ujęcie tradycyjne związane ze wskaźnikami jest znanym i już stosowanym podejściem nie tylko na poziomie międzynarodowym i krajowym, ale też wojewódzkim. Niewątpliwie należy stosować i rozwijać metodologię wskaźników społecznych w polityce senioralnej i aktywnego starzenia się, jednocześnie stale oceniając krytycznie prawidłowe wykorzystanie i znaczenie wskaźników w powiązaniu z celami polityk z tego obszaru.

To uzupełniające rozwiązanie ewaluacji zaproponowane w niniejszym artykule może być przydatne, ale trzeba mieć świadomość jego ograniczeń. Starałam się też pokazać w tekście te ograniczenia, m.in.: brak odpowiednich danych do zastosowania określonych

procedur czy zbyt małe liczebności, a w konsekwencji problemy z wyborem właściwych metod i prawidłowym szacowaniem wyników. Mimo to, jest to na pewno kierunek, który można dalej rozwijać i sprawdzać, czy wyniki tych analiz faktycznie mogą być cennym uzupełnieniem podejścia wskaźnikowego. W artykule tym próbowałam wykazać, że włączenie pewnych informacji w różnego rodzaju badaniach (nie tylko Diagnozy Społecznej) dotyczących udziału starszych Polaków i Polek w aktywnościach finansowanych ze środków publicznych mogłoby być cennym źródłem analiz w ewaluacji polityki senioralnej. Starałam się też pokazać, jak przeprowadzić takie analizy, prezentując nie tylko wyniki, ale i procedury, które mogą być przykładem do replikacji lub innego wykorzystania przez zainteresowanych tą tematyką badaczy. Jeśli w badaniach są informacje o udziale w programach finansowanych ze środków publicznych, to można próbować wykorzystywać analizę stanów kontrfaktycznych. Można wtedy oceniać — niezależnie od tradycyjnego pomiaru efektów za pomocą wskaźników (zarówno zakładanych, jak i dodatkowych) lub specjalnie przeznaczonych do ewaluacji badań (niestety kosztownych) — czy mamy do czynienia z pozytywnymi, negatywnymi czy zerowymi skutkami działań podejmowanych z myślą o seniorach. Dzięki temu uzasadnienie decyzji w sprawie wydatkowania środków publicznych na te zakładane cele zyska dodatkową argumentację, adresowaną zresztą nie tylko do decydentów, przesądzającą o konieczności podejmowania lub może — właśnie — zaniechania określonych inicjatyw w zależności od tego, czy są one lub nie są skuteczne. Nie znaczy to, że metody proponowane w tekście mają zastąpić tradycyjne sposoby oceny rezultatów polityk dotyczących osób starszych, ale mogą być uzupełnieniem i dodatkowym głosem w dyskusji o ich kształtowaniu.

Bibliografia

ActivAge Consortium (2008). Overcoming the barriers and seizing the opportunities for active ageing policies in Europe. *International Social Science Journal*, vol. 58, nr 190, first published: December 2006), s. 617–631.

Bednarski, M., Szatur-Jaworska, B. (red.) (1999). *Wskaźniki społeczne jako narzędzia pomiaru skuteczności i efektywności polityki społecznej*. Warszawa: IPiSS.

Blackwell, M., Iacus, S., King, G., Porro, G. (2009). cem: Coarsened exact matching in Stata. *Stata Journal*, vol. 9, nr 4, s. 524.

Błędowski, P. (2012), Polityka wobec osób starych — cele i zasady. *Studia BAS*, nr 2 (30), s. 201–216.

Breza, M., Perek-Białas, J. (2014). *The Active Ageing Index and its extension to the regional level.* Host country paper, Peer Review in Poland: The Active Ageing Index and its extension to the regional level, Kraków. Available at: http://ec.europa.eu/social/BlobServlet?docId=12940&langId=en

Council of the European Union (2012). *Council Declaration on the European Year for Active Ageing and Solidarity between Generations (2012): The Way Forward*, 17468/12, SOC 992, SAN 322. Available at: http://europa.eu/ey2012/BlobServlet?docId=9231&langId=en

Golinowska, S. (red.) (2010). Wyzwania Małopolski w kontekście starzenia się społeczeństwa. Podejście strategiczne. *Małopolskie Studia Regionalne*, nr 2–3.

Górniak, J., Keler, K. (2007). Wskaźniki w ewaluacji ex-post programów publicznych. W: A. Haber (red.). *Ewaluacja ex-post. Teoria i praktyka badawcza* (s. 109–124). Warszawa: PARP.

Ho, D., Imai, K., King, G., Stuart, E. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, vol. 15, nr 3, s. 199–236.

Karpinska, K., Dykstra, P.A. (2014). The Active Ageing Index and its extension to the regional level: Discussion Paper. W: *European Commission Peer Review*. Brussels: European Commission.

Karpinska, K., Dykstra, P.A. (2015). *The Active Ageing Index and its extension to the regional level. Synthesis Report*. Brussels: European Commission.

Konarski, R., Kotnarowski, M. (2007). Zastosowanie metody propensity score matching w ewaluacji ex-post. W: A. Haber (red.). *Ewaluacja ex-post. Teoria i praktyka badawcza* (s. 183–207). Warszawa: PARP.

Kurowska, A. (2010), Metody eksperymentalne i quasi-eksperymentalne. W: B. Szatur-Jaworska (red.), *Ewaluacja w służbach społecznych* (s. 130–140). Warszawa: Mazowieckie Centrum Polityki Społecznej.

Kurowska, A. (2011). *Wskaźniki społeczne w polityce społecznej: historia, teoria i zastosowanie w praktyce*. Warszawa. Difin.

Martini, A. (2013). Różne oblicza randomizowanych prób kontrolnych. W: A. Haber, Z. Popis (red.), *(R)ewaluacja. Poszukiwanie nowych metod oceny efektów* (s. 63–78). Warszawa: PARP.

Perek-Białas, J. (2008). Refleksje na temat wykorzystania wskaźników o sytuacji starzejących się populacji — nowe wyzwania, nowe rozwiązania. W: J.T. Kowaleski, P. Szukalski (red.), *Pomyślne starzenie się w perspektywie nauk o pracy i polityce społecznej* (s. 109–125). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Perek-Białas, J. (2010). Sytuacja ludzi starych w państwach Europy Środkowej i Wschodniej w świetle wskaźników EUROSTATU. W: B. Balcerzak-Paradowska, A. Rączaszek (red. nauk.), *Międzynarodowa polityka społeczna — aspekty porównawcze* (s. 241–250). Warszawa–Katowice: Instytut Pracy i Spraw Socjalnych, Akademia Ekonomiczna w Katowicach.

Perek-Białas, J., Małodzińska, A., (2006). System wskaźników w ocenie polityki społecznej wobec ludzi starych-możliwości i ograniczenia. W: J.T. Kowaleski, P. Szukalski (red. nauk.), *Starość i starzenie się jako doświadczenie jednostek i zbiorowości ludzkich* (s. 141–149). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Perek-Białas, J., Mysińska, E. (2013). *Indeks aktywnego starzenia w ujęciu regionalnym*. Ekspertyza wykonana na zlecenie Departamentu Polityki Senioralnej, MPiPS, Warszawa, www.senior.gov.pl

Perek-Białas, J., Zwierzchowski, K. (2014). *Wskaźnik aktywnego starzenia w ujęciu regionalnym — 2014.* Ekspertyza wykonana na zlecenie Departamentu Polityki Senioralnej, MPiPS, Warszawa, www.senior.gov.pl

Portret generacji 50+ w Polsce i w Europie. Wyniki badania zdrowia, starzenia się i przechodzenia na emeryturę w Europie (SHARE) (2014). Warszawa: Instytut Badań Edukacyjnych.

Rosenbaum, P.R., Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, vol. 70, nr 1, s. 41–55.

Rubin, D.B., Thomas, N. (1996). Matching using estimated propensity scores: relating theory to practice. *Biometrics*, vol. 52, nr 1, s. 249–264.

Strawiński, P. (2008). Quasi-eksperymentalne metody ewaluacji. W: A. Haber (red.), *Środowisko i warsztat ewaluacji* (s. 193–220). Warszawa: PARP.

Stuart, E.A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, vol. 25, nr 1, s. 1–21.

Szatur-Jaworska, B. (2000). *Ludzie starzy i starość w polityce społecznej*. Warszawa: Oficyna Wydawnicza ASPRA-JR.

Szatur-Jaworska, B. (2015). Polityka senioralna w Polsce — analiza agendy. *Problemy Polityki Społecznej. Studia i Dyskusje*, nr 30 (3), s. 47–75. Dostępny w Internecie na www.problemypolitykispolecznej.pl [dostęp: 29.04.2016].

Trzciński, R. (2009). *Wykorzystanie techniki propensity score matching w badaniach ewaluacyjnych*. Warszawa: PARP.

Theiss, M. (2012). Trafna polityka lokalna? Przykład polityki społecznej wobec ludzi starych. *Problemy Polityki Społecznej. Studia i Dyskusje*, nr 7 (17), s. 149–165. Dostępny w Internecie: http://recesja.icm.edu.pl/ips/problemyps/pps17/PPS17-Theiss.pdf [dostęp: 15.05.2015].

UNECE/ European Commission (2015). *Active Ageing Index 2014: Analytical Report*, Report prepared by Asghar Zaidi of Centre for Research on Ageing, University of Southampton and David Stanton, under contract with United Nations Economic Commission for Europe (Geneva), co-funded by the European Commission's Directorate General for Employment, Social Affairs and Inclusion (Brussels).

Walker, A.C. (2002). A strategy for active ageing. *International Social Security Review*, vol. 55, nr 1, s. 121–38.

Walker, A.C., Maltby, T. (2012). Active ageing: a strategic policy solution to demographic ageing in the European Union. *International Journal of Social Welfare*, vol. 21, supl. 1, s. 117–130.

WHO (World Health Organization) (2002). *Active ageing: a policy framework*. http://www.who.int/ageing/publications/active_ageing/en/

Zaidi, A., Gasior, K., Hofmarcher, M.M., Lelkes, O., Marin, B., Rodrigues, R., Schmidt, A., Vanhuysse, P., Zolyomi, E. (2013). *Active Ageing Index 2012: Concept, Methodology and Final Results*. EC/UNECE, Active Ageing Index Project, UNECE Grant ECE/GC/2012/003, European Centre for Social Welfare Policy and Research, Vienna.

Akty normatywne

Uchwała nr 137 Rady Ministrów z dnia 24 sierpnia 2012 r. w sprawie ustanowienia Rządowego Programu na rzecz Aktywności Społecznej Osób Starszych na lata 2012–2013 (Monitor Polski 2012 r., poz. 642).

Uchwała nr 237 Rady Ministrów z dnia 24 grudnia 2013 r. w sprawie ustanowienia Rządowego Programu na rzecz Aktywności Społecznej Osób Starszych na lata 2014–2020 (Monitor Polski 2014 r., poz. 52).

Uchwała nr 238 Rady Ministrów z dnia 24 grudnia 2013 r. w sprawie przyjęcia dokumentu Założenia Długofalowej Polityki Senioralnej w Polsce na lata 2014–2020 (Monitor Polski 2014 r., poz. 118).

Strony internetowe

http://www.helpage.org

http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=76287845

http://www.senior.gov.pl

http://ec.europa.eu/social/main...

Summary

The goal of the article is to present the evaluation methods of senior policy in Poland. The article discusses manner in which senior policy in Poland is presented in official documents as well as describes it with use of standard indicators at the national and regional level. This traditional approach to policy description is extended in the article by the counterfactual analysis — an analytical method well present in public policy analysis but not used in the field of senior policy in Poland. Counterfactual analysis presented in the article is based on the data from Social Diagnosis 2015. It is concluded in the article that constant monitoring of senior policy outcomes with the use of counterfactual analysis is needed in Poland.

Key words: senior policy, old people, active ageing index, counterfactual analysis, coarsened exact matching

Cytowanie

Jolanta Perek-Białas (2016), *Tradycyjne i uzupełniające podejścia do ewaluacji efektów polityki senioralnej w Polsce*, „Problemy Polityki Społecznej. Studia i Dyskusje" nr 34(3), s. 91–111. Dostępny w Internecie na www.problemypolitykispolecznej.pl [dostęp: dzień, miesiąc, rok]