

RAPORT TEMATYCZNY Z BADANIA
ANEKS DO RAPORTU

Uwarunkowania decyzji edukacyjnych

Wyniki pierwszej rundy
badania panelowego
gospodarstw domowych



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



IBE  *entuzjaści
edukacji*

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Redakcja merytoryczna:

prof. dr hab. Małgorzata Rószkiewicz

dr Katarzyna Saczuk

Recenzenci:

prof. dr hab. Jarosław Górniak, Uniwersytet Jagielloński

dr Mikołaj Jasiński, Uniwersytet Warszawski

dr Jacek Liwiński, Uniwersytet Warszawski

dr Irena Topińska, Centrum Analiz Społeczno-Ekonomicznych CASE

Autorzy:

mgr Zuzanna Brzozowska

Beata Koń

prof. dr hab. Irena Elżbieta Kotowska

prof. SGH dr hab. Tomasz Kuszewski

dr Iga Magda

dr Barbara Minkiewicz

prof. dr hab. Tomasz Panek

dr Jolanta Perek - Białas

dr Anna Ruzik - Sierdzińska

dr Katarzyna Saczuk

prof. dr hab. Tomasz Szapiro

dr Przemysław Szufel

dr Jan Zwierchowski

Pomoc analityczna:

Marta Borawska

Paweł Ekk - Cierniakowski

Aleksandra Łagan

Redakcja językowa:

Elżbieta Tarnowska

Wydawca:

Instytut Badań Edukacyjnych

ul. Górczewska 8

01 - 180 Warszawa

tel. (22) 241 71 00; www.ibe.edu.pl

Skład:

PoPrzecinku (www.poprzecinku.com) - Karol Szczawiński, Aleksandra Gajda

© Copyright by: Instytut Badań Edukacyjnych, Warszawa 2014

Publikacja została wydrukowana na papierze ekologicznym.

Publikacja opracowana w ramach projektu systemowego: Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego, współfinansowanego przez Unię Europejską ze środków Europejskiego Funduszu Społecznego, realizowanego przez Instytut Badań Edukacyjnych.

ISBN 978-83-61-693

Egzemplarz bezpłatny

Spis treści

Aneks A - Metody analizy danych z próby

A.I. Blok I	4
A.I.1. Koncepcja analizy biografii edukacyjnych	4
A.I.2. Operacjonalizacja ścieżek edukacyjnych na podstawie danych z próby	11
A.I.2.1. Dane	11
A.I.2.2. Operacjonalizacja definicji ścieżek	11
A.I.3. Modelowanie uwarunkowań wyborów edukacyjnych	15
A.I.3.1. Konstrukcja zbioru danych	15
A.I.3.2. Metody analiz – eksploracyjna analiza czynnikowa	17
A.I.3.3. Metody analiz – wielowartościowa regresja logistyczna	22
A.II. Blok II	23
A.II.1. Narzędzia badawcze i metody analizy danych z próby	23
A.II.1.1. Testy statystyczne	24
A.II.1.2. Modele ekonometryczne	25
A.III. Blok III	25
A.III.1. Wprowadzenie	25
A.III.2. Współzależności pomiędzy wykluczeniem społecznym a edukacją	26
A.III.2.1. Cel i przedmiot analizy	26
A.III.3. Zakres analizy	28
A.III.3.1. Wykluczenie edukacyjne	28
A.III.3.2. Wpływ wykluczenia edukacyjnego na wykluczenie społeczne	29
A.III.4. Stan wiedzy	30
A.III.4.1. Edukacja i wykluczenie społeczne dzieci	30
A.III.4.2. Edukacja i wykluczenie społeczne dorosłych	35
A.III.4.3. Nakłady publiczne na edukację a ubóstwo gospodarstw domowych	37
A.III.5. Statystyczne metody analizy wykluczenia z edukacji	38
A.III.5.1. Zasięg wykluczenia z edukacji oraz nierówności w dostępie do kształcenia	38
A.III.5.2. Analiza oddziaływania różnych czynników na wykluczenie z edukacji – analiza czynnikowa	48
A.III.5.3. Analiza ścieżki oddziaływania czynników na wykluczenie z edukacji	51
A.III.5.4. Pomiar wykluczenia społecznego	57
A.III.5.5. Analiza wpływu niskiego poziomu wykształcenia na wykluczenie w wyróżnionych obszarach życia	62

A.III.6.	Wewnątrz- i międzypokoleniowe zależności kształtowania kapitału ludzkiego w gospodarstwach domowych	66
A.III.6.1.	Cel i przedmiot analizy	66
A.III.6.2.	Zakres analizy	69
A.III.6.3.	Stan wiedzy	71
A.III.6.4.	Metody badawcze	75
A.III.6.5.	Spodziewane wyniki i korzyści analityczne	76
A.IV. Blok IV		77
A.IV.1.	Wprowadzenie	77
A.IV.1.1.	Rynek edukacyjny w kontekście narzędzia badawczego	77
A.IV.1.2.	Pytania i hipotezy badawcze skutecznie weryfikowane metodami symulacji komputerowej	80
A.IV.1.3.	Założenia symulacji rynków edukacyjnych	80
A.IV.2.	Symulacja wieloagentowa	82
A.IV.2.1.	Podejście symulacyjne w modelowaniu systemów	83
A.IV.2.2.	Symulacja wieloagentowa	86
A.IV.2.3.	Wieloagentowa analiza rynków edukacyjnych	91
A.IV.2.4.	Kalibracja wieloagentowego modelu systemu edukacyjnego – dane UDE w badaniu symulacyjnym	93
A.IV.2.5.	Eksperymenty symulacyjne	94

Aneks B – Zestawienia tabelaryczne, szczegółowe wyniki analiz

B.I.Blok I: Wyniki – uwarunkowania biografii edukacyjnych.....	97	
B.II.Blok II: Wyniki	99	
B.III.Blok III: Wyniki	100	
B.IV.Blok IV: Parametryzacja modelu wieloagentowego	117	
B.IV.1.	Wartości skalibrowanych parametrów symulacyjnego wieloagentowego modelu rynku edukacji wyższej	117
B.IV.2.	Rozważane scenariusze eksperymentów symulacyjnych	119

Aneks A - Metody analizy danych z próby

A.I. Blok I

A.I.1. Koncepcja analizy biografii edukacyjnych

Biografie edukacyjne respondentów w wieku 15-65 lat, rozumiane jako sekwencja stanów edukacyjnych i czas pozostawania w tych stanach, są analizowane na podstawie ścieżki edukacyjnej. Podstawą do wyodrębnienia ścieżek edukacyjnych było podejście stosowane w badaniach GUS (2005, 2013). Zostało ono zmodyfikowane stosownie do rodzaju danych służących wyodrębnieniu ścieżek edukacyjnych (dane indywidualne o biografiach edukacyjnych) oraz zmian systemu edukacji w Polsce określających zmieniające się możliwości kształcenia. Poprzez stan edukacyjny rozumie się pozostawanie w systemie kształcenia na określonym poziomie, będące wynikiem podejmowanych decyzji edukacyjnych w przebiegu życia respondenta. W analizie ograniczamy się do formalnego systemu edukacji¹. Ścieżkę edukacyjną określa więc układ stanów edukacyjnych zdefiniowanych według określonych kryteriów dotyczących osiągania kolejnych poziomów kształcenia formalnego. Konieczne jest zatem zaproponowanie kryteriów wyodrębniania tych ścieżek. Nie jest to jednak tożsame z ustaleniem poziomu wykształcenia respondentów, ale ze sposobem dojścia do wykształcenia, którym się respondent legitymuje; np. w badaniu ścieżek edukacyjnych realizowanym przez GUS (2013, s. 35) ścieżkę edukacyjną zdefiniowano jako uporządkowaną w czasie sekwencję typów szkół, do których respondent uczęszczał i ukończył. W badaniu tym wyróżniono 6 ścieżek oraz kategorię *inne*, do której zaliczono odmienne niż zdefiniowane ścieżki kształcenia.

W naszym podejściu badawczym wyróżniono ścieżki zasadnicze i wariantowe. Ścieżka kształcenia (zasadnicza i wariantowa) jest to uporządkowana w czasie sekwencja decyzji edukacyjnych dotyczących wyboru szkoły, do której badany uczęszczał i którą ukończył, przy czym decyzję edukacyjną mogły podejmować osoby badane lub w ich imieniu (dotyczy dzieci na niższych poziomach edukacji) rodzice albo opiekunowie. Ścieżka zasadnicza obejmuje edukację w szkołach publicznych. Identyfikacja zasadniczych ścieżek kształcenia dokonano na podstawie odpowiedzi na pytania ankiety dotyczące sposobu dojścia do najwyższego uzyskanego wykształcenia. Tabela A.I.1. przedstawia propozycję wyróżnienia ścieżek zasadniczych, zaś tabela A.I.2. zawiera kryteria wyodrębniania ścieżek wariantowych. W tabeli A.I.3. umieszczono opis innych sposobów dojścia do określonego poziomu wykształcenia, które nie są ujęte w tabelach A.I.1. oraz A.I.2.

¹ Edukacja formalna obejmuje edukację szkolną i pozaszkolną, na którą składa się obowiązek szkolny i obowiązek nauki oraz edukację ponadobowiązkową, której rodzaj i forma kształcenia wynikają z kształtu systemu edukacji. Edukacja nieformalna to każda, świadomie zorganizowana, działalność oświatowa, która odbywa się poza systemem kształcenia formalnego.

Tabela A.I.1.

Zasadnicze ścieżki edukacyjne – propozycja typologii i wskaźników

Ścieżka	Opis ścieżki zasadniczej	Wskaźnik – odsetek osób z taką biografią edukacyjną
SZ1 Ścieżka zasadnicza 1 [SP (8) lub SP (6)-G]	Nauka zakończona na 8-klasowej szkole podstawowej lub na gimnazjum po 6-klasowej SP; wszystkie szkoły dla młodzieży publiczne, nauka przebiega bez przerw i wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 rozpoczęcia nauki, wiek 14-15 lat – zakończenie nauki, okres nauki 8-9 lat).	
SZ2 Ścieżka zasadnicza 2 [SP (8) lub SP (6)-G] – ZSZ	Po SZ1 osoba kończy publiczną zasadniczą szkołę zawodową dla młodzieży dwu- lub trzyletnią, nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 16-17 lat – zakończenie nauki, okres nauki: 11 lat, 12 lat).	
SZ3 Ścieżka zasadnicza 3 [SP (8) lub SP (6)-G] -LO	Po SZ1 osoba kończy naukę w publicznym liceum ogólnokształcącym lub technikum dla młodzieży, bez matury, nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 18-20 lat – zakończenie nauki, okres nauki: 12 lat lub 13 lat).	
SZ4 Ścieżka zasadnicza 4 [SP (8) lub SP (6)-G] -LO+M lub [SP (8) lub SP (6)-G] -TECH+M	Po SZ1 osoba kończy naukę w publicznym lub niepublicznym liceum ogólnokształcącym lub technikum dla młodzieży i zdaje maturę; nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 18-20 lat – zakończenie nauki, okres nauki: 12 lat lub 13 lat).	
SZ5 Ścieżka zasadnicza 6 [SP (8) lub SP (6)-G] -LO+M – LIC	Po SZ4 osoba kończy naukę na studiach licencjackich stacjonarnych w uczelni publicznej; nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów/studentów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 21-23 lata – zakończenie nauki, okres nauki: 15 lub 18 lat).	
SZ6 Ścieżka zasadnicza 6 [SP (8) lub SP (6)-G] -LO+M – LIC – MGR lub [SP (8) lub SP (6)-G] -LO+M – MGR	Po SZ5 osoba kończy naukę na studiach magisterskich stacjonarnych w uczelni publicznej. Po SZ4 osoba kończy naukę na studiach magisterskich lub równoważnych (kierunki medyczne) w uczelni publicznej; nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów/studentów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 23-26 lat – zakończenie nauki, okres nauki: 17 lub 20 lat).	

Tabela A.I.2.

Wariantowe ścieżki edukacyjne – propozycja typologii i wskaźników

Ścieżka zasadnicza	Ścieżka wariantowa i jej opis	Wskaźnik 1 – odsetek osób z taką biografią edukacyjną	Wskaźnik 2 – odsetek osób z taką ścieżką wariantową wśród osób realizujących odpowiednią zasadniczą ścieżkę edukacyjną
SZ1	SZ1W1 Nauka zakończona na 8-klasowej szkole podstawowej lub na gimnazjum po 6-klasowej SP publicznej lub niepublicznej; przynajmniej jedna szkoła niepubliczna.		
	SZ1W2 Nauka zakończona na 8-klasowej szkole podstawowej lub na gimnazjum po 6-klasowej SP publicznej lub niepublicznej; przynajmniej jedna szkoła dla dorosłych.		
SZ2	SZ2W21 Po SZ1 (w tym ścieżki wariantowe) osoba kończy publiczną lub niepubliczną zasadniczą szkołę zawodową dla młodzieży.		
	SZ2W1 Po SZ1 (w tym ścieżki wariantowe) osoba kończy publiczną lub niepubliczną zasadniczą szkołę zawodową ucząc się zawodu jako młodociany pracownik.		
SZ3	SZ3W1 Po SZ1 (w tym ścieżki wariantowe) osoba kończy naukę w niepublicznym liceum ogólnokształcącym lub technikum dla młodzieży, bez matury.		
	SZ3W2 Po SZ1 (w tym ścieżki wariantowe) osoba kończy naukę w publicznym lub niepublicznym liceum ogólnokształcącym lub technikum dla dorosłych, bez matury.		

Ścieżka zasadnicza	Ścieżka wariantowa i jej opis	Wskaźnik 1 – odsetek osób z taką biografią edukacyjną	Wskaźnik 2 – odsetek osób z taką ścieżką wariantową wśród osób realizujących odpowiednią zasadniczą ścieżkę edukacyjną
SZ4	SZ4W1 Po SZ1 (w tym ścieżki wariantowe) osoba kończy naukę w niepublicznym liceum ogólnokształcącym lub technikum dla młodzieży i zdaje maturę.		
	SZ4W2 Po SZ1 (w tym ścieżki wariantowe) osoba kończy naukę w publicznym lub niepublicznym liceum ogólnokształcącym lub technikum dla dorosłych i zdaje maturę.		
SZ5	SZ5W1 Po SZ4 (w tym ścieżki wariantowe) osoba kończy naukę na studiach licencjackich stacjonarnych lub niestacjonarnych w uczelni niepublicznej.		
SZ6	SZ6W1 Po SZ5 (w tym ścieżki wariantowe) osoba kończy naukę na studiach magisterskich stacjonarnych lub niestacjonarnych w uczelni niepublicznej.		

Tabela A.I.3.

Inne ścieżki edukacyjne

1. Ukończenie technikum po ZSZ w szkołach dla młodzieży lub dorosłych, publicznych lub niepublicznych;
2. Ukończenie innych niż wymienione jako zasadnicze ścieżki kształcenia szkół średnich, np. liceum technicznego, liceum profilowanego;
3. Ukończenie na jakimś etapie edukacji szkoły specjalnej;
4. Ukończenie szkoły policealnej;
5. Ukończenie studiów podyplomowych;
6. Osoby, które nie ukończyły danego szczebla kształcenia, w tym osoby uczące się na danym etapie edukacji.

W raporcie *Diagnoza i hipotezy badawcze* przyjęto założenie, że głównym podmiotem, który podejmuje decyzje edukacyjne, jest gospodarstwo domowe i zaproponowano listę uwarunkowań, które wpływają na te decyzje.

Wyróżniono mianowicie:

1. uwarunkowania obiektywne (np. sieć szkół w miejscu zamieszkania, koszt studiów) i subiektywne (aspiracje, wartości, postawy wobec kształcenia się i mobilności),
2. uwarunkowania wewnętrzne (skład gospodarstwa domowego, wykształcenie rodziców i inne cechy środowiska rodzinnego) i zewnętrzne w stosunku do gospodarstwa domowego (otoczenie społeczne, uwarunkowania lokalne oraz systemowe).

Odwołując się do tego podziału, uwarunkowania decyzji edukacyjnych można ująć następująco:

Tabela A.I.4.

Typologia uwarunkowań decyzji edukacyjnych na poziomie gospodarstwa domowego (GD)

Rodzaj uwarunkowań	Obiektywne	Subiektywne
Wewnętrzne	Skład GD, liczba osób; Zasoby materialne GD (warunki mieszkaniowe i wyposażenie); Status społeczny członków GD (także status na rynku pracy, wiedza i umiejętności).	Postawy i przekonania członków GD – wartości związane z kształceniem, aspiracje edukacyjne, zawodowe i materialne członków GD, postawy wobec mobilności, planowana liczba dzieci, percepcja systemu edukacyjnego i pracy zawodowej; Kapitał społeczny wewnątrz rodziny (silne więzi, częstotliwość kontaktów, zaufanie).
Zewnętrzne (otoczenie społeczne GD)	Sieć znajomości i potencjał w niej nagromadzony przekładający się na decyzje edukacyjne: – osoba, która może pomóc dziecku w lekcjach, – pracodawca, który chce inwestować w rozwój kompetencji, – znajomości ułatwiające znalezienie pracy.	

Wyodrębnione grupy czynników mogą być opisane za pomocą informacji pozyskanych w badaniu ankietowym. Potencjalny zbiór informacji do wykorzystania w analizach zawiera poniższe zestawienie pytań ankiety zamieszczonych w kwestionariuszu gospodarstwa domowego i kwestionariuszu indywidualnym.

Tabela A.I.5.

Pytania o uwarunkowania wyborów edukacyjnych

Kwestionariusz gospodarstwa domowego

Pytania w kwestionariuszu skierowane do głowy GD lub osoby najlepiej zorientowanej w sprawach GD

Dział 1. GD: charakterystyka osób mieszkających w GD

Wiek (pytanie 1.5; 1.6).

(15 i więcej lat) Stan cywilno-prawny (1.19).

Czy z powodu problemów zdrowotnych ma ograniczoną zdolność wykonywania codziennych czynności, właściwych dla wieku, trwającą 6 miesięcy lub dłużej? (1.20).

Czy posiada ważne orzeczenie ustalające niepełnosprawność, niezdolność do pracy lub inwalidztwo? (1.21). (3-18) Czy z powodu stanu zdrowia posiada orzeczenie lub opinię poradni psychologiczno-pedagogicznej o potrzebie kształcenia specjalnego? (1.22).

Status obecności w GD (1.31).

Aktywność zawodowa w bieżącym tygodniu (1.36).

Dział 2. Sytuacja dochodowa i sposób gospodarowania dochodami

Dochód netto na rękę w zł w poprzednim miesiącu w GD? (2.3).

Dochód netto na rękę w zł w GD w całym 2012 roku? (2.4).

Wartość pozostałych dochodów niepieniężnych ze wszystkich źródeł wszystkich członków GD w całym 2012 roku i w poprzednim miesiącu (2.7).

Dział 3. Wybory edukacyjne i wydatki dotyczące dzieci

Czy w GD są dzieci urodzone w 1998 roku lub później? (3.0).

Jeśli małe dziecko nie uczęszcza do żłobka/przedszkola, to w jaki sposób ma zapewnioną opiekę? Tu możliwa: bezpłatna opieka innych członków GD – więzi (3.5).

Gdzie znajduje się szkoła/przedszkole/żłobek? (3.17) Tu możliwa lokalizacja: poza miejscowością zamieszkania.

Czy dziecko korzysta z opieki przed lub po zajęciach? (3.22) Możliwości: płatna opiekunka (finanse), dziadkowie (więzi rodzinne), osoby spoza gospodarstwa (sieci społeczne).

Jak często spędza z dzieckiem czas w określony sposób? (3.41; 3.42).

Czy członkowie rodziny pomagają dziecku w lekcjach/nauce? (3.43).

Czy jakieś osoby spoza gospodarstwa regularnie pomagają dziecku w lekcjach/nauce? (3.44).

Czy dziecko pomaga – komuś z GD i spoza GD? (3.45; 3.46).

Ocena wyników w nauce (3.52; 3.53).

Kwestionariusz indywidualny

Dział I. Biografia edukacyjna

Poziom wykształcenia rodziców, gdy miał 10 lat (M_I.11).

Czy w czasie, gdy miał około 10 lat, sytuacja materialna w domu w porównaniu do innych, znanych rodzin była lepsza/podobna/gorsza? (M_I.12)

Ile książek (liczba w przybliżeniu) było w miejscu, w którym mieszkał, gdy miał 10 lat? (M_I.13)

Kto miał główny wpływ na decyzję w sprawie wyboru szkoły, w której podjął naukę na najwyższym ukończonym poziomie edukacji? (M_I.14)

Ocena ważności powodów, które miały wpływ na wybór szkoły na najwyższym ukończonym poziomie? (M_I.15)

Którym jest dzieckiem? (M_I.17)

Ile miał rodzeństwa? (M_I.18)

(dom rodzinny) Czy o jego domu rodzinnym można powiedzieć, że były/są w nim silne więzi rodzinne?

Że wspólnie spędzało/spędza się czas wolny? Że położony był/jest nacisk na pracę i naukę? (M_I.19)

(dom rodzinny) Czy rodzice lub osoba zajmująca się jego wychowaniem interesowali/interesują się jego edukacją? (M_I.20)

(dom rodzinny) Co było/jest głównym wyznacznikiem sukcesu w domu rodzinnym? (M_I.21)

Uwagi: Ponieważ zagadnienia, które są przewidziane w wywiadach indywidualnych członków GD w wieku 15-65 w kolejnych rundach, w pierwszej rundzie wyczerpuje w całości kwestionariusz modułu I (Kalendarz retrospektywny), nie są one powtarzane w kwestionariuszu indywidualnym dla I rundy.

W naszej analizie ograniczymy się do respondentów w wieku 15-65 lat, przy czym ze szczególną uwagą będziemy się zajmować osobami dorosłymi, czyli osobami w wieku 20-65 lat. Analiza ścieżek edukacyjnych oraz ich uwarunkowań zostanie przeprowadzona na podstawie następującej procedury:

1. operacjonalizacja zaproponowanej typologii ścieżek edukacyjnych;
2. wyodrębnienie grup respondentów, których przebieg biografii edukacyjnej charakteryzuje ścieżka SZI dla $i=1,2,\dots,r$. W jej wyniku uzyskamy podział populacji N respondentów na r grup o licznosci N_i . Uzyskamy zatem informację o tym, jak często były realizowane określone ścieżki (N_i/N);
3. analiza opisowa respondentów według wyodrębnionych ścieżek oraz cech respondentów je realizujących, cech ich gospodarstwa domowego oraz cech ich domu rodzinnego mająca na celu ukazanie kontekstu społeczno-ekonomicznego wyborów edukacyjnych. Ten etap analiz opisowych, nie tylko pokaże rozkład populacji respondentów według przebiegu ich edukacji charakteryzowanej za pomocą typu ścieżki edukacyjnej, ale także umożliwi wstępne rozpoznanie zmiennych, które mogły opisywać czynniki wpływające na realizację określonej ścieżki SZI;
4. identyfikacja głównych uwarunkowań decyzji edukacyjnych poprzez estymację wielowartościowych modeli logistycznych.

A.1.2. Operacjonalizacja ścieżek edukacyjnych na podstawie danych z próby

A.1.2.1. Dane

Do wyodrębnienia grup osób realizujących zdefiniowane ścieżki wykorzystane zostały dane z kwestionariusza gospodarstwa (zbiór *P*) o ukończonym wykształceniu respondentów oraz dane z kwestionariusza indywidualnego I rundy badania - moduł I. Kalendarz retrospektywny (zbiór *I*).

Ponieważ dane w zbiorze *P* pochodzą od głowy gospodarstwa lub innej osoby dobrze poinformowanej o sytuacji wszystkich członków gospodarstwa, a w zbiorze *I* od respondentów, mogą występować rozbieżności dotyczące wykształcenia badanych. Z zestawienia danych z obu kwestionariuszy wynika, że rozbieżności dotyczące wykształcenia dotyczą około 2% respondentów (mierzonych zarówno liczbą rekordów, jak i w danych ważonych). Ponieważ nie ma możliwości zweryfikowania danych w żadnym z kwestionariuszy, osoby, dla których odnotowano rozbieżności, zostały wykluczone z analiz.

W celu zwiększenia liczby rekordów do analiz ścieżek wykształcenia wykorzystano wyłącznie dane o latach rozpoczęcia i zakończenia edukacji. W zbiorze danych znajdują się również te dotyczące miesięcy, ale względnie duża liczba respondentów, biorąc pod uwagę organizację polskiego systemu edukacji (tj. ujęcie edukacji formalnej w lata szkolne rozpoczynające się na większości etapów najczęściej we wrześniu, czasem w lutym, a kończące - w czerwcu, czasem styczniu), podawała inne miesiące rozpoczęcia edukacji niż wrzesień i luty oraz inne jej zakończenia niż maj, czerwiec czy styczeń. Gdyby wykorzystać dane o miesiącach, należałoby te rekordy odrzucić.

A.1.2.2. Operacjonalizacja definicji ścieżek

Analiza zdarzeń. Sposób zdefiniowania (poprzez dodanie szczegółowych warunków sposobu osiągania kolejnych etapów edukacji – np. w szkole publicznej, bez przerw, warunki na czas nauki w szkole danego typu) sprawia, że opisanie ścieżek edukacyjnych za pomocą zdefiniowanej przestrzeni stanów byłoby mocno nienaturalne i utrudnione technicznie. Ponadto opis zachowań respondentów standardowymi narzędziami sekwencji zdarzeń, nawet przy stosunkowo krótkich sekwencjach stanów, mógłby się okazać zbyt skomplikowany (za dużo możliwych sekwencji), a dane z badania dalece niewystarczające do wykorzystania narzędzi statystycznych. Z tego powodu zdecydowano się sprawdzić podane przez respondentów historie edukacyjne pod kątem zadanych ścieżek i przypisać im numery ścieżek, które odpowiadają podanym wyżej definicjom. Oznacza to, że na tym etapie w danych analizowane są pewne wybrane z góry ścieżki, a nie poszukuje się ścieżek w podanych historiach edukacyjnych. Standardowa analiza sekwencji zdarzeń zostanie wykorzystana do analizy danych z badania; przy inaczej zdefiniowanych ścieżkach edukacyjnych – na kolejnym etapie analiz wyników badania.

Definicje ścieżek. Aby wykorzystać większą liczbę rekordów, rozluźniono założenia w definicjach ścieżek w stosunku do pierwotnie określonych (powyżej). Poniżej podano najważniejsze założenia i uproszczenia w stosunku do zdefiniowanych ścieżek.

Techniczny opis poszczególnych ścieżek, na podstawie danych z próby, zestawiono w tabeli A.I.6.

Wykształcenie. Dla wszystkich ścieżek wykształcenie ukończone przyjęto na podstawie zmiennej P23 z kwestionariusza gospodarstwa; z zestawień wykluczone zostały osoby, dla których dane o wykształceniu z kwestionariusza gospodarstwa były niespójne z danymi z kalendarza retrospektywnego.

Przerwy w edukacji. Przy definiowaniu wszystkich ścieżek za przerwę w nauce uznano przerwę trwającą dłużej niż rok (na podstawie lat rozpoczęcia i zakończenia poszczególnych etapów edukacji). Takie ujęcie pozwala elastyczniej podejść do kolejnych etapów edukacji oraz uwzględnić, dla wyższych etapów edukacji, możliwość rozpoczynania edukacji w semestrze letnim. Ponadto pozwala ono częściowo zniwelować błędy popełniane przez respondentów (niesprawdzone przez ankieterów), przy raportowaniu kolejnych etapów edukacji – skala przerw w edukacji wynikająca z badania, również w przypadku osób ze względu na wiek objętych obowiązkiem szkolnym, nie znajduje odzwierciedlenia w innych źródłach danych.

Wiek zakończenia nauki. W technicznych definicjach ścieżek wiek zakończenia edukacji potraktowano wynikowo (tzn. nie był kontrolowany). Definiowano jedynie wiek rozpoczęcia edukacji, przerwy w edukacji oraz czas jej trwania na poszczególnych poziomach.

Ścieżki a kontynuacja nauki. Ścieżki zostały zdefiniowane w sposób, który sugeruje opis pewnego dokonanego procesu. Tymczasem w próbie, ze względu na ograniczenie wieku respondentów (od 15 lat), znajduje się duża grupa osób, które są w trakcie nabywania edukacji formalnej, co w naturalny sposób rodzi pytanie, jak takie osoby traktować. Gdyby literalnie przyjąć podane definicje ścieżek, należałoby wszystkie osoby kontynuujące edukację wykluczyć z 6 zasadniczych ścieżek – czas ich edukacji jest dłuższy niż właściwy dla poziomu edukacji, który takie osoby ukończyły, a jeśli chodzi o poziom edukacji na którym się kształcą, nie wiadomo czy zostanie on osiągnięty. Wydaje się natomiast, że wykluczenie tych osób z analizy – tzn. uwzględnienie ich tylko w kategorii SZ7 – (zwłaszcza w przypadku osób poniżej 25 lat) mogłoby zaburzać wyniki, szczególnie dotyczące najmłodszych grup wiekowych. Z tego powodu zdecydowano się potraktować osoby kontynuujące naukę tak, jak wszystkie pozostałe. Przypisane im ścieżki odpowiadają najwyższemu osiągniętemu przez te osoby dotychczas poziomowi wykształcenia (jeśli spełnione są pozostałe warunki w definicji ścieżek).

Tabela A.I.6.

Ścieżki edukacyjne wyodrębnione na podstawie danych z próby

SZ1	
Ścieżka zasadnicza 1 [SP (8) lub SP (6)-G]	
Opis ścieżki	Nauka zakończona na 8-klasowej szkole podstawowej lub na gimnazjum po 6-klasowej SP; wszystkie szkoły dla młodzieży publiczne, nauka przebiega bez przerw i wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 rozpoczęcia nauki, wiek 14-15 lat – zakończenie nauki, okres nauki 8-9 lat).
Operacjonalizacja	Ukończona: i) szkoła podstawowa (założony czas nauki w szkole podstawowej, jeśli respondent nie uczęszczał do gimnazjum – 7, 8 lat lub 9 lat, ze względu na sposób kodowania – w kalendarzu retrospektywnym szkoła podstawowa została zakodowana razem z zerówką; nie ma możliwości odróżnienia kiedy respondent chodził do zerówki, a kiedy nie) lub ii) gimnazjum (założony czas nauki - 6 lub 7, ze względu na zerówkę, + 3 lub 4 lata); rozpoczęcie w wieku 5 (ze względu na zerówkę), 6 lub 7 lat. UWAGA: ukończenie szkoły podstawowej lub gimnazjum, zgodnie z powyższym opisem, zostało identycznie zakodowane we wszystkich ścieżkach.
SZ2	
Ścieżka zasadnicza 2 [SP (8) lub SP (6)-G] – ZSZ	
Opis ścieżki	Po SZ1 osoba kończy publiczną zasadniczą szkołą zawodową dla młodzieży dwu- lub trzyletnią, nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 16-17 lat – zakończenie nauki, okres nauki: 11 lat, 12 lat).
Operacjonalizacja	Ukończona zasadnicza szkoła zawodowa bez przerwy po szkole podstawowej lub gimnazjum. Założony czas nauki w szkole zawodowej – od 2 do 5 lat.
SZ3	
Ścieżka zasadnicza 3 [SP (8) lub SP (6)-G] -LO	
Opis ścieżki	Po SZ1 osoba kończy naukę w publicznym liceum ogólnokształcącym lub technikum dla młodzieży, bez matury, nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 18-20 lat – zakończenie nauki, okres nauki: 12 lat lub 13 lat).
Operacjonalizacja	Ukończone LO bez matury, liceum profilowane bez matury (w kalendarzu retrospektywnym nie ma rozróżnienia na LO i liceum profilowane – nie ma sposobu, aby sprawdzić dane z kwestionariusza gospodarstwa) lub technikum bez matury, bez przerwy po szkole podstawowej, gimnazjum lub szkole zasadniczej zawodowej. Założony czas nauki w szkołach, liceach i technikum – 3 (licea były kiedyś 3-letnie), 4 lub 5 lat. W definicji ścieżki dopuszczono również szkoły zasadnicze zawodowe jako etap przed szkołami średnimi z maturą. Założony czas nauki w szkole zasadniczej zawodowej, jeśli respondent do takiej uczęszczał – od 2 do 5 lat.

SZ4

Ścieżka zasadnicza 4 [SP (8) lub SP (6)-G] -LO+M lub [SP (8) lub SP (6)-G] -TECH+M

Opis ścieżki	Po SZ1 osoba kończy naukę w publicznym lub niepublicznym liceum ogólnokształcącym lub technikum dla młodzieży i zdaje maturę; nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 18-20 lat – zakończenie nauki, okres nauki: 12 lat lub 13 lat).
Operacjonalizacja	Ukończone LO z maturą, liceum profilowane z maturą (w kalendarzu retrospektywnym nie ma rozróżnienia na LO i liceum profilowane – nie ma sposobu, aby sprawdzić dane z kwestionariusza gospodarstwa) lub technikum z maturą, bez przerwy po szkole podstawowej, gimnazjum lub szkole zasadniczej zawodowej. Założony czas nauki w szkołach, liceach i technikum – 3 (licea były kiedyś 3-letnie), 4 lub 5 lat. W definicji ścieżki dopuszczono również szkoły zasadnicze zawodowe jako etap przed szkołami średnimi z maturą. Założony czas nauki w szkole zasadniczej zawodowej, jeśli respondent do takiej uczęszczał – od 2 do 5 lat.

SZ5

Ścieżka zasadnicza 6 [SP (8) lub SP (6)-G] -LO+M – LIC

Opis ścieżki	Po SZ4 osoba kończy naukę na studiach licencjackich stacjonarnych w uczelni publicznej; nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów/studentów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 21-23 lata – zakończenie nauki, okres nauki: 15 lub 18 lat).
Operacjonalizacja	Osoby z dyplomem licencjata lub inżyniera (w kalendarzu retrospektywnym nie ma rozróżnienia na studia licencjackie i inżynierskie – ujęte są razem). Wcześniej ukończone LO, liceum profilowane lub technikum. Wcześniej ukończona szkoła podstawowa lub gimnazjum. W definicji ścieżki dopuszczono również szkoły policealne lub kolegia jako etap przed szkołami wyższymi oraz szkoły zasadnicze zawodowe jako etap przed szkołami średnimi z maturą. Założony czas nauki w szkołach, liceach i technikum – 3 do 5 lat; czas nauki na studiach licencjackich – 2 do 4 lat; w szkole zasadniczej zawodowej, jeśli respondent do takiej uczęszczał – od 2 do 5 lat. Brak warunków na czas nauki w szkołach policealnych, jeśli respondent do takich uczęszczał.

SZ6

Ścieżka zasadnicza 6 [SP (8) lub SP (6)-G] -LO+M – LIC – MGR lub [SP (8) lub SP (6)-G] -LO+M – MGR

Opis ścieżki	Po SZ5 osoba kończy naukę na studiach magisterskich stacjonarnych w uczelni publicznej lub po SZ4 osoba kończy naukę na studiach magisterskich lub równoważnych (kierunki medyczne) w uczelni publicznej; nauka przebiega bez przerw, wiek osób odpowiada wiekowi uczniów/studentów realizujących taką ścieżkę edukacyjną (wiek 6-7 lat – rozpoczęcie nauki, wiek 23-26 lat – zakończenie nauki, okres nauki: 17 lub 20 lat).
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Operacjonalizacja	Osoby z dyplomem magistra lub lekarza (w kalendarzu retrospektywnym nie ma rozróżnienia – ujęte są razem) po ukończonych studiach licencjackich lub inżynierskich lub ukończonym LO, liceum profilowanym lub technikum. Wcześniej ukończona szkoła podstawowa lub gimnazjum. W definicji ścieżki dopuszczono również szkoły policealne lub kolegia jako etap przed szkołami wyższymi oraz szkoły zasadnicze zawodowe jako etap przed szkołami średnimi z maturą. Założony czas nauki w szkołach liceach i technikach – 3 do 5 lat; czas nauki na studiach licencjackich – 2 do 4 lat; na studiach magisterskich – 2 do 6 lat; w szkole zasadniczej zawodowej, jeśli respondent do takiej uczęszczał – od 2 do 5 lat. Brak warunków na czas nauki w szkołach policealnych, jeśli respondent do takich uczęszczał.
-------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

SZ7

Pozostałe sposoby realizacji edukacji

Opis ścieżki	Wszystkie osoby, które realizowały inne ścieżki niż zdefiniowane w SZ1- SZ6.
Operacjonalizacja	Wszystkie osoby, które realizowały inne ścieżki niż zdefiniowane w SZ1-SZ6; wykluczone osoby, dla których niespójne są dane dotyczące wykształcenia z kwestionariusza gospodarstwa i kalendarza retrospektywnego oraz rekordy niekompletne ze względu na zmienne definiujące ścieżki.

A.1.3. Modelowanie uwarunkowań wyborów edukacyjnych

A.1.3.1. Konstrukcja zbioru danych

Oryginalna baza danych posłużyła do utworzenia zbioru wykorzystywanego w analizie modelowej wyborów edukacyjnych. Poniżej przedstawione są kolejne kroki przetwarzania bazy danych wyjściowych na bazę danych do obliczeń.

Dane wyeliminowane

Z wyjściowej bazy danych usunięto rekordy dotyczące:

- osób, które kontynuują edukację,
- osób, które realizowały ścieżkę 7,
- osób, które miały braki danych w ścieżkach,
- osób, które zrealizowały ścieżkę 1 (szkoła podstawowa i gimnazjum), ponieważ obecnie prawie nikt nie kończy edukacji na tym poziomie, a ponadto pytania dotyczące kryteriów wyboru szkoły nie zostały tym osobom zadane.

Ostatecznie uzyskano liczbę obserwacji $N= 23479$ (12497 kobiet i 10982 mężczyzn).

Wszystkie analizy modelowe przeprowadzono osobno dla kobiet i mężczyzn.

Tabela A.I.7.

Rekodowanie zmiennych

Zmienna	Pierwotnie	Po zrekodowaniu	Nazwa zmiennej po zrekodowaniu
M11_A, B: Wykształcenie rodziców	17 kategorii	4 kategorie	edu_m_rec (matka) edu_o_rec (ojciec)
M13: Książki	6 kategorii	4 kategorie	książki
M15_	5 kategorii	5 kategorii w kolejności: 1. zupełnie nieważne, 2. raczej nieważne, 3. ani ważne ani nieważne, 4. raczej ważne, 5. bardzo ważne	M15_znaczenie powodów wyboru szkoły
M18: liczba rodzeństwa	0-16 (zmienna ciągła)	0-4+ (zmienna na skali porządkowej)	rodz
M17: kolejność wśród rodzeństwa	1-17 (zmienna ciągła)	1. najstarszy – 2. średni – 3. najmłodszy (zmienna na skali porządkowej); jedynacy przekodowani na brak danych, bo dla nich ta zmienna nie ma sensu	kolej
M20: zainteresowanie opiekunów edukacją	3 kategorie	3 kategorie w kolejności: 1. zbyt małe, 2. odpowiednie, 3. zbyt duże	M20
P24: rok uzyskania najwyższego wykształcenia	zmienna interwałowa	Porządkowa, 3 kat: 1. przed 1990, 2. 1990-99, 3. 2000-13	koniec
Ścieżki edukacyjne	7 kategorii (6 + <i>inne</i>)	4 kat. szkoła podstawowa i gimnazjum, zasadnicza szkoła zawodowa, szkoła średnia (liceum + technikum), studia (licencjat + magisterium); <i>inne</i> nieuwzględnione; szkoła podstawowa i gimnazjum wyrzucone z analiz	ściezka

Tabela A.I.8.

Nowe zmienne

Zmienna	Kodowanie	Cel
homog: Homogamia rodziców	-1: matka niższe wykształcenie niż ojciec 0: matka i ojciec takie samo wykształcenie (kategoria referencyjna w modelu) 1: matka wyższe wykształcenie niż ojciec	Do modelu regresji: wykształcenie rodziców silnie skorelowane – włączenie do modelu wykształcenia jednego z rodziców i informacji o homogamii pozwala uwzględnić informacje o wykształceniu obojga rodziców bez naruszania założeń modelu.

A.1.3.2. Metody analiz – eksploracyjna analiza czynnikowa

Analiza czynnikowa miała na celu zidentyfikowanie uwarunkowań wewnętrznych decyzji edukacyjnych o charakterze subiektywnym oraz ograniczenie liczby zmiennych, które mogą być uwzględnione w modelach regresji logistycznej.

Wśród uwarunkowań subiektywnych kształtujących wybory edukacyjne można wyróżnić takie, które dotyczą kapitału rodzinnego, czyli systemu wartości i relacji w rodzinie, oraz związane z cechami respondenta, czyli postawami respondenta wobec własnego rozwoju i kształcenia się, jego zdolnościami. Eksploracja tej dwuwymiarowej struktury uwarunkowań subiektywnych może być dokonana na podstawie zmiennych z przygotowanej bazy danych. Schemat conceptualny operacjonalizacji poszczególnych wymiarów przedstawia się następująco:

Uwarunkowania wewnętrzne subiektywne

System wartości i relacji w rodzinie	Postawy respondenta wobec własnego rozwoju i kształcenia się, jego zdolności
<p>Cechy domu rodzinnego respondenta (M19_):</p> <ul style="list-style-type: none">a) były silne więzi rodzinne;b) wspólnie spędzało się czas wolny;c) położony był nacisk na pracę i naukę. <p>Główny wyznacznik sukcesu w domu rodzinnym (M21_):</p> <ul style="list-style-type: none">d) rodzina;e) prestiżowa praca;f) wysoki status materialny;g) dobre wykształcenie. <p>Stopień zainteresowania rodziców edukacją (zrekodowana M20).</p> <p><i>Uwaga: pominięto kategorie 'żadne z powyższych', 'nie dotyczy', 'nie wiem/nie pamiętam'.</i></p>	<p>Kryteria wyboru szkoły (najwyższy poziom wykształcenia) – zrekodowana M15_:</p> <ul style="list-style-type: none">a) odległość od miejsca zamieszkania/łatwość dojazdu;b) prestiż szkoły;c) jakość nauczania, wysoki poziom;d) profil zgodny z zainteresowaniami respondenta;e) możliwość znalezienia dobrej pracy po skończeniu szkoły;f) to, czy do szkoły chodzą znajomi;g) to, czy do szkoły chodzi/ło rodzeństwo. <p><i>Uwaga: pominięto kategorię 'inne' (M15_08).</i></p>

Do zbioru zmiennych wskaźnikowych dotyczących relacji w rodzinie włączono też zmienną M14: *Kto decydował o wyborze szkoły?*, jednak po wstępnych analizach ostatecznie zdecydowano o jej pominięciu.

Wyboru zmiennych do eksploracyjnej analizy czynnikowej dokonano na podstawie tablicy korelacji polichorycznych, które stosuje się w przypadku zmiennych porządkowych o krótkiej skali². Braki danych nie zostały uwzględnione przy liczeniu korelacji (zostały pominięte, nie przypisano im żadnej wartości).

² Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443-460.

Tabela A.I.9.

Tablica korelacji polichorycznych – kobiety

	M15_01	M15_02	M15_03	M15_04	M15_05	M15_06	M15_07	M19_1	M19_2	M19_3	M20	M21_1	M21_2	M21_3	M21_4
M15_01	1.00														
M15_02	0.35	1.00													
M15_03	0.28	0.82	1.00												
M15_04	0.09	0.42	0.49	1.00											
M15_05	0.17	0.37	0.41	0.54	1.00										
M15_06	0.31	0.13	0.10	-0.07	0.04	1.00									
M15_07	0.29	0.18	0.14	-0.07	-0.02	0.64	1.00								
M19_1	0.08	0.20	0.19	0.19	0.12	0.03	0.10	1.00							
M19_2	0.05	0.15	0.17	0.18	0.11	0.01	0.07	0.89	1.00						
M19_3	0.07	0.22	0.23	0.16	0.10	-0.02	0.04	0.54	0.55	1.00					
M20	-0.04	0.11	0.12	0.12	0.05	-0.08	-0.08	0.37	0.34	0.40	1.00				
M21_1	0.11	0.06	0.07	0.06	0.04	0.06	0.10	0.57	0.50	0.23	0.06	1.00			
M21_2	0.01	0.12	0.09	0.07	0.13	0.06	0.09	-0.04	0.00	0.14	0.04	-0.25	1.00		
M21_3	0.01	0.09	0.09	0.04	0.05	0.06	0.07	-0.15	-0.12	0.05	0.00	-0.15	0.51	1.00	
M21_4	-0.06	0.20	0.23	0.19	0.10	-0.13	-0.13	0.15	0.16	0.29	0.21	-0.30	0.27	0.38	1.00

Tabela A.I.10.

Tablica korelacji polichorycznych – mężczyźni

	M15_01	M15_02	M15_03	M15_04	M15_05	M15_06	M15_07	M19_1	M19_2	M19_3	M20	M21_1	M21_2	M21_3	M21_4
M15_01	1.00														
M15_02	0.40	1.00													
M15_03	0.35	0.82	1.00												
M15_04	0.13	0.39	0.44	1.00											
M15_05	0.16	0.34	0.37	0.58	1.00										
M15_06	0.34	0.19	0.16	-0.03	0.05	1.00									
M15_07	0.32	0.24	0.23	-0.11	-0.08	0.60	1.00								
M19_1	0.09	0.18	0.20	0.17	0.11	0.05	0.06	1.00							
M19_2	0.09	0.17	0.18	0.16	0.12	0.00	0.06	0.87	1.00						
M19_3	0.10	0.24	0.28	0.16	0.12	0.01	0.08	0.58	0.57	1.00					
M20	-0.01	0.11	0.12	0.14	0.06	-0.05	-0.08	0.37	0.32	0.38	1.00				
M21_1	0.11	0.10	0.10	0.04	0.00	0.05	0.10	0.52	0.50	0.26	0.09	1.00			
M21_2	0.00	0.09	0.09	0.06	0.11	0.06	0.08	0.03	0.04	0.15	0.04	-0.29	1.00		
M21_3	-0.01	0.06	0.08	0.04	0.06	0.00	0.02	-0.12	-0.16	0.03	0.00	-0.23	0.42	1.00	
M21_4	-0.04	0.20	0.24	0.17	0.13	-0.09	-0.08	0.12	0.12	0.29	0.21	-0.23	0.28	0.40	1.00

Na podstawie wyników korelacji polichorocznnych włączono do analizy czynnikowej wszystkie rozpatrywane zmienne. Oszacowań dokonano metodą *principal axis factoring*³, stosując rotację Varimax. Test Kaisera-Meyera-Olkina (KMO) wykazał, że macierz korelacji jest co najmniej wystarczająco adekwatna i stosowanie analizy czynnikowej jest uzasadnione (Tabela A.I.11).

Tabela A.I.11.

Wartości współczynnika KMO

Zmienne	kobiety	mężczyźni
M15_01dojazd	0.82	0.85
M15_02prestiz	0.67	0.68
M15_03poziom	0.67	0.68
M15_04zainteres	0.77	0.73
M15_05praca	0.76	0.71
M15_06znajomi	0.60	0.63
M15_07rodz	0.61	0.64
M19_1wiezi_rodz	0.64	0.65
M19_2wspolny_czas	0.64	0.65
M19_3nauka_praca	0.80	0.81
M20_zainteres	0.76	0.76
M21_1sukces_rodzina	0.71	0.74
M21_2sukces_praca	0.62	0.64
M21_3sukces_statusmat	0.59	0.63
M21_4sukces_wykszt	0.69	0.71
Ogółem	0.68	0.69

Stosując kryterium Kaisera, przyjęto za istotne cztery czynniki (Tabela A.I.12).

Tabela A.I.12.

Kryterium Kaisera

	Wartości własne	
	kobiety	mężczyźni
PA1	3.53	3.59
PA2	2.38	2.34
PA3	2.05	2.05
PA4	1.63	1.57
PA5	0.94	0.90
PA6	0.79	0.81
PA7	0.71	0.66

³ Fabrigar, L. R. Wegener, D. T. MacCallum, R. C. Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299.

PA8	0.65	0.64
PA9	0.57	0.59
PA10	0.47	0.45
PA11	0.43	0.41
PA12	0.34	0.36
PA13	0.25	0.34
PA14	0.17	0.17
PA15	0.10	0.12

Wyniki eksploracyjnej analizy czynnikowej dla kobiet i mężczyzn są bardzo podobne (por. poniższe tabele).

Tabela A.I.13.

Wyniki analizy czynnikowej dla kobiet (N=12 497)

	PA1	PA2	PA3	PA4
M15_01dojazd	0.01	0.28	0.39	-0.05
M15_02prestiz	0.11	0.80	0.22	0.10
M15_03poziom	0.11	0.86	0.15	0.09
M15_04zainteres	0.14	0.62	-0.11	0.04
M15_05praca	0.06	0.55	-0.01	0.04
M15_06znajomi	-0.03	0.00	0.77	0.02
M15_07rodz	0.04	0.00	0.81	0.04
M19_1wiezi_rodz	0.95	0.12	0.08	-0.15
M19_2wspolny_czas	0.90	0.09	0.04	-0.09
M19_3nauka_praca	0.63	0.15	0.00	0.19
M20zainteres	0.42	0.08	-0.12	0.13
M21_1sukces_rodzina	0.48	0.03	0.15	-0.43
M21_2sukces_praca	0.02	0.05	0.09	0.62
M21_3sukces_statusmat	-0.06	0.03	0.08	0.64
M21_4sukces_wykszt	0.21	0.19	-0.20	0.60
SS loadings	2.60	2.22	1.58	1.46
Proportion Var	0.17	0.15	0.11	0.10
Cumulative Var	0.17	0.32	0.43	0.52
Proportion Explained	0.33	0.28	0.20	0.19
Root mean square of the residuals			0.04	
Fit based upon off diagonal values			0.97	
Measures of factor score adequacy				
Correlation of scores with factors	0.98	0.92	0.89	0.84
Multiple R square of scores with factors	0.95	0.85	0.79	0.71
Minimum correlation of possible factor scores	0.90	0.70	0.58	0.43

Tabela A.I.14.

Wyniki analizy czynnikowej dla mężczyzn (N=10 982)

	PA1	PA2	PA3	PA4
M15_01dojazd	0.04	0.30	0.44	-0.06
M15_02prestiz	0.13	0.74	0.35	0.09
M15_03poziom	0.15	0.78	0.30	0.11
M15_04zainteres	0.12	0.66	-0.13	0.04
M15_05praca	0.06	0.57	-0.06	0.06
M15_06znajomi	-0.02	0.02	0.67	0.00
M15_07rodz	0.03	-0.05	0.84	0.03
M19_1wiezi_rodz	0.93	0.08	0.06	-0.10
M19_2wspolny_czas	0.89	0.08	0.04	-0.11
M19_3nauka_praca	0.66	0.16	0.06	0.19
M20zainteres	0.41	0.10	-0.10	0.12
M21_1sukces_rodzina	0.48	0.04	0.12	-0.46
M21_2sukces_praca	0.05	0.03	0.07	0.58
M21_3sukces_statusmat	-0.08	0.03	0.02	0.61
M21_4sukces_wykszt	0.19	0.19	-0.12	0.61
SS loadings	2.60	2.10	1.64	1.40
Proportion Var	0.17	0.14	0.11	0.09
Cumulative Var	0.17	0.31	0.42	0.52
Proportion Explained	0.34	0.27	0.21	0.18
Cumulative proportion	0.34	0.61	0.82	1.00
Root mean square of the residuals			0.04	
Fit based upon off diagonal values			0.98	
Measures of factor score adequacy				
Correlation of scores with factors	0.96	0.91	0.89	0.83
Multiple R square of scores with factors	0.92	0.82	0.80	0.70
Minimum correlation of possible factor scores	0.84	0.64	0.60	0.39

Zidentyfikowano zatem cztery wymiary subiektywnych uwarunkowań wewnętrznych, na które składają się głównie (wymiaru wymienione według „ważności”; po średniku wymieniono zmienne istotne, ale znacznie mniej niż te, podane przed średnikiem):

1. relacje w domu rodzinnym (PA1): silne więzi rodzinne, spędzanie wspólnie czasu w domu rodzinnym; nacisk na naukę i pracę, rodzina jako wyznacznik sukcesu w domu rodzinnym oraz odpowiednie zainteresowanie rodziców edukacją (M19_1, M19_2; M19_3, M21_1 i M20).

2. nastawienie respondenta na samorealizację i wykształcenie (PA2): prestiż szkoły, poziom nauczania, możliwość znalezienia dobrej pracy i profil zgodny z zainteresowaniami jako kryteria wyboru szkoły; łatwość dojazdu do szkoły (M15_03, M15_02, M15_05 i M15_04; M15_01).
3. doświadczenia rodziny/znajomych (PA3): chodzenie do szkoły rodziców/rodzeństwa lub znajomych; dobry dojazd jako kryteria wyboru szkoły (M15_07 i M15_06; M15_01).
4. nastawienie na sukces w domu rodzinnym (PA4): prestiżowa praca, wysoki status materialny i dobre wykształcenie jako wyznaczniki sukcesu w domu rodzinnym, niewymienienie rodziny jako wyznacznika sukcesu (M21_2, M21_3, M21_4, M21_1).

A.1.3.3. Metody analiz – wielowartościowa regresja logistyczna

Rozpatrujemy zmienną losową Y , która może przyjąć jedną z kategorii $j=1, 2, \dots, J$. W modelu regresji logistycznej dla zmiennej zależnej o wielu kategoriach (*multinomial logit model*) zakłada się, że logarytm ilorazu szans (*log odds*) zmiennych zależnych Y_i wyraża się modelem liniowym:

$$\eta_{ij} = P(Y_i = j) = \log(p_{ij}/p_{i,J}) = \alpha_j + x_i \beta_j$$

gdzie α_j – stała, a β_j – wektor współczynników regresji przy zmiennych objaśniających x_i dla $j=1, 2, \dots, J-1$ ⁴.

W naszych analizach zmienna Y przyjmuje kategorie odpowiadające wyodrębnionym ścieżkom edukacyjnym. Model regresji logistycznej dla zmiennej zależnej o wielu kategoriach opisuje wybór ścieżki edukacyjnej przez $i=1, \dots, N$ respondentów, zatem $P(Y_i=j)$ oznacza prawdopodobieństwo wyboru ścieżki j przez respondenta i , a wektor zmiennych x_i zawiera zmienne, które mogą wpływać na wybór określonej ścieżki edukacyjnej.

Ustalono, że możliwe kategorie zmiennej objaśnianej to następujące ścieżki edukacyjne: ukończenie edukacji na poziomie zasadniczej szkoły zawodowej (SZ2), ukończenie edukacji na poziomie szkoły średniej (czyli ścieżki SZ3 i SZ4) i ukończenie edukacji na poziomie szkoły wyższej (czyli ścieżki SZ5 i SZ6). Modele oszacowano dla następujących kombinacji zmiennej objaśnianej: szanse skończenia szkoły średniej vs. zasadniczej szkoły zawodowej, szanse skończenia studiów vs. zasadniczej szkoły zawodowej oraz szanse skończenia studiów vs. szkoły średniej.

Zbiór zmiennych objaśniających zawierał zmienne charakteryzujące uwarunkowania wewnętrzne subiektywne reprezentowane przez zmienne z analizy czynnikowej (zmienne PA) oraz uwarunkowania obiektywne reprezentowane przez zmienne z tabeli A.1.7., a także zmienną M12 dotyczącą sytuacji materialnej w domu rodzinnym respondenta w okresie, gdy miał 10 lat w porównaniu z sytuacją w innych rodzinach. Dokonano oszacowań sześciu modeli dla kobiet i mężczyzn. Wybór ostatecznej postaci modelu dokonany był na podstawie wartości miernika AIC (Akaike Information Criterion). Poniżej podane są kolejne wersje modeli.

⁴ Rodriguez, G. *Generalized Linear Model*, rozdz. 6.2, Office of Population Research, Princeton University, <http://data.princeton.edu/wws509/notes/c6.pdf>.

Najpierw oszacowano modele:

model 1: $\text{edu_m_rec} + \text{homog} + \text{rodz} + \text{kolej} + \text{koniec} + \text{ksiazki2} + \text{M12sytuacja_mat}$

model 2: $\text{edu_o_rec} + \text{homog} + \text{rodz} + \text{kolej} + \text{koniec} + \text{ksiazki2} + \text{M12sytuacja_mat}$

Wybrano model 1, bowiem lepsze wyniki daje uwzględnienie wykształcenia matki niż ojca.

Następnie oszacowano modele:

model 3: $\text{edu_m_rec} + \text{homog} + \text{rodz} + \text{kolej} + \text{koniec} + \text{ksiazki2} + \text{M12sytuacja_mat} + \text{PA1} + \text{PA2} + \text{PA3} + \text{PA4} + \text{PA5}$

model 4: $\text{model 3} + \text{edu_m_rec} * \text{homog}$

Model 4 jest nieistotnie lepszy niż model 3.

Kolejną wersję stanowił **model 5:**

$\text{model 3} + \text{edu_m_rec} * \text{koniec}$.

Okazał się on nieco lepszy niż model 3 (i model 4).

Oprócz tego oszacowano **model 6** w postaci:

$\text{model 5} + \text{ksiazki2} * \text{koniec} + \text{rodz} * \text{koniec} + \text{PA1} * \text{koniec} + \text{PA2} * \text{koniec} + \text{PA3} * \text{koniec} + \text{PA4} * \text{koniec} + \text{PA5} * \text{koniec}$

Jest to model zawierający interakcje zmiennych objaśniających ze zmienną opisującą rok uzyskania najwyższego wykształcenia. Jednak według kryterium AIC model ten jest gorszy niż model 5. Ponadto w modelu dla mężczyzn wszystkie oszacowania parametrów przy dodanych zmiennych okazały się nieistotne, zaś w modelu dla kobiet istotne były jedynie oceny parametrów przy zmiennych $\text{PA2} * \text{koniec} + \text{PA3} * \text{koniec} + \text{PA4} * \text{koniec}$.

Do analizy porównawczej dla kobiet i mężczyzn wybrano zatem oszacowany **model 5**.

A.II. Blok II

A.II.1. Narzędzia badawcze i metody analizy danych z próby

Nasze podejście metodologiczne obejmuje:

- statystyki opisowe prezentujące podstawowe przekroje interesujących nas zagadnień (tj. częstość uczestnictwa w edukacji formalnej, nieformalnej i pozaformalnej, budżety gospodarstw domowych w powiązaniu z edukacją [w tym w przekrojach demograficzno – społecznych i regionalnych; są one możliwe, biorąc pod uwagę liczebność próby]), korelacje uczestnictwa w różnych typach edukacji);

- testy statystyczne służące sprawdzeniu istotności różnic w charakterystykach aktywności edukacyjnej między kobietami a mężczyznami oraz osobami w różnym wieku;
- modele ekonometryczne: modele ze zmienną binarną (probitowe/logitowe) opisujące determinanty uczestnictwa w danym typie edukacji oraz determinanty łączenia różnych rodzajów kształcenia;
- analizę jakościową służącą podsumowaniu uzyskanych wyników; wykorzystuje ona dane i oszacowania z punktów (1) i (2) oraz wspierana jest teorią i wynikami badań empirycznych z literatury.

A.II.1.1. Testy statystyczne

Przeprowadzone zostały testy parametryczne t-Studenta oraz testy nieparametryczne zgodności rozkładów dla zmiennych jakościowych o małej liczebności porównujące dwie grupy (płeć) oraz więcej niż dwie grupy (wiek). Stosowanie testów nieparametrycznych wymaga mniej ostrych założeń niż przy testach parametrycznych kosztem wykorzystania mniejszego zakresu informacji zawartych w danych z próby.

Statystyka testu U Manna-Whitneya ma postać:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R$$

R (suma rang n_1, n_2) oznacza liczebność w badanych grupach.

Dla próby większej niż 20 stosuje się poniższy wzór przy założeniu, że rozkład U jest w przybliżeniu normalny:

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 [n_1 + (n_2 + 1)]}{12}}}$$

gdzie:

- $U = R1 - n_1(n_1 + 1)/2$,
- $R1$ – suma rang elementów z pierwszej próby,
- $n_1 n_2$ – licznosci odpowiednio pierwszej i drugiej próby.

Test U Manna-Whitneya nie wymaga równoliczności grup, rozkładu normalnego czy homogenicznych wariancji.

Statystyka testu Kruskalla-Wallisa ma postać:

$$H = \frac{12}{n(n+1)} \sum_i \frac{R_i^2}{n_i} - 3(n+1)$$

- R_i - suma rang w kolejnych próbach,
- n - liczba obserwacji we wszystkich próbach,
- n_i - liczebność poszczególnej próby.

A.II.1.2. Modele ekonometryczne

Modele z binarną zmienną objaśnianą (modele logitowe/probitowe, modele regresji logistycznej) umożliwiają oszacowanie prawdopodobieństwa wystąpienia danej cechy/zjawiska, np. prawdopodobieństwo uczestnictwa w danym typie edukacji jako funkcja cech danej jednostki (osoby). Modele probitowe i logitowe różnią się jedynie funkcją prawdopodobieństwa (Greene, 2003; Verbeek, 2004). Niech:

$$\mathbb{P}(X_i = k) = F(x_i' \beta)$$

gdzie x_i' i β są wektorami obserwowanych charakterystyk, a $F(x)$ oznacza dystrybuantę rozkładu. Funkcję logistyczną wyraża wzór:

$$F(x) = \frac{\exp(x)}{1 + \exp(x)}$$

Obie funkcje mają podobny kształt i w zasadzie stosuje się je wymiennie, a w praktycznych zastosowaniach najczęściej oba rodzaje modeli dają podobne wyniki.

A.III. Blok III

A.III.1. Wprowadzenie

Wykształcenie oraz kompetencje są jednymi z podstawowych wyznaczników szans życiowych. Uczenie się dotyczy przy tym nie tylko pewnej fazy życia. Jest to proces trwający całe życie, obejmujący także wszystkie pozaszkolne i nieformalne możliwości kształcenia. Z tego też powodu edukację, w tym permanentne podnoszenie kompetencji i nabywanie kwalifikacji (kształcenie dorosłych), traktuje się jako jeden z podstawowych czynników przeciwdziałających wykluczeniu społecznemu. Z drugiej natomiast strony, edukacja może być jedną ze sfer wykluczenia społecznego (Federowicz, Sitek, 2011; Górniak, 2007). Ponadto wykluczenie z edukacji stanowi, co potwierdzają wyniki badań empirycznych (por. np. Rada Monitoringu Społecznego, 2013), jeden z podstawowych czynników wykluczenia społecznego w wielu jego wymiarach.

W prezentowanym opracowaniu zostały przedstawione metody identyfikacji przyczyn (czynników) wykluczenia w obszarze szeroko rozumianej edukacji oraz metody analizy

wpływu niskiego poziomu wykształcenia na wykluczenie społeczne w różnych jego wymiarach.

W ramach wykluczenia edukacyjnego ocenie będą poddane: jego zasięg oraz nierówności w dostępie do kształcenia dzieci, młodzieży i dorosłych w obszarze szeroko rozumianej edukacji. Do ich oceny zaproponowano zastosowanie indeksów zasięgu wykluczenia edukacyjnego oraz współczynnika Giniego. Ponadto omówiono wykorzystanie analizy korespondencji do analizy porównawczej profili osób wykluczonych i niewykluczonych z edukacji na różnych jej poziomach.

Jako narzędzie umożliwiające rozpoznanie siły oddziaływania wyróżnionych w badaniu czynników na wykluczenie z edukacji na jej różnych poziomach oraz rozpoznanie powiązań między tymi czynnikami przedstawiono model ścieżki. Dla wstępnego rozpoznania zależności pomiędzy czynnikami wpływającymi na wykluczenie z edukacji zaproponowano zastosowanie analizy czynnikowej.

W ramach analizy wpływu wykluczenia edukacyjnego na inne obszary wykluczenia społecznego została najpierw przedstawiona metoda oceny stopnia wykluczenia społecznego, w wyróżnionych w badaniu jego obszarach za pomocą agregatowych indeksów wykluczenia społecznego. Natomiast do zbadania wpływu niskiego poziomu wykształcenia na stopień wykluczenia społecznego ogółem (oraz na stopień wykluczenia w wyróżnionych jego wymiarach) zaproponowano wykorzystanie metody dopasowania opartej o indeks skłonności.

A.III.2. Współzależności pomiędzy wykluczeniem społecznym a edukacją

A.III.2.1. Cel i przedmiot analizy

Celem analizy w ramach prezentowanego tematu badawczego jest identyfikacja przyczyn (czynników) wykluczenia w obszarze szeroko rozumianej edukacji oraz wpływ niskiego poziomu wykształcenia i braku kompetencji na wykluczenie społeczne w różnych jego wymiarach.

Termin *wykluczenie społeczne* po raz pierwszy został użyty przez francuskiego ministra Dobrobytu Społecznego Rene Lenoir'a (1974). Użył on go w stosunku do osób uznanych za nieprzystosowane do życia w społeczeństwie industrialnym, żyjących na marginesie społeczeństwa oraz osób, które nie są objęte żadnym systemem ubezpieczeń. W europejskiej polityce społecznej kategoria wykluczenia społecznego po raz pierwszy wystąpiła w oficjalnym dokumencie Komisji Europejskiej dotyczącym programu walki z ubóstwem w 1990 roku (*Commission of the European Communities*, 1990).

Kategoria wykluczenia społecznego nie jest jednoznaczna i może być definiowana w różny sposób (Silver, 1994). Ma ona charakter wielowymiarowy i znacznie wychodzi poza brak środków pieniężnych oraz zasobów materialnych, odnosząc się także do innych ograniczeń, które nie pozwalają jednostce (osobie, rodzinie, gospodarstwu domowemu, grupie społecznej) żyć na poziomie akceptowalnym w danym kraju (Panek, 2011). Pojęcie wykluczenia społecznego odnosi się do sytuacji, w której jednostki, z różnych przyczyn, pozbawione są możliwości pełnego uczestni-

czenia w życiu społecznym. Nieuczestniczenie to nie jest przy tym wynikiem wyboru jednostki, lecz przeszkód jakie ona napotyka (*Social exclusion...*, 1997).

Burchard, Grand oraz Piachaud (2002) wskazali cztery obszary (aspekty), w których jednostka może podlegać wykluczeniu społecznemu:

- konsumpcja – jednostka podlega wykluczeniu ze względu na niewystarczające dochody. Brak wystarczających dochodów uniemożliwia lub znacznie ogranicza zarówno dostęp do podstawowych dóbr zapewniających przeżycie (wyżywienie, ubranie, mieszkanie), jak i innych dóbr i usług zapewniających zaspokojenie podstawowych na danym etapie potrzeb, zaspokajanych przez większość społeczeństwa (edukacja, opieka medyczna itd.);
- produkcja – jednostka podlega wykluczeniu ze względu na bycie bezrobotną oraz niemożność podniesienia kwalifikacji pozwalających na uzyskanie pracy. Konsekwencją tej sytuacji jest wykluczenie z konsumpcji;
- zaangażowanie polityczne – jednostka podlega wykluczeniu ze względu na ograniczanie jej biernych lub czynnych praw wyborczych. Sytuacja ta wynika zarówno z braku dostępu do informacji, jak i możliwości niezbędnych do zaangażowania się w życie polityczne oraz niewykorzystanie możliwości decydowania o istotnych dla społeczności lokalnej problemach;
- integracja społeczna – jednostka podlega wykluczeniu ze względu na nie-wchodzenie w kontakty z innymi członkami społeczeństwa. Stanowi to konsekwencję słabo rozwiniętych relacji społecznych i rodzinnych. W przypadku integracji społecznej często wskazywany jest jeszcze aspekt kulturowy wykluczenia społecznego związany z brakiem kontaktu z kulturą. Może to prowadzić do stopniowej utraty zdolności skutecznego porozumiewania się z otoczeniem na skutek niezrozumienia sensu i znaczenia pewnych symboli, w ramach których odbywa się komunikacja społeczna.

Poszczególne wymiary wykluczenia często nakładają się na siebie, pogłębiając tym samym marginalizację jednostek (Gore, Flueiredo, 1997).

W pracach ekspertów Zespołu Zadaniowego ds. Reintegracji Społecznej utworzonego przy Ministerstwie Gospodarki, Pracy i Polityki Społecznej, wykluczenie społeczne zdefiniowano jako „sytuację uniemożliwiającą lub znacznie utrudniającą jednostce lub grupie zgodne z prawem pełnienie ról społecznych, korzystanie z dóbr publicznych i infrastruktury społecznej, gromadzenie zasobów i zdobywanie dochodów w godny sposób” (Golinowska, Broda-Wysocki, 2005, s. 46). Jak podkreślają autorzy definicji, należy zwrócić uwagę na fakt występowania w niej trzech ważnych elementów:

- sytuację wykluczającą będącą splotem czynników lub też warunków wykluczających – odpowiedź na pytanie, co wyklucza;
- jednostkę (osobę lub grupę) znajdującą się w sytuacji wykluczającej – odpowiedź na pytanie, kto jest wykluczany;

- społeczne funkcjonowanie oraz korzystanie z zasobów publicznych (usług, infrastruktury itd.) i zabezpieczanie własnej egzystencji w godny sposób (zdobywanie dochodów i gromadzenie zasobów), które w wyniku sytuacji wykluczającej jest znacznie utrudnione lub wręcz uniemożliwione.

Przez wykluczenie edukacyjne rozumie się najczęściej nierówności w dostępie do kształcenia oraz różne mechanizmy prowadzące do tych nierówności (Białecki, 2010; Federowicz i Sitek, 2011). Nierówności edukacyjne dotyczą dwóch obszarów, a mianowicie nierówności warunków (np. różnice w poziomie zamożności, różnice związane ze środowiskiem rodzinnym, w poziomie wykształcenia rodziców, w poziomie szkół w miejscu zamieszkania) oraz nierówności wyników (różnice w wynikach w nauce, w poziomie testów kompetencji, w proporcjach osób studiujących ze względu na miejsce zamieszkania, poziom wykształcenia rodziców, czy też poziom zamożności).

A.III.3. Zakres analizy

A.III.3.1. Wykluczenie edukacyjne

W ramach wykluczenia edukacyjnego analizie zostały poddane dwa obszary. Pierwszy z nich dotyczy nierówności w dostępie do kształcenia dzieci i młodzieży oraz dorosłych w obszarze szeroko rozumianej edukacji (wychowanie przedszkolne, szkoły podstawowe i gimnazjum, szkoły ponadgimnazjalne, szkoły wyższe, doksztalcanie). Nierówności edukacyjne stanowią podstawową przeszkodę w budowaniu spójności społecznej przez edukację (Federowicz, Sitek, 2011). Został oceniony zasięg wykluczenia edukacyjnego oraz stopień nierówności w dostępie do kształcenia różnych grup typologicznych osób. W wyniku tej analizy zostały także określone profile społeczno-demograficzne wykluczanych z edukacji na różnych jej poziomach oraz porównane profile osób wykluczonych i niewykluczonych z edukacji na tych poziomach.

Drugi z obszarów będzie dotyczył rozpoznania siły oddziaływania wyróżnionych czynników na wykluczenie z edukacji na różnych poziomach oraz powiązań między tymi czynnikami. Identyfikacja barier i mechanizmów wykluczenia w ramach systemu edukacji może wskazać obszary oddziaływania polityki edukacyjnej nakierowanej na jego zmniejszenie.

W analizie przyczyn wykluczenia w obszarze edukacji jako punkt wyjścia zostały wzięte pod uwagę dwa podstawowe sposoby wyjaśniania tego wykluczenia, a mianowicie:

- wykluczenie jest efektem określonych zachowań jednostek wynikających z ich indywidualnych predyspozycji;
- wykluczenie jest spowodowane niemożliwymi do przewyżczenia przez jednostkę różnymi barierami społecznymi (pochodzenie społeczne, miejsce zamieszkania, status materialny, dostęp do edukacji itd.).

Czynniki powodujące wykluczenie z edukacji można podzielić na cztery zasadnicze grupy (Federowicz, Sitek, 2011; Górniak, 2007; Kozarzewski, 2008):

- społeczno-kulturowe: wyniesione z domu niedocenywanie edukacji (rodzice nie widzą sensu edukacji, co się przekłada na dzieci, nie mają ambicji, aby ich dzieci osiągnęły wyższe poziomy wykształcenia, nie motywują dzieci do nauki, nie chwalać dzieci za sukcesy ani nie dopilnowują dzieci w nauce), środowiskowe wzory drogi zawodowej (poziom edukacji rodziców), wzory wczesnej samodzielności;
- bariery psychologiczne: niska samoocena, brak pewności i wiary w siebie;
- bariery ekonomiczne: bariery związane z szeroko rozumianą sytuacją materialną wyrażaną zarówno przez wskaźniki monetarne (poziom dochodów), jak i pozamonetarne (symptomy złej sytuacji dochodowej wskazujące na niemożność zaspokojenia określonych potrzeb ze względów finansowych: problemy z zakupem podręczników – pomimo „wyprawek szkolnych” i świadczeń rodzinnych, inne wydatki związane ze szkołą przekraczające możliwości finansowe gospodarstw – zwłaszcza, gdy uczy się kilkoro dzieci – jak np. brak pieniędzy na dojazdy i pomoce szkolne), niemożność liczenia na pomoc rodziców w odrabianiu lekcji ze względu na ich zaabsorbowanie sprawami bytowymi, konieczność podjęcia pracy (praca sezonowa w czasie wakacji u rodziny lub sąsiadów, sezonowe zbiory ziół, ślimaków, runa leśnego), opieka nad młodszym rodzeństwem, brak zrozumienia ze strony nauczycieli i rówieśników, trudności w nawiązywaniu kontaktów z rówieśnikami (narażenie się na wstyd i upokorzenie związane z różnymi brakami materialnymi), brak własnego miejsca do nauki w domu, niedożywienie, brak właściwego ubrania, odmowy uczestnictwa w wycieczkach szkolnych oraz różnych formach wypoczynku wakacyjnego z przyczyn finansowych, „zamykanie” w domu z przyczyn finansowych, wycofywanie się z różnych form aktywności, które mogłyby rozszerzyć horyzonty myślowe i wzbogacić zasób doświadczeń;
- inne czynniki: problemy zdrowotne, niepełnosprawność, wczesne rodzicielstwo, patologie w rodzinie (nadużycie alkoholu, zażywanie narkotyków, przemoc w rodzinie, konflikt z prawem któregoś z członków rodziny), brak możliwości dojazdu do odpowiednich szkół z innych przyczyn niż finansowe (np. zbyt daleka odległość do szkoły).

A.III.3.2. Wpływ wykluczenia edukacyjnego na wykluczenie społeczne

Analiza wpływu wykluczenia edukacyjnego (niskiego poziomu wykształcenia⁵) dotyczyła następujących jego obszarów:

- zasobność materialna (ubóstwo rozumiane wielowymiarowo: dochody, trudności finansowe w zaspokojeniu potrzeb żywnościowych, zdrowotnych, mieszkaniowych, kultury i wypoczynku, korzystania z nowoczesnych technologii, itd.);
- rynek pracy (bezrobocie, bierność zawodowa);

⁵ W prowadzonym badaniu empirycznym nie jest badany poziom kompetencji poza kompetencjami cywilizacyjnymi.

- ważne aspekty życia społecznego i publicznego: brak działania dla społeczności (zrzeszanie się i pełnienie funkcji w organizacjach, wspólne działania i praca dla innych) oraz brak aktywności wyborczej.

Za osoby o niskim poziomie wykształcenia są uważane osoby w wieku 25-64 lata, posiadające ukończone co najwyżej gimnazjum oraz w wieku 18-24 lata, posiadające ukończone co najwyżej gimnazjum i niekontynuujące nauki. Takie rozwiązanie przyjmowane jest w analizach porównawczych wykluczenia społecznego prowadzonych w ramach Unii Europejskiej (Atkinson, Cantillon, Marlier, Nolan, 2002). Ponadto wyróżnione zostały dwie dodatkowe grupy osób podlegające wykluczeniu edukacyjnemu: osoby z wykształceniem zasadniczym zawodowym (cechujące się wysokim stopniem bezrobocia) oraz osoby niepracujące i niekontynuujące nauki.

A.III.4. Stan wiedzy

Współzależności pomiędzy wykluczeniem społecznym a edukacją są wielorakie. Edukacja ma istotny wpływ na rozwój dzieci i młodzieży. Prawo do nauki należy do ważnych praw określonych w Konwencji o prawach dziecka (Convention on the Rights of the Child) – przyjętej przez Zgromadzenie Ogólne ONZ 20 listopada 1989 roku (art. 28 konwencji). Z tej perspektywy wykluczenie edukacyjne jest rozumiane jako brak dostępu do nauki, ale też, co szczególnie widoczne jest np. w badaniach prowadzonych przez OECD, zróżnicowaniu lub segmentacji dostępu do edukacji i wyników edukacyjnych. Edukacja może być zatem źródłem wykluczenia nie tylko dzieci, ale też może dalej skutkować wykluczeniem społecznym osób dorosłych. Dlatego w analizie przedmiotu można znaleźć podział wątków dyskusji na dwie grupy: edukację i wykluczenie społeczne dzieci oraz edukację i wykluczenie społeczne dorosłych (np. Klasen, 1998). Dodatkowo, można też spojrzeć na edukację, jako dobro publiczne, które jest konsumowane przez gospodarstwa domowe. Sposób dystrybucji usług edukacyjnych może być zatem też widziany, jako jedno z narzędzi zmniejszania nierówności dochodowych w społeczeństwie.

A.III.4.1. Edukacja i wykluczenie społeczne dzieci

Wykluczenie społeczne dzieci w dużym stopniu wynika z wykluczenia społecznego ich rodzin czy gospodarstw domowych. Jakość życia dzieci, w wielu wymiarach, powiązana jest z dostępem i możliwością uczestnictwa w edukacji. Zaproponowany przez OECD (Organization for Economic Co-operation, 2009) zestaw wskaźników jakości życia dzieci w wielu przypadkach odnosi się – bezpośrednio lub pośrednio – do edukacji oraz warunkowań szkolnych (tabela A.III.1.).

Tabela A.III.1.

Wskaźniki jakości życia dzieci OECD oraz pozycja Polski w poszczególnych obszarach jakości życia

Obszar	Wskaźniki	Pozycja Polski (na 30 krajów)
Dobrostan materialny	Przeciętny dochód do dyspozycji Dzieci w ubogich rodzinach Deprywacja edukacyjna	28
Mieszkanie i środowisko	Przepełnienie w domu Złe warunki środowiska	22
Dobrostan edukacyjny	Średni wynik w czytaniu Nierówność w czytaniu Poziom NEET wśród młodych	8
Zdrowie i bezpieczeństwo	Niska waga urodzeniowa Umieralność niemowląt Odsetek dzieci karmionych piersią Odsetek szczepień (krztusiec) Odsetek szczepień (świnka) Aktywność fizyczna Wskaźnik umieralności Wskaźnik samobójstw	15
Zachowania ryzykowne	Palenie papierosów Picie alkoholu Ciąże nastolatków	20
Jakość życia szkolnego	Zastraszanie w szkole Lubienie szkoły	15

Uwaga: W drugiej kolumnie wskaźniki powiązane z uczeniem się lub środowiskiem szkolnym przedstawione zostały pogrubioną czcionką.

Źródło: Organization for Economic Co-operation (2009).

Bazując na tak określonym zestawie wskaźników, OECD opracowało ranking 30 krajów pod względem jakości życia dzieci. Pozycja Polski w tym rankingu, przedstawionym w tabeli 1., wskazuje na główne obszary wykluczenia społecznego dzieci: sytuację materialną gospodarstw domowych, warunki mieszkaniowe, a także zachowania ryzykowne odnotowywane wśród młodzieży. Poniżej omówione zostały wybrane aspekty jakości życia dzieci w Polsce w oparciu o dostępne badania.

W przypadku jakości życia w obszarze edukacji, poza standardowymi miarami odnoszącymi się do sytuacji dochodowej rodzin oraz ubóstwa, w rankingu OECD uwzględniony został wskaźnik deprywacji edukacyjnej definiowanej jako posiadanie mniej niż czterech z ośmiu podstawowych przedmiotów, tj. biurka do nauki, cichego miejsca do pracy, komputera do odrabiania prac domowych, oprogramowania edukacyjnego, dostępu do internetu, kalkulatora, słownika oraz podręczników szkolnych. Oszacowanie wartości wskaźnika bazuje na wynikach badania PISA obejmującego grupę 15-latków. Według danych z 2006 roku, 21 na 1000 polskich piętnastolatków nie ma dostępu do co najmniej czterech wymienionych przedmiotów – jest to poniżej średniej OECD wynoszącej 35, ale to zapewnia nam 21 miejsce na 30 analizowanych krajów (Organization for Economic Co-operation, 2009, s.36). Badanie UDE

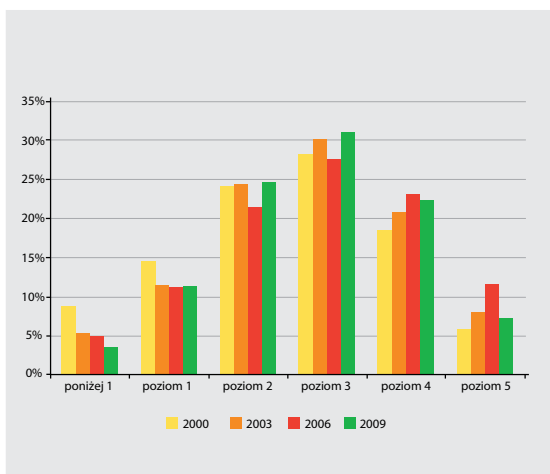
umożliwia poszerzenie tej analizy o uwzględnienie takich cech jak wiek dzieci (co nie było do tej pory możliwe) lub miejsce zamieszkania, a także o monitorowanie zmian zachodzących w czasie.

Jednym z wymiarów jakości życia dzieci są warunki mieszkaniowe, w tym zbyt duże zagęszczenie mieszkania (tj. liczba osób w gospodarstwie domowym jest wyższa niż liczba pokoiów). OECD (Organization for Economic Co-operation, 2009) na podstawie wyników badania EU-SILC szacuje, że Polska należy do krajów o najwyższym poziomie zagęszczenia mieszkań, gdzie 74% dzieci i młodzieży w wieku 0-17 lat mieszka w takich warunkach.

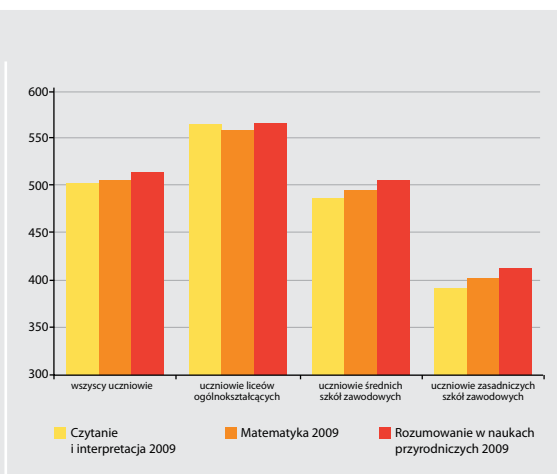
Skutkiem wykluczenia z edukacji są również wyniki osiągnięte przez dzieci i młodzież, w tym również poziom osiągniętych przez nie kompetencji. Wyniki badania PISA wskazują na zróżnicowanie poziomu kompetencji zarówno pomiędzy krajami, jak i wewnątrz krajów. Dotyczy to również Polski. Warto zauważyć, że poprawa średnich wyników badania PISA odnotowywana w Polsce wynika głównie z ograniczenia odsetka 15-latków osiągających najniższe poziomy umiejętności, co ilustruje rysunek A.III.1. Jednocześnie obserwowane jest zróżnicowanie wyników uczniów w zależności od typu szkoły ponadgimnazjalnej – znacząco niższe wyniki osiągają uczniowie szkół zasadniczych zawodowych (rysunek A.III.2.).

Niskie osiągnięcia na początkowym etapie kształcenia, po zakończeniu formalnej czy obowiązkowej edukacji, skutkują w późniejszych etapach przebiegu życia ryzykiem bierności edukacyjnej i zawodowej.

Rysunek A.III.1. Odsetek 15-latków według poziomów umiejętności czytania i interpretacji w badaniu PISA 2000-2009



Rysunek A.III.2. Średni wynik uczniów poszczególnych typów szkół w poszczególnych dziedzinach badania PISA 2009

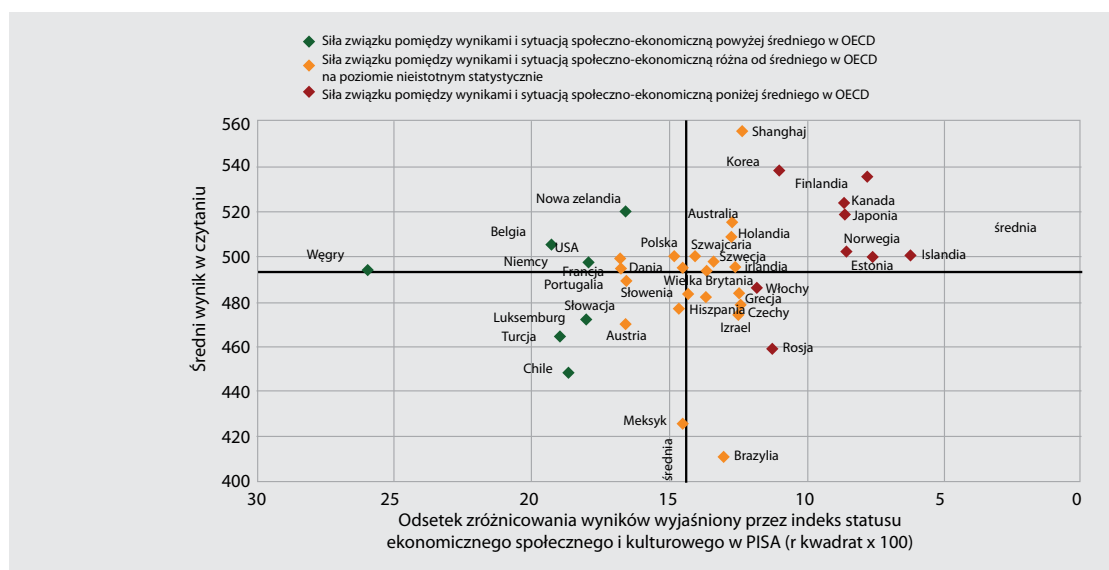


Źródło: Federowicz (2009).

Zróźnicowanie wyników edukacyjnych widoczne jest również w wynikach sprawdzianu szóstoklasisty oraz wynikach egzaminu gimnazjalnego. W perspektywie ostatnich lat dynamicznie wzrasta polaryzacja gimnazjów w dużych miastach, co oznacza wczesne pojawianie się progów selekcyjnych w edukacji (Federowicz, Sitek, 2011).

Jednym z ważnych kierunków prowadzonych badań jest sprawdzenie w jakim stopniu systemy edukacji zmniejszają ryzyko wykluczenia wynikające między innymi z uwarunkowań społeczno-ekonomicznych gospodarstw domowych poprzez działania na rzecz ograniczania zróźnicowania wyników edukacyjnych, a szczególnie zmniejszania odsetka uczniów osiągających najniższe wyniki. System edukacji może niwelować wpływ statusu społeczno-ekonomicznego rodziny na osiągnięcia dzieci, chociaż nadal widoczny jest duży związek pomiędzy tymi zjawiskami. Analizy mające na celu określenie roli polityki edukacyjnej w niwelowaniu różnic osiągniętych wyników prowadzi m.in. OECD w oparciu o wyniki badania PISA (np. Organization for Economic Co-operation, 2012a). Wynika z nich, że najlepiej oceniane systemy edukacji łączą wysoką jakość i równe traktowanie uczniów. Polska należy do krajów, w których poziom zróźnicowania wyników badania PISA w czytaniu jest wyjaśniony przez sytuację społeczno-ekonomiczną gospodarstwa domowego na poziomie zbliżonym do średniej OECD – około 15% zmienności wyników w czytaniu jest wyjaśniane przez indeks zróźnicowania społeczno-ekonomicznego (por. rysunek A.III.3.). Badanie UDE pozwoli na poszerzenie wnioskowania w zakresie powiązań bierności edukacyjnej i zawodowej z sytuacją w gospodarstwie domowym, a także historią edukacyjną i zawodową.

Rysunek A.III.3. Siła związku pomiędzy osiągniętym wynikiem w czytaniu a zróźnicowaniem społeczno-ekonomicznym (PISA 2009)



Źródło: Organization for Economic Co-operation (2012a).

Poza osiąganymi wynikami wykluczenie edukacyjne przejawia się również w dostępie do edukacji, począwszy od jej najwcześniejszych etapów. J. Brandsma (2000) wskazuje, że uczestnictwo dzieci w edukacji przedszkolnej ma pozytywne, chociaż niewielkie znaczenie dla osiągniętych później wyników. Należy dodać, że wraz z wiekiem wpływ

ten się zmniejsza. Dla wyrównywania szans edukacyjnych istotne znaczenie ma jakość edukacji przedszkolnej i programów nauczania. Wnioski płynące z prowadzonych badań wskazują również, że inwestycje we wczesną edukację są efektywne kosztowo i przyczyniają się do rozwoju umiejętności, w tym szczególnie umiejętności językowych, a także kompetencji społecznych. Wyniki badania PISA z 2009 roku wskazują, że średnio w krajach OECD różnica w umiejętnościach czytania pomiędzy uczniami, którzy uczestniczyli w edukacji przedszkolnej przez okres dłuższy niż rok oraz tymi, którzy nie uczestniczyli w niej w ogóle, wynosiła 54 punkty. Zgodnie z szacunkami OECD (Organization for Economic Co-operation, 2011a) wydłużenie okresu uczestnictwa w edukacji przedszkolnej stanowi czynnik, który w największym stopniu wpływa na poprawę kompetencji 15-latków. Dane statystyczne dla Polski wskazują na dynamiczny wzrost odsetka dzieci w wieku 3-5 lat objętych wychowaniem przedszkolnym. Niemniej jednak luka pomiędzy uczestnictwem w tym typie edukacji dzieci zamieszkałych w miastach i na wsi pozostaje bardzo wysoka (tabela A.III.2.). Dla zamknięcia tej luki niezbędne jest pogłębienie wiedzy na temat determinant uczestnictwa dzieci w wychowaniu przedszkolnym z perspektywy gospodarstw domowych.

Tabela A.III.2.

Odsetek dzieci w wieku 3-5 lat objętych wychowaniem przedszkolnym

	2008	2009	2010	2011	2012
dzieci w miastach	70,5	75,9	78,8	84,0	83,7
dzieci na wsi	28,5	37,5	41,0	49,3	50,8
różnica	42,0	38,4	37,8	34,7	32,9

Źródło: Bank Danych Lokalnych GUS.

Wykluczenie edukacyjne dotyczy również uczestnictwa w edukacji na dalszych etapach życia. Wymiarem tego wykluczenia jest stosowany przez OECD w analizach wskaźnik NEET (Not in Education, Employment, or Training) dla młodzieży w wieku 15-18 lat. Klasen (1998) wymienia również inne wskaźniki z tego obszaru: odsetek osób nieuczestniczących w edukacji w wieku opuszczania systemu edukacji, odsetek osób nieuczestniczących w edukacji w wieku 17 lat czy odsetek absolwentów kończących pierwszy program edukacji później w relacji do ludności w teoretycznym wieku, w którym program ten powinien zostać ukończony. Wskazuje on na istotne zróżnicowanie tych miar w krajach OECD. Odsetki wykluczenia z edukacji są wysokie nie tylko w krajach o niskim rozwoju gospodarczym (Turcja, Hiszpania, Węgry), ale też wśród niektórych bogatszych krajów (USA, Wielka Brytania, Francja).

Wśród czynników związanych ze stanem zdrowia warto odnieść się do aktywności fizycznej, która również zależy od czynników związanych z edukacją i rodziną, w tym dostępu do odpowiedniej infrastruktury (np. boiska szkolne, tzw. „orliki”), a także uczestnictwa w dodatkowych zajęciach związanych z aktywnością fizyczną, organizowanych przez rodzinę bądź też wspólnej, rodzinnej aktywności fizycznej. Wyniki badania *Health Behaviour in School-aged Children* (HBSC; World Health Organization, 2008) przeprowadzonego w latach 2005-2006 wskazują, że 17% polskich dzieci ćwiczy regularnie lub intensywnie (w porównaniu do średniej OECD wynoszącej 20%). Badanie HBSC moni-

torowało także palenie i picie alkoholu wśród młodzieży – 16% 15-latków w Polsce pali przynajmniej raz w tygodniu (jest to na poziomie średniej w krajach OECD), a 24,2% dzieci w wieku 13-16 lat upiło się przynajmniej 2 razy (powyżej średniej w OECD). Natomiast nieco poniżej średniej w OECD kształtuje się liczba ciąż wśród nastolatek (według danych World Development Indicators z 2008 roku, 14,5 na 1000 dziewcząt w wieku 15-19 lat zachodzi w ciążę: World Bank, 2008).

Polskie dzieci raczej lubią szkołę (21,1% uczniów czyli mniej niż średnia w OECD wskazywało w badaniu HBSC, że szkoły nie lubi). Również zastraszanie w polskich szkołach odnotowywane jest rzadziej niż przeciętnie w OECD (według tego samego badania 9,6% dzieci w Polsce doświadczyło zastraszania; średnia w OECD wynosi 11%).

Wykluczenie edukacyjne wynikać może również z indywidualnych charakterystyk uczniów. Warto zwrócić uwagę na dzieci i młodzież z niepełnosprawnościami, które narażone są na większe ryzyko wykluczenia edukacyjnego. Chociaż w Polsce możliwe jest uczestnictwo tej grupy w edukacji otwartej, na kolejnych etapach edukacyjnych wyraźnie widać rosnący odsetek uczniów w szkołach specjalnych. W konsekwencji osoby z niepełnosprawnościami cechują się znacznie mniej korzystną strukturą wykształcenia (Federowicz, Sitek, 2011).

Podsumowując, w ciągu ostatniej dekady w literaturze istotnie rozwinął się wątek związany z oceną powiązań pomiędzy edukacją a wykluczeniem społecznym i sytuacją dzieci. Wiele z czynników i zjawisk związanych z wykluczeniem edukacyjnym wynika zarówno z sytuacji w rodzinie, jak i sytuacji w szkole. Analizy międzynarodowe wskazują, że polskie dzieci i młodzież znajdują się w wielu obszarach w gorszej sytuacji niż ich rówieśnicy z innych krajów OECD. Pogłębienie wiedzy na temat tych zjawisk jest potrzebne, aby zrozumieć mechanizmy, które wpływają na zmniejszanie lub narastanie wykluczenia społecznego i edukacyjnego dzieci i młodzieży.

A.III.4.2. Edukacja i wykluczenie społeczne dorosłych

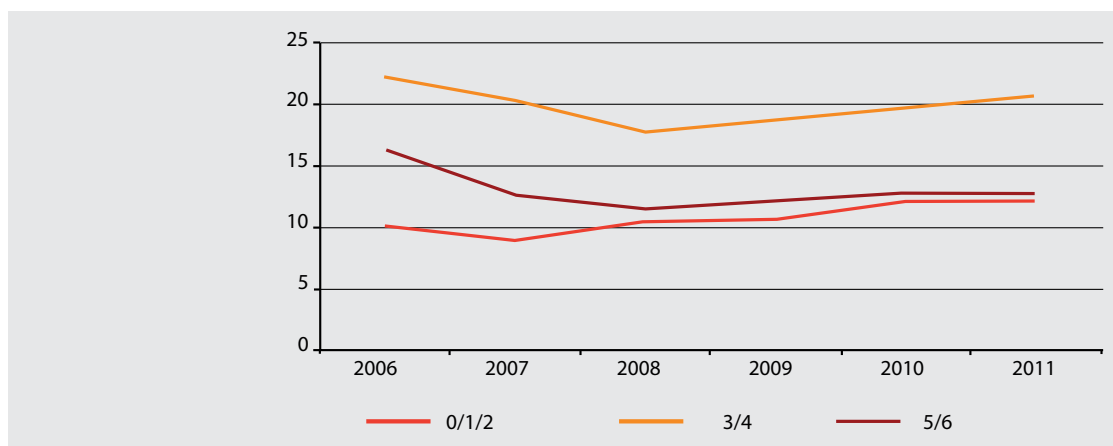
Zjawiska związane z wykluczeniem społecznym dzieci mają również wpływ na późniejsze wykluczenie społeczne dorosłych. Jednym z podstawowych kanałów transmisji wykluczenia społecznego i nierówności jest rodzina. W Polsce wykształcenie osiągnięte przez młode osoby zależy w znacznym stopniu od wykształcenia rodziców (potwierdzają to np. wyniki badania GUS *Kształcenie Dorosłych* 2006, 2009). Chociaż możemy w Polsce zaobserwować silny związek pomiędzy wykształceniem kolejnych pokoleń (współczynnik determinacji R^2 wynosi 28%), to jak podkreślają autorzy raportu – *Społeczeństwo w drodze do wiedzy. Raport o stanie edukacji 2011* (Federowicz, Sitek, 2011), w Polsce nie mamy pełnych danych pozwalających dokładnie prześledzić nierówności edukacyjne na kolejnych etapach kariery edukacyjnej (s. 202).

Ważną perspektywą badawczą jest również wchodzenie osób młodych na rynek pracy. J. Brandsma (2000) podkreśla, że proces ten ulega zmianom – dzisiejsze ścieżki wchodzenia osób młodych na rynek pracy są różnorodne, na co wskazują również, wykorzystując analizę historii edukacyjnych i zawodowych, Quintini i Manfredi (2009). Absolwenci o niskich kwalifikacjach mają trudności z wejściem na rynek pracy i jej rozpoczęciem. Są również bardziej podatni na skutki pogarszającej się sytuacji gospodarczej. Obecnie

w Polsce nie są prowadzone badania, które pozwoliłyby na analizę ścieżek wchodzenia młodych na rynek pracy i identyfikację tych, które skutkują ryzykiem wykluczenia społecznego.

Ryzyko wykluczenia jest jednak widoczne między innymi przez kształtowanie się wskaźników NEET, tj. osób które się nie uczą, nie pracują i nie przygotowują się do zawodu. W Polsce wyraźnie widać znacząco większy poziom tego wskaźnika wśród osób młodych z wykształceniem na poziomach ISCED 3 lub 4 (odpowiadających szkole zawodowej, technikum lub liceum), na co wskazuje rysunek A.III.4. Wskaźnik ten, począwszy od rozpoczęcia globalnego kryzysu i spowolnienia gospodarczego (tj. od 2008 roku), również najsilniej wzrasta w tej populacji osób młodych.

Rysunek A.III.4. Poziom NEET w Polsce w zależności od poziomu wykształcenia (ISCED); osoby w wieku 15-29 lat w latach 2006-2011



Źródło: Organization for Economic Co-operation (2013).

Wykształcenie determinuje również sytuację na rynku pracy osób dorosłych. Wśród osób o niskich poziomach wykształcenia obserwujemy zarówno niższe wskaźniki zatrudnienia, jak również wyższe stopy bezrobocia, co związane jest głównie z niedopasowaniem ich kompetencji do zapotrzebowania rynku pracy. Wykształcenie determinuje również poziom wynagrodzenia, co nazwane jest „premią z edukacji”. Wszystkie wspomniane zjawiska są szeroko omawiane w literaturze zarówno krajowej, jak i zagranicznej (por. Boeri, van Ours, 2011; Bukowski, 2010; European Commission, 2010; Górniak, 2013; Organization for Economic Co-operation, 2012b).

W kontekście wykluczenia edukacyjnego warto zwrócić także uwagę na to, że wykształcenie ma również wpływ na dalszą aktywność edukacyjną. Osoby o niskim poziomie wykształcenia mają również niższe szanse na uczestnictwo w uczeniu się przez całe życie, jako osoby dorosłe (Chłoń-Domińczak, Lis, 2013). Wynika to zarówno z braku ich własnej inicjatywy, jak również inicjatywy ich pracodawców (Brandsma, 2000).

Wykształcenie wpływa jednak nie tylko na sytuację osób dorosłych na rynku pracy. Obserwowane są również zależności pomiędzy poziomem wykształcenia a stanem zdrowia populacji. OECD (2011b) wskazuje, że wśród dorosłych z niższym wykształceniem we wszystkich krajach OECD częściej występuje otyłość niż wśród osób z wykształce-

niem wyższym. Podobnie, wraz ze wzrostem wykształcenia, spada odsetek osób palących papierosy.

Badania odnoszące się do analizy uwarunkowań wykluczenia społecznego i edukacyjnego osób dorosłych sięgają również do oceny tego, na ile wykluczenie takie wynika z dzieciństwa wzorców w rodzinie. Badania prowadzone na danych panelowych w Wielkiej Brytanii (Machin, 1998) pozwalają zauważyć, że doświadczenie ubóstwa i wykluczenia społecznego w dzieciństwie zwiększa ryzyko wykluczenia na etapie życia dorosłego, obejmującego ryzyko bezrobocia, a nawet przestępstw. John Hobcraft (1998) dodaje również, że wykluczenie w dorosłym życiu dotyka osoby, które osiągały złe wyniki edukacyjne, przebywały w domach dziecka, jak również cechowały się agresją. Ryzyko bezdomności jest również większe wśród osób żyjących w rodzinach, w których jeden z rodziców nie jest rodzicem biologicznym. Ryzyko porażki edukacyjnej zwiększa się wraz z brakiem zainteresowania rodziców szkołą oraz ubóstwem w dzieciństwie. Złe wyniki edukacyjne w okresie nauki w szkole zwiększają również ryzyko bezrobocia. Badania jakościowe prowadzone w Polsce (Tarkowska, 2007) pokazują również, że wśród rodzin byłych pracowników PGR pojawiają się wzorce dziedziczenia biedy czy utrudnionego startu w dorosłość ze względu na ograniczenia wynikające z sytuacji w domu rodzinnym.

Podsumowując, można wskazać, że osoby które w dzieciństwie były wykluczone edukacyjnie są narażone na ryzyko ubóstwa i bezrobocia częściej i dłużej niż inni. Poza tymi czynnikami wykluczenia, osoby takie często cechują się gorszym stanem zdrowia, gorszym dostępem do mieszkań, żywności czy opieki zdrowotnej (Atkinson i Hills, 1998).

A.III.4.3. Nakłady publiczne na edukację a ubóstwo gospodarstw domowych

Jak już wspomniano wcześniej, dobrze zaprojektowany system edukacji może w wielu przypadkach przeciwdziałać ryzyku wykluczenia społecznego. Finansowanie edukacji ze środków publicznych daje szansę i możliwości dostępu do tych usług dzieciom i młodzieży bez względu na ich sytuację społeczno-ekonomiczną. Organization for Economic Co-operation (2011b) szacuje wpływ wartości usług związanych z edukacją publiczną na zmniejszenie nierówności dochodowych gospodarstw domowych. Polska należy do krajów, gdzie publiczna edukacja w sposób znaczący wpływa na zwiększenie szacowanego dochodu gospodarstw domowych uwzględniającego wartość usług publicznych w niższych grupach kwintylowych (tabela A.III.3.). Dotyczy to szczególnie systemu oświaty, w mniejszym stopniu szkolnictwa wyższego, gdyż wartość usług publicznych w tym przypadku częściej trafia do gospodarstw domowych o wyższym statusie społeczno-ekonomicznym i dochodach.

Tabela A.III.3.

Efekt zwiększenia dochodów gospodarstw domowych wynikający z korzyści płynących z publicznych usług edukacyjnych w grupach kwintylowych: Polska oraz średnia OECD, 2007

Procent dochodu gospodarstw domowych w danej grupie kwintylowej odpowiadającego równowartości otrzymanych publicznych usług edukacyjnych:						
	kwintyl 1	kwintyl 2	kwintyl 3	kwintyl 4	kwintyl 5	razem
Polska	40	21	14	9	5	12
średnia OECD	31	19	14	11	6	12

Źródło: Organization for Economic Co-operation (2011b).

A.III.5. Statystyczne metody analizy wykluczenia z edukacji

A.III.5.1. Zasięg wykluczenia z edukacji oraz nierówności w dostępie do kształcenia

Specyfikacja profili społeczno-demograficznych wykluczonych z edukacji

Punktem wyjścia analizy nierówności w dostępie do kształcenia dzieci i młodzieży oraz dorosłych na różnych jego etapach, ocena skali tych wykluczeń oraz zidentyfikowanie podstawowych charakterystyk osób im podlegających. Uzyskane informacje pozwalają na określenie profili społeczno-demograficznych wykluczonych z edukacji na różnych jej etapach z uwzględnieniem ich specyfiki.

W procedurze tworzenia profili osób wykluczonych z edukacji na różnych jej etapach, zostały uwzględnione zarówno charakterystyki osób wykluczonych, jak i charakterystyki ich gospodarstw domowych. W ramach charakterystyk osób wykluczonych włączono do analiz następujące cechy:

- poziom wykształcenia rodziców (wyższy z poziomów wykształcenia rodziców),
- klasa miejscowości zamieszkania,
- region (województwo, typ powiatu),
- płeć,
- stan zdrowia,
- status społeczno-zawodowy (pracownicy, pracujący na własny rachunek, rolnicy, emeryci, renciści, bezrobotni, inni bierni zawodowo).

Jako cechy charakterystyczne gospodarstw domowych, do których należą wykluczeni z edukacji, zostały wyróżnione:

- zamożność gospodarstwa domowego (należenie do najniższej grupy kwintalowej dochodów, pozostawanie w skrajnym ubóstwie),

- typ gospodarstwa domowego wyodrębniony na podstawie liczby rodzin i typu rodziny biologicznej (gospodarstwa domowe wielodzietne, gospodarstwa domowe rodzin niepełnych),
- grupa społeczno-ekonomiczna (gospodarstwa domowe pracownicze, rolników, emerytów, rencistów, pracujących na własny rachunek, utrzymujących się z niezarobkowych źródeł innych niż emerytura i renta),
- wyposażenie gospodarstwa domowego w dobra kultury,
- występowanie patologii w rodzinie.

Profile wykluczonych z edukacji i zasięg wykluczenia z edukacji

Zasięg wykluczenia z edukacji na danym poziomie oraz profile wykluczenia z edukacji są analizowane za pomocą agregatowego indeksu zasięgu wykluczenia:

$$H^w = \frac{n_w}{n}, \quad (\text{A.III.1})$$

gdzie:

- n – liczba badanych osób dla danego etapu edukacji,
- n_w – liczba osób wykluczonych z danego etapu edukacji.

Indeks ten spełnia warunek dekompozycji mówiący, że dany indeks dla całej badanej populacji może być obliczany jako średnia ważona z indeksów dla tego samego typu podpopulacji (na przykład: podpopulacji osób wykluczonych z danego etapu edukacji, wyróżnionych ze względu na klasę miejscowości zamieszkania):

$$H^w = \frac{1}{n} \sum_{s=1}^r H_s^w \cdot n_s, \quad (\text{A.III.2})$$

gdzie:

- H_s – wartość indeksu H dla s -tej podpopulacji,
- n_s – liczba osób w s -tej podpopulacji.

Inną pożądaną w badaniu profili wykluczonych z edukacji własnością indeksu jest jego zgodność w subpopulacji. Gdy wartość indeksu dla jakiegokolwiek subpopulacji wzrośnie, wartość indeksu dla całej populacji także wzrośnie przy braku zmian wartości tego indeksu w pozostałych podpopulacjach.

Nierówności w dostępie do kształcenia

Badaną populację osób wykluczonych z edukacji na danym poziomie rozdziela się na rozłączne podpopulacje (wyróżnione ze względu na podstawowe charakterystyki osób podlegających temu wykluczeniu) i dla każdej z nich oblicza niezależnie wartości indeksów zasięgu wykluczenia z edukacji na danym etapie, a następnie porównuje te wartości między sobą oraz z wartością indeksu ogółem dla całej populacji. Przeprowa-

dzona analiza porównawcza pozwala na ocenę nierówności w dostępie do kształcenia na różnych etapach wyróżnionych w badaniu podpopulacji osób.

Nierówności w dostępie do edukacji na poszczególnych jej etapach są także analizowane na podstawie stopnia odchylenia rozkładu wykluczonych z edukacji od rozkładu egalitarnego dostępu do edukacji, tzn. takiego rozkładu, w którym każda podpopulacja osób wyróżnionych ze względu na dane kryterium (np. klasę miejscowości zamieszkania) ma taki sam dostęp do danego typu edukacji. Taki sam dostęp do edukacji na danym jej etapie oznacza identyczne odsetki wykluczonych z edukacji w podpopulacjach.

W literaturze przedmiotu istnieje wiele propozycji pomiaru nierówności (Panek, 2011). Szczególnie ważną rolę w określaniu nierówności odgrywa krzywa (funkcja) Lorenza. Argumentami funkcji Lorenza w tym badaniu będą skumulowane odsetki osób w wyróżnionych grupach typologicznych uporządkowane według niemalejących odsetków wykluczonych z edukacji, a wartościami – skumulowane odsetki wykluczonych z edukacji.

Najbardziej popularną miarą stopnia nierówności, opartą na skali odstępów funkcji Lorenza od linii egalitarnej, jest współczynnik Giniego. Dla oszacowania wartości współczynnika Giniego, dysponując danymi dla wyróżnionych grup typologicznych osób, można użyć następującego wzoru:

$$G = \frac{2}{2n^2w} = \sum_{s=1}^r \sum_{z=1}^r |w_s - w_z| n_s n_z, \quad (\text{A.III.3})$$

gdzie:

- w_s, w_z – odsetki osób wykluczonych z edukacji na danym poziomie odpowiednio w s -tej i w z -tej grupie typologicznej osób,
- w – odsetek osób wykluczonych z edukacji na danym poziomie w całej populacji.

Analiza porównawcza profili wykluczonych i niewykluczonych z edukacji

Kolejnym etapem analizy jest analiza porównawcza profili wykluczonych i niewykluczonych z edukacji, oddzielna dla różnych ich etapów. Wykorzystana została w niej jedna z metod analizy czynnikowej, a mianowicie analiza korespondencji.

Wprowadzenie

Analiza korespondencji jest metodą analizy czynnikowej, która prezentuje związki pomiędzy obiektami przestrzennymi (w tym badaniu zbiorowością osób wykluczonych z danego etapu edukacji) oraz pomiędzy zmiennymi charakteryzującymi te obiekty (w tym badaniu – wariantami charakterystyk badanych osób) przede wszystkim w formie graficznej. Posiada ona szereg własności, których nie mają inne metody czynnikowe. Jako jedyna metoda analizy czynnikowej daje możliwość umieszczenia w jednym czynnikowym układzie odniesienia zarówno punktów reprezentujących zmienne, jak i punktów reprezentujących obiekty, co pozwala wykryć strukturalne związki pomiędzy zmiennymi i obiektami, a tym samym znacznie zwiększyć możliwości interpretacyjne. Analiza korespondencji umożliwia zarówno analizę danych jakościowych, jak i ilościowych. Nie stawia także żadnych wymagań co do liczebności obserwacji.

Analiza korespondencji była rozwijana równolegle w ośrodkach naukowych w wielu krajach (Beh, 2004). Znaczący wkład w rozwój analizy korespondencji miały prace zespołu naukowców kierowanego przez Jean Paula Benzéciego w latach 60-tych we Francji (1973a i 1973b). Nadali oni analizie korespondencji formę geometryczną oraz najczęściej aktualnie stosowaną w praktyce jej nazwę⁶. Szerokie omówienie analizy korespondencji w języku polskim znajduje się na przykład w pracach Tomasza Panka (1987, 2008) oraz Agnieszki Stanimir (2005).

Model analizy korespondencji

Najogólniejszym punktem wyjścia analizy korespondencji (podejście to będziemy nazywali nieklasyczną analizą korespondencji) jest macierz danych wejściowych o postaci⁷:

$$X=[x_{ji}], \quad x_{ji} \geq 0; \quad j=1,2,\dots,m; i=1,2,\dots,n \quad (\text{A.III.4})$$

gdzie:

- x_{ji} – wartość j -tej zmiennej w i -tym obiekcie.

Należy zauważyć, że elementy macierzy (A.III.4) mogą być dowolnymi danymi nieujemnymi. Wiersze macierzy (A.III.4) mogą być interpretowane w ujęciu geometrycznym jako współrzędne m punktów-zmiennych w n -wymiarowej przestrzeni obiektów R^n . Natomiast kolumny tej macierzy są w ujęciu geometrycznym współrzędnymi punktów-obiektów w m -wymiarowej przestrzeni zmiennych R^m .

W podejściu klasycznym macierzą danych wejściowych jest dwuwymiarowa tablica kontyngencji:

$$N=[n_{ji}], \quad (\text{A.III.5})$$

gdzie:

- n_{ji} – liczebność jednostek obserwacji posiadających jednocześnie j -tą kategorię pierwszej z charakteryzujących je zmiennych i i -tą kategorię drugiej ze zmiennych.

Tutaj zastosowano nieklasyczne ujęcie analizy korespondencji, chociaż można także wykorzystać wielowymiarowe podejście nieklasyczne zaproponowane przez C. Burta (1950) operujące, w miejsce dwuwymiarowej tablicy kontyngencji, wielowymiarową tablicą kontyngencji (tzw. macierzą Burta). W przypadku stosowania nieklasycznej analizy korespondencji zmienne wejściowe poddawane są standaryzacji, najczęściej przy zastosowaniu przekształcenia ilorazowego lub unitaryzacji. Należy pamiętać, że zmienne po standaryzacji w dalszym ciągu nie mogą przyjmować wartości ujemnych.

⁶ W polskiej literaturze przedmiotu analiza korespondencji występuje także jako analiza odpowiedniości czy też analiza powiązań.

⁷ Możliwe jest również operowanie macierzą danych o elementach, poprzez wprowadzenie do analizy elementu czasu t ($t=1,2,\dots,T$). W celu uproszczenia zapisu pominiemy w dalszych rozważaniach wymiar czasu.

Kolejnym krokiem jest przekształcenie macierzy danych wejściowych w macierz częstości względnych, zwaną też macierzą korespondencji, poprzez podzielenie każdego jej elementu przez sumę wartości jej elementów:

$$P=[p_{ji}], \quad j=1,2,\dots,m; i=1,2,\dots,n \quad (\text{A.III.6})$$

gdzie:

$$p_{ji} = \frac{z_{ji}}{\sum_{j=1}^m \sum_{i=1}^n z_{ji}} \quad (\text{A.III.7})$$

przy czym:

- z_{ji} – wystandaryzowana wartość j -tej zmiennej w i -tym obiekcie.

Na podstawie macierzy P wyznaczamy tzw. macierze profili. Macierz profili wierszowych R uzyskujemy, dzieląc każdą częstość w wierszu macierzy P przez sumę wszystkich częstości w tym wierszu:

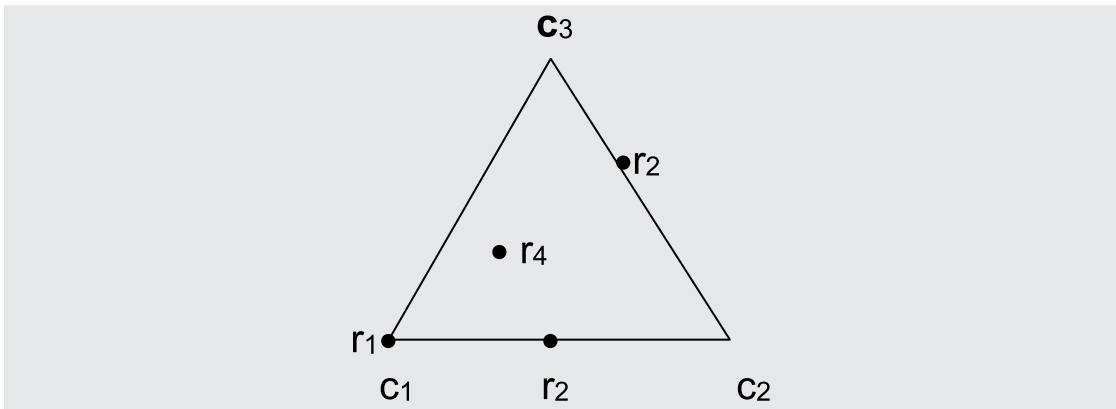
$$R=[r_{ji}] = \left[\frac{p_{ji}}{p_j} \right], \quad j=1,2,\dots,m; i=1,2,\dots,n \quad (\text{A.III.8})$$

Elementy profili wierszowych mogą być traktowane jako współrzędne wektorów wierszowych (zmiennych) w n -wymiarowej przestrzeni euklidesowej R^n . Punkty te leżą w ściśle określonym obszarze przestrzeni R^n , gdyż elementy każdego wiersza sumują się do jedności. Obszarem tym jest $n-1$ wymiarowy sympleks o n wierzchołkach⁸ w postaci: $(1,0,0,\dots,0)$, $(0,1,0,\dots,0)$, $(0,0,0,\dots,1)$. Każdy z wierzchołków jest wyznaczany przez n współrzędnych. Wynika z tego, że zbiór wszystkich m punktów (reprezentujących zmienne w ujęciu nieklasycznym) umieszczonych pierwotnie w przestrzeni n -wymiarowej można przedstawić bez żadnych zniekształceń w przestrzeni R^{n-1} .

Przykładowo, gdy analizujemy macierz składającą się z czterech wierszy reprezentujących zmienne oraz z trzech kolumn reprezentujących obiekty, możemy punkty reprezentujące profile wierszowe przedstawić na płaszczyźnie (czyli w przestrzeni o jeden wymiar mniejszej niż liczba obiektów) wewnątrz trójkąta o trzech wierzchołkach postaci $(1, 0, 0)$, $(0, 1, 0)$ i $(0, 0, 1)$, czyli trójkąta rozpiętego na jednostkowych wektorach (rysunek A.III.5.).

⁸ Sympleksem n -wymiarowym o $n+1$ wierzchołkach, które są punktami n -wymiarowej przestrzeni, nazywany jest najmniejszy zbiór wypukły zawierający te punkty.

Rysunek A.III.5. Punkty reprezentujące profile wierszowe (zmienne) w trójkącie wyznaczonym przez objekty



Źródło: Opracowanie własne.

Natomiast dzieląc każdą z częstości w kolumnie macierzy P przez sumę wszystkich częstości w tej kolumnie, otrzymujemy macierz profili kolumn C :

$$C = [c_{ji}] = \left[\frac{P_{ji}}{P_{\cdot j}} \right], \quad j=1,2,\dots,m; i=1,2,\dots,n \quad (\text{A.III.9})$$

Elementy profili kolumnowych są wtedy współrzędnymi wektorów kolumnowych (objektów) w m -wymiarowej przestrzeni euklidesowej R^m . Punkty te (reprezentujące objekty) są położone w obszarze wyznaczonym przez $m-1$ – wymiarowy sympleks o m wierzchołkach, czyli można je również przedstawić w przestrzeni R^{m-1} .

Wierzchołki sympleksów mają duże znaczenie w interpretacji tzw. map percepcji, czyli wykresów, na których wiersze i kolumny przedstawione są w postaci punktów na płaszczyźnie wyznaczonej przez dwie pierwsze osie czynnikowe. Punkty te są rzutami prostopadłymi n -wymiarowych wierszy i m -wymiarowych kolumn na tę płaszczyznę.

W nieklasycznej analizie korespondencji wierzchołki sympleksów odpowiadają hipotetycznym obiektom, dla których tylko jedna zmienna przyjmuje wartość różną od zera, a pozostałe zmienne przyjmują wartość zero. Można tym samym przyjąć, że im bliżej danego wierzchołka leży punkt reprezentujący obiekt, tym większy wpływ na charakterystykę tego obiektu ma zmienna odpowiadająca wierzchołkowi.

Częstości brzegowe, odpowiednio wierszy (r) i kolumn (c), w macierzach profili R i C są średnimi profilami kolumnowymi i wierszowymi (centrum kolumnowym albo wierszowym). Punkty reprezentowane przez przeciętne profile wierszowe i kolumnowe nazywane są centroidami i leżą w środku układu współrzędnych.

Odległości pomiędzy profilami wierszowymi (punktami reprezentującymi zmienne) w przestrzeni R^n , wylicza się za pomocą ważonej metryki euklidesowej, gdzie wagami są częstości brzegowe kolumn o postaci:

$$d^2(\mathbf{r}_j, \mathbf{r}_{j'}) = \sum_{i=1}^n \frac{1}{p_{.i}} \left(\frac{p_{ji}}{p_j} - \frac{p_{j'i}}{p_{j'}} \right)^2 \quad j, j' = 1, 2, \dots, m; j \neq j'. \quad (\text{A.III.10})$$

Analogicznie określamy odległości pomiędzy profilami kolumnowymi (punktami reprezentującymi obiekty w przestrzeni R^m) za pomocą metryki euklidesowej, gdzie wagami są częstości brzegowe wierszy:

$$d^2(\mathbf{c}_i, \mathbf{c}_{i'}) = \sum_{j=1}^m \frac{1}{p_j} \left(\frac{p_{ji}}{p_{.i}} - \frac{p_{j'i'}}{p_{.i'}} \right)^2 \quad i, i' = 1, 2, \dots, n; i \neq i'. \quad (\text{A.III.11})$$

Analiza odległości pomiędzy profilami wierszowymi (kolumnowymi) jest tożsama z analizą odległości profili wierszowych (kolumnowych) od średnich profili wierszowych (kolumnowych).

Odległości (A.III.10); (A.III.11) są jednocześnie odległościami chi-kwadrat. Z odległościami chi-kwadrat związane jest pojęcie inercji. Jest to miara zróżnicowania elementów w macierzy danych wejściowych. Całkowita inercja macierzy określa stopień dyspersji profili wierszowych (kolumnowych) względem odpowiadających im centroid, czyli wskazuje, jak bardzo poszczególne profile wierszowe (kolumnowe) różnią się od odpowiadającego im średniego profilu.

Inercja dla wierszy jest równa inercji dla kolumn i jednocześnie równa inercji całkowitej. Inercja posiada też interpretację geometryczną – jako miara rozproszenia punktów reprezentujących profile w wielowymiarowej przestrzeni. Kiedy wartość inercji równa jest zero, to punkty reprezentujące profile wierszowe (kolumnowe) skupiają się w początku układu współrzędnych. Odpowiada to sytuacji, w której wszystkie profile wierszowe (kolumnowe) są takie same. Czym większa wartość inercji, tym większe rozproszenie punktów reprezentujących profile w stosunku do środka układu współrzędnych. Maksymalna wartość całkowitej inercji jest równa $s = \min(m-1, n-1)$.

Dotychczas przedstawione rozwiązania umożliwiają jedynie oddzielne analizy profili wierszowych i kolumnowych. Podstawowym celem analizy korespondencji jest natomiast ich jednoczesna analiza, najczęściej w formie prezentacji graficznej. W tym celu macierz P zostaje przekształcona w macierz A nazywaną macierzą różnic wystandaryzowanych, czyli ważonych odchyłeń profili od centrum wierszowego i kolumnowego o postaci:

$$A = [a_{ji}] \quad j=1, 2, \dots, m; i=1, 2, \dots, n \quad (\text{A.III.12})$$

gdzie:

$$a_{ji} = \frac{p_{ji} - p_j \cdot p_{.i}}{\sqrt{p_j \cdot p_{.i}}} \quad (\text{A.III.13})$$

Stanowi ona podstawę ostatecznej fazy analizy korespondencji.

Suma kwadratów elementów macierzy A jest jednocześnie równa inercji całkowitej układu:

$$\sum_{j=1}^m \sum_{i=1}^n a_{ij}^2 = \lambda. \quad (\text{A.III.14})$$

Przekształcenie macierzy P w macierz A ma charakter symetryczny względem wierszy i kolumn. Umożliwia to znalezienie przestrzeni czynnikowej (wektorów kierunkowych osi czynnikowych) na bazie m -punktów reprezentujących zmienne w przestrzeni n -wymiarowej (analiza względem kolumn macierzy A) lub też na podstawie n -punktów reprezentujących obiekty w przestrzeni m -wymiarowej (analiza względem wierszy macierzy A). Właśnie symetryczna standaryzacja wejściowej macierzy danych pozwala nie tylko na ustalenie struktury czynnikowej zmiennych oraz położenia obiektów w nowym czynnikowym układzie odniesienia, podobnie jak ma to miejsce w pozostałych metodach analizy czynnikowej, lecz także na ustalenie struktury czynnikowej obiektów oraz położenia zmiennych w tym samym układzie odniesienia, co nie jest możliwe w żadnej innej metodzie analizy czynnikowej. Innymi słowy, przeprowadzając analizę korespondencji, szuka się wspólnego, ortogonalnego układu odniesienia [podprzestrzeni o wymiarze $s = \min(m-1, n-1)$] dla punktów reprezentujących profile wierszowe i profile kolumnowe.

Analiza korespondencji jest metodą dekompozycji wartości inercji całkowitej. Kolejne osie czynnikowe szukane są w taki sposób, aby odpowiadające im wymiary wyjaśniały jak największą część całkowitej inercji. W praktyce poszukując optymalnej przestrzeni wspólnej dla profili wierszowych i kolumnowych, najczęściej korzysta się z metody rozkładu macierzy A według wartości osobliwych (Greenacre, 1984; Stanimir, 2005).

Jednoczesne umieszczenie punktów reprezentujących obie kategorie zmiennych w jednym czynnikowym układzie współrzędnych powoduje duże zagęszczenie punktów, a stąd trudności interpretacji przeprowadzanej analizy. Często korzystna jest zatem niezależna analiza konfiguracji punktów reprezentujących zmienne albo obiekty. Współrzędne profili kolumnowych są wtedy wyliczane na podstawie macierzy profili wierszowych, a współrzędne profili wierszowych – na podstawie macierzy profili kolumnowych. Noszą one wtedy nazwę współrzędnych standardowych i są obliczane poprzez podzielenie ich współrzędnych głównych przez odpowiadające im wartości osobliwe.

Wybór przestrzeni czynnikowej

Często stosowaną w analizie korespondencji metodą analizy danych jest analiza konfiguracji punktów reprezentujących zmienne lub obiekty na rysunku. Gdy odtworzymy odległości pomiędzy punktami reprezentującymi obiekty albo zmienne w przestrzeni o maksymalnym wymiarze, odtwarzamy pierwotne konfiguracje punktów bez żadnych zniekształceń, tzn. zostają zachowane kąty między wektorami i odległości wektorów reprezentujące profile wierszowe (kolumnowe), a co za tym idzie, również odległości między punktami. Sytuacja ta jest tożsama z brakiem strat informacyjnych o badanym zjawisku przy przejściu od konfiguracji punktów reprezentujących wiersze (kolumny) w macierzy wyjściowej (umieszczonych w przestrzeni R^n/R^m) do konfiguracji tych punktów w przestrzeni czynnikowej R . Każde zmniejszenie maksymalnego wymiaru przestrzeni czynnikowej powoduje zniekształcenie konfiguracji wyjściowej punktów, co oznacza stratę informacji o badanym zjawisku.

Operowanie przestrzenią czynnikową o wymiarze niższym od maksymalnego powoduje konieczność oceny jakości odwzorowania punktów (stopnia zniekształcenia wyjściowej konfiguracji punktów przy ich rzutowaniu na podprzestrzeń o mniejszym wymiarze) w kolejnych wymiarach przestrzeni czynnikowej oraz w danej podprzestrzeni R^s ($s < S$).

Jakość odwzorowania poszczególnych punktów (wierszy lub kolumn) w danym wymiarze przestrzeni czynnikowej mierzona jest poprzez stosunek kwadratu odległości danego punktu w tym wymiarze od środka układu osi czynnikowych do kwadratu odległości tego punktu w przestrzeni czynnikowej o maksymalnym wymiarze od środka układu czynnikowego. Stosunek ten jest tożsamy ze stosunkiem udziału danego wymiaru w inercji punktu (kwadrat korelacji wektorów wierszowych (kolumnowych) z daną osią czynnikową), czyli określa, jaka część inercji punktu jest wyjaśniana przez dany wymiar. Czym wyższa wartość jakości, tym lepiej dany wymiar opisuje punkt. Suma wartości jakości odwzorowania danego punktu na wszystkich wymiarach przestrzeni czynnikowej równa jest jeden.

Znaczenie poszczególnych punktów reprezentujących obiekty i zmienne w tworzeniu poszczególnych wymiarów przestrzeni czynnikowej jest mierzone udziałem tych punktów w inercji tych wymiarów. Im wyższy relatywnie jest ten udział, tym ważniejszy jest analizowany punkt w definiowaniu danego wymiaru. Suma tych udziałów dla wszystkich punktów łącznie jest dla każdego z wymiarów przestrzeni czynnikowej równa jedności.

Interpretacja wyników

W wyniku zastosowania analizy korespondencji uzyskujemy wyniki, które mogą być prezentowane w dwojaki sposób, przy czym oba ujęcia nawzajem się uzupełniają, a mianowicie:

- w formie tabel, w których zawarte są współrzędne poszczególnych zmiennych i obiektów względem kolejnych osi czynnikowych, na podstawie których obliczane są odległości pomiędzy punktami reprezentującymi obiekty i zmienne;
- w formie graficznej, w postaci wykresów sporządzonych na podstawie tych tabel, będących dwumiarowymi przekrojami przestrzeni czynnikowych. Powstają w ten sposób mapy percepcji.

Możemy tworzyć dwa rodzaje wykresów. Na pierwszym z nich prezentowany jest jeden z typów punktów (wiersze albo kolumny) za pomocą współrzędnych standardowych. Drugi rodzaj wykresów prezentuje na tym samym wykresie obydwa typy punktów za pomocą współrzędnych głównych.

Analiza układów punktów reprezentujących zmienne lub obiekty w ujęciu geometrycznym pozwala na wizualne uchwycenie ogólnych prawidłowości w układach tych zmiennych lub obiektów. Musimy jednak w tym przypadku dokonać wyboru płaszczyzny, na którą będzie rzutowana konfiguracja punktów z przestrzeni czynnikowej o maksymalnym wymiarze. Powinna to być płaszczyzna, która w najmniejszym stopniu zniekształca konfigurację wektorów umieszczoną w przestrzeni czynnikowej o maksymalnym wymiarze. Taką płaszczyznę tworzy przekrój przestrzeni czynnikowej według dwóch pierwszych osi.

Analizując uzyskane wyniki rozpatrujemy następujące typy konfiguracji punktów:

- położenie punktów względem środka układu współrzędnych (centrum rzutowania),
- położenie względem siebie punktów odpowiadających obiektom albo zmiennym,
- położenie względem siebie punktów reprezentujących obiekty względem punktów reprezentujących zmienne i *vice versa*.

Punkty położone, w porównaniu z innymi punktami, blisko początku czynnikowego układu odniesienia (reprezentującego profil przeciętny) mają profile bliskie profilowi przeciętnemu, a położone relatywnie daleko – znacznie się od niego różniące. Jednocześnie im punkty te są położone dalej od początku układu odniesienia, tym mają większy wkład w tworzenie przestrzeni czynnikowej (większe zasoby informacyjne o badanym zjawisku).

Analiza położenia punktów zmiennych względem początku czynnikowego układu odniesienia wskazuje na stopień ich zróżnicowania w badanych obiektach na tle pozostałych rozpatrywanych zmiennych. Małe oddalenie zmiennej od początku układu punktu świadczy o niewielkim jej zróżnicowaniu w badanych obiektach w porównaniu ze zróżnicowaniem innych zmiennych. Duża odległość punktu zmiennej od początku układu czynnikowego wskazuje na jej silne zróżnicowanie w badanych obiektach w porównaniu z innymi zmiennymi.

Analiza odległości punktu obiektu od początku układu czynnikowego, będąca wypadkową odległości pomiędzy odpowiadającymi sobie zmiennymi w danym obiekcie i w obiekcie przeciętnym (centroidzie), pozwala na ocenę relatywnego (w stosunku do innych obiektów) stopnia nietypowości danego obiektu ze względu na strukturę opisujących go zmiennych.

Bliskie położenie punktów zmiennych w przestrzeni czynnikowej oznacza, że zmienne te kształtują się podobnie w badanych obiektach. Natomiast bliskie położenie punktów reprezentujących obiekty, wskazuje na podobieństwo struktury opisujących je zmiennych.

W analizie korespondencji nie można co prawda interpretować odległości pomiędzy punktami zmiennymi i punktami – obiektami, ale można interpretować położenie punktu zmiennej względem konfiguracji punktów obiektów lub położenie punktu obiektu względem konfiguracji punktów zmiennych. Przykładowo, bliskie położenie punktów zmiennych w stosunku do danego punktu obiektu wskazuje na to, że właśnie ze względu na wartości zmiennych reprezentowanych przez te punkty, dany obiekt wyróżnia się wśród innych badanych obiektów.

A.III.5.2. Analiza oddziaływania różnych czynników na wykluczenie z edukacji – analiza czynnikowa

Do wstępnego rozpoznania zależności pomiędzy czynnikami wpływającymi na wykluczenie z edukacji zastosowano analizę czynnikową. Jej wyniki zostały następnie wykorzystane do sformułowania postaci modeli ścieżek oddziaływania tych czynników na wykluczenie z edukacji.

Wprowadzenie

Celem analizy czynnikowej jest wyodrębnienie nowego zbioru hipotecznych zmiennych zwanych czynnikami, mniej licznego niż zbiór zmiennych obserwowalnych, charakteryzujących badane zjawisko. Zakłada się w niej, że każdą zmienną obserwowalną (wejściową) można przedstawić jako kombinację liniową pewnej liczby nieobserwowalnych zmiennych zwanych czynnikami, wspólnych dla całego zbioru zmiennych wejściowych oraz jednego nieobserwowalnego czynnika swoistego dla tej zmiennej. Jednocześnie czynniki wspólne i czynniki swoiste nie są ze sobą skorelowane. Założenia te dają podstawę do podziału wariancji każdej zmiennej wejściowej (będącej miarą jej zasobu informacyjnego o badanym zjawisku) na wariancję wyjaśnianą przez czynniki wspólne i wariancję wyjaśnianą przez czynnik swoisty. Analiza czynnikowa pozwala na znalezienie takiego zbioru czynników wspólnych oraz określenie ich relacji ze zmiennymi obserwowalnymi, który umożliwia wyjaśnienie struktury powiązań między zmiennymi obserwowalnymi.

Analiza czynnikowa była rozwijana w różnych ośrodkach naukowych (Anderson, Rubin, 1956; Harman, 1960; Morrison 1990; Thurston, 1945). W polskiej literaturze jest szeroko prezentowana m.in. w pracach Pluty (1986), Zakrzewskiej (1994), Jajugi (1993), Panka (2008), Ostasiewiczza (2012).

Model analizy czynnikowej

Zakładamy, że każda ze zmiennych wejściowych ma rozkład normalny. Każdą ze zmiennych wejściowych, po ich uprzednim wystandaryzowaniu, przedstawia się tutaj jako kombinację liniową zmiennych nieobserwowalnych, zwanych czynnikami wspólnymi (głównymi) przenoszącymi informacje wspólne dla zmiennych wejściowych, oraz pojedynczej zmiennej, zwanej czynnikiem swoistym (specyficznym) przenoszącym informacje zawarte wyłącznie w tej zmiennej wejściowej i niepowielane przez inne zmienne wejściowe.

Zależność tę można przedstawić następująco (Harman, 1960, s. 13):

$$Z=WF+DU \quad (\text{A.III.15})$$

gdzie:

- $Z=[z_{ij}]$ – wystandaryzowana macierz obserwacji ($m \times n$), przy czym z_{ij} jest wartością wystandaryzowanej j -tej zmiennej w i -tym obiekcie;
- $W=[w_{jl}]$ – macierz ładunków czynnikowych czynników wspólnych ($m \times s$), przy czym w_{jl} jest ładunkiem czynnikowym l -tego czynnika wspólnego dla j -tej zmiennej;

- $F=[f_{ij}]$ – macierz czynników wspólnych ($s \times n$), przy czym f_{ij} jest wartością i -tego czynnika wspólnego w j -tym obiekcie;
- $D=[diag(d)]$ – macierz diagonalna ładunków czynnikowych czynników swoistych ($m \times m$), przy czym d_j jest ładunkiem czynnikowym j -tego czynnika swoistego;
- $U=[u_{ij}]$ – macierz czynników swoistych ($m \times n$), przy czym u_{ij} jest wartością j -tego czynnika swoistego w i -tym obiekcie.

Zakłada się, że zarówno czynniki wspólne, jak i czynniki swoiste nie są między sobą skorelowane. Innymi słowy, wariancję (zasób informacyjny) każdej ze zmiennych wejściowych można rozłożyć na wariancję wyjaśnianą przez czynniki wspólne (zasoby informacyjne danej zmiennej wejściowej wspólne z zasobami informacyjnymi innych zmiennych wejściowych) oraz przez czynnik swoisty (zasoby informacyjne tej zmiennej wejściowej nie powielane przez inne zmienne wejściowe):

$$S^2(z_j) = h_j^2 + d_j^2 = \sum_{i=1}^s w_{ji}^2 + d_j^2 = 1 \quad j=1,2,\dots,m \quad (\text{A.III.16})$$

gdzie:

- h_j^2 – zasoby zmienności wspólnej j -tej zmiennej,
- d_j^2 – zasoby zmienności swoistej j -tej zmiennej.

W analizie czynnikowej dąży się do eliminacji wpływu czynników swoistych na rzecz czynników wspólnych, co jest równoważne z minimalizacją wpływu na kształtowanie się wartości zmiennych wejściowych wszystkich innych zmiennych poza czynnikami wspólnymi. Eliminacja tego wpływu odbywa się poprzez zastąpienie w macierzy korelacji R współczynników korelacji znajdujących się na głównej przekątnej (współczynników korelacji, których wartości są równe 1) zasobami zmienności wspólnej (wartościami najczęściej mniejszymi od 1). Uzyskujemy w ten sposób tzw. zredukowaną macierz korelacji o postaci:

$$\tilde{R} = \begin{cases} \tilde{r}_{jj'} = r_{jj'}, \text{ dla } j \neq j' \\ \tilde{r}_{jj'} = h_j^2, \text{ dla } j = j' \end{cases} \quad j, j' = 1, 2, \dots, m \quad (\text{A.III.17})$$

Zasoby zmienności wspólnej mogą być szacowane za pomocą różnych metod (Zakrzewska, 1994).

Ostatecznie układ równań (A.III.15) sprowadza się do postaci:

$$Z = WF \quad (\text{A.III.18})$$

Zredukowana macierz korelacji stanowi podstawę do wyznaczenia ładunków czynnikowych w kolejnych równaniach modelu (A.III.18) na podstawie zależności:

$$\tilde{R} = W W^T \quad (\text{A.III.19})$$

Istnieje szereg metod wyznaczania ładunków czynnikowych. Ich opis możemy znaleźć m.in. w opracowaniach Kima i Muellera (1978), Zakrzewskiej (1994) oraz pracy pod redakcją Gatnara i Walesiaka (2004).

Metody rotacji czynników

Uzyskany w wyniku stosowania przedstawionych wcześniej metod szacunku ładunków czynnikowych zbiór czynników nie jest jedynym możliwym układem czynnikowym przy ustalonej wcześniej wielkości zasobów zmienności wspólnej. Tę samą zredukowaną macierz korelacji możemy odtworzyć za pomocą nieskończenie wielu różnych macierzy ładunków czynnikowych otrzymanych w wyniku obrotu (rotacji) układu osi czynnikowych dookoła swojego początku, przy czym, co należy jeszcze raz podkreślić, udział czynników w wyjaśnianiu wspólnej wariancji (suma ich zasobów informacyjnych) nie ulega w wyniku rotacji zmianie.

Podstawowym celem rotacji układu osi czynnikowych jest uzyskanie jak najprostszej interpretacji poszczególnych czynników. Kryteria prostoty czynnikowej zmiennych, tzw. prostej struktury ładunków czynnikowych, zostały sformułowane przez Thurston'a (1945). Prowadzą one do wyodrębnienia rozłącznych grup zmiennych wejściowych, z których każda zawiera zmienne o wysokich ładunkach dla jednego czynnika oraz bliskie zero dla pozostałych czynników.

Interpretacja wyników

Czynniki przenoszą cały zasób zmienności wspólnej zmiennych wejściowych i są ze sobą nieskorelowane. Pozwala to przyjąć jako miarę zasobów informacyjnych dotyczących zmienności wspólnej zmiennych wejściowych przenoszonych przez l -ty czynnik wyrażenie:

$$h_l = \frac{\lambda}{\sum_{l=1}^s \lambda_l} \quad \text{A.III.20}$$

gdzie:

- λ_l – wariancja l -tego czynnika.

Tym samym udział skumulowanych zasobów informacyjnych pierwszych r czynników wyznaczamy jako:

$$H_r = \sum_{l=1}^r h_l \quad \text{(A.III.21)}$$

Ładunki czynnikowe są jednocześnie, ze względu na standaryzację zmiennych wejściowych, współczynnikami korelacji pomiędzy zmiennymi wejściowymi i czynnikami:

$$W_{jl} = r_{jl}, \quad j=1,2,\dots,m; l=1,2,\dots,s. \quad \text{(A.III.22)}$$

Kwadraty współczynników korelacji są zarazem współczynnikami determinacji pomiędzy zmiennymi wejściowymi i czynnikami. Określają, jakie odsetki wariancji wspólnej zmiennych wejściowych przenoszą poszczególne czynniki. Mogą tym samym służyć

nie tylko do interpretacji poszczególnych czynników, ale także do wyodrębnienia grup zmiennych wejściowych silnie ze sobą skorelowanych.

A.III.5.3. Analiza ścieżki oddziaływania czynników na wykluczenie z edukacji

Wprowadzenie

Najczęściej stosowanymi w analizach związków przyczynowych pewnego zespołu cech charakteryzujących badane zjawisko są modele regresji wielorakiej. Także w przypadku analiz wpływu czynników powodujących wykluczenie z edukacji stosowano ten typ modeli (np. Białecki, 2010; Kozarzewski, 2008). Modele te posiadają jednak znaczące mankamenty, pozostawiając bez odpowiedzi szereg problemów związanych z analizą przyczynowości. Przede wszystkim interpretacja relacji pomiędzy zmienną objaśnianą (tu wykluczeniem z edukacji na danym jej etapie) a zmiennymi predykcyjnymi wprowadzonymi do modelu (czynnikami powodującymi wykluczenie z edukacji) zostaje zawężone do pojęcia współzmienności między nimi. Najczęściej w analizie wariacyjnej modeli regresji wykorzystuje się pojęcie tzw. wariacji wyjaśnionej zmiennej objaśnianej przez zmienne predykcyjne, sugerując kierunkowe zależności pomiędzy zmiennymi występującymi w modelu. Ponadto w modelach regresji występuje zbyt uproszczona struktura relacji pomiędzy zmiennymi – zakłada się, że zmienne predykcyjne oddziałują na zmienną objaśnianą wyłącznie w sposób bezpośredni. W rzeczywistości sytuacja jest o wiele bardziej skomplikowana. Rozważane w danej analizie zmienne predykcyjne oddziałują na zmienną objaśnianą nie tylko bezpośrednio, lecz także w sposób pośredni, przez inne zmienne predykcyjne.

Powyższych mankamentów nie posiadają modele ścieżkowe wprowadzone na początku ubiegłego wieku przez Wrighta (1921, 1934). Modele *ścieżkowe* są szeroko prezentowane w literaturze obcojęzycznej m. in. w pracach Duncana (1975) i Loehlina (1987), a literaturze polskojęzycznej m.in. w opracowaniach Gaula i Machowskiego (2004) oraz Ostasiewicza (2012).

Modele te umożliwiają weryfikację hipotez dotyczących struktury zależności przyczynowych w określonym zbiorze potencjalnych czynników wykluczających z edukacji (zmiennych) poprzez obserwację bezpośrednich i pośrednich (przez inne zmienne predykcyjne) ich oddziaływań na to wykluczenie. Nazwa modele ścieżkowe pochodzi stąd, że analizowane są ścieżki zależności zmiennej objaśnianej od zmiennych predykcyjnych. W modelach ścieżkowych zakłada się relacje przyczynowe między zmiennymi, które następnie podlegają testowaniu. W przypadku, gdy nie ma podstaw do odrzucenia założonego modelu teoretycznego powiązań między zmiennymi, możemy interpretować założone relacje pomiędzy predyktorami i zmiennymi objaśnianymi (zarówno bezpośrednio, jak i pośrednio) jako przyczynowe i zastąpić hierarchizację predyktorów ze względu na siłę ich oddziaływania na zmienną objaśnianą uzyskiwaną w modelach regresji wielorakiej hierarchizacją predyktorów opartą na porównaniach wielkości ich efektów przyczynowych (bezpśrednich i pośrednich). Analiza ścieżkowa umożliwia również upraszczanie zaproponowanego teoretycznego modelu, poprzez eliminację z niego powiązań pomiędzy zmiennymi, dla których efekt związków przyczynowych jest bliski zeru.

Sformułowanie modelu ścieżki

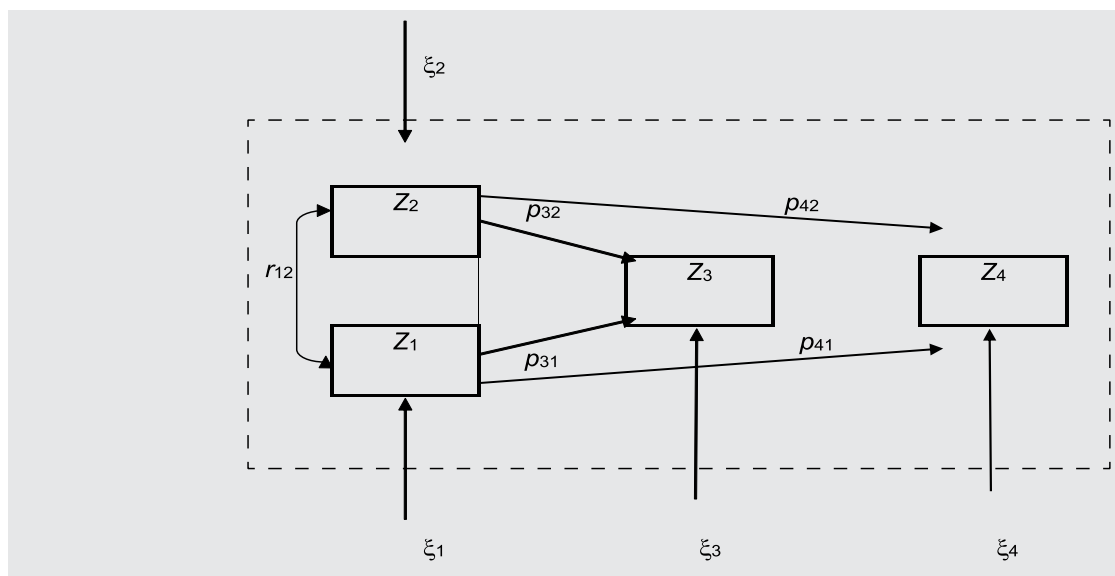
Model ścieżki możemy przestawić zarówno w formie graficznej, w postaci diagramu, jak i alternatywnie w postaci układu równań regresji.

Diagram ścieżki

Konstrukcja diagramu ścieżki nie jest niezbędna dla przeprowadzenia analizy ścieżki, ale znacznie ułatwia prezentację założonej w modelu struktury relacji przyczynowych.

Konstrukcję diagramu ścieżki wyjaśniono na przykładzie prostego modelu, dla którego diagram ścieżkowy został przedstawiony na rysunku A.III.6.⁹

Rysunek A.III.6. Diagram ścieżkowy



S. Wright sformułował trzy podstawowe zasady tworzenia ścieżki (Ostasiewicz, 2012, s. 276):

- ścieżkę stanowi sekwencja strzałek łączących symbole cech. Jeżeli w sekwencji strzałek wystąpiła strzałka do przodu (z lewej do prawej), to po niej nie może wystąpić strzałka do tyłu;
- ścieżka może zawierać tylko jeden raz jeden kierunek ze strzałki dwustronnej;
- ta sama zmienna nie może wystąpić dwa razy w tej samej ścieżce, czyli ścieżka nie może zawierać pętli.

W modelu ścieżkowym, a tym samym w reprezentującym go diagramie, wyróżnia się dwa typy zmiennych: egzogeniczne i endogeniczne. Zmienne egzogeniczne są zmienne, których zmienność wyjaśniana jest czynnikami nierozpatrywanymi w modelu (których zmienność nie jest wyjaśniana relacjami ze zmiennymi występującymi

⁹ W tym miejscu nie podajemy postaci diagramów ścieżkowych, które zostały ostatecznie zastosowane w analizie empirycznej.

w modelu). W przykładzie zmiennymi takimi są zmienne Z_1 i Z_2 ¹⁰. Natomiast zmiennymi endogenicznymi są te zmienne, których zmienność wyjaśniana jest oddziaływaniem innych zmiennych z modelu (w przykładzie są to zmienne Z_3 i Z_4). Zmienne te nazywane są także zmiennymi pośredniczącymi, gdyż obok oddziaływania na inne zmienne w modelu pośredniczą również we wpływie poprzedzających je w łańcuchu przyczynowo-skutkowym predyktorów na zmienną objaśnianą. Zakładane relacje pomiędzy zmiennymi reprezentują proste strzałki. Nad strzałkami łączącymi zmienne znajdują się współczynniki ścieżkowe określające siłę i kierunek zależności jednej zmiennej od drugiej. Pierwszy z indeksów oznacza zmienną zależną, a drugi zmienną determinującą jej zmienność. Stąd też kolejność indeksów przy współczynnikach jest ważna. Przykładowo współczynnik p_{32} określa kierunek i siłę zależności zmiennej Z_3 od zmiennej Z_2 . Na rysunku występuje także strzałka dwustronna w postaci łuku zakończony dwoma grotami reprezentująca korelację pomiędzy zmiennymi Z_1 i Z_2 . Tego typu korelacje traktowane są jako dane i pozostają poza ścieżkową analizą modelu. Na rysunku, oprócz zmiennych endogenicznych i egzogenicznych uwzględnionych w modelu (obwiedzonego przerywaną linią), znajdują się jeszcze zmienne resztowe reprezentujące czynniki losowe, czyli zmienne nieobjęte analizą, ale wpływające na zmienność zmiennych uwzględnionych w modelu.

Wartości współczynników ścieżkowych szacowane są na podstawie układu tzw. fundamentalnych równań analizy ścieżkowej o postaci:

$$r_{ij} = \sum_q p_{iq} r_{iq}, \quad (\text{A.III.23})$$

gdzie:

- q przebiega przez wszystkie zmienne, z których ścieżki bezpośrednio lub pośrednio prowadzą od j -tej do i -tej zmiennej;
- r_{ij} – współczynnik korelacji pomiędzy i -tą i j -tą zmienną.

Model ścieżki w postaci układu równań regresyjnych

Przedstawiony w sposób graficzny model możemy opisać w postaci równań regresyjnych. Równania w układzie równań konstruowane są dla każdej ścieżki oddziaływania uwzględnionej w modelu. Przyjmuje się przy tym następujące założenia co do modelu (Gaul i Machowski, 2004):

- relacje pomiędzy zmiennymi uwzględnionymi w modelu są przyczynowe, liniowe oraz addytywne;
- zmienne resztowe w modelu nie są ze sobą skorelowane i nie są skorelowane ze zmiennymi poprzedzającymi je w modelu;
- skorelowanie zmiennych egzogenicznych traktowane jest jako dane, a nie spowodowane wspólną przyczyną, i nie poddawane jest analizie.

¹⁰ Operujemy wystandaryzowanymi danymi, co ułatwia interpretację parametrów występujących w równaniach regresji opisujących model ścieżki.

Należy podkreślić, że rozważany model tworzy system rekursywny, w którym nie występują sprzężenia zwrotne, gdzie poszczególne zmienne mogłyby oddziaływać na same siebie¹¹.

W omawianym przykładzie model, przy założeniu, że zmienne w nim występujące zostały wystandaryzowane, określają cztery równania:

$$Z_1 = \xi_1, \quad (\text{A.III.24})$$

$$Z_2 = \xi_2, \quad (\text{A.III.25})$$

$$Z_3 = p_{31}Z_1 + p_{32}Z_2 + \xi_3, \quad (\text{A.III.26})$$

$$Z_4 = p_{41}Z_1 + p_{42}Z_2 + p_{43}Z_3 + \xi_4. \quad (\text{A.III.27})$$

Argumentami w równaniach dla zmiennych endogenicznych mogą być inne zmienne endogeniczne, zmienne egzogeniczne oraz zmienne resztowe (równania A.III.26 i A.III.27). Natomiast w równaniach dla zmiennych egzogenicznych argumentami mogą być tylko zmienne resztowe (równania A.III.24 i A.III.25). Estymacja i interpretacja modelu jest nazywana właśnie analizą ścieżki. Rozwiązywanie układu równań modelu jest tożsame z rozwiązywaniem kolejnych jego równań metodą najmniejszych kwadratów.

Ponieważ korzystamy z danych wystandaryzowanych, to współczynniki regresji występujące w równaniach modelu są wystandaryzowanymi współczynnikami beta, zwanymi także wagami beta, i są równe współczynnikom ścieżki:

$$\beta_{ij.k} = p_{ij}, \quad (\text{A.III.28})$$

gdzie:

- $\beta_{ij.k}$ – współczynnik beta oceniający zależność i -tej zmiennej od j -tej zmiennej, przy eliminacji wpływu k -tej zmiennej.

Wpływ czynników losowych (reprezentujących zmienne nie objęte analizą) wpływających na zmienność zmiennych uwzględnionych w modelu obliczamy na podstawie wzoru:

$$\xi_i = \sqrt{1 - R_{i.1, \dots, l}^2} \quad (\text{A.III.29})$$

gdzie:

- $R_{i.1, \dots, l}^2$ – współczynnik determinacji wielorakiej określający łączny efekt oddziaływania na i -tą zmienną wszystkich zmiennych uwzględnionych w modelu (poza czynnikami losowymi).

¹¹ Przedstawiany model ma charakter ogólny. Przyjęte w nim niektóre założenia dotyczące na przykład addytywności relacji pomiędzy zmiennymi, czy też braku występowania sprzężeń zwrotnych mogą zostać uchylone.

Wprowadzając do modelu zmienne porządkowe w celu ominięcia problemów związanych z założeniem o równości interwałów pomiędzy kolejnymi wartościami zmiennej (gdy za kolejne wartości zmiennej uznajemy wartości różniące się o jednostkę skali) lub o liniowości związku pomiędzy zmiennymi, możemy przedstawić je w postaci zmiennych instrumentalnych. Zmienne instrumentalne zawierają pełną informację co do przynależności badanych osób do kategorii (wariantów) zmiennej pierwotnej, a ich zastosowanie pozwala ominąć wskazane powyżej problemy¹².

Liczba zmiennych instrumentalnych reprezentujących daną zmienną jest o jeden mniejsza od liczby wariantów (kategorii) tej zmiennej. Przykładowo zmienną porządkową *poziom wykształcenia rodziców* posiadającą 4 warianty poziomu tego wykształcenia rozбивa się na 3 zmienne instrumentalne. Schemat tego rozбивa pokazuje tabela A.III.4.

Tabela A.III.4.

Rozбивie zmiennej porządkowej „poziom wykształcenia rodziców” na zmienne instrumentalne

Poziom wykształcenia rodziców X_1	Zmienne instrumentalne		
	X_{1A}	X_{1B}	X_{1C}
podstawowe lub niższe	1	1	1
średnie bez matury	0	1	1
średnie z maturą	0	0	1
wyższe	0	0	0

Rozбивając zmienną X_1 na zmienne instrumentalne X_{1A} , X_{1B} i X_{1C} przechodzi się od przykładowego modelu:

$$X_{WI} = \alpha_0 + \alpha_{WI1} X_1 \quad (\text{A.III.30})$$

do modelu:

$$X_{WI} = \alpha_0 + \alpha_{WI1A} X_{1A} + \alpha_{WI1B} X_{1B} + \alpha_{WI1C} X_{1C} \cdot \quad (\text{A.III.31})$$

Wartość współczynnika regresji α_{WI1A} informuje nas o oczekiwanej zmianie (spadku lub wzroście – w zależności od znaku tego współczynnika) zmiennej zależnej (ryzyka wykluczenia z l -tego danego poziomu edukacji) przy przechodzeniu z poziomu wykształcenia rodziców ponadgimnazjalnego bez matury do poziomu co najwyżej podstawowego, współczynnika α_{WI1B} przy przechodzeniu z poziomu wykształcenia średniego z maturą na poziom ponadgimnazjalny bez matury, a współczynnika α_{WI1C} przy przechodzeniu z poziomu wykształcenia wyższego do poziomu średniego z maturą.

Każda ze zmiennych instrumentalnych jest zmienną dychotomiczną, co skutkuje tym, że unika się problemu równości interwałów, jak i nieliniowości.

¹² Szerzej o problemach związanych z wprowadzaniem do modelu ścieżki zmiennych porządkowych traktuje np. praca R. P. Boyla (1978).

Testowanie modelu

Dla weryfikacji poprawności zaproponowanego modelu można zastosować test adekwatności Kima i Kohouta (1975). Opiera się on na porównaniu naszego teoretycznego modelu z modelem, w którym występują wszystkie możliwe bezpośrednie relacje przyczynowe między zmiennymi odtwarzające z definicji wyjściową macierz korelacji pomiędzy zmiennymi.

Weryfikuje się hipotezę, że usunięcie połączeń bezpośrednich (występujących w modelu pełnym) w rozważanym modelu nie zmniejsza znacząco wariacji wyjaśnianej zmiennymi endogenicznymi.

Statystyka testująca powyższą hipotezę ma następującą postać:

$$L = n \sum_i \log \left(\frac{(1 - R_{si}^2)/d_{si}}{(1 - R_{fi}^2)/d_{fi}} \right), \quad (\text{A.III.32})$$

gdzie:

- $R_{\bar{n}_i}^2, R_{si}^2$ – współczynniki determinacji wielorakiej określające wpływ na i -tą zmienną endogeniczną odpowiednio wszystkich poprzedzających ją predykatów z modelu pełnego oraz poprzedzających ją predykatów z modelu teoretycznego, z którymi ma bezpośrednie połączenie;
- $d_{\bar{n}_i}, d_{si}$ – liczba stopni swobody resztowych sum kwadratów z odpowiednich równań regresji dla i -tej zmiennej dla modelu pełnego i modelu teoretycznego.

Statystyka L posiada rozkład chi-kwadrat z liczbą stopni swobody:

$$df = \sum d_{si} - \sum d_{fi},$$

czyli równej liczbie ścieżek pominiętych w modelu testowanym w porównaniu do modelu pełnego.

Interpretacja modelu

Po zaakceptowaniu naszego hipotetycznego modelu (pozytywnej jego weryfikacji) można przejść do jego interpretacji, czyli oceny ogólnego wpływu poszczególnych zmiennych niezależnych i pośredniczących na zmienność zmiennych zależnych opisywanych przez kolejne równania regresji. Siłę wpływu ogólnego danej i -tej zmiennej (niezależnej lub pośredniczącej) na j -tą zmienną określają wartości współczynników korelacji r'_{ij} reprodukowanych na podstawie współczynników ścieżki (oraz ewentualnie na podstawie korelacji traktowanych jako danych r_{ij}). Ogólny wpływ danej i -tej zmiennej na j -tą zmienną można rozbić na trzy elementy, co jest tożsame z dekompozycją korelacji odtworzonych na trzy składowe, a mianowicie (Panek, 1999):

- wpływ bezpośredni danej zmiennej równy wartości współczynnika ścieżki prowadzącej od i -tej zmiennej do j -tej zmiennej (p_{ji});

- wpływ pośredni danej zmiennej na zmienne pośredniczące znajdujące się bliżej końcowej (j -tej) zmiennej w ciągu przyczynowym równy sumie iloczynów wartości współczynników wszystkich ścieżek dla wszystkich łańcuchów przyczynowych prowadzących od zmiennej i -tej do zmiennej j -tej (u_{ij});
- wpływ nieprzyczynowy (wspólny i pozorny) wynikający z wzajemnych korelacji rozważanych zmiennych z innymi zmiennymi modelu (w_{ij}) i równy różnicy pomiędzy korelacją reprodukowaną r'_{ij} , a sumą wpływu bezpośredniego i pośredniego i -tej zmiennej na j -tą zmienną.

Efekt wspólny jest związany z faktem, że część korelacji pomiędzy zmienną zależną a zmienną pośredniczącą jest rezultatem ich późniejszej korelacji z inną zmienną egzogeniczną modelu. Pozorna część wpływu ogólnego jest natomiast odbiciem wspólnej przyczyny zmienności i -tej i j -tej zmiennej.

A.III.5.4.

Pomiar wykluczenia społecznego

Punktem wyjścia dla pomiaru wykluczenia społecznego jest zdefiniowanie jego wymiarów, a następnie wybór zmiennych będących symptomami wykluczenia w poszczególnych wymiarach.

Symptomy wykluczenia społecznego mierzone są najczęściej w skali nominalnej i mają charakter binarny. W efekcie można wyłącznie określać występowanie lub niewystępowanie danego symptomu wykluczenia u badanej osoby, czyli ustalić, czy badana osoba podlega, czy też nie podlega wykluczeniu ze względu na dany symptom. W sytuacji, gdy pomiar symptomu wykluczenia jest możliwy na skalach mocniejszych, tj. na przynajmniej skalach porządkowych, możemy dodatkowo mierzyć inne aspekty wykluczenia danej osoby ze względu na dany symptom deprywacji. Tutaj poddany analizie zostanie zasięg i głębokość wykluczenia.

Miarą oceniającą zasięg wykluczenia osób ze względu na j -ty symptom wykluczenia w h -tym jego wymiarze jest stopa wykluczenia, czyli odsetek osób podlegających wykluczeniu ze względu na j -ty symptom w h -tym wymiarze:

$$H_{hj}^w = \frac{n_{hj}^w}{n}, \quad (\text{A.III.33})$$

gdzie:

- n_{hj}^w – liczba osób podlegających wykluczeniu w badanej zbiorowości ze względu na j -ty symptom w h -tym wymiarze.

W przypadku, w którym symptom wykluczenia mierzony jest na skali binarnej, dana osoba podlega wykluczeniu, gdy charakteryzuje się tym symptomem. W sytuacji, gdy symptom wykluczenia mierzony jest na skali mocniejszej (przynajmniej porządkowej), osoba podlega wykluczeniu ze względu na ten symptom, jeśli wartość tego symptomu dla danej osoby przekracza ustaloną wartość progową. W tej drugiej z przedstawionych sytuacji oceniana będzie dodatkowo głębokość wykluczenia badanych osób ze względu na dany symptom wykluczenia.

Wstępnie porządkuje się wartości, jakie może przyjmować dany symptom wykluczenia, według rosnącego stopnia zagrożenia wykluczeniem (zakładając, że symptomy wykluczenia mają charakter destymulant). Miarą oceniającą głębokość wykluczenia ze względu na j -ty symptom wykluczenia w h -tym wymiarze wykluczenia jest indeks luki wykluczenia osób jemu podlegających ze względu na dany symptom o postaci:

$$I_{hj}^w = \frac{1}{n_{hj}^w} \sum_{i=1}^{n_{hj}^w} \left(\frac{z_{hj}^* - z_{i,hj}}{z_{hj}^*} \right), \quad (\text{A.III.34})$$

gdzie:

- z_{hj}^* – wartość progowa j -tego symptomu wykluczenia w h -tym wymiarze, czyli wartość j -tego symptomu wykluczenia w h -tym wymiarze, powyżej której dana osoba nie podlega już wykluczeniu;
- $z_{i,hj}$ – poziom wykluczenia i -tej osoby ze względu na j -ty symptom wykluczenia w h -tym wymiarze.

W celu oceny zasięgu i głębokości wykluczenia w poszczególnych jego wymiarach przyjmuje się, że zagrożenie wykluczeniem osoby rośnie wraz ze wzrostem liczby występujących u niej symptomów wykluczenia w tych wymiarach. Następnie porządkujemy liczbę symptomów wykluczenia w danym wymiarze według malejącego stopnia wykluczenia (od największej liczby symptomów wykluczenia, która może występować w danym wymiarze do braku symptomów wykluczenia). W kolejnym kroku definiuje się dla każdego wymiaru wykluczenia zmienną z_h , przyporządkowując uporządkowanej malejąco liczbie symptomów wykluczenia kolejne liczby naturalne ($z_h = 0, 1, 2, \dots, k_h$). Indeks mierzący zasięg deprivacji materialnej w h -tym wymiarze jest stopa wykluczenia, czyli odsetek jednostek podlegających wykluczeniu w h -tym wymiarze:

$$H_h^w = \frac{n_h^w}{n} \quad (\text{A.III.35})$$

gdzie:

- n_h^w – liczba osób podlegających wykluczeniu w h -tym wymiarze, czyli liczba osób charakteryzujących się przynajmniej taką liczbą symptomów wykluczenia, przy której uważamy, że podlega ona wykluczeniu w h -tym wymiarze.

Ocenę głębokości wykluczenia w h -tym jej wymiarze stanowi indeks luki wykluczenia osób podlegających wykluczeniu w h -tym wymiarze definiowany następująco:

$$I_h^w = \frac{1}{n_h^w} \sum_{i=1}^{n_h^w} \left(\frac{z_h^* - z_{i,h}}{z_h^*} \right), \quad (\text{A.III.36})$$

gdzie:

- $z_{i,h}$ – wartość zmiennej z w h -tym wymiarze u i -tej osoby;
- z_h^* – granica wykluczenia w h -tym wymiarze, czyli górna granica liczby symptomów wykluczenia, przy której osoba jeszcze nie podlega wykluczeniu.

Ogólna ocena wykluczenia (we wszystkich jego wymiarach łącznie) wymaga analizy liczby wymiarów, w których badane osoby podlegają wykluczeniu.

W celu konstrukcji indeksów agregatowych oceniających zasięg i głębokość wykluczenia we wszystkich wyróżnionych jego wymiarach łącznie porządkujemy liczbę wymiarów wykluczenia według malejącego stopnia wykluczenia (od największej liczby wymiarów wykluczenia, w których badana osoba może podlegać wykluczeniu, do niewystępowania wykluczenia w żadnym z jego wymiarów). Następnie definiuje się zmienną, przyporządkowując uporządkowanej malejąco liczbie wymiarów wykluczenia kolejne liczby naturalne ($z=0,1,2,\dots,h$). W kolejnym kroku musimy ustalić wartość granicy wykluczenia, czyli górną granicę liczby wymiarów wykluczenia, przy której osoba jeszcze nie podlega wykluczeniu.

Indeksem oceniającym zasięg wykluczenia jest stopa wykluczenia, czyli odsetek osób podlegających wykluczeniu:

$$H^w = \frac{n^w}{n} \quad (\text{A.III.37})$$

gdzie:

- n^w – liczba osób podlegających wykluczeniu w badanej zbiorowości.

Miarą oceniającą głębokość wykluczenia jest indeks luki wykluczenia osób podlegających wykluczeniu, który definiujemy następująco:

$$I^w = \frac{1}{n^w} \sum_{i=1}^{n^w} \left(\frac{z^* - z_i}{z^*} \right), \quad (\text{A.III.38})$$

gdzie:

- z_i – liczba wymiarów wykluczenia u i -tej osoby,
- z^* – granica wykluczenia.

Pomiar wykluczenia społecznego w wyróżnionych w badaniu obszarach

W przypadku wykluczenia społecznego reprezentowanego przez pojedyncze symptomy (sytuacja dochodowa, rynek pracy) miara oceniająca zasięg wykluczenia w danym wymiarze (h), czyli stopa wykluczenia w tym wymiarze, przyjmuje postać:

$$H_h^w = \frac{n_{hj}^w}{n}, \quad j=1; h=1,2, \quad (\text{A.III.39})$$

gdzie:

- n_{hj}^w – liczba osób podlegających wykluczeniu w badanej zbiorowości ze względu na j -ty symptom wykluczenia w h -tym obszarze.

Osoby wykluczone w tych wymiarach to osoby, dla których symptom wykluczenia w tym obszarze przyjmuje wartość wyższą niż wartość progowa wykluczenia równa 0.

W celu szacunku głębokości wykluczenia społecznego w obszarach sytuacji dochodowej oraz rynku pracy przyporządkowuje się rangi kolejnym wariantom symptomu wykluczenia w danym obszarze po uporządkowaniu wariantów danego symptomu według malejącego stopnia wykluczenia ($c_{hj}=1, \dots, (k-1)_{hj}, k_{hj}$).

Następnie dla każdego respondenta podlegającego wykluczeniu społecznemu w danym jego wymiarze obliczamy wskaźnik luki wykluczenia w tym wymiarze:

$$u_{n,i} = \frac{(c_{hj}^* = k_{hj} - 1) - (c_{hji} - 1)}{c_{hj}^* = k_{hj} - 1}, \quad h=1,2; i=1,2, \dots, n_{hj}, \quad (\text{A.III.40})$$

gdzie:

- $c_{hj}^* = k_{hj}$ – ranga przyporządkowana wariantowi j -tego symptomu wykluczenia w h -tym wymiarze wykluczenia społecznego, przy którym wykluczenie w tym wymiarze już nie występuje;
- c_{hji} – ranga przyporządkowana wariantowi j -tego symptomu wykluczenia w h -tym wymiarze wykluczenia, którym charakteryzuje się i -ty respondent.

Indeks luki wykluczenia dla całej badanej populacji (osób podlegających wykluczeniu społecznemu w danym wymiarze) w h -tym jego wymiarze przyjmuje następującą postać:

$$I_h^w = \frac{\sum_{i=1}^{n_h^w} u_{h,i}}{n_h^w}, \quad h=1,2. \quad (\text{A.III.41})$$

W obszarze kompetencji cywilizacyjnych, w którym wyróżniono dwa symptomy wykluczenia o charakterze binarnym, stopa wykluczenia obliczana jest według wzoru:

$$H_h^w = \frac{n_h^w}{n}, \quad h=3, \quad (\text{A.III.42})$$

gdzie:

- n_h^w – liczba osób podlegających wykluczeniu w wymiarze kompetencji cywilizacyjnych, czyli liczba osób, które charakteryzują się przynajmniej jednym z wyróżnionych symptomów wykluczenia.

Dla oceny głębokości wykluczenia społecznego w obszarze kompetencji cywilizacyjnych definiuje się zmienną przyjmującą wartości równe kolejnym możliwym liczebnościom wystąpienia symptomów wykluczenia ($z_h=0,1,\dots,(k-1)_h$). Następnie, po uprzednim uporządkowaniu wartości tej zmiennej według malejącego stopnia wykluczenia, przyporządkowuje się jej rangi ($c_h=1,\dots,(k-1)_h, k_h$).

W kolejnym kroku obliczamy dla każdego respondenta podlegającego wykluczeniu społecznemu w wymiarze kompetencji cywilizacyjnych wskaźnik luki wykluczenia:

$$u_{h,i}^w = \frac{(c_h^* = k_h - 1) - (c_{h,i} - 1)}{c_h^* = k_h - 1}, \quad h=3; i=1,2,\dots,n_h, \quad (\text{A.III.43})$$

gdzie:

- $c_{h,i}$ – ranga przyporządkowana i -temu respondentowi w wymiarze kompetencji cywilizacyjnych;
- $c_h^* = k_h$ – ranga przyporządkowana wartości zmiennej z_h odpowiadająca sytuacji, gdy nie występuje żaden symptom wykluczenia w wymiarze kompetencji cywilizacyjnych.

Indeks luki wykluczenia w obszarze kompetencji cywilizacyjnych osób podlegających wykluczeniu w tym wymiarze wykluczenia społecznego obliczany jest według wzoru:

$$I_h^w = \frac{\sum_{i=1}^{n_h^w} u_{h,i}^w}{n_h^w}, \quad h=3. \quad (\text{A.III.44})$$

W celu konstrukcji agregatowych indeksów oceniających zasięg i głębokość wykluczenia społecznego we wszystkich jego wymiarach łącznie definiuje się zmienną przyjmującą wartości równe kolejnym możliwym liczebnościom wymiarów wykluczenia społecznego, w których występuje wykluczenie ($z=0,1,\dots,k-1$). Następnie, po uporządkowaniu wartości zmiennej według malejącej liczbą obszarów wykluczenia społecznego, przyporządkowuje się jej wartościom rangi ($c=1,2,\dots,k-1,k$). Przyjmuje się jednocześnie, że dana osoba nie podlega wykluczeniu społecznemu, jeżeli nie podlega ona wykluczeniu w żadnym z jego wymiarów.

Indeks oceniający zasięg wykluczenia społecznego (stopa wykluczenia społecznego) przyjmuje postać:

$$H^w = \frac{n^w}{n}, \quad (\text{A.III.45})$$

gdzie:

- n_w – liczba osób podlegających wykluczeniu przynajmniej w jednym z wyróżnionych wymiarów wykluczenia społecznego.

W celu oceny głębokości wykluczenia obliczamy dla każdego respondenta wskaźnik luki wykluczenia społecznego:

$$u_i = \frac{(c^* = k - 1) - (c_i - 1)}{c^* = k - 1}, \quad i=1,2,\dots,n, \quad (\text{A.III.46})$$

gdzie:

- c_i – ranga przyporządkowana i -temu respondentowi;
- $c^* = k$ – ranga przyporządkowana wartości zmiennej z odpowiadającą sytuacji, gdy nie występuje wykluczenie w żadnym z wyróżnionych wymiarów wykluczenia społecznego.

Indeks luki wykluczenia społecznego w populacji osób podlegających temu wykluczeniu ma postać:

$$I^w = \frac{\sum_{i=1}^{n^w} u_i^w}{n^w}, \quad (\text{A.III.47})$$

A.III.5.5. Analiza wpływu niskiego poziomu wykształcenia na wykluczenie w wyróżnionych obszarach życia

Wprowadzenie

Do badania wpływu niskiego poziomu wykształcenia na stopień wykluczenia społecznego ogółem (oraz na stopień wykluczenia w wyróżnionych jego wymiarach) wykorzystana zostanie metoda dopasowania (ang. *matching*) (Heckman, Ichmura, Todd, 1998) oparta o indeks skłonności (ang. *propensity score*) (Rosenbaum, Rubin, 1983).

Celem analizy będzie oszacowanie prawdopodobieństwa wystąpienia hipotetycznej sytuacji wykluczenia społecznego, w wyróżnionych jego wymiarach dla osób wykluczonych edukacyjnie (gdyby nie były one wykluczone edukacyjnie) oraz jego porównanie z prawdopodobieństwem wykluczenia społecznego tych osób w rzeczywistym, zaobserwowanym stanie ich edukacyjnego wykluczenia. Uśrednienie tych jednostkowych porównań (efektów jednostkowych) doprowadzi do oszacowania przeciętnego wpływu (efektu) wykluczenia edukacyjnego na prawdopodobieństwo wykluczenia w innych

wymiarach wykluczenia społecznego dla osób wykluczonych edukacyjnie. Efekt taki w literaturze anglojęzycznej nazywa się *average treatment effect on the treated (ATT)*. W tym badaniu jest to średni efekt wpływu wykluczenia edukacyjnego osób nim objętych na wykluczenie społeczne w innych wymiarach.

Wykorzystana w analizie metoda dopasowania sprowadza się do identyfikacji osób o podobnych, obserwowalnych charakterystykach w grupach osób wykluczonych i niewykluczonych edukacyjnie i porównywania ich prawdopodobieństw wykluczenia społecznego w wyróżnionych w badaniu jego wymiarach. Dla każdej osoby wykluczonej edukacyjnie zostanie zidentyfikowana z góry ustalona liczba „najbardziej podobnych” do niej osób z grupy osób niewykluczonych edukacyjnie (liczba tzw. osób-„połączeń”). Obserwowalne charakterystyki osób- „połączeń” posłużą za źródło informacji o hipotetycznej sytuacji (prawdopodobieństwie wykluczenia społecznego) osoby wykluczonej edukacyjnie, gdyby nie była ona wykluczona edukacyjnie.

Przyjęte w modelu oznaczenia

Niech I oznacza próbę losową osób będącą przedmiotem analizy. Zbiór ten składa się z elementów i , gdzie $i=1,2,\dots,n$.

Niech T oznacza zmienną definiującą wykluczenie z edukacji, tj. uzyskanie przez osobę wykształcenia gimnazjalnego lub niższego i niekontynuowanie edukacji. Zmienna T przyjmuje dwie wartości: 1 oznacza wystąpienie wykluczenia z edukacji, a 0 – brak wykluczenia z edukacji.

Zbiór I podzielimy na dwa podzbiory: osób wykluczonych edukacyjnie – I_1 oraz osób niewykluczonych edukacyjnie – I_0 , których liczebności w próbie wynoszą odpowiednio n_1 i n_0 . Niech X oznacza wektor obserwowalnych charakterystyk osób, wpływających na ich wykluczenie edukacyjne.

Przedmiotem badania będzie zmienna $Y(X,t)$, której wartości należą do przedziału $[0,1]$, oznaczająca prawdopodobieństwo wykluczenia społecznego w jednym z wyróżnionych jego wymiarów. Prawdopodobieństwo to jest funkcją charakterystyk X_i i -tej osoby oraz poziomu wykształcenia ewentualnego decydującego o wykluczeniu edukacyjnym $t \in T = \{0,1\}$.

Niech $m \in \mathbb{N}$ oznacza z góry ustaloną liczbę „połączeń”. Dla każdej osoby z grupy wykluczonych edukacyjnie szukamy m najbardziej podobnych pod względem charakterystyk wpływających na ich wykluczenie edukacyjne osób z grupy niewykluczonych edukacyjnie. Każda osoba z grupy niewykluczonych edukacyjnie będzie mogła stanowić potencjalne „połączenie” dla poszczególnych osób wykluczonych edukacyjnie. Do oszacowania wpływu wykluczenia edukacyjnego na wykluczenie społeczne wykorzystany zostanie model metody dopasowania „ze zwracaniem” – dla każdej osoby wykluczonej edukacyjnie będziemy niezależnie szukać zbioru m najbardziej podobnych do niej osób z grupy niewykluczonych edukacyjnie, co oznacza, że każda osoba niewykluczona edukacyjnie będzie mogła stanowić „połączenie” dla więcej niż jednej osoby wykluczonej edukacyjnie. Przez $K_m(i)$ oznaczymy liczbę osób wykluczonych edukacyjnie, dla których dana osoba $i \in I_0$ (czyli niewykluczona edukacyjnie) stanowi „połączenie”, przy ustalonej liczbie m poszukiwanych „połączeń”. $K_m(i)$ oznacza tym samym, ile razy dana osoba

niewykluczona edukacyjnie stanowi „połączenie” dla osób wykluczonych edukacyjnie w modelu metody dopasowania.

Założenia analizy

Pierwszym założeniem omawianej analizy jest istnienie wpływu wykluczenia edukacyjnego (T) na prawdopodobieństwo wystąpienia wykluczenia społecznego w innych jego obszarach (Y). Dla oszacowania siły tego wpływu należy znać wartości zmiennej Y dla osób wykluczonych edukacyjnie oraz dla tych samych osób w hipotetycznym stanie, w którym mają one wykształcenie wyższe niż gimnazjalne (nie są wykluczone edukacyjnie).

Tradycyjnym statystycznym narzędziem służącym poznaniu nieobserwowalnych, hipotetycznych stanów i wyciągania na tej podstawie wniosków o charakterze przyczynowo-skutkowym jest kontrolowany eksperyment z losowym przydziałem interesującego badacza oddziaływania. W podejściu eksperymentalnym do uzyskania informacji o zachowaniu się konkretnej osoby w nieobserwowalnych dla niej stanach wykorzystuje się informacje z pozostałej części próby/populacji objętej oddziaływaniem alternatywnym. Dzięki losowemu przydziałowi oddziaływania osoby objęte jego różnymi poziomami nie różnią się od siebie w żaden systematyczny sposób, a tym samym nie ma powodu sądzić, że na poszczególne poziomy oddziaływania reagują przeciętnie w różny sposób. Założenie to można zapisać jako:

$$Y(x, t) \perp T \quad (\text{A.III.48})$$

W kontekście badania wpływu poziomu wykształcenia na wykluczenie społeczne w jego innych obszarach podejście eksperymentalne jest bezużyteczne. Nie jest bowiem możliwe zorganizowanie eksperymentu, w którym poziom wykształcenia byłby losowo przydzielany do osób.

W rzeczywistości powyższe założenie również nie jest spełnione, gdyż osoby lepiej wykorzystujące korzyści wynikające z edukacji (lepiej wykorzystujące nabyte umiejętności w życiu zawodowym, etc.), posiadają większą skłonność do kontynuowania nauki. W związku z tym proste porównanie sytuacji osób wykluczonych edukacyjnie z osobami niewykluczonymi edukacyjnie, tj. z osobami lepiej wykształconymi, będzie obciążone estymatorem wpływu wykluczenia z edukacji na wykluczenie społeczne w innych jego wymiarach (efekt wyższego poziomu wykształcenia będzie przeszacowany).

Przy założeniu, że cechy wpływające na wykluczenie edukacyjne (lub charakterystyki z nimi silnie skorelowane) są obserwowalne, można wyeliminować obciążenie związane z auto-selekcją osób (Rosenbaum, Rubin, 1983). Założenie o niezależności przydziału wykształcenia od jego wpływu na wykluczenie społeczne można zastąpić założeniem warunkowej niezależności:

$$Y(x, t) \perp T \mid X \quad (\text{A.III.49})$$

W praktyce modelowania wektor obserwowalnych charakterystyk X zwykle podlega redukcji do zmiennej jednowymiarowej zwanej indeksem skłonności (ang. *propensity score*), który w omawianym badaniu będzie warunkowym prawdopodobieństwem

wystąpienia wykluczenia edukacyjnego przy danych wartościach wektora obserwowalnych charakterystyk X :

$$p(X) = P(T = 1|X) \quad (\text{A.III.50})$$

Oparcie modelu o indeks skłonności pozwala usunąć obciążenie związane z nielosowym przydziałem wykształcenia w porównywalnym lub nawet wyższym stopniu niż oparcie o cały wektor obserwowalnych charakterystyk (Rosenbaum i Rubin, 1983) oraz zwiększa efektywność oszacowań (Hahn, 1998). Jeśli bowiem spełnione jest założenie (A.III.49), to spełnione jest również założenie:

$$Y(x, t) \perp T|p(X) \quad (\text{A.III.51})$$

Kolejnym założeniem w omawianym modelu jest niezmienność wartości szacowanego efektu na poziomie jednostek w zależności od przydziału poziomu wykształcenia dla innych osób. (W literaturze anglojęzycznej założenie to nazywa się *stable unit-treatment value assumption* – *SUTVA*). W badaniach nad wpływem wykształcenia na wykluczenie społeczne osób założenie to jest zwykle niespełnione. Gdyby bowiem hipotetycznie zwiększyć poziom wykształcenia wszystkim osobom wykluczonym edukacyjnie, to średni efekt wykształcenia okazałby się niższy niż efekt oszacowany, zmieniłaby się bowiem sytuacja osób niewykluczonych edukacyjnie (np. w wymiarze rynku pracy: zmniejszenie podaży pracy nisko-wykwalfikowanej, obniżenie prestiżu dyplomu, etc. ze względu na wyższą konkurencję na rynku pracy wysoko-wykwalfikowanej). Założenie to przypomina, że w modelu ekonometrycznym szacujemy jedynie brzegowy efekt wpływu poziomu wykształcenia na wykluczenie społeczne w różnych jego wymiarach.

Budowa modelu efektów wpływu wykluczenia edukacyjnego na prawdopodobieństwo (ryzyko) wykluczenia społecznego w wyróżnionych wymiarach

W pierwszym etapie zostały oszacowane wartości indeksu skłonności ($p[X]$ – prawdopodobieństwa wykluczenia edukacyjnego) dla wszystkich osób w próbie w oparciu o wartości zaobserwowanych u nich charakterystyk. Charakterystyki te powinny w możliwie wysokim stopniu wyjaśniać prawdopodobieństwo wykluczenia edukacyjnego, ale jednocześnie nie mogą być to zmienne, na które wykluczenie edukacyjne oddziałuje zwrotnie, jeśli model ma zachować interpretację przyczynowo-skutkową. Dla przykładu – zarówno klasa miejscowości urodzenia, jak i klasa miejscowości zamieszkania danej osoby zostały skorelowane z prawdopodobieństwem wystąpienia wykluczenia edukacyjnego, jednak klasa miejscowości zamieszkania osób będzie zależeć od osiągniętego przez nie poziomu wykształcenia (ze względu na migrację osób wykształconych do dużych miast i pozostawanie osób wykluczonych edukacyjnie na wsi i w małych miastach) i nie może być użyta do szacowania indeksu skłonności. Indeks skłonności zostanie oszacowany przy pomocy modelu regresji probitowej (wartości teoretyczne funkcji regresji posiadają interpretację prawdopodobieństw).

Do oszacowania efektu wpływu wykluczenia edukacyjnego na fakt wykluczenia społecznego w innych jego wymiarach dla osób wykluczonych edukacyjnie posłuży estymator metody dopasowania (ang. *matching*) postaci (Imbens, Abadie, 2006):

$$\widehat{ATT} = \frac{1}{n_1} \sigma_{i=1}^n (T_i - (1 - T_i) \frac{K_m(i)}{m}) \hat{Y}_i \quad (\text{A.III.52})$$

gdzie:

- \hat{Y}_i – oszacowane przy pomocy modelu regresji probitowej wartości prawdopodobieństwa wykluczenia społecznego i -tej osoby w jednym z wyróżnionych jego wymiarów.

Bliskość osób, tzn. ich podobieństwo ze względu na skłonność do wykluczenia społecznego, będzie definiowana odległością euklidesową w jednowymiarowej przestrzeni indeksu skłonności (ang. *propensity score*).

Parametr oszacowany według wzoru (A.III.52) stanowi ocenę średniego wpływu wykluczenia edukacyjnego na prawdopodobieństwo wystąpienia wykluczenia w innych wyróżnionych jego wymiarach dla osób wykluczonych edukacyjnie. Zostanie on oszacowany oddzielnie dla każdego z wyróżnionych wymiarów wykluczenia społecznego, w których może dojść do wykluczenia na skutek niskiego poziomu wykształcenia.

Parametry modeli efektu wpływu wykluczenia edukacyjnego na wykluczenie społeczne w innych jego wymiarach zostały oszacowane z użyciem pakietu Stata 12 i dodatkowego oprogramowania match (Imbens, Abadie, Drukker, Leber Herr, 2001). Błędy standardowe szacunku zostały obliczone metodą analityczną zaproponowaną przez Imbensa (Imbens, Abadie, 2006).

A.III.6. Wewnętrz- i międzypokoleniowe zależności kształtowania kapitału ludzkiego w gospodarstwach domowych

A.III.6.1. Cel i przedmiot analizy

Celem analizy w tym temacie badawczym jest identyfikacja czynników wpływających na kształtowanie kapitału ludzkiego w ujęciu międzypokoleniowym oraz wewnątrzpokoleniowym.

Międzypokoleniowa transmisja kapitału ludzkiego odnosi się do przekazania indywidualnych kompetencji, cech, zachowań oraz efektów z rodziców na dzieci. W literaturze ekonomicznej najczęściej analizowane są takie elementy jak: transmisja poziomu wykształcenia, dochodów i zarobków, zasobności, decyzji dotyczących dzietności oraz korzystania z transferów społecznych (tzw. dziedziczenie biedy). W krajach o silnym dziedziczeniu międzypokoleniowym dzieci w dużym stopniu replikują strukturę społeczno-ekonomiczną pokolenia rodziców, co skutkuje niską mobilnością społeczną (Lochner, 2007).

W literaturze często podnoszony jest wątek dziedziczenia wzorców międzypokoleniowych. Bellani (2013) wskazuje na dwa główne kanały transmisji prowadzące do replikacji struktur społeczno-ekonomicznych. Pierwszy z nich to uwarunkowania monetarne, czyli brak środków wśród uboższych rodzin dla inwestycji w zdrowie i edukację dzieci (Becker, Tomes, 1979; Banerjee, Newman, 1994; Loury, 1981; Carneiro, Heckman, 2002). Drugim kanałem są czynniki o charakterze niemonetarnym, które w sposób negatywny wpływają na sytuację dzieci w przyszłości, takie jak niski poziom umiejętności czy samotne rodzicielstwo (Corcoran i Adams, 1997). Odziedziczone lub nabyte w środowisku

rodzinnym cechy, takie jak poziom inteligencji czy kompetencje społeczne, mają wpływ na dobrostan osób w dorosłości, co pokazują ostatnie badania dotyczące wpływu indywidualnych charakterystyk, w tym ogólnych i zawodowych umiejętności, na osiągnięcia osób (np. Anger, Heineck [2010]; Bjrkklund i in. [2010]; Black i in. [2009]; Heineck, Anger [2010]; Heckman, Rubinstein [2001]; Heckman i in. [2006]). Badania podejmowane w ostatnich latach wskazują również na istnienie bezpośredniego powiązania wykształcenia rodziców i dzieci, szczególnie w przypadku rodziców z wyższym wykształceniem, którzy preferują, aby również ich dzieci osiągnęły ten poziom wykształcenia, podejmując działania mające na celu socjalizację dzieci do specyficznego (właściwego dla nich) modelu kulturowego (np. Bisin, Verdier, 2000).

Najczęściej wymienianym w literaturze modelem zakładającym znaczący wpływ rodziców na kapitał ludzki i osiągnięcia dzieci na rynku pracy przez początkowe inwestycje w ich kapitał ludzki jest model Beckera i Tomesa (Becker, Tomes, 1979, 1986). Sformułowali oni teoretyczne podstawy modelu międzypokoleniowych zależności kształtowania kapitału ludzkiego, w tym wpływu pokolenia rodziców (oraz wcześniejszych pokoleń) na edukację oraz dochody dzieci w zakresie transmisji zarobków, aktywów oraz konsumpcji między pokoleniami. Model ten zakłada, że rodzice nie tylko przekazują dzieciom spadek, ale także wpływają na ich zarobki w dorosłym życiu przez inwestycje w ich umiejętności, zdrowie, uczenie się, motywację i inne charakterystyki.

Do charakterystyk tych zaliczamy m.in. nakłady (środków oraz czasu) rodzin na edukację, które zależą nie tylko od potencjału dzieci, ale także od dochodów, preferencji rodziców, a także liczby dzieci w rodzinie. Ważnym czynnikiem są również publiczne wydatki na edukację i szerzej rozumiany kapitał ludzki dzieci gromadzony w ramach otoczenia edukacyjnego i społecznego, w którym funkcjonują.

Ponieważ dla większości osób wynagrodzenie z pracy jest głównym źródłem dochodu, rodzice wpływają na sytuację ekonomiczną swoich dzieci przez wpływ na ich potencjalne zarobki. Becker i Tomes w swoim modelu zakładają, że zamożność rodziców nie jest jedynym czynnikiem wpływającym na mobilność międzypokoleniową, rozumianą przede wszystkim jako poprawa statusu edukacyjnego kolejnych pokoleń. Mobilność międzypokoleniowa obserwowana jest również w przypadku biedniejszych rodzin, skłonnych do finansowania inwestycji w dzieci, w celu uzyskania przez nie awansu edukacyjnego. Poziom mobilności międzypokoleniowej zależy również od liczby dzieci w rodzinie. Większa liczba dzieci ogranicza kwoty inwestycji w każde z nich, jeżeli są one finansowane przez rodzinę. W konsekwencji ujemna korelacja pomiędzy wielkością gospodarstwa domowego i dochodami rodziców ogranicza mobilność międzypokoleniową.

Model Beckera-Tomesa można również wykorzystać do analizy dziedziczenia wzorców edukacyjnych dla wielu pokoleń. Lindahl, Palme, Massih i Sjögren (2012) stosują to podejście do badania związków pomiędzy wykształceniem i dochodami dziadków oraz wnuków, znajdując istotne powiązania międzypokoleniowe łącznie dla czterech badanych generacji w oparciu o empiryczne dane pochodzące z rejestrów w Szwecji.

W przypadku edukacji w modelu BT związek pomiędzy edukacją dzieci i rodziców można zapisać jako:

$$s_t = \alpha + \beta s_{t-1} + \rho e_{t-1} + u_t$$

tzn. edukacja (s) pokolenia dzieci t jest addytywną funkcją liniową edukacji w pokoleniu rodziców $t-1$, nieobserwowanych zdolności i „spadku” w postaci określonych zachowań i wzorców kulturowych (e) oraz błędu (u), który ujmuje szoki zewnętrzne mające wpływ na s_t . Założenia u_t nie jest skorelowane z e_{t-1} oraz s_{t-1} . Współczynnik β powinien przyjmować wartości dodatnie ze względu na założony w modelu zwrot z inwestycji rodziców w edukację dzieci. Model BT zakłada również transmisję „spadku” pomiędzy pokoleniami, który jest procesem o charakterze autoregresyjnym:

$$e_t = \alpha + \lambda e_{t-1} + v_t$$

W tym przypadku zakłada się, że błąd losowy v_t jest nieskorelowany z e_{t-1} , u_t oraz s_{t-1} .

Drugim występującym w literaturze modelem rozwoju kapitału ludzkiego jest model, który skupia się na decyzjach dotyczących własnego rozwoju edukacyjnego, podejmowanych przez osoby, szczególnie młode, decydujące się na naukę na poziomie ponadgimnazjalnym oraz wyższym, w zależności od ich początkowego „wyposażenia” wynikającego z inwestycji rodziców. Model ten uwzględnia zatem zarówno pewne aspekty międzypokoleniowego dziedziczenia kapitału ludzkiego, jak i uwarunkowań wewnątrzpokoleniowych. Podejście to stosowane jest w analizach opartych o modele regresji, odnoszących się do wyborów edukacyjnych w zależności od osiągniętych wyników na egzaminach i testach, sytuacji rodzinnej, najbliższego otoczenia, a także charakterystyk grupy rówieśniczej (Haveman, Wolfe, 1995; Rainey, Murova, 2004; Tobias, 2003). Zanim młode osoby osiągną wiek, w którym rozpoczynają naukę na poziomie wyższym, można już zauważyć istotne różnicowania kompetencji pomiędzy nimi. Potwierdzają to m.in. wyniki Międzynarodowego Badania Kompetencji Osób Dorosłych (Organization for Economic Co-operation and Development, 2013a; Rynko, 2013). Różnice te wynikają ze zróżnicowania odziedziczonych zdolności oraz kapitału ludzkiego nabytego w poprzednich latach. Z kolei kapitał ludzki, którym te osoby dysponują odzwierciedla nakład czasu oraz innych zasobów dokonany przez rodziców, nauczycieli, rodzeństwo i samego zainteresowanego.

Ważnym przedmiotem analiz, co wynika z przedstawionego powyżej modelu decyzji edukacyjnych, są również uwarunkowania wewnątrzpokoleniowe, czy inaczej – wpływ rówieśników na podejmowane decyzje dotyczące dalszej nauki i akumulowania kapitału ludzkiego. Warto tutaj zwrócić uwagę na analizy, które odnoszą się do relacji dotyczącej osiągnięć rówieśników w przeszłości jako czynnika wyjaśniającego obecne osiągnięcia (Burke, Sass 2008; Hanushek, Kain, Markman, Rivikin, 2003, Vigdor, Nechyba, 2007). Wewnątrzpokoleniowe relacje związane z kształtowaniem kapitału ludzkiego są analizowane również w kontekście relacji pomiędzy partnerami (małżonkami) (Chong, Hongbin, Pak Wai, Junsen, 2009).

A.III.6.2. Zakres analizy

Analizy w ramach obszaru badawczego dotyczą uwarunkowań międzypokoleniowych odnoszących się do kształtowania kapitału ludzkiego w gospodarstwach domowych oraz uwarunkowań wewnątrzpokoleniowych.

Uwarunkowania międzypokoleniowe

W obszarze uwarunkowań międzypokoleniowych analiza obejmuje przede wszystkim międzypokoleniową transmisję kapitału ludzkiego w ramach gospodarstwa domowego. Statystyki GUS (zob. Główny Urząd Statystyczny, 2009) pokazują silny związek wykształcenia dzieci z wykształceniem rodziców. Celem badania jest ukazanie szerszego wpływu poprzedniego pokolenia na wybory edukacyjne następnych generacji. Analizy obejmują, w zależności od dostępnych danych, trzy kolejne pokolenia (dziadków, rodziców oraz dzieci).

Zakres transmisji międzypokoleniowej został przeprowadzony odrębnie dla grup osób o różnym poziomie wykształcenia (wyższe, średnie i zawodowe, podstawowe i niższe) dla lepszej identyfikacji czynników mających wpływ na mobilność edukacyjną (lub jej brak). Ponadto analizy obejmują różne grupy wieku (kohorty), co pozwoliło na ocenę, na ile w czasie zmieniały się w Polsce uwarunkowania mobilności międzypokoleniowej (analizy takie dla Wielkiej Brytanii przeprowadził anp. Machin [2004]).

Czynniki uwzględnione w analizach objęły¹³ (por. m.in. Bellani, 2013):

- liczbę lat w edukacji rodziców,
- wykształcenie rodziców,
- zawód rodziców,
- dochody gospodarstwa domowego,
- wyposażenie gospodarstwa domowego w książki,
- wartości wyznawane w rodzinie (znaczenie edukacji i kształcenia),
- inne czynniki (wczesne rodzicielstwo, niepełnosprawność).

Dla pogłębienia analizy uwarunkowań międzypokoleniowego dziedziczenia kapitału ludzkiego, zostały również uwzględnione zmienne instrumentalne:

- charakterystyka miejsca zamieszkania w okresie młodości (wieś/miasto),
- sytuacja na rynku pracy,
- charakterystyka systemu edukacji (w tym uwzględnienie zmiany systemu edukacji w 1999 roku).

W ramach przeprowadzonych analiz zostały oszacowane macierze przejścia (*transition matrices*) dotyczące poziomu wykształcenia, rodzaju wykształcenia (zawodowe/ogólne) oraz osiągniętych dochodów dla wybranych generacji, zgodnie z podejściem stosowanym m.in. przez Lindahl i in. (2012).

¹³ Pełny zakres analiz będzie możliwy po drugiej rundzie badania ze względu na zakres danych ujętych w kwestionariuszu. W pierwszej rundzie analizy obejmują cechy takie jak: wykształcenie rodziców, charakterystyki domu rodzinnego oraz motywacje wyborów edukacyjnych.

Przeprowadzone analizy miały na celu weryfikację empiryczną modelu Beckera-Tomesa dla Polski (w oparciu o modele regresji MNK oraz IV¹⁴).

W ramach analiz dokonana również została ocena wpływu kolejności urodzenia na uczestnictwo w edukacji (w tym na wykształcenie oraz długość okresu uczenia się) oraz sytuację na rynku pracy (w tym poziom dochodów) przy jednoczesnym uwzględnieniu sytuacji społeczno-ekonomicznej rodziny (Behrman, Taubman, 1986).

Odrębnie rozpatrywaną kwestią są inwestycje w kapitał ludzki ponoszone na poziomie gospodarstwa domowego (por. np. Organization for Economic Co-operation and Development, 2011), z czym wiąże się konieczność dokonywania wyborów w ramach gospodarstwa. W szczególności dokonano analizy relacji pomiędzy aktywnością edukacyjną dorosłych oraz dzieci i młodzieży zamieszkałych w gospodarstwach domowych oraz roli i skali transferów edukacyjnych w ramach gospodarstwa domowego rozumianych jako czas poświęcony przez starsze pokolenia na uczenie się dzieci i młodzieży.

Uwarunkowania wewnątrzpokoleniowe

W ramach tego obszaru rozważano siłę wpływu grup rówieśniczych na aktywność edukacyjną oraz wybory edukacyjne osób. Ze względu na zakres badania i sposób doboru próby analiza nie uwzględniała stosowanych w literaturze metod odnoszących się do oceny grupy rówieśniczej w ramach np. jednego roku nauki w konkretnej placówce edukacyjnej (np. Vigdor, Nechyba, 2007; Carrell i in., 2009).

Zastosowane narzędzie badawcze pozwoliło na przeprowadzenie analiz dotyczących wpływu innych osób oraz znaczenia wyborów edukacyjnych znajomych dotyczących decyzji wyboru szkoły lub uczelni, w której analizowane grupy osiągnęły najwyższy poziom wykształcenia. Analizy te prowadzone są łącznie dla wszystkich respondentów, a także w podziale na grupy (według wieku, płci, poziomu wykształcenia, miejsca zamieszkania). Pozwoliło to na lepszą identyfikację zróżnicowania wpływu grup rówieśniczych na podejmowanie decyzji. Zastosowane metody analizy (modele regresji) pozwoliły na ocenę znaczenia grupy rówieśniczej oraz innych czynników na dokonane wybory edukacyjne.

Dla osób dorosłych dokonane zostaną analizy mobilności wewnątrzgeneracyjnej w ramach klas społecznych – wykorzystane zostanie podejście zaproponowane do analizy mobilności pomiędzy klasami przez Keller (2011), w którym jako czynnik wyjaśniający mobilność pomiędzy klasami uwzględniony będzie poziom wykształcenia. Wyniki przedstawione przez Keller (2011) wskazują na istotne statystycznie znaczenie wykształcenia jako czynnika wpływającego na pozycję klasową. Jednocześnie wyższe wykształcenie zwiększa prawdopodobieństwo awansu klasowego.

Ponadto analizowane będą związki pomiędzy wykształceniem małżonków a poziomem zarobków (por. Aström, 2011) oraz uczestnictwem w rynku pracy (por. Papps, 2010).

Analizy te będą przeprowadzone po II rundzie badania.

¹⁴ Metoda IV będzie zastosowana po drugiej rundzie badania w odniesieniu do dochodów osób dorosłych.

W perspektywie ostatnich kilkunastu lat dokonany został znaczący postęp zarówno w obszarze teorii, jak i badań empirycznych odnoszących się do mobilności międzypokoleniowej. Black i Devereux (2010) dokonują podsumowania zarówno w zakresie stosowanych i prezentowanych w literaturze metod badawczych, jak i obecnych w niej wyników. W swoim przeglądzie uwzględniają zarówno międzypokoleniową transmisję zarobków, jak i edukacji (wykształcenia). W ostatnich latach badania w coraz większym stopniu dotyczą identyfikacji mechanizmów i przyczyn powodujących międzypokoleniową transmisję dochodów i wykształcenia, które mogą wyjaśnić obserwowane korelacje. Stosowane metody bazują w szczególności na ocenie korelacji efektów pomiędzy rodzeństwem, uwzględniając również efekty zamieszkania w jednej okolicy (sąsiedztwie). Celem tego typu analiz jest identyfikacja efektów genetycznych oraz wynikających z uwarunkowań środowiskowych. Inne metody opierają się o analizy wyników osiągniętych przez bliźniaki, jak również dzieci adoptowane. Black i Devereux (2010) pokazują, że wyniki badań wskazują na znaczenie zarówno czynników genetycznych, jak i środowiskowych w transmisji dochodów i wykształcenia. Podkreślają również potrzebę dalszych badań w tym obszarze. Kolejnym stosowanym w literaturze modelem analiz jest uwzględnienie naturalnych eksperymentów oraz metody zmiennych instrumentalnych. W literaturze analizowane są między innymi: wpływ różnego typu programów wsparcia dochodowego dzieci i rodziców, jak również wpływ szoków dochodów rodziców, związanych ze zmianą ich statusu na rynku pracy na kapitał ludzki ich dzieci. Wnioski, jakie płyną z tych analiz, wskazują, że szok dotyczący jednego pokolenia ma również wpływ na sytuację kolejnych generacji.

Do analiz związków przyczynowych pomiędzy wykształceniem dzieci i rodziców najczęściej stosowana jest metoda zmiennych instrumentalnych. W tym przypadku wykształcenie rodziców podlega zewnętrznemu szokowi, na przykład na skutek zmian polityki. Zmiana ta skutkuje potem również na dzieci. Black, Devereux i Salvanes (2005) stosują tę metodę, analizując dane pochodzące z danych rejestrowych w Norwegii, wnioskuje oni, że dowody na występowanie przyczynowej relacji pomiędzy wykształceniem dzieci i rodziców są mało znaczące. Oreopoulos, Page i Stevens (2006) stosują podobną metodę do analizy danych w Stanach Zjednoczonych. Ich wyniki wskazują, że wydłużenie okresu uczenia się przez rodziców o jeden rok, skutkuje zmniejszeniem prawdopodobieństwa powtarzania klasy przez dzieci od 2 do 7 punktów procentowych. Inne badania (Ceroneiro, Meghir, Parey, 2007; Chevalier, 2004; Maurin, Mc Nally, 2008) wskazują na występujący związek przyczynowy pomiędzy wykształceniem rodziców i efektami osiąganymi przez dzieci.

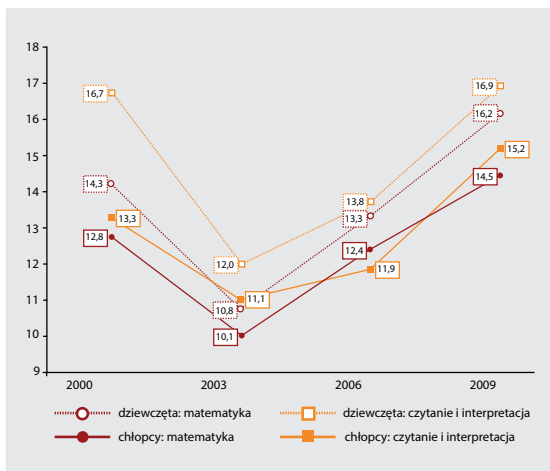
Prowadzone w ostatnich latach badania poszerzają również zakres analizowanych relacji pomiędzy pokoleniami, uwzględniając nie tylko dochody czy poziom wykształcenia. Z perspektywy badania UDE warto zwrócić uwagę na analizy odnoszące się do zawodów oraz stanowisk pracy. W kilku badaniach dotyczących tego obszaru identyfikowane są silne korelacje pomiędzy ojcami i synami. Hellerstein i Morrill (2008) pokazują, że w obecnej populacji około 30% synów i 20% córek pracuje w tym samym zawodzie, co ich ojcowie, chociaż odsetek ten zależy od tego, jak szczegółowo definiowany jest zawód. W przypadku analiz, w których zawody definiowane są szerzej, w oparciu o ocenę ich prestiżu (np. Ermish i Francesconi, 2002), wnioskuje się, że korelacja międzypokoleniowa wynosi od 0.4 do 0.75 w przypadku par ojców – dziecko oraz od 0.3 do 0.5 w przypadku par matki – dziecko.

Dotychczas prezentowane analizy dotyczące międzypokoleniowych i wewnątrzpokoleniowych uwarunkowań dziedziczenia kapitału ludzkiego w Polsce prowadzone były głównie w oparciu o wyniki badania PISA, a także o wyniki badania *Kształcenie dorosłych* prowadzonego przez GUS.

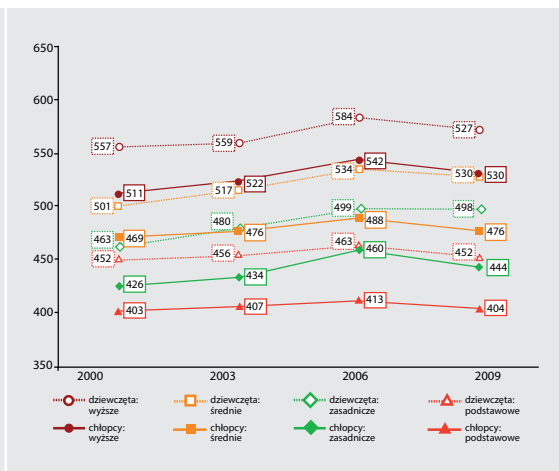
Analizy przeprowadzone w oparciu o wyniki badań PISA w latach 2000-2009 (Federowicz, 2009) miały na celu ocenę, w jaki sposób wprowadzenie gimnazjów do systemu szkolnego wpłynęło na zależność między wykształceniem rodziców a osiągnięciami uczniów oraz w jakim stopniu wykształcenie rodziców „wyjaśnia” (w sensie analizy statystycznej) zróżnicowanie wyników uczniów. Otrzymane wyniki wskazują, że istnieje związek pomiędzy wykształceniem rodziców a osiągnięciami uczniów. Związek ten widoczny był w roku 2000, a więc okresie sprzed wdrożenia reformy systemu edukacji wprowadzającej gimnazja. Jednym z celów reformy z 1999 roku miało być zwiększenie roli szkół i systemu edukacji w wyrównywaniu szans. Wyniki z 2003 roku wskazują, że ten cel został osiągnięty, tj. zmniejszyły się zależności wyników uczniów i wykształcenia ich rodziców. Niemniej jednak w kolejnych latach rola wykształcenia rodziców w wyjaśnianiu zmienności wyników uczniów wzrasta (rysunek A.III.7). Analizy prowadzone w podgrupach (rysunek A.III.8) wskazują, że chłopcy pochodzący z domów, w których rodzice mają wykształcenie wyższe, osiągają podobne wyniki do dziewcząt, których rodzice mają wykształcenie średnie, co wskazuje na to, że płeć ma znaczenie w zakresie aspiracji edukacyjnych uczniów.

Podsumowując, przez przeprowadzone analizy Federowicz (2009) wskazuje, że rodzice, którym zależało na wykształceniu dzieci, coraz większą uwagę zaczęli poświęcać dobremu ich przygotowaniu do egzaminu gimnazjalnego, rozumiejąc, że jego wyniki decydują o możliwościach wyboru szkoły ponadgimnazjalnej. Również sam wybór szkoły tego szczebla coraz silniej zależy od zasobów materialnych i kapitału kulturowego rodziców dziecka. Sprzyja temu obserwowane rozwarstwianie się liceów ogólnokształcących.

Rysunek A.III.7. Zróżnicowanie wyników piętnastoletnich dziewcząt i chłopców w zależności od wykształcenia rodziców. Odsetek zróżnicowania wyników wyjaśnionego przez wykształcenie rodziców



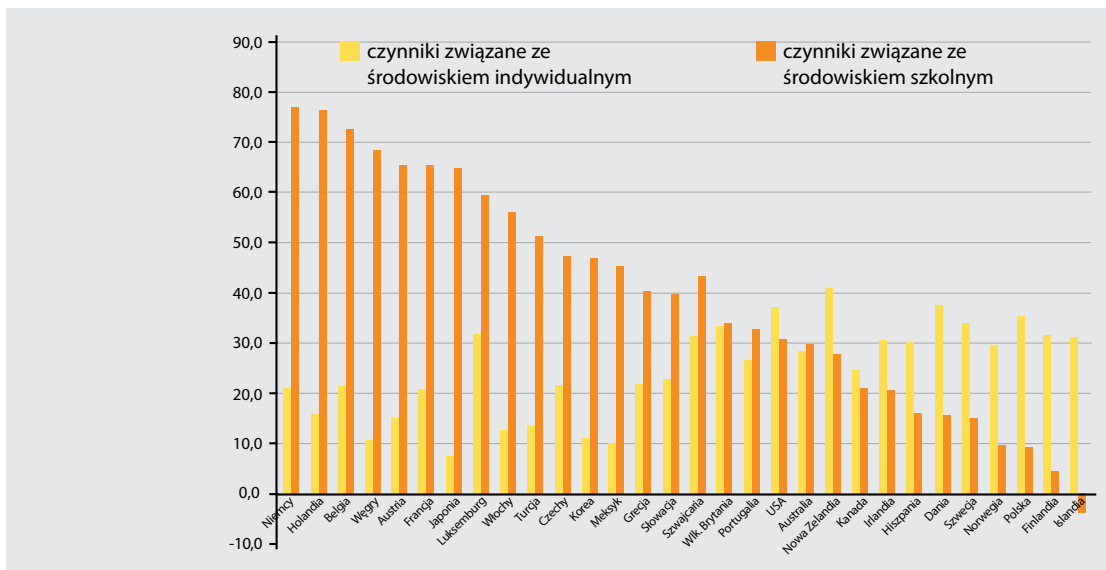
Rysunek A.III.8. Czytanie i interpretacja: wyniki piętnastoletnich dziewcząt i chłopców w kategoriach wykształcenia rodziców



Źródło: Federowicz (2009).

W porównaniach międzynarodowych prowadzonych przez OECD (2010a) Polska należy do krajów, w których związek pomiędzy wykształceniem rodziców a wynikami osiąganymi przez ich dzieci jest relatywnie wysoki, podczas gdy wpływ środowiska szkolnego – niski. Oznacza to, że w przypadku młodego pokolenia obserwujemy w Polsce stosunkowo wysoki poziom międzypokoleniowego dziedziczenia kapitału ludzkiego.

Rysunek A.III.9. Wpływ środowiska rodzinnego (indywidualnego) oraz szkolnego na wyniki testu PISA



Źródło: OECD (2010a).

Oceny wpływu rodziny na wyniki osiągnięte przez uczniów z wykorzystaniem danych dokonał również Herbst (2002). W zaproponowanym modelu wyjaśnia zmienność wyników sprawdzianów 6-klasisty na poziomie gminy przez szereg zmiennych obejmujących m.in. oszacowaną przeciętną liczbę lat nauki dorosłych mieszkańców gminy. Zmienna ta jest traktowana jako *czynnik rodzinny*. Uzyskane wyniki wskazują, że czynnik rodziny (transfer międzypokoleniowy) jest najważniejszą determinantą jakości kształcenia na poziomie gminy. Zwiększenie o 1 rok przeciętnej liczby lat nauki oznacza, *ceteris paribus*, przeciętną poprawę wyników o około 3%. Chociaż analizy te prowadzone są w oparciu o statystyki gminne, a nie dane indywidualne, wnioski potwierdzają wyniki analiz z badania PISA.

Wyniki badania *Kształcenie dorosłych* z 2011 roku (GUS, 2013a) wskazują, że wykształcenie rodziców jest jedną ze zmiennych społecznych najsilniej determinujących poziom wykształcenia osiąganego przez dzieci. Wśród osób w wieku 18-69 lat, które uczestniczyły w edukacji formalnej (w ciągu ostatnich 12 miesięcy przed przeprowadzeniem badania) oraz posiadały wykształcenie wyższe (w momencie badania) największą liczbą było osób, których matki miały wykształcenie policealne lub średnie (45,3%). Osoby z tej grupy miały najczęściej ojca legitymującego się wykształceniem zasadniczym zawodowym (36,7%). Również osoby, które miały wykształcenie średnie ogólnokształcące i kontynuowały naukę w systemie formalnym miały zazwyczaj matki z wykształceniem policealnym lub średnim (38,7%) oraz ojców z wykształceniem zasadniczym zawodowym (38,2%). Wy-

kształcenie rodziców było wyznaczane dla okresu, kiedy respondent był nastolatkiem. Wyniki te potwierdzają, iż dzieci uzyskują zazwyczaj wyższy poziom wykształcenia niż ich rodzice, jednak jednocześnie struktura wykształcenia w kolejnych pokoleniach jest replikowana, co oznacza brak mobilności edukacyjnej pomiędzy pokoleniami.

Aspiracje edukacyjne rodziców wobec dzieci były natomiast analizowane w ramach badania GUS *Wybory ścieżki kształcenia a sytuacja zawodowa Polaków* przeprowadzonego w III kwartale 2011 roku w zakresie badania modułowego w ramach gospodarstw domowych uczestniczących w badaniu budżetów gospodarstw domowych (GUS, 2013b). Jego wyniki wskazują, że wykształcenia co najmniej na poziomie średnim oczekiwali od swoich dzieci 92,1% rodziców (w miastach 94,5%, a na wsiach 89,2%). Prawie 3/4 – 74,5% przebadanych osób pragnęłoby, aby ich dzieci osiągnęły wykształcenie wyższe. Jest to trwale zjawisko – w 2004 roku wyniki podobnego badania wskazywały, że odsetek ten był nawet nieco wyższy i wynosił 76,9%. W badaniu minimalna była grupa osób oczekujących od swoich dzieci niższego wykształcenia niż średnie.

Wyniki cytowanego badania pozwalają również na ocenę związków pomiędzy wykształceniem rodziców i ich dzieci. Wynika z niego, że związek ten jest bardzo silny (GUS, 2013b). Osiągnięte przez ojca wykształcenie najczęściej skutkowało osiągnięciem podobnego lub o stopień wyższego poziomu wykształcenia przez dzieci. Miało to miejsce w szczególności na wsiach. Wśród osób, których ojciec nie posiadał żadnego wykształcenia, ponad 40% osiągnęło wykształcenie najwyższe zasadnicze zawodowe (w miastach 5,9%, na wsiach 44,5%), a prawie 1/3 średnie lub policealne, z czego dwa razy więcej osób w miastach – 36,5%. Warto zwrócić uwagę na dość wysoki odsetek osób – 30%, których ojciec nie posiadał wykształcenia, a które ukończyły studia wyższe (w miastach – 46,2%, na wsiach – 5,5%). W gronie badanych, których ojciec ukończył najwyższe gimnazjum, największy udział mieli absolwenci szkół zasadniczych zawodowych – 36,7%, średnich lub policealnych – ponad 31%. Wśród osób, których ojciec ukończył szkołę zasadniczą zawodową – prawie połowa zdobyła wykształcenie średnie lub policealne, a 1/5 – wyższe. Dzieci ojców z wykształceniem średnim i policealnym oraz wyższym w większości powielają ścieżkę edukacyjną tego rodzica. W przypadku średniego wykształcenia ojca: 37% powtórzyło jego drogę, a ponad połowa ukończyła studia wyższe; natomiast w przypadku wyższego: 1/5 osób ukończyła szkołę średnią, a prawie 80% osób studia wyższe. W przypadku wykształcenia matek obserwowane są podobne tendencje. Oznacza to, że najczęściej dzieci powielają drogę edukacyjną rodziców lub kończą edukację o jeden szczebel wyżej od nich.

Wyniki *Międzynarodowego Badania Kompetencji Osób Dorosłych (PIAAC)* również potwierdzają występowanie w Polsce, podobnie jak w innych krajach OECD, znaczącego powiązania osiągniętych umiejętności i wykształcenia rodziców. Polacy, którzy mają najlepiej wykształconych rodziców, nie odstają od średniej OECD (wyznaczonej dla osób, których rodzice mają wykształcenie wyższe), a wśród najmłodszych prezentują nawet wyższy poziom rozumienia tekstu. Wyniki Polaków, których rodzice mieli niższe poziomy wykształcenia są wyraźnie gorsze, także w odniesieniu do przeciętnych wyników osób o analogicznym pochodzeniu w krajach OECD (Rynko, 2013). Na podstawie danych PIAAC trudno jest wyjaśnić niskie poziomy umiejętności osób w starszym wieku ze względu na zbyt małą liczebność próby.

Na tle międzynarodowym Polska cechuje się średnim poziomem korelacji wykształcenia rodziców i dzieci. Hertz i in. (2007) opracowali ranking krajów pod względem korelacji poziomów wykształcenia (od najwyższej do najniższej korelacji) na podstawie analiz z lat 1994-2004 osób w wieku 20-69 lat dla 42 krajów. Polskę, ze współczynnikiem korelacji wynoszącym 0,43, plasują na 19 miejscu.

W przeprowadzonym przeglądzie literatury nie zidentyfikowano prowadzonych dla Polski analiz dotyczących mobilności wewnątrzpokoleniowej.

A.III.6.4. Metody badawcze

Zastosowane metody badawcze objęły większość metod stosowanych w podobnego typu wymienionych i opisanych powyżej badaniach. Poniżej przedstawiona jest krótka charakterystyka tych metod w odniesieniu do badania UDE.

Statystyki opisowe

Zastosowanie statystyk opisowych miało na celu opis badanych subpopulacji pod względem ich głównych cech oraz obserwowanych rozkładów. Statystyki przedstawione zostały zarówno w postaci opisu tabelarycznego, jak i graficznego (rysunki). Zastosowano również podstawowe miary rozkładu analizowanych cech.

Do podstawowych opisywanych cech należą: rozkład wykształcenia populacji oraz liczby lat w edukacji ze względu na grupy wieku, płeć, miejsce zamieszkania, rozkład zarobków, a także zawodów populacji ze względu na grupy wieku, płeć, miejsce zamieszkania.

Ze szczególną uwagą opisano rozkłady poziomów wykształcenia osób w podziale na grupy odpowiadające poziomowi wykształcenia rodziców, co pozwoliło na wstępną ocenę powiązania wykształcenia dzieci i rodziców.

Analiza korelacji

Kolejną zastosowaną metodą analiz jest analiza korelacji, której celem było określenie związków pomiędzy badanymi cechami w poszczególnych pokoleniach, w szczególności korelacja wykształcenia oraz liczby lat w edukacji pokoleń rodziców z wybranymi cechami pokoleń dzieci, w tym przede wszystkim: wykształceniem, osiąganymi dochodami czy aktywnością edukacyjną.

Prawdopodobieństwa przejścia

Kolejną zastosowaną metodą statystyczną jest oszacowanie prawdopodobieństw przejścia mierzonych jako warunkowe prawdopodobieństwa osiągnięcia określonego poziomu wykształcenia przez dzieci w zależności od poziomu wykształcenia rodziców.

Macierz prawdopodobieństw przejścia będzie miała postać:

Tabela A.III.5. Macierz prawdopodobieństw przejścia

Wykształcenie rodziców (pokolenie 1) \ Wykształcenie dzieci (pokolenie 2)		Gimnazjalne i niższe	Zasadnicze zawodowe	Średnie	Wyższe	Razem P_j l. obs i
Gimnazjalne i niższe	P_{1j} P_{1j}/P_j					
Zasadnicze zawodowe	P_{2j} P_{2j}/P_j					
Średnie	P_{3j} P_{3j}/P_j					
Wyższe	P_{4j} P_{4j}/P_j					
Razem	P_j l. obs j					

Prawdopodobieństwa przejścia pozwolą na oszacowanie zarówno stabilności osiągniętych poziomów wykształcenia przez kolejne pokolenia, jak również potencjalnego nieliniowego rozkładu tych prawdopodobieństw w zależności od poziomu wykształcenia dla pokolenia rodziców. Przykładowo, analizy prowadzone przez Lindahl i in. (2012) wskazują na taką nieliniowość w Szwecji dla pokolenia rodziców posiadających wyższe poziomy wykształcenia.

W przypadku osiągniętych zarobków analiza prawdopodobieństw dotyczyć będzie kwintylowego rozkładu wynagrodzeń (po II rundzie badania).

Do zmiennych instrumentalnych stosowanych w modelach dotyczących międzypokoleniowej transmisji kapitału ludzkiego zalicza się między innymi: objęcie pokolenia rodziców zmianą w obszarze polityki edukacyjnej (np. zmiana zakresu obowiązku szkolnego, reforma podstawy programowej itp.) lub też określone uwarunkowania domowe, np. wykształcenie pokolenia dziadków lub pradiadków. Zmienną instrumentalną mogą być również uwarunkowania lokalne.

Model ścieżki

W zadaniu model ścieżki będzie wykorzystany do analizy uwarunkowań awansu edukacyjnego pomiędzy pokoleniami w celu identyfikacji znaczenia tych czynników, które wpływają na osiągnięcie danego poziomu wykształcenia przez osoby dorosłe.

A.III.6.5. Spodziewane wyniki i korzyści analityczne

Przeprowadzenie analiz w oparciu o przyjęte założenia badawcze oraz wybrane metody, pozwoli na pełniejszą niż dzisiaj stosowaną ocenę zjawiska międzypokoleniowej transmisji kapitału ludzkiego w Polsce. Dotychczasowe analizy obejmowały ograniczoną populację (badanie PISA – młodzież 15-letnia, badanie SHARE – osoby w wieku 50+). Ze względu na to, że w badaniu UDE mamy do czynienia z populacją osób w wieku 15-64 lata (w kwestionariuszu indywidualnym), będzie można w pełniejszy sposób dokonać oceny skali transmisji międzypokoleniowej, a także wpływu reform

i zmian zachodzących w systemie edukacji na kształtowanie kapitału ludzkiego w przypadku poszczególnych pokoleń.

Istotną wartością dodaną prowadzonych analiz będzie również przeprowadzenie analiz dotyczących wewnątrzpokoleniowych uwarunkowań kształtowania kapitału ludzkiego, w tym mobilności społecznej i zawodowej osób, co w znaczący sposób zwiększy stan wiedzy w tym obszarze.

Przeprowadzone analizy pozwoliły na weryfikację przyjętych hipotez badawczych odnoszących się do dziedziczenia wzorców kariery edukacyjnej oraz zawodów, awansu edukacyjnego kolejnych pokoleń, a także oceny zależności wewnątrzpokoleniowych, dotyczących m.in. współzależności kształtowania kapitału ludzkiego pomiędzy współmałżonkami oraz w grupach rówieśniczych.

A.IV.Blok IV

A.IV.1. Wprowadzenie

Celem aneksu jest przedstawienie informacji o proponowanej metodzie badawczej, która została zastosowana w przygotowaniu raportu z realizacji dwóch zadań (9 i 10) w projekcie *Uwarunkowania decyzji edukacyjnych*. Zadania badawcze są zatytułowane następująco: *Wpływ krajowej i lokalnej polityki edukacyjnej na decyzje edukacyjne* oraz *Symulacyjna analiza skutków zmian polityki finansowania w systemie edukacyjnym dla decyzji edukacyjnych gospodarstw domowych*.

W niniejszej części przedstawiono podstawowe założenia modelu symulacyjnego, który był skonstruowany w oparciu o wyniki badania ankietowego uwarunkowań decyzji edukacyjnych (UDE). Wyniki posłużyły do opracowania dyskryminacji zbioru gospodarstw domowych ze względu na cechy społeczno-demograficzne, sposób podejmowania decyzji edukacyjnych i postać tych decyzji oraz do skalibrowania wartości parametrów zdefiniowanych w modelu symulacyjnym. Opis podzielono na trzy części. Najpierw przedstawiono charakterystyczne cechy warunkujące złożoność systemu edukacyjnego, które stanowią o podstawach motywacji wyboru narzędzia badawczego. Następnie przedstawiono hipotezy badawcze dopuszczalne i możliwe do weryfikacji z zastosowaniem opisywanego narzędzia. Ostatecznie zaproponowano i szerzej opisano symulację wieloagentową - jako narzędzie pozwalające uwzględnić złożoność systemu edukacyjnego oraz dane empiryczne z badania UDE.

A.IV.1.1. Rynek edukacyjny w kontekście narzędzia badawczego

Analiza skutków zmian polityki finansowania rynku edukacyjnego wymaga stworzenia narzędzia badawczego, które uwzględni charakterystyczne cechy systemu edukacyjnego, a następnie umożliwi analizę wpływu polityki edukacyjnej na indywidualne decyzje heterogenicznych gospodarstw domowych. W analizie polityki edukacyjnej należy uwzględnić szereg cech specyficznych dla rynku edukacyjnego:

- heterogeniczność odbiorców usług edukacyjnych (zróżnicowanie przestrzenne, społeczne, demograficzne, dochodowe, ze względu na zdolności);
- heterogeniczność i hierarchię instytucji stanowiących prawo i zarządzających edukacją (administracja centralna, organy samorządu terytorialnego, fundacje, osoby fizyczne);
- heterogeniczność instytucji edukacyjnych oraz oferowanych przez nie usług (według typów instytucji, rozmieszczenia przestrzennego, jakości oraz rodzaju wykształcenia);
- uwarunkowania decyzji o kształceniu bezpłatnym albo płatnym, w miejscu zamieszkania albo poza miejscem zamieszkania, stacjonarnym albo niestacjonarnym i wpływie różnych czynników na formowanie kapitału ludzkiego;
- zależności międzypokoleniowe (dziedziczenie wykształcenia, finansowanie edukacji dzieci przez rodziców);
- współzależność decyzji edukacyjnych (decyzje podejmowane przez jedne podmioty wpływają na przyszłą sytuację decyzyjną innych podmiotów) oraz sprzężenie zwrotne polityki edukacyjnej (wpływ polityki edukacyjnej na gospodarkę, której rozwój wpływa na sytuację decyzyjną podmiotów).

Liczne opracowania naukowe wskazują, że zróżnicowanie cech podmiotów na rynku ma istotny wpływ na opłacalność ich decyzji. W klasycznym równaniu Mincera (por. Mincer, 1970, 1974) założono, że na wynagrodzenie wpływa liczba lat wykształcenia oraz doświadczenie zawodowe. W analizie polityki edukacyjnej należy jednak uwzględnić indywidualne zróżnicowanie podmiotów, jak wykształcenie pobierających oraz jakość oferowanej usługi. Wyniki analiz przeprowadzonych przez Sewell i Shah (1967) wskazują, że status społeczny ma istotny wpływ na osiągnięcia edukacyjne podmiotów. Jednocześnie gospodarstwa domowe charakteryzują się różną wrażliwością na prowadzoną politykę edukacyjną. W szczególności badania empiryczne przeprowadzone przez Kane (1994) oraz Cameron i Heckman (1998) wskazują, że programy finansowania czesnego mają największe znaczenie dla decyzji edukacyjnych osób o niskich i średnich dochodach. Nie są skuteczne dla osób o bardzo niskich dochodach.

Kolejną kwestią, którą należy uwzględnić w analizie polityki edukacyjnej, jest zróżnicowanie jakości kształcenia i jego wpływ na efekt. Speakman i Welch (2006) oraz Hanushek (2006) dokonują przeglądu badań empirycznych i zauważają, że jakość nauczania ma bardzo istotny wpływ na wielkość wynagrodzeń absolwentów. W szczególności badania różnic w poziomie wynagrodzeń pomiędzy białymi a czarnymi mieszkańcami Stanów Zjednoczonych wskazują, że różnice te mogą być aż w połowie wyjaśnione jakością szkoły, do której te osoby uczęszczały. Jakość instytucji dydaktycznych jest również ważnym czynnikiem wpływającym na wynagrodzenie absolwentów. Tym samym jakość kształcenia determinuje stopę zwrotu z edukacji, np. Rizzuto i Wachtel (1980) oraz Nechyba (1990).

Szczególną cechą polskiego rynku edukacyjnego na poziomie wyższym jest znaczny udział kształcenia w trybie niestacjonarnym. Jak podaje GUS (2012), w Polsce studenci

niestacjonarni stanowią aż 45% wszystkich studentów. Zgodnie z raportem OECD (2013) Polska ma największy odsetek studentów zaocznych wśród krajów OECD. Osoby te, z reguły łączą pracę ze studiowaniem. GUS (2011) podaje, że 69% studentów zaocznych łączy pracę ze studiami, podczas gdy w przypadku studentów dziennych odsetek ten wynosi 5%. Stąd studia zaoczne z reguły prowadzą do mniejszego wzrostu wiedzy niż studia dzienne, ale jednocześnie ich absolwenci wkraczają na rynek pracy posiadając doświadczenie zawodowe, którego nie mają absolwenci studiów dziennych. Analiza polityki edukacyjnej powinna uwzględniać te specyficzne uwarunkowania polskiego rynku.

Kolejną cechą rynku edukacyjnego jest rola rodziców w decyzjach edukacyjnych dzieci. Po pierwsze, zarobki rodziców, które z kolei są skorelowane z ich wykształceniem, determinują ograniczenie budżetowe gospodarstwa domowego i przez to wpływają na dostępny zbiór decyzji o wyborze studiów. Po drugie, w gospodarstwie domowym decyzje mogą zapadać na podstawie procesów decyzyjnych przebiegających w różny sposób (por. Browning i in., 2011). W rodzinie występuje też przekazywanie wzorców edukacyjnych. Badania empiryczne wskazują na występowanie pozytywnej korelacji pomiędzy wykształceniem dzieci i wykształceniem rodziców, szczególnie wykształceniem matki – np. Behrman i Rosenzweig (2002). Wyniki badań dla Polski opisane przez Sztanderską (2004) potwierdzają pozytywną korelację pomiędzy decyzją o studiowaniu dzieci a wykształceniem rodziców.

Ostatnią cechą wyróżniającą rynek edukacyjny jest współzależność decyzji na tym rynku. Współzależność jest związana z konkurencją o zasoby edukacyjne, wpływem decyzji wyborów edukacyjnych na rozwój gospodarczy oraz istniejącymi w modelowanych procesach licznymi sprzężeniami zwrotnymi.

Decyzje jednych podmiotów wpływają na sytuację decyzyjną pozostałych. Po pierwsze, decyzja o studiowaniu podjęta przez jedną osobę wpływa na możliwości studiowania innych (przy założeniu ograniczonej liczby miejsc). Po drugie, szereg publikacji wskazuje na ścisły związek pomiędzy wykształceniem a dobrobytem i wzrostem gospodarczym zarówno na świecie (m.in. meta-analiza badań dokonana przez Lange i Topel 2006 oraz w Polsce – Sztanderska i Minkiewicz 2007). Stąd w analizie polityki edukacyjnej należy rozpatrywać sprzężenie zwrotne w systemie edukacyjnym. Prowadzona polityka edukacyjna będzie miała istotny wpływ na środki budżetowe w przyszłości również te, które będzie można przeznaczyć na finansowanie edukacji.

Przeprowadzone rozważania wskazują na to, że model rynku edukacyjnego powinien posiadać szereg cech, aby umożliwić rzeczywistą ocenę polityki edukacyjnej. W szczególności powinien uwzględniać istnienie w systemie heterogenicznych podmiotów oraz istniejącą współzależność decyzji na tym rynku. Szczegółowe wymagania odnośnie narzędzia dla modelowania rynku edukacyjnego zostały przedstawione w następującej sekcji.

A.IV.1.2. **Pytania i hipotezy badawcze skutecznie weryfikowane metodami symulacji komputerowej**

Symulacyjny model rynku edukacyjnego zostanie skonstruowany w sposób pozwalający odpowiedzieć m.in. na następujące pytania badawcze:

- Kiedy zmiany w strukturze finansowania prowadzą do istotnych zmian wyborów edukacyjnych dokonywanych przez gospodarstwa domowe, t.j. jakie są uwarunkowania skuteczności regulacji systemu edukacyjnego?
- Jakie są różnice w następstwach regulacji rynku edukacyjnego w horyzoncie czasowym jednego pokolenia, a jakie w horyzoncie czasowym wielu pokoleń (decyzje edukacyjne młodych ludzi wpływają na kształt gospodarstw domowych, które zostaną przez nich założone, zatem skutkiem dzisiejszej polityki edukacyjnej jest również sytuacja decyzyjna gospodarstw domowych w przyszłych pokoleniach)?
- Kiedy zmiana regulacji systemu edukacyjnego może się okazać przeciwnie skuteczna (chodzi np. o sytuację, gdy skutkiem regulacji w systemie edukacyjnym jest wzrost bezrobocia)?
- Jakie są różnice w skutkach wprowadzenia doraźnych instrumentów finansowych stymulujących zachowania podmiotów w systemie edukacyjnym? Jaką korzyść odnosi indywidualnie, a jaką społeczeństwo, z ustanowienia stypendiów socjalnych w porównaniu z przeznaczeniem tych samych środków na stypendia naukowe?
- Jakie są skutki wprowadzania innego podziału środków na kształcenie pomiędzy szkoły publiczne i niepubliczne oraz innego finansowania czesnego stacjonarnych i niestacjonarnych uczestników procesu kształcenia?
- Jakie mogą być skutki uzależnienia finansowania czesnego od rodzaju szkoły lub kierunku studiów (techniczne, ekonomiczne, pedagogiczne) zamiast np. od formy własności szkoły (publiczna, niepubliczna)?
- Jak zróżnicowanie indywidualnych cech podmiotów (gospodarstw domowych – wariacja dochodów, uczelni – zróżnicowanie jakości kształcenia) w systemie edukacyjnym determinuje skuteczność zmian polityki edukacyjnej?

A.IV.1.3. **Założenia symulacji rynków edukacyjnych**

Celem niniejszej sekcji jest podsumowanie przytoczonych rozważań na temat cech rynku edukacyjnego oraz określenie cech, jakie musi spełniać modelowe narzędzie analityczne do badania polityki edukacyjnej.

Rozważania wskazują na konieczność uwzględnienia w modelu:

- indywidualnych, heterogenicznych cech konsumentów usług edukacyjnych na wszystkich etapach procesu kształcenia;

- heterogeniczności i hierarchii w procesie podejmowania decyzji o kształcie systemu edukacji;
- różnych form kształcenia przez instytucje edukacyjne o różnym poziomie jakości;
- sytuacji decyzyjnej indywidualnie na poziomie każdego podmiotu, którego zachowanie podlega modelowaniu;
- zależności społecznych, demograficznych, rodzinnych i innych charakteryzujących modelowane systemy zarówno na poziomie mikro, jak i na poziomie makro.

Ponadto jest konieczna możliwość takiego skonstruowania modelu symulacyjnego, aby w analizie polityki edukacyjnej było możliwe wykorzystanie danych z badania ankietowego przeprowadzanego w projekcie *Uwarunkowania decyzji edukacyjnych*. Nie jest wykluczone korzystanie z wyników innych badań związanych z edukacją i opublikowanych przez różne instytucje, np. *Wybory ścieżki kształcenia a sytuacja zawodowa Polaków* (2013), *Ścieżki rozwoju edukacyjnego młodzieży – szkoły pogimnazjalne* (2013) czy standardowych opracowań Głównego Urzędu Statystycznego z serii *Oświata i wychowanie* oraz *Szkoły wyższe i ich finanse*.

Podjęciem, które jest na tyle elastyczne, że pozwala na uwzględnienie powyższych założeń jest symulacja wieloagentowa – metoda polegająca na modelowaniu systemu na poziomie mikro jako zbioru niezależnych elementów – agentów. W modelach makroekonomicznych budowanych jako systemy współdziałających indywidualności skali mikro wyróżnia się dwa podejścia. W jednym z nich, chronologicznie wcześniejszym w historii makromodelowania, globalne makrostruktury wynikowe powstają jako proste sumowanie zachowań podmiotów opisanych w skali mikro. W podejściu późniejszym makrostruktury powstają jako złożenie działań podmiotów ze skali mikro. Ich cechy i zachowanie nie mogą być rozpatrywane jako prosta suma elementów składowych. Całość posiada cechy nieistniejące w żadnym elemencie składowym. Te cechy są skutkiem współdziałania, współzależności i heterogeniczności poszczególnych elementów składowych systemu. To drugie podejście do makromodelowania jest typowe dla modeli wieloagentowych. Modele te, pozwalają więc uwzględniać zróżnicowanie indywidualności, interakcje między nimi i obserwować pojawiające się efekty sieciowe. W modelu wieloagentowym pojedynczy agenci są autonomiczni, ale powiązani z otoczeniem. Twórca modelu określa ich charakterystyki i wprowadza ich do wirtualnego świata, który działa według określonych wcześniej reguł. W przeprowadzanej symulacji działania makrosystemu ów wirtualny świat może żyć swoim własnym życiem, bez żadnej ingerencji ze strony twórcy modelu. Wynika stąd, że w prawidłowo skonstruowanym modelu wieloagentowym ingerencja w działający model nie jest konieczna, ale jest dopuszczalna. W najbardziej ortodoksyjnym podejściu do modelowania wieloagentowego, twórca modelu nie zakłada wprost żadnych zależności w skali makroekonomicznej. Mają być one wynikiem działań podejmowanych przez poszczególnych agentów i interakcji między tymi agentami.

Praca z modelem wieloagentowym polega na prowadzeniu eksperymentów i badaniu dynamiki systemu. Analiza modelu wieloagentowego może być prowadzona zarówno w ujęciu mikro, mezo jak i makro – jest możliwa dowolna agregacja wyników w zależności od hipotez badawczych. Farmer i Foley (2009) podkreślają, że symulacja wieloagen-

towa jest narzędziem umożliwiającym analizę złożonych systemów ekonomicznych, a w szczególności pozwala na analizę systemów przy zmianie zasad ich funkcjonowania oraz dynamiki powiązań w tych systemach. Trzeba zaznaczyć, że podejście mikrosymulacyjne było stosowane w badaniach edukacyjnych dużo wcześniej niż pojawienie się modeli wieloagentowych (por. O'Donoghue, 1999).

Reasumując, model wieloagentowy pozwoli na jego parametryzację w skali mikro w zakresie indywidualnych decyzji edukacyjnych oraz analizę skutków regulacji tego systemu w skali makro. Tym samym model umożliwi poszukiwanie głównych czynników kształtujących proces rozwoju kompetencji w skali makro. Analiza symulacyjna umożliwia elastyczny i sekwencyjny przegląd różnych zestawów takich czynników, co pozwoli na modelowanie m.in. stanu i użytkowania zasobów służących rozwojowi kompetencji członków polskiego społeczeństwa we wszystkich grupach wieku zarówno w wymiarze indywidualnym, jak zagregowanym do skali całej gospodarki. Podstawą prac badawczych prowadzonych w omawianych zadaniach są akty legislacyjne stanowiące ustrój edukacyjny na szczeblu centralnym, wybrane regulacje jednostek samorządu terytorialnego stanowiące zasady na określonym terytorium w ramach istniejącego centralnego systemu edukacyjnego oraz wybrane regulacje jednostek edukacyjnych dotyczące opłat, a także wewnętrzne analizy systemów regulacyjnych różnej skali oraz wyniki innych zespołów badawczych w projekcie.

W szczególności model umożliwi analizę różnych sposobów prowadzenia polityki edukacyjnej:

- publiczno-prywatne finansowanie edukacji,
- systemy stypendiów socjalno-naukowych,
- badanie wpływu regulacji rynku edukacyjnego na jego dynamikę w ujęciu wielopokoleniowym.

W otoczeniu międzynarodowym można zaobserwować postępujące odchodzenie od pełnego finansowania czesnego przez państwo w krajach OECD (por. Marcucci, Johnstone, 2007). Temat zmian systemu finansowania edukacji pojawia się również w polskim systemie edukacyjnym – proponowana metodologia pozwoli na wielokryterialną analizę skutków takich regulacji dla różnych podmiotów. Modele wieloagentowe, a właściwie jeden z elementów posługiwania się nimi, czyli analiza wrażliwości wyników na zmiany wartości parametrów modelu, pomogą w zrozumieniu skutków przyjęcia różnych unormowań prawnych.

A.IV.2. Symulacja wieloagentowa

Celem niniejszego rozdziału jest przedstawienie założeń oraz scenariusza zastosowania symulacji wieloagentowej do analizy polityki w obrębie rynku edukacyjnego.

W literaturze z zakresu analizy symulacyjnej autorzy stosują podejście systemowe, m.in. Law (2007), Gilbert i Troitzsch (2005) oraz Miller i Page (2009). Dla potrzeb analizy polityki edukacyjnej, ze względu na objaśnioną wcześniej złożoność zagadnienia, również zostanie zastosowane podejście systemowe. W niniejszym raporcie system jest rozumiany jako zestaw elementów wzajemnie powiązanych w sposób bezpośredni lub pośredni (por. Ackoff, 1971). Stąd w konstrukcji symulacyjnego modelu rynku edukacyjnego jest stosowane podejście systemowe, gdzie na wymienionym rynku zostały wyodrębnione

niezależne, wzajemnie powiązane obiekty – indywidua gospodarujące: gospodarstwa domowe, jednostki świadczące usługi edukacyjne, firmy oraz hierarchiczny regulator.

Niniejszy rozdział składa się z pięciu części. Najpierw, w podrozdziale A.IV.2.1, przedstawiono podejście symulacyjne i jego zastosowanie do modelowania systemów i eksperymentowania z modelami systemów złożonych. Opis ten obejmuje proces modelowania symulacyjnego, a ponadto szczególnie w nim uwzględniono symulację systemów zdarzeń dyskretnych. Następnie, w podrozdziale A.IV.2.2, opis podejścia symulacyjnego i symulacji dyskretnych został uszczegółowiony o symulację wieloagentową. Podano m.in. definicję agenta oraz symulacji wieloagentowej. W podrozdziale A.IV.2.3 naszkicowano wstępne założenia konceptualnego modelu systemu edukacyjnego. Część A.IV.2.4 poświęcono krótkiemu omówieniu problemów kalibracji złożonego modelu symulacyjnego, a w części ostatniej zarysowano sposób prowadzenia eksperymentów symulacyjnych.

A.IV.2.1. Podejście symulacyjne w modelowaniu systemów

Celem podrozdziału jest przedstawienie podejścia symulacyjnego do analizy systemów. W kolejnym rozdziale wywód ten zostanie rozszerzony o podejście wieloagentowe i symulację systemów wieloagentowych.

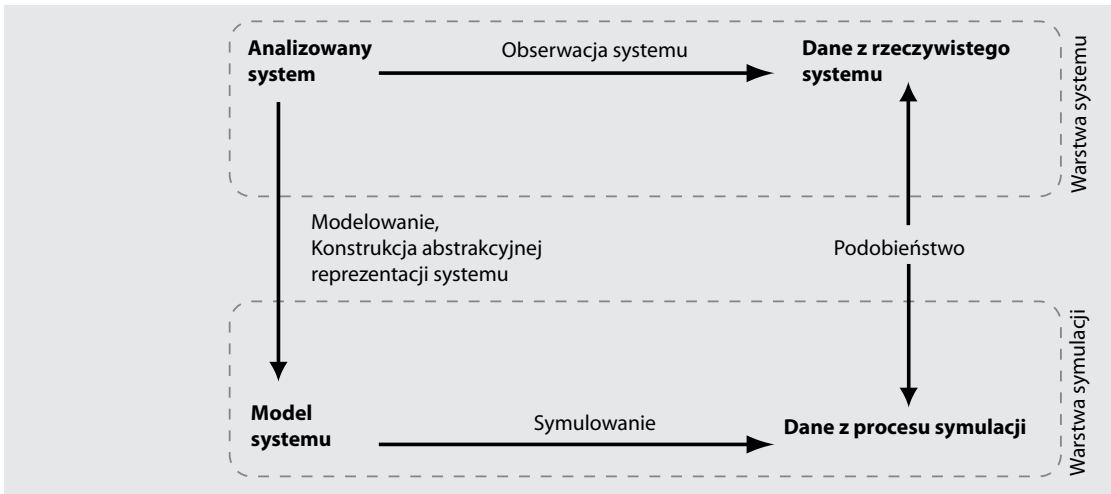
Przedmiotem analizy symulacyjnej jest system rozumiany zgodnie z koncepcją analizy systemowej – jako zestaw powiązanych i wyodrębnionych z otoczenia elementów. Tym samym symulacja jest zestawem technik do analizy dynamicznego zachowania modelu systemu (por. Law, 2007). Tym samym przyjmujemy, że celem modelowania symulacyjnego jest eksperymentowanie z utworzonym modelem systemu. Model symulacyjny jest takim sposobem opisu badanej rzeczywistości, który pozwala badać numerycznie analizowane wielkości wyjściowe zależnie od wielkości wejściowych.

W analizie rynku edukacyjnego zostanie zastosowane takie podejście, zgodnie z którym badanie symulacyjne nie daje bezpośredniego rozwiązania problemu, lecz pozwala porównywać i oceniać różne scenariusze (por. Krajewski, Ritzman, 2002).

Law (2007) definiuje symulację Monte Carlo jako postępowanie zmierzające do zastosowania liczb losowych do rozwiązywania problemów stochastycznych lub deterministycznych. Metody Monte Carlo, obok szeregu zastosowań m.in. w statystyce, inżynierii, finansach, są również stosowane w symulacji systemów zdarzeń dyskretnych.

Gilbert i Troitzsch (2005) zauważają, że w trakcie analizy symulacyjnej dane z modelu symulacyjnego powinny być porównywane z danymi z systemu rzeczywistego, por. rysunek A.IV.1. Stąd też dane z badania empirycznego mogą zarówno posłużyć do kalibracji modelu wieloagentowego, jak i do późniejszego sprawdzenia poprawności działania tego modelu.

Rysunek A.IV.1. Dane z modelu symulacyjnego służą do konstrukcji modelu systemu. Dane te są również porównywane z danymi z rzeczywistego systemu



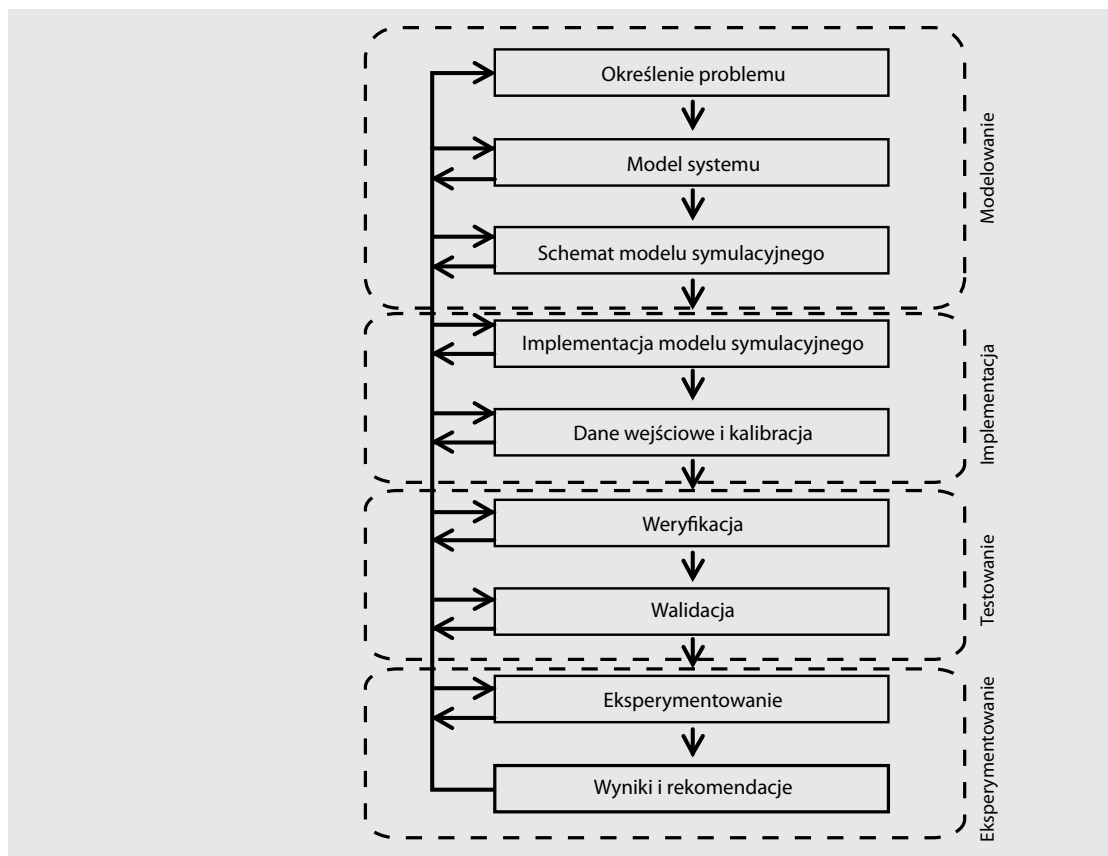
Źródło: Szufel (2012) opracowanie własne na podst. Gilbert i Troitzsch (2005).

Punktem wyjścia w realizowanym zadaniu badawczym jest analizowany system. Na tym systemie (oryginalie) przeprowadzane są obserwacje zmieniających się jego stanów i stawiane są hipotezy dotyczące zachodzących procesów powodujących zmiany stanów oryginału. Na podstawie posiadanej wiedzy i zebranych o oryginalie informacji budowany jest model systemu. W testowych symulacjach wykonywanych na modelu generowane są stany modelu, które są porównywane ze stanami oryginału. Po osiągnięciu akceptowanego poziomu podobieństwa stanów systemu i stanów modelu ten ostatni jest gotowy do przeprowadzania analiz symulacyjnych. Takie analizy symulacyjne wybiegają swoimi założeniami poza kiedykolwiek obserwowane stany systemu. Zajmujący się analizami symulacyjnymi ufają, że jeśli model w sytuacjach testowych zachowywał się podobnie jak oryginał, to i dla innych założonych stanów i reguł zmian stanów zachowa się podobnie do oryginału. Z tej ufności wynika możliwość wnioskowania o zmianach systemu (oryginału) na podstawie obserwowania zmian w modelu.

W literaturze analiza modelu symulacyjnego jest usystematyzowana. Opis etapów pracy z modelem symulacyjnym nosi nazwę procesu symulacji (Bennett, 1995), dzieli się on na trzy etapy: konstruowania modelu, testowania oraz przeprowadzania obliczeń symulacyjnych. Etap konstrukcji modelu symulacyjnego składa się z jego sformułowania (w tym określenia problemu i hipotez badawczych) oraz implementacji w postaci programu komputerowego. Etap testowania składa się z dwóch faz: fazy weryfikacji polegającej na sprawdzeniu, czy model jest prawidłowo oprogramowany oraz fazy walidacji polegającej na testowaniu zgodności zachowań modelu z jego założeniami. Po skonstruowaniu i przetestowaniu modelu można przystąpić do eksperymentów z modelem. Podobnie proces symulacji jest definiowany przez Schroedera (1993). Uwzględnia on definicję problemu poprzedzającą tworzenie modelu oraz zaleca tworzenie schematu modelu. Następnie model jest oprogramowany, zbierane są dane służące do jego kalibracji, model jest weryfikowany i są przeprowadzane eksperymenty. Analogiczne podejście do prezentacji procesu symulacji proponują Gilbert i Troitzsch (2005). Zauważają

oni, że metodologia badania symulacyjnego składa się z następujących etapów: (1) określenie zakresu modelowanego systemu, (2) obserwacja modelowanego systemu w celu określenia parametrów początkowych, (3) poczynienie założeń i konstrukcja modelu, (4) weryfikacja modelu (ang. *model verification*) – sprawdzenie, czy model działa zgodnie z założeniami, (5) walidacja modelu (ang. *model validation*) – sprawdzenie, czy model zachowuje się podobnie do rzeczywistego systemu, (6) analiza wrażliwości – sprawdzenie wpływu zmian parametrów początkowych na zachowanie modelu. Powyższe podejście zostanie zastosowane przy konstrukcji narzędzia do symulacji w ramach badania uwarunkowań decyzji edukacyjnych.

Rysunek A.IV.2. Proces analizy symulacyjnej, który zostanie przyjęty w ramach badania UDE

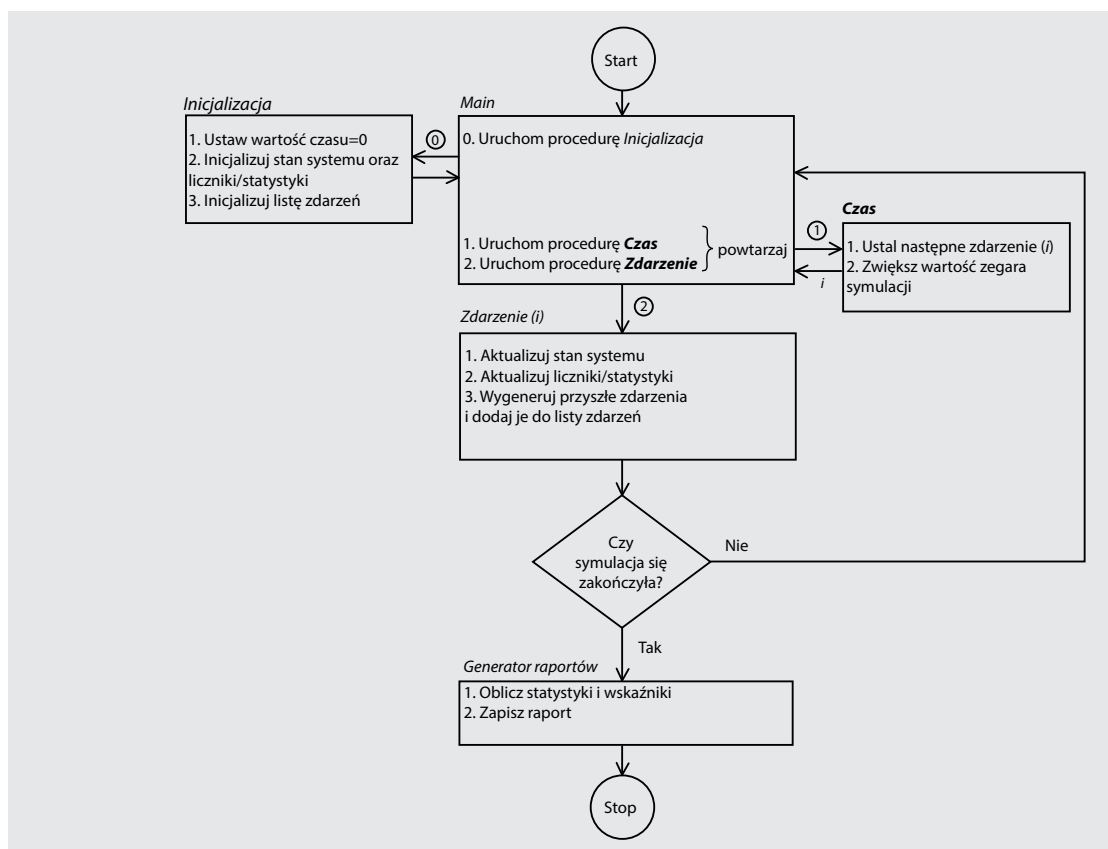


Źródło: Szufel (2012).

W przyjętym procesie budowy modelu symulacyjnego rynku edukacyjnego można wyróżnić cztery etapy: modelowania, implementacji, testowania oraz eksperymentowania (por. rysunek A.IV.2.). Etap modelowania obejmuje określenie problemu, stworzenie modelu systemu oraz stworzenie schematu oprogramowania modelu symulacyjnego. Etap implementacji obejmuje stworzenie kodu źródłowego oraz kalibrację modelu na podstawie dostępnych danych. Etap testowania obejmuje sprawdzenie, czy model działa zgodnie z założeniami poczynionymi wcześniej (weryfikacja) oraz sprawdzenie, czy model systemu zachowuje się podobnie do systemu rzeczywistego (walidacja). W końcu etap eksperymentowania obejmuje zebranie wyników i ich analizę. Rysunek

A.IV.2. uwzględnia proces iteracyjno-przystosowy wytwarzania oprogramowania (por. Larman, Basili, 2003). Wieloagentowa implementacja modelu systemu z reguły jest przeprowadzana małymi krokami, a poprawność działania przed przejściem do kolejnych kroków jest sprawdzana w krokach poprzednich. Zastosowanie procesu iteracyjnego przystosowego powoduje, że każdy z kroków procesu konstrukcji modelu symulacyjnego przedstawionego na rysunku A.IV.2. zakłada możliwość dowolnego cofania się do poprzednich etapów.

Rysunek A.IV.3. Proces symulacji dyskretnej. Kolejne zdarzenia są przetwarzane po ukończeniu zdarzeń poprzedzających. W symulacji wieloagentowej zdarzenia dotyczą działań podejmowanych przez agentów w systemie



Źródło: opracowanie własne na podst. Law (2007).

Jak wcześniej wspomniano, analiza polityki prowadzonej w ramach systemu edukacyjnego zostanie w ramach badania UDE przeprowadzona przez symulację wieloagentową, która zostanie scharakteryzowana w następnym podrozdziale.

A.IV.2.2. Symulacja wieloagentowa

Symulacyjna analiza systemu edukacji i uwarunkowań decyzji edukacyjnych powinna uwzględniać złożoność tego systemu, a w szczególności zróżnicowanie indywidualnych cech podmiotów w tym systemie oraz wpływ tego zróżnicowania na podejmowane przez nie decyzje. Ponadto model symulacyjny powinien uwzględniać relacje międzypokoleniowe na rynku edukacyjnym, takie jak finansowanie edukacji dzieci przez gospodarstwo domowe oraz dziedziczenie wzorców edukacyjnych. W końcu, model powinien również uwzględniać wpływ decyzji jednych podmiotów na sytuację decyzyjną innych - przykładowo decyzja uczelni o oferowaniu nowego kierunku studiów o zadanym programie kształcenia i spodziewanej jakości kształcenia ma wpływ na sytuację decyzyjną i wybory edukacyjne potencjalnych studentów.

Metodą analizy, która pozwala na analizę systemów złożonych (ang. *complex systems*, por. Holland, 1992), jest symulacja wieloagentowa. Macal i North (2009) definiują symulację wieloagentową (ang. *agent-based modelling and simulation* – ABMS), jako „...sposób modelowania złożonych systemów składających się z niezależnych agentów...”. W przypadku modelu systemu edukacyjnego agentami mogą być konsumenci usług edukacyjnych wyodrębnieni z innego rodzaju agentów, jakimi są gospodarstwa domowe. Agentami są również podmioty oferujące podaż usług edukacyjnych, czyli instytucje edukacyjne wszelkich możliwych typów i na wszelkich możliwych poziomach. Agentami są również organy administracji samorządowej i administracji centralnej podejmujące decyzje w zakresie regulowania rynku edukacyjnego. Ogólniej biorąc, agentami w modelu mogą być pojedyncze indywiduala, ale można również definiować jako agentów podmioty kolektywne. Dozwolone są konstrukcje hierarchiczne, tzn. pojedynczy agent należący do pewnej klasy może składać się z wielu agentów należących do innej klasy. W zależności od potrzeb agentem może być również zasób przyrodniczy albo fizyczny. Najogólniej mówiąc, z technicznego punktu widzenia agent jest wiązką danych i reguł opisujących jego zachowania w różnych sytuacjach. Agenci mogą być autonomiczni albo ze sobą powiązani. Decyzja o rodzajach klas agentów, liczebności klas oraz rodzajach powiązań między elementami systemu należy do konstruktora modelu. Z przytoczonych cech wynika, że modelowanie wieloagentowe jest bardzo elastyczną techniką i model tworzony w jego konwencji można dopasować do każdego problemu badawczego.

Skonstruowane narzędzie pozwoli badać skutki zmian w polityce edukacyjnej zarówno w ujęciu jednego, jak i wielu pokoleń. Badanie jednopokoleniowe jest rozumiane jako sprawdzenie tego, jak zmiany zasad polityki wpłyną na osoby, których zmiany te dotyczą. Symulacja wieloagentowa pozwala także uwzględnić wielopokoleniowe skutki regulacji systemu edukacyjnego. W szczególności zostanie uwzględniony fakt, że decyzje edukacyjne młodych ludzi wpływają na kształt gospodarstw domowych, które zostaną przez nich utworzone i tym samym wpływają na sytuację decyzyjną gospodarstw domowych w kolejnych pokoleniach.

Symulacja wieloagentowa jest rodzajem symulacji systemów, którego cechą wyróżniającą jest podział modelu na niezależne elementy – agentów. Z przytoczonych wcześniej przykładów wynika, że pojęcie agenta w literaturze nie jest jednoznaczne. Macal i North (2006) dokonują przeglądu

literatury i przytaczają następującą definicję agenta: agentem jest niezależnie zachowujący się komponent dowolnego rodzaju (por. Bonabeau, 2002). Niezależność polega na tym, że agenta można wyodrębnić z systemu. Zachowanie się jest tutaj rozumiane jako podejmowanie decyzji przez agenta. Macal i North (2006) zauważają, że druga grupa podejść wymaga, aby zachowania komponentu były adaptacyjne, czyli żeby mógł być uznany za agenta, to powinniśmy uczyć się na podstawie stanów środowiska, a następnie odpowiednio dopasowywać swoje zachowania. Gilbert i Troitzsch (2005) definiują agenta ogólnie jako program komputerowy lub część programu komputerowego, która może być uznana za działającą samodzielnie oraz reprezentuje indywiduum, organizację, państwo lub innego aktora społecznego.

Macal i North (2006) zauważają, że agentowi należy przypisać następujący zestaw cech: (1) identyfikowalność, (2) środowisko, (3) podporządkowanie realizacji celu, (4) niezależność, (5) adaptacyjność. Agent jest identyfikowalnym konkretnym bytem (ang. *entity*) z zestawem cech i reguł określających jego zachowanie oraz możliwości podejmowania decyzji. Dla agenta można jednoznacznie określić granice jego definicji – można jednoznacznie stwierdzić, czy coś jest częścią agenta, czy też nie jest. Agent jest umieszczony wewnątrz środowiska, w którym prowadzi interakcje z innymi agentami oraz może reagować na zmiany tego środowiska. Agenci mają zdolność do analizy cech innych agentów i dostosowania do nich swoich decyzji. Decyzje agenta mogą być podporządkowane realizacji wcześniej ustalonych celów – np. maksymalizacji funkcji użyteczności. Podejmując decyzję, agent ocenia poszczególne warianty pod względem stopnia realizacji przez niego jego celów. Agent jest niezależny i sam kieruje swoimi poczynaniami w ramach środowiska, w którym się znajduje. Agent zapamiętuje reakcje środowiska na jego zachowania i dostosowuje do nich swoje przyszłe reguły decyzyjne. Agent może być wyposażony w reguły określające zasady zmian reguł jego postępowania. Stąd z kolei wynika, że w modelowaniu wieloagentowych systemów ekonomicznych można zrezygnować z założenia o doskonałej racjonalności podmiotów gospodarujących na rzecz jakiegokolwiek racjonalności ograniczonej.

Cechą wspólną przedstawionych wyżej definicji agenta jest analogia do definicji elementu systemu w podejściu systemowym. Dalej przez agenta będziemy rozumieć element modelowanego systemu, który może prowadzić interakcje z innymi agentami. Macal i North (2006) definiują agenta jako „byt prowadzący interakcje ze swoim otoczeniem określony poprzez następujący zestaw cech: atrybuty, reguły zachowania, pamięć, zasoby, złożony proces podejmowania decyzji, umiejętność uczenia się, dostosowywania reguł decyzyjnych do zmieniającego się środowiska”. Znaczenie słowa *byt* jest tu tożsame z definicją pojęcia *element* w analizie systemowej oraz pojęcia *obiekt* w analizie obiektowej. Dokonując adaptacji definicji podanej przez Macal i North (2006), przyjęto definicję agenta jako elementu systemu prowadzącego interakcje ze swoim otoczeniem, określonego poprzez wybrane cechy z następującego zestawu: atrybuty, reguły zachowania, pamięć, zasoby, proces podejmowania decyzji, umiejętność uczenia się, dostosowywania reguł decyzyjnych do zmieniającego się otoczenia. W definicji tej poprzez otoczenie elementu systemu rozumiemy inne elementy, w tym innych agentów oraz otoczenie systemu.

Przez symulację wieloagentową będziemy rozumieć symulację systemu prowadzoną za pomocą komputera, w której w modelu systemu występują agenci. Gilbert i Troitzsch (2005) definiują modelowanie wieloagentowe (ang. *agent-based modeling*) jako metodę numeryczną, która umożliwia badaczowi tworzenie, analizowanie i eksperymentowanie z modelami składającymi się z agentów. Tym samym proces konstrukcji symulacji wieloagentowej i pracy z nią będziemy nazywać modelowaniem wieloagentowym.

Axtell (2000) zauważa, że w implementacji informatycznej wieloagentowego modelu systemu agenci są reprezentowani jako obiekty. Pojęcie *obiektu* jest tu rozumiane jako abstrakcja programistyczna łącząca stan z zachowaniem (por. Coad, Yourdon, 1994). Dlatego dla każdego agenta należy określić zestaw stanów oraz zasady zachowania się. Zarówno stany, jak i zachowania się mogą być prywatne lub publiczne. Prywatne stany to stany dostępne tylko danemu agentowi, stany publiczne są dostępne dla całego środowiska, w tym innych agentów. Podobnie zachowania publiczne dotyczą innych agentów (np. przeprowadzenie transakcji), a zachowania prywatne dotyczą wyłącznie agenta (np. ocena użyteczności jakiegoś zachowania).

Axtell (2000) wyróżnia następujące cztery cechy charakterystyczne modelowania wieloagentowego: heterogeniczność podmiotów, możliwość uwzględnienia lokalizacji agentów (fizycznej lub w sieci społecznej), możliwość modelowania ograniczonej racjonalności oraz możliwość pełnej analizy danych z procesu symulacji. Modelowanie heterogeniczności agentów jest bezpośrednio możliwe poprzez stosowanie podejścia obiektowego, które zakłada, że dla każdego obiektu jest określony jego stan. Częścią opisu stanu może być lokalizacja geograficzna agenta oraz referencje do innych agentów – umożliwia to tworzenie sieci społecznych. Racjonalność agenta jest modelowana poprzez tworzenie metod (zachowań) obiektu, stąd może być w modelu dowolnie kształtowana. Ponieważ implementacja wieloagentowego modelu symulacyjnego jest programem komputerowym, to istnieje możliwość dowolnego zbierania danych z procesu symulacji. Axtell (2000) podkreśla, że uwzględnienie przedstawionych czterech cech w analizie agentowej jest łatwe i naturalne, natomiast w analizie matematycznej może prowadzić do skomplikowania modelu uniemożliwiającego jego rozwiązanie.

Macal i North (2009) zauważają, że modelowanie systemów, w którym występuje adaptacja (ang. *complex adaptive systems*) charakteryzuje się następującymi trzema cechami: (1) nieliniowością, gdy zasoby są wymieniane pomiędzy agentami w sposób, w którym nie można ich dodawać; (2) zróżnicowaniem agentów, które może postępować z czasem symulacji; (3) agregacją występującą wtedy, gdy grupa agentów jest w pewnych ujęciach postrzegana jako jeden agent (tzn. badanie zachowań ludzi a badanie społeczeństwa). Ponadto w systemach z adaptacją można wyróżnić następujące cztery mechanizmy: (1) przepływ informacji i zasobów pomiędzy agentami odbywający się w taki sposób, że mogą one być przekazywane kolejnym agentom; (2) mechanizm oznaczeń (*tagging*) pozwalający agentom identyfikować wybrane cechy innych agentów; (3) mechanizm ujęcia w modelu zachowań agentów formalnie lub nieformalnie występujących w świecie rzeczywistym; (4) blokowy mechanizm konstrukcji agentów (*building block mechanism*), w którym agent może pełnić więcej niż jedną rolę w systemie. North i Macal (2009) zwracają uwagę na to, że przedstawione wyżej mechanizmy są współzależne: agregacja powoduje,

że wielu agentów może występować jako jeden, budowa agentów z elementów oznacza, że jeden agent może zachowywać się jak zbiór agentów.

Przedstawione cechy symulacji wieloagentowej wskazują na jej szczególną przydatność w modelowaniu systemów edukacyjnych. Podejście wieloagentowe umożliwia uwzględnienie heterogeniczności podmiotów na rynku edukacyjnym, dynamiki systemu edukacyjnego oraz pozwala na dowolne modelowanie procesu decyzyjnego agentów, a w szczególności uwzględnienie współzależności decyzji w systemie edukacyjnym. Cechy te przesądziły o wyborze symulacji wieloagentowej do modelowania dynamiki systemu edukacyjnego.

Podejście do modelowania systemów ekonomicznych jako całości składającej się z wielu agentów, pomiędzy którymi dochodzi do różnego rodzaju transakcji, znajduje potwierdzenie w teorii ekonomii. Literatura ekonomiczna wskazuje, że zachowania jednostek w skali mikro mają wpływ na kształt gospodarki widziany w skali makro (por. Schelling, 1978). Tym samym jest możliwe modelowanie gospodarki w skali mikro – na poziomie agentów, ich zachowań i sieci społecznych, a następnie analiza stanu modelu reprezentującego gospodarkę w skali makro. Holland (1995) zauważa, że klasyczne modele ekonomiczne zakładają pełną racjonalność decydentów, podczas gdy w rzeczywistych systemach gospodarczych podejmowanie w pełni racjonalnych decyzji nie jest możliwe. Racjonalność agentów może być ograniczona poprzez zakres dostępnych dla nich informacji, koszt pozyskania pełnej informacji i jej analizy, czy też ich możliwości przetwarzania danych. Dlatego Holland (1995) wskazuje na podejście wieloagentowe jako metodę uwzględnienia ograniczonej racjonalności w analizie systemów. Tesfatsion (2002) dokonuje obszernego przeglądu zastosowań metod symulacji wieloagentowej do modelowania systemów ekonomicznych oraz zauważa, że obecne zwiększenie dostępności mocy obliczeniowej oraz istnienie narzędzi do symulacji komputerowej pozwala na budowę modeli symulacyjnych w skali mikro. Modelując zachowania agentów w skali mikro, można wyciągać wnioski odnośnie systemu w skali makro.

Ze względu na specyfikę modelowania wieloagentowego modele te są dokumentowane w odmienny sposób niż modele DSGE bądź modele ekonometryczne. Dokumentacja modelu wieloagentowego składa się z następujących elementów:

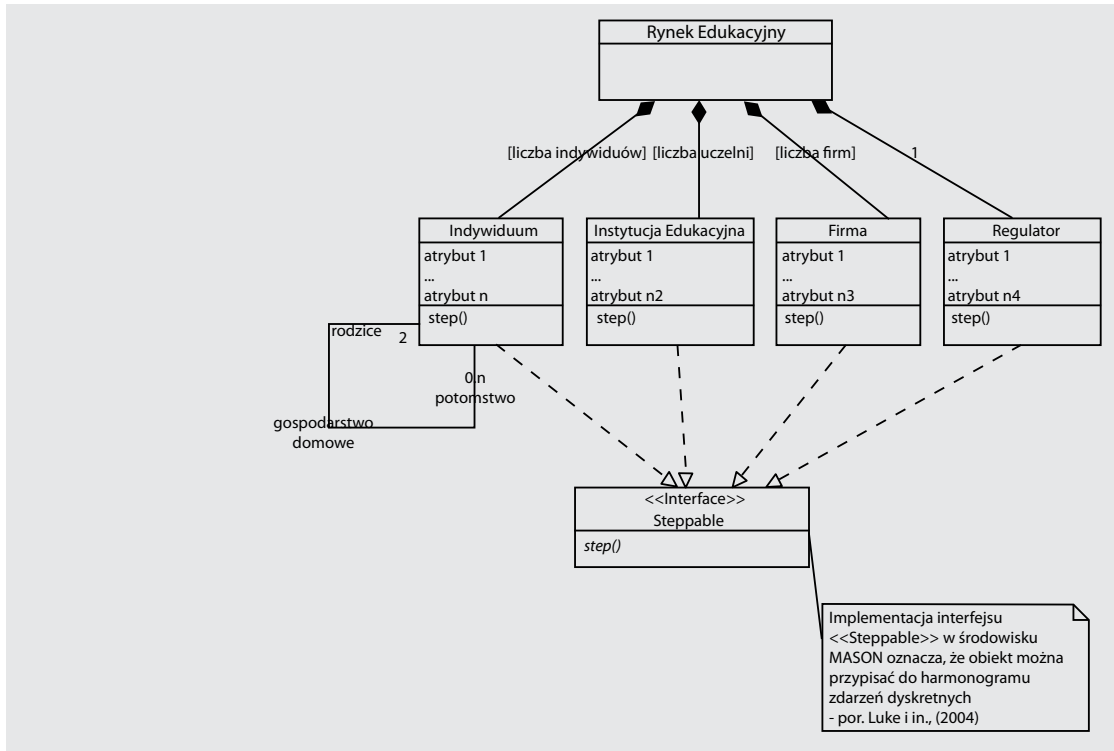
- modelu konceptualnego systemu,
- diagramów języka UML – por. OMG (2011)
 - diagramów klas – do reprezentacji klas agentów, ich atrybutów, metod oraz powiązań pomiędzy nimi;
 - diagramów czynności – do reprezentacji dynamiki systemu symulacyjnego;
 - diagramów stanów – do reprezentacji procesów decyzyjnych i zachowań pojedynczych agentów,
- pseudokodu,
- kodu źródłowego zaimplementowanego modelu w języku programowania, w którym model ten został zaimplementowany, np. w Javie.

A.IV.2.3. Wieloagentowa analiza rynków edukacyjnych

Punktem wyjścia analizy polityki regulacyjnej w ujęciu wieloagentowym jest rynek edukacyjny postrzegany w ujęciu systemowym. W badaniu zostanie przyjęte, że system edukacyjny składa się z czterech (lub w wyniku zmian założeń modelu – większej liczby) klas podmiotów: indywidualów (które tworzą gospodarstwa domowe), instytucji świadczących usługi edukacyjne, firm oraz regulatorów. Celem indywidualów jest nabycie umiejętności funkcjonowania w systemie społecznym i gospodarczym. Indywidua tworzą gospodarstwa domowe. Indywidua mogą podejmować decyzję o rozpoczęciu bądź kontynuowaniu nauki oraz o wyborze poziomu kształcenia i typu instytucji edukacyjnej. Rozważając decyzję o podjęciu nauki, indywiduum albo dostosowuje się do wymogów obowiązującego prawa, albo porównuje potencjalne korzyści z ponoszonymi kosztami oraz uwzględnia ograniczenia budżetowe. Celem instytucji edukacyjnej jest realizacja misji edukacyjnej oraz w przypadku instytucji niepublicznych – maksymalizacja zysku. Decyzje instytucji edukacyjnych obejmują wszystkie aspekty ich działalności, np. w przypadku szkół wyższych: sposób rekrutacji, tworzenie kierunków studiów, alokację dostępnych zasobów (kadrowych, finansowych, lokalowych) pomiędzy tymi kierunkami. Celem firm jest maksymalizacja zysku oraz utrzymanie pozycji i przetrwanie na rynku, na którym działają. Działania firm na rynku edukacyjnym obejmują decyzje związane z pozyskiwaniem pracowników oraz podnoszeniem ich poziomu kapitału ludzkiego. Celem regulatorów rynku jest rozwój gospodarczy, maksymalizacja dobrobytu społecznego oraz ograniczanie poziomu nierówności społecznych. Decyzje regulatorów obejmują wielkość i strukturę finansowania edukacji ze środków publicznych, sposób alokacji tych środków między instytucjami edukacyjnymi oraz strukturę organizacji rynku edukacyjnego (wymogi formalne, minima programowe).

Szkic prezentujący diagram klas dla modelu wieloagentowego został przedstawiony na rysunku A.IV.4. Opis informatyczny rzeczywistego rynku edukacyjnego odpowiada opisowi stworzonemu w oparciu o podejście systemowe. Dlatego w opisie wieloagentowego modelu systemu edukacyjnego można wyróżnić cztery klasy obiektów odpowiadające czterem klasom elementów w systemie edukacyjnym: indywidua, instytucja edukacyjna, firma, regulator. Dla Indywiduów model uwzględnia relacje międzypokoleniową dzieci-rodzice. Relacje pomiędzy indywiduami a instytucjami edukacyjnymi, firmami oraz regulatorem zostały na tym rysunku pominięte.

Rysunek A.IV.4. W modelowaniu wieloagentowym systemu edukacyjnego zostali uwzględnieni agenci reprezentujący poszczególne klasy podmiotów w systemie edukacyjnym. Klasy w modelu wieloagentowym mogą być przedstawiane w postaci diagramu UML



Źródło: Opracowanie własne.

Dla poprawnego przeprowadzenia analiz symulacyjnych zachowań dowolnego systemu podstawowe znaczenie mają następujące kwestie:

- zdefiniowanie systemu podlegającego modelowaniu,
- wybór adekwatnego narzędzia analizy – modelu systemu,
- organizacja i pozyskanie zbioru danych wejściowych,
- ustalenie warunków początkowych analizy i kalibrowanie parametrów modelu,
- ustalenie wiązki scenariuszy symulacyjnych i eksperymentowanie z modelem,
- przeprowadzenie obliczeń i analiza wrażliwości wyników modelowania na zmiany parametrów symulacji,
- zasadna interpretacja wyników obliczeń wykonanych z pomocą modelu,
- uogólnienie wyników obliczeń do rekomendacji dla decydentów sterujących systemem (oryginałem).

W przeprowadzanym badaniu systemem oryginalnym (oryginałem) jest system ekonomiczno-społeczny Polski, czyli gospodarka narodowa produkująca bądź oferująca dobra oraz usługi materialne i niematerialne w połączeniu ze zbiorem gospodarstw domowych usytuowanych na terytorium Polski oraz z przeszłymi i obecnymi decyzjami regulatora systemu. Szczególną uwagę zajmuje w tym sys-

temie sfera materialna i sfera zachowań podmiotów, które są związane z procesem edukacyjnym.

W rozpatrywanym przypadku regulatorem są agendy władzy państwowej i samorządowej, te zarówno ustawodawcze, jak i wykonawcze. Członkowie gospodarstw domowych świadczą w systemie usługi dostarczania pracy i kapitału, a z drugiej strony są konsumentami dóbr i usług materialnych oraz niematerialnych pojawiających się w systemie. Usługi edukacyjne są jednymi z oferowanych i konsumowanych usług. Są usługami o tyle szczególnymi, iż korzystanie z niektórych z nich jest obowiązkowe (szkolnictwo na poziomach od podstawowego do średniego) dlatego, że fakt wykorzystania usługi edukacyjnej zmienia charakterystykę podmiotu korzystającego z tej usługi. Konsument usług po przejściu każdego kolejnego etapu obecności w systemie edukacyjnym ma inne umiejętności, kwalifikacje i kompetencje. Wynika stąd, że przeprowadzane analizy muszą być analizami przeprowadzonymi dla skali makroekonomicznej i brać pod uwagę zarówno zróżnicowanie modelowanych podmiotów świadczących usługi edukacyjne, jak i zróżnicowanie zbioru korzystających z usług edukacyjnych. W tychże analizach musi istnieć także możliwość odwzorowania skutków decyzji regulatora na szczeblu krajowym i lokalnym. Narzędzie modelowe musi również umożliwić przeprowadzenie analiz dynamicznych dla wielu przedziałów czasowych, ponieważ nieuzasadnione jest przyjęcie założenia, że usługa edukacyjna jest świadczona w przedziale czasu, którego długość można zaniedbać.

A.IV.2.4. Kalibracja wieloagentowego modelu systemu edukacyjnego – dane UDE w badaniu symulacyjnym

Dane zebrane w badaniu UDE pozwalają na kalibrację systemu edukacyjnego w ten sposób, że w modelu symulacyjnym zostały uwzględnione charakterystyczne cechy polskiego rynku edukacyjnego oraz relacje występujące na tym rynku.

Dane zebrane w badaniu ankietowym UDE zostały uwzględnione w modelu symulacyjnym przy zastosowaniu poniższych metod:

- Obok modelowania ekonometrycznego może być przeprowadzana symulacja stochastyczna, która pozwoli na uwzględnienie nieznanego rozkładu składnika losowego; w ten sposób model symulacyjny może uwzględnić niepewność prognoz z modeli ekonometrycznych.
- Klastrowanie respondentów na grupy, a następnie dopasowywanie teoretycznych rozkładów zmiennych losowych w obrębie każdego klastra. Np. podział respondentów ze względu na formę studiów (stacjonarne/niestacjonarne) oraz rodzaj otrzymanego wykształcenia (techniczne/nietechniczne) i następnie dopasowywanie rozkładów wybranych parametrów (np. poziomu wynagrodzeń) do danych empirycznych. Proces dopasowania danych empirycznych do rozkładu teoretycznego zostanie zrealizowany poprzez wykonanie po kolei szeregu testów statystycznych badających zgodność danych empirycznych z popularnymi rozkładami teoretycznymi

(testy: Kołmogorowa-Smirnowa, Andersona-Darlinga, chi-kwadrat), a następnie wybór rozkładu, dla którego poziom dopasowania jest największy.

- Losowanie wartości bezpośrednio ze zbioru danych w trakcie symulacji.

W modelu symulacyjnym została uwzględniona, w razie potrzeby, zależność pomiędzy posiadanym wykształceniem a funkcjonowaniem na rynku pracy, a w szczególności następujące zależności:

- wykształcenie a wiek podjęcia pierwszej pracy,
- wykształcenie a okresy aktywności zawodowej,
- dochody gospodarstwa domowego a wykształcenie jego członków,
- wpływ rodzaju i formy studiów na uzyskiwane zarobki.

W modelu symulacyjnym mogą zostać uwzględnione dowolne relacje związane z tworzeniem związków i gospodarstw domowych oraz relacją pokoleń. W szczególności w modelu zostały uwzględnione następujące informacje o relacjach pomiędzy podmiotami w systemie edukacyjnym zebrane w ramach badania UDE:

- finansowanie edukacji dzieci z dochodów gospodarstwa domowego (ograniczenie budżetowe gospodarstwa wpływa na wybory edukacyjne dzieci, modelowanie przerwania lub niepodjęcia nauki z powodu niewystarczających dochodów gospodarstwa domowego),
- dziedziczenie wzorców edukacyjnych (korelacja wykształcenia dzieci i rodziców),
- tworzenie związków i gospodarstw domowych – uwzględnienie wpływu czynników edukacyjnych na proces łączenia się w związki, a następnie przekazywanie wzorców wykształcenia potomstwu oraz finansowanie tego wykształcenia.

Ponadto w modelu mogą zostać uwzględnione dane demograficzne, a w szczególności prognozy występowania różnych grup wiekowych. Model symulacyjny umożliwia również uwzględnienie trendów zwiększającego się udziału osób podejmujących studia wyższe oraz emigracji.

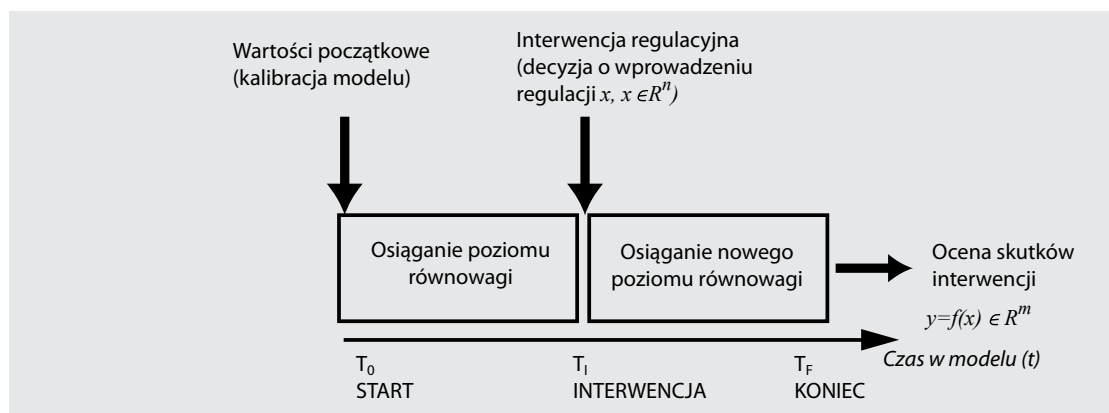
A.IV.2.5. Eksperymenty symulacyjne

Analiza przedstawionych w sekcji A.IV.1.2 pytań badawczych odbyła się poprzez przeprowadzenie eksperymentów symulacyjnych (por. Law, 2007). Eksperymentowanie na skalibrowanym modelu symulacyjnym polega na porównywaniu stanu, w jakim znajduje się model systemu edukacyjnego bez zmian polityki finansowania (tj. utrzymania obecnej polityki finansowania edukacji) z różnymi potencjalnymi scenariuszami zmian tej polityki. Model symulacyjny jest modelem stochastycz-

nym, dlatego dla każdego rozważanego sposobu regulacji tego systemu zostały wykonane wielokrotnie powtarzane symulacje.

Po wielokrotnym powtórzeniu symulacji dla różnych polityk finansowania edukacji zostały porównywane skutki wprowadzania modelowanych regulacji. Przykładowe miary skutków polityki finansowania edukacji mogą być następujące: poziom dochodu, poziom zróżnicowania nierówności społecznych, poziom bezrobocia, poziom wydatku na edukację z budżetu państwa.

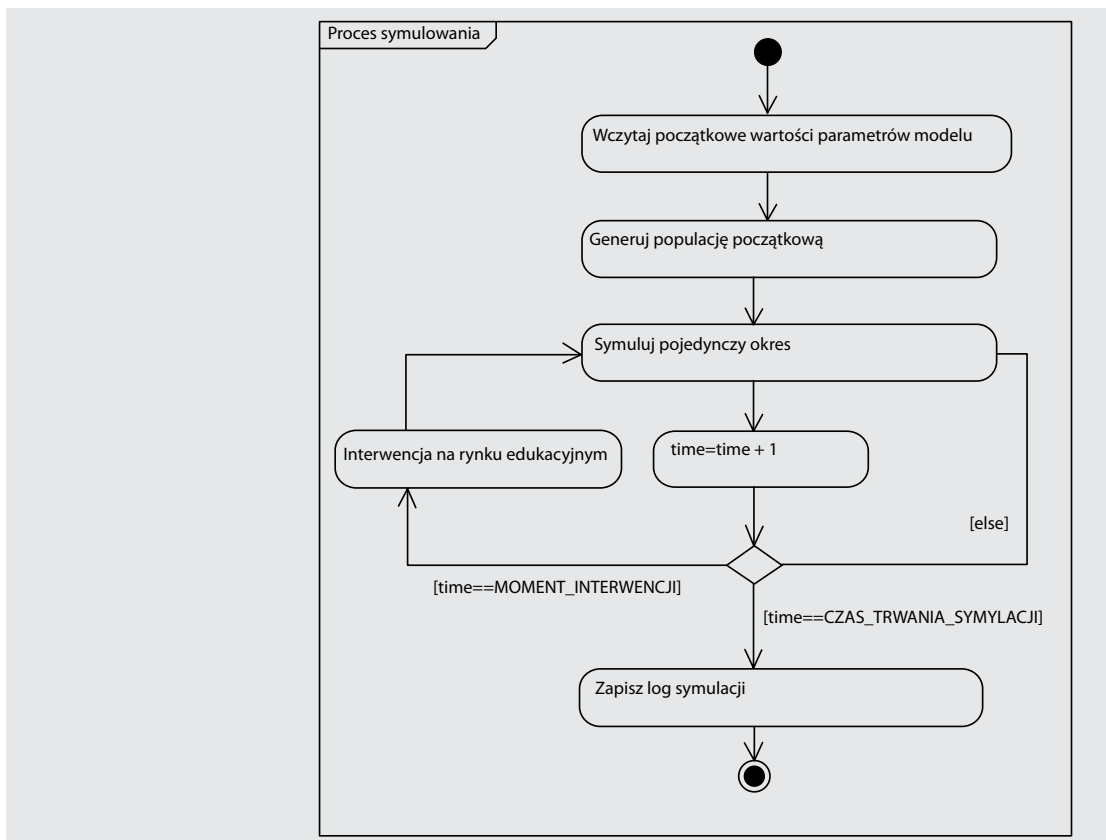
Rysunek A.IV.5. Schemat przebiegu pojedynczej symulacji dla pojedynczego eksperymentu. Aby uzyskać oszacowania zadeklarowanych wielkości wynikowych symulacja musi być wielokrotnie powtórzona



Źródło: Opracowanie własne.

Proces symulowania dynamiki w wieloagentowym systemie edukacji wyższej został przedstawiony na rysunku A.IV.6. Po inicjalizacji parametrów modelu zostaje wygenerowana początkowa populacja indywidualów. Generacja populacji początkowej wiąże się z przypisaniem indywidualom ich indywidualnych cech. Cechy te zostały wylosowane z rozkładów, których parametry zostały oszacowane na podstawie danych empirycznych z badania UDE (por. poprzedni podrozdział). Następnie będzie symulowany stan systemu w każdym kolejnym okresie (por. rysunek A.IV.3). W badaniu symulacyjnym mogą być rozważane dowolne interpretowalne merytorycznie scenariusze.

Rysunek A.IV.6. Proces symulacji interwencji (zmiany reżimu) w wieloagentowym modelu systemu edukacyjnego



Źródło: Szufel (2012).

Skutki polityki finansowania wyliczone przy pomocy modelu stochastycznego są zmiennymi losowymi, dlatego ich porównywanie wymaga zastosowania odpowiednich testów statystycznych (np. test Welcha).

Aneks B – Zestawienia tabelaryczne, szczegółowe wyniki analiz

B.I. Blok I: Wyniki – uwarunkowania biografii edukacyjnych

Tabela B.I.1.

Wyniki estymacji modelu regresji logistycznej dla zmiennej zależnej o wielu kategoriach, kobiety (N=12497)

	szkoła średnia vs. szkoła zawodowa	studia vs. szkoła zawodowa	studia vs. szkoła średnia
Wyraz wolny	0.062(0.145)	-2.53***(0.215)	-2.592***(0.186)
Wykształcenie matki (ref: podstawowe)			
zasadnicze zawodowe	-0.027(0.085)	-0.018(0.178)	0.01(0.172)
średnie	1.275***(0.156)	2.981***(0.185)	1.706***(0.134)
wyższe	1.097***(0.358)	3.631****(0.375)	2.534****(0.24)
Homogamia rodziców (ref: rodzice takie samo wykształcenie)			
matka gorzej wykształcona od ojca	0.256****(0.078)	0.665****(0.107)	0.409****(0.091)
matka lepiej wykształcona od ojca	-0.386****(0.117)	-1.051****(0.135)	-0.665****(0.09)
Liczba rodzeństwa (ref: 0=jedynak)			
1	-0.041(0.129)	-0.037(0.16)	0.003(0.12)
2	-0.219(0.13)	-0.342*(0.166)	-0.122(0.129)
3	-0.443***(0.137)	-0.919****(0.183)	-0.476***(0.149)
4+	-0.694****(0.136)	-1.273****(0.19)	-0.579****(0.159)
Kolejność wśród rodzeństwa (ref: najstarszy)			
środkowy	-0.246****(0.073)	-0.301***(0.11)	-0.055(0.098)
najmłodszy	-0.133(0.069)	-0.231*(0.093)	-0.098(0.076)
Okres zakończenia edukacji (ref: przed 1990)			
1990-1999	-0.098(0.1)	0.57***(0.184)	0.668****(0.18)
2000-2013	0.362***(0.132)	2.454****(0.167)	2.091****(0.152)
Liczba książek w domu rodzinnym (ref: 0-10)			
11-25 książek	0.282****(0.069)	0.548****(0.138)	0.267*(0.133)
26-100 książek	0.649****(0.072)	1.399****(0.132)	0.75****(0.125)
ponad 100 książek	0.835****(0.108)	1.986****(0.157)	1.151****(0.136)
Sytuacja materialna domu rodzinnego w porównaniu do innych (ref: podobna)			
gorsza	-0.066(0.079)	-0.1(0.131)	-0.034(0.12)
lepszta	0.079(0.089)	-0.095(0.117)	-0.174(0.093)
Zmienne z analizy czynnikowej			
relacje w domu rodzinnym	0.081***(0.027)	0.202****(0.039)	0.121****(0.034)

	szkoła średnia vs. szkoła zawodowa	studia vs. szkoła zawodowa	studia vs. szkoła średnia
nastawienie na samorealizację, wykształcenie	0.337*** (0.027)	0.762*** (0.04)	0.425*** (0.034)
korzystanie z doświadczeń rodziny i znajomych	-0.215*** (0.027)	-0.59*** (0.038)	-0.375*** (0.031)
ambicje domu rodzinnego	0.27*** (0.028)	0.497*** (0.037)	0.227*** (0.03)
Wykształcenie matki * okres zakończenia edukacji (ref: podstawowe * przed 1990)	0.057 (0.151)	0.012 (0.277)	-0.045 (0.269)
zasadnicze zawodowe*1990-1999	0.355 (0.249)	0.144 (0.304)	-0.211 (0.23)
średnie*1990-1999	1.186 (1.121)	1.955 (1.108)	0.769 (0.447)
wyższe*1990-1999	0.65*** (0.175)	0.482 (0.247)	-0.167 (0.225)
zasadnicze zawodowe*2000-2013	0.936*** (0.276)	0.247 (0.3)	-0.689*** (0.193)
średnie*2000-2013	1.485 (0.812)	0.3 (0.815)	-1.185*** (0.306)
wyższe*2000-2013	0.062 (0.145)	-2.53*** (0.215)	-2.592*** (0.186)

ocena istotna przy poziomie istotności * 0,1; **0,05; *** 0,001

Tabela B.I.2.

Wyniki estymacji wielo-wartościowego modelu regresji logistycznej, mężczyźni (N=10982)

	szkoła średnia vs. szkoła zawodowa	studia vs. szkoła zawodowa	studia vs. szkoła średnia
Wyraz wolny	-1.04*** (0.145)	-3.405*** (0.245)	-2.365*** (0.231)
Wykształcenie matki (ref: podstawowe)			
zasadnicze zawodowe	0.477*** (0.094)	0.014 (0.22)	-0.463* (0.221)
średnie	1.545*** (0.137)	2.296*** (0.195)	0.751*** (0.174)
wyższe	1.428*** (0.314)	3.283*** (0.368)	1.855*** (0.312)
Homogamia rodziców (ref: rodzice takie samo wykształcenie)	0*** (0)	0*** (0)	0*** (0)
matka gorzej wykształcona od ojca	0.493*** (0.082)	0.929*** (0.119)	0.436*** (0.11)
matka lepiej wykształcona od ojca	-0.571*** (0.105)	-1.485*** (0.133)	-0.913*** (0.105)
Liczba rodzeństwa (ref: 0=jedynak)			
1	-0.022 (0.121)	0.012 (0.163)	0.034 (0.139)
2	-0.118 (0.125)	-0.423* (0.174)	-0.305* (0.153)
3	-0.194 (0.136)	-0.606** (0.204)	-0.412* (0.185)
4+	-0.761*** (0.14)	-1.112*** (0.219)	-0.35 (0.205)
Kolejność wśród rodzeństwa (ref: najstarszy)			
środkowy	-0.277*** (0.081)	-0.284* (0.132)	-0.006 (0.124)
najmłodszy	-0.191** (0.07)	-0.263* (0.102)	-0.072 (0.091)

Okres zakończenia edukacji (ref: przed 1990)			
1990-1999	-0.056(0.121)	0.301(0.227)	0.357(0.235)
2000-2013	0.918*** (0.145)	1.74*** (0.215)	0.822*** (0.207)
Liczba książek w domu rodzinnym (ref: 0-10)			
11-25 książek	0.312*** (0.078)	0.693*** (0.173)	0.381* (0.174)
26-100 książek	0.569*** (0.08)	1.238*** (0.167)	0.669*** (0.166)
ponad 100 książek	0.984*** (0.113)	2.293*** (0.185)	1.31*** (0.174)
Sytuacja materialna domu rodzinnego w porównaniu do innych (ref: podobna)			
gorsza	-0.034(0.098)	0.015(0.167)	0.049(0.161)
lepszta	0.006(0.099)	-0.349*(0.138)	-0.355** (0.118)
Zmienne z analizy czynnikowej			
relacje w domu rodzinnym	0.145*** (0.03)	0.284*** (0.05)	0.139** (0.047)
nastawienie na samorealizację, wykształcenie	0.287*** (0.029)	0.772*** (0.045)	0.485*** (0.041)
korzystanie z doświadczeń rodziny i znajomych	-0.023(0.028)	-0.289*** (0.042)	-0.266*** (0.038)
ambicje domu rodzinnego	0.288*** (0.029)	0.537*** (0.041)	0.25*** (0.036)
Wykształcenie matki * okres zakończenia edukacji (ref: podstawowe * przed 1990)			
zasadnicze zawodowe*1990-1999	-0.154(0.169)	-0.549(0.368)	-0.395(0.373)
średnie*1990-1999	-0.063(0.205)	0.111(0.31)	0.174(0.295)
wyższe*1990-1999	1.264(0.653)	1.574*(0.702)	0.31(0.465)
zasadnicze zawodowe*2000-2013	-0.113(0.183)	0.328(0.307)	0.441(0.298)
średnie*2000-2013	0.187(0.222)	0.686*(0.293)	0.499(0.255)
wyższe*2000-2013	0.887(0.518)	0.798(0.562)	-0.088(0.376)

ocena istotna przy poziomie istotności * 0,1; **0,05; *** 0,001

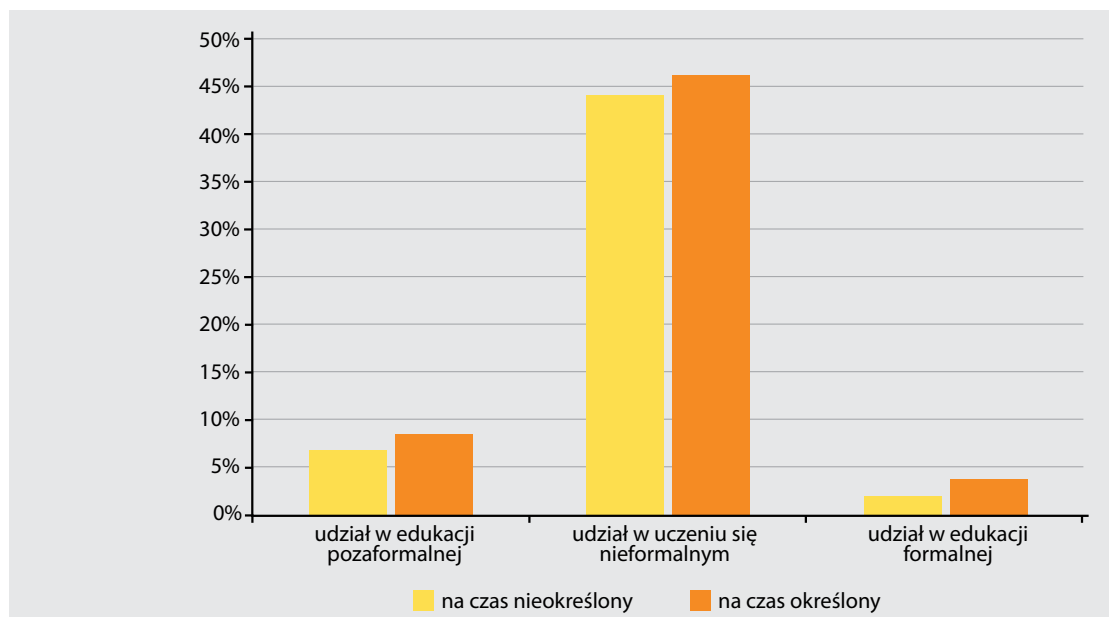
B.II. Blok II: Wyniki

Tabela B.II.1.

Udział w poszczególnych formach edukacji wg wielkości miejsca zamieszkania

	formalna	pozaformalna	nieformalna
gmina wiejska	1.5%	3.5%	32.7%
miasto do 10 tys. mieszkańców	2.7%	4.1%	39.2%
miasto 10.000-19.999 mieszkańców	2.1%	5.1%	39.9%
miasto 20.000-49.999 mieszkańców	2.4%	5.8%	33.6%
miasto 50.000-99.999 mieszkańców	1.8%	5.0%	38.3%
miasto 100.000-199.999 mieszkańców	2.8%	7.8%	42.4%
miasto 200.000-499.999 mieszkańców	3.2%	7.3%	45.0%
miasto 500.000-999.999 mieszkańców	4.2%	12.4%	59.9%
Warszawa	5.0%	10.1%	46.2%

Rysunek B.II.1. Udział respondentów w różnych typach edukacji wg typu umowy o pracę



Źródło: Obliczenia własne.

B.III. Blok III: Wyniki

Tabela B.III.1.

Wartości własne i stopień wyjaśniania inercji przez czynniki w analizie porównawczej profili wykluczonych i niewykluczonych z edukacji na jej różnych poziomach

Czynniki	Wartości własne czynników	Stopień wyjaśniania inercji	Skumulowany stopień wyjaśniania inercji
F ₁	0,2947	0,7926	0,7926
F ₂	0,1249	0,1423	0,9349
F ₃	0,0673	0,0414	0,9763
F ₄	0,0510	0,0237	1,0000

Tabela B.III.2.

Jakość odwzorowania w dwuwymiarowej przestrzeni czynnikowej oraz znaczenie podzbiorowości wykluczonych i niewykluczonych z edukacji (obiektów) na różnych jej poziomach, w tworzeniu przestrzeni czynnikowej

Obiekty - podzbiorowości wykluczonych i niewykluczonych	Jakość odwzorowania w dwuwymiarowej przestrzeni czynnikowej	Znaczenie obiektów w tworzeniu dwuwymiarowej przestrzeni czynnikowej
Wykluczeni z edukacji:		
szkoła podstawowa i gimnazjum ¹	0,843	0,100
szkoła ponadgimnazjalna bez matury ²	0,962	0,214
szkoła ponadgimnazjalna z maturą	0,962	0,231
szkoła wyższa ³	0,216	0,036
Niewykluczeni z edukacji	0,989	0,419

¹Gimnazjum dotyczy osób w wieku 16-28 lat, a szkoła podstawowa osób w wieku 29-65 lat.

²Osoby w wieku 23-65 lat.

³Osoby w wieku 19-65 lat.

Tabela B.III.3.

Jakość odwzorowania w dwuwymiarowej przestrzeni czynnikowej oraz znaczenie charakterystyk osób wykluczonych z edukacji na różnych jej poziomach w tworzeniu przestrzeni czynnikowej

Zmienne – charakterystyki osób wykluczonych i niewykluczonych	Jakość odwzorowania w dwuwymiarowej przestrzeni czynnikowej	Znaczenie obiektów w tworzeniu dwuwymiarowej przestrzeni czynnikowej
Poziom wykształcenia rodziców¹⁾:		
- podstawowe lub niższe	0,955	0,074
- gimnazjum	0,964	0,063
- zasadnicze zawodowe	0,987	0,015
- średnie	0,946	0,027
- wyższe	0,974	0,128

Zmienne – charakterystyki osób wykluczonych i niewykluczonych	Jakość odwzorowania w dwuwymiarowej przestrzeni czynnikowej	Znaczenie obiektów w tworzeniu dwuwymiarowej przestrzeni czynnikowej
Klasa miejscowości zamieszkania:		
- wieś	0,972	0,007
- miasto poniżej 20 tys. mieszkańców	0,916	0,002
- miasto 20-100 tys. mieszkańców	0,943	0,001
- miasto 100-200 tys. mieszkańców	0,921	0,005
- miasto 200-500 tys. mieszkańców	0,999	0,019
- miasto powyżej 500 tys. mieszkańców	0,984	0,045
Województwo:		
- dolnośląskie	0,023	0,004
- kujawsko-pomorskie	0,949	0,010
- lubelskie	0,575	0,002
- lubuskie	0,854	0,012
- łódzkie	0,237	0,001
- małopolskie	0,854	0,006
- mazowieckie	0,995	0,013
- opolskie	0,607	0,004
- podkarpackie	0,629	0,002
- podlaskie	0,547	0,003
- pomorskie	0,469	0,004
- śląskie	0,778	0,002
- świętokrzyskie	0,666	0,001
- warmińsko-mazurskie	0,686	0,003
- wielkopolskie	0,166	0,002
- zachodnio-pomorskie	0,420	0,002
Płeć:		
- kobieta	1,000	0,009
- mężczyzna	0,996	0,003
Niepełnosprawność:		
- ograniczona zdolność wykonywania codziennych czynności trwających dłużej niż?	0,932	0,095
- brak ograniczeń	0,995	0,001
Aktywność ekonomiczna:		
- pracujący	0,988	0,006
- bezrobotny	0,946	0,013
- bierny zawodowo	0,704	0,021

Zmienne – charakterystyki osób wykluczonych i niewykluczonych	Jakość odwzorowania w dwuwymiarowej przestrzeni czynnikowej	Znaczenie obiektów w tworzeniu dwuwymiarowej przestrzeni czynnikowej
Zamożność gospodarstwa domowego		
- grupa kwintylowa rozkładu dochodów ekwiwalentnych:		
1	0,986	0,032
2	0,979	0,010
3	0,623	0,002
4	0,849	0,012
5	0,973	0,054
- skrajne ubóstwo:		
sfera skrajnego ubóstwa	0,982	0,044
poza sferę skrajnego ubóstwa	0,989	0,002
Typ gospodarstwa domowego (liczba rodzin i typ rodziny biologicznej)		
- jednorodzinne:		
małżeństwa bez dzieci	0,949	0,038
małżeństwa z 1 dzieckiem	0,987	0,010
małżeństwa z 2 dziećmi	0,938	0,001
małżeństwa z 3 i więcej dzieci rodziny niepełne	0,996	0,014
- wielorodzinne	0,955	0,019
- nierodzinne:	0,979	0,007
jednoosobowe	0,996	0,034
wieloosobowe	0,112	0,002
Grupa społeczno-ekonomiczna gospodarstwa domowego:		
- pracownicy	0,996	0,003
- rolnicy	0,947	0,013
- pracujący na własny rachunek	0,913	0,010
- emeryci	0,756	0,003
- renciści	0,827	0,044
- utrzymujący się z niezarobkowych źródeł innych niż emerytura i renta	0,991	0,038

1) Wyższy z poziomów wykształcenia rodziców.

Tabela B.III.4.

Zredukowana macierz korelacji

Zmienne	Współczynniki korelacji											
	X_1	X_2	X_3	X_4	X_5	X_6	X_{12}	X_8	X_9	X_{10}	X_{11}	
Poziom wykształcenia rodziców – X_1	0,684											
Wzorce kulturowe – X_2	0,579	0,569										
Wyznaczniki sukcesu w domu rodzinnym:												
rodzina – X_3	-0,077	-0,049	0,638									
prestżowa praca – X_4	0,045	0,007	-0,235	0,668								
wysoki status materialny – X_5	0,100	0,038	-0,181	0,507	0,638							
dobrze wykształcenie – X_6	0,409	0,340	-0,247	0,288	0,391	0,664						
Motywacje wyborów edukacyjnych:												
odległość od szkoły – X_{12}	-0,065	-0,067	0,111	0,008	-0,003	-0,027	0,586					
prestż szkoły – X_8	0,216	0,174	0,081	0,105	0,075	0,259	0,384	0,744				
jakość nauczania – X_9	0,244	0,210	0,079	0,093	0,095	0,306	0,325	0,820	0,767			
profil związany z zainteresowaniami – X_{10}	0,254	0,219	0,064	0,085	0,071	0,29	0,141	0,454	0,526	0,682		
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}	0,088	0,076	0,036	0,140	0,104	0,200	0,197	0,392	0,446	0,605	0,626	
Zamożność domu rodzinnego – X_{18}	-0,304	-0,284	-0,041	-0,094	-0,128	-0,188	0,026	-0,106	-0,109	-0,097	-0,052	
Charakterystyka domu rodzinnego:												
silne więzi rodzinne – X_{13}	0,070	0,120	0,567	0,020	-0,116	0,187	0,106	0,208	0,211	0,211	0,148	
wspólne spędzanie wolnego czasu – X_{14}	0,094	0,128	0,541	0,005	-0,103	0,203	0,093	0,167	0,181	0,192	0,144	
nacisk na pracę i naukę – X_{15}	0,210	0,221	0,294	0,160	0,060	0,362	0,082	0,255	0,291	0,236	0,181	
Uczęszczanie do przedszkola – X_{16}	0,398	0,290	-0,025	0,038	0,047	0,201	-0,020	0,087	0,101	0,159	0,060	
Uczęszczanie na zajęcia dodatkowe – X_{17}	0,423	0,414	-0,113	0,017	0,038	0,399	-0,133	0,150	0,192	0,249	0,053	
Zainteresowanie rodziców edukacją dziecka – X_7	0,256	0,267	0,054	0,066	0,005	0,345	-0,006	0,137	0,159	0,186	0,102	
Klasa miejscowości zamieszkania – X_{19}	0,332	0,302	-0,110	-0,001	0,038	0,075	-0,096	0,044	0,054	0,074	0,048	
Poziom rozwoju województwa zamieszkania – X_{20}	0,037	0,065	0,002	0,028	0,024	-0,024	0,016	0,023	0,007	0,005	0,024	
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	0,022	0,051	-0,008	0,008	0,024	-0,042	0,008	0,026	0,018	0,020	0,027	
Wczesne rodzicielstwo – X_{22}	0,292	0,162	-0,030	0,016	0,034	0,166	0,020	0,099	0,112	0,109	0,045	

Tabela B.III.4.

Zredukowana macierz korelacji (c.d)

Zmienne	Współczynniki korelacji										
	X_{18}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_7	X_{19}	X_{20}	X_{21}	X_{22}
Poziom wykształcenia rodziców – X_1											
Wzorce kulturowe – X_2											
Wyznaczniki sukcesu w domu rodzinnym:											
rodzina – X_3											
prestżowa praca – X_4											
wysoki status materialny – X_5											
dobrze wykształcenie – X_6											

Zmienne	Współczynniki korelacji										
	X_{18}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_7	X_{19}	X_{20}	X_{21}	X_{22}
Motywacje wyborów edukacyjnych:											
odległość od szkoły – X_{12}											
prestż szkoły – X_8											
jakość nauczania – X_9											
profil związany z zainteresowaniami – X_{10}											
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}											
Zamożność domu rodzinnego – X_{18}	0,251										
Charakterystyka domu rodzinnego:											
silne więzi rodzinne – X_{13}	-0,191	0,836									
wspólne spędzanie wolnego czasu – X_{14}	-0,209	0,856	0,822								
nacisk na pracę i naukę – X_{15}	-0,192	0,580	0,580	0,620							
Uczęszczanie do przedszkola – X_{16}	-0,138	0,033	0,043	0,130	0,393						
Uczęszczanie na zajęcia dodatkowe – X_{17}	-0,146	0,023	0,028	0,164	0,301	0,511					
Zainteresowanie rodziców edukacją dziecka – X_7	-0,216	0,305	0,289	0,420	0,156	0,263	0,398				
Klasa miejscowości zamieszkania – X_{19}	-0,054	-0,086	-0,084	-0,002	0,175	0,223	0,058	0,420			
Poziom rozwoju województwa zamieszkania – X_{20}	-0,008	-0,008	-0,024	-0,018	0,043	-0,022	-0,025	0,244	0,906		
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	0,001	0,003	-0,020	-0,016	0,036	-0,020	-0,021	0,263	0,928	0,948	
Wczesne rodzicielstwo – X_{22}	-0,061	0,042	0,056	0,037	0,222	0,148	0,097	0,004	-0,008	-0,001	0,479

Tabela B.III.5.

Wartości miary Kaisera-Mayera-Olkina (KMO)

Zmienne	Wartości miary KMO
Poziom wykształcenia rodziców – X_1	0,8174
Wzorce kulturowe – X_2	0,8598
Wyznaczniki sukcesu w domu rodzinnym:	
rodzina – X_3	0,6799
prestżowa praca – X_4	0,5681
wysoki status materialny – X_5	0,5264
dobre wykształcenie – X_6	0,7454
Zainteresowanie rodziców edukacją dziecka – X_7	0,8801
Motywacje wyborów edukacyjnych:	
prestż szkoły – X_8	0,7263
jakość nauczania – X_9	0,7456
profil związany z zainteresowaniami – X_{10}	0,8046
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}	0,7556
odległość od szkoły – X_{12}	0,8137

Zmienne	Wartości miary KMO
Charakterystyka domu rodzinnego:	
silne więzi rodzinne – X_{13}	0,7173
wspólne spędzanie wolnego czasu – X_{14}	0,7524
nacisk na pracę i naukę – X_{15}	0,9016
Uczęszczanie do przedszkola – X_{16}	0,8847
Uczęszczanie na zajęcia dodatkowe – X_{17}	0,8566
Zamożność gospodarstwa domowego – X_{18}	0,8335
Klasa miejscowości zamieszkania – X_{19}	0,8116
Poziom rozwoju województwa zamieszkania – X_{20}	0,4991
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	0,4978
Wczesne rodzicielstwo – X_{22}	0,8082

Tabela B.III.6.

Wartości własne czynników wspólnych oraz ich udziały w wariancji wyjaśnianej przez czynniki wspólne

Czynniki	Wartości własne czynników	Udziały czynników w wariancji wspólnej w proc.	Skumulowane udziały czynników w wariancji wspólnej w proc.
F_1	3,03180	13,78	13,78
F_2	2,97483	13,52	27,30
F_3	2,81289	12,79	40,09
F_4	2,06534	9,39	49,48
F_5	1,90841	8,67	58,15
F_6	1,09190	4,96	63,11

Tabela B.III.7.

Wyniki estymacji modeli ścieżek oddziaływania wybranych czynników na wykluczenie z edukacji

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{2A,1A}$	0,070	0,017	4,21	0,000
$\beta_{2A,1B}$	0,159	0,013	12,43	0,000
$\beta_{2A,1C}$	0,032	0,005	6,21	0,000
$\beta_{2B,1A}$	0,008	0,015	0,6	0,569
$\beta_{2B,1B}$	0,244	0,015	16,5	0,000
$\beta_{2B,1C}$	0,092	0,010	8,9	0,000

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{2C,1A}$	-0,022	0,011	-2,1	0,038
$\beta_{2C,1B}$	0,202	0,016	12,4	0,000
$\beta_{2C,1C}$	0,251	0,017	14,6	0,000
$\beta_{2D,1A}$	-0,031	0,009	-3,3	0,001
$\beta_{2D,1B}$	0,108	0,015	7,4	0,000
$\beta_{2D,1C}$	0,319	0,023	14,1	0,000
$\beta_{2E,1A}$	-0,017	0,009	-1,8	0,076
$\beta_{2E,1B}$	0,050	0,013	4,0	0,000
$\beta_{2E,1C}$	0,252	0,025	10,1	0,000
$\beta_{2F,1A}$	0,008	0,015	0,6	0,569
$\beta_{2F,1B}$	0,244	0,015	16,5	0,000
$\beta_{2F,1C}$	0,092	0,010	8,9	0,000
$\beta_{4,1A}$	0,007	0,013	0,5	0,607
$\beta_{4,1B}$	0,013	0,016	0,8	0,399
$\beta_{,1C}$	0,004	0,015	0,3	0,789
$\beta_{4,15}$	0,056	0,012	4,8	0,000
$\beta_{,4}$	0,273	0,017	16,5	0,000
$\beta_{5,6}$	0,149	0,013	11,5	0,000
$\beta_{5,15}$	-0,015	0,011	-1,4	0,177
$\beta_{6,1A}$	0,003	0,012	0,2	0,828
$\beta_{6,1B}$	0,131	0,015	8,5	0,000
$\beta_{6,1C}$	0,128	0,016	8,0	0,000
$\beta_{6,4}$	0,169	0,014	12,2	0,000
$\beta_{6,15}$	0,134	0,009	14,2	0,000
$\beta_{9,1A}$	-0,006	0,013	-0,5	0,622
$\beta_{9,1B}$	0,110	0,015	7,3	0,000
$\beta_{9,1C}$	0,060	0,013	4,5	0,000
$\beta_{9,15}$	0,182	0,013	14,4	0,000
$\beta_{10,1A}$	0,017	0,011	1,54	0,123
$\beta_{10,1B}$	0,059	0,013	4,39	0,000
$\beta_{10,1C}$	0,011	0,012	0,96	0,336
$\beta_{10,9}$	0,409	0,018	23,31	0,000
$\beta_{10,15}$	0,071	0,012	5,83	0,000
$\beta_{11,1A}$	0,023	0,010	2,26	0,024
$\beta_{11,1B}$	-0,035	0,012	-2,89	0,004
$\beta_{11,1C}$	-0,025	0,013	-1,97	0,049
$\beta_{11,4}$	0,060	0,011	5,43	0,000

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{11,5}$	0,010	0,011	0,94	0,346
$\beta_{11,6}$	0,007	0,011	0,63	0,527
$\beta_{11,9}$	0,196	0,017	11,52	0,000
$\beta_{11,10}$	0,419	0,015	27,35	0,000
$\beta_{11,15}$	0,018	0,010	1,82	0,069
$\beta_{13,3}$	0,270	0,015	18,5	0,000
$\beta_{14,3}$	0,113	0,125	9,0	0,000
$\beta_{14,13}$	0,573	0,018	32,3	0,000
$\beta_{16,1A}$	0,011	0,014	0,8	0,405
$\beta_{16,1B}$	0,119	0,016	7,5	0,000
$\beta_{16,1C}$	0,091	0,014	6,4	0,000
$\beta_{17,1A}$	-0,009	0,012	-0,8	0,428
$\beta_{17,1B}$	0,154	0,016	9,6	0,000
$\beta_{17,1C}$	0,139	0,018	7,8	0,000
$\beta_{21,1A}$	0,025	0,011	2,3	0,025
$\beta_{21,1B}$	0,011	0,013	0,8	0,401
$\beta_{21,1C}$	-0,021	0,012	-1,8	0,077
$\beta_{21,20}$	0,636	0,010	63,2	0,000
$\beta_{w0,1A}$	-0,027	0,016	-1,66	0,098
$\beta_{w0,1B}$	-0,007	0,013	-0,56	0,575
$\beta_{w0,1C}$	0,014	0,010	1,43	0,154
$\beta_{w0,2A}$	-0,034	0,022	-1,54	0,124
$\beta_{w0,2B}$	-0,018	0,017	-1,05	0,292
$\beta_{w0,2C}$	0,021	0,017	1,27	0,203
$\beta_{w0,2D}$	-0,020	0,012	-1,67	0,094
$\beta_{w0,2E}$	-0,001	0,004	-0,21	0,831
$\beta_{w0,3}$	0,029	0,015	1,92	0,055
$\beta_{w0,4}$	0,022	0,018	1,27	0,203
$\beta_{w0,5}$	-0,028	0,010	-2,80	0,005
$\beta_{w0,6}$	-0,005	0,008	-0,66	0,510
$\beta_{w0,8}$	0,026	0,011	2,48	0,013
$\beta_{w0,9}$	0,004	0,009	0,41	0,683
$\beta_{w0,10}$	-0,034	0,012	-2,92	0,004
$\beta_{w0,11}$	-0,039	0,009	-4,39	0,000
$\beta_{w0,13}$	-0,021	0,024	-0,88	0,377
$\beta_{w0,14}$	-0,033	0,024	-1,37	0,171
$\beta_{w0,15}$	-0,034	0,020	-1,69	0,090

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{W0,16}$	-0,031	0,014	-2,27	0,023
$\beta_{W0,17}$	-0,002	0,009	-0,25	0,801
$\beta_{W0,19A}$	0,000	0,017	0,02	0,986
$\beta_{W0,19B}$	0,013	0,020	0,66	0,508
$\beta_{W0,19C}$	-0,011	0,020	-0,55	0,583
$\beta_{W0,19D}$	-0,005	0,016	-0,33	0,739
$\beta_{W0,19E}$	0,008	0,012	0,66	0,509
$\beta_{W0,20}$	0,004	0,013	0,33	0,741
$\beta_{W0,21}$	0,001	0,017	0,08	0,937
$\beta_{W0,22}$	0,017	0,014	1,25	0,212
$\beta_{W1,1A}$	-0,088	0,015	-5,97	0,000
$\beta_{W1,1B}$	-0,017	0,011	-1,47	0,141
$\beta_{W1,1C}$	0,033	0,009	3,70	0,000
$\beta_{W1,2A}$	-0,095	0,020	-4,75	0,000
$\beta_{W1,2B}$	-0,032	0,016	-2,03	0,043
$\beta_{W1,2C}$	-0,016	0,011	-1,50	0,133
$\beta_{W1,2D}$	-0,012	0,008	-1,62	0,105
$\beta_{W1,2E}$	0,007	0,007	1,06	0,288
$\beta_{W1,3}$	0,028	0,011	2,52	0,012
$\beta_{W1,4}$	-0,003	0,012	-0,25	0,805
$\beta_{W1,5}$	-0,017	0,010	-1,65	0,099
$\beta_{W1,6}$	-0,039	0,008	-4,65	0,000
$\beta_{W1,8}$	0,035	0,012	3,01	0,003
$\beta_{W1,9}$	-0,015	0,012	-1,26	0,207
$\beta_{W1,10}$	-0,139	0,012	-12,03	0,000
$\beta_{W1,11}$	-0,114	0,011	-10,78	0,000
$\beta_{W1,13}$	-0,030	0,018	-1,63	0,104
$\beta_{W1,14}$	-0,001	0,018	-0,08	0,939
$\beta_{W1,15}$	-0,095	0,017	-5,50	0,000
$\beta_{W1,16}$	-0,017	0,013	-1,36	0,173
$\beta_{W1,17}$	-0,025	0,008	-3,27	0,001
$\beta_{W1,19A}$	-0,031	0,018	-1,69	0,091
$\beta_{W1,19B}$	0,037	0,020	1,85	0,065
$\beta_{W1,19C}$	-0,033	0,018	-1,80	0,073
$\beta_{W1,19D}$	-0,021	0,017	-1,20	0,229
$\beta_{W1,19E}$	0,023	0,013	1,81	0,071
$\beta_{W1,20}$	-0,008	0,016	-0,52	0,604

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{W1,21}$	-0,003	0,016	-0,19	0,851
$\beta_{W1,22}$	0,007	0,014	0,51	0,607
$\beta_{W2,1A}$	0,029	0,013	2,19	0,029
$\beta_{W2,1B}$	-0,157	0,013	-11,74	0,000
$\beta_{W2,1C}$	0,004	0,009	0,48	0,629
$\beta_{W2,2A}$	-0,012	0,015	-0,75	0,452
$\beta_{W2,2B}$	-0,077	0,016	-4,90	0,000
$\beta_{W2,2C}$	-0,045	0,012	-3,88	0,000
$\beta_{W2,2D}$	0,015	0,010	1,50	0,133
$\beta_{W2,2E}$	0,014	0,008	1,80	0,073
$\beta_{W2,3}$	0,011	0,011	1,01	0,314
$\beta_{W2,4}$	0,015	0,012	1,25	0,211
$\beta_{W2,5}$	0,000	0,011	-0,02	0,985
$\beta_{W2,6}$	-0,122	0,011	-11,00	0,000
$\beta_{W2,8}$	-0,090	0,017	-5,28	0,000
$\beta_{W2,9}$	-0,081	0,018	-4,55	0,000
$\beta_{W2,10}$	0,008	0,013	0,58	0,562
$\beta_{W2,11}$	0,070	0,013	5,63	0,000
$\beta_{W2,13}$	0,021	0,015	1,45	0,148
$\beta_{W2,14}$	-0,004	0,014	-0,26	0,791
$\beta_{W2,15}$	-0,022	0,013	-1,72	0,085
$\beta_{W2,16}$	-0,025	0,011	-2,23	0,026
$\beta_{W2,17}$	-0,121	0,009	-13,45	0,000
$\beta_{W2,19A}$	-0,003	0,018	-0,16	0,875
$\beta_{W2,19B}$	-0,019	0,019	-0,98	0,329
$\beta_{W2,19C}$	-0,013	0,019	-0,66	0,507
$\beta_{W2,19D}$	-0,014	0,019	-0,77	0,441
$\beta_{W2,19E}$	0,014	0,014	1,00	0,316
$\beta_{W2,20}$	-0,043	0,017	-2,57	0,010
$\beta_{W2,21}$	0,046	0,016	2,98	0,003
$\beta_{W2,22}$	0,109	0,012	9,20	0,000
$\beta_{W3,1A}$	0,050	0,012	4,06	0,000
$\beta_{W3,1B}$	0,015	0,017	0,90	0,371
$\beta_{W3,1C}$	-0,076	0,014	-5,25	0,000
$\beta_{W3,2A}$	0,039	0,014	2,86	0,004
$\beta_{W3,2B}$	0,024	0,016	1,52	0,129
$\beta_{W3,2C}$	-0,036	0,016	-2,24	0,025

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{W3,2D}$	-0,018	0,014	-1,24	0,215
$\beta_{W3,2E}$	-0,024	0,013	-1,87	0,062
$\beta_{W3,3}$	0,010	0,012	0,82	0,415
$\beta_{W3,4}$	-0,010	0,013	-0,79	0,429
$\beta_{W3,5}$	0,000	0,012	0,04	0,969
$\beta_{W3,6}$	-0,002	0,014	-0,13	0,893
$\beta_{W3,8}$	-0,001	0,018	-0,08	0,936
$\beta_{W3,9}$	0,005	0,019	0,28	0,778
$\beta_{W3,10}$	-0,037	0,015	-2,50	0,012
$\beta_{W3,11}$	0,004	0,013	0,26	0,792
$\beta_{W3,13}$	-0,018	0,014	-1,29	0,198
$\beta_{W3,14}$	0,017	0,015	1,15	0,250
$\beta_{W3,15}$	0,054	0,012	4,68	0,000
$\beta_{W3,16}$	0,020	0,012	1,73	0,084
$\beta_{W3,17}$	-0,019	0,013	-1,53	0,127
$\beta_{W3,19A}$	-0,014	0,019	-0,74	0,457
$\beta_{W3,19B}$	-0,006	0,020	-0,30	0,763
$\beta_{W3,19C}$	0,029	0,023	1,30	0,195
$\beta_{W3,19D}$	0,005	0,023	0,20	0,844
$\beta_{W3,19E}$	0,011	0,016	0,68	0,496
$\beta_{W3,20}$	0,012	0,019	0,63	0,527
$\beta_{W3,21}$	-0,031	0,016	-1,96	0,050
$\beta_{W3,22}$	-0,035	0,013	-2,82	0,005
$\beta_{W4,1A}$	-0,026	0,010	-2,62	0,009
$\beta_{W4,1B}$	0,149	0,014	10,40	0,000
$\beta_{W4,1C}$	0,045	0,014	3,24	0,001
$\beta_{W4,2A}$	0,024	0,010	2,44	0,015
$\beta_{W4,2B}$	0,069	0,013	5,43	0,000
$\beta_{W4,2C}$	0,085	0,015	5,82	0,000
$\beta_{W4,2D}$	0,012	0,015	0,80	0,426
$\beta_{W4,2E}$	-0,004	0,014	-0,31	0,758
$\beta_{W4,3}$	-0,035	0,010	-3,45	0,001
$\beta_{W4,4}$	-0,003	0,011	-0,28	0,779
$\beta_{W4,5}$	0,006	0,011	0,52	0,600
$\beta_{W4,6}$	0,137	0,012	11,21	0,000
$\beta_{W4,8}$	0,068	0,015	4,50	0,000
$\beta_{W4,9}$	0,080	0,015	5,19	0,000

Współczynnik ścieżki	Oszacowania współczynników ścieżki	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
$\beta_{W4,10}$	0,098	0,012	8,37	0,000
$\beta_{W4,11}$	-0,007	0,013	-0,53	0,599
$\beta_{W4,13}$	0,013	0,012	1,11	0,266
$\beta_{W4,14}$	-0,012	0,012	-0,99	0,322
$\beta_{W4,15}$	0,020	0,009	2,19	0,029
$\beta_{W4,16}$	0,016	0,011	1,50	0,135
$\beta_{W4,17}$	0,142	0,012	12,09	0,000
$\beta_{W4,19A}$	0,029	0,017	1,73	0,084
$\beta_{W4,19B}$	0,005	0,019	0,28	0,780
$\beta_{W4,19C}$	0,002	0,020	0,10	0,924
$\beta_{W4,19D}$	0,021	0,020	1,01	0,313
$\beta_{W4,19E}$	-0,034	0,016	-2,15	0,032
$\beta_{W4,20}$	0,031	0,017	1,86	0,062
$\beta_{W4,21}$	-0,011	0,014	-0,75	0,453
$\beta_{W4,22}$	-0,076	0,010	-7,29	0,000

Tabela B.III.8.

Oszacowania parametrów modelu regresji probitowej wyróżnionych zmiennych dla wykształcenia podstawowego/gimnazjalnego

Zmienne	Oszacowania współczynników regresji	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
Poziom wykształcenia rodziców – X_1	0,244	0,072	3,37	0,001
Wzorce kulturowe – X_2	0,002	0,052	0,04	0,966
Wyznaczniki sukcesu w domu rodzinnym:				
rodzina – X_3	-0,125	0,051	-2,43	0,015
prestżowa praca – X_4	-0,107	0,046	-2,30	0,022
wysoki status materialny – X_5	0,062	0,058	1,06	0,288
dobre wykształcenie – X_6	0,096	0,053	1,83	0,067
Motywacje wyborów edukacyjnych:				
odległość – X_7				
prestż szkoły – X_8	0,060	0,033	1,83	0,067
jakość nauczania – X_9	-0,200	0,053	-3,76	0,000
profil związany z zainteresowaniami – X_{10}	0,049	0,054	0,92	0,358
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}	0,099	0,035	2,82	0,005
	0,159	0,033	4,81	0,000

Zmienne	Oszacowania współczynników regresji	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
Charakterystyka domu rodzinnego:				
silne więzi rodzinne – X_{13}	0,079	0,046	1,72	0,085
wspólne spędzanie wolnego czasu – X_{14}	0,037	0,046	0,80	0,424
nacisk na pracę i naukę – X_{15}	0,080	0,040	1,99	0,046
Uczęszczanie do przedszkola – X_{16}	0,102	0,046	2,21	0,027
Uczęszczanie na zajęcia dodatkowe – X_{17}	-0,024	0,050	-0,48	0,632
Klasa miejscowości zamieszkania – X_{19}	0,002	0,052	0,03	0,974
Poziom rozwoju województwa zamieszkania – X_{20}	0,006	0,053	0,11	0,914
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	-0,018	0,056	-0,33	0,744
Wczesne rodzicielstwo – X_{22}	0,125	0,045	2,75	0,006
Zamożność domu rodzinnego	-0,062	0,034	-1,83	0,067
Zainteresowanie rodziców edukacją	0,031	0,034	0,93	0,352
Wiek	0,189	0,058	3,25	0,001
Niepełnosprawność lub choroba przewlekła	-0,218	0,040	-5,44	0,000
Stała	2,161	0,115	18,74	0,000

Tabela B.III.9.

Oszacowania parametrów modelu regresji probitowej wyróżnionych zmiennych dla wykształcenia ponadgimnazjalnego bez matury

Zmienne	Oszacowania współczynników regresji	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
Poziom wykształcenia rodziców – X_1	0,276	0,050	5,55	0,000
Wzorce kulturowe – X_2	0,176	0,039	4,46	0,000
Wyznaczniki sukcesu w domu rodzinnym:				
rodzina – X_3	-0,068	0,031	-2,16	0,031
prestżowa praca – X_4	-0,032	0,031	-1,04	0,298
wysoki status materialny – X_5	0,043	0,032	1,35	0,177
dobrze wykształcenie – X_6	0,136	0,038	3,54	0,000

Zmienne	Oszacowania współczynników regresji	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
Motywacje wyborów edukacyjnych:				
odległość – X_7	0,016	0,016	0,77	0,439
prestż szkoły – X_8	-0,131	-0,131	-3,31	0,001
jakość nauczania – X_9	0,032	0,032	0,82	0,411
profil związany z zainteresowaniami – X_{10}	0,251	0,251	10,48	0,000
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}	0,253	0,253	10,30	0,000
Charakterystyka domu rodzinnego:				
silne więzi rodzinne – X_{13}	0,042	0,030	1,40	0,162
wspólne spędzanie wolnego czasu – X_{14}	0,004	0,030	0,12	0,906
nacisk na pracę i naukę – X_{15}	0,098	0,026	3,73	0,000
Uczęszczanie do przedszkola – X_{16}	0,024	0,030	0,80	0,422
Uczęszczanie na zajęcia dodatkowe – X_{17}	0,052	0,039	1,33	0,182
Klasa miejscowości zamieszkania – X_{19}	0,036	0,035	1,03	0,303
Poziom rozwoju województwa zamieszkania – X_{20}	0,033	0,041	0,80	0,424
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	-0,009	0,036	-0,25	0,804
Wczesne rodzicielstwo – X_{22}	0,079	0,032	2,44	0,015
Zamożność domu rodzinnego	-0,045	0,029	-1,56	0,119
Zainteresowanie rodziców edukacją	0,077	0,026	3,00	0,003
Wiek	0,083	0,037	2,22	0,027
Niepełnosprawność lub choroba przewlekła	-0,221	0,030	-7,30	0,000
Stała	1,776	0,079	22,53	0,000

Tabela B.III.10.

Oszacowania parametrów modelu regresji probitowej wyróżnionych zmiennych dla wykształcenia średniego z maturą

Zmienne	Oszacowania współczynników regresji	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
Poziom wykształcenia rodziców – X_1	0,409	0,026	15,46	0,000
Wzorce kulturowe – X_2	0,192	0,025	7,61	0,000
Wyznaczniki sukcesu w domu rodzinnym:				
rodzina – X_3	-0,038	0,021	-1,83	0,067
prestżowa praca – X_4	-0,044	0,020	-2,18	0,029
wysoki status materialny – X_5	0,006	0,020	0,28	0,779
dobre wykształcenie – X_6	0,235	0,022	10,48	0,000
Motywacje wyborów edukacyjnych:				
odległość – X_7	-0,060	0,016	-3,70	0,000
prestż szkoły – X_8	0,104	0,032	3,29	0,001
jakość nauczania – X_9	0,162	0,032	5,06	0,000
profil związany z zainteresowaniami – X_{10}	0,085	0,023	3,68	0,000
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}	-0,003	0,023	-0,15	0,881
Charakterystyka domu rodzinnego:				
silne więzi rodzinne – X_{13}	-0,025	0,025	-0,98	0,325
wspólne spędzanie wolnego czasu – X_{14}	0,010	0,025	0,40	0,690
nacisk na pracę i naukę – X_{15}	0,099	0,020	4,96	0,000
Uczęszczanie do przedszkola – X_{16}	0,028	0,019	1,44	0,150
Uczęszczanie na zajęcia dodatkowe – X_{17}	0,288	0,024	11,95	0,000
Klasa miejscowości zamieszkania – X_{19}	0,028	0,022	1,24	0,213
Poziom rozwoju województwa zamieszkania – X_{20}	0,082	0,030	2,76	0,006
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	-0,079	0,025	-3,18	0,001
Wczesne rodzicielstwo – X_{22}	0,137	0,021	6,60	0,000
Zamożność domu rodzinnego	-0,020	0,020	-1,00	0,318
Zainteresowanie rodziców edukacją	-0,003	0,021	-0,14	0,889
Wiek	-0,094	0,023	-3,99	0,000
Niepełnosprawność lub choroba przewlekła	-0,191	0,028	-6,70	0,000
Stała	0,649	0,055	11,78	0,000

Tabela B.III.11.

Oszacowania parametrów modelu regresji probitowej wyróżnionych zmiennych dla wykształcenia wyższego

Zmienne	Oszacowania współczynników regresji	Błędy standardowe szacunku	Statystyka t-Studenta	Krytyczny poziom istotności
Poziom wykształcenia rodziców – X_1	0,355	0,024	14,89	0,000
Wzorce kulturowe – X_2	0,212	0,024	8,89	0,000
Wyznaczniki sukcesu w domu rodzinnym:				
rodzina – X_3	-0,070	0,020	-3,40	0,001
prestżowa praca – X_4	-0,014	0,020	-0,71	0,477
wysoki status materialny – X_5	0,013	0,019	0,71	0,476
dobre wykształcenie – X_6	0,207	0,020	10,31	0,000
Motywacje wyborów edukacyjnych:				
odległość – X_7	-0,049	0,016	-3,08	0,000
prestż szkoły – X_8	0,141	0,029	4,80	0,000
jakość nauczania – X_9	0,149	0,031	4,75	0,000
profil związany z zainteresowaniami – X_{10}	0,172	0,025	6,89	0,000
możliwość znalezienia dobrej pracy po zakończeniu szkoły – X_{11}	-0,012	0,025	-0,49	0,627
Charakterystyka domu rodzinnego:				
silne więzi rodzinne – X_{13}	0,021	0,027	0,78	0,438
wspólne spędzanie wolnego czasu – X_{14}	-0,026	0,028	-0,93	0,351
nacisk na pracę i naukę – X_{15}	0,056	0,021	2,64	0,008
Uczęszczanie do przedszkola – X_{16}	0,007	0,020	0,36	0,717
Uczęszczanie na zajęcia dodatkowe – X_{17}	0,223	0,020	11,36	0,000
Klasa miejscowości zamieszkania – X_{19}	-0,013	0,021	-0,62	0,536
Poziom rozwoju województwa zamieszkania – X_{20}	0,053	0,030	1,79	0,074
Czynniki instytucjonalne wspierające edukację w województwie zamieszkania – X_{21}	-0,039	0,025	-1,57	0,116
Wczesne rodzicielstwo – X_{22}	0,192	0,020	9,53	0,000
Zamożność domu rodzinnego	-0,025	0,019	-1,30	0,195
Zainteresowanie rodziców edukacją	-0,021	0,021	-1,02	0,309
Wiek	0,076	0,023	3,31	0,001
Niepełnosprawność lub choroba przewlekła	-0,151	0,033	-4,54	0,000
Stała	-0,372	0,053	-6,97	0,000

B.IV.Blok IV: Parametryzacja modelu wieloagentowego

B.IV.1. Wartości skalibrowanych parametrów symulacyjnego wieloagentowego modelu rynku edukacji wyższej

Dane z przeprowadzonych ankiet posłużyły do przeprowadzenia kalibracji parametrów w sposób zgodny z opisaną metodologią oraz wynikami analiz. Poniżej przedstawiono wykaz skalibrowanych wartości. Zgodnie z podanymi założeniami model symulacyjny został skonstruowany w języku Java z zastosowaniem biblioteki symulacyjnej MASON (por. Luke i in., 2004). Dlatego wartości skalibrowanych zmiennych zostały przedstawione wraz z typami zmiennych w języku Java.

```
//odsetek czasu poświęcanego na pracę w trakcie
//studiów niestacjonarnych
//te zarobki mogą być przeznaczone na sfinansowanie czesnego
double partTimeWorkShare = 0.5;

//wiek posiadania potomstwa – zakres min
int parentsAgeMin = 20;
//wiek posiadania potomstwa – zakres max
int parentsAgeMax = 41;

//maksymalny wiek – indywidua powyżej tego wieku są usuwane
//z modelu i w ich miejsce rodzą się nowe
//z analizy wyłączono system emerytalny
int maxAge = 65;

//liczba indywiduów w danej grupie wiekowej
int ageGroupSize = 200;
//liczba indywiduuów = ageGroupSize * maxAge

//odsetek absolwentów studiów w populacji początkowej
double probInitialAgentHasGraduated = 0.8;

//typy uczelni/sposobów kształcenia w modelu
// w języku Java stosuje się dla list numerowanie
//typu 0-based (zaczynając do zera).
int[] universityTypes = {
    0, //0 = publiczny stacjonarny (x1)
    1, //1 = publiczny niestacjonarny (x2)
    2, //2 = prywatny stacjonarny (x3)
    3 //3 = prywatny niestacjonarny (x4)

};

//liczba miejsc dla studentów na poszczególnych uczelniach
int[] universityCapacities = {60,60,500,500};
```

```

//przyjęto log-normalny rozkład wynagrodzeń
//osoby bez wyższego wykształcenia wartość oczekiwana
double wagesLnEv = 7.45;
//osoby bez ez wyższego wykształcenia odch stand.
double wagesLnStdDev = 0.60;

// rozkład wynagrodzeń absolwentów poszczególnych uczelni

//efekt multiplikatywny dla poszczególnych uczelni
double[] universityWagesLnEv
= {7.680, 7.600, 7.530, 7.530 };
double[] universityWagesLnStdDev
= {0.700, 0.675, 0.650, 0.625 };
//efekt addytywny dla poszczególnych uczelni
double [] universityWagesAdditiveEffectEv
= {750, 500, 300, 200 };
double [] universityWagesAdditiveEffectStd
= {250, 166, 100, 66 };

//czesne
double[] universityFee = {2000, 2000, 2000, 2000};
//stałe koszty studiowania (niezależne od czesnego)
double[] universityConstCost = {500, 500, 500, 500};

//czas na początkowy rozruch modelu
//(do momentu stabilizacji jego parametrów)
double financingRegimeChangeTime = 400;

//sposób finansowanie czesnego przez regulatora
//w okresie początkowym
//wartości określone dla poszczególnych typów uczelni
double[] typeFinancing_0 = {1.0, 0.0 ,0.0 ,0.0};
//sposób finansowania czesnego przez regulatora po
//zmianie reżimu
double[] typeFinancing_t = typeFinancing_0;

//poziom dochodów gospodarstwa domowego kwalifikujący
//do przyznania stypendium socjalnego
double socialStipendIncomeLevel = 3000;

//stypendia socjalne przeznaczone na sfinansowanie
//kształcenia – określone dla poszczególnych typów uczelni
double[] typeSocialStipend_0 = {0, 0 ,0 ,0};
//sposób finansowania czesnego przez regulatora
//po zmianie reżimu regulacyjnego
double[] typeSocialStipend_t = typeSocialStipend_0;

//stałe koszty utrzymania gospodarstwa domowego
double fixedFamilyCosts = 1000;

```

```

//Korelacja zdolności dzieci i rodziców
double parentsChildrenIqCorr = 0.7;
//korelacja zdolności i początkowego kapitału ludzkiego
double iqHqCorr = 0.7;

//stopa dyskontowa dla obliczania opłacalności (npv)
//decyzji o podjęciu studiów
private double npv_wacc = 0.03;

```

B.IV.2. Rozważane scenariusze eksperymentów symulacyjnych

W modelu symulacyjnym są rozważane następujące poziomy finansowania czesnego:

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
1	1	1	1
0.25	0	0	0
0.25	0	0	0.25
0.25	0	0.25	0
0.25	0	0.25	0.25
0.25	0.25	0	0
0.25	0.25	0	0.25
0.25	0.25	0.25	0
0.25	0.25	0.25	0.25
0.5	0	0	0
0.5	0	0	0.25
0.5	0	0	0.5
0.5	0	0.25	0
0.5	0	0.25	0.25
0.5	0	0.25	0.5
0.5	0	0.5	0
0.5	0	0.5	0.25
0.5	0	0.5	0.5
0.5	0.25	0	0
0.5	0.25	0	0.25
0.5	0.25	0	0.5
0.5	0.25	0.25	0
0.5	0.25	0.25	0.25
0.5	0.25	0.25	0.5
0.5	0.25	0.5	0

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
0.5	0.25	0.5	0.25
0.5	0.25	0.5	0.5
0.5	0.5	0	0
0.5	0.5	0	0.25
0.5	0.5	0	0.5
0.5	0.5	0.25	0
0.5	0.5	0.25	0.25
0.5	0.5	0.25	0.5
0.5	0.5	0.5	0
0.5	0.5	0.5	0.25
0.5	0.5	0.5	0.5
0.75	0	0	0
0.75	0	0	0.25
0.75	0	0	0.5
0.75	0	0	0.75
0.75	0	0.25	0
0.75	0	0.25	0.25
0.75	0	0.25	0.5
0.75	0	0.25	0.75
0.75	0	0.5	0
0.75	0	0.5	0.25
0.75	0	0.5	0.5
0.75	0	0.5	0.75
0.75	0	0.75	0
0.75	0	0.75	0.25
0.75	0	0.75	0.5
0.75	0	0.75	0.75
0.75	0.25	0	0
0.75	0.25	0	0.25
0.75	0.25	0	0.5
0.75	0.25	0	0.75
0.75	0.25	0.25	0
0.75	0.25	0.25	0.25
0.75	0.25	0.25	0.5
0.75	0.25	0.25	0.75
0.75	0.25	0.5	0

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
0.75	0.25	0.5	0.25
0.75	0.25	0.5	0.5
0.75	0.25	0.5	0.75
0.75	0.25	0.75	0
0.75	0.25	0.75	0.25
0.75	0.25	0.75	0.5
0.75	0.25	0.75	0.75
0.75	0.5	0	0
0.75	0.5	0	0.25
0.75	0.5	0	0.5
0.75	0.5	0	0.75
0.75	0.5	0.25	0
0.75	0.5	0.25	0.25
0.75	0.5	0.25	0.5
0.75	0.5	0.25	0.75
0.75	0.5	0.5	0
0.75	0.5	0.5	0.25
0.75	0.5	0.5	0.5
0.75	0.5	0.5	0.75
0.75	0.5	0.75	0
0.75	0.5	0.75	0.25
0.75	0.5	0.75	0.5
0.75	0.5	0.75	0.75
0.75	0.75	0	0
0.75	0.75	0	0.25
0.75	0.75	0	0.5
0.75	0.75	0	0.75
0.75	0.75	0.25	0
0.75	0.75	0.25	0.25
0.75	0.75	0.25	0.5
0.75	0.75	0.25	0.75
0.75	0.75	0.5	0
0.75	0.75	0.5	0.25
0.75	0.75	0.5	0.5
0.75	0.75	0.5	0.75
0.75	0.75	0.75	0

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
0.75	0.75	0.75	0.25
0.75	0.75	0.75	0.5
0.75	0.75	0.75	0.75
1	0	0	0
1	0	0	0.25
1	0	0	0.5
1	0	0	0.75
1	0	0	1
1	0	0.25	0
1	0	0.25	0.25
1	0	0.25	0.5
1	0	0.25	0.75
1	0	0.25	1
1	0	0.5	0
1	0	0.5	0.25
1	0	0.5	0.5
1	0	0.5	0.75
1	0	0.5	1
1	0	0.75	0
1	0	0.75	0.25
1	0	0.75	0.5
1	0	0.75	0.75
1	0	0.75	1
1	0	1	0
1	0	1	0.25
1	0	1	0.5
1	0	1	0.75
1	0	1	1
1	0.25	0	0
1	0.25	0	0.25
1	0.25	0	0.5
1	0.25	0	0.75
1	0.25	0	1
1	0.25	0.25	0
1	0.25	0.25	0.25
1	0.25	0.25	0.5

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
1	0.25	0.25	0.75
1	0.25	0.25	1
1	0.25	0.5	0
1	0.25	0.5	0.25
1	0.25	0.5	0.5
1	0.25	0.5	0.75
1	0.25	0.5	1
1	0.25	0.75	0
1	0.25	0.75	0.25
1	0.25	0.75	0.5
1	0.25	0.75	0.75
1	0.25	0.75	1
1	0.25	1	0
1	0.25	1	0.25
1	0.25	1	0.5
1	0.25	1	0.75
1	0.25	1	1
1	0.5	0	0
1	0.5	0	0.25
1	0.5	0	0.5
1	0.5	0	0.75
1	0.5	0	1
1	0.5	0.25	0
1	0.5	0.25	0.25
1	0.5	0.25	0.5
1	0.5	0.25	0.75
1	0.5	0.25	1
1	0.5	0.5	0
1	0.5	0.5	0.25
1	0.5	0.5	0.5
1	0.5	0.5	0.75
1	0.5	0.5	1
1	0.5	0.75	0
1	0.5	0.75	0.25
1	0.5	0.75	0.5
1	0.5	0.75	0.75

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
1	0.5	0.75	1
1	0.5	1	0
1	0.5	1	0.25
1	0.5	1	0.5
1	0.5	1	0.75
1	0.5	1	1
1	0.75	0	0
1	0.75	0	0.25
1	0.75	0	0.5
1	0.75	0	0.75
1	0.75	0	1
1	0.75	0.25	0
1	0.75	0.25	0.25
1	0.75	0.25	0.5
1	0.75	0.25	0.75
1	0.75	0.25	1
1	0.75	0.5	0
1	0.75	0.5	0.25
1	0.75	0.5	0.5
1	0.75	0.5	0.75
1	0.75	0.5	1
1	0.75	0.75	0
1	0.75	0.75	0.25
1	0.75	0.75	0.5
1	0.75	0.75	0.75
1	0.75	0.75	1
1	0.75	1	0
1	0.75	1	0.25
1	0.75	1	0.5
1	0.75	1	0.75
1	0.75	1	1
1	1	0	0
1	1	0	0.25
1	1	0	0.5
1	1	0	0.75
1	1	0	1

Publiczne Stacjonarne	Publiczne Niestacjonarne	Niepubliczne Stacjonarne	Niepubliczne Niestacjonarne
Poziom finansowania x1	Poziom finansowania x2	Poziom finansowania x3	Poziom finansowania x4
1	1	0.25	0
1	1	0.25	0.25
1	1	0.25	0.5
1	1	0.25	0.75
1	1	0.25	1
1	1	0.5	0
1	1	0.5	0.25
1	1	0.5	0.5
1	1	0.5	0.75
1	1	0.5	1
1	1	0.75	0
1	1	0.75	0.25
1	1	0.75	0.5
1	1	0.75	0.75
1	1	0.75	1
1	1	1	0
1	1	1	0.25
1	1	1	0.5
1	1	1	0.75

Instytut Badań Edukacyjnych

Głównym zadaniem Instytutu jest prowadzenie badań, analiz i prac przydatnych w rozwoju polityki i praktyki edukacyjnej.

Instytut zatrudnia ponad 150 badaczy zajmujących się edukacją – pedagogów, socjologów, psychologów, ekonomistów, politologów i przedstawicieli innych dyscyplin naukowych – wybitnych specjalistów w swoich dziedzinach, o różnorodnych doświadczeniach zawodowych, które obejmują, oprócz badań naukowych, także pracę dydaktyczną, doświadczenie w administracji publicznej czy działalność w organizacjach pozarządowych.

Instytut w Polsce uczestniczy w realizacji międzynarodowych projektów badawczych w tym PIAAC, PISA, TALIS, ESLC, SHARE, TIMSS i PIRLS oraz projektów systemowych współfinansowanych przez Unię Europejską ze środków Europejskiego Funduszu Społecznego.



www.ibe.edu.pl

Instytut Badań Edukacyjnych

ul. Górczewska 8, 01-180 Warszawa | tel. +48 22 241 71 00 | ibe@ibe.edu.pl | www.ibe.edu.pl

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.