

Wpływ zróżnicowania wieku uczniów na szacowanie gimnazjalnych wskaźników edukacyjnej wartości dodanej

ALEKSANDRA JASIŃSKA-MACIĄŻEK

Instytut Badań Edukacyjnych*

Badania pokazują, że nie tylko poziom osiągnięć szkolnych, lecz także ich przyrost zależy od wieku uczniów. Znaczenie ma zarówno miesiąc urodzenia osób z tego samego rocznika, jak również fakt bycia młodszym lub starszym od dominującej grupy wiekowej. Wskaźniki edukacyjnej wartości dodanej (EWD) wykorzystywane są w ewaluacji pracy szkół. Istotą ich szacowania jest pomiar względnych przyrostów osiągnięć. W polskich modelach EWD nie wykorzystuje się jednak informacji o wieku uczniów, przez co wyliczone wskaźniki mogą być obciążone. W artykule przedstawiono wyniki badania szacującego wielkość tego obciążenia. Do analiz wykorzystano dane egzaminacyjne polskich gimnazjalistów z czterech sesji egzaminacyjnych (przeprowadzonych w latach szkolnych 2009/2010–2012/2013). Wyniki pokazały, że uwzględnienie w modelach EWD informacji o wieku uczniów w bardzo niewielkim stopniu zmienia oszacowania wskaźników, czyli że wspomniane obciążenie było w ostatnich latach znikome.

SŁOWA KLUCZOWE: edukacyjna wartość dodana, trafność, wiek uczniów, wcześniejszy start szkolny, opóźniony start szkolny, powtarzanie klasy, przyrost osiągnięć szkolnych.

System egzaminów zewnętrznych istnieje w Polsce od 2002 r. Obok oceniania stopnia opanowania przez uczniów wymagań programowych na potrzeby celów selekcyjnych czy diagnozy indywidualnej oraz monitorowania procesów zachodzących w systemie oświaty, pełni on także funkcję ewaluacyjną (Dolata i Szaleniec, 2012). Wykorzystywanie wyników egzaminacyjnych do ewaluacji wewnętrznej i zewnętrznej

wymaga jednak właściwego ich przekształcenia, ponieważ liczba punktów zdobytych w teście w dużej mierze zależy od cech, na które szkoła nie ma wpływu, takich jak: inteligencja uczniów, uprzednie osiągnięcia i doświadczenia szkolne, status społeczno-ekonomiczny rodziny, aspiracje rodziców, cechy środowiska lokalnego (Dolata i in., 2013; 2014). Jedną z metod, która umożliwia wydzielenie z wyników egzaminacyjnych tego, co od szkoły zależy, jest metoda edukacyjnej wartości dodanej (EWD).

Modele EWD służą do oszacowania względnych postępów uczniów w danym cyklu kształcenia. Co ważne – postępów,

Artykuł powstał w ramach projektu „Rozwój metody edukacyjnej wartości dodanej na potrzeby wzmocnienia ewaluacyjnej funkcji egzaminów zewnętrznych” realizowanego w Instytucie Badań Edukacyjnych i współfinansowanego ze środków Europejskiego Funduszu Społecznego w ramach Programu Operacyjnego Kapitał Ludzki 2007–2013. Priorytet III: Wysoka jakość systemu oświaty.

© Instytut Badań Edukacyjnych

* Adres: Pracownia Edukacyjnej Wartości Dodanej, ul. Górczewska 8, 01-180 Warszawa. E-mail: a.jasinska@ibe.edu.pl

które są związane z pracą szkoły. Podstawą do ich wyliczenia są co najmniej dwa pomiary osiągnięć szkolnych: pierwszy, przeprowadzony na początku, drugi – pod koniec okresu nauczania, będącego przedmiotem zainteresowania (OECD, 2008). Modele EWD mogą się jednak różnić zakresem dodatkowych zmiennych (tzw. zmiennych kontekstowych), które są w nich uwzględniane. Najprostsze modele EWD nie uwzględniają żadnych zmiennych kontekstowych, wykorzystują jedynie wyniki pomiaru osiągnięć szkolnych. W tych modelach zakłada się, że wpływ ważnych czynników niezależnych od szkoły na osiągane wyniki, jest kontrolowany przez uwzględnienie informacji o uprzednich osiągnięciach ucznia. Założenie to jest słuszne w przypadku czynników, które mają znaczenie przede wszystkim dla poziomu osiągnięć, a nie ich przyrostów, a także czynników, których oddziaływanie nie zmienia się w czasie między pomiarami osiągnięć.

Potrzeba włączenia do modelu zmiennych kontekstowych wynika z przeświadczenia, że w ten sposób można lepiej oczyścić wskaźnik EWD z wpływu tych zmiennych, które są niezależne od szkoły. Jeśli wartość wskaźnika EWD, rozumianego jako miara efektywności nauczania, byłaby powiązana z tym, na co szkoła nie ma wpływu, należałoby stwierdzić, że wskaźnik taki jest obciążony, częściowo bowiem zdaje sprawę z tego, czego z założenia nie miał mierzyć.

Można wyróżnić dwie grupy modeli, które uwzględniają dodatkowe zmienne (Raudenbush i Willms, 1995). Pierwsza grupa modeli uwzględnia tylko cechy indywidualne uczniów, takie jak np. płeć, wiek, pochodzenie społeczne. Pokazują one, w której szkole uczeń ma największe szanse na osiągnięcie wysokiego wyniku w teście, niezależnie od tego, w jakim stopniu będzie on efektem działania szkoły, a w jakim – korzystnego środowiska, w którym uczeń pracuje. Modele z drugiej grupy uwzględniają dodatkowo zmienne kontekstowe,

charakteryzujące otoczenie szkoły (np. średni poziom osiągnięć uczniów w szkole czy poziom bezrobocia w gminie). W modelach EWD nie uwzględnia się natomiast zmiennych charakteryzujących np. wysiłek nauczycieli czy organizację procesu nauczania. One bowiem składają się na istotę tego, co w założeniu mają mierzyć te wskaźniki.

W polskich modelach EWD¹, obok danych o uprzednich osiągnięciach, uwzględnia się informację o płci, posiadaniu zaświadczenia o dysleksji oraz o standardowym lub wydłużonym toku kształcenia (Żółtak, 2013). Nie włącza się do modeli informacji o wieku uczniów, chociaż badania dowodzą, że ma on znaczenie nie tylko dla poziomu osiągnięć, lecz także dla ich przyrostu (Ding i Davison, 2005; Hutchison i Sharp, 1999; Wen, Bulotsky-Shearer, Hahs-Vaughn i Korfmacher, 2012). Na potrzebę uwzględniania w modelach EWD informacji o wieku uczniów zwracał uwagę już kilka lat temu międzynarodowy zespół ekspertów opracowujący raport o metodzie EWD (OECD, 2008). Doświadczenia niektórych krajów także wskazywały na potrzebę włączenia tej zmiennej do modeli EWD (Hægeland, Kirkeboen, Raaum i Salvanes, 2005; Ray, 2006). Przykładem modeli, w których uwzględniono zmienną opisującą wiek uczniów, są angielskie kontekstowe modele edukacyjnej wartości dodanej² (Ray, McCormack i Evans, 2009).

Podstawowym celem tego artykułu jest odpowiedź na pytanie, czy nieuwzględnienie w polskich modelach EWD informacji o wieku uczniów obciąża wskaźniki wyliczane dla szkół. Jeśli tak, to jak duży jest to

¹ Modele te są liczone dla gimnazjów, liceów i techników, z wykluczeniem m.in. szkół dla dorosłych, szkół specjalnych, przyszpitalnych, znajdujących się przy ośrodkach wychowawczych.

² Informacja ta, obok wielu innych zmiennych kontekstowych, była włączana do modelu jako względny wiek ucznia w miesiącach w obrębie roku urodzenia. Nie uwzględniano dodatkowych zmiennych na oznaczenie uczniów starszych i młodszych od głównej kohorty, ponieważ uznano, że jest ich zbyt mało.

problem, i czy wobec tego warto uwzględniać w modelach EWD tę informację? W prezentowanych analizach skupiono się na wskaźnikach dla gimnazjów.

Podjęty w artykule problem ma także znaczenie bardziej ogólne. Zaprezentowane wyniki pokazują, jak wygląda zależność przyrostów osiągnięć szkolnych od wieku uczniów.

Źródła zróżnicowania wieku uczniów na poziomie danej klasy

Stwierdzenie, że uczniowie tej samej klasy różnią się wiekiem, może wydawać się zaskakujące. Prawo oświatowe wyznacza bowiem wiek rozpoczęcia nauki szkolnej na rok kalendarzowy, w którym dziecko kończy określoną liczbę lat, a uczniowie uczą się w klasach utworzonych na podstawie kryterium roku urodzenia. Tymczasem mamy do czynienia z czterema głównymi źródłami wariacji wieku biologicznego uczniów z tych samych klas: (a) wcześniejszy start szkolny; (b) odroczenie startu szkolnego; (c) brak promocji do kolejnej klasy; (d) przyjęta definicja rozpoczęcia realizacji obowiązku szkolnego, która skutkuje zróżnicowaniem wieku uczniów sięgającym 12 miesięcy.

Wcześniejsze rozpoczęcie nauki szkolnej przez dziecko jest możliwe za zgodą dyrektora (na wniosek rodziców). Jest ona uzależniona od opinii poradni psychologiczno-pedagogicznej wydanej na podstawie diagnozy dojrzałości psychofizycznej dziecka (art. 16 ust 2 ustawy z dnia 7 września 1991 r. o systemie oświaty). Decyzja o wcześniejszym posłaniu dziecka do szkoły jest więc związana z czynnikami, które mają duże znaczenie dla jego przyszłych wyników. Ta pozytywna selekcja może prowadzić do obserwowania średnio wyższych osiągnięć szkolnych w grupie uczniów młodszych niż główna kohorta wiekowa.

Zjawiskiem odwrotnym jest odroczenie startu szkolnego. Jest ono możliwe

w przypadkach „uzasadnionych ważnymi przyczynami” i może nastąpić tylko na rok³ (art. 16, ust. 3 ustawy). Przyczyny te najczęściej negatywnie wpływają na szansę osiągnięcia sukcesu edukacyjnego (np. brak gotowości szkolnej). Mamy tu więc do czynienia z selekcją negatywną. Uczniowie, którym odroczone start szkolny, są starsi od swoich klasowych kolegów, jednak mają ukończone tyle samo lat nauki szkolnej.

Starsi od głównej kohorty wiekowej będą także uczniowie, którzy w poprzednich latach nauki nie uzyskali promocji do następnej klasy. Nie tylko wiek odróżnia ich od klasowych kolegów i koleżanek, lecz także to, że decyzja o powtarzaniu nauki w tej samej klasie została podjęta na podstawie ich niskich wyników lub zbyt częstej absencji na zajęciach. Prowadzi to do oczywistego związku faktu bycia starszym ze względu na brak promocji z osiągnięciami szkolnymi.

Ostatnie źródło wariacji wieku uczniów to zróżnicowanie w obrębie głównej kohorty. W kontekście analiz związków z osiągnięciami, w odróżnieniu od wyżej omówionych, miesiąc urodzenia można uznać za losowy (Dolata i in., 2013; Dolata i Pokropek, 2012). To, w którym miesiącu urodził się uczeń, raczej nie zależało od jego genów, statusu rodziny itp. Fakt ten będzie miał znaczenie dla interpretacji modelowanych związków. Przedstawione w artykule modele nie pozwolą w przypadku nielosowych źródeł zróżnicowania wieku uczniów (pierwsze trzy wymienione) na interpretację tych związków w kategoriach przyczynowo-skutkowych.

Wiek uczniów a osiągnięcia szkolne

Wiele badań pokazuje, że uczniowie starsi, spośród urodzonych w tym samym roku i uczęszczających do tej samej klasy,

³ Chyba że dziecko ma orzeczenie o potrzebie kształcenia specjalnego. W takiej sytuacji możliwe jest odroczenie obowiązku szkolnego do 10. roku życia.

przewyższają młodszych w poziomie osiągnięć szkolnych (Lee i Fish, 2010; Morrison, Griffith i Alberts, 1997; Sharp, 2002; Sweetland i De Simone, 1987). Efekt ten obserwuje się zarówno w odniesieniu do różnych umiejętności przedmiotowych (Bell i Daniels, 1990; Hutchison i Sharp, 1999; Konarzewski, 2013; Martin, Foels, Clanton i Moon, 2004; Smith, 2009), jak i szerzej rozumianych wyników nauczania, takich jak motywacja, zaangażowanie w naukę lub stosunek do szkoły i nauki (Martin, 2009). Niektórzy badacze skupiają uwagę na sytuacji uczniów najmłodszych w klasie wśród dzieci urodzonych w tym samym roku i wykazują, że częściej doświadczają one niepowodzeń szkolnych, takich jak nieuzyskanie promocji do następnej klasy lub uzyskanie opinii o specjalnych potrzebach edukacyjnych (Langer, Kalk i Searls, 1984; Martin i in., 2004; May i Kundert, 1995; Verachtert, De Fraine, Onghena i Ghesquière, 2010). Wyniki te wspierają hipotezę, że wiek uczniów ma znaczenie dla osiągnięć szkolnych. Można jednak znaleźć takie analizy, które nie potwierdzają istnienia wspomnianej zależności (Dietz i Wilson, 1985). Przegląd badań prowadzonych na różnych poziomach nauczania pokazuje prawidłowość, że w kolejnych latach nauki szkolnej siła korelacji osiągnięć z wiekiem dla głównej kohorty wiekowej spada (Smith, 2009). Sytuacja ta dotyczy wielu krajów (Lee i Fish, 2010). Wyniki badań dość spójnie pokazują, że spadek siły zależności osiągnięć od wieku uczniów jest większy w pierwszych latach nauki, niż kiedy uczniowie są starsi (Bell i Daniels, 1990; Hutchison i Sharp, 1999; Langer i in., 1984; Verachtert i in., 2010). Tym, co różni owe doniesienia, to moment, w którym zależność ta przestaje być istotna. Niektóre publikacje pokazują, że przewaga uczniów starszych nad młodszymi z tej samej klasy znika już po drugiej lub trzeciej klasie (Crone i Whitehurst, 1999; Konarzewski, 2013; Stipek i Byler, 2001), inne donoszą, że efekt ten pozostaje istotny w wieku 10 lat (Smith, 2009), a przestaje mieć znaczenie

dopiero w wieku 12 (Hutchison i Sharp, 1999) lub 17 lat (Langer i in., 1984).

Analizy prowadzone przez badaczy na całym świecie pokazują, że o ile efekt wieku dla uczniów z głównej kohorty jest raczej niewielki, o tyle zróżnicowanie wiekowe wynikające z wcześniejszego lub późniejszego posłania dziecka do szkoły oraz z braku promocji do następnej klasy, jest wyraźnie skorelowane z późniejszymi osiągnięciami. Wynika to z faktu, że przyspieszenie lub odroczenie startu szkolnego najczęściej jest bezpośrednio związane z cechami uczniów mającymi znaczenie dla osiągnięć szkolnych.

Zagadnienia przyspieszonego i opóźnionego startu szkolnego są często rozpatrywane w kontekście problemu gotowości szkolnej (Carlton i Winsler, 1999; Martin, 2009). Zwolennicy odroczenia obowiązku szkolnego widzą w nim szansę na nabycie przez ucznia odpowiedniego poziomu rozwoju poznawczego i społecznego, który umożliwi mu odnoszenie szkolnych sukcesów. Wyniki badań pokazują natomiast, że uczniowie, którzy zostali posłani do szkoły później niż ich rocznikowi rówieśnicy, mają niższe lub porównywalne osiągnięcia szkolne jak koledzy z tej samej klasy (Martin, 2009; May i Kundert, 1995; Morrison i in., 1997). Nie stwierdzono także odroczonego w czasie korzyści edukacyjnych i pozaszkolnych wynikających z opóźnionego startu szkolnego (Cameron i Wilson, 1990; Lincove i Painter, 2006). Mimo że uczniowie ci są starsi niż koledzy z tej samej klasy, pod tym względem nie mają nad nimi przewagi.

Na drugim krańcu tego zjawiska znajdują się uczniowie o przyspieszonym starcie szkolnym. Ci zazwyczaj mają istotnie wyższe wyniki niż ich starsi koledzy z tej samej klasy (Mayer i Knutson, 1999). Wynika to z faktu, że w tej nielicznej grupie wcześniej rozpoczynających naukę szkolną są uczniowie wyselekcjonowani pod względem cech, które są powiązane z późniejszymi sukcesami szkolnymi.

Analizy poświęcone kwestii braku promocji do następnej klasy przemawiają za negatywnym związkiem tego zjawiska z osiągnięciami. Metaanaliza wyników badań z lat 90. XX w., przeprowadzona przez Shane'a Jimersona (2001) pokazała, że uczniowie, którzy powtarzali naukę na którymś etapie, mieli niższe osiągnięcia szkolne niż ich koledzy z tej samej klasy, choć nie wszystkie wyniki analizowanych badań to potwierdziły. Co więcej, według innych badań efekt braku promocji jest negatywny niezależnie od płci, klasy ani wieku ucznia (Martin, 2009).

Omawiane efekty mogą zależeć od stosowanych metod nauczania i przyjętych rozwiązań w danym systemie edukacji. Jednak wyniki badań prowadzonych w Polsce doprowadzają do podobnych wniosków, jak prace zagranicznych naukowców. Dane zebrane podczas badania mającego na celu skonstruowanie Testu umiejętności na starcie szkolnym pokazały, że wśród sześciu- i siedmioletków dzieci starsze osiągały wyższe wyniki na skalach umiejętności czytania, pisanie i matematycznych, niż ich młodsi koledzy będący na tym samym etapie nauczania (Karwowski i Dziedziewicz, 2012). W badaniach standaryzacyjnych Skali gotowości szkolnej stwierdzono istotne efekty dla wieku, ale tylko w przypadku chłopców na dwóch spośród sześciu skal: Samodzielności i Umiejętności szkolnych (Frydrychowicz, Koźniewska, Matuszewski i Zwierzyńska, 2006). Słabe zależności mogą jednak wynikać ze specyfiki zastosowanego narzędzia, które jest przeznaczone dla nauczycieli, a oceny nauczycielskie mają zwykle niższą rzetelność niż standaryzowane testy umiejętności.

W ogólnopolskich badaniach prowadzonych z udziałem uczniów rozpoczynających czwartą klasę szkoły podstawowej również potwierdzono, że starsi spośród urodzonych w tym samym roku, mają większą szansę uzyskania wyższych wyników w testach osiągnięć szkolnych (Dolata i in., 2014;

Jasińska-Maciążek i Modzelewski, 2014). W badaniach tych stwierdzono także silny negatywny efekt dla uczniów starszych niż główna kohorta (wyniki niższe o ponad 2/3 odchylenia standardowego niż w dominującej grupie wiekowej). Natomiast uczniowie młodsi niż główna kohorta uzyskali wyniki na porównywalnym poziomie jak główna grupa wiekowa (Dolata i in., 2014) lub charakteryzowali się trochę wyższym poziomem osiągnięć z zakresu umiejętności matematycznych i świadomości językowej, ale nie umiejętności czytania (Jasińska-Maciążek i Modzelewski, 2014). Opisywane efekty nie były duże, choć silniejsze niż w zależności, które w Polsce obserwuje się w przypadku danych pochodzących ze sprawdzianu po szóstej klasie szkoły podstawowej lub egzaminu gimnazjalnego (Dolata i Pokropek, 2012).

Kierunek związku wieku uczniów z wynikami nauczania zmienia się, jeśli zamiast poziomu osiągnięć weźmiemy pod uwagę ich przyrost. Wyniki badań podłużnych zgodnie pokazują, że wśród uczniów młodszych obserwujemy większe przyrosty osiągnięć szkolnych niż wśród starszych (Ding i Davison, 2005; Hutchison i Sharp, 1999; Wei, Blackorby i Schiller, 2011; Wen i in., 2012), a szczególnie szybkie tempo ich rozwoju stwierdza się w początkowych latach nauki (Ding i Davison, 2005; Wen i in., 2012). Okazuje się zatem, że młodsi uczniowie z czasem doganiają starszych, a przewaga wynikająca z różnicy wieku biologicznego traci na znaczeniu. Wyniki te są spójne ze wspomnianymi rezultatami badań przekrojowych, w których stwierdzono spadek związku osiągnięć szkolnych z wiekiem uczniów w kolejnych latach nauki. W polskich badaniach podłużnych prowadzonych z udziałem gimnazjalistów również zaobserwowano podobną prawidłowość. Ponieważ wykorzystano w nich wyniki dwóch różnych (niezrównanych ze sobą) testów, w modelach skupiono się na względnych przyrostach osiągnięć. Pokazały one, że

wśród uczniów urodzonych w tym samym roku i uczęszczających do tej samej klasy, młodsi uzyskiwali wyższe względne przyrosty osiągnięć szkolnych podczas trzech lat nauki w gimnazjum, choć stwierdzony efekt nie był duży (Dolata i in., 2013).

Uczniowie, którzy później rozpoczęli naukę w gimnazjum, niż wynikałoby to z ich daty urodzenia, są starsi od ich kolegów z głównej kohorty wiekowej. Przyczyny odroczonego startu szkolnego lub nieuzywania promocji do następnej klasy na wcześniejszych etapach nauki mogą być związane z trudnościami w uczeniu się. Czy w grupie tej możemy spodziewać się innej wielkości przyrostów osiągnięć szkolnych niż wśród rówieśników tych uczniów? Niektóre badania wskazują na rekompensujący charakter specjalnych programów edukacyjnych, dzięki którym uczniowie z niekorzystnych środowisk osiągają wyższe przyrosty kompetencji językowych niż inni, mimo niższego początkowego poziomu osiągnięć w tym zakresie (Wen i in., 2012). Inne natomiast pokazują, że uczniowie mający trudności z nauką rozwijają się pod względem osiągnięć szkolnych w podobnym tempie jak ich rówieśnicy bez takich problemów (Ding i Davison, 2005). Wyniki wspomnianych polskich badań pokazały, że uczniowie ci charakteryzują się niższymi względnymi przyrostami osiągnięć szkolnych z przedmiotów humanistycznych niż uczniowie z głównej kohorty wiekowej, choć efekt jest znacznie słabszy niż w przypadku poziomu osiągnięć (Dolata i in., 2013). Względne przyrosty osiągnięć szkolnych z przedmiotów matematycznych były porównywalne w obu grupach.

Próba analizy lub rozstrzygnięcia, jakie mechanizmy są odpowiedzialne za obserwowane związki wieku uczniów z ich osiągnięciami, wykracza poza ramy tego artykułu. Czytelnik zainteresowany takimi dociekaniem znajdzie hipotezy warte rozpatrzenia w pracach innych autorów (Konarzewski, 2013; Martin i in., 2004;

Sykes, Bell i Vidal Rodeiro, 2009; Verachtert i in., 2010).

Przytoczone wyniki badań pokazały, że przyrost osiągnięć szkolnych jest uzależniony od wieku uczniów. Z punktu widzenia szacowania wskaźników EWD dla gimnazjów najważniejsze są pytania o to, jak silną zależność obserwujemy na tym etapie nauczania, oraz czy międzyszkolne zróżnicowanie tej cechy uczniów jest na tyle duże, by przełożyło się na znaczące zmiany we wskaźnikach EWD, jeśli będą one wyliczane z uwzględnieniem informacji o wieku uczniów. Pytania te są przedmiotem analiz przedstawionych w artykule.

Metoda

Wykorzystane dane

Analizy przeprowadzono z wykorzystaniem danych pochodzących z systemu egzaminów zewnętrznych. Wykorzystano połączone dla uczniów wyniki sprawdzianu po szóstej klasie szkoły podstawowej oraz wyniki egzaminu gimnazjalnego. W celu zbadania stabilności oszacowanych efektów w analizach uwzględniono dane dla czterech kolejnych kohort uczniów zdających egzamin gimnazjalny w latach: 2010, 2011, 2012 i 2013. W bazie połączonych wyników egzaminacyjnych znaleźli się uczniowie o standardowym, trzyletnim cyklu kształcenia w gimnazjum oraz o wydłużonym, czteroletnim cyklu kształcenia.

Wyniki egzaminów zewnętrznych zostały wyskalowane w programie Mplus z zastosowaniem dwuparametrycznego modelu IRT dla zadań ocenianych dychotomicznie oraz modelu odpowiedzi stopniowanej (*graded response model*, GRM) dla zadań punktowanych na dłuższych skalach. Zastosowano metodę estymacji brzegowej największej wiarygodności (*marginal maximum likelihood*, MML). Za wskaźnik poziomu osiągnięć szkolnych przyjęto oszacowania *expected*

a posteriori (EAP). W przypadku egzaminu gimnazjalnego osobno skalowano wyniki dla części matematyczno-przyrodniczej (wykorzystano zadania z testu matematyczno-przyrodniczego, a od 2012 r. – z testu z matematyki oraz testu przyrodniczego), osobno z humanistycznej (wykorzystano zadania z testu humanistycznego, a od 2012 r. – z testu z języka polskiego oraz testu wiedzy o społeczeństwie i historii). Wyniki testów przedstawiono na skali o średniej 100 i odchyleniu standardowym 15 w populacji uczniów (zob. Żółtak, 2013).

Zmienna wyjaśniana:

względne przyrosty osiągnięć szkolnych

W celu wyliczenia względnych przyrostów osiągnięć szkolnych w pierwszym kroku oszacowano modele EWD dla części humanistycznej i matematyczno-przyrodniczej dla każdej z czterech analizowanych sesji egzaminacyjnych (łącznie 8 modeli). Wykorzystano dwupoziomowe modele mieszanych efektów, uwzględniające podział na szkoły z efektem losowym dla stałej regresji. Analogiczne modele są liczone na potrzeby publikacji wskaźników EWD w Polsce (Żółtak, 2013)⁴. Wskaźniki EWD szkoły są wyliczane z takich modeli jako Bayesowskie predykcje *a posteriori* realizacji efektu losowego dla stałej, a ich wartość jest interpretowana jako EWD szkoły. Wzór modelu w bardzo ogólnej postaci wygląda następująco:

$$y_{ij} = w^k(x_{ij}) + \beta Z_{ij} + u_j + r_{ij} \quad (1)$$

Zmienną zależną (y_{ij}) jest wynik egzaminu gimnazjalnego z części humanistycznej lub z części matematyczno-przyrodniczej. Podstawową zmienną niezależną jest wynik sprawdzianu (x_{ij}). Relacja między nimi jest

wymodelowana z wykorzystaniem wielomianu k -tego stopnia (w^k), ponieważ nie jest ona prostoliniowa. Parametry wielomianu są osobno szacowane dla uczniów o standardowym (trzyletnim) cyklu kształcenia i osobno dla uczniów o wydłużonym o rok cyklu kształcenia. W modelowaniu uwzględniono także niektóre z dodatkowych zmiennych pochodzących z baz danych egzaminacyjnych: płeć oraz posiadanie zaświadczenia o dysleksji (na sprawdzianie, na egzaminie gimnazjalnym oraz interakcję między nimi). Wektor zmiennych kontrolnych opisujących i -tego ucznia w j -tej szkole został oznaczony jako Z_{ij} , a wektor parametrów (efektów stałych) związanych z tymi zmiennymi – jako β . W równaniu modelu opisano też dwa efekty losowe: realizację efektu losowego dla j -tej szkoły (u_j), która jest interpretowana jako EWD szkoły, oraz błąd losowy z poziomu indywidualnego (r_{ij}), czyli resztę indywidualną i -tego ucznia w j -tej szkole. Oba efekty losowe składają się na tę część wariancji zmiennej zależnej, której nie udało się wyjaśnić przez uwzględnione w modelu zmienne niezależne. Suma tych efektów, będąca resztami z modelu EWD, jest natomiast równoważna różnicy między uzyskanym wynikiem egzaminu (wartością zmiennej zależnej) a wynikiem przewidywanym na podstawie zmiennych niezależnych uwzględnionych w modelu (przewidywaniem z części stałej modelu).

W drugim kroku wyliczono opisane powyżej reszty (suma reszty indywidualnej i reszty z poziomu szkoły) dla każdego z 8 modeli EWD, które mogą być interpretowane jako względny przyrost osiągnięć szkolnych uczniów (czy uczniowie osiągnęli na egzaminie gimnazjalnym wyniki wyższe, czy niższe niż średnio inni uczniowie o takich samych wynikach na sprawdzianie i innych charakterystykach). Są one podstawową zmienną zależną wykorzystywaną w prezentowanych w artykule analizach. Analizy te zostały przeprowadzone w środowisku R, z wykorzystaniem funkcji `lmer` z pakietu `lme4`.

⁴ Z tym zastrzeżeniem, że publicznie dostępne wskaźniki EWD są wyliczane z tzw. modeli trzyletnich, w których uwzględnia się łącznie dane z trzech kolejnych sesji egzaminacyjnych. W przedstawianych tu analizach w jednym modelu wykorzystano dane z jednej sesji egzaminacyjnej.

Strategia, w której zdecydowano się modelować zależność względnych przyrostów osiągnięć od wieku uczniów, a nie wprowadzać bezpośrednio zmienne opisujące wiek uczniów do modeli EWD, została zastosowana z kilku powodów. Po pierwsze dla metody EWD zrozumienie tego, w jaki sposób przyrosty (a nie poziom) osiągnięć szkolnych zależą od zmiennych będących przedmiotem zainteresowania, jest kluczowe, a taki schemat analizy umożliwia bardziej czytelne zobrazowanie tej zależności⁵. Co więcej, z punktu widzenia analizowanego problemu najważniejsze jest to, ile wariacji osiągnięć możemy wyjaśnić przez włączenie do modelu EWD informacji o wieku uczniów ponad to, co jest już wyjaśnione przez zmienne uwzględnione w modelu (m.in. wyniki sprawdzianu po szóstej klasie). Przyjęty schemat analizy umożliwił odpowiedź na tak postawione pytanie. Dodatkowo strategia ta pozwoliła na stworzenie prostszych modeli, co przekłada się na przejrzystość interpretacji.

Zmienne wyjaśniające

Opóźnione rozpoczęcie nauki w gimnazjum. Zmienna ta przyjmuje wartość 1, jeśli uczeń w momencie rozpoczęcia nauki w gimnazjum był starszy o rok lub dwa lata od uczniów z głównej kohorty wiekowej. W pozostałych przypadkach przyjmuje wartość 0. W grupie tej znaleźli się uczniowie, którzy mieli opóźniony start szkolny lub nie uzyskali promocji do kolejnej klasy w szkole podstawowej. Dane nie umożliwiają rozdzielania tych dwóch sytuacji.

Przyspieszone rozpoczęcie nauki w gimnazjum. Zmienna przyjmująca wartość 1, jeśli uczeń w momencie rozpoczęcia nauki w gimnazjum był młodszy o rok od uczniów z głównej kohorty wiekowej. W pozostałych przypadkach przyjmuje wartość 0.

Wiek ucznia. Wiek uczniów został ustalony na podstawie daty urodzenia (z baz danych egzaminacyjnych). Dla tej zmiennej stwierdzono obecność braków danych w połączonych zbiorach z wynikami egzaminacyjnymi. Obserwacje z brakami danych dla daty urodzenia (0,02–0,28% w zależności od kohorty) oraz przypadki odstające (ok. 0,02%) zostały usunięte z analiz. Na podstawie daty urodzenia uczniów utworzono wiele wskaźników opisujących ich wiek, np.:

- Wiek ucznia w miesiącach – wyliczony na podstawie daty urodzenia w taki sposób, że przyjmuje wartość 0 dla uczniów z głównej kohorty wiekowej, o trzyletnim cyklu kształcenia, urodzonych w grudniu (czyli najmłodszych z głównej kohorty).
- Względny wiek ucznia w miesiącach – wyliczony na podstawie daty urodzenia i przekształcony osobno w grupie uczniów o trzyletnim cyklu kształcenia i osobno w grupie uczniów o czteroletnim cyklu kształcenia (czyli starszych o rok), tak aby przyjmował wartość 0 dla uczniów urodzonych w grudniu w głównej kohorcie wiekowej dla trzyletniego i czteroletniego cyklu kształcenia. Wartość wskaźnika pokazuje więc różnicę wieku wyrażonego w miesiącach w stosunku do uczniów urodzonych w grudniu o takiej samej długości cyklu kształcenia w gimnazjum.

W Tabeli 1 przedstawiono rozkład roku urodzenia uczniów zdających egzamin gimnazjalny w 2013 r. wraz z informacją, którą grupę wieku rozpoczęcia nauki w gimnazjum reprezentowali uczniowie. Kursywą zaznaczono te grupy uczniów, które zostały uwzględnione w analizach. Dla trzech

⁵ Gdyby wiek uczniów wpływał tylko na poziom osiągnięć szkolnych (a nie na ich przyrost), wyniki na sprawdzianie po szóstej klasie pozwoliłyby w wystarczającym stopniu kontrolować jego znaczenie (wyniki sprawdzianu także zależą od wieku) i nie zaobserwowano by związku z przyrostami osiągnięć. Jeśli jednak okaże się, że opisane powyżej reszty z modelu EWD (wyniki na egzaminie gimnazjalnym przy kontroli wyników na sprawdzianie) zależą od wieku uczniów, oznaczać to będzie, że ma on znaczenie nie tylko dla poziomu, lecz także dla przyrostów osiągnięć szkolnych.

Tabela 1

Rozkład roku urodzenia uczniów w podziale na cykl kształcenia wśród kohorty uczniów zdających egzamin gimnazjalny w 2013 r.

Grupa uczniów ze względu na rok urodzenia	Rok urodzenia	Trzyletni cykl kształcenia		Rok urodzenia	Czteroletni cykl kształcenia		Procent populacji
		<i>n</i>	%		<i>n</i>	%	
Opóźniony start nauki w gimnazjum o 4 lata	1993	8	0	1992	2	0,02	0,00
Opóźniony start nauki w gimnazjum o 3 lata	1994	50	0,01	1993	5	0,04	0,02
Opóźniony start nauki w gimnazjum o 2 lata	1995	535	0,15	1994	101	0,84	0,18
Opóźniony start nauki w gimnazjum o 1 rok	1996	5 196	1,49	1995	1 211	10,07	1,78
Główna kohorta wiekowa	1997	339 977	97,61	1996	10 680	88,81	97,32
Przyspieszony start nauki w gimnazjum o 1 rok	1998	2 517	0,72	1997	27	0,22	0,71
Przyspieszony start nauki w gimnazjum o 2 lata	1999	5	0				0,00
Razem		348 288	100		12 026	100	100

pozostałych kohort rozkłady są podobne, dlatego nie będą tu prezentowane.

Cykl kształcenia. Zmienna przyjmuje wartość 1 dla uczniów o czteroletnim cyklu kształcenia w gimnazjum. Są to przede wszystkim osoby raz powtarzające klasę podczas nauki w gimnazjum. Dla uczniów o standardowym, trzyletnim cyklu kształcenia zmienna przyjmuje wartość 0.

Międzyszkolne zróżnicowanie wieku uczniów

Jeżeli poziom i przyrost osiągnięć szkolnych zależy od wieku uczniów, to pewni uczniowie mogą mieć, wynikającą z wieku, przewagę nad innymi. Nie oznacza to jednak, że wskaźniki opisujące wyniki nauczania w szkołach są zniekształcone przez te zależności. Aby można było mówić o mającym praktyczne znaczenie obciążeniu takich wskaźników, szkoły musiałyby różnić się między sobą pod względem wieku uczniów

lub liczby uczniów o opóźnionym lub przyspieszonym starcie nauki w danej szkole. Jeśli rozkład tych zmiennych byłby we wszystkich szkołach taki sam, żadna placówka nie uzyskiwałaby przewagi na podstawie wynikającej z korzystnego dla osiągnięć wieku uczniów. Dlatego, aby uzupełnić charakterystykę zmiennych niezależnych, w Tabeli 2 przedstawiono wyniki analiz pokazujących, jak bardzo szkoły różnią się między sobą pod kątem wieku uczniów. W tym celu dokonano dekompozycji wariancji na część, którą można przypisać podziałowi uczniów na szkoły oraz wariancję wewnątrzszkolną. Wykorzystano modele wielopoziomowych regresji liniowych z efektem losowym dla stałej, uwzględniające podział uczniów na szkoły (Raudenbush i Bryk, 2002). Zmienną zależną w tych modelach był wiek uczniów wyrażony w miesiącach. Modele nie uwzględniały żadnych zmiennych niezależnych (były to tzw. modele puste). Analizy wykonano za pomocą oprogramowania Stata 13.1 (procedura xtmixed), wykorzystując metodę

Tabela 2

*Zróźnicowanie międzyszkolne wieku uczniów wyrażonego w miesiącach. Oszacowania efektów losowych z trzypoziomowych modeli z efektami losowymi dla stałej**

Parametr	Rok sesji egzaminacyjnej			
	2013	2012	2011	2010
Wariancja efektów szkół	1,464 (0,047)	1,339 (0,044)	1,001 (0,035)	1,217 (0,040)
Wariancja na poziomie ucznia	19,556 (0,047)	19,152 (0,045)	19,795 (0,046)	20,083 (0,045)
Zróźnicowanie międzyszkolne	6,96%	6,53%	4,85%	5,71%
Liczba uczniów	360 244	374 285	387 909	408 599
Liczba szkół	6 415	6 419	6 418	6 362

* W nawiasach podano wartości błędów standardowych.

estymacji największej wiarygodności (*maximum likelihood*).

Analiza międzyszkolnego zróźnicowania wieku uczniów pokazała, że szkoły istotnie różnią się między sobą pod tym kątem, przy czym podział na szkoły wyjaśnia 5–7% wariancji wieku uczniów (wskaźnik liczony jako stosunek wariancji efektu losowego na poziomie szkoły do wariancji całkowitej, wyrażony w procentach). Oznacza to, że międzyszkolne zróźnicowanie wieku uczniów na poziomie gimnazjów nie jest obecnie duże, choć jest ok. dwukrotnie większe niż w przypadku szkół podstawowych, gdzie wynosi ono ok. 1–3%, jak wynika z przytaczanych

wcześniejszych badań (Dolata i in., 2014; Jasińska i Modzelewski, 2013).

Międzyszkolne zróźnicowanie opóźnionego i przyspieszonego startu szkolnego przedstawiono za pomocą statystyk opisowych dla zmiennych opisujących odsetek takich uczniów w szkołach. Wyniki przedstawiono w Tabeli 3. Pokazują one, że istnieje zróźnicowanie międzyszkolne w tym zakresie, choć średni odsetek uczniów o opóźnionym lub przyspieszonym starcie szkolnym jest niewielki (zwłaszcza w przypadku przyspieszonego startu), a w co najmniej 25% szkół nie ma takich uczniów.

Wyniki analiz pokazały, że istnieje

Tabela 3

Statystyki opisowe międzyszkolnego zróźnicowania udziału uczniów o opóźnionym i przyspieszonym starcie szkolnym

Statystyka	Procent uczniów o opóźnionym starcie nauki w gimnazjum				Procent uczniów o przyspieszonym starcie nauki w gimnazjum			
	Sesja egzaminacyjna				Sesja egzaminacyjna			
	2013	2012	2011	2010	2013	2012	2011	2010
Średnia	2,34	2,16	2,22	2,06	0,77	0,89	0,85	0,79
Odch. st.	5,76	5,62	5,81	5,46	2,27	2,70	2,13	1,81
25. percentyl	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Mediana	2,34	2,16	2,22	2,06	0,77	0,89	0,85	0,79
75. percentyl	2,70	2,54	2,56	2,47	0,64	0,99	1,00	1,02
99. percentyl	26,32	26,67	26,67	22,22	9,09	9,09	9,09	8,82
Liczba szkół	6 415	6 419	6 418	6 362	6 415	6 419	6 418	6 362

międzyszkolne zróżnicowanie zmiennych związanych z wiekiem uczniów. Mimo że nie jest ono duże, nie powinno być bagatelizowane w badaniach nad optymalną postacią modeli EWD.

Schemat analizy

Przeprowadzone analizy zmierzały do zwerifikowania, czy wariację niewyjaśnioną – zwłaszcza na poziomie szkół – z modeli EWD można częściowo wyjaśnić na podstawie informacji o wieku uczniów. Jeśli taka sytuacja ma miejsce, to wskaźniki EWD obciążałoby częściowe uzależnienie od cech uczniów, na które szkoła nie ma wpływu. Przeprowadzone analizy miały także pokazać, w jaki sposób względne przyrosty osiągnięć, będące podstawą wyliczania miar EWD, zależą od wieku uczniów, a tym samym, na czym może polegać obciążenie wskaźników, jeśli w modelach informacja o wieku nie jest uwzględniana. W tym celu dla każdej z czterech sesji egzaminacyjnych oraz dwóch zmiennych zależnych (względnych przyrostów osiągnięć z zakresu matematyczno-przyrodniczego i humanistycznego) badano związek względnych przyrostów osiągnięć szkolnych z wiekiem uczniów. Analizy przeprowadzono osobno dla każdej z czterech sesji egzaminacyjnych, żeby zbadać, czy zaobserwowane efekty są stabilne w kolejnych latach, co świadczyłoby o tym, że mamy do czynienia z ogólnymi prawidłowościami, niezależnymi od egzaminu czy specyficznego rocznika uczniów.

W modelach tych wykorzystano względny wiek ucznia w miesiącach, ponieważ przyrosty osiągnięć szkolnych zostały wyliczone pod kontrolą zmiennej opisującej cykl kształcenia (trzy- lub czteroletni). Dlatego względne przyrosty osiągnięć odnoszą się do różnic w ramach danego cyklu, niezależnie od tego, że w rzeczywistości uczniowie, którzy spędzili w gimnazjum cztery lata są o rok starsi od uczniów o standardowym, trzyletnim cyklu kształcenia. Następnie weryfikowano, czy

obserwuje się istotne efekty dla zmiennych opisujących przynależność do grupy uczniów o opóźnionym (o rok lub dwa lata) i przyspieszonym starcie. Szukano zatem odpowiedzi na pytanie, czy przyrosty osiągnięć w tych grupach są różne od przyrostów osiągnięć w głównej kohorcie wiekowej, po wytrąceniu znaczenia wieku wyrażonego w miesiącach. Następnie zbadano, czy zależność względnych przyrostów od wieku przebiega tak samo wśród podgrupy uczniów z głównej kohorty, wśród uczniów o przyspieszonym, opóźnionym o rok i opóźnionym o dwa lata starcie nauki w gimnazjum (czyli badano interakcję względnego wieku w miesiącach ze zmiennymi dychotomicznymi opisującymi przynależność do poszczególnych grup). Weryfikowano także, czy zachodzi potrzeba wyodrębniania w modelu dwóch zmiennych opisujących opóźniony o rok i o dwa lata start nauki w gimnazjum, czy można uprościć model, włączając tylko jedną zmienną opisującą efekt rozpoczęcia nauki w gimnazjum później, niż by to wynikało z roku urodzenia. W kolejnym kroku sprawdzano, czy zależność względnych przyrostów osiągnięć od zmiennych opisujących wiek uczniów przebiega tak samo wśród uczniów o standardowym, trzyletnim, jak i o czteroletnim cyklu kształcenia. W tym celu do modelu dodano efekty interakcyjne dla zmiennych opisujących zależność przyrostów od wieku i zmiennej dychotomicznej oznaczającej uczniów o czteroletnim cyklu kształcenia.

Testowanie potrzeby uwzględnienia w modelu zmiennych dychotomicznych odzwierciedlających selekcję uczniów do poszczególnych grup (przyspieszony, opóźniony start nauki w gimnazjum, czteroletni cykl kształcenia) oraz zmiennych interakcyjnych modelujących różną siłę zależności przyrostów osiągnięć od wieku uczniów, jest niezbędne dla prawidłowego oszacowania związku między wiekiem uczniów w miesiącach a względnymi przyrostami osiągnięć. Przynależność uczniów do wspomnianych

grup nie ma bowiem charakteru losowego, a co więcej, jest z definicji powiązana z osiągnięciami szkolnymi. Uwzględnienie w modelu zmiennych opisujących przydział do grup (jeśli okaże się to konieczne), a także interakcji między przynależnością do grupy a wiekiem uczniów, umożliwi oszacowanie parametrów selekcji i różnej postaci zależności w analizowanych grupach uczniów.

Podczas prowadzonych analiz konieczne było podjęcie wielu decyzji dotyczących tego, czy warto pozostawić w modelu niektóre z testowanych zmiennych. Z punktu widzenia badanego problemu bardziej interesujące są ogólne prawidłowości pozwalające opisać relację między względnymi przyrostami osiągnięć a wiekiem uczniów niż ich skutki w konkretnej kohorcie (por. Deming, 1953), wyniki z poszczególnych sesji egzaminacyjnych kolejnych roczników uczniów potraktowano jako realizacje danych pochodzących z superpopulacji uczniów kończących gimnazjum w Polsce (Malec, 2008). Dlatego decyzję o pozostawieniu w modelu danej zmiennej podejmowano na podstawie testu istotności statystycznej. Dodatkowo wsparto się analizą współczynnika pseudo R^2 jako miarą odsetka wyjaśnianej wariancji zmiennej zależnej przez dodaną do modelu zmienną niezależną (Domański i Pokropek, 2011). Należy jednak pamiętać, że dla modeli wielopoziomowych miara ta wykazuje pewną niestabilność oszacowań (Snijders i Bosker, 2012). Wskaźnika tego używano przede wszystkim do oceny, z którego poziomu analizy wyjaśniana jest wariancja po dodaniu analizowanych zmiennych niezależnych.

Drugi etap analiz polegał na porównaniu wskaźników EWD wyliczonych z modeli trzyletnich bez uwzględnienia wieku (realizacja u_j z modelu opisanego równaniem 2) ze wskaźnikami z modeli uwzględniających dodatkowo informację o wieku (wartość u_j z modelu opisanego równaniem 3), zdefiniowaną za pomocą takiej postaci funkcyjnej, jaka okazała się najlepsza na poprzednim

etapie analiz (β_2 to wektor parametrów związanych ze zmiennymi opisującymi wiek ucznia i przynależność do grup o opóźnionym i przyspieszonym starcie nauki, a $WIEK_{ij}$ to wektor zmiennych niezależnych opisujących wiek i -tego ucznia w j -tej szkole). Na tym etapie posłużono się modelami trzyletnimi, czyli uwzględniającymi dane z trzech kolejnych sesji egzaminacyjnych, ponieważ na podstawie takich modeli liczone są publikowane wskaźniki szkół. W równaniach (2) i (3) oznaczono to zmiennymi: rok pisania sprawdzianu ($rok_{s_{ij}}$) oraz rok pisania egzaminu gimnazjalnego ($rok_{g_{ij}}$) i -tego ucznia w j -tej szkole. Pozostałe oznaczenia są analogiczne do zastosowanych w równaniu (1).

$$y_{ij} = w^k \left(x_{ij}, rok_{s_{ij}}, rok_{g_{ij}} \right) + \beta Z_{ij} + u_j + r_{ij} \quad (2)$$

$$y_{ij} = w^k \left(x_{ij}, rok_{s_{ij}}, rok_{g_{ij}} \right) + \beta Z_{ij} + u_j + r_{ij} \quad (3)$$

Celem tych analiz było zbadanie, o ile dodanie informacji o wieku uczniów zmieniłoby oszacowania wskaźników. Do oszacowania wskaźników EWD posłużono się dokładnie tymi samymi funkcjami i metodami estymacji, które są wykorzystywane do wyliczania wskaźników trzyletnich na potrzeby ich publikacji (Żółtak, 2013)⁶.

Wyniki

Zależność względnych przyrostów osiągnięć od wieku uczniów

Jako pierwsze zostaną przedstawione wyniki analiz, w których zmienną wyjaśnianą były względne przyrosty osiągnięć z przedmiotów matematyczno-przyrodniczych; jako

⁶ Autorka dziękuje Tomaszowi Żółtakowi za udostępnienie skryptów do liczenia modeli EWD. Dzięki temu możliwe było zastosowanie porównywalnej metodologii wyliczania modeli do wykorzystywanej na potrzeby publikacji wskaźników EWD.

kolejne – wyniki dla przedmiotów humanistycznych. Ze względu na ograniczoną objętość artykułu zaprezentowane zostaną modele uwzględniające te zmienne, które mają znaczenie dla przewidywania względnych przyrostów osiągnięć. Wyniki pozostałych zostaną tylko skomentowane.

Przedmioty matematyczno-przyrodnicze

Pierwsze pytanie dotyczy tego, czy informacja o wieku i przynależności do grup o przyspieszonym lub opóźnionym starcie nauki w gimnazjum pozwala przewidywać względne przyrosty osiągnięć, które są podstawą wyliczania efektów dla szkół. Okazuje się, że informacja ta, uwzględniona

w modelu przez zestaw zmiennych wymienionych w Tabeli 4, pozwala wyjaśnić 0,66–0,85% wariancji względnych przyrostów osiągnięć z przedmiotów matematyczno-przyrodniczych. Wskazuje na to współczynnik pseudo R^2 (Domański i Pokropek, 2011). Nie jest to dużo, jednak warto zwrócić uwagę, że wyjaśniana jest przede wszystkim wariancja z poziomu międzyszkolnego. Dodanie wspomnianej informacji do modelu pozwala wyjaśnić ok. 2,5–3% zróżnicowania międzyszkolnego względnych przyrostów osiągnięć szkolnych. Może to już przekładać się na pewne różnice w oszacowaniach wskaźników efektywności nauczania dla szkół w zależności od tego,

Tabela 4

Zależność względnych przyrostów osiągnięć szkolnych z przedmiotów matematyczno-przyrodniczych od wieku uczniów. Wyniki dwupoziomowych regresji z efektami losowymi dla stałej

Rok sesji egzaminacyjnej	2013	2012	2011	2010
Oszacowanie efektów stałych				
Wiek (względny)	-0,104* (0,004)	-0,119* (0,004)	-0,094* (0,004)	-0,122* (0,004)
Opóźniony start	-1,586* (0,128)	-1,636* (0,132)	-2,360* (0,128)	-2,291* (0,122)
Przyspieszony start	2,354* (0,179)	2,428* (0,167)	2,555* (0,169)	2,178* (0,157)
Czteroletni cykl kszt.	0,049 (0,087)	0,044 (0,090)	0,060 (0,087)	0,139 (0,072)
Interakcja: opóźniony start x czteroletni cykl	1,343* (0,285)	1,773* (0,291)	1,997* (0,294)	0,881* (0,262)
Stała	0,626* (0,049)	0,694* (0,046)	0,572* (0,049)	0,733* (0,043)
Oszacowanie efektów losowych				
Wariancja efektu na poziomie szkół	9,491 (0,209)	7,531 (0,173)	9,241 (0,204)	6,605 (0,152)
Wariancja efektu na poz. uczniów	75,521 (0,180)	77,413 (0,181)	79,580 (0,182)	70,877 (0,158)
Pseudo R^2	0,66%	0,68%	0,67%	0,83%
Pseudo R^2 (poz. szkół)	2,71%	2,74%	2,58%	3,07%
Pseudo R^2 (poz. uczniów)	0,40%	0,47%	0,44%	0,61%
Podsumowanie				
Liczba uczniów	360 232	374 105	387 738	408 372
Liczba szkół	6 415	6 418	6 418	6 362
Log likelihood	-1 296 090,7	-1 349 821,6	-1 404 697,9	-1 455 028,2

W nawiasach podano błędy standardowe.

* Efekty istotne statystycznie na poziomie $p < 0,05$.

czy uwzględnimy podczas ich wyznaczania informację o wieku uczniów, czy nie.

Przeprowadzone analizy potwierdziły, że względny przyrost osiągnięć szkolnych z przedmiotów matematyczno-przyrodniczych jest negatywnie związany z wiekiem uczniów. Obserwowany efekt ma podobną wielkość w kolejnych latach, choć nie jest on silny. Uczniowie młodszy o 12 miesięcy osiągają średnio o trochę ponad 1 punkt egzaminacyjny (na skali o odchyleniu standardowym 15) większy przyrost osiągnięć niż uczniowie starsi. Zależność ta ma zasadniczo taką samą siłę wśród uczniów z głównej kohorty wiekowej, jak wśród uczniów o przyspieszonym lub opóźnionym starcie nauki w gimnazjum, a także wśród uczniów o trzyletnim i czteroletnim cyklu kształcenia. Zostało to zbadane za pomocą testowania istotności efektów interakcyjnych zmiennej „względny wiek w miesiącach” ze zmiennymi dychotomicznymi określającymi przynależność do wspomnianych grup⁷.

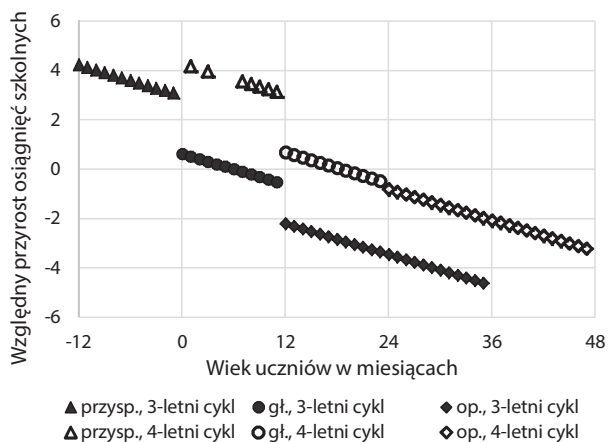
Wyniki analiz pokazały istnienie stabilnych i znaczących efektów dla przyspieszonego i opóźnionego startu nauki w gimnazjum. Uczniowie, którzy rozpoczęli naukę w gimnazjum będąc o rok młodszymi niż dominująca kohorta, uzyskali średnio większe o ponad 2 punkty (ok. 1/6 odchylenia standardowego) względne przyrosty osiągnięć z przedmiotów matematyczno-przyrodniczych, niż uczniowie rozpoczynający

naukę w gimnazjum w typowym wieku. Co ważne, efekt ten nie wynika z różnicy w wieku biologicznym, który jest modelowany przez zmienną *względny wiek uczniów w miesiącach*. To konsekwencja faktu, że w grupie uczniów o przyspieszonym starcie nauki w gimnazjum znaleźli się uczniowie o cechach, które są powiązane z tempem przyrostów osiągnięć. Dla pełnego obrazu należy też dodać, że nie stwierdzono istotnej różnicy między wielkością tego efektu w grupach uczniów o trzyletnim i czteroletnim toku kształcenia.

Uczniowie, którzy mieli natomiast opóźniony start nauki w gimnazjum oraz kształcili się w standardowym trzyletnim cyklu, uczynili średnio mniejsze o 1,6–2,4 punktu egzaminacyjnego (ok. 1/8 odchylenia standardowego) postępy niż uczniowie z głównej kohorty ponad to, co wynikało z różnic w wieku w miesiącach. Różnica ta jest mniejsza dla uczniów z czteroletniego cyklu kształcenia, na co wskazuje wartość efektu interakcyjnego między zmiennymi *opóźniony start* i *czteroletni cykl kształcenia*. Dla lat 2013, 2012 i 2011 efekty się praktycznie znoszą, sprawiając, że wśród uczniów o czteroletnim cyklu kształcenia zależność względnych przyrostów od wieku w grupach uczniów z głównej kohorty i uczniów o opóźnionym starcie jest zasadniczo liniowa bez wyraźnych uskoków. W 2010 r. natomiast uczniowie o opóźnionym starcie wśród uczniów o czteroletnim cyklu kształcenia uzyskali przyrosty osiągnięć szkolnych z przedmiotów matematyczno-przyrodniczych o ok. 1,4 punktu niższe (od -2,291 do +0,881).

W analizach weryfikowano także, czy powinien być modelowany oddzielny efekt dla opóźnionego o rok startu nauki w gimnazjum i oddzielny dla opóźnionego startu o dwa lata. Innymi słowy, sprawdzano, czy zależność w grupach uczniów o opóźnionym o rok i dwa lata starcie należy modelować za pomocą regresji przedziałami liniowej (z uskokiem między tymi dwiema kategoriami),

⁷ Efekty te były testowane przez oddzielne włączanie ich do modeli uwzględniających informację o wieku i przynależności do wymienianych grup. Okazały się one nieistotne statystycznie (przy założonym poziomie istotności $p < 0,05$), z wyjątkiem interakcji zmiennej wiek ze zmienną określającą uczniów o opóźnionym o rok starcie nauki w gimnazjum wśród uczniów o trzyletnim toku kształcenia, ale tylko dla 2010 r., oraz tej samej interakcji, ale wśród uczniów o czteroletnim toku kształcenia dla 2013 r. W wymienionych dwóch podgrupach zależność była trochę silniejsza. Ponieważ rezultaty te nie zyskały potwierdzenia w innych latach, uznano je za specyficzne dla danej sesji egzaminacyjnej, mogące być dziełem przypadku, i z tego powodu nieznaczące dla określenia ogólnej postaci analizowanej zależności.



Rysunek 1. Zależność względnych przyrostów osiągnięć szkolnych z przedmiotów matematyczno-przyrodniczych (przewidywań z części stałej modelu) od wieku uczniów w miesiącach dla kohorty zdającej egzamin gimnazjalny w 2013 r.

czy można ją uprościć do zależności liniowej. Analizy pokazały, że wielkość uskoğu między dwiema grupami dla wartości przewidywanych wynosi średnio poniżej 0,5 punktu, a tylko wśród uczniów o czteroletnim toku kształcenia zdających egzamin w 2010 r. przyjmuje on wartość powyżej 1 punktu. Ze względu na niewielkie wielkości tych efektów uznano, że lepszy będzie model prostszy, w którym zależność względnych przyrostów od wieku w obu tych grupach modelujemy za pomocą funkcji liniowej.

Komentarza może jeszcze wymagać fakt, że względne przyrosty osiągnięć uczniów o czteroletnim cyklu kształcenia średnio nie różnią się od względnych przyrostów uczniów o trzyletnim cyklu kształcenia. Wskazują na to nieistotne efekty dla zmiennej *czteroletni cykl kształcenia* przedstawione w Tabeli 4. Rezultat ten nie dowodzi jednak, że wśród uczniów o czteroletnim cyklu kształcenia obserwujemy taki sam wzrost osiągnięć, jak wśród uczniów, którzy w gimnazjum uczyli się trzy lata. Wynika to z faktu, że w modelach EWD, na podstawie których wyliczane były względne przyrosty osiągnięć uczniów, uwzględniana jest

zmienna opisująca czteroletni cykl kształcenia. Zatem przyrosty osiągnięć dla uczniów o wydłużonym toku kształcenia są wyznaczane na tle innych uczniów, którzy w gimnazjum kształcili się cztery lata.

W celu lepszego zobrazowania wyników przedstawionych w Tabeli 4, zależność dla przykładowego rocznika została także pokazana na Rysunku 1. Przedstawiono na nim przewidywania z części stałej modelu, w którym względne przyrosty osiągnięć szkolnych (oś pionowa) przywołujemy zestawem zmiennych niezależnych wymienionych w Tabeli 4. Oś pozioma opisuje wiek uczniów wyrażony w miesiącach, gdzie 0 to wiek najmłodszych uczniów (czyli urodzonych w grudniu) w głównej kohorcie wiekowej o trzyletnim cyklu kształcenia. Na wykresie osobno zobrazowano zależność dla uczniów o standardowym, trzyletnim cyklu kształcenia i osobno dla uczniów o czteroletnim cyklu kształcenia oraz osobno dla uczniów o przyspieszonym, opóźnionym starcie nauki i uczniów z głównej kohorty wiekowej. Na potrzeby wykresu posłużono się zmienną oznaczającą rzeczywisty wiek w miesiącach, tak

więc wartości dla uczniów o czteroletnim cyklu kształcenia mieszczą się w przedziale przesuniętym o 12 miesięcy w stronę wartości wyższych (są to bowiem uczniowie o 12 miesięcy starsi, którzy spędzili w gimnazjum dodatkowy rok).

Przedmioty humanistyczne

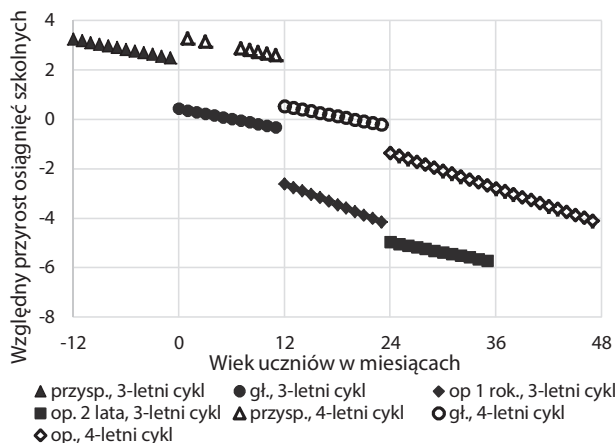
W przypadku przyrostów osiągnięć szkolnych z przedmiotów humanistycznych informacja o wieku uczniów pozwala wyjaśnić podobny odsetek ich wariancji, jak w przypadku przedmiotów matematyczno-przyrodniczych. Pseudo R^2 dla całego modelu wynosi ok. 0,6–0,8%. Zmienne opisujące wiek uczniów oraz przynależność do grup o opóźnionym lub przyspieszonym starcie nauki w gimnazjum wyjaśniają, podobnie jak w poprzednio opisanych modelach, przede wszystkim wariancję między-szkolną (pseudo R^2 na tym poziomie waha się od 2,5 do 3,6%).

Zależność względnych przyrostów osiągnięć szkolnych z przedmiotów humanistycznych od wieku uczniów jest mniej stabilna w kolejnych latach niż zależność dla przedmiotów matematyczno-przyrodniczych. Dodatkowo przeprowadzone analizy pokazały, że do jej opisu potrzeba trochę bardziej złożonego modelu. Względne przyrosty osiągnięć szkolnych z przedmiotów humanistycznych są negatywnie związane z wiekiem uczniów, choć efekt jest trochę słabszy niż w przypadku przedmiotów matematyczno-przyrodniczych. Uczniowie starsi o 12 miesięcy uzyskują średnio niższe względne przyrosty osiągnięć od prawie 0,5 do niecałego 1 punktu egzaminacyjnego na skali o odchyleniu standardowym równym 15. Nie stwierdzono różnic w sile tej zależności między grupami uczniów o trzy- i czteroletnim cyklu kształcenia, a także w grupie uczniów z głównej kohorty i uczniów o przyspieszonym starcie nauki w gimnazjum.

Uczniowie, który rozpoczęli naukę

w gimnazjum o rok wcześniej, zrobili natomiast trochę większe postępy podczas nauki w tej szkole – ich względne przyrosty osiągnięć z przedmiotów humanistycznych są o ok. 2 punkty większe od uczniów z głównej kohorty. Efekt jest zasadniczo taki sam dla uczniów kształconych zarówno w trzy-, jak i czteroletnim cyklu, ponieważ poza sesją egzaminacyjną z 2010 r. nie stwierdzono istotnych interakcji między zmiennymi opisującymi przyspieszony start i czteroletni tok kształcenia. W 2010 r. efekt przyspieszonego startu dla uczniów z czteroletniego cyklu kształcenia był silniejszy, jednak ponieważ wystąpił tylko w tej sesji egzaminacyjnej, a dodatkowo dotyczył niecałego 0,01% uczniów, uznano go za nieznaczący dla opisywanej zależności.

Zależność względnych przyrostów osiągnięć od wieku w grupie uczniów, którzy rozpoczęli naukę w gimnazjum później niż główna kohorta, okazała się bardziej złożona. Stwierdzono silniejszy związek przyrostów osiągnięć od wieku wśród uczniów o opóźnionym o rok starcie nauki w gimnazjum, zwłaszcza w grupie o trzyletnim cyklu kształcenia. Dodatkowe analizy pokazały, że w grupie tej należy modelować oddzielny efekt dla opóźnionego o rok startu nauki w gimnazjum i oddzielny dla startu opóźnionego o dwa lata. Inny obraz zależności odnotowano wśród uczniów o czteroletnim cyklu kształcenia dla grupy o opóźnionym starcie nauki w gimnazjum. Efekty dla opóźnionego startu o dwa lata okazały się bardzo niestabilne w kolejnych latach, a obserwowane różnice można z dużym prawdopodobieństwem przypisać małej liczebności tej podgrupy badanych (stanowią oni 0,03% próby). Aby zapobiec wpływowi losowych wahań, ostatecznie podjęto decyzję o modelowaniu regresją liniową zależności przyrostów od wieku w grupie uczniów o czteroletnim cyklu kształcenia i opóźnionym starcie szkolnym, bez wyróżniania dodatkowych



Rysunek 2. Zależność względnych przyrostów osiągnięć szkolnych z przedmiotów humanistycznych (przewidywań z części stałej modelu) od wieku uczniów w miesiącach dla kohorty zdającej egzamin gimnazjalny w 2013 r.

efektów na start opóźniony o rok i dwa lata⁸. Rysunek 2, na którym przedstawiono wyniki modelu dla przykładowego rocznika, zobrazuje ostateczną postać modelu.

Wśród uczniów o opóźnionym o rok starcie nauki w gimnazjum, którzy kształcili się w trzyletnim cyklu, obserwujemy średnio mniejsze o 2,3–3,3 punkty egzaminacyjne względne przyrosty osiągnięć w porównaniu do uczniów z głównej grupy wiekowej⁹. Dla uczniów z czteroletniego cyklu kształcenia różnica ta, podobnie jak w przypadku przedmiotów matematyczno-przyrodniczych, okazała się wyraźnie mniejsza (wynosiła 1,1–2,3). Uczniowie kształceni według

trzyletniego cyklu, którzy rozpoczęli naukę w gimnazjum o dwa lata później, uzyskali jeszcze o 0,7–1,1 punktu mniejsze przyrosty osiągnięć (poza różnicą wynikającą z wieku) od uczniów o starcie opóźnionym o rok. Dodatkowo stwierdzono istnienie efektów interakcyjnych między wiekiem wyrażonym w miesiącach a zmiennymi dychotomicznymi opisującymi przynależność do grupy uczniów o opóźnionym o rok starcie nauki w gimnazjum spośród kształconych w trzyletnim cyklu (efekt nieistotny tylko w 2013 r.) oraz do grupy uczniów o opóźnionym starcie wśród uczniów kształconych w czteroletnim cyklu (efekt istotny w latach 2011 i 2010). W grupach tych zależność względnych przyrostów osiągnięć od wieku uczniów jest silniejsza.

Wskaźniki EWD uwzględniające i nieuwzględniające informację o wieku uczniów

Przedstawione wyniki pokazały, że informacja o wieku uczniów i przynależności do grup o opóźnionym lub przyspieszonym starcie

⁸ Model taki porównywano z modelem, w którym szacowano łączny efekt dla startu opóźnionego o dwa lata w grupie uczniów o trzy- i czteroletnim cyklu kształcenia. Modele te okazały się tak samo dobrze dopasowane do danych (na podstawie porównania wartości *log likelihood*). O wyborze modelu dla uczniów o czteroletnim cyklu kształcenia przesądziła więc jego zbliżona postać do modelu dla danych z przedmiotów matematyczno-przyrodniczych.

⁹ Jest to wielkość różnicy pomiędzy najstarszymi uczniami z głównej kohorty a najmłodszymi uczniami spośród uczniów o starcie opóźnionym o rok wyliczona na podstawie oszacowanych parametrów regresji.

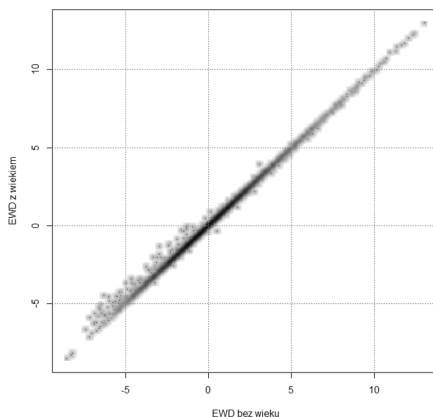
Tabela 5

Zależność względnych przyrostów osiągnięć szkolnych z przedmiotów humanistycznych od wieku uczniów. Wyniki dwupoziomowych regresji z efektami losowymi dla stałej

Rok sesji egzaminacyjnej	2013	2012	2011	2010
Oszacowanie efektów stałych				
Wiek (względny)	-0,068* (0,005)	-0,074* (0,005)	-0,047* (0,004)	-0,040* (0,004)
Przyspieszony start	1,989* (0,188)	1,817* (0,174)	2,174* (0,175)	1,852* (0,177)
Start opóźniony o rok wśród uczniów o trzyletnim cyklu	-1,362* (0,625)	-0,832 (0,633)	-1,640* (0,619)	-1,123 (0,623)
Interakcja: start opóźniony o rok wśród uczniów o trzyletnim cyklu x wiek (względny)	-0,072 (0,037)	-0,129* (0,038)	-0,135* (0,037)	-0,163* (0,037)
Start opóźniony o dwa lata wśród uczniów o trzyletnim cyklu	-3,759* (0,414)	-4,610* (0,414)	-5,358* (0,403)	-5,891* (0,461)
Czteroletni cykl kształcenia	0,099 (0,092)	0,132 (0,093)	0,172 (0,090)	0,195* (0,081)
Start opóźniony wśród uczniów o czteroletnim cyklu	-0,459 (1,026)	-1,025 (1,044)	0,061 (1,024)	-0,133 (1,018)
Interakcja: start opóźniony wśród uczniów o czteroletnim cyklu x wiek (względny)	-0,051 (0,057)	-0,030 (0,057)	-0,150* (0,056)	-0,173* (0,056)
Stała	0,441* (0,050)	0,475* (0,049)	0,339* (0,053)	0,294* (0,050)
Oszacowanie efektów losowych				
Wariancja efektu na poziomie szkół	9,267 (0,209)	8,757 (0,198)	11,697 (0,252)	9,207 (0,206)
Wariancja efektu na poziomie uczniów	83,092 (0,198)	83,270 (0,194)	84,876 (0,194)	89,688 (0,200)
Pseudo R^2	0,61%	0,70%	0,79%	0,59%
Pseudo R^2 (poziom szkół)	3,10%	3,56%	3,09%	2,45%
Pseudo R^2 (poziom uczniów)	0,33%	0,39%	0,46%	0,39%
Podsumowanie				
Liczba uczniów	360 218	374 254	387 800	408 551
Liczba szkół	6 415	6 419	6 418	6 362
Log likelihood	-1 312 939,2	-1 364 232,2	-1 417 896,5	-1 504 006,9

W nawiasach podano błędy standardowe.

* Efekty istotne statystycznie na poziomie $p < 0,05$.



Rysunek 3. Porównanie trzyletnich wskaźników EWD dla przedmiotów matematyczno-przyrodniczych (2010–2012) wyliczonych z modeli: tradycyjnego (bez wieku) oraz uwzględniającego informację o wieku uczniów.

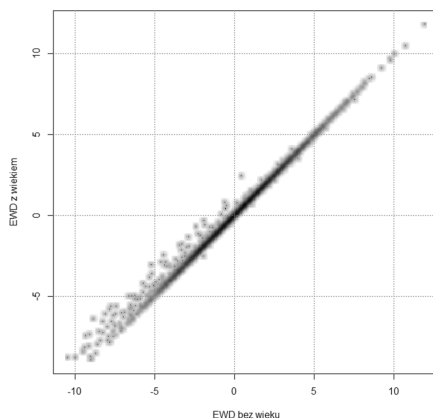
nauki w gimnazjum pozwala wyjaśnić część wariancji względnych przyrostów osiągnięć przede wszystkim z poziomu międzyszkolnego. Może to skutkować zmianą oszacowań wskaźników EWD, jeśli w modelach uwzględnilibyśmy tę informację. Omówione wyniki przybliżyły także, na czym mogłoby polegać obciążenie owych wskaźników, pokazały bowiem, jacy uczniowie mają większą szansę na większy przyrost osiągnięć. Szkoły, które mają więcej takich uczniów, mogą mieć zawyżone oszacowania EWD. W celu odpowiedzi na pytanie, czy rzeczywiście tak jest, wyliczono wskaźniki EWD uwzględniające informację o wieku uczniów i porównano je ze wskaźnikami liczonymi tak, jak obecnie publikowane miary.

Policzono modele EWD dla dwóch okresów trzyletnich, które poza standardowym zestawem zmiennych niezależnych (uwzględnianych w modelach liczonych na potrzeby publikacji wskaźników) uwzględniały zmienne wymienione w Tabelach 4 i 5 (odpowiednio dla modeli matematyczno-przyrodniczych i humanistycznych). Modele te zakładały, że efekty związane z wiekiem uczniów mogą przebiegać w inny sposób

w kolejnych latach (dodano efekty interakcyjne między rokiem pisania egzaminu a zmiennymi uwzględnionymi w modelu) i były one estymowane na podstawie danych z trzech sesji egzaminacyjnych. Wyniki porównania wskaźników EWD z uwzględnieniem i bez uwzględnienia wieku przedstawiono na Rysunku 3 (dla wskaźników za okres 2011–2013 wyniki są analogiczne, dlatego nie zostały pokazane).

Porównanie to pokazuje, że dodanie do modelu informacji o wieku uczniów w bardzo niewielkim stopniu przyczynia się do zmiany oszacowań punktowych wskaźników. Korelacja Pearsona między dotychczas publikowanymi wskaźnikami a wskaźnikami uwzględniającymi dodatkowo informację o wieku uczniów i ich przynależności do grup o opóźnionym lub przyspieszonym starcie, wynosi 0,999. Widzimy też, że niewielkie różnice w oszacowaniach wskaźników występują przede wszystkim wśród wartości niższych.

Wyniki są podobne w przypadku wskaźników z przedmiotów humanistycznych. Na Rysunku 4 pokazano porównanie wskaźników za lata 2010–2012 (dla wskaźników za lata



Rysunek 4. Porównanie trzyletnich wskaźników EWD dla przedmiotów humanistycznych (2010–2012) wyliczonych z modeli: tradycyjnego (bez wieku) oraz uwzględniającego informację o wieku uczniów.

2011–2013 uzyskano takie same rezultaty). Wskaźniki wyliczone z modeli z uwzględnieniem informacji o wieku uczniów i przynależności do grup o opóźnionym i przyspieszonym starciu tylko nieznacznie różnią się od wskaźników wyliczonych za pomocą dotychczas stosowanego modelu. Korelacja między nimi jest tylko trochę niższa niż w przypadku przedmiotów matematyczno-przyrodniczych i wynosi 0,998.

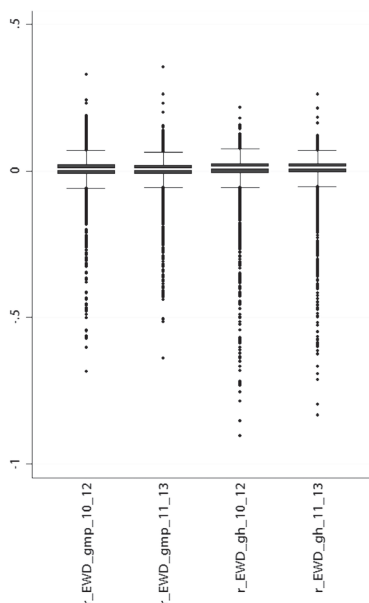
Wielkość zaobserwowanych różnic między wartościami standardowych trzyletnich wskaźników EWD oraz wartościami trzyletnich wskaźników uwzględniających informację o wieku uczniów została dodatkowo przedstawiona na wykresie skrzynkowym (Rysunek 5). Różnice te zostały wyrażone w jednostkach odchylenia standardowego wskaźników EWD.

Wykresy skrzynkowe dobrze pokazują, że dla ok. 99% szkół (obszar pomiędzy wąsami) wartości wskaźników z obu modeli nie różnią się więcej niż 0,13 odchylenia standardowego dotychczas stosowanych wskaźników EWD, w tym dla ok. 50% szkół wielkość różnicy nie przekracza 0,04 odchylenia standardowego. Rozkład tych różnic jest lewoskośny, oznacza to, że dla większości

szkół wartości standardowych wskaźników EWD są nieznacznie wyższe niż wskaźników uwzględniających informacje o wieku uczniów. Jednak wyraźne jednostkowe różnice (przypadki odstające po lewej stronie rozkładu) obserwuje się w sytuacji, w której wartość standardowego wskaźnika jest zaniżona w stosunku do wskaźnika uwzględniającego informację o wieku uczniów. Różnice te nie przekraczają jednak 1 odchylenia standardowego. Wyniki te pokazują, że dla zdecydowanej większości szkół (ponad 99%) uwzględnienie w modelach informacji o wieku uczniów nie przełożyłoby się na zauważalne różnice w wartościach wskaźników. Znaczące różnice dotyczyłyby tylko jednostkowych przypadków.

Podsumowanie

Przedstawione w artykule analizy miały na celu znalezienie odpowiedzi na pytanie, czy nieuwzględnienie podczas wyliczania wskaźników EWD informacji o wieku uczniów oraz ich przynależności do grupy o przyspieszonym lub opóźnionym starciu nauki w gimnazjum obciąża oszacowania wskaźników dla szkół. Włączenie tej informacji



Rysunek 5. Różnice wartości wskaźników EWD wyliczonych z modeli standardowych i modeli z uwzględnieniem informacji o wieku uczniów (gmp – przedmioty matematyczno-przyrodnicze, gh – humanistyczne).

do modeli pozwoliło na zredukowanie niewielkiej części wariancji niewyjaśnionej (o ok. 0,7 %). Przede wszystkim pozwoliło na wyjaśnienie wariancji międzyszkolnej (o ok. 3 %), co ma większe znaczenie dla oszacowań wskaźników EWD. W artykule omówiono, na czym mogłoby polegać obciążenie wskaźników, dzięki przedstawieniu, jacy uczniowie (ze względu na ich wiek lub przynależność do wspomnianych grup) mają większą szansę na większy przyrost osiągnięć szkolnych. Obecnie jednak szkoły w niewielkim stopniu różnią się między sobą pod względem wieku swoich uczniów czy posiadania uczniów o przyspieszonym lub opóźnionym starcie nauki w gimnazjum. Dlatego, mimo że względne przyrosty osiągnięć szkolnych zależą od wieku uczniów i przyspieszonego lub opóźnionego startu szkolnego, wskaźniki EWD wyliczone z modeli uwzględniających informację o przynależności uczniów do tych

grup i ich wieku praktycznie nie różnią się od wyliczanych z pominięciem tych danych. Korelacja między tymi wskaźnikami wyniosła 0,999 dla przedmiotów matematyczno-przyrodniczych i 0,998 dla przedmiotów humanistycznych. Przedstawione wyniki pokazały, że dla zdecydowanej większości szkół uwzględnienie w modelach informacji o wieku uczniów nie przełożyłoby się na zauważalne różnice w wartościach wskaźników.

To ważny wniosek dla metodologii szacowania wskaźników EWD. Wyniki te pokazały bowiem, że obciążenie wynikające z nieuwzględnienia w obecnie stosowanych modelach EWD informacji o wieku uczniów jest zaniedbywalnie małe. Biorąc jednak pod uwagę fakt, że dodanie tych zmiennych do modeli nie niesie za sobą dodatkowych kosztów oprócz potrzeby zbudowania bardziej złożonych modeli (dane

te są dostępne w bazach, z których korzysta się podczas wyliczania wskaźników), można rozważyć rozbudowanie modeli o tę informację w przyszłości, co wpłynie na ich poprawę w niewielkim stopniu.

Istnieje też potrzeba monitorowania analizowanych tu procesów, bowiem już od przysłego roku różnicowanie międzyszkolne oraz wewnątrzszkolne wieku gimnazjalistów może przybrać na sile w związku ze sposobem wdrażania zmian związanych z reformą obniżającą wiek obowiązku szkolnego¹⁰. W 2015 r. pierwsi uczniowie z tzw. okresu przejściowego realizacji zaleceń ustawy przystąpią do sprawdzianu, a trzy lata później do egzaminu gimnazjalnego. W okresie przejściowym naukę w pierwszej klasie rozpoczynają dzieci sześć- i siedmioletnie, co przełoży się na naturalny wzrost różnicowania wieku uczniów tych samych klas. Wzrośnie też różnicowanie międzyszkolne tej cechy, ponieważ odsetek sześciolatków w pierwszych klasach zależy m.in. od działań gmin i samych szkół, a także od nastrojów panujących w środowisku lokalnym. Przed wprowadzeniem reformy w 2008 r. średni odsetek sześciolatków w pierwszych klasach w szkołach wynosił niecały 1%. W kolejnych latach odsetek ten wzrastał do poziomu 18% w 2013 r.¹¹. Zwiększył się nie tylko udział dzieci sześciolatków w pierwszych klasach, lecz także różnicowanie międzyszkolne odsetka sześciolatków w szkołach. W 2009 r. odchylenie standardowe procentu

sześciolatków w pierwszych klasach w szkołach wynosiło 3,63 i rosło w kolejnych latach, osiągając w 2013 r. wartość 19,67.

Dane te pokazują, że w okresie przejściowym szanse na uzyskanie dobrego indywidualnego wyniku przez ucznia na egzaminie, będą w większym niż dotychczas stopniu zależały od jego wieku. Z drugiej zaś strony, wzrost różnicowania międzyszkolnego ze względu na tę cechę może spowodować realne obciążenie wskaźników EWD, jeśli przy ich wyznaczaniu nie zostanie uwzględniona informacja o wieku uczniów. Mimo że obecnie nie ma potrzeby rozbudowywania gimnazjalnych modeli EWD o zmienne opisujące wiek uczniów, w najbliższych latach może się to zmienić. Różnicowanie między- i wewnątrzszkolne wieku uczniów w okresie przejściowym będzie bardziej zależało od działań samej szkoły. W konsekwencji uwzględnienie tej informacji w modelowaniu będzie, ze statystycznego punktu widzenia, jeszcze bardziej problematyczne.

Ostatnią kwestią wymagającą omówienia jest charakter wykorzystanych do modelowania testów egzaminacyjnych. Sprawdzian po szóstej klasie szkoły podstawowej jest testem względnie łatwym, o czym świadczą lewoskośne rozkłady wyników¹². Konsekwencją tego jest gorsze różnicowanie uczniów o wyższym poziomie osiągnięć (tzw. efekt sufitowy). Egzamin gimnazjalny natomiast zwykle lepiej różnicuje tych uczniów. Może to zniekształcać profil analizowanych zależności przyrostów osiągnięć szkolnych od wieku. Ponieważ poziom osiągnięć jest pozytywnie związany z wiekiem uczniów, obecność efektu sufitowego na sprawdzianie oraz jego brak (lub mniejsze nasilenie) na egzaminie gimnazjalnym może powodować oszacowanie większych względnych

¹⁰ Ustawą z dnia 19 marca 2009 r. o zmianie ustawy o systemie oświaty oraz o zmianie niektórych innych ustaw (Dz.U. nr 56, poz. 458 z późn. zm) wprowadzono obowiązek szkolny dla sześciolatków. Wcześniej obowiązek szkolny obejmował dzieci siedmioletnie. Kolejne zmiany odraczały rocznik uczniów, który zostanie objęty obowiązkiem rozpoczęcia nauki w wieku 6 lat, wydłużano okres przejściowy, w którym rodzice mieli prawo posłania dziecka w tym wieku do pierwszej klasy. Jeśli nie zajdą kolejne zmiany, to rok szkolny 2015/2016 będzie ostatnim, w którym do pierwszych klas pójść równocześnie dzieci sześć- i siedmioletnie.

¹¹ Obliczenia własne na podstawie danych z Systemu Informacji Oświatowej ze spisu wrzesniowego.

¹² Charakterystyka rozkładów wyników jest dostępna m.in. w sprawozdaniach Centralnej Komisji Egzaminacyjnej zamieszczonych na stronie <http://www.cke.edu.pl/index.php/sprawdzian-left/o-sprawdzianie-do-2014/88-sprawdzian/36-informacje-o-wynikach>.

przyrostów osiągnięć wśród uczniów starszych, co z kolei prowadziłyby do osłabienia analizowanego związku wieku z przyrostem osiągnięć. Kształty rozkładów wyników egzaminów zmieniają się jednak między latami, a przedstawione w artykule efekty dla kolejnych modeli jednorocznych są dość stabilne, co pokazuje, że jeśli mamy do czynienia ze zniekształceniem tej zależności, to nie jest ono duże.

Literatura

- Bell, J. F. i Daniels, S. (1990). Are summer-born children disadvantaged?. The birthdate effect in education. *Oxford Review of Education*, 16(1), 67–80.
- Cameron, M. B. i Wilson, B. J. (1990). The effects of chronological age, gender, and delay of entry on academic achievement and retention: implications for academic redshirting. *Psychology in the Schools*, 27(3), 260–263.
- Carlton, M. P. i Winsler, A. (1999). School readiness: the need for a paradigm shift. *School Psychology Review*, 28(3), 338–352.
- Crone, D. A. i Whitehurst, G. J. (1999). Age and schooling effects on emergent literacy and early reading skills. *Journal of Educational Psychology*, 91(4), 604–614.
- Deming, E. W. (1953). On the distinction between enumerative and analytic surveys. *Journal of the American Statistical Association*, 48, 244–255.
- Dietz, C. i Wilson, B. J. (1985). Beginning school age and academic achievement. *Psychology in the Schools*, 22(1), 93–94.
- Ding, C. S. i Davison, M. L. (2005). A longitudinal study of math achievement gains for initially low achieving students. *Contemporary Educational Psychology*, 30(1), 81–95.
- Dolata, R. i Pokropek, A. (2012). Czy warto urodzić się w styczniu? Wiek biologiczny a wyniki egzaminacyjne. W: B. Niemierko i M. K. Szmigiel (red.), *Regionalne i lokalne diagnozy edukacyjne. Materiały XVIII Konferencji Diagnostyki Edukacyjnej* (s. 52–58). Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.
- Dolata, R. i Szaleniec, H. (2012). Funkcje krajowych egzaminów w systemie edukacji. *Polityka Społeczna* (numer tematyczny 1), 37–41.
- Dolata, R., Hawrot, A., Humenny, G., Jasińska, A., Koniewski, M., Majkut, P. i Żółtak, T. (2013). *Trafność metody edukacyjnej wartości dodanej dla gimnazjów*. Warszawa: Instytut Badań Edukacyjnych.
- Dolata, R., Hawrot, A., Humenny, G., Jasińska-Maciążek, A., Koniewski, M. i Majkut, P. (2014). *Kontekstowy model oceny efektywności nauczania po pierwszym etapie edukacyjnym*. Warszawa: Instytut Badań Edukacyjnych.
- Domanski, H. i Pokropek, A. (2011). *Podziały terytorialne, globalizacja a nierówności społeczne: wprowadzenie do modeli wielopoziomowych*. Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii PAN.
- Frydrychowicz, A., Koźniewska, E., Matuszewski, A. i Zwierzyńska, E. (2006). *Skala gotowości szkolnej. Podręcznik*. Warszawa: Centrum Metodyczne Pomocy Psychologiczno-Pedagogicznej.
- Hægeland, T., Kirkebøen, L., Raau, O. i Salvanes, K. (2005). *School performance indicators for Oslo, Reports 2005/36*. Oslo: Statistics Norway.
- Hutchison, D. i Sharp, C. (1999). A lasting legacy? The persistence of season of birth effects. Referat wygłoszony podczas the British Educational Research Association Conference, University of Sussex, Brighton. Pobrano z <https://www.nfer.ac.uk/publications/44411/44411.pdf>
- Jasińska, A. i Modzelewski, M. (2013). Międzyszkolne zróżnicowanie wyników nauczania po pierwszym etapie kształcenia. W: B. Niemierko i M. K. Szmigiel (red.), *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych* (s. 165–178). Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.
- Jasińska-Maciążek, A. i Modzelewski, M. (2014). Modele analizy zróżnicowania wyników nauczania. W: R. Dolata (red.), *Czy szkoła ma znaczenie? Analiza zróżnicowania efektywności nauczania na pierwszym etapie edukacyjnym* (s. 65–92). Warszawa: Instytut Badań Edukacyjnych.
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: implications for practice in the 21st century. *School Psychology Review*, 30(3), 420–437.
- Karwowski, M. i Dziedziewicz, D. (2012). *Test umiejętności na starcie szkolnym. Podręcznik*. Warszawa: Instytut Badań Edukacyjnych.
- Konarzewski, K. (2013). Wiek startu szkolnego a osiągnięcia w nauce w okresie wczesnoszkolnym. *Edukacja*, 124(4), 5–19.
- Langer, P., Kalk, J. M. i Searls, D. (1984). Age of admission and trends in achievement: a comparison of Blacks and Caucasians. *American Educational Research Journal*, 21(1), 61–78.

- Lee, J. i Fish, R. M. (2010). International and interstate gaps in value-added math achievement: multilevel instrumental variable analysis of age effect and grade effect. *American Journal of Education*, 117(1), 109–137.
- Lincove, J. A. i Painter, G. (2006). Does the age that children start kindergarten matter? Evidence of long-term educational and social outcomes. *Educational Evaluation and Policy Analysis*, 28(2), 153–179. doi:10.3102/01623737028002153
- Malec, D. J. (2008). Superpopulation. W: P. J. Lavrakas (red.), *Encyclopedia of survey research methods* (t. 2, s. 855–856). Thousand Oaks: Sage.
- Martin, A. J. (2009). Age appropriateness and motivation, engagement, and performance in high school: effects of age within cohort, grade retention, and delayed school entry. *Journal of Educational Psychology*, 101(1), 101–114.
- Martin, R. P., Foels, P., Clanton, G. i Moon, K. (2004). Season of birth is related to child retention rates, achievement, and rate of diagnosis of specific LD. *Journal of Learning Disabilities*, 37(4), 307–317.
- May, D. C. i Kundert, D. K. (1995). Does delayed school entry reduce later grade retentions and use of special education services? *Remedial and Special Education*, 16(5), 288–94.
- Mayer, S. E. i Knutson, D. (1999). Does the timing of school affect how much children learn? W: S. E. Mayer i P. E. Peterson (red.), *Earning and learning: how schools matter* (s. 79–102). Washington: Brookings Institution Press.
- Morrison, F. J. F. J., Griffith, E. M. i Alberts, D. M. (1997). Nature-nurture in the classroom: entrance age, school readiness, and learning in children. *Developmental Psychology*, 33(2), 254–262.
- Organisation for Economic Co-operation and Development (2008). *Measuring improvements in learning outcomes: best practices to assess the value-added of schools*. Paris: OECD Publishing.
- Raudenbush, S. W. i Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Thousand Oaks: Sage.
- Raudenbush, S. W. i Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307–335.
- Ray, A. (2006). *School value added measures in England*. Pobrano z <http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/eOrderingDownload/RW85.pdf>.
- Ray, A., McCormack, T. i Evans, H. (2009). Value-added in English schools. *Education Finance and Policy*, 4(4), 415–438.
- Sharp, C. (2002). School starting age: European policy and recent research. Referat wygłoszone podczas Seminarium LGA “When Should Our Children Start School?”, Londyn.
- Smith, J. (2009). Can regression discontinuity help answer an age-old question in education? The effect of age on elementary and secondary school achievement. *The B.E. Journal of Economic Analysis i Policy*, 9(1). doi: 10.2202/1935-1682.2221
- Snijders, T. A. B. i Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (wyd. 2). Los Angeles: Sage.
- Stipek, D. i Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Applied Developmental Psychology*, 22(2), 175–189.
- Sweetland, J. D. i De Simone, P. A. (1987). Age of entry, sex, and academic achievement in elementary school children. *Psychology in the Schools*, 24(4), 406–412.
- Sykes, E., Bell, J. i Vidal Rodeiro, C. (2009). *Birth-date effects: a review of the literature from 1990-on*. Pobrano z <http://www.cambridgeassessment.org.uk/Images/109784-birthdate-effects-a-review-of-the-literature-from-1990-on.pdf>
- Verachtert, P., De Fraine, B., Onghena, P. i Ghesquière, P. (2010). Season of birth and school success in the early years of primary education. *Oxford Review of Education*, 36(3), 285–306.
- Wei, X., Blackorby, J. i Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, 78(1), 89–106.
- Wen, X., Bulotsky-Shearer, R. J., Hahs-Vaughn, D. L. i Korfmacher, J. (2012). Head Start program quality: examination of classroom quality and parent involvement in predicting children’s vocabulary, literacy, and mathematics achievement trajectories. *Early Childhood Research Quarterly*, 27(4), 640–653.
- Żółtak, T. (2013). *Statystyczne modelowanie wskaźników edukacyjnej wartości dodanej - podsumowanie polskich doświadczeń* [Raport tematyczny z badania]. Warszawa: Instytut Badań Edukacyjnych. Pobrano z http://2013.ewd.edu.pl/modele/modele_2013_uzup.pdf