



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE  **EDUKACYJNA
WARTOŚĆ
DODANA**

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



RAPORT TEMATYCZNY Z BADANIA

Analiza procesów różnicowania się gimnazjów i szkół podstawowych

Edycja 2012



Autorzy:

dr hab. Roman Dolata, prof. IBE

Aleksandra Jasińska

Michał Modzelewski

Raport pod tytułem „Wykorzystanie egzaminów krajowych jako instrumentu polityki oświatowej na przykładzie procesu różnicowania się gimnazjów w dużych miastach” ukazał się w miesięczniku Polityka Społeczna, Nr tematyczny 1, 2012.

Wydawca:

Instytut Badań Edukacyjnych

ul. Górczewska 8

01-180 Warszawa

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

Instytut Badań Edukacyjnych, Warszawa, 2012

Publikacja opracowana w ramach projektu systemowego: *Badania dotyczące rozwoju metodologii szacowania wskaźnika edukacyjnej wartości dodanej*, współfinansowanego przez Unię Europejską ze środków Europejskiego Funduszu Społecznego, realizowanego przez Instytut Badań Edukacyjnych.

Egzemplarz bezpłatny

Spis Treści

Problem.....	4
Proces różnicowanie się gimnazjów ze względu na osiągnięcia	5
Wykorzystane dane i metoda analizy	7
Zróżnicowanie międzyszkolne w perspektywie porównawczej.....	8
Zróżnicowanie międzyszkolne gimnazjów w świetle wyników egzaminu	9
Różnicowanie się gimnazjów w perspektywie lokalnej.....	12
Konkluzje	16
Bibliografia	17

Problem

Czy polskiej oświacie potrzebne są krajowe egzaminy wykorzystujące standaryzowane testy osiągnięć szkolnych? Gdy w debacie publicznej lub profesjonalnej pada to pytanie, zwykle uwaga dyskutantów koncentruje się na jakości używanych testów, zwrotnym wpływie egzaminów na nauczanie (tzw. nauczaniu pod testy) lub selekcyjnej funkcji egzaminów. W tym tekście chcielibyśmy zająć się innymi aspektami tego elementu systemu oświatowego. Egzaminy krajowe mogą mianowicie dostarczać informacji o zachodzących w systemie oświaty procesach (funkcja monitorująca) oraz mogą być instrumentem zarządzania tymi procesami (funkcja ewaluacyjna). Funkcje te mogą być realizowane zarówno na poziomie krajowym jak i lokalnym. Funkcję ewaluacyjną egzaminów można bezpośrednio łączyć z ideą *accountability* rozumianą jako obarczenie instytucji publicznych, w tym władz lokalnych i szkół, odpowiedzialnością za skutki swoich działań. Idea *accountability* nabiera w polityce oświatowej szczególnego znaczenia w sytuacji, gdy szkoły mają coraz więcej autonomii i równocześnie wprowadzane są na poziomie krajowym lub lokalnym elementy polityki rynkowej (Waslander, Pater, van der Weide 2010). Decentralizacja i urynkowanie oświaty może sprawiać, że w systemie zachodzą może wiele spontanicznych procesów (Rado 2010). Na przykładzie jednego z takich procesów - różnicowania się w miastach gimnazjów ze względu na osiągnięcia szkolne (Dolata 2008, 2010) – zobaczymy jak egzaminy krajowe mogą pełnić funkcję monitorującą i ewaluacyjną.

Proces różnicowanie się gimnazjów ze względu na osiągnięcia

W Polsce wdrożono w 1999 roku jedną z bardziej masywnych reform edukacyjnych w Europie. Dwa najważniejsze jej elementy to wprowadzenie gimnazjów i krajowych egzaminów wykorzystujących standaryzowane testy osiągnięć szkolnych. Zgodnie z założeniami reformy z 1999 r. po 6 latach nauki w szkole podstawowej uczniowie mieli być bez selekcji lokowani w 3-letnich gimnazjach. Stworzenie gimnazjów miało zapewnić lepsze dostosowanie działań wychowawczych do potrzeb rozwojowych 12-15-latków. Czy podział etapu jednolitego kształcenia na szkołę podstawową i gimnazjum nie uruchomił jednak ukrytych procesów selekcyjnych polegających z jednej strony na segregacji uczniów na progu gimnazjum ze względu na uprzednie osiągnięcia, z drugiej na skorelowany z nią proces różnicowania się gimnazjów ze względu na efektywność kształcenia? Za hipotezą mówiącą, że proces taki mógłby zajść przemawiają poza samym wprowadzeniem gimnazjum dodatkowe przesłanki. Pierwsza z nich to zbiegnięcie się wprowadzenia gimnazjów z silnym spadkowym trendem liczebności szkolnej populacji. Druga to osłabienie urzędowej rejonizacji szkół. Trzecia w końcu to wzrost popularności w Polsce neoliberalnych, quasi-rynkowych pomysłów na poprawę jakości szkół (Dolata 2008, Potulicka, Rutkowiak 2010). Wszystkie te przesłanki ze szczególnym nasileniem występują w dużych miastach, co pozwala sformułować hipotezę, że interesujący nas proces selekcji zachodzić będzie szczególnie intensywnie w tym właśnie segmencie systemu szkół gimnazjalnych. Jednym z kluczowych wskaźników procesów selekcji szkolnych jest międzyszkolne zróżnicowanie wyników nauczania. Na nim skupimy uwagę.

Zanim przejdziemy do analizy interesującego nas zjawiska, by wykazać użyteczność danych pochodzących z krajowych egzaminów do jego monitorowania, zapoznamy się z kontekstem teoretycznym. Zaczniemy od rozważań pojęciowych. W jakich ramach pojęciowych umieścić badania nad różnicowaniem się gimnazjów? G. Therborn (2006) proponuje ciekawą sieć pojęć pozwalających opisywać różne aspekty społecznych nierówności we współczesnym świecie. Mechanizmy powstawania nierówności to różnicowanie się osiągnięć (Therborn używa pojęcia „wyprzedzanie”), wykluczanie społeczne, hierarchizacja i eksploatacja. Dziedziny nierówności to życie i zdrowie, wolność i uznanie społeczne oraz zasoby materialne i symboliczne. Nierówności możemy analizować w obrębie małych populacji lokalnych, całego społeczeństwa lub systemu globalnego. Jednostką analizy nierówności może być pojedynczy człowiek, grupa społeczna lub społeczeństwo. Tę ostatnią listę warto uzupełnić o jeszcze jeden element – instytucję. Korzystając z tych pojęć możemy różnicowanie się szkół ze względu na wyniki nauczania opisać jako proces hierarchizacji populacji - z założenia - jednolitych szkół w skali całego społeczeństwa lub jego segmentów (np. szkoły wielkomiejskie lub szkoły w mieście X) ze względu na dobra symboliczne. Ten proces hierarchizacji szkół sam w sobie wytwarza społeczne nierówności (np. poprzez tabele ligowe), ale jest również interesujący z punktu widzenia

potencjalnego wzmocnienia procesu różnicowania się osiągnięć zarówno na poziomie jednostek (wzrost wariacji wyników) jak i grupowym (wzrost dystansów międzygrupowych ze względu na osiągnięcia szkolne).

Kontynuując analizy teoretyczne rozważmy, jakie mechanizmy mogą odpowiadać za zróżnicowanie szkół ze względu na wyniki nauczania? Dla zrozumienia interesującego nas procesu kluczowe znaczenie mogą mieć segregacje przestrzenne (Jargowsky 1996, Mayer 2002). Miarą nasilenia tych segregacji jest to, w jakim stopniu dany podział przestrzenny wyjaśnia zmienność danej cechy. Dla zróżnicowania szkół ze względu na wyniki nauczania znaczenie ma przede wszystkim segregacja ze względu na wykształcenie rodziców uczniów i jego korelaty (przekaz genetyczny, kapitał kulturowy i społeczny). Jeżeli dany system szkół funkcjonuje w społeczeństwie o wysokim poziomie segregacji przestrzennych, to mimo jednolitości systemu oświaty, wskaźnik zróżnicowania międzyszkolnego będzie wysoki.

Kolejny mechanizm prowadzący potencjalnie do różnicowania się szkół to procesy autoselekcji na progu szkoły. Wiążą się one z funkcjonowaniem quasi-rynków edukacyjnych (Gorard, Fitz 2000), zwanych inaczej rynkami publicznymi (Woods, Bagley, Glatter 1998). Wyniki badań wskazują, że mechanizmy rynkowe, wprowadzone często pod hasłem rodzicielskiego prawa do wyboru szkoły, uruchamiają na progach szkolnych procesy autoselekcji, które w pewnych warunkach nasilają procesy różnicowania się szkół ze względu na skład społeczny i osiągnięcia (Ball 1993, Goldstein, Noden 2003, Murawska 2004). Nie zawsze jednak mechanizmy rynkowe prowadzą do segregacji i wzrostu zróżnicowania. Wyniki analiz tego problemu nie dają spójnego obrazu (Waslander, Pater, van der Weide 2010). W Polsce badania nad funkcjonowaniem rynków publicznych w oświacie nie są częste (Herczyński i Herbst 2005; Dolata 2008), ale uzyskane wyniki raczej potwierdzają, że procesy autoselekcji związane ze swobodnym wyborem szkoły zwiększają międzyszkolne zróżnicowanie.

Następny potencjalny mechanizm powstawiania interesującego nas fenomenu to procesy selekcji na progu gimnazjum. Gdy w systemie oświaty w danym punkcie kariery szkolnej funkcjonuje próg selekcyjny, w oczywisty sposób przyczynia się to do silnego zróżnicowania składu szkół „za progiem” ze względu na uprzednie osiągnięcia szkolne i zasoby z tym powiązane. Zróżnicowanie na wejściu w oczywisty sposób prowadzi do zróżnicowania na wyjściu.

Również tzw. efekt rówieśników może przyczyniać się do zróżnicowania szkół ze względu na osiągnięcia. Wpływ rówieśników jest przedmiotem licznych badań, głównie w kontekście problemów segregacji w oświacie i skuteczności różnego typu innowacji dydaktycznych wykorzystujących zasoby rówieśnicze (Wilkinson i in. 2000, Angrist, Lang 2002, Markman, Hanushek, Kain, Rivkin 2003; Dolata 2008). Otrzymywane rezultaty nie zawsze są jednak jednoznaczne, pojawiają się doniesienia wskazujące, że efekt rówieśników jest artefaktem statystycznym (tzw. efekt fantomowy: Harker, Tymms 2004, Pokropek 2011).

Na koniec należy wskazać na mechanizm różnicowania się efektywności nauczania. Prawdopodobnie szkoły mające dobrą opinię na rynku przyciągają nie

tylko dobrych uczniów, ale z biegiem czasu również lepszych nauczycieli, dodatkowe środki finansowe i rzeczowe. To może prowadzić do międzyszkolnego zróżnicowania efektywności nauczania.

Wykorzystane dane i metoda analizy

W analizach wykorzystano dane krajowe z części humanistycznej egzaminu gimnazjalnego dla lat 2002-2012 oraz, pomocniczo, analogiczne dane ze sprawdzianu w klasie VI szkoły podstawowej. Dla planowanych analiz ważne znaczenie mają własności psychometryczne testów egzaminacyjnych. Testy gimnazjalne są standaryzowanymi testami ogólnokrajowymi tworzonymi z wykorzystaniem klasycznej teorii testów. Kolejne edycje nie są zrównywane, a wyniki są przedstawiane w postaci surowej liczby punktów. W związku z tym przed analizami wyniki surowe dla każdego roku znormalizowano (metoda Bloma) i wystandaryzowano (średnia w kraju 100, odchylenie standardowe 15), a następnie z analizy wykluczono dane pochodzące ze szkół mniejszych niż 5 uczniów (na poziomie ostatniej klasy danej szkoły).

W zakresie trafności jedynym gwarantem stałości treściowej testów jest dość szczegółowy plan testów. W 2012 roku ten plan w znaczący sposób uległ zmianie, co może zaburzać wyniki analizy trendu w zakresie zróżnicowania międzyszkolnego. Drugi kluczowy aspekt jakości testów, trafność, też pozostawia wiele do życzenia. Wartości współczynnika rzetelności alfa Cronbacha dla części humanistycznej testów gimnazjalnych¹ w latach 2002-2012 wahały się, najmniej rzetelny był test w 2002 r. (0,78), najbardziej rzetelny w 2012 (0,89). Jak na testy wysokiej stawki rzetelność jest zbyt niska, ale dla celów naszej analizy wystarczająca. Występują znaczące wahania w rzetelności, ale nie są one skorelowane z trendem w zakresie zróżnicowania międzyszkolnego.

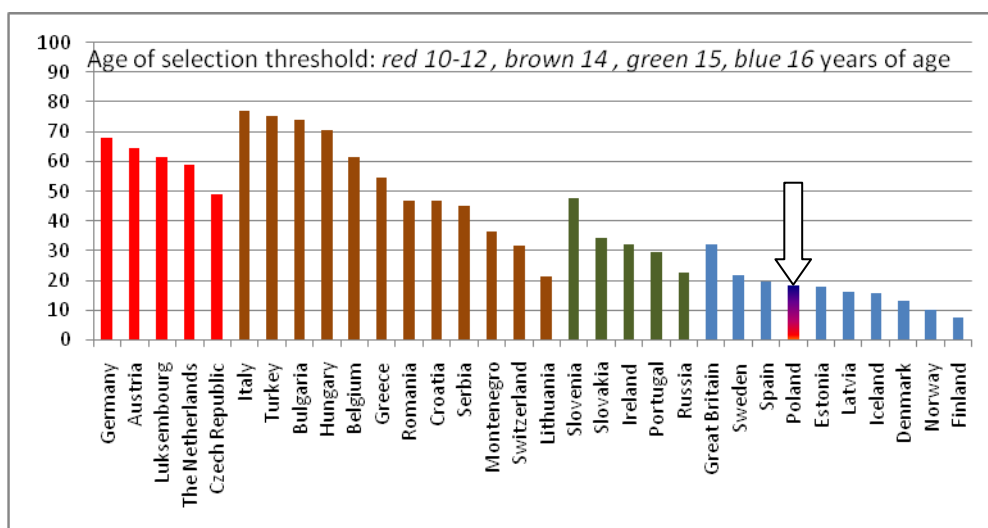
Obliczenie wskaźnika międzyszkolnego zróżnicowania wyników nauczania to klasyczny problem dekompozycji wariancji. W zaprezentowanych w tym artykule analizach najpierw podzielono gimnazja ze względu na lokalizację na 4 grupy: wieś, miasta do 20 tysięcy, miasta od 20 do 100 tysięcy i miasta ponad 100 tysięcy mieszkańców. Następnie oddzielnie dla każdej grupy gimnazjów całkowitą wariancję wyników testu egzaminacyjnego rozłożono na trzy części: wariancję wewnątrzszkolną, wariancję międzyszkolną i wariancję między gminami/dzielnicami. Na trzecim poziomie analizy wybrano gminy/dzielnice, ponieważ organem prowadzącym dla gimnazjów są te właśnie jednostki samorządowe. Następnie wariancję międzyszkolną podzielono przez wariancję całkowitą, a wynik wyrażono w procentach. Interesujący nas zatem wskaźnik mówi

¹ Na potrzeby analizy rzetelności zadania wielopunktowe i oceniane wielokryterialnie dychotomizowano po medianie. Stosowany w raportach CKE sposób liczenia rzetelności traktujący poszczególne kryteria oceny wykonania zadań złożonych jak niezależne pozycje testowe, jest niepoprawny.

nam, jaki odsetek wariancji całkowitej to wariancja międzyszkolna. Choć na ogólnym poziomie oszacowanie wskaźnika jawi się prosto, to w praktyce statystycznej mamy bardzo wiele algorytmów dekompozycji wariancji. W analizach użyto trzypoziomowych modeli pustych, a składniki wariancji szacowano metodą *full maximum likelihood* przy użyciu specjalistycznego oprogramowania do analiz hierarchicznych HLM 6.0. W literaturze statystycznej znajdujemy ostrzeżenie, że ta metoda może zawyżać oszacowania wariancji na wyższych poziomach analizy, jednak badacze-praktycy polecają ten właśnie sposób estymacji wariancji (Domański, Pokropek 2011). Trzeba podkreślić, że choć różne metody szacowania wskaźnika zróżnicowania międzyszkolnego przynoszą trochę inne oszacowania interesującej nas statystyki, to obraz trendu czasowego zawsze jest taki sam.

Zróżnicowanie międzyszkolne w perspektywie porównawczej

Zanim przedstawimy wyniki analiz na podstawie danych egzaminacyjnych, przejrzymy się interesującemu nas zjawisku w perspektywie porównawczej na podstawie wyników badania PISA 2009. W analizach wykorzystano szacunki wskaźnika międzyszkolnego zróżnicowania wyników z zakresu czytania podane w oficjalnym raporcie OECD. Poniższa tabela pokazuje wartości wskaźnika dla krajów europejskich. Kraje pogrupowano ze względu na wiek, w którym w systemie szkolnym występuje pierwszy, formalny próg selekcyjny, po który następuje zróżnicowanie ścieżek edukacyjnych, głównie na ścieżkę kształcenia akademickiego i zawodowego.



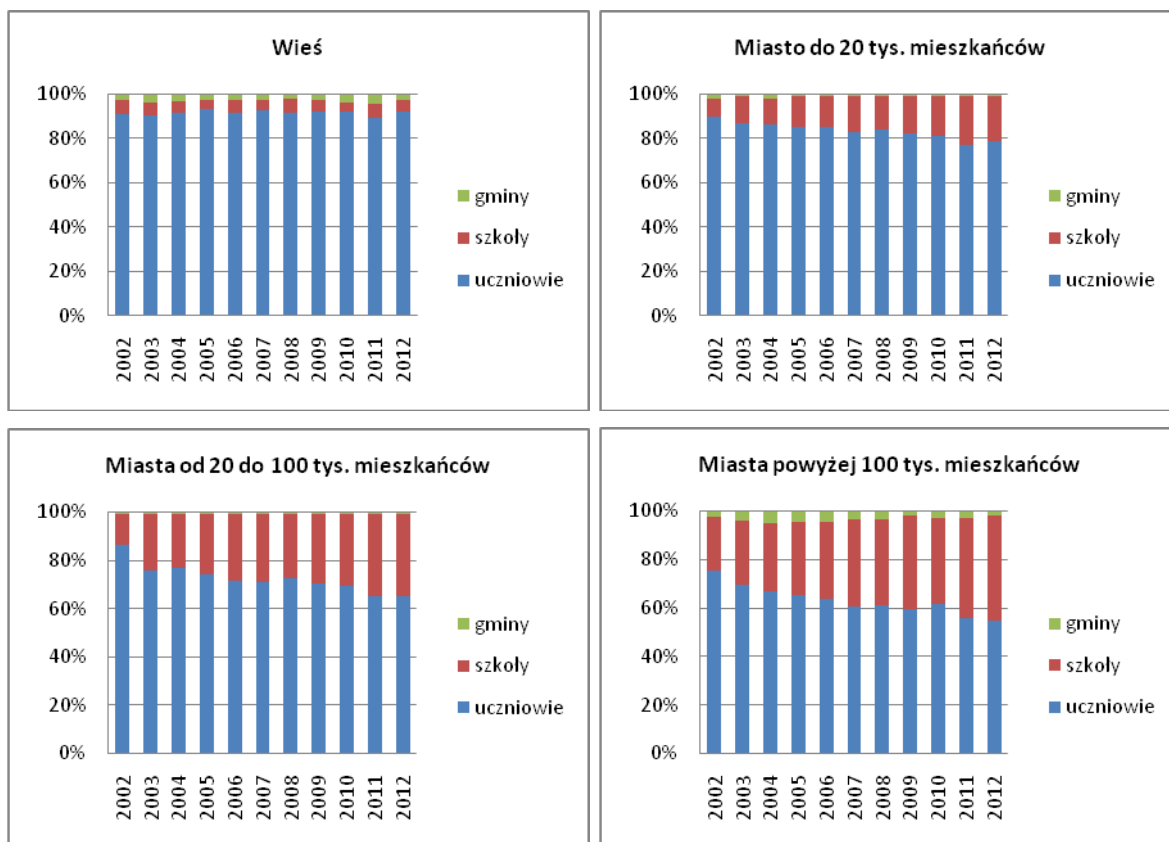
Wykres 1. Zróżnicowanie międzyszkolne wyników testu czytania. Wariancja międzyszkolna jako procent wariancji całkowitej w krajach, PISA 2009

Wykres wyraźnie pokazuje silną i dość oczywistą zależność: jeżeli badanie przeprowadzono przed progiem selekcyjnym (PISA obejmuje piętnastolatki), międzyszkolne zróżnicowanie wyników testowych jest znacznie niższe, jeżeli po progu, jest znacząco wyższe. Polska, dzięki wprowadzeniu w 1999 roku gimnazjum, lokuje się w towarzystwie krajów o niskim poziomie zróżnicowania międzyszkolnego. Znalezienie się w grupie krajów, w których po II wojnie światowej podjęto reformy szkolnictwa wydłużające czas jednolitego kształcenia (np. głośna reforma szwedzka firmowana przez T. Husena wprowadzająca jednolitą szkołę 9-letnią), to raczej powód do zadowolenia. Zadowolenie mąci jednak świadomość, że już na poziomie I klasy szkoły ponadgimnazjalnej poziom interesującego nas zróżnicowania dramatycznie wzrośnie. Wskaźnik wyliczony dla Polski na podstawie wyników PISA 2000 (piętnastolatki uczyły się wtedy w Polsce jeszcze w I klasie szkół ponadpodstawowych) lokował nas w grupie krajów o najwyższym poziomie międzyszkolnego zróżnicowania wyników nauczania. Obraz ten potwierdziły wyniki z opcji narodowej badania PISA 2006 (opcja narodowa objęła uczniów klas I i II szkół ponadgimnazjalnych).

Zróżnicowanie międzyszkolne gimnazjów w świetle wyników egzaminu

Jak wygląda interesujące nas zjawisko w świetle wyników egzaminu gimnazjalnego? W analizach ograniczono się do wyników w części humanistycznej egzaminu gimnazjalnego. Jest to podyktowane faktem, że w wypadku części matematyczno-przyrodniczej obserwujemy bardziej złożony – choć nie sprzeczny z wynikami niżej referowanymi – obraz procesów różnicowania gimnazjów (patrz: Dolata 2008).

Poniższe wykresy pokazują poziom i dynamikę procesu różnicowania się gimnazjów. Oszacowania składowych wariacji zostały przedstawione w podziale gimnazjów ze względu na lokalizację: tereny wiejskie i trzy kategorie wielkości miast. Podział wyników ze względu na lokalizację gimnazjum wyjaśnia bardzo niewielką część wariacji całkowitej, a wariacje całkowite wyników egzaminu gimnazjalnego w wyróżnionych segmentach gimnazjów niewiele się od siebie różnią. Można więc za podstawę procentowania przyjąć wariację całkowitą oddzielnie szacowaną dla każdej lokalizacji

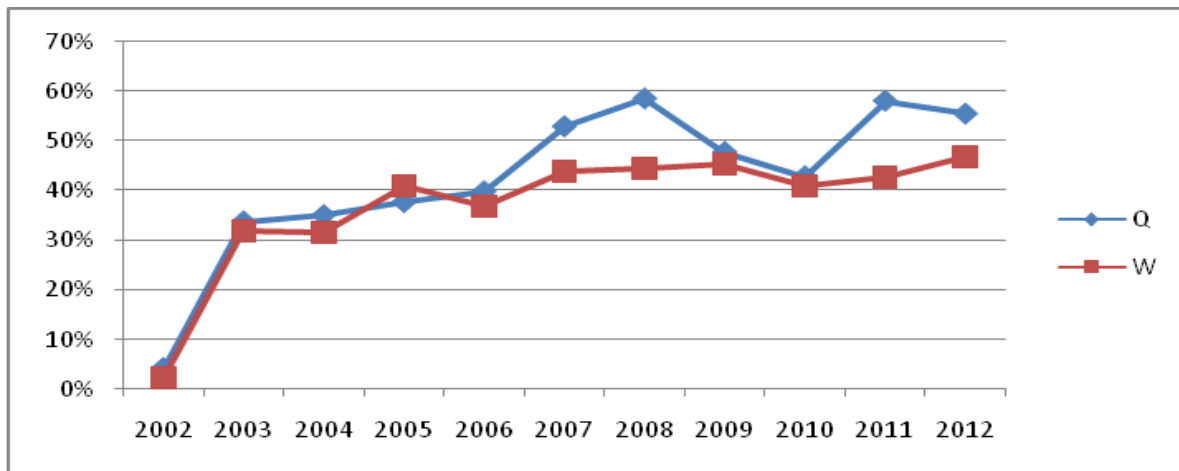


Wykres 2. Oszacowania udziału wariacji międzygminnej, międzyszkolnej i wewnątrzszkolnej w wariacji całkowitej wyniku egzaminu gimnazjalnego, część humanistyczna. Analiza dynamiki ze względu na lokalizację gimnazjum

Zanim przejdziemy do różnicowania międzyszkolnego wyników egzaminu, należy zwrócić uwagę, że wariacja międzygminna stanowi niewielką część wariacji całkowitej wyników egzaminacyjnych i nie notuje się w tym zakresie żadnego znaczącego trendu.

Teraz można zająć się już wariacją międzyszkolną. W 2002 roku oszacowany poziom różnicowania międzyszkolnego w czterech analizowanych segmentach gimnazjów wynosił od 6% w wypadku gimnazjów zlokalizowanych na wsi, do 22% dla grupy gimnazjów w dużych miastach. Z biegiem lat na wsi nie notujemy znaczących zmian, ale im wyższy poziom urbanizacji, tym silniejsza dynamika wzrostowa różnicowania się szkół. Proces różnicowania się ze względu na wyniki kształcenia gimnazjów zlokalizowanych w dużych miastach ma już bardzo dużą dynamikę. Tak więc analizy wyników egzaminacyjnych wskazują, że w dużych miastach mamy do czynienia z silnym i rosnącym różnicowaniem gimnazjów. W 2012 roku oszacowana wariacja międzyszkolna stanowi już ponad 40% wariacji całkowitej.

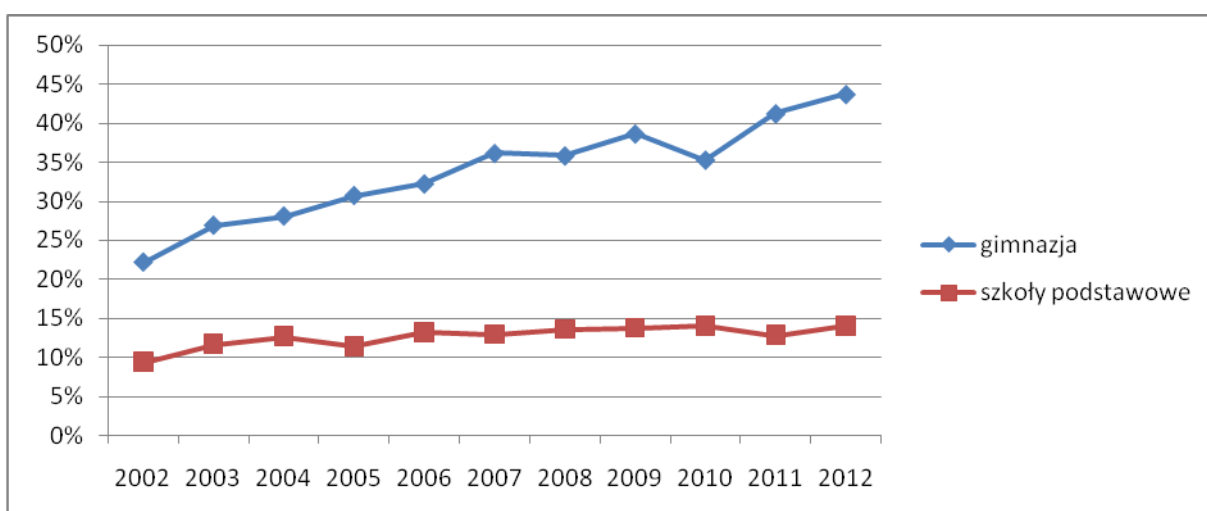
Analizy dla poszczególnych dużych miast potwierdzają ten obraz. Poniższy wykres przedstawia dynamikę procesu różnicowania gimnazjów ze względu na wyniki nauczania w dwóch przykładowych miastach. Zostały one symbolicznie oznaczone literami Q i W.



Wykres 3. Dynamika procesu różnicowania gimnazjów ze względu na wyniki nauczania w dwóch przykładowych miastach

Wykres wskazuje, że w przypadku konkretnych miast możemy obserwować bardziej złożony obraz dynamiki analizowanego zjawiska, ale ogólny wzrostowy wzorzec trendu występuje we wszystkich największych miastach w Polsce.

Aby zrozumieć proces różnicowania się gimnazjów warto porównać jego dynamikę z analogicznym zjawiskiem na poziomie szkoły podstawowej. W analizach wykorzystano dane ze sprawdzianu w klasie VI dla dużych miast. Poniższy wykres pokazuje zróżnicowanie segmentu gimnazjów zlokalizowanych w dużych miastach na tle zróżnicowania segmentu szkół podstawowych o analogicznej lokalizacji. Wyniki te pozwalają pośrednio zweryfikować hipotezę wpływu segregacji przestrzennej w dużych miastach (gettyzacji) na proces różnicowania się gimnazjów.



Wykres 4. Porównanie dynamiki różnicowania się gimnazjów w dużych miastach z analogicznym procesem na poziomie szkoły podstawowej

Wartość zróżnicowania gimnazjów w 2012 roku trzeba porównać z poziomem zróżnicowania szkół podstawowych trzy lata wcześniej, czyli w 2009 roku. To właśnie ci uczniowie trzy lata później przystąpili do egzaminu gimnazjalnego. 2012. Oszacowanie wskaźnika dla szkół podstawowych wynosi 14%, czyli o około 30 punktów procentowych mniej, niż w wypadku gimnazjów! Ponieważ szkoły podstawowe są bardziej „podatne” na wpływ procesów segregacji przestrzennej (mniejsze rejony działania) niż gimnazja, wskazuje to na niewielki udział tego czynnika w procesie różnicowania się wielkomiejskich gimnazjów.

Jak widzimy, analizy z wykorzystaniem wyników krajowego systemu egzaminacyjnego pozwoliły odkryć i opisać niezwykle ważny z punktu widzenia polityki oświatowej proces. W miastach, szczególnie dużych, obserwujemy dynamiczny proces różnicowania się miejskich systemów gimnazjalnych ze względu na wyniki kształcenia. Silne zróżnicowanie ze względu na wyniki kształcenia gimnazjów w dużych miastach nie da się wyjaśnić samoistnie zachodzącą segregacją przestrzenną. Innymi słowy, wprowadzenie w 1999 roku gimnazjów uruchomiło w miastach ukryty proces selekcyjny na poziomie 12 roku życia (gdy dotrze do gimnazjów fala uczniów po obniżeniu wieku szkolnego, próg selekcyjny obniży się do 11 roku życia!). To niezwykle ważne odkrycie dla badacza procesów edukacyjnych stanowi przykład wielkiej użyteczności krajowego systemu egzaminacyjnego z punktu widzenia funkcji monitorującej.

Co oznacza to odkrycie dla krajowej polityki oświatowej? Ocena zależy od przesłanek ideologicznych i głębszego jeszcze zrozumienia samego procesu różnicowania się gimnazjów, ale jedno jest pewne – zjawisko to powinno być przedmiotem dyskusji zarówno na poziomie polityki krajowej, jaki i polityk lokalnych. Władze samorządowe i same szkoły powinny mieć dostęp do informacji o nasileniu tego zjawiska w danej gminie, mieście lub dzielnicy miasta. W ten sposób przechodzimy do drugiej części tekstu, w której omówimy instrumenty, jakimi dysponują władze lokalne, by monitorować ten proces.

Różnicowanie się gimnazjów w perspektywie lokalnej

Monitorowanie różnicowania się gimnazjów na poziomie lokalnym wymaga dostępu do odpowiednio przetworzonych danych egzaminacyjnych. Sposób komunikowania wyników przez okręgowe komisje egzaminacyjne niestety nie ułatwia śledzenia interesującego nas procesu, ogranicza się bowiem do prezentacji wyników surowych. Jednak dzięki pracom nad metodą edukacyjnej wartości dodanej prowadzonym w ramach przywoływanego już w tym artykule projektu (patrz przypis 1), zarówno szkoły jak i władze lokalne uzyskują dostęp do wyników analiz, które zdecydowanie lepiej obrazują interesujące nas zjawisko. Zanim na przykładzie wybranego miasta przyjrzymy się instrumentom analitycznym będącymi do dyspozycji władz lokalnych, musimy się zapoznać z polską implementacją metody edukacyjnej wartości dodanej (EWD).

Metoda ta pozwala na podstawie odpowiednio przetworzonych danych egzaminacyjnych uzyskać miary efektywności nauczania, czyli postępu, jaki robią uczniowie w trakcie nauki w danej szkole. Metoda EWD dostarcza bardziej przydatnych dla polityki oświatowej - w stosunku do wyników końcowych szkoły - wskaźników efektywności pracy szkoły (OECD, 2008). Efektywność nauczania w danej szkole często utożsamiana jest z jej wynikami egzaminacyjnymi. Gdyby wszystkie szkoły w pracowały z uczniami o tym samym poziomie uprzednich osiągnięć szkolnych, tych samych zasobach rodzinnych i dysponowały tymi samymi zasobami materialnymi, faktycznie wyniki egzaminu końcowego byłyby dobrą miarą efektywności. Jednakże tak nie jest. Metoda EWD pozwala uwzględnić zróżnicowanie szkół ze względu na zasoby „na wejściu”. Najczęściej miarą zasobów na wejściu są wyniki testów mierzących uprzednie osiągnięcia szkolne.

W 2005 grupa ekspertów pracująca pod egidą Centralnej Komisji Egzaminacyjnej rozpoczęła pracę nad metodami szacowania wskaźników EWD przydatnych zarówno w ewaluacji wewnątrzszkolnej jak i interesującej nas ewaluacji na poziomie lokalnym. W tej chwili użytkownicy mają do dyspozycji dwa podstawowe narzędzia. Pierwsze z nich to tzw. Kalkulator EWD. To prosta aplikacja komputerowa pozwalająca szkołom szacować wskaźniki EWD dla oddziałów klasowych lub innych podgrup uczniów. Drugie narzędzie to prezentowane graficznie, dostępne publicznie on-line trzyletnie ewaluacyjne wskaźniki egzaminacyjne dla gimnazjów, liceów ogólnokształcących i techników (www.ewd.edu.pl). Wskaźniki mają charakter względny, czyli służą do porównywania szkół z populacją gimnazjów w Polsce. Można wybrać też inne układy odniesienia: gimnazja w województwie, powiecie lub gminie/dzielnicy. W szacowaniu wskaźników egzaminacyjnych wykorzystuje się wyniki z trzech kolejnych lat. Obliczanie trzyletnich wskaźników uzasadnione jest z kilku powodów. Najważniejszy to fakt, że wyniki egzaminacyjne, jak każdego innego testu, obarczone są niepewnością pomiarową. Im większą pulą danych dysponujemy, tym mniejszy wpływ tej niepewności na szacowanie wskaźników dla szkoły. Używając wyników z trzech kolejnych lat, zyskuje się mniej więcej trzy razy więcej danych, a duża ilość informacji oznacza większą precyzję szacowania. Ewaluacja szkoły nie jest dzięki temu uzależniona od losowych czynników, zwiększających lub zmniejszających wartość wskaźników.

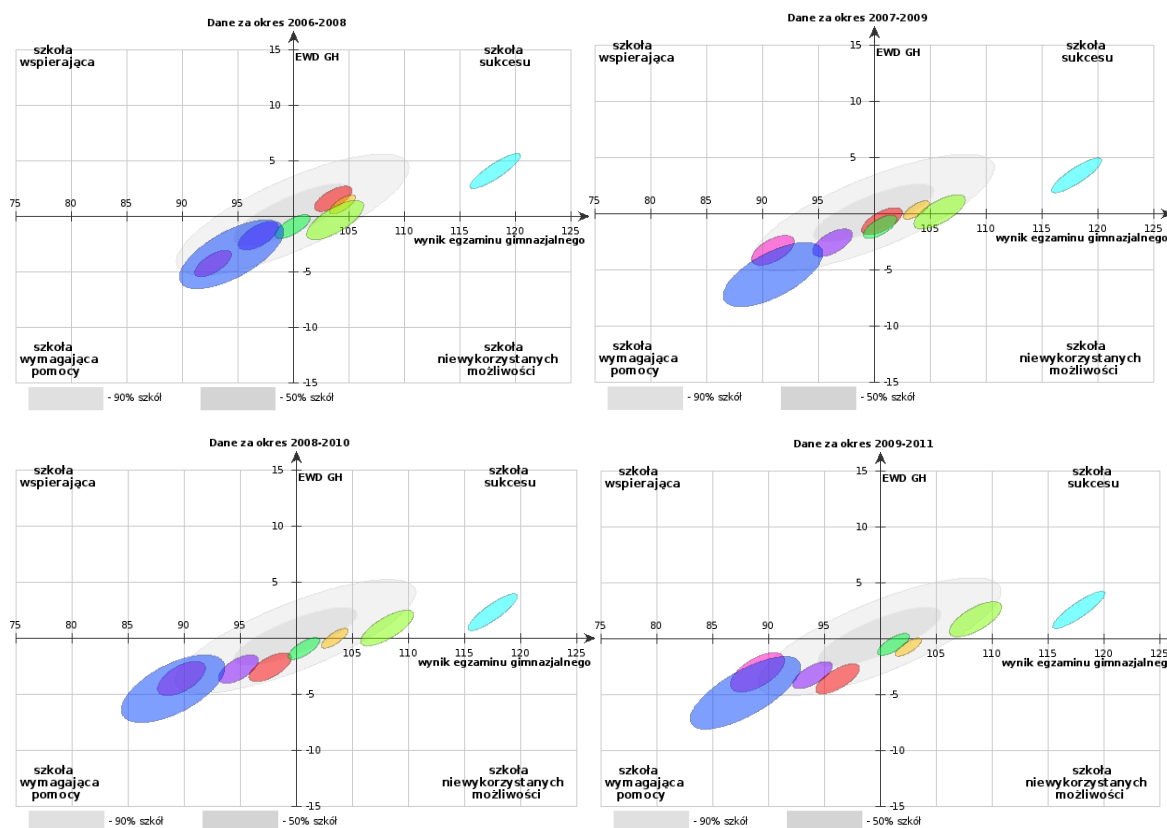
Ponieważ surowe wyniki egzaminów nie są porównywalne między kolejnymi edycjami egzaminu, normalizuje się je i co roku standaryzuje tak, by średnia krajowa wyniosła 100 punktów, a odchylenie standardowe wyniosło 15 punktów. Taka procedura umożliwia wygodne wykorzystywanie w ewaluacji szkół wyników z różnych lat. Znormalizowana i wystandaryzowana skala wyników uczniowskich służy zarówno do prezentacji wyniku egzaminacyjnego danego gimnazjum, jak i do prezentacji wartości wskaźnika edukacyjnej wartości dodanej. W wypadku EWD skala ma swój środek w punkcie 0. Wynik EWD równy np. 5 pkt, oznacza, że przeciętny uczeń tego gimnazjum otrzymał na egzaminie gimnazjalnym o 5 punktów więcej na skali wyników indywidualnych, niż by to wynikało z jego rezultatu na sprawdzianie po szkole podstawowej.

Na podstawie wyników egzaminacyjnych analizowanych w złożonych statystycznych modelach wejście-wyście szkoła jest lokowana w dwuwymiarowej przestrzeni ewaluacyjnej. Używany do estymacji wskaźników model statystyczny to analiza regresji z efektami losowymi. Używane są też metody bayesowskie i techniki kurczenia (shrinkage). Oś pozioma tej przestrzeni to średni, trzyletni wynik egzaminacyjny w szkole, oś pionowa to miara EWD. Wynik egzaminacyjny i wskaźnik EWD to miary komplementarne, łącznie dają pełniejszą informację o szkole, niż każda z osobna. Ze względu na statystyczną i pomiarową niepewność szacowania, którą obarczone są wyznaczone punktowo wskaźniki egzaminacyjne, pozycja szkoły na wykresach przedstawiana jest za pomocą elipsy obrazującej 95% obszar ufności. Wyniki są prezentowane dla gimnazjów, dla których w okresie trzyletnim system egzaminacyjny dysponuje wynikami dla nie mniej niż 30 uczniów. Im więcej wyników uwzględnia się w szacowaniu wskaźników i im mniejsze zróżnicowanie wewnątrzszkolne, tym większa pewność a tym samym mniej rozległy obszar ufności (elips).

Dla ewaluacji szkół ważny jest trend czasowy. Czy znalezienie się w latach 2006-2008 blisko punktu (100;0) jest etapem w wędrówce ku szkole o wysokich wynikach końcowych i wysokiej efektywności nauczania, czy też punktem w drodze ku niepokojącej kategorii szkół o niskich wynikach i EWD? Pokazują to wyniki analiz w kolejnych latach. W tej chwili gimnazja mogą być analizowane w latach 2006-2011. Dla liceów i techników dostępne są wskaźniki dwuletnie za lata 2010-2011.

Po tym wprowadzeniu zaprezentowane zostaną dostępne on-line wyniki dla wybranego miasta. Miasto A to średniej wielkości miasto, około 60 tys. mieszkańców. Średni wynik egzaminu gimnazjalnego w latach 2002-2011 jest dość stabilny i bliski średniej krajowej, obserwujemy jednak bardzo wysoki poziom zróżnicowania gimnazjów i bardzo dużą dynamikę tego procesu.

Na wykresach przedstawiono wskaźniki (w oryginale oznaczone kolorowymi elipsami) dla ośmiu szkół gimnazjalnych dla kolejnych lat (3-letnie wskaźniki ruchome obejmujące lata 2006-2011).



Wykres 5. Analiza systemu gimnazjów w mieście A. Pozycja gimnazjów w latach 2006-2011 na wymiarach wyników egzaminacyjnych i EWD, na podstawie części humanistycznej egzaminu

Na wykresach znajdujemy potwierdzenie dużego zróżnicowania szkół ze względu na średni wynik egzaminu na zakończenie gimnazjum. W pierwszym okresie (2006-2008) najlepsze ze względu na wyniki egzaminacyjne gimnazjum uzyskało średni wynik o ponad jedno odchylenie standardowe powyżej średniej krajowej, natomiast gimnazja najsłabsze pod tym względem lokowały się zdecydowanie poniżej średniej krajowej. Zróżnicowanie ze względu na EWD nie jest już tak duże, ale i tak notujemy znaczące różnice między szkołami. Bardziej interesujący obraz wyłania się jednak z porównania kolejnych wykresów, co umożliwia spojrzenie na dynamikę opisywanego zjawiska. Obraz ten jest dość spójny i pokazuje coraz większe różnicowanie się szkół ze względu na średni wynik egzaminacyjny. Szkoły coraz bardziej oddalają się od siebie, tworzą coraz bardziej rozciągnięty „peleton oświatowy”. Główna oś różnicowania gimnazjów w mieście A to wyniki końcowe, ale również na osi EWD obserwujemy znaczące rozproszenie.

Analiza dynamiki wskaźników w kolejnych latach dla poszczególnych gimnazjów pokazuje, że w większości szkół zmiana średnich wyników egzaminacyjnych była związana ze zmianą efektywności nauczania. Na przykład, jedno z gimnazjów w latach 2006-2008 osiągnęło wyniki egzaminacyjne powyżej średniej krajowej przy przeciętnej efektywności nauczania, a w kolejnych latach zanotowało zarówno

wzrost efektywności nauczania jak i wzrost wyników egzaminacyjnych. Nie jest to jednak bezwarunkową regułą.

Analiza ewaluacyjnych wskaźników egzaminacyjnych może być z pewnością inspirująca dla władz miasta A. Czy gimnazja oferują tę samą szansę na dobre wykształcenie? Zdecydowanie nie. Gdyby tylko w grę wchodziła segregacja na progu gimnazjum ze względu na uprzednie osiągnięcia, można by utrzymywać, że zasada równości szans nie jest zagrożona. Słabsi uczą się ze słabszymi, lepsi uczniowie z lepszymi. Nikt nie traci, nikt nie zyskuje (choć jak pisaliśmy we wprowadzeniu teoretycznym sam proces hierarchizacji gimnazjów nie jest społecznie „niewinny”). Tak jednak nie jest. Różnica na wymiarze EWD między najbardziej i najmniej efektywnym gimnazjum wynosi ok. 5-6 punktów, czyli 1/3 odchylenia standardowego. O tyle właśnie rosną lub spadają szanse na dobry wynik egzaminacyjny w zależności od tego, do którego gimnazjum w tym mieście trafi uczeń.

Konkluzje

Krajowy system egzaminacyjny dostarcza wartościowych danych pozwalających monitorować i ewaluować ważne procesy oświatowe zarówno na poziomie centralnym jak i lokalnym. Dobrym przykładem potencjału analitycznego tkwiącego w danych egzaminacyjnych jest proces różnicowania się miejskich systemów gimnazjalnych. Odkrycie, że w miastach, szczególnie dużych, obserwujemy dynamiczny proces różnicowania się gimnazjów ze względu na wyniki kształcenia jest niezwykle ważne dla polityki edukacyjnej państwa. Pozwala na opartą na faktach ocenę skutków wprowadzenia gimnazjów. System egzaminacyjny dostarcza równocześnie danych pozwalających monitorować ten proces na poziomie lokalnym. Samorządy mają dostęp do wyników pozwalających nie tylko monitorować proces różnicowania gimnazjów na swoim terenie działania, ale dzięki metodzie EWD mogą też ocenić skutki tego procesu z punktu widzenia efektywności pracy poszczególnych szkół. Jeżeli władze lokalne podjęłyby działania zmierzające do zmniejszania skali nierówności edukacyjnych, mają równocześnie instrument pozwalający ewaluować ich skutki. W obecnej chwili wydaje się, że główną barierą w wykorzystaniu danych egzaminacyjnych są kompetencje analityczne potencjalnych użytkowników. Jednak coraz większe zainteresowanie po stronie samorządów dobrą informacją o wynikach kształcenia oraz z drugiej strony wzbogacenie oferty szkoleniowej w zakresie wykorzystania wyników egzaminacyjnych (Jakość oświaty ... 2011) pozwala na umiarkowany optymizm. Ważne jest jednak, by centralne władze oświatowe wysłały czytelny sygnał, iż traktują funkcję monitorującą i ewaluacyjną krajowych egzaminów szkolnych jako ich rację bytu. Nie po to jako społeczeństwo wydajemy niemało grosza publicznego na ten system, by tylko „zaetykietkować” uczniowskie osiągnięcia.

Bibliografia

- Angrist J., Lang K. (2002) How important are classroom peer effects? Evidence from Boston's METCO program. National Bureau of Economic Research, www.nber.org/papers.
- Ball, S., J. (1993) Education Markets, Choice and Social Class: The Market as a Class Strategy in the UK and USA. *British Journal of Sociology of Education*, Vol. 14, Issue 1, pp. 3-20.
- Dolata R. (2008) *Szkoła-segregacje-nierówności*. Warszawa, Wydawnictwa Uniwersytetu Warszawskiego.
- Dolata R. (2010) Międzyszkolne zróżnicowanie wyników nauczania w szkołach podstawowych i gimnazjach. W: J. Łukasik, I. Nowosad, M., J. Szymański (red.) *Edukacja. Równość czy jakość?* Toruń, Wydawnictwo Adam Marszałek.
- Domański H., Pokropek A. (2011) Podziały terytorialne, globalizacja a nierówności społeczne. Wprowadzenie do modeli wielopoziomowych. Warszawa, Wydawnictwa IFiS PAN.
- Goldstein H., Noden P. (2003) Modeling social segregation. *Oxford Review of Education*. Vol. 29, no. 2, pp. 225-237.
- Gorard S., Fitz J. (2000) Markets and stratification: a view from England and Wales. *Educational Policy*, vol. 14, no. 3, pp. 405-428.
- Harker R., Tymms (2004) The effects of student composition on school outcomes. *School Effectiveness and School Improvement*. Vol. 15, no. 2, s.177-199.
- Herczyński J., Herbst M. (2005) *School Choice and Student Achievement. Evidence from Poland*. Warsaw University.
- Jakość oświaty jako efekt zarządzania strategicznego* (2011) Warszawa, ORE.
- Jargowsky P., A. (1996) Take the money and run: Economic segregation in U.S. metropolita areas. *American Sociological Review*. Nr 61, s. 984-998.
- Markman J., M., Hanushek E., A., Kain J., F., Rivkin S., G. (2003) Does peer ability affect student achievement? *Journal of Applied Econometrics*, vol. 18(5), s. 527-544.
- Mayer S., E. (2002) How economic segregation affects children's educational attainment. *Social Forces*. Vol. 81, Issue 1, s. 153-177.
- Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*. (2008) OECD Publishing.
- Murawska B. (2004) *Segregacje na progu szkoły*. Warszawa, Wydawnictwo Instytutu Spraw Publicznych.
- Pokropek A. (2011) *Efekt rówieśników w kształceniu szkolnym*, niepublikowany doktorat, Wydział Pedagogiczny, Uniwersytet Warszawski.

- Potulicka E., Rutkowiak J. (2010) *Neoliberalne uwikłania edukacji*. Kraków, Impuls.
- Rado P. (2010) *Governing decentralized education systems. Systemic change in South Eastern Europe*. Budapest, Open Society Foundations.
- Riodan C. (2004) *Equality and achievement*. Upper Saddle River, Pearson Prentice Hall.
- Therborn G. (2006) *Meaning, Mechanisms, Patterns, and Forces; An Introduction. W: Inequalities of the World. New theoretical frameworks, multiple empirical approaches*. Verso.
- Waslander, S., C. Pater, M. van der Weide (2010) *Markets in Education: An Analytical Review of empirical research on market mechanisms in education*. OECD Education Working Papers, No. 52, OECD Publishing.
- Wilkinson I., A., G., Hattie J., A., Parr J., M., Townsend M., A., R., Fung J., Ussher Ch., Thrupp M., Lauder H., Robinson T. (2000) *Influence of peer effects on learning outcomes: A review of the literature*. Auckland Uniservices Limited.
- Woods P., Bagley C., Glatter R. (1988) *School choice and competition: markets in the public interest?* London and New York, Routledge.