

WYKORZYSTANIE KRAJOWYCH EGZAMINÓW JAKO INSTRUMENTU POLITYKI OŚWIATOWEJ NA PRZYKŁADZIE PROCESU RÓZNICOWANIA SIĘ GIMNAZJÓW W DUŻYCH MIASTACH*

Roman Dolata
Aleksandra Jasińska
Michał Modzelewski

Instytut Badań Edukacyjnych w Warszawie

PROBLEM

Czy polskiej oświacie potrzebne są krajowe egzaminy wykorzystujące standaryzowane testy osiągnięć szkolnych? Gdy w debacie publicznej lub profesjonalnej pada to pytanie, uwaga dyskutantów koncentruje się zazwyczaj na jakości używanych testów, zwrotnym wpływie egzaminów na nauczanie (tzw. nauczaniu pod testy) lub selekcyjnej funkcji egzaminów.

* Prace nad rozwojem metody edukacyjnej wartości dodanej i analizy procesu różnicowania się gimnazjów są od 2008 r. systematycznie prowadzone w ramach projektu *Badania dotyczące rozwoju metodologii szacowania wskaźnika edukacyjnej wartości dodanej*. Od 2007 r. projekt realizowała Centralna Komisja Egzaminacyjna, a od 1.09.2012 r. – Instytut Badań Edukacyjnych. Projekt współfinansowany ze środków EFS UE w ramach PO KL.

W artykule chcielibyśmy zająć się innymi aspektami tego elementu systemu oświatowego. Egzaminy krajowe mogą bowiem dostarczać informacji o zachodzących w systemie oświaty procesach (funkcja monitorująca); mogą też 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 osiągają coraz większą autonomię, a na poziomie krajowym lub lokalnym wprowadzane są równocześnie elementy polityki rynkowej (Waslander i in. 2010).

Decentralizacja i urynkowienie oświaty może doprowadzić do wielu spontanicznych procesów wewnątrz systemu (Rado 2010). Na przykładzie jednego z takich procesów – różnicowania się w miastach gimnazjów pod względem osiągnięć szkolnych (Dolata 2008, 2010) – omówimy funkcję monitorującą i ewaluacyjną krajowych egzaminów gimnazjalnych.

PROCES RÓZNICOWANIA SIĘ GIMNAZJÓW POD WZGLĘDEM OSIĄGNIĘĆ

W 1999 r. wdrożono w Polsce jedną z bardziej masywnych reform edukacyjnych w skali całej Europy. Dwoma najważniejszymi jej elementami było wprowadzenie gimnazjów i krajowych egzaminów wykorzystujących standaryzowane testy osiągnięć szkolnych. Zgodnie z założeniami tej reformy po 6 latach nauki w szkole podstawowej uczniowie mieli być bez selekcji lokowani w trzyletnich 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 segregacyjnych, polegających z jednej strony na selekcji uczniów na progu gimnazjum, ze względu na ich uprzednie osiągnięcia, a z drugiej na skorelowanym z nią procesie różnicowania się gimnazjów pod względem efektywności kształcenia?

Za hipotezą potwierdzającą taki bieg wydarzeń przemawiają – poza samym wprowadzeniem gimnazjum – dodatkowe przesłanki. Pierwszą z nich jest zbiegnięcie się w czasie wprowadzenia gimnazjów z silnym spadkowym trendem liczebności szkolnej populacji, drugą – osłabienie urzędowej rejonizacji szkół, a trzecią – wzrost popularności w Polsce neoliberalnych, quasi-rynkowych pomysłów na poprawę jakości szkół (Dolata 2008; Potulicka, Rutkowiak 2010). Zjawiska przez nie opisywane występują ze szczególnym nasileniem 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 różnicowanie wyników nauczania. Na nim skupimy uwagę.

Zanim przejdziemy do analizy interesującego nas zjawiska, by wykażać użyteczność danych pochodzących z krajowych egzaminów do jego monitorowania, omówimy jego kontekst teoretyczny. 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 społecznie, hierarchizacja i eksploatacja. Dziedzina nierówności są: 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 pod względem wyników nauczania opisać jako proces hierarchizacji populacji szkół – z założenia jednolitych – ze względu na dobra symboliczne. Ten proces hierarchizacji szkół bezpośrednio wytwarza społeczne nierówności (np. poprzez rankingi szkół), jest jednak interesujący z punktu widzenia potencjalnego wzmacniania procesu różnicowania się osiągnięć zarówno na poziomie jednostek (wzrost wariacji wyników), jak i grup (wzrost dystansów międzygrupowych ze względu na osiągnięcia szkolne).

Kontynuując analizy teoretyczne, rozważmy, jakie mechanizmy mogą odpowiadać za różnicowanie 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 stopień, w jakim dany podział przestrzenny (np. regiony kraju lub rejony miasta) wyjaśnia zmienność danej cechy. Dla różnicowania szkół ze względu na wyniki nauczania znaczenie ma przede wszystkim segregacja przestrzenna 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 segre-

gacji przestrzennej ze względu na wymienione wcześniej charakterystyki, 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 i in. 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 różnicowanie się szkół pod względem składu społecznego i osiągnięć (Ball 1993; Goldstein, Noden 2003; Murawska 2004).

Nie zawsze jednak mechanizmy rynkowe prowadzą do różnicowania szkół. Wyniki analiz tego problemu nie dają spójnego obrazu (Waslander i in. 2010). W Polsce badania nad funkcjonowaniem rynków publicznych w oświacie nie są częste (Herczyński, Herbst 2005; Dolata 2008), ale uzyskane wyniki raczej potwierdzają, że procesy autoselekcji związane ze swobodnym wyborem szkoły zwiększają międzyszkolne różnicowanie.

Następny potencjalny mechanizm powstawania 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 on do silnego różnicowania składu osobowego szkół „za progiem” ze względu na uprzednie osiągnięcia szkolne uczniów i zasoby z tym powiązane. Różnicowanie na wejściu w oczywisty sposób prowadzi do różnicowania na wyjściu.

Również tzw. efekt rówieśników może przyczyniać się do różnicowania szkół ze względu na osiągnięcia. Wpływ rówieśników jest przedmiotem licznych badań, głównie dotyczących problemów segregacji w oświacie i skuteczności różnego typu innowacji dydaktycznych wykorzystujących zasoby rówieśnicze (Wilkinson i in. 2000; Angrist, Lang 2002; Markman i in. 2003; Dolata 2008). Otrzymywane rezultaty nie zawsze są 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. Szkoły mające dobrą opinię na rynku przyciągają dobrych uczniów, a z biegiem czasu również lepszych nauczycieli, dodatkowe środki finansowe i rzeczowe. Może to prowadzić do międzyszkolnego różnicowania efektywności nauczania.

DANE I METODA ANALIZY

W analizach wykorzystano dane ogólnokrajowe z części humanistycznej egzaminu gimnazjalnego dla lat 2002–2012 oraz, pomocniczo, analogiczne dane ze sprawdzianu w klasie VI szkoły podstawowej. Dla przeprowadzonych analiz ważne znaczenie miały własności psychometryczne testów egzaminacyjnych.

Testy gimnazjalne są standaryzowanymi testami ogólnokrajowymi tworzonymi z wykorzystaniem klasycznej teorii testów. Kolejne ich edycje nie są systemowo 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). Następnie z analizy wykluczono dane pochodzące ze szkół, w których na poziomie ostatniej klasy uczyło się mniej niż pięciu uczniów.

W zakresie trafności jedynym gwarantem stałości mierzonego przez egzaminy konstruktów jest dość szczegółowy plan testów. W 2012 r. uległ on jednak znacznej zmianie, co może zaburzać wyniki analizy trendu w zakresie różnicowania międzyszkolnego. W analizach prowadzonych dla 2012 r. wykorzystywano łączny wynik z testu humanistycznego i testu z języka polskiego. Drugi kluczowy aspekt jakości testów, rzetelność, również pozostawia wiele do życzenia. Wartości współczynnika rzetelności alfa Cronbacha dla części humanistycznej testów gimnazjalnych¹ w latach 2002–2012 wahały się – najmniej rzetelny był test z 2002 r. (0,78), najbardziej rzetelny z 2012 r. (0,89). Jak na testy wysokiej stawki rzetelność jest zbyt niska, ale do celów naszej analizy wystarczająca. Występują znaczące wahania w rzetelności, ale nie są one skorelowane z trendem w zakresie różnicowania międzyszkolnego.

Obliczenie wskaźnika międzyszkolnego różnicowania wyników nauczania to klasyczny problem dekompozycji wariacji. W zaprezen-

towanych w tym artykule analizach gimnazja podzielono najpierw na 4 grupy ze względu na ich lokalizację: wieś, miasta do 20 tys., miasta od 20 tys. do 100 tys. i miasta ponad 100 tys. 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 zazwyczaj te jednostki samorządowe.

W kolejnym kroku trzy składowe wariancje podzielono przez wariancję całkowitą, a wynik wyrażono w procentach. Dla dalszych analiz kluczowy jest odsetek wariancji międzyszkolnej w stosunku do wariancji całkowitej. Choć na ogólnym poziomie oszacowanie wskaźnika jest dość proste, to w praktyce statystycznej dysponujemy bardzo wieloma algorytmami dekompozycji wariancji. W analizach użyto trzydziestu 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.

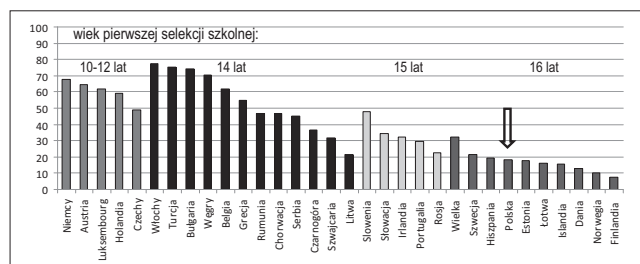
ZRÓŻNICOWANIE MIĘDZYSZKOLNE W PERSPEKTYWIE PORÓWNAWCZEJ

Zanim przedstawimy wyniki analiz danych egzaminacyjnych, przyjrzymy się interesującemu nas zjawisku w perspektywie porównawczej, bazując na wynikach badania PISA (OECD 2010). W analizach OECD wykorzystano szacunki wskaźnika międzyszkolnego zróżnicowania wyników z zakresu czytania. 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 którym następuje zróżnicowanie ścieżek edukacyjnych, głównie na ścieżkę kształcenia akademickiego i zawodowego.

Wykres 1 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 r. gimnazjum, znajduje się wśród krajów o niskim poziomie zróżnicowania międzyszkolnego. Znalezie-

nie 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. Burzy je jednak świadomość, że już na poziomie I klasy szkoły ponadgimnazjalnej poziom interesującego nas zróżnicowania dramatycznie wzrasta. Wskaźnik wyliczony dla Polski na podstawie wyników PISA 2000 (piętnastolatki uczyły się wówczas w I klasie szkół ponadpodstawowych) lokował nas w grupie krajów o najwyższym poziomie międzyszkolnego zróżnicowania wyników nauczania. Obraz ten potwierdziły wyniki z opcji krajowej badania PISA 2006 (opcja krajowa objęła uczniów klas I i II szkół ponadgimnazjalnych).

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

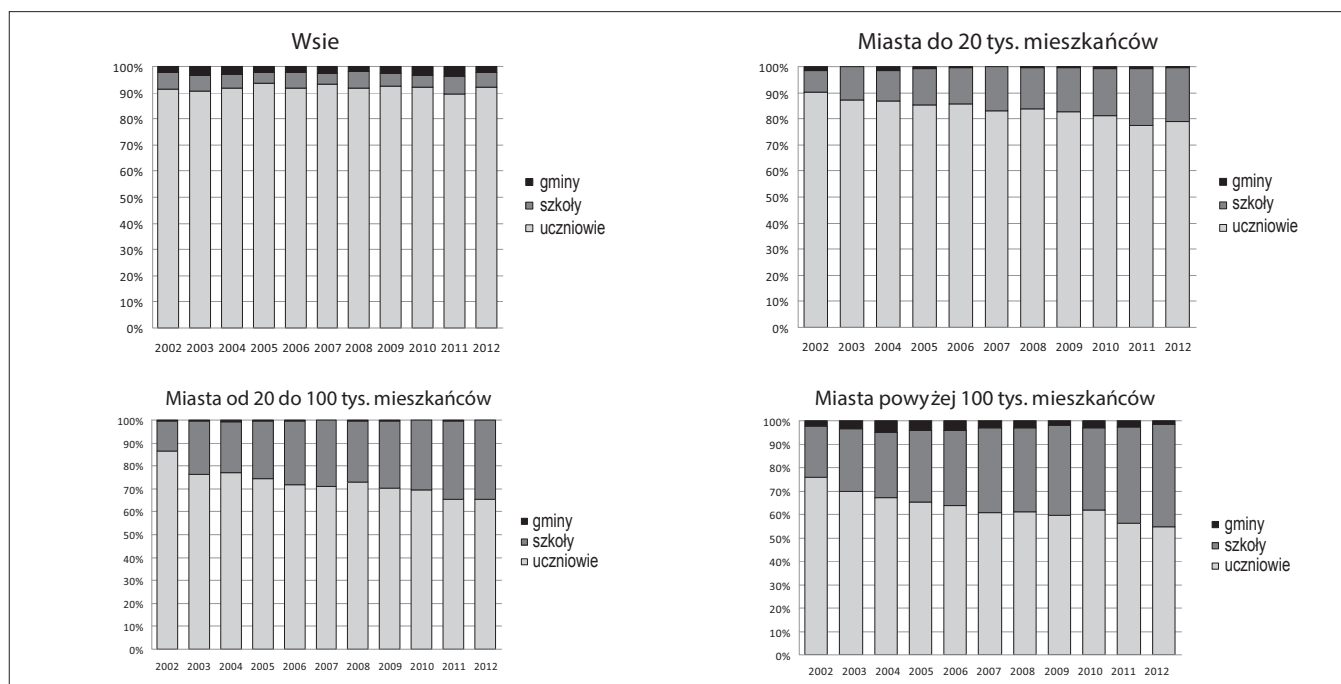


ZRÓŻNICOWANIE MIĘDZYSZKOLNE GIMNAZJÓW W ŚWIELE WYNIKÓW EGZAMINU

Jak wygląda interesujące nas zjawisko w świetle wyników egzaminu gimnazjalnego? W analizach ograniczono się do wyników z 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 wariancji zostały przedstawione w podziale gimnazjów ze względu na lokalizację: teryny wiejskie i trzy kategorie wielkości miast. Taki podział wyników wyjaśnia niewielką część wariancji całkowitej, a wariancje całkowite wyników egzaminu gimnazjalnego w wyróżnionych segmentach gimnazjów niewiele się od siebie różnią. Za podstawę procentowania można zatem przyjąć wariancję całkowitą oddzielnie szacowaną dla każdej lokalizacji.

Wykres 2. Oszacowania udziału wariancji międzygminnej, międzyszkolnej i wewnątrzszkolnej w wariancji całkowitej wyniku egzaminu gimnazjalnego, w części humanistycznej. Analiza dynamiki ze względu na lokalizację gimnazjum



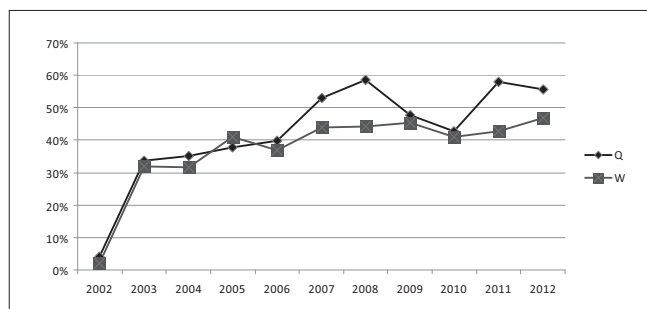
Zanim przejdziemy do zróżnicowania międzyszkolnego wyników egzaminu, należy zwrócić uwagę, że wariancja międzygimnna stanowi niewielką część wariancji całkowitej wyników egzaminacyjnych i, co więcej, nie notuje się w tym zakresie żadnego znaczącego trendu.

Inaczej sytuacja prezentuje się w przypadku wariancji międzyszkolnej. W 2002 r. oszacowany poziom zróżnicowania międzyszkolnego w czterech analizowanych segmentach gimnazjów wynosił od 6% dla 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 silniejszy wzrostowy trend różnicowania się szkół.

Proces różnicowania się pod względem wyników kształcenia gimnazjów zlokalizowanych w dużych miastach ma już bardzo dużą dynamikę. Analizy wyników egzaminacyjnych wskazują, że w dużych miastach mamy do czynienia z silnym i rosnącym zróżnicowaniem gimnazjów. W 2012 r. oszacowana wariancja międzyszkolna stanowiła już ponad 40% wariancji całkowitej.

Analizy dla poszczególnych dużych miast potwierdzają ten obraz. Wykres 3 przedstawia dynamikę procesu różnicowania się gimnazjów pod względem wyników nauczania w dwóch przykładowych miastach. Zostały one symbolicznie oznaczone literami Q i W.

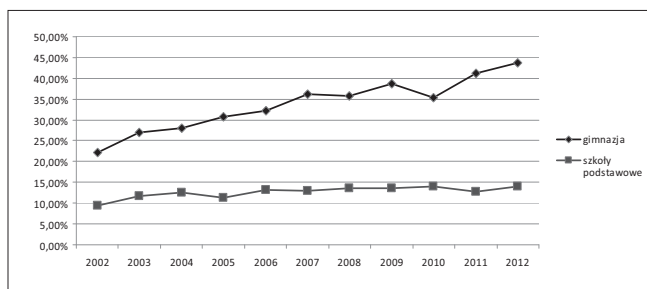
Wykres 3. *Dynamika procesu różnicowania się gimnazjów pod względem wyników nauczania w dwóch przykładowych miastach*



W przypadku konkretnych miast możemy obserwować bardziej złożony obraz dynamiki analizowanego zjawiska, ale ogólny wzrostowy wzorzec trendu występuje w prawie 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. Wykres 4 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 na proces różnicowania się gimnazjów.

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



Poziom zróżnicowania gimnazjów pod względem wyników nauczania w 2012 r. należy porównywać z poziomem tego wskaźnika dla szkół podstawowych trzy lata wcześniej (czyli w 2009 r.). To właśnie

uczniowie kończący edukację w szkole podstawowej w 2009 r. trzy lata później przystąpili do egzaminu gimnazjalnego. Oszacowanie wskaźnika dla szkół podstawowych wynosi 14%, czyli o około 30 pkt. proc. mniej niż w wypadku gimnazjów! Z uwagi na to, że szkoły podstawowe są bardziej podatne na wpływ procesów segregacji przestrzennej (mniejsze rejony działania) niż gimnazja, tak duża różnica wartości tych wskaźników wskazuje na to, że segregacja przestrzenna ma niewielki udział w procesie różnicowania się wielkomiejskich gimnazjów.

Analizy prowadzone 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 pod względem wyników kształcenia. Nie da się tego zjawiska wyjaśnić samoistnie zachodzącą segregacją przestrzenną. Innymi słowy, wprowadzenie w 1999 r. gimnazjów uruchomiło w miastach ukryty proces selekcyjny na poziomie 12. roku życia (gdy do gimnazjów dotrze fala uczniów po obniżeniu wieku szkolnego, próg selekcyjny obniży się do 11. roku życia!). To ważne odkrycie stanowi przykład wielkiej użyteczności danych z krajowego systemu egzaminacyjnego.

Co ono oznacza dla krajowej polityki oświatowej? Ocena zależy od przesłanek ideologicznych i przyszłych możliwości bardziej wnikliwego zrozumienia samego procesu różnicowania się gimnazjów. Jedno jest pewne – zjawisko to powinno być przedmiotem dyskusji zarówno na poziomie polityki krajowej, jak i polityk lokalnych. Władze samorządowe i same szkoły powinny mieć dostęp do informacji o nasileniu tego zjawiska w danej gminie, mieście lub dzielnicy. Warto zastanowić się, jakimi instrumentami dysponują władze lokalne, by monitorować ten proces. Temu zagadnieniu poświęcamy drugą część artykułu.

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. Dzięki pracom nad metodą edukacyjnej wartości dodanej (EWD) zarówno szkoły, jak i władze lokalne uzyskują dostęp do wyników analiz, które zdecydowanie lepiej obrazują interesujące nas zjawisko.

Zanim przyjrzymy się instrumentom analitycznym będącymi do dyspozycji władz lokalnych, musimy się zapoznać z polską implementacją metody edukacyjnej wartości dodanej.

EWD pozwala na podstawie odpowiednio przetworzonych danych egzaminacyjnych uzyskać miary efektywności nauczania, czyli poziomu, jaki robią uczniowie w trakcie nauki w danej szkole. Metoda ta dostarcza bardziej przydatnych dla polityki oświatowej (w porównaniu z wynikami końcowymi szkoły) wskaźników efektywności pracy szkoły (OECD 2008).

Efektywność nauczania w danej szkole często utożsamiana jest z jej wynikami egzaminacyjnymi. Gdyby wszystkie szkoły pracowały z uczniami o tym samym poziomie uprzednich osiągnięć szkolnych i ogólnych zdolności poznawczych oraz wyposażonych w identyczne zasoby rodzinne, wtedy wyniki egzaminu końcowego byłyby dobrą miarą efektywności nauczania. Tak jednak nie jest. Metoda EWD pozwala uwzględnić zróżnicowanie zasobów szkół „na wejściu”. Najczęstszą miarą takich zasobów są wyniki testów mierzących uprzednie osiągnięcia szkolne.

W 2005 r. grupa ekspertów pracujących 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. Obecnie 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, ogólnodostępne w Internecie trzyletnie ewaluacyjne wskaź-

niki egzaminacyjne, m.in. dla gimnazjów². Wskaźniki te 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 wskaźników dla takiego przedziału czasowego uzasadnione jest z kilku powodów. Najważniejszy wiąże się z niepewnością pomiarową, którą obarczone są wyniki każdego testu, również egzaminu gimnazjalnego. Im większą pulą danych dysponujemy, tym mniejszy wpływ tej niepewności na szacowanie wskaźników dla szkoły. Korzystając z wyników z trzech kolejnych lat, można uzyskać w przybliżeniu trzy razy więcej danych, a duża ilość informacji wiąże się z większą precyzją 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 standardyzuje tak, by średnia krajowa wyniosła 100 punktów, a odchylenie standardowe 15 punktów. Taka procedura umożliwia wygodne wykorzystywanie w ewaluacji szkół wyników z różnych lat. Znormalizowana i wystandaryzowana skala 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. Wartość EWD równa np. 5 pkt. oznacza, że przeciętny uczeń tego gimnazjum otrzymał na egzaminie gimnazjalnym o 5 punktów więcej na wystandaryzowanej skali wyników indywidualnych, niż by to wynikało z jego rezultatu zdobytego na sprawdzianie szóstoklasistów.

Na podstawie wyników egzaminacyjnych analizowanych w złożonych statystycznych modelach wejście-wyjście szkoła jest lokowana w dwuwymiarowej przestrzeni ewaluacyjnej. Do estymacji wskaźników stosuje się analizę regresji z efektami losowymi. Używane są też metody bayesowskie i techniki kurczenia (*shrinkage*). Oś pozioma w tej przestrzeni to średni, trzyletni wynik egzaminacyjny w szkole, oś

pionowa oznacza wartość EWD. Wynik egzaminacyjny i wskaźnik EWD to miary komplementarne; łącznie dają pełniejszą informację o pracy szkoły, 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-procentowy obszar ufności. Wyniki są prezentowane dla gimnazjów z co najmniej 30 absolwentami w trzyletnim okresie. Im więcej wyników uwzględnia się w szacowaniu wskaźników i im mniejsze zróżnicowanie wewnętrzne, tym większa pewność, a tym samym mniej rozległy obszar ufności.

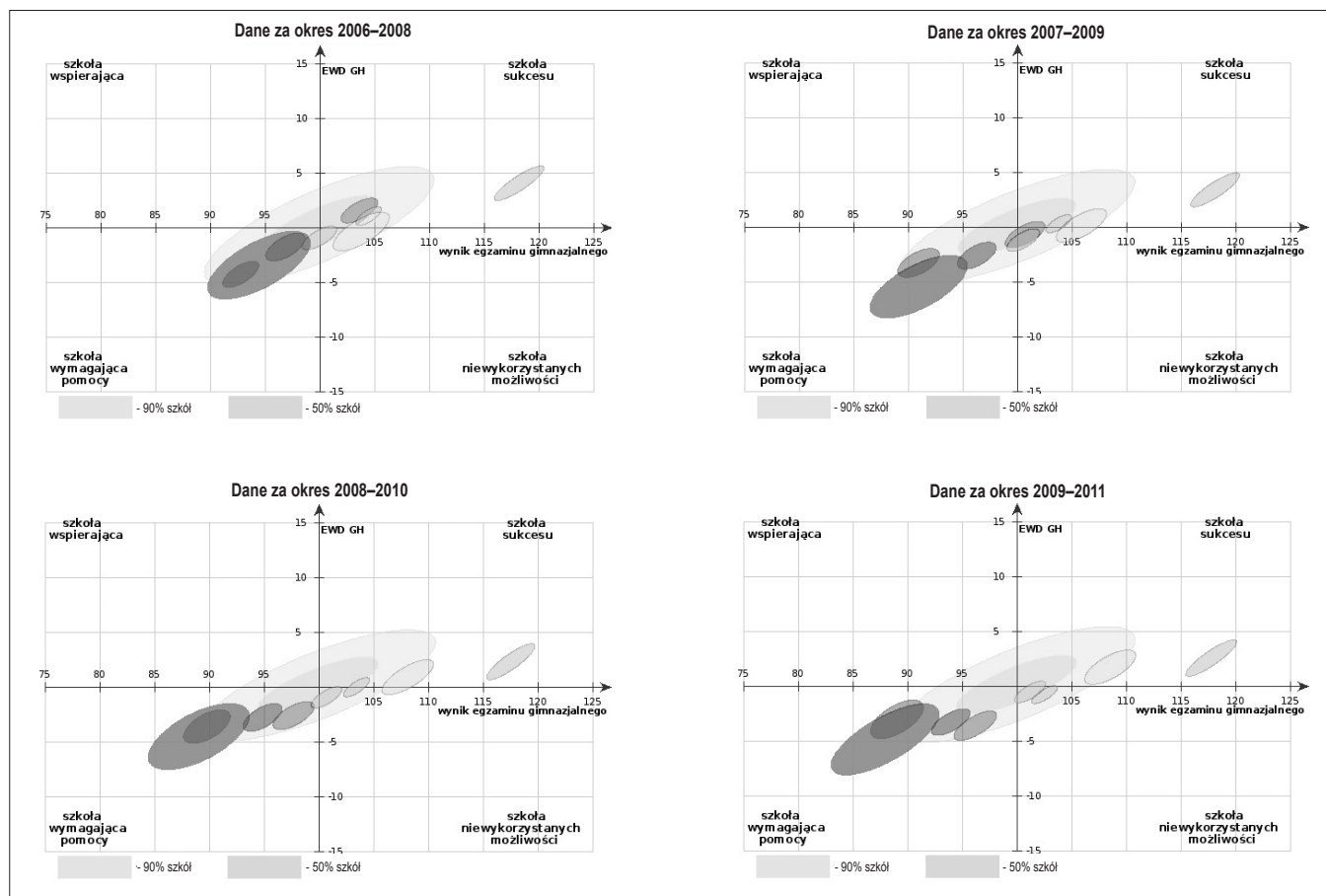
Dla ewaluacji szkół ważny jest trend czasowy. Czy znalezienie się w latach 2006–2008 blisko punktu (100;0) jest etapem w wędrówce ku szkole o wysokich wynikach końcowych i wysokiej efektywności nauczania, czy też punktem w drodze ku niepokojącej kategorii szkół o niskich wynikach i EWD? Odpowiedzi na to pytanie pozwalają udzielić analizy prowadzone dla kolejnych trzyletnich okresów³.

Poniżej prezentujemy wyniki analiz dla wybranego miasta. Miasto A to średniej wielkości miejscowość, którą zamieszkuje około 60 tys. osób. 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 wykresie 5 przedstawiono wskaźniki (w oryginale oznaczone kolorowymi elipsami) dla ośmiu szkół gimnazjalnych w kolejnych latach (trzyletnie wskaźniki ruchome obejmujące lata 2006–2011).

Na wykresach znajdujemy potwierdzenie dużego zróżnicowania szkół ze względu na średni wynik egzaminu kończącego gimnazjum. Główna oś różnicowania gimnazjów w mieście A to wyniki końcowe, ale również na osi EWD obserwujemy znaczące rozproszenie. W pierwszym okresie (2006–2008) gimnazjum, które zanotowało najlepsze wyniki egzaminacyjne, uzyskało średni wynik o ponad jedno odchylenie standardowe powyżej średniej krajowej, natomiast gimnazja najgorsze 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.

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 gimnazjalnego



Bardziej interesujący obraz wyłania się jednak z dynamicznego ujęcia opisywanego zjawiska, co możemy zaobserwować, porównując kolejne wykresy. Obraz ten jest dość spójny i pokazuje coraz większe różnicowanie się szkół pod względem średniego wyniku egzaminacyjnego. Szkoły coraz bardziej oddalają się od siebie, tworzą coraz bardziej rozciągnięty „peleton oświatowy”.

Analiza dynamiki wskaźników za kolejne lata 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 bezwarunkowa reguła.

Analiza ewaluacyjnych wskaźników egzaminacyjnych może być z pewnością inspirująca dla władz miasta A. Czy gimnazja w tym mieście 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 pisałiśmy 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 około 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 pod względem wyników kształcenia jest niezwykle ważne dla polityki edukacyjnej państwa. Pozwala na opartą na faktach ocenę skutków wprowadzenia gimnazjów.

System egzaminacyjny dostarcza równocześnie danych umożliwiających monitorowanie tego procesu na poziomie lokalnym. Dzięki metodzie EWD samorządy mogą też ocenić skutki tego procesu z punktu widzenia efektywności pracy poszczególnych szkół. Jeżeli władze lokalne zdecydują się podjąć działania zmierzające do zmniejszenia skali nierówności edukacyjnych, zyskują instrument pozwalający ewaluować ich skutki.

W obecnej chwili wydaje się, że główną barierą w wykorzystaniu danych egzaminacyjnych są kompetencje analityczne potencjalnych ich użytkowników. Rosnące zainteresowanie po stronie samorządów rzetelną informacją o wynikach kształcenia oraz wzbogacanie oferty szkoleniowej w zakresie wykorzystania wyników egzaminacyjnych (ORE 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 jedynie „etykietkować” uczniowskie osiągnięcia.

SUMMARY

Apart from diagnostic, selective and formative functions, the national examination system can provide valuable information about processes that operate within a school system (monitoring function) as well as constitute an instrument to manage these processes (evaluative function). In this article we illustrate the analytical potential of national examination data for monitoring and evaluation of educational processes by analyzing process of differentiation of lower secondary schools. Analyses of lower secondary school exam results show that within cities, especially big ones, we are dealing with a dynamic process of schools differentiation in regards to educational outcomes. This discovery is of particular importance for state educational policies. Article also describes the possibilities of using national examination data in the form of exam scores and educational value-added indexes to monitor and evaluate processes at the local level.

- 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.
- ² Patrz www.ewd.edu.pl. Dostępne są również wskaźniki dla liceów ogólnokształcących i techników.
- ³ W tej chwili dostępne dane pozwalają analizować gimnazja w latach 2006–2011. Dla liceów i techników dostępne są wskaźniki dwuletnie za lata 2010–2011.

LITERATURA

- 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” 14(1), s. 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: Wydawnictwo IFiS PAN.
- Goldstein H., Noden P. (2003), *Modeling social segregation*, „Oxford Review of Education” 29(2), s. 225–237.
- Gorard S., Fitz J. (2000), *Markets and stratification: a view from England and Wales*, „Educational Policy” 14(3), s. 405–428.
- Harker R., Tymms P. (2004), *The effects of student composition on school outcomes*, „School Effectiveness and School Improvement” 15(2), s. 177–199.
- Herczyński J., Herbst M. (2005), *School choice and student achievement. Evidence from Poland*, Warsaw: Warsaw University.
- Jargowsky P.A. (1996), *Take the money and run: economic segregation in U.S. metropolitan areas*, „American Sociological Review” 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” 18(5), s. 527–544.
- Mayer S.E. (2002), *How economic segregation affects children's educational attainment*, „Social Forces” 81(1), s. 153–177.
- Murawska B. (2004), *Segregacje na progu szkoły*, Warszawa: Wydawnictwo Instytutu Spraw Publicznych.
- ORE (2011), *Jakość oświaty jako efekt zarządzania strategicznego*, Warszawa.
- OECD (2010), *PISA 2009 results: What Students know and can do: student performance in reading, mathematics and science*, Paris.
- OECD (2008), *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*, Paris.
- Pokropek A. (2011), *Efekt rówieśników w kształceniu szkolnym*, niepublikowany doktorat, Warszawa: 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.
- Therborn G. (2006), *Meaning, mechanisms, patterns, and forces: an introduction*, w: *Inequalities of the world. New theoretical frameworks, multiple empirical approaches*, Verso.
- Waslander S., Pater C., van der Weide M. (2010), *Markets in education: an analytical review of empirical research on market mechanisms in education*, OECD Education Working Papers, No. 52, Paris.
- 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.