

Maciej Jakubowski, Paweł Sakowski

Wydział Nauk Ekonomicznych

Uniwersytet Warszawski¹

Wpływ wielkości klasy na wyniki uczniów szkół podstawowych.

I. Wprowadzenie.

Zmniejszenie liczby uczniów w klasie jest propozycją, która w wielu krajach stanowi centralny element debat dotyczących reformy oświaty. Jednocześnie jest to częsty w ekonomii edukacji temat badań. Dla zwykłych obywateli, nauczycieli oraz polityków popularność tej propozycji wynika zapewne z jej prostoty, a dla ekonomistów zarówno z dużej dostępności danych, bowiem wiele systemów egzaminacyjnych zbiera dane na poziomie klasy zawierające przynajmniej liczbę uczniów i ich średni wynik, jak i ze względu na intrygujące problemy metodologiczne. Badania nad tzw. „class size effect” stanowią swoisty poligon doświadczalny umożliwiający testowanie nowych pomysłów metodologicznych. Na świecie powstało kilkaset badań, które pośrednio lub bezpośrednio analizują tę kwestię. W Polsce znane nam jest kilka prac, które jedynie pośrednio jej dotyczą (por. Białecki, Haman, 2003; Herbst, Herczyński, 2002; Śleszyński, 2002).

Powszechnym przekonaniem jest to, że mniejsza liczba uczniów w klasie powinna sprzyjać jakości nauczania, ponieważ nauczyciele mogą więcej czasu poświęcić pojedynczym osobom, łatwiej utrzymać im porządek w klasie, a także kontakt z rodzicami. Tak więc zmniejszenie klas w szkołach publicznych powinno zaowocować podniesieniem osiągnięć uczniów. Taka wizja wydaje się atrakcyjna wielu rodzicom i politykom, bowiem stanowi proste rozwiązanie, którego pozytywne efekty wydają się oczywiste. W wielu krajach programy redukcji liczby uczniów w klasie popierane są też przez związki zawodowe nauczycieli ze względu na to, że ich wprowadzenie wymagałoby zatrudnienia większej liczby pedagogów, a ich praca stałaby

¹ Kontakt: email: mjakubowski@uw.edu.pl, tel. 0 602 748 740.

się zapewne mniej uciążliwa. Presja społeczna i polityczna spowodowała, że przez ostatnie 30 lat w USA i Francji przeciętna wielkość klas spadła z ok. 30 do 20. Wśród ekonomistów nie ma zgody, co do efektów tych programów. Wielu uważa je za kosztowne propozycje, które nie przynoszą żadnych rezultatów, kiedy inni podkreślają ich wysoką efektywność (por. dyskusja między dwoma znanymi badaczami: Hanushek, 2003, oraz Krueger, 2003).

W Polsce zmniejszenie klas jest rzadko dyskutowaną kwestią². Samorządy walcząc o obniżenie kosztów oświatowych często zmuszają szkoły do tworzenia dużych oddziałów, przekraczających 30 osób. W istocie przeciętna liczba uczniów w klasie różni się znacząco nie tylko między wsią (ok. 18) a miastem (ok. 24), ale również między gminami i szkołami. Brak standardów oświatowych powoduje, że samorządy mają tu stosunkowo dużą dowolność i jedynie Kuratoria, powołując się na nie do końca określone wymogi utrzymania odpowiedniej liczby uczniów np. na zajęciach wychowania fizycznego, ograniczają zapędy niektórych samorządowców do tworzenia oddziałów daleko przekraczających 30 osób. Jednak nawet groźby Kuratora mogą nie skutkować, gdy samorząd stwierdzi, że nie posiada zasobów na utrzymanie mniejszych klas. Znacznie większe przeciętnie klasy gimnazjalne pokazują, jakie zapewne byłyby rozmiary klas w szkołach podstawowych, gdyby samorządów nie ograniczały np. trudności w likwidacji mniejszych placówek. Dane empiryczne pokazują, że większość szkół podstawowych tworzy nowe oddziały, gdy rozmiar istniejących przekracza 28-29 osób.

Pomijając powszechne przekonania trudno przedstawić argumenty za tym, że klasy nie powinny przekraczać pewnej wielkości, bowiem nie jest pewne czy ich rozmiar negatywnie wpływa na osiągnięcia uczniów. Trudności metodologiczne powodują znaczne różnice w rezultatach badań prowadzonych nawet na tych samych próbach. Jedni ekonomiści twierdzą, że wielkość klasy ma pomijalny z praktycznego punktu widzenia wpływ na wyniki, inni, że w pewnych warunkach większa liczba uczniów w klasie może sprzyjać podnoszeniu osiągnięć uczniów. Efekt ten związany jest z tzw. „peer effects” polegającymi np. na większym prawdopodobieństwie znalezienia ucznia o podobnej wiedzy i zdolnościach w większej klasie (por. Dobbelsteen i in., 2002). Tak więc nie proste intuicje mogą być w tym wypadku mylące, a teorie psychologiczne dają argumenty za tym, że zmniejszanie klas poniżej pewnego poziomu nie przyniesie spodziewanych korzyści.

² W pracy pojęcia „klasa” i „oddział” stosowane są równoważnie, co jest zgodne z potocznym rozumieniem tych terminów. W każdym przypadku chodzi o grupę uczniów, która wspólnie uczęszcza na większość zajęć w szkole.

Związek przyczynowo-skutkowy między liczbą uczniów w klasie a ich osiągnięciami jest empirycznie bardzo trudny do badania. Rzeczywista liczba uczniów w klasie zależy w dużym stopniu od lokalizacji szkoły, niezależnych wyborów lokalnych władz, dyrektorów i rodziców. Nawet klasy w tej samej szkole mogą różnić się znacząco pod względem liczby uczniów. Różnice te często skorelowane są z cechami uczniów, ich rodzin, nauczycieli, otoczenia społecznego oraz ekonomicznego. Swoboda wyboru może powodować np., że szkoły lepsze, bardziej popularne będą miały większe klasy, a szkoły słabsze mniejsze. Możliwe też, że szkoły z dziećmi, których rodzice przykładają większą wagę do ich edukacji, będą miały mniejsze klasy, podobnie jak placówki zdobywające dodatkowe środki z innych niż publiczne źródeł (lub też finansowane przez zamożniejsze samorzady). Wewnątrz szkół często tworzone są klasy z mniej zdolnymi uczniami, które mogą być mniejsze, jeśli dyrektor szkoły w ten sposób chce zwiększyć szanse edukacyjne umieszczonych w nich uczniów, lub też większe, gdy gromadzi się w nich uczniów, którzy nie rokują wg dyrektora i nauczycieli szans na podniesienie ich umiejętności. W Polsce wielkość klasy zależy głównie od rozmiaru miejscowości, w której położona jest szkoła, a także od niemożliwych do przewidzenia decyzji samorządów i dyrektorów, często segregujących uczniów o różnych zdolnościach, czy pochodzeniu społecznym, do klas różnej wielkości.

Powyższa dyskusja pokazuje, że liczba uczniów w klasie jest efektem wielu niemożliwych do kontrolowania wyborów. Jest ona też silnie skorelowana z cechami społecznymi i ekonomicznymi, które z kolei oddziałują na wyniki uczniów. Stąd, w równaniach „edukacyjnej funkcji produkcji” liczba uczniów w klasie jest najczęściej zmienną endogeniczną, której szacowanie przez zwykłą regresję metodą najmniejszych kwadratów daje obciążone rezultaty. Przy czym, jak sugerują powyższe przykłady, kierunek obciążenia nie jest możliwy do przewidzenia.

Sposobem na uniknięcie problemów związanych z silną korelacją wielkości klasy z innymi czynnikami wpływającymi na osiągnięcia uczniów jest wykorzystanie jedynie wariacji powodowanej przez losowe zmiany w liczebnościach klas. Inaczej mówiąc, oparcie się na metodach eksperymentalnych i quasi-eksperymentalnych. O ile eksperymenty są w ekonomii edukacji bardzo rzadkie, to metody quasi-eksperymentalne opierające się na technice zmiennych instrumentalnych są dość częste. Kluczowy w tym przypadku jest wybór zmiennej instrumentalnej, której użycie daje nieobciążone oszacowania tylko pod warunkiem, że jest ona skorelowana jedynie z wynikami uczniów i wielkością klas. W badaniach nad wpływem wielkości klasy często wykorzystywanym instrumentem jest przeciętna liczba uczniów w

klasach w szkole, którego użycie zapobiega obciążeniom wynikającym z segregacji uczniów o różnych zdolnościach między klasami o różnych wielkościach wewnątrz szkoły (Akerhielm, 1995). Nie zapobiega to jednak obciążeniom wynikającym z segregacji między szkołami.

Jedną z metod stosowanych w tym przypadku jest włączenie do równań ze zmienną instrumentalną efektów stałych pojedynczych szkół lub ich większych grup odpowiadających np. obszarowi danego samorządu, czy też różnie definiowanej społeczności lokalnej. Takie podejście wymaga odpowiedniego zasobu danych. Przykładowo, Caroline Hoxby analizuje dane z kilkunastu lat, gdzie efekty stałe szacowane są dla „school district” (Hoxby 2000). W innym badaniu wykorzystano natomiast dane dla dwóch kolejnych kohort uczniów tej samej szkoły zebrane w jednej edycji międzynarodowych testów TIMSS (Wößmann, 2003).

Istotą badań tego typu jest założenie, że wariancja zmiennej instrumentalnej ma źródła losowe. W dwóch najbardziej znanych pracach opierających się na tej metodologii dla badania wpływu wielkości klasy określono dwa różne źródła takiej wariancji: zmiany w populacji uczniów powodowane zmianą w liczbie urodzeń (liczebności całego rocznika na danym terenie) oraz zmiany wielkości klas ze względu na stworzenie nowej klasy w szkole. Obydwie metody wykorzystwała Caroline Hoxby w przytoczonej już pracy, szacując wpływ wielkości klasy na podstawie wariancji powodowanej zmianą w wielkości populacji kolejnych roczników w danym „school district” na przestrzeni kilkunastu lat, a także powodowanej zmianą w liczbie klas w poszczególnych szkołach. W ten sposób stwierdziła, że wielkość klasy nie ma istotnego z praktycznego punktu widzenia wpływu na wyniki uczniów.

Nieco inną metodę zastosowali Joshua Angrist i Victor Lavy w znanej pracy, odwołującej się do tzw. zasady Majmonidesa obowiązującej szkoły publiczne w Izraelu. Według tej zasady nie można prowadzić klas z więcej niż 40 uczniami, a więc szkoły z 80 i 81 uczniami powinny różnić się pod względem liczby i przeciętnej wielkości klas. Angrist i Lavy przyjęli jako instrument funkcję, która na podstawie liczby uczniów w szkole przewiduje liczbę i rozmiar klas. Metoda ta ma w istocie charakter tzw. „regression discontinuity design” i opiera się na założeniu, że szkoły o bardzo podobnej liczbie uczniów, lecz z różną liczbą klas, mają niemal identyczne charakterystyki. Takie założenie wynika z obserwacji, że zmienna określająca tzw. „SES” szkoły („Socio-Economic Status”) jest bardzo silnie skorelowana zarówno z wynikami, jak i liczbą uczniów w szkole. Włączając tę zmienną do równania wraz ze zmienną instrumentalną oddającą nieciągłość w przeciętnej wielkości klas daje możliwość oszacowania nieobciążonego efektu wpływu wielkości klasy na wyniki uczniów. Stosując

takie podejście Angrist i Lavy stwierdzili, że liczba uczniów w klasie wpływa negatywnie na wyniki (Angrist, Lavy, 1999).

W badaniu przedstawionym poniżej zastosowano dwie, podobne do opisanych powyżej metody celem oszacowania wpływu liczby uczniów w klasie na wyniki szkół podstawowych w Polsce. Analizie poddano wyniki egzaminów zewnętrznych z lat 2002-2004. Pierwsza zastosowana metoda wykorzystuje przeciętną liczbę uczniów w klasach w danej szkole jako instrument dla rzeczywistej wielkości klas, dzięki czemu możliwe jest wyłączenie wpływu wewnątrzszkolnej segregacji uczniów i dodatkowo szacuje efekty stałe szkół. W podejściu tym analizowana jest zatem wariancja w wielkości klas powodowana losowymi zmianami w liczebnościach kolejnych roczników w danej szkole. W drugiej metodzie wykorzystano fakt, że większość szkół w Polsce tworzy nowe klasy, kiedy liczba uczniów w klasach zbliża się do określonego poziomu. Stworzenie nowej klasy powoduje znaczący spadek w przeciętnej wielkości klas tworząc losowe źródło wariancji. W tym przypadku zastosowano dwie zmienne instrumentalne. Jedną z nich jest podobnie jak poprzednio przeciętna liczba uczniów w klasie w szkole, a drugą funkcja przewidującą liczbę uczniów w klasach na podstawie liczby uczniów w szkole i przyjętej zasady określającej maksymalną wielkość klasy. Poniżej omówiono różnice między dwoma podejściami. Warto jednak zaznaczyć, że obydwa podejścia w określonych warunkach dają quasi-eksperymentalne oszacowanie interesujących nas zależności. Stąd ich wykorzystanie na tym samym zbiorze danych daje wartościową możliwość porównania, potwierdzenia lub też podważenia, uzyskanych rezultatów.

W dalszej części rozdziału przedstawiono po kolei: charakterystykę zbioru danych, metodologię badania oraz uzyskane rezultaty wraz z interpretacją.

II. Dane.

W badaniu wykorzystano dane o wynikach uczniów na sprawdzianie szóstoklasistów w szkołach podstawowych w latach 2002-2004. Sprawdzian jest egzaminem obowiązkowym, wielopredmiotowym, którego celem jest sprawdzenie wiedzy uczniów pod kątem kontynuacji ich nauki w gimnazjum. Przeprowadzają go Okręgowe Komisje Egzaminacyjne, na identycznych zasadach w całej Polsce, a oceniają zewnętrzni wobec szkoły egzaminatorzy. Zdają go niemal wszyscy uczący się według standardowego programu nauczania. Wyjątki są bardzo rzadkie.

W badaniu wykorzystano wyniki z województwa mazowieckiego, przy czym Okręgowa Komisja Egzaminacyjna w Warszawie udostępniła jedynie średnie na poziomie klas. Trzeba podkreślić, że województwo mazowieckie jest niezwykle zróżnicowane, przy czym takie charakterystyki jak lokalizacja szkoły, zasobność i wykształcenie mieszkańców, dostęp do kultury itp. są silnie skorelowane zarówno z wielkością klas, jak i osiągnięciami uczniów.

Dane uzyskane z OKE w Warszawie zawierały liczbę uczniów w klasie i jej średni wynik, a także nazwę i kod szkoły oraz gminy, na podstawie których połączono wyniki szkół między latami, a także oznaczono szkoły nietypowe. Liczbę uczniów w szkole otrzymano jako sumę liczby uczniów we wszystkich jej klasach, co bliskie jest rzeczywistości ze względu na niewielką liczbę uczniów, którzy do sprawdzianu nie podeszli³.

Przekazany przez OKE zbiór zawierał 9462 klasy. Na podstawie nazw szkół odrzucono z niego 304 klasy w szkołach prywatnych oraz 324 klasy w szkołach przyszpitalnych, specjalnych, muzycznych, sportowych oraz dla dorosłych. Pozostały 8834 klasy w typowych szkołach publicznych, dla których zasady regulujące wielkość klas są podobne i które stanowiły obiekt zainteresowania w tym badaniu. Z tych obserwacji odrzucono 143 klasy zawierające wg danych OKE ponad 40 uczniów, co można traktować jako pomyłkę przy kodowaniu, bowiem ich legalne prowadzenie nie jest możliwe, a także 27 klas z mniej niż 10 uczniami w szkołach z więcej niż 20 uczniami, bowiem można przypuszczać, że to klasy specjalne tworzone przy zwykłych szkołach. W ostatecznym zbiorze znalazły się 8664 klasy. Tabela poniżej zawiera opis statystyczny wybranych zmiennych dla tego zbioru danych. Na nim przeprowadzono analizy, których rezultaty omówiono poniżej. Analizy te powtórzono jednak także dla pełnego zbioru przekazanego przez OKE. Nie zmieniło to charakteru uzyskanych rezultatów, co świadczy o ich odporności na dobór próby.

³ Więcej informacji o wynikach i samym egzaminie znaleźć można w oficjalnych raportach OKE w Warszawie dostępnych na www.oke.waw.pl.

Tabela 1. Statystyki opisowe zmiennych wykorzystanych w analizie

	cała próba			przedziały nieciągłości*					
	2002-2004			29 +/-5			29 +/-4		
liczba uczniów w klasie	21.05 (6.55)			24.03 (5.12)			24.18 (5.23)		
liczba uczniów w szkole	63.49 (49.17)			56.59 (23.47)			50.38 (23.27)		
liczba klas	8664			1867			1511		
liczba szkół	1657			658			561		
	2002 rok			2003 rok			2004 rok		
	miasta bez Warszawy			miasta bez Warszawy			miasta bez Warszawy		
	cała próba	Warszawy	wieś	cała próba	Warszawy	wieś	cała próba	Warszawy	wieś
liczba uczniów w klasie	21.02 (6.59)	24.95 (3.93)	17.77 (6.91)	21.15 (6.65)	25.17 (4.02)	17.93 (6.99)	20.98 (6.38)	23.96 (5.15)	18.01 (6.73)
średni wynik klasy	30.36 (2.89)	30.41 (2.83)	29.76 (2.88)	27.56 (3.00)	28.13 (2.55)	26.48 (2.87)	25.92 (3.43)	26.15 (3.06)	24.68 (3.09)
liczba klas	2891	651	1206	2903	648	1218	2852	700	1124

W nawiasach podano błędy standardowe.

* Przedziały nieciągłości zawierają obserwacje dla liczby uczniów w szkole między:

- dla 29 +/-5 : 24-34, 53-63, 82-92,

- dla 29 +/-4 : 25-33, 54-62, 83-91.

III. Metodologia.

Celem przyjętej strategii empirycznej jest uniknięcie obciążenia wynikającego z endogeniczności zmiennej określającej liczbę uczniów w klasie. Zmienną zależną w każdym modelu jest średni wynik klasy. Ogólny model opisuje równanie poniżej:

$$T_{kst} = \alpha + \beta_C c_{kst} + \mathbf{K}'_{kst} \boldsymbol{\beta}_K + \mathbf{S}'_{st} \boldsymbol{\beta}_S + \mathbf{Y}'_t \boldsymbol{\varphi}_t + \varepsilon_{kst} \quad (1)$$

gdzie zmienną objaśnianą T_{kst} jest średni wynik w klasie, zmienna c_{kst} oznacza logarytm faktycznej liczebności klasy, wektor \mathbf{K}_{kst} zawiera zmienne na poziomie klasy (k), wektor \mathbf{S}_{st} zmienne na poziomie szkoły (s), \mathbf{Y}_t jest wektorem zmiennych binarnych dla lat obserwacji (t). Celem jest oszacowanie parametru β_C . Podobnie jak w pracy Hoxby (2000) faktyczną liczbę uczniów w klasie poddano transformacji logarytmicznej, ponieważ zmiana liczby uczniów o jedną osobę jest proporcjonalnie większa w przypadku mniejszych klas. Redukcja piętnastoosobowej klasy o pięciu uczniów może mieć potencjalnie większy efekt niż redukcja klasy trzydziestoosobowej o tę samą liczbę uczniów. Uwzględnienie w równaniu logarytmu naturalnego liczby uczniów w klasie ma na celu ocenę wpływu procentowych zmian tej zmiennej.

Taka specyfikacja równania określającego czynniki wpływające na osiągnięcia uczniów (czyli tzw. „edukacyjnej funkcji produkcji”) i jej estymacja metodą najmniejszych kwadratów (dalej MNK) daje obciążone oceny parametru β_C , nawet gdy równanie uwzględniałoby bogaty zestaw zmiennych kontrolnych, łącznie z nakładami szkolnymi z lat poprzednich (por. Todd, Wolpin, 2003). Spowodowane jest to faktem, że wielkość klasy jest skorelowana z wieloma niemierzalnymi bądź nieobserwowalnymi czynnikami, a to z kolei może być źródłem endogeniczności wielkości klasy. W takim przypadku regresory są skorelowane z błędem losowym, co oznacza, że estymatory uzyskane MNK są obciążone i niezgodne. Powyżej wskazywaliśmy, że potencjalnymi źródłami endogeniczności są efekty sortowania uczniów o różnych umiejętnościach i pochodzeniu społecznym, między szkołami i wewnątrz szkół, do klas o różnych liczebnościach.

Problemy te zdają się potwierdzać sprzeczne rezultaty badań opierających się na estymacji edukacyjnej funkcji produkcji, które dla niektórych stanowią dowód na brak zależności przyczynowo-skutkowej między zasobami szkolnymi a osiągnięciami uczniów (por.

Hanushek, 2003). Według wielu badaczy takie wnioski są jednak nieuprawnione, o ile nie opieramy się na badaniach potencjalnie dających możliwość oszacowania nieobciążonych ocen efektów zasobów szkolnych. Takimi badania są eksperymenty, w których możemy porównać wyniki uczniów losowo przydzielonych do klas różnej wielkości, z których najbardziej znanym przykładem jest projekt STAR przeprowadzony w stanie Tennessee w USA. Zaawansowana metodologicznie analiza danych uzyskanych w tym eksperymencie wskazuje na to, że mniejsze klasy sprzyjają podnoszeniu osiągnięć uczniów (Krueger, 2003). Inni badacze sugerują jednak, że badania eksperymentalne prowadzone na dużą skalę dają także obciążone rezultaty, w tym wypadku m.in. ze względu na zmiany w składzie analizowanych grup i niekontrolowanym wpływie sytuacji eksperymentalnej na postawy nauczycieli (por. Hoxby, 2000).

Stąd zainteresowanie badaczy budzą badania quasi-eksperymentalne, czy też tzw. naturalne eksperymenty, które analizowane są najczęściej poprzez wykorzystanie zmiennych instrumentalnych i podwójnej metody najmniejszych kwadratów. Ich zastosowanie jest ograniczone przez dość wąski zbiór potencjalnych instrumentów, których zły dobór może w rezultacie dawać bardziej obciążone estymatory (Angrist, Krueger, 2001). W niniejszym badaniu jako zmienną instrumentalną dla c_{kst} wykorzystano logarytm naturalny średniej wielkości klasy w danej kohorcie w danej szkole. Podejście takie zostało już zastosowane wcześniej w podobnych badaniach (Akerhielm, 1995; Wößmann, 2003). Instrument ten pozwala uniknąć obciążenia wynikającego z sortowania uczniów o różnych zdolnościach do klas różnej wielkości w danej szkole. W sytuacji, kiedy nie występuje sortowanie uczniów między szkołami bądź, kiedy zmienne w równaniu w pełni kontrolują to zjawisko, instrument taki uważa się za zmienną egzogeniczną.

Pierwszym krokiem w stosowaniu podwójnej metody najmniejszych kwadratów jest estymacja równania:

$$c_{kst} = \varpi + \beta_{\bar{c}} \bar{c}_{kst} + \mathbf{K}'_{kst} \boldsymbol{\beta}_{\mathbf{K}} + \mathbf{S}'_{st} \boldsymbol{\beta}_{\mathbf{S}} + \mathbf{Y}_t \boldsymbol{\varphi}_t + \eta_{kst} \quad (2)$$

gdzie logarytm liczebności klasy c_{kst} jest wyjaśniany przez logarytm średniej wielkości klasy w kohorcie \bar{c}_{kst} i inne zmienne egzogeniczne. W drugim kroku, podstawiając wartości teoretyczne $\hat{c}_{kst} = c_{kst} - \eta_{kst}$ z pierwszego równania, szacujemy następującą regresję:

$$T_{kst} = \alpha + \beta_{\hat{c}} \hat{c}_{kst} + \mathbf{K}'_{kst} \boldsymbol{\beta}_{\mathbf{K}} + \mathbf{S}'_{st} \boldsymbol{\beta}_{\mathbf{S}} + \mathbf{Y}_t \boldsymbol{\varphi}_t + \varepsilon_{kst} \quad (3)$$

gdzie przy spełnieniu założeń β_c jest nieobciążonym estymatorem efektu wielkości klasy. Stosując powyższe rozwiązanie dla danych indywidualnych Karen Akerhielm zaobserwowała dość silny ujemny efekt wielkości klasy (Akerhielm, 1995).

Średnia liczebność klasy może być wykorzystana jako instrument tylko w przypadku, kiedy różnice między szkołami są w pełni wyjaśniane przez zmienne kontrolne. Taka sytuacja ma rzadko kiedy miejsce. Przykładowo, niezwykle rzadko dostępne są zmienne charakteryzujące decyzje rodziców dotyczące miejsca zamieszkania lub wyboru szkoły, a jeśli są one powiązane z wielkością klas i zdolnościami dziecka lub motywacją rodziców do jego kształcenia, to uzyskane oceny parametrów będą obciążone. Jednym ze sposobów wyjścia z tej sytuacji jest uwzględnienie w estymowanym równaniu efektów stałych szkół lub rejonów, o których zakłada się, że mają podobny charakter. Przykładem takiego rozwiązania mogą być badania przeprowadzone na danych pochodzących z projektu TIMSS, w którym testy były rozwiązywane przez wszystkich uczniów w dwóch następujących po sobie rocznikach w każdej szkole w całej próbie. Wykorzystując ten fakt Ludger Wößmann i Martin West zastosowali podwójną metodę najmniejszych kwadratów ze średnią wielkością klasy w szkole jako instrumentem, estymując dodatkowo efekty stałe dla szkół w celu uwzględnienia różnic między nimi (Wößmann, West, 2002; Wößmann, 2003). Zdaniem autorów strategia taka pozwala na otrzymanie quasi-eksperymentalnych oszacowań wpływu wielkości klasy, ponieważ pod uwagę brana jest jedynie wariancja liczebności klasy między rocznikami w tej samej szkole, na co z kolei wpływ ma głównie fluktuacja w wielkości roczników.

W pierwszej strategii empirycznej w niniejszym badaniu zastosowano podejście zaproponowane przez Wößmanna i Westa. Analizowane dane zawierają obserwacje z trzech lat, co umożliwia uwzględnienie efektów stałych szkół dla uniknięcia obciążenia wynikającego z sortowania uczniów między nimi, a równoczesne zastosowanie przeciętnej wielkości klasy w danym roczniku w szkole, jako zmiennej instrumentalnej, pozwala uniknąć obciążenia powodowanego sortowaniem uczniów wewnątrz szkół. Ze względu na brak dodatkowych zmiennych kontrolnych w równaniach uwzględniono jedynie wyniki uczniów, wielkość klas oraz efekty stałe szkół (wektor zmiennych zerojedynkowych \mathbf{D}_s przyjmujących 1 dla obserwacji z danej szkoły i 0 w przeciwnym przypadku):

$$T_{kst} = \alpha + \beta_c \hat{c}_{kst} + \mathbf{D}'_s \boldsymbol{\beta}_s + \mathbf{Y}'_t \boldsymbol{\theta}_t + \varepsilon_{kst} \quad (4)$$

Zastosowanie podwójnej metody najmniejszych kwadratów ze średnią wielkością klasy w roczniku w danej szkole i uwzględnieniem zmiennych zerojedynkowych dla szkół umożliwia skoncentrowanie się na tej części wariancji liczebności klasy, która wynika ze zmian w liczebności kohorty między latami w tej samej szkole. Według Wößmanna i Westa tę część wariancji można określić jako egzogeniczną (Wößmann, West, 2002). Jednak duże zmiany w liczbie uczniów w szkole mogą być powodowane czynnikami nielosowymi, np. rosnącą popularnością placówki lub migracją. Stąd powyższy model oszacowano także dla próby, gdzie liczba klas pomiędzy latami zmieniła się nie więcej niż o +1.

W drugiej strategii empirycznej wykorzystano fakt, że większość szkół w Polsce tworzy nowe klasy, o ile przeciętna ich wielkość zbliża się lub przekracza określony próg. Stworzenie nowej klasy powoduje drastyczny spadek przeciętnej liczby uczniów w klasach w danej szkole, co tworzy źródło losowej wariancji, o ile decyzja o stworzeniu nowej klasy wynika jedynie z fluktuacji w liczbie uczniów między rocznikami. Aby zastosować tę metodę należy obliczyć funkcję według wzoru poniżej:

$$C_{st}^* = f(E_{st}) = \frac{E_{st}}{\text{int}\left(\frac{E_{st}-1}{C^{\max}}\right) + 1} \quad (5)$$

a jej logarytm naturalny wykorzystać jako instrument dla C_{kst} , czyli rzeczywistej liczby uczniów w klasie:

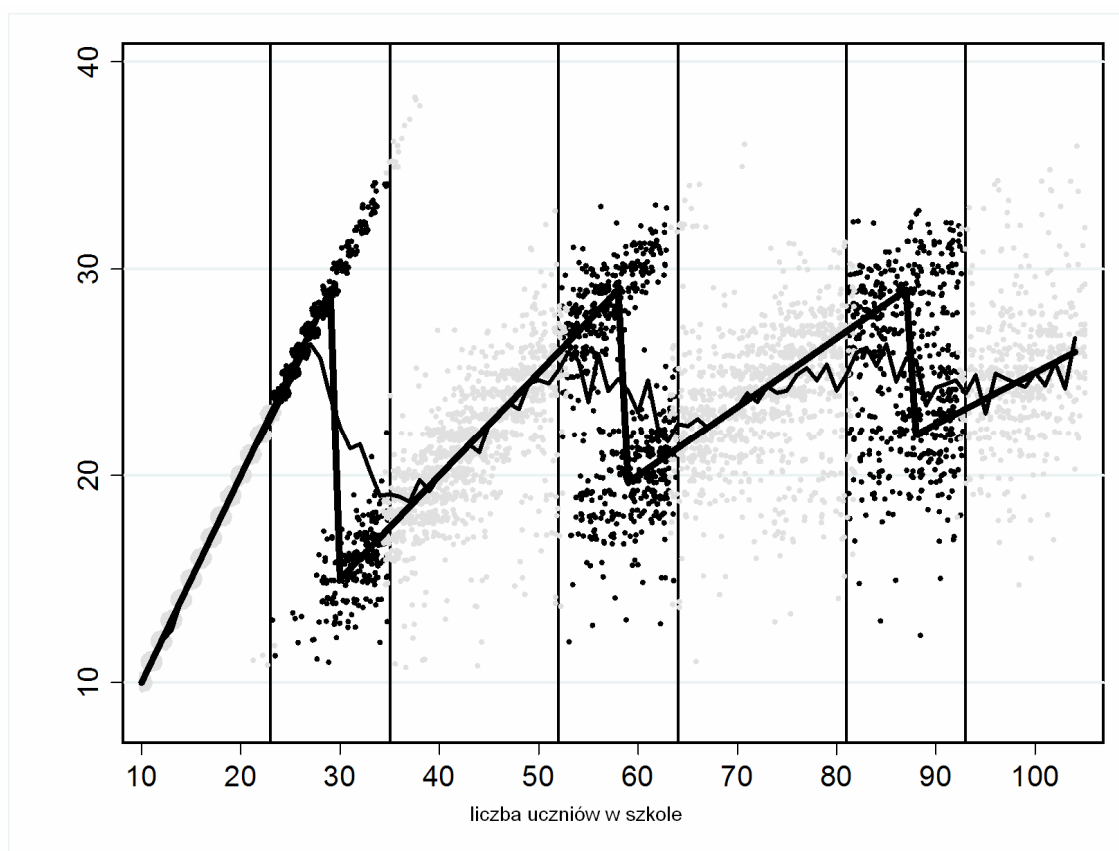
$$c_{st}^* = \ln(C_{st}^*) \quad (6)$$

gdzie C_{st}^* oznacza przewidywanie wielkości klas dla danej liczby uczniów w szkole E_{st} oraz reguły określającej jako C^{\max} maksymalną liczbę uczniów w klasie. Idąc za autorami tej metody równania (5) i (6) w dalszej części pracy określamy jako „teoretyczną funkcję” wielkości klasy (por. Angrist, Lavy, 1999). Wykorzystanie wyniku tej funkcji C_{st}^* jako instrumentu dla faktycznej wielkości klasy jest przykładem tzw. „regression-discontinuity design” w wersji rozmytej (“fuzzy version”), bowiem niektóre szkoły dodają nowe klasy wcześniej a niektóre później (Lee, 2005).

Sukces podejścia Angrista i Lavy’ego opierał się na znalezieniu prawnego przepisu określającego maksymalną wielkość klasy, którego powszechne stosowanie tworzy losową wariancję. Inaczej mówiąc, Angrist i Lavy znaleźli „naturalny eksperyment” wynikający z

zastosowania regulacji prawnych. W rzeczywistości trudno oczekiwać, że w szkoły będą stosować tę zasadę bezwarunkowo. Jednak dla oszacowania nieobciążonego wpływu wielkości klasy na wyniki wystarczy założenie, że większość szkół stosuje się do tej zasady, czego efektem jest to, że dla liczby uczniów poniżej pewnego progu większość szkół ma o jedną klasę mniej niż większość szkół powyżej tego progu. Mimo, że w Polsce nie istnieje przepis prawny ściśle określający maksymalną wielkość klasy, to dane empiryczne pokazują, że mamy do czynienia z dokładnie taką sytuacją. Ukazuje to wykres poniżej.

Wykres 1. Przewidywane i rzeczywiste wielkości klas a liczba uczniów w szkole.



Na wykresie na osi poziomej oznaczono liczbę uczniów w szkole a na osi pionowej liczbę uczniów w klasie. Gruba linia oznacza przewidywaną na podstawie funkcji teoretycznej średnią wielkość klasy, a cienką linię przeciętną rzeczywistych średnich wielkości klas dla szkół o tej samej liczbie uczniów. Pojedyncze punkty to obserwacje w analizowanym zbiorze

danych⁴. Widać, że większość szkół tworzy nowe klasy, gdy przeciętna zbliża się do 28-29 uczniów w klasie. Trzy przedziały, które wyraźniej przedstawiono na wykresie, to tzw. przedziały nieciągłości („discontinuity samples”) zawierające szkoły w okolicach progu, przy którym tworzone są nowe klasy. Przedziały na wykresie zawierają placówki z liczbą uczniów w szkole w granicach 24-34, 53-63 oraz 82-92, a więc dla wielokrotności 29 i +5 uczniów. Widać, że w przedziałach tych przeciętna liczba uczniów w klasach maleje, a poza nimi rośnie. Wykres ten jest bardzo zbliżony do przedstawionego w pracy Angrista i Lavy’ego, co sugeruje, że podobna metoda może być zastosowana dla danych polskich, mimo że formalna, prawna reguła określająca maksymalną wielkość klasy w Polsce nie istnieje⁵. Na wykresie widać też, że przeciętna wielkość klasy jest bliska przewidywanej przez teoretyczną funkcję poza przedziałami nieciągłości. W istocie, te dwa, zastosowane w tej pracy instrumenty, są poza przedziałami nieciągłości niemal identyczne i narażone na podobne problemy, przez co szacowanie wpływu wielkości klasy na podstawie wariancji w całej próbie wymaga uwzględnienia efektów stałych szkół. Jednak w przedziałach nieciągłości, gdzie wariancja powodowana jest głównie tworzeniem nowych klas, obydwa instrumenty mogą dawać nieobciążone oszacowania⁶.

Warto w tym miejscu podkreślić różnice między opisanymi powyżej zmiennymi instrumentalnymi. Są one identyczne, o ile wszystkie szkoły posiadają taką samą liczbę klas. W przeciwnym wypadku, a jest tak przede wszystkim w przedziałach nieciągłości, działają odmiennie. W przedziałach tych wykorzystanie przeciętnej liczby uczniów w klasie jako instrumentu powoduje, że porównujemy szkoły podobnej wielkości, lecz z inną liczbą klas. Jeżeli przy tym za zmienną kontrolną posłuży liczba uczniów w szkole, tak jak to zostało uczynione w tym badaniu, to metoda ta daje w istocie oszacowania wynikające z różnic w wynikach między szkołami tej samej wielkości, lecz różniących się liczbą klas. Przy założeniu, że w przedziałach nieciągłości o tym czy stworzyć nową klasę decyduje głównie czynnik losowy, oszacowania te będą nieobciążone. W innym wypadku, gdy nowe klasy tworzą wcześniej szkoły o charakterystykach skorelowanych z wynikami uczniów, oszacowania te będą obciążone. Wykorzystując teoretyczną funkcję klasy jako instrument,

⁴ Do faktycznej liczby uczniów dodano na wykresie niewielki czynnik losowy dla uniknięcia pokrywania się obserwacji.

⁵ Funkcja teoretyczna skonstruowana dla 28 lub 29 uczniów jako hipotetycznego maksimum wyjaśnia ok. 60% wariancji w wielkości klas. Podobną moc wyjaśniającą miał instrument wykorzystany w badaniu szkół w Izraelu (Angrist, Lavy, 1999).

⁶ W istocie powinno się korzystać z wariancji jedynie w punktach nieciągłości, jednak w takim przypadku uzyskanie precyzyjnych oszacowań wymagałoby niezwykle licznej próby, trudnej do uzyskania w rzeczywistości (por. Hoxby 2000).

pomijamy wariancję powodowaną przez szkoły, które „zbyt wcześnie” lub „zbyt późno” tworzą nowe klasy. Ma to sens, gdy możemy być pewni, że szkoły, które nie stosują się do zasady określającej próg tworzenia nowej klasy są nietypowe. Można jednak przypuszczać, że w rzeczywistości tak nie jest. Dla przykładu wyobraźmy sobie trzy szkoły o liczbie uczniów odpowiednio 56, 57 i 58. Załóżmy, że najmniejsza z nich stworzyła nową klasę, bowiem początkowo zapisało się do niej 3 uczniów więcej, jednak ostatecznie przeszli oni do innej placówki. Stosując teoretyczną funkcję jako instrument pomijamy w takim przypadku wartościową wariancję, czego nie robimy używając jako instrumentu przeciętnej liczby uczniów w klasie. Jednak w tym drugim przypadku bierzemy także pod uwagę wariancję, która może być powodowana świadomymi decyzjami szkół.

Wyboru między opisanymi powyżej metodami i zmiennymi instrumentalnymi nie można oprzeć na danych empirycznych. W istocie, podobnie jak w wielu innych przypadkach, gdzie stosowane są zmienne instrumentalne, wybór ten oparty jest na trudnych do sprawdzenia założeniach, które odwołują się do wiedzy na temat danego systemu oświatowego. Stąd, poniżej zaprezentowano wyniki dla obydwu metod i obydwu zmiennych instrumentalnych, co daje możliwość porównania rezultatów. Dodatkowo, oszacowania wykonane zostały dla różnych przedziałów nieufności tak, aby uniezależnić wyniki od ich doboru.

IV. Wyniki i ich omówienie.

W tabeli poniżej przedstawiono wyniki analiz przeprowadzonych osobno dla danych z każdego roku. W nawiasach podano błędy standardowe, skorygowane ze względu na korelację między obserwacjami w jednej szkole. Jedna, dwie i trzy gwiazdki oznaczają istotność statystyczną współczynników odpowiednio na poziomie 0.10, 0.05 oraz 0.01. Rezultaty regresji oszacowaną MNK, bez zmiennych kontrolnych, pokazują, że wielkość klasy pozytywnie wpływa na osiągnięcia uczniów. Dodanie liczby uczniów w szkole obniża oceny współczynników wpływu wielkości klasy, jednak wciąż pozostają one pozytywne, przy tym widać, że liczba uczniów w szkole jest także pozytywnie skorelowana z osiągnięciami. Oszacowania uzyskane przez podwójną MNK ze zmiennymi instrumentalnymi dają także pozytywne lub nieistotne statystycznie oceny.

Tabela 2. Rezultaty dla prób z poszczególnych lat.

zmienna wyjaśniana: średni wynik klasy na sprawdzianie						
	2002		2003		2004	
zmienne wyjaśniające: ↓	(1)	(2)	(3)	(4)	(5)	(6)
MNK						
logarytm liczby uczniów w klasie	.475*** (.131)	.164 (.159)	1.315*** (.125)	.515*** (.147)	1.560*** (.157)	.478** (.188)
liczba uczniów w szkole	—	.004*** (.001)	—	.012*** (.001)	—	.016*** (.001)
R^2	0.0041	0.0078	0.0364	0.0670	0.0328	0.0660
2-MNK - Zmienna instrumentalna: przeciętna liczba uczniów w klasie w szkole						
logarytm liczby uczniów w klasie	.218 (.134)	-.246 (.164)	1.167*** (.127)	.235 (.152)	1.396*** (.161)	.145 (.194)
liczba uczniów w szkole	—	.006*** (.001)	—	.014*** (.001)	—	.017*** (.001)
R^2	0.0028	0.0055	0.0359	0.0658	0.0324	0.0650
2-MNK - Zmienna instrumentalna: funkcja teoretyczna						
logarytm liczby uczniów w klasie	.212 (.142)	-.354 (.181)	1.347*** (.140)	.234 (.176)	1.724*** (.173)	.349* (.217)
liczba uczniów w szkole	—	.006*** (.0013)	—	.014*** (.001)	—	.017*** (.001)
R^2	0.0028	0.0042	0.0364	0.0658	0.0324	0.0658
N	2891		2903		2852	

W nawiasach podano błędy standardowe, skorygowane ze względu na korelację między obserwacjami w jednej szkole. Jedna, dwie i trzy gwiazdki oznaczają istotność statystyczną współczynników odpowiednio na poziomie 0.10, 0.05 oraz 0.01.

Powyższe rezultaty są obciążone ze względu na selekcję uczniów między szkołami. W tabeli 3 zaprezentowano wyniki regresji na całej próbie z 3 lat, gdzie przedstawiono także oszacowania z efektami stałymi szkół⁷. Kolumna (1) zawiera rezultaty regresji bez efektów stałych, wskazujące na pozytywną zależność między wielkością klasy a wynikami uczniów, zarówno w przypadku metody najmniejszych kwadratów, jak i przy zastosowaniu obydwu zmiennych instrumentalnych. Rezultaty te stanowią w istocie powtórzenie wyników z poprzedniej tabeli, dla regresji na danych z pojedynczych lat. Należy je porównać do kluczowych rezultatów regresji z efektami stałymi szkół przedstawionymi w kolumnie (2). Dla oszacowań MNK wpływ wielkości klasy na wyniki jest istotny statystycznie i pozytywny, jednak dla obu zmiennych instrumentalnych jest on istotny statystycznie, lecz negatywny. Należy podkreślić zmianę znaku w ocenie parametrów.

Wiarygodność rezultatów przedstawionych w tej tabeli opiera się na założeniu o losowości zmian w wielkości populacji uczniów między latami. Założenie to nie zawsze jest spełnione, a szczególnie „podejrzane” mogą być znaczne zmiany w liczbie uczniów w danej szkole. Zmiany te mogą być powodowane dołączeniem nowych klas ze zlikwidowanych szkół, rosnącą popularnością danej placówki, położeniem szkoły w rejonie znacznej migracji ludności itp. Z tego powodu, aby wykluczyć potencjalne źródła nielosowej wariancji, powtórzono wszystkie oszacowania na próbie ograniczonej do szkół, gdzie między kolejnymi latami liczba klas nie zmieniła się lub zmieniła się o jeden (na plus lub minus). W ten sposób ograniczono źródło wariancji w wielkości klas do niewielkich zmian w liczbie uczniów w szkole, które można przypisać losowym zmianom w liczbie urodzeń w roczniku w danej miejscowości lub okręgu szkolnym. Wyniki regresji na tak zawężonej próbie przedstawiono w kolumnach (3) i (4), przy czym należy porównać kluczowe rezultaty z kolumny (4) z oszacowaniami na całej próbie z kolumny (2). Widać, że przyjęte ograniczenia nie spowodowały znaczącej zmiany w ocenach parametrów. Przeprowadzono także analizę regresji dla prób ograniczonych do 20% i 30% zmian w populacji uczniów w danej szkole, jednak wyniki były bardzo podobne i nie zaprezentowano ich poniżej.

⁷ We wszystkich regresjach na danych łączonych z 3 lat wprowadzono zmienne zerojedynkowe dla oznaczenia wyników z danego roku, które dla większej klarowności pominięto w tabelach. Współczynniki przy tych zmiennych były statystycznie istotne i prawidłowo odzwierciedlały spadek średniej dla całej populacji w kolejnych latach.

Tabela 3. Rezultaty dla próby łącznej z 3 lat.

zmienna wyjaśniana: średni wynik klasy na sprawdzianie				
	Pełna próba		Próba z ograniczeniami *	
zmienne wyjaśniające: ↓	(1)	(2)	(3)	(4)
MNK				
logarytm liczby uczniów w klasie	1.119*** (.079)	.919*** (.150)	1.113*** (0.125)	.797*** (.247)
zmienne zerojedynkowe dla szkół	nie	tak	nie	tak
R^2	0.2736	0.5628	0.2703	0.6522
2-MNK - Zmienna instrumentalna: przeciętna liczba uczniów w klasie w szkole				
logarytm liczby uczniów w klasie	.930*** (.081)	-.583*** (.173)	0.941*** (.127)	-.644*** (.234)
zmienne zerojedynkowe dla szkół	nie	tak	nie	tak
R^2	0.2732	0.5566	0.2699	0.6477
2-MNK - Zmienna instrumentalna: funkcja teoretyczna				
logarytm liczby uczniów w klasie	1.086*** (.087)	-.755*** (.227)	1.104*** (0.141)	-.746*** (.297)
zmienne zerojedynkowe dla szkół	nie	tak	nie	tak
R^2	0.2736	0.5551	0.2703	0.6470
N	8664		8414	

W nawiasach podano błędy standardowe, skorygowane ze względu na korelację między obserwacjami w jednej szkole. Jedna, dwie i trzy gwiazdki oznaczają istotność statystyczną współczynników odpowiednio na poziomie 0.10, 0.05 oraz 0.01.

* Próba dla szkół, gdzie liczba klas z roku na rok nie zmieniła się o więcej niż +/-1.

Celem drugiej strategii empirycznej było wykorzystanie wariacji powodowanej przez tworzenie nowych klas. W tabeli poniżej zaprezentowano wyniki oszacowań dla progów tworzenia nowych klas określonego jako 29 i jego wielokrotności, przy czym różnie zdefiniowano szerokość przedziałów. W kolumnach (1) i (2) przedstawiono rezultaty dla szerokości przedziału określonej jako +/-5 uczniów, a w kolumnach (3) i (4) dla +/-4 uczniów. Podobne rezultaty uzyskano dla analiz powtórzonych dla progów równych 28 oraz różnych szerokości przedziału. Trzeba jednak podkreślić, że zmniejszanie szerokości przedziałów, choć zwiększa prawdopodobieństwo, że analizowana wariacja ma losowe źródło, to obniża precyzję oszacowania, co w tym przypadku prowadziło do uzyskania rezultatów o niskiej wiarygodności statystycznej. Natomiast poszerzanie przedziałów podważa sens metody, która może być stosowana jedynie w pobliżu „nieciągłości” powodowanych tworzeniem lub likwidacją klas. Ponadto w badaniu wykorzystano obserwacje leżące jedynie w pierwszych trzech przedziałach nieciągłości (dla progów równych 29, 58 i 87), bowiem w kolejnych przedziałach nie widać było klarownej zależności między liczbą uczniów w szkole a progiem, dla którego tworzone są nowe klasy.

Rezultaty umieszczone w tabeli poniżej pokazują, że dla zwykłej regresji MNK większość oszacowań sugeruje brak zależności między liczbą uczniów w klasach a wynikami. Zastosowanie zmiennych instrumentalnych, przy wprowadzeniu jako zmiennej kontrolnej liczby uczniów w szkole, prowadzi do rezultatów wskazujących na negatywny wpływ wielkości klas, co widać w kolumnach (2) i (4). Warto odnotować, że zastosowanie jako zmiennej instrumentalnej teoretycznej funkcji przewidującej wielkość klasy prowadzi do znacznie większego błędu standardowego oszacowanych współczynników. Jedynie dla przeciętnej liczby uczniów w szkole jako zmiennej instrumentalnej uzyskane rezultaty są istotne statystycznie, choć precyzja oszacowania jest niższa niż w wypadku pierwszej strategii empirycznej z efektami stałymi szkół. Widać też, że wraz ze zmniejszaniem szerokości przedziałów nieciągłości precyzja ta maleje. Oszacowanie wpływu wielkości klasy na podstawie wariacji tylko w punktach nieciągłości wymagałoby znacznie liczniejszej próby.

Tabela 4. Rezultaty dla przedziałów nieciągłości.

zmienna wyjaśniana: średni wynik klasy na sprawdzianie				
	29 +/-5 24-34, 53-63, 82-92		29 +/-4 25-33, 54-62, 83-91	
zmienne wyjaśniające: ↓	(1)	(2)	(3)	(4)
MNK				
logarytm liczby uczniów w klasie	.478 (.304)	.013 (.299)	.346 (.326)	-.0204 (.320)
liczba uczniów w szkole	—	.0302*** (.003)	—	.0289*** (.003)
R^2	0.2669	0.3053	0.2973	0.3317
2-MNK, Zmienna instrumentalna: przeciętna liczba uczniów w klasie w szkole				
logarytm liczby uczniów w klasie	-.073 (.321)	-.632** (.317)	-.139 (.344)	-.576* (.338)
liczba uczniów w szkole	—	.031*** (.003)	—	.029*** (.003)
R^2	0.2656	0.3036	0.2963	0.3304
2-MNK – Zmienna instrumentalna: funkcja teoretyczna				
logarytm liczby uczniów w klasie	2.107** (.978)	-.469 (1.044)	2.251* (1.202)	-.242 (1.268)
liczba uczniów w szkole	—	.031*** (.003)	—	.0292*** (.004)
R^2	0.2656	0.3036	0.2963	0.3304
N	1867		1511	

Przedstawione rezultaty przekonują, że wzrost liczby uczniów w klasie ma negatywny wpływ na wyniki uczniów. Zwiększenie liczby uczniów o 10% obniża przeciętnie średni wynik klasy o ok. 0.05-0.08 punktów. To mniej niż 5% odchylenia standardowego średnich wyników klas. W innych studiach wykorzystujących podobną metodologię uzyskano nieco silniejsze efekty (por. Angrist, Lavy, 1999; Krueger, 2003; Urquiola, 2000), należy też jednak podkreślić, że niektóre z nich wskazywały na brak interesującej nas zależności (por. Hoxby 2000). Można stwierdzić, że oszacowane efekty są zbyt małe, aby miały praktyczne znaczenie. Byłoby to jednak błędem, bowiem nie znamy czynników, które są w Polsce powodem zróżnicowania wyników szkół, a przez to nie wiemy jakie znaczenie praktyczne mają różnice punktowe między uczniami, klasami i szkołami. Być może większość wariacji powodowana jest różnicami w środowisku społecznym uczniów, niezależnymi od polityki oświatowej, a charakterystyki, które możemy kontrolować mają znacznie mniejszy wpływ. Stąd, konstatacja, że uzyskane tu rezultaty nie mają praktycznego znaczenia, nie ma większego sensu, o ile nie jest umieszczona w odpowiednim kontekście, stanowiącym zainteresowanie polityki oświatowej. Podobnie jak inne stwierdzenia, że dany efekt jest „niewielki”, pojawiające się często w Polsce odnośnie czynników wpływających na wyniki egzaminów zewnętrznych.

Uzyskane rezultaty można zestawić z różnicą między wynikami uczniów w gminach wiejskich i miejskich. We wszystkich latach uczniowie szkół miejskich uzyskali wyższe rezultaty niż uczniowie szkół wiejskich, a różnice w osiągnięciach są istotne statystycznie i potwierdzają je m.in. wyniki międzynarodowych testów PISA. Jednym z celów stawianych przed reformą oświaty było ich zmniejszenie, a proponowanym rozwiązaniem była likwidacja mniejszych szkół, tworzenie większych klas i inwestowanie zaoszczędzonych w ten sposób zasobów w inne nakłady sprzyjające podnoszeniu jakości nauczania (por. Levitas, Golinowska, Herczyński, 2001, jako przykład takiego podejścia). Nie wiemy, jakie byłyby efekty inwestowania w „inne” nakłady, ponieważ brakuje w Polsce takich badań. Możemy jednak ocenić, w jaki sposób zwiększenie wielkości klas na terenach wiejskich wpłynęłoby na różnicę między wynikami uczniów wiejskich i miejskich. Autorzy wspomnianego przed chwilą raportu za jeden z sukcesów reformy oświaty uznają to, że przeciętna liczba uczniów w nowo utworzonych gimnazjach była początkowo wyższa na wsi niż w miastach. Można zastanowić się, jaki byłby wpływ na osiągnięcia uczniów zmian, dzięki którym klasy na wsi osiągnęłyby przeciętną wielkość klas w miastach.

Przedstawione powyżej rezultaty sugerują, że efekt zwiększenia liczby uczniów w klasie o 10% wynosi od -0.08 do -0.05 punktów. Mając to na uwadze można z pomocą danych zamieszczonych w dolnej części tabeli 1 łatwo obliczyć, że zwiększenie wielkości klas na wsi do poziomu wielkości klas w miastach spowodowałoby zwiększenie różnicy między osiągnięciami uczniów wiejskich i miejskich o 13 do 47 procent, w zależności od roku i przyjętej oceny efektu wpływu wielkości klasy na wyniki. W kalkulacjach tych pominięto klasy ze szkół warszawskich, które z oczywistych względów są wyjątkowe i znacząco odstają nie tylko od średnich dla Mazowsza, ale i dla całej Polski. Pomijając ich wyniki średnie w badanej próbie są podobne do ogólnopolskich. Przy tym zarówno różnice między wynikami na wsi i w miastach, jak i zmiany w tych różnicach spowodowane hipotetyczną zmianą wielkości klas są istotne statystycznie. Można więc stwierdzić, że oszacowany efekt wpływu wielkości klasy, choć wydaje się niewielki jako procent odchylenia standardowego wyników, to nabiera znaczenia w odniesieniu do realnych problemów polityki oświatowej.

V. Podsumowanie.

W pracy tej poddano analizie oddziaływanie liczby uczniów w klasie na wyniki uzyskane przez nich podczas sprawdzianu szóstoklasistów szkół podstawowych w latach 2002-2004. Ze względu na to, że wielkość klasy jest zmienną, która jest silnie skorelowana nie tylko z osiągnięciami uczniów, ale i z cechami je warunkującymi, oszacowanie jej wpływu zwykłą metodą regresji prowadzi do uzyskania obciążonych rezultatów. Niezbędne jest zatem wykorzystanie danych eksperymentalnych lub „naturalnych eksperymentów”. W pracy tej przedstawiono wyniki badań opierających się na dwóch quasi-eksperymentalnych metodach, których zastosowanie znacząco zmienia ocenę wpływu wielkości klasy w porównaniu ze zwykłą metodą regresji. Uzyskane rezultaty sugerują, że wzrost liczby uczniów w klasie negatywnie wpływa na osiągnięcia uczniów, wbrew pozornej, pozytywnej korelacji między wielkością klasy a uzyskanym wynikiem. Rezultaty badania odporne są na drobne zmiany w metodzie i zakresie poddanej analizie próby, a ich wiarygodność potwierdza zgodność wyników uzyskanych przez dwie różne, quasi-eksperymentalne strategie empiryczne.

Odniesienie uzyskanych rezultatów do różnicy między wynikami uczniów na wsi i w miastach pokazuje, że zwiększenie wielkości klas wiejskich może spowodować dalsze,

znaczące jej poszerzenie. Nie ulega wątpliwości, że utrzymywanie mniejszych klas i niewielkich szkół samorządowych jest kosztowne, jednak uzyskane tu rezultaty sugerują, że likwidowanie mniejszych placówek i tworzenie szkół zbiorczych z bardziej licznymi klasami może obniżać wyniki uczniów. Efekt ten należy mieć na uwadze analizując zmiany zachodzące obecnie w oświacie. Koncentrowanie się jedynie na stronie kosztowej może paradoksalnie prowadzić do obniżenia efektywności całego systemu.

Przedstawione tu badanie stanowi przykład analizy, którą można by wykorzystać przy projektowaniu polityki oświatowej. Przykład ten pokazuje jednak, że badania oświatowe wymagają zastosowania odpowiedniej metodologii, a proste analizy, opierające się na standardowych narzędziach statystycznych prowadzą często do stwierdzenia pozornych zależności. Wnioskowanie na podstawie współczynników korelacji lub analizy regresji z potencjalnie endogenicznymi zmiennymi prowadzi do obciążonych, mało wiarygodnych rezultatów, które nie powinny stanowić podstawy dla polityki oświatowej. Ponadto, praktyczne wykorzystanie uzyskanych tu rezultatów wymaga przeprowadzenia ilościowych analiz efektów innych inwestycji i programów interesujących z punktu widzenia polityki oświatowej. Stworzyłoby to możliwość porównania ich efektywności, tworząc wiedzę niezbędną dla podejmowania świadomych decyzji o przyszłości szkół w Polsce.

Bibliografia

- Akerhielm Karen. 1995. Does Class Size Matter? *Economics of Education Review*, XIV, 229-241.
- Angrist Joshua, Krueger Alan. 2001. *Instrumental Variables and the Search for Identification: from Supply and Demand to Natural Experiments*. Princeton University, Working Paper 455.
- Angrist Joshua, Lavy Victor. 1999. Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics*, 533-575.
- Białycki Ireneusz, Haman Jacek. 2003. Program międzynarodowej oceny umiejętności uczniów OECD/PISA. Wyniki polskie - raport z badań. IFiS PAN, Warszawa.
- Dobbelsteen S., Levin J., Oosterbeek H. 2002. The Causal Effect of Class Size on Scholastic Achievement: Distinguishing the Pure Class Size Effect from the Effect of Changes in Class Composition. *Oxford Bulletin of Economics and Statistics*, 64, s. 17-38.
- Hanushek Eric A. 2003. The Failure of Input-based Schooling Policies. *Economic Journal* 113, February, s. F64-F98.
- Herczyński Jan, Herbst Mikołaj. 2002. Pierwsza odsłona. Fundacja Klub Obywatelski.
- Hoxby Caroline Minter. 2000. The effects of class size and composition on student achievement: new evidence from natural population variation. *Quarterly Journal of Economics*, 115(4), 1239-1285.
- Krueger Alan. 2003. Economic Considerations and Class Size. *Economic Journal* 113 (485): F34-F63.
- Lee M. 2005. *Micro-Econometrics for Policy, Program and Treatment Effects*, Series: Advanced Texts in Econometrics, Oxford University Press
- Levitas Anthony, Golinowska Stanisława, Herczyński Jan. 2001. *Improving Rural Education in Poland*. CASE - Center for Economic and Social Studies, Warsaw.
- Śleszyński P. 2002. Ekonomiczne uwarunkowania wyników sprawdzianu szóstoklasistów i egzaminu gimnazjalnego przeprowadzonego wiosną 2002 roku. Raport dla Ministerstwa Edukacji Narodowej.
- Todd P., Wolpin K. 2003. On the Specification and Estimation of the Production Function of Cognitive Achievement. *Economic Journal* 113 (485).
- Urquiola M. 2000. Identifying class size effects in developing countries: Evidence from rural schools in Bolivia. Development Research Group The World Bank.
- Wößmann Ludger. 2003. European Education Production Functions: What Makes a Difference for Student Achievement in Europe? *European Economy, Economic Papers* No. 190, European Commission, Brussels.
- Wößmann Ludger, West Martin. 2002. *Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS*. Harvard University, Program on Education Policy and Governance Research Paper, PEPG/02-02.