THE ECONOMIC JOURNAL

The Economic Journal, **119** (*October*), 1243–1269. doi: 10.1111/j.1468-0297.2009.02271.x. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

DOES IMMIGRATION AFFECT THE LONG-TERM EDUCATIONAL OUTCOMES OF NATIVES? QUASI-EXPERIMENTAL EVIDENCE*

Eric D. Gould, Victor Lavy and M. Daniele Paserman

This article uses the mass migration wave to Israel in the 1990s to examine the impact of immigrant concentration during elementary school on the long-term academic outcomes of native students in high school. The results suggest that the overall presence of immigrants in a grade had an adverse effect on the chances of passing the high school matriculation exam, which is necessary to attend college. This result is robust to a variety of alternative specifications. We also perform a 'placebo' analysis which shows that the high school outcomes of natives are affected only by the immigrant concentration in their own 5th grade class.

An increasingly important issue faced by Western countries at the turn of the 21st century is the social tension created by growing numbers of immigrants from less developed countries. The effect of immigration on the local labour market has received considerable attention in the literature, but little is known about the impact of immigration on the school system. This article examines the impact of immigrant concentration during elementary school on the long-term academic outcomes of native students in high school. The focus on the long-term outcomes of natives, seven years after their exposure to the 'treatment', is one of the distinguishing features of our work.

Our analysis exploits a unique opportunity to identify the effect of immigrant children in the classroom using the mass migration wave from the former Soviet Union (FSU) to Israel in the early 1990s. The sheer size of the immigration wave, nearly 20% of the total population within a decade, produced large variation in the absorption level of recent immigrants across schools throughout the country. The raw data show that native students who attended a school with a high concentration of immigrants in 5th grade typically had lower achievements in high school. This correlation is not surprising, given that immigrants tended to locate in poorer areas where housing costs were lower. Therefore, identification of the causal effect of immigrants on their native peers must account for the endogenous placement of immigrants into areas that are more likely to be populated with lower-achieving native students, regardless of the local level of immigrant concentration.

* We appreciate helpful comments by Joshua Angrist, Esther Duflo, Caroline Hoxby, Chris Taber, Steve Pischke, three anonymous referees and seminar participants at the Hebrew University, Bocconi University, the Bank of Israel, Ben-Gurion University, LSE, the NBER Education Meetings, the University of Helsinki, the Institute for International Economic Studies in Stockholm, Northwestern University, the University of Chicago Graduate School of Business, UC Berkeley, Stanford University, the conference on 'Inequality, Minorities and Social Exclusion' at Bar-Ilan University, the CEPR conference on 'Inequality in Labour Markets', and the European Summer Symposium in Labour Economics. Vadim Krimsky and Tamar Roth provided expert research assistance. All errors are our own.

To control for this possibility, our empirical strategy exploits random variation in the number of immigrants across grades within the same school. For example, conditional on the total number of immigrants in a given school in grades 4 to 6, the actual number of immigrants in grade 5 can be considered as being determined solely by random factors such as variation in the year of birth among the pool of immigrant children in the school district. This natural experiment is best thought of as the random drawing of balls from an urn in which one third of the balls are marked with the number '4', one third with the number '5', and one third with the number '6'. Conditional on the total number of balls drawn, the number of '5' balls is as good as randomly assigned. Conditioning on the total number of immigrants in the advanced grades in elementary school removes much of the heterogeneity across schools that may confound our estimates: schools that absorbed the same number of immigrants are likely to be similar among themselves, not only in their propensity to absorb immigrants but also in other characteristics.

The analysis uses administrative panel data on school enrollment and test scores for each 5th grade child from the 1993-4 school year until the 2000-1 school year, the year this cohort was scheduled to graduate from high school. The fraction of immigrants in 5th grade is strongly negatively associated with dropout rates and high school matriculation rates. However, when we control for the total number of immigrants in grades 4 to 6 and on the total number of children in grade 5, we find that 5th grade immigrant concentration has a marginally significant negative effect on matriculation rates and a small and statistically insignificant effect on dropout rates. Specifically, an increase of 10 percentage points in the immigrant concentration in the 5th grade raises a native student's dropout rate by (a statistically insignificant) 0.3–0.4 percentage points, relative to an average of 5.4% in our sample; it lowers a native student's matriculation rate by 1.5–1.8 percentage points, relative to an average of about 61% in our sample. This is roughly equivalent to a third of the total gap in matriculation rates between native and immigrant students. For comparison purposes, the magnitude of the effect is similar to the estimated effect of reducing a father's education by three quarters of a year, or changing the school's socioeconomic index by two-thirds of a standard deviation.

A set of balancing tests indicate that the fraction of immigrants in 5th grade is not systematically related to school and student background characteristics, once we control for the number of immigrants in grades 4 to 6 and the total number of children in grade 5. Hence, we infer that the observed relationship between 5th grade immigrant concentration and high school outcomes is not spurious.

To check further whether our results are picking up unobserved characteristics of the school, we perform a set of 'placebo' regressions where we regress the high school outcomes of 5th graders in 1994 on the immigrant concentration in the first and second grades in the same school in 1994. The results from this analysis show statistically insignificant effects for the placebo's treatment effect. That is, the high school achievements of children who were in 5th grade in 1994 are affected only by the fraction of immigrants in 5th grade but not by the fraction of immigrants in first and second grades. This 'placebo' analysis strengthens the interpretation that our estimates are indeed picking up the effect of immigrant concentration in the student's grade, and not the effect of other unobserved school characteristics.

Furthermore, we find that the negative effect of immigrants on native outcomes is larger for native students from a more disadvantaged socio-economic background, and that the effects are generally non-linear – they are stronger at lower levels of immigrant concentration (going from 0 to 10% of the class) than at higher levels of concentration (going from 10% to 20% of the class). These results indicate that the negative impact of immigrants is mitigated when their concentration is high, suggesting that the integration of immigrants is easier when they are present in sufficiently large numbers.

In addition, we address the possible endogenous placement of immigrant students across grades with an IV analysis. In particular, we note that a high fraction of immigrants are enrolled in a grade that is different from what would have been expected given their dates of birth. This raises the concern that administrators or parents may endogenously place immigrants in a given grade, depending on the quality of the existing native students. We therefore instrument for the actual fraction of immigrants in 5th grade with the *predicted* fraction of immigrants. The predicted fraction of immigrants in the 5th grade is based entirely on exogenous information (the date of birth of each student), and therefore it is not subject to any endogenous placement decisions across grades by administrators or parents. Our IV results point to a stronger adverse effect of immigrant concentration on native outcomes, but the estimates are not significantly different from the OLS coefficients.

The rest of the article is structured as follows: the next Section reviews the existing literature that is most closely related to our work. Section 2 presents a brief background of the immigration episode we are exploiting and of the institutional setting of Israeli schools, and describes our data. Section 3 describes our empirical framework. Section 4 presents our main results, while Section 5 looks at additional specifications which examine non-linear effects, differential effects across socio-economic groups, and the placebo analysis. Section 6 presents our IV analysis. Section 7 examines the effect of elementary school immigrant concentration on intermediate outcomes, and Section 8 concludes.

1. Literature Review

While there is a vast literature on the effects of immigration on native labour market outcomes, the question of whether immigration affects natives' educational outcomes has received relatively little attention. Exceptions include Betts (1998), who examines whether immigration reduces the contemporaneous high school graduation rate of natives, and Hoxby (1998) and Borjas (2007), who look at whether immigrants crowd-out natives from slots in college and graduate programmes. Also, Betts and Fairlie (2003) investigate whether immigration in California induced native flight from public to private schools. These papers tend to find a small crowding-out effect of immigrants on natives' educational outcomes, although the effect is sometimes limited to specific ethnic groups.

Other studies that have looked directly at peer effects in an educational setting include Boozer and Cacciola (2001), Hanushek *et al.* (2003), Lefgren (2004) and Burke and Sass (2008). Sacerdote (2001) and Zimmermann (2003) study peer effects in a university setting and find that academic and social outcomes of college students are affected by the quality of randomly assigned roommates. In terms of identification

strategy, the papers closest to our article are those of Hoxby (2000*a*) and Ammermueller and Pischke (2006). Hoxby exploits idiosyncratic variation in gender and race composition of adjacent cohorts in Texas public schools and, in a wide range of specifications, she finds that children's elementary school test scores are affected by those of their peers, with intra-race peer effects appearing to be particularly strong. Ammermueller and Pischke instead use variation in the peer composition across different classrooms of the same grade in six European countries, and find sizeable estimates of peer effects. Hanushek *et al.* (2003) examine a similar type of peer effect but use a different empirical strategy. In particular, they control for individual fixed-effects and exploit variation in one's peers which stems from individuals moving between classes or schools. This strategy has the advantage of completely controlling for fixed factors such as unobserved ability and family characteristics but could be hindered if individuals move for reasons correlated with unobserved shocks which also affect school outcomes.

Closely related to our article is the literature on desegregation which examines the effect of minority students on the achievements of the other students. Recent papers have adopted three main approaches to address the endogenous sorting and selection issues in this context: using experimental or quasi-experimental variation in exposure (Angrist and Lang, 2004; Guryan, 2004); aggregating to a level at which sorting is arguably reduced or eliminated (Evans *et al.*, 1992; Cutler and Glaeser, 1997; Card and Rothstein, 2006); and using within-school variation in minority exposure. This latter approach includes Hoxby (2000*a*) and Hanushek *et al.* (2002), who exploit variation in the racial and ethnic composition of each cohort within the same school. Similar to our research design, these studies assume that while students may sort across schools on the basis of long-run factors like the average racial composition, choices are unaffected by cohort-specific variation in these factors. Both papers find very large exposure effects.[1]

There are similarities and differences between the desegregation literature and our study. Both look at the effect of disadvantaged students on the outcomes of other students, mostly from a more advantaged background. The immigrants in our study actually tend to have relatively well-educated parents but, as we show later, they exhibit other signs of economic distress (high unemployment, high incidence of single-parenthood and low rates of home ownership). The language barrier represents an important difference between our study and the desegregation literature.

In contrast to all of these papers, our focus is on the effect of immigrants on natives and we examine long-run effects rather than the contemporaneous effects of peers. Our findings, in combination with recent findings regarding the benefits of placing

[1] Angrist and Lang study the effects of Metco, a long-running desegregation programme that sends minority children out of the Boston public school district to a large and affluent suburban district. Overall, they find little evidence that the test scores of non-Metco students are affected by the presence of disadvantaged peers, although there seems to be some evidence of a negative effect on the test scores of non-Metco minority students. Guryan conducts a quasi-experimental analysis of the effect of school segregation on black dropout rates, using variation in the scope and timing of major court ordered desegregation plans in the 1970s and 1980s, and finds a modest but statistically significant effect of policies that reduced relative black exposure to black schoolmates. Card and Rothstein find robust evidence that the black-white test score gap is higher in more segregated cities. They report that neighbourhood segregation has a consistently negative impact but that school segregation has no independent effect.

immigrants in better schooling environments, allow for a fuller assessment of the optimal policy regarding the placement of immigrants across schools.[2]

2. Background and Data

2.1. *Immigration to Israel in the 1990s*

The mass migration from the former Soviet Union to Israel since 1989 can be viewed as a unique opportunity for studying a wide range of economic and social phenomena. More than one million immigrants moved to Israel since 1989, increasing its population by a remarkable 20% within a decade. The migration wave was also highly concentrated, with approximately 330,000 immigrants arriving in 1990 and 1991. The fall of the Berlin wall in November 1989 was the trigger that started the whole process, as a number of factors combined to induce migration. First, the Soviet Union lifted its emigration restrictions. Second, the political instability and the uncertain economic climate in the beginning of the 1990s greatly increased the incentives to migrate. Third, Israel was one of the few viable options, as it imposed no entry restrictions and no lengthy waiting periods. Along with the mass migration from the former Soviet Union, the early 1990s also saw the exodus of 15,000 Ethiopian Jews in the context of Operation Solomon in May 1991 and increasing immigration rates from other countries, mainly the US and Argentina.

Not surprisingly, there are large differences in the family background of immigrants from different source countries and the overall immigrant population is quite different from native Israelis in terms of family background. These differences are displayed in Table 1, which shows the background characteristics of 5th grade students in 1994, by immigrant status and by country of origin, based on the Ministry of Education administrative data that is used throughout the article. We define as immigrants all children who were born outside of Israel and immigrated after January 1st, 1989. The Table reveals that immigrant children from the former Soviet Union generally have parents who are more educated than the parents of native Israeli children (a little more than one additional year of education for each parent), and come from smaller families (1.11 siblings versus 2.41 for native Israelis). In contrast, Ethiopian immigrant children come from much larger families (an average of 4.28 siblings) and have parents with only one or two years of education. The vast majority of immigrants during this period came from the former Soviet Union, so the characteristics of the overall immigrant population are dominated by the characteristics of the Soviet immigrants.

While it appears that native Israeli children were exposed to a new immigrant population with more favourable family background characteristics, it is also true that the new immigrant population displayed several signs of social and economic distress. Table 1 presents characteristics of households with children between the ages of 8 and

[2] Recent years have seen a flurry of research that uses exogenous variation created by true or natural experiments to study the effects of peers and the environment at the neighbourhood level on a variety of educational, health, and economic outcomes (Katz *et al.*, 2001; Edin *et al.*, 2003; Oreopoulos, 2003; Jacob, 2004; Weinberg *et al.*, 2004, Goux and Maurin, 2007, Sanbonmatsu *et al.*, 2006). In the Israeli context, Gould *et al.*, (2004*a*) found that Ethiopians who arrived during Operation Solomon generally benefited from attending better schools.

Table 1

Background Characteristics of Immigrants and Natives

		Immigrants from:			
	All Immigrants	Former Soviet Union	Ethiopia	Other	Native Israelis
(*a*) *Ministry of Education data*					
Father's years of schooling	11.45	13.04	1.80	14.21	12.03
	(5.49)	(3.62)	(3.30)	(4.46)	(3.56)
Mother's years of schooling	11.58	13.25	1.39	13.72	12.06
	(5.31)	(3.26)	(2.90)	(3.70)	(3.41)
Number of siblings	1.54	1.11	4.28	1.98	2.41
	(1.73)	(1.16)	(2.61)	(1.35)	(1.57)
Number of children	7,159	5,477	945	737	54,700
(*b*) *Census 1995 data*					
Father's years of schooling	12.69	13.51	2.57	13.50	11.67
	(4.50)	(3.20)	(4.61)	(4.32)	(3.83)
Mother's years of schooling	12.53	13.45	1.92	13.40	11.59
	(4.52)	(3.09)	(4.20)	(3.95)	(3.98)
Female headed household	0.176	0.179	0.248	0.093	0.085
	(0.381)	(0.383)	(0.432)	(0.290)	(0.278)
Home ownership	0.533	0.561	0.482	0.338	0.693
	(0.499)	(0.496)	(0.500)	(0.473)	(0.461)
Household head unemployed	0.071	0.067	0.096	0.086	0.037
	(0.256)	(0.252)	(0.295)	(0.281)	(0.189)
Household monthly income	3,895.85	4,258.76	751.01	3,497.91	5,899.36
	(3,911.46)	(3,738.91)	(1,184.43)	(5,371.44)	(6,963.88)
Number of households	8,910	7,321	737	852	68,725

Notes. The Ministry of Education sample includes the universe of children enrolled in 5th grade during the 1993-4 school year. The Census data includes all households with at least one child between the ages of 8 and 12 in 1995.

12 drawn from the 1995 Israeli Census.[3] Compared to native Israeli children, Table 1 reveals that immigrant children came from poorer households which tend to rent rather than own their own homes, and where the head of the household is much more likely to be female and/or unemployed. Again, the Ethiopian immigrants display much larger signs of distress compared to the Russians but both groups of immigrants are less well-off compared to the native Israeli population. So, while it is true that the overall immigration wave during this period flooded the country with immigrants who were generally well-educated, the overall picture reveals that these immigrants displayed several signs of socio-economic difficulties, which made them disadvantaged in many respects relative to the native population.[4] Therefore, despite the unique characteristics of this mass migration episode, immigrants to Israel share characteristics that are common to immigrants in all Western societies, making our results relevant beyond the Israeli context.

[3] The 'long questionnaire' of the Israeli Census was administered to 20% of Israeli households.

[4] For a more comprehensive analysis of the economic assimilation of the 1990s immigrants and their impact on the Israeli labour market, see Eckstein and Weiss (2002; 2004) and Cohen-Goldner and Paserman (2004; 2006).

2.2. *The Israeli School System*

The Israeli school system is composed of three stages: elementary school (grades 1–6), middle school (grades 7–9), and high school (grades 10–12). Within the Jewish sector, there are two types of schools: secular and modern orthodox. The only private schools in Israel belong to the ultra-orthodox sector and it is extremely unlikely that a non ultra-orthodox child would attend such a school. Most Russian immigrants came with a weak religious background and, therefore, attended secular schools. The Ethiopian immigrants were mainly absorbed in the modern orthodox schools.

Students in Israel generally attend their local elementary school and are placed in the same class from grades 1 to 6. There is very low mobility between schools during the primary and middle school years and any such transfer requires special permission from the local school authority. In contrast, the high school attended by each student is determined by the preferences of students and schools, which in many cases are free to accept and reject students. Some high schools provide a general academic education while others may be devoted to music and the arts. Others may be vocationally-oriented for trades and occupations including practical engineering, electronics etc. This setup generally produces a large change in the peer environment of each student as he/she begins high school. As a result, a native student's exposure to immigrants from primary school is not necessarily strongly correlated to his/her exposure in high school.

It should be noted that immigrants are enrolled in schools according to the same rules as natives. They are placed in regular classes with native students but most receive extra language training and additional instruction in subjects as needed for a few years after arrival. In some cases, immigrants are taken out of the classroom for a brief period of time during the day to receive individualised instruction. Schools are funded mostly at the national level. School funding depends on a formula that takes into account the number of immigrants and the number of disadvantaged students, so that these extra costs were at least partially funded by the central government.

Zoning laws are quite strictly enforced in Israel, hence mobility between schools can only occur if it is associated with residential mobility. We cannot entirely rule out the possibility of natives moving out in response to a high concentration of immigrants in their school. However, it is also true that many schools with significant levels of Russian immigrants responded by developing special programmes in mathematics and science and these programmes attracted attention from many native students with the same interests.

Given these features of the school system, the fraction of immigrants in grade 5, the key explanatory variable in our regressions, can be interpreted as a proxy for the total exposure to immigrant children throughout the elementary school years. Using data on the date of immigration, we calculate that by the end of 5th grade, native students had been exposed to immigrants for an average of 31.7 school months (slightly more than 3 school years).[5] Interestingly, the average date of arrival of immigrants is nearly completely uncorrelated with the fraction of immigrants in the school. It is *not* the case that in schools with a high proportion of immigrants, native students also had been

[5] Length of exposure for a given immigrant is simply calculated as the total number of school months (i.e., excluding July and August) from the month of immigration to September 1994. The underlying assumption is that an immigrant always attended the same school from the moment they immigrated.

exposed to immigrants for a longer period of time. This reinforces our assumption that the immigrant concentration in 5th grade is simply capturing the total exposure to immigrants in elementary school and is not confounded by other factors.

2.3. *Data*

To examine whether the influx of this immigration wave into elementary schools influenced the high school outcomes of native Israeli children several years later, we link detailed information on each child's school environment in the 5th grade to his or her achievements throughout high school. The data come from administrative records collected by Israel's Ministry of Education on the universe of Israeli elementary schools in the 1993-4 school year, the first year a computerised file is available. The data are based on reports from school authorities to the Ministry of Education at the beginning of the school year. The file contains an individual identifier, a school and class identifier, and detailed demographic information on all immigrant pupils in grades 1 to 6, and all pupils in grades 1, 2 and 5. Note that we do not have any data on native children enrolled in grades 4 and 6 in 1993/4. This prevents us from using a traditional identification strategy based on multiple cohorts and school fixed-effects. In practice, though, we show that our estimation strategy is similar to a fixed-effects estimator.

We are able to exactly identify pupils by their country of origin and date of immigration (month and year), so we know precisely who is a native Israeli student and how many immigrant children were present in the 5th grade class of each elementary school. We focus our attention on the 5th grade cohort in 1993–4 for two reasons. First, this is one cohort for which we have complete data on native Israeli pupils. Second, 5th grade students in the 1993–4 school year had enough time within our sample period (which ends with the 2000–1 school year) to finish high school if they progressed through the system without repetition. The data set also contains the birth date of each child, which is mapped into each child's *predicted* grade level using the enrollment cutoff dates for that cohort.[6] The predicted grade level will be used in our instrumental variable strategy.

We link the elementary school records to individual data on high school enrollment and matriculation exam outcomes in the 1998–9, 1999–2000, and 2000–1 school years. Therefore, we are able to follow each native Israeli student from 5th grade in 1993 all the way through the advanced stages of high school. We study two important high school outcomes: dropping out before completing 12th grade and passing the matriculation exams. The latter outcome is particularly important in Israel since it is required to attend college. Similar high school matriculation exams are found in many countries and in some states in the US. Examples include the French Baccalaureate, the German Certificate of Maturity (Abitur), the Italian Diploma di Maturità, the New York State Regents examinations and the recently instituted Massachusetts Comprehensive Assessment System.

[6] In Israel, children mostly enter first grade in September of the calendar year in which they turn six years old, 'mostly' because the relevant threshold dates are based on the Hebrew calendar. For example, the first grade class of September 2007 was composed of children born between the 1st of Tevet 5761 (December 27th, 2000) and the 30th of Kislev 5762 (December 15th, 2001). For conversion between Hebrew dates and Gregorian dates, see http://www.hebcal.com.

Table 2 presents summary statistics for various measures of the 5th grade learning environment faced by native Israeli students. We classify the learning environment variables into two groups: 'peer characteristics' and 'school characteristics'. The 'peer characteristics' describe the family background of the native Israeli 5th grade students in the school. The 'school characteristics' contain additional aspects measured at the school level: average scores for the 4th and 5th grade students on the 1991 standardised mathematics and verbal tests[7] and a 1991 socio-economic index of students in each elementary school.[8] All of the 'school characteristics' are taken from the 1991 school year, which is two years before the sample of native students used in Table 2 attended 5th grade.

Summary statistics are presented in four columns in Table 2, where the sample used in each column is defined by the total number of immigrants absorbed into the elementary school in grades 4 to 6. For example, the first column presents summary statistics for only the native Israeli 5th graders who attended an elementary school in 1993–4 where the total number of post-1989 immigrants in grades 4 to 6 was between 1 and 10. The next column uses a range of immigrants in the 4th to 6th grades between 11 and 20, and the next two columns use ranges of 21–30 and 31–40 respectively. The last column shows the summary statistics for all schools that had between 1 and 40 immigrants in grades 4 to 6.

A close inspection of the numbers in Table 2 reveals why it is important to break down the sample in this manner. When comparing the means across columns, a clear pattern emerges: schools with larger numbers of total immigrants in grades 4 to 6 are increasingly more disadvantaged in terms of the characteristics of the native Israeli students. For example, as the number of immigrants in the school increases, the parental education of native students falls, the percent of students with origins from Asia and Africa (which is widely considered to be a disadvantaged ethnic group in Israel) increases, the average mathematics and verbal scores of the elementary school fall, and the socio-economic index of the school falls. In addition, immigrants represent only 5.39% of the 5th grade class when the total number of immigrants in 4th to 6th grades is between 1 and 10 but increases to 19.48% when there are 31–40 immigrants in the same three grades. Overall, larger numbers of immigrants in grades 4 to 6 are associated with larger concentrations of immigrants in 5th grade, more disadvantaged native Israeli students in terms of their family background and lower-performing schools in terms of student test scores in previous years. This pattern is not very surprising since new immigrants were not placed or directed to areas where the best schools and most-educated parents are likely to be. Instead, immigrants typically started out in areas with lower-cost housing and, therefore, native Israelis from disadvantaged

[7] In June of 1991, near the end of the school year, all fourth and fifth grade students were given achievement tests designed to measure mathematics and Hebrew reading skills. The scores used here consist of the average score on some of the basic and all of the more advanced questions in the test, so that scores are scaled from 1 to 100. The low average and the high variation of scores on this achievement test generated considerable public controversy in Israel, leading to the abandonment of the national testing programme in 1992. This is the same data used in Angrist and Lavy (1999).

[8] The socioeconomic index is based on a function of the pupils' fathers' education and continent of birth, and of family size. The raw index is recorded as the percentage of students in the school who come from what is defined to be a disadvantaged background. Our socioeconomic status index is obtained by standardising this variable and multiplying it by –1, so that high socioeconomic status schools have a high value of the index.

Table 2

Summary Statistics

	Number of immigrants in grades 4 through 6 in 1995				
	1–10	11–20	21–30	31–40	All (1–40)
Number of schools	271	244	164	124	803
Total number of children in grade 5	12,593	13,175	9,573	7,005	42,346
Average number of children in grade 5	50.49	62.02	69.96	71.44	61.21
Average fraction immigrants in grade 5	0.0539	0.1022	0.1397	0.1948	0.1079
Peer characteristics in (non-immigrants)					
Father's years of schooling	12.62	12.24	11.97	11.91	12.24
	(3.66)	(3.53)	(3.32)	(3.34)	(3.50)
Mother's years of schooling	12.65	12.23	12.04	11.83	12.25
	(3.47)	(3.37)	(3.13)	(3.29)	(3.35)
Number of siblings	2.43	2.45	2.37	2.31	2.40
	(1.69)	(1.54)	(1.52)	(1.42)	(1.57)
Fraction with Parents from Asia-Africa	0.240	0.290	0.317	0.312	0.285
	(0.427)	(0.454)	(0.465)	(0.463)	(0.451)
Class Size	30.08	31.09	32.49	33.04	31.43
	(6.04)	(5.74)	(5.16)	(5.36)	(5.74)
School characteristics from 1991					
Standardised Math Score in 1991	0.347	0.164	0.115	0.102	0.197
	(0.772)	(0.837)	(0.802)	(0.814)	(0.813)
Standardised Verbal Score in 1991	0.410	0.144	0.023	0.011	0.174
	(0.769)	(0.799)	(0.739)	(0.808)	(0.795)
Socioeconomic Index in 1991	0.397	0.194	0.155	0.129	0.235
	(0.749)	(0.817)	(0.811)	(0.814)	(0.803)
Outcome Variables (non-immigrants)					
Fraction dropping out before completing 12th grade	0.047	0.055	0.059	0.058	0.054
	(0.211)	(0.229)	(0.236)	(0.233)	(0.226)
Fraction eligible for high school matriculation	0.642	0.611	0.605	0.584	0.614
	(0.480)	(0.488)	(0.489)	(0.493)	(0.487)
Outcome Variables (immigrants)					
Fraction dropping out before completing 12th grade	0.068	0.069	0.074	0.094	0.078
	(0.252)	(0.254)	(0.262)	(0.292)	(0.268)
Fraction eligible for high school matriculation	0.615	0.586	0.556	0.523	0.561
	(0.487)	(0.493)	(0.497)	(0.500)	(0.496)

Notes. Authors' calculations based on Ministry of Education data

backgrounds and lower-performing schools were more likely to attend schools with a higher concentration of immigrants.

Finally, the lower two panels of Table 2 present the high school outcomes of native Israeli students and immigrants in 5th grade during the 1993–4 school year, according to the number of immigrants in grades 4–6. Not surprisingly, the Table indicates that native students who attended schools with a larger proportion of immigrants in 5th grade also had lower outcomes in high school: they dropped out of high school more often and were less likely to pass the matriculation exams. However, given the correlations highlighted in the top part of the Table, this pattern clearly does not have any causal interpretation.[9]

[9] Immigrants who attended schools with a high number of immigrants also tended to have lower outcomes. Importantly, immigrant students have markedly lower academic outcomes: their dropout rate stands at 7.8% (as opposed to 5.4% among natives), and their matriculation rate is 56.1% (61.4% among natives).

3. Identification Strategy

3.1. *Theoretical Framework*

In order to frame our empirical analysis, we now present a very stylised model of the education production function. For simplicity, assume that educational output at the end of high school (Y) depends on an input in elementary school (X_1) and an input in high school (X_2). The education production function can be written as:

$$Y = f(X_1, X_2) + \varepsilon$$

In our case, the input at the time of elementary school is the fraction of immigrants in 5th grade, which can be thought of as a proxy for the exposure of immigrants throughout elementary school (as discussed above). The total derivative of high school outcomes with respect to fraction of immigrants in 5th grade can be written as:

$$\frac{\mathrm{d}Y}{\mathrm{d}X_1} = \frac{\partial f}{\partial X_1} + \frac{\partial f}{\partial X_2}\frac{\partial X_2}{\partial X_1}.$$

Our identification strategy, which relies on quasi-random variation in the fraction of immigrants in 5th grade, allows us to obtain an unbiased estimate of this total effect. We cannot, however, disentangle the direct effect of immigrant concentration in elementary school from the indirect effect that operates through inputs in high school (even though we later provide some evidence on this). Similarly, even an ideal randomised experiment would be unable to identify separately *endogenous* peer effects from *contextual* effects in the terminology of Manski (1993). Nevertheless, the reduced form parameter that we can estimate is an important causal parameter of interest, as it is one that can be directly manipulated by policy.

3.2. *Empirical Framework*

In order to identify the causal link of the immigrant concentration on the outcomes of natives, our strategy is to exploit the variation in the number of immigrants in 5th grade conditional on the total number of immigrant students in grades 4 to 6. As Table 2 clearly shows, immigrant concentration in 5th grade is not randomly assigned with respect to background characteristics. Schools that absorbed large numbers of immigrants are different in terms of student and school characteristics relative to schools that absorbed fewer immigrants. However, conditional on the total number of immigrants in grades 4 to 6, the fraction of immigrants in grade 5 is determined solely by random variation in the grade distribution among the pool of immigrant children in a school district, and can therefore be thought of as randomly assigned. This is our key identifying assumption, which leads us to estimate the following equation:

$$\begin{aligned}(\text{High School Outcome})_{ij} = {} & \beta(\text{Fraction immigrants in 5th grade})_j \\ & + f\begin{bmatrix}(\text{no. of of immigrants in grades 4 to 6})_j, \\ (\text{no. of of children in grade 5})_j, \\ (\text{Family Background})_{ij}\end{bmatrix} \\ & + u_{ij}\end{aligned}$$

where i denotes individuals, j denotes schools, and $f(\cdot)$ is a flexible function of its arguments.

There are several features worth noting in this equation. First, the key explanatory variable is the fraction of immigrants at the *grade* level in the entire school and not within the student's actual class. It is possible that most of the peer effect is indeed generated at the class level but, since students are probably not placed into classes at random within a grade, using immigrant concentration at the class level may be subject to bias.[10] Using the class level treatment would add precision to the estimates but since we do not have an instrument for student placement across classes, we have chosen to sacrifice some precision in order to obtain a consistent estimate of the peer effect, which is still obtainable even if most of the effect is produced at the class level.

Second, the key explanatory variable is the *fraction* of immigrants in grade 5: the effect of one immigrant child in a class of 15 students may be different from the effect of one immigrant in a class of 30. Therefore, even holding constant the number of immigrants in grades 4 to 6, the *fraction* of immigrants varies across schools also because of variation in school size. Hence, it will be important to control for the total number of children in grade 5 as well as for the total number of immigrants in grades 4 to 6.

Third, given the crucial role played by our conditioning variables, we would like to ensure that our results are not sensitive to the particular parametric form in which these variables enter the regression equation.[11] Ideally, we could obtain a fully non-parametric estimate of the parameter of interest by calculating the effect of immigrant concentration separately for every possible combination of our two key conditioning variables (the number of immigrants in grades 4 to 6 and the total number of children in grade 5) and then averaging over these estimates. In practice, this approach is not feasible because the two key conditioning covariates are essentially continuous. Instead, we adopt a number of different methods to control for these two variables. The simplest approach is simply to include the number of immigrants in grades 4 to 6 and the total number of children in grade 5 linearly in the regression equation. Alternatively, one may also include higher order terms of the two key conditioning variables and interactions between them. However, these relatively naïve approaches may not be effective in entirely removing all the correlation between school characteristics and the fraction of immigrants in grade 5.

Therefore, we adopt a specification that includes elements of the non-parametric approach, but allows us to obtain estimates using a standard linear regression. Specifically, we first divide the sample into eight different intervals defined by the number of immigrants in grades 4 to 6 in a given school: the first interval includes schools with 1 to 5 immigrants in grades 4 to 6, the second interval includes schools with 6 to 10 immigrants in grades 4 to 6 and so on. We then include a separate dummy for each one of these intervals and the interactions of these dummy variables with both the key

[10] One possibility would be to instrument the class-level immigrant concentration with the grade level immigrant concentration. However, it can be shown that if class size is constant within a school, this IV estimate is numerically identical to the OLS estimate with grade-level immigrant concentration on the right hand side. In practice, class size tends to have very little variation within a school, so we see little value added in reporting this estimate.

[11] What we describe is essentially a version of the *control function* approach for estimation in the presence of selection on observables (Heckman and Robb, 1985).

conditioning variables (as well as higher order terms and interactions in what turns out to be our preferred specification) and with all of the family background characteristics. This specification is equivalent to running separate regressions within each interval, but constraining the coefficient on the treatment variable to be the same across intervals.

Formally, our preferred regression equation is represented by:

$$(\text{High School Outcome})_{ij} = \beta(\text{Fraction immigrants in grade 5})_j + \sum_{m=1}^{8} \gamma_{m'}(d_{jm} Z_{ij}) + u_{ij},$$

$$d_{jm} = \begin{cases} 1 & \text{if } 5(m-1)+1 \leqslant IMMIG46_j \leqslant 5m \\ 0 & \text{otherwise} \end{cases},$$

$$Z_{ij} = \begin{pmatrix} 1, IMMIG46_j, NKIDS5_j, IMMIG46_j^2, NKIDS5_j^2, \\ IMMIG46_j \times NKIDS5_j, X_{ij} \end{pmatrix}',$$

where *IMMIG46* represents the number of immigrants in grades 4 to 6, and *NKIDS5* is the total number of children in grade 5. Since the fraction of immigrants in 5th grade does not vary at the individual level, the standard errors are adjusted for clustering at the elementary school level. We refer to this specification as 'OLS with stratification', because of the stratification of the sample on the number of immigrants in grades 4 to 6. We will later show that this specification is indeed effective at removing the correlation between fraction of immigrants in grade 5 and school characteristics.[12]

4. OLS Results

4.1. *Balancing Tests*

To determine whether the data supports our identification strategy, we first examine the relationship between the immigrant concentration in grade 5 and the observable characteristics of native Israeli children in the same class. Each coefficient in Table 3 represents the estimate from a single regression for each of the various measures of the native student's background and elementary school environment on the fraction of immigrants in 5th grade.[13] The first column presents a naïve specification where no additional controls are added to each regression. The coefficients in the first column of Table 3 are large and highly significant and show that a larger immigrant concentration in 5th grade is highly correlated with socio-economic factors (parental education, number of siblings, etc.) which may also be affecting high school outcomes.

The second column of Table 3 shows that controlling for the total number of students in grades 5 and also for the number of immigrants in grades 4 to 6 in the same

[12] We emphasise again that we cannot carry out a conventional estimation with school fixed effects, because we only have one full cohort of native students which we observe from 5th to 12th grade. There is however a strong similarity between our estimation and the traditional within estimator in the regression of $(y_{ijt} - \bar{y}_j)$ on $(x_{ijt} - \bar{x}_j)$. The within point estimate is also the coefficient on x_{ijt} in the regression of y_{ijt} on x_{ijt} and $\bar{x}_j$. Hence, our estimator differs from the within estimator in that (*a*) we control for the number of immigrants in grades 4 to 6 and the total number of children in grade 5 separately, rather than controlling for the mean of the fraction of immigrants across all three grades (which we cannot observe); and (*b*) we allow the control function to be a flexible polynomial, rather than a simple linear function.

[13] The regressions are at the individual student level, with standard errors adjusted for clustering at the school level.

Table 3

Balancing Tests for the Actual Fraction of Immigrants

	(1)	(2)	(3)	(4)	(5)
Dependent variable:					
Mother's years of schooling	−5.257***	−0.581	−0.178	−0.283	0. 071
	(0.675)	(1.084)	(1.135)	(1.070)	(1.136)
Father's years of schooling	−5.402***	−1.253	−0.517	−0.700	−0.255
	(0.691)	(1.095)	(1.182)	(1.118)	(1.183)
Number of Siblings	1.252***	1.100*	0.542	1.288**	0.588
	(0.415)	(0.621)	(0.685)	(0.630)	(0.678)
Fraction Asia-Africa	0.530***	0.144	−0.018	0.052	−0.063
	(0.069)	(0.107)	(0.122)	(0.112)	(0.121)
Standardised maths score in 1991	−3.170***	−1.669***	−0.810	−1.334**	−0.698
	(0.400)	(0.611)	(0.658)	(0.609)	(0.654)
Standardised verbal score in 1991	−3.477***	−1.593***	−0.692	−1.077*	−0.341
	(0.385)	(0.590)	(0.641)	(0.598)	(0.631)
Socioeconomic index in 1991	−3.789***	−2.306***	−0.966	−1.903***	−0.642
	(0.374)	(0.592)	(0.648)	(0.619)	(0.644)
Class size in 1994	−6.570**	0.545	6.343*	−4.744	4.092
	(2.648)	(3.638)	(3.570)	(3.623)	(3.600)
Controls for number of immigrants in grades 4 to 6 and number of children in grade 5	No	Linear	Second order terms and interaction	Linear within each cell	Second order terms and interaction within each cell
Stratification	No	No	No	Yes	Yes
Number of observations	42,346	42,346	42,346	42,346	42,346
Number of schools	803	803	803	803	803

Note. Entries in the Table represent the coefficients from separate regressions of the relevant dependent variable on the fraction immigrants in 5th grade.
In the regressions with stratification, the whole sample was stratified into eight cells based on the total number of immigrants in grades 4 to 6 (1 to 5, 6 to 10, 11 to 15 etc.) and the regression includes cell dummies interacted with a constant and the full set of control variables. See text for details.
Robust standard errors (adjusted for clustering at the school level) are in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

school reduces the size and significance of most of these correlations but does not entirely remove it. When second order terms and interactions for these controls are entered into the specification (column 3), the correlations are reduced virtually to zero. The last two columns of Table 3 include dummy variables for each interval of students in grades 4 to 6 interacted with the key conditioning variables (just linearly in column 4, and with higher order terms in column 5). Remarkably, these columns show that the correlation between immigrant concentration in 5th grade and the student's characteristics is virtually wiped out if we control for the total number of immigrants in grades 4 to 6 in a very flexible, non-linear manner.[14]

Therefore, the results in Table 3 justify the identification strategy outlined above. Once the analysis is flexibly conditioned on a tight interval of total immigrant students in grades 4 to 6, the number of immigrant students that happen to be in grade five can

[14] Appendix Table 2 (available online as supporting information) shows the balancing tests within each interval. Out of the 64 coefficients, only six are significant at the 10% significance level or lower. The other 58 coefficients point to no systematic relationship between observable characteristics and the fraction of immigrants in 5th grade. Furthermore, the six coefficients which do turn out to be statistically significant do not tell a consistent story about whether students are positively or negatively selected into schools with larger fractions of immigrants in the 5th grade.

be considered random in the sense of being uncorrelated with the other characteristics of the native 5th grade students. That is, the 'treatment' is balanced after we condition the sample in this manner. This type of balancing test does not necessarily provide a proof of random assignment, as the assumption requires there to be no correlation between the fraction of immigrants in 5th grade and both observable and *unobservable* background characteristics. However, the lack of a significant relationship between the fraction of immigrants in 5th grade and observable characteristics suggests that unobservable characteristics are also unlikely to be correlated with the fraction of immigrants in 5th grade. Overall, the stark contrast between the 'naïve' coefficients in the first column of Table 3 versus those in column 5 illustrates the extent to which our identification strategy can reduce any potential bias stemming from the non-random selection of immigrants across schools.[15]

4.2. *Main OLS results*

Our main results are presented in Table 4. The first column simply regresses each outcome variable on the fraction of immigrant students in each native child's 5th grade class, without any additional controls. For both outcomes, the coefficient on the proportion of immigrants in 5th grade is large and highly significant. A 10 percentage point increase in the immigrant share is associated with a 1.3 percentage point increase in the dropout rate and a 5.7 percentage point decrease in the matriculation rate. The coefficients are reduced dramatically in size and significance in the second column, which controls for the individual and school characteristics described in Table 1.

Adding a linear control for the number of immigrants in the school from grades 4 to 6 in column (3) has little effect on the size of the coefficient on the high school matriculation rate but it lowers somewhat the coefficient on the dropout rate, which becomes insignificant. The coefficient on the matriculation rate is essentially unaffected when we include the square terms of the number of immigrants in grades 4 to 6 and the total number of children in grade 5 and their interaction (column 4), while the coefficient on the dropout rate becomes even smaller.

The last two columns in Table 4 employ our preferred identification strategy (which we call the 'stratified' model) by controlling for the narrowly defined intervals of immigrant students in grades 4 to 6. The specifications in these two columns differ only in that column 5 controls linearly for the key conditioning variables within each interval, while column 6 also includes squared terms and interactions. The latter specification, which is the most flexible, will be used throughout the rest of the article.[16] The results in these two columns tell a very similar story, and point in the direction of an adverse effect of immigrants on the rate of matriculation but no effect on dropout rates. The point estimates imply that an increase of 10 percentage points in

[15] The balancing tests also address the concern that our results may be driven by endogenous mobility of native students (and in particular high socio-economic status ones) in response to a large concentration of immigrants in their grade. If this were the case, we would have observed a negative correlation between immigrant concentration and native characteristics.

[16] An alternative approach would be to estimate a separate treatment effect within each interval (1–5 immigrants in grades 4 to 6, 6 to immigrants etc.) and then average across intervals. The results of this approach are presented in Appendix Table 1 (available online as supporting information).

Table 4

OLS Estimates of the Effect of Immigration on Natives' Long-term Educational Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
High School Outcome:						
Dropped out before completing 12th grade	0.133*** [0.022]	0.037* [0.022]	0.029 [0.031]	0.011 [0.034]	0.040 [0.032]	0.027 [0.035]
Passed HS matriculation exam	−0.566*** [0.065]	−0.134*** [0.050]	−0.145** [0.071]	−0.155** [0.078]	−0.177*** [0.072]	−0.148* [0.079]
Controls for individual and school characteristics	No	Yes	Yes	Yes	Yes	Yes
Controls for number of immigrants in grades 4 to 6 and number of children in grade 5	No	No	Linear	Second order terms and interaction	Linear within each cell	Second order terms and interaction within each cell
Stratification on number of immigrants in grades 4 to 6.	No	No	No	No	Yes	Yes
Number of observations	42,346	42,346	42,346	42,346	42,346	42,346
Number of schools	803	803	803	803	803	803

Note. Entries in the Table represent the coefficients from separate regressions of the relevant dependent variable on the fraction immigrants in 5th grade. Individual controls: mother's years of schooling, father's years of schooling, dummy indicators for whether parents' schooling is missing, number of siblings, a dummy for whether the number of siblings is missing, age in 1994, a gender dummy. School controls: a dummy for whether the elementary school is religious, class size in 1994, mathematics scores in 1991, verbal scores in 1991 and school socioeconomic index in 1991. In the regressions with stratification, the whole sample was stratified into eight cells based on the total number of immigrants in grades 4 to 6 (1 to 5, 6 to 10, 11 to 15 etc.) and the regression includes cell dummies interacted with a constant and the full set of control variables. See text for details. Robust standard errors (adjusted for clustering at the school level) are in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

the immigrant concentration in the 5th grade raises a native student's dropout rate by (a statistically insignificant) 0.3-0.4 percentage points, relative to an average of 5.4% in our sample; and lowers a native student's matriculation rate by 1.5-1.8 percentage points, relative to an average of about 61% in our sample.[17] This is roughly equivalent to a third of the total gap in matriculation rates between native and immigrant students (5.3 percentage points, see Table 2). For comparison purposes, one additional year of father's education is associated with a 1.9 percentage points increase in matriculation rates; a one standard deviation change in the elementary school's socio-economic index is associated with a 2.4 percentage points increase in matriculation rates; and a one standard deviation change in 1991 verbal test scores is associated with a 3.9 percentage points increase in matriculation rates.[18]

Overall, comparing the results across the columns of Table 4 reveals that most of the observed correlation between immigrant exposure in 5th grade and future outcomes in

[17] The standard deviation of the fraction of immigrants in grade 5 in our sample is 0.0903. We often report coefficient magnitudes in terms of a 10 point increase in the fraction of immigrants, which we think is more intuitive. To convert this into the effect of a one-standard deviation increase in the fraction of immigrants, simply multiply by 0.9.

[18] We note that schools may mitigate the adverse impact of immigration by reallocating resources towards weaker students. However, from a policy perspective, the parameter of interest is the overall effect of immigrant concentration, since it is unlikely that any school would not react at all to an influx of immigrants.

high school is due to unobserved characteristics of the student and the school. This pattern shows that the strong correlation between native outcomes and immigrant concentration is mostly spurious, and highlights the value-added of our strategy for identifying the causal effect. Moreover, the similarity of the results between columns 3 and 6 suggests that controlling for the individual and school characteristics appears to be sufficient to identify the causal effect.

5. Robustness Checks

5.1. *Placebo Regressions*

One may still be concerned that the estimates from our OLS analysis may be picking up a spurious correlation between unobserved school characteristics and the fraction of immigrants. Therefore, to test our identification strategy, we run a series of 'placebo' regressions: instead of using the actual fraction of immigrants in grade 5, we use as the treatment variable the fraction of immigrants in grades 1 and 2 in the 1993/4 school year. If the results in Table 4 were due to unobserved school characteristics that are correlated with the immigrant concentration at the school level, even after our flexible conditioning strategy, we would expect that the long-term outcomes of fifth graders would also be affected by the fraction of immigrants in the earlier grades. This scenario is clearly not supported by the placebo regressions in Table 5, which compares our previous results with those obtained using the immigrant concentration in grades 1 and 2 as a placebo treatment. The results are presented for the simplest non-stratified specification (analogous to column 3 in Table 4) and for the most flexible stratified specification (analogous to column 6 in Table 4). Both specifications control for the total number of immigrants in grades 4 to 6 and the total number of students in grade 5. In both cases, the coefficient on the placebo is always small and insignificant. In other words, the high school achievements of children who were in 5th grade in 1994 are affected only by the fraction of immigrants in 5th grade and are uncorrelated with the fraction of immigrants in first and second grades. In addition, although the coefficients on the dropout rate in Table 4 were not significant, they are still much larger in magnitude than the placebo estimates. Overall, the stark contrast between our main analysis and placebo results provides strong evidence that our results are not spuriously driven by unobserved school characteristics.

5.2. *Non-Linear Effects*

This Section examines whether the treatment effect of the immigrant concentration is linear or non-linear. Table 6 presents results from the Stratified-OLS specification (analogous to column 6 in Table 4) after adding the square of the immigrant concentration as an additional treatment variable. Similar to Table 4, there is no significant linear or non-linear effect of immigrants on the dropout rate. However, the effect of immigrants on the matriculation rate of natives appears to be non-linear – the coefficient on the linear term is negative while the coefficient on the quadratic term is positive. Furthermore, the non-linearity of the effect is quite severe – increasing the immigrant concentration from 0 to 10% lowers a native student's chances of passing

Table 5

Placebo Regressions

	No Stratification Controls for number of immigrants in grades 4 to 6 and number of children in grade 5: linear			Stratification Controls for number of immigrants in grades 4 to 6 and number of children in grade 5: second order polynomial and interaction within each cell		
	Treatment variable:			Treatment variable:		
	Actual treatment: fraction immigrants in 5th grade	'Placebo': fraction immigrants in 1st grade	Placebo: fraction immigrants in 2nd grade	Actual treatment: fraction immigrants in 5th grade	'Placebo': fraction immigrants in 1st grade	Placebo: fraction immigrants in 2nd grade
Dropped out before completing 12th grade	0.029	−0.001	−0.003	0.027	−0.001	−0.004
	(0.031)	(0.005)	(0.005)	(0.035)	(0.005)	(0.005)
Passed HS matriculation exam	−0.145**	0.001	0.013	−0.148*	0.004	0.012
	(0.071)	(0.014)	(0.014)	(0.079)	(0.014)	(0.013)
Number of observations	42,346	40,535	41,171	42,346	40,535	41,171
Number of schools	803	771	781	803	771	781

Notes. Entries in the Table represent the coefficients on the treatment variable. The treatment variable is the fraction of immigrants in either 5th, 1st, or 2nd grade, depending on the column. Individual controls: mother's years of schooling, father's years of schooling, dummy indicators for whether parents' schooling is missing, number of siblings, a dummy for whether the number of siblings is missing, age in 1994, a gender dummy. School controls: a dummy for whether the elementary school is religious, class size in 1994, mathematics scores in 1991, verbal scores in 1991 and school socioeconomic index in 1991.
In the regressions with stratification, the whole sample was stratified into eight cells based on the total number of immigrants in grades 4 to 6 (1 to 5, 6 to 10, 11 to 15 etc.) and the regression includes cell dummies interacted with a constant and the full set of control variables. See text for details.
Robust standard errors (adjusted for clustering at the school level) are in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

the matriculation rate by 4.2 percentage points, while going from a concentration of 10% immigrants to 20% reduces the matriculation rate by a further 1.9 percentage points. Thus, the results indicate that the adverse effect of immigrants on the matriculation rate of natives is higher at lower levels of immigrant concentration.

There are several possible explanations for this pattern of results. One possibility is that there are returns to scale in the ability of schools to absorb immigrants and to adapt to their needs. For example, if the immigrant concentration is sufficiently large, it may be cost-effective to hire additional language teachers and set up extra classes specifically for immigrants. The immigrants themselves may find it easier to integrate and may require fewer resources if there is a sufficient number of peers with a similar background. Or, with a higher concentration of immigrants, there may be scope to redirect resources towards those classes with greater need (e.g., allocating the better teachers to classes or grades with a high fraction of immigrants). Finally, this pattern could result from heterogeneity in the types of immigrants and natives found in schools with different immigrant concentrations.[19]

[19] In Appendix Table 3 (available online as supporting information), we explore the robustness of our results to alternative specifications: using a finer stratification or a two-dimensional stratification and using the *number* of immigrants in grade 5 as the explanatory variable, rather than the fraction of immigrants. Overall, the results are generally robust in size and significance across these alternative specifications.

Table 6

Nonlinear Effects

		No Stratification Controls for number of immigrants in grades 4 to 6 and number of children in grade 5: linear	Stratification Controls for number of immigrants in grades 4 to 6 and number of children in grade 5: second order polynomial and interaction within each cell
Dropped out before completing 12th grade	Fraction immigrants in 5th grade	0.064	0.100
		(0.061)	(0.069)
	Fraction immigrants in 5th grade squared (÷ 100)	−0.101	−0.219
		(0.140)	(0.168)
	Change in outcome when going from 0% to 10% immigrants	0.0054	0.0078
		(0.0049)	(0.0055)
	Change in outcome when going from 10% to 20% immigrants	0.0034	0.0034
		(0.0032)	(0.0036)
Passed HS matriculation exam	Fraction immigrants in 5th grade	−0.387***	−0.532***
		(0.145)	(0.161)
	Fraction immigrants in 5th grade squared (÷ 100)	0.703**	1.145***
		(0.311)	(0.375)
	Change in outcome when going from 0% to 10% immigrants	−0.0318***	−0.0418***
		(0.0118)	(0.0128)
	Change in outcome when going from 10% to 20% immigrants	−0.0177**	−0.0188**
		(0.0075)	(0.0079)
No. of obs.		42,346	42,346
No. of schools		803	803

Notes. Entries in the first two rows represent the coefficients on the fraction of immigrants in 5th grade and its square from a single regression model. For the list of controls and the definition of 'stratification,' see Table 4. Robust standard errors (adjusted for clustering at the school level) are in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

5.3. *Heterogeneous Effects*

Up to now, we have only attempted to estimate the effect of immigrants on the overall population of natives. However, the effect of immigrants on the outcomes of natives may vary with the background characteristics of the native student. We now examine the effect of immigrants on disadvantaged native students who are possibly more likely to be affected by the presence of immigrants in the classroom.[20] We use two measures to identify disadvantaged native students. The first is based on parental schooling, while the second is based on ethnic background. Native students are defined as coming from a low socio-economic status if their parents have fewer than 10 years of schooling, or come from an Africa-Asian ethnicity. The regressions are run separately for 'low' and 'high' status natives in Table 7.

The results in Table 7 once again show that there is no effect of immigrants on dropout rates and this is now shown to be true for native students from either a high or low socioeconomic background. However, the significant results found in Table 4 for the matriculation rate are now shown to be exclusively due to the strong adverse effect of immigrants on natives from a low socio-economic background. The coefficients for

[20] A number of papers have shown that the outcomes of disadvantaged students are more responsive to changes in school resources (Angrist and Lavy, 1999; Krueger, 1999; Hoxby, 2000*b*; Jepsen and Rivkin, 2009).

Table 7

Heterogeneous Effects – High and Low Socioeconomic Status Students

		No Stratification Controls for number of immigrants in grades 4 to 6 and number of children in grade 5: linear			Stratification Controls for number of immigrants in grades 4 to 6 and number of children in grade 5: second order polynomial and interaction within each cell		
		Ethnicity	Father's education	Mother's education	Ethnicity	Father's education	Mother's education
(*a*) *Low Status*	Dropped out before	0.066	0.023	0.010	0.069	0.064	0.023
(Ethnicity Asia-Africa,	completing 12th grade	(0.049)	(0.075)	(0.067)	(0.056)	(0.084)	(0.076)
parent's education < 10)	Passed HS	−0.326***	−0.314**	−0.274**	0.364***	−0.323**	−0.272*
	matriculation exam	(0.100)	(0.134)	(0.132)	(0.119)	(0.149)	(0.154)
	Number of observations	12,062	8,269	7,873	12,062	8,269	7,873
(*b*) *High Status*	Dropped out before	0.007	0.027	0.034	0.009	0.012	0.030
(Ethnicity not Asia-Africa,	completing 12th grade	(0.036)	(0.030)	(0.033)	(0.039)	(0.033)	(0.037)
parent's education ⩾ 10)	Passed HS	−0.044	−0.115	−0.117	−0.027	−0.092	−0.116
	matriculation exam	(0.081)	(0.076)	(0.076)	(0.090)	(0.083)	(0.083)
	Number of observations	30,284	34,077	34,473	30,284	34,077	34,473

Note. Each cell in the Table presents the estimated effect of the fraction of immigrants in 5th grade on the relevant dependent variable from separate models. The top panel presents results for low socioeconomic status students, the bottom panel presents results for high socioeconomic status students. For the list of controls and the definition of 'stratification', see Table 4. Robust standard errors (adjusted for clustering at the school level) are in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

disadvantaged students are roughly double the size of the coefficients found in Table 4 for the full sample. Therefore, the estimated effect for the general population is rather muted by mixing a large adverse effect on disadvantaged students with no effect on the rest of the native population.

6. IV Results

Our analysis so far is based on the assumption that the endogenous placement of immigrants in the 5th grade across schools is eliminated once we control for the total number of immigrants in the school from grade 4 to 6. This assumption is supported by the balancing tests in Table 3. However, given that a school has a particular number of immigrants in grades 4 to 6, it is possible that the placement of immigrants across grades within the same school may not be entirely random. This concern is based on the high rate of holding immigrant students back a grade. For example, among the immigrant children who should have attended 5th grade based on their date of birth and the enrollment rules in place for that cohort, 55% were in fact enrolled in grade 5, while 38% were in grade 4. If a school tends to hold immigrant students back so that more of them are placed in the grade with better or worse native students compared to the adjacent grade, then our previous results could be biased.[21] Measurement error in the grade enrollment information of immigrants could also bias the results in the previous Section towards zero.[22]

To account for this possibility, we adapt our identification strategy by predicting how many immigrants are supposed to be in grade 5 within a school based on their date of birth and use the *predicted* fraction of immigrants in grade 5 as an instrument for the actual immigrant concentration in grade 5 within a school.[23]

Since the predicted fraction of immigrants in the 5th grade is based entirely on exogenous information (the date of birth of each student), it should not be subject to any endogenous placement decisions across grades by administrators or parents. The key identifying assumption for the IV strategy is that, conditional on the actual number of immigrants in grades 4 to 6 and the total number of children in grade 5, the predicted fraction of immigrants does not affect outcomes. This assumption could be violated if a school with a higher predicted number of immigrants in 5th grade, conditioning on the actual number of immigrants in 5th grade, pursues different policies than a school with a

[21] The decision on which grade to put an immigrant child in would typically be made by the school counsellor or psychologist, subject to the parents' consent.

[22] Even with the best administrative data, some students' grades may be recorded incorrectly, or there may be movements across grades in the middle of the school year. If measurement error is of the classical type, this would lead to attenuation bias. Moreover, attenuation bias could be exacerbated by controlling for variables such as the number of immigrants in grades 4 to 6, since the standard deviation in this latter variable is probably larger than the standard deviation in the number of grade 5 immigrants; hence, much of the true signal in the explanatory variable would be swept away by the control variable, leaving mostly noise.

[23] Since we do not have data on native children who attend grades 4 and 6, the *predicted* fraction of immigrants is calculated as:

$$z = 100 \times \frac{\text{Predicted number of immigrants in grade 5}}{\text{Predicted no. of immigrants in grade 5} + \text{Actual no. of natives in grade 5}}.$$

Similarly, the *predicted* number of children in grade 5 is calculated as the sum of the *predicted* number of immigrants plus the *actual* number of natives in grade 5.

lower predicted number of immigrants, such as diverting more resources to the 4th grade. The second key condition is that the instrument (predicted fraction of immigrants) is correlated with the endogenous regressor (the actual fraction of immigrants), and this should be satisfied almost trivially (this is confirmed by the results of the first stage regressions). Appendix Table 4 (available online as supporting information) presents balancing tests for our instrument, and shows that the predicted fraction of immigrants is uncorrelated with the personal and school characteristics of natives, after conditioning on the same control function used in the OLS analysis. Since we have one instrument for one endogenous variable, the regressions are exactly identified.

Table 8 reports the IV results in a manner similar to Table 4 for the OLS analysis. We focus first on the third row in the Table, which presents the key statistics from the first stage regressions. Not surprisingly, there is a very strong first stage relationship between the predicted and the actual fraction of immigrants in grade 5, and there is no reason to fear a weak instrument problem. Shea's R^2 shows that our instrument explains about 70% of the variation in the actual immigrant concentration with limited controls (the first two columns), and around 30% of the variation with the extended set of controls in the remaining columns. The F-statistic from the first stage is always highly significant (always above 180). Therefore, the predicted immigrant concentration is a strong instrument for the actual immigrant concentration.

Going from the first to the last column of Table 8, the results show once again that there is a high correlation between the immigrant concentration in 5th grade and high school outcomes but most of this correlation is explained by controlling for heterogeneity across schools. In general, the IV results are larger and more significant than the OLS coefficients in Table 4. This pattern is true for both high school outcomes. For example, the last column indicates that a 10 percentage point increase in the fraction of immigrants in 5th grade raises the dropout rate of native students by 0.8 percentage points and lowers the individual matriculation rate by 2.8 percentage points.

These findings may indicate that the OLS estimates are biased towards zero because of the endogenous holding-back of immigrants within a school. However, the results might be stronger with IV due to measurement error in the treatment variable. In any case, the IV coefficients are generally not significantly different from the OLS coefficients (they are very similar to the magnitudes obtained for natives from a lower socioeconomic background in Table 7). Therefore, the overall conclusion from the IV analysis is that our main OLS results are robust to the potential endogenous placement of children across schools and measurement error.

7. Intermediate Outcomes

We now try to shed light on the mechanism that may drive the relationship between the 5th grade immigrant concentration and high school outcomes. It is possible that the acquisition of basic learning skills is hindered by the presence of a high concentration of immigrants in elementary school; on the other hand, it is also possible that exposure to immigrants in elementary school increases the exposure to immigrants in later grades – which may be driving the estimated effect on high school outcomes.

To address these issues, we estimate the effect of immigrant concentration in elementary school on two intermediate outcomes: the quality of the high school attended

Table 8

IV Estimates of the Effect of Immigration on Natives' Long-term Educational Outcomes Instrument: Predicted Fraction Immigrants in Grade 5

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome variable:						
Dropped out before completing 12th grade	0.167*** (0.025)	0.053** (0.024)	0.071 (0.047)	0.049 (0.057)	0.108** (0.055)	0.082 (0.061)
Passed HS matriculation exam	−0.673*** (0.076)	−0.141** (0.059)	−0.178 (0.118)	−0.220 (0.151)	−0.302** (0.130)	−0.279* (0.155)
[Shea Partial R^2, First stage F]	[0.73, 1517.7]	[0.70, 1377.1]	[0.39, 286.1]	[0.29, 181.5]	[0.35, 241.0]	[0.29, 181.3]
Controls for individual and school characteristics	No	Yes	Yes	Yes	Yes	Yes
Controls for number of immigrants in grades 4 to 6 and number of children in grade 5	No	No	Linear	Second order terms and interaction	Linear within each cell	Second order terms and interaction within each cell
Stratification on number of immigrants in grades 4 to 6.	No	No	No	No	Yes	Yes
Number of observations	42,346	42,346	42,346	42,346	42,346	42,346
Number of schools	803	803	803	803	803	803

Note. Entries in the Table represent the IV coefficients from separate regressions of the relevant dependent variable on the fraction immigrants in 5th grade. The instrument is the predicted fraction of immigrants in grade 5 based on dates of birth and enrollment rules in place for that cohort. For the list of control variables, Table 4.

In the regressions with stratification, the whole sample was stratified into eight cells based on the total number of immigrants in grades 4 to 6 (1 to 5, 6 to 10, 11 to 15 etc.) and the regression includes cell dummies interacted with a constant and the full set of control variables. See text for details.

Robust standard errors (adjusted for clustering at the school level) in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

Table 9

The Effect of Fraction Immigrants in 5th Grade on Intermediate Outcomes

	No stratification	Stratification
High School Quality (high school matriculation rate in 1996-8)	−0.033	−0.128**
	(0.055)	(0.056)
Number of observations, number of schools	37,808, 802	37,808, 802
Fraction immigrants in 10th grade	0.192***	0.165***
	(0.046)	(0.049)
Number of observations, number of schools	41,683, 803	41,683, 803
Controls for individual and school characteristics	Yes	Yes
Controls for number of immigrants in grades 4 to 6 and number of children in grade 5	Linear	Second order terms and interaction within each cell
Stratification on number of immigrants in grades 4 to 6.	No	Yes

Note. Entries in the Table represent the coefficients from separate regressions of the relevant dependent variable on the fraction immigrants in 5th grade. Individual controls: mother's years of schooling, father's years of schooling, dummy indicators for whether parents' schooling is missing, number of siblings, a dummy for whether the number of siblings is missing, age in 1994, a gender dummy. School controls: a dummy for whether the elementary school is religious, class size in 1994, mathematics scores in 1991, verbal scores in 1991 and school socioeconomic index in 1991. In the regressions with stratification, the whole sample was stratified into eight cells based on the total number of immigrants in grades 4 to 6 (1 to 5, 6 to 10, 11 to 15 etc.) and the regression includes cell dummies interacted with a constant and the full set of control variables. See text for details. Robust standard errors (adjusted for clustering at the school level) in parentheses. One, two and three stars indicate statistical significance at the 10%, 5% and 1% level respectively.

(measured by the average matriculation rate in the high school between 1996 and 1998, which is at least two years before the cohort under analysis reached 10th grade); and the fraction of immigrants in 10th grade.

The results are presented in Table 9. The first column controls linearly for the number of immigrants in grades 4 to 6 and the number of children in grade 5, while the second column adopts the most flexible stratified specification (analogous to column 6 in Table 4). We find that immigrant concentration in 5th grade is positively correlated with immigrant concentration in 10th grade and negatively correlated with high school quality, although this result is sensitive to the specification used. A 10 point increase in the fraction of immigrants in 5th grade leads to a 1.65 point increase in the fraction of immigrants in 10th grade and a 1.28 point drop in the average matriculation rate of the high school attended.

The coefficient in a regression of the high school matriculation rate on high school quality is 0.744. If we assume that this can be taken as a causal effect, this implies that high school quality can account for about 63% of the total effect of immigrant concentration in the 5th grade on the high school matriculation rate.[24] Similarly, we calculate that the fraction of immigrants in 10th grade can account for about 31% of the total effect. Of course, these back of the envelope estimates should be taken with caution, since we do not have an estimate of the causal relationship between the intermediate and final outcomes. However, they suggest that immigrant concentration

[24] A 10 point increase in the fraction of immigrants in 5th grade leads to a 0.0128 drop in high school quality: $0.744 \times 0.0128 = 0.0095$, which is about 63% of 0.015, the estimated effect of a 10 point increase in 5th grade fraction of immigrants on the matriculation rate (Table 4, column 6).

in 5th grade affects high school outcomes both through a direct effect and through its effect on intermediate outcomes.

8. Conclusion

This article represents one of the first attempts to study the consequences of natives and immigrants interacting in the same social and learning environment, while paying particular attention to issues of identification and causality. In addition, this article breaks from the existing literature on peer effects within schools by looking at the long-term effects on natives in high school, rather than focusing on contemporaneous effects. After accounting for the endogenous placement of immigrants, both across schools and between grades within schools, we find that the fraction of immigrants in 5th grade has a substantial adverse effect on the passing rate for native students on their matriculation exams. Passing the matriculation exam is needed in order to attend college, so this is an important outcome in the Israeli context. Since students are typically kept in the same classes throughout elementary school, it is important to note that the immigrant concentration in 5th grade in Israel is a good proxy for the total exposure throughout elementary school.

These results stand in contrast with some of the recent estimates in the US literature on desegregation, which found little or no impact of desegregation programmes on the outcomes of white students (Angrist and Lang, 2004; Guryan, 2004). One possible explanation for the larger estimated effects in our study is that the average number of immigrant children in each class in our sample is particularly large. Also, given that the average class size in Israel is also relatively high, it is possible that the migration wave crowded out school resources that were already strained.

In many respects, these findings can be of general interest beyond the local Israeli context, even though we must keep in mind the differences between this particular migration wave and other similar events (migration waves in other western countries, or desegregation in US schools). The empirical episode studied here examines how immigrants from a relatively educated and skilled background from the former Soviet Union affect the human capital outcomes of natives. Despite the high levels of parental education, our analysis shows that the parents of these immigrants exhibit many troubling socio-economic patterns which could lead to a crowding out of school resources or generate other forms of negative peer effects on the outcomes of natives. The lessons learned from this analysis are becoming increasingly relevant to Western countries that are expected to absorb growing numbers of immigrants from Central and Eastern Europe (especially following the enlargement of the European Union), and others who are debating whether to move toward a more skill-biased immigration policy.

Our findings may also have implications for income inequality. The most important dividing line in schooling in the US is high school completion. In Israel and in many European countries, the crucial threshold is passing the matriculation exams, which opens the door to a college education. Angrist and Lavy (forthcoming) estimated the returns to a year of schooling in Israel at about 8.3%, while holders of the Israeli matriculation certificate earn a further 24%. The effect of immigrants on the matriculation outcomes of natives may therefore have severe implications for the income of disadvantaged natives as well as for income inequality. This is especially true in light of

our finding that disadvantaged students are more likely to be adversely affected by the presence of immigrants in the classroom.

Technical Appendix

Additional supporting information may be found in the online version of this article and on the RES website:
Appendix Table 1: Stratified OLS Regressions for the Outcomes of Native Israelis – Detailed Results
Appendix Table 2: Balancing Tests for the Actual Fraction of Immigrants
Appendix Table 3: Robustness Tests
Appendix Table 4: Balancing Tests for the Predicted Fraction of Immigrants
Appendix Table 5: Stratified IV Regressions for Outcomes for Native Israelis – Detailed Results
Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Hebrew University and Shalem Center;
Hebrew University and Royal Holloway University of London;
Boston University and Hebrew University

Submitted: 20 February 2007
Accepted: 31 July 2008

References

Ammermueller, A., and Pischke, J-S. (2006). 'Peer effects in European primary schools: evidence .from PIRLS', Centre for the Economics of Education, Discussion Paper No. 65 May.
Angrist, J.D. and Lang, K. (2004). 'Does school integration generate peer effects? Evidence from Boston's Metco Program', *American Economic Review*, vol. 94(5) (December), pp. 1613–34.
Angrist, J.D. and Lavy, V. (1999). 'Using Maimonides' rule to estimate the effect of class size on scholastic achievement', *Quarterly Journal of Economics*, vol. 114(2) (May), pp. 533–75.
Angrist, J.D. and Lavy, V. (forthcoming). 'The effect of high stakes high school achievement awards evidence from a group randomized trial', *American Economic Review*.
Betts, J.R. (1998) 'Educational crowding out: do immigrants affect the educational attainment of American minorities?', in (D.S. Hamermesh and F.D. Bean, eds.), *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, pp. 253–81, New York: Russell Sage Foundation.
Betts, J.R. and Fairlie, R.W. (2003). 'Does immigration induce "native flight" from public schools into private schools', *Journal of Public Economics*, vol. 87(5-6) (May), pp. 987–1012.
Boozer, M.A. and Cacciola, S.E. (2001). 'Inside the ''black box'' of Project STAR: estimation of peer effects using experimental data', Economic Growth Center Discussion Paper 832, Yale University, June.
Borjas, G.J. (2007). 'Do foreign students crowd out native students from graduate programs?', in (P.E. Stephan and R.G. Ehrenberg, eds.), *Science and the University*, pp. 134–49, Madison, WI: University of Wisconsin Press.
Burke, M.A. and Sass, T.R. (2008). 'Classroom peer effects and student achievement', Federal Reserve Bank of Boston Working Paper 08-5, June.
Card D. and Rothstein, J. (2007). 'Racial segregation and the black-white test score gap', *Journal of Public Economics*, vol. 91(11-12) (December), pp. 2158–84.
Cohen-Goldner, S. and Paserman, M.D. (2006). 'Mass migration to Israel and natives' transitions from employment', *Industrial and Labor Relations Review*, vol. 59(4) (July), pp. 630–52.
Cohen-Goldner, S. and Paserman, M.D. (2004). 'The dynamic impact of immigration on natives' labor market outcomes: evidence from Israel', CEPR Discussion Paper 4640, September.

Cutler, D.M. and Glaeser, E.L. (1997). 'Are ghettos good or bad?', *Quarterly Journal of Economics*, vol. 112, pp. 827–72.
Eckstein, Z. and Weiss, Y. (2002). 'The integration of immigrants from the former Soviet Union in the Israeli labor market', in (Ben-Bassat, ed.), *The Israeli Economy, 1985-1998: From Government Intervention to Market Economics, Essays in Memory of Prof. Michael Bruno*, Cambridge MA: MIT Press.
Eckstein, Z. and Weiss, Y. (2004). 'On the wage growth of immigrants: Israel 1990-2000', *Journal of the European Economic Association*, vol. 2(4) (June), pp. 665–95.
Edin, P-A., Fredriksson, P. and Åslund, O. (2003). 'Ethnic enclaves and the economic success of immigrants – evidence from a natural experiment', *Quarterly Journal of Economics*, vol. 118, pp. 329–57.
Evans, W.N., Oates, W.E. and Schwab, R.M. (1992). 'Measuring peer group effects: a study of teenage behavior', *Journal of Political Economy*, vol. 100(5), pp. 966–91.
Gould, E.D., Lavy, V. and Paserman, M.D. (2004*a*). 'Immigrating to opportunity: estimating the effects of school quality using a natural experiment on Ethiopians in Israel', *Quarterly Journal of Economics*, vol. 119 (2), pp. 489–526.
Gould, E.D. Lavy, V. and Paserman, M.D. (2004*b*). 'Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence', NBER Working Paper 10844, October.
Goux, D. and Maurin, E. (2007). 'Close neighbours matter: neighbourhood effects on early performance at school', Economic Journal, vol. 117(523), (October), pp. 1193–215.
Guryan, J. (2004). 'Desegregation and black dropout rates', *American Economic Review*, vol. 94(4) (September), pp. 919–43.
Hanushek, E.A., Kain, J.F. and Rivkin, S.G. (2002). 'New evidence about Brown v. Board of Education: the complex effects of school racial composition on achievement', National Bureau of Economic Research: Working Paper no. 8741.
Hanushek, E.A Kain, J.F., Markman, J.M. and Rivkin, S.G. (2003). 'Does peer ability affect student achievement?', *Journal of Applied Econometrics*, vol. 18(5), (September/October), pp. 527–44.
Heckman, J.J. and Robb, R. Jr. (1985). 'Alternative methods for evaluating the impact of interventions: an overview', in (J.J. Heckman and B. Singer, eds.), *Longitudinal Analysis of Labor Market Data.*, pp. 156–245, Cambridge: Cambridge University Press.
Hoxby, C.M. (1998). 'Do immigrants crowd disadvantaged American natives out of higher education?', in (D.S. Hamermesh and F.D. Bean, eds.), *Help or Hindrance? The Economic Implications of Immigration for African Americans*, pp. ?, New York: Russell Sage Foundation,.
Hoxby, C.M. (2000*a*). 'Peer effects in the classroom: learning from gender and race variation', NBER Working Paper No. 7867, August.
Hoxby, C.M. (2000*b*). 'The effects of class size on student achievement: new evidence from population variation', *Quarterly Journal of Economics*, vol. 115(4) (November), pp. 1239–85.
Jacob, B.A. (2004). 'Public housing, housing vouchers and student achievement: evidence from public housing demolitions in Chicago', *American Economic Review*, vol. 94(1) (March), pp. 233–58.
Jepsen, C. and Rivkin, S. (2009). 'Class size reduction and student achievement: the potential tradeoff between teacher quality and class size', *Journal of Human Resources*, vol. 44(1), (Winter), pp. 223–50.
Katz, L.F., Kling, J.R. and Liebman, J.B. (2001). 'Moving to opportunity in Boston: early results from a randomized mobility experiment', *Quarterly Journal of Economics*, vol. 116, pp. 607–54.
Krueger, A.B. (1999). 'Experimental estimates of education production functions', *Quarterly Journal of Economics*, vol. 114, pp. 497–532.
Lefgren, L. (2004). 'Educational peer effects and the Chicago public schools', *Journal of Urban Economics*, vol. 56(2) (September), pp. 169–91.
Manski, C.F. (1993). 'Identification of endogenous social effects: the reflection problem', *Review of Economic Studies*, vol. 60(3), (July), pp. 531–42.
Oreopoulos, P. (2003). 'The long-run consequences of living in a poor neighborhood', *Quarterly Journal of Economics*, vol. 118, pp. 1533–75.
Sacerdote, B. (2001). 'Peer effects with random assignment: results for Dartmouth roommates', *Quarterly Journal of Economics*, vol. 116, pp. 681–704.
Sanbonmatsu, L., Kling, J., Duncan G. and Brooks-Gunn, J. (2006). 'Neighborhoods and academic achievement: results from the Moving to Opportunity experiment', *Journal of Human Resources*, vol. 41(4), (Fall), pp. 649–91.
Vigdor, J. and Nechyba, T. (2004). 'Peer effects in elementary school: learning from "apparent" random assignment', mimeo., Duke University, October.
Weinberg, B.A., Reagan, P.B. and Yankow, J.J. (2004). 'Do neighborhoods affect hours worked: evidence from longitudinal data', *Journal of Labor Economics*, vol. 22 (4), pp. 891–924.
Zimmerman, D.J. (2003). 'Peer effects in academic outcomes: evidence from a natural experiment', *Review of Economics and Statistics*, vol. 85, pp. 9–23.