Effects of Free Choice Among Public Schools

VICTOR LAVY

Hebrew University, Royal Holloway University of London, CEPR and NBER

In this paper, I investigate the impact of a programme in Tel-Aviv, Israel, that terminated an existing inter-district busing integration programme and allowed students free choice among public schools. The identification is based on difference-in-differences and regression discontinuity designs that yield various alternative comparison groups drawn from untreated tangent neighbourhoods and adjacent cities. Across identification methods and comparison groups, the results consistently suggest that choice significantly reduces the drop-out rate and increases the cognitive achievements of high-school students. It also improves behavioural outcomes such as teacher–student relationships and students' social acclimation and satisfaction at school, and reduces the level of violence and classroom disruption.

1. INTRODUCTION

This paper presents an analysis of the impacts of school choice among public schools on students' cognitive achievements and behavioural outcomes. The analysis is based on a school choice programme that is very similar to recent school choice reforms in the United States, which are the result of federal court decisions terminating race-based bussing plans that had been in effect for decades. Well-known examples are the choice programmes in Seattle (1999) and in Mecklenburg County, North Carolina (2002).[1] The Tel-Aviv School Choice Program (hereafter, TASCP) studied in this paper had an identical policy benchmark, whereby the assignment of students to secondary schools before the reform was motivated and guided by social and ethnic integration and included bussing of some students across the city's schooling districts. The 1994 programme terminated the previous system and granted students choice among schools in and outside their school district.

During the experimental phase (the first 2 years) of the programme, it was implemented only in schooling district 9, the city's largest. Focusing on this period, I use administrative data to follow students from the moment of school choice (the beginning of middle school) to the end of high school and estimate the impact of school choice on students' outcomes, including the drop-out rate and success in high school matriculation exams. The latter are key determinants of post-secondary schooling and market wages in Israel. I then provide empirical evidence on the effect on several behavioural outcomes, such as discipline and violence in the classroom, student–teacher relationships and students' social acclimation in school. Some of these outcomes can also be viewed as mediating factors of the effect of choice.

---

1. Many other cities including Nashville, Oklahoma City, Denver, Wilmington, and Cleveland replaced busing with school choice. Other examples include the Pinellas County, FL, Montclair, NJ and Cambridge, MA.

I use two different identification strategies. Both are based on the special geographical location of district 9. On its West side district 9 borders three of the other eight school districts (6, 7, and 8) of the city, whereas on its East side it borders two adjacent cities that belong to the same metropolitan area, Givataim and Ramat-Gan (hereafter, GR). South of district 9 is Holon, another large city which is part of the same metropolitan area. The gradual implementation of the programme makes districts 6–8 a potentially appropriate comparison group. Similarly, either GR or Holon can also be a comparison group because they did not introduce school choice before or after the Tel-Aviv programme. The downside is that districts 6–8 had marginally worse pre-programme mean pupil outcomes (though similar characteristics) relative to district 9, while GR and Holon had better outcomes and characteristics than district 9. On the positive side, however, the differences in characteristics were stable before and after the choice programme, as were the mean outcomes of the potential comparison groups, lending an opportunity for a promising difference-in-differences estimation strategy that exploits panel data on affected and unaffected cohorts. Remarkably, all three comparison groups yield almost identical treatment estimates.

The second identification strategy that I use is an RD design that is based on a sample of pupils drawn from a narrow band around the municipal border between GR and district 9.[2] Similar to Black (1999), limiting the sample to observations within such a narrow bandwidth yields a sample that is balanced in the constant observable and unobservable characteristics of treatment and control units. I use this RD-natural experiment framework jointly with the before and after panel data in difference-in-differences estimation. The findings obtained using this RD method are very similar to the treatment estimates based on either of the three alternative comparison groups and all of district 9 students used for the difference-in-differences estimation. This suggests that the sharp reduction in the drop-out rate and the significant improvement in matriculation outcomes can be interpreted as a causal effect of the choice programme.

The second part of the paper identifies the effect of school choice on behavioural outcomes such as disruption and violence in class, student–teacher relationships and students' social acclimation in class and overall satisfaction with school. These outcomes are based on using a unique national survey administered to middle and primary-school students. The effects of choice on these behavioural outcomes are interesting in their own right, as exemplified by numerous studies that highlight their central role in school choice decisions (see, e.g., Hoxby, 1998; Black, 1999; Cullen, Jacob, and Levitt, 2006; Kane, Riegg, and Staiger, 2006; Imberman, forthcoming) and in teachers' transfer and quit decisions (see, e.g., Boyd *et al*., 2003; Hanushek, Kain, and Rivkin, 2004). However, the effect of choice on some of these factors can be viewed as a mediating channel through which choice affects cognitive outcomes.

In studying the effect of choice on behavioural outcomes I am able to exploit an additional identification strategy based on longitudinal data. I assemble this data using the fact that I observe students in two different school environments, primary-school without school choice and middle school with school choice. In this case, I generate student fixed effects estimates that reflect how a change in available choices as a result of the student's transition from primary to middle school is associated with changes in behavioural outcomes. The evidence shows that school choice in Tel-Aviv lowered the level of violence and classroom disruption, improved teacher–student relationships and increased students' social acclimation and satisfaction at school.

---

2. Districts 6–7 are not appropriate for such an RD strategy because its number of pupils per cohort is very small and the sample of students that reside close to the border with district 9 is even smaller.

As noted above, the background and the structure of the Tel-Aviv choice programme are very similar to the 2002 Mecklenburg County, North Carolina, school choice programme, which recently received academic attention. Hastings, Kane, and Staiger (2005) estimated the role of proximity and of mean test score increases in shaping parental preferences for school characteristics, whereas Hastings, Kane, and Staiger (2006) estimated the effect of attending a first-choice school on students' test scores, and report that it is not associated with improvements in any academic outcomes. There is, however, an earlier relevant literature regarding choice programmes in the United States that allowed specific groups to attend private or charter schools. Among the first of these studies, Rouse (1998) evaluated the effect of the Milwaukee Parental Choice Program. Others are Mayer *et al.* (2002), Angrist, Bettinger, Bloom, King, and Kremer (2002), Angrist, Bettinger, and Kremer (2006), Krueger and Zhu (2004), Cullen, Jacob, and Levitt (2005), and Hoxby (2002). Some programmes allowed public school students to apply to magnet schools and to public schools outside of their neighbourhood (Cullen *et al.*, 2006). Several studies looked at housing markets as conveying the effect of a potentially informative, indirect form of school choice, and established a relationship between housing markets and school quality or productivity (Black, 1999; Hoxby, 2000; Rothstein, 2006).[3]

The rest of the paper is structured as follows: Section 2 presents the background and details of TASCP and gives some preliminary information about the pattern of choice. Section 3 describes the data, and Section 4 presents the identification strategy and the estimates of the choice programme's effects on academic achievements. Section 5 presents evidence on the effect of choice on the behavioural outcomes and mobility rates of students and Section 6 concludes.

2. THE TEL-AVIV SCHOOL-CHOICE PROGRAM

In May 1994, the Israeli Ministry of Education approved TASCP as a 2-year experiment to be implemented in the city's 9th district. It was the first-choice programme in the country since the 1968 education reform that enacted compulsory integration in grades 7–9.[4] TASCP was a response to parents' dissatisfaction with students' outcomes and with the rigid lack of school choice. Its objectives were to give disadvantaged students access to better schools, facilitate a better match between students and schools, and motivate school productivity improvements through competition. The 9th schooling district included 16 public primary schools–12 secular and 4 religious. Until 1994, the graduates of five of the secular primary schools were bussed to one of five secondary schools in districts 1–5 in north Tel-Aviv (about 36% of the districts' pupils) and a few more of the districts' pupils (5%) were enrolled in charter schools outside the district (Tel-Aviv Educational Authority, 1994). The graduates of the other seven secular primary schools were assigned to one of the three secondary schools within district 9.[5] In May 1994, the education board of Tel-Aviv announced that as of September 1994 this system would

---

3. Several recent studies examine the effect of general school choice reforms on school performance, for example Ahlin (2003) and Sandstrom and Berstrom (2002) in Sweden; Bradley, Johnes, and Millington (2001) and Gibbons, Machin, and Silva (2008), in the United Kingdom; Hsieh and Urquiola (2003) in Chile; and Fiske and Ladd (2000) in New Zealand.

4. The 1968 reform established a three-tier structure of schooling: primary (grades 1–6), middle (7–9), and high school (10–12). The reform established neighbourhood school zoning as the basis of primary enrolment and of the integration and bussing of students out of their neighbourhoods in middle school. In Tel-Aviv, most middle schools were part of six-year high schools and there were several high schools who offered only the higher grades (10th–12th).

5. These schools were located on the same campus but they were very different in terms of their curriculum of studies and programmes offered to students. For example, one included low and high tech vocational schooling.

be replaced by free choice for the incoming 7th graders in the district, while older cohorts would continue with the old system. The structure of choice was as follows. At the end of sixth grade each student was asked to rank his preference among the five schools in his choice set, which consisted of the district's three secondary schools and two out of district schools (in districts 1–5 which were the same schools to which students were bussed before the programme). The choice set varied among students in accordance with the primary school they attended (Tel-Aviv Educational Authority, 1995). In the event of excess demand for a particular school, students were assigned to schools in a manner that maintained a socioeconomic balance matching the respective makeup of the city.[6] The city opened choice information centres and ran workshops to parents and pupils, and high schools had open days to provide additional information to the incoming 7th grade cohorts (Tel-Aviv Educational Authority, 1996). City reports indicate that in the programme's first year, 90% of students received their first choice and others their second. In the second year the respective first-choice rate was even higher,[7] since 2003 excess demand was resolved by lottery. Another relevant factor was an expansion of the supply of middle-school classes as four high schools, two in district 9 and two in the city's north districts, who had only the higher grades (10th–12th), were expanded at the commencement of the reform to include also the middle-school grades. Despite these changes, over time the choice programme led to the expansion of some high schools and to the contraction of others (one school was even closed due to declining enrolment). Enrolment in the city's schools was also affected by the stricter enforcement of the Ministry's rule that pupils were not allowed to attend schools outside of Tel-Aviv. Schools who enjoyed expanded enrolment gained more resources as their budget was determined according to enrolment. Some additional resources were targeted to all schools in the city for the purpose of tracking and assisting underperforming students at the beginning of middle school (for these details and more, see Heiman and Shapira, 1998, 2002).

The choice programme was accompanied by a decision that all the city's post-primary schools would be six grade structures that included the middle (7th–9th) and higher grades (10th–12th) as part of the same school. Most of the city's post-primary schools were already such structures and only four schools had to be expanded to include the middle-school grades. This allowed the city in practice to cancel the admission process at the end of 9th grade and to introduce the concept of "persistence" whereby students automatically enrolled into 10th grade in the same school in which they completed their middle-school education. This important component of the reorganization of the school system in Tel-Aviv, which took place throughout the city at the same time, very much limited the ability of schools to select students to their higher grades based on academic performance. The explicit default became that pupils could progress through their secondary education in the same school they chose in 7th grade. To prevent any student having this default option, a school had to gain an explicit approval of a special city committee that granted it only in cases of pupils with severe behavioural problems and never on the grounds of poor academic performance. This policy change most likely explains a large part of the dramatic decline in the pupil transfer rate in 9th grade, from about 50% before the choice programme to about 15% following it. This decline was achieved despite stubborn resistance by some high achieving high schools to the policy that forbade them selecting their students based on academic ability. However, schools were given much more autonomy in pedagogy and in the expansion of academic programmes and they received additional funding to improve physical infrastructure.

---

6. Siblings in the same school and school capacity were also used as criteria to balance enrollment.
7. The Tel-Aviv Educational Authority (1999). More related evidence is provided in Levy, Levy and Libman (1996, 1997)

In 1996, the experiment was expanded to district 8, in 1998 to district 7, and in the following year to the rest of the city (Tel-Aviv Education Authority, 2001). During the first 4 years of the programme, two evaluation teams provided useful and important insights with respect to the educational and social changes that took place in schools and among teachers, students, and parents. Heiman and Shapira (1998, 2002) provide detailed summaries of the programme and the changes observed over the years. The short- and long-term causal impact of the programme, however, has not been studied.

3. THE DATA

The data I use in this study comes from administrative records of the Ministry of Education on the universe of Israeli primary schools during the 1992–1994 school years. The files contain an individual identifier, a school and class identifier, and the following family-background variables: fathers' and mothers' years of schooling, number of siblings, gender, immigration status (= 1 if arrived in the country during the previous 5 years, in line with the Ministry of Education's official definition) and family ethnic origin (Asia/Africa, Europe/America or Israel) and the students' home addresses. Data on distances from the students' homes to the municipal border between Tel-Aviv and GR were obtained from the Central Bureau of Statistics. The three cohorts on which I focus in this study had sufficient time within the sample period (which ends with the 2000/2001 school year) to finish high school if they progressed through the system without repeating classes.

I link the primary-school records to individual data on high-school enrolment and matriculation-exam outcomes in the 1998/99 through the 2001/02 school years. This allows monitoring each student from the end of 6th grade (in 1992, 1993, or 1994) to the advanced stages of high school. As outcomes I use an indicator of dropping out before completing 12th grade, an additional indicator for matriculation (*Bagrut*) eligibility,[8] credit-weighted average score on the matriculation exams, number of matriculation credits, number of matriculation credits in science subjects and number of matriculation subjects at honours level. Several of these outcomes are used to screen and select students for prestigious universities and desired academic programmes such as medicine, engineering, and computer science.

Columns 1–2 of Table 1 present summary statistics for the cohort that completed primary school in June 1994 (the first enrolled in the choice programme) in Tel-Aviv and in district 9. A comparison of column 2 with column 1 and the resulting *t*-statistics reported in column 6 indicate that district 9 students had lower socioeconomic characteristics than other students in Tel-Aviv. For instance, they had a lower level of parental schooling, larger family size, a higher proportion of students with Asian/African origins and a lower proportion with European/American origins. Similar results are obtained when using the cohort that completed primary school in June 1993, which was the last cohort before the onset of the choice programme.

---

8. Matriculation eligibility is ascertained by passing a series of national exams in core and elective subjects, most taken in 12th grade. Students choose to be tested at various proficiency levels, each test awarding 1–5 credit units per subject depending on difficulty. A minimum of 20 credit units is required to qualify for a matriculation certificate, which is received by about half of all high-school seniors. Similar high-school matriculation exams are found in many countries and in some US states. Examples include the French Baccalaureate, the German Certificate of Maturity, the Italian Diploma di Maturità, the New York State Regents examinations and the recently instituted Massachusetts Comprehensive Assessment System.

TABLE 1

Student mean characteristics and mean equality t-tests by location: Tel-Aviv, district 9, districts 6–8, Givataim and Ramat-Gan, and Holon, 1994

	Sample						Mean equality t-stats			
	Tel-Aviv districts:			Givataim & Ramat Gan	Holon	*t*-value: (2) vs. (1)	*t*-value: (2) vs. (3)	*t*-value: (2) vs. (4)	*t*-value: (2) vs. (5)
	All	9	6–8						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Students' characteristics									
Fathers' years of schooling	12·049	10·336	9·755	12·097	11·635	−3·070	−1·173	−4·704	−3·554
Mothers' years of schooling	12·127	10·566	10·128	12·364	11·880	−3·058	−0·802	−4·660	−3·521
Number of siblings	1·948	2·145	2·278	1·910	2·121	1·627	0·425	2·286	0·227
Gender (male = 1)	0·490	0·475	0·538	0·526	0·495	−0·855	2·364	−2·212	−0·787
Immigration status	0·054	0·052	0·090	0·077	0·083	−0·209	1·253	−1·580	−1·633
Country of origin–Israel	0·630	0·659	0·530	0·609	0·542	1·253	−3·270	1·711	3·491
Country of origin–Asia/Africa	0·157	0·196	0·219	0·194	0·231	2·140	0·823	0·073	−1·632
Country of origin–Europe/America	0·096	0·047	0·042	0·103	0·084	−3·580	−0·329	−3·861	−2·489
Number of students	2809	791	576	1842	2426	3600	1367	2633	3217

Notes: The samples include only students in secular state schools. Religious Jewish schools and Arab schools are excluded.

4. IDENTIFICATION STRATEGY

4.1. *Using late enrolled neighbouring school districts as a comparison group*

Due to the gradual implementation of the choice programme, the school districts that joined the programme 2 years after school district 9 can be used as a comparison group. Because all the schools in districts 1–5 were included in the choice sets of students in district 9, only districts 6–8 could serve as a comparison group. Districts 6 and 8 are adjacent to district 9 but their sample of students is too small and therefore I consider district 7 as well to be part of the potential comparison group. All these three districts are part of the South of the city, geographically adjacent or near district 9 (see Map 1), and their population is much more similar to that of district 9's than that of the North of the city. This is demonstrated in Table 1, columns 2 and 3: districts 6–8 students are very similar in mean characteristics to district 9 students (t-statistics for these differences are presented in column 7). For example, the fathers' and mothers' years of schooling differences are 0.58 (t-value $= 1.17$) and 0.44 (t-value $= 0.80$), respectively, relative to respective district 9's means of 10.3 and 10.6. Another example of the close similarity between the two groups is reflected in the composition of students by ethnic origin: the difference in the proportion of students from Asia/Africa is -0.02 (t-value $= 0.82$) relative to a mean of 0.196 in district 9, and the difference in the proportion of students from
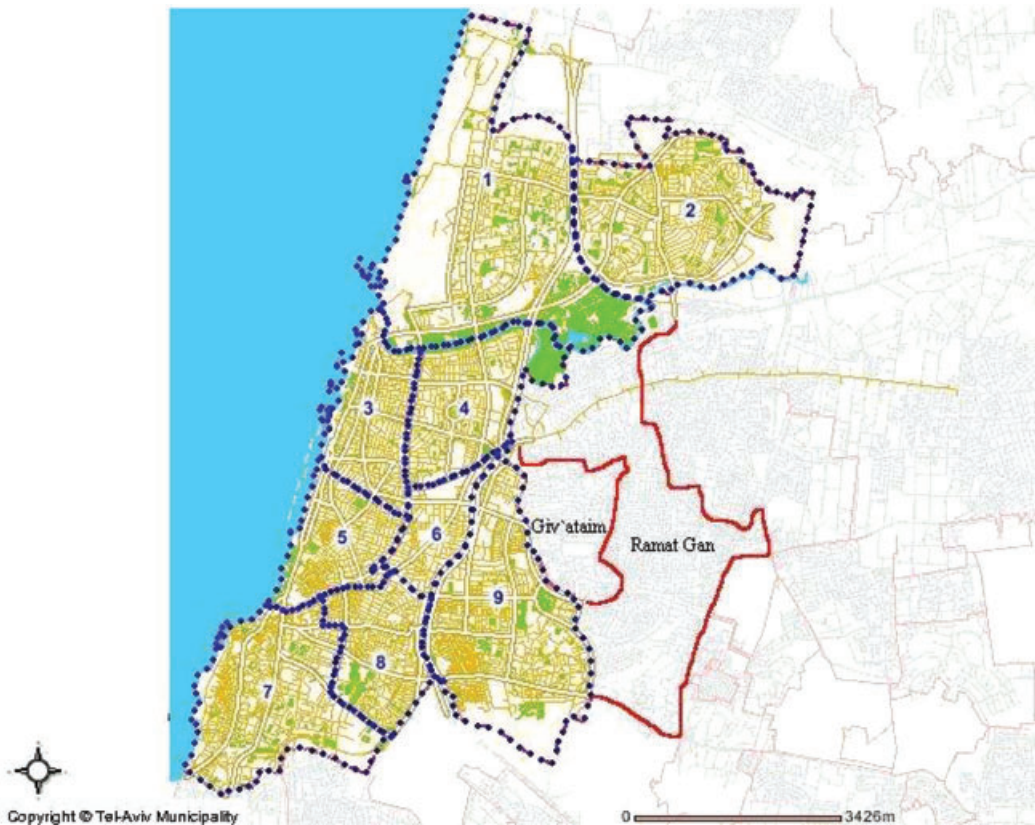
MAP 1

Tel-Aviv city, school districts 1–9, and the cities Giv'ataim and Ramat-Gan

Europe/America is −0.005 (*t*-value = 0.33) relative to 0.047 in district 9. The 1992 and 1993 cohorts are equally well balanced (results shown in the online Appendix Table A1) which indicates stability in the composition of students in both groups over the 1993–1994 cohorts.

Therefore, the first identification approach that I apply in this paper is based on a contrast between district 9 and districts 6–8, before and after the programme was implemented. I use data on pre- and post-programme cohorts (panel data) in a difference-in-differences framework that removes any remaining time invariant heterogeneity across treated and control groups. Because this DID estimation compares two consecutive cohorts, and because the programme was implemented immediately after it was announced, it is reasonable to assume that the remaining differences were constant within this narrow time range. A concern with this DID approach, however, is that the immediately prior cohort that I use as a control group might be affected through spillover effects at the school level. As these students will be attending the same schools as the treated students, peer effects or competitive effects on school productivity might impact the untreated students as well. A useful way to check that the results are not biased by such spillover effects is to test whether there are significant treatment effects when using two previous cohorts for estimating DID models. Such falsification tests are also useful to test for the effect of omitted time varying factors. I therefore exploit the presence of multiple control groups formed by successive cohorts not exposed to the choice programme (the 1992 and 1993 6th grade cohorts) to conduct falsification tests for spillover effects and for spurious treatment effects.[9]

4.2. *Using adjacent cities as a comparison group*

Tel-Aviv is part of a metropolitan area whose core region includes five major cities. District 9 includes the city's southeastern neighbourhoods (see Map 1) and is tangent to two of the neighbouring cities: Givataim and Ramat-Gan (referred to as GR). GR have independent and separate education systems and therefore were not part of the school choice reform of Tel-Aviv.[10] The metropolitan geography of district 9 and the adjacent cities raises the possibility of using GR students as a comparison group for district 9. However, as shown in Table 1, column 4, GR students are very different in mean characteristics from district 9 students (*t*-statistics for these differences are presented in column 8). However, these differences are very stable as they are similar in 1992 and 1993 as well (online Appendix Table A1). The solution, therefore, to the pre-programme imbalances is to use data on pre- and post-programme cohorts (panel data) in a difference-in-differences framework that removes time invariant heterogeneity across treated and control groups. I therefore use the DID method and apply it to the sample composed of district 9 and GR students.

Holon is another city adjacent to Tel-Aviv (South) and it is very close to district 9. It is, however, more similar to district 9 in its characteristics (see columns 5 and 9 in Table 1) than GR. The evidence that I will show below will demonstrate that the results based on Holon as a comparison group are identical to those based on GR as a comparison group. Furthermore, and even more striking, both GR and Holon based estimates are almost identical to the evidence based on using districts 6–8 as a comparison group. The fact that two alternative sets of DID

9. See Heckman and Hotz (1989) and Rosenbaum (1987). Duflo (2001) applied a similar idea using the difference between untreated cohorts across different treated and untreated regions as a falsification test. An illustration of these general issues in a different setting is presented in Galiani, Gertler, and Schargrodsky (2005).

10. Givataim, Ramat-Gan, and Holon high-school enrolment system before the inception of the TASCP was based on zoning and it has not changed since, nor have these cities undergone any other major educational reform since 1994.
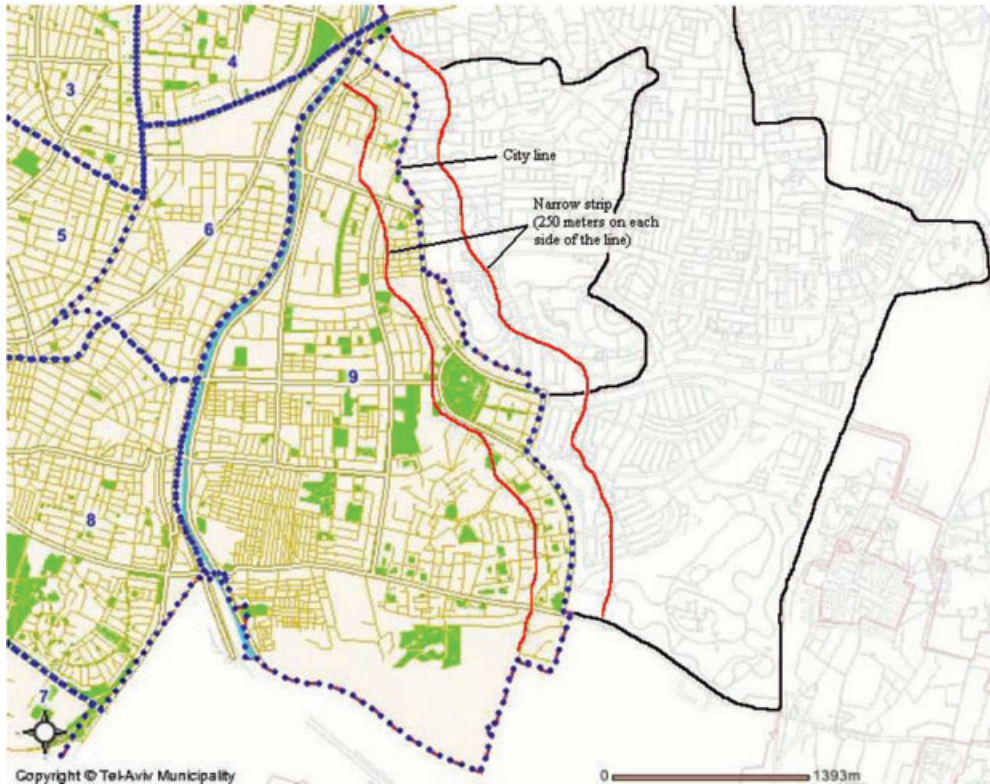
MAP 2

Tel-Aviv school district 9 and tangent neighbourhoods of Giv'ataim and Ramat Gan
Notes: The thin lines approximately draw the band.

estimates, one that is based on a comparison group that has much better characteristics and outcomes (GR or Holon) than the treated group and a second that is based on a comparison group that has marginally worse characteristics and outcomes (districts 6–8), yield exactly the same results is reassuring given the possibility that the DID estimates are biased because of regression to the mean or due to differential time trends in unobserved heterogeneity between treatment and control.

4.3. *Using adjacent neighbourhoods as a comparison group*

A regression discontinuity design that limits the sample, in a manner similar to Black (1999), to observations within a narrow band around the municipal border between district 9 and GR may eliminate the imbalances observed in columns 2 and 4 of Table 1, because proximity of residence may be paralleled by similarity in other characteristics.[11] Indeed, the physical and other characteristics of the communities within this strip (e.g., type and average size of homes)

---

11. In Black (1999), school quality varies across school zoning boundaries and these differences are capitalized into housing prices, because they affect where households choose to live. In marked contrast, the RD strategy that is proposed here in the district 9/GR setting is based on the assumption that households' preferences lead them to live close to the border and they are indifferent about being on one side or the other.

are identical, as are zoning laws and municipal (kind of property) taxes which are determined by the central government. But presumably, there might still be some differences, such as the political affiliation of the mayor, for example. The concern remains then that such remaining differences may confound the effect of the programme. As above, the use of data on pre- and post-programme cohorts in a difference-in-differences framework will remove such time invariant heterogeneity across treated and control groups.

For the RD natural-experiment method, I define samples based on drawing symmetric bands around the municipal border, starting from 250 metres on each side and increasing gradually (Map 2 presents an example of two such symmetric bands). As will be shown below, contrary to the large imbalances found when comparing all of district 9 and GR, the natural experiment samples based on narrow bands around the municipal border yield perfectly balanced treatment and control groups.

Table 2 presents detailed descriptive statistics and balancing tests for equality of the means of the treated and the comparison groups, for samples based on bandwidths of 250 metres, and 500 metres. Results are shown for the pre- (1993) and post- (1994) cohorts of treatment.[12] All 16 estimates of the treatment–control differences in 1994 are not statistically different from zero and in most cases they are also very small. For example, the fathers' and mothers' years of schooling differences in 1994 are -0.458 (s.e. $= 0.959$) and -0.229 (s.e. $= 0.923$), respectively, relative to respective means of 11.6. The 1993 cohort is equally well balanced, except that a gap in the proportion of immigrants can be observed. This difference is likely random because, as will be shown in the next section, it is paralleled by small and insignificant pre-programme treatment–control differences in outcomes. Note again that the means and treatment–control contrast for 1993 are similar to the respective evidence for 1994 which suggests a stable composition of students in both groups over the 1993–1994 cohorts. It is therefore safe to conclude that the treatment–control contrast in the 250 bandwidth sample truly reflects a natural experiment that can be used to identify the general effect of the choice programme. However, this RD sample might be too small to allow precise estimation of treatment effect. I therefore present in columns 3–4 of Table 2 balancing evidence based on a bandwidth of 500 metres. The treatment and control group are still well balanced: none of the 16 estimates is statistically different from zero and most difference estimates are also small. This sample has an additional advantage of being much larger (more than twice) than that used in columns 1 or 2 and therefore it is more likely to yield more precise estimates.

4.4. *Estimation*

I first present a controlled comparison of treated and untreated students using cross-section samples of pre- and post-treatment cohorts based on the following regression:

$$y_{ijt} = x_{ijt}\beta + Z_j d + u_{ijt} \tag{1}$$

where y_{ijt} is the i-th student's outcome in school j and year t; x_{ijt} is a vector of the same student's characteristics; Z_j is the treatment indicator (which equals 1 for district 9 students) and d is the treatment effect. I will estimate the equation using three samples, each corresponding to one of the comparison groups. The first sample pools district 9 with districts 6–8, the second pools district 9 and GR (or Holon) and the third is the natural-experiment sample.

In addition, I use the before-and-after cross-section data as stacked panel data that permits regression analysis with controls for primary-school fixed effects. Therefore, I will estimate

---

12. Missing data on exact addresses for GR in 1992 does not permit a similar analysis for 1992.

TABLE 2

Descriptive statistics and balancing tests–natural experiment samples

Dependent variable	Up to 250 m		Up to 500 m	
	1993 (1)	1994 (2)	1993 (3)	1994 (4)
Fathers' years of schooling				
Control group mean	11·955	11·569	11·706	11·732
Treatment–control difference	−0·271	−0·458	−0·274	−0·939
	(0·640)	(0·959)	(0·643)	(0·690)
Mothers' years of schooling				
Control group mean	11·735	11·608	11·765	12·074
Treatment–control difference	−0·023	−0·229	−0·185	−0·891
	(0·622)	(0·923)	(0·616)	(0·663)
Number of siblings				
Control group mean	1·985	2·026	2·086	1·955
Treatment–control difference	0·190	0·137	0·084	0·164
	(0·147)	(0·287)	(0·142)	(0·242)
Gender (male = 1)				
Control group mean	0·515	0·526	0·543	0·522
Treatment–control difference	0·039	−0·031	0·022	−0·054
	(0·050)	(0·048)	(0·038)	(0·038)
Immigration status				
Control group mean	0·025	0·026	0·058	0·030
Treatment–control difference	0·082	0·037	0·021	0·028
	(0·039)	(0·043)	(0·026)	(0·027)
Country of origin–Israel				
Control group mean	0·630	0·664	0·601	0·635
Treatment–control difference	−0·048	−0·027	0·022	−0·001
	(0·056)	(0·064)	(0·033)	(0·048)
Country of origin–Asia/Africa				
Control group mean	0·190	0·159	0·220	0·197
Treatment–control difference	0·036	−0·007	−0·013	−0·023
	(0·041)	(0·044)	(0·036)	(0·035)
Country of origin–Europe/America				
Control group mean	0·115	0·112	0·097	0·106
Treatment–control difference	−0·053	−0·023	−0·030	−0·031
	(0·032)	(0·030)	(0·027)	(0·026)
Number of students				
Control group	200	232	464	471
Treatment group	177	190	329	344

Notes: Standard errors in parentheses are adjusted for primary-school level clustering. Sample is limited to schools that appear both before- and- after treatment in each of the subsamples that are used in the difference-in-differences estimates of Table 3. The natural experiment samples contain pupils who reside in tangent neighbourhoods within a 250- or 500-m band on both sides of the city border.

stacked models using 3 (or just 2) years of cross-section data combined. The treatment indicator Z_{jt} is now defined as the interaction between a dummy for the year 1994 and the district 9 indicator, as follows:

$$y_{ijt} = \mu_j + \pi_t + x_{ijt}\beta + Z_{jt}d + \varepsilon_{ijt} \tag{2}$$

where μ_j is the primary school fixed effect and π_t is a year (i.e., 1992, 1993, and 1994) effect. Apart from providing a check on the precision of the 1992–1993 vs. 1994 contrast in treatment

effects, equation (2) may be seen as a framework for the control of omitted school effects that correlate with treatment status. The validity of this control, however, depends on the validity of an additive conditional mean function as a specification for potential outcomes in the absence of treatment.

5. RESULTS

5.1. *Evidence based on using districts 6–8 as a comparison group*

Columns 1–3 of Table 3 present the results for three cohorts, 1992–1994. There are six panels of results in the table, one for each of the six outcomes. The estimates presented in columns 1–2 in the first row of each of the panels show that district 9 students have better high-school outcomes than districts 6–8 students before the programme started (1992–93). The outcome levels (first row in each panel) and treatment–control simple mean differences (second row in each panel) are remarkably similar in both years. For example, the unconditional mean drop-out rates in district 9 in 1992 (18.1%) and in 1993 (19.3%) are approximately a third lower than the corresponding rates in districts 6–8. The mean matriculation rates in district 9 in 1992 and 1993 (43.6 and 44.6%, respectively) exceed those of districts 6–8 by more than 45%. Similar differences are observed in the other outcomes presented in the table. However, controlling for students' characteristics (levels of maternal and paternal education, number of siblings, gender, immigrant status, and ethnicity) greatly reduces these baseline differences. The treatment–control conditional mean difference in the drop-out rate in 1992, for example, is −6.6% as against a simple mean difference of −10.4%. The corresponding matriculation rate unconditional difference was 15.3% while the respective conditional difference was 9.8%. This pattern recurs in all six outcomes, suggesting that a third or more of the observed outcome differences are explained by observed differences in characteristics.

Column 3 in Table 3 presents the respective cross-section estimates for the cohort that was exposed to the programme. Comparing the simple treatment–control mean differences and the controlled differences of the 1994 cohort with those of the two pre-programme cohorts reveals a large relative improvement in district 9 students' outcomes. The magnitude of improvement implied by the comparison of the simple differences is very similar to that based on the controlled differences. The DID estimates based on the use of these cross-sections, do an even better job of demonstrating this important similarity and provide a concise summary of these results.

Column 4 presents DID estimates when all three cross-sections are used as stacked panel data. I also estimate DID models when only the 1992 or the 1993 cohorts are included as baseline and the results are unchanged. Therefore, I present and discuss only the results where both years were used as a baseline. The specification reported in the second row of each panel (in column 4), includes year dummies and school fixed effects.[13] The specification reported in the third row of each panel (in column 4) includes the students' characteristics as well as the year dummies and school fixed effects. The control variable coefficients in this model are constraint to have the same coefficients across treatment and control group and over time.

The DID estimates closely resemble the difference in simple mean differences as well as the difference in controlled differences presented in columns 1–3, respectively. They are significant for all outcomes except for the number of science credits, for which the point estimates are

13. The difference-in-differences estimates that are simply the difference between the treatment and control group differences at the two time periods (the mean of 1994 minus the mean of 1992 and 1993) are presented in the online Appendix Table A2. These estimates are generally lower than the difference and difference estimates that are obtained from the regressions that include school fixed effects and are presented in the second row of each panel of Table 3.

TABLE 3

Simple, controlled and difference-in-differences estimated effects of school choice: Tel-Aviv district 9 vs. Tel-Aviv districts 6–8, and Tel-Aviv district 9 vs. Givataim and Ramat-Gan

	Tel-Aviv district 9 vs. Tel-Aviv districts 6–8			DID		Tel-Aviv district 9 vs. Givataim and Ramat Gan			DID	
				True	False				True	False
	1992	1993	1994	92/3 vs. 94	92 vs. 93	1992	1993	1994	92/3 vs. 94	92 vs. 93
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Drop-out										
Control group mean	0.285	0.310	0.269			0.090	0.087	0.099		
Treatment–control difference	−0.104	−0.117	−0.141	−0.056	−0.022	0.089	0.107	0.028	−0.068	0.013
	(0.040)	(0.029)	(0.039)	(0.028)	(0.032)	(0.033)	(0.024)	(0.026)	(0.016)	(0.019)
Treatment–control controlled difference	−0.066	−0.099	−0.122	−0.058	−0.028	0.043	0.051	−0.009	−0.080	0.019
	(0.023)	(0.029)	(0.033)	(0.026)	(0.030)	(0.028)	(0.021)	(0.017)	(0.015)	(0.018)
Eligible for Bagrut										
Control group mean	0.283	0.305	0.319			0.605	0.635	0.621		
Treatment–control difference	0.153	0.141	0.201	0.073	−0.007	−0.167	−0.189	−0.101	0.073	−0.015
	(0.069)	(0.054)	(0.054)	(0.032)	(0.036)	(0.063)	(0.043)	(0.045)	(0.025)	(0.029)
Treatment–control controlled difference	0.098	0.107	0.171	0.062	−0.008	0.013	−0.047	−0.005	0.054	−0.043
	(0.042)	(0.038)	(0.042)	(0.031)	(0.035)	(0.051)	(0.025)	(0.030)	(0.024)	(0.028)
Average score										
Control group mean	49.628	49.972	52.310			75.595	76.816	75.934		
Treatment–control difference	10.216	10.104	14.389	6.575	0.340	−15.600	−16.761	−9.235	6.729	−0.540
	(4.125)	(3.960)	(3.967)	(2.403)	(2.768)	(3.943)	(3.004)	(3.467)	(1.598)	(1.861)
Treatment–control controlled difference	6.123	7.195	11.859	6.444	0.309	−4.965	−7.639	−1.732	6.797	−2.328
	(2.119)	(3.055)	(2.785)	(2.265)	(2.573)	(2.780)	(1.951)	(2.016)	(1.465)	(1.706)
Number of Bagrut credits										
Control group mean	12.870	12.194	13.222			20.001	20.762	20.305		
Treatment–control difference	2.801	3.612	4.352	1.763	1.049	−4.293	−4.967	−2.731	1.866	−0.520
	(1.238)	(1.310)	(1.146)	(0.742)	(0.847)	(1.135)	(0.983)	(1.018)	(0.519)	(0.602)
Treatment–control controlled difference	1.578	2.903	3.796	1.649	1.050	−1.005	−1.534	−0.522	1.714	−0.842
	(0.723)	(0.980)	(0.905)	(0.709)	(0.794)	(0.823)	(0.617)	(0.768)	(0.490)	(0.566)

Continued

TABLE 3
Continued

	Tel-Aviv district 9 vs. Tel-Aviv districts 6–8			DID		Tel-Aviv district 9 vs. Givataim and Ramat Gan			DID	
	1992	1993	1994	True	False	1992	1993	1994	True	False
Dependent variable				92/3 vs. 94	92 vs. 93				92/3 vs. 94	92 vs. 93
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of science credits										
Control group mean	1.196	0.932	1.195			3.077	3.025	2.887		
Treatment–control difference	0.386	0.560	0.534	0.170	0.245	−1.488	−1.529	−1.157	0.317	0.019
	(0.385)	(0.267)	(0.305)	(0.212)	(0.237)	(0.392)	(0.262)	(0.279)	(0.200)	(0.230)
Treatment–control controlled difference	0.071	0.429	0.519	0.177	0.253	−0.262	−0.585	−0.275	0.301	−0.182
	(0.235)	(0.189)	(0.183)	(0.207)	(0.230)	(0.245)	(0.184)	(0.158)	(0.193)	(0.224)
Number of honours level subjects										
Control group mean	1.220	1.083	1.206			2.265	2.377	2.342		
Treatment–control difference	0.278	0.469	0.569	0.258	0.227	−0.764	−0.827	−0.567	0.224	−0.049
	(0.176)	(0.184)	(0.166)	(0.096)	(0.108)	(0.164)	(0.150)	(0.150)	(0.077)	(0.088)
Treatment–control controlled difference	0.100	0.354	0.493	0.227	0.227	−0.244	−0.290	−0.177	0.186	−0.096
	(0.104)	(0.136)	(0.119)	(0.092)	(0.101)	(0.106)	(0.097)	(0.101)	(0.073)	(0.083)
Number of students										
Control group	601	532	558	1,691	1133	1,815	1,863	1,842	5,520	3,678
Treatment group	761	829	791	2,381	1590	758	827	791	2376	1585

Notes: Standard errors in parentheses are adjusted for (primary) school level clustering in the level estimates for each year. In the difference in difference specifications, the outcome in question is regressed on treatment interacted with a dummy for 1994 (this is the coefficient reported), and other controls: year and primary-school fixed effects, levels of maternal and paternal education, number of siblings, gender, immigrant status, and ethnicity. The sample is limited to schools that appear both before and after treatment in each of the difference in difference subsamples.

positive but imprecise. For example, they suggest that the programme led to a 5.8% point decline in the drop-out rate, to a 6.2% point increase in the matriculation rate and to a 6.4-point increase in the average score. They also show a significant and sharp increase in the number of honours level subjects. It is important to note that these estimated effects are not sensitive at all to whether the regressions include student-level individual controls. This is probably due to the absorption by the school fixed effects of most of the variance among students in their background characteristics.[14]

Two of the outcomes, drop-out rate and eligibility for Bagrut, are binary dependent variables. Therefore, for these outcomes I estimated the various models of Table 3 using probit models instead of the linear probability model. The marginal effects based on the probit estimates (evaluated at the mean of the variables) are presented in online Appendix Table A3. These estimates are very similar to those presented in the first two panels of Table 3, suggesting that the treatment effect estimates are robust to the probit specification. For example, in the simple difference in differences estimates in Table A3, column 4 (second row), the probit estimate of the effect on the matriculation rate is 0.074 while the respective estimate in Table 3 is 0.073. The respective estimates for the controlled DID are 0.066 and 0.062. Similar comparability is seen between the probit and the linear probability estimates of the effect on the drop-out rate.

Column 5 in Table 3 presents DID treatment estimates based on the 1992 and 1993 cohorts. As explained above, this controlled experiment is a type of falsification test for spurious treatment effects, because neither cohort was exposed to treatment, and for spillover effects on the baseline cohort. Indeed, all the estimates in this column are small and not statistically different from zero. For example, the estimate for the matriculation rate is -0.008 (s.e. $= 0.035$) vs. 0.062 (s.e. $= 0.031$) for the respective estimate in column 4. The evidence that the pre-treatment cohort was unaffected is also reassuring confirmation that the choice programme did not have spillover effects on older pupils who did not have the choice option.

The estimated effect on drop-out out rate is larger than the estimated effect on all other outcomes. The former effect is about 35% of the pre-programme drop-out rate in district 9 while the effect on the other outcomes is about 10–12% of the pre-programme mean of these outcomes in district 9. Except for the drop-out rate, these effect sizes are actually smaller or similar than the effect of recent high-school interventions in Israel, for example, remedial education (Lavy and Schlosser, 2005), students' monetary incentives (Angrist and Lavy, 2009) or teachers' incentives (Lavy, 2009). It could very well be that the effect on drop-out rate is much larger because of the city's new policy (adopted at the same time as school choice) of forbidding schools to force pupils to continue their schooling elsewhere on the grounds of poor academic ability. This new concept of "persistence" (described in detail in Section 2) which has lowered the mobility rate of students during secondary schooling could be responsible as well for some of the decline in the drop-out rate.

---

14. I should note that peer effects or competitive effects on school productivity can also affect the cohort immediately prior to treatment. However, this is very unlikely given that all the DID estimated effect is due to an increase in the mean outcomes of the affected cohort without any change in the mean outcomes of the cohort immediately prior. Second, as I have noted above, the estimates based on the cohort 2 years prior (at t−2) to the programme are completely identical to those based on using the cohort immediately prior as a control group. This result is reassuring as I expect spillover over effects should be much weaker as I use older cohorts. Furthermore, as shown below the falsification tests using prior cohorts seem to indicate that the pre-treatment cohort was unaffected.

5.2. *Evidence based on adjacent cities as a comparison group*

Columns 6–10 of Table 3 replicate the regression estimates presented in columns 1–5, using GR as a comparison group. The estimates presented in columns 6–7 show that district 9 students have much lower high-school outcomes than GR students before the programme started. The outcome levels and treatment-control differences are remarkably similar in both years. For example, the mean matriculation rates in GR in 1992 and 1993 (60.3 and 63.6%, respectively) exceed those of district 9 by 40%. Similar differences are observed in the other outcomes presented in the table. However, controlling for students' characteristics greatly reduces these baseline differences. The treatment–control conditional mean difference in the drop-out rate in 1992, for example, is 4.4% as against a simple mean difference of 8.9%. The corresponding matriculation rate unconditional difference is −16.5% while the respective conditional difference is 1.5%. This pattern recurs in all six outcomes, suggesting that much of the observed outcome differences are explained by observed differences in characteristics.

Comparing the simple treatment–control mean differences and the conditional differences of the 1994 cohort (column 8) with those of the two pre-programme cohorts, reveals a large relative improvement in district 9 students' outcomes. The magnitude of improvement implied by the comparison of the simple differences is very similar to that based on the conditional differences. The DID estimates based on use of these cross-sections, again do an even better job of demonstrating this important similarity and provide a concise summary of these results. Column 9 presents DID estimates when all three cross-sections are used. Here, again I also attempt versions of DID estimation with only 1992 or only 1993 as a baseline and the results are not sensitive at all to any such variation.[15] The DID estimates are significant for all outcomes except for the number of science credits, for which the point estimates are positive but imprecise. The remarkable result is that these estimates are identical to those based on using districts 6–8 as a comparison group (and presented in column 4 of Table 3). For example, they suggest that the programme led to a 5.6% point increase in the matriculation rate vs. the 6.2% improvement reported in column 4 for the same outcome. The two respective estimates for the improvement in the average score are 7.218 and 6.444. Recall that district 9 had better pre-programme mean outcomes than districts 6–8 but much worse pre-programme mean outcomes than GR, yet identical programme effects based on these two alternative comparison groups are obtained. Thus, the DID estimates reported in column 9 support the interpretation of the results reported in column 4 as causally related to treatment.

Column 10 in Table 3 presents the falsification test of treatment estimates based on the 1992 and 1993 cohorts. All these estimates are small and not statistically different from zero. The fact that the estimates in column 10 have opposite signs compared to those in column 9 may raise concern of reversion to the mean, but the estimates in column 10 are much smaller than those in column 9 while their standard errors are similar. Moreover, close scrutiny of the levels reported in columns 6–8 reveals that the reversal of signs in column 10 is driven almost exclusively by minor fluctuations in GR, and not in Tel-Aviv, which then recede the following year. In contrast, the DID estimates in column 9 are driven predominantly by changes in Tel-Aviv that can be attributed to treatment. The evidence reported in column 4 also reduce the

---

15. It is important to note again that these estimated effects are not sensitive at all to whether the regressions include student-level individual controls. This is probably due to the absorption by the school fixed effects of most of the variance among students in their background characteristics. On average, 6% of students in district 9 and a similar rate of those in GR enrolled in schools outside of their city of residence. Excluding these students from the sample does not change the results at all.

likelihood of mean reversion because the pre-programme mean of outcomes in the districts 6–8 were lower than those of district 9 while in GR they were higher.

To assess once more how sensitive the results are to the use of GR as a comparison group, I replace it with a respective sample of students from Holon. The results are presented in online Appendix Table A4 (balancing tests) and Table A5 (estimates of effect on outcomes). Remarkably, the programme effect estimates in column 4 of Table A5 are very similar to those presented in columns 4 and 9 of Table 3 and the falsification estimates in column 5 are all not significantly different from zero or have the opposite sign.

The similarity in results obtained based on three different comparison groups is also useful in addressing the concern that the impact estimates do not represent the effect of the choice programme but rather the effect of the change of secondary schools to six grade structures. Although only four secondary schools had to be modified so as to comply with the new structure, we cannot rule out *a priori* that it does not account for some of the average change in outcomes in district 9 that followed the choice programme. However, we should note that since the change in structure of secondary schools was implemented in all the city's schools at the same time, it is enough to assume that the consolidation policy (of middle and high schools grades in one school) had identical additive effects on district 9 and on districts 6–8. However, the evidence that the programme effect estimates based on the neighbouring cities as a comparison group (GR or Holon) are identical to those based on districts 6–8 as a comparison group suggests that the assumption of additivity is redundant because both GR and Holon did not make any change in the structure of their secondary schools during the period of analysis.

5.3. *Evidence based on the natural-experiment samples*

Columns 1–2 of Table 4 present estimates based on the cross-section results of the 1993 and 1994 cohorts, based on the respective samples that correspond to the 250-m bandwidth. The corresponding difference-in-differences estimates are presented in column 3.[16] Column 1 shows that in 1993 there are no systematic treatment–control differences in mean outcomes; the observed differences have different signs across outcomes and they are not statistically significant, especially after conditioning on covariates. This close resemblance in pre-programme outcomes parallels the almost identical means of students' characteristics of the two groups shown in Table 2. However, the 1994 differences show an advantage for the treatment group that is statistically significant in some cases. The DID estimates also indicate positive programme effects that are significant for four of the six outcomes. For example, the DID estimates suggest that the matriculation rate went up by 13.8% and that the average score increased by 9.3 points, both changes being statistically different from zero at the 5% level of significance. The effect on the drop-out rate is negative and on the number of science units it is positive, but these two effects are not precisely measured. This could be due to the fact that the samples within the band are too small to pick up the effects with precision. In an attempt to enlarge the sample, the estimates presented in columns 4–6 are based on a bandwidth of 500 metres. It can be seen from Table 2 that even after such an enlargement of the band the treated and control samples are very balanced in characteristics and therefore I can be quite confident that this resemblance still reflects similarity in unobserved characteristics. The treatment effect estimates based on this larger sample are much more precise and are also very similar to those reported in Table 3. For example, the effect on the drop-out rate is −0.064 vs. −0.080 in the

---

16. As noted earlier, missing data on exact addresses for GR in 1992 do not permit a similar analysis for 1992.

TABLE 4

Simple, controlled and difference-in-differences estimated effects of school choice: evidence based on the natural experiment samples

Dependent variable	Up to 250 m		DID	Up to 500 m		DID
	1993 (1)	1994 (2)	93 vs. 94 (3)	1993 (4)	1994 (5)	93 vs. 94 (6)
Drop-out						
Control group mean	0·065	0·086	.	0·065	0·091	. . .
Treatment−control difference	0·076	0·024	−0·058	0·090	0·028	−0·070
	(0·049)	(0·046)	(0·042)	(0·034)	(0·036)	(0·030)
Treatment−control *controlled* dif.	0·021	0·007	−0·030	0·051	−0·008	−0·064
	(0·048)	(0·038)	(0·042)	(0·028)	(0·028)	(0·029)
Eligible for Bagrut						
Control group mean	0·605	0·534	. . .	0·638	0·582	. . .
Treatment−control difference	−0·108	0·060	0·173	−0·121	−0·015	0·115
	(0·106)	(0·082)	(0·068)	(0·083)	(0·072)	(0·048)
Treatment−control *controlled* dif.	0·021	0·148	0·138	−0·018	0·077	0·095
	(0·083)	(0·060)	(0·070)	(0·048)	(0·051)	(0·048)
Average score						
Control group mean	78·885	75·604	. . .	78·677	75·723	. . .
Treatment−control difference	−15·111	−2·989	12·075	−14·528	−5·494	9·738
	(6·638)	(5·496)	(4·174)	(4·802)	(5·268)	(2·983)
Treatment−control *controlled* dif.	−6·379	1·648	9·302	−7·813	0·444	8·681
	(4·293)	(3·331)	(4·098)	(2·160)	(3·128)	(2·879)
Number of Bagrut credits						
Control group mean	21·265	19·828	. . .	21·334	20·138	. . .
Treatment−control difference	−3·841	−0·570	3·324	−3·717	−1·557	2·478
	(1·812)	(1·484)	(1·385)	(1·426)	(1·452)	(0·982)
Treatment−control *controlled* dif.	−0·882	0·587	2·373	−1·213	0·138	1·800
	(1·454)	(1·110)	(1·386)	(0·927)	(1·138)	(0·965)
Number of science credits						
Control group mean	3·375	2·741	. . .	3·142	2·894	. . .
Treatment−control difference	−1·522	−0·278	1·178	−1·264	−0·772	0·542
	(0·639)	(0·633)	(0·548)	(0·486)	(0·491)	(0·389)
Treatment−control *controlled* dif.	−0·370	0·246	0·645	−0·498	−0·194	0·309
	(0·388)	(0·342)	(0·542)	(0·320)	(0·266)	(0·374)
Number of honours level subjects						
Control group mean	2·540	2·246	. . .	2·513	2·357	. . .
Treatment−control difference	−0·760	−0·146	0·607	−0·726	−0·371	0·396
	(0·295)	(0·266)	(0·208)	(0·231)	(0·236)	(0·151)
Treatment−control *controlled* dif.	−0·290	0·128	0·482	−0·321	−0·067	0·282
	(0·215)	(0·154)	(0·206)	(0·149)	(0·167)	(0·147)
Number of students						
Control group	200	232	432	464	471	935
Treatment group	177	190	367	329	344	673

Notes: The natural experiment samples contain pupils who reside in tangent neighbourhoods within a 250- or 500-m band on both sides of the city border. Standard errors in parentheses are adjusted for (primary) school level clustering. In the difference-in-difference specification, the outcome in question is regressed on treatment interacted with a dummy for 1994 (this is the coefficient reported), and other controls: year and primary-school fixed effects, levels of maternal and paternal education, number of siblings, gender, immigrant status, and ethnicity. The sample is limited to schools that appear both before and after treatment in the difference-in-differences subsamples.

full GR sample, the effect on credit units is 1.800 vs. 1.714 in the full sample. This similarity is striking given that the natural-experiment sample size is only about 40% of the full sample.

In summary, I present above four alternative sets of estimates of the effect of school choice on students' cognitive achievement, which all exhibit the same direction and magnitude of treatment effect.

6. EFFECT OF SCHOOL CHOICE ON BEHAVIOURAL OUTCOMES

The results reported above show that the school choice programme leads to an improvement in schooling attainment and in the scholastic achievements of students. In this Section I explore whether school choice improves a few behavioural non-cognitive outcomes. Using a rich set of behavioural outcomes of primary and middle-school students I examine whether school choice alters several aspects of the school climate, such as student discipline and violence, student–teacher relationships and student's social acclimation and overall satisfaction at school. Several studies have highlighted the central role of these outcomes in school choice decisions (see, e.g., Hoxby, 1998; Black, 1999; Kane *et al*., 2006; Cullen *et al*., 2006). In addition, regardless of their possible role as desired outcomes, the effects of school choice on these behavioural outcomes are interesting as potential mediating factors or channels of effects on test scores.

6.1. *School "climate": discipline and violence*

Allowing students to attend their preferred school by means of free school choice may also improve school climate by enhancing students' discipline and by lowering classroom violence. To examine this issue, I use unique data consisting of students' responses to a questionnaire concerning the classroom learning environment and their study efforts. The questionnaire is part of a primary and middle-school survey called Growth and Effectiveness Measures for Schools (GEMS) (meizav in Hebrew) from the years 2002–2003. The GEMS includes a series of tests and questionnaires administered by the Division of Evaluation and Measurement of the Ministry of Education.[17] The GEMS is administered at the mid-term of each school year to a representative 1-in-2 sample of all elementary and middle schools in Israel, so that each school participates in GEMS once every 2 years. Student data include the responses of 5th through 9th grade students to questionnaires. In principle, all students except those in special education classes are tested and have questionnaires administered to them. The rate of questionnaire completion is roughly 91%.

The student questionnaire includes 71 questions addressing various aspects of school and the learning environment. I focus on one section of the questionnaire, which addresses issues related to these aspects. In this section students are asked to rate the extent to which they agree with a series of statements on a 6-point scale ranging from "strongly disagree" to "strongly agree". To reduce measurement error and increase precision, I also group the individual items into aggregate categories and average across responses. The student questionnaire data were linked to student administrative records (identical in structure to the data used for high-school students) that include student background characteristics.

I investigate the effect of school choice on students' discipline and violence by focusing on the following four items in the students' questionnaire: "the classroom is frequently noisy and

---

17. The GEMS are not administered for school accountability purposes and only aggregated results at the district level are published. For more information on the GEMS, see the Division of Evaluation and Measurement website (in Hebrew): http://cms.education.gov.il/educationcms/units/rama/odotrama/odot.htm.

not conducive to learning". "there are many fights among students in my classroom", "I was involved in violence many times this year (physical fights)", and "sometimes I am scared to go to school because there are violent students".[18] The empirical strategy for identifying the effect of choice on these four items appeals once again to the fact that exposure to the choice programme varied between Tel-Aviv and its neighbouring cities and that in Tel-Aviv only middle-school students could choose their school while primary school students had to attend their neighbourhood school. In other words, I contrast the difference in response of students in primary schools and middle schools in Tel-Aviv to the respective difference in primary and middle-school responses in the neighbouring cities of GR and Holon. This approach amounts to a difference-in-differences identification strategy where the treatment variable is an indicator of being a middle-school student in Tel-Aviv. Since by 2002 the choice programme was already implemented in the whole city I include all Tel-Aviv students in the sample and not just those from district 9. More formally I estimate the following equation:

$$y_{igct} = \mu_g + \lambda_c + \pi_t + X_{igct}\beta + Z_{gc}d + \varepsilon_{igct}. \tag{3}$$

The dependent variable, y_{igct}, is the response of student i in grade g, city c and period t to an item in the questionnaire; π_t are year effects (i.e., 2002 and 2003) and X_{igct} is defined as a vector of the student's characteristics; Z_{gc} is the treatment indicator, which equals 1 for Tel-Aviv's middle-school students, and d is the treatment effect. I also included grade (5th–9th) effects, μ_g, and city fixed effects, λ_c, in the equation. In addition, I estimate a version of equation (3) based on separate samples of primary and middle schools in order to measure the first differences (between Tel-Aviv and GR and Holon) that are used to construct the second difference.

An interesting aspect of the Tel-Aviv choice programme is that it allows testing the validity of this difference-in-differences (or fixed effect) identification strategy. Students in 5th and 6th grade were not exposed to the programme so variation in the response to the questionnaire items across these grades should not differ systematically across cities. Similarly, students in 7th–9th grade in Tel-Aviv were all exposed to the choice programme so again any variation between these grades should not differ between Tel-Aviv and its neighbouring cities. These controlled experiments serve as falsification tests of spurious treatment effect estimates, by exploiting the presence of multiple control groups formed by successive cohorts that were not exposed to the choice programme.

Panel A of Table 5 reports results for the four student questionnaire items related to discipline and violence as well as for the mean of these items. Columns 1–2 present the mean response of 5th–6th grade students in GR/Holon and in Tel-Aviv, respectively, and columns 3–4 the respective means of 7th–9th grade students. Columns 5–6 present the DID estimates and their standard errors based on estimating equation (3). They suggest that school choice significantly reduces the level of disruption and violence. For example, the estimate on "noisy classroom" is −0.156 (s.e. = 0.083) and on "many fights among students" is −0.216 (s.e. = 0.120). Note that the question that relates to the student's own violent behaviour is also negative and significant (estimate = −0.152, s.e. = 0.054) which strengthens the conclusion that school choice alters the behaviour of students, making them less disruptive and less violent in class. No effect is found for the perceived sense of safety in school although the average effect on all four items is still positive and significant. The reduced levels of violence and classroom disruption not only imply a better learning environment, but might also impact students' achievements indirectly by improving their motivation, concentration and other non-cognitive factors that are important for learning.

18. In constructing the mean, all variables are transformed so that high values indicate a less violent environment.

TABLE 5

Difference-in-differences estimates of the effect of school choice on school violence and discipline: Tel-Aviv vs. Givataim and Ramat-Gan (GR) and Holon

Questionnaire item	Before "treatment"		After "treatment"		Controlled DID	
	GR and Holon (1)	Tel-Aviv (2)	GR and Holon (3)	Tel-Aviv (4)	Estimate (5)	s.e. (6)
Panel A: True DID						
	5th–6th grades		7th–9th grades			
The classroom is frequently noisy and not conducive to learning	4·804	4·906	4·924	4·867	−0·156	(0·083)
There are many fights among students in my class	3·617	3·711	3·417	3·273	−0·216	(0·120)
I was involved in violent incidents many times this year (physical fights)	1·888	2·026	1·772	1·742	−0·152	(0·054)
Sometimes I am scared to go to school because there are violent students	1·841	1·788	1·602	1·552	0·001	(0·066)
Mean of violence related questions	3·958	3·885	4·068	4·141	0·137	(0·066)
Panel B: DID falsification I						
	5th grade		6th grade			
The classroom is frequently noisy and not conducive to learning	4·771	4·901	4·835	4·912	−0·045	(0·069)
There are many fights among students in my class	3·758	3·791	3·479	3·632	0·142	(0·112)
I was involved in violent incidents many times this year (physical fights)	1·912	2·086	1·865	1·967	−0·049	(0·069)
Sometimes I am scared to go to school because there are violent students	1·970	1·957	1·716	1·623	−0·052	(0·073)
Mean of violence related questions	3·895	3·807	4·020	3·962	0·010	(0·059)
Panel C: DID falsification II						
	7th grade		8th–9th grades			
The classroom is frequently noisy and not conducive to learning	4·927	4·813	4·921	4·926	0·115	(0·076)
There are many fights among students in my class	3·449	3·204	3·383	3·348	0·213	(0·148)
I was involved in violent incidents many times this year (physical fights)	1·841	1·732	1·699	1·751	0·178	(0·065)
Sometimes I am scared to go to school because there are violent students	1·625	1·593	1·577	1·507	−0·035	(0·065)
Mean of violence related questions	4·036	4·163	4·102	4·117	−0·118	(0·075)

Notes: The sample includes meizav responders from the years 2002 and 2003. Regressions include grade effects, city effects and a common time trend. In addition, controls are included for parental education, gender, number of siblings, immigrant status, and ethnic origin. Standard errors are clustered at the school level.

Panels B–C in Table 5 present DID falsification estimates when the contrast between treatment and control is between primary school grades (Panel B) and between middle-school grades (Panel C). All these estimates are very small, some are even negative, and none are significantly different from zero. For example, the mean estimate for all questions

based on 6th vs. 5th grade (panel B) and 8th–9th vs. 7th grade (panel C) are 0.010 and −0.118, respectively, compared with the true treatment effect estimate of 0.137 (panel A). These falsification or placebo treatment estimates add credibility to the identification based on DID strategy (Panel A).

6.2. *Improved student–teacher relationships*

Five items from the student questionnaire can be used to examine whether school choice makes an impact by altering the relationship between students and teachers. Overall, the results suggest positive treatment effects, even though the estimates are significant for only three of five items. This evidence is presented in Table 6. School choice reduces the incidence of students being rude to teachers (estimate = −0.430, s.e. = 0.172), and of teachers being rude to students (estimate = −0.156, s.e. = 0.094). School choice also positively impacts the students' perception of having access to help from teachers when needed. The effects on "good relationship" and "mutual respect" between students and teachers are positive though they have relatively large standard errors. The average effect on all five items is, nevertheless, positive and significantly different from zero. The falsification estimates presented in Panels B and C are again very small and not statistically significant.

6.3. *Students' social acclimation and satisfaction at school*

Allowing students to attend their preferred school by means of free school choice may also be reflected in better social acclimation of students in school and in their improved overall satisfaction from school. This analysis is based on two items in the questionnaire: "I feel well adjusted socially in my class" and "I am generally well off at school" which can essentially be seen as a measure of the students' opinions about their classroom and school. The estimates in Panel A of Table 7 show that school choice has a positive effect on social acclimation as seen for both items. The estimates on "social adjustment" and on "feeling comfortable" are 0.112 and 0.295, respectively, and both are significantly different from zero. The effect on the average of these items is 0.205 (s.e. = 0.062). It is important to note that these positive DID estimates reflect positive differences among middle-school students between Tel-Aviv and GR/Holon while among primary school students (5th–6th grade) there are slight differences in favour of GR/Holon. This strengthens the interpretation of the DID estimates as a causal effect of the choice programme in Tel-Aviv.

Panels B–C in Table 5 present DID estimates when the contrast between treatment and control is between primary school grades (Panel B) and between middle-school grades (Panel C). Unlike the estimates presented in Panel A of this table, these falsification treatment estimates are again very small and not statistically significant.

6.4. *More evidence on effect on behavioural outcomes from student fixed effect estimates*

The structure of the GEMS allows following a sample of students from elementary schools (at 5th or 6th grade in 2002, 2003) to middle schools (at grade 7th, 8th, or 9th in 2004 or 2005).[19] I take advantage of this feature and construct a longitudinal dataset at the student

---

19. Specifically, I follow students who were in 5th or 6th grade in 2002 or 2003 to their respective grades in 2004 and 2005. I am therefore able to follow about 50% of the sample of elementary school students reported in columns 1–4 of Tables 5–7, finding roughly 80% of them. I did not link between datasets from consecutive years because almost all localities were sampled once every 2 years.

TABLE 6

Difference-in-differences estimates of the effect of: school choice on student–teacher relationships: Tel-Aviv vs. Givataim and Ramat-Gan (GR) and Holon

Questionnaire item	Before "treatment"		After "treatment"		Controlled DID	
	GR and Holon (1)	Tel-Aviv (2)	GR and Holon (3)	Tel-Aviv (4)	Estimate (5)	s.e. (6)
Panel A: True DID						
	5th–6th grades		7th–9th grades			
Students are frequently rude to the teachers	3·860	4·117	4·332	4·153	−0·430	(0·172)
Sometimes the teachers treat me in an insulting or hurtful way	2·844	2·957	3·152	3·103	−0·156	(0·094)
There are good relationships between the teachers and the students	4·451	4·320	3·764	3·760	0·116	(0·085)
There is mutual respect between the teachers and the students	4·435	4·313	3·745	3·749	0·110	(0·093)
When I have a problem I have someone to turn to at school (teachers, advisor)	4·908	4·771	4·290	4·376	0·205	(0·084)
Mean of student–teacher relationship questions	4·214	4·059	3·663	3·725	0·206	(0·083)
Panel B: DID falsification I						
	5th grade		6th grade			
Students are frequently rude to the teachers	3·749	4·025	3·968	4·206	−0·013	(0·101)
Sometimes the teachers treat me in an insulting or hurtful way	2·814	2·916	2·874	2·997	0·051	(0·086)
There are good relationships between the teachers and the students	4·599	4·385	4·306	4·256	0·167	(0·090)
There is mutual respect between the teachers and the students	4·557	4·390	4·316	4·238	0·092	(0·078)
When I have a problem I have someone to turn to at school (teachers, advisor)	5·046	4·903	4·774	4·644	0·018	(0·085)
Mean of student–teacher relationship questions	4·324	4·138	4·107	3·982	0·052	(0·068)
Panel C: DID falsification II						
	7th grade		8th–9th grades			
Students are frequently rude to the teachers	4·257	4·049	4·413	4·266	0·029	(0·111)
Sometimes the teachers treat me in an insulting or hurtful way	3·059	2·964	3·252	3·253	0·097	(0·125)
There are good relationships between the teachers and the students	3·871	3·851	3·648	3·662	0·035	(0·118)
There is mutual respect between the teachers and the students	3·891	3·854	3·589	3·635	0·087	(0·120)
When I have a problem I have someone to turn to at school (teachers, advisor)	4·407	4·522	4·166	4·218	−0·085	(0·126)
Mean of student–teacher relationship questions	3·771	3·841	3·548	3·599	−0·014	(0·099)

Notes: The sample includes meizav responders from the years 2002 and 2003. Regressions include grade effects, city effects, and a common time trend. In addition, controls are included for parental education, gender, number of siblings, immigrant status and ethnic origin. Standard errors are clustered at the school level.

level to examine how changes in students' assessments of their classroom environment and of their own behaviour are associated with the change to school choice (due to their transition

TABLE 7

Difference-in-differences estimates of the effect of school choice on students' social acclimation: Tel-Aviv vs. Givataim and Ramat-Gan (GR) and Holon

Questionnaire item	Before "treatment"		After "treatment"		Controlled DID	
	GR and Holon (1)	Tel-Aviv (2)	GR and Holon (3)	Tel-Aviv (4)	Estimate (5)	s.e. (6)
Panel A: True DID						
	5th–6th grades		7th–9th grades			
I feel well adjusted socially in my class	5·180	5·152	5·038	5·121	0·112	(0.055)
I am generally comfortable at school	5·198	5·084	4·764	4·956	0·295	(0.081)
Mean of inter-student relationship questions	5·189	5·114	4·901	5·038	0·205	(0.062)
Panel B: DID falsification I						
	5th grade		6th grade			
I feel well adjusted socially in my class	5·211	5·178	5·149	5·127	0·017	(0.062)
I am generally comfortable at school	5·255	5·140	5·142	5·030	−0·001	(0.074)
Mean of inter-student relationship questions	5·233	5·154	5·145	5·075	0·009	(0.063)
Panel C: DID falsification II						
	7th grade		8th–9th grades			
I feel well adjusted socially in my class	5·061	5·147	5·014	5·094	−0·014	(0.081)
I am generally comfortable at school	4·880	5·057	4·640	4·846	0·015	(0.096)
Mean of inter-student relationship questions.	4·970	5·100	4·828	4·970	0·002	(0.083)

Notes: Sample includes meizav responders from the years 2002 and 2003. Regressions include grade effects, city effects, and a common time trend. In addition, controls are included for parental education, gender, immigrant status and ethnic origin. Standard errors are robust and clustered at the school level.

from elementary school to middle school).[20] I estimate the following first difference equation by differencing out two relationships like equation (3) for each student (one for middle school and one for elementary schools):

$$y_{igct}^{ms} - y_{igct}^{ps} = \alpha_g^{ps} + \lambda_g^{ms} + \pi_t + x_{igct}\beta + z_{igct}d + \Delta\varepsilon_{igct} \qquad (4)$$

where *ps* denotes primary school and *ms* denotes middle school. A student fixed effect is differenced out from this equation and controls for student's background characteristics and a primary school fixed effect are added. The results are virtually identical when these controls are omitted from the regression or when also adding middle-school fixed effects.

Table 8 presents student fixed effects estimates of the effect of school choice on the learning and classroom environment. For example, focusing on the average effects, the estimate for the effect on classroom disruption and violence in the student fixed effects model is 0.127, vs. the respective school fixed effect estimate reported in Table 5 of 0.137. The respective two estimates for the effect on social acclimation of students at school and in their improved overall satisfaction from school are 0.245 (Table 8) and 0.205 (Table 7). This remarkable similarity in the comparison of the school fixed effects and student fixed effects estimates is also seen for

---

20. I cannot perform a similar analysis with test scores because the longitudinal data do not include students that are observed at both 5th and 8th grades, and these are the only grades at which the cognitive tests are administered.

TABLE 8

Pupils' fixed effect estimates of the effect of school choice on behavioural outcomes: Tel-Aviv vs. Givataim and Ramat-Gan (GR) and Holon

Questionnaire item	Before "treatment"		After "treatment"		DID	
	GR and Holon (1)	Tel-Aviv (2)	GR and Holon (3)	Tel-Aviv (4)	Estimate (5)	s.e. (6)
Panel A: Students' social acclimation						
	5th–6th grades		7th–9th grades			
I feel well adjusted socially in my class	5·536	4·201	5·428	4·206	0·112	(0·089)
I am generally comfortable at school	5·467	4·322	5·035	4·251	0·361	(0·094)
Mean of inter-student relationship questions	5·500	4·255	5·230	4·231	0·245	(0·077)
Panel B: School violence						
	5th–6th grades		7th–9th grades			
The classroom is frequently noisy and not conducive to learning	4·373	5·890	4·484	5·819	−0·183	(0·093)
There are many fights among students in my class	3·715	3·592	3·439	3·100	−0·216	(0·112)
I was involved in violent incidents many times this year (physical fights)	1·667	2·629	1·481	2·305	−0·139	(0·092)
Sometimes I am scared to go to school because there are violent students	1·643	2·201	1·492	2·111	0·061	(0·094)
Mean of violence related questions	4·146	3·413	4·274	3·667	0·127	(0·061)
Panel C: Student–teacher relationships						
	5th–6th grades		7th–9th grades			
Students are frequently rude to the teachers	3·806	4·258	4·255	4·303	−0·403	(0·111)
Sometimes the teachers treat me in an insulting or hurtful way	2·682	3·262	3·022	3·502	−0·101	(0·116)
There are good relationships between the teachers and the students	4·443	4·248	3·698	3·780	0·278	(0·099)
There is mutual respect between the teachers and the students	4·469	4·119	3·725	3·674	0·300	(0·100)
When I have a problem I have someone to turn to at school (teachers, advisor)	4·872	4·685	4·318	4·596	0·466	(0·113)
Mean of student–teacher relationship questions	4·256	3·899	3·690	3·649	0·316	(0·069)

Notes: Clustered standard errors (by primary school) are reported in parentheses. The difference-in-differences specification is uncontrolled, and includes individual fixed effects.

the student–teacher relationship questions; the respective two estimates for the average effect are 0.316 (Table 8) and 0.206 (Table 6). I view these results as supporting the informational content of the survey data as valid measures of the behavioural outcomes I explore and as additional supporting evidence for the causal interpretation of the effects of the school choice programme on the classroom environment.

7. CONCLUSIONS

This paper examines the effect of a programme that replaces zoning and forced bussing with a degree of unrestricted school choice as the determinant of students' enrolment in public high

schools. The first part of the paper focuses on the general effect of choice of public schools on cognitive outcomes at the end of high school. The results suggest that school choice leads to large improvements in attainment and in the quality of the curriculum of study, as reflected in a reduced drop-out rate, a higher matriculation rate and a higher mean score on the matriculation exams. The improvements are also noticeable in quality measures of the programme of studies, such as credits in science subjects and the mean number of subjects studied at the honours level.

The results are robust to a variety of estimation strategies and samples. The positive effect of school choice in TASCP is in line with Rouse (1998) who finds, based both on a quasi-experimental control group and a randomized comparison group, that students in the Milwaukee Parental Choice Program had faster maths score gains but similar reading score gains. However, the Milwaukee programme provided vouchers to low-income students to attend private schools while the Tel-Aviv programme allowed choice among public schools, both in and out of district. The positive effect of choice found in Tel-Aviv is also consistent with evidence presented in Hoxby (2000) of the positive effect of Tiebout choice options on test scores and negative effects on spending, and in Bayer and McMillan (2005) of strong performance responses of schools to competitiveness in their local environment. However, it contrasts with Rothstein (2006), who finds little effect of choice on students' academic achievements and no evidence that school choice improves average school effectiveness, and with Cullen *et al*. (2006) who find that the Chicago Public School Choice Program had no effect on the academic achievements of affected students.

The second part of the paper investigates the effects of school choice on pupils' behavioural outcomes based on a unique survey data on the school and classroom environment. School choice led to lower levels of classroom disruption and violence, to better student–teacher relationships and to better social acclimation of pupils in school. These results are similar to the evidence that school choice in Chicago reduced crime and violence of students (Cullen *et al*., 2006). This evidence is also consistent with recent findings that suggest that charter schools in the United States, which provide a form of school choice, improve student discipline and attendance (Imberman, forthcoming) and with evidence that discipline and safety drive many parents' decisions to enrol their children in charter schools (Weiher and Tedin, 2002).

In many respects, these findings are of general interest far beyond the local Israeli context, because the benchmark policy is not unusual and because the Tel-Aviv reform unambiguously changed the amount of choice for a set of households. The reform structure bears great resemblance to many recent programmes in US schooling districts where free choice has replaced, under court order, zoning and cross-district busing. Furthermore, the lessons learned from an analysis of the general effects of choice, both on academic and behavioural outcomes, are also increasingly relevant to European and to some developing countries that have implemented comprehensive, system wide school choice programmes.

REFERENCES

AHLIN, A. (2003), "Does School Competition Matter? Effects of a Large-Scale Choice Reform on Student Performance" (Mimeo, Uppsala University).

ANGRIST, J., BETTINGER, E., BLOOM, E., KING, E. and KREMER, M. (2002), "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment", *American Economic Review*, **92**, 1535–1558.

ANGRIST, J., BETTINGER, E. and KREMER, M. (2006), "Long-Term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia", *American Economic Review*, **96**, 847–862.

ANGRIST, J. and LAVY, V. (2009), "The Effect of High-Stakes High School Achievement Awards: Evidence from a Randomized Trial", *American Economic Review*, **99**, 1384–1414.

BAYER, P. and McMILLAN, R. (2005), "Choice and Competition in Local Education Markets" (NBER Working Paper No. 11802).

BLACK, S. (1999), "Do Better Schools Matter? Parental Valuation of Elementary Education", *Quarterly Journal of Economics,* **114**, 577–599.

BOYD, D., LAUKFORD, H., LOEB, S. and WYCKOFF, J. (2003), "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets" (NBER Working Paper No. 9878).

BRADLEY, S., JOHNES, G. and MILLINGTON, J. (2001), "The Effect of Competition on the Efficiency of Secondary Schools in England", *European Journal of Operational Research,* **135**, 545–568.

CHANG-TAI, H. and URQUIOLA, M. (2003), "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program" (NBER Working Paper No. 10008).

CULLEN, J. B., JACOB, B. A. and LEVITT, S. D. (2005), "The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools", *Journal of Public Economics,* **89** (5–6), 729–760.

CULLEN, J. B., JACOB, B. A. and LEVITT, S. D. (2006), "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries", *Econometrica,* **74** (5), 1191–1230.

ESTHER, D. (2001), "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *American Economic Review,* **91** (4), 795–813.

FISKE, E. and LADD, H. F. (2000), *When Schools Compete: A Cautionary Tale* (Brookings Institution Press).

SEBASTIAN, G., GERTLER P. and SCHARGRODSKY, E. (2005), "Water for Life: The Impact of the Privatization of Water Supply on Child Mortality", *Journal of Political Economy,* **113**, 83–120.

GIBBONS, S., MACHIN S. and SILVA, O. (2008), "Choice, Competition and Pupil Achievement", *Journal of the European Economic Association,* **6** (4), 912–947.

HANUSHEK, E. A., KAIN, J. F. and RIVKIN, S. G. (2004), "Why Public Schools Lose Teachers", *Journal of Human Resources,* **39** (2), 326–354.

HASTINGS, J., KANE, T. J. and STAIGER, D. O. (2005), "Parental Preferences and School Competition: Evidence from a Public School Choice Program" (NBER Working Paper No. 11805).

HASTINGS, J., KANE T. J. and STAIGER, D. O. (2006), "Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery" (National Bureau of Economic Research Working Paper No. 12145).

HECKMAN, J. J. and HOTZ, V. J. (1989), "Choosing Among Alternative Non Experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training", *Journal of the American Statistical Association,* **84** (408), 862–874.

HEIMAN, P. and SHAPIRA, R. (1998), "Restructureing the Tradeoff in an Education System: Choice and Autonomy–The Tel-Aviv Vision" (Tel-Aviv University) (in Hebrew).

HEIMAN, P. and SHAPIRA, R. (2002), "Parental Choice in Autonomous Schools as a Stategy of Restructuring Schooling Systems" (Mimeo, Tel-Aviv University) (in Hebrew).

HOXBY, C. M. (1998), "When Parents Can Choose, What Do They Choose? The Effects of School Choice on Curriculum and Atmosphere", in S. Mayer and P. Peterson (eds.) *When Schools Make a Difference* (Washington, DC: The Brookings Institution Press).

HOXBY, C. M. (2000), "Does Competition among Public Schools Benefit Students and Taxpayers?", *American Economic Review,* **90**, 1209–1238.

HOXBY, C. M. (2002), "School Choice and School Productivity (Or Could School Choice Be a Tide That Lifts all Boats?)" (NBER Working Paper No. 8873).

HSIEH, C. T. and URQUIOLA, M. (2003), "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide Voucher Program"(NBER Working Paper No. 10008).

IMBERMAN, S. A. (forthcoming) "Achievement and Behavior in Charter Schools: Drawing a More Complete Picture", *The Review of Economics and Statistics*.

KANE, T. J., RIEGG, S. K. and STAIGER, D. O. (2006), "School Quality, Neighborhoods and Housing Prices", *American Law and Economics Review,* **8** (2), 183–212.

KRUEGER, A. and ZHU, P. (2004), "Another Look at the New York City School Voucher Experiment", *American Behavioral Scientist,* **47** (5), 718–728.

LAVY, V. (2006), "From Forced Busing to Free Choice in Public Schools: Quasi-Experimental Evidence of Individual and General Effects" (NBER Working Paper No. 11969).

LAVY, V. (2009), "Performance Pay and Teachers' Effort, Productivity and Grading Ethics", *American Economic Review,* **99** (5).

LAVY, V. and SCHLOSSER, A. (2005), "Targeted Remedial Education for Under-Performing Teenagers: Costs and Benefits", *Journal of Labor Economics,* **23** (4), 839–874.

LEVY, A., LEVY, K. and LIBMAN, M. (1996 and 1997), "An Evaluation of the Tel-Aviv School Choice Program" (Tel-Aviv University, School of Education) (in Hebrew).

MAYER, D. P., PETERSON, P. E., MYERS, D. E., CLARK TUTTLE, C. and HOWELL, W. G. (2002), *School Choice in New York City after Three Years: An Evaluation of the School Choice Scholarships Program* (Washington, DC: Mathematica).

ROSENBAUM, P. R. (1987), "The Role of a Second Control Group in an Observational Study", *Statistical Science,* **3**, 292–306.

ROTHSTEIN, J. M. (2006), "Good Principals or Good Peers? Parental Valuations of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions", *American Economic Review,* **96** (4), 1333–1350.

ROUSE, C. E. (1998), "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program", *Quarterly Journal of Economics,* **118**, 553–602.

Tel-Aviv Educational Authority (1994), "An Organizational and Educational Program to Develop the Post Primary Schooling in Tel-Aviv" (in Hebrew).

Tel-Aviv Educational Athority (1995), "Memorandum, Summary Statistics from the School Choice Administration, 1994–95", 28 May (in Hebrew).

Tel-Aviv Educational Authority (1996), "Response to a Petition to the High Court of Justice Against the Choice Program" (in Hebrew).

Tel-Aviv Educational Authority (1999), "Evaluation Report of the Choice Program" (in Hebrew).

Tel-Aviv Educational Authority (2001), "Tracking Student Mobility In Tel-Aviv" (in Hebrew).

WEIHER, G. R. and TEDIN, K. L. (2002), "Does Choice Lead to Racially Distinctive Schools? Charter Schools and Household Preferences", *Journal of Policy Analysis and Management,* **21** (1), 79.