The Long-Term Spillover Effects of Changes in the Return to Schooling*

Ran Abramitzky
Stanford University
and NBER

Victor Lavy
Hebrew University,
University of Warwick
and NBER

Santiago Pérez
University of
California, Davis
and NBER

April 2020

Abstract

We study the spillover effects of a pay reform that substantially increased the returns to schooling in kibbutzim, socialist-oriented communities in Israel. In the late 1990s, kibbutzim reformed their decades-long policy of equal income sharing to one of market-based wages. We show that this reform, which induced kibbutz students to improve their academic achievements during high school, spilled over to non-kibbutz members who attended the same schools. In the short run, peers improved their high school outcomes and shifted to courses with higher financial returns. In the long run, peers completed more years of post-secondary schooling and increased their earnings.

* We thank Jaime Arellano-Bover, Alvaro Calderón, Arun Chandrasekhar, Raj Chetty, Giacomo De Giorgi, Nathaniel Hendren, Matt Jackson, Magne Mogstad, Karthik Muralidharan, John Pencavel, Emmanuel Saez, Tom Zohar, Gabriel Zucman, seminar participants at Pontificia Universidad Católica de Chile, Hebrew University, Warwick, and Stanford, and participants at the CEPR Labor Economics Conference at LSE 2016, the All California Labor Economics Conference at Stanford 2017, and GRIPS 2018 conference in Tokyo for useful discussions and suggestions. We thank Hadar Avivi, Elior Cohen, and Nadav Kunievsky for excellent research assistance. Lavy acknowledges financial support from the European Research Council through ERC Advance Grant 323439 and from CAGE.

1. Introduction

We study the short- and long-term spillover effects of a change in the returns to schooling. Starting in the late 1990s, kibbutzim (socialist-oriented communities in Israel) reformed their decades-long policy of equal income sharing to one of market-based wages. In reformed kibbutzim, members were allowed to keep a substantial fraction of their earnings for themselves, substantially increasing the financial returns to schooling. In an earlier study, Abramitzky and Lavy (2014) found that this pay reform led to significant gains in academic achievements of kibbutz high school students. In this paper, we shift attention to the short and long-term spillover effects of this reform, on their high-school peers who live in non-collective communities through post-secondary schooling and into the labor market.[1] The spillover could occur through classroom peer effects, or through the peers learning about the large increases or decreases in kibbutz classmates' family incomes that resulted from the reform.

Our identification strategy takes advantage of the fact that some kibbutzim reformed earlier than others and that some grades (school-cohorts) had students from early reformed kibbutzim and some grades did not. We identify spillover effects using a difference-in-differences approach, comparing the peers of students from kibbutzim that reformed early to the peers of students from kibbutzim that reformed late, before and after the implementation of the early reforms. Our identification assumption is that in the absence of the reforms, the outcomes of the peers of students from early reformed kibbutzim would not have been systematically different than the outcomes of peers of students from kibbutzim that reformed late. We provide evidence that peers of students from kibbutz that reformed early and peers of students from kibbutz that reformed late were similar in their observable characteristics and pre-reform schooling outcomes, both in terms of baseline levels and pre-reform trends.

We start by using administrative records collected by the Israeli Ministry of Education to study the effects of the reform on short-term schooling outcomes of kibbutz peers. We find that peers of early reformers improved their high school performance. The high school completion rate increased by 1.6 percentage points (relative to an already very high baseline of 95.5%), average matriculation exams scores went up by 2.8 points (baseline of 70.9 points) and the matriculation and university qualified matriculation rates increased by 9 and 9.5 percentage points, respectively (baselines of 61 and 58%, respectively). In line with the results in Abramitzky and Lavy (2014), we find that some of these short-run effects are stronger on males and on students from relatively low socioeconomic status.

[1] Such long-term analysis is not feasible for kibbutz members since their earnings were not reported until recently in the administrative data sources.

We then combine the high school records with National Social Security administrative data to examine the spillover effects of the reform on longer-term outcomes (when students were in their early 30s) such as whether they attained post-secondary schooling, their employment status and whether they received unemployment benefits, and their earnings. Treated peers experienced economically meaningful gains in terms of post-secondary schooling attainment. These gains were mainly in university schooling, which requires a matriculation certificate, and not in academic colleges schooling, a lower quality tier of academic institutions in Israel. University enrollment of treated peers increased by 9.5 percentage points and completed years of university schooling increased by 0.5. Moreover, we find an 8 percent increase in annual earnings and a 1.5% decline in the probability of receiving unemployment benefits. These improvements in labor market outcomes are consistent with the higher levels of post-secondary schooling attained by treated peers.

Our findings suggest sizable spillover effects of the pay reforms. Specifically, the size of the spillover effects on high school performance that we document is at least 50% of the size of the direct effects on kibbutz students. Such large spillover effects are comparable with those measured in other contexts (see Online Appendix Table A.1).

Our main contribution is to empirically show that changes in the returns to schooling can influence both directly and indirectly affected students. Standard models of optimal human capital accumulation (Becker 1967, Ben-Porath 1967, Weiss 1995) predict that the returns to education positively affect an individual own's level of investment in schooling. However, if a student's own effort is a complement to other students' effort, increases in the returns to schooling can also have indirect effects. Our paper is the first to show empirically that such spillovers occur in practice, and that they can have both short- and long-term consequences for students.

Our study also contributes to the literature on the spillover effects of social programs. These studies have looked at spillover effects in the context of retirement decisions (Duflo and Saez, 2003), health interventions (Miguel and Kremer, 2004), conditional cash transfers (Angelucci and De Giorgi, 2009; Lalive and Cattaneo, 2009), active employment programs (Crépon, Duflo, Gurgand, Rathelot and Zamora, 2013), program participation (Dahl, Løken, and Mogstad, 2014), mass layoffs (Gathmann, Helm and Schönberg, 2016), and expanding access to college education (Bianchi, 2016).

Closer to our study is the literature on the spillover effects of interventions in the context of school learning. Bobonis and Finan (2009) study the spillover effects of a program that subsidized school enrollment and find sizable spillover effects on ineligible students. Similarly, Alderman, Kim and Orazem (1999) and Kremer, Miguel and Thornton (2009) study programs targeted at improving the schooling outcomes of girls, and show that these programs resulted in sizeable spillovers for boys (who were not eligible). Moreira (2019) studies the spillover effects of having a classmate who received an honorary mention in Brazil's Math Olympiad. Finally, Joensen and Nielsen (2018) look at the spillover effects that older siblings' educational choices have on their younger siblings. While these studies (with the exception of Joensen and Nielsen 2018) focus on short-term spillovers, our paper explores both the short- and long-term gains.

More broadly, the pay reform can be interpreted as a sharp decrease in the marginal tax rate faced by kibbutz's members. While there is a substantial literature studying the effects of such changes on the labor supply of directly affected individuals, there is more limited evidence on their impacts on not explicitly targeted individuals. Hence, from a public economics perspective, this paper sheds light on how redistribution can affect investments in human capital –and hence labor supply decisions- of individuals not explicitly targeted by redistribution.

The rest of the paper is structured as follows. Section 2 presents the background of kibbutzim and the pay reform, and of the Israeli high school system. Section 3 describes the data and sample restrictions. Section 4 presents the empirical framework and identification strategy. Section 5 presents the results on the effect of the reform on high school outcomes. Section 6 presents the results on the long-term effects on post-high school education and labor market outcomes. Section 7 discusses possible mechanisms and concludes.

2. Brief background

a. Kibbutzim and the pay reform

Kibbutzim are voluntary communities that have provided their members with a high degree of income equality for almost a century.[2] Traditionally, all kibbutzim were based on full income sharing between members. Each member of a kibbutz was paid an equal wage, regardless of her economic contribution to the community. Specifically, there were no monetary returns to schooling in the kibbutz, as members earned the same regardless of their education levels.

Unlike American communes, kibbutzim are not isolated from the Israeli society as a whole, and their members are well aware of their outside options (Abramitzky 2011). While kibbutzim never accounted for a large proportion of the Israeli population (currently less than 2%), they have exerted a disproportionate influence on the rest of the Israeli society.[3] Kibbutzim are usually located close to cities and their members often have family outside of the kibbutz. Crucial to our setting,

[2] For a more detailed background on kibbutzim and the pay reform, see Abramitzky (2018).

[3] As described in Abramitzky (2018), "Kibbutzniks were held in high esteem in Israeli society, both before and after the establishment of the state. They had high economic, social, and military status, and had a disproportionate impact on the ideological, political, and military leadership of Israel."

kibbutz-born children typically attend school outside their kibbutz, where they interact with members of other kibbutzim and residents of surrounding villages and towns.

The episode that we study is a pay reform that kibbutzim in Israel adopted beginning in 1998. These pay reforms were a response to changing external pressures and circumstances facing kibbutzim, including a decline in world prices of agricultural goods, bad financial management, and a high-tech boom during the mid-1990s, which increased members' outside options considerably. Most notably, the 1985 stabilization program in Israel following a few years of high inflation, which raised interest rates dramatically and left many kibbutzim with huge debts they could not repay.

In reformed kibbutzim, members' wages reflected market wages so that members were allowed to keep a substantial fraction of their earnings for themselves. For members who worked outside their kibbutzim (about a fourth of all members), market wages were the wages they received from their employers. For members who worked inside, market wages were based on the wages of non-kibbutz workers of similar occupation, education, skills, and experience. A kibbutz 'tax' was deducted from members' gross wages to guarantee older members and low wage earners in the kibbutz a minimum wage.

A survey of three thousand kibbutz members conducted by Pilat Institute in 2004 revealed large wage differences by occupation and education. For example, a director of a kibbutz sector (e.g., the agricultural sector or industry sector) might earn close to 30,000 NIS (about US$8,000 per month), and members in leading positions such as the main secretary (chairman) and the treasurer of the kibbutz earned over 15,000 NIS (about $4,000) per month. Over 80 percent of members holding such positions have academic degrees. In contrast, a member working as a menial laborer in the kitchen or in the laundry, without a post high school academic education, earned less than 4,000 NIS (about $1,000) per month.[4]

The move from equal sharing to differential pay signaled strongly to young adults in the kibbutzim an increase in the financial rewards to human capital. This increase in the return to skills was noticeable within a family, as students' parents experienced a decrease or increase in their earnings depending on their skills. In particular, the reform caused substantial stress in those whose incomes declined after its introduction. For instance, Yuval Albashan, one of the founders of Yedid

[4] A more recent survey in 2009 that included 180 kibbutzim that reformed their pay structures again revealed large pay gaps within kibbutzim. The survey looked only at members who worked inside kibbutzim; it provided data on the monthly wages of 120 different occupations. The highest gross monthly income recorded in the survey was 17,500 NIS ($4,600) and the lowest, 4,100 NIS ($1,080). This range suggests large income inequality, which would most likely be even higher if the wages of the members employed outside the kibbutz were taken into account. This information is provided in the daily newspaper Haaretz [in Hebrew], Sept. 17, 2009, **www.haaretz.co.il/hasite/objects/pages/PrintArticle.jhtml** ?**itemNo=1115205**.

(an Israeli NGO), was quoted saying that in 2008 alone there were 746 requests for help by members in their fifties and sixties whose kibbutz privatized.[5]

Furthermore, the pay reform has been the most discussed topic in kibbutzim since the reforms started. The new productivity-based sharing rules were hotly debated and voted on by members in kibbutzim; booklets elaborating on the reforms were distributed to all members; and the reforms also received substantial attention in the media both in Israel and abroad. The pay reform frustrated many kibbutz members, especially the older generation.[6] Further details on the pay reform are provided in Abramitzky and Lavy (2014).

b. High school and post high school schooling in Israel

Israeli high school students are enrolled either in an academic track leading to a matriculation certificate (*bagrut*) or in an alternative track leading only to a high school diploma. The *bagrut* is completed by passing a series of national exams in core and elective subjects taken by the students between 10th and 12th grade. Thus, *bagrut* certificates are typically obtained at the end of senior year (twelfth grade) or later. Similar high school matriculation exams are found in many countries and in some states in the United States. Examples include the French Baccalaureate, the German Certificate of Maturity, the Italian Diploma di Maturità, and the New York State Regents examinations.

Students choose to be tested at various proficiency levels, with each test awarding one to five credit units per subject, depending on difficulty. Some subjects are mandatory and many must be taken for at least three units. Advanced level subjects are those subjects taken at a level of four or five credit units. A minimum of 20 credit units is required to qualify for a *bagrut* certificate, though some university study programs require more, and students must also satisfy distribution requirements. About 52 percent of all high school seniors received a matriculation certificate in the 1999 and 2000 cohorts (Israel Ministry of Education 2001). Roughly 60 percent of those who took at least one *bagrut* subject test ended up receiving a *bagrut* certificate.

After completing high school, students can decide to continue their studies in various post-secondary schooling institutions. The post high school schooling system in Israel includes seven universities (one of which confers only graduate and PhD degrees), and over 50 colleges that confer

[5] From Arnon Lapidot, an article in the online newspaper ynet, March 12, 2009, http://mynetkibbutz.co.il/article/140474.

[6] For instance, a member of a reformed kibbutz (Gesher Haziv) said, "I had helped pay for their education, and they had much better jobs. Change was inevitable, but it could be a little fairer to everyone all around. I put thirty-two years into this place. I have nothing to show for it. I am a simple grunt in an assembly plant".

academic undergraduate degrees (some of these also give master's degrees).[7] All universities require a *bagrut* diploma for enrollment. Most academic colleges also require a *bagrut*, though some look at specific *bagrut* diploma components without requiring full certification. For a given field of study, it is typically more difficult to be admitted to a university than to a college. Hence, we expect improvements in outcomes related to the *bagrut* to translate into improvements in post-secondary schooling outcomes and, in particular, to university-related outcomes. The national university enrollment rates for the cohort of graduating seniors in 1995 (through 2003) was 27.6 percent and the respective rate for academic colleges was 8.5 percent.[8]

3. Data and Sample Restrictions

a. High school outcomes

The first part of our empirical analysis is based on administrative records collected by the Israeli Ministry of Education. In these records, we observe the schooling outcomes of students starting high school from 1994 to 2000. Each record contains individual level and class identifiers, as well as demographic information on background characteristics of the students. Importantly, the demographic information includes the home address of each student, allowing us to identify which of them resided in a kibbutz by the start of 10th grade, the first year of high school.

We focus on the following schooling outcomes that are available for all the sample years: an indicator for whether the student graduated from high school, the average score in the matriculation exams, an indicator for whether the student received a matriculation certificate (*bagrut*) and an indicator for whether the student received a matriculation certificate that meets university entrance requirements. Because these outcomes are measured at the end of high school, we only observe them once for each student in our sample. About 15 percent of the students in the sample did not take the matriculation exams. These students get zero values in the average matriculation score. The other three high school outcomes that we use - matriculation status, matriculation status that meets university entrance requirements, and the high school completion indicator - do not require such imputation. Finally, note that high school completion was above 95% in our sample of schools in the pre-reform period (and hence there is limited room for improvement along this margin).

---

[7] A 1991 reform sharply increased the supply of postsecondary schooling in Israel by creating publicly funded regional and professional colleges.

[8] These data are from the Israel Central Bureau of Statistics, Report on Post-Secondary Schooling of High School Graduates in 1989–1995 (available at: http://www.cbs.gov.il/publications/h_education02/h_education_h.htm).

To identify students from early and late reformed kibbutzim, we merged the student level data with kibbutz level data collected by the Institute for Research of the Kibbutz and the Cooperative Idea (Getz 1998-2004). These data include several characteristics of kibbutzim, including whether they adopted the pay reform and its date of implementation.

b. Post-high school outcomes

We combine the data on high-school outcomes with annual data on post-secondary schooling and economic outcomes in adulthood. To do so, we link students from their schools to their post-secondary outcomes using administrative data provided by Israel's National Insurance Institute (NII).

In these data, we observe two sets of outcomes for each of the students in our sample. First, we observe post-secondary schooling attainment, including the type of post-secondary schooling institution attended, if any, and the number of years of schooling completed in each type of institution. The post-secondary schooling outcomes of interest are indicators of ever having enrolled in a university or in an academic college, and the number of years of schooling completed in these two types of academic institutions. Even after accounting for compulsory military service, we expect most students who enrolled in post-high school education, including those who continued schooling beyond undergraduate studies, to have graduated by age 30.

Second, we observe year-by-year labor market outcomes from high school graduation to 2014, including employment status, information on unemployment benefits and annual earnings in the formal sector. Individual earnings data come from the Israel Tax Authority (ITA). Filing tax forms in Israel is compulsory only for individuals with non-zero self-employment earnings but ITA has information on annual gross earnings from salaried and non-salaried employment, which they transfer annually to NII, including number of months of work in a given year. Using these data, NII produces an annual series of total annual earnings from salaried work and self-employment. Following NII practice, individuals with positive (non-zero) number months of work and zero or missing value for earnings are assigned zero earnings. We were allowed restricted access to these data in the NII protected lab in Jerusalem.

c. Sample restrictions

Throughout the analysis, we restrict the analysis to schools and grades that satisfy the following set of conditions: (1) school has a positive number of students in every sample year (1995 to 2000), (2) school has at least two students from either early (1998-2000) or late (2003-2004) reformed kibbutzim, both before (1995-1996) and after the early reforms (1999-2000), and (3) grade has a positive number of students from early reformed kibbutzim and/or a positive number of students from late reformed kibbutzim. The goal of these restrictions is to capture the set of schools that are typically attended by students from early and late reformed kibbutzim. In addition, although some of the peers of early reformers and late reformers are kibbutz members from different kibbutzim - for instance, from kibbutzim that never reformed-, we further restrict the sample of peers to non-kibbutz members. In the robustness section of the paper, we assess the sensitivity of the results to these sample restrictions.

4. Empirical Strategy

To identify the spillover effects of the pay reform, our baseline strategy takes advantage of differences in the timing of the implementation of the reform in a difference-in-differences (DID) framework. Our first difference compares non-kibbutz members in grades with students from kibbutzim that reformed early (1998-2000) to non-kibbutz members in grades with students from kibbutzim that reformed late (2003-2004). Our second difference compares cohorts of students who started high school before (1995-1996) and after (1999-2000) the implementation of the early reforms. We start from this model since it enables us to compare more directly the magnitudes of the spillover and the direct effects.

In our baseline empirical exercise, we estimate the following regression:

$$Y_{isc} = \alpha_s + \alpha_c + \beta_1 Treated_{sc} + \beta_2 Treated_{sc} X\ After_c + \varepsilon_{isc} \qquad (1)$$

where Y_{isc} is an outcome of student i in cohort c in school s, α_s are school fixed effects, α_c are cohort fixed effects (for students starting school in 1995, 1996, 1999 and 2000), $Treated_{sc}$[9] is an indicator variable that captures whether a student is exposed to peers from early reformed kibbutzim, and $Treated_{sc}\ X\ After_c$ is the interaction of interest, indicating if a student was exposed to early reformers and attended school in the post-reform period. We also estimate a version of equation 1 in which we add a vector of student's background characteristics including gender, mother's years of education, father's years of education, number of siblings and ethnicity indicators. In all the regressions throughout the paper, we cluster the standard errors at the school level.

We define treatment status at the grade (school-cohort) level, based on a student's peers in 10^{th} grade, the *first year* of high school (high school in Israel includes grades 10^{th} to 12^{th}). We

[9] Note that because treatment status is defined at the grade and not at the school level, the treatment indicator is not perfectly correlated with the school fixed effects.

choose the grade rather than the class as the level of analysis since classes are potentially endogenous, as parents and school authorities may have discretion in placing students in different classes within a grade (Hoxby 2000, Lavy and Schlosser 2011). We note that this is not a very restrictive compromise because in our baseline sample there is a very high correlation (above 0.7) between treatment status defined at the grade and treatment status defined at the class level.[10] Similarly, we define treatment status based on a student's peers on the first year of high school since subsequent changes (for instance, students who move to a different school or who drop out) might also be endogenous.

In our baseline exercise, a grade is defined as treated if the number of students from early reformed kibbutzim is greater than zero. Our comparison group comprises grades in which the number of students from early reformed kibbutzim is zero, but the number of students from late reformed kibbutzim is positive.[11] Note that while we define a grade as being treated if the number of students from early reformed kibbutzim is one or more, the average grade has many more than one treated student: the average number of directly treated students in a grade is 19, which represents about 15% of the typical grade.

While our base specification has the advantage of enabling us to compare the direct and the spillover effects, it has the disadvantage that it does not permit to test for differences in the *intensity* of treatment. Hence, in addition to this baseline model, we also estimate the following two models:

$$\begin{aligned} Y_{isc} = \alpha_s + \alpha_c + \beta_1 ShareEarlyReformers_{sc} \\ + \beta_2 ShareEarlyReformersX\ After_c + \varepsilon_{isc} \end{aligned} \quad (2)$$

where $ShareEarlyReformers$ is the proportion of students from early reformed kibbutzim in a student's grade (measured at the beginning of 10th grade, the first year of high school), and:

$$Y_{isc} = \alpha_s + \alpha_c + \sum_{q=1}^{4} \beta_q TreatedQ_{sc} + \sum_{q=1}^{4} \beta_q TreatedQ_{sc} X After_c + \varepsilon_{isc} \quad (3)$$

where $TreatedQ$ are indicators corresponding to different quartiles of the proportion of students from early reformed kibbutzim in a student's grade (the omitted category are grades with no early reformers).

[10] Not surprisingly given this high correlation, we show in the robustness section that the results are similar if we instrument treatment status defined at the class level with treatment status defined at the grade level.

[11] In the robustness section, we report an alternative specification in which we use the same control group but we define a grade as being treated if the number of students from early reformed kibbutzim is greater than zero *and* the number of students from late reformed kibbutzim is zero.

Online Appendix Table A.2 presents the sample of schools, grades and students that we use in our baseline analysis. In total, our baseline sample includes students from 31 high schools in Israel. Our pre-treatment sample includes a total of 3,177 students and our post-treatment sample includes 4,529 students. There are a total of 61 grades in the pre-treatment period, with 48 in the treatment and 13 in the control group. The number of grades in the post-treatment period is 62, out of which 52 are in the treatment and 10 in the control group. The average grade size in the sample is approximately 125. Because we define a grade as treated if there is at least one student from an early-reformed kibbutz, the larger the grade the more likely there would be at least one such student. However, our results are unchanged if we include grade size as an additional control variable (Online Appendix Table A.19).

Our identification assumption is that the exact timing of the reform is orthogonal to the potential outcomes of the peers of students in early and late reformed kibbutzim. In other words, we assume that in the absence of the reforms, the outcomes of the peers of students from early reformed kibbutzim would not have been systematically different than the outcomes of peers of students from kibbutzim that reformed late.

We provide evidence that this assumption is plausible. First, we show that students with peers from early reformed kibbutzim were similar to students with peers from late reformed kibbutzim, both in terms of their background characteristics and in terms of their schooling outcomes before the reform. While our identification strategy only requires parallel trends in the outcomes, it is reassuring that even the *levels* of the outcomes were similar in the pre-reform period. Second, we show that peers of early reformers were on a similar time trend to peers of late reformers in the pre-reform period. Third, we show that there is no evidence of sorting of peers -based on observable characteristics - as a result of the reform. That is, we do not find evidence of a differential change in the pool of peers of early reformers after the reform. Fourth, we perform a placebo exercise assuming that the reform happened on an earlier year and find no effects.

Peers of early reformers are similar to peers of late reformers. In Online Appendix Table A.3, we show that the peers of early reformers are similar to the peers of late reformers. In columns 1 and 5, we report the mean and standard deviation (in parentheses) of each of the student's background characteristics and outcomes, before and after the early reforms. In columns 2, 3, 6 and 7 we display the mean and standard deviation of each of these variables, separately for treatment and control students and before and after the early reforms. In columns 4 and 8, we report the estimated coefficient and standard error (in parentheses) in a regression of each of the variables on a treatment indicator and cohort fixed effects. In particular, we estimate for each of the background characteristics (X_{isc}) and separately for the pre and post-reform periods the following regressions:

$$X_{isc} = \alpha_c + \beta Treated_{sc} + \varepsilon_{isc} \quad (4)$$

Similarly, we estimate for each of the schooling outcomes, Y_{isc}:

$$Y_{isc} = \alpha_c + \beta Treated_{sc} + \varepsilon_{isc} \quad (5)$$

Panel A of Online Appendix Table A.3 shows that the background characteristics of students and their families are similar in both groups, both before and after the early reforms. Father's years of schooling are lower in the control group in the pre-treatment period, a difference of 0.6 years. Only mother's years of schooling in the pre-treatment period is significantly different across the two groups. The differences in parental years of schooling between the treatment and control groups becomes smaller and not significant in the post-treatment period, 0.39 for the fathers and 0.53 for mothers. We also note that, all our results are robust to controlling for parental years of schooling.

Differences in the average number of siblings are small and not statistically significant, both in the pre- and the post-treatment periods. On average, students have between 2.2 and 2.6 siblings in both groups. The treatment and control groups are also similar with respect to their ethnic origins. The more salient difference among the two is that students in the control group are 5 percentage points more likely to belong to the Asia-Africa ethnic group. Finally, students in the control and treatment groups are also relatively similar in terms of average family income.

Panel B of Online Appendix Table A.3 shows that, consistent with the small differences in background characteristics, schooling outcomes are similar across the two groups in the pre-reform period. The rate of students graduating from high school is 0.7 percentage points smaller in the treatment group, relative to a mean of 95%. The mean matriculation score is also similar across the two groups, a difference of 0.6 points in favor of the treatment group. The fraction of students obtaining a matriculation certification is slightly higher in the treatment group, both for the regular and the university qualified. None of these differences are statistically significant at the conventional levels.

Peers of early reformers and peers of late reformers had a similar time trend. We next test if the outcomes of the treatment and the control groups had similar trends prior to the implementation of the early reforms. In Figure 1, we estimate a version of equation (1) in which the treatment indicator is interacted with a series of cohort dummies corresponding to students starting 10th grade in each of these years. In addition to the cohorts that we include in the baseline analysis (1995, 1996, 1999 and 2000), we also include students who started high school in 1994, 1997 and 1998 so as to compare the treatment and the control groups for the full period for which data is available. Note that students who started high school in1997 and 1998 were partially exposed to spillovers for most of their high school years (since more than half of the early reforms took place in 1998). Specifically, we estimate:

$$Y_{isc} = \alpha_s + \alpha_c + \sum_{c=1994}^{2000} \beta_c Treated_{sc} + \varepsilon_{isc} \qquad (6)$$

Figure 1 shows the estimates of β_c from these regressions, focusing on the high school outcomes. The omitted category are students in the control group who started high school in 1994. Regardless of the outcome we consider, the interaction between the treatment indicators and the cohort dummies is insignificant for students starting high school in 1994, 1995 or 1996.

No sorting across schools as a result of the reform. One possible violation of our identification assumption is the endogenous sorting of students across schools as a result of the reform. This sorting might have happened for two reasons. First, students from kibbutzim that reformed early might have decided to enroll in better quality schools after the reform. Note that because our analysis includes school fixed effects, for this type of sorting to bias our results students from early reformed kibbutzim must have switched to schools on a better time trend. Second, the prospects of sharing a school with early reformers might have attracted a better pool of peers in the post-reform period to those schools typically attended by early reformers. In this case, our estimation strategy would be capturing a compositional change in the group of peers rather than spillover effects from the pay reform.

A number of features of our setting and empirical strategy make this concern less worrisome. First, note that we define treatment status based on the first year of high school. Hence, if there were any sorting, it would have needed to occur before students actually started high school. Second, we define treatment status at the grade (school-year) level, which rules out sorting occurring at the class level. Third, note that by restricting the sample to schools attended by kibbutz members both before and after the early reforms, we largely rule out the effects being driven by kibbutz students attending a different set of schools after the reforms.

Yet, the possibility of sorting is a threat to our identification strategy. We provide two pieces of evidence that suggest that this sorting did not occur. First, we document that early reformers did not switch to a different set of schools in response to the reform. In practice, most students living in the same kibbutz also attend the same high school. Collapsing our data at the kibbutz-year level, we find that in 76% percent of the cases all the students in the kibbutz attended the exact same high school, and that the average share of students attending the largest school within a kibbutz-year is 95%. Indeed, the median number of schools per kibbutz-year is 1 and in 88% of the kibbutz-years students attended at most two different schools. Moreover, in 97% of the kibbutz-years, the most attended school was the same as in the previous year. Importantly for our identification strategy, we do not observe any systematic pattern of school switching before and after the early reforms. The mean and median number of schools remains similar in kibbutz that reformed earlier. In addition, the share of students who attend the largest school is also stable. These findings are consistent with the fact that, unlike in the US context, there is very little mobility between schools in the Israeli educational system (Lavy and Schlosser, 2011).

Second, there is no evidence of a systematic change in the observable background characteristics of peers after the early reform. To formally test for this possibility, we regress each of the background characteristics on a treatment indicator and an interaction between the treatment indicator and a post indicator. If students from earlier reformed kibbutzim were not systematically sorting across schools as a result of the reform, then we should not find any differential change in the background characteristics of their peers, relative to the control group. More precisely, we estimate:

$$X_{isc} = \alpha_c + \beta_1 Treated_{sc} + \beta_2 Treated_{sc}\ X\ After_c + \varepsilon_{isc} \qquad (7)$$

where X_{isc} corresponds to a background characteristic of student *i* in school *s* in cohort *c*. In the absence of sorting, we expect to find that $\beta_2 = 0$.

Table A.4 in the Online Appendix shows the results of estimating this specification for each of the background characteristics that we observe in our data. Peers appear to look slightly worse in terms of parental educational background in the post-reform period relative to the control group, but better in terms of family income. All the other differences are small and statistically non-significant.

Spillovers outside the classroom and anticipation effects. We cannot rule out that students who did not share a grade with early reformers still knew about the pay reforms happening in some kibbutzim. If information about the pay reform was equally salient irrespective of sharing a grade with an early reformer, then our estimates will just capture the peer effects component of the overall spillover effects (and will likely be biased downwards). In addition, students from kibbutzim that reformed late might have increased their effort in anticipation to the late reforms. Note, however, that in our design we focus on late reforms that took place at least three years after the early reforms, making such anticipation less plausible. In Abramitzky and Lavy (2014), we empirically document the lack of anticipation effects among students in late reformed kibbutzim.

Why not using peers of never-reformers in the baseline specification. Students from kibbutzim that never reformed were significantly different than those from either early or late reformed kibbutzim in terms of high school outcomes (Abramitzky and Lavy 2014). In Table A.5 in the Online Appendix, we show that adding peers of never reformers to the control group worsens the balancing between the treatment and the control groups in the pre-reform period. Specifically, pre-reform high-school outcomes in the control group become slightly better than in the treatment group (although for none of the high-school outcomes in the pre-reform period we can reject the hypotheses that the outcomes are the same). While difference-in-differences identification strategy does not strictly require such balancing (i.e. we only need to assume parallel trends), we believe that the parallel trends assumption becomes less plausible if we add never reformers. We show, however, that the results are very similar when including these students to the control group sample (Table A.8 in the Online Appendix).

5. Short-Term Effects on High School Outcomes

a. Basic results

In Panel A of Table 1, we present the results of estimating equation (1) using the high school outcomes as dependent variables. We report two main specifications for each of the high school outcomes. In the first row, we report the simple DID, without any further controls other than the school fixed-effects. In the second row, we include students' background characteristics as additional controls. In each of the rows, we show the estimated coefficient of interest corresponding to the treated group in the post-reform period.

The table shows a positive coefficient on all the schooling outcomes (columns 1 to 4). First, the fraction of students completing high school increases by approximately 1.6 percentage points, relative to an already high mean completion rate of 95% (column 1), implying a 2 percent improvement. Second, the mean matriculation score increases by 2.3 points, relative to a mean matriculation score of 70 points (column 2), effectively a 3 percent increase. Note that these effect sizes are relatively small and that neither of the previous estimates is precisely measured.

We next report our estimated effects on the probability of obtaining a matriculation certificate (column 3) and of obtaining a university-qualified matriculation certificate (column 4). We find an increase of 7.8 percentage points in the matriculation rate, relative to a pre-reform level around 61%, a 13 percent improvement. The increase is of similar magnitude in the university-qualified matriculation, although the pre-reform mean is lower in this case (57%). Note that as a result of the balancing documented in Table 2, the point estimates exhibit little sensitivity to controlling for student's background characteristics (row 1 versus row 2 estimates).

The positive impact on high-school outcomes holds when we estimate aggregate treatment impacts, using a summary index instead of individual outcomes to account for multiple inference (Kling et al. 2007). In column 5, we present the results of a specification that uses this summary index measure as the dependent variable. This index is computed as the equally weighted average of each of the high-school outcomes' z-scores. The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation. Thus, each component of the index for the control group has mean 0 and standard deviation 1. The results using this summary measure also indicate an overall improvement in high-school performance.

One testable implication of our identification assumption is that we should not find any effects of the reforms on unaffected cohorts, i.e. students who attended school before the early reforms. To directly test this implication, in Panel B of Table 1, we report the results of a placebo exercise in which we estimate the same DID specification as in equation (1), but assuming that the early reforms happened in 1996 instead of 1998. In particular, we compare students in grades with students from kibbutzim that reformed early to grades without early reformers, before (1994-1995) and after (1996-1997) the *placebo* reforms.

Reassuringly, the point estimates in this exercise are all small in magnitude, some of opposite sign, relative to the estimates in Panel A and none of them is statistically significant. Together with the lack of any pre-reform time trends documented in Figure 1, this exercise provides further support to our assumption that the outcomes of peers in the treatment group would have been similar on average to those in the control group in the absence of the reform.

Table A.6 in the Online Appendix shows that, in addition to improving their high-school outcomes, peers changed the type of subjects that they took during high school. Specifically, peers increased the number of credit units in English, math and sciences. Completing five credit units in these subjects (which is equivalent to enrolling in honor level classes in the US) is often required in Israel for admission to fields of study such as Engineering, Computer Science and Economics. Hence, this finding is consistent with students switching towards subjects with relatively high financial returns.

How large are the spillover effects relative to the direct effect on students from kibbutzim. A simple comparison between the size of the spillover effects we document here and the direct effect on kibbutz students reported in Abramitzky and Lavy (2014) suggests that the effects are of similar magnitude.

However, Abramitzky and Lavy (2014) might have *underestimated* the direct effects. Specifically, that paper did not take into account the fact that students from early reformed kibbutzim (the "treatment group" in Abramitzky and Lavy 2014) were often in the same grades as students from late reformed kibbutzim (the "control group" in Abramitzky and Lavy 2014). If students whose kibbutz reformed late were affected by their peers from kibbutzim that reformed early through spillover effects, this would have led to a downward bias in the direct effects estimated in Abramitzky and Lavy (2014).

To test this possibility, we replicated the results of Abramitzky and Lavy (2014) using a sample that *excludes* those grades with students from *both* early and late reformed kibbutzim. By focusing on this sample, we obtain a 'cleaner' control group in which we shut down the possibility of within grade spillovers from early to late reformers. Using this restricted sample, we find (Panel B of Table A.6) that the direct effects are in all cases at least twice as large as in the baseline sample of Abramitzky and Lavy (2014). This suggests the spillover effects could be at most *half the size* of the direct effects. We note, however, that this result is based on a much smaller sample of about one fourth the size of the original sample (because we drop all grades that have a positive number of both early and late reformers). Moreover, we have worse balancing between the treatment and control groups than in the full sample, although the main DID parallel trend assumption does hold for that smaller sample. Overall, we conclude that while the indirect effect is sizeable, it is likely much smaller than the direct effect.

While smaller than the direct effects on kibbutz students, the spillover effects that we document are sizeable: the kibbutz students that were directly affected experienced an increase of 0.3 in their index of high school outcomes while their non-kibbutz peers in high school had their respective index increase by half of that, about 0.15. We note that our estimates of spillover effects are within the bounds of estimates presented in recent related papers (see Table A.1 in the Online Appendix). For example, Duflo and Saez (2003) find spillover effects of an information treatment that are similar in size to their estimated direct effects. Similarly, Bobonis and Finan (2009) find that spillover effects resulted in doubling the direct effect of a school enrollment subsidy in Mexico. Miguel and Kremer (2004) estimate spillover effect on school peers that are of similar magnitude to the effect on students who directly received the treatment (deworming drugs), demonstrating that deworming creates large epidemiological externalities. Finally, in Angelucci and Di Giorgi (2009) study of the spillover effects of Progresa, the increase in the consumption of the ineligible in treated villages is about half of the increase of directly treated individuals.

b. Accounting for differences in treatment intensity

Table 2 shows the results of estimating the two models that account for differences in treatment intensity. When estimating equation (2), we find that the effect is generally larger when there are more students who are treated, but the results are not statistically significant (with the exception of the high school completion outcome). When estimating equation (3), we find that the effect is bigger when peers are exposed to more students (top three quartiles vs. 1st quartile), but we do not find differences between each of upper three quartiles. This is possibly a matter of power, because our sample size is relatively small.

Indeed, when estimating these specifications in an expanded sample that includes peers of students from kibbutzim that never reformed in the control group, we find that the effects are in general larger when more students are treated. Tables A.8 and A.9 present the estimation results based on this expand sample, both when using our baseline specification (using a 0/1 treatment) and when using the specification that account for differences in treatment intensity. First, note that the baseline results (Table A.8) are very similar to the results presented in Table 1. When estimating the model in equation (6), we find that in most outcomes the effects are larger when the proportion of directly treated students in the grade is larger. Similarly, when we use indicators for different quartiles of the distribution of early reformers as measures of treatment intensity, we find that the estimated effect of the upper two quartiles is larger and more precise relative to the effect of the below median quartiles, which is small and statistically insignificant.

Interestingly, however, we continue to find in this expanded sample that the effect does not seem to be fully monotonic across the different quartiles. In particular, in some cases the effect does not continue to increase (and sometimes decreases) once the share of kibbutz students is very large (in the top quartile). There is a natural (although suggestive) interpretation for this pattern: Once kibbutz students become a significant fraction of the grade, they are more likely to interact with one another rather than with their non-kibbutz peers, thus reducing the scope for spillover effects. Such a pattern is consistent with Carrell et al (Econometrica 2013), which finds that once low-academic-ability students are a large enough fraction of the class, they are more likely to interact among themselves rather than with the high-academic-ability students (thus limiting the scope for positive spillovers).

c. Robustness Checks

We next assess the sensitivity of our main results on high school outcomes to: (1) using only post-treatment cross-sectional variation, (2) including additional control variables, (3) alternative sample restrictions, and (4) school-time specific shocks.

First, in the third row of Panel A of Table 1, we show that the results are similar when we estimate a cross-sectional regression using only the post early reforms cohorts. This finding implies that the DID estimates are driven by improvements in the treatment group rather than by a decline in performance of the control group. This panel also shows that the pre-reform outcomes were very close in both groups: none of the pre-treatment differences in outcomes are statistically significant. Consistent with this pattern, Online Appendix Table A.10 shows that the results are similar when not including school fixed effects in our baseline specification.

Second, in Table A.11 in we show that the results are similar when we add a student's family average earnings in 2000-2002 as an additional control in the DID estimation. We prefer a multi-year average because it is more likely to be correlated with the permanent level of family resources. Note that performing this exercise was not possible in Abramitzky and Lavy (2014), since family income cannot be properly measured among families who live in the kibbutz.

Third, in Online Appendix Table A.12 we show that the results are similar when we implement an instrumental variables strategy in which we instrument a class-level treatment indicator with the treatment indicator defined at the grade level. The validity of this instrument rests on the assumption that cohort-to-cohort changes in the exposure to students from reformed kibbutzim is random conditional on school fixed effect that account for any confounding factors. This is a reasonable assumption because within a short period of time it is safe to assume that students from adjacent cohorts in a given school have similar characteristics and face the same school environment, except for the fact that one cohort has more students from reformed kibbutzim due to purely random factors. We note that the reduced form effect of this instrument is exactly the grade level treatment effect that we presented above. Secondly, note that within a school the proportion of students from reformed kibbutzim in a grade is highly correlated with the students from reformed kibbutzim in a class, which forms the first stage regression in this 2SLS set up.

We next explore the sensitivity of our results to the sample restrictions and to different definitions of the treatment and control groups. In Table A.14 we present results from two alternative samples. In the first, we restrict the analysis to schools and grades that have at least three students from either early (1998-2000) or late (2003-2004) reformed kibbutzim, both before (1995-1996) and after the early reforms (1999-2000). In the second sample we require at least 6 students. We jump from 3 to 6 students because there are no schools with 4 or 5 such students. Remarkably, the estimates we obtain from these two smaller samples are very similar to the estimates obtained when the restriction is at least 2 students. For example, the effect on high school completion is 0.018 in the 2+ and 6+ samples. The effect on matriculation certification is 0.088 and 0.079, respectively. These similarities are obtained even though the sample size declines by 18 percent.

We next report a specification in which we keep the same students in the control group but drop from the treatment group all the grades with students from both early and late reformed kibbutzim. That is, we compare grades with early reformers but no late reformers to grades with late reformers but no early reformers, before and after the implementation of the early reforms. We report the results of this exercise, as well as the corresponding balancing and sample size Tables in Online Appendix Tables A.15, A.16. and A.17. The results are similar to those in our main specification both qualitatively and quantitatively. Note, however, that the sample size goes down reflecting the more stringent definition of the treatment group.

Finally, we assess the robustness of our results to time and school-specific shocks correlated with the presence of early reformers in the grade. There are a number of reasons why such shocks are unlikely to explain our results. First, our findings (discussed below) that some of the effects are larger for males and for students whose parents are less educated largely rule out grade level factors that affect all students equally, such as improvements in schooling infrastructure, changes in teaching practices or in the composition of teachers.[12]

Second, since we define treatment at the school-year level, our sample includes schools that have grades in both the treatment and control groups, both in the pre and in the post-reform cohorts. In a robustness check, we exploit this feature of the data and estimate our model with a restricted sample that includes *only* schools that have at least one grade in the control group. To have the largest possible sample for this robustness check, we use in the control group peers of students from any kibbutzim. More explicitly, the treatment group in this robustness check includes peers of students from kibbutzim that reformed early and the control group includes peers of students from kibbutzim that reformed late or never reformed. The estimates from this model are presented in Online Appendix Table A.18. The estimates are remarkably similar to our main high school results presented in Table 1. For example, the effect on the summary index in Table A.18 is 0.118 (se=0.076) and in Table 4 it is 0.153 (se=0.058). These results practically rule out the possibility that a school-specific shock is driving our results.

d. Heterogeneous effects

Differences between men and women. Previous research on the effectiveness of schooling interventions has shown differences in the responsiveness of men relative to women (for example, Angrist and Lavy 2009). To test for this possibility, in Panel (i) of Table 3, we stratify the sample based on the gender of students. In the last row of each panel of Table 3, we include the p-value for the hypotheses that the effects are equal across genders.

[12] We noted earlier in the paper that the high schools that are attended by students from kibbutzim are regional schools where over 80 percent of the students are from non-collective localities in the area. Kibbutzim do not own these schools and they have no influence the composition of the schools' staff. Some of these schools belong to the regional or district authority and some belong to private or non-profit organizations. Note also that laying off teachers, even if ineffective, is almost impossible in Israel education system because of the tenure system and the strong influence of the teachers' union.

All the point estimates suggest that the effects are larger among men, although in some cases the male-female differences are not large and we cannot reject the equality between the coefficients. In the case of high school graduation and mean matriculation scores, differences between men and women are large and statistically significant. The high school graduation rate of men goes up by 3.8 percentage points, but barely changes among women. Similarly, we find an increase of 5.6 points in the mean matriculation score of men, but no such change among women.

Differences between men and women in the estimated effect on obtaining a matriculation certification or a university-qualified matriculation are smaller and not statistically significant. The proportion of male students obtaining a matriculation certification goes up by 12.2 percentage points, relative to 5.3 percentage points among women. The proportion of men obtaining a university-qualified matriculation goes up by 10.9 percentage points and by 7.6 percentage points among women.[13] We note that obtaining a matriculation certificate has important long-term consequences for students as it is a gateway to higher education, especially in research universities.[14]

Differences by social background of students. In Panel (ii) of Table 5, we stratify the sample based on the education of the student's mother (below and above median education). We find that the effects are larger among students with below median mother education, and in some cases insignificant for the high socio-economic status families. This pattern is also consistent with the findings in Abramitzky and Lavy (2014) on the direct effects of the reform on kibbutz students. In this case, we cannot reject - at the 10 percent level - the hypothesis that the effects on high-school completion, mean matriculation score and the summary index are the same across the two subgroups.

6. Long-Term Effects

We next analyze whether the improvements observed during high school resulted in long-term gains in educational and labor market outcomes. To do so, we link our schooling data to

[13] The fact that the direct effect on students from kibbutzim is larger for males (Abramitzky and Lavy 2014) suggests that being exposed to male students from a kibbutz should have stronger effects. Unfortunately, it is not possible to separately identify the effects of being exposed to male kibbutz students from the effects of being exposed to female kibbutz students, since both are highly collinear. Specifically, out of the 100 treated grades in our sample, 94 have both male and female directly treated students. There is also a very high correlation in the *share* of directly treated male and female students within a grade. Collapsing our data at the school-year level, the correlation between the share of males and the share of females within a grade is 0.88.

[14] Angrist and Lavy (2009) describe the high school matriculation certificate as arguably marking "the dividing line between the working class and the middle class."

administrative annual data from the Israeli National Insurance Institute (the equivalent of the social security administration in other countries) spanning the 2000-2014 period. The youngest individuals in our data started high school in 2000, so they were around 31 years old by 2014. Hence, in our main exercise we focus on educational and labor market outcomes by age 31, which is the latest age at which we are able to observe the youngest individuals in our sample.

a. Post-secondary schooling

In Table 4, we start by looking at the spillover effects on post-secondary schooling. In columns 1 and 2, we test whether treated peers: (1) were more likely to enroll in any post-secondary schooling at some point from high-school graduation and up to 12 years after high school completion and (2) completed more years of post-secondary schooling up to 12 years since high school graduation. In columns 3 and 4, we repeat the analysis but focusing instead on university enrollment and years of university schooling. In columns 5 and 6 we provide the respective estimates for academic colleges.

On average, peers of early reformers are approximately 10 percentage points more likely to have been enrolled in university schooling 12 years after high school completion. On the intensive margin, students complete 0.53 additional years of university schooling, relative to a mean of around 1.7 years. Note that in the section on high-school outcomes, we found a 9.5 percentage points increase in the probability of obtaining a university-qualified matriculation. When focusing on post-secondary schooling, we find a similar increase in the likelihood of university attendance. This similarity suggests that most of those who obtained a university qualified matriculation indeed enrolled in university education.

The increase in both enrollment and years of university education is accompanied by a shift away from academic colleges. In particular, students are 4 percentage points less likely to enroll in academic colleges and complete 0.12 fewer years of academic colleges education. This decrease suggests a shift away from lower into higher quality of post-secondary schooling. As already discussed, for a given field of study, it is typically more difficult to be admitted to a university than to an academic college. Moreover, universities offer a premium in the labor market relative to academic colleges (Caplan et al 2009).[15]

[15] We note that this result is very different than the direct effect on students from reformed kibbutzim, for whom the post-secondary gain was an increase in academic colleges with zero effect on university schooling (Abramitzky and Lavy 2014). This different pattern of the margin at which we find a positive treatment effect could result from the higher average high school outcomes of the peers relative to the kibbutz students. These higher achievements enable the peers to be admitted to better higher education institutions, in particular universities, and to highly demanded fields of study such as medicine and computer science. Note that, similar

In Panels a) and b) of Figure 2, we measure the treatment effect for each year since high school graduation –starting in year 3, after students have completed the mandatory military service - and trace the dynamic pattern for each of the post-secondary schooling outcomes. To do so, we run a separate regression for each of the outcomes and for each of the years since high school graduation. We then plot the coefficients of these regressions around a 95% confidence interval. Note that both the ever-enrolled variable and the years of schooling are cumulative variables. Hence, we expected the effects to be either flat or increasing over time.

We find that the effect on enrollment is flat after five years. This pattern likely reflects the fact that students who do not enroll in post-secondary schooling in the first five years are unlikely to return to school later in life. In contrast, the effect on years of schooling accumulates over time. Although most of the increase happens in the first five years, the effect seems to be increasing even after 12 years since graduation. The fact that the increase keeps accumulating even 12 years after high school graduation suggests that measuring outcomes too close to high-school graduation might underestimate the long-term effects.

The substitution over time from (typically lower quality) academic colleges into (typically higher quality) university can be seen graphically in Panels a) and b) of Figure 1.[16] The divergence starts early on, suggesting differences in the initial choice of academic institutions and accumulate over time as students spend time in these institutions. By year 12 after high school graduation, students had accumulated 0.5 extra years of university education and 0.12 less years of academic college education.

b. Labor market: employment and earnings in adulthood

We expect this increase in both the quality and the quantity of education to result in better labor market outcomes in adulthood. In Table 5, we estimate the long-term spillover effects now focusing on labor market outcomes. In column 1, our dependent variable is an indicator that takes a value of one if the individual was employed at least 6 months in a given year. In the second column the dependent variable is the number of months of work in a given year. In the third column, our dependent variable is annual earnings measured in 2009 Israeli NIS. In all columns in this table, we focus on labor market outcomes measured 12 years after high school graduation.

We find a positive but small and insignificant effect on employment, on either of the two employment measures that we use. The mean employment rate is 85% in the pre-reform period and

to the regressions on high school outcomes and regardless of the specific outcome that we consider, the point estimates barely change as we control for students' background characteristics.

[16] Academic colleges in Israel are mainly public teaching (non-research) institutions.

it is practically unchanged following the reform. However, we document an increase in annual earnings of about 6988 Israeli New Shekels (NIS)–in 2009 prices -, which is equivalent to $1742[17], relative to mean earnings of approximately 73,000 NIS. The estimated effect on earnings appears to operate through higher paying jobs because we do not find any effect on employment.

The estimates presented in column 4 show that the spillover effect had lowered the unemployment rate, the duration of unemployment spells and the annual average of unemployment benefits in the treated group. These improvements can be consistent with the zero effect we find on the employment indicators if the duration of unemployment is short enough so that they are not associated with a change in the annual indicators of employment.

In Panels c) and d) of Figure 2, we repeat the year-by-year analysis but now focusing on the two main labor market outcomes (employment and annual earnings). The figure shows the estimated effects by years since graduation from high school. We find an increasing pattern in both employment and earnings. As treated students spent more years on average in the schooling system and appear on average to start working later, we expect the effect on earnings to increase as students accumulate labor market experience. Indeed, we find that the effects are initially small and become significantly different from zero by the end of our sample period. The effects on earnings become significantly different from zero about after 9-11 years from high school graduation, a similar dynamic pattern as in Chetty et al (2016) study on the Moving to Opportunity experiment. Similarly, the effect on employment is initially negative, then it increases for a few years and it levels offs thereafter.

We estimate that students exposed to peers from early reformed kibbutzim increased their years of university education by about 0.5 and decreased their years of college education by 0.12 years. Combining the earnings and years of schooling effects, we can compute the returns to schooling that, given the observed increase in schooling, would rationalize the size of the earnings effects. The mean annual earnings 11 years after graduation for individuals in our sample is approximately 73000 Israeli NIS. Hence, an increase of 6988 NIS represents approximately an 9.0% increase. If the increase in the years of schooling would have been the only channel through which individuals increased their long-term earnings, then the return to a year of university or college education would have needed to be such that: Return X 0.54=9.0%. Hence, the observed

[17] 1 Israeli NIS was worth 0.25 US dollars in 2009.

simultaneous increase in earnings and schooling is consistent with a return to one extra year of post-secondary schooling of 16%.[18]

This calculation suggests that an important fraction of the increase in earnings was due to the increase in post-secondary educational attainment. However, the improved matriculation outcomes can account for part of the increase in earnings independently of their effect on university years of schooling. Particularly important is the matriculation rate where evidence suggests that a matriculation diploma is rewarded in the labor market by a return beyond its effect on post-secondary schooling. For example, Angrist and Lavy (2009) estimates that *bagrut* holders earn 13 percent more than other individuals with exactly 12 years of schooling. Similarly, the quality improvements in the matriculation study program, reflected partly by a higher average score, are also rewarded in the labor market beyond their effect on post-secondary schooling (Caplan et al (2006)).[19]

c. Alternative specifications and robustness

In Tables A.19 and A.20, we show that, similar to the results on high school outcomes, the results for university and college schooling and labor market outcomes are similar when controlling for average family income in the regressions. In Table A.21, we show that the results also hold when we estimate aggregate treatment effects using a summary index for post-secondary schooling and labor market outcomes**.**

A natural question about the above estimated effect on earnings is whether it captures the permanent long-term effects. First, note that we measure the effect on earnings when individuals already completed their post-secondary schooling. Second, based on a sample of older cohorts, we find that earnings at age 30-35 is a strong predictor of earnings at an older age. Yet, it is important to note that earnings have larger variation over time than other personal outcomes. To get a better indication about the permanency of the effect on earnings, we estimated the effect on the percentile rank of individuals in the respective distribution of their cohort (at the national level). There is no direct evidence that suggests that rank forecast is more stable than earnings or log earnings.

[18] Recent estimates of the rate of return to a year of university schooling in Israel ranges from 12 to 16 percent. Frish (2009) exploit changes in compulsory schooling laws and obtain IV estimates that are much larger than the OLS Mincerian estimates. Navon (2006) estimate that the return to an MA degree (two years of schooling) is 30 percent.

[19] Caplan et al (2006) demonstrate that earnings in Israel is highly positively correlated with the quality of post-secondary schooling (colleges versus universities and higher versus lower quality universities). For example, this study shows that earnings are much higher for graduates of Tel Aviv, Jerusalem and the Technion Universities relative to graduates from the other four universities in the country. Admission to the top universities is of course positively correlated with the high school matriculation outcomes.

However, recent papers in the intergenerational mobility literature provide some indirect evidence that is relevant to this issue. These studies have shown that movements across ranks in the income distribution are uncorrelated with parental income conditional on rank at age 30; in contrast, movement in log earnings are correlated with parental income conditional on log income at age 30.[20]

Table A.22 in the Online Appendix presents estimates of the effect of the program on percentile rank of earnings, where the rank is computed separately for each cohort based on their percentile in the national income distribution. The estimates are fully consistent with the estimated effects on earnings that are presented in Table 5. After 12 years from high school graduation, the spillover effects moved treated individuals by about 4 percentile ranks in the national income distribution.

d. Heterogeneous effects

Differences between men and women. In Panel (ii) of Table 4 we present the estimated effects on post-secondary schooling outcomes by gender. When focusing on overall post-secondary schooling attainment (without distinguishing universities from academic colleges), we find that the spillover effects are stronger for men. In particular, men are 7.7 percentage points more likely to enroll in any type of post-secondary education (the point estimate is close to zero among women) and they complete 0.475 of extra years of schooling (relative to only 0.052 among women). However, there is a similar increase in the likelihood of university enrollment (10.1 percentage points for men, 8.1 percentage points for women) and in the number of years of university education completed (0.579 for men, 0.442 for women). This similarity is consistent with the similar increase in the likelihood of obtaining a university qualified matriculation among men and women documented above. Overall, these two findings suggest that the gains for women are mostly concentrated in the intensive margin of schooling quality. Women shifted away from schooling in colleges towards the more selective research universities, which offer a wider range of major choices (including areas absent from colleges such as life sciences and humanities). Men also moved from colleges to universities, but unlike women they also completed more years of post-graduate schooling.

[20] For example, Nybom and Stuhler (2016) show with data from Sweden that the relationship between a child's income rank and their parental income rank stabilizes by around age 30; in contrast, the relationship in log earnings is less stable. Chetty et al (2016) find a similar pattern in the US tax data, reporting that percentile ranks predict well where children of different economic backgrounds will fall in the income distribution later in life. Using instead log earnings leads to inferior predictions because of the growth path expansions at the top of the income distribution.

Panel (ii) of Table 7, presents the estimated effects on long-term labor market outcomes by gender. Here, the patterns are more mixed but overall suggest an improvement for both males and females: The point estimate of the effect on employment is larger for men than for women, but the effect on earnings is slightly larger for women than for men (although both point estimates are statistically indistinguishable and the estimates become quite imprecise once we split the sample by gender). Women also seem to have a stronger response in the unemployment benefits margin. Overall, the improvement in both men's and women's labor market outcomes is expected given that both groups increased their years of university education.

Differences by socio-economic background of students. The evidence regarding long term post-secondary outcomes in Table 6 (columns 3-4) reveals that both groups experienced a similar increase in university years of schooling (0.48 and 0.6, respectively) but for the low mother's education group this improvement was offset by a decline in college years of schooling (-0.31). This suggest that the spillover effects induced disadvantaged students to move from lower to higher level of post-secondary schooling with smaller gains at the extensive schooling margin for this group. Finally, we note that the gain in earnings is actually larger for the high mother's education group, although again we cannot reject the hypothesis that the two estimates are equal.

7. Discussion and Conclusions

We studied the spillover effects to non-kibbutz members of a reform that increased the returns to schooling of kibbutz students. To do so, we compared the high school and post-secondary schooling outcomes of peers of students from early and late reformed kibbutz, before and after the early reforms. In the short-run, students exposed to early reformers improved their high school outcomes and shifted to courses with potentially higher financial returns. In the long run, these students completed more years of university education and had better labor market outcomes in adulthood.

Our data and setting do not allow us to separately identify which precise channel of social interaction contributed to these estimated effects. A first potential explanation is that peers of kibbutz students might have benefited from standard peer effects taking place in the classroom. For instance, peers of kibbutz students might have decided to put more effort due to a "competition effect" (Ching-Huei et al 2017) or to avoid falling to the bottom of their classes (Tincani 2015, Hopkins and Kornienko 2004, and Kuziemko et al. 2014). Peers might have been more affected by such competition and ranking considerations than kibbutz students themselves because the latter grew up in an egalitarian community where competition was less common and peer rivalry was against the equal sharing ideology. Peers also might have benefited from direct *learning* spillovers from their classmates (as kibbutz students improved their schooling performance). Alternatively, the improvement in high school performance among kibbutz students might have freed teachers' time to the advantage of their peers. While some studies have found classroom peer effects that are on average relatively small (see for instance, Angrist, 2014 and Feld and Zölitz, 2017), it is unclear whether such effects will be small or large in our context.

A second channel that is potentially relevant in our setting is that the reform might have increased the salience of the relationship between school effort and financial success. The pay reform was hotly debated within the kibbutz and was an important source of distress, as many kibbutz students experienced actual declines in their family income. In other words, witnessing that the family of one of your classmates lost income because of lack of education likely increased the salience of the link between schooling effort and financial success (and might have constituted a powerful incentive to study). Indeed, a number of studies has shown that students (particularly those from low SES backgrounds) may have not been be fully aware about the returns to schooling in the labor market when making their schooling choices (see for instance Orepoulos 2007, Jensen 2010 and Baker et al 2017).[21]

The information salience hypothesis also provides an alternative explanation to the non-linearity of the effect. As we discussed, one natural explanation for this pattern is that kibbutz students are increasingly more likely to interact with one another (rather than with their non-kibbutz peers) once they constitute a significant fraction of the grade (as in Carrell, Sacerdote and West 2013). Specifically, one implication of the information salience hypothesis is that a relatively small fraction of treated students in the classroom might have been enough to generate large spillovers. This is because, in a standard model of information transmission (Banerjee, Chandrasekhar, Duflo and Jackso, 2013), the probability that an individual receives information grows at an exponential rate with the number of initially informed individuals in the network.[22]

The large direct and indirect response to changes in the returns to schooling in the Israeli context stands in contrast to the more muted response that has been documented in the US context (Altonji et al 2012). One potential explanation for this difference is that the direct monetary costs of acquiring skills are much lower in Israel than in the US, and that these costs have been shown to be an important driver of schooling decisions (Dynarski 2003). More broadly, the pay reform can

[21] For instance, Boneva and Rauh (2017) show that low SES students perceive the returns to education to be lower than high SES students.

[22] To illustrate this point formally, assume that there are n directly treated individuals in a grade, and that there is a probability p that each of them shares information with another student. For a given student, the probability of interacting with at least one treated student (the probability of "contagion") will be equal to: $1-(1-p)^n$. This expression converges to one at an exponential rate. For instance, if the probability p of interaction is 0.5, then it only takes 4 directly treated students for the probability of contagion to be above 90%.

be interpreted as a sharp decrease in the marginal tax rate faced by kibbutz members. Such changes might affect both the human capital accumulation of those directly affected and of those not directly affected through spillover effects. For instance, if there are complementarities in production, changes in the tax schedule that affect only some individuals might indirectly also affect others.

References

R. Abramitzky. Lessons from the Kibbutz on the Equality–Incentives Trade-off. *The Journal of Economic Perspectives*, 25(1):185–207, 2011.

R. Abramitzky. *The Mystery of the Kibbutz: Egalitarian Principles in a Capitalist World*. Princeton University Press, 2018.

R. Abramitzky and V. Lavy. How responsive is investment in schooling to changes in redistributive policies and in returns? *Econometrica*, 82(4):1241–1272, 2014.

J. G. Altonji, P. Bharadwaj, and F. Lange. Changes in the characteristics of american youth: Implications for adult outcomes. *Journal of Labor Economics*, 30(4):783–828, 2012.

M. Angelucci and G. De Giorgi. Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption? *The American Economic Review*, pages 486–508, 2009.

J. Angrist and V. Lavy. The effects of high stakes high school achievement awards: Evidence from a randomized trial. *The American Economic Review*, 99(4):1384–1414, 2009.

J. D. Angrist. The perils of peer effects. *Labour Economics*, 30:98–108, 2014.

R. Baker, E. Bettinger, B. Jacob, and I. Marinescu. The effect of labor market information on community college students' major choice. Technical report, National Bureau of Economic Research, 2017.

A. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson. The diffusion of microfinance. *Science (New York, N.Y.)*, 341(6144):1236498, July 2013. ISSN 1095-9203. doi: 10.1126/science.1236498.

G. S. Becker. *Human capital and the personal distribution of income: An analytical approach*. Number 1. Institute of Public Administration, 1967.

Y. Ben-Porath. The production of human capital and the life cycle of earnings. *Journal of political economy*, 75(4, Part 1):352–365, 1967.

N. Bianchi. The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in STEM Majors. Jan. 2016. URL `https://www.scholars.northwestern.edu/en/publications/the-indirect-effects-of-educati`

G. J. Bobonis and F. Finan. Neighborhood peer effects in secondary school enrollment decisions. *The Review of Economics and Statistics*, 91(4):695–716, 2009.

T. Boneva and C. Rauh. Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. 2017.

G. Brunello, E. Crivellaro, and L. Rocco. Lost in transition? *Economics of Transition*, 20(4): 637–676, 2012.

T. Caplan, O. Furman, D. Romanov, and N. Zussman. The quality of Israeli academic institutions: What the wages of graduates tell about it? *Samuel Neaman Institution for Advanced Studies in Science and Technology. Haifa, Israel: Technion-Israel Institute of Technology*, 2006.

S. E. Carrell, B. I. Sacerdote, and J. E. West. From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81 (3):855–882, May 2013. ISSN 1468-0262. doi: 10.3982/ECTA10168. URL `http://onlinelibrary.wiley.com/doi/10.3982/ECTA10168/abstract`.

C.-H. Chen, V. Law, and W.-Y. Chen. The effects of peer competition-based science learning game on secondary students' performance, achievement goals, and perceived ability. *Interactive Learning Environments*, 26(2):235–244, 2018.

R. Chetty, J. N. Friedman, and J. E. Rockoff. Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, 104(9): 2633–2679, 2014a.

R. Chetty, N. Hendren, P. Kline, and E. Saez. Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129 (4):1553–1623, 2014b.

R. Chetty, N. Hendren, and L. F. Katz. The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *The American Economic Review*, 106(4):855–902, 2016.

T. Cornelissen, C. Dustmann, and U. Schönberg. Peer effects in the workplace. *American Economic Review*, 107(2):425–56, 2017.

B. Crépon, E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora. Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment*. *Quarterly Journal of Economics*, 128(2), 2013.

G. B. Dahl, K. V. Løken, and M. Mogstad. Peer Effects in Program Participation. *American Economic Review*, 104(7):2049–2074, July 2014. ISSN 0002-8282. doi: 10.1257/aer.104.7.2049. URL `https://www.aeaweb.org/articles?id=10.1257/aer.104.7.2049`.

E. Duflo, E. Saez, and others. The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *The Quarterly Journal of Economics*, 118(3):815–842, 2003.

S. M. Dynarski. Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review*, 93(1):279–288, 2003.

J. Feld and U. Zölitz. Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, 35(2):387–428, 2017.

R. Frish. The economic returns to schooling in Israel. *Israel Economic Review*, 7(1):113–141, 2009.

C. Gathmann, I. Helm, and U. Schönberg. Spillover effects of mass layoffs. *Journal of the European Economic Association*, 2016.

S. Getz. Surveys of changes in kibbutzim. *Institute for Research of the Kibbutz and the Cooperative Idea, University of Haifa, Reports*, 2004, 1998.

E. Hopkins and T. Kornienko. Running to keep in the same place: Consumer choice as a game of status. *American Economic Review*, 94(4):1085–1107, 2004.

C. Hoxby. Peer effects in the classroom: Learning from gender and race variation. Technical Report working paper 7867, National Bureau of Economic Research, 2000.

R. Jensen. The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2), 2010.

J. S. Joensen and H. S. Nielsen. Spillovers in education choice. *Journal of Public Economics*, 157:158–183, 2018.

J. R. Kling, J. B. Liebman, and L. F. Katz. Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119, 2007.

M. Kremer, E. Miguel, and R. Thornton. Incentives to learn. *The Review of Economics and Statistics*, 91(3):437–456, 2009.

I. Kuziemko, R. W. Buell, T. Reich, and M. I. Norton. “last-place aversion”: Evidence and redistributive implications. *The Quarterly Journal of Economics*, 129(1):105–149, 2014.

R. Lalive and M. A. Cattaneo. Social interactions and schooling decisions. *The Review of Economics and Statistics*, 91(3):457–477, 2009.

V. Lavy and A. Schlosser. Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2):1–33, 2011.

M. McGuigan, S. McNally, and G. Wyness. Student Awareness of Costs and Benefits of Educational Decisions: Effects of an Information Campaign. *Journal of Human Capital*, 10(4): 482–519, 2016.

E. Miguel and M. Kremer. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217, 2004.

D. Moreira. Success spills over. 2019.

R. Murphy and F. Weinhardt. Top of the class: The importance of ordinal rank. Technical report, National Bureau of Economic Research, 2018.

G. Navon. Human Capital Heterogeneity: University Choice and Wages. Technical Report 9708, University Library of Munich, Germany, Mar. 2006. URL `https://ideas.repec.org/p/pra/mprapa/9708.html`.

M. Nybom and J. Stuhler. Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *Journal of Human Resources*, 51(1):239–268, 2016.

P. Oreopoulos. Do dropouts drop out too soon? wealth, health and happiness from compulsory schooling. *Journal of public Economics*, 91(11-12):2213–2229, 2007.

M. Tincani. Heterogeneous peer effects and rank concerns: Theory and evidence. 2017.

A. Weiss. Human capital vs. signalling explanations of wages. *Journal of Economic perspectives*, 9(4):133–154, 1995.

Table 1: Short-Term Effects on High-School Outcomes

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
A. Short-Term Effects					
i. Simple diff-in-diff (N=7698)					
Treated X After	0.016	2.383	0.077	0.082	0.130
	(0.009)	(1.985)	(0.033)	(0.031)	(0.062)
ii. Controlled diff-in-diff (N=7698)					
Treated X After	0.018	2.759	0.088	0.094	0.149
	(0.008)	(1.838)	(0.030)	(0.029)	(0.057)
iii. Cross-sectional regression					
Treatment-control diff., before (N=3174)	-0.007	0.754	0.002	0.009	0.005
	(0.012)	(1.827)	(0.051)	(0.057)	(0.084)
Treatment-control diff., after (N=4524)	0.003	2.540	0.076	0.087	0.118
	(0.009)	(0.964)	(0.038)	(0.050)	(0.063)
B. Placebo Timing					
i. Simple diff-in-diff (N=5424)					
Treated X After	-0.012	-0.785	0.018	0.005	-0.012
	(0.014)	(1.651)	(0.033)	(0.023)	(0.061)
ii. Controlled diff-in-diff (N=5424)					
Treated X After	-0.007	-0.306	0.026	0.013	0.007
	(0.013)	(1.534)	(0.031)	(0.021)	(0.055)
iii. Cross-sectional regression					
Treatment-control diff., before (N=2463)	-0.005	0.162	-0.046	-0.043	-0.050
	(0.016)	(1.548)	(0.052)	(0.058)	(0.089)
Treatment-control diff., after (N=2961)	0.000	1.668	0.018	0.014	0.038
	(0.029)	(3.056)	(0.073)	(0.078)	(0.148)

Note: The dependent variable in column 1 is an indicator of whether the student completed high school; in column 2 it is her mean score in the matriculation exams; in column 3 it is an indicator of whether she received a matriculation certificate; in column 4 it is an indicator of whether she received a matriculation certificate that satisfies the requirements for university study; in column 5 is the summary index based on the outcomes in columns 1 to 4. In Panel A, the sample includes all the students (excluding kibbutzim members themselves) who attended schools with a positive number of either early or late reformed kibbutzim residents in both the before (1995/1996) and the after (1999/2000) periods. The first two rows of Panel A presents the estimated coefficients of interest in difference-in-differences regressions, comparing students in treated and untreated grades who are treated (10th grade in 1999/2000) and untreated (10th grade in 1995/1996). A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim. The simple difference-in-differences regressions include only cohort dummies and school fixed effects. The second panel of the table shows the controlled difference-in-differences, which also includes the following students demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries). The third row of Panel A shows the estimated effects using only the before (1995/1996) cohorts and using only the after (1999/2000) cohorts. Panel B reports the results of a placebo experiment in which we assume the early reforms happened in 1996 instead of 1998. We then use data from 1994-1995 and 1996-1997 to compare treated to control grades, before (1994-1995) and after (1996-1997) the placebo reforms. Standard errors clustered at the school level and presented in parentheses.

Table 2: Short-Term Effects on High-School Outcomes, by Intensity of Exposure

	High School Completion	Mean Matriculation Score	Matriculation Certification	University Qualified Matriculation	Summary Index
	(1)	(2)	(3)	(4)	(5)
A. Share of early reformers					
i. Simple diff-in-diff (N=7698)					
Share early reformers X After	0.120	8.422	0.053	0.058	0.310
	(0.053)	(9.217)	(0.169)	(0.175)	(0.313)
ii. Controlled diff-in-diff (N=7698)					
Share early reformers X After	0.106	7.123	0.039	0.037	0.259
	(0.052)	(9.039)	(0.153)	(0.165)	(0.297)
B. Categorical					
i. Simple diff-in-diff (N=7698)					
1st quartile X after	0.004	1.038	0.029	0.050	0.057
	(0.013)	(3.205)	(0.060)	(0.053)	(0.103)
2nd quartile X after	0.009	3.032	0.129	0.127	0.179
	(0.012)	(2.404)	(0.059)	(0.055)	(0.095)
3rd quartile X after	0.025	3.176	0.041	0.051	0.116
	(0.018)	(2.499)	(0.035)	(0.035)	(0.076)
4th quartile X after	0.029	1.257	0.063	0.054	0.112
	(0.011)	(2.517)	(0.036)	(0.038)	(0.075)
ii. Controlled diff-in-diff (N=7698)					
1st quartile X after	0.008	1.808	0.049	0.071	0.093
	(0.013)	(2.870)	(0.050)	(0.043)	(0.089)
2nd quartile X after	0.011	3.459	0.142	0.141	0.200
	(0.010)	(1.922)	(0.044)	(0.041)	(0.071)
3rd quartile X after	0.026	3.636	0.057	0.068	0.140
	(0.019)	(2.601)	(0.033)	(0.033)	(0.080)
4th quartile X after	0.029	1.144	0.063	0.053	0.110
	(0.012)	(2.374)	(0.032)	(0.035)	(0.071)

Note: The dependent variables in this table are the same as in Table 1. In the first row of Panel A, we report the estimated coefficients of interest in difference-in-differences regressions using the share of early reformers in the grade as our measure of treatment intensity, and including only cohort dummies and school fixed effects. In the second row of Panel B, we instead replace the treatment indicator with four dummies corresponding to quartiles of the share of early reformers on the grade. The controlled difference-in-differences rows also include the following students demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries).

Table 3: Short-Term Effects on High-School Outcomes, by Gender and Mother's Education

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index
i. Sample Stratification by Gender					
Male (N=3917)	0.038	5.663	0.122	0.109	0.242
	(0.016)	(1.844)	(0.033)	(0.032)	(0.055)
Female (N=3781)	-0.002	-0.438	0.053	0.076	0.059
	(0.014)	(2.234)	(0.052)	(0.049)	(0.088)
p-value	0.096	0.001	0.260	0.564	0.034
ii. Sample Stratification by Mother's Education					
Low (N=4156)	0.022	3.609	0.121	0.123	0.202
	(0.015)	(2.057)	(0.034)	(0.035)	(0.068)
High (N=3542)	0.015	1.751	0.045	0.053	0.093
	(0.012)	(2.253)	(0.041)	(0.037)	(0.078)
p-value	0.755	0.325	0.076	0.061	0.163

Note: This table presents the same results as in Table 1 but estimated separately for males and females (panel i) and for low and high mother's education (panel ii). We also report the p-value corresponding to the null hypothesis that the effects are the same in both subsamples.

Table 4: Long-Term Effects on Post-Secondary Schooling Outcomes

	All post secondary		University		College	
	Enrollment (1)	Years of schooling (2)	Enrollment (3)	Years of schooling (4)	Enrollment (5)	Years of schooling (6)
i. Full sample (N=7555)						
Simple diff-in-diff	0.040	0.235	0.087	0.487	-0.043	-0.165
	(0.021)	(0.191)	(0.034)	(0.183)	(0.036)	(0.113)
Controlled diff-in-diff	0.045	0.279	0.094	0.527	-0.043	-0.164
	(0.020)	(0.183)	(0.034)	(0.182)	(0.037)	(0.117)
ii. Stratification by gender						
Male (N=3851)	0.077	0.475	0.101	0.579	-0.037	-0.107
	(0.030)	(0.203)	(0.033)	(0.206)	(0.051)	(0.130)
Female (N=3704)	0.009	0.052	0.081	0.442	-0.054	-0.247
	(0.024)	(0.227)	(0.057)	(0.235)	(0.035)	(0.131)
p-value	0.038	0.082	0.380	0.330	0.391	0.224
iii. Stratification by mother's education						
Low (N=4062)	0.033	0.135	0.082	0.478	-0.068	-0.311
	(0.035)	(0.161)	(0.029)	(0.161)	(0.061)	(0.169)
High (N=3493)	0.053	0.467	0.101	0.609	-0.012	0.024
	(0.031)	(0.231)	(0.044)	(0.209)	(0.027)	(0.090)
p-value	0.334	0.119	0.359	0.309	0.200	0.040

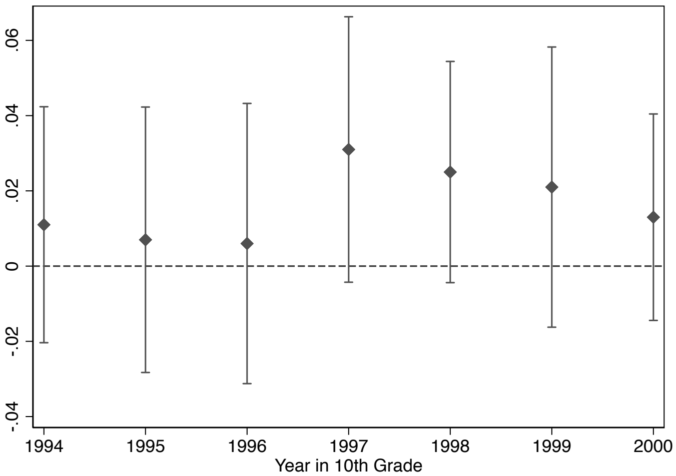
Note: The dependent variables in columns 1 and 2 are an indicator whether a student ever enrolled in any post-secondary education, and the total years of schooling in any post-secondary education 13 years after high-school graduation; In columns 3 and 4 these are an indicator whether a student ever enrolled in a university, and total years of schooling in university 13 years after high-school graduation; In columns 5 and 6 the dependent variables are an indicator whether a student ever enrolled in a college, and total years of schooling in college 13 years after high-school graduation. We also report the p-value corresponding to the null hypothesis that the effects are the same in both subsamples. Standard errors clustered at the school level and presented in parentheses.

Table 5: Long-Term Effects on Labor Market Outcomes

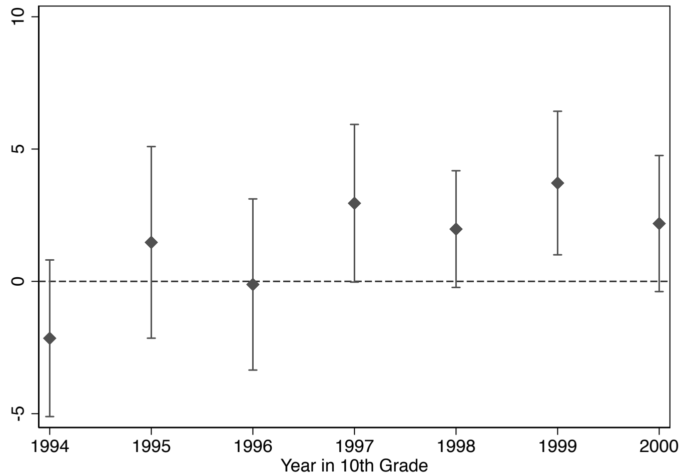
	Labor market			Unemployment benefits		
	Employ-ment	Work-months	Earn-ings	Unemployed indicator	Total benefits	Number of months
	(1)	(2)	(3)	(4)	(5)	(6)
i. Full sample (N=7546)						
Simple diff-in-diff	0.022	0.277	7614.4	-0.014	-293.2	-0.089
	(0.016)	(0.209)	(3536.5)	(0.009)	(132.4)	(0.039)
Controlled diff-in-diff	0.027	0.322	6988.8	-0.015	-309.2	-0.089
	(0.016)	(0.223)	(3518.5)	(0.009)	(134.6)	(0.040)
ii. Stratification by gender						
Male (N=3847)	0.043	0.508	6919.1	-0.006	-112.6	-0.043
	(0.026)	(0.324)	(6419.1)	(0.012)	(174.7)	(0.042)
Female (N=3699)	0.009	0.150	8283.7	-0.024	-512.7	-0.141
	(0.023)	(0.238)	(4283.7)	(0.014)	(269.4)	(0.066)
p-value	0.163	0.186	0.429	0.164	0.106	0.105
iii. Stratification by mother's education						
Low (N=4057)	0.004	-0.017	3940.5	0.001	-95.9	-0.064
	(0.022)	(0.254)	(4037.7)	(0.015)	(208.1)	(0.064)
High (N=3489)	0.057	0.793	11847.1	-0.032	-545.7	-0.110
	(0.020)	(0.325)	(5331.6)	(0.010)	(159.9)	(0.038)
p-value	0.037	0.024	0.118	0.033	0.043	0.268

Note: The dependent variables in columns 1, 2 and 3 are an indicator of whether the student was in the labor force, number of work months and her annual earnings in 2009 Israeli NIS 12 years after high-school graduation; In columns 4, 5 and 6 these are an indicator whether the student is entitled to unemployment benefits, number of months receiving unemployment benefits and total unemployment benefits in 2009 Israeli NIS in year 2012. We also report the p-value corresponding to the null hypothesis that the effects are the same in both subsamples. Standard errors clustered at the school level and presented in parentheses.

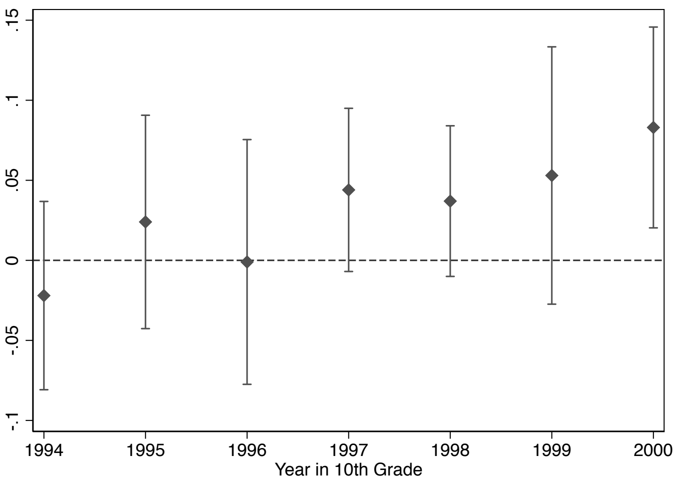
Figure 1: Treatment-Control Differences in High-School Outcomes



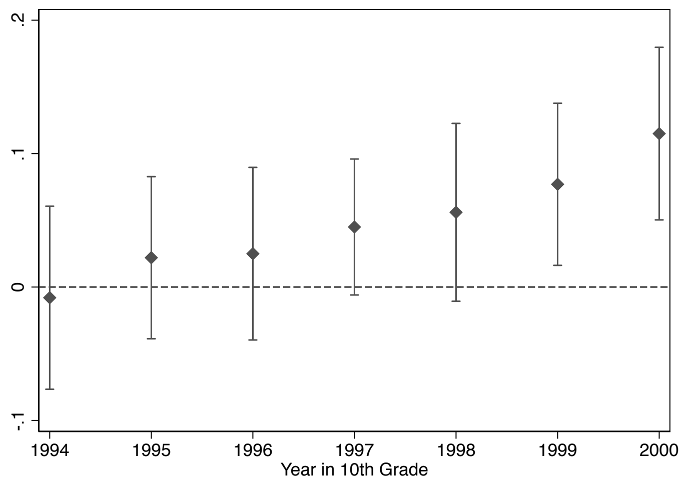
(a) High school graduation



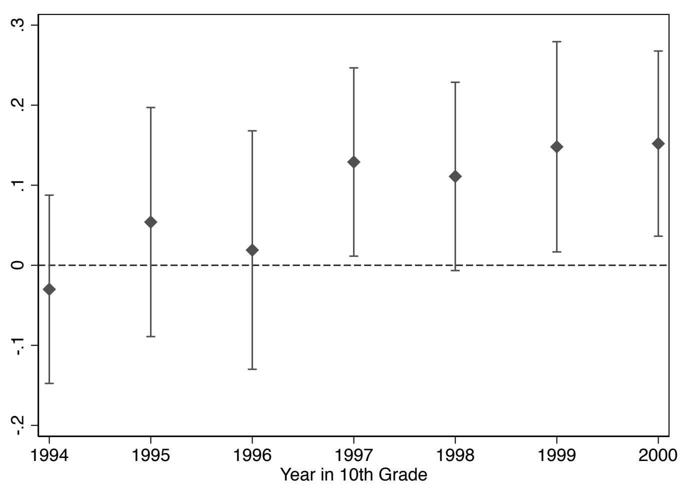
(b) Mean matriculation score



(c) Matriculation certification



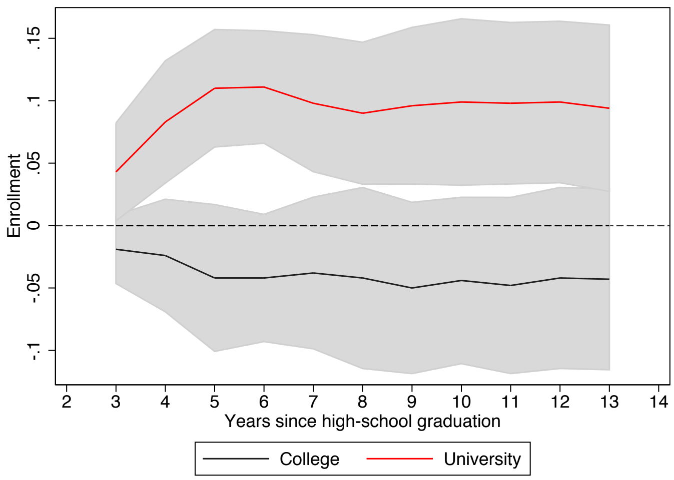
(d) University qualified matriculation



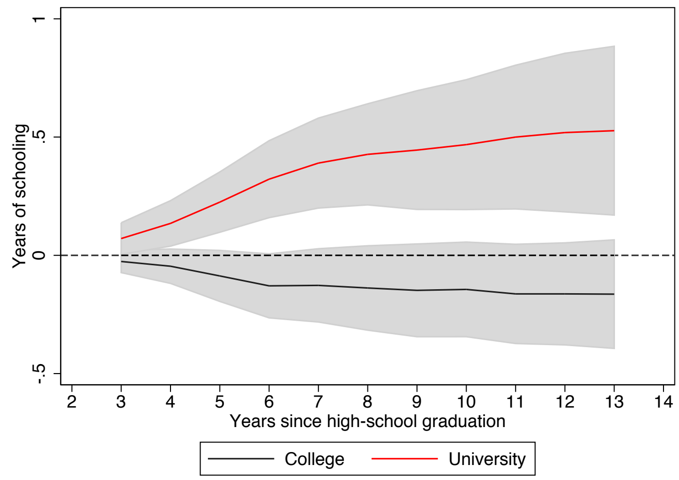
(e) Summary Index

Note: The dependent variable in panel (a) is an indicator of whether the student completed high school; in panel (b) it is her mean score in the matriculation exams; in panel (c) it is an indicator of whether she received a matriculation certificate; in panel (d) it is an indicator of whether she received a matriculation certificate that satisfies the requirements for university study; in panel (e) is the summary index based on the outcomes in panels (a) to (d). The sample includes all the students (excluding kibbutzim members themselves) who started high school from 1994 to 2000 and who were in grades with a positive number of either early or late reformed kibbutzim residents. A grade (school/year combination) is defined as treated if it includes students from early reformed

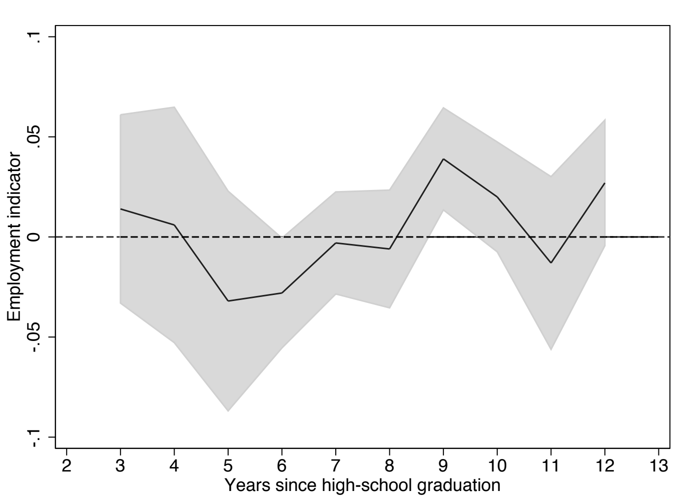
Figure 2: Long Term Effects on Post-Secondary Schooling and Labor Market Outcomes, by Years Since High-School Graduation



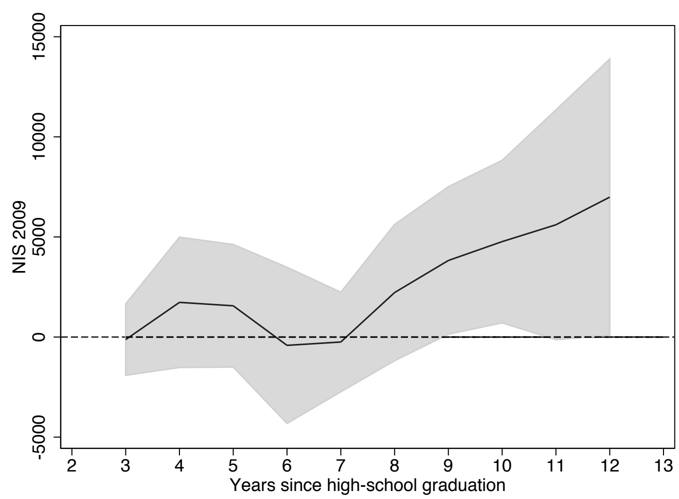
(a) Enrollment



(b) Years of schooling



(c) Employment indicator



(d) Annual earnings

Note: We plot the estimated effects from 3 to 13 years after high-school graduation. The dependent variable on panel (a) is an indicator that takes a value of one if the student was ever enrolled in post-secondary schooling by the corresponding year. The dependent variable in panel (b) is the years of post-secondary schooling completed by the corresponding year. The dependent variable on panel (c) is an indicator that takes a value of one if the student was part of the labor force in the corresponding year. The dependent variable in panel (d) are annual earnings in 2009 Israeli NIS in the corresponding year.

Online Appendix - Not for publication

Table A1: Comparison between direct and spillover effects

Article	**Program**	**Outcome**	**Indirectly treated group**	**Spillovers as % of direct effect**
Alderman, Kim and Orazem (World Bank Economic Review, 1999)	Providing subsidies for girls to enroll in private schools	School Enrollment	Boys	60-100
Angelucci and De Giorgi (AER, 2009)	PROGRESA Conditional cash transfer	Food consumption	Ineligible individuals in treated villages	50
Bobonis and Finan (RESTAT, 2009)	PROGRESA Conditional cash transfer	School Enrollment	Ineligible individuals in treated villages	50-80
Dahl, Loken, and Mogstad, (AER, 2014) Duflo and Saez (QJE, 2003)	Paid paternity leave Invitation to attend a benefits information fair	Take up Enrollment in retirement plan	Coworkers and siblings of eligible workers Coworkers of those compelled to attend the fair	11 100
Joensen and Nielsen (JPubE, 2018)	Lowering the cost to enroll in high school STEM courses	Enrollment in STEM courses	Siblings of treated individuals	30-50
Kremer, Miguel and Thornton (RESTAT, 2009)	Merit scholarship for girls	Test scores	Boys	30-50
Lalive and Cattaneo (RESTAT, 2009)	PROGRESA Conditional cash transfer	School enrollment	Ineligible individuals in treated villages	30-50
Moreira (unpublished, 2019)	Receiving an honorable mention in Math Olympiad	Academic performance	Classmates of winners	20

Notes: This table provides examples of studies documenting sizable spillover effects of social programs.

Table A2: Sample Size

	Full		Treated		Control	
	Before	After	Before	After	Before	After
Number of Schools	31	31	.	.	.	.
Number of Grades (school/years)	61	62	48	52	13	10
Number of Students						
I. Peers	3177	4529	2052	3379	1125	1150
II. Kibbutzniks						
i. Early reformers	999	905	999	905	0	0
ii. Late reformers	502	487	390	400	112	87

Note: This table shows the number of schools and number of treatment and control grades in our baseline sample. A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim.

Table A3: Descriptive Statistics, Balancing and Post-Treatment Differences

	10th Grade Students in 1995 and 1996				10th Grade Students in 1999 and 2000			
	Full	Treatment	Control	Difference	Full	Treatment	Control	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Background characteristics								
Male Indicator	0.512	0.499	0.536	-0.036	0.507	0.505	0.512	-0.010
	(0.500)	(0.500)	(0.499)	(0.018)	(0.500)	(0.500)	(0.500)	(0.021)
Father Years of Schooling	13.449	13.683	13.022	0.658	13.653	13.753	13.358	0.392
	(3.459)	(3.251)	(3.773)	(0.559)	(3.459)	(3.096)	(3.554)	(0.506)
Mother Years of Schooling	13.523	13.735	13.136	0.604	13.926	14.060	13.535	0.535
	(3.114)	(2.964)	(3.336)	(0.385)	(3.114)	(2.893)	(3.231)	(0.402)
Number of Siblings	2.449	2.374	2.587	-0.213	2.300	2.321	2.238	0.086
	(1.361)	(1.213)	(1.587)	(0.329)	(1.361)	(1.165)	(1.134)	(0.129)
Asia-Africa Ethnicity	0.212	0.195	0.244	-0.050	0.184	0.176	0.209	-0.036
	(0.409)	(0.396)	(0.429)	(0.063)	(0.409)	(0.381)	(0.407)	(0.061)
Europe-America Ethnicity	0.213	0.225	0.192	0.032	0.192	0.201	0.166	0.033
	(0.410)	(0.418)	(0.394)	(0.023)	(0.410)	(0.401)	(0.372)	(0.018)
Other Ethnicity	0.005	0.004	0.008	-0.004	0.009	0.010	0.006	0.005
	(0.073)	(0.062)	(0.089)	(0.003)	(0.073)	(0.100)	(0.078)	(0.004)
Former Soviet Union Ethnicity	0.053	0.047	0.064	-0.018	0.056	0.056	0.055	0.004
	(0.224)	(0.211)	(0.245)	(0.038)	(0.224)	(0.230)	(0.228)	(0.018)
Ethiopia Ethnicity	0.007	0.002	0.015	-0.012	0.017	0.012	0.032	-0.019
	(0.083)	(0.049)	(0.122)	(0.007)	(0.083)	(0.108)	(0.177)	(0.016)
Family Income	24.117	23.322	25.674	-2.404	24.018	23.965	24.179	-0.156
	(24.898)	(24.716)	(25.191)	(3.808)	(20.952)	(20.841)	(21.295)	(3.399)
B. High School Outcomes								
High School Completion	0.955	0.953	0.960	-0.007	0.961	0.962	0.958	0.003
	(0.207)	(0.212)	(0.196)	(0.013)	(0.207)	(0.190)	(0.200)	(0.009)
Mean Matriculation Score	70.892	71.150	70.421	0.777	74.010	74.665	72.084	2.551
	(21.609)	(21.711)	(21.423)	(1.851)	(21.609)	(19.830)	(19.977)	(0.961)
Matriculation Certification	0.616	0.616	0.614	0.002	0.687	0.707	0.630	0.076
	(0.487)	(0.486)	(0.487)	(0.051)	(0.487)	(0.455)	(0.483)	(0.038)
University Qualified Matriculation	0.575	0.578	0.570	0.009	0.632	0.654	0.568	0.088
	(0.494)	(0.494)	(0.495)	(0.057)	(0.494)	(0.476)	(0.496)	(0.050)
Observations	3177	2052	1125		4529	3379	1150	

	10th Grade Students in 1995 and 1996				10th Grade Students in 1999 and 2000			
	Full	Treatment	Control	Difference	Full	Treatment	Control	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C. Post-secondary education								
Post-secondary years of schooling	2.969	2.928	3.045	-0.124	2.958	2.996	2.843	0.162
	(2.539)	(2.543)	(2.531)	(0.344)	(2.464)	(2.471)	(2.438)	(0.215)
Post-secondary Enrollment	0.705	0.695	0.724	-0.030	0.713	0.719	0.696	0.023
	(0.456)	(0.461)	(0.447)	(0.049)	(0.452)	(0.450)	(0.460)	(0.044)
University years of schooling	1.554	1.478	1.698	-0.236	1.400	1.448	1.253	0.215
	(2.426)	(2.381)	(2.503)	(0.303)	(2.273)	(2.299)	(2.187)	(0.181)
University Enrollment	0.366	0.348	0.399	-0.054	0.345	0.354	0.319	0.039
	(0.482)	(0.477)	(0.490)	(0.067)	(0.475)	(0.478)	(0.466)	(0.049)
College years of schooling	0.981	1.015	0.917	0.103	1.141	1.141	1.140	0.002
	(1.625)	(1.654)	(1.567)	(0.078)	(1.668)	(1.672)	(1.656)	(0.118)
College Enrollment	0.332	0.342	0.313	0.031	0.390	0.391	0.388	0.003
	(0.471)	(0.475)	(0.464)	(0.023)	(0.488)	(0.488)	(0.488)	(0.032)
D. Labor market								
Employment	0.775	0.767	0.789	-0.021	0.793	0.790	0.801	-0.005
	(0.418)	(0.423)	(0.408)	(0.015)	(0.405)	(0.407)	(0.400)	(0.012)
Annual earnings	7.787	7.636	8.075	-0.451	7.728	7.769	7.602	0.303
	(7.719)	(7.960)	(7.233)	(0.459)	(7.135)	(7.078)	(7.304)	(0.519)
Unempyomnet indicator	0.063	0.067	0.056	0.010	0.052	0.051	0.055	-0.006
	(0.243)	(0.249)	(0.231)	(0.007)	(0.222)	(0.221)	(0.227)	(0.010)
Number of months of UI benefits	0.207	0.229	0.164	0.067	0.172	0.167	0.186	-0.033
	(0.896)	(0.961)	(0.754)	(0.030)	(0.818)	(0.803)	(0.862)	(0.047)
Total unemployment benefits	861.025	941.263	708.420	241.263	576.043	567.320	602.229	-88.927
	(3924.410)	(4135.073)	(3485.447)	(125.897)	(2949.807)	(2931.195)	(3006.102)	(156.030)
Observations	3177	2052	1125		4529	3379	1150	

Note: Columns 1 and 5 present means and standard deviations (in parentheses) of background characteristics and outcomes of students before and after the early reforms. Columns 2, 3, 6 and 7 present the means and standard deviations for students in treatment and control grades for affected (1999-2000) and unaffected (1995-1996) cohorts of 10th graders. Columns 4 and 8 present the differences between treatment and control grades, controlling for cohort fixed effects. Family income and annual earnings are in ten thousands NIS. Standard errors of these differences clustered at the school level are given in parentheses.

Table A4: No Change in Background Characteristics of Peers as a Result of the Reform

	Treated X after (1)
i. Full Sample (N=7706)	
Male Indicator	0.029
	(0.030)
Father Years of Schooling	-0.174
	(0.180)
Mother Years of Schooling	0.018
	(0.201)
Number of Siblings	0.378
	(0.385)
Europe-America Ethnicity	0.009
	(0.017)
Other Ethnicity	0.010
	(0.005)
Former Soviet Union Ethnicity	0.001
	(0.017)
Ethiopia Ethnicity	-0.015
	(0.011)
Family income	2.0593
	(1.0245)

Note: Each row corresponds to a separate regression for each of the student's background characteristics on an interaction between the treatm
and an indicator corresponding to cohorts who started school after the early reforms (1999/2000), as described in the main text.

Table A5: Descriptive Statistics, Balancing and Post-Treatment Differences, Expanded Control Group

	10th Grade Students in 1995 and 1996				10th Grade Students in 1999 and 2000			
	Full	Treatment	Control	Difference	Full	Treatment	Control	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Background characteristics								
Male Indicator	0.452	0.456	0.448	0.006	0.450	0.459	0.433	0.028
	(0.498)	(0.498)	(0.497)	(0.060)	(0.498)	(0.498)	(0.496)	(0.080)
Father Years of Schooling	13.397	13.650	13.180	0.467	13.395	13.765	12.697	1.058
	(3.871)	(3.706)	(3.995)	(0.408)	(4.018)	(3.497)	(4.772)	(0.826)
Mother Years of Schooling	13.320	13.588	13.091	0.494	13.494	13.884	12.762	1.105
	(3.410)	(3.318)	(3.471)	(0.330)	(3.615)	(3.077)	(4.361)	(0.788)
Number of Siblings	2.642	2.496	2.767	-0.271	2.693	2.746	2.594	0.155
	(1.593)	(1.429)	(1.711)	(0.199)	(1.606)	(1.622)	(1.570)	(0.254)
Asia-Africa Ethnicity	0.234	0.205	0.259	-0.053	0.197	0.195	0.200	-0.005
	(0.424)	(0.404)	(0.438)	(0.027)	(0.397)	(0.396)	(0.400)	(0.028)
Europe-America Ethnicity	0.214	0.232	0.198	0.036	0.195	0.199	0.186	0.014
	(0.410)	(0.422)	(0.399)	(0.022)	(0.396)	(0.400)	(0.390)	(0.018)
Other Ethnicity	0.011	0.010	0.011	-0.002	0.016	0.016	0.014	0.002
	(0.103)	(0.099)	(0.106)	(0.003)	(0.124)	(0.127)	(0.117)	(0.005)
Former Soviet Union Ethnicity	0.065	0.054	0.075	-0.020	0.071	0.065	0.083	-0.019
	(0.247)	(0.226)	(0.263)	(0.026)	(0.257)	(0.246)	(0.276)	(0.022)
Ethiopia Ethnicity	0.009	0.004	0.013	-0.009	0.018	0.015	0.023	-0.008
	(0.094)	(0.063)	(0.114)	(0.003)	(0.133)	(0.123)	(0.150)	(0.009)
B. High School Outcomes								
High School Completion	0.951	0.942	0.958	-0.017	0.956	0.954	0.961	-0.007
	(0.217)	(0.234)	(0.201)	(0.010)	(0.205)	(0.210)	(0.193)	(0.009)
Mean Matriculation Score	70.394	69.591	71.082	-1.719	73.304	73.385	73.153	0.236
	(22.667)	(23.703)	(21.720)	(1.877)	(21.258)	(22.119)	(19.538)	(2.018)
Matriculation Certification	0.630	0.610	0.647	-0.040	0.682	0.694	0.659	0.035
	(0.483)	(0.488)	(0.478)	(0.044)	(0.466)	(0.461)	(0.474)	(0.047)
University Qualified Matriculation	0.587	0.565	0.605	-0.043	0.627	0.641	0.600	0.040
	(0.492)	(0.496)	(0.489)	(0.045)	(0.484)	(0.480)	(0.490)	(0.046)
Observations	8045	3711	4334		9471	6183	3288	

Note: Columns 1 and 5 present means and standard deviations (in parentheses) of characteristics and outcomes of students before and after the early reforms. Columns 2, 3, 6 and 7 present the means and standard deviations for students in treatment and control grades for affected (1999-2000) and unaffected (1995-1996) cohorts of 10th graders. Columns 4 and 8 present the differences between treatment and control grades, controlling for cohort fixed effects. Standard errors of these differences clustered at the school level are given in parentheses.

Table A6: Short-Term Effects on Type of Subjects Taken in High School

	# of Credit Units Received in Bagrut (1)	# of Credit Units in English (2)	# of Credit Units in Math (3)	# of Subjects in High School (4)	# of Non-Science Subjects in High School (5)	# of Science Subjects in High School (6)
i. Simple diff-in-diff (N=7435)						
Treated X After	1.722	0.128	0.241	0.593	-0.175	0.768
	(0.795)	(0.038)	(0.097)	(0.276)	(0.161)	(0.305)
ii. Controlled diff-in-diff (N=7435)						
Treated X After	1.845	0.155	0.261	0.616	-0.205	0.821
	(0.748)	(0.035)	(0.083)	(0.256)	(0.175)	(0.293)
Mean dependent variable	23.045	4.215	3.190	8.001	4.642	3.359

Note: The first panel of the table presents the estimated coefficients of interest in difference-in-differences regressions, comparing students in treated and untreated grades who are treated (10th grade in 1999/2000) and untreated (10th grade in 1995/1996). A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim. The simple difference-in-differences regressions include only cohort dummies and school fixed effects. The second panel of the table shows the controlled difference-in-differences, which also includes the following students demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries).

Table A7: Direct Effects on Kibbutz Students

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
A. All grades					
i. Full Sample (N=3349)					
Simple difference-in-differences	0.033	3.617	0.048	0.060	0.135
	(0.017)	(1.625)	(0.035)	(0.037)	(0.056)
Controlled difference-in-differences	0.033	3.546	0.049	0.060	0.134
	(0.017)	(1.605)	(0.035)	(0.036)	(0.054)
B. No grades with both early/late reformers					
Simple difference-in-differences	0.038	9.333	0.138	0.173	0.311
	(0.030)	(3.197)	(0.061)	(0.059)	(0.105)
Controlled difference-in-differences	0.038	8.968	0.127	0.155	0.293
	(0.031)	(3.048)	(0.065)	(0.060)	(0.101)
N=963					

Notes: The dependent variable in column 1 is an indicator of whether the student completed high school; in column 2 it is her mean score in the matriculation exams; in column 3 it is an indicator of whether she received a matriculation certificate; in column 4 it is an indicator of whether she received a matriculation certificate that satisfies the requirements for university study; the dependent variable in column 5 is an outcome index that receives the following values: 0 if the student drops out of school, 1 if the student graduates without receiving matriculation certification, 2 if the student receives a matriculation certification, and 3 if the student receives a matriculation certification that is university qualified. The simple difference-in-differences regressions include only cohort dummies and kibbutz fixed effects. The controlled difference-in-differences regressions include cohort dummies, kibbutz fixed effects, and the following student's demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries).

Table A8: Short-Term Effects on High-School Outcomes, Expanded Control Group

	High School Completion	Mean Matriculation Score	Matriculation Certification	University Qualified Matriculation	Summary Index
	(1)	(2)	(3)	(4)	(5)
i. Simple diff-in-diff (N=17516)					
Treated X After	0.012	2.453	0.086	0.092	0.138
	(0.009)	(1.277)	(0.026)	(0.025)	(0.045)
ii. Controlled diff-in-diff (N=17516)					
Treated X After	0.009	1.886	0.071	0.075	0.112
	(0.009)	(1.287)	(0.023)	(0.024)	(0.044)
iii. Cross-sectional regression					
Treatment-control diff., before (N=8045)	-0.017	-1.719	-0.040	-0.043	-0.087
	(0.010)	(1.877)	(0.044)	(0.045)	(0.077)
Treatment-control diff., after (N=9471)	-0.007	0.236	0.035	0.040	0.032
	(0.009)	(2.018)	(0.047)	(0.046)	(0.083)

Note: The dependent variable in column 1 is an indicator of whether the student completed high school; in column 2 it is her mean score in the matriculation exams; in column 3 it is an indicator of whether she received a matriculation certificate; in column 4 it is an indicator of whether she received a matriculation certificate that satisfies the requirements for university study; in column 5 is the summary index based on the outcomes in columns 1 to 4. The sample includes all the students (excluding kibbutzim members themselves) who attended schools with a positive number of either early or late/never reformed kibbutzim residents in both the before (1995/1996) and the after (1999/2000) periods. The first two rows present the estimated coefficients of interest in difference-in-differences regressions, comparing students in treated and untreated grades who are treated (10th grade in 1999/2000) and untreated (10th grade in 1995/1996). A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim. The simple difference-in-differences regressions include only cohort dummies and school fixed effects. The second panel of the table shows the controlled difference-in-differences, which also includes the following students demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries). The third row shows the estimated effects using only the before (1995/1996) cohorts and using only the after (1999/2000) cohorts.

Table A9: Short-Term Effects on High-School Outcomes, by Intensity of Exposure, Expanded Control Group

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
A. Share of early reformers					
i. Simple diff-in-diff (N=17516)					
Share early reformers X After	0.122	12.514	0.293	0.288	0.618
	(0.043)	(7.296)	(0.139)	(0.149)	(0.251)
ii. Controlled diff-in-diff (N=17516)					
Share early reformers X After	0.102	9.037	0.209	0.188	0.453
	(0.043)	(7.397)	(0.127)	(0.142)	(0.246)
B. Categorical					
i. Simple diff-in-diff (N=17516)					
1st quartile X after	0.012	1.777	0.046	0.060	0.092
	(0.014)	(2.194)	(0.040)	(0.037)	(0.079)
2nd quartile X after	-0.021	-0.002	0.062	0.088	0.050
	(0.026)	(2.227)	(0.051)	(0.053)	(0.099)
3rd quartile X after	0.015	3.697	0.130	0.127	0.200
	(0.011)	(1.475)	(0.044)	(0.042)	(0.066)
4th quartile X after	0.035	3.631	0.094	0.091	0.187
	(0.010)	(1.949)	(0.032)	(0.034)	(0.058)
ii. Controlled diff-in-diff (N=17516)					
1st quartile X after	0.009	1.229	0.033	0.042	0.066
	(0.014)	(2.165)	(0.039)	(0.038)	(0.078)
2nd quartile X after	-0.021	-0.127	0.057	0.084	0.044
	(0.023)	(1.842)	(0.046)	(0.048)	(0.079)
3rd quartile X after	0.012	2.884	0.111	0.104	0.163
	(0.011)	(1.358)	(0.039)	(0.036)	(0.059)
4th quartile X after	0.030	2.793	0.073	0.066	0.147
	(0.011)	(2.052)	(0.029)	(0.034)	(0.061)

Note: The dependent variables in this table are the same as in table 1. We replace the treatment indicator with four dummies corresponding to quartiles of the share of early reformers on the grade. Each rows corresponds to the estimated coefficient of interest in a difference-in-differences regression.

Table A10: Short-Term Effects on High-school Outcomes, No School Fixed Effects

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
i. Full Sample (N=7698)					
Simple diff-in-diff	0.010	1.774	0.075	0.079	0.113
	(0.011)	(1.953)	(0.034)	(0.033)	(0.064)
Controlled diff-in-diff	0.015	2.391	0.089	0.097	0.142
	(0.009)	(1.772)	(0.029)	(0.029)	(0.054)

Note: This table replicates the results in table 1 without including school fixed effects to the regression.

Table A11: Short-Term Effects on High-School Outcomes, Controlling for Family Income

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Sum-mary Index (5)
i. Simple diff-in-diff (N=7178)					
Treated X After	0.016	2.387	0.078	0.083	0.131
	(0.009)	(2.001)	(0.033)	(0.031)	(0.062)
ii. Controlled diff-in-diff (N=7178)					
Treated X After	0.020	2.695	0.077	0.079	0.137
	(0.009)	(1.997)	(0.032)	(0.030)	(0.061)
iii. Cross-sectional regression					
Treatment-control diff., before (N=2956)	-0.007	0.777	0.002	0.009	0.006
	(0.013)	(1.851)	(0.051)	(0.057)	(0.085)
Treatment-control diff., after (N=4222)	0.003	2.551	0.076	0.088	0.118
	(0.009)	(0.961)	(0.038)	(0.050)	(0.063)
Mean dependent variable	0.955	70.892	0.616	0.575	-0.009

Note: This table replicates the results in Table 1 adding family income as an additional control variable. Sample is restricted to students whose parents had no missing earnings data.

Table A12: Short-Term Effects on High-School Outcomes, Instrumental Variables Model

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
i. Simple diff-in-diff (N=7706)					
Grade-level treatment	0.016	2.387	0.078	0.083	0.135
	(0.009)	(2.001)	(0.033)	(0.031)	(0.064)
Class-level treatment IV	0.018	2.840	0.089	0.096	0.156
	(0.013)	(2.503)	(0.073)	(0.065)	(0.111)
ii. Controlled diff-in-diff (N=7706)					
Grade-level treatment	0.018	2.759	0.088	0.094	0.153
	(0.008)	(1.838)	(0.030)	(0.029)	(0.058)
Class-level treatment IV	0.021	3.338	0.102	0.110	0.181
	(0.011)	(2.194)	(0.065)	(0.057)	(0.094)
Mean dependent variable	0.955	70.892	0.616	0.575	-0.087

Note: This table reports an exercise in which we instrument a class-level treatment indicator with the grade-level indicator. More precisely, we define a treatment indicator that takes a value of 1 if there is a positive number of early reformers in the class, and a treatment indicator that takes a value of 1 if there is a positive number of early reformers in the grade, as well as their respective interactions with an indicator corresponding to the treated cohorts. The table presents the estimated coefficients of interest in a difference-in-differences regressions comparing students in treated and untreated classes who are treated (10th grade in 1999/2000) and untreated (10th grade in 1995/1996). The outcome variables are the same as in Table 1.

Table A13: Short-Term Effects on High-School Outcomes, by Minimum Number of Peers from Reformed Kibbutzim

	High School Completion	Mean Matriculation Score	Matriculation Certification	University Qualified Matriculation	Summary Index
	(1)	(2)	(3)	(4)	(5)
A. At least 3 kibbutzniks					
i. Simple diff-in-diff (N=7130)					
Treated X After	0.019	3.235	0.085	0.089	0.152
	(0.009)	(2.057)	(0.034)	(0.032)	(0.063)
ii. Controlled diff-in-diff (N=7130)					
Treated X After	0.021	3.597	0.096	0.100	0.170
	(0.009)	(1.892)	(0.029)	(0.030)	(0.058)
Mean dependent variable	0.956	71.141	0.622	0.582	0.000
B. At least 6 kibbutzniks					
i. Simple diff-in-diff (N=6344)					
Treated X After	0.016	1.291	0.072	0.072	0.109
	(0.009)	(1.776)	(0.039)	(0.035)	(0.064)
ii. Controlled diff-in-diff (N=6344)					
Treated X After	0.018	1.661	0.079	0.077	0.122
	(0.009)	(1.635)	(0.035)	(0.032)	(0.059)
Mean dependent variable	0.954	70.712	0.621	0.584	-0.008

Note: This table replicates the results in Table 1 using two alternative samples. In the first panel, the sample is restricted to grades with at least 3 students from reformed kibbutzim. In the second panel, the sample is restricted to grades with at least 6 students from reformed kibbutzim.

Table A14: Descriptive Statistics: Treatment indicator (1 if Early Reformed > 0 and Late Reformed $= 0$)

	Full		Treated		Control	
	Before	After	Before	After	Before	After
Number of Schools	18	16	.	.	.	.
Number of Grades (school/years)	27	27	14	17	13	10
Number of Students						
I. Peers	1675	2285	550	1135	1125	1150
II. Kibbutzniks						
i. Early reformers	175	232	175	232	0	0
ii. Late reformers	112	87	0	0	112	87

Note: A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim. Kibbutzniks peers are those who share a grade with kibbutz members from early or late reformed kibbutzim.

Table A15: Descriptive Statistics, Balancing and Post-Treatment Differences (1 if Early Reformed >0 and Late Reformed $=0$)

	10th Grade Students in 1995 and 1996				10th Grade Students in 1999 and 2000			
	Full	Treat-ment	Con-trol	Differ-ence	Full	Treat-ment	Con-trol	Differ-ence
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Student's characteristics								
Male Indicator	0.512	0.476	0.536	-0.060	0.507	0.514	0.512	-0.004
	(0.500)	(0.500)	(0.499)	(0.018)	(0.500)	(0.500)	(0.500)	(0.025)
Father Years of Schooling	13.449	14.403	13.022	1.372	13.653	14.372	13.358	1.000
	(3.459)	(2.965)	(3.773)	(0.610)	(3.459)	(3.038)	(3.554)	(0.623)
Mother Years of Schooling	13.523	14.381	13.136	1.240	13.926	14.461	13.535	0.933
	(3.114)	(2.695)	(3.336)	(0.411)	(3.114)	(2.811)	(3.231)	(0.488)
Number of Siblings	2.449	2.224	2.587	-0.360	2.300	2.342	2.238	0.112
	(1.361)	(1.122)	(1.587)	(0.380)	(1.361)	(1.102)	(1.134)	(0.226)
Asia-Africa Ethnicity	0.212	0.138	0.244	-0.104	0.184	0.145	0.209	-0.067
	(0.409)	(0.345)	(0.429)	(0.058)	(0.409)	(0.353)	(0.407)	(0.062)
Europe-America Ethnicity	0.213	0.264	0.192	0.072	0.192	0.229	0.166	0.062
	(0.410)	(0.441)	(0.394)	(0.030)	(0.410)	(0.420)	(0.372)	(0.030)
Other Ethnicity	0.005	0.004	0.008	-0.004	0.009	0.017	0.006	0.011
	(0.073)	(0.060)	(0.089)	(0.003)	(0.073)	(0.128)	(0.078)	(0.006)
Former Soviet Union Ethnicity	0.053	0.071	0.064	0.008	0.056	0.043	0.055	-0.010
	(0.224)	(0.257)	(0.245)	(0.050)	(0.224)	(0.203)	(0.228)	(0.019)
Ethiopia Ethnicity	0.007	0.000	0.015	-0.016	0.017	0.009	0.032	-0.023
	(0.083)	(0.000)	(0.122)	(0.008)	(0.083)	(0.093)	(0.177)	(0.016)
B. High School Outcomes								
High School Completion	0.955	0.949	0.960	-0.011	0.961	0.952	0.958	-0.007
	(0.207)	(0.220)	(0.196)	(0.015)	(0.207)	(0.215)	(0.200)	(0.013)
Mean Matriculation Score	70.89	72.71	70.42	2.229	74.01	74.26	72.08	2.187
	(21.60)	(20.94)	(21.42)	(2.319)	(21.60)	(20.68)	(19.97)	(1.692)
Matriculation Certification	0.616	0.662	0.614	0.047	0.687	0.699	0.630	0.069
	(0.487)	(0.474)	(0.487)	(0.059)	(0.487)	(0.459)	(0.483)	(0.045)
University Qualified Matriculation	0.575	0.633	0.570	0.062	0.632	0.659	0.568	0.093
	(0.494)	(0.483)	(0.495)	(0.063)	(0.494)	(0.474)	(0.496)	(0.060)
Observations	3177	550	1125		4529	1135	1150	

Note: Columns 1 and 5 present means and standard deviations (in parentheses) of background characteristics and outcomes of students before and after the early reforms. Columns 2, 3, 6 and 7 present the means and standard deviations for students in treatment and control grades for affected (1999-2000) and unaffected (1995-1996) cohorts of 10th graders. Columns 4 and 8 present the differences between treatment and control grades, controlling for cohort fixed effects. The treatment group is defined as being comprised by grades in which the number of students from early reformed kibbutzim is greater than zero and the number of students from late reformed kibbutzim is equal to zero.

Table A16: Short-Term Effects on High-School Outcomes (1 if Early Reformed $>$ 0 and Late Reformed $=$ 0)

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
i. Full Sample (N=3957)					
Simple diff-in-diff	0.010	0.582	0.031	0.044	0.058
	(0.012)	(2.081)	(0.033)	(0.032)	(0.067)
Controlled diff-in-diff	0.009	1.114	0.051	0.065	0.084
	(0.011)	(1.759)	(0.031)	(0.032)	(0.059)

Note: This table replicates the results in Table 1 using the alternative definition of treatment as described in the previous table.

Table A17: Short-Term Effects on High-School Outcomes, Schools with both Treatment and Control Grades

	High School Completion (1)	Mean Matriculation Score (2)	Matriculation Certification (3)	University Qualified Matriculation (4)	Summary Index (5)
i. Simple diff-in-diff (N=12340)					
Treated X After	0.016	3.011	0.083	0.100	0.152
	(0.015)	(2.052)	(0.042)	(0.039)	(0.079)
ii. Controlled diff-in-diff (N=12340)					
Treated X After	0.013	2.328	0.063	0.075	0.118
	(0.015)	(2.027)	(0.039)	(0.037)	(0.076)
iii. Cross-sectional regression					
Treatment-control diff., before (N=5915)	-0.022	-0.947	-0.005	-0.021	-0.054
	(0.013)	(2.179)	(0.060)	(0.063)	(0.102)
Treatment-control diff., after (N=6425)	-0.007	1.185	0.067	0.065	0.073
	(0.011)	(2.429)	(0.054)	(0.055)	(0.098)

Note: The dependent variable in column 1 is an indicator of whether the student completed high school; in column 2 it is her mean score in the matriculation exams; in column 3 it is an indicator of whether she received a matriculation certificate; in column 4 it is an indicator of whether she received a matriculation certificate that satisfies the requirements for university study; in column 5 is the summary index based on the outcomes in columns 1 to 4. The sample includes all the students (excluding kibbutzim members themselves) who attended schools with a positive number of either early or late/never reformed kibbutzim residents in both the before (1995/1996) and the after (1999/2000) periods. Sample is restricted to schools that have both treated and control grades throughout the sample period. The first two rows of Panel A presents the estimated coefficients of interest in difference-in-differences regressions, comparing students in treated and untreated grades who are treated (10th grade in 1999/2000) and untreated (10th grade in 1995/1996). A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim. The simple difference-in-differences regressions include only cohort dummies and school fixed effects. The second panel of the table shows the controlled difference-in-differences, which also includes the following students demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries). The third row of Panel A shows the estimated effects using only the before (1995/1996) cohorts and using only the after (1999/2000) cohorts.

Table A18: Short-Term Effects on High-School Outcomes, Controlling for Grade Size

	High School Completion	Mean Matriculation Score	Matriculation Certification	University Qualified Matriculation	Summary Index
	(1)	(2)	(3)	(4)	(5)
A. Short-Term Effects					
i. Simple diff-in-diff (N=7698)					
Treated X After	0.016	2.383	0.077	0.082	0.130
	(0.009)	(1.985)	(0.033)	(0.031)	(0.062)
ii. Controlled diff-in-diff (N=7698)					
Treated X After	0.021	2.628	0.082	0.088	0.145
	(0.009)	(2.202)	(0.036)	(0.037)	(0.068)

Note: The dependent variable in column 1 is an indicator of whether the student completed high school; in column 2 it is her mean score in the matriculation exams; in column 3 it is an indicator of whether she received a matriculation certificate; in column 4 it is an indicator of whether she received a matriculation certificate that satisfies the requirements for university study; in column 5 is the summary index based on the outcomes in columns 1 to 4. In Panel A, the sample includes all the students (excluding kibbutzim members themselves) who attended schools with a positive number of either early or late reformed kibbutzim residents in both the before (1995/1996) and the after (1999/2000) periods. The first two rows of Panel A presents the estimated coefficients of interest in difference-in-differences regressions, comparing students in treated and untreated grades who are treated (10th grade in 1999/2000) and untreated (10th grade in 1995/1996). A grade (school/year combination) is defined as treated if it includes students from early reformed kibbutzim. The simple difference-in-differences regressions include only cohort dummies and school fixed effects. The second panel of the table shows the controlled difference-in-differences, which also includes the following students demographic controls: gender, father's and mother's education, number of siblings, a set of ethnic dummies (origin from Africa/Asia, Europe/America, immigrants from FSU, Ethiopia and other countries). The third row of Panel A shows the estimated effects using only the before (1995/1996) cohorts and using only the after (1999/2000) cohorts. Panel B reports the results of a placebo experiment in which we assume the early reforms happened in 1996 instead of 1998. We then use data from 1994-1995 and 1996-1997 to compare treated to control grades, before (1994-1995) and after (1996-1997) the placebo reforms. Standard errors clustered at the school level and presented in parentheses.

Table A19: Long-Term Effects on Post-Secondary Schooling Outcomes, Controlling for Family Income

	All post secondary		University		College	
	Enroll-ment	Years od schooling	Enroll-ment	Years of schooling	Enroll-ment	Years of schooling
	(1)	(2)	(3)	(4)	(5)	(6)
i. Full sample (N=7178)						
Simple diff-in-diff	0.040	0.235	0.087	0.487	-0.043	-0.165
	(0.021)	(0.191)	(0.034)	(0.183)	(0.036)	(0.113)
Controlled diff-in-diff	0.046	0.263	0.093	0.494	-0.038	-0.149
	(0.021)	(0.206)	(0.040)	(0.208)	(0.037)	(0.122)

Note: This table replicates the results in Table 4 adding family income as an additional control variable.

Table A20: Long-Term effects on Labor Market Outcomes, Controlling for Family Income

	Labor market			Unemployment benefits		
	Employ-ment	Work-months	Earn-ings	Unemployed indicator	Total benefits	Number of months
	(1)	(2)	(3)	(4)	(5)	(6)
i. Full sample (N=7169)						
Simple diff-in-diff	0.022	0.277	7614.4	-0.014	-293.2	-0.089
	(0.016)	(0.209)	(3536.5)	(0.009)	(132.4)	(0.039)
Controlled diff-in-diff	0.028	0.331	6376.8	-0.018	-353.9	-0.096
	(0.017)	(0.226)	(3231)	(0.009)	(128.8)	(0.039)

Note: This table replicates the results in Table 5 adding family income as an additional control variable.

Table A21: Long-Term Effects on Summary Index

	Post-secondary and labor market outcomes	University and labor market outcomes
	(1)	(2)
i. Full sample (N=7555)		
Simple diff-in-diff	0.086	0.104
	(0.031)	(0.031)
Controlled diff-in-diff	0.091	0.108
	(0.033)	(0.032)
ii. Stratification by gender		
Male (N=3151)	0.112	0.130
	(0.026)	(0.031)
Female (N=3072)	0.065	0.096
	(0.071)	(0.066)
iii. Stratification by mother's education		
Low (N=3219)	0.030	0.068
	(0.040)	(0.038)
High (N=3004)	0.159	0.170
	(0.040)	(0.037)

Note: The full sample includes students that have at least 2 peers in a grade from reformed kibbutzim. Standard errors clustered at the school level and presented in parentheses.

Table A22: Effects on Percentile Ranking of Annual Earnings

	Percentile Ranking in National Income DIstribution
	(1)
i. Full sample (N=7524)	
Simple difference-in-differences	4.27
	(1.62)
Controlled difference-in-differences	4.08
	(1.75)

Note: In this table, we replace the income variable with the percentile ranking of an individual in the national income distribution.