Effects and Mechanisms of CEOs Quality in Public Education*

Short Title: The Effects of CEOs Quality in Public Education

Victor Lavy, Genia Rachkovski and Adi Boiko

Abstract

`CEOs' of public schools in many countries have authority and responsibilities that can greatly affect the quality of schooling. This paper estimates the impact of education CEOs on student outcomes in Israeli elementary schools. We estimate CEO quality in two ways- once using schools that don't switch CEOs and once using schools that do. We show that switches are exogenous and are not correlated with potential outcomes. CEOs' quality positively affects students' test scores and behavioural outcomes, with pronounced effects for disadvantaged schools. Potential mechanisms show that high-quality CEOs lead to improvements in school priorities, working procedures, and violence reduction.

Keywords: Value Added, CEO Quality, Student Outcomes, Student Behaviour, Test Scores, Public Schools, School Violence, School Principal, School Resources

JEL: **J24** Human Capital, **J53** Labor–Management Relations **J45** Public Sector Labor Markets **I20** Education and Research Institutions: General, **I21** Analysis of Education **I22** Education and Research Institutions: Government Policy **H75** State and Local Government: Health • Education • Welfare • Public Pensions

* Corresponding author: Victor Lavy, Economics, The Hebrew University of Jerusalem, Mt. Scopus, 9190501 Jerusalem, Israel & University of Warwick, United Kingdom. Email: msvictor@mail.huji.ac.il.

We benefited from comments and suggestions by Emma Duchini, Naomi Hausman, Rolant Rathelet, Moshe Shayo, and participants in seminars at the Hebrew University, University of Warwick, University College London, the Madrid Education Conference, CESifo Economics of Education Network 2017 Conference, and the 2017 NBER Summer Institute Conference in Personnel Economics. We thank the Israeli Ministry of Education for providing data for this project. We thank Elior Cohen for providing excellent research assistance during the early stages of this study. The first author acknowledges financial support for this project from the European Research Council through ERC Advance Grant 323439, CAGE at Warwick and the Falk Institute in Jerusalem.

1. Introduction

Recent studies examined the relationship between management quality and productivity in the public sector. Di Liberto (2015), Bloom et al. (2015), and Bryson et al. (2018) have extended the World Management Survey methodology to measure management practices in schools and present descriptive evidence about the quality of management by school principals and education outcomes in several countries. They show that better quality management is strongly associated with better educational outcomes, but they cannot establish causality in this relationship. Branch et al. (2012), Coelli and Green (2012), and Dhuey and Smith (2014) study school principals' leadership roles and measure the correlation with school outcomes. Bloom, Propper et al. (2015), McCormack et al. (2014), and Rasul and Rogger (2018) study management practices in public sector institutions and other government agencies. In this paper, we focus on measuring the effect of CEOs in public education on the education outcomes of students and schools and the channels of these effects.[1] School CEOs lead a cluster of schools within a district or a local school authority. As in many other countries, school CEOs in Israel have wide responsibilities that affect schooling quality and output. For example, they are the direct managers of school principals and draw their authority from the district or regional education director. Since this format of CEOs is widespread in many other countries, the evidence we present in this paper has relevant and important external validity.

The link between management quality and the effectiveness/productivity of schools has not been studied as carefully as, for example, the impact of teachers' value added (Rivkin et al., 2005; Rothstein, 2010; Jacob et al., 2010, 2016; Chetty et al., 2014a,b). This paper provides empirical evidence about the causal effect of management quality in public education on students' cognitive and non-cognitive schooling outcomes and explores relevant mechanisms. We measure CEOs' quality by a measure similar to teachers' value added based on students' test scores. In a second step, we estimate the effect on students' cognitive and behavioural outcomes. Our measures of CEOs quality are based on panel data of schools and not students, where each school is observed three times, and therefore it is

[1] Superintendent is the title used in the US and in some other countries. In the UK the title used is Chief Education Officer, in Canada it is Director of Education.

not as detailed and refined as measures of teachers' value added in recent studies. Yet our study makes significant progress in measuring management quality and its causal effect in education and, more generally, in other public sector activities.

The role and responsibilities of schools' CEOs vary across countries, but it is common for them to set the tone, chart the course of the district or sub-district, and work closely with the school board or district board of trustees. In many European countries, including Israel, school CEOs are responsible for hiring and supervising school principals and teachers, including deciding to grant tenure to novice teachers. It is the CEO's job to evaluate school principals' performances and ensure that they are effective leaders, working with teachers to serve the needs of students and meet the district's goals. The CEO must respond to demands from all other constituencies in the district or sub-district: the teachers, students, parents, and the community. In some countries, CEOs can affect the district's allocation of financial and human resources. Therefore, a high-quality CEO can have an important effect on the quality of a school and student's academic achievements.

Since it is plausible that struggling schools/districts are more likely to be allocated to a good CEO, non-random matching of schools and CEOs can lead to biased measures of CEOs' quality. We provide natural-experimental evidence showing that this concern is irrelevant in the Israeli context and in the sample and period we examine. We exploit a quasi-random pairing of CEOs and schools that results from a Ministry of Education rule, under which the CEOs 'supervision area' must change every 3-5 years, leading to schools being reshuffled across CEOs' clusters. Implementing this regulation is closely linked to CEOs' retirement and sometimes, though not very often, to the promotion of CEOs within the Ministry of Education. Such exits lead to new CEOs' entry and some reshuffling of schools across existing CEOs. The entire cluster of schools of a retiring CEO is often transferred as a group to a different, veteran or novice CEO. Clearly, there should be no concern for sorting in such cases, and we show evidence that supports our claim of a random matching of schools and CEOs. We show that the probability of CEO turnover is unrelated to school characteristics and outcomes. We show that the

quality of a new CEO is unrelated to the characteristics and mean outcomes of the school. We also show that the quality of the outgoing and the incoming CEO at a given school are not correlated.[2]

We estimate the effects of CEOs' quality based on two different methods and using an innovative split-sample approach. In the first method, we derive estimates of CEO quality based on the CEO fixed effect in a school production equation. These fixed effects are estimated based on the sample of schools that retain the same CEO throughout the period. Separate estimates are produced for each sub-period of the sample with no structure imposed on the data, even though the influence of a CEO may increase over time as the effects of prior CEOs wane. However, since we observe only two years of a given CEO regime, we prefer to generate two independent quality measures and assess their correlation. Then, these CEO quality estimates are used as explanatory variables in achievement regressions for the sample of schools that switched CEOs to confirm they are meaningful.

In the second method, we exploit CEOs' turnover and measure their quality based on schools that switched to a new CEO. We regress within school test score change between two periods on the changes in characteristics and a CEO fixed effect. This CEO quality estimate is then used as a treatment in outcomes' regressions using the sample of schools that did not switch CEO.

It is important to note that in both methods, the schools used in the first step (where we measure CEO quality) are not used in the second (where we measure the effect of CEO quality) and vice versa. This guarantees that we use an out-of-sample measure of CEO quality. However, these methods have some limitations discussed in section 5 of the paper. Yet, the fact that the ranking of quality of CEOs obtained from these two methods is highly positively correlated and the fact that the estimated effect on students' outcomes based on these two alternative measures of CEO quality is very similar is somewhat reassuring that these concerns are not serious limitations in our context.

[2] Interrelationships between schools clustered in a district may overlap with the 'patch' that the CEO is responsible for, maybe a concern. For example, such interrelation may result from common catchment areas, leading to the movement of pupils across schools within the set for which a CEO is responsible. However, catchment areas are specific to each school and do not overlap across schools or districts. But any such related examples are dealt with in our model by the fact that schools are moving across CEOs and that these movements are not correlated with school's characteristics or outcomes or with other reasons such as transport patterns changing or changes in enrolment.

As learning outcomes, we use test scores in national exams in primary and middle schools in Israel in math, Hebrew, and English. We also examine effects on students' behavioural outcomes, particularly bullying and violence in school, and students' social skills and behaviour. Finally, as potential mechanisms for the effect of CEO quality we focus on several management practices and working procedures at their schools.

The results show that the quality of CEOs has positive and significant effects on students' academic achievements. As measured in the first method, a one standard deviation improvement in management-supervision quality increases students' test scores in math, English, and Hebrew by 0.04 standard deviation. The respective estimated effect size based on the second method of measuring CEO quality is 0.05. The similar effect size of the two measures of quality that we use, even though they are based on different methods and implemented in non-overlapping samples of schools, enhances the validity of the methods we use to measure CEO quality and the causal interpretation of the findings.

We also estimated the effect of the CEO's quality using two alternative estimation strategies. In the first strategy, we implement a Bayes shrinkage estimation strategy and construct an unbiased measure of CEO quality that accounts for noise in the measurement. In the second strategy we use a two-step bootstrapping algorithm to account for estimating CEOs' quality in the first step, and adjust their estimated standard errors. The results are relatively unchanged when using these two alternative strategies.

We also find that the effect is non-linear, being higher for the highest quality CEOs and that the effect is stronger for the weaker schools of the cluster for each CEO. When the analysis is done separately for each subject, the treatment effect is on the same scale. Interestingly, the estimated effect of CEO quality on student outcomes seems to be stronger for female CEOs compared to their male counterparts, but this difference is not significant.

In this study, we also explore additional school-level outcomes that can be viewed as channels or mechanisms of the effects of CEOs. We find that schools with CEOs of higher quality have more focused priorities, more clearly defined working procedures, and a better climate and environment with lower levels of violence and bullying among students. We find no discernible effect on school resources, which can be viewed as a placebo effect because funding is determined centrally at the Ministry of

Education based on clearly defined criteria. Finally, we also find that higher-quality CEOs are more likely to replace an existing school principal.

This paper also contributes to several other literatures. First, and more generally, we contribute to the emerging literature investigating management practices in public sector institutions (For example, see McCormack et al. (2014), Rasul and Rogger (2018), and Bloom, Propper et al. (2015)). Secondly, this paper is related to recent research on the effect of school principals and leadership (see Branch et al. (2012), Clark et al. (2009), Béteille et al. (2012), and Horng et al. (2010)). Some of the duties of CEOs in Europe and Israel are similar to those of a principal or "school leader" in many places in the US, which suggests that we can think of our results as informing the literature on principal value-added (e.g. Grissom et al. (2015)). Finally, we link to work on teachers' value-added and the effect on students' academic achievement (see Rockoff (2004), Rivkin et al. (2005), Jacob (2010), Rothstein (2010), Bacher-Hicks et al. (2014), and Chetty et al. (2014b)).

The rest of the paper is organized as follows. Section 2 presents the institutional background of the CEOs (superintendents) system in Israel and elsewhere. Section 3 describes the data and provides descriptive statistics. Section 4 discusses the quasi-random pairing of schools and CEOs. Section 5 presents the empirical framework and identification strategy, and section 6 the results regarding the effect on students' outcomes. Section 7 presents evidence on the mechanisms of the effect of CEO quality. Section 8 offers a summary and some conclusions.

2. Context and Background

There are management and supervision services over schools in nearly all countries.[3] Their key role is to monitor the quality of education, i.e., schools and teachers, and support their improvement, creating two distinct but complementary tasks: on the one hand, to control and evaluate and, on the other hand, to advise and support teachers and head-teachers. CEOs (in some countries, they are called supervisors) are based outside of the school at a local or regional government body overseeing public schools. Each

[3] Much of the material in this section is based on "Reforming School Supervision for Quality Improvement", United Nations, International Institute for Educational Planning, UNESCO 2007.

CEO is assigned several schools, which they visit yearly. The role and powers of the CEO vary between countries; however, they include a common core that includes: supervising the implementation of government education policy, regulations, and national curriculum, advising on teaching methods, supervising, guiding and assessing teachers on probation, liaising with and advising headmasters, guiding schools in responding to patterns of problems and needs, and reporting on periodic inspections. For example, some countries, like Spain, separate the administrative from the pedagogic tasks of supervisors. Other countries tend to separate control and support roles. This has been the case mainly in countries with strong school-based management practices, like New Zealand and the UK. Some countries adopted a management-supervision approach directed towards the school and less towards individual teachers.

A noticeable example is the OSTED school audit system in the UK. Similarly, in Germany, inspection visits provide evidence (such as school quality, problems, and possible school development) for accountability and discussions with teachers and head teachers to facilitate school improvement. They report the results to the respective school supervisory authority responsible for advising the school on its action plan regarding improvement measures and assessing the results. In France, CEOs provide regular reports relating to issues based on particular fieldwork and unique school visits. In a few countries, schools use self-assessment instead of external supervision. For example, Finland follows this model, and teachers are entrusted with quality control. Other countries use a model that combines internal school evaluation and assessment with external supervision (New Zealand, Wales, Australia, and Chile).

School CEOs in Israel share similar responsibilities to their counterparts in many countries, especially in Europe. While a school principal is responsible for the school's day-to-day administration, the school CEO has various activities and responsibilities similar to what is observed in many other countries. The CEO oversees the implementation of the school action plan, which sets the educational objectives for the school under the national curriculum framework. She is responsible for the hiring, placement, and transferring of teachers, assessing teachers' performance, providing educational guidance and know-how to school principals and teachers, and deciding about the tenure and dismissal of teachers. To carry out these tasks, the CEO visits every school a few times a year, meets with the

school principal and members of the senior leadership team, attends classes (mainly for novice teachers), and holds confidential meetings with teachers and staff. The CEO's responsibility is to encourage teachers and school principals to improve their classroom instruction and ensure that they work within the norm, policies, and codes of practice of the Ministry of Education and the law. The CEO should monitor training and teaching progress and ensure that novice teachers receive appropriate career guidance. The school staff views the CEO as an educational pedagogical authority. The CEO submits a report to the school district director following each school visit. Supervisees receive relevant parts of the report and can comment on their assessment and on the support provided and discuss any identified problems.

CEOs are part of the staff of regional directors. Each is in charge of several schools, and every 3 to 5 years, they rotate to a different "supervision zone", or, if a CEO leaves the position, there is a general reallocation of the schools they supervised. We exploit in this paper this rotation of CEOs to identify their causal effect on children's learning and behaviour. Therefore, a more detailed description of CEO rotation appears in a later section.

3. Data

The data used in this study is based on the Growth and Effectiveness Measures for Schools (GEMS - Meizav in Hebrew) datasets for 2002-2005. The GEMS was administered for the first time in 2002, including a series of tests and questionnaires administered by the Division of Evaluation and Measurement of the Ministry of Education.[4] The GEMS is administered towards the end (from mid-May to mid-June) of each school year to a representative 1-in-2 sample of all elementary and middle schools in Israel so that each school participates in GEMS once every two years. The GEMS data includes test scores for fifth-grade (primary school) and eighth-grade (middle school) students in math, science, Hebrew, and English. In principle, all students except those in special education classes are

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

tested, and the proportion of students tested is above 90%. The raw test scores are graded on a 1-to-100 scale which we transform into z-scores to facilitate the interpretation of the results.

The test scores for 2002-2005 are linked to student administrative records collected by the Ministry of Education. The administrative records include student demographics, which we use to construct all students' background characteristics. Using the linked datasets, we create a panel of elementary schools with test scores for 2002-2005. The sample is restricted to Jewish public schools that follow the same national curriculum and participate in the GEMS national testing. We exclude the religious Orthodox Jewish and Arab schools for these reasons. There are 939 elementary schools with test score data. Since every school is sampled once every two years, we have two observations of the same school for more than 90% of the schools.

The GEMS also includes interviews with all teachers and the school principal. The questionnaire for 'home class' teachers of all classes included questions about instruction time in each subject and the total instruction time per week.[5] Lavy (2020) finds a positive impact of these measures on student academic achievements. We use teachers' responses to these items to compute the school average for fifth-grade instruction time in each subject.[6]

The school's principal questionnaire includes questions on pedagogical and management practices in school. However, as a public sector and as an educational system, it is clear that these questions do not relate too closely to management practices used in the broader economics management literature. They are not measures of incentives, monitoring, or organization of production. They serve more as outcome measures of management practices rather than details of the practices, yet they do capture channels through which CEOs can impact schooling quality. We attempted to choose the

[5] A 'home class' teacher in primary school in Israel teaches most weekly sessions of his class, and has additional duties such as taking attendance registers, acting as an intermediary in cases of conflict, collating other teachers' impressions of the class and of individual students in preparation for the quarterly report, liaising with parents and various other administrative tasks.

[6] Though there was very little difference between or among fifth-grade classes in a school in these time inputs, we still prefer to use the school-level mean per grade to avoid any biases that might be caused by sorting of students into certain classrooms and setting time allocations for given academic subjects according to those students' particular strengths and weaknesses. In any case, the grade- and class-level measures of these time inputs are very highly correlated.

questions that relate to management practices and are similar to those in Bloom and Van Reenen (2007). We use the following questions from the principal questionnaire, which are: (1) "Did the school evaluate its environment and discipline?" and (2) "Does the school have interventions related to values, norms and discipline?". We also use items from the GEMS student questionnaire that address various aspects of the school and its learning environment. We concentrate on the section of the questionnaire that provides information on student satisfaction in school and on the violent behaviour of other students, both of which, previous research indicates, can impact students' academic performance (Lavy and Schlosser, 2011). In this section, students are asked to rate the extent to which they agree with a series of statements on a six-point scale ranging from "strongly disagree" to "strongly agree". These items include: (1) "There are many fights among students in my classroom"; (2) "Sometimes I'm scared to go to school because there are violent students"; (3) "This year I was involved in many fights"; (4) "When I have a problem at school there is always someone I can turn to (from the teaching staff)"; (5) "I am satisfied in school"; and (6) "I feel well-adjusted socially in my class". We transformed students' responses to these items into standardized z-scores.[7]

Descriptive statistics about our data are presented in Table 1. The sample in 2002-2005 includes about 800 schools, all in the Jewish non-religious primary school system, and about 50 plus CEOs, 70-80% of which are female. The number of schools per CEO is about 16, and it is marginally higher for female CEOs. The table presents summary statistics on the entry and exit of CEOs.

4. The Quasi-Random Pairing of Schools and CEOs

Israel's public education system has seven geographical regions, each headed by a regional director who reports directly to the general director of the Ministry of Education. All schools within each of the seven regions are under the directorate of the regional director. Each region has several CEOs, and each is in charge of several schools. Except for the largest five cities in the country, all other

[7] We experimented with binary versions of these variables denoting above median answers and also using them linearly and the results were not different, confirming that the evidence regarding these variables is not a feature of this particular transformation of the data.

localities have only one CEO who supervises all schools in the locality.[8] Each such CEO manages a few geographically close localities. This is evident from the information presented in Maps 1-6 in the online appendix. Each map presents the geographical distribution of the schools by CEO in each of the seven regions in 2005. The spatial clustering of the schools of each CEO reveals that the matching of schools and CEOs is based on geographic considerations. About 77% of schools in our sample are in municipalities with a single CEO. About 82% of these schools have a CEO who oversees schools in more than one municipality. Every CEO change in these municipalities involves all schools, minimizing the scope for selective matching between CEOs and schools in this context. This assignment mechanism of CEOs to schools rules out the possibility that the pairing of schools and CEOs is based on endogenous matching.

Another structural feature of assigning CEOs to schools in Israel that helps rule out the potential for endogenous sorting of CEOs across schools is the policy to reassign CEOs every 3-5 years.[9] This is often implemented in conjunction with the departures of CEOs due to retirement – rarely due to promotion. Our panel data for 2002-2005 includes 771 schools that appear in the sample each of the four years. There are three potential transition points, one for every two adjacent years, and 402 schools (52%) switched CEO at least once within this period. This transition rate implies that schools receive a new CEO every 5-6 years. The entire cluster of schools of a retiring CEO is often transferred as a group to a novice or veteran CEO. 15% of the changes in the pairing of schools and CEOs result from such reassignment, and an additional 61% of the transitions result from a transfer of at least 5 schools from a retired or promoted CEO to a different CEO. Clearly, when the reassignment involves such a large number of schools as in these cases, we can confidently rule out any endogenous sorting that might lead to selection bias in the measure of CEO quality and the treatment effects estimates. The reassignment of one or two schools to a new or veteran CEO accounts for only 9% of all reassigned schools (40 out

[8] The five largest cities are Jerusalem, Tel Aviv, Haifa, Beer-Sheba, and Rishon-Lezion. The first three have 3 CEOs and the other two cities have 2. The distribution of schools of each CEO in these large cities is also clustered geographically, shown in Maps 1-6 in the online appendix. The schools that are included in each CEO cluster are mostly the same over time, particularly when there is a change of CEO.

[9] See this source for details: https://he.wikipedia.org/wiki/%D7%9E%D7%A4%D7%A7%D7%97_(%D7%97%D7%99%D7%A0%D7%95%D7%9A). The Wikipedia text is in Hebrew.

of 459). The reassignment of three or more schools to a new or veteran CEO accounts for 91% of all reassigned schools. In fact, out of the 40 schools reassigned as a single or a couple of schools in the reassigned cluster, the reassignments of 25 schools seem to be driven by spatial reasons, as the reassigned schools are geographically closer to the receiving CEO's cluster of schools. These patterns provide strong supportive evidence against endogenous sorting in these transitions

Evidence for the lack of endogenous sorting of CEOs can also be seen in Figures 1-3, which are representative examples of turnover patterns of CEOs following the retirement of one or more CEOs. Figure 1 presents the changes in the Central educational district following the retirement of two CEOs. All of Ruth's (this and other names in the paper are fictitious) schools were transferred to Rebecca (a new CEO), and all of Amalia's schools were transferred to Hannah (a veteran CEO). Rebecca also received ten schools from Johanna. The transfer as a bloc of such a large number of schools between pairs of CEOs suggests no sorting involved. However, Figure 1 also shows transfers of smaller numbers of schools between CEOs, in most cases 1 or 2 schools, which may involve sorting. In Figure 2, Judith, a CEO in the South district, retired, and her 13 schools were distributed between 4 other CEOs. Figure 3 shows transitions in the Haifa district from 2003-2004. There are 8 pairwise transfers, five of them involving a move of 5 or 6 schools. In most cases, the number of outgoing and incoming schools of a CEO is equal or almost equal.

5. CEO Quality in the Education Production Function

We specify a general education production function on which we base our methods of measuring CEO quality. The production function includes an input that captures CEO quality in addition to the standard school and student-level inputs:

$$Y_{js} = \alpha_j + \gamma O_s + \beta X_j + \Lambda LA_j + u_{js} \quad (1)$$

where Y_{js} is the average learning outcome of students in school j assigned to CEO s. We use test scores in national exams in primary (5th grade) and middle school (8th grade) in Israel in math, Hebrew, and English as learning outcomes. We also examine the effects on students' behavioural outcomes, particularly bullying and violence in school and students' social skills and behaviour. As potential

mechanisms for the effect of CEOs' quality, we focus on several management practices and working procedures at their schools. α_j is a school fixed effect, O_s captures the quality of CEO s, X_j is a set of potentially time-varying covariates of school j (including students' characteristics, school enrolment, number of classes), LA_j is a vector of lagged test scores in Hebrew and math, and u_{js} is a mean zero residual.

The parameter of interest is γ, but O_s is unobserved and has to be estimated. Since CEO value-added might be measured with error, it can be correlated with u_{js} and, therefore, with school specific outcome shocks which are part of the error term in equation (1). To avoid this problem, we use the same approach as in recent studies of teachers' value-added. For example, Jacob et al. (2010), where the estimation of the value-added of a student's teacher does not incorporate information from that student's cohort. In our case, this approach implies that the estimation of the value-added of a school's CEO does not incorporate information from that school. Stated differently, the quality of a school's CEO is estimated out of sample for every school for every period. Explicitly, in the first method we use to estimate CEO quality, we measure quality based on 'non-switchers' (schools that **did not** change their CEO during the study period). In the second step, we estimate the effect of quality on the outcomes using the sample of 'switchers' (schools that **did c**hange their CEO during the study period). In the second method that we use to estimate CEO quality, we measure quality based on 'switchers'. In the second step, we estimate its effect using the 'non-switchers' sample.

Both methods I and II rely on the assumption that schools that are part of a CEO's cluster of schools do not share any other common factors that can correlate with their residual test scores after controlling for school covariates (method I) or their residual growth in test scores (method II). Both methods are similar in spirit to the models used to estimate individual managers' effect on corporate behaviour and performance (Bertrand and Schoar, 2003).

We show below that CEOs' quality derived based on methods I and II is highly correlated. Furthermore, since method I yields two measures of quality for each CEO (one for each year), we examine the correlation between the two and find that they are highly correlated, a sign of the consistency of the measure we estimate. Finally, as we will show below, the measures of CEO quality obtained from method I and method II lead in the second step of each method to very similar estimated

effects on students' outcomes. In the following two subsections, we explain and demonstrate in more detail both methods.

5.1 Method I: Measuring CEOs' Quality Based on 'Non-Switchers.'

The first method (I) we use to measure CEOs' quality is based on the non-switchers' sample (schools that stayed with the same CEO during the two years surveyed). Next, we estimate an education production function similar to equation (1) above while adding a CEO fixed effect instead of O_s (i.e. including a dummy variable for each CEO) and using these estimated effects as measures of CEOs' quality.

The practical representation of this specification is that a vector of CEOs' fixed effects is included in equation (1) and the estimates of these fixed effects are the measure of CEOs' quality[10],

$$Y_{jst} = O_{si} + \beta X_{jt} + \Lambda LA_j + \delta S_{jt} + u_{jst} \quad (2)$$

where Y_{jst} is the average learning outcome of students in school j that is assigned to CEO s in period t. We estimate this regression separately for each sub-period: period one includes 2002 and 2003, and period two includes 2004 and 2005. Each school is included once in each of these regressions because the test scores data is available for each school in one of the two years, in each of these two two-year periods. We ensure that the CEOs' fixed effects pick up only the CEO's quality by including many students' and school-level controls, particularly lagged school-level mean test scores, as done in the above-referenced studies on teachers' value-added. The student controls (the X_{jt} vector) include an indicator for male students, the number of siblings, an immigrant indicator, the father's and mother's years of schooling, and six indicators of the student's ethnicity. The school-level controls include the 1991 school standardized test scores of 4th and 5th grade in Hebrew and math (LA_j), school enrolment, and the number of classes (S_{jt}). We also add year and district-fixed effects. Standard errors are clustered at the CEO period level. Including the CEO fixed effect is feasible because more than several schools share the same CEO.

[10] This approach is identical to the method used in Rivkin et al. (2005) and Jacob et al. (2010) who measured teachers' value added as teachers fixed effects and also very similar to the method used in Rothstein (2010) and Chetty et al. (2014b) in a teachers' value added context as well.

The means and standard deviations for these variables are presented in Table 2's columns 1 and 3 for each of the two periods. Period 0 is 2002-2003. Period 1 is 2004-2005. Table 2, columns 2 and 4 present the regression estimates of equation (2), once based on the first period (2002 and 2003) data and the second time based on the second period (2004 and 2005) data. The point estimates of the control variables have the expected signs: boys have 0.10-0.13 standard deviation lower test scores than girls, the number of siblings is negatively correlated with test scores, immigrants and students from Asia-Africa and Ethiopian ethnic backgrounds have a negative correlation with test scores, and parental schooling levels and ethnicity from Europe, America, Israel, and the former Soviet Union have a positive correlation with test scores. The estimates of the school average lagged test scores are much weaker because the school test score means are highly correlated with the school means of students' characteristics that are also included in these regressions. From each regression, we recover the CEOs' estimated fixed effects. Therefore, this step yields two estimated quality measures for each CEO, corresponding to the two sub-periods. We standardized the CEOs' fixed effects as a z-scores distribution with a mean zero and a unit standard deviation. Figure 4 displays the distributions of these two standardized quality measures. The sample includes 47 CEOs that are in charge of 497 schools. The two distributions look similar, and their equality is not rejected in a two-sample KS test, though we note that it is not easy to detect small differences in a small sample. The correlation coefficient between the pairwise value-added measures of each CEO is 0.62. Note that there is a large variation in CEO's quality, which allows for estimating its impact on schools' academic and non-academic outcomes.

Next, we assess how sensitive these two quality estimates (based on method I) are to the fact we use as controls lagged school mean test scores from testing a decade ago. We reiterate that these controls are intended to ensure that the CEO fixed effect does not capture some unobserved school-specific characteristics. Although short lags would have been better for this purpose, the 1991 scores that we use do just as well because they are strong and precise predictors of the 2002-2005 test scores. To substantiate this explanation with evidence, we compare the 1991 test scores' predictive power to the more recent test scores from 2002 to 2003. The R^2 of a regression of the 2002-2003 test score in Hebrew and math on the 2004-05 test scores is 0.42. The R^2 when the 1991 test scores replace the 2002-2003 scores is 0.41.

Furthermore, when the 1991 test scores are included jointly with the 2002-2003 test scores in a regression on the 2004-05 test scores, both lagged test scores have positive and similar coefficients. In a similar validation exercise, we test whether adding the controls for the 2002-2003 exam scores (using them as lagged scores) would have significantly added to the prediction power of our results. If it won't, that would indicate that our 1991 lagged scores are strong and precise predictors of more recent lagged scores. Following the procedure of estimating the bias using observable characteristics performed in Chetty et al. (2014a), we find the predicted value generated from a regression of test scores on the 2002-2003 lagged test scores after residualizing all variables with respect to the baseline control vector. Indeed, we find that the coefficient is less than 0.01. The magnitude of forecast bias due to selection on 2002-2003 lagged scores is very small, indicating that a large fraction of the variation in test scores that project onto the more recent lagged scores is captured by the 1991 lagged test scores and the other controls, signalling the predictive strength of our lagged 1991 scores as the more recent 2002-2003 lagged scores don't provide much prediction beyond them.[11]

We further demonstrate in Figure 5 the similarity in the predictive power of the 1991 test scores and the more recent 2002-2003 test scores. The figure displays the distributions of estimated standardized quality of CEOs based on 2004-2005, once when using 1991 lagged scores and once while using 2002-2003 lagged scores as controls. The two distributions in Figure 5 look identical, and a two-sample KS test cannot reject their equality. This supports the validity of our quality estimates as capturing CEO's quality and not something else about schools or students.

In the second step of method I, we estimate the effect of these CEOs' quality on students' outcomes,

$$Y_{jst} = \gamma \hat{O}_{js1\text{-}t} + \alpha_j + \beta X_{jt} + \Lambda LA_j + \delta S_{jt} + u_{jst} \quad (3).$$

The sample in this second step estimation includes schools that **did** change their CEO ('switchers') and in which their old and new CEOs are of the CEOs for which we estimated their quality ('fixed effect') in the first step. α_j is a school fixed effect, $\hat{O}_{js1\text{-}t}$ is the standard normal of the CEO quality measure, estimated in equation (2), using the 'Non-Switchers' sample in the opposite time period. Using a sample

[11] We don't present these results in the paper and they are available from the authors upon request.

of schools in the first step and a completely different sample in the second step (**no overlap between the two samples**) guarantees that the CEO fixed effect estimation does not reflect specific characteristics of the schools included in the sample in the second step. This 'out of sample' procedure of estimating the CEO fixed effect is crucial even though we estimate the role of CEOs of education districts in a framework where we can control for observable and unobservable differences across schools. We construct a CEO-school-matched panel data set for this analysis, where we track CEOs across different schools over time. After controlling for schools' time-varying characteristics, this allows us to estimate how much of the unexplained variation in schools' average outcomes can be attributed to CEO quality (fixed effect).

Remaining concerns. The methodology and results regarding this first method of estimating CEO quality are subject to some concerns, and we discuss them here. First, even though we have shown that the lagged test scores from a decade earlier as a primary control are highly predictive of current achievement, they still might fail to capture very recent school influences that could introduce bias. Second, the possibility of correlated shocks remains a threat even though we find that CEO clusters within a school district are not distinguishable in terms of observables. An example of a cluster-specific unobserved shock is a new entry of a school to an existing cluster managed by the CEO. Both the CEO switch and the addition to the cluster are shocks that may well affect both incumbent and entering schools as the CEO must reallocate time and attention. In poor-performing clusters, the CEO may struggle to work with the new school, while in well-performing clusters with stable schools, the CEO may devote much more effort to the new entrant. One way we deal with this issue is by using the quality measure estimated using the 'Non-Switchers' sample in the opposite period to estimate the effect of CEO quality on outcomes for the 'Switchers' sample (We discuss this in more detail later on in the paper), this limits the potential for correlated shocks. Combining this choice with the finding that the value-added measures are not affected much by adding the period 0 score controls and that we include time-varying school and student characteristics controls, such as parental education, alleviates some of the concerns related to potentially correlated shocks.

Furthermore, the fact that the number of schools that each CEO has to supervise is different might make it difficult to compare their effects. To deal with this concern, we show in online appendix

Table A1 column 1 that the results are virtually unchanged if we remove the top and bottom deciles of the CEOs with the largest and smallest number of schools. In column 2 of Table A1, we show that there doesn't appear to be a heterogeneous effect of the main outcome by the CEO number of schools (or when using a split sample of above and below the median number of schools per CEO). Another potential concern is attributing effects to the CEO in a context where the CEO plays an important role in choosing school principals. For example, school principals in the cluster might influence the performance of cluster entrants, and a previous CEO may have selected those school principals. However, even if this is a likely scenario, it can be viewed as another channel of the effect of a CEO: selecting school principals that might lead to spill-over effects of other school principals in the CEO's cluster. We will show below that a new CEO leads to a higher turnover of school principals.

5.2 Method II: Measuring CEOs' Quality Based on Switchers

The second method (II) reverses the role of 'switchers' and 'non-switchers': In the first step, we exploit CEOs turnover, using schools that were assigned a new CEO ('switchers'), and measure quality as the CEO 'fixed effect' in a regression of the change in these schools' mean learning outcomes as a dependent variable. We specify a school-level regression of the change in the average performance of students in a school:

$$Y_{j1} - Y_{j0} = O_{js1} + \beta\,(X_{j1} - X_{j0}) + \delta\,(S_{j1} - S_{j0}) + \Lambda\, LA_j + u_j \qquad (4)$$

where the dependent variable is the change in the average learning outcome of students between period 0 and period 1. O_{js1} is the CEO fixed effect of the new CEO s in period 1 of school j. X_{j0} and S_{j0} are the students and school's mean characteristics in the first period, and X_{j1} and S_{j1} are the respective means in the second period. $(X_{j1} - X_{j0})$ and $(S_{j1} - S_{j0})$ are the changes in students' average and the school's characteristics. LA_j is the lagged test scores in math and Hebrew. As for lagged test scores, we use the same 1991 data described and discussed in the previous subsection. To estimate CEOs' quality based on the 'Switchers' sample, we estimate the vector O_{js1}, which is the mean residual change in test scores (quality) of the schools that are part of the cluster of CEO s in period 1.

We combine all observed data throughout the study period, 2002-2005, to estimate CEO's quality using method II. Table 3 presents the estimates of equation (4). The X_j vector includes an indicator for male students, the number of siblings, an immigrant indicator, the father's and mother's years of schooling, six student's ethnicity indicators, school enrolment, and the number of classes. The means and standard deviations for these variables are presented in column 1, and the estimates appear in column 2. Most of the estimated coefficients of the right-hand side variables are not different from zero except for the two parental education variables. This pattern is most likely a result of the stability of the characteristics of a school within this short period. From this regression we recover the CEOs' fixed effects. These are then standardized as a z-scores distribution with a mean zero and standard deviation of one. Figure 6 displays the distribution of the standardized quality measure of 44 CEOs that are in charge of 226 schools.[12] It is clear that there is less variation in this measure than in the first measure displayed in Figure 4 but, as will be shown, it is large enough to permit estimation of its effect on students' outcomes and receives estimates that are very similar to those obtained using the first quality measure.

We note two implicit assumptions in the model of estimating CEO's quality using method II. First, the fifth-grade students tested in the later years were third-grade students under the previous CEO. We assume that the change in 5th-grade scores provides a good measure of CEO quality, even though learning is likely a cumulative process. Secondly, even though it might matter whether the tested year is the first or second of the CEO's tenure in the school, we do not distinguish between these two possibilities. We think this is a reasonable assumption as CEO's tenure in a given school is about five years and therefore, any systematic decline or improvement in achievement during the first two years is likely to be similar.

Finally, we compare the similarity in the CEO quality measures based on method I and method II. The correlation coefficient between the two measures, based on the intersection of the two samples

[12] We also estimated the quality regression where student-individual data replaces school means in outcome test scores and student characteristics. The distribution of the CEOs' quality values looks identical to that displayed in Figure 6.

(all CEOs for which we computed quality measures I and II), is 0.61. It is very reassuring that these two alternative measures, based on very different methods and samples, are highly correlated.

In the second step of method II, we estimate the effect of the quality of these CEOs, estimated in the first step using the `switchers' sample, on mean student outcomes of the sample of the `non-switchers', the schools that were not assigned a new CEO,

$$Y_{jst} = \gamma \hat{O}_{jst} + \beta X_{jt} + \Lambda LA_j + \delta S_{jt} + u_{jst} \quad (5).$$

$\hat{O}_{jst}$ is the standard normal of the CEO quality measure, estimated in equation (4). All other variables definitions and out-of-sample logic are similar to the previous subsection.

Are Switching Schools and CEOs Quasi-Randomly Matched? Are 'switchers' and 'non-switchers' observationally equivalent?

In this sub-section, we show that the likelihood that a school experienced a match with a new CEO is unrelated to the CEO's quality or the school's characteristics. This is an expected result, given the details in section 2 that clarified that changes in CEO-school paring are related to spatial-geographical considerations or the CEO's personal circumstances (retirement), not to the relationship between the school (characteristics, performance) and the quality of its CEO.

In Table 4, we present summary statistics for the variables used in the analysis. Column 1 presents the means for the sample of all schools that **did not** change CEO during the period of the study, and column 2 presents the respective means for the schools that **did** change CEO during this period. Panel A includes 13 variables of school means for the characteristics of the school and its students. Panel B includes schools' mean of its students' test scores in the 2002-2003 (period 0) national standardized exams in math, Hebrew, and English. The 497 schools that stayed with the same CEO had 63,979 students and 47 CEOs during the study period. The 222 schools that switched CEO between 2002 and 2004 or 2003 and 2005 had 55 CEOs. Finally, in column 3, we present the balancing tests between columns 1 and 2, which test a correlation between the probability of a school changing CEO and its observable characteristics. There are 16 parameter estimates; only one is statistically different from zero. Concerning the proportion of students of Ethiopian ethnicity, the means are very small, 1.8%

for the non-switchers and 1% for the switchers, but the difference is statistically significant. However, it should be noted that the two groups are statistically indistinguishable in terms of their parental years of schooling and in terms of the average test scores in each of the three subjects, both in terms of the absolute differences and their statistical significance. In appendix Table A2 we also present additional balancing tests relating to test scores in the 1991 national standardized exams in math and Hebrew, as well as a rich set of school practices and outcomes based on responses to questionnaires of primary school principals, teachers, and students. In total, only 5 of these 51 measures are significantly different between the two groups at the 10% level, and all of them are very small in magnitude compared to the means, likely a random event. We view this evidence as suggestive of no particular selection pattern in the probability of changing CEO and definitely no correlation with students' socio-economic background and test scores.

Next we estimate whether the quality of the incoming CEO is correlated with observable school characteristics and outcomes. In column 4, we present estimates of these balancing tests based on quality derived from method II. The sample includes 301 schools that did not change CEO, and for which CEO quality could be estimated in the first stage using the sample of schools that did change CEO. Only one of the 16 estimates is statistically different from zero, ethnicity Ethiopian. In column 5, we present the estimates based on quality derived from method I. The sample includes only the 82 schools (10,249 students and 29 CEOs) that switched CEO between 2002 and 2004 or between 2003 and 2005, and for which CEO quality could be estimated in the first stage using the sample of schools that did not change CEO. Four of the 16 estimates in column 5 are statistically different from zero: ethnicity Israel and ethnicity Europe-America are positively and negatively correlated, respectively, with CEO quality. These two ethnic groups have higher socio-economic backgrounds than the other ethnic groups, so we expect they will have the same direction of selection. But we see that these two ethnic indicators have the opposite sign of their correlation with CEO quality, one positive and the other negative, suggesting that their significant imbalance does not reflect a systematic selection pattern in the assignment of schools to CEOs. The evidence clearly indicates that this quality measure is not systematically correlated with students' and schools' observed background variables. Particularly reassuring is the lack of any correlation with students' parental education and with lagged test scores in panel B. In Panel C

of Table 4 we also show that the correlation of the estimated CEO quality between the outgoing and incoming CEOs of a school is 0.005, providing further evidence against any pattern of selection. Another point that underscores the relatively marginal importance of these imbalances is that our method I for estimating the effect of quality on school outcomes includes a school-fixed effect. This specification is feasible because we observe each school with two different CEOs. Finally, we note that even though the sample we use in method I consist of only 82 schools, it still yields meaningful variation in CEO quality. This evidence is presented in Table 5. For example, the minimum z-score of CEO quality is -3.467 SD, and the maximum is 2.213 SD.

6. Estimated Impact of CEOs' Quality

We estimate the effect of CEOs' quality twice for each method, first using school-level means for all variables and secondly with student-level data. Using student microdata allows estimation with a much larger sample, potentially leading to more precise estimates. But more importantly, it also allows estimation of treatment heterogeneity by students' characteristics: particularly important in this context is the heterogeneity of the effect of CEO quality by students' socio-economic background.

The first row of Table 6 presents estimations with school-level data of the effect of CEO quality on students' achievement in math, English, and Hebrew. In columns 1-3 we report results on CEO quality based on method II. In column 1 the regression includes a year effect as a control; in column 2, subject-fixed effects are added as a control; in column 3, schools' and mean students' characteristics are added as well. The estimates in columns 1-3 are positive, but only the estimate in column 3, 0.058 (se=0.03), is significantly different from zero at the 10% significance level. Since CEO quality and the test scores are standardized to mean zero and a unit standard deviation, this estimated effect implies that one standard deviation increase in CEO quality increases test scores in the three subjects by 5.8% of a standard deviation of the test score distribution. The respective estimates using student-level data (presented in the second row) are very similar, and the estimate in column 3 is 0.051, very close to the estimate obtained with school means, but it is less precisely measured.

In column 4 of Table 6, we present the estimates of the effect of CEO quality that we derive using method I. We note again that here we use the sample of 'switchers', schools that changed CEO

between two periods. Therefore, these schools will have a different CEO in each period, so their CEO quality ($\hat{O}_{jst}$) will be period (t) specific. We prefer using the CEO quality measure from the "opposite" period, i.e. using the 2002-03 measure of CEO Quality for a CEO where that measure is related to a school's outcomes in 2004-05 and vice versa. We think this 'conservative' approach is preferred over using the common element of quality of CEO from both estimates, especially if there are time-specific shocks to the cluster of schools that affect the measures of CEO quality in specific periods.

The GEMS data provide panel data on schools where each school participates in two rounds of national testing, between which it experiences a change in its CEO. Schools that participated in the 2002 testing round were also included in 2004. Similarly, schools included in the 2003 testing are also included in 2005. We stack this panel data so that α_j can be estimated as a school-fixed effect, meaning that we regress the change in school test scores on the change in CEO fixed effects. The advantage of this school fixed effect model (equivalent to a difference equation at the school level) is that it controls for omitted time-invariant variables biases that could potentially correlate to CEOs' quality. The within-school estimate (regressions with schools fixed effect) presented in column 4 is 0.038, but it is only marginally significantly different from zero. The estimate in the second row of column 4 (based on student-level data and clustering standard errors at the CEO by year level) is similar, 0.040, but it is much more precisely estimated (se=0.015).[13]

Remarkably, the two measures of CEO quality yield a similar estimated effect size, especially when comparing the estimates using the micro students' data: 0.051 using method II versus 0.040 while using method I. Both are statistically different from zero, and we cannot reject the hypothesis that they are not different from each other.

Evidence from a Sample of Middle Schools: Since the middle school system in Israel (grades 7-9) shares the same model of CEOs as that of primary schools, we replicated the analysis presented above

[13] As a robustness check we also estimated this model while restricting the sample to include only schools where changes in CEOs between the two periods are strictly according to spatial consideration so sorting of CEO's is very unlikely. We first used a sample that excluded the 15 schools that were reassigned as a single-school or two-schools cluster and their reassignment is not consistent with spatial considerations. The estimates are presented in online appendix Table A3. Moreover, results are also robust to the exclusion of all the 40 schools that were reassigned as a single-school or two-schools cluster. These estimates are presented in online appendix Table A4.

with the data for middle schools. We used the CEO quality based on method I because the sample of middle schools is much smaller than the primary schools' sample: only 15 schools had changed their CEO from 2002-2003 to 2004-2005. During the two periods, 13 different CEOs were the CEOs of these schools. Using the students' level data, the estimated effect of quality from equation 2 for middle schools is presented in Table A5 in the online appendix. The estimated effect is 0.048, very close to the corresponding estimate based on the sample of primary schools (0.040, se=0.016) but much less precisely estimated (se=0.055), probably due to the much smaller sample of schools, 15 versus 82.

Robustness Checks: We estimated the effect of the CEO's quality using two alternative estimation strategies. In the first strategy, we implement a Bayes shrinkage estimation strategy and construct an unbiased measure of CEO quality that accounts for noise in the measurement. Using this approach, the noisy measure of CEO quality is multiplied by an estimate of its reliability. The reliability of a noisy measure is the ratio of signal variance to signal variance plus noise variance. Thus, less reliable measures are shrunk back toward the mean of the distribution of the CEO quality measure.[14] In the second strategy, we use a two-step bootstrapping algorithm to account for the estimation of CEOs' quality in the first step and adjust their estimated standard errors.[15]

In these alternative estimation strategies, we use both measures of CEO quality, based on methods I and II. We present estimates based on the third specification (of column 3 in Table 6) for Method II. The results are presented in Table 7. The two alternative estimation strategies yield similar

[14] Following Morris (1983) and the teacher value added literature (for example, Kane and Staiger (2008) we construct the EB shrinkage factor for CEO *i* by the ratio of signal variance to signal variance plus noise variance of CEO *i*. In a similar way to the teacher value added literature, we assume that the measure of CEO bias includes an error component. Thus, estimating CEOs' effects on students' test scores enables separation of the signal variance (variance of CEOs' effects) and noise variance of CEO *i* (variance of the residuals for CEO *i*). The EB estimate for each CEO is a weighted average of the CEO estimated effect and the mean of CEO estimates, where the weight is the EB shrinkage factor. Implementing this methodology, the less reliable estimates of CEO quality (those with a large variation in estimated residuals) are shrunk towards the mean of CEO estimates.

[15] The bootstrap estimates of the standard errors are constructed as follows. In a first step, a random sample with replacement is drawn from each CEOs' schools. A new measurement of CEO quality for each CEO is created, based on the new sample of schools. In a second step, the effect of these new quality measures on student test scores in 5th grade is estimated (based on the preferred specification presented in Table 6) and the coefficients are stored. This process of two-step bootstrap sampling and estimation is repeated 1,000 times. The standard deviations in the sample of 1,000 observations of coefficient estimates from the second step are the bootstrap standard errors of the estimated effects of CEO quality.

results to the corresponding estimates presented in Table 6. In addition, the standard errors of the bootstrapping algorithm are almost identical to those in Table 6. Accounting for the fact that empirical Bayes estimates are smaller in absolute values than the initial estimates because of the shrinkage procedure (before standardization, the pre-shrinkage mean value of CEO quality is -0.17 while its post-shrinkage value is 0.013), we compare the elasticities of CEO quality effects at their mean values which yield comparable outcomes. We view this finding that the two completely different methods of measuring CEO quality, implemented while using two non-overlapping samples of schools, yield the same point estimates of the effect on students' outcomes as supporting the credibility of the methods we use to measure CEO quality and the causal interpretation of the findings.

In the paper, we use the number of students and classes in each school as part of the school characteristics control variables, even though we believe that those variables cannot be influenced by the school CEO, as they are decided by fixed school enrolment zones which the CEOs do not affect, we wanted to alleviate any concern of these variables turning into bad controls. In online Appendix table A6, we run our main analysis without those two control variables. The results are not affected by this omission in magnitude and significance level.[16] Furthermore, to test the robustness of our results to additional control variables, we add to our analysis, as control variables, available data from 1998 on the number of students and the socio-economic decile of the school, as determined by the Ministry of Education, which were available for about 94% of the schools in our sample. The results are robust to the addition of these controls and are even stronger (e.g. 0.073 with se=0.032 for quality method II and 0.040 with se=0.017 for quality method I for the school-level results with full controls).

Given that our sample of CEOs is relatively small, we wanted to examine how sensitive our results are to outliers. To examine this, we remove, for both methods, CEOs for which the estimated quality was more than 1.5 SDs away from the mean[17]. We then implement both methods without those CEOs. For the quality method I, 3 CEOs were dropped from the estimation (-3.57, -2.77, and 2.08 SDs

[16] In a separate analysis, we find that the number of students or classes is not affected by CEO quality, further demonstrating the unlikelihood of these variables turning into bad controls, results are not presented in the paper and are available from the author.

[17] Setting the cut-off anywhere in the range between 1.22 and 2 SDs away from the mean would not have changed the CEOs chosen to be dropped from the estimation, making the CEOs identified as outliers the only natural choice for that distinction.

away from the mean), and the effect of quality estimated was 0.024 with se=0.016 and 0.036 with se=0.011 for the school level and student levels results with full controls respectively. For quality method II, also 3 CEOs were dropped from the estimation (-2.43, -2.33, and 2.25 SDs away from the mean), and the effect of quality estimated was 0.045 with se=0.02 for both the school level and student levels results with full controls. We can observe that without the highest and lowest quality CEOs, the estimates are only slightly smaller than our main outcomes in Table 6, and the effect remains significant or close to significance in all the specifications.

Finally, Chetty et al. (2014a) study the best linear unbiased prediction forecasts, showing that if the estimates are unbiased, the coefficient of teacher value-added should be equal to 1 in out-of-sample regressions. In this paper, our main focus is not on a similar analysis of the best linear unbiased prediction for our measures, partly due to the smaller sample size of CEOs available, and rather attempt to present supporting evidence relating to the importance of CEOs in the education system for student outcomes. Nevertheless, when following the procedures as in Chetty et al. (2014a) and after adjusting for the re-standardizations of our quality measures and student outcomes, we find that our second stage (out of sample) coefficients are 0.734 and 0.921 for our preferred individual sample specification for quality methods I and II respectively. Our coefficients are not equal to 1, as expected with the smaller sample size of CEOs, and perhaps potentially due to the non-linear connection between CEO quality and student test scores and the differing properties of our two samples (panel data vs. cross-section), all which might not allow for a completely unbiased prediction of CEO quality. Still, the fact that our coefficients are not far from 1 (especially for method II where we estimate quality using panel data) provides supportive evidence for the validity of our estimators.

Comparing the Effect Size of CEO Quality to Other Related Treatments: It is useful to benchmark the effect sizes in columns 3 and 4 of Table 6 against the effect of teachers' added value. For example, Rockoff (2004) and Rivkin et al. (2005) suggest that a one standard deviation increase in teacher quality improves student math scores by about 0.1 standard deviations. Aaronson et al. (2007) find similar results using high school data. Branch et al. (2012) report that a one standard deviation increase in principal leadership (i.e., a principal in the top 16% of the quality distribution) leads to a 0.05 standard

deviation gain in test scores of all students in the school. Chetty et al. (2014b) find that one standard deviation improvement in teacher value-added in a single grade in primary school in New York City raises the probability of college attendance at age 20 by 0.82 percentage points, relative to a sample mean of 37%. Compared to other schooling interventions, it would require one additional hour of instruction per week in math, Hebrew, and English (a 25% increase) to achieve the same effect as a one standard deviation increase in CEO value-added (Lavy, 2015; Lavy, 2020).

Effect Size Heterogeneity: Since 80% of the CEOs in our sample are female, we report in Panel B of Table 6 the estimates from regressions when the sample is restricted to schools with female CEOs. The estimated effect based on the school-level means and student-level regressions is 0.061 (se=0.032) and 0.053 (se=0.034), respectively, when using quality method II. The estimated effect when using the quality method I is 0.048 (se=0.022) and 0.053 (se=0.011) based on the school-level means and student-level regressions, respectively. These estimates are larger than when male CEOs were also included in the sample. However, they are not significantly different from the full sample estimates presented in panel A of Table 6. The estimates in Table 7 panel B based on alternative estimation strategies present similar evidence.

In Panel C of Table 6, we present evidence based on sub-samples of students by parental education. Father's or mother's years of schooling are good proxies for students' socio-economic backgrounds. Research on the causal impact of school inputs suggests that students from a disadvantaged background benefit more from smaller class sizes, remedial education, higher quality peers, and teacher quality (For examples, see Lavy et al. (2022), and Lavy et al. (2012)). Our estimates show no such differences concerning the effect of CEO management quality in public education. The estimates in Table 7 panel C based on alternative estimation strategies show similar evidence.

When looking separately by each of the three subjects at the effects of CEO quality on student test scores, we find slightly stronger effects on math scores when using quality method II. When using quality method I, the effects are similar across the three subjects. Estimates are presented in online appendix Table A7. Interestingly, the correlation of CEO value-added across the three subjects is very high for both the sample of schools that did not change CEO and those that did, ranging between 0.783

and 0.849, see Table A8 in the online appendix for details—providing further support for the robustness of our value-added measures.

Non-Linear Effect of CEO Quality: To check for potential non-linearity in the effect of CEO quality, we report in Table 8 estimates where we divide the distribution of CEO quality into ranges from low to high. First, we divide the distribution to above and below the median of CEO quality. Using CEO quality based on method II, the estimate for the above-median CEO quality indicator is 0.082 (se=0.036), and using quality derived with method I yields a marginally lower estimate, 0.065 (se=0.036). Secondly, we divide the distribution into quartiles of CEO quality. The estimates for the three upper quartile quality indicators are increasing monotonically: with quality measures using method II the estimates are 0.080, 0.099, and 0.135. With quality measures using method I they are 0.031, 0.055, and 0.130. The pattern in both cases is similar and suggests a non-linear effect of CEO quality with a monotonically increasing effect. The estimated effect of the upper quartile of ability (0.135 for method II, 0.130 for method I) is statistically different from zero at the five percent level of significance. Still, we cannot distinguish them statistically from the effects of the 2nd and 3rd quartiles of ability. Note also that the effect of the highest quartile is identical when using either of the two alternative quality measures.

Clearly, the effect of CEOs' quality is non-linear, and it increases with quality. It is natural and interesting to compare this pattern of the non-linear effect with that of the effect of teachers' value-added, but we did not find studies that provide such evidence.

7. Mechanisms of the Effect of CEO Quality

The results reported above show that schools exhibit higher achievements when they have a higher-quality CEO. This section explores several potential mechanisms through which CEOs' quality may affect their students' academic achievement. We use a rich set of school practices and outcomes

based on responses to questionnaires of primary school principals, teachers, and students.[18] We focus on items that relate to the classroom and school environment (student questionnaire), on school activities and programs aimed at improving school climate and students' norms, and on school resources (principal questionnaire), school procedures, and teachers' on and off the job training (teacher questionnaire). We also group outcomes into eight categories to obtain a more general picture of the possible mechanisms and increase statistical power. We analyse each category by creating a category-specific average effect. This allows us to control for the potential problem of over-rejection of the null hypothesis due to multiple inferences. Because different outcomes have different data scales, simply averaging the estimators for the treatment effect is not likely to produce a meaningful statistic. To address this concern, we follow the summary-index approach per Kling et al. (2007). The average effect of multiple outcomes is the average of z-scores of each outcome variable. This summary index is a special case of the z-score and is identical to the mean effect size of treatment if there are no missing values.[19] In general, the sign of the summary index reveals information on the direction of the aggregate impact of a class of outcomes, and the more the summary index deviates from zero, the stronger is the implied aggregate effect.

We are aware, of course, that we are not able to measure all the relevant mechanisms, and we cannot rule out the possibility that other mechanisms are in place, but the analysis presented in this

---

[18] See H. Jerome Freiberg (2005) and J. Barry Fraser (2012) for recent reviews of the educational research literature about the validity of students' and teachers' assessments of the classroom environment and their associations with students' achievements.

[19] In the regression specification this approach yields standardized estimators as follows: the treatment effects for K outcomes are aggregated and reflected in a single standard normal statistic,

$$\tau = \frac{1}{K}\sum_{k}\frac{\boldsymbol{\beta}_{1\mathrm{k}}}{\sigma_{k_C}}, \quad k = 1, \dots, K$$

where $\boldsymbol{\beta}_{1k}$ indicates the average treatment effect for outcome k and σ_{k_C} denotes the standard deviation of the k^{th} control outcome. Having included the covariates, the K average treatment effects ($\boldsymbol{\beta_1}$) and sample variances can be easily acquired through a linear regression. By doing so, the above equation can be thought of as a point estimator representing a collection of standardized treatment effects. However, this paper also takes into account the covariance of effects and therefore adapts a seemingly unrelated regression (Zellner, 1963; Kling et al., 2007):

$$\mathbf{Y} = \boldsymbol{I_K}\otimes(\boldsymbol{T} \quad \boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{\upsilon}$$

Where $\boldsymbol{T}$ is the treatment indicator(s), and $\boldsymbol{X}$ consists of controlled regressors as well as a constant term.

section provides important insights regarding some possible mediating factors that drive the positive effect of the school CEO on students' achievements.[20] Our hypothesis is that if the effects of the CEO quality are partially driven by a particular mediating factor, observing a significant effect of the CEO quality on this factor provides some evidence for the validity of this hypothesis.[21] We use in this section CEO quality based on method I as our measure of CEO quality because its estimation is based on a sample where each school is observed twice, each time with a different CEO, which allows for the inclusion of schools' fixed effects. This within-school estimation controls for any time-invariant school heterogeneity.

***School Priorities, Working Procedures and Resources*:** In Table 9 we present evidence regarding three potential channels for the effect of CEO quality: priority setting by schools and working procedures, and school resources. Three items in the teachers' questionnaire ("school has clearly defined priorities", "teachers are involved in setting school priorities", and "school has clearly defined working procedures") reflect management practices that can affect the allocation and use of school resources and therefore be conducive or harmful to learning and achievement. Clearly the scope, responsibilities and management directives of a CEO can affect these school factors by the frequent interaction with school principals and teachers. In panel A we present treatment effect estimates for these outcomes, the mean of which is very high, over 5 on a scale of 1-6. We use three different specifications, the preferred one is the within-school estimation (presented in column 4). The first two rows in panel A show that CEOs of higher quality are associated with more focused school priorities, involving teachers in their setting, and more clearly defined working procedures. Panel B of Table 9 provides evidence on the

[20] A further limitation is that we cannot identify the causal effect of the mechanisms on outcomes because the former are numerous and we have only one potential instrument.

[21] Lavy and Schlosser (2011) who use the same data, show in online Appendix Table 5 that all indicators of the quality of the classroom environment, as described by the students, are highly correlated with students' academic performances even after controlling for school fixed effects and students' background characteristics. For example, they report that lower levels of classroom disruption and violence, better inter-student relationships, and a higher quality of interaction between teachers and students are all positively associated with students' test scores. Though they do not provide a causal interpretation to these correlations, their results suggest that students' assessments of their classroom environment have a high informational content and that these mechanisms, as pointed out in the educational literature, might play an important role in student's learning process.

impact of the CEO's quality on school resources, including outcomes that are less likely to be affected by the CEO, instruction budget per class, length of the school week, and total instruction time on different subjects. National or regional educational authorities determine these school inputs, and the CEO should not be able to influence them. Indeed, we find no significant effect on these inputs. We also find no significant effects of CEO quality on instruction time for different subjects and on the time students allocate to homework, these results are presented in Table A9 in the online appendix.

***School climate programs and outcomes*:** From two items in the school principal questionnaire we define an indicator of whether, in the current year, the school assessed its climate and norms and whether it had interventions aiming at improving norms, values, and discipline of students. In panel A of Table 10, we report estimates from regressions when each of these two indicators is the dependent variable. 71% of the schools in the sample had an assessment of discipline, violence, and norms in school in the current year, while 29% of schools had interventions targeted at improving these aspects of the school environment. CEOs' quality has a positive and statistically significant effect on both of these outcomes. The average effect presented in the third row of panel A shows a similar positive effect (0.262, se=0.077).

In panel B of Table 10, we present estimates of the effect of CEOs' quality on classroom and school violence. This analysis is based on the following items from the students' questionnaire: (1) "There are many fights among students in my classroom."; (2) "This year I was involved in many fights"; (3) "Sometimes I'm scared to go to school because there are violent students." The estimates reported in panel B columns 2–4 of Table 10 suggest that a higher CEO quality significantly lowers the level of violence in school. This effect is evident in each of the three items and their average effect. For example, the estimate for the effect of CEO quality on students' reports regarding the level of violence in the classroom is −0.035 (se = 0.018). The average effect of these three items is more precise than the estimates for the individual items. The average estimate is −0.033 (se = 0.012). Overall, these results suggest that having a higher quality CEO improves the safety climate in school by lowering the incidence of fights, increasing the safety of students, and lowering their anxiety about attending school.

Three items in the student's questionnaire ("When I have a problem at school, there is always someone I can turn to (from the teaching staff)", "I feel well-adjusted socially in my class," and "I am satisfied in school") reflect school environment and climate, in particular the relationships between students and teachers and the quality of inter-student relationships. These factors can be conducive or harmful to learning and achievement. For example, being well-adjusted and accepted socially among classroom peers may improve a student's self-confidence, self-image, motivation, and other non-cognitive attributes that might be essential for effective learning.[22] In panel C we present treatment effect estimates for these outcomes, the means of which are very high, over 5 on a scale of 1-6. The within school estimates show that higher CEO quality improves only the first outcome, reflecting student-teacher relationships, but not the others. The estimates of the effect on social and school satisfaction are positive but small and not statistically different from zero, and so is the average effect.

Scholastic Programs and Teachers Training: In Table 11, we present evidence on three additional potential channels for the effect of CEO quality: school scholastic programs in Hebrew, math, and English, and two forms of teachers' training in each of these three subjects, in service on the job training and external courses. The subject-specific programs include additional instructional resources for a given subject or improvements to teaching methods and practices. The effect of CEO quality on scholastic programs is positive and significant in English but not in programs for the other two subjects.

The effect of CEO quality on teachers' training outside of schools is practically zero, as seen by the estimated effect on external training in each of the three subjects. The estimated effect of on-the-job training is positive and significant. In math, it is positive but not precisely measured. In English, it is negative but not different from zero. The average effect is positive but imprecise, suggesting that the overall effect is negligible, but the effect on the composition of the in-school training is meaningful.

[22] Table A10 in the online appendix provides further evidence on the relationship between school climate and violence, and students' scores. Even though this evidence cannot be interpreted as strictly causal, the large and statistical within school conditional correlations with students' test scores provide additional support for the idea that school climate and violence are mechanisms through which the schools' CEOs affect students' outcomes.

Angrist and Lavy (2001) estimated a large effect of such on-the-job in-school training on students' test scores.

***Changing School's Principal*:** In Table 12 we present evidence on another important channel for the effect of CEO quality, changing the school principal. About 15% of the schools in the sample change their principal every year, implying that the average tenure of a school principal is about 6-7 years. In columns 1-3 of panel A we present estimates of the effect of an indicator of change in the school CEO on the probability of a school changing its principal. In this estimation, we use the full sample of schools because we observe whether and when they experienced a change in CEO. The likelihood of a school changing its principal is about 6% lower in the year a new CEO steps in, and it is 9% higher the following year. When the contemporary and one-year lagged effects are included jointly, the estimates change marginally, though the pattern is the same. In columns 4-9 of panel A, we present the estimates for each year separately. The 2004-2005 academic year results are most interesting because we can estimate the contemporaneous effect and the one-year and two-year lag effects. The pattern that emerges is striking: the likelihood of a change in school principal in a new CEO's first year is lower by 7.6%; higher by 14% in the second year, and zero in the third year. These estimates suggest that a change in school management is associated with the engagement of a new CEO. In panel B of Table 12, we further explore the connection between CEO quality and the probability of changing a principle. In column 3 of panel B, we see that an increase of one standard deviation in CEO quality increases the probability of changing a school principle by 9.3 percentage points, which is almost 28% of the mean share of schools changing principals after a CEO change. Further emphasizing the potential importance of this channel for the effects of CEO quality. It is possible that the other mechanisms discussed above operate through the principal rather than through other channels. However, since all the mechanisms are endogenous, including the change of school principal, we cannot study the interlinkages between them. Nonetheless, even if the change in school principal mediates the reduced form effects on the channels we examined above, it is clear that it is the consequence of the decision of the CEO to change the school principal and her actions.

***CEOs as Troubleshooters/Problem Solvers*:** In Table 13, we present evidence on another potential channel for the effect of CEO quality. While the view that one CEO may benefit all the schools they

lead is certainly a valid one, a possible alternative is that CEOs may disproportionably focus their attention on the weaker/more troublesome schools and act as troubleshooters or problem solvers. To investigate this hypothesis, in Table 13, we divide our sample into above and below-median schools, for each CEO, on several potential measures which can represent relatively weaker/troublesome schools in the pre-treatment period. These potential measures are (1) Scores on the national standardized exams in period 0 and 1991; (2) Parent education levels; (3) "Madad Tipuah" - the socio-economic decile of the school, as determined by the ministry of education in 1998. We find that the effects of CEO quality are heterogeneous, with the effects being much stronger for the relatively weaker/troublesome schools of each CEO. Suggesting a potential mechanism of effect through increased attention to weaker schools and troubleshooting or solving issues that tend to arise more in those schools.[23]

8. Concluding Remarks

This is the first paper that measures the causal effect of the quality of management supervision in public education. As in many other countries, the structure of the education system in Israel includes a CEO who is in charge of a cluster of about 15 schools over which they have extensive responsibility and authority. We exploit quasi-random turnover of CEOs over time to measure their quality in terms of test score gains in English, Hebrew, and math in primary schools in Israel. This turnover in the system is largely dictated by a routine that rotates CEOs across schools every 3-5 years, as well as by other naturally occurring events such as retirements. We show that geographical considerations largely determine which schools are included in a CEO's cluster of schools. We further demonstrate that the turnover of CEOs is unrelated to schools' or students' potential outcomes. We measure CEOs' quality similarly to recent measures of teachers' quality, including controls for school-level lagged outcomes. We construct two alternative quality measures: the first based on the common academic achievement level of all schools that share the same CEO, and the second based on the common growth in academic achievement shared by all schools that were assigned to the same new CEO. We then show that both measures of CEOs' quality are uncorrelated with students' and schools' characteristics or lagged scores.

[23] We would like to thank an anonymous referee for suggesting this idea.

We find that CEO quality positively impacts primary school students' math, Hebrew, and English scores. Based on the first quality measure, a one standard deviation improvement in CEO quality leads to an increase of approximately 0.04 standard deviations in test scores. The effect is similar for students from lower and higher socioeconomic backgrounds; it is highly non-linear, increasing sharply for CEOs in the highest quartile of the quality distribution and larger but not statistically significantly so for female CEOs. We obtain similar results when using the second measure of quality. We explore several mechanisms for these effects and find that higher-quality CEOs are associated with more focused school priorities and clearly defined working procedures. We do not find an effect on school resources – as funding is determined centrally at the Ministry of Education and do not find an effect on teachers' on-the-job and external training. Schools with higher quality CEOs are more likely to address school climate, violence, and bullying and implement interventions that lead to lower violence in school and higher social school satisfaction among students. It is interesting to note that this channel of effect of reducing violence and bullying in school is consistent with the strong discipline 'no excuse' philosophy of charter schools (Dobbie and Fryer, 2013; Angrist et al., 2013) which also was effective in improving school outcomes when implemented in non-charter public schools (Fryer, 2014).

Two additional remarks point to the relevance of our findings for public policy. First, we note that the effect size of CEO quality may imply that it is potentially cost-effective because an increase in the quality of one manager can improve the outcomes of thousands of students. Therefore, while still being cautious pending an investigation into the costs of improving CEO quality, investing in the quality of these school CEOs is very compelling from a resource perspective, compared to investing in other school inputs such as teacher quality or instructional time. Second, about 70% of all CEOs are women, very different to the relative scarcity of women in leadership roles in other public sector management roles as well as in the private sector. Our findings suggest that the quality of women as CEOs does not fall short of that of men in the same position, raising the policy concern of why there are so few women in leadership roles even in the public sector, and how allocative efficiency can be improved in this regard in the labour market.

University of Warwick, UK, *Hebrew University*, *Israel* & NBER
Northwestern University, USA
Hebrew University, Israel

Additional Supporting Information may be found in the online version of this article:

Online Appendix

Replication Package

References

Angrist, J. D., Pathak, P. A. and Walters, C. R. (2013). `Explaining charter school effectiveness', *American Economic Journal: Applied Economics*, vol. 5(4), pp. 1-27.

Angrist, J. and Lavy, V. (2001). `The Effect of Teachers' Training on Student Achievements', *Journal of Labor Economics*, vol. 19(2), pp. 343-369.

Aaronson, D., Barrow, L. and Sander, W. (2007). `Teachers and student achievement in the Chicago public high schools', *Journal of labor Economics*, vol. 25(1), pp. 95-135.

Bacher-Hicks, A., Kane T. J. and Staiger D. O. (2014): `Validating Teacher Effect Estimates Using Changes in Teacher Assignments in Los Angeles', Working paper 20657, National Bureau of Economic Research.

Bertrand, M. and Schoar, A. (2003). `Managing with style: The effect of managers on firm policies', *The Quarterly journal of economics*, vol. 118(4), pp. 1169-1208.

Bloom, N., Propper, C., Seiler, S. and Van Reenen, J. (2015). `The impact of competition on management quality: evidence from public hospitals', *The Review of Economic Studies*, vol. 82(2), pp. 457-489.

Bloom, N. and Van Reenen, J. (2007). `Measuring and explaining management practices across firms and countries', *The quarterly journal of Economics*, vol. 122(4), pp. 1351-1408.

Bloom, N., Lemos, R., Sadun, R. and Van Reenen, J. (2015). `Does management matter in schools?', *The Economic Journal*, vol. 125(584), pp. 647-674.

Branch, G. F., Hanushek E. A. and Steven G. Rivkin S. G. (2012). 'Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals', Working Paper 17803, National Bureau of Economic Research.

Bryson, Alex, Lucy Stokes and David Wilkinson (2018) `Can HRM Improve Schools' Performance?', Discussion Paper 11348. IZA Institute of Labor Economics.

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014a). `Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates', *American economic review*, vol. 104(9), pp. 2593-2632.

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014b). `Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood', *American economic review*, vol. 104(9), pp. 2633-79.

Clark, D., Martorell P. and Rockoff J. E. (2009). 'School Principals and School Performance', CALDER Working Paper 38. Urban Institute (NJ1).

Coelli, M. and Green, D. A. (2012). `Leadership effects: School principals and student outcomes', *Economics of Education Review*, vol. 31(1), pp. 92-109.

Di Liberto, A., Schivardi, F. and Sulis, G. (2015). `Managerial practices and student performance', *Economic Policy*, vol. 30(84), pp. 683-728.

Dobbie, W. and Fryer Jr, R. G. (2013). `Getting beneath the veil of effective schools: Evidence from New York City', *American Economic Journal: Applied Economics*, vol. 5(4), pp. 28-60.

Dhuey, E. and Smith, J. (2014). `How important are school principals in the production of student achievement?', *Canadian Journal of Economics/Revue canadienne d'économique*, vol. 47(2), pp. 634-663.

Fraser, B. J. (2012). Classroom environment. Routledge.

Freiberg, H. J. (2005). School climate: Measuring, improving and sustaining healthy learning environments. Routledge.

Fryer Jr, R. G. (2014). `Injecting charter school best practices into traditional public schools: Evidence from field experiments', *The Quarterly Journal of Economics*, vol. 129(3), pp. 1355-1407.

Grissom, J. A., Kalogrides, D. and Loeb, S. (2015). `Using student test scores to measure principal performance', *Educational evaluation and policy analysis*, vol. 37(1), pp. 3-28.

Horng, E. L., Klasik, D. and Loeb, S. (2010). `Principal's time use and school effectiveness', *American journal of education*, vol. 116(4), pp. 491-523.

Rasul, I. and Rogger, D. (2018). `Management of bureaucrats and public service delivery: Evidence from the nigerian civil service', *The Economic Journal*, vol. 128(608), pp. 413-446.

Jacob, B. A., Lefgren, L. and Sims, D. P. (2010). `The persistence of teacher-induced learning', *Journal of Human resources*, vol. 45(4), pp. 915-943.

Kane, T. J., and Staiger, D. O. (2008). `Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation', Working Paper 14607, National Bureau of Economic Research.

Kling, J. R., Liebman, J. B. and Katz, L. F. (2007). `Experimental analysis of neighborhood effects', *Econometrica*, vol. 75(1), pp. 83-119.

Lavy, V. (2015). `Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries', *The Economic Journal*, vol. 125(588), pp. 397-424.

Lavy V. (2020). `Expanding School Resources and Increasing Time on Task: Effects on Students' Academic and Non-Cognitive Outcomes', *Journal of the European Economic Association*, vol. 18(1), pp. 232-265.

Lavy, V., Paserman, M. D. and Schlosser, A. (2012). `Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom', *The Economic Journal*, vol. 122(559), pp. 208-237.

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

Lavy, V., Kott, A. and Rachkovski, G. (2022). `Does Remedial Education in Late Childhood Pay Off After All? Long-Run Consequences for University Schooling, Labor Market Outcomes, and Intergenerational Mobility', *Journal of Labor Economics*, vol. 40(1), pp. 239-282.

McCormack, J., Propper, C. and Smith, S. (2014). 'Herding cats? Management and university performance', *The Economic Journal,* vol. 124 (578), pp. 534–564.

Morris, C. N. (1983). `Parametric empirical Bayes inference: theory and applications', *Journal of the American statistical Association*, vol. 78(381), pp. 47-55.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005). `Teachers, schools, and academic achievement', *Econometrica*, vol. 73(2), pp. 417-458.

Rockoff, J. E. (2004). `The impact of individual teachers on student achievement: Evidence from panel data', *American economic review*, vol. 94(2), pp. 247-252.

Rothstein, J. (2010). `Teacher quality in educational production: Tracking, decay, and student achievement', *The Quarterly Journal of Economics*, vol. 125(1), pp. 175-214.

Zellner, A. (1963). `Estimators for seemingly unrelated regression equations: Some exact finite sample results', *Journal of the American Statistical Association*, vol. 58(304), pp. 977-992.

Table 1. *Number of Schools and Number of CEOs*

	Year			
	2002	2003	2004	2005
A. All CEOs				
Number of CEOs	54	56	57	54
Number of Schools	810	807	803	797
Number of Schools per CEO	16.1	15.5	15.1	15.9
	(4.92)	(5.67)	(5.52)	(6.10)
Male Sample				
Number of CEOs	14	15	13	10
Number of Schools	163	163	150	108
Number of Schools per CEO	13.9	12.9	13.6	13.5
	(4.43)	(5.11)	(4.72)	(4.20)
Female Sample				
Number of CEOs	40	41	44	44
Number of Schools	647	644	653	689
Number of Schools per CEO	16.9	16.4	15.5	16.4
	(4.90)	(5.64)	(5.71)	(6.38)
B. CEOs with Turnover of Schools				
Number of CEOs	-	35	21	26
Number of Schools	-	563	309	448
Number of Schools per CEO	-	16.7	15.8	18.4
		(5.24)	(3.95)	(5.35)
C. CEOs with Net Turnover of Schools > 0				
Number of CEOs	-	15	9	11
Number of Schools	-	277	135	234
Number of Schools per CEO	-	18.9	16.0	22.1
		(4.03)	(3.64)	(4.48)
D. CEOs with Net Turnover of Schools < 0				
Number of CEOs	-	18	10	11
Number of Schools	-	238	142	135
Number of Schools per CEO	-	14.1	15.0	14.3
		(4.76)	(4.22)	(3.69)
E. Final Analysis Sample of CEOs				
Number of CEOs	27	29	29	26
Number of Schools	428	461	461	445
Number of Schools per CEO	16.7	16.6	16.7	18.0
	(5.40)	(5.24)	(4.75)	(5.72)
F. Retirement of Veteran CEOs and Entry of New CEOs				
Number of New CEOs	8	6	5	4
Number of Schools per New CEO	13.1	12.5	11.5	13.9
	(6.53)	(6.72)	(6.94)	(7.36)
Number of Retiring CEOs (at year's end)	4	4	7	-
Number of Schools per Retiring CEO	13.6	12.3	13.7	-
	(3.69)	(6.00)	(5.91)	
Average number of New CEOs' Schools from Retired CEOs	-	2.5	5.6	12.0
	-	(3.02)	(4.83)	(7.57)

Notes: Standard deviations are presented in parenthesis. The sample includes all schools in the Jewish non-religious primary school system.

Table 2. *Regressions for Estimating CEOs' Quality, Method I*

	Period 0: 2002-2003		Period 1: 2004-2005	
	Mean (SD)	Regression Coefficient	Mean (SD)	Regression Coefficient
	(1)	(2)	(3)	(4)
School Average Characteristics				
Male	0.508	-0.130	0.504	-0.106
	(0.500)	(0.013)	(0.500)	(0.012)
Number of Siblings	2.115	-0.050	2.114	-0.049
	(1.194)	(0.004)	(1.204)	(0.006)
Immigrant	0.147	-0.093	0.118	-0.219
	(0.354)	(0.083)	(0.322)	(0.120)
Father's Education	12.625	0.033	12.613	0.026
	(3.542)	(0.002)	(3.641)	(0.003)
Mother's Education	12.968	0.048	12.965	0.035
	(3.067)	(0.003)	(3.271)	(0.003)
Ethnicity - Asia Africa	0.146	-0.031	0.096	-0.045
	(0.353)	(0.013)	(0.295)	(0.019)
Ethnicity - Europe America	0.159	0.063	0.154	0.082
	(0.365)	(0.016)	(0.361)	(0.014)
Ethnicity - Ethiopia	0.019	-0.223	0.017	-0.130
	(0.137)	(0.074)	(0.130)	(0.045)
Ethnicity - Former USSR	0.109	0.011	0.086	0.252
	(0.312)	(0.099)	(0.281)	(0.125)
Ethnicity - Other	0.032	0.101	0.027	0.239
	(0.176)	(0.088)	(0.163)	(0.126)
School's Number of Classes	15.757	0.001	15.846	0.001
	(3.837)	(0.009)	(3.983)	(0.009)
School's Number of Students	461.931	0.000	458.217	0.000
	(144.421)	(0.000)	(147.471)	(0.000)
School Average Lagged Z-Score				
1991 4th Grade Math	0.017	0.013	0.018	-0.008
	(0.876)	(0.031)	(0.863)	(0.031)
1991 4th Grade Verbal	0.002	0.018	0.009	0.048
	(0.864)	(0.030)	(0.852)	(0.034)
1991 5th Grade Math	0.028	-0.044	0.029	0.003
	(0.882)	(0.023)	(0.876)	(0.023)
1991 5th Grade Verbal	0.005	0.041	0.009	0.013
	(0.869)	(0.027)	(0.860)	(0.029)
Year FE	-	Yes	-	Yes
Subject FE	-	Yes	-	Yes
District FE	-	Yes	-	Yes
CEO FE	-	Yes	-	Yes
Number of Students	32,805	32,805	31,174	31,174
Number of Schools	497	497	497	497
Number of CEOs	47	47	47	47

Notes: In columns (1) and (3) standard deviations are presented in parenthesis. In columns (2) and (4) standard errors in parenthesis are clustered at the CEO level. Schools characteristics include number of classes, number of students and the number of schools under the school's CEO. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Sample includes stacked math, Hebrew and English tests scores. Test scores are standardized with 0 mean and unit sd.

Table 3. *Regressions for Estimating CEO's Quality, Method II*

	Mean (SD)	Regression Coefficient
	(1)	(2)
School Characteristics		
Δ Male	0.001	-0.019
	(0.105)	(0.189)
Δ Number of Siblings	-0.034	0.026
	(0.380)	(0.081)
Δ Immigrant	-0.031	0.236
	(0.078)	(2.618)
Δ Father's Education	0.043	0.048
	(0.872)	(0.027)
Δ Mother's Education	-0.024	-0.082
	(0.716)	(0.041)
Δ Ethnicity - Asia Africa	-0.052	-3.287
	(0.072)	(13.185)
Δ Ethnicity - Europe America	-0.013	-3.244
	(0.089)	(13.122)
Δ Ethnicity - Ethiopia	-0.002	-2.939
	(0.021)	(13.266)
Δ Ethnicity - Former USSR	-0.027	-3.169
	(0.071)	(13.702)
Δ Ethnicity - Other	-0.004	-4.088
	(0.034)	(13.627)
Δ School's Number of Classes	-0.416	0.021
	(2.364)	(0.018)
Δ School's Number of Students	-13.633	0.000
	(67.449)	(0.001)
School Average Lagged Z-Score		
1991 4th Grade Math	-0.113	-0.012
	(0.900)	(0.041)
1991 4th Grade Verbal	-0.079	0.036
	(0.936)	(0.047)
1991 5th Grade Math	-0.130	0.007
	(0.907)	(0.037)
1991 5th Grade Verbal	-0.077	-0.041
	(0.939)	(0.038)
Year FE	-	Yes
Subject FE	-	Yes
District FE	-	Yes
CEO FE	-	Yes
Number of Students (Period 1)	13,605	13,605
Number of Schools	226	226
Number of CEOs	44	44

Notes: In column (1) standard deviations are presented in parenthesis. In column (2) standard errors in parenthesis are clustered at the superintendent level. Δ represents the change in a variable between the two periods. Schools characteristics include number of classes, number of students and the number of schools under the school's superintendent. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Sample includes stacked math, Hebrew and English tests scores. Test scores are standardized with 0 mean and unit sd.

Table 4. *Balancing Tests on Probability of Schools Changing CEO and on CEOs' Quality*

	Changing CEO			CEO Quality	
	Mean of Schools that did not change CEO	Mean of Schools that did change CEO	Probability of Changing CEO	Method II	Method I
	(1)	(2)	(3)	(4)	(5)
A. Student and School Characteristics					
Male	0.506	0.502	-0.004	0.009	0.009*
	(0.500)	(0.500)	(0.004)	(0.006)	(0.005)
Number of siblings	2.115	2.127	0.013	0.032	0.008
	(1.199)	(1.164)	(0.035)	(0.092)	(0.031)
Immigrant	0.133	0.139	0.006	0.004	-0.007
	(0.043)	(0.346)	(0.010)	(0.012)	(0.017)
Mother's years of education	12.966	13.033	0.067	-0.262	0.153
	(3.168)	(3.184)	(0.114)	(0.224)	(0.179)
Father's years of education	12.619	12.680	0.061	-0.301	0.202
	(3.591)	(3.654)	(0.129)	(0.233)	(0.223)
Ethnicity - Israel	0.576	0.587	0.011	-0.027	0.036*
	(0.494)	(0.492)	(0.013)	(0.022)	(0.020)
Ethnicity - Asia Africa	0.122	0.114	-0.007	0.006	-0.006
	(0.327)	(0.318)	(0.005)	(0.007)	(0.009)
Ethnicity - Europe America	0.157	0.154	-0.003	0.000	-0.020*
	(0.363)	(0.361)	(0.006)	(0.014)	(0.006)
Ethnicity - Ethiopia	0.018	0.010	-0.009*	0.018*	-0.002
	(0.134)	(0.097)	(0.003)	(0.006)	(0.003)
Ethnicity - Former USSR	0.098	0.099	0.001	0.007	0.000
	(0.297)	(0.299)	(0.010)	(0.015)	(0.014)
Ethnicity - Other	0.030	0.036	0.007	-0.004	-0.008
	(0.170)	(0.187)	(0.003)	(0.004)	(0.008)
School's Number of Students	460.121	451.097	-9.022	29.176	-24.821*
	(145.927)	(152.871)	(13.432)	(19.242)	(14.677)
School's Number of Classes	15.800	15.668	-0.132	0.714	-0.638
	(3.909)	(4.185)	(0.357)	(0.537)	(0.423)
B. Tests' Z-scores 2002-2003 (Period 0)					
Math	-0.001	0.010	0.011	0.052	0.002
	(1.004)	(0.986)	(0.025)	(0.077)	(0.017)
Hebrew	0.006	-0.005	-0.011	-0.009	-0.035
	(0.998)	(0.998)	(0.026)	(0.053)	(0.041)
English	-0.002	0.011	0.013	0.000	-0.028
	(0.999)	(1.005)	(0.034)	(0.089)	(0.074)
C. Correlation Between Previous and Current School's CEO Quality Based on Method II					
			$SI(VA)_{t+1} = \underset{(0.197)}{-0.153} + \underset{(0.222)}{0.0916} * SI(VA)_t$		
			$R^2 = 0.005$		
Number of Schools	497	222	719	301	82
Number of CEOs	47	55	67	30	29

Notes: In columns (1)-(2) standard deviations are presented in parenthesis. In column (3) standard errors in parenthesis are adjusted for school level clustering. In columns (4)-(5) standard errors in parenthesis are adjusted for CEO level clustering. All regressions include district dummies. In Panel C standard errors in parenthesis are adjusted for CEO level clustering. Number of schools and CEOs refer to Panels A-B. All schools are in the Jewish non-religious primary school system. CEOs VA are standardized with 0 mean and unit sd. * indicates significant level at 10% or lower.

Table 5. *Distribution of the Change of CEO's Quality Within Schools, Method I*

	Mean Change	SD	Min	Median	Max	Schools
	(1)	(2)	(4)	(5)	(6)	(7)
Full Sample	0.091	1.341	-3.467	0.228	2.213	82
	(0.154)					
Female CEOs Subsample	0.383	1.342	-3.467	0.572	2.213	64
	(0.177)					

Notes: Robust standard errors are presented in parenthesis. All 82 schools that are included in this sample are in the Jewish non-religious primary school system. The sample includes stacked math, Hebrew and English tests scores. CEOs Quality is standardized with 0 mean and unit sd.

Table 6. *The Effect of CEO's Quality on Tests Z-Scores*

	Quality Method II			Quality Method I
	(1)	(2)	(3)	(4)
A. Full Sample				
Schools' Sample Regressions	0.037	0.037	0.058	0.038
	(0.041)	(0.041)	(0.030)	(0.019)
Students' Sample Regressions	0.035	0.035	0.051	0.040
	(0.041)	(0.041)	(0.032)	(0.016)
B. Female CEOs Sample				
Schools' Sample Regressions	0.035	0.035	0.061	0.048
	(0.043)	(0.043)	(0.032)	(0.022)
Students' Sample Regressions	0.034	0.034	0.053	0.053
	(0.043)	(0.043)	(0.034)	(0.011)
C. Full Sample - By Parental Education (Student Level)				
Below Median Father's Years of Schooling	0.046	0.046	0.050	0.047
	(0.042)	(0.042)	(0.036)	(0.027)
Above Median Father's Years of Schooling	0.066	0.066	0.060	0.031
	(0.028)	(0.028)	(0.028)	(0.012)
Year FE	Yes	Yes	Yes	Yes
Subject FE	No	Yes	Yes	Yes
Schools' and Students' Characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes
Number of Students	39,915	39,915	39,852	10,190
Number of Schools	301	301	301	82
Number of CEOs	30	30	30	29

Notes: Standard errors in parenthesis are clustered at the CEO-by-year level. Schools characteristics include number of classes, number of students and the number of schools included in a CEO's cluster. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Sample includes stacked math, Hebrew and English tests scores. Test scores and CEOs Quality are standardized with 0 mean and unit sd.

Table 7. *The Effect of CEO's Quality on Tests Z-Scores - Alternative Estimation Strategies*

	Bayes Corrected Estimates		Two-Step Bootstrapped Standard Errors	
	Method II (1)	Method I (2)	Method II (3)	Method I (4)
A. Full Sample				
Students' Sample Regressions	0.050	0.038	0.051	0.040
	(0.034)	(0.017)	(0.018)	(0.015)
B. Female CEOs Sample				
Students' Sample Regressions	0.052	0.053	0.053	0.053
	(0.036)	(0.012)	(0.018)	(0.017)
C. Full Sample - By Parental Education (Student Level)				
Below Median Father's Years of Schooling	0.049	0.043	0.050	0.047
	(0.038)	(0.029)	(0.019)	(0.022)
Above Median Father's Years of Schooling	0.063	0.032	0.060	0.031
	(0.029)	(0.013)	(0.018)	(0.020)
Year FE	Yes	Yes	Yes	Yes
Subject FE	Yes	Yes	Yes	Yes
Schools' and Students' Characteristics	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes
Number of Students	39,852	10,190	39,852	10,190
Number of Schools	301	82	301	82
Number of CEOs	30	29	30	29

Notes: Standard errors in columns (1)-(2) are clustered at the CEO-by-year level. Standard errors in columns (3)-(4) are bootstrapped using the following two-step procedure: First, for each school in each period, we draw a random sample of students with replacement (In method I we draw a random sample of schools with replacement). The CEO quality regression is than estimated using this random sample. The second stage regression uses a random sample of students with replacement. This process is repeated a 1,000 times, and the bootstrap standard errors are computed using the sample of these 1,000 second-stage coefficient estimates. Schools characteristics include number of classes, number of students and the number of schools under the school's CEO. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Sample includes stacked math, Hebrew and English tests scores. Test scores and CEOs Quality are standardized with 0 mean and unit sd.

Table 8. *Non Linear Effects of CEO's Quality*

	Quality Method II	Quality Method I
	(1)	(2)
A. Median		
Above Median VA	0.082	0.065
	(0.036)	(0.036)
B. Quartiles		
2nd VA Quartile	0.080	0.031
	(0.051)	(0.038)
3rd VA Quartile	0.099	0.055
	(0.056)	(0.039)
4th VA Quartile	0.135	0.130
	(0.057)	(0.045)
Year FE	Yes	Yes
Subject FE	Yes	Yes
Schools' and Students' Characteristics	Yes	Yes
School FE	No	Yes
Number of Students	39,852	10,190
Number of Schools	301	82
Number of CEOs	30	29

Notes: Standard errors in parenthesis are clustered at the CEO-by-year level. Schools characteristics include number of classes, number of students and the number of schools included in the CEO's cluster. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Sample includes stacked math, Hebrew and English tests scores. Test scores and CEOs quality are standardized with 0 mean and unit sd.

Table 9. *The Effect of CEO's Quality on Working Procedures and School Resources*

	Mean	Specification		
	(1)	(2)	(3)	(4)
A. School's Priorities and Working Procedures (Rank 1-6)				
Clearly Defined Priorities	5.419	0.231	0.199	0.230
	(0.462)	(0.122)	(0.101)	(0.103)
Teachers Involved in Setting School Priorities	4.657	0.095	0.107	0.330
	(0.635)	(0.076)	(0.085)	(0.173)
Clearly Defined Working Procedures	5.461	0.179	0.149	0.102
	(0.494)	(0.140)	(0.122)	(0.121)
Average Effect	5.229	0.159	0.145	0.144
	(0.456)	(0.101)	(0.087)	(0.115)
B. School's Resources (Weekly Instruction Hours)				
Instruction Budget per Class	44.981	-0.048	-0.090	-0.015
	(6.364)	(0.131)	(0.055)	(0.055)
Length of the School Week	33.931	-0.057	-0.024	-0.082
	(3.062)	(0.124)	(0.061)	(0.047)
Instruction Hours of Math, Science and English	14.186	-0.012	-0.017	-0.038
	(1.937)	(0.135)	(0.098)	(0.088)
Average Effect	31.032	-0.039	-0.044	-0.045
	(3.397)	(0.125)	(0.057)	(0.046)
Year FE		Yes	Yes	Yes
Student and School Characteristics		No	Yes	Yes
School FE		No	No	Yes
Number of Schools	82	82	82	82
Number of CEOs	29	29	29	29

Notes: In column (1) standard deviations are presented in parenthesis. In columns (2)-(5) standard errors in parenthesis are adjusted for CEO level clustering. An average effect is an equally weighted average of the rest of its panel's questions' values in column (1), and z-scores in columns (2)-(4). Panel A is drawn from teachers' survey data and non-index variables are originally on a scale of 1-6. Panel B is drawn from administrative data on school funding. Schools characteristics include number of classes, number of students and the number of schools included in the CEO's cluster. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. CEOs Quality is standardized with 0 mean and unit sd. Outcomes and students' characteristics are school means. The dependent variable is a standardized transformation (mean zero, unit standard deviation) of the raw variable.

Table 10. *The Effect of CEO's Quality on School Environment and Climate*

	Mean	Specification		
	(1)	(2)	(3)	(4)
A. School Climate and Norms Indicators (0/1)				
School Evaluates its Environment, Climate and Discipline	0.707	0.267	0.264	0.250
	(0.456)	(0.070)	(0.093)	(0.074)
Interventions Related to School Environment and Discipline	0.291	0.097	0.335	0.493
	(0.456)	(0.075)	(0.099)	(0.224)
Average Effect	0.549	0.203	0.253	0.262
	(0.385)	(0.054)	(0.070)	(0.077)
B. Violence in School based on Students' Assessment (Rank 1-6)				
"There are many fights among students in my classroom"	3.593	-0.040	-0.042	-0.035
	(1.497)	(0.023)	(0.022)	(0.018)
"This year I was involved in many fights"	1.941	0.014	0.006	-0.034
	(1.389)	(0.021)	(0.016)	(0.016)
"Sometimes I'm scared to go to school because there are violent students"	1.938	-0.012	0.002	-0.024
	(1.483)	(0.019)	(0.013)	(0.012)
Average Effect	2.492	-0.013	-0.012	-0.033
	(1.031)	(0.014)	(0.010)	(0.012)
C. Students' Satisfcation (Rank 1-6)				
"There's someone in the teaching staff I can turn to"	5.091	0.012	0.007	0.036
	(1.324)	(0.017)	(0.016)	(0.008)
"I am staisfied in school"	5.252	0.003	0.010	0.011
	(1.147)	(0.019)	(0.017)	(0.017)
"I'm socially staisfied in school"	5.246	-0.001	-0.002	0.004
	(1.195)	(0.009)	(0.009)	(0.013)
Average Effect	5.193	0.005	0.006	0.019
	(0.926)	(0.017)	(0.016)	(0.014)
Year FE		Yes	Yes	Yes
Student and School Characterisics		No	Yes	Yes
School FE		No	No	Yes
Number of Schools	82	82	82	82
Number of CEOs	29	29	29	29

Notes: In column (1) standard deviations are presented in parenthesis. In columns (2)-(5) standard errors in parenthesis are adjusted for CEO level clustering. An average effect is an equally weighted average of the rest of its panel's questions' values in column (1), and z-scores in columns (2)-(4) Panel A is drawn from principals' survey data, outcomes and students' characteristics are school means and non-index variables originally take value of 0 or 1. Panels B and C are drawn from students' survey data, outcomes and students' characteristics are individual level data and non-index variables are on a scale of 1-6. Schools characteristics include number of classes, number of students and the number of schools included in the CEO's cluster. Student and school characteristics also include share of boys in the class in the violence related dependent variables. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. CEO Quality is standardized with 0 mean and unit sd. The dependent variable is a standardized transformation (mean zero, unit standard deviation) of the raw variable.

Table 11. *The Effect of CEO's Quality on School Scholastic Programs and Teachers' In and Out of School Training*

	Mean	Specification		
	(1)	(2)	(3)	(4)
A. School has an Academic Program (0/1) in:				
Hebrew	0.333	-0.123	0.047	0.274
	(0.473)	(0.104)	(0.129)	(0.294)
Math	0.265	-0.220	-0.063	-0.116
	(0.443)	(0.098)	(0.111)	(0.432)
English	0.222	-0.116	-0.116	0.472
	(0.418)	(0.100)	(0.147)	(0.297)
Average Effect	0.274	-0.153	-0.044	0.210
	(0.363)	(0.085)	(0.094)	(0.294)
B. Teachers Receive Out of School Training (0/1) in:				
Hebrew	0.465	0.030	0.087	0.043
	(0.201)	(0.078)	(0.092)	(0.149)
Math	0.372	0.096	0.096	0.021
	(0.191)	(0.062)	(0.092)	(0.106)
English	0.183	0.099	-0.022	0.035
	(0.259)	(0.150)	(0.113)	(0.083)
Average Effect	0.342	0.075	0.064	0.050
	(0.134)	(0.053)	(0.055)	(0.077)
C. Teachers Receive In-School Training (0/1) in:				
Hebrew	0.507	0.037	0.064	0.313
	(0.259)	(0.084)	(0.079)	(0.120)
Math	0.380	0.033	0.117	0.192
	(0.259)	(0.079)	(0.104)	(0.122)
English	0.188	-0.044	-0.172	-0.151
	(0.347)	(0.080)	(0.078)	(0.102)
Average Effect	0.358	-0.008	-0.005	0.024
	(0.162)	(0.014)	(0.016)	(0.020)
Year FE		Yes	Yes	Yes
Student and School Characteristics		No	Yes	Yes
School FE		No	No	Yes
Number of Schools	82	82	82	82
Number of CEOs	29	29	29	29
Number of Teachers	2,236	2,236	2,236	2,236

Notes: In column (1) standard deviations are presented in parenthesis. In columns (2)-(5) standard errors in parenthesis are adjusted for CEO level clustering. An average effect is an equally weighted average of the rest of its panel's questions' values in column (1), and z-scores in columns (2)-(4) The three panels are drawn from teachers' survey data and non-index variables originally take value of 0 or 1. Schools characteristics include number of classes, number of students and the number of schools included in the CEO's cluster. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. CEOs Quality is standardized with 0 mean and unit sd. Outcomes and students' characteristics are school means. The dependent variable is a standardized transformation (mean zero, unit standard deviation) of the raw variable.

Table 12. *The Effect of Change of CEO's and of CEO Quality on Change of School Principal*

A. The Effect of Change of CEO's on Change of School Principal

Years Since Change of CEO	Full Sample Period			2002-2003	2003-2004		2004-2005		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First Year	-0.044		-0.061	0.001	-0.049		-0.076		
	(0.021)		(0.028)	(0.033)	(0.055)		(0.032)		
Second Year		0.091	0.081			0.062		0.139	
		(0.028)	(0.028)			(0.035)		(0.050)	
Third Year									0.007
									(0.033)
Share of Schools Changing Principal	0.15	0.15	0.15	0.15	0.17	0.17	0.14	0.14	0.14
Number of Schools	738	721	721	723	721	714	716	708	701

B. The Effect of CEO's Quality on the Probability of Change of School Principal

	Share of Schools Changing Principal	CEO Quality Method I	CEO Quality Method I
	(1)	(2)	(3)
	0.333	0.067	0.093
	(0.474)	(0.030)	(0.037)
Year FE		Yes	Yes
Schools' and Students' Characteristics		No	Yes

Notes: In panel A, the sample includes schools that changed CEO once or less throughout the sample period (~90% of all schools). Standard errors are presented in parenthesis. All schools are in the Jewish non-religious primary school system. In column 4 the change of school principal is between 2002-2003. In columns 5-6 the change in school principal is between 2003-2004. In columns 7-9 the change in school principal is between 2004-2005. In panel B, the sample includes all schools that changed CEO at least once. Column 1 shows the share of schools which changed principals, in the sample period, after changing a CEO. Columns 2 and 3 show the outcomes of a school level regression of the probability of changing a school principal, after changing a CEO, on CEO quality. Schools characteristics include number of classes, number of students and the number of schools included in a CEO's cluster. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Sample includes stacked math, Hebrew and English tests scores. CEOs Quality is standardized with 0 mean and

Table 13. *The Effect of CEO's Quality on Tests Z-Scores- Above and Below Median for Each CEO*

	Quality Method II School Sample Regressions	
	Above Median Schools (1)	Below Median Schools (2)
Full Sample Main Score	0.058 (0.030)	
Median Variable - School Average Characteristics		
Test Scores Period 0	0.027 (0.033)	0.079 (0.034)
Test Scores 1991	0.012 (0.050)	0.083 (0.024)
Family Education Period 0	0.057 (0.034)	0.064 (0.036)
Socio Economic Decile of the School, as Determined by the Ministry of Education in 1998	0.047 (0.029)	0.074 (0.032)
Year FE	Yes	Yes
Subject FE	Yes	Yes
Schools' and Students' Characteristics	Yes	Yes
School FE	No	No
Number of CEOs	30	30

Notes: We divide our sample to above and below median schools, for each CEO, by each of the variables in the table. Then we regress CEO quality on test scores for each of the groups separately. Column (1) and (2) are the above median and below median sample of schools accordingly. Standard errors in parenthesis are clustered at the CEO-by-year level. Test scores in period 0 represent an average of the Math, Hebrew and English scores in the national standardized tests. Test scores in year 1991 represent an average of the Math and Hebrew scores from the 1991 standardized tests and family education represent the average number of years of total parental education for each school. Schools characteristics include number of classes, number of students and the number of schools included in a CEO's cluster. Students' characteristics include gender, ethnic origin, parents' education, number of siblings and immigrant status. All schools are in the Jewish non-religious primary school system. Number of schools for each column depends on the specific variable. Sample includes stacked math, Hebrew and English tests scores. Test scores and CEOs Quality are standardized with 0 mean and unit sd.