The Long Run Human Capital and Economic Consequences of High-Stakes Examinations*

Victor Lavy
University of Warwick, Hebrew University of Jerusalem and NBER

Avraham Ebenstein
Hebrew University of Jerusalem

Sefi Roth
Royal Holloway, University of London

March 2014

Abstract

Cognitive performance during high-stakes exams can be affected by random disturbances that, even if transitory, may have permanent consequences for long-term schooling attainment and labor market outcomes. We evaluate this hypothesis among Israeli high school students who took the high school high stakes matriculation exams between 2000 and 2002. As a source of random (transitory) shocks to high-stakes matriculation test scores we use exposure to ambient air pollution during the day of the exam. We find a significant and negative relationship between PM2.5 exposure during each of these examinations and a student's exam score, and between average exposure across the exams and a student's composite matriculation score. Using 2SLS with PM2.5 as our instrument for a student's composite score, we find that scores have a significant long-term effect on post-secondary educational attainment and earnings during adulthood. The results suggest that random disturbances during high-stakes examinations can have long-term consequences for schooling and labor force outcomes, and highlight the drawbacks of using high-stakes examinations in university admissions.

*Corresponding author: msvictor@huji.ac.il. Excellent research assistance was provided by Eyal Frank, Michael Freedman, Michal Hodor, Susan Schwartz, and Ben Raven. We thank seminar participants at Hebrew University, University of London Royal Holloway, University of Warwick, for useful comments and suggestions. We thank Israel's National Insurance Institute (NII) for allowing restricted access to data in the NII protected research lab. Research of the first author is supported by European Research Council (ERC) Advance Grant No. 323439.

I. Introduction

Cognitive performance is critical to scholastic achievement and to successful performance in many occupations. Cognitive acuity can be affected temporarily by a variety of factors, including the intake of caffeine, nicotine, and sleep deprivation (Jarvis 1993, Angus, Heslegrave and Myles 1985). Factors that induce variation in the cognitive performance of students during high-stakes exams may be of particular interest because they potentially affect long-term schooling attainment, and therefore can have permanent effects on wages and earnings. In this paper, we examine the potential long-term effect of transitory disturbances to cognitive performance during high-stakes exit exams in Israeli high schools. The exams are known as a the *Bagrut* and are a critical component of Israel's college admissions system, acting as a gatekeeper for the most selective schools, similar to the role played by Scholastic Aptitude Tests in the United States or A-levels in England and other European countries. Access to college majors is also determined by *Bagrut* performance, with many lucrative professional programs requiring minimum overall average scores for admission, such as law and medicine. As a consequence, *Bagrut* scores can affect an individual's entire academic career, and subsequent labor market outcomes.

Although many countries use high-stakes testing to rank students for college admission, the consequences of this policy are largely unknown. Does having a particularly good or bad performance on a high-stakes examination have long-term consequences for test takers, after accounting for a student's cognitive ability? Insofar as there are permanent wage consequences to variation induced by completely random shocks to student performance on a test like the *Bagrut*, it suggests that the use of high-stakes testing as a primary method for ranking students may be unfair. Aggregate welfare may also be reduced by relying too heavily on examinations that provide noisy measures of student quality, since it may lead to poor matching between students and occupations, and an inefficient allocation of labor. Recent debate over the planned redesign of the SATs has been in part motivated by concerns that the current version is highly random and does not represent a fair measure of student quality (New York Times 2014).[1] In spite of a dearth of evidence regarding the consequences of these tests, they are used extensively globally to rank students and allocate opportunity.

Assessing the consequences of using high-stakes examinations for ranking students is challenging. First, large data samples are generally not available with standardized test scores and wages during adulthood for a representative population.[2] Second, since higher-ability students presumably perform better on high-stakes tests, it is difficult to separately distinguish the return to cognitive ability from the return to doing well on the examination. One candidate solution is to examine the consequences of fluctuations in a random component affecting performance on these tests. Provided pollution has an effect on cognitive acuity and test scores, this generates plausibly random variation in a given student's outcome, creating a unique opportunity to assess whether these tests have long-term consequences. This enables direct measurement of the return to the component of a student's score which is related entirely to luck.

We use fluctuations in air quality during the day of a *Bagrut* exam as a source of random variation in cognitive performance and in exam outcomes. In Lavy, Ebenstein and Roth (2013), we identify and estimate the reduced form causal impact of various pollutants on cognitive performance. In this paper, we use fine particulate matter ($PM_{2.5}$) as an instrumental variable for *Bagrut* test scores because $PM_{2.5}$ is considered particularly harmful to human health and cognition, and it is monitored closely by the Israeli monitoring system. We observe the same student at multiple test administrations during each year of high school, allowing us to simultaneously control for time-invariant features of localities, schools and individual students. Overall, adding student fixed effects does not change the results very much in comparison to evidence based on school fixed effect regressions. In the latter, the identification assumption is that variation

[1] In a recent discussion of the planned revisions to the SATs, the president of the College Board stated that "only 20 percent…see the college-admission tests as a fair measure of the work their students have done." (New York Times 2014)

[2] Note that in the United States, Educational Testing Service (ETS) is notoriously private and no scholarship (to our knowledge) has been carried out linking SAT scores to adult outcomes for even small subsets of the population. For military recruits, the ASVAB has been made available but it is unclear how relevant this is for other sub-populations (Cawley et al. 2001).

across students from the same school in the dates of *Bagrut* exams is not correlated with potential outcomes. This is of course a very plausible assumption because dates of national *Bagrut* exams are determined by the Ministry of Education and students choose their *Bagrut* study program years before the dates of exams are determined. In support of this assumption, we provide balancing estimates that demonstrate that, conditioned on locality or school, pollution does not appear to be correlated with observable features of the student.

In our empirical work we examine a data set of merged high school high-stakes exit exams, pollution data for the universe of Israeli test takers during 2000-2002, and schooling attainment and earnings at adulthood. Several features of the context make it ideal for analysis. First, the rigorous nature of the *Bagrut* tests and the precise scoring of the exams provide a context to analyze a potential link between cognition and air pollution, even if there are only modest declines in cognitive performance due to pollution. Second, Israel's small size and well-developed pollution monitoring system implies that most of its *Bagrut* testing locations are near a station where we observe precise levels of pollution concentration. Third, Israel's national registration system allows us to carefully match the universe of students who take the *Bagrut*, their completed post-high school education, and their wages in adulthood. Therefore, Israel's *Bagrut* represents a unique opportunity to examine the consequences of high-stakes examinations on long run academic and economic outcomes in a context where we can evaluate both the relationship between air pollution and cognitive performance, and a more critical policy-relevant question regarding the consequences of these examinations for labor market outcomes.

In the first part of our analysis, we focus on random fluctuations in performance on high stakes exams by examining the first stage impact of fine particulate matter on exam outcomes. We find that a 10 unit increase in the ambient concentration of fine particulate matter ($PM_{2.5}$) as measured by the Air Quality Index (AQI) reduces *Bagrut* test scores by .46 points, or roughly 1.9% of a standard deviation of the *Bagrut* (sd=23.7). Alternatively, relative to a day with average air quality, a one standard deviation increase in the $PM_{2.5}$ AQI value (sd=22.81) is associated with a .65 point decrease in score, or 2.8% of a standard deviation. We also find that our results are largely driven by poor performance of test takers on extremely polluted days, with an AQI reading above 101 for $PM_{2.5}$ associated with a decline in test score of 1.95 points, or 8.2% of a standard deviation. These results suggest that modest pollution levels have only a marginal impact, but extremely polluted days can have much larger impacts, suggesting a non-linearity in pollution's relationship with cognitive performance.

In light of the 'first stage' relationship between *Bagrut* outcomes and exposure to air pollution, we then attempt to answer our main research question: what are the consequences of having a good or bad *Bagrut* score on the long run academic and economic achievement of a student, after controlling for student quality? We focus on a set of academic outcomes that should be particularly sensitive to the *Bagrut* score. Our primary academic outcomes are a student's eligibility for higher education (matriculation certification), a student's composite *Bagrut* score, a student's enrollment in post-secondary education, and a student's actual completed years of post-secondary education. We find clear evidence that completely random variation in *Bagrut* scores has an impact on the two 'aggregate' *Bagrut* outcomes: the overall composite score and the probability of receiving a matriculation certificate. Furthermore, exposure to pollution during the *Bagrut* continues to affect all post-secondary academic outcomes ten years later. We estimate that an additional 10 units of $PM_{2.5}$ (AQI) is associated with 0.03% percent decline in *Bagrut* matriculation certification, a .023% decline in a student's composite score, a .019% decline in enrollment in post-secondary education, and a 0.15 decline in years of education at a university. Furthermore, we find wage impacts of pollution exposure during *Bagrut* exams: having an additional 10 units of $PM_{2.5}$ is associated with a 109 Israeli shekel ($29) decline in monthly income, roughly a 2% decline in average income, suggesting non-trivial consequences for students who take the exam on a day with very severe air pollution.

We then exploit the strong first-stage relationship between pollution and *Bagrut* composite scores to estimate the economic return to each additional point on the *Bagrut*. Using 2SLS, we estimate that each point is worth (in 2010) 66 shekels ($18) in additional monthly earnings. Since the standard deviation on *Bagrut* composite scores is roughly 24 points, this implies that there are significant wage consequences to the exam, even for relatively small deviations in one's score. We examine mechanisms for this result by examining our other academic outcomes, treating the *Bagrut* composite score as the endogenous regressor and using pollution as our instrument. We find that each additional instrumented point increases the probability of receiving a matriculation certificate by 2 percentage points, enrollment rates in post-secondary academic schooling by 1.9 percentage points, and post-secondary education by .092 years. This suggests that the mechanism for the *Bagrut*'s impact on student outcomes is through the posited channel of affecting a student's prospects for post-secondary education.

Note that since the *Bagrut* composite score directly affects the post-secondary education options available to a student, 2SLS models of the return to post-secondary education using pollution as an instrument will be biased by the omission of the *Bagrut* composite score. However, in order to benchmark our results and assess how reasonable they are in terms of magnitude, we overlook this bias in order to generate 2SLS estimates of the implied return to education exploiting the first-stage relationship between $PM_{2.5}$ and post-secondary schooling. We estimate that an additional post-secondary year of schooling is worth 707 shekels ($191) per month, an implied return to college education of 14%. This rate is only marginally higher relative to existing estimates found in Israel and elsewhere for the return to post-secondary education (Frish 2009, Oreopoulos and Petronijevic 2013, Angrist and Chen 2011). This gap most likely reflects the independent effect of the average *Bagrut* score and of a matriculation diploma on earnings.[3] Our results are similar to existing estimates of the return to post-secondary education, , and represents supporting evidence that the magnitude of our estimated effects of $PM_{2.5}$ on schooling and economic outcomes are reasonable.

We then examine the heterogeneity of pollution's effect on scores and wages across sub-populations in Israel. We find that our first-stage relationship is strongest among populations of Israel which have the highest rates of asthma, suggesting a physiological mechanism for the observed relationship between pollution and exam outcomes. Our estimated effects are four times larger for boys than girls, twice as large for weaker students relative to stronger students, and a fifth larger for students from lower SES relative to students from high SES background. We then examine the impact of pollution on long run

[3] For example, Angrist and Lavy (2009) estimate that *Bagrut* holders earn 13 percent more than other individuals with exactly 12 years of schooling.

outcomes for these groups. We estimate that a point on the *Bagrut* is worth more for boys than girls (78 shekels versus 59 shekels), for stronger students than weaker students (124 shekels versus 80 shekels), and for higher SES students relative to lower SES students (105 shekels versus 56 shekels). These magnitudes suggest that the return to an extra point is very substantial, especially for already-strong students or students from privileged backgrounds, who presumably can capitalize on the opportunity of gaining admission to a longer academic programs or professions that require long (and poorly paid) internships, like law or medicine.

Our analysis highlights a major drawback of using high-stakes examinations to rank students. If completely random variation in scores can still matter ten years after a student completes high school, this suggests that placing too much weight on high-stakes exams like the *Bagrut* may not be consistent with meritocratic principles. A second implication of this finding is that by temporarily lowering the productivity of human capital, high pollution levels lead to allocative inefficiency as students with higher human capital may be assigned a lower rank than their less qualified peers. This may lead to inefficient allocation of workers across occupations, and possibly a less productive workforce. The results highlight the danger in assigning too much weight to a student's performance on a high-stakes exam, rather than their overall academic record. Our results also contribute to existing evidence that transitory random shocks that are unrelated to levels of human capital can have long-term implications for earnings (Oreopoulos et al. 2012). Our study also highlights the role of luck in determining wages, a pattern noted by other scholars (Bertrand and Mulinathan 2001).

The rest of the paper is laid out as follows. In the second section, we present relevant background information on air pollution and cognition, and on the controversial use of high-stakes examinations in college admissions both in Israel and abroad. In Section III, we present our empirical results. We conclude in Section IV.

II. Background and Data

A. Air pollution and Cognitive Performance

The existing literature provides compelling evidence that cognition may be affected by pollution, as a result of pollution's effect on the respiratory system. Researchers have documented that short-term acute exposure to particulate matter leads to increased risk of illness including heart disease, stroke, and lung cancer and increased hospitalization rates (Pope et al. 1995, Dockery and Pope 1996; Schlenker and Walker 2011, Chay and Greenstone 2003). Exposure to fine particulate matter is particularly dangerous since these small particles penetrate deep in to the lungs effecting blood flow and oxygen circulation, which may also affect other aspects of human life, and in particular, cognition (Pope and Dockery 2006). Since the brain consumes a large fraction of the oxygen needs of the body, any deterioration in oxygen quality can in theory affect cognition (Clark and Sokoloff 1999, Calderón-Garcidueñas et al. 2008). Air pollution can also impact the nervous system, leading to symptoms such as memory disturbance, fatigue and blurred vision (Kampa and Castanas 2007), and may also impact labor productivity (Graff Zivin and Neidell 2011). In a related paper (Lavy, Ebenstein and Roth, 2013), we explore this hypothesis in the Israeli context, examining the relationship between test scores and exposure to particulate matter, carbon monoxide, and other pollutants. In this paper, we focus our attention exclusively on particulate matter ($PM_{2.5}$): a complex mixture of solid and liquid microscopic droplets found in the air that consists of various components including acids, metals, dust particles, organic chemicals and allergens. This is a natural choice since it is arguably the most harmful pollutant for human health, and presumably most likely to affect cognition.

B. High-Stakes Examinations in Israel and Abroad

Since the Scholastic Aptitude Test's (SAT) first administration in 1926, it has been taken by millions of test-takers and has been used to rank students applying for college in the United States, and similar tests are used around the globe. However, the use of these tests is extremely controversial. Numerous concerns have been voiced by both popular and academic sources, including allegations of racial bias, arguments that test prep courses give privileged students an unfair advantage, and suggestions that the test places too much emotional stress on students.[4] Nevertheless, the SAT remains a critical component of college admissions in the US and, just as the *Bagrut* does in the Israeli educational system. The great weight placed on an exam given on one particular day has the benefit of being a cost-effective way of comparing students across schools with a similar metric, but may also represent a noisy measure of student quality. Many factors can affect student performance that are unrelated to cognitive ability, including how a student slept the previous night, whether the testing room has a comfortable temperature, and potentially, exposure to ambient air pollution. In light of the great weight placed on test scores in admissions processes at many elite schools, it is worth knowing whether (a) these scores are sensitive to random shocks and (b) whether bad draws have long-run consequences. Since this would be an extremely challenging question to address in the US, where SAT score data is fiercely guarded and generally not available for matching to adult outcomes, the Israeli *Bagrut* represents a novel opportunity to examine this question.

Passing all the Israeli *Bagrut* exams is required for receiving a matriculation certificate, a pre-requisite for university admission and one of the most economically important education milestones for Israeli students. Students complete the matriculation process by passing a series of national exams (*Bagrut* tests) in core and elective subjects following tenth and eleventh grade, and then complete a larger set of exams following twelfth grade. The exam focuses on seven mandatory subjects and one or more elective subjects, allowing us to observe students completing exams with separate grades for each examination and over a range of subjects. About 52 percent of high-school graduates in 2002 and 46 percent of the overall cohort received matriculation certificates, a figure that is publicized in the national media in light of the great significance of passing the *Bagrut* in the Israeli educational system

Students are admitted to university programs on the basis of their average *Bagrut* scores and a separate psychometric examination. Each university ranks applicants according to the same formula, thus producing an index based on a weighted average of the student's average score on all his or her *Bagrut*

[4] In 2001, the President of University of California famously threatened to remove the SAT requirement for admission, leading to a re-design of the examination and the introduction of a writing section. However, the writing section later came under fire for rewarding students simply for lengthy essays (Winerip 2005).

exams and the psychometric examination. This ranking determines students' eligibility for university admission, and even which major they can choose within the university. Therefore, pollution levels can affect students' university schooling by affecting their probability of passing *Bagrut* exams, and also by affecting the average score in these exams. The first channel will affect eligibility for university admission while the second will affect which programs (or majors) will be available to the student.

III. Evaluating the Consequences of High-Stakes Examinations

A. Data

Our data set is generated by combining Israeli test score data with air pollution and meteorological data for 2000-2002, and subsequent higher education and earnings outcomes from 2010-2011. The *Bagrut* exam information and demographic information for each test taker are provided by the Israeli Ministry of Education. These files also contain rich demographic information on the student and the student's family, such as parental education level, number of siblings, country of origin, and ethnicity. For each exam, we also know the exact date of the test and the precise location of the testing site, allowing us to assign pollution measures to each test administration. Our pollution data are taken from files published by the Israeli Ministry of Environmental Protection, which reports daily mean readings of particular matter less than 2.5 microns in width, or $PM_{2.5}$ ($\mu g/m^3$) at 139 monitoring stations throughout Israel for the sample period (see Figure 1). Readings are taken at 5 minute intervals and averaged over the course of the day. Each test site, which is at the student's school, is assigned the average pollution reading for all monitoring stations within 2.5 kilometers of the city limits of the school. Since Israeli cities are not very large, we generally are taking readings from stations very close to the schools. While we ideally would have a measure of pollution inside the test room, the air quality inside a test site is presumed to be highly correlated with the ambient reading outdoors (Branis et al. 2005). Schools that had no monitoring station within the city limits or 2.5 kilometers of the city limits were dropped from the sample.[5] These monitoring stations also record temperature and

[5] Since Israel's population is densely concentrated in several metropolitan areas, this led to the dropping of less than 5% of schools.

relative humidity, which are also assigned in a similar manner to pollution and are used as control variables. We use the daily average reading of pollution, temperature, and humidity at the monitoring stations in our analysis. The pollution measure is then converted into units of Air Quality Index (AQI) using a formula specified by the US EPA (see Table A1).

The post-secondary education system in Israel consist of eight universities that grant PhDs (as well as other degrees), approximately 50 academic colleges which offer undergraduate degrees (of which a very limited subset which offer masters degrees), and a set of non-university institutions of higher education that confer teaching and vocational certificates.[6] All universities and most academic and teachers' colleges require a *Bagrut* certificate for enrollment, while entry requirements for other post-secondary education establishments vary. For a given field of study, it is typically more difficult to be admitted to a university than to an academic college, since universities are both more prestigious and more heavily subsidized by the government.

Our information on post-secondary enrollment and earnings is taken from administrative records provided by the National Insurance Institute of Israel (NII). In order to facilitate the analysis presented here, the NII Research and Planning Division constructed an extract containing indicators of post-secondary enrollment, the number of years of post-secondary schooling, annual earnings, and number of months employed among all individuals in our study. This file was then merged with our original sample and analyzed at a secure research lab at the NII headquarters in Jerusalem, Israel. Our administrative sample includes information for each student on whether they have ever enrolled in higher education for each of the aforementioned institution types, and the number of years they were enrolled in higher education by institution type. This information is available for all students, and is recorded as of the 2010-2011 academic year. The youngest cohorts in our sample are already 28 years old at this time, implying that even after

[6] Practical engineering colleges run two- and three-year programs awarding degrees (or certificates) in fields like electronics, computers, and industrial production. An additional two years of study in an engineering school is required in order to complete a BSc in engineering.

accounting for compulsory military service, most students who enrolled in post-secondary education, including those who continued on to graduate school, will have graduated by 2010-2011.[7]

The summary statistics for our sample are presented in Table 1 in two panels; Panel A reports sample means of our exam-level, and Panel B reports sample means of our student-level data. The sample is composed of 415,219 examinations taken by 55,873 students at 712 schools throughout Israel. In columns (2) and (3) we stratify the sample by sex, and in columns (4) and (5), we stratify by a measure of achievement known as the *Magen* score. The *Magen* score is calculated using the student's performance over the course of the school-year, and on exam similar to the *Bagrut*, making it a natural candidate for stratifying the sample by student quality.[8] As shown in Table 1, for each *Bagrut* examination we observe the exam score, the pollution the day of the exam ($PM_{2.5}$), and the average temperature and humidity that day.[9] The table reveals that students face average pollution levels (AQI) that appear balanced along observables, with similar average readings among boys and girls (59.5 versus 59.9), and among higher/lower achievement students (60.0 versus 59.5). The sample means also reflect that girls perform better than boys, and student with higher *Magen* scores also have higher *Bagrut* scores.

In Panel B, we report our student-level means, which includes demographic information on the student, the education of both parents, and the student's earnings in 2010. Since our analysis of long-term economic outcomes will rely on school fixed effects, it is particularly important that we are able to include this rich set of control variables. The sample means also reveal several interesting patterns, including the higher achievement of girls: roughly 71% of girls receive a matriculation certificate, compared to only 64% of boys. Interestingly, however, girls earn lower earnings than their male counterparts. Boys earn on average 5,697 New Israeli Shekels (NIS) versus 4,848 for girls ($1≈ 3.75NIS). In columns (3) and (4), we observe higher rates of matriculation certification (91% versus 48%) and wages (5,506 versus 5,026) in the group of high achievement students, consistent with our expectations. Almost two thirds (63%) of a high school

[7] Boys serve for three years in the military and girls for two (longer if they take a commission).
[8] The date on which the *Magen* exam is given is unavailable, precluding a direct analysis of these scores.
[9] These variables are not shown due to space considerations but are used in regressions available in the appendix.

cohort are enrolled in post-secondary studies, 27% in universities, and 25% in academic colleges. The sample also reveals that we are able to match the entire universe of student test takers with their long-term outcomes, a particularly desirable feature of our data relative to panel data sets that face attrition.

B. Empirical Strategy

Our estimation strategy is relatively straightforward. We estimate OLS models where we examine the partial correlation between pollution and test scores, and 2SLS models where we exploit the first-stage relationship between pollution and test scores to estimate the economic return to success on the *Bagrut*. For identification, we rely on the panel structure of the data and the repeated nature of the *Bagrut* exam. Since we observe the exact location of the test, we can include city or school fixed effects. Since we observe the students taking exams following each grade, we can include student fixed effects. Formally, the models we estimate are of the following form:

$$(1)\ R_{ist} = \beta_0 X_{it} + \beta_1 POL_{st} + \beta_2 Temp_{st} + \beta_3 RH_{st} + M_t + L_l + I_i + \varepsilon_{ist}$$

where R_{ist} is the test score of student *i* at school *s* at time *t*; X_{it} is a vector of individual characteristics possibly related to test outcomes, such as parental education[10]; POL_{st} is our measure of air pollution ($PM_{2.5}$) at school *s* at time *t*; $Temp_{st}$ is the mean temperature[11] at school *s* at time *t*; RH_{st} is the relative humidity measure at school *s* at time *t*; M_t and L_l are month and exam proficiency level fixed effects respectively; I_i is our fixed effect for the individual; and e_{ist} is an idiosyncratic error term. Our analysis of the long-term consequences of pollution exposure is estimated in a similar manner, but since the analysis is carried out at the student-level, we only include school fixed effects.

[10] Our results with individual fixed effects exclude individual controls.

[11] In the empirical analysis, we include linear and quadratic terms in both temperature and humidity, and linear and quadratic interaction terms of the two variables. We also experimented with higher order polynomial of these variables and the results remained unchanged.

The key identifying assumption for inferring a causal relationship between pollution and test scores estimated by equation (1), and the use of pollution as a valid instrument for a student's score, is that unobserved determinants of student's test scores are uncorrelated with ambient pollution. Without any fixed effects to absorb unobserved variation in schools or individuals, this assumption is likely violated since it is likely that pollution is correlated with time invariant features of a testing location or a particular student. For example, if poorer schools are located in more polluted parts of cities, OLS will likely overstate the causal link between pollution and test scores. Conversely, if schools in denser (and wealthier) cities have more pollution exposure, OLS might understate the true cost of pollution, as it is mitigated by other compensating factors (e.g. tutoring). More generally, endogenous sorting across schools, heterogeneity in avoidance behavior, or measurement error in assigning pollution exposure to individuals will all bias results that do not properly account for unobserved factors correlated with both our outcome of interest and ambient pollution (Moretti and Neidell 2011). In our setup, since we account for time-invariant features of schools and students with fixed effects, the challenge relevant to our estimation is to account for omitted variables that are varying over time but are potentially correlated with pollution and *Bagrut* scores. For example, if weather or traffic the day of the exam is correlated with pollution, our fixed effects models will fail to identify the true effect. In our empirical analysis, we include controls for time-varying factors that could be contemporaneous with pollution, such as daily temperature and relative humidity, but of course it is untestable whether there are factors that are unobserved that are both correlated with pollution and *Bagrut* exam scores. As such, we conduct a rich set of robustness checks and placebo tests which are available in the Appendix.

C. First Stage Estimates: Pollution Effects on High Stakes Exam Outcomes

In Table 2, we report our baseline results of the relationship between the AQI values for $PM_{2.5}$ and *Bagrut* test scores. In columns (1) and (2) of Panel A, we report the correlation between *Bagrut* scores and a continuous measure of $PM_{2.5}$ (AQI) using OLS without city, school or student fixed effects. In column (1), we estimate that a 1 unit increase of $PM_{2.5}$ is associated with a 0.06 points decrease in a student's test score and the estimate is nearly unchanged by adding controls for parental education, gender, temperature, relative humidity and dummies for the month of the exam and difficulty of the exam (column 2). In columns (3)-(5), we take advantage of the panel structure of our data and include city, school, and student fixed effects, respectively. These account for variation in time-invariant unobserved heterogeneity that could be correlated with ambient pollution. The estimates from a regression with city or school fixed effects in columns (3) and (4), are somewhat larger, with estimated coefficients of -0.082 and -0.069 respectively. Adding student fixed effects weakens the results slightly, with our preferred estimate indicating that a 1 unit increase in $PM_{2.5}$ is associated with a 0.046 (sd=0.007) decline in the *Bagrut* score. This estimate implies that a test score on an exam on a day with average pollution (AQI=59.74) will be lower relative to an exam taken on a day with the minimum pollution level (AQI=10.1) by 0.10 (.046*(59.7-10.1)/22.8) standard deviations. In Panel B, we perform a similar analysis but replace our continuous measure of pollution with a dichotomous indicator for whether the test occurred on a day classified as having "moderate" or "poor" air quality, defined as AQI in the range 51-100 or above 101 respectively. The results are qualitatively similar to the results using the continuous measure for $PM_{2.5}$ and indicate that the relationship is of significant consequence for test takers.[12]

As a frame of reference for evaluating the magnitude of our measured effects, it is worth noting that the effect of $PM_{2.5}$ on *Bagrut* scores for the 99th percentile of exposure in our sample (AQI=137) is very large and implies a decline of roughly a sixth (.149) of a standard deviation relative to an average exposure day. This effect is similar to the estimated effect of reducing class size from 31 to 25 students (Angrist and Lavy 1999) and larger than the test scores gains associated with paying teachers large financial bonuses

[12] In Table A2, we exploit the fact that we know the exact time of day that the examination was administered, and consider whether our pollutants have different effects at different times of day. While the majority of our sample is given a 9AM examination time, roughly 40% of examinations are given after 12PM. We posit that fine particulate matter, which is generated from sandstorms and coal-burning plants, will affect students throughout the day in a similar manner at all hours of the day (or night). As shown in the table, our coefficient estimates are relatively similar for both afternoon and morning examinations. In our preferred student fixed effect specification, we find that having poor air quality for an afternoon exam is associated with a .045 point and .054 point decline per unit of AQI respectively. Likewise, our results using the dichotomous measure are similar; we observe a 3.16 point decline in the student's *Bagrut* score for days with very high AQI in afternoon exams, which is about 20% larger than the coefficient for morning exams.

based on their students' test scores (Lavy 2009). Unfortunately, days with elevated levels of particulate matter are not unusual in Israel and in neighboring countries in the Middle East, as they are often the result of sand storms that originate in the Sahara desert and are relatively common in the spring and summer months (Bell et al. 2008).

D. Placebo Tests

In this section, we perform a set of placebo tests where we examine the relationship between air pollution on days *other than* the actual exam and exam scores. In Table 3, we examine whether there is a correlation between pollution from the day of the previous *Bagrut* and the score on the following exam. Note that since students take the *Bagrut* exams over a short period of time, this will generally be a pollution reading taken within several days of the exam. As shown in Panel A of Table 3, the correlation between *Bagrut* outcomes is weak relative to the correlation with the actual exam. While some of the specifications are statistically significant, our preferred specification with student fixed effects are all statistically insignificant. For example, in our estimates using our threshold measure with student fixed effects, the impact of $PM_{2.5}$ during the previous exam is a .78 point *increase* in the student's score, and the result is not statistically significant. This can be compared to our main result using the $PM_{2.5}$ reading from the day of the *Bagrut*, where poor air quality reduces scores by 1.9 points (significant at the 1% level).

In Panel B, we perform a similar exercise but use the air pollution reading on the date exactly one year before the exam. For the continuous measure of pollution, column 5 indicates a negative and statistically insignificant relationship between $PM_{2.5}$ and test scores. For our dichotomous measure of pollution, we observe a correlation between exams and $PM_{2.5}$ in the previous year when we include no fixed effects: having a day classified as polluted in the previous year is associated with a 2.8 point decline in scores in models with controls, even though there should be no relationship. This underscores the importance of including fixed effects to absorb a time-invariant correlation between pollution and student quality, and suggests that more polluted areas have lower exam scores in general. Once we include student fixed effects in our models, the correlation between $PM_{2.5}$ from the previous year and the *Bagrut* score declines to 1.15 points, and it is not statistically different from zero (at the 5% level).

In Figure 2, we examine the impact of $PM_{2.5}$ three days prior to the exam, the day of the exam, and three days following the exam on test scores. As shown in the figure, the main effects of $PM_{2.5}$ are concentrated on the day of the exam, and no significant relationship between pollution readings and the exam score is observed for days before and after the exam. The figure indicates that the coefficient on pollution *the day of* the exam is much larger and more negative than the other days: an additional 100 units of AQI is associated with a 0.2 point decrease in student scores, and the coefficient estimates are small and positive on the days before and after the exam.

E. Pollution and Long-Term Schooling Attainment and Economic Outcomes

In this section, we examine the relationship between $PM_{2.5}$ exposure during the series of *Bagrut* exams and (1) two aggregate measures of academic achievements in the *Bagrut* exams program and (2) two measures of post-secondary schooling attainment. In these models using data at the student-level, student fixed effects are of course infeasible and we rely primarily on models where we control for observables at the individual level and school fixed effects. Identification is based on the assumption that variation across students from the same school in the dates of *Bagrut* exams is not correlated with potential outcomes. This is of course a very plausible assumption because dates of national *Bagrut* exams are determined by the Ministry of Education and students choose their *Bagrut* study program years before the dates of exams are determined. In support of this assumption, we note the results of balancing exercises presented in Table A3, where we demonstrate that, conditioned on city or school, pollution does not appear to be correlated with observable features of the student. Note that many of the estimated balancing coefficients estimates are very small, practically zero, and not significantly different from zero. We also note that in earlier results presented in Table 2 the estimated effect of both pollution measures were not systematically sensitive to adding student fixed effects.

In Table 4, we examine the relationship between $PM_{2.5}$ and our main set of academic outcomes. In the first two rows, we report the outcomes most directly related to performance on the *Bagrut*: receiving a matriculation certificate, and the *Bagrut* composite score. These are formed from a combination of a student's test scores, and have significant relevance for later study. Receiving a matriculation certificate is of paramount importance, since this is a pre-requisite for post-secondary college and university studies. However, the *Bagrut* composite score can also influence the range of programs accessible to a student. As highlighted in the table, both exhibit a robust negative relationship with air quality. An additional 10 units of $PM_{2.5}$ (AQI) is associated with a 1.64 reduction in a student's composite score and a .033% decline in certification in our preferred estimation strategy which include school fixed effects.

In rows 3 and 4, we examine the relationship between pollution exposure and two longitudinal educational outcomes: enrollment in post-secondary institution (1=yes), and years of post-secondary schooling attained. The results demonstrate that these academic outcomes are affected by pollution exposure. In particular, enrollment rates in higher education decline by .030% when a student is exposed to an additional 10 units of $PM_{2.5}$, schooling declines by .015 years. All estimates are statistically significant at the 5% level, and suggest that taking *Bagrut* exams in highly polluted days can have long-lasting effects on completed schooling attainment.

In the fifth row, we present the reduced form effect of $PM_{2.5}$ on average monthly earnings. In our preferred specification with school fixed effects in column 3, we estimate that a student is exposed during the *Bagrut* exams dates to an additional 10 units of $PM_{2.5}$ is associated with an average monthly earnings decline at age 28 of 109 shekels ($29). This estimate is also precisely estimated, with a T statistic greater than three.

In Table 5, we exploit the first-stage relationship between pollution and the *Bagrut* Composite score to examine the long-term consequences of the examination in a 2SLS context. In Panel A, we estimate the return to an additional point on the *Bagrut* composite score. In the first row, we reproduce the relationship between the *Bagrut* composite score and $PM_{2.5}$ shown in Table 4 that is used here as our first-stage. Exploiting the relationship between scores and pollution, we find using 2SLS that an additional point is worth an estimated 45 shekels in monthly earnings in models with city fixed effects, and is worth 66 shekels in our preferred specification including school fixed effects. Since the standard deviation of the *Bagrut* composite score is roughly 24 points, these estimates imply that even small deviations from a student's "average" score can have significant consequences on adult income.

In Panel B, we use the first-stage relationship between pollution and the *Bagrut* composite score to examine the mechanisms underlying the strong relationship between scores and earnings. Since the *Bagrut* composite score is an important factor in gaining admission into lucrative courses of study, it is logical to examine whether the instrumented score is correlated with subsequent outcomes. As shown in Panel B, we find that each additional instrumented point increases the probability of receiving a matriculation certificate by 2 percentage points, enrollment rates in post-secondary schooling by 1.9 percentage points, and post-secondary educational attainment by .092 years. This indicates that an additional *Bagrut* point has a tremendous economic value to a student, and even random variation in scores can have important consequences for a student's future attainment of post-secondary schooling.

In Panel C, we exploit the relationship between pollution exposure and the *Bagrut* composite score to estimate the return to an additional year of post-secondary schooling. It is worth noting that this strategy does identify 'cleanly' the rate of return to schooling since the *Bagrut* score can directly affect earnings, and therefore its omission implies a violation of the exclusion restriction. However, as way of benchmarking our results and checking whether our returns estimates are reasonable, we wish to compute the return to education and compare our estimates to those found in the existing literature. Treating post-secondary schooling as the endogenous regressor and $PM_{2.5}$ as the instrument, we estimate using 2SLS that each additional year of post-secondary schooling is worth 707 ($191) shekels. Relative to the average monthly earnings of 5,084 shekels ($1,374) in our sample of young adults, this means that fortunate students who have less pollution on their *Bagrut* exams dates will have long-term economic benefits to their good luck. This estimate implies a rate of return to college education of 14%, which is very similar to existing estimates found elsewhere for the return to post-secondary education. This is somewhat higher in comparison with recent estimates in Israel and elsewhere. For example, Angrist and Chen exploit variation in veteran status and the GI Bill to estimate a return to education of roughly 9% (Angrist and Chen 2011).

While our IV approach will not be valid if pollution exposure on exam dates leads to a permanent diminution of intellectual ability, in light of the highly temporary nature of pollution in Israel, such as sandstorms, we think it is unlikely that our effects are picking up such an effect. In Table A4, we present evidence that exposure to pollution in our data is only related to temporary variation in performance. We find that a student's average pollution exposure during the *Bagrut* period (May-July) in 11th grade has no correlation with his or her average *Bagrut* scores on exams taken in May-July of 12th grade, suggesting that the effects we estimate are capturing the consequences of short-term random shocks, rather than reflecting long-term cognitive diminution.

We examine possible mechanisms for our results by examining how pollution affects the probability of a student matriculating at different types of post-secondary institutions. If our results are operating through a mechanism in which the *Bagrut* is a gatekeeper to lucrative occupations, we should find that our results are driven by large estimated effects for universities, and milder effects for academic colleges. In fact, it may be that for students who attend technical schools, there is no financial value to passing the *Bagrut*, insofar as they purse a profession of a technical nature. This could similarly be true for students planning to be small business managers, which is common in Israel, especially among the Israeli-Arab population, who generally have more limited access to lucrative professions.[13] As reported in Table A5, this is indeed the case, with our effects significant and negative only for the probability of attending a university. In fact, interestingly, the impact of pollution is *positive* (though imprecisely measured) for the less competitive programs, such as teacher's colleges and semi-engineering programs, possibly due to students being shifted out of universities or academic colleges and into these less selective programs.

[13] Willis and Rosen (1979) find that, in a sample of World War II veterans, comparative advantage dictates whether people sort into higher education. This is consistent with our findings, which indicate that there is almost no marginal value of academic achievement for the lower ability students.

F. Heterogeneity

In this section, we examine heterogeneity in the relationship between random variation in the average *Bagrut* score and long-term schooling and economic outcomes. We stratify our data by comparing three groups: boys and girls, academically stronger and weaker students, and students from high and low socioeconomic background. We begin with an analysis of the first stage relationship between pollution and test scores in the three sub-samples. Our motivation for this stratification is twofold. First, we have a prior that asthma rates are higher among boys and among those of lower SES. Insofar as we observe larger effects for these groups, it sheds light on a possible mechanism for our finding, and implies that our result is being driven partly by a physiological mechanism. Second, in terms of fairness, it is worth considering how these exams affect different students. Since these exams are often the gatekeeper for prized occupations in Israel, it is worth investigating how different students are able to capitalize on these forms of achievement.

i. Heterogeneity in the First Stage Relationship

In this section, we examine whether there is heterogeneity in the first stage relationship between scores and pollution by building on a set of stylized facts regarding which groups would be most sensitive to poor air quality, taken from an extensive medical literature. First, Israeli boys are more likely to be asthmatic than Israeli girls (Laor et al. 1993). Second, children of lower economic status are known to have higher rates of asthma and respiratory illnesses (Eriksson et al. 2006, Basagana et al. 2004). Third, Laor et al. (1993) also found that *Ashkenazic* Jews (ethnic origin from America and Europe) have 63% higher incidence of these illnesses than *Sephardim* (ethnic origin from Africa and Asia). This gives a rich set of potential comparisons for gauging whether asthma (or other respiratory illnesses) is a mechanism for the observed reduced form relationship between pollution and exam outcomes.

We examine our results separately by sex, student quality, and student background (SES) in Table 6. Our models are estimated in the same manner as those reported in Table 2, with demographic controls and city or school or student fixed effects included in each specification. We also separately estimate treatment effects using both our continuous and dichotomous (AQI≥101) measure of pollution. In Panel A, we estimate that boys are between 2 and 4 times more sensitive to pollution than girls. We posit that the difference could be partly generated by the different asthma rates in these cohorts. Another possibility is that male students are more likely to be affected by small cognitive decline and distraction, consistent with higher rates of Attention Deficit Disorder in males (Biederman et al. 2002). As shown in Panel B, a similar pattern is found when comparing between academically weaker and stronger students. When we stratify the students by whether their *Magen* score is above or below the median, our estimated treatment effects for $PM_{2.5}$ are more than two times larger among those classified as low achievers. For this group we estimate that a 10 unit increase in $PM_{2.5}$ is associated with a .061 point decline versus only a .028 point impact among higher quality students with larger effects among the former than the latter in Panel B. The results also indicate a similarity between results estimated with school and student fixed effects, which is reassuring since our results on earnings are estimated at the student level and rely on school fixed effects. Lastly, in Panel C, we stratify the sample by SES. It may be that poorer families are more affected by air pollution as well, due to lower ability to engage in compensating behavior (Neidell 2004). Poorer children also have higher incidence of asthma (Basagana et al. 2004, Eriksson et al. 2006). Since we have no direct measure of income for the student's parents, we proxy SES with the student's father's education, with high SES being students with a father above the median value in our sample We again find evidence in support of an interpretation that the relationship between pollution and exam performance is driven by a physiological mechanism.

In Table A6, we examine the relationship between air pollution exposure and the probability of failing a *Bagrut* exam for each of these sub-populations. The results indicate that boys are more sensitive to $PM_{2.5}$ than girls, lower quality students are more sensitive than stronger students, and students from lower SES backgrounds are more sensitive than those from higher SES backgrounds. In particular, raising the fraction of days with very polluted air by 10 percentage points is associated with a .57 percentage point increase for boys in the chance of failing a particular *Bagrut* in models with student fixed effects. Girls appear largely unaffected, with the increased chance of not passing being statistically indistinguishable from zero. The gap is even more striking for student with low *Magen* scores: a 10 percentage point increase in the fraction of days with very polluted air is associated with a .59 percentage point increase in failure probability. Similar results are found when students are stratified by parental SES, with very polluted days increasing the probability of failure among those of low SES and having an insignificant effect among higher SES children.[14]

ii. Heterogeneity in the Long Run Consequences of the *Bagrut*

In this section, we examine the long term schooling and economic consequences of the *Bagrut* for different sub-populations in Israel. We repeat the group breakdowns used earlier, stratifying students by sex, quality, and socio-economic background. In columns 1 and 2 of Panel A, we estimate the return to an additional point on the *Bagrut* using 2SLS, where the *Bagrut* composite score is treated as the endogenous regressor and $PM_{2.5}$ is the instrument. Our results by student sex are reported in columns 1 and 2, and indicate that the return to an additional point is roughly 60% higher for boys than girls: 78 shekels vs 59 shekels ($21 vs $16). One explanation is that women choose less lucrative fields of study than men, even when they have similar qualifications, as a result of gender attitudes in Israel. It is also worth noting that although female labor force participation rates are relatively similar to the US, Israeli women have much higher fertility than their American counterparts.[15] This may lead Israeli women to choose less lucrative professions than men and often work part time, which would be reflected in a lower payoff per additional year of higher education. In our context, this is plausible, since many Israelis work in government jobs which are lower-wage, offer more flexible work schedules, and have generous maternity leave policies. A

[14] In Lavy, Ebenstein and Roth (2013), we exploit the unique ethnic heterogeneity of the Jewish population in Israel to estimate models for sub-populations. The primary distinction is between Sephardic Jews of Middle Eastern and North African origin, and Ashkenazic Jews who are from Eastern Europe and Russia. We find that the impact of air pollution is larger among Ashkenazic Jews relative to Sephardic Jews. Since Ashkenazic Jews has lower rates of asthma and respiratory conditions (Laor et al. 1993), the results are consistent with our prior that $PM_{2.5}$ which affects the respiratory system will have a larger impact on weaker groups who are more sensitive to poor air quality.

[15]Average fertility rates in Israel are 3.0, roughly 50% higher than the US rate of 2.0 (World Bank, 2010). However, employment rates are relatively similar. Among women 25-45, the employment rates among Israeli men and women were 80% and 61% respectively (Israel 1995 census), compared to rates in the US of 86% and 69% (US 2010 census).

second explanation could be discrimination against women in the labor market, resulting in a lower payoff to an additional year of schooling.

In columns 3 and 4, we find larger returns to a point among higher achievement students. Specifically, stronger students experience a 124 shekel return to each point, compared to only a 80 shekel return among lower quality students ($34 vs $22). We offer two explanations for this finding. First, it may be related to the instrument we are using; insofar as our estimate is a local average treatment effect where the disturbance to a student's true potential is relatively small, the estimated return to an additional point on the *Bagrut* will be larger among those who could participate in lucrative occupations. For weaker students, pollution is not affecting their already-low chance of being accepted into a very lucrative profession. A second explanation is that there are heterogeneous returns to different types of higher education. So, for example, the return to an additional year of studying economics might be different than the return to philosophy. Since Israeli majors have different standards for admission, with humanities having lower standards generally, our estimates may be picking up the differences between the return to different majors, which provide an avenue to different occupations. Unfortunately, our data do not contain information on area of study, precluding further examination into this hypothesis.

In columns 5 and 6, we observe very large differences between students of high and low SES. The return to an additional point is 105 shekels ($28) among high SES, and roughly half that amount for low SES (56 shekels or $15). Interestingly, this gap is even larger than when we had previously split the sample by *Magen* score in columns 1 and 2. This suggests that coming from a wealthier background raises the return to education significantly, and in a more dramatic way than even stratifying by student quality. One possible explanation is that parental income enables students to undertake longer and more costly academic paths, but results in them landing ultimately in more lucrative positions. Having a non-binding funding constraint could be a partial explanation for the higher return to higher education. Another explanation is that credentials and connections are complements, so students with greater social capital *and* qualifications can capitalize on their qualifications more than students from less privileged background.

In Panel B of Table 7, we examine the mechanisms for the aforementioned results by estimating 2SLS models where we instrument for a point with pollution and treat the *Bagrut* composite score as our endogenous regressor. We repeat our earlier analysis performed on the overall sample, and examine 3 channels for *Bagrut* scores to influence long run economic outcomes: through affecting the probability of receiving a matriculation certificate, enrollment rates in post-secondary institutions, and through its effect on total completed post-secondary education. Interestingly, we find that girls, weaker students, and lower SES background students are *more* affected by each additional instrumented point on the *Bagrut* than boys, stronger students, or higher SES background students. We interpret this as evidence that the stakes of each point is higher for students with lower labor-force attachment or less economic advantage – not necessarily economically in terms of the consequences for wages, but in terms of their likelihood of pursuing post-secondary education.

For example, if male students are committed to participating in a lucrative profession, they will perhaps be less dissuaded by a poor score; in Israel, students can attend a lower-ranked school in their field of study of choice, if they are denied admission in the top programs in Tel Aviv or Jerusalem. Likewise, stronger students and students from high SES background may proceed with their intended course of study in spite of a bad score. Greater commitment to the labor force or greater access financial resources may result in matriculation at campuses in other cities, such as Be'er Sheva, where students generally cannot live at home and need to pay for dormitories but often have lower *Bagrut score* thresholds for admission. Having sufficient economic resources could also facilitate a student retaking the exam in a subsequent year, lowering the stakes of a single bad outcome for students from privileged backgrounds.[16] Our analysis of mechanisms highlights the disruptive effect of a poor *Bagrut* outcome on female and weaker/low SES

[16] Vigdor and Clotfelter found this to be important in the US context, where students from wealthier backgrounds are more likely to retake the SAT (2003). However, since Israeli high-school graduates immediately begin a period of military service (3 years for boys and 2 years for girls), retaking the exam is only possible with a long delay. Therefore, retaking the exam is relatively uncommon but could explain in part the weaker relationship between instrumented *Bagrut* scores and post-secondary education for students from high SES.

students, and is further evidence that reliance on the *Bagrut* has questionable properties in terms of supporting a meritocratic environment.

In Table 8, we examine heterogeneity in the estimated return to education by sub-population in Israel using 2SLS, with pollution again as our instrument. Note that since pollution affects scores as well, this will not satisfy the exclusion restriction, but is worth exploring nonetheless to assess the economic magnitude of our estimated effects. In Panel A, where we stratify the sample by sex, we estimate that an additional 10 units of $PM_{2.5}$ reduces male and female post-secondary schooling by .17 and .14 years respectively. We estimate that the return to an additional year of schooling is 888 shekels and 564 shekels respectively ($240 versus $152), suggesting that male students are more able to capitalize on post-secondary education, possibly due to the choice of more lucrative majors and professions, discrimination in the labor force, or due to their stronger labor-force attachment. We also find that stronger students are able to capitalize more from higher education: the wage return to post-secondary schooling is nearly twice as high among stronger students, with each year increasing wages by 1,131 shekels per month for strong students and only by 698 for weaker students. This pattern is even more extreme when we consider students stratified by SES: an additional year is worth 1,264 shekels to a student of high SES background, more than twice the return to low SES students (580 shekels). Similar to our discussion of the return to a point on the *Bagrut*, this highlights the interplay between achievement and status: the results indicate that the return to post-secondary education is largest among those most able to leverage this achievement, highlighting an additional avenue by which high stakes examinations can affect the wage distribution and wage inequality.

VII. Conclusion

This paper has examined the relationship between transitory shocks to performance in high stakes exams and their long term consequences for determining college schooling attainment and earnings. We exploited variation in ambient air pollution during high-stakes examinations as an instrumental variable, and demonstrated that pollution affects student test scores on very important high-school exit exams, which in turn affects post-secondary schooling and long run earnings. Our analysis consisted of two parts. First, using a large sample of Israeli high-school *Bagrut* examinations (2000-2002), we presented evidence that there is a robust negative relationship between academic outcomes and random fluctuations in ambient pollution concentrations. We also find that among Israeli sub-populations with higher rates of asthma and respiratory illnesses, our estimated treatment effects for $PM_{2.5}$ are larger, suggesting that physiological impairment is a potential mechanism for our findings. Second, using matched data between students and their adult earnings, we find that ambient pollution exposure during the *Bagrut* has long-term consequences; students exposed to high levels of pollution during the *Bagrut* are less likely to receive a matriculation certificate, have fewer years of post-secondary schooling, and have lower earnings when they are observed 8-10 years after high school graduation.

We argue that this is evidence that placing so much weight on *Bagrut* scores may reduce welfare of some students and cause allocative inefficiencies. As a result this will negatively affect the overall economy, as the mis-ranking of students due to variability in pollution exposure could result in bad assignment of workers to different occupations and lead to reduced labor productivity.[17] Our findings also suggest that the exams may be generating random variation in people's opportunities, consistent with recent concerns voiced by officials in the US regarding the reliance of the SATs for college admissions (Lewin 2014).

[17] As an implication of these findings, it may also be that the gains from improving air quality may be underestimated by a narrow focus on health impacts. Insofar as air pollution may lead to reduced cognitive performance, the consequences of pollution may be relevant for a variety of everyday activities that require mental acuity. Traffic accidents, injuries in the workplace, and reduced worker productivity may all be the byproduct of reduced cognitive performance. As such, the results presented here highlight a channel by which the consequences of pollution are vastly understated by a narrow focus on the immediate and acute health consequences, and suggest that improvements in air quality may yield tremendous benefits in welfare.

References

Angrist, J.D., and S.H. Chen. 2011. "Schooling and the Vietnam-Era GI Bill: Evidence from the Draft Lottery." *American Economic Journal: Applied Economics*, 3(2): 96-118.

Angrist, J. and Lavy, V. 1999. “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement.” *Quarterly Journal of Economics*, 114(2), 533–575.

Angus, R. and R. Heslegrave, and W. Stewart Myles. 1985. “Effects of Prolonged Sleep Deprivation, With and Without Chronic Physical Exercise, on Mood and Performance.” *Psychophysiology*, 22(3):276-282.

Basagaña, X., S. Jordi, K. Manolis, Zock, J., Duran-Tauleria, Enric, J.D., Burney, P. and Josep Maria Anto. 2004. “Socioeconomic status and asthma prevalence in young adults: the European Community Respiratory Health Survey.” *American Journal of Epidemiology*, 160: 178–188.

Bell, M., JK and Lin Z. 2008. “The effect of sandstorms and air pollution on cause-specific hospital admissions in Taipei, Taiwan.” *Occupational and Environmental Medicine*, 65: 104-111.

Bertrand, M., and S. Mullainathan. 2001. “Are CEO’s Rewarded for Luck? The Ones Without Principals Are.” *Quarterly Journal of Economics* 116(3):901-932.

Biederman, J.M., E. Faraone, S. Braaten, E. Doyle, A. Spencer, T. Wilens, T. Frazier, E. and M.A. Johnson. 2002. "Influence of Gender on Attention Deficit Hyperactivity Disorder in Children Referred to a Psychiatric Clinic." *The American Journal of Psychiatry*, 159: 36–42.

Braniš, M., Ř. Pavla, and M. Domasová. 2005. "The effect of outdoor air and indoor human activity on mass concentrations of PM10, PM2.5, and PM1 in a classroom." *Environmental Research*, 99: 143–149.

Calderón-Garcidueñas, L., Mora-Tiscareño, A., Ontiveros, E., Gómez-Garza, G., Barragán-Mejía, G., Broadway, J., Chapman, S., Valencia-Salazar, G., Jewells, V., Maronpot, RR., Henríquez-Roldán, C., Pérez-Guillé, B., Torres-Jardón, R., Herrit, L., Brooks, D., Osnaya-Brizuela, N., Monroy, M., González-Maciel, A., Reynoso-Robles, R., Villarreal-Calderon, R., Solt. A., and R. Engle. 2008. "Air pollution, cognitive deficits and brain abnormalities: a pilot study with children and dogs." *Brain and Cognition*, 68(2): 117-127.

Cawley, J, J. Heckman and E. Vytlacil. 2001. “Three observations on wages and cognitive ability.” *Labour Economics* 8:419-442.

Chay, Kenneth Y and Michael Greenstone. 2003. “The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession.” *Quarterly Journal of Economics* 118(3):1121-116

Clark, D. and L. Sokoloff. 1999. "Circulation and energy metabolism of the brain," in G. Siegel, B. Agranoff, R. Albers, S. Fisher and M. Uhler, eds., Basic Neurochemistry. Molecular, Cellular and Medical Aspects, Lippincott-Raven, pp. 637–670.

Dockery, W., and C. Pope. 1996. “Epidemiology of Acute Health Effects: Summary of Time Series Studies.” in R. Wilson and J. Spengler, eds., Particles in Our Air, Harvard University Press, 123-148.

Eriksson K., Hedlund U., and E. Ronmark. 2006. "Socio-economic status is related to incidence of asthma and respiratory symptoms in adults." *European Respiratory Journal*, 28(2): 303-310.

Frisch, R. 2009. “The Economic Returns to Schooling in Israel”. *Israel Economic Review* (7)1:113–141.

Graff Z., J. S. and M. J. Neidell. 2011. "The Impact of Pollution on Worker Productivity." NBER Working Paper #17004, forthcoming, American Economic Reviervis, M. 1993. “Does caffeine intake enhance absolute levels of cognitive performance?” Psychopharmacology 110: 45-52.

Lewin, Tamar. “A New SAT Aims to Realign With Schoolwork”. *The New York Times* 5 March 2014; accessed at: http://www.nytimes.com/2014/03/06/education/major-changes-in-sat-announced-by-college-board.html?_r=0. 10 March 2014.

Kampa, M., and E. Castanas. 2007. "Human health effects of air pollution." *Environmental Pollution*, 151: 362-367.

Laor, A., Cohen, L., and Y Danon. 1993. "Effects of time, sex, ethnic origin, and area of residence on prevalence of asthma in Israeli adolescents." *British Medical Journal*, 307(6908): 841-844.

Lavy, V. 2009. “Performance Pay and Teachers’ Effort, Productivity and Grading Ethics”, *American Economic Review*, 99(5): 1979-2011.

Lavy, V., A. Ebenstein and S. Roth, “The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance”, Mimeo 2013.

Moretti, E. and M. Neidell. 2011. "Pollution, Health, and Avoidance Behavior Evidence from the Ports of Los Angeles." *The Journal of Human Resources*, 46(1): 154-175.

Neidell, M. 2004. "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *Journal of Health Economics*, 23: 1209–1236.

Oreopoulos, P., von Wachter, T. and Heisz, A. 2012 “The Short-and Long-Term Career Effects of Graduating in a Recession.” *American Economic Journal: Applied Economics*, 4(1): 1-29.

Oreopoulos, P. and U. Petronijevic, 2013. "Making College Worth It: A Review of Research on the Returns to Higher Education," NBER Working Papers 19053.

Pope, C. III, Bates, D., and M. Raizenne. 1995. "Health Effects of Particulate Air Pollution: Time for Reassessment?" *Environmental Health Perspectives*, 103: 472-480.

Pope, C. III, and D. Dockery. 2006. "Critical Review–Health effects of fine particulate air pollution: Lines that connect." *Journal of the Air and Waste Management Association*, 56: 709-742.

Vigdor, Jacob L. and C. Clotfelter. 2003. “Retaking the SAT”. *Journal of Human Resources* 38:1-33.

Schlenker, W. and W. R. Walker. 2011. "Airports, Air Pollution, and Contemporaneous Health." NBER Working Paper #17684.

Winerip, M. 2005. "SAT Essay Test Rewards Length and Ignores Errors". New York Times, May 5th.
Willis, R. and S. Rosen. 1979. “Education and Self-selection.” *Journal of Political Economy* 87(5):S7-S36.

Table 1

Summary Statistics: Particulate Matter Exposure and Israeli *Bagrut* Scores

Variable	All (1)	By Sex: Boys (2)	By Sex: Girls (3)	By *Magen* Score (Course Grade[1]): Low Scores (4)	By *Magen* Score (Course Grade[1]): High Scores (5)
Panel A: Exam-Level Data					
Pollution Measures					
$PM_{2.5}$	21.05	20.89	21.18	21.15	20.96
($\mu g/m^3$)	(10.86)	(10.57)	(11.10)	(10.88)	(10.87)
$PM_{2.5}$	59.74	59.47	59.98	60.01	59.51
(AQI Index)	(22.81)	(22.50)	(23.08)	(22.89)	(22.75)
$PM_{2.5}$	0.05	0.05	0.05	0.05	0.05
(AQI ≥101)	(0.21)	(0.21)	(0.22)	(0.22)	(0.21)
Examination Outcomes					
Bagrut Exam Score	70.76	68.91	72.33	53.22	77.10
(1-100 points)	(23.74)	(24.86)	(22.64)	(30.69)	(22.18)
Failed a *Bagrut* Exam	0.19	0.21	0.17	0.33	0.04
(1=yes)	(0.39)	(0.41)	(0.37)	(0.47)	(0.19)
Magen Score	75.45	73.27	77.30	64.09	86.93
(1-100 points)	(21.37)	(22.50)	(20.19)	(23.25)	(10.47)
Climate Controls					
Temperature	23.81	23.81	23.82	23.84	23.83
(celsius)	(2.61)	(2.61)	(2.62)	(2.66)	(2.50)
Relative Humidity	50.90	50.86	50.94	50.98	50.95
(percent saturation)	(14.71)	(14.52)	(14.87)	(15.08)	(14.35)
Observations	415,219	190,410	224,809	206,571	204,527
Panel B: Student-Level Data					
Demographic Information					
Mother's Education	11.44	11.60	11.30	10.79	12.08
(years)	(5.04)	(5.09)	(5.00)	(4.87)	(5.13)
Father's Education	11.62	11.83	11.44	10.85	12.39
(years)	(5.03)	(5.02)	(5.03)	(4.84)	(5.10)
Number of Siblings	2.02	1.95	2.07	2.03	2.00
	(1.58)	(1.49)	(1.65)	(1.61)	(1.55)
Bagrut Outcomes and Matriculation Certification Rates					
Bagrut Composite Score	70.76	68.91	72.33	53.22	77.10
	(23.74)	(24.86)	(22.64)	(30.69)	(22.18)

Matriculation Certification Rate	0.68	0.64	0.71	0.48	0.91
	(0.47)	(0.48)	(0.45)	(0.50)	(0.28)
Post-Secondary Enrollment Rates					
Any Post-Secondary	0.631	0.602	0.656	0.475	0.821
University	0.274	0.258	0.289	0.115	0.469
Academic Colleges	0.248	0.253	0.244	0.244	0.253
Teacher & Semi-eng.	0.070	0.063	0.076	0.078	0.059
Other[2]	0.046	0.036	0.055	0.046	0.047
Post-Secondary Schooling in Years					
Any Post-Secondary	2.25	2.05	2.42	1.45	3.23
	(2.15)	(2.10)	(2.18)	(1.86)	(2.08)
University	1.03	0.95	1.10	0.35	1.85
	(1.90)	(1.83)	(1.95)	(1.13)	(2.28)
Academic Colleges	0.83	0.80	0.85	0.73	0.95
	(1.47)	(1.44)	(1.50)	(1.38)	(1.57)
Teachers & Semi-engineering	0.26	0.21	0.31	0.25	0.27
	(0.87)	(0.68)	(1.00)	(0.82)	(0.92)
Adult Earnings					
Monthly Wages ([3]NIS 2010)	5,084	5,531	4,699	4,867	5,352
	(4,515)	(5,198)	(3,788)	(4,053)	(5,013)
Observations	55,796	26,158	29,638	30,668	25,128

Notes : Standard deviations are in parentheses. In Panel A, each observation represents a *Bagrut* exam. The measure of pollution is particulate matter smaller than 2.5 microns, or $PM_{2.5}$. The AQI value for each reading is calculated from a formula that converts micrograms (µg/m3) into a 1-500 index value. Relative humidity is the amount of moisture in the air as a share of what the air can hold at that temperature. [1] low and high subsamples were based on being above or below the median of a student's average *Magen* score over all subjects. In Panel B, each observation represents a student. The pollution and climate controls are averages over a student's exams. Receiving a matriculation certificate is determined by a combination of the student's average *Bagrut* and *Magen* score. [2]The other programs include technical schools, non-academic colleges, and smaller schools. [3]Wages are repoted in monthly New Israeli Shekels ($1=3.6 NIS) and are taken for 2010 from the students who took Bagrut examinations between 2000 and 2002. The schooling and wage outcomes were made available by the Israeli National Insurance Institute (Bituach Leumi). Each student's record contains whether they matriculated at a post-secondary institution, and the number of years they were enrolled at the institution.

Table 2

Pooled OLS and Fixed Effect Models of Particulate Matter's Impact on *Bagrut* Test Scores

	Pooled OLS		Fixed Effects		
	No Controls (1)	Controls (2)	City (3)	School (4)	Student (5)
Panel A: Air Quality Index (continuous measure)					
$PM_{2.5}$	-0.06	-0.065	-0.082	-0.069	-0.046
(AQI Index)	(0.015)	(0.011)	(0.008)	(0.007)	(0.007)
Female (1=yes)		3.22	3.30	2.72	
		(0.34)	(0.34)	(0.22)	
Mother's Education		0.165	0.141	0.112	
		(0.063)	(0.062)	(0.034)	
Father's Education		0.410	0.396	0.241	
		(0.061)	(0.058)	(0.033)	
R-squared	0.003	0.042	0.046	0.145	0.493
Observations	415,219	380,435	380,435	380,435	380,435
Panel B: Air Quality Index above Threshold Value					
$PM_{2.5}$	-3.00	-2.63	-2.75	-2.68	-1.95
(AQI ≥101)	(1.54)	(1.03)	(0.84)	(0.70)	(0.74)
Female (1=yes)		3.19	3.25	2.68	
		(0.340)	(0.337)	(0.219)	
Mother's Education		0.158	0.143	0.111	
		(0.064)	(0.063)	(0.035)	
Father's Education		0.409	0.396	0.241	
		(0.061)	(0.058)	(0.033)	
R-squared	0.001	0.040	0.043	0.143	0.492
Observations	415,219	380,435	380,435	380,435	380,435

Notes : All regressions include suppressed controls for temperature and humidity on the exam date, which are included as linear and quadratic terms in each, and linear and quadratic interaction terms of the two variables. The coefficients in Panel A are reported per 100 units of AQI. Standard errors are clustered at the school level, are heteroskedastic-consistent, and are reported below the coefficients in parentheses.

Table 3

Placebo Tests Measuring the Relationship between the *Bagrut* Score and Particulate Matter Exposure on Irrelevant Days

	Pooled OLS		Fixed Effects		
	No Controls (1)	Controls (2)	City (3)	School (4)	Student (5)
Panel A: Pollutant Level from Previous Exam					
$PM_{2.5}$	-0.02	-0.035	-0.049	-0.034	-0.005
(AQI Index)	(0.013)	(0.010)	(0.007)	(0.006)	(0.006)
$PM_{2.5}$	-1.10	-0.48	-0.66	-0.29	0.78
(AQI ≥101)	(1.37)	(0.87)	(0.78)	(0.68)	(0.71)
Panel B: Pollutant Level from Previous Year					
$PM_{2.5}$	-0.008	-0.033	-0.027	-0.014	-0.006
(AQI Index)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)
$PM_{2.5}$	-2.78	-2.89	-1.03	-0.75	-1.15
(AQI ≥101)	(0.81)	(0.68)	(0.76)	(0.73)	(0.69)

Notes : Each cell in the table represents a separate regression. The regressions are estimated in the same manner as those reported in Table 2, but actual pollution on the day of the examination is replaced by the pollution recorded during the student's previous exam (one to several days prior) or the pollutant level from the previous year. The coefficients are reported per 100 units of AQI. Standard errors are clustered at the school level, are heteroskedastic-consistent, and are reported below the coefficients in parentheses.

Table 4

Particulate Matter, Post-Secondary Education and Adult Earnings

	Pooled OLS	Fixed Effects	
	Controls (1)	City (2)	School (3)
Bagrut Composite Score	-6.66 (0.84)	-26.57 (1.28)	-16.38 (1.85)
Matriculation Certification	-0.234 (0.021)	-0.534 (0.031)	-0.335 (0.047)
Enrolled in Post Secondary Institution (1=yes)	-0.093 (0.021)	-0.497 (0.030)	-0.307 (0.039)
Completed Years of Post-secondary Education	-0.671 (0.087)	-2.361 (0.126)	-1.518 (0.181)
Average Monthly Earnings (NIS)	-1,548 (326)	-1,199 (331)	-1,093 (344)

Notes: Each cell in the table represents a separate regression. The table reports the relationship between average $PM_{2.5}$(AQI) during the *Bagrut* and the listed outcome, estimated using the student-level sample described in Table 1. All regressions include suppressed controls for average temperature and humidity during the *Bagrut*, mother's and father's years of schooling, sex, and age in 2010. The coefficients are reported per 100 units of AQI. Standard errors are clustered at the school level, are heteroskedastic-consistent, and are reported below the coefficients in parentheses.

Table 5

The Economic and Academic Return to the *Bagrut* Composite Score

	Pooled OLS	Fixed Effects	
	Controls	City	School
	(1)	(2)	(3)
Panel A: Effect of the Bagrut Composite Score on Adult Earnings using $PM_{2.5}$ (AQI) as an IV			
First Stage	-6.66	-26.57	-16.38
	(0.84)	(1.27)	(1.85)
Reduced Form	-1,528	-1,203	-1,073
	(324)	(327)	(341)
2SLS	229	45	66
	(147)	(13)	(21)
Panel B: Effect of the Bagrut Composite Score on Follow Up Academic Outcomes using $PM_{2.5}$ (AQI) as an IV			
Matriculation Certification	0.034	0.020	0.020
	(0.011)	(0.002)	(0.002)
Enrolled in Post Secondary Institution (1=yes)	0.016	0.019	0.019
	(0.006)	(0.002)	(0.002)
Completed Years of Post-secondary Education	0.105	0.089	0.092
	(0.026)	(0.063)	(0.009)
Panel C: Estimated Return to Post-Secondary Education using $PM_{2.5}$ (AQI) as an IV			
First Stage	-0.67	-2.36	-1.52
	(0.09)	(0.13)	(0.18)
Reduced Form	-1548	-1199	-1093
	(326)	(331)	(344)
2SLS	2278	509	707
	(1343)	(139)	(219)

Notes: Each cell in the table represents a separate regression. The regressions are estimated in the same manner as those reported in Table 4. In Panel A, we present 2SLS models of the relationship between the *Bagrut* Composite Score and Adult Earnings using $PM_{2.5}$(AQI) as an instrumental variable. In Panel B, we present 2SLS models of the relationship between *Bagrut* Composite Score and other academic outcomes, using $PM_{2.5}$(AQI) as an instrumental variable. In Panel C, we estimate the implied return to post-secondary schooling using $PM_{2.5}$(AQI) as the instrument. Standard errors are clustered at the school level, are heteroskedastic-consistent, and are reported below the coefficients in parentheses.

Table 6

Heterogeneity in the Impact of $PM_{2.5}$ on *Bagrut* Test Scores

	$PM_{2.5}$ (AQI Index)		$PM_{2.5}$ (AQI >=101)	
	School	Student	School	Student
	(1)	(2)	(3)	(4)
Panel A: By Sex				
Boys	-0.10	-0.08	-5.33	-4.10
	(0.01)	(0.01)	(0.82)	(0.87)
Girls	-0.04	-0.02	-0.66	-0.38
	(0.01)	(0.01)	(0.80)	(0.83)
Panel B: By Student Quality				
Low Achievement Students	-0.08	-0.06	-3.86	-3.49
	(0.01)	(0.01)	(1.04)	(1.10)
High Achievement Students	-0.03	-0.03	-0.93	-0.76
	(0.01)	(0.01)	(0.57)	(0.68)
Panel C: By Socio-Economic Status (SES)				
Low SES	-0.07	-0.05	-2.76	-2.07
	(0.01)	(0.01)	(0.74)	(0.77)
High SES	-0.06	-0.04	-2.40	-1.66
	(0.01)	(0.01)	(0.86)	(0.93)

Notes: Each cell in the table represents a separate regression. The regressions are estimated in the same manners as those reported in Table 2. Student quality is determined by whether the student's average *Magen* score was above or below the median. High SES is composed of children whose father was above the median level of education in the sample. The column title reports whether fixed effects are included at the school or student level. Coefficients are reported per 100 units of AQI. Standard errors are heteroskedastic-consistent and are reported below the coefficients in parentheses.

Table 7

Heterogeneity in the Economic and Academic Return to the *Bagrut* Composite Score

	By Sex		By Student Quality		By Socio-Economic Status	
	Boys	Girls	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of the Bagrut Composite Score on Adults Earnings using PM $_{2.5}$ (AQI) as an IV						
First Stage	-19.72	-13.69	-6.61	-12.15	-13.22	-21.10
	(2.65)	(1.94)	(1.84)	(1.65)	(1.98)	(2.52)
Reduced Form	-1,540	-810	-530	-1,510	-747	-2,205
	(495)	(421)	(352)	(622)	(288)	(734)
2SLS	78	59	80	124	56	105
	(27)	(30)	(58)	(51)	(24)	(32)
Panel B: Effect of the Bagrut Composite Score on Follow up Academic Outcomes						
Matriculation Certification	0.017	0.025	0.032	0.011	0.023	0.015
	(0.002)	(0.003)	(0.008)	(0.004)	(0.003)	(0.002)
Enrolled in Post Secondary Institution (1=yes)	0.018	0.020	0.029	0.019	0.021	0.013
	(0.002)	(0.003)	(0.008)	(0.004)	(0.003)	(0.002)
Completed Years of Post-secondary Education	0.087	0.108	0.108	0.118	0.096	0.085
	(0.010)	(0.014)	(0.030)	(0.015)	(0.013)	(0.010)

Notes : Each cell in the table represents a separate regression. The regressions are estimated in the same manner as those reported in Table 4. Student quality is determined by whether the student's average Magen score was above or below the median. High SES is defined as children whose father was above the median level of education.

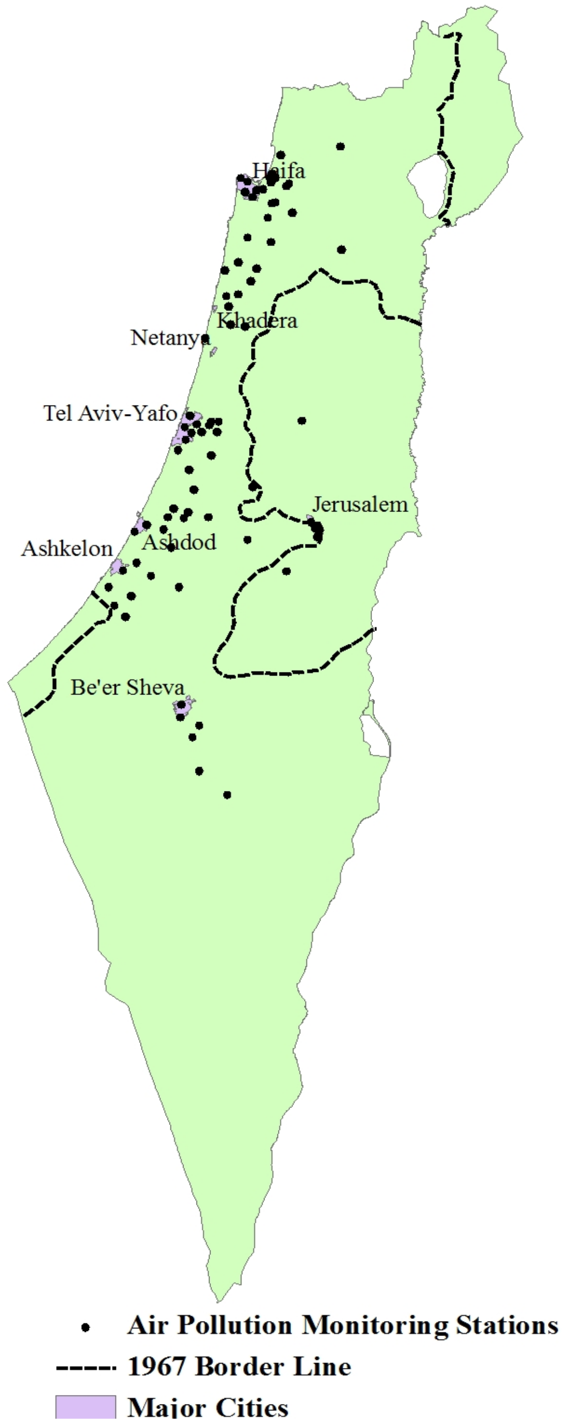
Table 8

Heterogeneity in the Estimated Return to Post-Secondary Education

	First stage (1)	Reduced Form (2)	2SLS (3)
Panel A: By Sex			
Boys	-1.73	-1,559	888
	(0.218)	(498)	(296)
Girls	-1.44	-839	564
	(0.227)	(428)	(265)
Panel B: By Student Quality			
Low Achievement Students	-0.76	-521	698
	(0.155)	(355)	(484)
High Achievement Students	-1.33	-1,568	1,131
	(0.270)	(617)	(454)
Panel C: By Socio-Economic Status (SES)			
Low SES	-1.29	-743	580
	(0.192)	(288)	(235)
High SES	-1.74	-2,251	1,264
	(0.286)	(747)	(405)

Notes: Each cell in the table represents a separate regression. Standard errors are clustered by school. The results here are estimated in the same manner as those reported in Table 3 with school fixed effects. Each estimate is in terms of a single year of additional post-secondary schooling, and the instrument is $PM_{2.5}$(AQI). Student quality is determined by whether the student's average *Magen* score was above or below the median. High SES is defined as children whose father was above the median level of education. Standard errors are heteroskedastic-consistent and are reported below the coefficients in parentheses.

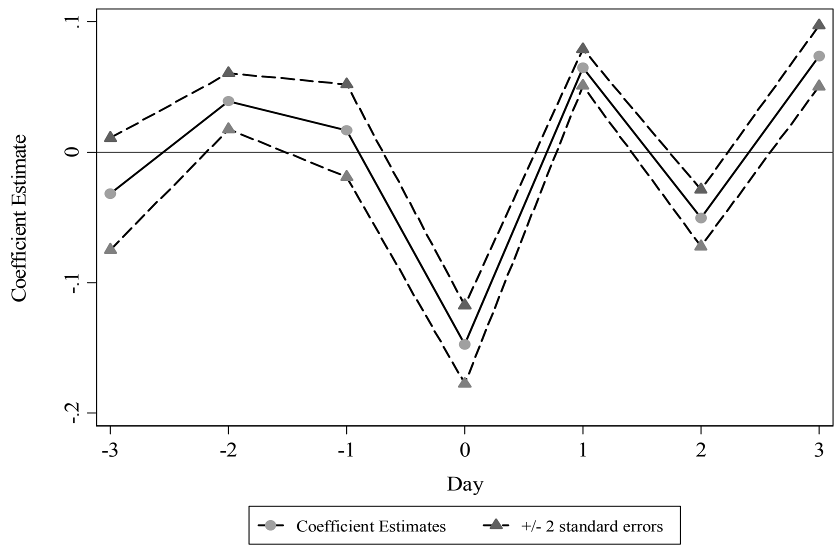
Figure 1
Locations of Major Cities and Air Quality Monitoring Stations in Israel



Notes : The boundaries of Israel are reported in the plot, with the main cities shaded in.

Figure 2

Impact of $PM_{2.5}$ on Test Scores in the Days Pre and Post Examination



Notes : The figure plots the coefficients from a regression of *Bagrut* test scores on $PM_{2.5}$ AQI readings in the days prior to, the day of (Day=0), and the days following the examination. Standard errors are clustered by school.

Table A1

Breakpoints for $PM_{2.5}$ ($\mu g/m^3$) and AQI Index Categories

$PM_{2.5}$ (μg/m3)	AQI Index Value	Category
0.0 - 15.4	0 - 50	Good
15.5 -40.4	51 - 100	Moderate
40.5 - 65.4	101 - 150	Unhealthy for Sensitive Groups
65.5 - 150.4	151 - 200	Unhealthy
150.5 - 250.4	201 - 300	Very unhealthy
250.5 - 350.4	301 - 400	Hazardous
350.5 - 500.4	401 - 500	Hazardous

Note : Standards are determined by the United States Environmental Protection Agency.

Table A2

Pooled OLS and Fixed Effect Models of Pollutant Matter on Afternoon and Morning Test Scores

	Pooled OLS		Fixed Effects		
	No Controls (1)	Controls (2)	City (3)	School (4)	Student (5)
Panel A: Afternoon Examinations					
$PM_{2.5}$(AQI)	-0.019	-0.079	-0.116	-0.082	-0.045
	(0.017)	(0.015)	(0.013)	(0.010)	(0.013)
$PM_{2.5}$(AQI $\geq$101)	-2.94	-3.11	-3.04	-2.56	-3.16
	(2.11)	(2.01)	(1.95)	(1.35)	(1.42)
Observations	162,912	148,026	148,026	148,026	148,026
Panel B: Morning Examinations					
$PM_{2.5}$(AQI)	-0.074	-0.067	-0.074	-0.066	-0.054
	(0.016)	(0.013)	(0.009)	(0.008)	(0.010)
$PM_{2.5}$(AQI $\geq$101)	-2.93	-3.13	-2.97	-3.13	-2.46
	(1.38)	(1.11)	(0.87)	(0.73)	(0.97)
Observations	252,307	232,409	232,409	232,409	232,409

Notes : Each cell in the table represents a separate regression. The regressions are estimated in the same manners as those presented in Table 2. The examinations that are given at 12PM or later are classified as afternoon exams.

Table A3

Balancing Tests: Assessing the Relationship between Students' Characteristics and Pollution

Variable	Pooled OLS (1)	School Fixed Effects (2)
Female (1=yes)	0.00	0.10
	(0.00)	(0.00)
Father's Education	0.10	0.40
	(1.00)	(0.50)
Mother's Education	0.30	-0.10
	(1.00)	(0.60)
Number of Siblings	0.60	0.30
	(0.30)	(0.10)
Ashkenazi (1=yes)	0.00	0.00
	(0.00)	(0.00)
Sephardi (1=yes)	0.00	0.00
	(0.00)	(0.00)
Father Born in Israel (1=yes)	0.00	0.00
	(0.00)	(0.00)
Observations	54,294	54,294

Notes : Each cell in the table represents a separate regression, where the dependent variable is $PM_{2.5}$(AQI) and the independent variable is the coviate listed in the row. The regressions are estimated in the same manners as those presented in Table 2.

Table A4

Relationship Between Particulate Matter Exposure During Previous Exams and Average *Bagrut* Scores at Conclusion of 12th Grade

	Pooled OLS		Fixed Effects	
	No controls	Controls	City	School
	(1)	(2)	(3)	(4)
Panel A: All Students				
	-0.80	0.90	-0.40	1.70
	(2.90)	(2.80)	(3.50)	(2.10)
Panel B: By Sex				
Boys	-0.90	0.30	-2.40	-0.70
	(3.40)	(3.50)	(4.50)	(2.80)
Girls	-1.20	1.30	0.90	4.00
	(2.80)	(3.00)	(3.60)	(2.40)
Panel C: By Student Quality				
Low Achievement Students	2.60	3.30	0.20	2.60
	(2.50)	(2.50)	(3.30)	(2.30)
High Achievement Students	1.30	1.40	1.10	2.30
	(1.10)	(1.10)	(1.60)	(1.30)
Panel D: By Socio-Economic Status (SES)				
Low SES	-2.10	0.80	1.00	1.30
	(2.90)	(3.00)	(3.50)	(2.30)
High SES	1.10	0.10	-1.30	2.20
	(3.00)	(2.80)	(4.10)	(2.80)

Notes: Each cell in the table represents a separate regression. The regressions are estimated in the same manners as those presented in Table 2. Student quality is determined by whether the student's average *Magen* score was above or below the median. High SES is defined as children whose father was above the median level of education. Standard errors are heteroskedastic-consistent and are reported below the coefficients in parentheses.

Table A5

Particulate Matter's Impact on Post-Secondary Schooling by Type

	LHS: Enrolled in Post-Secondary Institution (1=yes)		LHS: Completed Years of Post-Secondary Education	
	City (1)	School (2)	City (3)	School (4)
All Post-Secondary Institutions	-0.495 (0.069)	-0.298 (0.040)	-2.353 (0.306)	-1.527 (0.182)
Universities	-0.537 (0.070)	-0.369 (0.044)	-2.225 (0.292)	-1.568 (0.184)
Academic Colleges	-0.086 (0.042)	0.019 (0.031)	-0.411 (0.130)	-0.036 (0.102)
Teacher and Semi-engineering	0.039 (0.030)	0.008 (0.019)	0.119 (0.070)	0.058 (0.049)

Notes : Each cell in the table represents a separate regression. In each regression, the dependent variable is either enrollment (columns 1 and 2) or years of schooling (columns 3 and 4) at the listed academic type. The depdent variable is the average $PM_{2.5}$ (AQI) exposure during the student's *Bagrut* examinations. The regressions are estimated with the same controls as those presented in Table 4, and the coefficients are reported per 100 units of $PM_{2.5}$ (AQI). The column title reports whether fixed effects are included at the city or school level. Standard errors are heteroskedastic-consistent and are reported below the coefficients in parentheses.

Table A6

Particulate Matter's Impact on Failing a Bagrut Exam
by Sex and Student Quality

	$PM_{2.5}$ (AQI Index)		$PM_{2.5}$ (AQI >=101)	
	School	Student	School	Student
	(1)	(2)	(3)	(4)
Panel A: All Students				
	0.018	-0.017	0.036	0.024
	(0.009)	(0.010)	(0.010)	(0.010)
Panel B: By Sex				
Boys	0.064	0.021	0.077	0.057
	(0.013)	(0.014)	(0.015)	(0.015)
Girls	-0.017	-0.046	0.004	-0.001
	(0.010)	(0.011)	(0.010)	(0.010)
Panel C: By Student Quality				
Low Achievement Students	0.026	0.001	0.066	0.059
	(0.015)	(0.017)	(0.016)	(0.016)
High Achievement Students	-0.035	-0.041	0.000	-0.002
	(0.005)	(0.007)	(0.006)	(0.007)
Panel D: By Socio-Economic Status (SES)				
Low SES	0.026	-0.011	0.042	0.031
	(0.011)	(0.012)	(0.011)	(0.011)
High SES	0.002	-0.027	0.027	0.014
	(0.008)	(0.010)	(0.012)	(0.012)

Notes: Each cell in the table represents a separate regression. The regressions are estimated in the same manners as those presented in Table 2. Student quality is determined by whether the student's average Magen score was above or below the median.High SES is defined as children whose father was above the median level of education. Standard errors are heteroskedastic-consistent and are reported below the coefficients in parentheses.