Getting Lucky: The Long-Term Consequences of Exam Luck

By Fanny Landaud, Éric Maurin, Barton Willage, and Alexander Willén*

Abstract

This paper studies the impact of exam luck on individuals' education and economic success. We leverage unique features of the Norwegian education system that produce random variations in the content of the exams taken by students at the end of high school. Lucky students are asked to take exams in the subjects they are most comfortable with, and we show that this generates very significant improvements in both their high school GPA and diploma probability. Subsequently, exam luck generates substantial and persistent wage differentials across otherwise identical individuals. These luck-induced wage differentials are of a similar magnitude as those generated by well-known education inputs, such as parental education or teacher quality. Exam luck impacts students' labor market outcomes mainly through increases in their GPA and ensuing improvements in their higher education outcomes; by contrast, luck-induced increases in diploma probability contribute to reducing time to high school graduation, but have little long-term consequences.

JEL Codes: D63, H52, I21, I23, I24, I26, J24, J31.

Keywords: luck, fairness, wage differentials, returns to education, high-stakes exams.

*Landaud: Norwegian School of Economics (fanny.landaud@nhh.no). Maurin: Paris School of Economics (eric.maurin@ens.fr). Willage: Louisiana State University (bwillage@lsu.edu). Willén: Norwegian School of Economics (alexander.willen@nhh.no). This project received financial support from the Research Council of Norway through its Centers of Excellence Scheme, FAIR project no. 262675, and from the NORFACE projects GUODLCCI and HuCIAW. The authors would like to thank Kjell Salvanes, Joshua Goodman, Michael Lovenheim, Richard Murphy, Martin Eckhoff Andresen, and Analisa Packham, as well as participants at the FAIR Virtual Workshop on the Economics of Education, for helpful discussions.

1 Introduction

Standardized screening and assessment tests are used frequently across the globe, both in the education sector and the labor market. These tests are designed to provide an objective measure of ability, and are heavily utilized for decisions related to university admission, internships, occupational licensing, and entrance into prestigious occupations. Performance on these tests therefore has important and lasting consequences not only for individuals' professional lives, but also for the broader economy due to the potential skill mismatch it induces. However, performance on test day depends on several factors that are difficult to control, and test results do not necessarily provide an accurate reflection of an individual's ability. Individuals with identical skills and abilities can end up with meaningfully different scores simply due to luck.

A primary source of luck in standardized screening and assessment tests is the assignment of questions or topics on tests. Specifically, only a small subset of possible questions are asked on an exam, and individuals with the same underlying ability may score differently depending on which questions are asked or which topics are tested. Lucky individuals will be asked questions on topics and chapters they are most comfortable with, and unlucky individuals will be asked questions on topics and chapters they know very little about. This luck component is likely particularly important in high school exit exams, as the results on these exams not only determine if students graduate from high school and become eligible for university, but also influence the high school GPA with which students apply to university (conditional on graduation).

Luck-induced variation in high school exit exams may have detrimental consequences for the individuals who end up being unlucky. In addition, it may have adverse effects on the allocation of talent across occupations, with negative effects on productivity. Understanding the extent and magnitude of the luck components in high school exit exams –and its effect on individual students– is therefore of great importance, as much from the point of view of social justice as from the point of view of economic efficiency. To the best of our knowledge, however, little is known about the role of luck and its two key components (i.e., the diploma component and the GPA component) on high school exit exams. Empirically, it is an extremely challenging set of questions to address, because it requires a setting in which high school students with identical underlying ability are randomly assigned to different high school exit exams (or exposed to different exam questions) that are more or less aligned with their academic strengths. In addition, it requires an institutional context in which such luck-induced variation does not affect the probability of obtaining a high school diploma and the probability of obtaining a high GPA in exactly the same way, such that the effects of the two components can be separately identified.

In this paper, we directly address these questions by exploiting a unique feature of the Norwegian education system that produces random variations in the exams taken by students at the end of high school. In the Norwegian context, the final evaluation of a student is based not only on teacher grades, but also on the grades obtained on a set of externally set and assessed written exams that are randomly drawn from the courses that students take. These exams generate exogenous variation in the probability of obtaining a good GPA as well as in the probability of obtaining a high school diploma across otherwise identical individuals. From the point of view of GPA, a good draw is a draw that expose students to exams in courses they are relatively strong in. From the point of view of graduation, a good draw is primarily one that minimizes the risk of receiving a failing grade in a subject, since graduation requires that one has no failing grade. One interesting feature of this set-up is that one draw may be better than another from the point of view of the probability of obtaining a good GPA, but not from the point of view of the probability of obtaining one's degree and vice versa.

We use these features of the Norwegian education system to construct and separately identify the effects of two luck components that depend exclusively on the random draw of exams –one that predicts students' GPA and one that predicts students' probability of earning a high school diploma at the end of the school year. Exploiting rich population-wide administrative data covering the universe of Norwegian students and much of their demographic, education, and labor market information, we examine the impact of our two luck measures on high school performance, college enrollment and performance, and long-term labor market outcomes. This is the first study to perform such an analysis and it generates several key results. First, we find that diploma luck has little long-term effect on students' education and labor market outcomes. This result is driven by the fact that students who are unlucky and fail to graduate respond by repeating the last year of high school and successfully graduate the following year. In this respect, the main effect of bad diploma luck is to increase time to high school graduation, such that the effect on graduation itself is small and of no detectable consequence in terms of long run outcomes. Second, we find that GPA luck not only has significant effects on students' exam scores and GPA, this luck component also has persistent impacts on students' long-run outcomes. Nine years after the exams, a one standard deviation increase in this luck component yields 0.7% higher annual market wages – similar to the effect of critical inputs in the education production function, such as parental education or teacher quality. This effect is equally strong for low-ability and high-ability students. Auxiliary analyses suggest that the main reason why GPA luck increases wages is that it increases the quality of time spent at university. Specifically, GPA luck enables students to enter more selective universities, meet peers of higher academic standing, and make academic choices that are more in line with their aspirations. Finally, under the assumption that luck affects students' labor market outcomes only insofar as it affects their actual high school GPA, it is possible to develop an instrumental variable design to estimate of the effect of high school GPA on students' long-term outcomes. This analysis suggests that high school GPA has a very large impact on students' long-term outcomes; a one standard deviation increase in high school GPA is associated with a 42% increase in market wage nine year after high school. In our sample, the IV estimate of the causal effect of GPA on wages appear to be much stronger than the usual OLS estimate (12%).

Taken together, the results from our analysis demonstrate that random variation in high school GPA generates long-term differences between individuals with the same initial level of ability, regardless of that initial level. We conclude that the use of standardized screening and assessment tests as a primary criterion for factors such as university admission, occupational licensing, entrance into occupations, scholarships, and internships, is likely sub-optimal in terms of fairness and maybe also in terms of efficiency.

Our paper contributes to the long-standing literature that explores the role luck plays in social and economic success (e.g., Audas, Barmby and Treble, 2004; Bertrand and Mullainathan, 2001; Frank, 2016; Jenter and Kanaan, 2015). The most studied form of luck is the birth lottery that allocates genes and early social environments to individuals (e.g., Black and Devereux, 2011; Mogstad and Torsvik, 2021). Using the words of Dworkin (1981), this birth lottery can be seen as an example of "brute luck" which is randomly distributed across individuals throughout their life, and many social policies are designed to counteract this form of luck and level the playing field (e.g., Cappelen et al., 2013; Konow, 2000).[1] One of the contributions of our paper is to show that the "brute luck" that inevitably determines the results of high-stakes exams can have long-run effects of the same order of magnitude as those of the birth lottery: a one SD difference in our measures of GPA luck generates differences in students' labor market outcomes that are not much different from the improvements generated by 60% of a SD increase in paternal education (or by 50% of a SD increase in maternal education). Identifying and measuring the role of luck is all the more important because the perception of its role determines people's attitudes towards redistribution and taxation (e.g., Alesina, Stantcheva and Teso, 2018; Lefgren, Sims and Stoddard, 2016).

Our paper also advances the burgeoning body of research examining the importance of exogenous

[1] Dworkin (1981) introduced the distinction between "brute luck" and "option luck", where option luck results from deliberate gambling choices.

shocks to the conditions under which high-stake exams are taken. So far, the literature has focused on the external conditions prevailing on the day of the test, whether in terms of outdoor temperature or the presence of pollutants or pollen in the atmosphere (e.g., Amanzadeh, Vesal and Ardestani, 2020; Bensnes, 2016; Ebenstein, Lavy and Roth, 2016; Garg, Jagnani and Taraz, 2020; Park, 2020). In addition, three papers have used the specific structure of Norwegian exams to study the effect of exam preparation (Bensnes, 2020; Falch, Nyhus and Strøm, 2014), or the effect of exam format (Andresen and Løkken, 2020). Similar to our setting, all these papers generally find that student performance is susceptible to exam conditions. However, our paper focuses not on the external conditions of the exam (or in the period leading up to the exam), but on another fundamental source of randomness, namely the content of the exams themselves. This source of randomness is different in nature, especially since it cannot be easily remedied by harmonizing exam preparation and exam format; it is at work in all outdoor conditions, in all climates, and can affect all students equally, whether or not they have health problems. We also complement this literature by simultaneously examining the two main channels through which test-taking may affect students' outcomes –the diploma channel and the GPA channel.

Our paper also contributes to, and complements, the literature on the effects of high school GPA on student's later-in-life outcomes. There is a well-established strand of education research that has identified a strong correlation between students' high school GPA and their subsequent performance, but it is still not clear whether these correlations should be interpreted as causal (e.g., Black, Cortes and Lincove, 2016; Cohn et al., 2004; Cyrenne and Chan, 2012; French et al., 2015). Using randomized shocks on the contents of exams as a source of identification, our findings suggest that high school GPA is in itself a very important source of later success for students, and that a main channel through which this effect operate is access to more selective (and preferred) universities. That access to more selective college programs has significant labor market effects are in line with recent studies in Europe and the US that have exploited admissions thresholds to study the effect of marginally better college programs on later-in-life outcomes (e.g., Clark and Del Bono, 2016; Deming et al., 2014; Heinesen, 2018; Hoekstra, 2009; Kirkeboen, Leuven and Mogstad, 2016; Kuuppelomäki et al., 2019; Öckert, 2010).

The rest of this paper is organized as follows: In Section 2, we provide an overview of the education system in Norway, as well as detailed information about the randomized exams that students take during high school; In Section 3, we introduce our data, discuss our sample, and outline our empirical method; In Section 4, we present the main results from our analysis, explore mechanisms, and provide a rich set of robustness tests and sensitivity analyses; In Section 5, we conclude and discuss policy recommendations.

2 Background

2.1 The Norwegian Education System

The Norwegian education system consists of 10 years of compulsory school starting the year children turn 6. Upon successful completion of compulsory school, all children have the right to attend 3 to 4 years of high school. Even though 95 percent of students choose to enroll in high school, only about 80 percent of each cohort ends up with a high school diploma. Education is free at all levels, including post-secondary school.

High school consists of two different tracks: a three-year academic track which provides students with direct access to higher education, and a four-year vocational track (two years in school followed by a two-year apprenticeship period) which results in a trade or journeyman's certificate. Approximately 50 percent of students choose to enroll in the vocational track, and 50 percent choose to enroll in the academic track. As very few vocational track students pursue higher education, this paper will focus exclusively on high school students who are enrolled in the academic track.

A range of universities and colleges offer higher education in Norway, and the majority of these are tuition-free public institutions. Eligibility for admission to these institutions is conditional on graduating from high school. The Norwegian Universities and Colleges Admission Service coordinates the admissions process. Students apply to specific programs and universities, and if the number of applications exceeds the number of seats, students are assigned almost exclusively based on high school GPA.[2] The more demand there is for a specific university, the higher is the minimum GPA required to gain admission to that university.

2.2 High School GPA, High School Diploma, and Randomized Exams

In the Norwegian context, a higher GPA provides access to a larger set of universities, and to higher quality universities. High school graduation and high school GPA are thus two decisive outcomes for high school students. In this subsection, we describe how they both depend not only on course grades, but also on grades obtained from randomized external exams taken throughout high school.

The exam subjects are chosen randomly, take place at the end of each school year, and the subjects are announced less than a week prior to the exam.[3] In terms of exam structure, students take between five and six exams throughout high school. In the first year, 20 percent of students are randomly selected

[2]There are also a few bonus points (related to factors such as age, gender, and military service experience), but the main determinant is the GPA. For more details, see Kirkeboen, Leuven and Mogstad (2016).

[3]Even is the delay is short, students can use these few days to prepare for exams (Bensnes, 2020). It can have the effect of mitigating the impact of being lucky (or unlucky) in the draw on subsequent exam results.

for either a written or oral exam in a randomly chosen subject. In the second year, all students take either a written or an oral exam in a randomly chosen subject. In the third and final year, all students take one written exam in Norwegian, two written exams in randomly chosen subjects, and one oral exam in a randomly chosen subject.[4] The randomization of students to subjects and type of test is delegated to the municipality. While the oral exams are locally designed and graded, the written exams are centrally designed and graded. Not all subjects that a student takes can be selected for written and oral exams, and this is known to the students prior to enrolling in high school.

An exam grade counts as much towards GPA as a course grade. High school GPA consists of the average of all of a given student's course grades and randomized exam grades in high school. The exams account for approximately 20 percent of the total number of elements that make up the GPA. Exams and courses are graded on a scale from 1 (worst) to 6 (best), where 1 constitutes a failing grade.

Successful graduation from academic high school, and eligibility to higher education, requires that the student passes all high school courses. This means that a student does not qualify for higher education in the event of receiving a grade of 1 in any of her courses. However, there is an exception to this rule that relates to a situation in which a student has failed the course, but been randomly drawn into and passed an exam in the subject. In such an event, the "pass" status from the randomized exam trumps the "fail" status from the course grade, and the student can receive a high school diploma and become eligible to enroll in higher education.

The selection of students into randomized exams therefore impacts students in two distinct ways. First, it impacts their diploma probability. Second, it impacts their GPA. However, it is not the same criterion that determines high school diploma and high school GPA; diploma is determined by the absence of failed courses (subject to the exception mentioned above) while GPA is a linear function of course grades and exam grades. Thus, a student may have a higher GPA than another student but still not obtain a high school diploma. More generally, the randomized exams at the end of the school year can have a very different impact on the GPA of the student and on the diploma probability of the student. This makes it possible to separately identify the effect of diploma luck and GPA luck on students' subsequent outcomes. It is this unique property of the Norwegian education system that will enable us, for the first time, to separately identify the effect of high school GPA luck and high school diploma luck on students' subsequent outcomes. Exam performance in the first and second year of high school may impact which courses students choose in the third year, what study specializations they select, and could even have

[4] Before 2008, students took two written exams in Norwegian, one written exam in a randomly chosen subject, and one oral exam in a randomly chosen subject.

an effect on dropout rates (see e.g., Andresen and Løkken, 2020; Hvidman and Sievertsen, 2021).[5] To avoid sample selection problems, we abstract away from randomized exams in the first two years of high school, and focus exclusively on the written exams in the third and final year.

3 Data and Method

3.1 Data

Our data come from national population-wide registers covering all Norwegian residents who were enrolled in the final year of high school between 2003 and 2009. A unique personal identifier enables us to follow individuals over time and across registers, such that we can construct a longitudinal panel covering the universe of students and much of their demographic, education, labor, and family background information. We obtain demographic characteristics from the central population register, we collect education information from the national education register, and we use income information from the tax register.

In terms of education data, we have information on high school GPA, diploma status, and whether the student has qualified for higher education. In addition, we have data on all courses that students take in high school, the grades they received in these courses, which courses students were randomized to take exams in, and which grades they received on those exams. Finally, we have information on enrollment in higher education, college major choice, and college degree completion.

With respect to labor market information, we have detailed information on both income as well as employment for the entire sample for each year up until 2018. Income is measured as pre-tax income (labor income and income from self-employment) including certain taxable government transfers (parental leave, sickness leave and unemployment benefits). Employment status (employed, unemployed, and not in the labor force) is defined based on the individual's status in the employment register. In our analysis, we focus on employment and income 9 years after high school graduation. Data limitations prevent us from exploring longer-term outcomes.

Concerning background characteristics, we have access to information on compulsory school GPA, age, sex, and municipality of residence. By exploiting unique family identifiers in the Norwegian registers, we are also able to link students to their parents and collect information on parents' age, educational attainment, and earnings.

[5]For example, Hvidman and Sievertsen (2021) exploits a Danish grade scale recoding reform that impacted students in the first year of high school to isolate the behavioral response to a change in the incentives associated with high-stakes exam grades. They find that individuals who experienced a reduction in GPA due to the reform exerted more effort and performed better in subsequent years.

In terms of sample selection, we restrict our analysis to students in the academic track who are enrolled in the final year of high school between 2003 and 2009. We further restrict the sample to those students who take at least one course in a subject that the student may be asked to do an exam in. This provides us with a sample size of approximately 130,000 individuals.

Table 1 provides descriptive statistics on all individuals in our sample. The table shows that 82% of the sample earn a high school diploma on time, 93% eventually earn a high school diploma, and almost 92% start college. 83% are ever employed, and 75% are employed 9 years after their third year of high school. Conditional on having earnings, log earnings in the first job are 12.2 NOK (approximately 200,000 NOK) and log earnings 9 years after taking 3$^{\text{rd}}$ year exams are 12.6 NOK (approximately 310,000 NOK).

3.2 Construction of Luck Variables

We use the unique setting in the Norwegian education system to construct two random variables that depend exclusively on the random draw of exams –one that predicts students' final GPA and one that predicts students' probability of completing high school. These variables enable us to separately identify the effect of high school diploma luck and high school GPA luck on students' short- and long-term education and labor market outcomes.

GPA Luck. To construct our GPA luck variable ($Luck_{GPA_i}$) we first generate, for each student i and course s, a measure of the score that student i can expect on an exam in subject s ($Exam^e_{i,s}$). We define $Exam^e_{i,s}$ as the average score obtained on the end-of-year exam in subject s by students (other than student i) who attended the same high school as student i, earned the same teacher assessment in the course on subject s as student i, and were randomly assigned to an exam in subject s. The underlying assumption is that teachers' assessments are good predictors for their students' exam results.

Building on this set of expected exam results, we then construct –for each student i and each possible combination (c) of exams[6] –a measure of the GPA that student i can expect if randomly assigned to that combination c of exams:

$$GPA^e_{i,c} = \frac{1}{S+K}\left(\sum_{s\in S} Course_{i,s} + \sum_{s\in c} Exam^e_{i,s}\right) \tag{1}$$

where S and K are the number of courses and exams that student i takes, and $Course_{i,s}$ is the score

[6] A given combination of exams (c) is drawn from all possible combinations of exams (C). The set of possible combinations C is determined by the number of exams taken (K) and the number of courses taken (S). For example, if $S = 10$ and $K = 3$, there are $\frac{10\times9\times8}{3\times2} = 120$ possible combinations.

that student i obtained from the teacher's assessment on course s. The expected total number of grade points is in parentheses. The first sum is the number of grade points from teachers' course assessments across all courses that the student takes. The second sum is the expected number of grade points from the randomly-assigned exams.

After having constructed $GPA^e_{i,c}$ for each student i and possible exam combination c, we subtract its average across all possible combinations c and scale the resulting difference by the standard deviation of its distribution across all possible combinations c.[7] This allows us to define a normalized version of variable $GPA^e_{i,c}$, which harmonizes the measure of luck in the student population. Denoting $c(i)$ the specific combination of exams that student i is randomly assigned to, we define our luck GPA variable as the value taken by the normalized version of $GPA^e_{i,c}$ when $c = c(i)$:

$$Luck_{GPA_i} = \frac{GPA^e_{i,c(i)} - \overline{GPA_i}}{SD_i(GPA)} \tag{2}$$

Diploma Luck. In Norway, at the end of each course, students receive a grade between 1 and 6 from their teachers. To graduate from high school and become eligible for higher education, a student must earn a minimum teacher grade of 2 in each course, except in courses that are randomly selected for a final exam. In these courses, students must earn a minimum grade of 2 in the final exam, and the teacher grade does not matter for passing the course. Thus, in these courses students can obtain a failing teacher grade but still pass the course as long as they obtain a minimum grade of 2 in the final exam. Put differently, when a student has a "fail" status from the course grade and a "pass" status from the final exam, the "pass" status from the final exam trumps the "fail" status from the course grade. Conversely, the "fail" status from a final exam trumps the "pass" status from a course grade.[8]

To define our diploma luck variable ($Luck_{Diploma_i}$), we first construct, for each student i and course s, a measure of the probability that student i obtains a "pass" grade (i.e., a grade of 2 or more) on the final exam in subject s ($D^e_{i,s}$). Specifically, we define $D^e_{i,s}$ as the proportion of students who obtained a "pass" on the final exam in course s among students (other than student i) who take a final exam in course s and who earned the same teacher assessment as student i in course s. Building on this set of measures, we then construct –for each student i and possible combination of exams c– a measure of the high school

[7]For each student i, the average value and standard deviation of $GPA^e_{i,c}$ across all possible c are given by $\overline{GPA_i} = \sum_{c\in C} P_c \times GPA^e_{i,c}$ and $SD_i(GPA) = \sqrt{\sum_{c\in C} P_c \left(GPA^e_{i,c}\right)^2 - \left(\overline{GPA_i}\right)^2}$. P_c is the probability of drawing a particular combination of exams as measured by the fraction of all students who draw that combination. As discussed in Section 2, not all subjects are used for the end-of-year exams, and some subjects have a higher probability of being drawn than others.

[8]As noted in Section 2, the randomization of students to exam subject is delegated to the municipality. Thus, student cannot simply choose to opt into an exam for a certain subject once they have realized they are failing a course.

diploma status that student i can expect if randomly assigned to that combination c of exams:

$$Diploma^e_{i,c} = \prod_{s \notin c} \mathbb{1}\left\{Course_{i,s} > 1\right\} \times \prod_{s \in c} D^e_{i,s} \tag{3}$$

The first product is equal to 1 if the student passed all courses for which they were not randomly assigned an end-of-year exam, where passing is earning a teacher assessment of 2 or more. The second product is the probability the student earns a two or more on all exams they were randomly assigned. A student earns a high school diploma if they pass all exams, and if they pass all courses for which they were not randomly asked to take an exam.

After having constructed $Diploma^e_{i,c}$ for each student i and possible combination c, we subtract its average across all possible combinations c and scale the resulting difference by the standard deviation of its distribution across all possible combinations c.[9] This allows us to obtain a normalized version of $Diploma^e_{i,c}$, which enables us to compare students with different portfolios of courses and different academic abilities. Denoting the specific combination of exams that student i is randomly assigned to $c(i)$, we define our luck diploma variable as the value taken by the normalized version of $Diploma^e_{i,c}$ when $c = c(i)$:

$$Luck_{Diploma_i} = \frac{Diploma^e_{i,c(i)} - \overline{Diploma_i}}{SD_i(Diploma)} \tag{4}$$

Appendix Figure A1 shows the distribution of our luck measures. Both measures follow a Gaussian-like distribution and, as expected, they appear to be evenly distributed around zero. In the remainder of the paper, we winsorize the top and bottom 0.1% of our luck measures to ensure that our results are not driven by a few outliers.[10] Both luck measures are positively correlated, but this correlation is far from perfect (0.4) and $Luck_{Gpa}$ only explains about 15% of the variation in $Luck_{Diploma}$. In this context, it is possible to simultaneously investigate the role of $Luck_{Gpa}$ and $Luck_{Diploma}$ for students' subsequent outcomes.

3.3 Empirical Method

After having constructed our luck variables for high school diploma and high school GPA, we leverage these variables to estimate the impact of high school diploma luck and high school GPA luck on students'

[9]For each student i, the average value and standard deviation of $Diploma^e_{i,c}$ across all possible c are given by $\overline{Diploma_i} = \sum_{c \in C} P_c \times Diploma^e_{i,c}$ and $SD_i(Diploma) = \sqrt{\sum_{c \in C} P_c \left(Diploma^e_{i,c}\right)^2 - \left(\overline{Diploma_i}\right)^2}$. P_c is the probability of drawing a particular combination of exams as measured by the fraction of all students who draw that combination.

[10]Subsection 4.6 checks that our results are robust to winsorizing.

short- and long-term education and labor market outcomes. Specifically, denoting Y_i the outcome of individual i, we estimate versions of the following regression model:

$$Y_i = \alpha + \beta_1 Luck_{Gpa_i} + \beta_2 Luck_{Diploma_i} + \eta_l + u_t + X_i\gamma + \epsilon_i \tag{5}$$

where the $Luck_{Gpa_i}$ and $Luck_{Diploma_i}$ terms represent our GPA and diploma luck variables while η_l and u_t represent full sets of high school fixed effects and year fixed effects. Equation 5 also contains a rich set of demographic controls (X_i). They include students' average high school course grade (linear and squared), average middle school GPA (linear and squared), sex, age (linear and squared), parents' age (linear and squared), parents' years of schooling (linear and squared), and parents log earnings. In Section 4, we provide evidence that our results are robust to using alternative sets of demographic controls. Standard errors will be clustered at the high school-by-year level (i.e., the random assignment level).

The coefficients of interest in Equation 5 are β_1 and β_2. They measure the impact of high school GPA luck and high school diploma luck, respectively. They are identified under the assumption that the luck variables are uncorrelated with unobserved determinants of students' outcomes (ϵ_i). In theory, the validity of this assumption follows directly from the fact that the two luck variables only depend on the combinations of exams to which students are assigned in each high school, which is random by design. In practice, it is possible to obtain suggestive evidence on the validity of this assumption by examining if the two luck variables are correlated with observed determinants of student outcomes (as measured in pre-assignment years). To this end, Table 2 shows results obtained from separately regressing the two luck variables on the grades assigned to students by teachers during the academic year (high school course grades), the grades assigned to students by teachers and their exam grades at the end of middle school (middle school GPA), and numerous sociodemographic variables (students' age and gender as well as parents' average age, education, and income). We also include the square term of each of the continuous variables.

Consistent with the random assignment assumption, the results in Table 2 demonstrate that there is very little correlation between the luck variables and observed student characteristics as measured in pre-assignment years. Specifically, only two of the 24 coefficients are statistically significant at the 10 percent level (the square term of high school course grades and gender), and none of the coefficients are economically meaningful. For both regressions, conventional F-tests cannot reject that all coefficients are jointly equal to zero.

4 Results

In this section, we present and discuss results on how exam luck affects students' short- and long-term education and labor market outcomes. In Section 4.1 we explore high school outcomes; in Section 4.2 we examine higher education outcomes to highlight the channels through which exam luck may impact career trajectories; in Section 4.3 we look at labor market outcomes. Following students from high school into the labor market enables us to trace the full effect of exam luck on students –from the immediate impact on exam grades to the long-run impact on labor market earnings– and provides us with a rich understanding of how exam luck impacts the human capital accumulation and labor market trajectory of students. All results in this section are based on estimations of Equation 5 and include a rich set of demographic controls, a full set of high school fixed effects, and a full set of year fixed effects. In Section 4.4, we probe the data further and explore effect heterogeneity across student gender and ability. In Section 4.5, we present results from an IV analysis where we use exam luck as a source of identification for the causal effect of students' high school GPA on their subsequent labor market outcomes. In Section 4.6, we document the robustness of our results to a range of falsification tests and sensitivity analyses.

4.1 High School Outcomes

The effect of exam luck on high school outcomes are shown in Table 3. The primary high school outcomes we examine are the students' exam grades (column 1), the students' GPA for the third year of high school (column 2), the students' overall high school GPA which is used to apply to universities and colleges (column 3),[11] a dummy variable indicating if the students receive on-time high school diplomas (column 4), and a dummy variable indicating if the students ever receive high school diplomas (column 5). Given the way our luck variables are constructed, we expect the first three outcomes to be impacted first and foremost by our GPA luck variable and the last two outcomes by our diploma luck variable.

The results in Table 3 are very much in line with this expectation. The first three columns reveal that both measures of luck have a statistically significant and economically meaningful effect on students' exam grades and high school GPA, but that it is the impact of GPA luck that is most important. For example, the result in column (1) reveals that a one SD increase in GPA luck leads to a 10% standard

[11]The overall high school GPA which is used to apply to universities and colleges includes teacher and exam grades during the three years of high school. As a consequence, the number of observations underlying the estimation in column (3) is smaller than that in column (2). The reason for this is that we only have GPA data available from year 2003. Thus, for the oldest cohorts in our sample, we are unable to calculate their entire high school GPA, and can only calculate their GPA for the last year of high school. To ensure that we obtain similar results to those in column (2) when we restrict our sample to this subsample, we have also re-estimated the regression underlying column (2) using the smaller sample. The results we obtain are not statistically different from those shown in Table 3 column (2).

deviation increase in students' exam grades while a one SD increase in diploma luck only generates a 4% standard deviation increase in exam grades. The impact of GPA luck is also considerably larger on students' high school GPA, whether this GPA corresponds to the third year alone or to the whole high school years. This last result means that GPA luck directly impacts the metric with which students apply to university and college, and that it may have important implications on individuals' careers not only in the short-run, but also in the long-run. We explore this in greater details below.

In terms of on-time diploma receipt, the results in column (4) show that both GPA luck and diploma luck causally impact students' probability of obtaining an on-time high school diploma. However, as expected, the impact of the diploma luck component is now considerably larger than that of the GPA luck component, with a coefficient that is about three times larger. Again, this is expected, as the diploma luck variable is precisely designed to predict students' probability of earning a high school diploma at the end of the school year, while the GPA luck variable is designed to predict students' GPA. In terms of magnitudes, a one standard deviation improvement in diploma luck leads to a 1.2-percentage point increase in the probability of receiving an on-time high school diploma, while a one standard deviation increase in GPA luck leads to a 0.4-percentage point increase in the probability of on-time diploma receipt.

Finally, the results in column (5) show that the effect on on-time diploma receipt extend to ever receiving a high school diploma as well. However, the magnitude of the effect is smaller, which suggests that many students who fail to secure an on-time diploma due to bad luck return to school to take supplemental classes and receive a diploma at a later time.

4.2 Higher Education Outcomes

The effect of high school exam luck on higher education outcomes are shown in Table 4. The primary outcomes we explore in this section are a dummy indicating if the students receive any college education (column 1), the average peer GPA at entry into college (column 2), the minimum peer GPA at entry into college (column 3), and the number of completed years in higher education (column 4). While the results in columns (1) and (4) provide us with information on the extensive margin effect of high school exam luck on higher education outcomes, the results in columns (2) and (3) provide proxy measures for the selectiveness and education quality that students are exposed to in college (conditional on going to college).

The results displayed in column (1) suggest that diploma luck has a small but non-trivial impact on the college enrollment decisions of individuals. Specifically, a one standard deviation increase in the diploma luck variable yields a 0.2 percentage point increase in the probability of receiving some college education. The GPA luck variable, on the other hand, has no effect on the decision of attending college. This result is consistent with our priors, as the diploma luck variable has a much greater effect on the probability that students obtain a high school diploma (which improves students' chances of qualifying for college), while the GPA luck variable has a bigger impact on the GPA with which students apply to college (which improves students' chances of qualifying for better programs and colleges). As such, we would expect a larger extensive margin effect of diploma luck, and a potentially larger intensive margin effect of GPA luck.

In terms of education quality, the results in columns (2) and (3) demonstrate that the GPA luck variable has a sizable impact on both the average peer GPA at entry into college and the minimum peer GPA at entry into college. Specifically, a one standard deviations change in GPA luck shifts the average peer GPA in college by 0.4 percent of a standard deviation, and the minimum peer GPA by 1 percent of a standard deviation. As admission to college is primarily based on high school GPA (conditional on receiving a high school diploma), this suggests that exam luck in high school enables students to "upgrade" their college quality through admission into more selective programs and universities with higher-ability peers. By design, a higher high school GPA also enables students to attend preferred programs and universities, i.e., programs and universities that students ranked higher when submitting their college applications. In this context, GPA luck may also enable students to attend programs and universities in which they have a comparative advantage (Kirkeboen, Leuven and Mogstad, 2016). In terms of diploma luck, we do not detect any impact on the college quality dimension. However, it is important to note that the extensive margin effect of diploma luck identified in column (1) means that there are compositional changes in terms of whom enter college as a function of this variable, and we must therefore be careful when interpreting the intensive margin quality effects of diploma luck in columns (2) and (3).

Finally, the results in column (4) demonstrate that there is no impact of GPA luck or diploma luck on the number of years completed in higher education. This has two important implications. First, it suggests that the quality upgrading that GPA luck contributes to is not offset by a potential reduction in educational attainments due to admission into more difficult schools and programs. Second, it implies that the enrollment effect generated by diploma luck is not permanent, in the sense that those who are induced to enroll because of diploma luck do not pursue in higher education until they complete their degree. Taken together, this implies that GPA luck is more likely to impact students' labor market outcomes once they have finished their education, as GPA luck has persistent effects on students' trajectories in higher education. By contrast, the education effects of diploma luck on students' high school diploma are partly offset by endogenous responses: students who fail to graduate on time due to bad luck at the exams take supplemental classes and manage to graduate latter in time; and students who manage to enroll in higher education due to diploma luck drop out before obtaining additional degrees.

4.3 Labor Market Outcomes

Understanding the impact of exam luck on the short- and long-run educational attainments of students is of great independent value. However, we are ultimately interested in understanding to what extent these effects translate into changes in the labor market opportunities of students once they have completed their human capital investments. To this end, we follow the affected students into the labor market and examine both their employment status as well as their earnings. These results are shown in Table 5. The data we have access to enable us to follow students up to 9 years after they have taken their third year high school exams, and we use this data to study a range of outcomes: the probability of ever having been employed (column 1), log annual labor income at the first job (column 2), the probability of being employed nine years after taking the tests (column 3), and log annual labor income measured nine years after graduating from high school (column 4).

In terms of extensive margin employment effects (columns 1 and 3), the results show that neither diploma luck nor GPA luck has a significant effect on employment. With respect to earnings, the table shows that GPA luck has a sizable impact both on the annual labor income at the first job the students secure (+0.9%, column 2), as well as on their annual labor income nine years after having taken their high school exit exams (+0.7%, column 4). With respect to diploma luck, we do not find a significant impact on earnings. This suggests that GPA luck –potentially through its impact on higher education quality and match– drives the long-term effects of exam luck on earnings.

To compare the wage differentials generated by exam luck to those generated by the birth lottery, we estimated the relationship between parental education and child's earnings. In our Norwegian sample, we find that a 1 SD increase in fathers' (mothers') years of education is associated with a 1.1% (1.4%) increase in children's annual labor income nine years after the exams. This suggests that luck GPA generates wage differences that are not much different from those generated by a 60% (50%) SD increase in paternal (maternal) education.To further put these effects in perspective, it is also possible to contrast them with the labor market impacts of well-known education inputs analyzed in the prior literature, such as teacher quality. Chetty, Friedman and Rockoff (2014) find that a one standard deviation increase in teacher value-added during one grade is associated with 1.3% higher annual earnings. Thus, luck at high-stakes high school exams has a similar impact on students career trajectories as a half standard deviation increase in teacher quality.

The lack of an extensive margin employment effect suggests that the identified earnings effects of high school exam luck may be operating through a change in the type of job individuals hold or in the type of firm they work at. To examine this in more detail, Appendix Table A1 studies the impact of exam luck on a number of key firm characteristics: the size of the firm, a dummy variable indicating if the firm is in the public or private sector, and the share of coworkers who have at least some college education (which is used as a proxy for occupational quality). While we find no evidence of differential sorting into the public sector or the size of the firm, we do find consistent evidence that GPA luck has a positive impact on the quality of the coworkers individuals are exposed to. While suggestive, this is consistent with the notion that the earnings effects are operating through an intensive margin quality effect in which students are able to "upgrade" to better occupations and jobs.[12]

4.4 Heterogeneity

In this section, we further probe the data and analyze potential heterogeneous effects of exam luck on the education and labor market outcomes of students by ability and gender.

The role of ability. Tables 6 and 7 show the main education and labor market effects stratified by students who are above or below median ability (as measured by the students' course grades). In terms of education outcomes (Table 6), we find suggestive evidence that the effects on diploma receipt load on students at the lower-end of the ability distribution, while the effects on exam grades and high school GPA are more equally distributed across high- and low-ability students. That individuals at the bottom of the ability distribution are more impacted by diploma luck than high ability students are consistent with the notion that high ability students are generally not at risk of failing to obtain a diploma.

With respect to labor market outcomes (Table 7), we find an impact of GPA luck on individual labor earnings both among those who have above median ability as well as those who have below median ability. The size of the coefficients are relatively similar across the two groups, and we are unable to rule out equality of coefficients through conventional t-tests. Table 7 further confirms that diploma luck does not translate into long-term wage gains, even when we focus on low ability students, i.e., on the students who are the most exposed to failing their diploma.

[12] In addition to education and labor market outcomes, we have also examined the potential impact of exam luck on other fundamental societal outcomes that have been shown to be affected by education interventions in prior literature: teenage pregnancies and marital behavior. However, we find little evidence to suggest that these outcomes are impacted by exam luck. Results are available upon request.

Gender differences. Appendix A2 and A3 show the main education and labor market results separately for boys and girls. The main take-away from this table is that exam luck –whether in terms of GPA luck or diploma luck– impacts boys and girls similarly. The one exception concerns diploma luck and on-time diploma receipt, where the effect is significantly larger for boys. Taken together, we interpret the results from these tables as indicating that there is little gender differences in the effect of high school exam luck on educational attainment or later-in-life labor market outcomes.

4.5 IV Estimation

In the previous section, we just showed that exam luck has a very significant impact on both the high school GPA and the labor market outcomes of high-ability students, but very little impact on their high school graduation probability. Hence, under the assumption that exam luck affects the labor market outcomes of high ability students only insofar as it impacts their high school GPA, it is possible to develop an instrumental variable approach to estimate the causal effect of high school GPA on the labor market outcomes of high ability students using exam luck as a source of identification.

In Panel A of Table 8, we follow this idea and provide results from using our measure of GPA luck as an instrument to identify the causal effects of high school GPA.[13] The results suggest that high school GPA has a very large impact on high-ability students' long-term outcomes. Specifically, a one standard deviation increase in high school GPA generates a 42% increase in their market wage nine years after high school. We have checked that this IV estimate is much higher than the corresponding OLS estimate (12%). Neglecting the endogeneity of high school GPA appears to lead to a significant underestimation of its effect on wages.

In panel B of Table 8, we replicate this exercise with the sample of low-ability students. We find similarly large effects among low-ability students as among high-ability students; a one standard deviation increase in high school GPA is associated with a 35% increase in market wage among low-ability students nine years after high school.

One potential problem with the Panel B estimates is that our GPA luck instrument has a first-stage effect not only on the high school GPA of low-ability students (as in the case of high ability students), but also on their probability of on-time graduation. Therefore, the Panel B estimates implicitly assume that we can neglect the effect that on-time graduation may have, as such, on long-term outcomes.

To overcome this issue, however, it is possible to use jointly our two measures of exam luck (i.e.,

[13]We checked that GPA luck does not impact the probability that high-ability students graduate on time; we also checked that we obtain similar results when we instrument high school GPA using GPA luck and diploma luck jointly.

GPA luck and diploma luck) as instruments to jointly identify the effect of high school GPA and the effect of on-time graduation on the labor market outcomes of low-ability students.[14] When we follow this approach, the GPA effect on long-term annual labor income ceases to be statistically significant at conventional levels, but it remains very similar to the effect estimated in Panel B (see Panel C of Table 8). This confirms the large effect of high school GPA on market wages. With respect to the effect of high school diploma, we find little evidence that it matters for individuals' labor market outcomes. This is consistent with our reduced-form analysis, in which we saw that diploma luck has, as such, very little effects on long-term outcomes.

Taken together, the results from this section point to sizable effects of high school GPA on students' long-run labor market outcomes. In interpreting the results from our IV estimation, we reiterate that they rely on the assumption that exam luck, as such, has no impact on long-term outcomes. If, for example, being lucky (unlucky) on exams was in itself a source of motivation (de-motivation), this exclusion restriction would be violated and our IV estimates of the causal effect of the GPA would likely be upward biased. Exploring these issues would be beyond the scope of this article.

4.6 Robustness and Sensitivity

Parameters β_1 and β_2 in Equation 5 are identified under the assumption that the luck variables are uncorrelated with unobserved determinants (ϵ_i) of students' outcomes. The results in Table 2 provide strong suggestive evidence in favor of this assumption. In this section, we probe the data further and explore the robustness of our results to a number of sensitivity checks and falsification tests. All of these results are provided in Appendix Table A4.

In Panel A, we explore the sensitivity of our results to using control variables selected with the double lasso procedure of Belloni, Chernozhukov and Hansen (2014). The idea behind this exercise it to obtain a more objective set of control variables that are outside the researchers' control. The results in Panel A demonstrate that using the control variables recommended by the double lasso approach does not generate coefficients that are statistically different from our main findings. This suggests that the findings we present in the paper are not driven by the particular set of control variables we use.

In Panel B, we show the p-values obtained from two sets of (1000) permutation tests in which we have randomly assigned GPA or diploma luck values to students, holding the distribution of values of

[14]Panel B of Table 6 show the relevance of these two instruments: they both have a very strong first stage effect on both high school GPA and on-time graduation. Furthermore, the predicted value of the two potentially endogenous regressors are clearly not collinear, since (as expected) high school GPA is mostly predicted by our measure of GPA luck and on-time graduation is mostly predicted by our measure of diploma luck.

these luck variables constant. The first row shows p-values that measure the probability that randomly assigning luck GPA values to students will generate point estimates at least as large as our baseline estimates. The second row report p-values that measure the probability that randomly assigning luck diploma values to students will generate point estimates at least as large as our baseline estimates. If our results were simply an artifact, we would expect these p-values to be large, but our permutation results provide clear evidence against this concern. In particular, less than 1% of the permutations for luck GPA generate wage coefficients of the same size as our main findings.

In Panel C, we relax the winsorization restrictions on our luck measures to ensure that our findings are not driven by the way in which we restrict the range of luck values. This panel demonstrates that our results are unaffected by this adjustment.

Individuals with a failing course grade (i.e., a teacher grade of 1) may be systematically different from students without a failing course grade. In particular, the luck measures we have constructed may impact these students differently. To ensure that our results are robust to eliminating this subset of students, we re-estimate our main results using only those students who do not have a failing course grade. Panel D presents these results, and shows that our findings are robust to this restriction.

Taken together, we interpret the evidence in Appendix Table A4 as providing strong additional support for our identifying assumption, thereby reinforcing the credibility of a causal interpretation of our findings.

5 Conclusion

There is a long-standing debate in the social sciences about the root causes of economic and social inequalities between individuals. The fundamental question is whether they reflect the fact that individuals do not all make the same choices or, more simply, that they are not all equally lucky, especially from the point of view of the family into which they were born and raised. In this paper, we contribute to this debate by showing that exam luck at key points in the school career can have long lasting effects on individuals' outcomes, of the same order of magnitude as the effects of the "brute" luck that assigns them more or less educated parents.

To reach these conclusions, we rely on a unique feature of the Norwegian educational system that produces random variation in the content of the high school exit exams taken by students. These exams generate exogenous variation in the probability of obtaining a good GPA, as well as in the probability of obtaining a high school diploma across otherwise identical individuals. From the point of view of GPA, a good draw is a draw that expose students to exams in courses they are relatively strong in. From the point of view of graduation, a good draw is primarily one that minimizes the risk of receiving a failing grade in a subject, since graduation requires that one has no failing grade.

We use these features of the Norwegian education system coupled with rich population-wide register data to construct and separately identify the effects of two luck components that depend exclusively on the random draw of exams – one that affect students' GPA and one that affect students' probability of earning a high school diploma at the end of the school year.

We find that diploma luck has little long-term effect on students' education and labor market outcomes. This result is driven by the fact that students who are unlucky and fail to graduate respond by repeating the last year of high school and successfully graduate the following year. In this respect, the main effect of bad diploma luck is to increase time to high school graduation, such that the effect on graduation itself is small and of no detectable consequence in terms of long-run outcomes. By contrast, we find that GPA luck has persistent impacts on students' long-run outcomes. Nine years after the exams, a one standard deviation increase in this luck component yields 0.7% higher annual market wages – similar to the effect of critical inputs in the education production function, such as teacher quality or parental education. In terms of mechanisms, we show that our results are consistent with the assumption that the GPA obtained in high school is, as such, a very important determinant of students' long-term outcomes, mainly through its effect on the quality of the higher education to which they can have access.

Taken together, our findings suggest that luck can have a very important impact on high-stakes test scores with very significant long-term consequences for all types of test-takers. These findings are important in their own right, but they also have important implications for the design of education systems. They show that by relying too heavily on high-stakes exams at a few key stages in students' educational careers, we run the risk of misclassifying a large number of students, resulting in an unfair allocation of students to different types of higher education and of young workers to different jobs and occupations. Our findings suggest promoting measures of student quality that are less random and subject to more frequent revision over time than those currently used in many countries. To the extent that the returns to education are likely lower for students whose trajectory is disrupted by luck (or misfortune), such reforms could also have the result of improving the overall skill level and productive efficiency of the workforce. We leave to future research the measurement of the potential efficiency gains that might be generated by a more accurate and fairer ranking of high school students.

References

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review*, 108(2): 521–554.

Amanzadeh, Naser, Mohammad Vesal, and Seyed Farshad Fatemi Ardestani. 2020. "The impact of short-term exposure to ambient air pollution on test scores in Iran." *Population and Environment*, 41(3): 253–285.

Andresen, Martin Eckhoff, and Sturla Andreas Løkken. 2020. "The Final straw: High school dropout for marginal students." University Library of Munich, Germany MPRA Paper 106265.

Audas, Rick, Tim Barmby, and John Treble. 2004. "Luck, Effort, and Reward in an Organizational Hierarchy." *Journal of Labor Economics*, 22(2): 379–395.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies*, 81(2): 608–650.

Bensnes, Simon Søbstad. 2016. "You sneeze, you lose:: The impact of pollen exposure on cognitive performance during high-stakes high school exams." *Journal of Health Economics*, 49: 1–13.

Bensnes, Simon Søbstad. 2020. "Scheduled to Gain: Short- and Longer-Run Educational Effects of Examination Scheduling." *The Scandinavian Journal of Economics*, 122(3): 879–910.

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Are CEOs Rewarded for Luck? The Ones Without Principals Are." *The Quarterly Journal of Economics*, 116(3): 901–932.

Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*, , ed. Orley C. Ashenfelter and David Card Vol. 4, 1487–1541. Elsevier.

Black, Sandra E., Kalena E. Cortes, and Jane Arnold Lincove. 2016. "Efficacy Versus Equity: What Happens When States Tinker With College Admissions in a Race-Blind Era?" *Educational Evaluation and Policy Analysis*, 38(2): 336–363.

Cappelen, Alexander W., James Konow, Erik Ø Sørensen, and Bertil Tungodden. 2013. "Just Luck: An Experimental Study of Risk-Taking and Fairness." *American Economic Review*, 103(4): 1398–1413.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review*, 104(9): 2633–2679.

Clark, Damon, and Emilia Del Bono. 2016. "The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom." *American Economic Journal: Applied Economics*, 8(1): 150–176.

Cohn, Elchanan, Sharon Cohn, Donald C. Balch, and James Bradley. 2004. "Determinants of undergraduate GPAs: SAT scores, high school GPA and high school rank." *Economics of Education Review*, 23(6): 577–586.

Cyrenne, Philippe, and Alan Chan. 2012. "High school grades and university performance: A case study." *Economics of Education Review*, 31(5): 524–542.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2014. "School Choice, School Quality, and Postsecondary Attainment." *American Economic Review*, 104(3): 991–1013.

Dworkin, Ronald. 1981. "What is Equality? Part 1: Equality of Welfare and What is Equality? Part 2: Equality of Resources." *Philosophy & Public Affairs*, 10(3 & 4): 185–246 & 283–345.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution." *American Economic Journal: Applied Economics*, 8(4): 36–65.

Falch, Torberg, Ole Henning Nyhus, and Bjarne Strøm. 2014. "Causal effects of mathematics." *Labour Economics*, 31: 174–187.

Frank, Robert H. 2016. *Success and Luck: Good Fortune and the Myth of Meritocracy.* Princeton University Press.

French, Michael T., Jenny F. Homer, Ioana Popovici, and Philip K. Robins. 2015. "What You Do in High School Matters: High School GPA, Educational Attainment, and Labor Market Earnings as a Young Adult." *Eastern Economic Journal*, 41(3): 370–386.

Garg, Teevrat, Maulik Jagnani, and Vis Taraz. 2020. "Temperature and Human Capital in India." *Journal of the Association of Environmental and Resource Economists*, 7(6): 1113–1150.

Heinesen, Eskil. 2018. “Admission to higher education programmes and student educational outcomes and earnings–Evidence from Denmark.” *Economics of Education Review*, 63: 1–19.

Hoekstra, Mark. 2009. “The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach.” *The Review of Economics and Statistics*, 91(4): 717–724.

Hvidman, Ulrik, and Hans Henrik Sievertsen. 2021. “High-Stakes Grades and Student Behavior.” *Journal of Human Resources*, 56(3): 821–849.

Jenter, Dirk, and Fadi Kanaan. 2015. “CEO Turnover and Relative Performance Evaluation.” *The Journal of Finance*, 70(5): 2155–2184.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. “Field of Study, Earnings, and Self-Selection.” *The Quarterly Journal of Economics*, 131(3): 1057–1111.

Konow, James. 2000. “Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions.” *American Economic Review*, 90(4): 1072–1091.

Kuuppelomäki, Tiina, Mika Kortelainen, Tuomo Suhonen, and Hanna Virtanen. 2019. “Does admission to elite engineering school make a difference?” VATT 127.

Lefgren, Lars J., David P. Sims, and Olga B. Stoddard. 2016. “Effort, luck, and voting for redistribution.” *Journal of Public Economics*, 143: 89–97.

Mogstad, Magne, and Gaute Torsvik. 2021. “Family Background, Neighborhoods and Intergenerational Mobility.” National Bureau of Economic Research 28874.

Park, R. Jisung. 2020. “Hot Temperature and High Stakes Performance.” *Journal of Human Resources*.

Öckert, Björn. 2010. “What’s the value of an acceptance letter? Using admissions data to estimate the return to college.” *Economics of Education Review*, 29(4): 504–516.

Table 1 – Summary Statistics

Variables	Mean	SD	Observations
Outcomes			
High school GPA in 3rd year	0.020	0.970	129917
Overall HS GPA in 3rd year	0.031	0.974	97094
On time HS diploma	0.820	0.384	129917
Ever HS diploma	0.930	0.256	129917
Any college	0.918	0.274	129917
Average peer GPA in college	0.074	0.630	110253
Minimum peer GPA in college	-1.888	1.260	110253
Number of completed years in HE	2.644	1.903	129917
Ever employed	0.827	0.378	129917
First job labor income (log)	12.226	0.907	107461
Employed 9 years after HS	0.745	0.436	129917
Labor income 9 years after HS (log)	12.648	0.631	96734
Demographics			
High school course grades	0.017	0.978	129917
Middle school GPA	0.022	0.991	129917
Female	0.549	0.498	129917
Age	19.093	0.877	129917
Parents' average age	48.195	4.814	129917
Parents' average years of education	13.701	2.996	129917
Parents' average log labor income	12.587	1.173	129917

NOTE: The table refers to the sample of students who enrolled for the first time in the final year of academic high school between 2003 and 2009, and who took at least one course in a subject where they could be assigned to a written exam. The table shows the means and standard deviations of the main outcome and baseline variables. Due to data constraints, statistics on the overall high school GPA in 3rd year are restricted to the students who were in their final year of academic high school between 2005 and 2009. Statistics on college peers' characteristics are conditional on enrolling in college. Statistics on individuals' labor incomes are conditional to being employed.

Table 2 – Balance Tests, Association between Luck and Baseline Characteristics

	Measures of Luck	
	Luck GPA	Luck diploma
High school course grades	-0.0058	0.0029
	(0.0036)	(0.0033)
High school course grades, squared	-0.0013	-0.0024*
	(0.0019)	(0.0014)
Middle school GPA	-0.0029	-0.0045
	(0.0036)	(0.0035)
Middle school GPA, squared	-0.0016	-0.0007
	(0.0014)	(0.0012)
Female	-0.0056	-0.0075*
	(0.0046)	(0.0044)
Age	-0.0026	-0.0040
	(0.0117)	(0.0109)
Age, squared	0.0000	0.0001
	(0.0002)	(0.0002)
Parents' average age	0.0034	0.0056
	(0.0063)	(0.0059)
Parents' average age, squared	-0.0000	-0.0001
	(0.0001)	(0.0001)
Parents' average years of education	0.0007	-0.0024
	(0.0029)	(0.0027)
Parents' average years of education, squared	-0.0001	0.0001
	(0.0001)	(0.0001)
Parents' average log earnings	-0.0006	0.0014
	(0.0018)	(0.0017)
F-statistic	1.245	1.296
Joint p-value	0.212	0.175
Mean	0.015	0.014
N	129917	129917

NOTE: The table refers to the same sample as Table 1. The first column shows the results of regressing our measure of GPA luck on a rich set of baseline demographic characteristics. The second column shows the results of regressing our measure of diploma luck on the same set of baseline demographic characteristics. Both regressions include high school and year fixed effects, and the F-tests of joint orthogonality control for these fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table 3 – Effect of Luck on High School Outcomes

	Outcomes				
	Exam grades in 3^{rd} year	High School GPA in 3^{rd} year	Overall HS GPA in 3^{rd} year	On time HS diploma	Ever HS diploma
Luck GPA	0.0988***	0.0192***	0.0114***	0.0046***	0.0018**
	(0.0025)	(0.0007)	(0.0011)	(0.0011)	(0.0008)
Luck diploma	0.0382***	0.0053***	0.0032***	0.0124***	0.0034***
	(0.0023)	(0.0007)	(0.0011)	(0.0012)	(0.0007)
Mean	0.000	0.020	0.031	0.820	0.930
N	129917	129917	97094	129917	129917

Note: The table refers to the same sample as Table 1. Each column corresponds to a specific regression, and reports the estimated impacts of our two measures of luck –luck GPA and luck diploma– on the dependent variable mentioned above. Due to data constraints, the estimated effects of luck GPA and luck diploma on the overall high school GPA in 3^{rd} year are restricted to the students who were in their final year of academic high school between 2005 and 2009. Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table 4 – Effect of Luck on Higher Education Outcomes

	Outcomes			
	Any college	Average peer GPA in college	Minimum peer GPA in college	Number of completed years in HE
Luck GPA	0.0001	0.0038**	0.0100**	0.0015
	(0.0009)	(0.0019)	(0.0050)	(0.0062)
Luck diploma	0.0016**	0.0005	-0.0050	0.0072
	(0.0008)	(0.0019)	(0.0051)	(0.0063)
Mean	0.918	0.074	-1.888	2.644
N	129917	110253	110253	129917

NOTE: The table refers to the same sample as Table 1. Each column corresponds to a specific regression, and reports the estimated impacts of our two measures of luck –luck GPA and luck diploma– on the dependent variable mentioned above. The estimated effects of luck GPA and luck diploma on college peers' characteristics are conditional on enrolling in college. Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table 5 – Effect of Luck on Labor Market Outcomes

	Outcomes			
	Ever employed	First job annual labor income (log)	Employed 9 years after HS	Annual labor income 9 years after HS (log)
Luck GPA	0.0008	0.0087**	0.0012	0.0067**
	(0.0015)	(0.0035)	(0.0017)	(0.0027)
Luck diploma	0.0002	0.0033	-0.0004	0.0025
	(0.0015)	(0.0036)	(0.0017)	(0.0028)
Mean	0.827	12.226	0.745	12.648
N	129917	107461	129917	96734

NOTE: The table refers to the same sample as Table 1. Each column corresponds to a specific regression, and reports the estimated impacts of our two measures of luck –luck GPA and luck diploma– on the dependent variable mentioned above. The estimated effects of luck GPA and luck diploma on individuals' labor incomes are conditional to being employed. Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table 6 – Effect of Luck on High School Outcomes, Heterogeneity by Ability Based on Course

	Outcomes				
	Exam grades in 3rd year	High School GPA in 3rd year	Overall HS GPA in 3rd year	On time HS diploma	Ever HS diploma
Panel A: High Ability, Above Median Course Grades					
Luck GPA	0.0921***	0.0186***	0.0095***	-0.0011	-0.0007*
	(0.0033)	(0.0009)	(0.0013)	(0.0010)	(0.0004)
Luck diploma	0.0380***	0.0054***	0.0045***	0.0024**	-0.0004
	(0.0032)	(0.0009)	(0.0014)	(0.0011)	(0.0004)
Mean	0.625	0.789	0.795	0.940	0.991
N	64940	64940	47997	64940	64940
Panel B: Low Ability, Below Median Course Grades					
Luck GPA	0.1066***	0.0199***	0.0134***	0.0110***	0.0049***
	(0.0035)	(0.0010)	(0.0017)	(0.0021)	(0.0016)
Luck diploma	0.0376***	0.0049***	0.0012	0.0237***	0.0074***
	(0.0034)	(0.0010)	(0.0017)	(0.0022)	(0.0015)
Mean	-0.625	-0.747	-0.717	0.701	0.869
N	64977	64977	49097	64977	64977

NOTE: The table report similar results as Table 3 separately on the sub-sample of students whose average course grade is above the sample median (Panel A), and on the sub-sample of students whose average course grade is below the sample median (Panel B). Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table 7 – Effect of Luck on Labor Market Outcomes, Heterogeneity by Ability Based on Course

	Outcomes			
	Ever employed	First job annual labor income (log)	Employed 9 years after HS	Annual labor income 9 years after HS (log)
Panel A: High Ability, Above Median Course Grades				
Luck GPA	-0.0002	0.0050	-0.0012	0.0071*
	(0.0020)	(0.0044)	(0.0023)	(0.0037)
Luck diploma	0.0004	0.0029	0.0005	0.0028
	(0.0021)	(0.0044)	(0.0023)	(0.0038)
Mean	0.816	12.419	0.744	12.720
N	64940	53014	64940	48317
Panel B: Low Ability, Below Median Course Grades				
Luck GPA	0.0019	0.0128**	0.0037	0.0065*
	(0.0021)	(0.0053)	(0.0025)	(0.0039)
Luck diploma	-0.0000	0.0032	-0.0017	0.0022
	(0.0021)	(0.0056)	(0.0025)	(0.0041)
Mean	0.838	12.038	0.745	12.576
N	64977	54447	64977	48417

NOTE: The table report similar results as Table 5 separately on the sub-sample of students whose average course grade is above the sample median (Panel A), and on the sub-sample of students whose average course grade is below the sample median (Panel B). Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

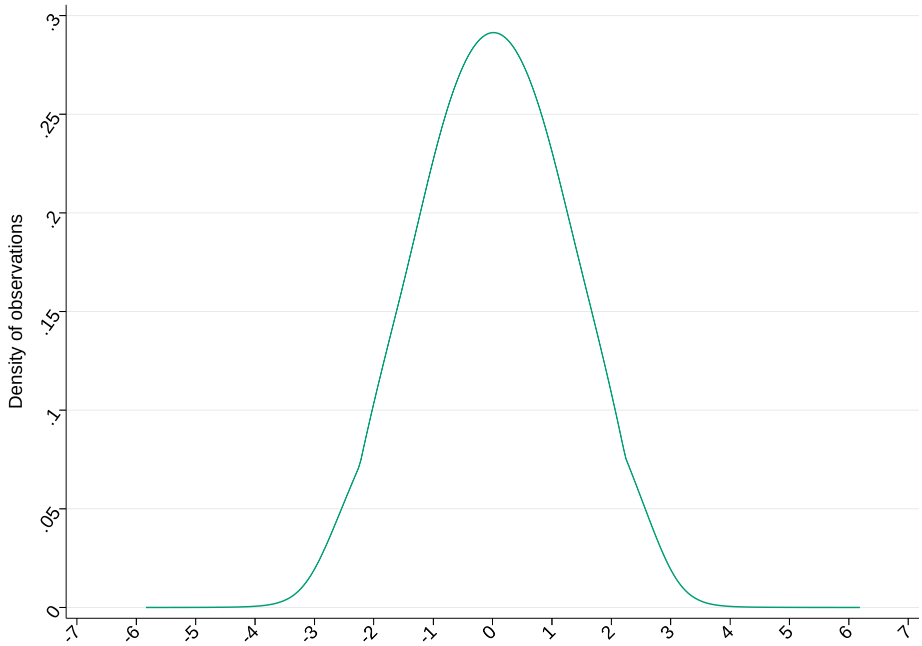
Table 8 – Two-Stage Least Squares Estimates, Effect of HS GPA and Diploma on Labor Market Outcomes by Ability Based on Course

	Outcomes			
	Ever employed	First job annual labor income (log)	Employed 9 years after HS	Annual labor income 9 years after HS (log)
Panel A: High Ability, Above Median Course Grades – Estimated Effect of GPA				
High school GPA in 3rd year	-0.0052	0.3045	-0.0463	0.4187**
	(0.0908)	(0.2050)	(0.0988)	(0.1769)
Mean	0.816	12.419	0.744	12.720
First stage F-statistics:				
Luck GPA on GPA in 3rd year	645			
N	64940	53014	64940	48317
Panel B: Low Ability, Below Median Course Grades – Estimated Effect of GPA				
High school GPA in 3rd year	0.0861	0.6633***	0.1366	0.3479**
	(0.0860)	(0.2294)	(0.1037)	(0.1681)
Mean	0.838	12.038	0.745	12.576
First stage F-statistics:				
Luck GPA on GPA in 3rd year	553			
N	64977	54447	64977	48417
Panel C: Low Ability, Below Median Course Grades – Estimated Effect of GPA and diploma				
High school GPA in 3rd year	0.1077	0.6956	0.2526	0.3319
	(0.1497)	(0.4307)	(0.1781)	(0.3306)
On time HS diploma	-0.0233	-0.0338	-0.1254	0.0166
	(0.1125)	(0.3281)	(0.1346)	(0.2589)
Mean	0.838	12.038	0.745	12.576
First stage F-statistics:				
Luck GPA & diploma on GPA in 3rd year	300			
First stage F-statistics:				
Luck GPA & diploma on HS diploma in 3rd year	118			
N	64977	54447	64977	48417

NOTE: The table refers to the same sample as Table 1. Panel A focuses the sub-sample of students whose average course grade is above the sample median, while Panel B and Panel C focus on the sub-sample of students whose average course grade is below the sample median. Each column corresponds to a specific dependent variable. Panel A and Panel B report the estimated impacts of students' GPA in 3rd grade, instrumented by our measure of luck GPA, on the dependent variable mentioned above. Panel C reports the estimated impacts of students' GPA and diploma in 3rd grade, instrumented by our measures of luck GPA and luck diploma, on the dependent variable mentioned above. The estimated effects on individuals' labor incomes are conditional to being employed. Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Appendix A

(a) Luck Gpa



(b) Luck Diploma

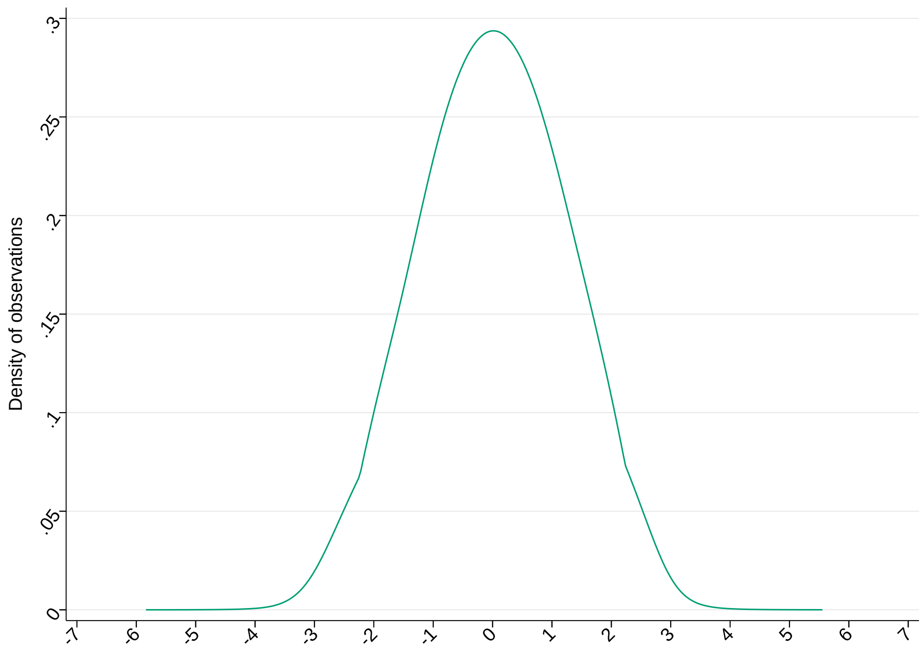


Figure A1 – Distribution of Students' GPA Luck and Diploma Luck

NOTE: The table refers to the same sample as Table 1. Figure A1a plots the distribution of our measure of GPA luck. Figure A1b plots the distribution of our measure of diploma luck.

Table A1 – Effect of Luck on Firm Characteristics

	Outcomes					
	First job			Job 9 years after HS		
	Public sector	Size	Coworkers with HE (%)	Public sector	Size	Coworkers with HE (%)
Luck GPA	-0.0009	-4.7713	0.0025**	0.0004	-8.9870	0.0024**
	(0.0020)	(10.8346)	(0.0011)	(0.0021)	(14.1715)	(0.0011)
Luck diploma	0.0004	6.3898	-0.0013	-0.0001	17.8601	-0.0005
	(0.0020)	(10.8268)	(0.0011)	(0.0021)	(14.1348)	(0.0012)
Mean	0.376	605.023	0.515	0.397	740.275	0.569
N	107461	107461	107461	96734	96734	96734

NOTE: The table refers to the same sample as Table 1, restricted to individuals who have ever been employed (columns 1 to 3), or to individuals who are employed nine years after taking the high school exit exams. Each column corresponds to a specific regression, and reports the estimated impacts of our two measures of luck –luck GPA and luck diploma– on the dependent variable mentioned above. Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table A2 – Effect of Luck on High School Outcomes, Heterogeneity by Gender

	Outcomes				
	Exam grades in 3rd year	High School GPA in 3rd year (log)	Overall HS GPA in 3rd year	On time HS diploma (log)	Ever HS diploma
Panel A: Girls					
Luck GPA	0.0940***	0.0182***	0.0090***	0.0058***	0.0020**
	(0.0031)	(0.0009)	(0.0014)	(0.0014)	(0.0009)
Luck diploma	0.0362***	0.0049***	0.0030**	0.0083***	0.0019**
	(0.0030)	(0.0009)	(0.0014)	(0.0014)	(0.0009)
Mean	0.115	0.153	0.155	0.854	0.945
N	71335	71335	53541	71335	71335
Panel B: Boys					
Luck GPA	0.1049***	0.0204***	0.0146***	0.0030	0.0014
	(0.0035)	(0.0010)	(0.0017)	(0.0019)	(0.0013)
Luck diploma	0.0403***	0.0058***	0.0033*	0.0174***	0.0053***
	(0.0034)	(0.0010)	(0.0017)	(0.0019)	(0.0013)
Mean	-0.140	-0.141	-0.122	0.779	0.911
N	58582	58582	43553	58582	58582

Note: The table report similar results as Table 3 separately on the sub-sample of girls (Panel A), and boys (Panel B). Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table A3 – Effect of Luck on Labor Market Outcomes, Heterogeneity by Gender

	Outcomes			
	Ever employed	First job annual labor income	Employed 9 years after HS	Annual labor income 9 years after HS
Panel A: Girls				
Luck GPA	0.0022	0.0089**	0.0015	0.0077**
	(0.0019)	(0.0042)	(0.0022)	(0.0033)
Luck diploma	0.0026	0.0030	0.0015	0.0020
	(0.0019)	(0.0042)	(0.0022)	(0.0035)
Mean	0.836	12.290	0.753	12.619
N	71335	59653	71335	53744
Panel B: Boys				
Luck GPA	-0.0012	0.0085	0.0006	0.0057
	(0.0023)	(0.0060)	(0.0025)	(0.0044)
Luck diploma	-0.0029	0.0027	-0.0028	0.0026
	(0.0023)	(0.0061)	(0.0026)	(0.0045)
Mean	0.816	12.147	0.734	12.685
N	58582	47808	58582	42990

NOTE: The table report similar results as Table 5 separately on the sub-sample of girls (Panel A), and on the sub-sample of boys (Panel B). Each regression includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.

Table A4 – Effect of Luck on Education and Labor Market Outcomes, Robustness Tests

	Outcomes										
	Exam grades in 3rd year	High school GPA in 3rd year	Overall HS GPA in 3rd year	On time HS diploma	Ever HS diploma	Any college	Minimum peer GPA in college	Ever employed	First job annual labor income (log)	Employed 9 years after HS	Annual labor income 9 years after HS (log)
Panel A: Controls for Students' Baseline Characteristics Selected by Double Lasso											
Luck GPA	0.0988***	0.0192***	0.0114***	0.0046***	0.0018**	0.0001	0.0100**	0.0008	0.0088**	0.0012	0.0067**
	(0.0025)	(0.0007)	(0.0011)	(0.0011)	(0.0008)	(0.0009)	(0.0049)	(0.0015)	(0.0035)	(0.0017)	(0.0027)
Luck diploma	0.0382***	0.0053***	0.0032***	0.0124***	0.0034***	0.0017**	-0.0050	0.0002	0.0035	-0.0004	0.0026
	(0.0023)	(0.0007)	(0.0011)	(0.0012)	(0.0007)	(0.0008)	(0.0051)	(0.0015)	(0.0036)	(0.0017)	(0.0028)
Panel B: P-values for Luck GPA Computed with a Permutation Test											
P-values for Luck GPA	0.000	0.000	0.000	0.000	0.018	0.879	0.029	0.541	0.007	0.428	0.011
P-values for Luck diploma	0.000	0.000	0.003	0.000	0.000	0.070	0.251	0.886	0.310	0.769	0.310
Panel C: Non-winsorized Measures of Luck											
Luck GPA	0.0963***	0.0187***	0.0112***	0.0043***	0.0018**	0.0001	0.0091*	0.0008	0.0082**	0.0011	0.0064**
	(0.0024)	(0.0007)	(0.0010)	(0.0011)	(0.0008)	(0.0009)	(0.0049)	(0.0014)	(0.0035)	(0.0017)	(0.0026)
Luck diploma	0.0365***	0.0050***	0.0031***	0.0122***	0.0033***	0.0016**	-0.0050	0.0000	0.0028	-0.0005	0.0024
	(0.0023)	(0.0007)	(0.0010)	(0.0012)	(0.0007)	(0.0008)	(0.0050)	(0.0015)	(0.0035)	(0.0017)	(0.0027)
Panel D: Excluding Students with a Failing Course Grade											
Luck GPA	0.0928***	0.0183***	0.0100***	0.0082***	0.0029***	0.0005	0.0103**	0.0010	0.0093**	0.0014	0.0059**
	(0.0026)	(0.0007)	(0.0011)	(0.0011)	(0.0006)	(0.0008)	(0.0052)	(0.0015)	(0.0037)	(0.0017)	(0.0028)
Luck diploma	0.0446***	0.0062***	0.0041***	0.0085***	0.0018***	0.0010	-0.0060	0.0001	0.0016	-0.0006	0.0023
	(0.0024)	(0.0007)	(0.0011)	(0.0011)	(0.0006)	(0.0008)	(0.0053)	(0.0015)	(0.0036)	(0.0018)	(0.0029)
Mean	0.103	0.164	0.155	0.875	0.964	0.947	-1.859	0.827	12.290	0.749	12.667
N	119385	119385	89601	119385	119385	119385	104534	119385	98720	119385	89435

NOTE: The table refers to the same sample as Table 1. The table report similar results as Table 3, 4 and 5. Panel A replicates the main analyses with a restricted set of baseline demographic controls selected by double lasso. Panel B report p-values from permutation tests for luck GPA or luck diploma. Panel C replicates the main analyses with non-winsorized measures of luck. Panel D focuses on a restricted sample which excludes the students who obtained a failing course grade. Each regression –expect for Panel A– includes a rich set of baseline demographic controls, as well as high school and year fixed effects. Standard errors clustered at the high school-by-year level are in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%.