Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

The Value of a High School GPA

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

Abstract

This paper provides novel evidence on the causal effect of high school Grade Point Average (GPA) on the human capital development and labor market trajectory of individuals. Causal identification is achieved by exploiting a unique feature of the Norwegian education system that produces exogenous variation in GPA among high school students. We find little effect on the number of completed years of higher education, but significant effects on the number and quality of higher education programs available to students after high school. Most importantly, we find persistent effects on students' long-run labor market outcomes, most notably market wage.

JEL Codes: I24, I26, J24, J31, D63

Keywords: Returns to education, high school GPA, high-stakes exams.

---

*Landaud: CNRS and CY Cergy Paris University (fanny.landaud@cyu.fr). Maurin: Paris School of Economics (eric.maurin@ens.fr). Willage: University of Colorado - Denver, and NBER (barton.willage@ucdenver.edu). Willén: Norwegian School of Economics, UCLS, CESifo, and IZA (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 ADRES 2022, ESEM 2022, EALE 2022, the FAIR Workshop on the Economics of Education, CREST, NTNU, Texas A&M University, and CY Cergy Paris University for helpful comments and discussions. This paper was first circulated under the title "Getting lucky: the Long Term Consequences of Exam Luck"

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

1 Introduction

High school GPA is one of the most frequently used metrics for identifying academic achievement, and it likely plays a pivotal role in determining an individual's future education and labor market outcomes. It is used in the admissions process to college, by grant offices when awarding scholarships, by career offices when allocating internships, and by employers when recruiting graduates. However, a lack of exogenous variations in GPA has prevented researchers from identifying the true value of a high school GPA.

This paper provides some of the first evidence on the impact of high school GPA on an individual's education and labor market outcomes, exploiting a unique feature of the Norwegian education system that produces exogenous variation in GPA among high school students with identical abilities. In Norway, high school GPA is based not only on the grades obtained in high school courses, but also on a set of randomized exams taken throughout high school. These exams are conducted in subjects chosen randomly by the municipality among the courses that the student takes, and are announced just days prior to the exams. Each exam counts as much as a course grade when calculating a student's GPA. This feature induces exogenous variation to a student's GPA depending on whether the subjects drawn for the exams correspond to the student's academic strengths or weaknesses.

We use the random exam feature of the Norwegian high school system to construct an instrument, GPA luck, which measures whether the students are allocated to exams in subjects corresponding to their academic strengths or weaknesses. This enables us to identify the causal effects of high school GPA. The instrument we construct is consistent with the general principles developed by Borusyak and Hull (2020), and our paper can be viewed as an empirical application of their method for the case where exposure of individuals to random shocks varies according to some of their pre-determined characteristics. We use the instrument to identify the impact of GPA on the student's long-term education and labor market outcomes.

Exploiting rich administrative data covering the universe of Norwegian senior high school students, we first confirm that our GPA luck measure is a strong predictor of high school GPA. We also demonstrate that our measure of GPA luck is unrelated to a rich set of baseline characteristics that may affect students' outcomes. We then show that GPA luck has little effect

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

on the number of years of higher education completed after high school, but significant effects on the number and selectivity of higher education programs that are available to students after high school. Finally, we show that GPA luck has persistent effects on students' long-run labor market outcomes. In particular, improvements in GPA luck produce significant increases in market wages eight years after the exams. Using GPA luck as an instrumental variable, we find that a one standard deviation increase in high school GPA generates a 23% increase in individuals' earnings eight years after the end of high school. The effect is most pronounced for students who pursue a short university education after high school (3 years or less). For this group of students, GPA luck increases the probability of accessing the most lucrative fields of study (such as Business and Science) and reduces the probability of ending up in the least lucrative ones (such as Humanities and Teaching). The fact that the wage effect of GPA luck is particularly strong for this group of students is consistent with the literature on the wage effects that field of study can have after high school (Kirkeboen, Leuven and Mogstad, 2016).

The main contribution of our paper is to provide some of the first evidence in the literature on the causal effect of high school GPA on the human capital development and labor market trajectory of individuals. It has long been asserted that high school GPA is a strong predictor of later-in-life outcomes,[1] and our paper is among the first to identify the extent to which this is indeed a cause-and-effect relationship.

We also contribute to the rich literature exploiting high school test score thresholds in regression discontinuity designs to examine the effect of high school diplomas, college admissions or scholarships (e.g., Clark and Martorell, 2014; Cohodes and Goodman, 2014; Goodman, 2008; Hastings, Neilson and Zimmerman, 2013; Heinesen, 2018; Hoekstra, 2009; Ockert, 2010; Ost, Pan and Webber, 2016; Zimmerman, 2014). We complement these papers by providing evidence on the overall effect of high school GPA on long-run educational and labor market outcomes. We also advance this literature by developing an instrument that affects high school student performance at all levels of the achievement distribution and, consequently, by providing estimates derived from the entire achievement distribution.

[1]E.g., Allensworth and Clark (2020); Black, Cortes and Lincove (2016); Galla et al. (2019); Geiser and Santelices (2007).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

Our paper also helps us better understand why high school GPA plays such an important role long after high school. This is not because it leads to additional years of university studies, but because it increases the possibility of enrolling in university programs that best corresponds to students' specific aspirations/talents. Whatever the field of study, a higher GPA also increases the probability of accessing more selective university programs, which can have a significant effect on wages; especially when the selectivity of the program attended by students is used by employers as a signal of their ability (e.g., Lang and Siniver, 2011; MacLeod et al., 2017).

Finally, there is a well-established literature on the role of luck in determining an individual's education and labor market success (e.g., Audas, Barmby and Treble, 2004; Bertrand and Mullainathan, 2001; Frank, 2016; Jenter and Kanaan, 2015). The most studied form of luck in this literature is the birth lottery, which allocates genes and early social environments to individuals (e.g., Black and Devereux, 2011; Mogstad and Torsvik, 2021). More related to our study, however, are the papers that have looked at exam luck in terms of the external conditions prevailing on test days, whether in terms of outdoor temperature, time of the day, or the presence of pollutants or pollen in the atmosphere (e.g., Amanzadeh, Vesal and Ardestani, 2020; Bensnes, 2016; Ebenstein, Lavy and Roth, 2016; Gaggero and Tommasi, 2020; Garg, Jagnani and Taraz, 2020; Park, 2022). We contribute to this literature by demonstrating the importance of another form of luck, linked to the content of the exams themselves, which always have an element of randomness impossible to anticipate by students.

2 Background

2.1 The Norwegian Education System

The Norwegian education system consists of 10 years of compulsory school from age 6, followed by 3 to 4 years of voluntary high school. Approximately 95% of students enroll in high school and 80% of each cohort ends up with a high school diploma. Education is free at all levels.

High school splits into an academic and a vocational track. Approximately 50% choose each track. We focus on students in the academic track. The reason is that the random exams for vocational track students depend on the chosen vocational field of study, and there is therefore

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

little random variation in the exams that these students face.

Several universities and colleges offer higher education in Norway. Admission is conditional on graduating from high school. The Norwegian Universities and Colleges Admission Service coordinates the admission process. Students apply to specific fields of study and universities, and if a school or a program is oversubscribed, students are assigned almost exclusively based on high school GPA. The more demand for a specific program (that is, a field of study within a given university), the higher the GPA required to gain admission to that program.

2.2 High School GPA and Randomized Exams

In Norway, high school GPA depends on a combination of course grades from teachers and exam grades. The exams take place at the end of each school year, the exam subjects are chosen randomly for each individual student, and the subjects are announced less than a week prior to the exam.[2] In terms of exam structure, students take between five and six exams throughout high school. In the first year, 20% of students are randomly selected 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 three written exams and one oral exam.

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 an effect on dropout rates (see e.g., Andresen and Løkken, 2020; Hvidman and Sievertsen, 2021). To avoid sample selection problems, we focus exclusively on the exams in the third and final year.

Before 2008, the exams in the final year of high school consisted of two written exams in Norwegian, one written exam in a randomly chosen subject and one oral exam in a randomly chosen subject. Since 2008, the final year exams consist of one written exam in Norwegian, two written exams in randomly chosen subjects and one oral exam in a randomly chosen subject.[3]

[2]Even if the delay is short, students can use these few days to prepare for exams (Bensnes, 2020). This can mitigate the impact of being lucky (or unlucky).

[3]Senior students take an average of nine courses, of which an average of five are included in the written exam draw. In addition to Norwegian, the most common written exam subjects

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

The randomization of students to subjects and types of tests is delegated to the municipality. While the written exams are designed and graded at the national level, the oral exams are designed and graded at the local level. In our analysis, we focus on the variation in GPA generated exclusively by the written exams, which removes any endogeneity issue driven by local exam designs.

An exam grade counts as much towards GPA as a course grade. High school GPA consists of the average grade of all of a given student's course grades and randomized exam grades in high school. Exams and courses are graded on a scale from 1 (worst) to 6 (best), where 1 constitutes a failing grade.

It should be noted that successful graduation from academic high school, and eligibility for higher education, requires that the student passes all high school courses. However, there is an exception to this rule if 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 receives a high school diploma.

The vast majority (92%) of senior students do not fail a course. For these students, a good exam draw unambiguously corresponds to a draw that matches their academic strengths. For the small fraction (8%) of senior students who failed a course during the year, the effects of the draw are more complex.[4] We focus on the 92% of senior students who did not fail a course during the year and for whom the definition of a good exam draw is unambiguous.

3 Data, Variables and Samples

3.1 Data Registers

Our data come from administrative registers covering all Norwegian residents who were enrolled in the final year of high school between 2005 and 2009. We follow these individuals over time

are English, mathematics, chemistry, biology, social sciences, physics, law, modern history, political science, and management.

[4]If they draw exams in their strong subjects, this potentially increases their GPA, but not their chance of graduating and of going to college. To have a chance at graduating and going to university, they must draw the subject they failed during the year and pass the exam.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

and across registers, such that we can construct a longitudinal panel covering the universe of students and their demographic, education, labor, and family background information.

In terms of education data, we have information on students' high school GPA, diploma status, and qualification for higher education. The data also includes details on all high school courses students take, the grades they receive, the courses they are randomized into taking exams in, and the exam grades. Additionally, information on higher education enrollment, college major choice, and degree completion is available.

From these data, we can construct a variable (denoted $Share_{i,t}$) identifying the proportion of university programs for which each student is eligible. Specifically, for each program s available in year t, we identify the minimum GPA (denoted $Min\_Gpa_{s,t}$) of the students enrolling in that program. For each student i, the variable $Share_{i,t}$ is equal to the proportion of programs whose $Min\_Gpa_{s,t}$ is below the GPA of student i. We also construct a variable measuring the selectivity (denoted $Select_{i,t}$) of the programs in which students enroll. To define $Select_{i,t}$, we ranked all higher education programs in year t based on their minimum student GPA ($Min\_Gpa_{s,t}$). If $Min\_Gpa_{s(i),t}$ denotes the minimum GPA of the first program in which student i enrolls after high school, the selectivity level of student i's first enrollment is defined as the percentile rank of $Min\_Gpa_{s(i),t}$, that is: $Selectivity\_level_{i,t} = Percent\_rank(Min\_Gpa_{s(i),t})$.

With respect to labor market information, we have detailed information on income and 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 is defined based on the individual's status in the employment register. In our analysis, we focus on employment and income eight years after high school graduation. Data limitations prevent us from exploring even longer-run outcomes.

Concerning background characteristics, we have information on compulsory school GPA, age, sex, and municipality of residence. We can also link students to their parents and collect information on parents' age, educational attainments, and earnings.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

3.2 The GPA Luck Instrument

In Norway, students' GPA at the end of high school depends not only on the course grades given by teachers, but also on the results obtained at the randomly drawn end-of-year exams. The draw produces an exogenous shock that is more or less favorable for students' GPA depending on whether the subjects drawn correspond to the students' academic strengths or weaknesses. We construct an instrument that uses the random nature of the end-of-year examinations to identify the causal effect of students' GPA on their outcomes.

We begin by constructing a measure of the score that student i can expect on an end-of-year exam in subject s if that subject is drawn (denoted $Exam^e_{i,s}$). Specifically, we define $Exam^e_{i,s}$ as the average score obtained on the end-of-year exam in subject s by students (other than i) who attended the same high school as i in the same year, earned the same teacher assessment in the course on subject s as i and were randomly assigned to an exam in subject s.[5]

Based on a complete set of predictions for exams taken at the end of senior year, we construct a measure of the expected contribution of these exams to the final GPA of student i if the student is randomly assigned to a specific combination c of exams at the end of senior year:

$$Z_{i,c} = \frac{1}{N} \sum_{s \in c} Exam^e_{i,s}, \tag{1}$$

where N represents the total number of senior year grades used to calculate i's GPA, that is $N = S + K$ where K is the number of courses and S the number of exams that student i takes. The sum on the right-hand-side represents the expected number of grade points from the randomly-assigned exams.

Denoting $c(i)$ the combination of subjects actually assigned to student i, the value taken by $Z_{i,c}$ when $c = c(i)$ is a predictor of the contribution of exams to student i's final GPA. We denote this predictor GPA_luck_i. This predictor varies randomly from one student to another, depending on subjects assignment, but this alone does not necessarily make it a good instrument for identifying the impact of the GPA on subsequent outcomes (see e.g., Borusyak and Hull,

[5] If $E_{i,s}$ represents the set of all these students, $Exam^e_{i,s}$ is the ratio between $\sum_{j \in E_{i,s}} Exam_{j,s}$ and the cardinal of $E_{i,s}$, where $Exam_{j,s}$ denotes the result of student j at end-of-year exam in subject s.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

2020). The problem comes from the fact that students' exposure to more or less favorable shocks (i.e., to stronger or weaker average $Exam^e_{i,s}$) is not randomly distributed, but depends on their initial academic level, as measured for example by the grades obtained during the school year. In short, randomizing exams is not exactly the same as randomizing the scores students receive on these exams.

As shown by Borusyak and Hull (2020), there is a simple way to neutralize the potential endogeneity bias that affects the use of an instrument constructed as GPA_luck_i. Specifically, it is sufficient to recenter GPA_luck_i, that is to consider its deviation with respect to the average value of $Z_{i,c}$ (denoted $\overline{GPA\_luck_i}$) across the set of all possible counterfactuals c,

$$\widetilde{GPA\_luck_i} = GPA\_luck_i - \overline{GPA\_luck_i}. \tag{2}$$

For each student i, $\overline{GPA\_luck_i}$ is calculated by assigning to each possible exam combination c a probability equal to the probability of occurrence of c actually observed in the student population.[6] We will show below that the recentered instrument $\widetilde{GPA\_luck_i}$ is both a strong predictor of students' GPA and is uncorrelated with students' predetermined characteristics. By recentering GPA_luck_i, we have eliminated its potentially endogenous component and isolated its purely random component.

As a robustness check, we examine an alternative version of the instrument constructed by redefining $Exam^e_{i,s}$ to represent the teacher assessment obtained by i in subject s. As discussed below, the first stage effect on high-school GPA becomes weaker, but remains highly statistically significant, when we use this version of the instrument.

3.3 Sample Selection and Descriptive Statistics

Our data include all full-time high school students enrolled in the final year of the academic high school for the first time between the academic year 2005-2006 and 2009-2010. We exclude the small share (less than 8%) of students who received a grade of 1 during the school year since the relationship between high school GPA and high school graduation is not monotonic

[6]Note that, in this formula, mandatory exams in Norwegian are subtracted out from $\widetilde{GPA\_luck_i}$, since such exams have a probability of occurrence of 1.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

for these students. This provides us with a sample of about 90,000 students. Table A1 provides descriptive statistics on all individuals in our sample. Figure A1 shows the distribution of our recentered GPA luck instrument $\widetilde{GPA\_luck}_i$. As expected, it follows a Gaussian-type distribution, evenly distributed around zero.

Table A2 demonstrates that there is very little correlation between the recentered GPA luck instrument and observed student characteristics measured in pre-assignment years. Furthermore, consistent with Borusyak and Hull (2020), the recentered instrument is uncorrelated with the expected instrument $\overline{GPA\_luck}_i$. When we regress our GPA luck variable on all the pre-assignment characteristics considered in Table A2, a conventional F-test cannot reject that the estimated coefficients are jointly equal to zero (P-value = 0.72).

4 Results

4.1 GPA luck and Students' Outcomes

Table 1 shows the results of regressing students' education and labor market outcomes on our measure of GPA luck. To be fully consistent with Borusyak and Hull (2020), the table shows the results of regressing student outcomes on the non-recentered instrument using the recentered instrument ($\widetilde{GPA\_luck}_i$) as an instrumental variable. This approach provides estimates that are similar to the OLS estimates obtained by directly regressing student outcomes on the recentered instrument.

Our regression model includes a full set of high school and year fixed effects as well as a rich set of demographic controls (including students' sex, age, average high school course grade, average middle school GPA, parents' age, years of schooling and log earnings). Standard errors are clustered at the high school-by-year level. In Appendix Table A3, we show that our results are robust to the use of a double lasso procedure for selecting control variables, and that the precision of our main estimates remains unchanged when we compute standard errors using a wild bootstrap procedure. We also provide results of permutation tests for the main first-stage and reduced-form regression models.

In Panel A of Table 1, we study the effect of GPA luck on students' high school outcomes.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

To facilitate the interpretation of the results, we divided our measure of GPA luck by its standard deviation over the sample, so that a one unit increase in GPA luck corresponds to a 1 SD improvement. The results confirm that GPA luck has a statistically significant effect on students' exam grades (first column), on their high school GPA (second column), on their probability of on-time graduation (third column), and on ever receiving a high school diploma (fourth column). Specifically, a one SD improvement in GPA luck leads to 14% of a standard deviation increase in exam grades, to 2.5% of a SD increase in GPA, to a 1.4 percentage points increase in the probability of on-time graduation, and to a 0.4 percentage point increase in the probability of ever receiving a high school diploma. The smaller effect on the probability of ever receiving a diploma suggests that many students who graduate on time due to exam luck would still have graduated a year later (after retaking classes) if they had not been lucky.

In Panel B of Table 1, we study the effect of GPA luck on students' higher education outcomes. The first column shows that GPA luck has no effect on the probability of ever going to university. This finding is consistent with the very small effect on the probability of ever graduating from high school documented in the fourth column of Panel A.

Based on this initial result, the remainder of the analysis in Panel B focuses on the subsample of students who went to university. First, we explore the effect of GPA luck on the share of higher education programs that the students qualify for (second column) and on the selectivity of the first higher education program in which they enroll (third column). The results show that a one SD increase in GPA luck is followed by a 0.1 percentage points increase in the share of higher education programs available to the students and by a 0.24 percentile rank increase in the selectivity of the higher education program in which they enroll. These results confirm that GPA luck broadens the range of possible choices for higher education enrollment and that students take advantage of this broader choice set to "upgrade" their college quality through enrollment into more selective programs and universities. We have also explored the impact of GPA luck on the likelihood of enrolling in each of the main fields of study (medicine, law, business, engineering, etc.). The estimated impacts are small and generally not statistically significant (see Appendix Table A4). However, as we will show below, these results mask significant heterogeneity among students depending on whether they pursue short

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

higher education programs (three years or less) or long higher education programs (more than three years). The result in the fourth column of Panel B demonstrates that there is no impact of GPA luck on the number of years completed in higher education. In the Online Appendix, we further show that the probability of completing more than three years of higher education (the Bachelor level) is not affected by GPA luck (see Table A5). GPA luck thus allows students to study in more selective programs, but this does not translate in changes in quantity of education. In particular, the quality upgrading that GPA luck contributes to is not offset by a potential reduction in educational attainment due to admission into more difficult programs.

In Panel C of Table 1, we study the effect of GPA luck on students' labor market outcomes. The first column shows that GPA luck has no significant effect on student's probability of ever having been employed. In light of this lack of employment effect, in the second column of Panel C we constrain our sample to the group of students who held at least one job in their lifetime, exploring the impact of GPA luck on the annual labor income at the first job the students secure. The result suggests that GPA luck has a sizable impact on the annual labor income at the first job the students secure.

Our data enables us to follow students up to eight years after they have taken their third-year high school exams. Using this information, the third column of Panel C shows that GPA luck has no significant effect on the probability of being employed eight years after the exams. The fourth column of Panel C therefore zooms in on the students who were employed 8 years after the exams, and it shows a significant effect on annual labor income. The magnitude of the effect on earnings is about the same eight years after the exams as it was at labor market entry.[7]

[7]In Table A3 in the online Appendix, we check that the effect of GPA_luck on annual labor income is very similar when we use the version of the recentered instrument constructed using teacher assessments (rather than the exam results of observably similar peers) to predict exam scores. Also, consistent with Borusyak and Hull (2020), the estimated effect of GPA_luck is unchanged when we no longer use an instrument, but simply add the *expected instrument* $\overline{GPA\_luck_i}$ as a control.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

4.2 Heterogeneity

Since GPA luck has no effect on the number of years completed in higher education, we examine whether the effect of GPA luck is different for the small group of students who never goes to university, for the group who completes three years or less of higher education, and for the group who completes more than three years of higher education (see Table 2). Interestingly, the first-stage effects are similar for all three groups, but the wage effects are mainly driven by students with three years or less of higher education. For these students, GPA luck leads to a significant drop in enrollments in Humanities or Teaching, and an increase in enrollments in more lucrative fields of study such as Science or Business (Appendix Table A4). This pattern of reallocation across programs is not observed for students who complete more than three years of higher education. While speculative, we believe that the lack of a program substitution effect contributes to explaining why the wage effect of GPA luck is weaker for them.

According to Kirkeboen, Leuven and Mogstad (2016), enrolling in Science (or Business) when the next best choice is Humanities (or Teaching) results in a two- to threefold increase in early wages. Assuming that the reallocation across fields of study induced by GPA luck also generate such wage multiplications, this reallocation can explain a large part of the wage effect observed for students with three or fewer years of higher education.[8]

In the Online Appendix, we also explore heterogeneity by student gender, parental income and parental education (Table A6). The first-stage effects of GPA luck on GPA are almost exactly the same in the different subgroups. The wage effects observed eight years after the exams also does not vary much by gender. However, wage effects tend to be higher for students whose parents are better educated or wealthier, perhaps because these families are better informed and

[8]It should be noted, however, that the estimates obtained by Kirkeboen, Leuven and Mogstad (2016) are derived from a sample consisting only of high school students who apply for at least two broad fields of study (out of ten), where the most preferred field needs to have an admission cutoff, and the next best alternative must have a lower cutoff (or no binding cutoff), a sample that in total represents less than 25% of university applicants. Our analyses use a much larger estimation sample (with no constraint on the choice of fields of study), so that the results obtained by Kirkeboen, Leuven and Mogstad (2016) do not necessarily apply to all our observations.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

better able to take advantage of a higher GPA to gain access to more lucrative fields of study.

We also examine effect heterogeneity by initial academic level, as measured by teacher assessments. When we stratify the sample based only on two groups (above and below the median of teacher assessments, Table A6), the first-stage effects as well as the wage effects of GPA luck are similar for both groups. However, when we stratify across five groups (Table A7), the wage effect tends to be stronger for the central groups around the median and much smaller for the groups at the tails of the distribution. As admission cut-offs tend to be concentrated around the median, this result is consistent with the idea that GPA luck impacts wage mainly through an increase in the number of accessible programs.

4.3 The Causal Effects of High-School GPA: IV Estimations

We have shown that GPA luck has a significant impact on the GPA of high school students but no impact on their probability of holding a job. Based on these results, we focus on the sample of former high school students who are employed after university or eight years after the exams (i.e., the same samples as Panel C of Table 1). Using this sample, we analyze the impact of high school GPA on wages using the recentered GPA luck variable as an instrument in a standard 2SLS approach. The identifying assumptions are that the recentered instrument has a direct effect on students' high school GPA (relevance criterion) and that it influences students' subsequent labor market outcomes only through its impact on their high school GPA (exclusion restriction). Insofar as we do not rule out the possibility that the effect of high school GPA may vary from one student to another, we must also assume that there are no students for whom an increase in exam luck might translate into lower exam results and lower GPA (monotonicity assumption). This monotonicity assumption is necessary if the estimated effect is to be interpreted as an average of heterogeneous causal effects.

The validity of the relevance criterion has already been established above. In terms of the exclusion restriction, this assumption would be violated if GPA luck in itself had a direct effect on subsequent labor market outcomes. It could be, for example, that students who experience a poor exam draw become more prone to depression and discouragement, which could harm student outcomes by reducing their ability to persevere. However, if this were the case, a poor

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

exam draw would also likely reduce students' propensity to repeat their final year of high school (as well as their probability of ever graduating from high school and entering university) if they fail the first attempt, which is not what we observe.

Finally, as pointed out by Angrist and Pischke (2009) as well as Frandsen, Lefgren and Leslie (2023), the fact that the first-stage effect of GPA luck has the same sign for all observable subgroups (as defined by gender, prior academic performance or social background) is consistent with the monotonicity assumption. The fact that the magnitude of the first-stage effect is almost exactly the same across all subgroups further suggests that the distribution of compliers across the different subgroups does not differ from that of the entire population.[9]

The results from our IV analysis are shown in Table 3. The table shows that high school GPA has a large impact on labor market earnings. Specifically, a one standard deviation increase in high school GPA generates a 23.2% increase in students' earnings at entry into the labor market and a 22.6% increase in earnings eight years after finishing high school.

To further facilitate the interpretation of the 2SLS results, Table 3 also reports the result from an OLS estimation of the same parameters. The OLS impact of a one SD increase in GPA on earnings is approximately 10% for entry wages and about 5% for wages earned eight years after high school. The difference between the IV and OLS estimates are not statistically significant, but they nonetheless suggest a negative correlation between unobserved determinants of high school graduates' performance on the labor market and the unobserved determinants of their high school GPA. Such an endogeneity bias could arise if, for example, it is easier to obtain a good GPA at a bad school (where teachers more easily give out good grades).

Another possible explanation for the differences between IV and OLS estimates is that the two estimates represent different weighted averages of marginal treatment effect. It is possible to explore this hypothesis by following Ishimaru (2022), and decomposing the gap between IV

[9]We also implement a test of the monotonicity assumption based on the results in Angrist and Imbens (1995) by checking that the CDF of the treatment given that the instrument is positive does not cross the CDF of the treatment given that the treatment is non-positive (see Appendix Figure A2). Comfortingly, the distribution of the gap between the two CDFs is very similar to the overall distribution of the GPA itself, consistent with the idea that the distribution of compliers across initial academic levels is similar to that of the overall student population.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

and OLS into the difference due to endogeneity bias and the difference due to the fact that IV and OLS do not place the same weights on students with different observable characteristics or with different levels of GPA. This approach leads to the conclusion that the gap between IV and OLS can be explained almost entirely by endogeneity bias (see Appendix Table A8).[10]

Estimated effects of student high school GPA on wages are economically large, consistent with existing evidence on how exogenous shifts in high school student performance can improve earnings (e.g., Ebenstein, Lavy and Roth, 2016, Kirkeboen, Leuven and Mogstad, 2016). For example, using pollution shocks as a source of identification, Ebenstein, Lavy and Roth (2016) show that a one SD increase in matriculation exam scores at the end of high school in Israel leads to an increase of over 20% in early career wages, close to our own estimates.

There are many mechanisms to explain the large effect of high school GPA on wages, and it is beyond the scope of this paper to explore them all in detail. The most direct channel is the increase in the number of fields of study through which students can pursue higher education. For the group of students that are applying to several different fields of study and are able to enter selective institutions for each of them, Kirkeboen, Leuven and Mogstad (2016) demonstrate that entering one's preferred field of study rather than the next best one can lead to a doubling or tripling of early career wages, and that these effects are highly underestimated by standard OLS estimators. Whether students apply to several fields of study or just one, our findings also show that a higher GPA increases the possibility of gaining access to more selective institutions, which can not only expose them to better professors and peers, but can also have wage effects when the selectivity of the program attended by students is used by employers as a signal of their ability.[11]

[10] It should also be noted that, in our context, IV estimates give more weight to students for whom the distribution of the instrument has the highest variance across all possible exam draws. Following Borusyak and Hull (2020), it is possible to obtain more conventional average treatment effects by rescaling the instrument by this variance. This approach leads to a rescaled IV estimate whose magnitude (31%) is not significantly different from that of the non-rescaled estimate (see Appendix Table A9).

[11] See for example Lang and Siniver (2011) or MacLeod et al. (2017). For further evidence on the signaling effect of college outcomes see, e.g., Hansen, Hvidman and Sievertsen (2021);

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

5 Conclusion

This paper provides some of the first evidence on the causal effect of high school GPA on human capital and labor market outcomes. Findings indicate that an improved GPA has little impact on receiving a high school diploma and years of higher education but significantly boosts enrollment in programs aligned with individual aspirations, leading to substantial increases in labor income. Notably, this effect is most pronounced for those opting for a short university education. For this group, an improve GPA generates an increase in enrollment in lucrative fields and reduced enrollment in less lucrative ones.

Our paper also contributes to the literature on the role of luck in the education process. While most countries do not have random assignment of students to examinations, the exact questions student experience are random. This paper therefore speaks more generally to the consequences of having a bad draw on the content of exams.

Kessler, Low and Sullivan (2019); Quadlin (2018).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

References

Allensworth, Elaine M, and Kallie Clark. 2020. "High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools." *Educational Researcher*, 49(3): 198–211.

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.

Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90(430): 431–442.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion. Economics Books*, Princeton University Press.

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

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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*, , ed. David Card and Orley Ashenfelter 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.

Borusyak, Kirill, and Peter Hull. 2020. "Non-Random Exposure to Exogenous Shocks: Theory and Applications." National Bureau of Economic Research Working Paper 27845.

Clark, Damon, and Paco Martorell. 2014. "The Signaling Value of a High School Diploma." *Journal of Political Economy*, 122(2): 282–318.

Cohodes, Sarah, and Joshua Goodman. 2014. "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy." *American Economic Journal: Applied Economics*, 6(4): 251–285.

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.

Frandsen, Brigham, Lars Lefgren, and Emily Leslie. 2023. "Judging Judge Fixed Effects." *American Economic Review*, 113(1): 253–77.

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

Gaggero, Alessio, and Denni Tommasi. 2020. "Time of Day, Cognitive Tasks and Efficiency Gains." Institute of Labor Economics (IZA) IZA Discussion Papers 13657.

Galla, Brian M, Elizabeth P Shulman, Benjamin D Plummer, Margo Gardner, Stephen J Hutt, J Parker Goyer, Sidney K D'Mello, Amy S Finn, and Angela L Duckworth. 2019. "Why high school grades are better predictors of on-time college graduation than

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

are admissions test scores: The roles of self-regulation and cognitive ability." *American Educational Research Journal*, 56(6): 2077–2115.

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.

Geiser, Saul, and Maria Veronica Santelices. 2007. "Validity of High-School Grades in Predicting Student Success beyond the Freshman Year: High-School Record vs. Standardized Tests as Indicators of Four-Year College Outcomes." *Center for studies in higher education, Research & Occasional Paper Series: CSHE. 6.07.*

Goodman, Joshua. 2008. "Who merits financial aid?: Massachusetts' Adams Scholarship." *Journal of Public Economics*, 92(10-11): 2121–2131.

Hansen, Anne Toft, Ulrik Hvidman, and Hans Henrik Sievertsen. 2021. "Grades and Employer Learning." *IZA Working Paper 14200.*

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman. 2013. "Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile." National Bureau of Economic Research Working Paper 19241.

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.

Ishimaru, Shoya. 2022. "Empirical Decomposition of the IV–OLS Gap with Heterogeneous and Nonlinear Effects." *The Review of Economics and Statistics*, 1–45.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

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

Kessler, Judd B, Corinne Low, and Colin D Sullivan. 2019. "Incentivized resume rating: Eliciting employer preferences without deception." *American Economic Review*, 109(11): 3713–44.

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.

Lang, Kevin, and Erez Siniver. 2011. "Why is an elite undergraduate education valuable? Evidence from Israel." *Labour Economics*, 18(6): 767–777.

MacLeod, W. Bentley, Evan Riehl, Juan E. Saavedra, and Miguel Urquiola. 2017. "The Big Sort: College Reputation and Labor Market Outcomes." *American Economic Journal: Applied Economics*, 9(3): 223–61.

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

Ockert, Bjorn. 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.

Ost, Ben, Weixiang Pan, and Douglas Webber. 2016. "The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies." *Journal of Labor Economics*, 36(3): 779–805.

Park, R. Jisung. 2022. "Hot Temperature and High-Stakes Performance." *Journal of Human Resources*, 57(2): 400–434.

Quadlin, Natasha. 2018. "The mark of a woman's record: Gender and academic performance in hiring." *American Sociological Review*, 83(2): 331–360.

Zimmerman, Seth. 2014. "The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies." *Journal of Labor Economics*, 32(4): 711–754.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

Table 1: Effects of GPA Luck on Later-in-life Outcomes

Panel A: High school Outcomes	Exam grades in 3^{rd} year	High school GPA	On time HS diploma	Ever HS diploma
GPA_luck	0.1359***	0.0247***	0.0141***	0.0039***
	(0.0026)	(0.0015)	(0.0011)	(0.0006)
Mean dep. var.	-0.000	-0.000	0.867	0.962
N	89638	89638	89638	89638
Panel B: Higher Education Outcomes	Ever higher education	Share of available HE programs	Selectivity of HE enrollment	Number of completed years in HE
GPA_luck	0.0003	0.0009***	0.2419**	0.0070
	(0.0007)	(0.0002)	(0.0999)	(0.0056)
Mean dep. var.	0.945	0.914	36.321	2.877
N	89638	84713	84713	89638
Panel C: Labor Market Outcomes	Ever employed	First job annual labor income (log)	Employed 8 years after the exams	Annual labor income 8 years after the exams (log)
GPA_luck	0.0005	0.0052*	0.0017	0.0050**
	(0.0012)	(0.0031)	(0.0014)	(0.0024)
Mean dep. var.	0.823	12.322	0.740	12.677
N	89638	73774	89638	66343

NOTES: The table refers to the sample of regular full-time high school students who enrolled for the first time in the final year of academic high school between 2005 and 2009, and who received no failing grade (course grade = 1) in the courses they took during the school year. Higher education outcomes on the share of available programs and on the selectivity of the first enrollment are restricted to the students who enrolled in college. Earnings are restricted to the students who obtained a first job (Panel C, second column) and who are employed 8 years after the exams (Panel C, fourth column). Each column reports the estimated impacts of GPA_luck—instrumented by its recentered counterpart $\widetilde{\text{GPA\_luck}}$. Measures of exam grades and high school GPA in third year are standardized to mean zero and unit variance. Each regression includes baseline controls (demographic characteristics, middle school GPA, and high school course grades), and 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%.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

Table 2: Effects of GPA Luck on Labor Market Outcomes, by Length of University Education

	High school GPA	Employed 8 years after the exams	Annual labor income 8 years after the exams (log)
Subsample with no higher education			
GPA_luck	0.0165**	-0.0074	0.0015
	(0.0075)	(0.0057)	(0.0089)
Mean dep. var.	-1.110	0.781	12.623
N	4925	4925	3847
Subsample with a maximum of 3 years of completed higher education			
GPA_luck	0.0248***	0.0028	0.0077**
	(0.0017)	(0.0019)	(0.0034)
Mean dep. var.	-0.192	0.727	12.559
N	54984	54984	39995
Subsample with more than 3 years of completed higher education			
GPA_luck	0.0258***	0.0007	-0.0009
	(0.0024)	(0.0025)	(0.0027)
Mean dep. var.	0.539	0.757	12.896
N	29729	29729	22501

Notes: The table reports the main results of Table 1, separately on the subsample of students who never enrolled in college (top panel), on the subsample of students who completed 3 years of higher education or less (middle panel), and on the subsample of students who completed more than 3 years of higher education (bottom panel). Each regression includes a set of baseline controls (demographic characteristics, middle school GPA, and high school course grades), 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%.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01422

Table 3: Effects of High School GPA on Annual Earnings: An Instrumental Variable Approach

	Outcomes			
	First job log annual labor income (2SLS)	First job log annual labor income (OLS)	Log annual labor income 8 years after the exams (2SLS)	Log annual labor income 8 years after the exams (OLS)
High school GPA	0.232*	0.100***	0.226*	0.046***
	(0.140)	(0.006)	(0.120)	(0.005)
N	73774	73774	66343	66343

NOTES: The table refers to the same sample as Table 1 Panel C. The first and third columns report the 2SLS estimates of the impact of students' GPA on the log of students' annual labor earnings at their first job or eight years after the exams, using $\widetilde{\text{GPA\_luck}}$ as instrument. The value of the F-statistics for the first stage regressions are 59 and 48, respectively. The second and fourth columns report the OLS estimates of the same parameters. Each regression includes baseline controls (demographic characteristics and middle school GPA), 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%.