# Rank, Sex, Drugs and Crime

## Supplementary appendix

Benjamin Elsner
IZA

Ingo E. Isphording
IZA

10 March 2017

**Abstract:** We show that a student's ordinal ability rank in a high-school cohort is an important determinant of engaging in risky behaviors. Using longitudinal data from representative US high schools, we find a strong negative effect of rank on the likelihood of smoking, drinking, having unprotected sex and engaging in physical fights. We further provide evidence that these results can be explained by sorting into peer groups and differences in career expectations. Students with a higher rank are less likely to be friends with other students who smoke and drink, while they have higher expectations towards their future educational attainment.

**JEL-Classification:** I12, I14, I21, I24

**Keywords:** Risky behavior, ability rank, peer effects, beliefs, expectations

---

We would like to thank the Editor, three anonymous referees, as well as Laura Argys, Sanni Nørgaard Breining, Scott Carrell, Arnaud Chevalier, Damon Clark, Deborah Cobb-Clark, Tine Mundbjerg Eriksen, Jason Fletcher, Hani Mansour, Dan Rees, Derek Stemple, Felix Weinhardt, Ulf Zölitz, as well as audiences at IZA, the Workshop Health.Skills.Education in Essen, EEA Mannheim, the workshop of the Copenhagen Education Network, U Aarhus, SOLE 2016, CU Denver, U Leicester and EALE 2016 for their helpful comments.

Corresponding author: Benjamin Elsner, IZA and CReAM. IZA - Institute of Labor Economics, Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. Phone, +49 (228) 3894-522, Fax +49 (228) 3894-510, E-mail elsner@iza.org

Ingo E. Isphording, IZA - Institute of Labor Economics, E-mail isphording@iza.org.

# Online Appendices

# A Summary statistics

Table 5 reports summary statistics for the main regressors.

Table 5: Descriptive statistics

| | Mean | (SD) | min | max |
|---|---|---|---|---|
| *Cognitive ability:* | | | | |
| Peabody Vocabulary Test | 100.46 | (14.65) | 13.00 | 139.00 |
| *Control variables:* | | | | |
| Age | 15.84 | (1.56) | 11.43 | 20.68 |
| Female | 0.52 | (0.50) | 0.00 | 1.00 |
| White | 0.56 | (0.50) | 0.00 | 1.00 |
| Asian | 0.07 | (0.25) | 0.00 | 1.00 |
| Black | 0.22 | (0.42) | 0.00 | 1.00 |
| Hispanic ancestry | 0.15 | (0.36) | 0.00 | 1.00 |
| Parental education: high school dropout | 0.14 | (0.35) | 0.00 | 1.00 |
| Parental education: high school | 0.25 | (0.43) | 0.00 | 1.00 |
| Parental education: some college | 0.24 | (0.43) | 0.00 | 1.00 |
| Parental education: college | 0.36 | (0.48) | 0.00 | 1.00 |
| Number of observations | 12528 | | | |

*Notes:* This table summarizes the mean, standard deviation and range of the control variables for all individuals whose risky behaviors are observed in wave II. The data source is the "in-home" sample of AddHealth.

# B Identifying variation in key variables

In our preferred specifications, we estimate the effect of a student's ordinal rank on various outcomes conditional on school-by-cohort fixed effects. Table 6 demonstrates that even in this specification there is considerable variation in the key variables. The numbers represent the variation in the main dependent and explanatory variables with various sets of fixed effects. In Column (1), we report the variation without any controls. In Columns (2) and (3), we first regress each variable on the respective fixed effects, and compute the variation of the residuals. In addition, in the second row, we condition on absolute ability. Numbers in square brackets report the share of remaining variation $(1 - R^2)$ after the fixed effects have been controlled for.

Perhaps surprisingly, adding school and cohort fixed effects, and even adding school-by-cohort fixed effects does not significantly reduce the remaining variation in ordinal rank. As shown in the second row, more than 50% of the variation in ordinal rank is left once absolute ability and school-by-cohort fixed effects are controlled for. This is the case because ranks are assigned *within* school cohorts, whereas the school-by-cohort fixed effects absorb all variation *between* school cohorts.

The variation in ordinal rank conditional on absolute ability is the identifying variation of our model. Without any fixed effects, the standard deviation of the ordinal rank conditional on ability is 0.16, which means that at a mean cohort size of 180 students and for a given level of absolute ability the ability rank would vary on average by $(0.16 \times 180 =)$ 28.8 absolute rank positions. Once we additionally control for separate school and cohort fixed effects, the within-school standard deviation of the rank conditional on ability becomes slightly smaller with 0.12. In the most demanding specification with school-by-cohort fixed effects, the standard deviation in rank is 0.115. For the remaining variables, school-by-grade fixed effects slightly lower the remaining variation.

In sum, these statistics show that a substantial degree of variation in the ordinal rank and all outcome variables remains even if ability and school-by-cohort fixed effects are controlled for.

Table 6: Variation in key variables after fixed effect transformations

| | (1) raw SD | (2) School *and* Cohort FE | (3) School $\times$ Cohort FE |
|---|---|---|---|
| Ordinal rank | 0.284 | 0.283 | 0.281 |
| | | [0.993] | [0.983] |
| Ordinal rank conditional on ability | 0.161 | 0.120 | 0.115 |
| | | [0.552] | [0.510] |
| Peabody score | 14.646 | 13.043 | 12.870 |
| | | [0.793] | [0.772] |
| Smoking (on at least 10 out of past 30 days) | 0.382 | 0.367 | 0.361 |
| | | [0.921] | [0.891] |
| Drinking (on at least 2 days per month past 12 months) | 0.387 | 0.374 | 0.368 |
| | | [0.936] | [0.905] |
| Marijuana use (at least once within past month) | 0.367 | 0.359 | 0.354 |
| | | [0.957] | [0.930] |
| Intercourse w/o birth control in past 6 months) | 0.281 | 0.277 | 0.274 |
| | | [0.968] | [0.945] |
| Stealing (at least once past 12 months) | 0.418 | 0.411 | 0.406 |
| | | [0.965] | [0.939] |
| Physical fight (at least once past 12 months) | 0.396 | 0.392 | 0.387 |
| | | [0.976] | [0.952] |
| Drug selling (at least once past 12 months) | 0.263 | 0.259 | 0.256 |
| | | [0.973] | [0.947] |

*Notes:* This table summarizes standard deviations of predicted variables after unweighted linear regressions of variables indicated on the left on sets of fixed effects: (1) no fixed effects, (2) separate school and fixed effects (3) school-by-cohort fixed effects. Numbers in square brackets report the share of remaining variation $(1 - R^2)$.

# C Robustness checks and threats to identification: summary

When discussing the main results in Section 4, we highlighted several sources of bias for our estimates, namely measurement error, omitted variables, misreporting of risky behaviors, attrition and strategic delay in school entry. In this section, we address these issues with a series of robustness checks and simulation exercises.

Table 7: Robustness checks

| | Baseline | Completed High-school | Mean age ±0.5 years | Share girls 40-60% |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Smoking:** | | | | |
| on at least 10 out of past 30 days | -0.124*** | -0.114** | -0.112** | -0.121** |
| | (0.045) | (0.047) | (0.054) | (0.052) |
| **Drinking:** | | | | |
| on at least 2 days per month past 12 months | -0.138*** | -0.110** | -0.150*** | -0.161*** |
| | (0.043) | (0.050) | (0.057) | (0.051) |
| **Marijuana use:** | | | | |
| at least once within past month | -0.072* | -0.071 | -0.063 | -0.094* |
| | (0.043) | (0.044) | (0.057) | (0.052) |
| **Sex:** | | | | |
| Intercourse w/o birth control in past 6 months | -0.055* | -0.028 | -0.045 | -0.055 |
| | (0.031) | (0.035) | (0.035) | (0.040) |
| **Crime:** | | | | |
| Stealing at least once past 12 months | -0.061 | -0.091* | -0.079 | -0.060 |
| | (0.045) | (0.054) | (0.050) | (0.055) |
| Physical fight at least once past 12 months | -0.138*** | -0.160*** | -0.117** | -0.121*** |
| | (0.037) | (0.041) | (0.048) | (0.044) |
| Drug selling at least once past 12 months | -0.052* | -0.072*** | -0.031 | -0.063** |
| | (0.027) | (0.028) | (0.034) | (0.031) |
| **Sample size** | 12528 | 9401 | 8156 | 8789 |

*Notes:* This table displays results of separate OLS regressions of binary variables indicating engagement in risky behavior on the ordinal rank. All regressions are unweighted. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the school level, are reported in parentheses. – All estimations include the same controls and fixed effects as in Table 2, Column (3). Column (1) reproduces the baseline results. In Column (2), the sample comprises all students who eventually completed high school. The sample in Column (3) only includes students born within 6 months before or after the cohort average. In Column (4), the sample only includes school cohorts with a relatively even gender balance (between 40-60 and 60-40).

## C.1 Measurement error

The estimates of the rank effect can potentially be biased by multiple sources of measurement error. Here, we summarize and discuss the results from extensive simulations aimed at quantifying the direction and magnitude of the bias from measurement error. A more detailed description of the simulations can be found in Appendix E.

**Sampling error** One source of measurement error is the fact that we only observe a student's percentile rank based on a sample of 30 out of 180 students in a school cohort on average. From this sampling ratio, critical observers might conclude that an insufficient number of students are sampled to meaningfully reconstruct the ordinal ranking and obtain an unbiased estimate. However, in Appendix E.2, we show that because students are *randomly* drawn from every school cohort, this sampling design results in classical measurement error in the rank variable, which means that the estimates represent a lower bound to the true effect. In addition, our simulations suggest that the attenuation bias converges to a moderate level as the sampling ratio becomes smaller. Even if we only sample 40 out of 400 students at random, we under-estimate the true effect by around 25%. At the average school cohort size of 180 students, we under-estimate the true effect by around 20%. Thus, the effects shown in Section 4 are around 20% lower than the true effects would be.

**Measurement error in the Peabody test score** A further potential source of measurement error lies in the measurement of the Peabody test score. For example, take a student who had a bad day when taking the Peabody test. This student's test score would be lower than his/her latent cognitive ability, meaning that we would assign this student a lower rank than the rank based on his/her latent ability. Even if such deviations between the test score and a student's latent ability are not systematic — i.e. unrelated to any student characteristics — they may bias the estimate of the rank effect. As we show in Appendix E.3, the extent of this bias depends the ability measure that underlies the ranking in the data-generating process. If the rank in the data-generating process — the rank that actually matters for a student's risky behavior — is based on the latent ability whereas we compute the rank based on the test score, we show that the estimates are biased towards zero. However, if what matters to students is the rank based on the test score, then measurement error in the test score does not bias the estimates.

**Gender stratification** The estimates could also be biased due to the gender stratification within school cohorts. From each school cohort, equal numbers of boys and girls were sampled, regardless of the underlying gender distribution in the population. Given that we observe the population gender distribution in the in-school sample of AddHealth, we can observe whether the effects significantly change if we only consider school cohorts with a relatively even gender balance. In Column 5 of Table 7, we only keep cohorts with a gender balance between 40-60 and 60-40. The results are very similar to the baseline, indicating the absence of measurement error due to gender stratification.

**Over-sampling of minorities** Finally, measurement error could arise from the over-sampling of minorities. If minorities systematically have a lower rank but are over-sampled, then non-minority students would be assigned a higher rank on average. To test whether this source of measurement error is important, we exploit the fact that the in-home sample has precise information on who has been over-sampled and subsequently compute the rank purely based on the randomly drawn core sample. The correlation between the rank with and without over-sampling is almost perfect ($\rho = 0.987$). In Appendix F.3, we re-estimate the baseline model with both definitions. The results remain virtually the same.

### C.2 Omitted variable bias

A further potential source of bias is omitted variables, i.e. variables that have an effect on the ordinal rank while having a direct impact on risky behaviors. Because all school-cohort variables are absorbed by the fixed effects, omitted variables have to vary at the individual level. In our context, important candidates for omitted variables are unobserved personality traits, influences from a student's environment (parents, teachers, peers, etc.) or tracking.

We distinguish between two types of omitted variables, namely those correlated with the ordinal rank conditional on absolute ability and those affecting the ability test score, which in turn affects the observed ordinal rank. Despite being a seemingly small distinction, both types of omitted variables lead to different biases, with the second type being a more severe source of bias than the first.

The first type is a variable that is correlated with the ordinal rank *conditional on absolute ability* and has a direct impact on the outcome. For example, teachers or parents may treat a student with a low rank differently from one with a high rank, whereby this difference in treatment may directly affect risky behaviors. While such a line of argumentation would fit the definition of an omitted variable, it describes a channel rather than an omitted variable. If parents' or teachers' behavior responds to a child's rank and the rank has been assigned exogenously, then these responses are a channel through which a high rank translates into less risky behavior. A similar argument can be made about tracking: if the number of slots in a higher track is fixed, then students with a higher rank are more likely to proceed to a higher track, which may have an impact on a student's risky behavior. Yet again, students proceed to a higher track *because* they have a higher rank and thus tracking is a mediating factor rather than a confounder. In Appendix E.4.1, we provide a more detailed discussion of this type of omitted variable bias.

The second type of omitted variable is one that is correlated with a student's absolute level of ability, while having a direct influence on risky behaviors. Because the ordinal rank is based on absolute ability, a variable correlated with ability may also affect the ordinal rank, thus biasing the estimated rank effect. One example of such an omitted variable is a student's motivation. A student who is more motivated may achieve a higher test score, while a higher motivation may also lead to less engagement in risky behaviors through other channels. Similar cases can be made for other personality traits. Another example is parental pressure, which may affect a student's test score while simultaneously influencing risky behaviors.

To assess the direction and magnitude of the omitted variable bias, we conduct a Monte Carlo experiment in which we assume a data-generating process where the outcome is determined by rank, absolute ability and an omitted factor that is also correlated with ability. The analysis yields two results. First, we obtain an unbiased estimate as long as the rank in the data-generating process is based on the test score. This is the case because the direct effect of the omitted variable is absorbed by the control for absolute ability. Second, our estimates are biased towards zero if students care about the ranking in terms of latent ability while we only observe a test score that is influenced by latent ability plus an omitted variable. This type of omitted variable bias is equivalent to a classical measurement error in the rank variable. To see this, consider two students with the same latent ability but different motivation. Student A, who has a higher motivation than student B, will have a higher test score and we will assign him/her a higher rank, while in the data-generating process both have the same ability rank. The fact that we assign some students a rank that is too high and others a rank that is too low attenuates the estimate. The magnitude of this bias depends on the extent to which the omitted variable affects the ability test score. In Appendix E.4.2, we provide a detailed account of the simulations.

### C.3 Misreporting

All measures for risky behaviors are self-reported, which is a potential source of bias in our estimates. The designers of AddHealth aimed at reducing the extent of misreporting of risky behaviors by using a so-called audio-CASI technique. Students listened to the questions through headphones and typed their answers into a computer. This method should eliminate peer pressure as a reason for misreporting, while it also avoids students having to disclose sensitive information to an unknown interviewer. Nonetheless, while this method can help to reduce misreporting, it may not fully eliminate it. For example, the descriptive statistics in Table 1 show a share of 7th and 8th graders having unprotected sex that may seem very large to some observers and could partially be the result of age-specific misreporting. In our regression, we eliminate the bias from age-specific misreporting through cohort fixed effects and controls for relative age. In fact, if misreporting is correlated with any of the control variables but not with rank, it does not bias our estimate of the rank effect.

However, misreporting can introduce a bias if it is systematically correlated with rank; for example, if highly-ranked students are more likely to under-report their engagement in risky behaviors. In Appendix G, we provide a more extensive discussion of the bias, showing that if students with a low rank over-report their risky behavior compared to students of high rank, this leads to an over-estimation of the rank effect. However, given the survey design and the application of audio-CASI, the correlation between rank and misreporting should be fairly small, such that our estimates would suffer from a small positive bias, if at all.

### C.4 Strategic delay of school entry

The central identifying assumption is that, conditional on school choice, being in one cohort or another is as good as random, which is the case if students and parents cannot influence the assignment into cohorts. This assumption may be violated if students have to repeat a grade or if parents strategically delay their children's school entry, allowing them to mature for one more year. Strategic delay of school entry — also called 'redshirting' — has become more common over time in the US. As shown by Deming and Dynarski (2008), 96 percent of schoolchildren in the US were enrolled at age 6 in 1968, whereas in 2005 this figure stood at 84 percent. Our regression partly accounts for the possible bias introduced by grade retention or redshirting, because we control for age in months. Nonetheless, the results could be biased if redshirted children systematically differ from those whose school entry was entirely determined by their birth date and the cut-off date of their school. To alleviate this concern, we restrict the sample to students who are at most 6 months older or younger than the cohort average. Because redshirted students would be more than 6 months older than the cohort average, they are excluded from the sample. The magnitude of the coefficients does not significantly decline compared to the baseline results. Some coefficients — notably those of risky sex and drug selling — are no longer statistically significant, although given that the point estimates are similar, the higher standard errors seem to be due to a smaller sample size rather than a bias in the estimates.

### C.5 Attrition

The baseline estimates are potentially biased by selective attrition. Between wave I and wave II, we lose 5,000 observations, which represents 28% of the sample. Attrition introduces a bias if it is selective: if low-ranked students are more likely to attrit from the sample, we would expect a downward bias in the results, whereas if higher-ranked students are more likely to attrit, we would expect an upward bias. To test whether selective attrition introduces a bias, we regress an attrition dummy on the ordinal rank, a quartic in absolute ability, individual controls, as well as school and cohort fixed effects. There is no evidence of systematic attrition. The coefficient of the ordinal rank is positive but statistically insignificant (coefficient 0.047, standard error 0.034, t-statistic 1.38).

A more subtle form of attrition could occur because we observe every cohort in a given survey wave at a different grade level. If the lowest-ranked students in every grade drop out, then grade 12 represents a much more positive selection of students than grade 7. This difference would be captured by cohort fixed effects if it was the same across all schools, although it would not be captured if dynamic attrition systematically differs between schools. To address this problem, we restrict the sample to students who report in wave IV that they finished high school, whereby we subsequently compute the rank based on this selected group, such that the rank measure is not contaminated by dynamic attrition. The results are displayed in Column (3) of Table 7. For most behaviors, the magnitude of the effect is the same as in the baseline (Column (1) of the same table). The statistical significance is lower for most coefficients, although the coefficients of smoking and sex remain significant at the 10% level and the coefficient of engagement in physical fights is significant at the 1% level.

### C.6 Ability influenced by prior rank

A further threat to identification could be that the ability test score based on which we compute the ordinal rank may be a function of a student's prior rank. We only observe students in wave I at the average age of 16, which means that we measure their ability a few years after they entered the school. If their ordinal rank at the time of school entry affected their measured ability, we could not interpret the estimate of $\gamma$ as being causal because the prior rank would lead to omitted variable bias. In Appendix F.1, we provide two theoretical arguments against omitted variable bias. In addition, we provide evidence that prior rank does not affect current ability by showing that the ordinal rank in wave I does not predict the Peabody test score in wave III.

# D Additional results

## D.1 Varying the intensity of risky behaviors

In this section, we analyze the extent to which the impact of the ordinal rank varies in the intensity of risky behaviors. For smoking, drinking, marijuana use and sex, we define binary indicators for different intensities, such as the onset of risky behaviors, regular or severe engagement (e.g. regular binge drinking). For delinquent behavior, we only observe the incidence, such that we cannot compute the intensity. Table 8 reports the results for regressions with school-by-cohort fixed effects. The intensities used in the main Tables 1 and 2 are marked with an asterisk.

Overall, the pattern of significant negative effects of the ordinal rank on probabilities of engaging in risky health behaviors holds across intensities. The relative importance of the ordinal rank at varying intensities appears to differ by risky behavior.

For smoking and marijuana use, we observe stronger effects for onset and regular smoking, but a smaller coefficient for severe smoking and marijuana consumption. For alcohol consumption, we do not observe an effect of the ordinal rank on the onset (ever drank alcohol yes/no), but distinctive negative effects on the severity. With respect to sex, we again observe the strongest influence on having had any sexual intercourse before, with weaker influences on recent or recent risky sexual intercourse.

Table 8: Rank effects with different intensities of risky behaviors

| | Mean | Coeff | Definition |
|---|---|---|---|
| **Smoking:** | | | |
| Ever | *mean = 62.44* | -0.129** | Ever smoked cigarettes yes/no |
| | | (0.062) | |
| Regularly* | *mean = 17.79* | -0.124*** | Smoked on at least 10 out of 30 last days |
| | | (0.045) | |
| Severe | *mean = 7.93* | -0.048* | Smoked >=10 daily cigarettes on at least 10 out of 30 last days |
| | | (0.028) | |
| **Drinking:** | | | |
| Ever | *mean = 64.77* | -0.067 | Ever drank alcohol yes/no |
| | | (0.053) | |
| Regularly* | *mean = 18.32* | -0.138*** | Drank on at least 2/3 days per month during last year |
| | | (0.043) | |
| Severe | *mean = 11.05* | -0.099*** | Got seriously drunk at least 2/3 days per month during last year |
| | | (0.033) | |
| **Marijuana:** | | | |
| Ever | *mean = 34.84* | -0.087 | Ever used Marijuana yes/no |
| | | (0.054) | |
| Recently* | *mean = 16.09* | -0.072* | Used marijuana during past 30 days |
| | | (0.043) | |
| Severe | *mean = 7.46* | -0.056* | Used marijuana at least 2 times during past 30 days |
| | | (0.029) | |
| **Sex:** | | | |
| Ever | *mean = 47.99* | -0.146*** | Had sexual intercourse before yes/no |
| | | (0.055) | |
| Recently | *mean = 33.01* | -0.103* | Had sexual intercourse past 6 months |
| | | (0.058) | |
| Recently w/o birth control* | *mean = 8.67* | -0.055* | Had sexual intercourse without birth control past 6 months |
| | | (0.031) | |

*Notes:* This table displays results of separate OLS regressions of binary variables indicating different intensities of risky behavior on the percentile rank. All regressions are unweighted. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ – Standard errors, clustered at the school level, are reported in parentheses. – The regression specification is the same as in Table 2, Column (3) including school-specific cohort fixed effects.

## D.2 Heterogeneous effects

In this section, we analyze heterogeneous effects of the ordinal rank on risky behaviors across various groups by interacting the ordinal rank with a group indicator. All regressions include school-by-cohort fixed effects and the same controls as the baseline specification. Table 9 reports the coefficients for the main effects — i.e. the coefficient of the ordinal rank — and the interaction between rank and the indicator.

Overall, we do not find a coherent pattern. In most cases, the interactions are imprecisely estimated, preventing us from identifying heterogeneous effects.

In Column (1), we analyze a potential non-linearity in the rank effect between students of high and low rank. Interaction effects between rank and an indicator for having a rank above the school-cohort median are small and insignificant. Therefore, our results do not support any non-linear rank effect on risky behaviors, which is in line with findings by Murphy and Weinhardt (2014), who find a virtually linear effect of a student's rank on test scores. While it would be interesting to test for more heterogeneous effects, our empirical setting does not provide us with the statistical power to do so.

We additionally analyze whether students in high-ability schools are more responsive to their ranking (Column (2)). We do not find any significant heterogeneity with the exception of alcohol consumption, for which the effect is significantly stronger in low-ability schools.

Gender differences (Column (3)) in the rank effect emerge for marijuana use, risky sex and stealing. In these cases, female risky behavior appears to react more strongly to the ordinal rank than that of male students.

We further analyze heterogeneity by family background (parental college degree, Column (4)) and race (Column (5)), finding some differences. In the case of alcohol consumption, children of college-educated parents are more affected by their rank than children of parents with less than a college degree. Non-white students are less affected than white students. The same pattern emerges for stealing. In the case of engagement in physical fights, both children of parents with college degree and non-white students are more likely affected by their rank status.

Table 9: Heterogeneous effects

| | (1) | | (2) | | (3) | | (4) | | (5) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Rank > .5 | | High abil. school | | Female | | Parents: College | | non-white | |
| | main | int | main | int | main | int | main | int | main | int |
| **Smoking:** | | | | | | | | | | |
| Regularly | -0.100* | -0.018 | -0.088* | -0.048 | -0.107** | -0.032 | -0.111** | -0.036 | -0.140** | 0.075 |
| | (0.054) | (0.022) | (0.047) | (0.030) | (0.046) | (0.021) | (0.049) | (0.029) | (0.043) | (0.025) |
| **Drinking:** | | | | | | | | | | |
| Regularly | -0.129** | -0.007 | -0.184*** | 0.060** | -0.127*** | -0.021 | -0.114*** | -0.066*** | -0.143*** | 0.024*** |
| | (0.057) | (0.025) | (0.046) | (0.024) | (0.046) | (0.024) | (0.044) | (0.024) | (0.044) | (0.024) |
| **Marijuana use:** | | | | | | | | | | |
| Recently | -0.110** | 0.029 | -0.083* | 0.014 | -0.050 | -0.041* | -0.064 | -0.020 | -0.073 | 0.004 |
| | (0.055) | (0.023) | (0.046) | (0.028) | (0.045) | (0.022) | (0.043) | (0.027) | (0.043) | (0.027) |
| **Sex:** | | | | | | | | | | |
| Risky intercourse | -0.048 | -0.005 | -0.037 | -0.024 | -0.034 | -0.040** | -0.058* | 0.008 | -0.048* | -0.036 |
| | (0.043) | (0.019) | (0.035) | (0.021) | (0.031) | (0.018) | (0.033) | (0.022) | (0.032) | (0.021) |
| **Crime:** | | | | | | | | | | |
| Stealing | -0.043 | -0.014 | -0.075 | 0.018 | -0.036 | -0.047* | -0.038 | -0.062** | -0.070 | 0.040** |
| | (0.058) | (0.026) | (0.049) | (0.028) | (0.050) | (0.026) | (0.048) | (0.028) | (0.047) | (0.031) |
| Physical fight | -0.127** | -0.008 | -0.119*** | -0.026 | -0.120*** | -0.035 | -0.116*** | -0.058** | -0.121*** | -0.081** |
| | (0.051) | (0.022) | (0.038) | (0.026) | (0.042) | (0.028) | (0.039) | (0.030) | (0.039) | (0.027) |
| Drug selling | -0.053 | 0.001 | -0.070** | 0.024 | -0.047 | -0.008 | -0.044 | -0.021 | -0.051 | -0.003 |
| | (0.035) | (0.015) | (0.030) | (0.018) | (0.031) | (0.020) | (0.027) | (0.020) | (0.027) | (0.016) |

*Notes:* This table displays results of OLS regressions of binary indicators indicating engagement in risky health behavior on the ordinal rank in a cohort and its interaction with group membership indicators. Columns *main* display the main effects, *int* display the coefficients of the interaction terms. All regressions are unweighted. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors, clustered at the school level, are reported in parentheses. The regression specification is similar to Table 2, Column (3) including school-specific cohort fixed effects.

# E Simulations: Measurement error and omitted variable bias

## E.1 Basic Monte Carlo experiment

The estimates presented in Section 4 are potentially biased due to measurement error in the rank variable or due to measurement error in the Peabody test scores that we use to construct the ordinal rank. Moreover, there could be unobservable factors that affect the ordinal rank as well as having a direct effect on the outcome. To shed light on the direction and magnitude of the bias arising from measurement error and omitted variables, we carry out an extensive series of Monte Carlo experiments.

In all simulations, we assume the following data-generating process (DGP):

$$y = -0.1r - 0.6a + \delta \tag{1}$$

where $r$ is a student's ability rank, $a$ is his/her ability and $\delta$ is a school-cohort specific intercept. For simplicity we drop all subscripts.

## E.2 Bias from sampling error in the rank variable

In the first Monte Carlo experiment, we quantify the bias from sampling error in the rank variable. Throughout our analysis, we compute the ability rank within a school cohort based on a random sample of 40 out of 180 students on average. Because we do not observe the full population, we inevitably assign some students a rank that is higher than their true rank in the population and others a rank that is lower. Given that the sample was drawn at random from the population, the average sampling error – i.e. the difference between the sample rank and the population rank – is zero, although the standard deviation of the sampling error is greater than zero. Therefore, the error from random sampling can be seen as a classical measurement error of a regressor, which attenuates the estimates.

To quantify the degree of attenuation bias, we conduct a Monte Carlo experiment with ten sets of 1,000 replications. In each replication, we construct a population of schools that have the same features as those in AddHealth. We draw from each school cohort a random sample of 40 students and estimate the model in Equation (1). In each set of replications, we assume a different size of the underlying school cohort. In the first set, the population school cohort size is 40, such that we sample the entire school cohort. In subsequent sets, we increase the population school cohort size in steps of 40 up to 400 students, in which case we only sample 10% of every school cohort.

In each replication, we construct a dataset of 500 school cohorts. To account for heterogeneity in the mean and variance of ability across school cohorts, we draw the ability distribution in two steps. We first draw for each school cohort the mean ability from a normal distribution $\delta \sim N(mean = 101, sd = 7)$, and the standard deviation of ability with $\sigma \sim N(mean = 12, sd = 2.5)$, and in a second step draw the ability of each student in a school cohort from the normal distribution $a \sim N(\delta, \sigma)$. Based on the ability distribution of every school cohort, we compute a student's rank in the population as well as the outcome $y$ using Equation (1). Finally, we draw from each population school cohort a random sample of 40 students, compute the ordinal rank based on this sample, estimate the model

$$y = \gamma r + \beta a + \delta + \varepsilon \tag{2}$$

1,000 times. To assess the bias from measurement error, we compare the average estimated $\hat{\gamma}$ from these 1,000 replications to the assumed true effect $\gamma = -0.1$.

Figure 3 plots the estimated coefficients as a function of the school cohort size $N$ in the underlying population. Overall, we can observe that the sampling error in the ordinal rank biases the estimates towards zero, while the bias peters out as the sampling ratio becomes smaller. At $N = 40$, when we sample the entire cohort there is no bias because the sample rank equals the population rank. At $N = 400$, when the sampling ratio is 10%, the estimated rank effect is smaller but still amounts to $\hat{\gamma} = -0.077$. The vertical dashed line represents the average population cohort size of 180 students. In the average school in AddHealth, we would under-estimate the effect of ordinal rank on risky behaviors by 20%.

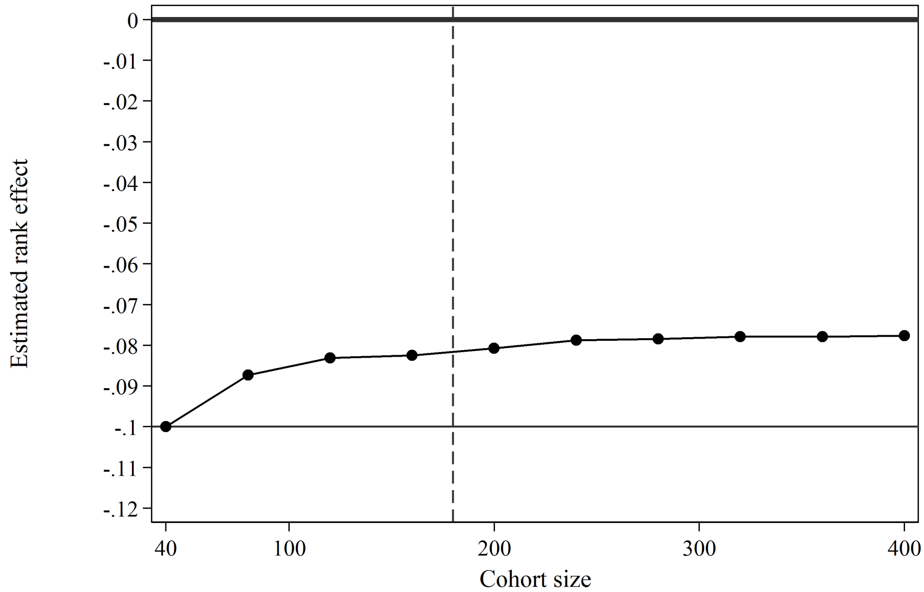


Figure 3: Measurement error due to stratified random sampling

*Notes:* This figure plots the simulated effect of rank on risky behaviors as a function of the school cohort size in the underlying population. Each point represents the average estimate from 1,000 replications, with 40 students randomly drawn per school cohort. The assumed true effect is $\gamma = -0.1$. The dashed vertical line represents the average school cohort size in AddHealth of 180 students.

## E.3 Bias from measurement error in the test score

A further potential source of bias is measurement error in the Peabody test score. Rather than measuring the pure ability $a$, the measured ability $\tilde{a}$ also includes an error component $z$,

$$\tilde{a} = a + \phi z, \tag{3}$$

with $\phi$ governing the strength of the measurement error. This error can occur either because the ability test score was computed wrongly or due to factors that affect a student's test performance yet are unrelated to rank.

### E.3.1 Measurement error in the test score but not in the rank

As long as $z$ is pure measurement error with mean zero unrelated to a student's ordinal rank, it will not affect our estimates of $\gamma$, because we control for the measured absolute ability $\tilde{a}$ in all regressions. To verify this, we conduct a Monte Carlo experiment in which we assume that $z$ is drawn from a standard normal distribution with mean zero and variance one. For simplicity, we also draw the latent ability $a$ from a standard normal distribution $a \sim N(0,1)$. We compute the percentile rank based on the measured ability using Equation (1) and assume that this ranking based on measured ability is the true rank in the data-generating process, i.e. the rank that affects student's engagement in risky behavior. We run 11 sets of 1,000 replications each, varying the measurement error $\phi$ in steps of 0.05 from 0 to 1. Each replication is based on a dataset of 500 simulated school cohorts with 40 students per school cohort. As shown in Figure 4, as long as the measurement error in the ability variable does not carry over to the ordinal ranking the estimates are unbiased.

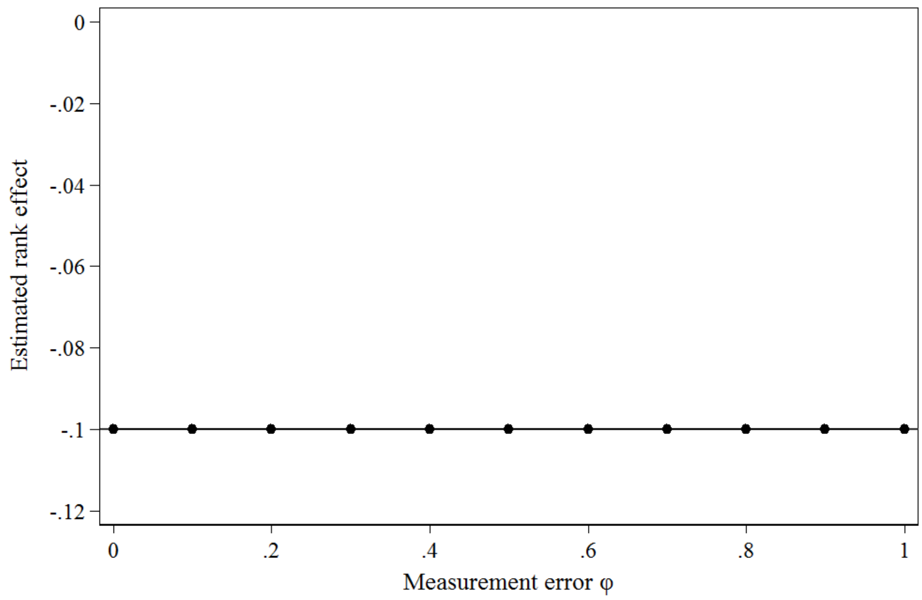


Figure 4: Simulations: measurement error in the Peabody test score

*Notes:* This figure plots the simulated effect of rank on risky behaviors as a function of the measurement error in the ability test score. Each point represents the average estimate from 1,000 replications. The assumed true effect is $\gamma = -0.1$.

### E.3.2 Measurement error in the test score carries over to rank

In a next step, we analyze the bias if the ability is measured with error and the measurement error carries over to the ordinal rank. This would be the case if a student's decisions are affected by a ranking based on latent ability $a$, while we observe the ability with a measurement error and base the ability rank on the mismeasured variable $\tilde{a}$.

Let $h(a)$ be the ability distribution of student $i$'s school cohort and $f(a_i, h(a)) \rightarrow [0,1]$ be the assignment function that maps student $i$'s absolute ability into an ability rank that ranges between zero and one. We now assume that the ranking that enters the DGP is based on $r = f(a_i, h(a))$, whereas we observe a ranking based on measured ability $\tilde{r} = f(\tilde{a_i}, h(\tilde{a}))$. Deriving this bias analytically is difficult, because the first step of the rank assignment turns the continuous ability distribution into a discrete ranking.[1]

Therefore, we assess the bias using Monte Carlo simulations. As before, each replication is based on 500 school cohorts, each with 40 students. Latent ability and the measurement error are independently drawn from a standard normal distribution and we compute the rank in the DGP based on the latent ability, whereas we estimate the model using the rank based on measured ability. We run 20 sets of simulations, varying the measurement error from 0 to 1 in steps of 0.05. Each set of simulations comprises 1,000 replications.

Figure 5 plots the estimated effect of ordinal rank as a function of the measurement error $\phi$. It shows that this measurement error biases the results towards zero. This suggests that our estimates are a lower bound to the true effect.

### E.4 Assessing omitted variable bias

A further potential source of bias could be omitted variables that have a direct effect on the outcome while being correlated with the rank and/or the Peabody test score. Examples of omitted variables are parental pressure, peer pressure, a student's intrinsic motivation or other personality traits. We assess two channels through which omitted variable bias can occur: one if the omitted variable is correlated with the ordinal rank conditional on absolute ability and one if the omitted variable affects measured ability, which in turn affects the ordinal rank. We assume the true model to be

[1] The extent to which measurement error in the ability variable carries over to the rank variable depends on the shape of the measured ability distribution $h(\tilde{a})$ and the size of the measurement error $z$. Take two students with latent abilities $a_2 > a_1$, but $z_1 > z_2$. Based on latent abilities, student 2 ranks higher than student 1. If $z_1$ is sufficiently large compared to $z_2$, then student 1 may rank higher than student 2 based on measured ability, $\tilde{a_1} > \tilde{a_2}$.

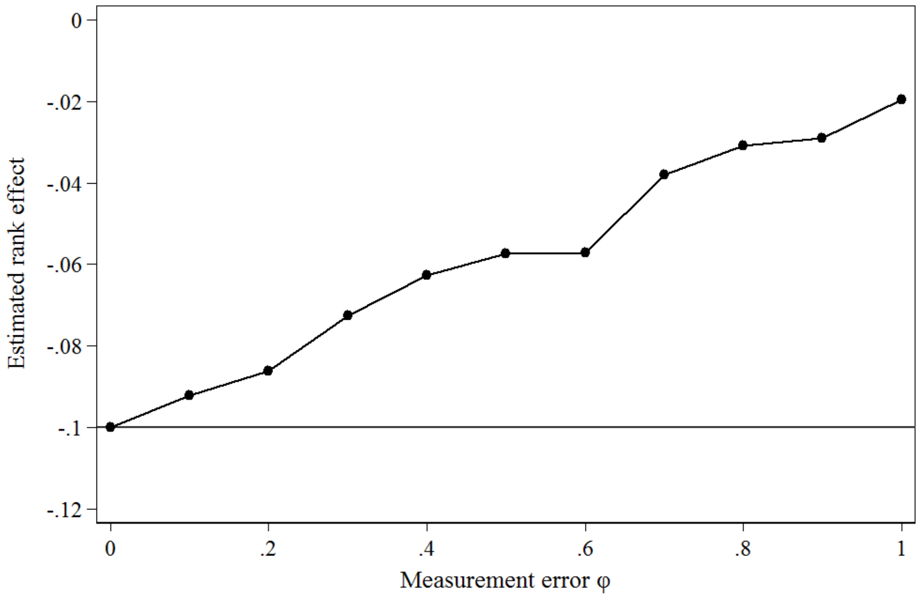


Figure 5: Simulations: measurement error in the Peabody test score that carries over to measurement error in the ordinal rank

*Notes:* This figure plots the simulated effect of rank on risky behaviors as a function of the measurement error in the ability test score. As opposed to Figure 4, the measurement error in the ability variable affects the ordinal ranking. Each point represents the average estimate from 1,000 replications. The assumed true effect is $\gamma = -0.1$.

$$y = \gamma r + \beta a + \rho z + \delta + \varepsilon, \tag{4}$$

where $\varepsilon$ is an i.i.d error term that is uncorrelated with the regressors. The variable $z$ is unobservable and has an impact on the outcome. For the simulations, we assume that $z$ has a negative impact on the outcome, $\rho < 0$.[2]

### E.4.1 Omitted variable correlated with rank

A first source of omitted variable bias could be a confounding factor that is correlated with the ordinal rank conditional on absolute ability and school-cohort fixed effects, i.e. $cov(z, r|a, \delta) \neq 0$. So far, we assumed that the assignment of the ordinal rank is quasi-random conditional on $a$ and $\delta$, allowing us to estimate a causal reduced-form relationship between ordinal rank and risky

[2]This would hold for many but not all potential confounders. For example, students who receive more support from their parents may be less likely to engage in risky behavior.

behaviors. This relationship can be the result of multiple channels. Many variables that come mind as being correlated with rank and having a direct effect on the outcomes represent these channels. For example, parents or teachers may support a child with a low rank more than a child with a high rank, whereby this support may directly affect the likelihood of engaging in risky behavior. However, given that parents or teachers respond *because* a child has been assigned a low rank, this response is not a confounder and does not lead to omitted variable bias.

Theoretically, there could exist omitted factors that affect a student's ordinal rank *conditional on his/her absolute ability*, but we find it difficult thinking of a factor that could plausibly classify as such. By contrast, a more likely — and perhaps more severe — source of omitted variable bias is unobserved variables that simultaneously affect the outcome and the ability test score and thus are also correlated with the ordinal rank. We will discuss these in the next section.

#### E.4.2 Omitted variable correlated with test score

In this section, we discuss two sources of bias whereby an omitted variable $z$ directly affects the outcome as in Equation (4), while at the same time affecting the measured ability test score $\tilde{a}$ as in Equation (3). There are several examples for an omitted variable $z$ that could fulfill these criteria:

- Personality traits. For instance, a student with a higher intrinsic motivation may achieve a higher test score, while motivation may also affect a student's engagement in risky behaviors. Similar cases can be made for more conscientious students, students with more grit, students who are less neurotic, etc.

- Parental or peer pressure. Parents may put pressure on a student to achieve a high test score, whereby this pressure may be correlated with other parental inputs that may directly affect the likelihood of children engaging in risky behaviors. Similarly, peers may influence a student's test performance and have a direct effect on risky behaviors.

While we argue that the setting of the Peabody test — a low-stakes test, with the results not being communicated to students, teachers or parents — makes it unlikely that parental pressure or peer pressure plays an important role, we cannot fully eliminate any of these factors as a source of bias. Therefore, we assess the direction and magnitude of the omitted variable bias for various strengths of the direct effect on the outcome, $\rho$, and the indirect effect through measured ability, $\phi$.

We make an important distinction between two DGPs. In one DGP, we assume that students know the measured test score $\tilde{a}$ and the corresponding rank $\tilde{r}(\tilde{a})$, and both variables determine the outcome in Equation (4). In the second case, we assume that students know their latent ability $a$ and the corresponding rank $r(a)$, whereas we estimate our model based on the measured test score $\tilde{a}$ and the corresponding rank $\tilde{r}(\tilde{a})$. In both cases we run simulations for three parameter values for the indirect effect $\phi \in \{0, 0.25, 0.5\}$, and for 20 parameter values of the direct effect $\rho \in [-1, 0]$. For each $(\phi, \rho)$-pair, we run 10,000 replications, assuming the direct effect of rank on the outcome to be $-0.1$ and the effect of ability on rank to be $-0.6$, as before. It should be noted that the ranges of $\phi$ and $\rho$ cover fairly extreme cases. For example, a direct effect of $\rho = -0.5$ means that the direct effect of $z$ on the outcome is five times as large as the effect of rank on the outcome.

**Case 1) Students base rank on measured ability score** In the first case, we assume that both the absolute ability and the ability rank in the DGP are based on the measured test score, i.e. $a = \tilde{a}$, and $r = \tilde{r}(\tilde{a})$. Figure 6 displays the simulation results. The values of the direct effect are displayed on the horizontal axis. The vertical axis shows the estimated rank effect based on 10,000 simulations. As can be seen, an omitted variable that affects both the test score and the outcome does not bias the estimated rank effect as long as the measured test score and the corresponding rank are what matters for students' outcomes. There is no bias because the entire direct effect of $z$ on the outcome is absorbed by the control for measured ability.

**Case 2) Students base rank on correct ability score** Matters are different when students base their decisions on their latent ability and the corresponding ability rank, whereas we compute the rank based on observed ability. Therefore, the ordinal rank in the DGP differs from the rank used in the estimation. In Figure 7, we show that this setting indeed leads to a bias in the estimates. The estimates are biased towards zero, such that we under-estimate the

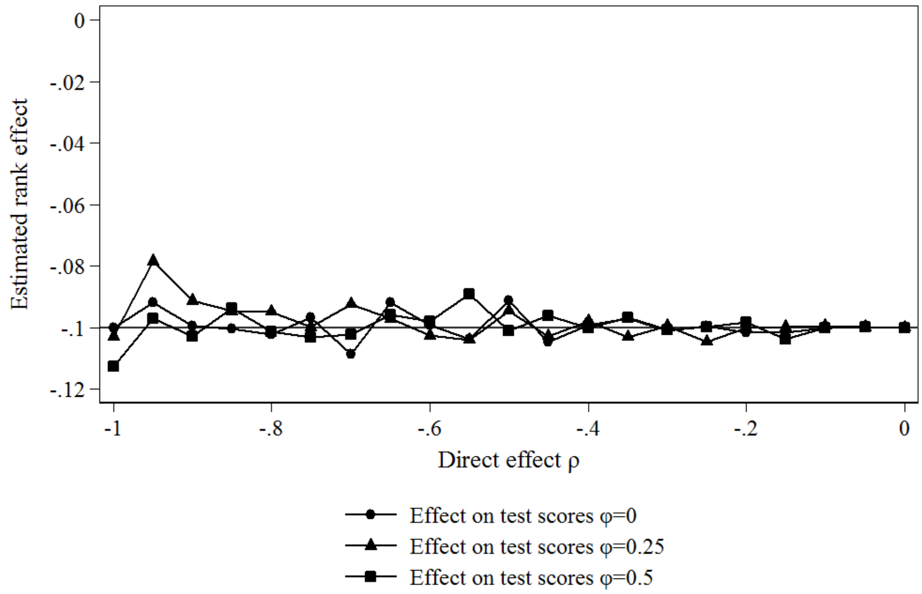


Figure 6: Simulations: omitted variable bias when students base their decisions on the measured ability score and the corresponding rank

*Notes:* This figure plots the simulated effect of rank on risky behaviors as a function of the direct effect of the omitted variable on the outcome, $\rho$. The three lines represent distinct values for $\phi = \{0, 0.25, 0.5\}$. Each point represents the average estimate from 10,000 replications. The assumed true effect is $\gamma = -0.1$.

true effect. The magnitude of the bias depends on the indirect effect. At $\phi = 0$, there is no omitted variable bias, because the omitted variable is uncorrelated with the regressors. If the indirect effect is greater than zero, the estimated rank effect is smaller in absolute value than the true effect. The simulation results become more volatile at very large negative values of $\rho$, although there is a clear pattern showing that the bias becomes larger the higher the indirect effect, whereas the bias does not increase in $\rho$. The result that the bias becomes larger with higher $\phi$ is consistent with the simulations shown in Figure 5, where we interpreted $z$ to be a measurement error. As the simulations show, the same bias occurs if we see $z$ as an omitted variable.

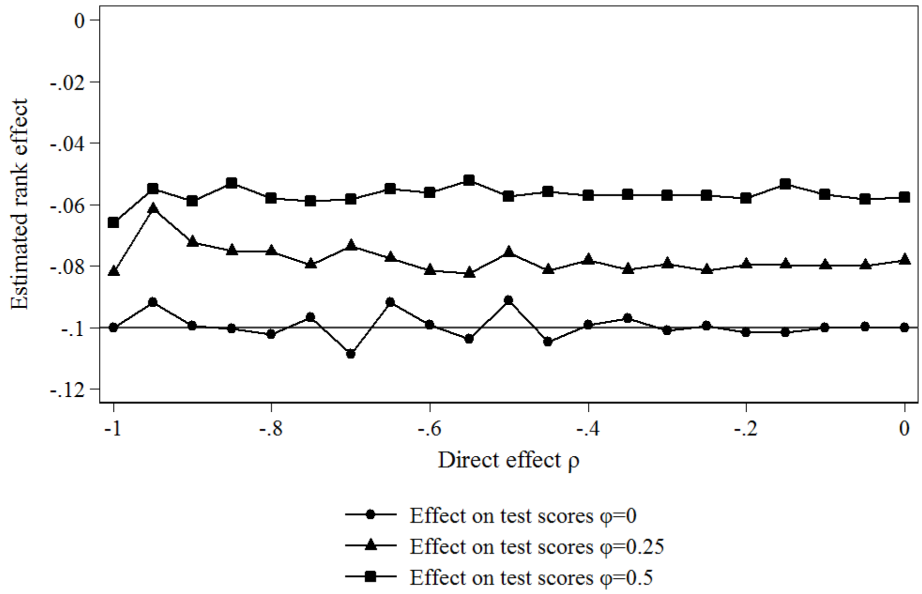


Figure 7: Simulations: omitted variable bias when students base their decisions on their latent ability and the corresponding rank

*Notes:* This figure plots the simulated effect of rank on risky behaviors as a function of the direct effect of the omitted variable on the outcome, $\rho$. The three lines represent distinct values for $\phi = \{0, 0.25, 0.5\}$. Each point represents the average estimate from 10,000 replications. The assumed true effect is $\gamma = -0.1$.

# F Further robustness checks

## F.1 Potential influence of prior rank and reverse causality

One potential threat to identification is that the Peabody score is a function of a student's rank in earlier years, which could cause omitted variable bias. There are at least two arguments against this type of omitted variable bias. First, even if the prior rank affects the ability score, it is unclear that this would affect the *ordinal* ability ranking at the time we measure cognitive ability. Suppose students that had we measure ability in $t = 1$ and suppose that a student with a higher rank in $t = 0$ will have a higher ability score in $t = 1$ than an otherwise-identical student with a lower rank in $t = 0$. Even if the cardinal difference in absolute ability between the two students is greater in $t = 1$ than in $t = 0$, the ordinal difference remains the same. A second argument — discussed in Section 2.2 — is that cognitive ability is formed during childhood and remains stable during adolescence.

To provide further evidence against omitted variable bias, we estimate the model in Equation (2) with school-by-cohort fixed effects, using as dependent variable the Peabody score in wave III of AddHealth. As shown in Column (1) of Table 10, the ordinal rank is not statistically significantly related to the ability score in wave III.

Table 10: Robustness check: prior rank and reverse causality

| | Ability | Moderate behavior | | | Severe Behavior | | |
|---|---|---|---|---|---|---|---|
| | Rank Wave 1 | Drinking | Smoking | Marijuana | Drinking | Smoking | Marijuana |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Peabody score in wave III | 1.063 | 2.366 | 0.803 | 0.524 | 0.262 | -1.687 | 6.186 |
| | (1.119) | (8.029) | (8.028) | (8.037) | (8.036) | (8.018) | (8.029) |
| Number of observations | 9,619 | 9,909 | 9,909 | 9,909 | 9,909 | 9,909 | 9,909 |

*Notes:* This table displays the results from a regression of the Peabody test score in wave III on the variables listed in the column headings. The regression controls for a fourth-order polynomial in absolute ability, individual characteristics and school-by-cohort fixed effects. Regressions are unweighted.

In Columns (2)-(7) of the same table, we provide evidence against reverse causality by regressing the ability score in wave III on previous risky behavior, controlling for a quartic in absolute ability, individual-level controls and school-by-cohort fixed effects. All coefficients are small and statistically insignificant and do not point at any systematic decrease in ability associated to previous risky behavior.

## F.2 Different controls for absolute ability

In the paper, we control for absolute ability with a fourth-order polynomial. By using a higher-order polynomial rather than a linear control, we want to capture the non-linear relationship between risky behaviors and absolute ability.

In Table 11 we assess the robustness of the results with respect to changes in the functional form of the ability controls. In Column (1), we reproduce the baseline results of a model with school-by-cohort fixed effects, as in Table 2, Column (3) in the paper. Alternatively, we control for ability with a linear, quadratic and cubic term, as well as a set of dummies for deciles of the global ability distribution. The latter is a very flexible way of controlling for absolute ability, but requires many degrees of freedom.

While there are some differences in point estimates between specifications, the overall pattern proves robust. Moreover, the specification controlling for a fourth-order polynomial in ability yields very similar results to the specification with decile dummies, indicating that the fourth-order polynomial fits the data well.

Table 11: Results with different controls for ability

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Smoking:** | | | | | |
| on at least 10 days (past 30 days) | -0.124*** | -0.144*** | -0.074* | -0.152*** | -0.121*** |
| *mean = 17.79%* | (0.045) | (0.040) | (0.040) | (0.044) | (0.041) |
| **Drinking:** | | | | | |
| on at least 2 days per month (past 12 months) | -0.138*** | -0.130*** | -0.099** | -0.147*** | -0.140*** |
| *mean = 18.32%* | (0.043) | (0.037) | (0.041) | (0.040) | (0.041) |
| **Marijuana use:** | | | | | |
| at least once (past month) | -0.072* | -0.094*** | -0.047 | -0.092** | -0.059 |
| *mean = 16.09%* | (0.043) | (0.032) | (0.038) | (0.041) | (0.037) |
| **Sex:** | | | | | |
| Intercourse w/o birth control (past 6 months) | -0.055* | -0.096*** | -0.057** | -0.068** | -0.050* |
| *mean = 8.67%* | (0.031) | (0.023) | (0.028) | (0.029) | (0.027) |
| **Crime:** | | | | | |
| Stealing at least once (past 12 months) | -0.061 | -0.081** | -0.025 | -0.051 | -0.067* |
| *mean = 22.63%* | (0.045) | (0.038) | (0.043) | (0.045) | (0.039) |
| Physical fight at least once (past 12 months) | -0.138*** | -0.144*** | -0.118*** | -0.138*** | -0.122*** |
| *mean = 19.53%* | (0.037) | (0.031) | (0.032) | (0.034) | (0.034) |
| Drug selling at least once (past 12 months) | -0.052* | -0.039 | -0.026 | -0.055** | -0.053** |
| *mean = 7.47%* | (0.027) | (0.026) | (0.025) | (0.028) | (0.025) |
| **Ability:** | | | | | |
| Baseline (Quartic) | Yes | No | No | No | No |
| Linear | No | Yes | No | No | No |
| Quadratic | No | No | Yes | No | No |
| Cubic | No | No | No | Yes | No |
| Decile indicators | No | No | No | No | Yes |

*Notes:* This table displays estimates from a model with school-by-cohort effects, with various controls for absolute ability. Column (1) reproduces the baseline effects from Table 2, Column (3) in the paper. In Columns (2)-(5), we control for ability with the functional form indicated at the bottom. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors, clustered at the school level, are reported in parentheses.

## F.3 Rank based on full sample vs rank based on core sample

A potential concern with the sample construction of AddHealth is that minorities have been over-sampled. If minority status is correlated with ability, then the over-sampling introduces a bias in the calculation of the ordinal rank. For example, if minorities have on average a lower ability, then we would assign higher ranks to majority students than they would have under random sampling.

To assess whether over-sampling indeed introduces a bias in the computation of the ordinal rank, we exploit that AddHealth flags people that are part of the randomly chosen core sample, or that have been sampled based on minority status. Figure 8 compares the ability distributions of the core sample and the the full sample. At first glance, both distributions look very similar.

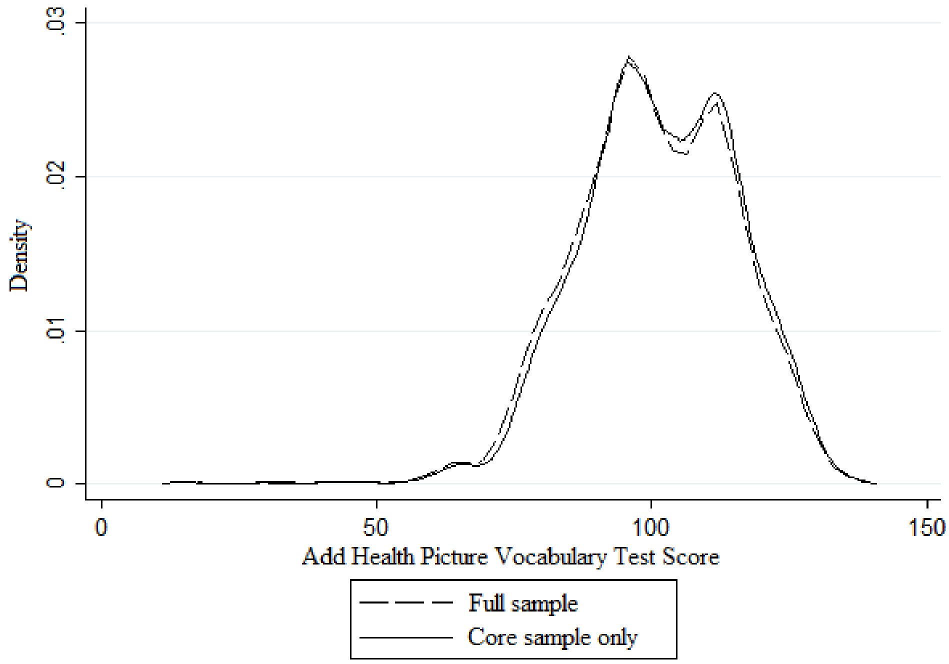


Figure 8: Kernel densities of ability by sample

*Notes:* This graph plots the distributions of the Peabody test score for the core sample and the full in-home sample.

It might be, however, that the distributions are more different within schools. Based on our baseline model with school-by-cohort fixed effects, we perform two tests to see whether the over-sampling biases the estimates. The results are presented in Table 12. Column (1) reproduces the baseline results from Table 2, Column (3), in the paper. As a first test, in Column (2), we estimate the model based on the core sample only. While the number of observations drops by about one third compared to the full sample, we find similar impacts on most outcomes.

As a second test, we use the full estimation sample, but calculate the ordinal rank based on the ability distribution of the core sample. This eliminates any bias from over-sampling. The results, displayed in Column (3), are very similar to those in shown in Column (1), suggesting that over-sampling does not introduce a substantial bias in the estimates.

Table 12: Results using alternative rank definitions

| | (1) | (2) | (3) |
|---|---|---|---|
| **Smoking:** | | | |
| on at least 10 days (past 30 days) | -0.124*** | -0.105* | -0.112*** |
| *mean = 17.79%* | (0.045) | (0.054) | (0.043) |
| **Drinking:** | | | |
| on at least 2 days per month (past 12 months) | -0.138*** | -0.128** | -0.139*** |
| *mean = 18.32%* | (0.043) | (0.054) | (0.041) |
| **Marijuana use:** | | | |
| at least once (past month) | -0.072* | -0.073 | -0.074* |
| *mean = 16.09%* | (0.043) | (0.054) | (0.040) |
| **Sex:** | | | |
| Intercourse w/o birth control (past 6 months) | -0.055* | -0.040 | -0.057* |
| *mean = 8.67%* | (0.031) | (0.034) | (0.030) |
| **Crime:** | | | |
| Stealing at least once (past 12 months) | -0.061 | -0.026 | -0.024 |
| *mean = 22.63%* | (0.045) | (0.054) | (0.045) |
| Physical fight at least once (past 12 months) | -0.138*** | -0.137*** | -0.151*** |
| *mean = 19.53%* | (0.037) | (0.042) | (0.034) |
| Drug selling at least once (past 12 months) | -0.052* | -0.033 | -0.047* |
| *mean = 7.47%* | (0.027) | (0.035) | (0.026) |
| Observations | 12528 | 8090 | 12498 |
| **Controls:** | | | |
| Individual controls | Yes | Yes | Yes |
| Individual ability | Yes | Yes | Yes |
| School × cohort fixed effects | Yes | Yes | Yes |

*Notes:* This table presents the results of robustness checks with different definitions of ordinal rank. Column (1) reproduces the baseline results from Table 2, Column (3). In Column (2), the sample is restricted to the AddHealth core sample. In Column (3), we use the full in-home sample but compute the rank based on the core sample only. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors, clustered at the school level, are reported in parentheses.

## G Misreporting

A potential source of bias is the misreporting of risky behaviors once it is systematically related to a student's ability rank.[3] For example, more highly-ranked students may face a greater stigma in admitting to smoke or drink and thus they may under-report their engagement. In the following, we assess under what conditions misreporting leads to an over- or under-estimation of the true effect. For this purpose, we view misreporting as a non-classical measurement error in the dependent variable. For simplicity, we consider risky behavior as a continuous variable.

Suppose that the true engagement in risky behavior is $y^*$ and assume the true model to be

$$y^* = \beta_0 + \beta_1 r + \beta_2 a + \delta + \varepsilon, \tag{5}$$

with $r$ being a person's ordinal rank and $a$ being his/her cognitive ability. However, due to misreporting, we observe a risky behavior $y = y^* + \eta$, that partly reflects the true risky behavior and partly misreporting $\eta$. Suppose further that misreporting is correlated with rank conditional on own ability and school-cohort fixed effects,

$$\eta = \gamma_0 + \gamma_1 r + \gamma_2 a + \delta + \rho, \tag{6}$$

with $\rho$ being an i.i.d error term. It can be shown that $\gamma_1$ is the bias from misreporting in the estimation of Equation (5) with non-classical measurement error in the dependent variable, i.e.[4]

$$\hat{\beta_1} - \beta_1 = \gamma_1. \tag{7}$$

From this result, we can determine the sign of the bias in the presence of misreporting. Assume that the true effect of ordinal rank on risky behaviors is negative, i.e. $\beta_1 < 0$. If highly-ranked students are more likely to under-report their risky behavior, i.e. if $\gamma_1 < 0$, then we would obtain a positive bias, i.e. in absolute value the estimated coefficient would be larger than the true effect, $|\hat{\beta_1}| > |\beta_1|$. If in turn highly-ranked students are more likely to over-report

[3] Misreporting could also depend on a student's ability, but this would not be a problem because we control for the absolute level of ability in all regressions.

[4] This follows from $\hat{\beta_1} = \frac{cov(r,y|a,\delta)}{Var(r|a,\delta)} = \frac{cov(r,y^*|a,\delta)}{Var(r|a,\delta)} + \frac{cov(r,\eta|a,\delta)}{Var(r|a,\delta)} = \beta_1 + \gamma_1$

their risky behavior, then we would under-estimate the true effect.

To assess the magnitude of the bias, it is helpful to consider the strength of the partial correlation of the rank with the misreporting, $\gamma_1$, relative to the true effect. If $\gamma_1$ is negative and in absolute value is 10% of the true effect $\beta_1$, then we would over-estimate the true effect by 10%. If the partial correlation $\gamma_1$ as large as the true effect $\beta_1$, then our estimated effect would be twice as large as the true effect. While we do not know the true extent of misreporting, its correlation with rank should not be too high, given that students gave the information on risky behaviors without their peers being present and the information was elicited in computer-assisted self-interviews, which gave students little incentive to misreport their behaviors.

## References

Deming, D. and Dynarski, S. (2008). The lengthening of childhood. *Journal of Economic Perspectives* 22: 71–92.

Murphy, R. and Weinhardt, F. (2014). Top of the class: The importance of ordinal rank. *CESifo Working Paper* 4815.