THE ECONOMIC JOURNAL

The Economic Journal, **133** (*January*), 281–322 https://doi.org/10.1093/ej/ueac039 © The Author(s) 2022. Published by Oxford University Press on behalf of Royal Economic Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Advance Access Publication Date: 6 July 2022

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

THE LONG-RUN EFFECTS OF PEERS ON MENTAL HEALTH*

Lukas Kiessling and Jonathan Norris

This paper studies how peers in school affect students' mental health. Guided by a theoretical framework, we find that increasing students' relative ranks in their cohorts by one standard deviation improves their mental health by 6% of a standard deviation conditional on own ability. These effects are more pronounced for low-ability students, persistent for at least 14 years and carry over to economic long-run outcomes. Moreover, we document a pronounced asymmetry: Students who receive negative rather than positive shocks react more strongly. Our findings therefore provide evidence on how the school environment can have long-lasting consequences for individuals' well-being.

Mental health is a growing concern with substantial costs for the economy both in the United States and around the world. In particular, the total costs of mental health disorders are estimated to be as high as 2.5% of the GDP in the United States and 3.5% in Europe (OECD, 2015). Many of these mental health issues can be traced back to symptoms during youth. About 20% of all adolescents suffer from diagnosable mental health disorders (Kessler *et al.*, 2007), and this number increased by a third between 2005 and 2014 (Mojtabai *et al.*, 2016). It is therefore important to understand the causes and long-term consequences of mental health disorders during school age.

In this paper, we study how a feature of the school environment—a student's rank among their school peers—affects mental health, and how these effects evolve over time. We motivate our analysis from a theoretical framework in which students are uncertain about their true ability. Students do, however, have beliefs about their ability that affect study effort. They update their beliefs based on the results of past efforts (e.g., what they have learned about their ability) and also based on information they receive about their ability, such as their rank in their cohort. Receiving negative information about their ability can then lead students to reduce their belief and in turn withdraw effort.

* Corresponding author: Jonathan Norris, Department of Economics, University of Strathclyde, 199 Cathedral Street, Glasgow G4 0QU, UK. Email: jonathan.norris@strath.ac.uk

This paper was received on 2 July 2021 and accepted on 27 June 2022. The Editor was Sule Alan.

The authors were granted an exemption to publish their data because access to the data is restricted. However, the authors provided the Journal with temporary access to the data, which allowed the Journal to run their codes. The codes are available on the Journal repository. The data and codes were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.6587421.

We thank Martin Andersen, Zvonimir Basic, Orla Doyle, Ben Elsner, Lorenz Goette, Johannes Haushofer, Matthias Heinz, Ingo Isphording, Sebastian Schneider and Matthias Sutter, as well as audiences at SOLE-EALE-AASLE World Conference 2020, Verein für Socialpolitik 2020, Essen Health Conference 2020, 2nd Workshop on Health and Inequality at Copenhagen Business School, and University of Regensburg for helpful comments and feedback. This research uses data from Add Health, a programme project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with co-operative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis. The views expressed in this paper are those of the authors and do not necessarily reflect the views of Comma Soft, the MPI or the University of Strathclyde.

We follow a rich psychological literature and interpret this belief-updating process as a key mechanism explaining how shocks in school affect mental health. In fact, leading cognitive (e.g., Beck, 1967) and attributional theories (Seligman, 1972) of depression as well as newer neurocognitive models (Clark *et al.*, 2009) emphasise the crucial role of incorrect beliefs as a source of depression. Moreover, many common depressive symptoms can be seen either as manifestations of these biased beliefs or as direct consequences thereof (De Quidt and Haushofer, 2016).

Based on this framework, we derive four predictions: First, negative shocks to performance in school decrease a student's belief about their own ability, while positive shocks increase this belief. Second, because students refrain from exerting effort once their beliefs about their own ability are sufficiently low, they do not receive any new informative signals, implying that negative shocks have stronger effects than positive shocks. Third, the consequences of shocks are more pronounced at the lower end of the ability distribution, as students attribute shocks relatively more to their own ability. Fourth, these shocks can have persistent effects over time.

To test these predictions, we exploit idiosyncratic variation in the ability composition of school peers as shocks to students' mental health. More specifically, for our preferred specification, we argue that conditional on school-by-grade fixed effects and students' own ability that ranks within a school-grade across schools are as good as random. We thus compare the mental health of students having the same ability and characteristics, but who happen to have different ranks due to differences in their peers' ability. We use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a survey of a representative sample of US adolescents in grades 7–12 who were first interviewed during the 1994/5 academic year (Harris, 2018). Importantly, Add Health repeatedly assesses students' mental health using a well-established measure to diagnose depressions (Center of Epidemiological Studies Depression Scale, CES-D; Radloff, 1977). Based on this measure, we further motivate our analysis by documenting that mental health seems to be malleable before the age of 20, i.e., when respondents are still in school, and show that mental health stays persistent into adulthood.

We find that shocks from students' ranks among their peers affect mental health. Increasing a student's rank by one standard deviation on average improves their mental health by 6% of a standard deviation (SD). This effect is large. It is of similar size as the difference in the mental health of children from college and non-college-educated households, or two-thirds of the difference between students raised by both parents rather than a single parent. Furthermore, our effect size amounts to about one-sixth to one-third of average effect sizes reported for pharmacological treatments and positive psychology interventions targeted at well-being and depressions (e.g., Mitte *et al.*, 2005; Turner *et al.*, 2008; Bolier *et al.*, 2013).[1] Note, however, that our effect stems from natural variation in the ability composition of school cohorts rather than from targeted treatments. We confirm our results with a series of robustness checks that examine functional form assumptions, probe the sensitivity of our estimates and assess several sources of measurement error. Throughout, we find strong and robust support for the central prediction of our model that shocks in school from variation in ranks among peers affect mental well-being.

Subsequently, we examine what is driving these effects. We find pronounced rank effects on students' beliefs about their relative intelligence and on expectations for whether they will attend

[1] Mitte *et al.* (2005) and Turner *et al.* (2008) conduct meta-analyses of anti-depressants and pharmacological treatments and find average effect sizes of about 0.31 SD. Similarly, Bolier *et al.* (2013) examines positive psychology interventions, e.g., trainings, exercises and therapies aimed at increasing positive feelings, and find effect sizes of 0.20–0.34 SD for different measures of depression and psychological well-being.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

college, supporting our interpretation of beliefs as a key mechanism for how shocks translate into mental health. We provide further support for our modelling choices by exploring specific facets of mental health underlying the CES-D measure. Grouping items of the CES-D scale using principal component analysis, we observe that rank effects mainly operate through a lack of positive attitudes and lack of motivation, rather than affecting loneliness and attributions to other external or social factors.

We then explore the second prediction of our theoretical framework, an asymmetry of positive and negative shocks. In line with our conjecture, we show that our results are driven by negative shocks, i.e., having a lower rank in one's school cohort than across students from all schools and cohorts, rather than positive shocks where students have a better rank than expected.

Moving to our third prediction, we document that rank effects are indeed larger at the lower end of the ability distribution. Students from the lowest decile of the ability distribution experience a rank effect that is about three times larger than the average effect, and that amounts to approximately 0.70 SD. The rank effect then slowly fades out for higher deciles.

Next, we explore how these heterogeneous effects in own ability evolve over time. Rank effects are persistent and last from adolescence to adulthood. In fact, we observe that rank effects for low-ability students are at least as large in Wave IV, 14 years after baseline when individuals in our dataset are 26–32 years old, as they are in Wave I, when they were adolescents, and fade out for higher deciles.

Finally, we investigate whether these effects carry over to other economic long-run outcomes. We find that better mental health in adolescence is highly predictive of better economic long-run outcomes in adulthood. While correlational, there are significant and sizeable associations of CES-D scores and college graduation, employment status, as well as income, ever being married, and ever being arrested conditional on a rich set of background characteristics. For example, a standard deviation worsening of mental health is associated with a decrease in income as an adult of approximately a third of the gender gap. These correlations therefore document the high predictive power of mental health status in youth for achievements later in life. We also estimate the causal effect of ranks among school peers on these long-run outcomes. Notably, we find the same heterogeneous pattern in ability that we observed for mental health, with low-ability individuals showing more pronounced effects than those having higher ability. Moreover, we estimate that these effects are partly mediated by mental health in adolescence.

Our study relates to the broad literature examining the causes and consequences of poor mental health. Previous research has shown that short-term variations in income induced by cash transfers (e.g., Baird *et al.*, 2013; Haushofer and Shapiro, 2016), better residential neighbourhoods (Katz *et al.*, 2001) or improved early-life circumstances (Adhvaryu *et al.*, 2019) relate to better psychological well-being in adolescence and adulthood. Yet, the relationship between economic outcomes and mental health is bi-directional as mental disorders can also be a cause of important economic outcomes. For instance, low mental health limits human capital accumulation (Currie and Stabile, 2006; Fletcher, 2010), and reduces employment and earnings (Bartel and Taubman, 1986; Frank and Gertler, 1991; Ettner *et al.*, 1997; Stewart *et al.*, 2003; Fletcher, 2014; Biasi *et al.*, 2021). We contribute to these previous studies by highlighting the consequences of the peer composition in schools for adolescents' psychological well-being in the short and long run. Thereby, we propose mental health as an important channel through which shocks in childhood and adolescence affect economic outcomes in adulthood.

This paper adds to an accumulating evidence base on the long-lasting effects of the school environment in general, and peers in particular. Although analysing the long-run effects of smaller

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

classes (Angrist and Lavy, 1999; Krueger and Whitmore, 2001; Chetty *et al.*, 2011; Angrist *et al.*, 2019) or better teachers (e.g., Chetty *et al.*, 2014; Rothstein, 2017) are long-standing and active fields of research, only recently have studies shed light on the long-term effects of peers during school. Carrell *et al.* (2018) document that having disruptive peers during childhood decreases earnings, while Olivetti *et al.* (2020) show that the labour supply of school peers' mothers affects females' labour force participation in adulthood. Furthermore, several studies suggest that peers in school have effects on educational attainment as well as major choices (as in, e.g., Gould *et al.*, 2009; Bifulco *et al.*, 2011; Black *et al.*, 2013; Bifulco *et al.*, 2014; Anelli and Peri, 2019) and the formation of non-cognitive skills (Bietenbeck, 2020). We expand this literature by tracking peers' short- and long-run effects on mental health, and simultaneously uncover similar patterns across a broad range of economic outcomes. By focusing on long-term effects for mental health, we also extend the small literature on contemporaneous peer effects in mental health, documenting large spillovers in some studies (e.g., Fowler and Christakis, 2008, on contagion in social networks) and more modest (Eisenberg *et al.*, 2013, focusing on spillovers among roommates) or zero effects in others (Zhang, 2018, on linear-in-means peer effects in classrooms).

Because we exploit a specific shock based on students' ability rank in their cohort, our study also relates to a growing literature on rank effects as a specific form of peer effects. This literature argues that ordinal ranks affect outcomes due to social comparisons and describes this as 'big fish in a little pond' effects (Festinger, 1954; Marsh, 1987). In particular, there is evidence that such comparisons affect the job satisfaction or general well-being of individuals (Luttmer, 2005; Brown *et al.*, 2008; Card *et al.*, 2012; Azmat and Iriberri, 2016). Recent papers have used these ideas in educational contexts similar to ours, to estimate the effect of ordinal ranks on educational outcomes (Elsner and Isphording, 2017; Murphy and Weinhardt, 2020; Delaney and Devereux, 2021; Elsner *et al.*, 2021), subsequent earnings (Denning *et al.*, 2021), risky behaviours and bullying (Elsner and Isphording, 2018; Comi *et al.*, 2021) and skill development (Pagani *et al.*, 2019).

While our paper shares a common empirical strategy and, in some cases, data with previous papers from the rank effects literature, our results and focus differ significantly. Perhaps closest to our paper is a study by Elsner and Isphording (2017), who study rank effects on educational outcomes using the same data source. We start from a similar baseline specification, but our results and focus differ significantly. Our analysis is motivated by a learning model in which students learn about their ability and update their beliefs in the process. In the presence of negative information shocks, belief updating can then translate into poorer mental health.

Based on this new mechanism, we are then able to identify a novel margin of rank effects on mental health, introduce asymmetric effects suggesting that negative shocks are more influential and trace the evolution of our effects over time. In addition, we show that they are more pronounced at the lower end of the ability distribution, with zero effects at the top end of the distribution. These patterns, particularly that there could be detrimental consequences, are consistent with our stylised belief-updating model, but cannot be rationalised with common 'big fish in a little pond' mechanisms. Thus, our results contrast previous interpretations that rank effects exclusively operate at the top end of the ability distribution through access to better schools, colleges, and jobs. We discuss this further in Subsection 5.7, summarising our proposed model and results in comparison to alternative mechanisms previously discussed in the literature.

Taken together, our results provide evidence on how the school environment can have long-lasting effects on mental health. We show that shocks not only affect immediate mental health—especially for students of low ability—but have persistent effects over at least 14 years. Therefore,

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

our results lend support to models that introduce mental health capital similar to general health (Grossman, 1972), or as a malleable skill in line with the growing evidence on the importance of non-cognitive skills in the development of children (Cunha and Heckman, 2008; Cunha *et al.*, 2010). Such models could explain why we observe long-lasting effects carrying over to educational and labour market outcomes, which is similar to findings for non-cognitive skills (e.g., as shown by Heckman *et al.*, 2006). Additionally, our results highlight two important points. First, ranks can have negative, not just beneficial, consequences that should be considered. Second, some features of school environments may be both unavoidable—such as ranks—and, at times, harmful, suggesting that mitigation efforts can form important policy tools to alleviate harm where the source is difficult to remove.

The remainder of this paper is structured as follows. In Section 1, we lay out a theoretical framework in which students are uncertain about their ability and base their learning process on noisy signals. We present the data we use to test our predictions in Section 2 and describe mental health patterns over the life cycle in Section 3. Our empirical strategy and main results for Wave I are presented in Sections 4 and 5, before we study the persistence of our results in Section 6. Finally, Section 7 concludes.

1. Theoretical Framework

Before we turn to our empirical analysis, we want to fix ideas and outline a stylised belief-updating model to highlight how beliefs could act as a particular channel for information shocks to affect students' mental health. The theoretical framework is motivated by cognitive (e.g., Beck, 1967) and attributional theories of depression (e.g., Seligman, 1972), which emphasise the crucial role of biased or incorrect beliefs as a source of depression. Furthermore, cognitive neuropsychological models of depression (Clark *et al.*, 2009; Roiser *et al.*, 2012) suggest a central role for negative affective biases. We closely follow De Quidt and Haushofer (2016), and conceptualise the belief about the returns to one's own ability as a key mechanism translating shocks to mental health.[2] In fact, many common symptoms of depression, such as pessimism, low self-esteem, lack of motivation or sadness, can be seen either as a manifestation of biased beliefs (as in the case of pessimism) or as a direct consequence thereof (as for the lack of motivation).

Our model builds on two features. First, students are uncertain about the return to effort and learn about their ability based on their performance in school (for evidence of students' imperfect knowledge about their ability and return to effort, see, e.g., R. Jensen, 2010; Zafar, 2011; Stinebrickner and Stinebrickner, 2012). Second, there are exogenous shocks affecting school performance and students use their performance to update their prior about their ability.[3] Thus, after receiving negative signals, students reduce their belief about their ability, while positive signals increase their beliefs. Yet, updating only occurs if students exert effort to receive a good grade. If they shirk, they will not attribute their educational success to their ability. Thus, if the prior about ability is sufficiently low, students may refrain from exerting any effort to avoid further negative signals. This implies that a low belief about one's own ability or a low mental health status may constitute an absorbing state, in which no further updating occurs.

[2] In their model, De Quidt and Haushofer (2016) focus on income shocks in developing countries and additionally consider non-food consumption, food, as well as sleep and other domains entering the utility function. Relative to their model, ours can be seen as a simplification, translating their model to an educational context.

[3] If teachers, e.g., grade their students on a curve, students' ranks among their peers can directly affect their GPA and school performance. In our empirical application, we will therefore exploit quasi-random variation in students' ranks among their peers as such information shocks.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

To formalise this intuition and derive more precise predictions, we adopt a simple educational production function in which 'school success' depends on own ability, time spent studying, the study intensity or study effort, as well as exogenous shocks. More specifically, let A_i denote a student's ability or return to effort which is drawn from some distribution F_A. For the ease of exposition, we keep the individual index i implicit and present the model for a single individual with ability A. Let y_t denote the 'school success' in period t. We specify the educational production function as

$$y_t = [Ae_t + \underline{A}(1 - e_t)]s_t + \epsilon_t,$$

in which school success depends on the amount the student studies, s_t, and given they invest time to study, their decision to exert high ($e_t = 1$) or low effort ($e_t = 0$).[4] High effort yields a return to studying equal to their ability A, while shirking yields a low return of $\underline{A}$, assumed to be known to the student. Moreover, school success is subject to exogenous shocks $\epsilon_t \sim N\left(0, \sigma^2_{\epsilon,t}\right)$. In the empirical part of our paper, we will use shifts in the rank of a student due to having better classmates as a shock to the school success y_t.[5]

Students are uncertain about their own ability and have priors or beliefs about their ability denoted by μ_t. Hence, from a student's perspective, their ability is a random variable $A \sim N\left(\mu_t, \sigma^2_{A,t-1}\right)$. We assume that students maximise their expected utility by allocating time between studying and leisure. While studying increases educational success, leisure also enters positively into the utility function. Their expected decision utility function is given by

$$EU\left(e_t, s_t, l_t | \mu_t\right) = \underbrace{\left[\mu_t e_t + \underline{A}\left(1 - e_t\right)\right]s_t}_{\text{Exp. school success}} + \underbrace{\phi\left(l_t\right)}_{\text{Utility from leisure}},$$

in which s_t and l_t denote study and leisure time, respectively; total time available is normalised to 1, such that $s_t + l_t = 1$, and $\phi'(\cdot) > 0$, $\phi'' \leq 0$. Expected school success depends on the prior about ability, μ_t, and the decision to exert effort, e_t, as described above.

Given this setup, a student's optimal effort decision is

$$e^*_t = \mathbb{1}\left\{\mu_t > \underline{A}\right\},$$

and, hence, we can replace $[\mu_t e^*_t + \underline{A}(1 - e^*_t)] = \max\{\mu_t, \underline{A}\}$. The optimal time spent studying therefore equals

$$s^*_t = 1 - \phi'^{-1}\left[\max\left\{\mu_t, \underline{A}\right\}\right],$$

and is increasing in perceived own ability.

We now want to characterise how students learn about their ability. Consider students who want to update their prior beliefs μ_{t-1} given they received a signal y_{t-1} about their own ability. If the students only exerted low study effort, $e_{t-1} = 0$, they do not learn new information about their ability A as studying yields a fixed return $\underline{A}$. However, if they exerted high effort ($e_{t-1} = 1$),

[4] Thus, students face decisions along two margins: they decide about a quantity of studying (e.g., how many hours are spent on studying rather than on leisure) and the quality of studying (focusing on homework or being constantly distracted by, e.g., their smartphones).

[5] As we will explain in detail in Section 4, we exploit the fact that schools only have limited size and that the ability composition of students varies across cohorts and schools. This implies that a student with a specific ability may be ranked highly in one cohort, but would only be a student in the middle of the ability distribution in another cohort. We use this variation in the ability distribution as an exogenous shock affecting students' beliefs after conditioning on own and peer ability.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

they can learn about their ability. In that case, we can rewrite y_{t-1} as follows:

$$y_{t-1} = As_{t-1} + \epsilon_{t-1} =: x_{t-1} + \epsilon_{t-1},$$

where $x_{t-1} = As_{t-1} \sim N(\mu_{t-1}s_{t-1}, \sigma^2_{A,t-1}s^2_{t-1})$. Given the signal about school success, y_{t-1}, the students try to learn about their ability, A. Using the new notation, they want to infer the expected value of x_{t-1} given y_{t-1}, i.e., the posterior $E[x_{t-1}|y_{t-1}]$:

$$\begin{aligned}
E[x_{t-1}|y_{t-1}] &= \frac{\sigma^2_\epsilon}{Var(x_{t-1}) + \sigma^2_\epsilon} E[x_{t-1}] + \frac{Var(x_{t-1})}{Var(x_{t-1}) + \sigma^2_\epsilon} y_{t-1} \\
&= E[x_{t-1}] + \frac{Var(x_{t-1})}{Var(x_{t-1}) + \sigma^2_\epsilon}(y_{t-1} - E[x_{t-1}]) \\
&= \mu_{t-1}s_{t-1} + \frac{\sigma^2_{A,t-1}}{\sigma^2_{A,t-1} + \sigma^2_\epsilon / s^2_{t-1}} \left[\left(A - \mu_{t-1}\right) s_{t-1} + \epsilon_{t-1} \right].
\end{aligned}$$

Hence, the corresponding posterior belief μ_t equals

$$\mu_t = \mu_{t-1} + \frac{\sigma^2_{A,t-1}}{\sigma^2_{A,t-1} + \sigma^2_\epsilon / s^2_{t-1}} \left[\left(A - \mu_{t-1}\right) + \frac{\epsilon_{t-1}}{s_{t-1}} \right].$$

Given that we study mental health through a belief mechanism, the posterior belief about one's own ability serves as a proxy for mental health as in De Quidt and Haushofer (2016). Next, we investigate how mental health changes through factors shifting this posterior belief.

Several results emerge. First, negative shocks ($\epsilon_{t-1} < 0$) decrease students' beliefs about their ability (i.e., μ_t decreases) and thus have detrimental effects on mental health, while positive shocks ($\epsilon_{t-1} > 0$) benefit mental health.

PREDICTION 1. *Positive shocks improve mental health, whereas negative shocks decrease mental health.*

Second, once a student's belief μ_t decreases below $\underline{A}$, the student withdraws effort and thus stops updating.[6] This implies that negative shocks may have more pronounced consequences relative to positive shocks, as negative shocks decrease the likelihood of a student receiving informative signals in the future. This asymmetry is also consistent with recent empirical evidence on the effect of performance feedback (see Villeval, 2020, for a recent survey of the literature). In response to achievement information, low-achieving students reduce their subsequent performance to a greater degree than high-achieving students increase their performance (Goulas and Megalokonomou, 2021).

PREDICTION 2. *There are asymmetric effects of positive and negative shocks, with the latter being more pronounced.*

Third, the students' study time, s_t, (weakly) decreases in the belief about their ability and low-ability students have lower priors μ_{t-1}. This implies that shocks have stronger effects for low-ability students as the term ${}^{\epsilon_{t-1}}/_{s_{t-1}}$ becomes larger.

[6] Barankay (2011) shows that workers receiving feedback about their rank (i.e., an information shock in our framework) are less productive and are less likely to return to work. Evidence in line with this mechanism in relation to mental well-being has also been found in the psychology literature. Kuppens *et al.* (2010) show that individuals with low self-esteem or depressions display high levels of emotional inertia in response to emotional fluctuations relative to individuals with normal levels of self-esteem and no depressions. Relatedly, Korn *et al.* (2014) document that depression is related to more pessimistic belief updating.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

PREDICTION 3. *The consequences of shocks are more pronounced for low-ability individuals.*

Fourth, given the lower propensity to update after receiving a negative shock and stronger effects for low-ability students, shocks have persistent effects over time and especially so for low-ability students with priors close to $\underline{A}$.

PREDICTION 4. *The effects of shocks are persistent over time. They are more pronounced for students with low ability.*

In summary, this stylised theoretical framework predicts that if beliefs are a key mechanism translating shocks to mental health—as suggested by leading psychological theories of depression—and students have imperfect knowledge about their ability, we should expect that negative shocks to students' school success decrease their beliefs about their ability and eventually their mental health, those effects are more pronounced in the lower part of the ability distribution, and persist over time. In the empirical part of our paper, we will test whether these predictions hold for mental health.[7]

2. Data

In order to test the predictions from the previous section, we use restricted data from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a longitudinal study of a set of representative middle and high schools in the United States. For our analysis, the Add Health dataset has several key features. First, it covers multiple cohorts within schools, which we need for our empirical strategy exploiting variation within schools across cohorts. Second, a representative set of students from each cohort is sampled. Third, students were first interviewed in 1994/5, when the majority of students were between 12 and 18 years old, and followed for five waves until 2016–18, when respondents were 36–42 years old. Hence, we can follow the development of adolescents' well-being well into adulthood. Fourth, the dataset has a standardised test of cognitive ability and repeated measures of an established mental health self-assessment, allowing us to trace the evolution of and effects on mental health over time. In the following, we discuss the mental health measure in more detail and defer the discussion of the ability measure to Section 4, where we discuss our empirical strategy.

2.1. *CES-D Scores as a Measure of Mental Health*

We assess mental health of students using the Center of Epidemiologic Studies Depression Scale (CES-D, Radloff, 1977), an established screening measure to test for depression and depressive disorders that is one of the most widely used instruments in psychiatric epidemiology. The CES-D consists of 19 symptoms (e.g., 'You felt sad') and asks respondents how often each symptom applied to them over the course of the past week. Responses are then rated on a scale from 0

[7] We acknowledge that we laid out a highly stylised model of mental health with the strong assumption of fixed ability over time. In principle, one could think of richer models that, e.g., allow effort to create stimulation for the brain eventually improving ability. In such a model, a negative shock and subsequent effort withdrawal would reduce ability. In the next period, the child would then have lower ability and likely an even lower rank, further amplifying the dynamics described in the model outlined above. We think that such a refined version of the model could generate further implications that are beyond the scope of the present paper (e.g., what happens in combination with policy interventions such as school counselling or when moving from primary to secondary school). Rather, we view our model as a framework to show how our empirical findings can be organised by a simplistic belief-updating model, following previous modelling attempts by De Quidt and Haushofer (2016).

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

('never or rarely') to 3 ('most of the time or all of the time') and aggregated to a final score ranging from 0 to 57, with higher scores indicating a higher propensity for depressive symptoms. Online Appendix Table A.1 presents all items of the CES-D score and Online Appendix Figure A.1 shows the distribution of CES-D scores in Wave I.

The CES-D scale is a widely used instrument to study mental health. It has been adopted to study how far an individual's mental health status spreads through a social network (Fowler and Christakis, 2008; Rosenquist *et al.*, 2011), the effect of mental health for educational attainment (Fletcher, 2008; 2010), and the consequences of wealth shocks (Schwandt, 2018), cash transfers (Haushofer and Shapiro, 2016) or religion (Fruehwirth *et al.*, 2019) on mental health. Moreover, a rich literature in psychology and psychiatric epidemiology has examined the concurrent validity (i.e., the extent to which the CES-D and a subsequent diagnosis coincide; e.g., Lewinsohn *et al.*, 1997), reliability, and internal consistency of the CES-D scale (e.g., Roberts *et al.*, 1990; Radloff, 1991), and it is frequently used in clinical practice (see, e.g., Murphy, 2011, for a review).

We further check the link between the CES-D scores in Add Health at Wave I with several measures related to mental well-being. Panel A of Online Appendix Table A.2 documents strong associations between CES-D scores measured in Add Health and receiving counselling in the past year, anxiety, self-esteem and suicidal ideation. These correlations are consistent with the validity of the scale for capturing mental health that has been established in the psychological literature.

In the main part of our analysis, we focus on the 19-item CES-D scale as a measure of mental health. Yet, later waves only administered a short scale comprised of a subset of the original items. Thus, when studying the persistence of our results, we scale the CES-D scores of later waves to obtain a comparable measure across waves. To perform the rescaling, we scale the nine (ten) item scales of Wave III (IV) by $^{19}/_{9}$ ($^{19}/_{10}$) to match the 19 item scale administered in Waves I and II. We confirm that these short scales are empirically comparable to the Wave I scale. First, using data from Wave I, Online Appendix Figure A.2 shows that the short and long versions of the CES-D scale are indeed highly correlated ($\rho = 0.95$ and 0.96 for the 9 and 10-item scale, respectively). Second, we repeat the associations between CES-D and Wave I counselling, anxiety, self-esteem, and suicidal ideation reported in panels B and C of the Online Appendix Table A.2. We find that both the Waves III and IV rescaled short versions continue to be highly correlated with these measures.

2.2. *Summary Statistics*

Table 1 presents summary statistics of our sample in Wave I. After dropping observations from schools with fewer than 20 students in total and 5 students per grade, we observe 18,459 students in Wave I, where 51% of these students are female and they are on average 15.6 years old. The majority of students are White (53%), about 20% report at least one foreign born parent, and 34% of all students come from college-educated households with average household incomes of $46,000. Moreover, the mean CES-D score in our sample is 11.3.

3. Stylised Facts on Mental Health Over the Life Cycle

We begin by documenting the evolution and persistence of mental health over the life cycle using the rich information from the Add Health study. In particular, we show that mental health manifests itself early in life and stays persistent over the life cycle. These stylised facts highlight

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 1. *Summary Statistics.*

	Mean	SD
Female	0.51	0.50
Age	15.63	1.70
White	0.53	0.50
Child of an immigrant	0.21	0.41
Number of older siblings	0.52	0.76
College-educated parents	0.34	0.47
Single-parent household	0.32	0.47
Household income (1000s)	46.33	45.88
Mental health (CES-D scores)	11.33	7.60
Ability (AHPVT scores)	100.17	14.67
Grade 7	0.13	0.34
Grade 8	0.13	0.34
Grade 9	0.18	0.38
Grade 10	0.20	0.40
Grade 11	0.19	0.39
Grade 12	0.16	0.37
Observations	18,459	–

Notes: This table presents summary statistics for the sample in Wave I of Add Health after dropping observations from schools (grades) with fewer than 20 (5) students.

the importance of studying the features of the school environment as determinants of mental health, thus motivating our subsequent analysis.

3.1. *Evolution of Mental Health over the Life Cycle*

Add Health data at Wave I covers respondents from different ages ranging from 11 to 19 years. Although there are several years between the data collections of different waves, there is a partial overlap in ages covered by different waves. This allows us to aggregate age-specific mental health measures across waves. We do this using the rescaled CES-D scores from later waves, as we defined in Section 2, because later waves employ a subset of the full scale (Online Appendix Figure A.4 replicates the results in this section using weights from OLS regressions in Wave I rather than rescaled CES-D scores with identical results). Moreover, we restrict the sample to those respondents whom we observe across Waves I, III, IV and V to present the evolution and persistence of mental health based on a balanced panel. To increase precision, we aggregate age groups into two-year bins and trace the evolution of mental health measured by CES-D scores from adolescence through midlife.

Figure 1 displays the average evolution of CES-D scores over time. We observe that CES-D scores increase, and hence mental health deteriorates until the age of 20, i.e., during the time when respondents are still in school, and stabilises afterwards. In the Online Appendix Figure A.3, we present analogous figures and differentiate the evolution of mental health by gender, ethnicity, parental education, and whether respondents were raised in a single-parent household. While there is cross-sectional variation in mental health with females, non-White people and respondents with lower socioeconomic status having higher CES-D scores and thus worse mental health, the evolution over the life cycle is similar across subgroups.[8] In particular, we observe the same steep increase in CES-D scores until age 20 and a relatively flat pattern afterwards for all subgroups.

[8] In Online Appendix Table A.3, we quantify these cross-sectional differences with regressions relating CES-D scores in Wave I to observable characteristics of students. Females have 0.25 SD higher CES-D scores than males, while differences between Whites and non-Whites, between children from college-educated and non-college-educated

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Fig. 1. *Evolution of Mental Health Over the Life Cycle. This figure presents mean CES-D scores at 2-year age bands after pooling over Waves I, III, IV and V, and controlling for survey wave effects. Higher CES-D scores correspond to worse mental health. The shaded area indicates 90% confidence intervals.*

This observed increase in the prevalence of depressive symptoms over the adolescent period is consistent with a number of observations and hypotheses from the behavioural sciences. Thapar *et al.* (2012) summarise the literature and describe four broad mechanisms giving rise to the increase in mental illnesses during adolescence: (*i*) family and genetic factors (e.g., inherited risk factors), (*ii*) psychosocial risk factors (e.g., exposure to stressful life events), (*iii*) gene-environment interplay (e.g., genetic pre-disposition with increasing sensitivity to adversity) and (*iv*) brain development and hormonal changes with the onset of puberty (e.g., development of emotional regulation during adolescence creating stronger emotional responses; see also Ahmed *et al.*, 2015).

Adverse life events may have an important role in adolescent depression because of heightened brain activity in the reward and danger-sensitive regions. Exposure to adversity does not necessarily lead to depression. However, there is considerable evidence in the literature that family adversity, bullying, peer rejection and a wide range of additional possible stressors, can prompt the onset of depression and have long-term consequences into adulthood (Thapar *et al.*, 2012; McCormick and Green, 2013). Groups at risk appear to be those who experience multiple adverse events, along with girls, in particular, who appear to exhibit greater differences in brain activity, which further enhances their risk (Thapar *et al.*, 2012). This is consistent with our observation that girls and groups likely to have experienced multiple adverse events (e.g., minorities and those with less educated parents) tend to have elevated depressive symptoms.

3.2. *Persistence of Mental Health Over Time*

While Figure 1 shows that average mental health remains relatively stable after the age of 20, it does not tell us about persistence on the individual level. We therefore provide further evidence on the persistence of mental health by studying the relation between CES-D scores in subsequent waves. Figure 2 presents the distribution of CES-D scores in Waves I, III, IV and V, plotted against the corresponding scores in the previous wave, including linear and non-parametric fits.

households, and between children raised in two-parent and single-parent families range between 6 and 10% of a standard deviation.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

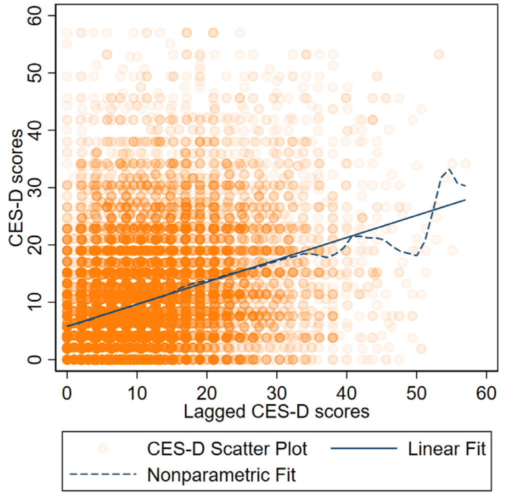

Fig. 2. *Persistence in CES-D Scores over Time. This figure presents a scatter plot of CES-D and lagged CES-D scores, as well as linear and non-parametric fits. We pool across Waves I, III, IV and V (note: Wave II is omitted because high school leavers in Wave I are not sampled during Wave II), resulting in a time lag of 7 (Waves I to III, and III to IV) and 9 years (Waves IV to V).*

We observe a strong autocorrelation of 0.39 in CES-D scores across waves. To put this number into perspective, we compare this autocorrelation to test–retest statistics of CES-D scales. For instance, Roberts *et al.* (1990) report one-month test–retest correlations of 0.49 and 0.60 for boys and girls. Given that the lags between the Add Health's waves are about seven to nine years, this persistence in mental health is remarkable.

The long-term persistence in mental health documented here is consistent with evidence in the psychological literature. Depressive symptoms during adolescence are linked with a high probability of depressive problems later in life (Thapar *et al.*, 2012). However, to our knowledge, very few studies have provided evidence based on nationally representative longitudinal data. In a representative sample of a cohort in New Zealand followed from early to midlife, Caspi *et al.* (2020) document that 59% of their sample experienced an onset of a mental health disorder by the close of adolescence. Moreover, those individuals who experienced an onset during adolescence also exhibited a greater number and variety of symptoms over time. Relatedly, Plana-Ripoll *et al.* (2019) report strong comorbidity over at least 15 years (i.e., an increased risk of developing further mental disorders after a first one) and Momen *et al.* (2020) extend these findings, highlighting an increase for subsequent medical conditions, independently of mental health disorders. Again, the strong persistence we observe is consistent with the limited evidence available.

Together, Figures 1 and 2 provide a first indication that the school environment may have a lasting effect on the mental health of students: Mental health seems to be particularly malleable during adolescence and displays a strong persistence over time.

4. Empirical Strategy

In order to test the predictions of the theoretical framework in Section 1, we aim at isolating a shock that may affect students' beliefs about their ability. Ideally, we want to lever a shock

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

that provides information about an individual's relative ability only, holding everything else constant. We use a particular peer effect—students' ordinal ranks in their school cohort—as such an information shock, which conditional on own and peer ability captures information about the relative standing within a cohort only. In the following, we first describe how we construct our main variable of interest, the ordinal rank of students, before we illustrate our identification strategy.

4.1. *Constructing Students' Ordinal Ranks*

We construct students' ordinal ranks based on an assessment of their cognitive ability, which is comparable across cohorts and schools. More specifically, we use the condensed version of the revised Peabody Picture Vocabulary Test (PPVT-R; Dunn and Dunn, 2007) that was administered as part of Wave I, and provides us with an objective, age-specific and standardised measure of ability. To administer the test, respondents matched progressively difficult words spoken by the interviewer to one of four pictures they thought best described the meaning of the word. An advantage is that this test could be efficiently implemented and did not require specific literacy skills to take part (Cheng and Udry, 2005).

The PPVT-R test is a well-known and established test for ability that has good internal consistency and reliability among child and adolescent populations (Beres *et al.*, 2000). It correlates with other common ability measures such as the verbal and full scale IQ components of the Wechsler Adult Intelligence Scale (Bell *et al.*, 2001). With our data, we show in Online Appendix Table A.5 that the PPVT-R correlates as expected with a range of educational and labour market outcomes. Furthermore, the PPVT-R scores are negatively correlated with Wave I CES-D scores (Online Appendix Figure A.5a) and highly stable over time (Online Appendix Figure A.5b).

To construct a student's ordinal rank, we first rank students based on their cognitive ability within their school cohort by assigning them an absolute ability rank.[9] Due to differing school and cohort sizes, we subsequently normalise the absolute rank to an ordinal rank by dividing by the school cohort size:

$$\text{Ordinal rank} = \frac{\text{Absolute rank} - 1}{\text{Cohort size} - 1}. \tag{1}$$

This results in an ordinal rank which assigns the value 1 to the highest-ranked student and 0 to the lowest-ranked student. Figure 3 illustrates how this ordinal rank varies with a student's ability. The average ordinal rank increases in a student's ability. Yet, as we are interested in estimating the effect of a student's ordinal rank on their mental health holding ability constant, we need sufficient variation in ranks for a given ability level. Figure 3 provides some evidence that this is indeed the case—for each ability decile in the global ability distribution, we observe sizeable variation in a student's local rank—but we will revisit this question in the following section after formalising our identification strategy.

A potential confound for the interpretation of our rank measure based on Add Health's Picture Vocabulary Test is that neither students nor teachers learn the results of this test.[10] Thus, the

[9] We assign the student with the lowest ability the rank 1 and then increase the absolute rank. Thus, the higher a student's absolute rank, the higher their ability. To define the absolute rank, we count the number of peers with a lower ability, implying that if two students have the same ability they are assigned an equal rank. We relax this definition in robustness checks.

[10] Alternatively, we could have used a student's GPA to calculate ranks. Yet, this measure would have considerable limitations. First, GPA may be comparable within a school cohort, but comparisons across cohorts and schools may be

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

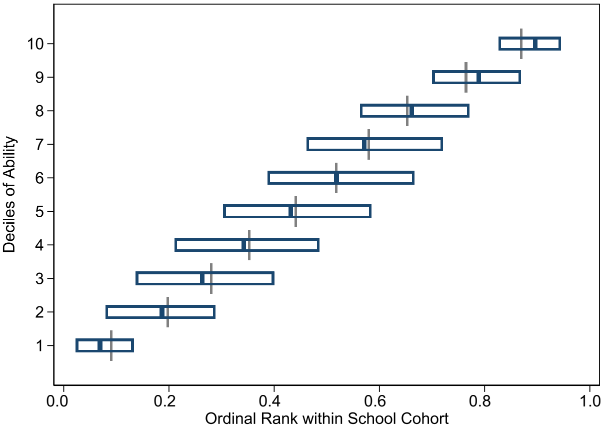

Fig. 3. *Variation in Students' Ordinal Ranks by Ability Decile. This figure presents the variation in ranks for each ability decile. In particular, for each decile, boxes illustrate the 25th, 50th and 75th percentiles of ordinal ranks, while grey lines indicate the mean ranks.*

question remains how salient is our rank measure. In Subsection 5.3, we provide evidence that ranks affect students' beliefs about their relative ability and college-going expectations, indicating that ranks are indeed salient to students.

Another concern is that ability was measured as part of Wave I, and hence could be determined simultaneously with students' ranks. Yet, evidence in the literature indicates that crystallised intelligence—as measured by the PPVT-R—is only malleable early in life and is considered stable from age 10 onward (Jensen, 1998). At the time of Add Health's Wave I, when students were on average 15.6 years old, our measure of cognitive ability can therefore be seen as pre-determined and unaffected by features of the school environment and the students' own or their parents' investments.

In order to provide empirical evidence that ability seems to be fixed, we exploit that the ability test was administered again in Wave III. Online Appendix Figure A.5b shows that the association between the two ability tests is near perfect. Yet, one might still be concerned with a potential spillover from ranks to ability. Hence, we also test for an effect of ordinal ranks in Wave I on ability in Wave III. As shown in Online Appendix Figure A.5c, when conditioning on Wave I ability, which we do in all of our analyses, the effect of ranks on ability in Wave III is essentially zero (the pattern is flat except in the extreme tails). This suggests that cognitive ability is rather stable and pre-determined at the time of observing our sample.

Furthermore, we might be concerned that depressive symptoms shift performance in the Picture Vocabulary Test, and therefore contaminate our ability measure and ranks. To examine such a concern, we extend the previous idea and regress ability in Wave III on CES-D scores in Wave I on the same controls—most importantly the polynomial in Wave I ability—as in our baseline specifications. Online Appendix Figure A.5d shows that we do not find any link between CES-D scores in Wave I and ability in Wave III conditional on ability in Wave I. We do recognise that brain development continues into adolescence and socio-emotional skills remain malleable.

difficult. Moreover, teachers have discretion about the grades of students potentially capturing confounding effects, and the students' GPA may be affected by classical peer effects.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

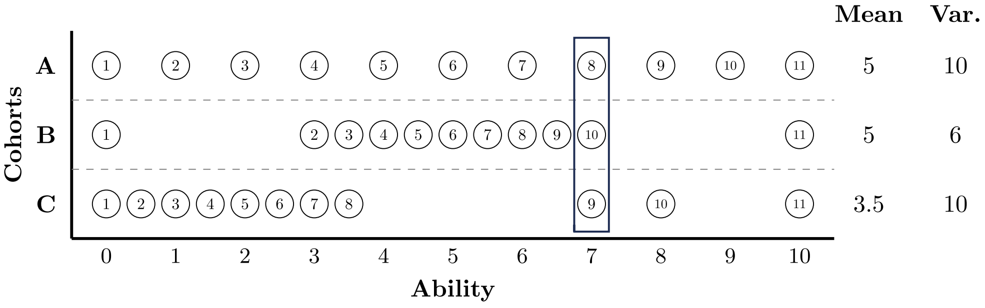

Fig. 4. *Illustrative Example of the Identifying Variation. This figure illustrates how variations in the ability distribution across cohorts allows us to identify rank effects. In these examples, we fixed the minimum and maximum of the ability distribution and allow either the mean or the variance of the ability distribution to differ across cohorts. Students are ranked according to their ability as illustrated by the numbers in the circles. A comparison of cohort A and B shows that holding the mean ability constant can give rise to different ranks for individuals of the same ability. A comparison of cohorts A and C illustrates that a variation in mean ability, but constant variance in ability, can also can give rise to different ranks. In general, any variation (i.e., not necessarily restricted to the first two moments) in the cohort composition can lead to a wide range of ranks for students of a given ability level. Empirically, we will therefore exploit any variation in the cohort composition, regardless of the source of variation.*

Thus, in later robustness checks, we include a control for personality in additional specifications where we control for potential determinants of mental health.

Finally, we observe a random sample of students in each school cohort, introducing additional sampling variation in our data. In Online Appendix F, we report results from a simulation study, showing that such sampling variation leads our estimates to be attenuated by approximately one-third if we observe only 10% of the students in a cohort. In our sample, for only 3.4% of students do we observe less than 10% of their school cohort, while on average 33.8% of all students in each school cohort are part of our sample.[11] Taken together, defining ordinal ranks using Add Health's Picture Vocabulary Test yields a measure that is based on pre-determined characteristics, salient to students and comparable across cohorts as well as schools.

4.2. *Exploiting Within School–Cross Cohort Variation in the Cohort Composition*

We aim to estimate the causal effect of a specific feature of the school environment: how does a student's ordinal rank in their cohort affect their mental health, holding both own and peer ability constant. Before we discuss our empirical strategy more formally, we want to provide some intuition for the identifying variation that we are exploiting. Consider three almost identical students having the same characteristics and, in particular, the same ability. The only difference between the students is that the composition of their peers in terms of ability differs, resulting in different ranks. Figure 4 presents an example of this identifying variation. In this example, the students of interest all have the same ability of 7 (on a 0 to 10 scale), but have different (absolute) ranks ranging from 8 to 10 depending on the ability composition of their peers. Empirically, we will compare the mental health of students having the same ability, but who just happen to

[11] For the majority of the schools, Add Health samples about 17 students from each grade and gender for the main sample, independently of the size of the cohort. In addition, there are 16 saturated schools, in which all students were interviewed.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

have different peers and therefore different ranks in their respective cohorts. In general, these differences in ranks can not only arise from differences in the mean and variance of peer ability, but may stem from any variation in the peer composition (i.e., not necessarily restricted to variation in the first two moments).

4.2.1. *Main specification*

The identifying variation illustrated in Figure 4 describes our basic identification strategy. We follow Hoxby (2000a,b) and exploit the idiosyncratic variation in the ability distribution across cohorts within the same school. This motivates the following empirical specification:

$$y_{ics} = \alpha\, rank_{ics} + f\left(a_{ics}\right) + \boldsymbol{X}_i'\beta + \boldsymbol{\theta}_{ics} + \epsilon_{ics}, \tag{2}$$

in which y_{ics} denotes the mental health of student i in cohort c and school s. $rank_{ics}$ is this student's ordinal rank within their cohort, as defined in (1), and $f\left(a_{ics}\right)$ denotes a flexible functional form of a student's own ability (in our application we use a fourth-order polynomial, but relax this in robustness checks). $\boldsymbol{X}_i$ corresponds to a vector of student characteristics which includes gender, age and age-squared, indicators for race or ethnicity (Asian, Black, Hispanic, Other), an indicator for being the child of an immigrant, the number of older siblings, indicators for their parents' highest degree (less than high school, high school/GED, some college, college degree, postgraduate degree), the log of household income and an indicator for being raised in a single-parent household. Finally, $\boldsymbol{\theta}_{ics}$ denotes a set of fixed effects—at a bare minimum school ($\boldsymbol{\delta}_s$) and cohort fixed effects ($\boldsymbol{\gamma}_c$)—to guide our identification as explained in the following. We cluster standard errors at the school level.

One obvious concern with (2) is that a student's ordinal rank is related to the average ability within the cohort. We aim to focus on the pure information shock and do not want our rank measure to be confounded by typical peer effects in ability. We therefore add the leave-one-out peer ability, $\overline{a}_{-ics}$, as an additional control variable, and control for $\boldsymbol{\theta}_{ics} = \lambda\overline{a}_{-ics} + \boldsymbol{\delta}_s + \boldsymbol{\gamma}_c$. In refinements, we add further linear-in-means peer effects and standard deviations in these peer characteristics, to capture other dimensions and potential non-linearities in peer effects.

An additional concern is that parents may select their children's schools based on trends in the school ability distribution (see, e.g., Rothstein, 2006; Hastings *et al.*, 2009, for evidence that parents prefer sending their children to schools with high ability peers). This would potentially bias our results. In a second specification, we therefore add school-specific cohort trends to capture this potential source of bias. In this case, we identify the rank effect from variation in the ability distribution within schools and across grades after taking school-specific linear trends into account (i.e., $\boldsymbol{\theta}_{ics} = \lambda\overline{a}_{-ics} + \boldsymbol{\delta}_s + \boldsymbol{\gamma}_c + c \times \boldsymbol{\delta}_s$).

Finally, in our third, and most restrictive specification, we introduce school-specific cohort fixed effects. Here we go a step further and control for any heterogeneity of a cohort in a given school. We do this by introducing school-specific grade (cohort) fixed effects, i.e., $\boldsymbol{\theta}_{ics} = \boldsymbol{\delta}_s \times \boldsymbol{\gamma}_c$, as in Murphy and Weinhardt (2020), and discussed in further detail in Denning *et al.* (2021). Using these school-by-grade fixed effects, we absorb any potential peer effects in terms of means, variances or any higher moment. In this case, to identify rank effects, we rely on the variation of students' ranks within their grade (cohort) compared to grades in other schools, after all observed and unobserved differences between school-specific grades are removed. We adopt this specification as our preferred specification, as it more clearly removes potential unobserved factors in school-grade groups that may correlate with rank and mental health.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

4.2.2. *Identifying assumption*

In order to identify the causal effect of ranks, α, the ordinal rank has to be as good as randomly assigned. More specifically, this means that we need to assume exogeneity of ranks conditional on a rich set of controls and fixed effects, that is,

$$E\left[\epsilon_{ics}|rank_{ics}, f\left(a_{ics}\right), \boldsymbol{X}_i, \boldsymbol{\theta}_{ics}\right] = 0.$$

In essence, this assumption implies that ϵ_{ics} is uncorrelated with a student's ordinal rank conditional on their own ability, individual characteristics and a set of cohort-level controls. In the first specification, we assume that these cohort-level controls are given by separate school and cohort fixed effects, as well as peer effects in student ability, and, in the second, we additionally capture school-specific time trends. Using these and individual controls, we compare students in the same school and cohort, with similar peers, and with the same observable characteristics and ability, but who happen to have different ranks.

Nonetheless, there might be other factors that potentially affect a student's mental health and rank that are unobservable to us. If such factors are present, this violates our exogeneity assumption, and hence prevents us from estimating unbiased rank effects. Therefore, our third specification with school-specific grade fixed effects absorbs all observable and unobservable differences between cohorts within and across schools, and is our preferred specification. As mentioned above, we then identify rank effects from variations in ranks within school cohorts or, more specifically, from combinations of different shapes of the ability distribution across school cohorts and own ability that define ordinal ranks. Below, we also report results of a sensitivity analysis that helps to assess how severe potential confounders would have to be to explain the effects we observe.

4.2.3. *Residual variation*

A natural question is how much variation is left in our rank variable after conditioning on our set of control variables and different fixed effects. We assess this variation in Online Appendix Table A.4. The standard variation in ranks without controls amounts to 0.28. However, since a student's rank and ability are positively correlated, as indicated by Figure 3, some part of the variation may be due to ability. Moreover, as our analysis will be focused on heterogeneous effects by ability decile, we need to ensure that there is sufficient variation in our variable of interest in each of the deciles.

To assess this condition, we calculate the residual variation in ranks after controlling for background characteristics and different sets of fixed effects, and we compare this to the raw standard deviation in ranks. Online Appendix Table A.4 shows that the raw standard deviation in ranks by decile varies between 0.09 and 0.18. Conditioning on school and grade fixed effects and our set of baseline controls reduces this variation to 0.07–0.12. Using school-specific grade fixed effects leaves a similar degree of variation at 0.07–0.11. Our rich set of controls and fixed effects leaves at least 40% of the raw variation. Thus, there remains substantial residual variation in ordinal ranks to study their causal effect on mental health.

4.2.4. *Balancing tests*

Finally, we perform balancing tests on our main treatment variable and other peer variables to provide evidence that the peer composition across cohorts within schools is indeed consistent with quasi-random peer assignment. Each cell of Table 2 presents a regression of our treatment variable of interest—the ordinal ranks of students in their cohort—or another variable that should be

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 2. *Balancing Tests.*

	Rank			Peer ability (std.)			SD (peer ability) (std.)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	−0.002	−0.001	−0.002	0.001	−0.002	0.000	−0.007	−0.005	0.002*
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.000)	(0.007)	(0.005)	(0.001)
White	−0.008***	−0.007**	−0.007**	0.003	−0.005	−0.003	0.019	0.024**	0.007**
	(0.003)	(0.003)	(0.003)	(0.011)	(0.006)	(0.002)	(0.013)	(0.012)	(0.003)
College-educated parents	−0.000	−0.000	0.000	0.003	0.003	−0.001	0.011	0.006	0.000
	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)	(0.001)	(0.009)	(0.008)	(0.002)
Raised by a single parent	−0.002	−0.002	−0.002	−0.003	0.001	0.000	0.016	0.014	0.001
	(0.002)	(0.002)	(0.001)	(0.005)	(0.004)	(0.001)	(0.011)	(0.009)	(0.001)
Household income (US$1,000)	−0.000	−0.000*	−0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household receives food stamps	−0.004	−0.003	−0.003	0.000	0.001	0.000	−0.023	−0.011	0.003
	(0.002)	(0.002)	(0.002)	(0.006)	(0.005)	(0.001)	(0.015)	(0.012)	(0.002)
Household size	0.000	−0.000	−0.000	−0.003	−0.001	−0.000	−0.003	−0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.000)	(0.004)	(0.003)	(0.001)
First-born child	−0.000	0.000	0.002	0.006	0.005	−0.000	−0.001	−0.007	−0.000
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.000)	(0.009)	(0.008)	(0.001)
Birth weight (ounces)	−0.000	−0.000	−0.000	0.000	0.000	−0.000	0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School and grade FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
School-specific trends	No	Yes	No	No	Yes	No	No	Yes	No
School × grade FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the school level. Each cell presents a separate regression of the variable in the column header (rank, standardised peer ability, or standard variation in peer ability) on the variable indicated at the beginning of each row. All specifications include controls for a fourth-order polynomial in ability and fixed effects as indicated at the bottom of the table.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

quasi-randomly assigned—peer ability and variation in peer ability—on pre-determined characteristics of students as well as a fourth-order polynomial in ability and one of the three sets of fixed effects θ_{ics}.

Consistent with quasi-random assignment of peers, we observe most characteristics are not related to our treatment variables. Only the indicator whether a student is White seems to be associated with a higher rank. Yet, given that this association does not hold for other quasi-randomly assigned peer variables, the number of tests performed is relatively high, and that the coefficient is small, amounting to less than one percentile score, we interpret the balancing check as consistent with quasi-random assignment of peers.

5. Results

How does a student's ordinal rank affect mental health? Our theoretical framework in Section 1 generates four predictions: First, we should observe that positive (negative) shocks, in our application proxied by ability ranks in the school cohort, benefit (worsen) a student's mental health. Second, negative shocks have more pronounced consequences than positive ones. Third, rank effects are predicted to be stronger at the lower end of the ability distribution, where students are more likely to withdraw their study effort in response to negative shocks. Finally, the framework suggests that these effects are persistent over time. In the following, we will test these predictions.

5.1. *Average Effect of Students' Ranks on Mental Health*

We begin by studying the average effect of a student's rank on their mental health. More specifically, we relate a student's mental health measured by CES-D scores to their ordinal rank based on our main specification in (2) with standard errors clustered at the school level. We present our results in Table 3 and Figure 5. Based on our first empirical specification controlling for separate school as well as cohort fixed effects, and ability peer effects, column (1) shows that higher ranks reduce CES-D scores, i.e., they improve the students' mental health. More specifically, moving a student from the 25th percentile to the 75th improves their mental health by 0.8 CES-D points; increasing a student's rank by 1 SD (0.28 percentiles) yields a 0.06 SD improvement in mental health.

How large are these effects relative to mental health gaps due to socioeconomic differences in mental health? In Online Appendix Table A.3, we present associations of several background characteristics on (standardised) CES-D scores. The estimated effect size is similar to the difference in mental health of children from college-educated and non-college-educated households, about two-thirds of the difference between White and non-White students, or the difference between students raised by a single parent and those raised by both parents.

Another point of comparison can be based on meta-analyses of the efficacy of anti-depressants or positive psychology interventions (Mitte *et al.*, 2005; Turner *et al.*, 2008; Bolier *et al.*, 2013): effect sizes of these interventions yield effects of 0.20–0.34 SD on outcomes such as psychological well-being, depression and subjective well-being. Given that these are targeted interventions, we consider the estimated effects of ordinal ranks as large. A fact that is all the more striking as it results from natural variation in the ability composition of school cohorts, conditions on a rich set of demographic characteristics, and removes variation from a rich set of fixed effects capturing systematic variations between schools and grades.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 3. *Average Effect of Ordinal Ranks on Mental Health.*

	Mental health (CES-D score)				
	(1)	(2)	(3)	(4)	(5)
Panel A. Baseline effects					
Rank	−1.60**	−1.59**	−1.64**	−1.62**	−1.70**
	(0.75)	(0.75)	(0.76)	(0.76)	(0.76)
Ability and controls	Yes	Yes	Yes	Yes	Yes
Ability peer effect (mean)	Yes	Yes	Yes	Yes	No
Further peer effects (mean)	No	Yes	Yes	Yes	No
Ability peer effect (SD)	No	No	Yes	Yes	No
Further peer effects (SD)	No	No	Yes	Yes	No
School and grade FEs	Yes	Yes	Yes	Yes	No
School-specific time trends	No	No	No	Yes	No
School × grade FEs	No	No	No	No	Yes
Mean CES-D score	11.3	11.3	11.3	11.3	11.3
Observations	18,459	18,459	18,459	18,459	18,459
R^2	0.109	0.109	0.110	0.117	0.128
Panel B. Standardised effects					
Rank (std.)	−0.06**	−0.06**	−0.06**	−0.06**	−0.06**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Panel C. Role of unobservables					
Oster's δ ($R^2_{max} = 1.3R^2$)	−1.05	−1.23	−1.65	−2.76	−2.39

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered at the school level. In Panel A, each coefficient presents a regression of CES-D scores (lower scores corresponding to better mental health) on an individual's percentile rank at the school level based on (2). We include a fourth-order polynomial in own ability, gender, ethnicity, age and age-squared, an indicator for being a child of an immigrant, the number of older siblings, parental education, logged household income and being raised by a single parent as control variables. Peer ability includes the leave-one-out mean and standard deviation of peer ability, peer controls comprises additional peer effect terms in gender, ethnicity and parental education. We present standardised effects of our main effect in Panel B. Panel C presents the results from a sensitivity analysis based on Oster (2019) and quantifies how severe selection based on unobservables would need to be for zero rank effects. To calculate δ, we follow Oster (2019) and assume a maximum R^2_{max} of 1.3 times the actual R^2.

In the remaining columns of Table 3 we adopt further specifications to investigate the robustness of this result. In particular, column (1) is restrictive in that it only allows for peer effects in ability. Yet, the literature on peer effects has identified a range of different peer characteristics that causally affect students' performance and thereby may also affect their mental health. Examples include the share of females, minorities or students with high socioeconomic status (Hoxby, 2000b; Lavy and Schlosser, 2011; Borbely *et al.*, 2021; Cools *et al.*, 2021). We add these additional peer effect terms in column (2). Furthermore, in column (3), we also add controls for the standard deviation in peer ability and other peer characteristics capturing potential non-linear peer effects. Our estimates show that the rank effect is robust to the inclusion of these additional peer effects and varies only slightly.

Our identification is based on quasi-random variation in peer ability across cohorts in a given school. Yet, if parents select schools for their children based on trends in the ability distribution, this might bias our results. To account for such factors that change within a school over time, we further include school-specific time trends in column (4), which neither change the size nor the statistical significance of our results.

In column (5), we adopt an even stricter empirical specification using grade-by-school fixed effects, and thus not only account for linear trends in the school-specific ability composition

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

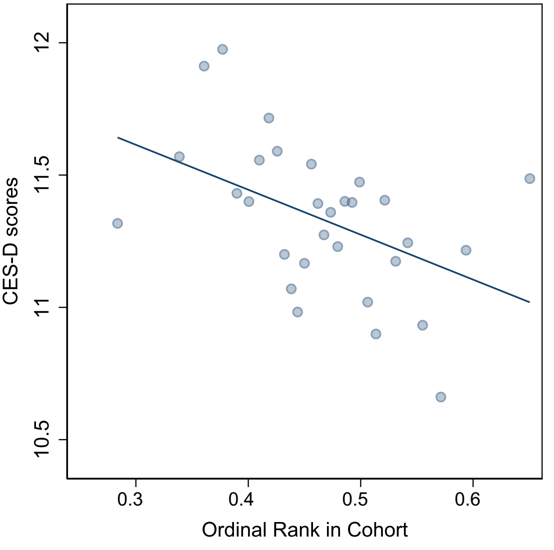

Fig. 5. *Average Effect of Ordinal Ranks on Mental Health. This figure presents the results from a regression of CES-D scores (lower scores corresponding to better mental health) on students' percentile ranks in their cohort (higher ranks correspond to higher relative ability) conditional on a fourth-order polynomial in own ability, gender, ethnicity, age and age-squared, being a child of an immigrant, number of older siblings, parental education, logged household income, and being raised by a single parent, as well as school and grade fixed effects as in column (5) of Table 3.*

over time, but for any trend. Additionally, this set of fixed effects accounts for all observed and unobserved peer effects and exploits variation within school-grades, to identify the effect of ordinal ranks on the mental health status of students. The coefficient of interest slightly increases in magnitude, but are wholly consistent with the other estimates. As this specification accounts more fully for potential contamination from unobserved peer effects, we focus on it moving forward as our preferred specification (also recommended by Denning *et al.*, 2021).

Taken together, the results from Table 3 document that students' ranks among their peers exert a causal effect on students' mental health measured by CES-D scores in line with the central prediction of our model. Decreasing (increasing) a student's rank by 1 SD causes an approximately 0.06 SD decrease (increase) in a their mental health, comparable in magnitude to effects of ranks on test scores (0.08 SD; Murphy and Weinhardt, 2020), and to the raw mental health difference between children with college and non-college-educated parents.

5.2. *Robustness Checks*

In this section, we report a series of additional analyses to probe the robustness of our finding.

5.2.1. *Nonlinearity in ability*

In our main specification, we adopt a fourth-order polynomial in Peabody scores to take the relation of mental health and ability into account (see also Online Appendix Figure A.5a on the relationship between ability and CES-D scores). Yet, one might be worried that this arbitrary choice drives our results. In Online Appendix Table B.2, we therefore examine different

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

polynomials up to a sixth order. We find that using linear or quadratic controls in ability increases our estimated coefficient on ordinal ranks and thus would strengthen our main result. The rank estimates stabilise for higher-order polynomials in ability. Moreover, we also include a specification with indicators for each level of the ability score, which non-parametrically controls for different ability levels, and find our results remain unchanged.

In a second set of specifications, we use a data-driven approach to select the (ability) control variables by employing a post-double selection (PDS) lasso (Belloni *et al.*, 2014). The PDS lasso penalises control variables, but allows valid inference on non-penalised treatment variables. We perform two such specifications, one in which we allow penalisation terms of an eight-order polynomial only, and one in which we additionally allow for penalisation of the set of baseline control variables (e.g., gender, race indicators, age). Both specifications penalise higher-order ability terms, leaving only a second-degree polynomial or a linear trend in ability, which suggests a relatively linear relationship of CES-D scores and ability, as also illustrated in Online Appendix Figure A.5a. More importantly, however, the estimated effects of ordinal ranks remain unaffected and, if anything, become more pronounced.

5.2.2. *Definition of peer groups*

We defined peer groups based on all students in a given cohort. Yet, evidence exists that students form friendships with similar peers (i.e., that friendship networks exhibit homophily; see McPherson *et al.*, 2001; Graham, 2015, for overviews over the literature) and that they systematically select their relevant peers from larger peer groups (Kiessling *et al.*, 2019). This raises the question whether the rank effects should be defined at a more local level. Hence, we explore how our results change when we allow for differential rank effects by finer subgroups. More specifically, we enrich our main specification and add a second rank calculated (*i*) within grade and gender, (*ii*) within grade and race, or (*iii*) within grade, gender and race. Online Appendix Figure B.1 presents the results from these specifications. As can be seen from this figure, the baseline effect remains similar across specifications and the additional ranks based on different peer group definitions have small and insignificant effects. Hence, calculating ranks within grades seems appropriate, and this suggests that, in our setting, the effects seem to stem from comparisons to all peers in a cohort rather than from a specific subgroup.

5.2.3. *Definition of ranks*

In our definition of ranks, we calculate absolute ranks based on the number of peers with a strictly lower ability as in Elsner and Isphording (2017), and Elsner and Isphording (2018). Yet, other definitions are conceivable with implications for the assignments of ranks for those students who are in the same school cohort and have the same measured ability. For instance, we could have assigned absolute ranks based on the number of peers with a lower or equal ability rather than a strictly lower ability. Alternatively, we could assign the mean of both methods to get at an average rank in case of ties. In Online Appendix Table B.3, we compare these different definitions of ranks and find that our estimates are robust to the precise definition of ranks.

5.2.4. *Personality and social interactions as potential confounds*

In Section 4, we provided some evidence that conditional on Wave I ability, it is unlikely that reverse causality contaminates our rank effect estimates. Yet, it is well known that non-cognitive skills such as self-control continue to develop during adolescence (Heckman and Mosso, 2014).

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

This could capture the impact of past shocks on ability beliefs, if past shocks also directly impact non-cognitive skill or do so indirectly through impacting past ability beliefs. Furthermore, empirical evidence indicates that positive peer relationships can improve mental health (Eisenberg *et al.*, 2013) and a positive link between friendship network centrality and perceptions of social climate (Alan *et al.*, 2021).

To check whether these factors mediate the estimated rank effect, we include conscientiousness as an important dimension of one's personality, as well as Bonacich centrality as a measure of students' popularity. We further control for indicators of whether students have named a best male friend who reciprocates as a friend and likewise for a best female friend. As shown in Online Appendix Table B.4, we find that our rank effect estimates remain entirely consistent with our baseline estimates, further alleviating concerns over the potential confounding channels discussed here.

5.2.5. *Role of unobservables*

Our identification strategy assumes that a student's rank is exogenous conditional on own ability and school and cohort fixed effects. It is reassuring that our findings remain nearly unaffected once we control for additional potential confounds, and when we adopt the stricter specification using school-specific cohort fixed effects. A more formal approach to test for the role of unobservables is to ask how severe selection based on unobservables would have to be to drive down the estimated rank effects to zero. In order to quantify this, we follow Oster (2019) and calculate δ, a measure for the degree of selection based on unobservables relative to observable characteristics. If δ is larger than one, this indicates that selection on unobservables would need to be at least as important as selection based on observables to explain our effects. As shown in Panel C of Table 3, the magnitude of δ is larger than one in all specifications. Given that we control for arguably the most important factors that could bias students' ordinal ranks and that may affect mental health through differences in the cohort composition, these numbers imply that we would have to be missing highly relevant variables in order for unobservables to give rise to our estimated effects. Hence, we conclude that unobservables are unlikely to drive our estimated rank effects.

5.2.6. *Sorting based on ranks*

One concern is that parents may select schools based on the rank that their children would have, violating our assumption that the rank is as good as randomly assigned. Yet, there is plenty of evidence that parents prefer sending their children to schools with high-ability peers (Rothstein, 2006; Hastings *et al.*, 2009; Burgess *et al.*, 2015; Jackson *et al.*, 2021). If this is the case, then this is not consistent with positive sorting based on ranks, as ranks and peer ability are inversely related.

Moreover, several patterns in our data suggest that sorting based on ranks is a minor concern. First, in Online Appendix Figure B.2 we show that even if parents sort into specific schools based on average peer ability, there remains high uncertainty about the resulting rank of a student with a given ability. This implies that sorting on average ability and rank is rather difficult. In addition, as we show in Online Appendix D, the size of the rank effect does not differ by average school ability, and is not driven by a specific subset of students or schools having certain characteristics.

Second, we assume that conditional on school and grade (school-by-grade) fixed effects as well as our baseline set of controls, the variation in ranks is as good as random. Yet, as shown

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

in our balancing checks in Table 2, pre-determined individual characteristics do not seem to be systematically related to ranks, average peer ability or the variation in average peer ability, indicating that sorting based on ranks or other peer characteristics is unlikely to explain our results.

5.2.7. *Heterogeneous effects of school ability distributions*

So far, we have accounted for variation at the school-by-grade level, individual background characteristics and flexibly controlling for ability. Yet, Booij *et al.* (2017) find heterogeneity across students' prior ability in how the distribution of classrooms in their data impacted students' outcomes. Denning *et al.* (2021) point out that these interactions may correlate with rank. Hence, omitting these factors will introduce a spurious correlation biasing our results. Their approach is to make comparisons across classes (school cohorts in our data) that have similar, but not identical, distributions by interacting ability with indicators for quartiles of the school mean ability and variance distributions.

We follow this approach and interact our ability controls with quartile indicators for a school's average ability (variance in student ability) relative to all other schools. Our results are reported in the Online Appendix Table B.5. We report three specifications for interacting students' ability with (*i*) school mean ability quartile indicators, (*ii*) school variance quartile indicators and (*iii*) both sets of interactions. Including these interactions, our results remain robust, despite losing some efficiency for the exhaustive set of interactions with both mean and variance indicators.

5.2.8. *Simulations to assess the role of measurement error*

The Add Health data have several sources of classical and non-classical measurement error. First, only a random subset of all students in each school is sampled introducing potential biases in our main variable as we observe only a fraction of the cohort. Second, our ability measure may suffer from measurement error that translates into a mis-measured rank. Third, although our analysis above suggests that unobservables are unlikely to drive our results, they potentially distort our estimates if there are omitted variables correlated with ability. Fourth, there may be sorting into different classrooms based on ability within school cohorts, which we cannot observe. This would imply that we calculate the students' ranks based on incorrect reference groups. Fifth, and finally, CES-D scores are aggregated from a small number of items that are scored on a scale from 0 to 3 rather than on a continuous scale.

We assess these concerns using a series of Monte Carlo simulations reported in Online Appendix F. We find that these various forms of measurement error lead to mild to moderate attenuation of our estimates and reduced efficiency. This implies that we are likely to underestimate the true causal effect of ranks on mental health.

5.3. *Beliefs as a Mechanism*

Our theoretical framework suggests that beliefs are a key mechanism for how shocks translate into mental health. While this is an untestable assumption, we provide empirical support for this modelling choice in Table 4. Specifically, Panel A (B) presents estimated rank effects (standardised rank effects) on three different beliefs and expectations: students' beliefs about their relative ability, whether students aspire to go to college, as well as students' expectations about attending college. All three beliefs are significantly affected by students' ranks in their cohorts. In fact, the standardised effect sizes are nearly identical in magnitude to the effects for

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 4. *Students' Ranks Affect their Beliefs.*

	Standardised beliefs		
	(1) Belief about relative ability	(2) Wants to attend college	(3) Expects to attend college
Panel A. Baseline effects			
Rank	0.26**	0.23***	0.21**
	(0.12)	(0.09)	(0.09)
Ability and controls	Yes	Yes	Yes
School × grade FEs	Yes	Yes	Yes
N	18,434	18,406	18,392
R^2	0.174	0.130	0.196
Panel B. Standardised effects			
Rank (std.)	0.07**	0.07***	0.06**
	(0.03)	(0.02)	(0.03)
Panel C. Mediation of rank effects on CES-D scores			
Share mediated	14.7%	15.8%	19.2%

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered at the school level. Each specification includes all controls as in our preferred baseline specification. Panel A presents specifications in which the dependent variable is students' standardised belief about how intelligent they feel compared to other people of their age (elicited on a 6-point Likert scale with 1 corresponding to 'moderately below average' and 6 corresponding to 'extremely above average') in column (1) and students' college-going aspirations, as well as students' expectation how likely they are to go to college as the dependent variables (elicited on a 5-point Likert scale with 1 corresponding to 'low' and 5 corresponding to 'high') in columns (2) and (3). Panel B expresses these estimates in terms of a 1SD change in ranks. Finally, Panel C calculates the share of the rank effect on CES-D scores that is mediated through each of the three belief variables.

CES-D scores in Table 3: a one standard deviation increase in ranks increases beliefs by 5–7% of a standard deviation.

Given these similar effect sizes, Panel C then asks how much of the original rank effect on CES-D scores is mediated by each of the three belief measures. To do this, we run an auxiliary regression, in which we enhance our main specification to estimate rank effects on CES-D scores by including the belief measure. We multiply the coefficient on the belief with the estimated rank effect from Panel A of Table 4. We then divide this through the original rank effect on CES-D scores shown in Table 3. While the resulting mediating effect cannot be interpreted as causal, it provides our best guess on how much of the original rank effect on mental health is mediated through each of the beliefs. We find that each of the beliefs mediates between 15% and 19% of the original effect, indicating beliefs as an important mechanism in line with our theoretical framework.

Recent results by Pagani *et al.* (2019) complement these findings. Using data on Italian high school students, they show that rank effects on personality traits seem to operate through beliefs in the form of perceived ability and academic motivation in line with the results presented in this section. We also check whether the rank effects translate into school performance effects. As also shown by Elsner and Isphording (2017), we document in Online Appendix Table B.1 that higher ranks indeed lead to a better grade point average (GPA). In addition, we document that they also decrease the likelihood of having been absent from school. These results indicate that adolescents respond to the ability ranks in terms of their school behaviour, and these effects seem to operate through students' beliefs.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 5. *Different Facets of Mental Health.*

	Principal components of CES-D scores (std.)			
	(1) Loneliness	(2) Lack of pos. attitude	(3) Lack of motivation	(4) External factors
Rank	−0.07	−0.24**	−0.16*	0.03
	(0.10)	(0.09)	(0.09)	(0.10)
Ability and controls	Yes	Yes	Yes	Yes
School × grade FEs	Yes	Yes	Yes	Yes
N	18,411	18,411	18,411	18,411
R^2	0.095	0.100	0.047	0.045

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the school level. The outcome is a standardised factor (with zero mean and a standard deviation of 1) from a principal component analysis of all 19 items of the CES-D scale. We include all base controls and school-grade fixed effects as in our baseline specification of column (5) in Table 3, and hence can compare the results to the standardised average effect in Panel B, column (5) of Table 3.

5.4. *Different Facets of Mental Health*

In our analysis, we use the CES-D score based on the sum of the single items as it is commonly used in the literature. While all of the items are related to depressive symptoms, they cover different facets. In order to shed more light onto which facet is driving our results, we perform a principal component analysis on the items. We then apply a Varimax rotation and predict factor scores. As shown in Online Appendix C, this results in four distinct factors corresponding to (*i*) loneliness, (*ii*) lack of positive attitudes, (*iii*) lack of motivation and (*iv*) external factors.

Table 5 presents regressions of these different facets on the ordinal rank of students. We find that our results are mainly driven by effects on factors capturing a lack of positive attitudes as well as a lack of motivation rather than loneliness or external factors. Consistent with our previous observation that rank effects operate through beliefs, we find that the effects are driven by mental health facets capturing the general mood and motivation of students rather than social exclusion and other external factors.

5.5. *Exploring Asymmetries in Shocks*

We have documented large effects from ordinal ranks on mental health. Yet, not all shocks are similar. Prediction 2 suggests that once a student experiences a negative shock, their mental health deteriorates and is more likely to remain in a poor condition. We now want to provide more evidence on the asymmetry of these effects. If our conjecture is right, we should observe that any effect is more pronounced for negative rather than positive shocks.

The previous economic literature on belief updating provides mixed evidence on potential asymmetries when it comes to positive and negative shocks. While some studies find support for the so-called good news–bad news effect (e.g., Eil and Rao, 2011; Zimmermann, 2020) in which people react to good news about themselves, but show less pronounced responses to negative signals, others find evidence of asymmetric updating in self-relevant domains along the lines of our conjecture (Ertac, 2011; Kuhnen, 2015; Coutts, 2018; Coffman *et al.*, 2021), or no asymmetry (e.g., Buser *et al.*, 2018). Taken together, there does not seem to be conclusive evidence on potential asymmetries for belief updating in the economics literature.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Correlational evidence from a literature in psychology focused on interactions of belief updating and mental health is in alignment with our prediction. This literature finds that depressed individuals have attentional biases for negative, but not positive, information (Gotlib *et al.*, 2004; Armstrong and Olatunji, 2012; Korn *et al.*, 2014). Moreover, recent neurocognitive theories of depression highlight the importance of negative information processing biases in the development of depression (Roiser *et al.*, 2012). We thus expect that negative signals, i.e., having a rank that is lower than one might expect, leads to stronger responses in mental health than positive shocks, as Prediction 2 suggests.

In order to differentiate between positive and negative shocks, we calculate the expected rank of students, independently of the ability composition of their local school cohort, and compare students having a local rank above or below this expected (global) rank.[12] The idea behind this is as follows: consider two students with identical ability. One of them is randomly assigned to better classmates, where they have a lower rank, whereas the other has worse peers and correspondingly a higher rank. We now investigate whether the effects of (local) ranks are more pronounced for those receiving negative rather than positive shocks. To do this, we calculate a rank measure similar to (1), but consider students from all schools attending a given grade. In other words, we calculate individual i's rank, $rank_{ic}$, among all students in a given cohort c, i.e., independently of their school s. We define student i receiving a negative shock if their rank in their local school cohort, $rank_{ics}$ is lower than the rank among all students in a given cohort, $rank_{ic}$:

$$\text{Negative shock} = \mathbb{1}\left\{rank_{ics} < rank_{ic}\right\}. \tag{3}$$

We then extend (2) by adding an indicator for negative shocks as well as the interaction of negative shocks and ranks. This allows us to study whether negative shocks differentially affect the mental health of students compared to positive shocks.

In Table 6, we study asymmetric responses using our definition of negative shocks from (3). Column (1) shows our baseline result of the first column of Table 3 that ranks significantly reduce CES-D scores. We then study the causal effect of receiving a negative shock on mental health, while abstracting from rank effects. Column (2) shows that negative shocks increase CES-D scores by 0.38 points, corresponding to 0.05 SD. In other words, negative shocks are detrimental to mental health.

Column (3) explores the interaction of ranks and negative shocks by regressing CES-D scores on an indicator for negative shocks and the interactions of ranks with indicators for positive and negative shocks. We find that once we account for ranks, the coefficient on negative shocks is more pronounced than in column (2) and increases CES-D scores by 0.63 (0.08 SD). Moreover, rank effects are approximately twice as large for students receiving negative shocks compared to those receiving positive shocks and similar to our baseline estimate in column (1), although the difference between positive and negative shocks is not significant at conventional levels ($p = 0.10$).

While these results support Prediction 2, one might be concerned about the saliency of these negative shocks. We check the robustness of our result by excluding students whose local rank is similar to their global rank. These students may or may not receive a negative shock, depending on small perturbations. Hence, we focus on those who receive a more pronounced positive or negative shock. We implement this by removing those students from our sample, whose

[12] If students have unbiased beliefs to begin with, this global rank would correspond to the prior in our theoretical framework.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 6. *Asymmetric Effects of Ordinal Ranks on Mental Health.*

	Mental Health (CES-D score)					
	Full sample			Trimmed sample ($\lvert\Delta_{rank}\rvert \geq 0.10\text{SD}(\Delta_{rank})$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rank	-1.70^{**}	–	–	-2.34^{***}	–	–
	(0.76)	–	–	(0.79)	–	–
Negative shock	–	0.38^{*}	0.63^{*}	–	0.51^{**}	0.69^{*}
	–	(0.20)	(0.33)	–	(0.24)	(0.36)
Rank × negative shock	–	–	-1.62^{*}	–	–	-2.21^{**}
	–	–	(0.97)	–	–	(1.06)
Rank × positive shock	–	–	-0.82	–	–	-1.37
	–	–	(0.96)	–	–	(1.04)
Ability	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
School × grade FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean CES-D score	11.3	11.3	11.3	11.3	11.3	11.3
p-value (no heterogeneity)	–	–	0.10	–	–	0.11
Observations	18,459	18,459	18,459	16,865	16,865	16,865
R^2	0.128	0.128	0.128	0.125	0.125	0.126

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the school level. We include all base controls and school-grade fixed effects, as in our baseline specification of column (5) in Table 3. Columns (1)–(3) use the full estimation sample, while columns (4)–(6) drop those students whose local and global rank are relatively similar (i.e., $\lvert\Delta_{rank}\rvert = \lvert rank_{ics} - rank_{ic}\rvert < 0.10\text{SD}(\Delta_{rank})$).

difference between the local and global rank is smaller than 10% of an SD, removing about 9% of our sample. Columns (4) through (6) show that the results from columns (1)–(3) become more pronounced once we trim the sample, indicating that our results indeed seem to stem from surprising (negative) shocks.

Finally, our definition of negative and positive shocks can just be an approximation, as we observe noisy measurements of both a student's ability as well as the presence of a positive or negative shock. *Ex ante*, the direction of the bias is not clear. While trimming the sample suggests that our results are indeed about surprising negative shocks, we study the role of measurement error in more detail. In a Monte Carlo simulation reported in Online Appendix F (see Simulation F), we find that the rank effect corresponding to positive shocks overestimates the true effect in the presence of small amounts of measurement error, but underestimates if measurement error becomes larger. More importantly, however, our coefficients of interest—on negative shocks and the interaction of ranks and negative shocks—are attenuated in the presence of measurement error, implying that we consistently underestimate the magnitude of these coefficients.

Overall, we conclude from these results that students update more strongly in the case of negative shocks. This asymmetry implies that rank effects on mental health do not seem to stem from positive shocks of unexpectedly being ranked highly, but rather from negative shocks.

5.6. *Heterogeneous Rank Effects*

Prediction 3 suggests that the shocks to students at the lower end of the distribution are larger than for higher-ability students. In our model, this occurs because, on the one hand, a diminished perception of ability decreases study time, which subsequently amplifies the consequences of

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

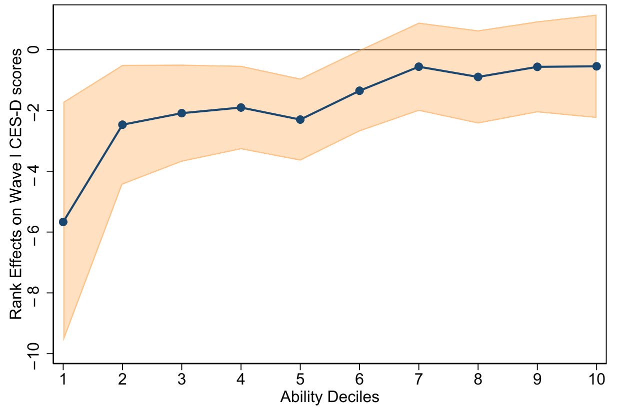

Fig. 6. *Effects of Ordinal Ranks by Ability Decile. This figure presents the effect of ordinal ranks by ability decile. We estimate the effects with our preferred baseline specification of controls and school-grade fixed effects enriched by interacting a student's rank with indicators for ability deciles. The shaded area indicates 90% confidence intervals clustered at the school level.*

exogenous shocks, and, on the other hand, this partly stems from low-ability students holding beliefs closer to the threshold where they withdraw their study effort. If this prediction is correct, we should observe stronger (weaker) rank effects for lower-ability (higher-ability) quantiles. We therefore enrich our main specification given in (2) by interacting the rank with indicators for each ability decile.

Figure 6 displays the results of this analysis graphically, while Online Appendix Table D.1 presents the corresponding regression estimates. We indeed find that rank effects are more pronounced at the lower end of the distribution. The ordinal rank reduces the CES-D score by 5.67 points when moving a student from the bottom to the top rank, which corresponds to 0.75 SD. This effect amounts to three times the average effect, and would suffice to move a student diagnosed with a moderate depression according to a threshold of 16 (Radloff, 1977) to the average CES-D score of 11.3 in our sample. While the point estimates are negative for all deciles, they are more pronounced at the lower end of the ability distribution, the estimated effects slowly fade out and are rather small, and not significant at the top end of the distribution (coefficient of −0.55 with a p-value of 0.60 for the tenth decile). Tests of equality between the coefficient from the lowest decile against the average of the other deciles yields a p-value of 0.06, and comparing effects of students with below and above median ability yields a p-value of 0.04. These results are therefore consistent with Prediction 3 of our theoretical framework that effects should be more pronounced for low-ability students.

One potential concern pointed out by Murphy and Weinhardt (2020) is that of multiplicative measurement error in ability. This could lead to bias, particularly in settings such as ours focusing on the tails of the ability distribution. Murphy and Weinhardt (2020) suggest percentilising the ability measure to a uniformly distributed measure—e.g., the percentilised rank in one's global cohort—and using this as the base for school-grade rank and as the ability control. Add Health provides a percentilised version of the PVT scores in three-month age bins. As a robustness check, in the Online Appendix Table E.2 we report rank estimates across ability deciles based

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

on this percentilised ability measure to construct ranks and as the control. Consistent with our previous results, we observe that lower ability students experience stronger rank effects.

5.6.1. *Further heterogeneities by other individual and school characteristics*

Although our theoretical framework is not aimed at providing predictions for specific sub-samples, these are nevertheless important for policymakers interested in targeting policies. In the Online Appendix D, we study two groups of heterogeneities—based on individual characteristics of students and based on school and cohort characteristics. We only observe limited heterogeneity with respect to sociodemographic characteristics or school-/cohort-level variables. The consequences of ranks, therefore, seem to affect the mental health of all students rather equally.

There are three caveats here to point out. First, extending our results on asymmetry, we find that adolescents living in poverty have a strong response to rank only when experiencing negative shocks. Also, their response appears more pronounced relative to those not living in poverty. This provides suggestive evidence that those who are more likely to experience multiple adverse events are more sensitive to negative shocks. Second, we also look at this asymmetry heterogeneity by a median split of our conscientiousness scale that we introduced in our robustness checks. The idea here is that if adverse events have lowered resiliency (as suggested by results of Pagani *et al.*, 2019), this would be a mechanism whereby negative shocks matter more, and that explains the stronger response to negative shocks by those in poverty. We find suggestive evidence that negative shocks are indeed more pronounced for those with low conscientiousness. Third, we also find that rank effects are stronger for those in early adolescence. Although rank continues to exert influence at later ages, this suggests that early shocks could be more important if those are critical periods of development where adolescents are more sensitive in their mental health. We further discuss these points in the Online Appendix D.

5.7. *Discussion of the Mechanisms*

The previous literature on rank effects often posits that being ranked highly or being a 'big fish in a little pond' opens up more opportunities regarding better schools, colleges and jobs. We would expect that results consistent with such a mechanism would be more pronounced at the top end of the ability distribution, where competition is fierce, and potentially be stronger for unexpected high ranks. The idea behind such a mechanism is that being ranked highly yields an option value of access to better schools.

We find the opposite: Rank effects on mental health are more pronounced at the lower end of the ability distribution, and seem to be driven by negative rather than positive shocks. Thus, our results on mental health are inconsistent with a 'big fish in a little pond' effect. Instead, we find that consistent with our stylised model that ranks seem to operate through beliefs, negative shocks, and withdrawal of study effort, and are stronger for low-ability students, consistent with a growing literature in psychology (e.g., Gotlib *et al.*, 2004; Roiser *et al.*, 2012; Korn *et al.*, 2014). That being said, we think that 'big fish in a little pond' mechanisms are likely important for outcomes other than mental health. In fact, when studying the long-run outcomes such as educational attainment in Subsection 6.2, we find positive effects of ranks both at the lower as well as the upper end of the ability distribution for educational attainment. This indicates that mental health is one of potentially several mechanisms affecting economic long-run outcomes, albeit a very important one.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

6. Persistence of Rank Effects

We have established that the ordinal rank exerts a causal effect on the mental health of students, and this effect is more pronounced for low-ability students. In a next step, we now want to explore the dynamics of these effects. To derive a prediction about expected patterns, we note two points. First, there is a significant association between ability and mental health (see Online Appendix Table A.3 and Online Appendix Figure A.5a). Second, students at the lower end of the distribution experience stronger effects, and they have a higher risk of becoming depressed as a result of negative shocks. Following our theoretical framework and evidence from the psychological and neuroscience literature (e.g., Holtzheimer and Mayberg, 2011), we think of depression as an absorbing state. If this is the case, we should observe that our effects are persistent for those at-risk students as in Prediction 4. Furthermore, we examine other economic long-run outcomes and ask about the role of mental health for these outcomes.

6.1. *Long-Run Effects on Mental Health*

In order to explore the long-run effects of ranks on mental health, we use CES-D scores elicited in each of the following waves. Specifically, the Add Health data allows us to trace the effects of ranks over time. We can study short-term persistence using Wave II looking at mental health one year after Wave I, in 1996, medium-term persistence using Wave III approximately seven years after the initial interview (2001/2), and long-term persistence using Wave IV, when respondents were adults aged 26–32. Similar to Subsection 5.6, we estimate our main specification (2), but study the heterogeneous effects of ordinal ranks by ability decile on measures of mental health in later waves. Unfortunately, not all waves conducted the 19-item version of the CES-D, but adopted a short version in Waves III and IV comprising a subset of the original items. To compare our estimates from all waves to the baseline, we scale the mental health measures from the short scales by $^{19}/_{9}$ (Wave III) and $^{19}/_{10}$ (Wave IV) to correspond to the same range from 0 to 57 as the full scale in Wave I.[13]

Figure 7 shows that the general pattern persists over time. Across all waves, the effect of ordinal ranks is significant and pronounced at the bottom of the ability distribution, and insignificant as well as smaller in magnitude for higher-ability deciles. Online Appendix Table E.1 quantifies these effects. We find that the significant effects for the lowest ability decile persist across Waves I to IV and amount to -5.67 to -11.64 CES-D points, and fade out for higher-ability deciles.[14] This pattern is strikingly similar from Wave I, when students are 12–18 years old, to Wave IV, when those students are adults of 26–32 years. We also repeat these results using the percentilised ability control, as we discussed in Subsection 5.6. We again find our results to remain highly consistent with those discussed here, concluding that rank effects are strong and persistence for those with low ability (see Online Appendix Table E.2).

[13] Andresen *et al.* (1994) validate the short version of the CES-D and shows that it is comparable to the longer version. In the context of our data, we show in Online Appendix Figure A.2 that the short and long versions of the CES-D are highly correlated in Wave I. We therefore use scaled version to compare our results to the baseline effects documented in Subsection 5.6. Scales based on fewer items reduce the efficiency of our estimates. Simulation F in Online Appendix F suggests that the standard errors on our variable of interest increases by about 5%. As an alternative to simple scaling, we also report results from a robustness check that predicts CES-D scores in later waves based on the correlation structure in Wave I.

[14] Tests of equality between the lowest and all other ability decile yields p-values of 0.06, 0.02, 0.05 and < 0.01 and comparing ranks effects for below and above median ability students yields p-values of 0.04, 0.23, 0.07 and 0.02 for Waves I, II, III and IV, respectively.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

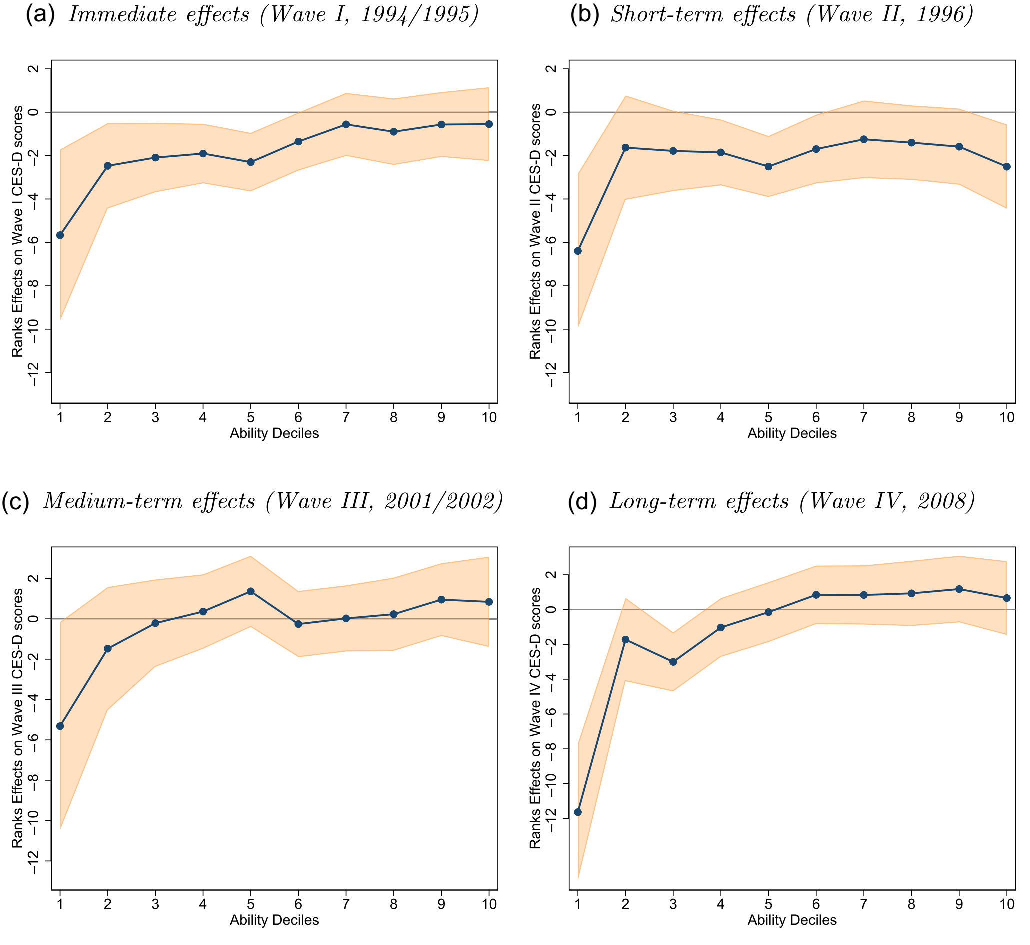

Fig. 7. *Persistent Effects of Ordinal Ranks by Ability Decile. This figure presents the effect of ordinal ranks by ability decile for each of the waves. We present the underlying regressions in Online Appendix Table E.1. The shaded area indicates 90% confidence intervals clustered at the school level.*

6.1.1. *Wave V (2016–18) results and attrition*

In principle, we could also lever data from Wave V, conducted about 23 years after the baseline in 2016–18. Performing the same analysis as for the previous waves, we do not observe any effects as shown in Online Appendix Table E.3. While this could suggest that the effects fade out 20 years after the initial shock, we think this is rather due to several problems with the data in Wave V. First, we face selective attrition in Wave V. While neither the outcome (CES-D scores) nor the treatment variables (ordinal rank and an indicator for negative shocks) are significantly related to attrition status in Waves II to IV, they are so in Wave V. In particular, those individuals who receive a negative shock, and who drive our results as shown in Subsection 5.5, are more likely to be missing in Wave V.

Second, in Subsection 5.4 we have shown that two of four facets—the lack of a positive attitude and a lack of motivation—drive our results. Yet, the CES-D instrument administered in Wave

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

V elicited only a subset of five items rather than the full CES-D scale, with only a single item that loads on these facets (cf. Online Appendix Tables A.1 and C.1), which reduces the power to detect similar effects as in the previous waves.

Finally, as we will show in the next subsection, we find the same pattern observed for mental health in a range of economic outcomes measured in Waves IV and V, suggesting that the effects indeed last for more than 14 years. We therefore think that the mental health measure available in Wave V does not allow us to extend our analysis. Note, however, that even if the effects indeed fade out between Waves IV and V, our results still show that having a low rank in school worsens mental health for low-ability students for at least 14 years.

6.1.2. *Predicting CES-D scores*

In our analysis, we used the available items in each wave and scaled the resulting total score up to correspond to the full scale as in Waves I and II. While this approach is transparent, it implicitly assumes that every available item has the same weight for the total score. Yet, as argued in Subsection 5.4, there are different facets captured by the CES-D scale inducing potentially different weights of the items. To acknowledge this unequal weighting, we construct a new set of outcome measures for each wave by regressing the total CES-D score in Wave I on the available items in the later wave (see Online Appendix Table A.1 for a list of items in each wave). In a second step, we then use the coefficients from these regressions as weights for the items in later waves, predict the total CES-D scores and replicate the regressions underlying Figure 7 using this predicted outcome measure. As shown in the Online Appendix Table E.4, the resulting estimates mirror the previous findings.

6.1.3. *Using clinical diagnoses rather than self-assessments*

CES-D scores are an established instrument to assess respondents' mental health that is widely used in clinical practice (Murphy, 2011). Yet, CES-D scores stem from a self-assessment. As an alternative measure, we construct an indicator whether a respondent was ever diagnosed with mental health disorders.[15] We present the same ability heterogeneity for this indicator in the Online Appendix Figure E.1 and the Online Appendix Table E.6, and observe the same pattern as for CES-D scores: rank effects are more pronounced for low-ability individuals and fade out with increasing ability.

These results are in line with Prediction 4, which suggested that once a negative shock reduces a student's belief in their ability sufficiently, they withdraw study effort and therefore avoid new signals. As a consequence, their belief about the returns to ability remain low, positive updating is less likely, and depressions are some form of absorbing states. In other words, their mental health remains in a poor state and negative shocks may trigger potential vicious cycles. Therefore, our results show that the school environment can have long-lasting effects on the mental well-being of students over the life cycle.

6.2. *Long-Run Effects on Economic Outcomes*

How do these long-run effects on mental health translate into other economic outcomes? Previous research suggests that worse mental health reduces educational attainment (e.g., Currie and

[15] We use information from Wave IV and V eliciting the age of the first diagnosis of depression to construct this indicator, and include an additional indicator for the wave the information stems from in the corresponding regressions. Furthermore, we drop observations where the first diagnosis occurred before Wave I.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Table 7. *Long-Run Outcomes, Mental Health, and Rank Effects.*

Dependent variable	(1) Association standardised CES-D scores	(2) Average rank effect	(3) Average rank effect (std.)	(4) p-value (H_0: No heterogeneity)	(5) Share of rank effect mediated by CES-D scores
$\mathbb{1}$\{College graduate\}	−0.05***	0.13***	0.04***	0.08	8%
	(0.00)	(0.04)	(0.01)	–	–
log(Personal income)	−0.10***	0.27**	0.08**	0.84	8%
	(0.01)	(0.12)	(0.03)	–	–
$\mathbb{1}$\{Currently employed\}	−0.02***	0.07*	0.02*	0.02	8%
	(0.00)	(0.04)	(0.01)	–	–
$\mathbb{1}$\{Ever married\}	−0.01***	0.05	0.02	0.37	5%
	(0.00)	(0.05)	(0.01)	–	–
$\mathbb{1}$\{Ever arrested\}	0.03***	−0.01	−0.00	0.00	78%
	(0.00)	(0.05)	(0.01)	–	–

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the school level. Column (1) presents the relationship of mental health measured by standardised CES-D scores in Wave I and several long-run economic outcomes as dependent variables. These specifications control for all characteristics as our baseline specification apart from the ordinal rank. The second column presents the effect of ranks on economic outcomes based on our main specification, while column (3) presents the corresponding estimates using a standardised rank measure. In column (4), we present p-values of tests of equality for rank effects of the lowest ability decile against the mean rank effect of the other deciles. Column (5) presents shares of these effect that are mediated by mental health. To obtain this share, we run an auxiliary regression where we add CES-D scores as an additional explanatory variable to the specification in column (2). We then calculate the share as the product of the coefficient from CES-D scores in this regression and the rank effect in Table 3, divided by the rank effect in column (2) of the present table. Outcomes are based on Wave V data appended by Wave IV data if the former data is missing. We add an additional indicator to control for the wave the outcome measure is from.

Stabile, 2006) and lowers employment as well as earnings (e.g., Fletcher, 2014), and has linked higher ranks to higher educational attainment (Elsner and Isphording, 2017) and income (Denning *et al.*, 2021). We conduct two analyses to add to these results. First, we assess the correlation between CES-D scores in Wave I and a range of long-term outcomes conditional on a rich set of background characteristics, to provide correlational evidence on the importance of mental health in youth for long-term outcomes. Second, we study the causal effect of ranks in Wave I on economic outcomes in adulthood. Together with our baseline estimates, we then can calculate how much of the long-run effects of rank are mediated through mental health.

In column (1) of Table 7, we show that a range of economic long-run outcomes—graduating from college, household income, and being employed—as well as important non-economic outcomes—ever being married or ever being arrested—are all significantly related to mental health measured by CES-D scores in Wave I. Importantly, these regressions control for a range of other individual characteristics and, most notably, a fourth-order polynomial in ability, as well as school-by-grade fixed effects as in our preferred baseline specification. Increasing CES-D scores by one standard deviation, i.e., worsening mental health, is associated with a 5 percentage points decrease in the probability of having a college degree, 11% lower income and being 2 percentage points less likely to be employed. In addition, we also find that those individuals with worse mental health in adolescence are also less likely to marry and are more likely to get arrested in adulthood. Although these associations are not necessarily causal, they highlight that having better mental health during adolescence predicts better economic and non-economic long-run outcomes.

We then present the effects of ordinal ranks during school on these outcomes. Being ranked higher during school significantly increases the probability of graduating from college, of being

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

employed, and it increases income. More specifically, an increase of one standard deviation in a student's ordinal rank in school increases their likelihood of graduating from college by 4 percentage points, employment by 2 percentage points, and income by 8%. The average results on college graduation mimic the effects found by Elsner and Isphording (2017), while our income results are about double the size to those reported in Denning *et al.* (2021).[16] One potential explanation for the last finding is that we can study income at a later point in life. If having a low rank sets people on different trajectories compared to those who have a high rank, this difference might increase over time explaining the effects that we observe.

For the remaining outcomes, the rank effect estimates have the sign we would expect, but we do not find evidence of ranks affecting these outcomes on average. However, as we have shown in the previous section, there exists a pronounced heterogeneity with respect to ability. Specifically, the consequences of ranks are more pronounced at the lower end of the ability distribution. In Figure 8, we therefore present the corresponding estimates for economic long-run outcomes, and report tests of equality between rank effects for the lowest ability decile against the mean rank effect of all other deciles in column (4) of Table 7. Strikingly, we observe the same qualitative pattern for all outcomes as for mental health: Rank effects are more pronounced for low-ability individuals and fade out with increasing ability.

These effects are sizeable. For the lowest ability decile, increasing the rank by one standard deviation (i.e., increasing the rank by 0.28) translates into a 9 percentage point higher propensity to obtain a college degree, increases income in adulthood by about 13%, employment by 10 percentage points and reduces the likelihood of ever being arrested by 10 percentage points. This confirms our expectations that ranks matter more for those of lower ability and highlights the importance of moving beyond average effects.

We next ask how much of these long-run effects are mediated by mental health in adolescence. Following Gelbach (2016), we decompose the rank effects for each economic long-run outcome (e_{ics}), and calculate how much of it operates through mental health (y_{ics}) as a mediator, or through other channels that cannot be attributed to mental health in adolescence (R_{ics}):

$$\frac{d\, e_{ics}}{d\, rank_{ics}} = \frac{\partial\, e_{ics}}{\partial\, y_{ics}} \frac{\partial\, y_{ics}}{\partial\, rank_{ics}} + R_{ics}.$$

While the resulting estimates only reflect causal estimates under very strong assumptions, they nevertheless are suggestive for the importance of mental health as a mediator in comparison to other channels. To operationalise this decomposition, we calculate the mediated effect as the product of the coefficient of the average rank on mental health in Wave I with the coefficient of mental health on long-run outcomes, which simultaneously controls for the ordinal rank. We then express this mediated effect as a share of the rank effect shown in column (2). Specifically, we estimate the following set of specifications (following the notation in (2))

$$\begin{aligned} y_{ics} &= \alpha^{y}\, rank_{ics} + f^{y}\left(a_{ics}\right) + \boldsymbol{X}_i'\beta^{y} + \boldsymbol{\theta}^{y}_{ics} + \epsilon^{y}_{ics} \\ e_{ics} &= \alpha^{e}\, rank_{ics} + f^{e}\left(a_{ics}\right) + \boldsymbol{X}_i'\beta^{e} + \boldsymbol{\theta}^{e}_{ics} + \epsilon^{e}_{ics} \\ e_{ics} &= \alpha^{m}\, rank_{ics} + \gamma^{m}\, y_{ics} + f^{m}\left(a_{ics}\right) + \boldsymbol{X}_i'\beta^{m} + \boldsymbol{\theta}^{m}_{ics} + \epsilon^{m}_{ics}, \end{aligned}$$

and calculate the mediated effect as a share of the rank effect as $\alpha^{y}\gamma^{m}/\alpha^{e}$.

[16] Elsner and Isphording (2017) also use Add Health data, but in contrast to them, we can lever data up to Wave V rather than Wave IV, where some individuals might still be enrolled in college. Denning *et al.* (2021) use administrative records for students in Texas and earnings earlier in life.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

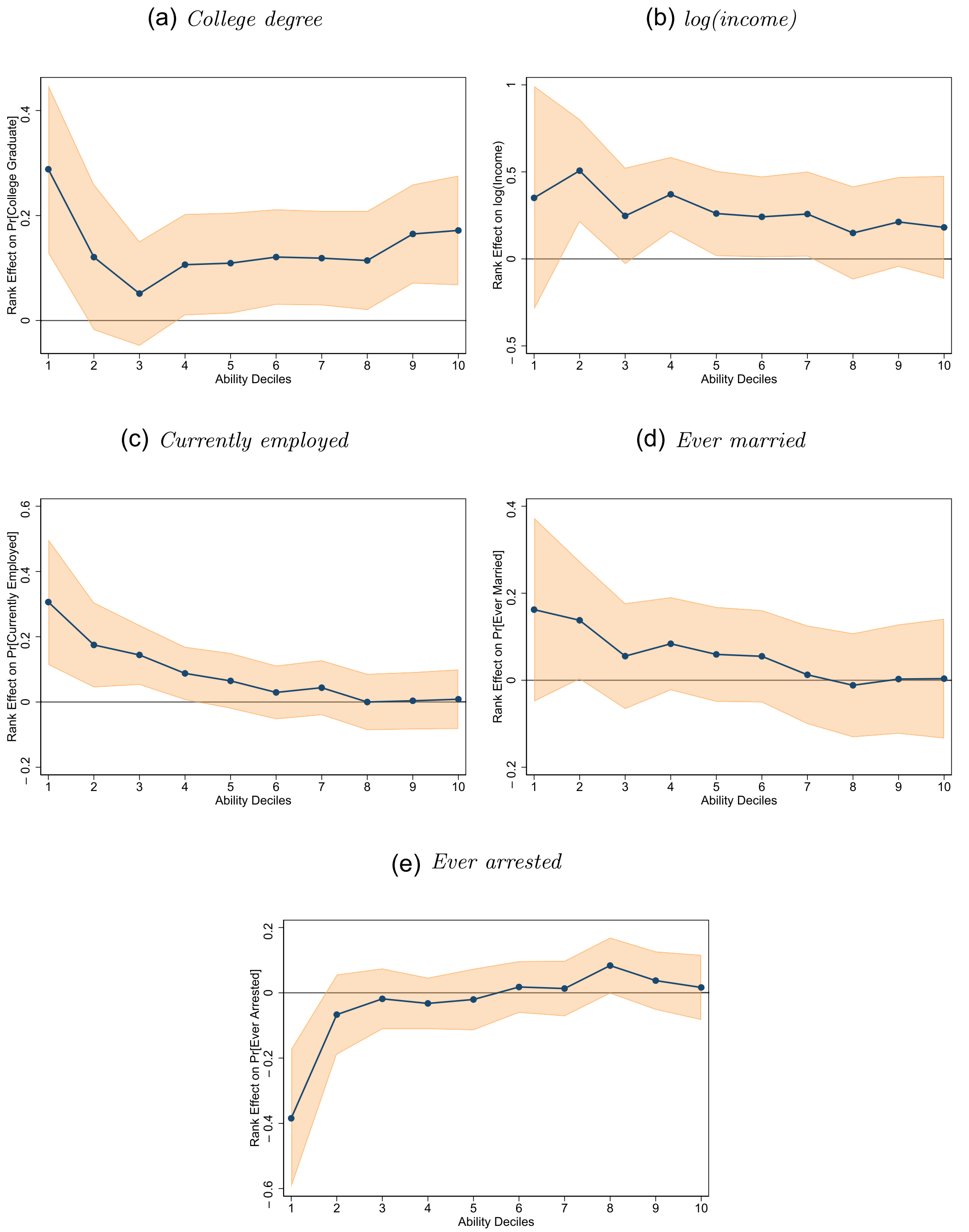

Fig. 8. *Effects of Ordinal Ranks by Ability Decile on Long-Run Outcomes. This figure presents the effect of ordinal ranks by ability decile on the outcome indicated in the title using our main specification for heterogeneous effects. The shaded area indicates 90% confidence intervals. Standard errors are clustered at the school level.*

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

The results displayed in column (5) of Table 7 suggest that we can attribute 5–8% of the rank effect for having a college degree, income, employment and marriage status mental health in adolescence. Interestingly, the decomposition attributes about 79% of the rank effects on ever being arrested to mental health. However, we caution against making conclusions on this dimension, as this particular result appears too strong and may relate to the fact the arrest measure is ever arrested since Wave I.

Overall, these results point in the same direction as our mental health results: experiencing negative shocks in school can have long-lasting negative consequences on many dimensions of life, and is particularly pronounced for at-risk students who struggle at school. Moreover, the strong association of mental health during youth and economic long-run outcomes, as well as the finding that long-run effects seem to be partly mediated by mental health in adolescence, both suggest long-lasting consequences from a poor mental health state during youth. These consequences affect an individual's economic and general well-being over the life cycle.

7. Conclusion

What are the lasting effects of the school environment in general, and peers more specifically, on the mental health of students? We provide evidence that mental health is malleable during adolescence and document its persistence over time. Investigating the causal effects of students' ordinal ranks among their peers on their mental health, we find that increasing a student's rank by one standard deviation improves their mental health by approximately 6% of a standard deviation. This effect is sizeable. It is comparable to rank effects estimated for test scores (Murphy and Weinhardt, 2020), amounts to approximately half of the effect of losing one's job on mental health (Marcus, 2013), and is about one-sixth to one-third of the effects from targeted medical treatments, or positive psychology interventions on depression and psychological well-being (Mitte *et al.*, 2005; Turner *et al.*, 2008; Bolier *et al.*, 2013). Moreover, these effects appear to operate through beliefs, are driven by negative rather than positive shocks, concentrated at the lower end of the ability distribution, and persist over time for at least 14 years. In addition, we find the same qualitative pattern in several economic outcomes measured in adulthood, and show that these effects are partly mediated by mental health during adolescence.

Given that rankings are an inherent feature of life, a natural question is how the negative consequences of ranks on mental health can be compensated. Our results and the existing literature point in three possible directions. First, investments into schools for additional counselling and a better school atmosphere may alleviate negative shocks in school. In fact, better support services in school are an important mechanism for a positive link between school spending and students' long-term economic outcomes (e.g., Jackson *et al.*, 2015), partly by enhancing students' socio-emotional development (Jackson *et al.*, 2020). Also, cognitive behavioural therapy—a leading therapy for depression—has, at least in part, a focus on correcting malformed beliefs (De Quidt and Haushofer, 2017). Given that our primary mechanism is about beliefs, a better provision of adequately trained counselling services may serve to moderate the consequences of negative shocks.

A second direction involves prevention. In particular, there is accumulating evidence that skill-enhancing interventions during childhood and youth can have long-lasting positive effects (Kautz *et al.*, 2014; Algan *et al.*, 2016; Kosse *et al.*, 2020; Sorrenti *et al.*, 2020). One skill of particular importance for the results documented in this paper is resilience. Promoting resilience through training and targeted interventions during childhood may not only compensate for the

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

immediate consequences of shocks, but protect students from being set on worse life trajectories. For instance, recent evidence by Alan *et al.* (2019) shows that grit, a non-cognitive skill associated with perseverance in response to negative shocks and higher achievement throughout life (Duckworth *et al.*, 2007), can be enhanced through targeted instruction in primary school. Eventually, such interventions could help to mediate the negative consequences of shocks in later life.

Third, the asymmetry of rank effects could provide a rationale for different classroom assignments within schools in an attempt to avoid the consequences of negative shocks. Yet, we want to caution against such assignment policies. As shown in Carrell *et al.* (2013), we do not understand the consequences of reassigning students well enough to design policies that strategically exploit peer effects. Moreover, other forms of peer effects may coexist, implying that the consequences of different assignment rules may be ambiguous (Kiessling *et al.*, 2021). As we show in Online Appendix D, even if schools employ tracking regimes that group low-ability students together, our results suggest rank effects remain, indicating a limited effectiveness of such assignment policies in our data.

We think that studying different causes of mental health and exploring strategies to cope with negative shocks are a fruitful area for future research. Given the rise of mental health issues in the developed world and the wide-spread prevalence in developing countries (for a review of the relationship of poverty and mental health, see Ridley *et al.*, 2020), policymakers have a high interest in understanding the causes of these issues to design policies that alleviate mental illnesses.

Comma Soft AG & Max Planck Institute for Research on Collective Goods, Germany
University of Strathclyde, UK

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

Online Appendix
Replication Package

References

Adhvaryu, A., Fenske, J. and Nyshadham, A. (2019). 'Early life circumstance and adult mental health', *Journal of Political Economy*, vol. 127(4), pp. 1516–49.
Ahmed, S.P., Bittencourt-Hewitt, A. and Sebastian, C.L. (2015). 'Neurocognitive bases of emotion regulation development in adolescence', *Developmental Cognitive Neuroscience*, vol. 15, pp. 11–25.
Alan, S., Boneva, T. and Ertac, S. (2019). 'Ever failed, try again, succeed better: Results from a randomized educational intervention on grit', *Quarterly Journal of Economics*, vol. 134(3), pp. 1121–62.
Alan, S., Kubilay, E., Bodur, E. and Mumcu, I. (2021). 'Social status in student networks and implications for perceived social climate in schools', Working Paper 9095, CESifo.
Algan, Y., Beasley, E., Vitaro, F. and Tremblay, R.E. (2016). 'The impact of non-cognitive skills training on academic and nonacademic trajectories: From childhood to early adulthood', Working paper, SciencesPo.
Andresen, E.M., Malmgren, J.A., Carter, W.B. and Patrick, D.L. (1994). 'Screening for depression in well older adults: Evaluation of a short form of the CES-D', *American Journal of Preventive Medicine*, vol. 10(2), pp. 77–84.
Anelli, M. and Peri, G. (2019). 'The effects of high school peers' gender on college major, college performance and income', Economic Journal, vol. 129(618), pp. 553–602.
Angrist, J.D. and Lavy, V. (1999). 'Using Maimonides' rule to estimate the effect of class size on scholastic achievement', *Quarterly Journal of Economics*, vol. 114(2), pp. 533–75.
Angrist, J.D., Lavy, V., Leder-Luis, J. and Shany, A. (2019). 'Maimonides' rule redux', *American Economic Review: Insights*, vol. 1(3), pp. 309–24.
Armstrong, T. and Olatunji, B.O. (2012). 'Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis', *Clinical Psychological Review*, vol. 32(8), pp. 704–23.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Azmat, G. and Iriberri, N. (2016). 'The provision of relative performance feedback: An analysis of performance and satisfaction', *Journal of Economics & Management Strategy*, vol. 25(1), pp. 77–110.

Baird, S., De Hoop, J. and Özler, B. (2013). 'Income shocks and adolescent mental health', *Journal of Human Resources*, vol. 48(2), pp. 370–403.

Barankay, I. (2011). 'Rankings and social tournaments: Evidence from a crowd-sourcing experiment', Working paper, University of Pennsylvania.

Bartel, A. and Taubman, P. (1986). 'Some economic and demographic consequences of mental illness', *Journal of Labor Economics*, vol. 4(2), pp. 243–56.

Beck, A.T. (1967). *Depression: Clinical, Experimental, and Theoretical Aspects*, Philadelphia: University of Pennsylvania Press.

Bell, N., Lassiter, K., Matthews, D. and Hutchinson, M. (2001). 'Comparison of the Peabody Picture Vocabulary Test—third edition and Wechsler Adult Intelligence Scale—third edition with university students', *Journal of Clinical Psychology*, vol. 57(3), pp. 417–22.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). 'Inference on treatment effects after selection among high-dimensional controls', *The Review of Economic Studies*, vol. 81(2), pp. 608–50.

Beres, K., Kaufman, A. and Perlman, M. (2000). 'Assessment of child intelligence', in (G. Goldstein and M. Hersen, eds.), *Handbook of Pyschological Assessment*, pp. 65–96, Amsterdam: Elsevier.

Biasi, B., Dahl, M.S. and Moser, P. (2021). 'Career effects of mental health', Working Paper 29031, NBER.

Bietenbeck, J. (2020). 'The long-term impacts of low-achieving childhood peers: Evidence from project star', *Journal of the European Economic Association*, vol. 18(1), pp. 392–426.

Bifulco, R., Fletcher, J.M., Oh, S.J. and Ross, S.L. (2014). 'Do high school peers have persistent effects on college attainment and other life outcomes?', *Labour Economics*, vol. 29, pp. 83–90.

Bifulco, R., Fletcher, J.M. and Ross, S.L. (2011). 'The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health', *American Economic Journal: Economic Policy*, vol. 3(1), pp. 25–53.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2013). 'Under pressure? The effect of peers on outcomes of young adults', *Journal of Labor Economics*, vol. 31(1), pp. 119–53.

Bolier, L., Haverman, M., Westerhof, G.J., Riper, H., Smit, F. and Bohlmeijer, E. (2013). 'Positive psychology interventions: A meta-analysis of randomized controlled studies', *BMC Public Health*, vol. 13, article ID 119.

Booij, A., Leuven, E. and Oosterbeek, H. (2017). 'Ability peer effects in university: Evidence from a randomized experiment', *The Review of Economic Studies*, vol. 84(2), pp. 547–78.

Borbely, D., Norris, J. and Romiti, A. (2021). 'Peer gender and schooling: Evidence from Ethiopia', Discussion paper 14439, Institute of Labor Economics (IZA).

Brown, G.D.A., Gardner, J., Oswald, A.J. and Qian, J. (2008). 'Does wage rank affect employees' well-being?', *Industrial Relations*, vol. 47(3), pp. 355–89.

Burgess, S., Greaves, E., Vignoles, A. and Wilson, D. (2015). 'What parents want: School preferences and school choice', Economic Journal, vol. 125(587), pp. 1262–89.

Buser, T., Gerhards, L. and Van der Weele, J. (2018). 'Responsiveness to feedback as a personal trait', *Journal of Risk and Uncertainty*, vol. 56(2), pp. 165–92.

Card, D., Mas, A., Moretti, E. and Saez, E. (2012). 'Inequality at work: The effect of peer salaries on job satisfaction', *American Economic Review*, vol. 102(6), pp. 2981–3003.

Carrell, S.E., Hoekstra, M. and Kuka, E. (2018). 'The long-run effects of disruptive peers', *American Economic Review*, vol. 108(11), pp. 3377–415.

Carrell, S.E., Sacerdote, B. and West, J. (2013). 'From natural variation to optimal policy? The importance of endogenous peer group formation', *Econometrica*, vol. 81(3), pp. 855–82.

Caspi, A., Houts, R.M., Ambler, A., Danese, A., Elliott, M.L., Hariri, A., Harrington, H., Hogan, S., Poulton, R., Ramrakha, S., Rasmussen, L.J.H., Reuben, A., Richmond-Rakerd, L., Sugden, K., Wertz, J., Williams, B.S. and Moffitt, T.E. (2020). 'Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study', *JAMA Network Open*, vol. 3(4), article ID e203221.

Cheng, M.M. and Udry, J.R. (2005). 'Sexual experiences of adolescents with low cognitive abilities in the US', *Journal of Developmental and Physical Disabilities*, vol. 17(2), pp. 155–72.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W. and Yagan, D. (2011). 'How does your kindergarten classroom affect your earnings? Evidence from Project Star', *Quarterly Journal of Economics*, vol. 126(4), pp. 1593–660.

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

Clark, L., Chamberlain, S.R. and Sahakian, B.J. (2009). 'Neurocognitive mechanisms in depression: Implications for treatment', *Annual Review of Neuroscience*, vol. 32(1), pp. 57–74.

Coffman, K.B., Ugalde Araya, P. and Zafar, B. (2021). 'A (dynamic) investigation of stereotypes, belief-updating, and behavior', Working Paper 29382, NBER.

Comi, S., Origo, F., Pagani, L. and Tonello, M. (2021). 'Last and furious: Relative position and school violence', *Journal of Economic Behavior & Organization*, vol. 188, pp. 736–56.

Cools, A., Fernández, R. and Patacchini, E. (2021). 'The asymmetric gender effects of high flyers', Working paper, NYU.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Coutts, A. (2018). 'Good news and bad news are still news: Experimental evidence on belief updating', *Experimental Economics*, vol. 22, pp. 369–95.
Cunha, F. and Heckman, J.J. (2008). 'Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation', *Journal of Human Resources*, vol. 43(4), pp. 738–82.
Cunha, F., Heckman, J.J. and Schennach, S.M. (2010). 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica*, vol. 78(3), pp. 883–931.
Currie, J. and Stabile, M. (2006). 'Child mental health and human capital accumulation: The case of ADHD', *Journal of Health Economics*, vol. 25(6), pp. 1094–118.
Delaney, J.M. and Devereux, P.J. (2021). 'High school rank in math and English and the gender gap in STEM', *Labour Economics*, vol. 69, article ID 101969.
Denning, J.T., Murphy, R. and Weinhardt, F. (2021). 'Class rank and long-run outcomes', *Review of Economics and Statistics*, published online ahead of print.
De Quidt, J. and Haushofer, J. (2016). 'Depression for economists', Working Paper 22973, NBER.
De Quidt, J. and Haushofer, J. (2017). 'Depression through the lens of economics: A research agenda', in (C.B. Barrett, M.R. Carter and J.-P. Chavas, eds.), *The Economics of Poverty Traps*, pp. 127–52, Chicago: University of Chicago Press.
Duckworth, A.L., Peterson, C., Matthews, M.D. and Kelly, D.R. (2007). 'Grit: Perseverance and passion for long-term goals', *Journal of Personality and Social Psychology*, vol. 92(6), pp. 1087–101.
Dunn, L.M. and Dunn, D.M. (2007). *The Peabody Picture Vocabulary Test*, 4th edn, Minneapolis, MN: NCS Pearson.
Eil, D. and Rao, J.M. (2011). 'The good news-bad news effect: Asymmetric processing of objective information about yourself', *American Economic Journal: Microeconomics*, vol. 3(2), pp. 114–38.
Eisenberg, D., Golberstein, E., Whitlock, J.L. and Downs, M.F. (2013). 'Social contagion of mental health: Evidence from college roommates', *Health Economics*, vol. 22(8), pp. 965–86.
Elsner, B. and Isphording, I.E. (2017). 'A big fish in a small pond: Ability rank and human capital investment', *Journal of Labor Economics*, vol. 35(3), pp. 787–828.
Elsner, B. and Isphording, I.E. (2018). 'Rank, sex, drugs, and crime', *Journal of Human Resources*, vol. 52(2), pp. 356–81.
Elsner, B., Isphording, I.E. and Zölitz, U. (2021). 'Achievement rank affects performance and major choices in college', Economic Journal, vol. 131, pp. 3182–206.
Ertac, S. (2011). 'Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback', *Journal of Economic Behavior & Organization*, vol. 80(3), pp. 532–45.
Ettner, S.L., Frank, R.G. and Kessler, R.C. (1997). 'The impact of psychiatric disorders on labor market outcomes', *Industrial and Labor Relations Review*, vol. 51(1), pp. 64–81.
Festinger, L. (1954). 'A theory of social comparison processes', *Human Relations*, vol. 7(2), pp. 117–40.
Fletcher, J.M. (2008). 'Adolescent depression: Diagnosis, treatment, and educational attainment', *Health Economics*, vol. 17(11), pp. 1215–35.
Fletcher, J.M. (2010). 'Adolescent depression and educational attainment: Results using sibling fixed effects', *Health Economics*, vol. 19(7), pp. 855–71.
Fletcher, J.M. (2014). 'The effects of childhood ADHD on adult labor market outcomes', *Health Economics*, vol. 23(2), pp. 159–81.
Fowler, J.H. and Christakis, N.A. (2008). 'Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham heart study', *BMJ*, vol. 337, article ID a2338.
Frank, R. and Gertler, P. (1991). 'An assessment of measurement error bias for estimating the effect of mental distress on income', *Journal of Human Resources*, vol. 26(1), pp. 154–64.
Fruehwirth, J.C., Iyer, S. and Zhang, A. (2019). 'Religion and depression in adolescence', *Journal of Political Economy*, vol. 127(3), pp. 1178–209.
Gelbach, J.B. (2016). 'When do covariates matter? And which ones? And how much?', *Journal of Labor Economics*, vol. 34, pp. 509–43.
Gotlib, I.H., Kranoperova, E., Yue, D.N. and Joorman, J. (2004). 'Attentional biases for negative interpersonal stimuli in clinical depression', *Journal of Abnormal Psychology*, vol. 113(1), pp. 121–35.
Goulas, S. and Megalokonomou, R. (2021). 'Knowing who you actually are: The effect of feedback on short- and longer-term outcomes', *Journal of Economic Behavior and Organization*, vol. 183, pp. 589–615.
Gould, E.D., Lavy, V. and Paserman, D.M. (2009). 'Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence', Economic Journal, vol. 119(540), pp. 1243–69.
Graham, B.S. (2015). 'Methods of identification in social networks', *Annual Review of Economics*, vol. 7(1), pp. 465–85.
Grossman, M. (1972). 'On the concept of health capital and the demand for health', *Journal of Political Economy*, vol. 80(2), pp. 223–55.
Harris, K.M. (2018). 'The national longitudinal study of adolescent to adult health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009; Wave V, 2016–2018 [machine-readable data file and documentation]', Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.
Hastings, J.S., Kane, T.J. and Staiger, D.O. (2009). 'Heterogeneous preferences and the efficacy of public school choice', Working paper, NBER.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Haushofer, J. and Shapiro, J. (2016). 'The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya', *Quarterly Journal of Economics*, vol. 131(4), pp. 1973–2042.

Heckman, J. and Mosso, S. (2014). 'The economics of human development and social mobility', *Annual Review of Economics*, vol. 6(1), pp. 689–733.

Heckman, J.J., Stixrud, J. and Urzua, S. (2006). 'The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior', *Journal of Labor Economics*, vol. 24(3), pp. 411–82.

Holtzheimer, P.E. and Mayberg, H.S. (2011). 'Stuck in a rut: Rethinking depression and its treatment', *Trends in Neurosciences*, vol. 34(1), pp. 1–9.

Hoxby, C.M. (2000a). 'The effects of class size on student achievement: New evidence from population variation', *Quarterly Journal of Economics*, vol. 115(4), pp. 1239–85.

Hoxby, C.M. (2000b). 'Peer effects in the classroom: Learning from gender and race variation', Working Paper 7867, NBER.

Jackson, C.K., Beuermann, D.W., Navarro-Sola, L. and Pardo, F. (2021). 'What is a good school, and can parents tell? Evidence on the multidimensionality of school output', Working Paper 25342, NBER.

Jackson, C.K., Johnson, R.C. and Persico, C. (2015). 'The effects of school spending on educational and economic outcomes: Evidence from school finance reforms', *Quarterly Journal of Economics*, vol. 131(1), pp. 157–218.

Jackson, C.K., Porter, S.C., Easton, J.Q., Blanchard, A. and Kiguel, S. (2020). 'School effects on socioemotional development, school-based arrests, and educational attainment', *American Economic Review: Insights*, vol. 2(4), pp. 491–508.

Jensen, A.R. (1998). *The g Factor: The Science of Mental Ability*, Westport, CT: Praeger.

Jensen, R. (2010). 'The (perceived) returns to education and the demand for schooling', *Quarterly Journal of Economics*, vol. 125(2), pp. 515–48.

Katz, L.F., Kling, J.R. and Liebman, J.B. (2001). 'Moving to opportunity in Boston: Early results of a randomized mobility experiment', *Quarterly Journal of Economics*, vol. 116(2), pp. 607–54.

Kautz, T., Heckman, J.J., Diris, R., Ter Weel, B. and Borghans, L. (2014). 'Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success', Working Paper 110, OECD Education.

Kessler, R.C., Angermeyer, M., Anthony, J.C., De Graaf, R., Demyttenaere, K., Gasquet, I., De Girolamo, G., Gluzman, S., Gureje, O., Haro, J.M., Kawakami, N., Karam, A., Levinson, D., Mora, M.E.M., Browne, M.A.O., Posada-Villa, J., Stein, D.J., Tsang, C.H.A., Aguilar-Gaxiola, S., Alonso, J., Lee, S., Heeringa, S., Pennell, B.E., Berglund, P., Gruber, M.J., Petukhova, M., Chatterji, S. and Üstün, T.B. (2007). 'Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's world mental health survey initiative', *World Psychiatry*, vol. 6(3), pp. 168–76.

Kiessling, L., Radbruch, J. and Schaube, S. (2019). 'Determinants of peer selection', Discussion paper, University of Bonn and University of Mannheim.

Kiessling, L., Radbruch, J. and Schaube, S. (2021). 'Self-selection of peers and performance', *Management Science*, published online ahead of print.

Korn, C.W., Sharot, T., Walter, H., Heekeren, H.R. and Dolan, R.J. (2014). 'Depression is related to an absence of optimistically biased belief updating about future life events', *Psychological Medicine*, vol. 44(3), pp. 579–92.

Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H. and Falk, A. (2020). 'The formation of prosociality: Causal evidence on the role of social environment', *Journal of Political Economy*, vol. 128(2), pp. 434–67.

Krueger, A.B. and Whitmore, D.M. (2001). 'The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project Star', Economic Journal, vol. 111(468), pp. 1–28.

Kuhnen, C.M. (2015). 'Asymmetric learning from financial information', *The Journal of Finance*, vol. 70(5), pp. 2029–62.

Kuppens, P., Allen, N.B. and Sheeber, L.B. (2010). 'Emotional inertia and psychological maladjustment', *Psychological Science*, vol. 21(7), pp. 984–91.

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

Lewinsohn, P.M., Seeley, J.R., Roberts, R.E. and Allen, N.B. (1997). 'Center for epidemiologic studies depression scale (CES-D) as a screening instrument for depression among community-residing older adults', *Psychology and Aging*, vol. 12(2), pp. 277–87.

Luttmer, E.F.P. (2005). 'Neighbors as negatives: Relative earnings and well-being', *Quarterly Journal of Economics*, vol. 120(3), pp. 963–1002.

Marcus, J. (2013). 'The effect of unemployment on the mental health of spouses—evidence from plant closures in Germany', *Journal of Health Economics*, vol. 32(3), pp. 546–58.

Marsh, H.W. (1987). 'The big-fish-little-pond effect on academic self-concept', *Journal of Educational Psychology*, vol. 79(3), pp. 280–95.

McCormick, C.M. and Green, M.R. (2013). 'From the stressed adolescent to the anxious and depressed adult: Investigations in rodent models', *Neuroscience*, vol. 249, pp. 242–57.

McPherson, M., Smith-Lovin, L. and Cook, J.M. (2001). 'Birds of a feather: Homophily in social networks', *Annual Review of Sociology*, vol. 27(1), pp. 415–44.

Mitte, K., Noack, P., Steil, R. and Hautzinger, M. (2005). 'A meta-analytic review of the efficacy of drug treatment in generalized anxiety disorder', *Journal of Clinical Psychopharmacology*, vol. 25(2), pp. 141–50.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023

Mojtabai, R., Olfson, M. and Han, B. (2016). 'National trends in the prevalence and treatment of depression in adolescents and young adults', *Pediatrics*, vol. 138(6), article ID e20161878.
Momen, N.C., Plana-Ripoll, O., Agerbo, E., Benros, M.E., Børglum, A.D., Christensen, M.K., Dalsgaard, S., Degenhardt, L., De Jonge, P., Debost, J.C.P., Fenger-Grøn, M., Gunn, J.M., Iburg, K.M., Kessing, L.V., Kessler, R.C., Laursen, T.M., Lim, C.C., Mors, O., Mortensen, P.B., Musliner, K.L., Nordentoft, M., Pedersen, C.B., Petersen, L.V., Ribe, A.R., Roest, A.M., Saha, S., Schork, A.J., Scott, K.M., Sievert, C., Sørensen, H.J., Stedman, T.J., Vestergaard, M., Vilhjalmsson, B., Werge, T., Weye, N., Whiteford, H.A., Prior, A. and McGrath, J.J. (2020). 'Association between mental disorders and subsequent medical conditions', *New England Journal of Medicine*, vol. 382(18), pp. 1721–31.
Murphy, J.M. (2011). 'Symptom scales and diagnostic schedules in adult psychiatry', in (M.T. Tsuang, M. Tohen and P.B. Jones, eds.), *Textbook of Psychiatric Epidemiology*, pp. 199–219, Chichester: Wiley Blackwell.
Murphy, R. and Weinhardt, F. (2020). 'Top of the class: The importance of ordinal rank', *Review of Economic Studies*, vol. 87(6), pp. 2777–826.
OECD. (2015). *Fit Mind, Fit Job*, Paris: OECD Publishing.
Olivetti, C., Patacchini, E. and Zenou, Y. (2020). 'Mothers, peers, and gender-role identity', *Journal of the European Economic Association*, vol. 18(1), pp. 266–301.
Oster, E. (2019). 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics*, vol. 37(2), pp. 187–204.
Pagani, L., Comi, S. and Origo, F. (2019). 'The effect of school rank on personality traits', *Journal of Human Resources*, published online ahead of print.
Plana-Ripoll, O., Pedersen, C.B., Holtz, Y., Benros, M.E., Dalsgaard, S., De Jonge, P., Fan, C.C., Degenhardt, L., Ganna, A., Greve, A.N., Gunn, J., Iburg, K.M., Kessing, L.V., Lee, B.K., Lim, C.C.W., Mors, O., Nordentoft, M., Prior, A., Roest, A.M., Saha, S., Schork, A., Scott, J.G., Scott, K.M., Stedman, T., Sørensen, H.J., Werge, T., Whiteford, H.A., Laursen, T.M., Agerbo, E., Kessler, R.C., Mortensen, P.B. and McGrath, J.J. (2019). 'Exploring comorbidity within mental disorders among a Danish national population', *JAMA Psychiatry*, vol. 76(3), pp. 259–70.
Radloff, L.S. (1977). 'The CES-D scale: A self report depression scale for research in the general population', *Applied Psychological Measurements*, vol. 1(3), pp. 385–401.
Radloff, L.S. (1991). 'The use of the center for epidemiologic studies depression scale in adolescents and young adults', *Journal of Youth and Adolescence*, vol. 20(2), pp. 149–66.
Ridley, M., Rao, G., Schilbach, F. and Patel, V. (2020). 'Poverty, depression, and anxiety: causal evidence and mechanisms', *Science*, vol. 370, article ID eaay0214.
Roberts, R.E., Andrews, J.A., Lewinsohn, P.M. and Hops, H. (1990). 'Assessment of depression in adolescents using the center for epidemiologic studies depression scale', *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, vol. 2(2), pp. 122–8.
Roiser, J.P., Elliott, R. and Sahakian, B.J. (2012). 'Cognitive mechanisms of treatment in depression', *Neuropsychopharmacology*, vol. 37, pp. 117–36.
Rosenquist, J.N., Fowler, J.H. and Christakis, N.A. (2011). 'Social network determinants of depression', *Molecular Psychiatry*, vol. 16, pp. 273–81.
Rothstein, J.M. (2006). 'Good principals or good peers? parental valuation of school characteristics, tiebout equilibrium, and the incentive effects of competition among jurisdictions', *American Economic Review*, vol. 96(4), pp. 1333–50.
Rothstein, J.M. (2017). 'Measuring the impacts of teachers: Comment', *American Economic Review*, vol. 107(6), pp. 1656–84.
Schwandt, H. (2018). 'Wealth shocks and health outcomes: Evidence from stock market fluctuations', *American Economic Journal: Applied Economics*, vol. 10(4), pp. 349–77.
Seligman, M.E.P. (1972). 'Learned helplessness', *Annual Review of Medicine*, vol. 23(1), pp. 407–12.
Sorrenti, G., Zölitz, U., Ribeaud, D. and Eisner, M. (2020). 'The causal impact of socio-emotional skills training on educational success', Working paper, Institute of Labor Economics (IZA).
Stewart, W.F., Ricci, J.A., Chee, E., Hahn, S.R. and Morganstein, D. (2003). 'Cost of lost productive work time among us workers with depression', *JAMA*, vol. 289(23), pp. 3135–44.
Stinebrickner, T. and Stinebrickner, R. (2012). 'Learning about academic ability and the college dropout decision', *Journal of Labor Economics*, vol. 30(4), pp. 707–48.
Thapar, A., Collishaw, S., Pine, D. and Thapar, A. (2012). 'Depression in adolescence', *The Lancet*, vol. 379(9820), pp. 1056–67.
Turner, E.H., Matthews, A.M., Linardatos, E., Tell, R.A. and Rosenthal, R. (2008). 'Selective publication of antidepressant trials and its influence on apparent efficacy', *New England Journal of Medicine*, vol. 358(3), pp. 252–60.
Villeval, M.C. (2020). *Performance Feedback and Peer Effects*, Essen: GLO.
Zafar, B. (2011). 'How do college students form expectations?', *Journal of Labor Economics*, vol. 29(2), pp. 301–48.
Zhang, A. (2018). 'Peer effects on mental health: Evidence from random assignment into classrooms', Working paper, SSRN.
Zimmermann, F. (2020). 'The dynamics of motivated beliefs', *American Economic Review*, vol. 110(2), pp. 337–61.

© The Author(s) 2023.

Downloaded from https://academic.oup.com/ej/article/133/649/281/6632959 by Max Planck Institute for Research on Collective Goods user on 06 March 2023