NET WORK LISER

Discussion Paper Series

IZA DP No. 18763

June 2026

Leaving School Early: A Mediation Analysis Linking Adolescent Mental Health Disorders to Early Adult Outcomes

Anna Adamecz
University College London,
ELTE KRTK and IZA@LISER

John Jerrim
University College London

Ágnes Szabó-Morvai
ELTE KRTK and
University of Debrecen

IZA DP No. 18763 June 2026

Leaving School Early: A Mediation Analysis Linking Adolescent Mental Health Disorders to Early Adult Outcomes*

Abstract

This paper examines the role of dropping out of school in the association between adolescent mental health disorders (MHDs) and subsequent early-life outcomes. Utilising an administrative panel dataset that links education, health, and employment records for half of the Hungarian school population, we track the life outcomes of a school cohort until age 22. Our findings indicate that adolescents diagnosed with an MHD between the ages of 14 and 16 are 5.8 percentage points (or 34%) more likely to drop out of secondary school compared to their peers, even after controlling for social background factors and educational performance. Furthermore, MHDs are associated with poorer early-life outcomes by age 22, including reduced employment rates, an increased likelihood of being neither in education nor employment, lower wages, and higher probabilities of motherhood, abortions, sexually transmitted diseases, and substance abuse. On average, approximately a third of these negative associations are mediated through school dropout but there are substantial differences across these outcomes. We conclude that educational policies aimed at facilitating the school-to-life transition should simultaneously address mental health disorders and dropout prevention.

JEL classification

I12, I29

Keywords

adolescent mental health, school dropout, school-to-work transition, NEET, abortion, substance abuse, sexually transmitted diseases, causal mediation analysis

Corresponding author

Anna Adamecz
a.adamecz-volgyi@ucl.ac.uk

* *Acknowledgements*: Anna Adamecz is grateful for the support from the Hungarian National Scientific Research Program (OTKA), Grant no. OTKA FK-138015. Ágnes Szabó-Morvai gratefully acknowledges financial support from the Hungarian National Scientific Research Program (OTKA), Grant no. FK131422, the Lendület programme of the Hungarian Academy of Sciences (LP2018-2/2018). Supported by the University of Debrecen Program for Scientific Publication. We thank Bálint Antal for research assistance.

1. Introduction

In recent years, especially following the COVID-19 pandemic, there has been increasing levels of reported mental health disorders (MHDs) amongst the younger population (see Foulkes & Andrews (2023) for a discussion). The World Health Organization (WHO, 2022) defines MHDs as clinically significant disturbances in an individual's cognition, emotional regulation, or behaviour. While this is obviously an important issue in its own right, it may also have important implications for young people's education and other life outcomes as well. Indeed, a collection of studies find that young people's mental health and educational outcomes are closely related (e.g. Van Poortvliet, 2024). Several studies have documented the link between adolescent MHDs and poorer outcomes during adulthood, for example, an increased probability of "NEET" (not in education, employment, or training) and unemployment (Rodwell et al., 2018), decreased wages (Philipson et al., 2020), early motherhood (Clayborne et al., 2019), and an increased prevalence of abortions (Jeon et al., 2024), sexually transmitted diseases (STDs) (Shrier et al., 2002), and substance abuse (Sibley et al., 2018).

A number of recent studies have suggested that education plays a mediating role in the relationship between adolescent MHDs and later life outcomes, in particular, wages and the probability of being NEET in early adulthood. Minh et al. (2023) found that around 15% of the total association between adolescent mental health and the probability of being NEET occurred through educational achievement. Plenty et al. (2021) and Hijdra et al. (2025) reach a similar conclusion using Swedish and Dutch data. In the United States, Johar & Truong (2014) explored how education mediated the link between depression during adolescence and wages during adulthood. Depression was associated with around a 10 to 15% wage penalty, though with the mediating effect of education (along with occupation and experience) being greater for women than men.

On the other hand, Veldman et al. (2022) and Fergusson and Woodward (2002) found that educational attainment did not mediate the association between depressive symptoms and NEET. The work of Dahmann & Schnitzlein (2019) in Germany also questions whether there is a causal relationship between education and mental health during adulthood, indicating that education also does not mediate the link between adolescent and adult mental ill-health.

Overall, the empirical evidence on the role that dropping out of school plays in the relationship between adolescent mental health and later life outcomes is inconclusive. Dropping out could

contribute to this association via two pathways. First, through *mediation effects*, where adolescent MHDs increase the probability of dropping out (Brännlund et al., 2017; Tong et al., 2023; von Simson et al., 2022) which in turn leads to poorer early-life outcomes (labour market outcomes: Campolieti et al. (2010); teen childbearing: Marcotte (2013); substance abuse: Reingle Gonzalez et al. (2016); and STDs: Mensch et al., (2020)). Second, through *interaction effects*, when dropping out intensifies the negative effects of adolescent MHDs on later outcomes, i.e. the strength of the association between adolescent MHDs and early adult outcomes varies depending on whether the young person drops out or remains in school.

The results of the existing literature on the mediating role of dropping out are not just mixed, but they also only consider a limited set of labour market outcomes. However, the social and economic costs of MHDs and dropping out might also come from other poor outcomes, including early motherhood, abortions, and risky behaviour. Furthermore, no work up to date has investigated both the mediation and interaction roles of dropping out of school at the same time, and more specifically for social and health outcomes like early motherhood, abortions, substance abuse, and STDs.

This paper hence further builds this aspect of the evidence base, drawing upon novel administrative data from Hungary. Specifically, we investigate how mental health diagnoses made between the ages of 14 and 16 are linked to earnings, employment, fertility, and risky behaviour outcomes at age 22, and the extent that these associations are mediated through the channel of school dropout. Thus, our first contribution to the literature is investigating the mediation and interaction roles of dropping out of school in the association between MHDs in adolescence and life outcomes at age 22. Based on the existing literature as presented above, we formulate two hypotheses. First, dropping out of school mediates the relationship between mental ill-health in adolescence and life outcomes, i.e., young people who suffered from mental ill-health in adolescence are more likely to drop out of school, and their increased probability of dropping out leads to poorer early-life outcomes.

Second, there is an interaction between adolescent MHDs and dropping out in the sense that MHDs have more negative effects on later life outcomes if someone has dropped out from school. This can happen because if dropped out, those with mental health problems spend less time around other people, get out of established routines, and spend their time on other activities (including engaging in risky health behaviours) compared to those still in school. We use the causal mediation framework to test these hypotheses. First, we confirm the mediation pathways between MHDs in adolescence, dropping out, and early-life outcomes using regression

analysis, and then we apply the four-way decomposition of mediation and interaction method of VanderWeele (2014).

Our second contribution is investigating these research questions for a comprehensive set of outcome variables including labour market outcomes (employment, wages, and NEET status), fertility (giving birth, abortions), and risky behaviour (STDs, substance abuse). To the best of our knowledge, this is the first paper linking adolescent mental health, dropping out as a mediator, and early adult measures of fertility outcomes and objective measures of risky behaviour. Looking at outcomes besides the labour market is important to map out the full social and economic costs of MHDs, dropping out, and their interaction.

2. Data

We use the Panel of Linked Administrative Data (Admin3) database, provided by the Databank of the ELTE Centre for Economic and Regional Studies (ELTE KRTK). The anonymised dataset links individual data from the National Insurance Fund Administration, the Hungarian State Treasury, the Educational Authority, the Ministry of Finance, and the National Tax and Customs Administration (Sebők, 2019). It covers data on school enrolment (the national administrative school census), labour market outcomes, and all inpatient and outpatient healthcare events (i.e., doctor visits and hospital treatments) for a random 50% of the population born before 1 Jan 2003 (people with a Social Security Number in 2003). This means about 50,000 people per birth year. Healthcare and education data are available from 2009 to 2017. For the cohort that we are looking at, the data are also linked to the National Basic Competencies Database, that provides information on centrally organised national exam scores and survey-collected data on family background. We focus on individuals born between June 1994 and May 1995, allowing us to observe mental health at ages 14–16 and outcomes up to age 22.

2.1 Mental health disorders

All healthcare event data points contain the International Statistical Classification of Diseases (ICD-10) codes of diagnoses given, which we use to define MHDs. Thus, MHDs are measured as getting inpatient or outpatient care for specific mental health disorders. We do not see directly if one had mental health problems, we observe healthcare utilisation and recorded diagnoses. For our main measure, we define MHD as having any diagnosis between ages 14 and 16 in the following categories: anxiety and stress-related disorders, psychotic disorders, mood disorders, personality disorders, and behavioural and neurodevelopmental disorders. The description and

ICD codes of these categories are presented in Table A1 in Appendix A. Overall, 5% of the sample was diagnosed with an MHD between ages 14 and 16 at least once (Table A2 in Appendix A).

2.2 Dropping out

We define dropping out as the event of leaving school without returning by Dec 2017 and earning a secondary degree. A dropout is thus a person who is not in school (and would not return to school until Dec 2017) and did not earn a secondary degree. The month of dropping out is defined as the next month right after being in school for the last time. In our sample, 17% of young people dropped out (Table A2 in Appendix A).

For this cohort, compulsory schooling lasted until the end of the academic year in which they reached the age of 18. Thus, most students dropped out of school between the ages of 18 and 20 (86 %, Figure S1_1 in the Supplementary material). Dropping out before this age was also possible if someone got married, had a child, or had severe illness or some other circumstances that made going to school impossible. The academic year lasts until June, so most dropouts dropped out in July (Figure S1_2 in the Supplementary material).

2.3 Early-life outcomes: outcome variables

Labour market outcomes. We investigate the following labour market outcomes: *NEET* (neither in school according to official administrative education enrolment data nor in employment according to the administrative employment records of the tax authority), *employed* (reported by the administrative employment records of the tax authority, share of months with employment history), *hours worked per week* if employed (as reported by employers to the tax authority), *log hourly wage* if employed (gross wage earned per hours worked according to the administrative wage records of the tax authority, reported by the employers). For those not in employment, the value of hours worked and wages are missing and thus those individuals are not included in the empirical models investigating these two outcomes.

Fertility outcomes. We use the ICD-10 codes to detect hospital in-patient and out-patient events of deliveries (ICD codes O6, O7, and O8) and induced surgical abortions (ICD code O04). The data covers most relevant events, for instance, we observe 86-93% of births and 95-98% of abortions reported in official statistics. Some of the birth records are missing because of children born in private hospitals, at home, or abroad, while some of the abortion records are missing due to abortions in private hospitals or abroad. Using these ICD codes we define two

fertility outcome variables: the probability of motherhood, which is one if a woman has given birth at least once by age 22 and zero otherwise; and the cumulative number of surgical abortions. Note that in Hungary all medically induced abortions are surgical as the abortion pill is not available.

Indicators of risky behaviour. We consider the cumulative number of STDs by age 22 and the monthly prevalence of *substance abuse diagnoses* as objective indicators of risky behaviour. *STDs* are coded as A51-A64, B20-B27, and Z21 in the ICD. *Substance abuse* diagnoses refer to psychoactive substance use (F10-19); findings of drugs and other substances, not normally found in blood (R78); poisoning by narcotics and psychodysleptics [hallucinogens] (T40); poisoning by psychotropic drugs (T43); sequelae of poisoning by drugs, medicaments, and biological substances (T96); poisoning by and exposure to alcohol, undetermined intent (Y15); and evidence of alcohol involvement determined by level of intoxication (Y91). The descriptive statistics of all outcome variables are presented in Table A2 in Appendix A.

2.4 Control variables

Age and gender are provided by the core administrative social security records. In our main models, we control for a rich set of family background variables that are reported in the NABC in take-home surveys completed by the parents. These are the following: *time spent in kindergarten* (pre-school typically between ages 3 and 6), which is a categorical variable capturing seven categories (0,<1 year; 1 year, 1-2 years, 2 years, 3 years, 4+ years); a binary variable that indicates any *financial problems of the family*, including receiving free meals or subsidised meals, receiving free school books, or regular child protection support due to a vulnerable social status; a binary variable indicating whether the student *lives with the mother only, the father only, or both*; a continuous variable showing *the number of persons living together in the household*; and the respondent's *father's and mother's highest educational attainment*.

Furthermore, we also control for centralised *math and reading test results*, taken in the spring of Grade 8 (at ages 14-15) and recorded in the NABC data. The test scores are assigned to quintiles, with a missing test score as a separate category. Grade 8 test scores are missing for 18% of the sample (Table A2 in Appendix A).

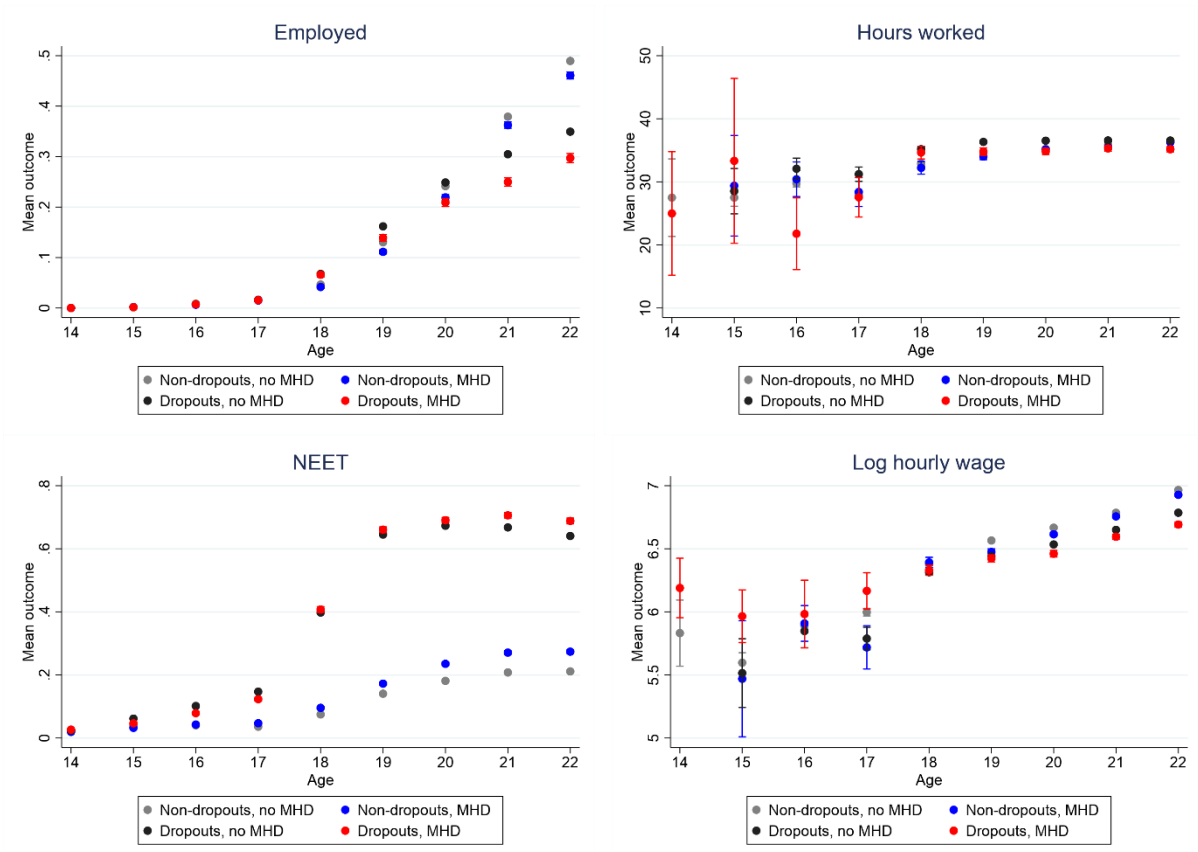
Missing test scores in the NABC and family background measures is not random, and is correlated with social background and ability: those with missing NABC data are likely the most disadvantaged students. NABC tests are low-stakes for students but high-stakes for

schools; thus, some schools might ask some low-ability students not to participate in order to increase school-level test scores. The background questionnaires are missing for 11% of the sample (Table A2 in Appendix A). We impute the missing values of background variables using the mean (for continuous variables) or the mode (for categorical variables) of the rest of the sample and generate indicator variables for the imputed missing values and include them as well in the regressions as suggested by Kézdi and Békés (2021).

3. Descriptive statistics

Figures 1, 2, and 3 compare medium-term outcomes by dropout and MHD status. The points indicate means and the vertical lines indicate 95% confidence intervals. Figure 1 illustrates the evolution of labour market outcomes by age across the four groups. Employment probabilities begin to diverge from age 20, with MHD status being more strongly associated with employment likelihood, whereas differences in hours worked are relatively minor across groups. Monthly and hourly wages also vary, with non-dropouts without an MHD earning, on average, 45% more than dropouts with an MHD. The probability of school attendance and being NEET is primarily influenced by dropout status.

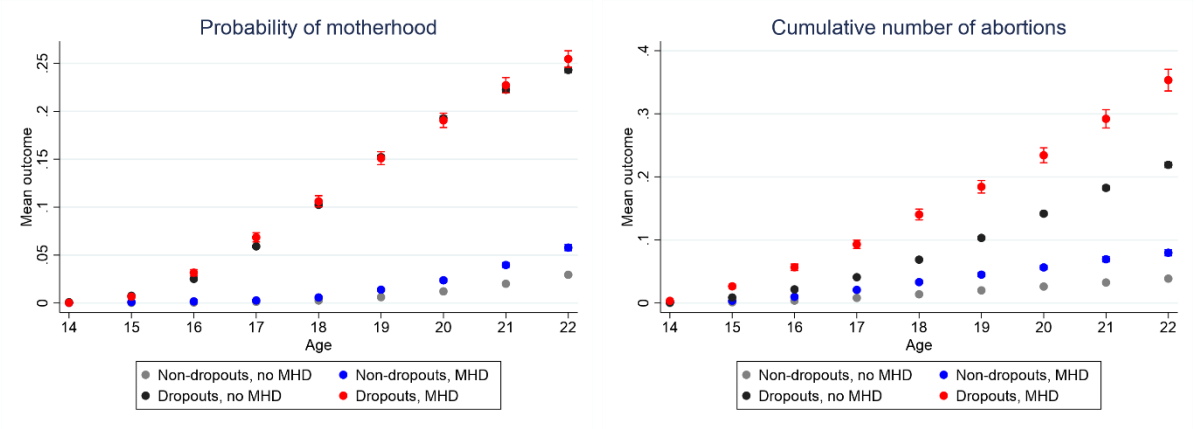
Figure 1: Labour market outcomes by age, dropout and MHD status



Source: Admin3. MHD: has been diagnosed with an MHD at least once between ages 14 and 16. Estimated means are plotted with their 95% confidence intervals, CIs are not visible when the standard errors are very low. Sample of those born between June 1994 and May 1995. No. of individuals: 51,801. Hours worked and wages are defined only for those who work.

Figure 2 illustrates that dropping out of school is a major factor associated with teenage and early fertility outcomes. There is a striking difference in the probability of motherhood between dropouts and non-dropouts, whereas the presence of an MHD plays a much smaller role in these outcomes. In contrast, when it comes to abortions, both dropout status and MHD are significant factors. Women who drop out and have an MHD are approximately eight times more likely to become mothers and have seven times as many abortions by age 22, compared to non-dropout women without an MHD.

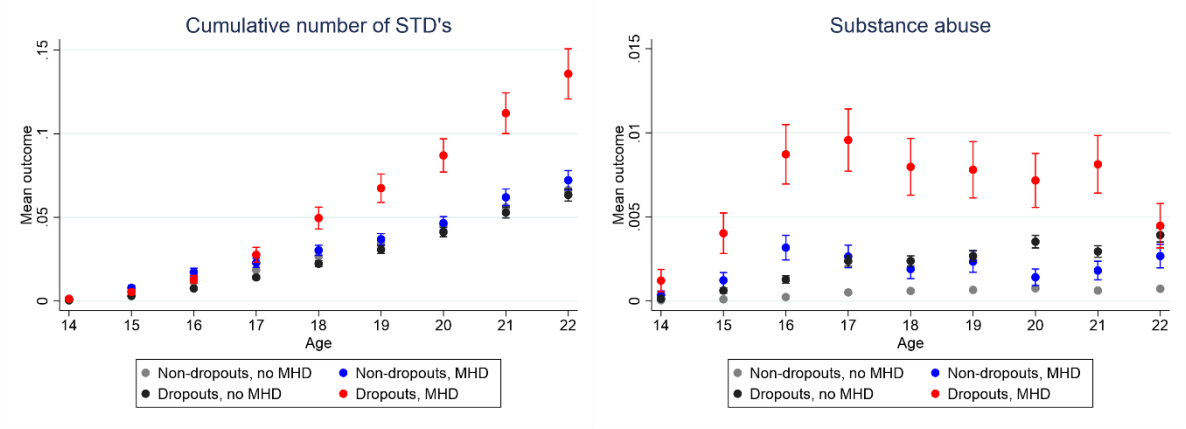
Figure 2: Fertility outcomes by age, dropout and MHD status



Source: Admin3. MHD: has been diagnosed with an MHD at least once between ages 14 and 16. Estimated means are plotted with their 95% confidence intervals, CI`s are not visible when the standard errors are very low. Sample of those born between June 1994 and May 1995. No. of individuals (women): 25,251.

Figure 3 shows the prevalence of risky behaviour-related diagnoses, substance abuse, and STDs. The monthly prevalence of substance abuse cases is 5 to 10 times higher among individuals who are both dropouts and had an MHD between the ages of 14 and 16, compared to the rest of the three groups. However, when either factor occurs alone, it is not associated with a drastically elevated incidence of substance abuse. We see a similar picture for STDs: dropouts with MHDs have about two times as many STD diagnoses between ages 18 and 22 than the other three groups.

Figure 3: Risky behaviour outcomes by age, dropout and MHD status



Source: Admin3. MHD: has been diagnosed with an MHD at least once between ages 14 and 16. Estimated means are plotted with their 95% confidence intervals, CIs are not visible when the standard errors are very low. Sample of those born between June 1994 and May 1995. No. of individuals: 51,801.

4. Empirical methods

4.1 Verifying the mediation pathways

We start the analysis by examining the statistical relationship between having at least one MHD diagnosis between ages 14 and 16 (which is our exposure variable in the context of causal mediation analysis), dropping out (the mediator), and early-life outcomes, to verify the mediation pathways.

The association between adolescent MHDs and dropping out. First, we investigate the role of MHDs in school dropout probability using probit models of the following form on a cross-sectional sample:

$$P(DROPOUT_i = 1 | \, MHD, X_i) = \Phi(\beta_1 \text{MHD}_i + \gamma' X_i + \varepsilon_i) \qquad (1)$$

The binary dependent variable $DROPOUT_i$ indicates whether the individual i is a school dropout at the age of 22. The main explanatory variable of interest is MHD_i, mental health disorder, which is equal to 1 if student i was diagnosed with a mental health disorder between ages 14 and 16 at least once and 0 otherwise. For our main results, we exclude diagnoses occurring after age 16, thus, if a diagnosis occurs after that, we include her as if she was never diagnosed. As a result, the MHD diagnoses precedes the dropout in time in each case. X_i is a matrix of explanatory variables as described in Section 2.4. We also provide a robustness check in the Supplementary material where we look at MHD diagnoses by age 17 that leads to similar results. We employ simple heteroscedasticity-robust standard errors in all models.

The association between adolescent MHDs and early-life outcomes. After exploring the role of MHDs in dropping out, we look at the association between MHDs and the eight early-life outcomes (Y_i) presented earlier. First, we estimate the following regression with and without including the set of control variables detailed above:

$$Y_i = \beta_1 + \beta_2 \mathrm{MHD}_i + \gamma' X_i + \varepsilon_i \quad (2)$$

Then, we extend Equation (2) by adding whether individual *i* is a dropout to the right-hand side:

$$Y_i = \beta_1 + \beta_2 \mathrm{MHD}_i + \beta_3 \mathrm{DROPOUT}_i + \gamma' X_i + \varepsilon_i \quad (3)$$

This specification investigates the contribution of dropping out to the statistical association between MHDs and early-life outcomes. We are interested in whether adolescent MHDs predict these outcomes even after controlling for dropping out of school. Lastly, we also test the interaction between MHDs and dropping out of school as

$$Y_i = \beta_1 + \beta_2 \mathrm{MHD}_i + \beta_3 \mathrm{DROPOUT}_i + \beta_4 \mathrm{DROPOUT}_i * MHD_i + \gamma' X_i + \varepsilon_i \quad (4)$$

Finally, for an easier interpretation, we re-estimate Equation 2 separately on the subsamples of non-dropouts and dropouts and separately investigate the role of MHDs in these two subsamples.[4]

4.2 Four-way decomposition analysis

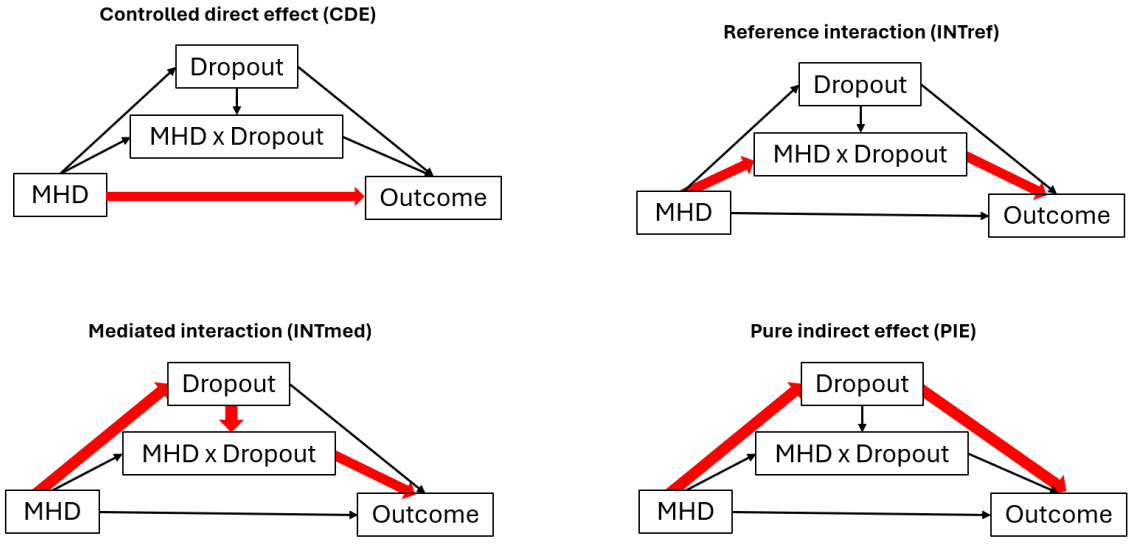
We use four-way decomposition causal mediation analysis (VanderWeele, 2014) to test the mediation and interaction role of dropping out in the relationship between MHDs in adolescence and early-life outcomes. The exposure is having at least one MHD diagnosis between ages 14 and 16, which is a binary variable, the mediator is dropping out of school, which is also a binary variable, while among the eight outcome variables, we have binary and continuous variables as well.

While traditional causal mediation designs assume that there is no interaction effect between the exposure and mediator (Imai et al., 2010), VanderWeele (2014) unifies the analysis of mediation and interaction effects. It decomposes the baseline association between having an MHD and the outcomes (total effect, TE) to four additive theoretical parts (these are never observed but could potentially be estimated in a sample on average, see Figure 4). (1) *Pure indirect effect (PIE)*. This effect is due to mediation only, and it stands for the mediation effect

[4] The full results of verifying the mediation pathways are presented in Tables S2_1-S2_9 in the Supplementary material.

of dropping out in the absence of interaction effects. (2) *Reference interaction effect (INTref)*. This effect is due to interaction only, and it refers to the interaction between MHD and dropping out in the theoretical absence of the exposure. (3) *Mediated interaction effect (INTmed)*. This is due to both mediation and interaction, and it accounts for both mediation and interaction effects when MHDs affect the probability of dropping out as well as the outcome. (4) *Controlled direct effect (CDE)*. This effect is due neither to mediation nor interaction. This is the effect of MHDs on the outcomes in the theoretical absence of the mediator.

Figure 4: The main effects we estimate using the four-way decomposition of VanderWeele (2014)



Notes: Figure edited by the authors. Original idea: Du et al. (2023)

We use part (1) of the decomposition to estimate what fraction of the total effect is mediated by dropping out:

$$Proportion\ Mediated\ (PM) = \frac{E(INTmed) + E(PIE)}{E(TE)}$$

Similarly, we use parts (2) and (3) to estimate the proportion of the total effect attributable to the interaction:

$$Proportion\ Attributable\ to\ Interaction\ (PI) = \frac{E(INTmed) + E(INTref)}{E(TE)}$$

E(.) denotes the average estimated effect for the population in the absence of confounding.

The four-way decomposition of mediation and interaction method of VanderWeele (2014) relies on four identification assumptions. First, we assume that the effect of MHDs on early-life outcomes is unconfounded conditional on our control variables, i.e. there is no such unobserved variable that would affect both MHDs and the outcomes. Second, the effect of dropping out on the outcome variables is unconfounded conditional on our control variables and MHDs. Third, the effect of MHDs on dropping out is unconfounded conditional on the control variables. Lastly, none of the confounders of the relationship between dropping out and the outcome variables are themselves affected by MHDs. These are strong assumptions, as neither MHDs nor dropping out are allocated randomly to individuals. However, we believe that as we measure MHDs between ages 14 and 16, most people drop out between ages 18 and 20, and we measure the outcome variables at age 22, we can at least separate the exposure, the mediator, and the outcome over time. Furthermore, we control for the individual characteristics mentioned most frequently as potential confounders on the MHD-dropout-outcome pathway (Minh et al., 2023).

4.3 Robustness checks

4.3.1 Sensitivity to potential unmeasured confounding in the four-way decomposition

To assess the robustness of our estimated four-way decomposition effects to potential unmeasured confounding, we conduct a sensitivity analysis using the approach of VanderWeele and Ding (2017). For each model, we compute E-values for the total effect and for the natural direct and indirect effects derived from the four-way decomposition (VanderWeele, 2014). For binary outcome variables (employed, NEET, motherhood), the E-value represents the minimum strength of association (on the risk-ratio scale) that an unmeasured confounder would need to have with both the exposure (MHDs) and the outcome, beyond the measured covariates, to fully explain away the observed effect. It is constructed as $E = RR + [RR \times (RR - 1)]^{0.5}$, where RR is the relative risk ratio of binary outcomes between the treated and the untreated group. Larger E-values indicate greater robustness to unmeasured confounding. Because the E-value framework applies to total and natural effects, we interpret these values for those components of the four-way decomposition; interaction components are not directly assessed.

For continuous outcome variables, E-values are computed using the approximation proposed by VanderWeele and Ding (2017, Table 2), which converts standardised mean differences to equivalent risk ratios as follows: $RR^* = e^{0.91 \times |d|}$, where d is the standardised (conditional)

mean difference in the outcome variable between the treated and the untreated group (i.e., the estimated effect that we want to test). We implement this adjustment manually to obtain comparable sensitivity estimates across outcome types. This allows a comparable interpretation of robustness to unmeasured confounding across outcome types, although the resulting E-values for continuous outcome variables are approximate and should be interpreted cautiously.

4.3.2 Sensitivity to measurement error in MHDs

Our exposure variable, MHD, captures only those individuals who went to see a doctor and got diagnosed, but this variable does not capture everybody with mental health problems. In our data about 5% of 14 to 16 year old teenagers get a diagnosis, as opposed to a recent estimation for Hungary of 15.8% based on national data used in the Semmelweis Plan (Kapócs, 2017). This estimation for Hungary is in line with UNICEF brief estimates that in Europe as a whole, 16.3% of 10–19-year-olds live with a mental disorder (UNICEF, 2021). According to a systematic literature review by Verhoog et al. (2024), those with lower socioeconomic backgrounds and unstable families are more likely to stay undiagnosed, which group is also more likely to drop out from school and have worse later outcomes. In general, this suggests that our estimated effect sizes are biased towards zero (attenuation bias due to measurement error), thus, our estimates are the lower bounds of the actual effects.

To assess the robustness of our results to this potential measurement error, we implement a regression-calibration sensitivity analysis for misclassification, a widely used approach that replaces the misclassified binary exposure with its posterior expected "true" value conditional on observed data and assumed sensitivity/specificity parameters (Carroll et al., 2006). MHDs are likely measured with imperfect sensitivity (under-detection) but high specificity. We therefore assumed plausible ranges for diagnostic sensitivity (Se $\in$ [0.3, 0.9]) and almost perfect specificity (Sp $\in$ [0.99]) based on prior evidence on underdiagnosis in adolescent populations (Merikangas et al., 2010). For each (Se, Sp) combination, we estimated the posterior expected probability of true disorder status, $E[A \mid A^*, C, Se, Sp]$, where A^* denotes the observed diagnosis and C the vector of covariates. This expectation was computed via Bayes' rule, combining the estimated $P(A^* = 1 \mid C)$ from a logistic model with the assumed misclassification parameters. The resulting posterior mean was substituted for the observed exposure in the four-way decomposition that we used above. This approach adjusts for classical exposure misclassification by replacing the binary exposure with its expected "true" value conditional on covariates, thereby attenuating bias toward the null that would arise from diagnostic under-ascertainment.

We re-estimate the four-way decomposition under each (Se, Sp) scenario and report the resulting total effect (TE) and proportion mediated (PM). Deviations from the baseline estimates provide a direct assessment of the extent to which exposure misclassification may attenuate the estimated effects and distort the relative contribution of the mediated pathway. As the correction relies on assumed misclassification parameters, the results should be interpreted as a sensitivity analysis rather than point identification.

4.3.3 Sensitivity to capturing MHDs between ages 14-17 instead of 14-16

Our baseline exposure captures MHDs diagnosed between ages 14 and 16, but some adolescents may receive a first diagnosis slightly later, and the choice of upper bound could in principle affect both who is classified as exposed and the timing of exposure relative to dropping out. To assess the sensitivity of our findings to this definitional choice, we re-estimate the four-way decomposition using an alternative exposure that captures any MHD diagnosis between ages 14 and 17.

5. Results

The main results of verifying the mediation pathways are presented in Appendix A while the detailed results are presented in the Supplementary material. Table A3 (Cols 1 and 2) shows that for those who had at least one MHD diagnosis between ages 14 and 16, the probability of dropping out is 12.7 percentage points (pp) higher (or 74.7% relative to the 17% sample mean). After controlling for a rich set of background characteristics and math and reading test score quintiles in specification (Cols 3 and 4), the probability of dropping out is 5.8 pp (34.1%) higher for those with an MHD diagnosis. This is a strong association between dropping out and having an MHD, providing evidence that our mediator is associated with the exposure variable.

In Table A4 in the Appendix, we report the associations between each outcome variable in turn and the exposure variable, having an MHD diagnosis. These results confirm that all outcome variables are strongly associated with having an MHD diagnosis, and these relationships have the expected signs and large magnitudes. These associations also prevail when we also control for whether they dropped out of school, suggesting both direct and mediated effects. The interaction term between having MHDs and being a dropout are significantly different from zero for five out of the eight outcome variables, supporting that applying causal mediation methods that assume no interaction effect would not have been ideal for our analysis.

Table 1 presents the results of the four-way causal mediation decomposition for labour market outcomes. The estimated total effects of having an MHD on labour market outcomes are similar

in magnitude to the naïve linear regression estimates discussed earlier. Individuals diagnosed with an MHD between ages 14 and 16 are 3.0 pp (6.5%) less likely to be employed, 5.7 pp (19.7%) more likely to be NEET, and earn 5.8% (6 log points) lower hourly wages than those without such diagnoses. The estimated total effect on hours worked is small and statistically insignificant.

School dropout mediates approximately 34.8% of the total effect of MHDs on employment, 34.5% on NEET status, and 26.7% on log hourly wages. Although the estimated proportion mediated for hours worked is relatively large (50.4%), it is not statistically significant. The estimated interaction components are generally small and often statistically insignificant, implying that the joint effect of MHDs and school dropout does not substantially differ from the sum of their individual effects. Overall, these results indicate that early school leaving is an important pathway linking adolescent mental health disorders to later labour market disadvantages.

Table 1: Causal mediation analysis of the association between MHDs and labour market outcomes: The role of dropping out of school

	Employed		NEET		Hours worked		Log hourly wage	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
TE: Total effect	-0.030	(-0.048,-0.018)***	0.057	(0.042,0.069)***	-0.225	(-0.660,0.246)	-0.060	(-0.081,-0.036)***
Proportions attributable to mediation and interaction								
PM: Proportion mediated	0.348	(0.180,0.663)***	0.345	(0.264,0.435)***	0.504	(-5.022,3.111)	0.267	(0.186,0.465)***
PI: Proportion attributable to interaction	0.099	(-0.966,0.831)	-0.072	(-0.537,0.324)	-3.417	(-22.251,42.105)	-0.984	(-1.971,-0.387)*
Four-way decomposition								
CDE: Controlled direct effect	-0.018	(-0.039,0.006)	0.042	(0.018,0.066)***	-0.948	(-1.851,-0.003)	-0.108	(-0.159,-0.057)***
INTref: Reference interaction effect	-0.003	(-0.030,0.027)	-0.003	(-0.030,0.021)	0.837	(-0.051,1.848)	0.063	(0.024,0.111)*
INTmed: Mediated interaction effect	0.000	(-0.003,0.003)	0.000	(-0.003,0.003)	-0.063	(-0.144,0.003)	-0.006	(-0.009,-0.003)*
PIE: Pure indirect effect	-0.012	(-0.015,-0.009)***	0.018	(0.015,0.024)***	-0.051	(-0.081,-0.027)***	-0.012	(-0.015,-0.009)***
Two-way decomposition								
TNIE: Total indirect effect	-0.012	(-0.015,-0.006)***	0.018	(0.015,0.024)***	-0.114	(-0.192,-0.042)***	-0.015	(-0.024,-0.012)***
PNDE: Pure direct effect	-0.021	(-0.042,-0.006)***	0.036	(0.027,0.051)***	-0.111	(-0.561,0.333)	-0.045	(-0.066,-0.021)***

Source: Admin3. The results of a four-way decomposition causal mediation analysis by VanderWeele (2014), operationalised via the CMAverse package in R (Shi et al., 2021). 95% bootstrapped confidence intervals in parentheses. Significance: 1% ***; 5% **; 10% *.

Table 2 presents the decomposition results for fertility outcomes at age 22. The total effect of having an MHD diagnosis between ages 14 and 16 is positive and statistically significant for both the probability of motherhood (2.1 pp or 14.0%) and the cumulative number of abortions (0.063 abortions or 39.4%). This indicates that adolescents with MHDs are more likely to become mothers and to have experienced abortions by early adulthood compared with their peers without such diagnoses.

For abortions, both the controlled direct effect (CDE = 0.156) and the indirect pathways are significant, while the reference interaction term is negative and statistically significant (INTref = –0.108). This pattern suggests that dropping out mediates a meaningful share of the link between adolescent MHDs and abortion outcomes, but that the two risk factors partly operate through overlapping pathways. The total indirect effect (TNIE = 0.018) indicates that dropout explains about 27% of the total association. The negative interaction term implies attenuation: once dropout status is accounted for, the additional effect of MHDs on abortion risk becomes smaller, consistent with dropout capturing part of the same disadvantage associated with mental health problems. Overall, these results indicate that school dropout is a key pathway linking adolescent mental health disorders to early fertility behaviours, mediating nearly half of the effect on motherhood and about one-quarter of the effect on abortion risk.

Table 2: Causal mediation analysis of the association between MHDs and fertility outcomes: The role of dropping out of school

	Probability of motherhood		Cumulative number of abortions	
	Estimate	95% CI	Estimate	95% CI
TE: Total effect	0.021	(0.012,0.030)***	0.063	(0.042,0.081)***
Proportions attributable to mediation and interaction				
PM: Proportion mediated	0.483	(0.303,0.759)***	0.267	(0.198,0.360)***
PI: Proportion attributable to interaction	0.030	(-0.978,1.251)	-1.590	(-2.229,-0.927)***
Four-way decomposition				
CDE: Controlled direct effect	0.009	(-0.015,0.036)	0.156	(0.093,0.216)***
INTref: Reference interaction effect	0.000	(-0.024,0.027)	-0.108	(-0.162,-0.057)***
INTmed: Mediated interaction effect	0.000	(-0.003,0.003)	0.009	(0.003,0.012)***
PIE: Pure indirect effect	0.009	(0.006,0.012)***	0.009	(0.006,0.012)***
Two-way decomposition				
TNIE: Total indirect effect	0.009	(0.006,0.012)***	0.018	(0.012,0.021)***
PNDE: Pure direct effect	0.009	(0.003,0.021)*	0.048	(0.030,0.063)***

Source: Admin3. The results of a four-way decomposition causal mediation analysis by VanderWeele (2014), operationalised via the CMAverse package in R (Shi et al., 2021). 95% bootstrapped confidence intervals in parentheses. Significance: 1% ***; 5% **; 10% *.

Table 3 presents the results for risky health behaviours. The total effect of having an adolescent MHD on the number of STD diagnoses is positive and statistically significant, indicating that individuals with an MHD between ages 14 and 16 have 0.021 (30.0%) more STD diagnoses by early adulthood compared to those without such diagnoses. The total effect for substance abuse is highly significant at 0.006 (60.0%), suggesting that MHDs modestly increase the risk of substance-related problems.

For STDs, both the controlled direct effect (CDE in Table 3) (0.066) and the reference interaction effect (INTref in Table 3) (–0.048) are statistically significant. This pattern indicates that MHDs exert a direct influence on the risk of STDs, while the negative interaction term suggests that the combined effect of MHDs and school dropout is smaller than the sum of their individual effects—consistent with partial overlap in the pathways through which the two risk factors operate. The mediated interaction effect (INTmed in Table 3) is small but significant (0.003), implying that the pathway involving both mediation and interaction plays a minor role. The pure indirect effect (PIE in Table 3) and total indirect effect (TNIE = 0.003) indicate that school dropout mediates a small portion of the relationship between adolescent MHDs and STD risk. The estimated proportion mediated (PM) is 18.6%, but it is not statistically significant.

For substance abuse, the decomposition shows a smaller total effect, largely explained by the pure direct effect (PNDE = 0.006), while the indirect and interaction components are minimal. The proportion mediated is modest but statistically significant (PM=10.5%), suggesting that school dropout accounts for a limited but detectable share of the association between adolescent MHDs and later substance use. Overall, these results indicate that leaving school early plays only a minor mediating role in linking adolescent mental health disorders to risky health behaviours, with most of the total effect being direct rather than mediated or interactive.

Table 3: Causal mediation analysis of the association between MHDs and risky behaviour outcomes: The role of dropping out of school

	STDs		Substance abuse	
	Estimate	95% CI	Estimate	95% CI
TE: Total effect	0.021	(0.003,0.039)*	0.006	(0.003,0.009)***
Proportions attributable to mediation and interaction				
PM: Proportion mediated	0.186	(-0.006,0.705)	0.105	(0.003,0.291)*
PI: Proportion attributable to interaction	-2.082	(-8.049,1.068)	0.321	(-1.161,2.514)
Four-way decomposition				
CDE: Controlled direct effect	0.066	(0.018,0.123)*	0.003	(-0.006,0.012)
INTref: Reference interaction effect	-0.048	(-0.102,-0.003)*	0.003	(-0.006,0.012)
INTmed: Mediated interaction effect	0.003	(0.000,0.006)*	0.000	(0.000,0.000)
PIE: Pure indirect effect	0.000	(0.000,0.000)	0.000	(0.000,0.000)***
Two-way decomposition				

TNIE: Total indirect effect	0.003	(0.000,0.009)*	0.000	(0.000,0.000)*
PNDE: Pure direct effect	0.018	(0.000,0.036)	0.006	(0.003,0.009)*

Source: Admin3. The results of a four-way decomposition causal mediation analysis by VanderWeele (2014), operationalised via the CMAverse package in R (Shi et al., 2021). 95% bootstrapped confidence intervals in parentheses. Significance: 1% ***; 5% **; 10% *.

5.1 Robustness checks

Table S3_1 in the Supplementary material presents the sensitivity analysis assessing the robustness of the estimated effects from the four-way decomposition to potential unmeasured confounding, following the E-value framework of VanderWeele & Ding (2017). The E-values indicate that the estimated effects are moderately robust to potential unmeasured confounding. For the main binary outcomes (employment, NEET, motherhood) total-effect E-values between 1.3 and 1.55 suggest that an unmeasured confounder would need to increase the risk of both adolescent MHDs and the outcomes by at least 30–50% to fully explain away the observed associations. The decomposed direct and indirect effects exhibit smaller E-values, implying that modest confounding could attenuate individual pathways even if the total effects remain stable. For continuous outcomes, approximate E-values derived from standardised mean differences show broadly similar robustness levels, though somewhat lower for hours worked and STDs, and these should be interpreted with caution given their approximate nature. Overall, the sensitivity analysis supports the credibility of the main findings, particularly the elevated risk of NEET status and reduced employment among adolescents with MHDs, while acknowledging some vulnerability of the pathway effects to unmeasured bias.

Table S3_2 in the Supplementary material presents the sensitivity of our results to potential misclassification in adolescent MHDs. Across all specifications, the estimated total effect remains broadly stable, suggesting that our main findings are not driven by plausible levels of measurement error. The proportion mediated shows somewhat more variation across scenarios, as expected given that misclassification affects the relative contribution of the mediated pathway. However, even under conservative assumptions with low sensitivity, the qualitative conclusions remain unchanged: a non-negligible share of the total effect operates through the mediator. Our findings are robust to a wide range of assumptions about exposure misclassification, although the precise magnitude of the mediated share should be interpreted with some caution.

Lastly, Table S3_3 shows that looking at MHDs between ages 14-17 instead of 14-16 leads to similar conclusions. The estimated total effects are highly stable across all eight outcomes: every total effect retains its sign and statistical significance, and the magnitudes change only

marginally (for example, the effect on employment moves from −0.030 to −0.039 and the effect on log hourly wages from −0.060 to −0.066). The proportion mediated by dropping out shifts modestly, declining somewhat for labour-market outcomes (e.g., from 34.8% to 25.8% for employment) and rising slightly for fertility and risky-behaviour outcomes (e.g., from 48.3% to 55.5% for motherhood), consistent with later diagnoses falling closer in time to the dropout decision. The four-way decomposition components are likewise consistent in sign and significance across both specifications. If anything, the larger diagnosis sample improves precision, bringing a few previously marginal estimates (the total effect on hours worked and the proportion mediated for STDs and substance abuse) to conventional significance levels. Overall, these results confirm that our main findings are not sensitive to the precise age window used to define adolescent mental health disorders.

5.2 Heterogeneity by gender

We explore the heterogeneity of our main results by gender in Table S4_1 in the Supplementary material. The total effects are consistently larger for men across most outcomes, indicating that adolescent MHDs have a stronger overall impact on men's labour market trajectories. However, the proportion mediated through school dropout is generally higher for women, particularly for NEET status and wages, suggesting that educational pathways play a relatively more important role in shaping women's outcomes. For other outcomes, the mediated share is estimated less precisely, especially where total effects are small. Overall, the main conclusions hold for both genders, while the relative importance of the dropout channel differs across them.

6. Discussion

This paper investigated the mediating and interaction role of dropping out of school in the relationship between adolescent mental health disorders and life outcomes in early adulthood. Using a comprehensive dataset that links administrative records on education, health, and employment for half of the Hungarian school population, we tracked individuals from ages 14 to 22. We find that adolescent MHDs are associated with higher dropout and poorer early-adult outcomes.

School dropout emerges as a key mechanism linking adolescent mental health disorders to early adult disadvantage, though its mediating importance varies across life domains. Dropout explains roughly one-third of the total effect of adolescent MHDs on employment, NEET status, and earnings, highlighting education as a major pathway through which early mental health problems translate into socioeconomic inequality. For fertility outcomes, dropout accounts for

about half of the effect on early motherhood and one-quarter of the effect on abortion, while the interaction terms suggest overlapping rather than mutually reinforcing mechanisms. In contrast, for risky health behaviours, the mediated shares are small and mostly insignificant, implying that the association between adolescent MHDs and outcomes such as STDs or substance abuse is driven primarily by direct, non-educational pathways.

Overall, these findings indicate that education serves as a central transmission channel for the long-term social and economic consequences of adolescent mental health disorders, while behavioural and health risks operate more directly. Early mental health problems thus set in motion multiple forms of disadvantage, but the extent to which these effects are shaped by schooling varies: strong for labour-market and family outcomes, weaker for health behaviours.

The consistent association between MHDs and school dropout underscores the need to address adolescent mental health not only as a clinical issue, but also as an educational and social one. Prior research has documented the detrimental effects of poor mental health on academic performance and retention (Van Poortvliet, 2024). Adolescents facing mental health challenges often struggle with absenteeism, concentration problems, and disengagement from school (Kearney et al., 2023). As our study suggests, addressing mental health disorders early could serve as a crucial intervention point, potentially reducing dropout and mitigating downstream inequalities in employment and family formation.

Our results are not without limitations. First, the data capture only clinically diagnosed mental health disorders, and individuals from lower socioeconomic backgrounds are less likely to receive a diagnosis (Verhoog et al., 2024) while being more prone to dropout, implying that we may underestimate the link between mental health problems and educational outcomes. Second, the unconfoundedness assumptions our analysis requires are strong: we cannot exploit exogenous variation in either adolescent MHDs or school dropout, so the estimates should be read as indicative of potential pathways rather than definitive causal effects. Reverse causality is nonetheless unlikely, since MHDs are measured at ages 14–16, dropout around age 18, and outcomes at age 22; but unobserved factors such as childhood trauma could simultaneously influence mental health, dropout, and later outcomes. Our sensitivity analyses addressing both unmeasured confounding and exposure misclassification suggest the main results are reasonably stable across a range of plausible assumptions, though identifying the causal effects of mental illness remains methodologically challenging.

The findings have important implications for education, health, and social policy. By identifying school dropout as an important pathway linking adolescent mental health disorders to poorer early-adult outcomes, the results point to the value of coordinated interventions that address both mental health and school engagement. Early identification of mental health problems, improved access to adolescent mental health services, and targeted retention policies could mitigate later negative outcomes. Given the social and fiscal costs associated with NEET status, early motherhood, abortions, and health-risk behaviours, these findings underscore the value of integrated strategies that support vulnerable students in remaining in school while addressing their mental health needs. Although the estimates rely on observational data, they point to clear policy levers for reducing long-term inequalities and improving population well-being.

Declaration of generative AI in scientific writing. During the preparation of this work, the authors used Claude (Opus 4.8) to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Brännlund, A., Strandh, M., & Nilsson, K. (2017). Mental-health and educational achievement: The link between poor mental-health and upper secondary school completion and grades. *Journal of Mental Health*, *26*(4), 318–325. https://doi.org/10.1080/09638237.2017.1294739

Campolieti, M., Fang, T., & Gunderson, M. (2010). Labour Market Outcomes and Skill Acquisition of High-School Dropouts. *Journal of Labor Research*, *31*(1), 39–52.

Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective, Second Edition* (2nd edn). Chapman and Hall/CRC. https://doi.org/10.1201/9781420010138

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(1), 72–79. https://doi.org/10.1016/j.jaac.2018.07.896

Dahmann, S. C., & Schnitzlein, D. D. (2019). No evidence for a protective effect of education on mental health. *Social Science & Medicine*, *241*, 112584. https://doi.org/10.1016/j.socscimed.2019.112584

Du, B., Li, Y., Xia, Y., Wu, S., Wei, Y., Wang, Z., Wang, S., Guo, J., Zhu, Q., Shen, T., & Chen, Q. (2023). The mediation and interaction of the obesity index between moderate-vigorous recreational physical activity and hypertension. *PLOS ONE*, *18*, e0296333. https://doi.org/10.1371/journal.pone.0296333

Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry*, *59*(3), 225–231. https://doi.org/10.1001/archpsyc.59.3.225

Foulkes, L., & Andrews, J. L. (2023). Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis. *New Ideas in Psychology*, *69*, 1–6. https://doi.org/10.1016/j.newideapsych.2023.101010

Hijdra, R., Oude Groeniger, J., Burdorf, A., & Schuring, M. (2025). The effect of mental health problems on having a 'neither in employment nor in education or training' period and the mediating role of high school dropout: A register-based study with a 14-year follow-up. *Journal of Epidemiology and Community Health*, jech-2024-222197. https://doi.org/10.1136/jech-2024-222197

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10.1037/a0020761

Jeon, N., Albogami, Y., Jung, S.-Y., Bussing, R., & Winterstein, A. G. (2024). Comparing pregnancy and pregnancy outcome rates between adolescents with and without pre-existing mental disorders. *PloS One*, *19*(3), e0296425. https://doi.org/10.1371/journal.pone.0296425

Johar, M., & Truong, J. (2014). Direct and indirect effect of depression in adolescence on adult wages. *Applied Economics*, *46*(36), 4431–4444. https://doi.org/10.1080/00036846.2014.962227

Kapócs, G. (2017). The Health Policy in Child and Adolescent Psychiatry in Hungary – a review of recent developments • New Medicine 1/2017 • Czytelnia Medyczna BORGIS. *New Medicine*. https://www.czytelniamedyczna.pl/5795%2Cthe-health-policy-in-child-and-adolescent-psychiatry-in-hungary-a-review-of-rec.html

Kearney, C. A., Dupont, R., Fensken, M., & Gonzálvez, C. (2023). School attendance problems and absenteeism as early warning signals: Review and implications for health-based protocols and school-based practices. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1253595

Kézdi, G., & Békés, G. (2021). *Data Analysis for Business, Economics, and Policy*. Cambridge University Press.

Marcotte, D. E. (2013). High school dropout and teen childbearing. *Economics of Education Review*, *34*(C), 258–268.

Mensch, B. S., Grant, M. J., Soler-Hampejsek, E., Kelly, C. A., Chalasani, S., & Hewett, P. C. (2020). Does schooling protect sexual health? The association between three measures of education and STIs among adolescents in Malawi. *Population Studies*, *74*(2), 241–261. https://doi.org/10.1080/00324728.2019.1656282

Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017

Minh, A., McLeod, C. B., Reijneveld, S. A., Veldman, K., van Zon, S. K. R., & Bültmann, U. (2023). The role of low educational attainment on the pathway from adolescent internalizing and externalizing problems to early adult labour market disconnection in the Dutch TRAILS cohort. *SSM - Population Health*, *21*, 101300. https://doi.org/10.1016/j.ssmph.2022.101300

Philipson, A., Alaie, I., Ssegonja, R., Imberg, H., Copeland, W., Möller, M., Hagberg, L., & Jonsson, U. (2020). Adolescent depression and subsequent earnings across early to middle adulthood: A 25-year longitudinal cohort study. *Epidemiology and Psychiatric Sciences*, *29*, e123. https://doi.org/10.1017/S2045796020000360

Plenty, S., Magnusson, C., & Låftman, S. B. (2021). Internalising and externalising problems during adolescence and the subsequent likelihood of being Not in Employment, Education or Training (NEET) among males and females: The mediating role of school performance. *SSM - Population Health*, *15*, 100873. https://doi.org/10.1016/j.ssmph.2021.100873

Reingle Gonzalez, J. M., Salas-Wright, C. P., Connell, N. M., Jetelina, K. K., Clipper, S. J., & Businelle, M. S. (2016). The long-term effects of school dropout and GED attainment on substance use disorders. *Drug and Alcohol Dependence*, *158*, 60–66. https://doi.org/10.1016/j.drugalcdep.2015.11.002

Rodwell, L., Romaniuk, H., Nilsen, W., Carlin, J. B., Lee, K. J., & Patton, G. C. (2018). Adolescent mental health and behavioural predictors of being NEET: A prospective

study of young adults not in employment, education, or training. *Psychological Medicine*, *48*(5), 861–871. https://doi.org/10.1017/S0033291717002434

Sebők, A. (2019). *The Panel of Linked Administrative Data of CERS Databank: A KRTK Adatbank Kapcsolt Allamigazgatási Paneladatbázisa* (No. 2; Budapest Working Papers on the Labour Market). Institute of Economics, Centre for Economic and Regional Studies.

Shi, B., Choirat, C., Coull, B. A., VanderWeele, T. J., & Valeri, L. (2021). CMAverse: A Suite of Functions for Reproducible Causal Mediation Analyses. *Epidemiology (Cambridge, Mass.)*, *32*(5), e20–e22. https://doi.org/10.1097/EDE.0000000000001378

Shrier, L. A., Harris, S. K., & Beardslee, W. R. (2002). Temporal Associations Between Depressive Symptoms and Self-reported Sexually Transmitted Disease Among Adolescents. *Archives of Pediatrics & Adolescent Medicine*, *156*(6), 599. https://doi.org/10.1001/archpedi.156.6.599

Sibley, M. H., Rohde, L. A., Swanson, J. M., Hechtman, L. T., Molina, B. S. G., Mitchell, J. T., Arnold, L. E., Caye, A., Kennedy, T. M., Roy, A., Stehli, A., & for the Multimodal Treatment Study of Children with ADHD (MTA) Cooperative Group. (2018). Late-Onset ADHD Reconsidered With Comprehensive Repeated Assessments Between Ages 10 and 25. *American Journal of Psychiatry*, *175*(2), 140–149. https://doi.org/10.1176/appi.ajp.2017.17030298

Tong, Y., Wang, S., Cao, L., Zhu, D., Wang, F., Xie, F., Zhang, X., Wang, G., & Su, P. (2023). School dropouts related to mental disorders: A systematic review and meta-analysis. *Asian Journal of Psychiatry*, *85*, 103622. https://doi.org/10.1016/j.ajp.2023.103622

UNICEF. (2021). *The State Of The World 'S Children 2021*. https://www.unicef.org/media/108121/file/SOWC-2021-Europe-regional-brief.pdf

Van Poortvliet, M. (2024). Child mental health and educational attainment: Longitudinal evidence from the UK. *SSM - Mental Health*, *5*, 100294. https://doi.org/10.1016/j.ssmmh.2023.100294

VanderWeele, T. J. (2014). A Unification of Mediation and Interaction: A 4-Way Decomposition. *Epidemiology*, *25*(5), 749–761. https://doi.org/10.1097/EDE.0000000000000121

VanderWeele, T. J., & Ding, P. (2017). Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of Internal Medicine*, *167*(4), 268–274. https://doi.org/10.7326/M16-2607

Veldman, K., Reijneveld, S. A., Hviid Andersen, J., Nøhr Winding, T., Labriola, M., Lund, T., & Bültmann, U. (2022). The timing and duration of depressive symptoms from adolescence to young adulthood and young adults' NEET status: The role of educational attainment. *Social Psychiatry and Psychiatric Epidemiology*, *57*(1), 83–93. https://doi.org/10.1007/s00127-021-02142-5

Verhoog, S., Eijgermans, D. G. M., Fang, Y., Bramer, W. M., Raat, H., & Jansen, W. (2024). Contextual determinants associated with children's and adolescents' mental health care utilization: A systematic review. *European Child & Adolescent Psychiatry*, *33*(7), 2051–2065. https://doi.org/10.1007/s00787-022-02077-5

von Simson, K., Brekke, I., & Hardoy, I. (2022). The Impact of Mental Health Problems in Adolescence on Educational Attainment. *Scandinavian Journal of Educational Research*, *66*(2), 306–320. https://doi.org/10.1080/00313831.2020.1869077

World Health Organization. (2022). *Mental disorders*. https://www.who.int/news-room/fact-sheets/detail/mental-disorders

World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems : tenth revision*. World Health Organization. https://apps.who.int/iris/handle/10665/42980

Appendix A

Table A1: Mental health disorder categories used in this paper

Category	Description	ICD codes
Anxiety and stress-related disorders	Includes anxiety disorders, obsessive-compulsive disorder, and stress-related and dissociative disorders.	F40–F44
Psychotic disorders	Severe disorders involving impaired perception of reality, including schizophrenia and related conditions.	F20–F29
Mood (affective) disorders	Includes depressive and bipolar disorders.	F30–F39
Personality disorders	Persistent patterns of behaviour and inner experience deviating from social norms.	F60–F69
Behavioural and neurodevelopmental disorders	Includes eating disorders, sleep disorders, physiological syndromes, and childhood-onset behavioural and developmental disorders (e.g. ADHD).	F50–F59, F90–F99

Source: World Health Organization (2004)

Table A2: Descriptive statistics

	Obs	Mean	SD
Female	51801	0.49	0.50
Dropping out	51801	0.17	0.38
Dropout age	8970	18.36	1.16

MHD between ages 14 and 16	51801	0.05	0.23
Time spent in kindergarten	51801	6.34	0.88
Time spent in kindergarten missing	51801	0.00	0.07
Financial problems in the family	51801	0.56	0.50
Financial problems in the family missing	51801	0.00	0.07
Lives with mother only	51801	0.15	0.35
Lives with father only	51801	0.02	0.14
Lives with mother and father	51801	0.82	0.39
Lives with parent missing	51801	0.05	0.22
Family size	51801	4.36	1.45
Family size missing	51801	0.01	0.11
Mother's level of education	51801	4.43	1.42
Father's level of education	51801	4.27	1.31
Background questionnaire missing	51801	0.11	0.31
Math test score: Q1	51801	0.12	0.33
Math test score: Q2	51801	0.18	0.38
Math test score: Q3	51801	0.18	0.38
Math test score: Q4	51801	0.18	0.38
Math test score: Q5	51801	0.18	0.38
Math test score: missing	51801	0.18	0.38
Reading test score: Q1	51801	0.12	0.33
Reading test score: Q2	51801	0.18	0.38
Reading test score: Q3	51801	0.18	0.38
Reading test score: Q4	51801	0.18	0.38
Reading test score: Q5	51801	0.18	0.38
Reading test score: missing	51801	0.18	0.38
Outcome variables			
Works at age 22	51801	0.46	0.43
NEET at age 22	51801	0.29	0.39
Hours worked at age 22, if employed	28218	35.82	8.36
Log hourly wage at age 22, if employed	28009	6.88	0.41
Cumulative number of STD diagnoses by age 22	51801	0.07	0.49
Monthly prevalence of substance abuse diagnoses at age 22	51801	0.01	0.08
Cumulative number of abortions by age 22	25251	0.16	0.54
Probability of motherhood at age 22	25251	0.15	0.36

Source: Admin3. Sample of those born between June 1994 and May 1995.

Table A3: The association between adolescent MHDs and dropping out – probit models

Dependent variables	Independent variables	(1)	(2)	(3)	(4)
		Coefficient	Marginal effect (dy/dx)	Coefficient	Marginal effect (dy/dx)
Dropping out	MHD[(a)]	0.502*** [0.026]	0.127*** [0.006]	0.328*** [0.031]	.058*** [0.005]
Controls		No	No	Yes	Yes
Sample		All	All	All	All

Source: Admin3. Sample of those born between June 1994 and May 1995. No. of obs: 51,801, Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *. Probit models according to Equation (1). The full output table is reported in Tables S2_1 in the Supplementary material. (a) at least one diagnosis between ages 14 and 16.

Table A4: Regression results for the exposure-outcome links

Dependent variables	Independent variables	(1)	(2)	(3)
Employed	MHD[(a)]	-0.033***	-0.019**	-0.019**
		[0.008]	[0.008]	[0.010]
	Dropout		-0.179***	-0.180***
			[0.006]	[0.006]
	Dropout * MHD			0.001
				[0.016]
NEET	MHD	0.065***	0.039***	0.038***
		[0.008]	[0.007]	[0.008]
	Dropout		0.324***	0.324***
			[0.005]	[0.006]
	Dropout * MHD			0.003
				[0.016]
Hours worked[(b)]	MHD	-0.034	-0.026	0.002
		[0.028]	[0.028]	[0.031]
	Dropout		-0.098***	-0.089***
			[0.018]	[0.019]
	Dropout * MHD			-0.116*
				[0.070]
Log hourly wage[(b)]	MHD	-0.070***	-0.054***	-0.037***
		[0.012]	[0.012]	[0.012]
	Dropout		-0.198***	-0.192***
			[0.008]	[0.008]
	Dropout * MHD			-0.071**
				[0.031]
Motherhood[(c)]	MHD	0.045***	0.019**	0.031***
		[0.010]	[0.010]	[0.009]
	Dropout		0.359***	0.363***
			[0.009]	[0.009]
	Dropout * MHD			-0.044*
				[0.027]
Abortion[(d)]	MHD	0.145***	0.119***	0.055***
		[0.021]	[0.020]	[0.015]
	Dropout		0.364***	0.342***
			[0.017]	[0.017]
	Dropout * MHD			0.237***
				[0.063]
STD[(d)]	MHD	0.026**	0.025**	0.008
		[0.011]	[0.011]	[0.011]
	Dropout		0.013*	0.008
			[0.007]	[0.007]
	Dropout * MHD			0.058**
				[0.029]

Substance abuse[e]	MHD	0.006***	0.005**	0.006**
		[0.002]	[0.002]	[0.002]
	Dropout		0.013***	0.013***
			[0.002]	[0.002]
	Dropout * MHD			-0.002
				[0.006]
Controls		Yes	Yes	Yes
Sample		All	All	All

Source: Admin3. Sample of those born between June 1994 and May 1995. No. of obs: 51,801, Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *. The full output tables are reported in Tables S2_2 to S2_9 in the Supplementary material. (a) at least one diagnosis between ages 14 and 16; (b) conditional on employment; (c) probability by age 22; (d) cumulative number by age 22; (e) average monthly prevalence at age 22.

Supplementary material for

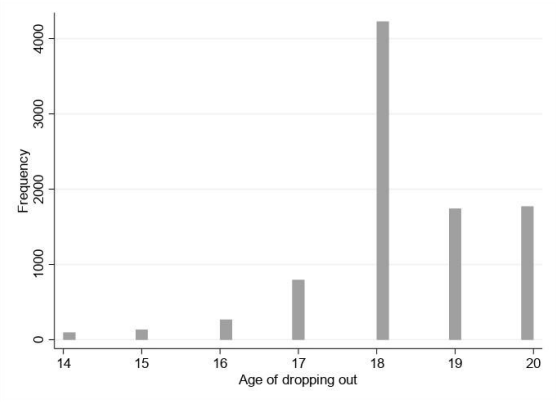
Leaving school early: a mediation analysis linking adolescent mental health disorders to early adult outcomes

Anna Adamecz[5], John Jerrim[6] and Ágnes Szabó-Morvai[7]

26 June 2026

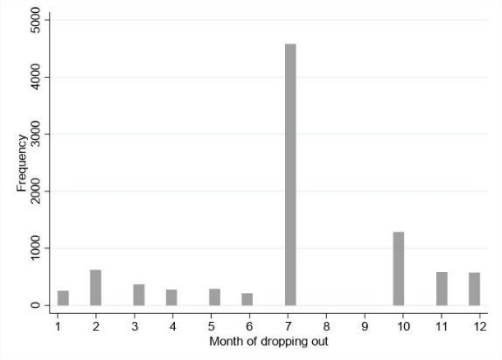
S1. Descriptive statistics and supporting information

Figure S1_1: The distribution of dropout age



Source: Admin 3. Sample of those born between June 1994 and May 1995. No. of dropouts: 8,970.

Figure S1_2: The distribution of dropout month



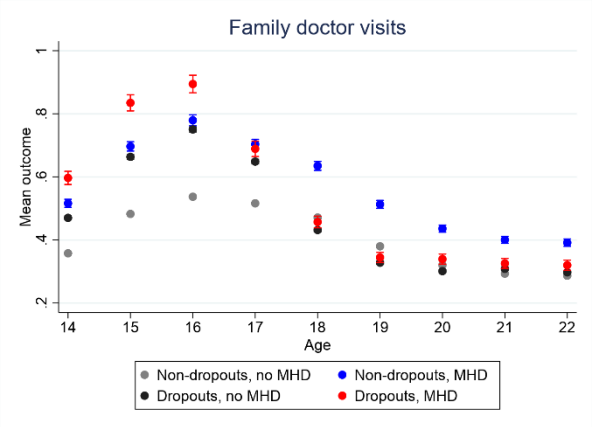
Source: Admin 3. Sample of those born between June 1994 and May 1995. No. of dropouts: 8,970.

[5] Corresponding author. UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA and Institute of Economics, HUN-REN Centre for Economic and Regional Studies, Toth Kalman u. 4, 1097 Budapest. E-mail: a.adamecz@ucl.ac.uk

[6] UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA. E-mail: j.jerrim@ucl.ac.uk

[7] Institute of Economics, Centre for Economic and Regional Studies, Toth Kalman u. 4, 1097 Budapest and University of Debrecen, Faculty of Economics and Business, Boszormenyi ut 138. 4032 Debrecen. E-mail: szabomorvai.agnes@krtk.hu

Figure S1_3: The number of family doctor visits by age, dropout and MHD status



Source: Admin 3. Sample of those born between June 1994 and May 1995. No. of dropouts: 8,970.

S2. Verification of the mediation pathways

The exposure-mediator link

Table S2_1: Dropout probability and mental health disorders – probit models

	(1) Dropping out	(2) Dropping out
MHD (at least one diagnosis between ages 14 and 16)	0.502***	0.328***
	[0.026]	[0.031]
Female		-0.035**
		[0.016]
Time spent in kindergarten		-0.045***
		[0.009]
Missing data on kindergarten		0.237***
		[0.087]
Financial problems in the family		0.230***
		[0.021]
Financial problems in the family (missing)		-0.048
		[0.099]
Lives with mother only		-0.301***
		[0.039]
Lives with father only		-0.217***
		[0.025]
Lives with mother (missing)		0.227*
		[0.128]
Lives with father (missing)		0.352***
		[0.127]
Family size		0.052***
		[0.013]
Family size (missing)		0.194***
		[0.071]
Mother's level of education		-0.180***
		[0.008]
Father's level of education		-0.144***
		[0.009]
Math test score: Q1		0.574***
		[0.200]
Math test score: Q2		0.465***
		[0.041]
Math test score: Q3		0.259***
		[0.040]
Math test score: Q4		0.137***
		[0.040]
Math test score: Q5		0.107***
		[0.041]
Reading test score: Q1		0.957***
		[0.201]
Reading test score: Q2		0.862***
		[0.047]
Reading test score: Q3		0.650***
		[0.046]

Reading test score: Q4		0.442***
		[0.045]
Reading test score: Q5		0.240***
		[0.046]
Constant	-0.974***	-0.382***
	[0.007]	[0.104]
Observations	51,801	51,801

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *. Probit models according to Equation (1).

The exposure-outcome links

Table S2_2: Employment and mental health disorders (the exposure-outcome link for employment)

	Outcome: Employed					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	-0.056***	-0.033***	-0.019**	-0.019**	-0.021**	-0.033**
	[0.008]	[0.008]	[0.008]	[0.010]	[0.010]	[0.013]
Dropout			-0.179***	-0.180***		
			[0.006]	[0.006]		
Dropout * MHD				0.001		
				[0.016]		
Time in kindergarten		0.015***	0.013***	0.013***	0.006**	0.025***
		[0.002]	[0.002]	[0.002]	[0.003]	[0.004]
Time in kindergarten (missing)		-0.039	-0.022	-0.022	-0.004	-0.069
		[0.028]	[0.028]	[0.028]	[0.034]	[0.045]
Financial problems in the family		-0.025***	-0.018***	-0.018***	-0.016***	-0.046***
		[0.004]	[0.004]	[0.004]	[0.005]	[0.014]
Financial problems in the family (missing)		-0.021	-0.017	-0.017	-0.030	-0.036
		[0.030]	[0.030]	[0.030]	[0.037]	[0.046]
Lives with mother		0.048***	0.033*	0.033*	0.032	0.062**
		[0.017]	[0.017]	[0.017]	[0.022]	[0.026]
Lives with father		0.063***	0.053***	0.053***	0.052**	0.056
		[0.021]	[0.021]	[0.021]	[0.026]	[0.035]
Lives with mother and father		0.060***	0.037**	0.037**	0.037*	0.052**
		[0.017]	[0.017]	[0.017]	[0.021]	[0.024]
Lives with mother (missing)		-0.054***	-0.029***	-0.029***	-0.022**	-0.014
		[0.009]	[0.008]	[0.008]	[0.010]	[0.015]
Family size		-0.010***	-0.007***	-0.007***	-0.003	-0.008**
		[0.002]	[0.002]	[0.002]	[0.002]	[0.003]
Family size (missing)		-0.041**	-0.035**	-0.035**	-0.031	-0.070**
		[0.017]	[0.017]	[0.017]	[0.019]	[0.030]
Female		-0.076***	-0.077***	-0.077***	-0.050***	-0.183***
		[0.004]	[0.004]	[0.004]	[0.004]	[0.008]

Mother's level of education		-0.012***	-0.019***	-0.019***	-0.025***	0.011**
		[0.002]	[0.002]	[0.002]	[0.002]	[0.004]
Father's level of education		-0.020***	-0.025***	-0.025***	-0.031***	0.008
		[0.002]	[0.002]	[0.002]	[0.002]	[0.005]
Math test score: Q1					-0.059	0.037
					[0.081]	[0.059]
Math test score: Q2		0.063	0.051	0.051		0.023
		[0.065]	[0.063]	[0.063]		[0.025]
Math test score: Q3		0.089	0.070	0.070	0.012	0.051**
		[0.065]	[0.063]	[0.063]	[0.007]	[0.025]
Math test score: Q4		0.074	0.052	0.052	-0.003	0.029
		[0.065]	[0.063]	[0.063]	[0.008]	[0.026]
Math test score: Q5		0.045	0.024	0.024	-0.030***	0.038
		[0.065]	[0.063]	[0.063]	[0.008]	[0.027]
Math test score: Q6		-0.025	-0.044	-0.044	-0.091***	
		[0.065]	[0.063]	[0.063]	[0.009]	
Reading test score: Q1					0.089	-0.095
					[0.081]	[0.061]
Reading test score: Q2		0.080	0.067	0.067	0.152***	0.011
		[0.065]	[0.063]	[0.063]	[0.009]	[0.029]
Reading test score: Q3		0.098	0.074	0.074	0.154***	0.017
		[0.065]	[0.063]	[0.063]	[0.008]	[0.029]
Reading test score: Q4		0.072	0.045	0.045	0.119***	0.020
		[0.065]	[0.063]	[0.063]	[0.007]	[0.029]
Reading test score: Q5		0.024	-0.006	-0.006	0.070***	-0.010
		[0.065]	[0.063]	[0.063]	[0.006]	[0.031]
Reading test score: Q6		-0.047	-0.075	-0.075		
		[0.065]	[0.063]	[0.063]		
Background questionnaire missing		-0.120***	-0.075***	-0.075***	-0.087***	-0.144***
		[0.008]	[0.008]	[0.008]	[0.010]	[0.016]
Constant	0.467***	0.478***	0.610***	0.610***	0.652***	0.268***
	[0.002]	[0.026]	[0.025]	[0.025]	[0.031]	[0.052]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts

Observations	51,801	51,801	51,801	51,801	42,831	8,970
R-squared	0.001	0.075	0.093	0.093	0.087	0.115

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

Table S2_3: NEET and mental health problems (the exposure-outcome link for NEET)

	Outcome: NEET					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	0.124***	0.065***	0.039***	0.038***	0.043***	0.029**
	[0.008]	[0.008]	[0.007]	[0.008]	[0.008]	[0.013]
Dropout			0.324***	0.324***		
			[0.005]	[0.006]		
Dropout * MHD				0.003		
				[0.016]		
Time spent in kindergarten		-0.018***	-0.014***	-0.014***	-0.010***	-0.025***
		[0.002]	[0.002]	[0.002]	[0.002]	[0.004]
Time spent in kindergarten (missing)		0.081***	0.050**	0.050**	0.042	0.079*
		[0.027]	[0.025]	[0.025]	[0.030]	[0.046]
Financial problems in the family		0.038***	0.026***	0.026***	0.025***	0.044***
		[0.004]	[0.004]	[0.004]	[0.004]	[0.014]
Financial problems in the family (missing)		-0.010	-0.018	-0.018	-0.044	0.043
		[0.029]	[0.026]	[0.026]	[0.030]	[0.046]
Lives with mother		-0.058***	-0.032**	-0.032**	-0.032*	-0.047*
		[0.017]	[0.016]	[0.016]	[0.019]	[0.027]
Lives with father		-0.058***	-0.040**	-0.040**	-0.040*	-0.041
		[0.020]	[0.019]	[0.019]	[0.022]	[0.035]
Lives with mother and father		-0.094***	-0.053***	-0.053***	-0.059***	-0.035
		[0.016]	[0.015]	[0.015]	[0.019]	[0.025]
Lives with mother (missing)		0.077***	0.032***	0.032***	0.040***	0.014
		[0.008]	[0.008]	[0.008]	[0.009]	[0.015]
Family size		0.012***	0.006***	0.006***	0.005**	0.008**
		[0.002]	[0.002]	[0.002]	[0.002]	[0.003]
Family size (missing)		0.025	0.014	0.014	-0.000	0.070**
		[0.017]	[0.015]	[0.015]	[0.018]	[0.030]
Female		0.049***	0.050***	0.050***	0.022***	0.183***
		[0.003]	[0.003]	[0.003]	[0.003]	[0.008]
Mother's level of education		-0.028***	-0.015***	-0.015***	-0.016***	-0.012***
		[0.002]	[0.002]	[0.002]	[0.002]	[0.004]
Father's level of education		-0.018***	-0.010***	-0.010***	-0.009***	-0.009*

		[0.002]	[0.002]	[0.002]	[0.002]	[0.005]
Math test score: Q1					-0.027	-0.042
					[0.069]	[0.061]
Math test score: Q2		-0.002	0.022	0.022		-0.027
		[0.060]	[0.055]	[0.055]		[0.026]
Math test score: Q3		-0.030	0.005	0.005	-0.015**	-0.056**
		[0.060]	[0.055]	[0.055]	[0.006]	[0.026]
Math test score: Q4		-0.041	-0.002	-0.002	-0.028***	-0.033
		[0.060]	[0.055]	[0.055]	[0.006]	[0.026]
Math test score: Q5		-0.052	-0.013	-0.013	-0.040***	-0.044
		[0.060]	[0.055]	[0.055]	[0.007]	[0.027]
Math test score: Q6		-0.049	-0.013	-0.013	-0.045***	
		[0.060]	[0.055]	[0.055]	[0.007]	
Reading test score: Q1					0.108	0.093
					[0.069]	[0.063]
Reading test score: Q2		-0.088	-0.065	-0.065	0.069***	-0.009
		[0.060]	[0.055]	[0.055]	[0.007]	[0.029]
Reading test score: Q3		-0.124**	-0.081	-0.081	0.045***	-0.012
		[0.060]	[0.055]	[0.055]	[0.006]	[0.029]
Reading test score: Q4		-0.146**	-0.098*	-0.098*	0.028***	-0.019
		[0.060]	[0.055]	[0.055]	[0.005]	[0.029]
Reading test score: Q5		-0.159***	-0.105*	-0.105*	0.021***	0.009
		[0.060]	[0.055]	[0.055]	[0.005]	[0.031]
Reading test score: Q6		-0.181***	-0.130**	-0.130**		
		[0.061]	[0.055]	[0.055]		
Background questionnaire missing		0.186***	0.106***	0.106***	0.075***	0.147***
		[0.008]	[0.007]	[0.007]	[0.009]	[0.016]
Constant	0.282***	0.718***	0.479***	0.479***	0.384***	0.719***
	[0.002]	[0.024]	[0.023]	[0.023]	[0.027]	[0.052]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts
Observations	51,801	51,801	51,801	51,801	42,831	8,970
R-squared	0.005	0.140	0.212	0.212	0.042	0.113

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

Table S2_4: Hours worked and mental health problems (the exposure-outcome link for hours worked)

	Hours worked (conditional on employment)					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	-0.059**	-0.034	-0.026	0.002	0.007	-0.127**
	[0.029]	[0.028]	[0.028]	[0.031]	[0.031]	[0.064]
Dropout			-0.098***	-0.089***		
			[0.018]	[0.019]		
Dropout * MHD				-0.116*		
				[0.070]		
Time spent in kindergarten		-0.000	-0.001	-0.000	-0.004	0.013
		[0.007]	[0.007]	[0.007]	[0.008]	[0.018]
Time spent in kindergarten (missing)		-0.166	-0.158	-0.159	-0.080	-0.362*
		[0.104]	[0.103]	[0.103]	[0.117]	[0.208]
Financial problems in the family		0.021	0.025*	0.024*	0.021	0.066
		[0.013]	[0.013]	[0.013]	[0.014]	[0.044]
Financial problems in the family (missing)		-0.043	-0.039	-0.036	0.002	-0.140
		[0.115]	[0.114]	[0.114]	[0.133]	[0.219]
Lives with mother		0.027	0.021	0.020	0.032	0.005
		[0.054]	[0.055]	[0.055]	[0.066]	[0.097]
Lives with father		-0.008	-0.013	-0.014	0.019	-0.150
		[0.065]	[0.065]	[0.065]	[0.076]	[0.132]
Lives with mother and father		0.074	0.064	0.063	0.080	-0.003
		[0.053]	[0.053]	[0.053]	[0.064]	[0.092]
Lives with mother (missing)		-0.142***	-0.129***	-0.130***	-0.144***	-0.059
		[0.029]	[0.030]	[0.030]	[0.035]	[0.057]
Family size		-0.005	-0.004	-0.004	-0.003	-0.004
		[0.004]	[0.004]	[0.004]	[0.005]	[0.008]
Family size (missing)		-0.124*	-0.122*	-0.123*	-0.118	-0.114
		[0.067]	[0.066]	[0.066]	[0.075]	[0.143]
Female		-0.182***	-0.186***	-0.186***	-0.179***	-0.227***
		[0.012]	[0.012]	[0.012]	[0.013]	[0.037]
Mother's level of education		-0.062***	-0.066***	-0.065***	-0.066***	-0.065***
		[0.005]	[0.005]	[0.005]	[0.006]	[0.015]
Father's level of education		-0.037***	-0.039***	-0.039***	-0.041***	-0.025
		[0.006]	[0.006]	[0.006]	[0.007]	[0.016]
Math test score: Q1					0.358*	0.272***
					[0.187]	[0.099]
Math test score: Q2		-0.262	-0.273	-0.274	0.086***	-0.044
		[0.170]	[0.169]	[0.169]	[0.027]	[0.070]
Math test score: Q3		-0.262	-0.275	-0.276	0.086***	-0.064
		[0.170]	[0.169]	[0.169]	[0.026]	[0.067]
Math test score: Q4		-0.256	-0.270	-0.270	0.099***	-0.137*
		[0.170]	[0.169]	[0.169]	[0.025]	[0.072]
Math test score: Q5		-0.268	-0.282*	-0.283*	0.075***	
		[0.171]	[0.169]	[0.169]	[0.025]	
Math test score: Q6		-0.350**	-0.364**	-0.365**		-0.128
		[0.171]	[0.170]	[0.170]		[0.090]
Reading test score: Q1		-0.345**	-0.344**	-0.344**	-0.369**	
		[0.171]	[0.169]	[0.170]	[0.187]	

Reading test score: Q2						0.366***
						[0.082]
Reading test score: Q3		0.006	0.002	0.002	-0.007	0.386***
		[0.018]	[0.018]	[0.018]	[0.019]	[0.096]
Reading test score: Q4		-0.015	-0.020	-0.021	-0.034	0.392***
		[0.020]	[0.020]	[0.020]	[0.021]	[0.100]
Reading test score: Q5		-0.058***	-0.065***	-0.065***	-0.081***	0.398***
		[0.022]	[0.022]	[0.022]	[0.023]	[0.112]
Reading test score: Q6		-0.187***	-0.193***	-0.193***	-0.211***	0.376***
		[0.027]	[0.027]	[0.027]	[0.028]	[0.131]
Background questionnaire missing		-0.087***	-0.067**	-0.067**	-0.117***	0.024
		[0.030]	[0.030]	[0.030]	[0.036]	[0.063]
Constant	0.003	0.783***	0.839***	0.839***	0.504***	0.016
	[0.006]	[0.186]	[0.186]	[0.186]	[0.093]	[0.197]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts
Observations	28,218	28,218	28,218	28,218	24,354	3,864
R-squared	0.000	0.036	0.037	0.037	0.040	0.029

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

Table S2_5: Log hourly wage and mental health problems (the exposure-outcome link for log hourly wage)

	Log hourly wage (conditional on employment)					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	-0.102***	-0.070***	-0.054***	-0.037***	-0.039***	-0.103***
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.029]
Dropout			-0.198***	-0.192***		
			[0.008]	[0.008]		
Dropout * MHD				-0.071**		
				[0.031]		
Time spent in kindergarten		-0.008**	-0.009***	-0.009***	-0.012***	0.010
		[0.003]	[0.003]	[0.003]	[0.003]	[0.009]
Time spent in kindergarten (missing)		-0.084**	-0.068*	-0.069*	-0.039	-0.188
		[0.042]	[0.040]	[0.040]	[0.036]	[0.117]
Financial problems in the family		-0.028***	-0.021***	-0.021***	-0.022***	-0.031
		[0.005]	[0.005]	[0.005]	[0.005]	[0.019]
Financial problems in the family (missing)		0.052	0.058	0.059	0.124***	-0.117
		[0.054]	[0.051]	[0.050]	[0.047]	[0.129]
Lives with mother		0.010	-0.003	-0.003	-0.009	0.013
		[0.023]	[0.022]	[0.022]	[0.024]	[0.050]
Lives with father		0.036	0.027	0.026	0.014	0.077
		[0.026]	[0.026]	[0.026]	[0.028]	[0.061]
Lives with mother and father		0.046**	0.025	0.024	0.017	0.060
		[0.022]	[0.021]	[0.021]	[0.023]	[0.047]
Lives with mother (missing)		-0.053***	-0.028**	-0.028**	-0.023*	-0.055*
		[0.012]	[0.012]	[0.012]	[0.012]	[0.029]
Family size		-0.007***	-0.004**	-0.004**	-0.003*	-0.006

		[0.002]	[0.002]	[0.002]	[0.002]	[0.007]
Family size (missing)		-0.059**	-0.054**	-0.055**	-0.053**	-0.057
		[0.025]	[0.024]	[0.024]	[0.026]	[0.061]
Female		-0.068***	-0.077***	-0.077***	-0.079***	-0.056***
		[0.005]	[0.005]	[0.005]	[0.005]	[0.016]
Mother's level of education		0.013***	0.007***	0.007***	0.007***	0.013*
		[0.002]	[0.002]	[0.002]	[0.002]	[0.007]
Father's level of education		0.012***	0.008***	0.008***	0.008***	0.008
		[0.002]	[0.002]	[0.002]	[0.003]	[0.008]
Math test score: Q1		-0.072	-0.044	-0.044		-0.216
		[0.098]	[0.096]	[0.096]		[0.148]
Math test score: Q2		-0.064***	-0.059***	-0.059***	-0.041	-0.033
		[0.011]	[0.011]	[0.011]	[0.103]	[0.039]
Math test score: Q3		-0.036***	-0.034***	-0.034***	-0.018	0.005
		[0.010]	[0.010]	[0.010]	[0.103]	[0.038]
Math test score: Q4		-0.020**	-0.020**	-0.020**	-0.002	0.009
		[0.010]	[0.010]	[0.010]	[0.103]	[0.040]
Math test score: Q5		-0.020**	-0.021**	-0.021**	-0.001	-0.017
		[0.009]	[0.009]	[0.009]	[0.103]	[0.041]
Math test score: Q6					0.020	
					[0.103]	
Reading test score: Q1						0.057
						[0.150]
Reading test score: Q2		0.039	0.039	0.039	0.044	-0.083**
		[0.098]	[0.096]	[0.096]	[0.103]	[0.038]
Reading test score: Q3		0.072	0.063	0.063	0.070	-0.057
		[0.098]	[0.096]	[0.096]	[0.103]	[0.038]
Reading test score: Q4		0.079	0.068	0.068	0.077	-0.070*
		[0.098]	[0.096]	[0.096]	[0.103]	[0.038]
Reading test score: Q5		0.099	0.086	0.086	0.095	-0.047
		[0.098]	[0.096]	[0.096]	[0.103]	[0.042]
Reading test score: Q6		0.110	0.097	0.098	0.105	
		[0.098]	[0.096]	[0.096]	[0.103]	
Background questionnaire missing		-0.103***	-0.061***	-0.061***	-0.030**	-0.150***
		[0.012]	[0.012]	[0.012]	[0.013]	[0.030]
Constant	6.880***	6.832***	6.919***	6.919***	6.919***	6.689***
	[0.002]	[0.103]	[0.100]	[0.100]	[0.034]	[0.093]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts
Observations	28,009	28,009	28,009	28,009	24,174	3,835
R-squared	0.003	0.049	0.072	0.073	0.028	0.043

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

Table S2_6: The probability of motherhood and mental health problems (the exposure-outcome link for the probability of motherhood

Table S2_7: Abortions and mental health problems (the exposure-outcome link for abortions)

	Cumulative number of abortions by age 22					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	0.186***	0.145***	0.119***	0.055***	0.056***	0.297***
	[0.022]	[0.021]	[0.020]	[0.015]	[0.015]	[0.061]
Dropout			0.364***	0.342***		
			[0.017]	[0.017]		
Dropout * MHD				0.237***		
				[0.063]		
Time spent in kindergarten		-0.006	-0.001	-0.001	0.002	-0.010
		[0.005]	[0.005]	[0.005]	[0.003]	[0.017]
Time spent in kindergarten (missing)		0.116	0.065	0.065	0.061	0.141
		[0.071]	[0.070]	[0.070]	[0.042]	[0.187]
Financial problems in the family		0.042***	0.027***	0.027***	0.018***	0.143***
		[0.006]	[0.006]	[0.006]	[0.005]	[0.041]
Financial problems in the family (missing)		-0.001	-0.009	-0.011	-0.043	0.059
		[0.070]	[0.065]	[0.066]	[0.040]	[0.181]
Lives with mother		-0.068*	-0.041	-0.037	-0.045	-0.070
		[0.039]	[0.038]	[0.038]	[0.031]	[0.097]
Lives with father		-0.029	-0.004	-0.003	-0.029	0.059
		[0.047]	[0.045]	[0.045]	[0.037]	[0.138]
Lives with mother and father		-0.099**	-0.057	-0.055	-0.073**	-0.034
		[0.039]	[0.038]	[0.037]	[0.031]	[0.093]
Lives with mother (missing)		0.112***	0.060***	0.060***	0.033**	0.127**
		[0.022]	[0.022]	[0.021]	[0.016]	[0.061]
Family size		0.014***	0.008**	0.009**	0.008***	0.009
		[0.003]	[0.004]	[0.004]	[0.002]	[0.010]
Family size (missing)		0.047	0.035	0.034	-0.001	0.196
		[0.043]	[0.041]	[0.041]	[0.026]	[0.135]
Female						
Mother's level of education		-0.030***	-0.016***	-0.016***	-0.015***	-0.022
		[0.003]	[0.003]	[0.003]	[0.002]	[0.015]
Father's level of education		-0.029***	-0.019***	-0.020***	-0.015***	-0.047***
		[0.003]	[0.003]	[0.003]	[0.002]	[0.017]
Math test score: Q1					-0.024	-0.212
					[0.057]	[0.298]
Math test score: Q2		0.108	0.133	0.138	0.022**	0.158
		[0.109]	[0.097]	[0.095]	[0.010]	[0.096]
Math test score: Q3		0.066	0.104	0.110	0.013	0.053
		[0.109]	[0.097]	[0.095]	[0.008]	[0.096]
Math test score: Q4		0.049	0.091	0.095	0.000	0.002
		[0.109]	[0.097]	[0.095]	[0.006]	[0.093]
Math test score: Q5		0.055	0.094	0.099	-0.006	0.122
		[0.109]	[0.097]	[0.095]	[0.005]	[0.098]
Math test score: Q6		0.062	0.096	0.101		

		[0.109]	[0.097]	[0.095]		
Reading test score: Q1		0.119	0.104	0.108	0.022	0.267
		[0.109]	[0.097]	[0.096]	[0.058]	[0.305]
Reading test score: Q2						-0.029
						[0.119]
Reading test score: Q3		-0.026*	0.003	0.002	-0.025**	0.085
		[0.015]	[0.015]	[0.015]	[0.012]	[0.115]
Reading test score: Q4		-0.065***	-0.025*	-0.026*	-0.044***	0.010
		[0.015]	[0.014]	[0.014]	[0.012]	[0.113]
Reading test score: Q5		-0.080***	-0.034**	-0.036**	-0.058***	0.005
		[0.015]	[0.015]	[0.015]	[0.012]	[0.118]
Reading test score: Q6		-0.087***	-0.043***	-0.044***	-0.068***	
		[0.016]	[0.015]	[0.015]	[0.012]	
Background questionnaire missing		0.170***	0.072***	0.071***	0.049***	0.165***
		[0.021]	[0.020]	[0.020]	[0.016]	[0.063]
Constant	0.151***	0.411***	0.164	0.162	0.268***	0.539***
	[0.003]	[0.121]	[0.110]	[0.108]	[0.040]	[0.191]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts
Observations	25,251	25,251	25,251	25,251	21,238	4,013
R-squared	0.007	0.078	0.123	0.125	0.036	0.026

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

Table S2_8: STD's and mental health problems (the exposure-outcome link for STDs)

	The cumulative number of STDs by age 22					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	0.029**	0.026**	0.025**	0.008	0.008	0.066**
	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.027]
Dropout			0.013*	0.008		
			[0.007]	[0.007]		
Dropout * MHD				0.058**		
				[0.029]		
Time spent in kindergarten		0.002	0.002	0.002	0.002	0.001
		[0.002]	[0.002]	[0.002]	[0.002]	[0.007]
Time spent in kindergarten (missing)		-0.051***	-0.053***	-0.052***	-0.039**	-0.085***
		[0.011]	[0.011]	[0.011]	[0.015]	[0.013]
Financial problems in the family		-0.003	-0.003	-0.003	0.000	-0.024
		[0.005]	[0.005]	[0.005]	[0.005]	[0.020]
Financial problems in the family (missing)		-0.017	-0.018	-0.018	0.002	-0.060***
		[0.022]	[0.022]	[0.022]	[0.030]	[0.023]
Lives with mother		-0.010	-0.008	-0.008	-0.013	0.001
		[0.018]	[0.018]	[0.018]	[0.020]	[0.037]
Lives with father		-0.025	-0.024	-0.023	-0.020	-0.041
		[0.020]	[0.020]	[0.020]	[0.023]	[0.035]
Lives with mother and father		-0.019	-0.017	-0.017	-0.019	-0.012
		[0.017]	[0.017]	[0.017]	[0.019]	[0.033]
Lives with mother (missing)		0.015	0.013	0.014	-0.002	0.040

		[0.011]	[0.011]	[0.011]	[0.009]	[0.028]
Family size		-0.001	-0.001	-0.001	-0.003**	0.004
		[0.001]	[0.001]	[0.001]	[0.002]	[0.003]
Family size (missing)		-0.014	-0.014	-0.014	0.002	-0.057***
		[0.017]	[0.017]	[0.017]	[0.021]	[0.021]
Female		0.023***	0.023***	0.023***	0.021***	0.034**
		[0.004]	[0.004]	[0.004]	[0.005]	[0.014]
Mother's level of education		0.008***	0.009***	0.009***	0.007***	0.022***
		[0.002]	[0.002]	[0.002]	[0.002]	[0.007]
Father's level of education		-0.002	-0.001	-0.001	0.001	-0.013*
		[0.002]	[0.002]	[0.002]	[0.002]	[0.007]
Math test score: Q1					0.021	-0.097
					[0.015]	[0.097]
Math test score: Q2		0.005	0.005	0.006		0.010
		[0.027]	[0.027]	[0.027]		[0.026]
Math test score: Q3		0.016	0.017	0.018	0.023***	-0.023
		[0.027]	[0.027]	[0.027]	[0.007]	[0.023]
Math test score: Q4		0.027	0.028	0.029	0.030***	0.004
		[0.027]	[0.027]	[0.027]	[0.008]	[0.026]
Math test score: Q5		0.013	0.014	0.015	0.013*	0.024
		[0.027]	[0.027]	[0.027]	[0.008]	[0.030]
Math test score: Q6		0.026	0.027	0.028	0.027***	
		[0.028]	[0.027]	[0.028]	[0.009]	
Reading test score: Q1					0.003	0.175*
					[0.016]	[0.098]
Reading test score: Q2		-0.025	-0.024	-0.025	0.007	0.054**
		[0.027]	[0.027]	[0.027]	[0.010]	[0.025]
Reading test score: Q3		-0.031	-0.029	-0.030	0.001	0.048**
		[0.027]	[0.027]	[0.027]	[0.008]	[0.021]
Reading test score: Q4		-0.027	-0.025	-0.026	0.003	0.072***
		[0.027]	[0.027]	[0.027]	[0.008]	[0.027]
Reading test score: Q5		-0.031	-0.029	-0.030	0.002	0.043*
		[0.027]	[0.027]	[0.027]	[0.007]	[0.025]
Reading test score: Q6		-0.034	-0.032	-0.032		
		[0.027]	[0.027]	[0.027]		
Background questionnaire missing		-0.003	-0.006	-0.007	0.004	-0.031
		[0.010]	[0.010]	[0.010]	[0.012]	[0.023]
Constant	0.070***	0.050*	0.040	0.041	0.019	-0.028
	[0.002]	[0.026]	[0.027]	[0.026]	[0.028]	[0.065]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts
Observations	51,801	51,801	51,801	51,801	42,831	8,970
R-squared	0.000	0.002	0.002	0.002	0.002	0.005

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

Table S2_9: Substance abuse and mental health problems (the exposure-outcome link for substance abuse)

	Substance abuse (average monthly prevalence at age 22)					
	(1)	(2)	(3)	(4)	(5)	(6)
MHD (at least one diagnosis between ages 14 and 16)	0.008***	0.006***	0.005**	0.006**	0.006**	0.003
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.005]
Dropout			0.013***	0.013***		
			[0.002]	[0.002]		
Dropout * MHD				-0.002		
				[0.006]		
Time spent in kindergarten		0.001**	0.001***	0.001***	0.000	0.003***
		[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
Time spent in kindergarten (missing)		0.004	0.003	0.003	0.001	0.010
		[0.008]	[0.008]	[0.008]	[0.007]	[0.020]
Financial problems in the family		0.002***	0.002**	0.002**	0.001*	0.004
		[0.001]	[0.001]	[0.001]	[0.001]	[0.004]
Financial problems in the family (missing)		0.003	0.003	0.003	-0.002	0.010
		[0.010]	[0.009]	[0.009]	[0.008]	[0.024]
Lives with mother		-0.009*	-0.008	-0.008	-0.009	-0.008
		[0.005]	[0.005]	[0.005]	[0.006]	[0.010]
Lives with father		-0.008	-0.008	-0.008	-0.006	-0.016
		[0.006]	[0.006]	[0.006]	[0.007]	[0.011]
Lives with mother and father		-0.010**	-0.009*	-0.009*	-0.010*	-0.006
		[0.005]	[0.005]	[0.005]	[0.006]	[0.009]
Lives with mother (missing)		0.004*	0.002	0.002	0.002	0.001
		[0.002]	[0.002]	[0.002]	[0.002]	[0.005]
Family size		-0.000	-0.000	-0.000	0.000	-0.001
		[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
Family size (missing)		0.008	0.008	0.008	0.003	0.024
		[0.006]	[0.006]	[0.006]	[0.005]	[0.016]
Female		-0.008***	-0.008***	-0.008***	-0.005***	-0.023***
		[0.001]	[0.001]	[0.001]	[0.001]	[0.003]
Mother's level of education		-0.000	0.000	0.000	0.000	-0.000
		[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
Father's level of education		-0.000	0.000	0.000	0.000	0.000
		[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
Math test score: Q1					0.034	0.008
					[0.032]	[0.008]
Math test score: Q2		-0.028	-0.027	-0.027		0.003
		[0.024]	[0.024]	[0.024]		[0.007]
Math test score: Q3		-0.027	-0.026	-0.026	0.001	0.004
		[0.024]	[0.024]	[0.024]	[0.001]	[0.007]
Math test score: Q4		-0.027	-0.026	-0.026	-0.001	0.017**
		[0.024]	[0.024]	[0.024]	[0.001]	[0.008]
Math test score: Q5		-0.028	-0.027	-0.027	-0.001	0.007
		[0.024]	[0.024]	[0.024]	[0.001]	[0.008]
Math test score: Q6		-0.030	-0.028	-0.028	-0.002	
		[0.024]	[0.024]	[0.024]	[0.001]	
Reading test score: Q1					-0.031	0.014*
					[0.031]	[0.008]

Reading test score: Q2		0.021	0.022	0.022	0.001	0.009
		[0.024]	[0.024]	[0.024]	[0.001]	[0.006]
Reading test score: Q3		0.021	0.023	0.023	0.000	0.017***
		[0.024]	[0.024]	[0.024]	[0.001]	[0.006]
Reading test score: Q4		0.021	0.023	0.023	0.000	0.012*
		[0.024]	[0.024]	[0.024]	[0.001]	[0.006]
Reading test score: Q5		0.021	0.023	0.023	0.000	0.020**
		[0.024]	[0.024]	[0.024]	[0.001]	[0.008]
Reading test score: Q6		0.020	0.022	0.022		
		[0.024]	[0.024]	[0.024]		
Background questionnaire missing		0.007***	0.003*	0.003*	0.002	0.005
		[0.002]	[0.002]	[0.002]	[0.002]	[0.006]
Constant	0.007***	0.021***	0.012**	0.012**	0.012*	-0.005
	[0.000]	[0.006]	[0.006]	[0.006]	[0.007]	[0.013]
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	Non-dropouts	Dropouts
Observations	51,801	51,801	51,801	51,801	42,831	8,970
R-squared	0.000	0.007	0.009	0.009	0.003	0.013

Source: Admin 3. Sample of those born between June 1994 and May 1995. Robust standard errors in brackets. Significance: 1% ***; 5% **; 10% *.

. Robustness checks

Table S3_1: Sensitivity to potential unmeasured confounding in the four-way decomposition

Employed (binary outcome variable)					
	Estimated	95% confidence interval of RR		Estimated	95% confidence interval of E-values
	RR	Lower	Upper	E-values	Upper
CDE	0.965	0.913	1.020	1.230	1.000
PNDE	0.958	0.925	0.992	1.257	1.096
TNDE	0.959	0.927	0.991	1.255	1.102
PNIE	0.977	0.971	0.983	1.179	1.151
TNIE	0.978	0.971	0.984	1.176	1.144
TE	0.937	0.905	0.969	1.336	1.212
NEET (binary outcome variable)					
	Estimated	95% confidence interval of RR		Estimated	95% confidence interval of E-values
	RR	Lower	Upper	E-values	Lower
CDE	1.101	1.036	1.171	1.436	1.227
PNDE	1.090	1.056	1.125	1.403	1.300
TNDE	1.091	1.057	1.125	1.406	1.304
PNIE	1.045	1.034	1.057	1.263	1.222
TNIE	1.046	1.034	1.058	1.266	1.223
TE	1.140	1.102	1.180	1.540	1.437
Motherhood (binary outcome variable)					
	Estimated	95% confidence interval of RR		Estimated	95% confidence interval of E-values
	RR	Lower	Upper	E-values	Lower
CDE	1.034	0.946	1.131	1.222	1.000
PNDE	1.037	1.002	1.072	1.231	1.052
TNDE	1.036	1.002	1.072	1.231	1.043
PNIE	1.034	1.026	1.043	1.223	1.187
TNIE	1.034	1.025	1.044	1.222	1.184
TE	1.072	1.035	1.111	1.350	1.225
Log hourly wage (continuous outcome variable)					
	Beta	SD	Approx. RR	Approx. E-value	
CDE	-0.108	0.407	1.272	1.861	

PNDE	-0.044	0.407	1.103	1.441
TNDE	-0.049	0.407	1.115	1.472
PNIE	-0.012	0.407	1.026	1.190
TNIE	-0.016	0.407	1.037	1.232
TE	-0.060	0.407	1.144	1.550
Hours worked (continuous outcome variable)				
	Beta	SD	Approx. RR	Approx. E-value
CDE	-0.949	8.359	1.109	1.456
PNDE	-0.112	8.359	1.012	1.124
TNDE	-0.176	8.359	1.019	1.160
PNIE	-0.050	8.359	1.006	1.080
TNIE	-0.114	8.359	1.012	1.125
TE	-0.226	8.359	1.025	1.185
Cumulative number of abortions (continuous outcome variable)				
	Beta	SD	Approx. RR	Approx. E-value
CDE	0.155	0.536	1.302	1.929
PNDE	0.047	0.536	1.082	1.381
TNDE	0.054	0.536	1.097	1.423
PNIE	0.009	0.536	1.015	1.141
TNIE	0.017	0.536	1.029	1.202
TE	0.064	0.536	1.114	1.470
Cumulative number of STDs (continuous outcome variable)				
	Beta	SD	Approx. RR	Approx. E-value
CDE	0.066	0.486	1.131	1.515
PNDE	0.017	0.486	1.033	1.219
TNDE	0.021	0.486	1.040	1.244
PNIE	0.000	0.486	1.001	1.031
TNIE	0.004	0.486	1.007	1.094
TE	0.021	0.486	1.041	1.247
Monthly prevalence of substance abuse (continuous outcome variable)				
	Beta	SD	Approx. RR	Approx. E-value
CDE	0.003	0.083	1.036	1.229
PNDE	0.005	0.083	1.059	1.310
TNDE	0.005	0.083	1.058	1.304

PNIE	0.001	0.083	1.008	1.101
TNIE	0.001	0.083	1.007	1.090
TE	0.006	0.083	1.067	1.333

Source: Admin3. Sensitivity analysis to potential unobserved confounding in the four-way decomposition following VanderWeele and Ding (2017), implemented via the cmsens() function in the CMAverse package in R (Shi et al., 2021). For each model, we computed E-values for the total, natural direct, and natural indirect effects from the four-way decomposition (VanderWeele, 2014). CDE = Controlled Direct Effect (effect of MHDs on the outcome if the mediator is fixed); PNDE = Pure Natural Direct Effect (direct effect not through the mediator, holding the mediator at its untreated level); TNDE = Total Natural Direct Effect (direct effect allowing for interaction with the mediator); PNIE = Pure Natural Indirect Effect (indirect effect through the mediator, assuming no interaction); TNIE = Total Natural Indirect Effect (indirect effect through the mediator, allowing for interaction); TE = Total Effect (sum of all direct, indirect, and interaction components).

For binary outcomes (employment, NEET, motherhood), the E-value represents the minimum strength of association (on the risk-ratio scale) that an unmeasured confounder would need to have with both adolescent mental health disorders (MHDs) and the outcome, beyond measured covariates, to fully explain away the observed effect. It is defined as $E = RR + \sqrt{[RR \times (RR - 1)]}$, where RR denotes the estimated risk ratio between exposed and unexposed groups; larger values indicate greater robustness. For continuous outcomes, we applied the approximation proposed by VanderWeele and Ding (2017, Table 2), converting standardized mean differences to risk-ratio equivalents ($RR \approx \exp(0.91 \times |d|)$, where $d = \beta / SD$). Beta stands for the estimated effects as reported in Tables 1, 2 and 3 in the main text. This yields approximate E-values that enable comparability across outcome types, though these should be interpreted with caution.

Table S3_2: Sensitivity to measurement error in the exposure (misclassification in MHDs)

Outcome	TE baseline	TE robustness range	PM baseline	PM robustness range	Interpretation
Employed	-0.0308	-0.0493 – -0.0374	34.8%	31.1 – 34.7%	very stable
NEET	0.0568	0.0677 – 0.0859	34.5%	33.6 – 42.6%	stable
Hours worked	-0.2233	-0.2994 – -0.2861	50.4%	39.6 – 46.8%	direction stable, imprecise
Log hourly wage	-0.0605	-0.0852 – -0.0732	26.7%	27.7 – 33.7%	stable
Motherhood	0.0210	0.0239 – 0.0411	48.3%	48.5 – 60.2%	stable, relatively high mediation
Abortions	0.0629	0.0779 – 0.1075	26.7%	27.0 – 31.9%	stable, somewhat larger effects
STDs	0.0213	0.0264 – 0.0316	18.6%	19.5 – 21.4%	stable
Substance abuse	0.0059	0.0072 – 0.0078	10.5%	9.4 – 10.2%	very stable

Source: Admin3. Sensitivity analysis to potential measurement error (misclassification) in the exposure (MHDs) in the four-way decomposition following VanderWeele and Ding (2017), implemented via the CMAverse package in R (Shi et al., 2021). The observed binary exposure (MHDs at age 16) is treated as a misclassified proxy for a latent true exposure. For each model, we assume values for sensitivity (Se) and specificity (Sp), fixing Sp at 0.99 and varying Se between 0.30 and 0.90. For each (Se, Sp) pair, we first estimate the probability of the observed exposure conditional on covariates, $P(A^* = 1 \mid C)$, using logistic regression. We then invert the misclassification relationship to recover the latent probability $P(A = 1 \mid C)$, and apply Bayes' rule to compute the posterior probability of the true exposure, $P(A = 1 \mid A^*, C)$. The observed exposure is replaced with this posterior probability, and the four-way decomposition is re-estimated using this corrected exposure. We report the Total Effect (TE) and the Proportion Mediated (PM), where PM captures the share of the total effect explained by the mediated pathway. More detailed results are available on request from the authors.

Table S3_3: Sensitivity to the age range of measuring MHDs (Having a diagnosis between ages 14-17 as opposed to 14-16 as in the main results)

	Employed		NEET		Hours worked		Log hourly wage	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
TE: Total effect	-0.039***	(-0.054, -0.027)	0.060***	(0.048, 0.075)	-0.417*	(-0.837, -0.024)	-0.066***	(-0.087, -0.045)
Proportions attributable to mediation and interaction								
PM: Proportion mediated	0.258***	(0.150, 0.402)	0.306***	(0.228, 0.390)	0.270*	(0.042, 1.746)	0.243***	(0.180, 0.351)
PI: Proportion attributable to interaction	0.450	(-0.132, 0.990)	0.189	(-0.171, 0.549)	-1.557	(-9.693, 0.684)	-0.597*	(-1.185, -0.054)
Four-way decomposition								
CDE: Controlled direct effect	-0.009	(-0.030, 0.012)	0.030*	(0.009, 0.054)	-1.014*	(-1.941, -0.111)	-0.093***	(-0.141, -0.045)
INTref: Reference interaction effect	-0.018	(-0.042, 0.006)	0.012	(-0.009, 0.036)	0.711	(-0.183, 1.533)	0.042*	(0.003, 0.084)
INTmed: Mediated interaction effect	0.000	(0.000, 0.003)	0.000	(-0.003, 0.000)	-0.060	(-0.129, 0.015)	-0.003*	(-0.006, 0.000)
PIE: Pure indirect effect	-0.012***	(-0.015, -0.009)	0.021***	(0.015, 0.024)	-0.054***	(-0.084, -0.030)	-0.012***	(-0.015, -0.009)
Two-way decomposition								
TNIE: Total indirect effect	-0.009***	(-0.012, -0.006)	0.018***	(0.015, 0.021)	-0.114***	(-0.192, -0.045)	-0.015***	(-0.021, -0.012)
PNDE: Pure direct effect	-0.027***	(-0.045, -0.015)	0.042***	(0.030, 0.054)	-0.306	(-0.711, 0.084)	-0.051***	(-0.069, -0.033)

	Probability of motherhood		Cumulative number of abortions		STDs		Substance abuse	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
TE: Total effect	0.018***	(0.009, 0.027)	0.054***	(0.033, 0.066)	0.021*	(0.003, 0.036)	0.009***	(0.003, 0.012)
Proportions attributable to mediation and interaction								
PM: Proportion mediated	0.555***	(0.339, 0.933)	0.288***	(0.213, 0.378)	0.186*	(0.030, 0.720)	0.132***	(0.051, 0.249)
PI: Proportion attributable to interaction	0.270	(-0.699, 1.743)	-1.365***	(-2.067, -0.663)	-2.064	(-7.563, 0.018)	-0.480	(-1.590, 0.384)
Four-way decomposition								
CDE: Controlled direct effect	0.003	(-0.021, 0.024)	0.114***	(0.063, 0.168)	0.066***	(0.024, 0.120)	0.012*	(0.003, 0.021)
INTref: Reference interaction effect	0.006	(-0.015, 0.027)	-0.078***	(-0.120, -0.030)	-0.048*	(-0.093, -0.006)	-0.003	(-0.015, 0.003)
INTmed: Mediated interaction effect	0.000	(-0.003, 0.000)	0.006***	(0.003, 0.009)	0.003*	(0.000, 0.006)	0.000	(0.000, 0.000)
PIE: Pure indirect effect	0.009***	(0.009, 0.012)	0.009***	(0.006, 0.012)	0.000	(0.000, 0.000)	0.000***	(0.000, 0.000)
Two-way decomposition								
TNIE: Total indirect effect	0.009***	(0.006, 0.012)	0.015***	(0.012, 0.018)	0.003*	(0.000, 0.006)	0.000***	(0.000, 0.003)

PNDE: Pure direct effect	0.009*	(0.000, 0.018)	0.039***	(0.021, 0.051)	0.018	(0.000, 0.033)	0.006***	(0.003, 0.012)

Source: Admin3. The results of a four-way decomposition causal mediation analysis by VanderWeele (2014), operationalised via the CMAverse package in R (Shi et al., 2021). 95% bootstrapped confidence intervals in parentheses. Significance: 1% ***; 5% **; 10% *.

S4. Heterogeneity by gender

Table S4_1: Causal mediation analysis of the association between MHDs and labour market outcomes: The role of dropping out of school. Results by gender

Outcome	Total effects (TE)		Proportion mediated (PM)	
	Men	Women	Men	Women
Employed	-0.053*** (-0.074, -0.032)	-0.013 (-0.035, 0.009)	0.160*** (0.076, 0.312)	0.881 (-17.616, 4.513)
NEET	0.075*** (0.053, 0.092)	0.042*** (0.024, 0.062)	0.252*** (0.161, 0.389)	0.457*** (0.293, 0.803)
Hours worked	-0.401 (-0.966, 0.250)	-0.072 (-0.701, 0.622)	0.230 (-1.193, 3.545)	2.043 (-6.195, 3.966)
Log hourly wage	-0.092*** (-0.129, -0.058)	-0.031** (-0.060, -0.007)	0.241*** (0.118, 0.384)	0.338** (0.131, 1.200)
STDs	0.036** (0.004, 0.066)	0.010 (-0.010, 0.035)	0.070 (-0.135, 0.432)	0.482 (-5.036, 6.502)
Substance abuse	0.008** (0.001, 0.015)	0.004* (0.000, 0.009)	0.192* (-0.041, 1.727)	-0.005 (-0.202, 0.319)

Source: Admin3. The results of a four-way decomposition causal mediation analysis by VanderWeele (2014), operationalized via the CMAverse package in R (Shi et al., 2021). 95% bootstrapped confidence intervals in parentheses. Significance: 1% ***; 5% **; 10% *.