Lockdown, Family Conflict, and Adolescent Mental Health in China

David Lander* & Jingyi Yu†

July 31, 2025

Abstract

This paper studies how COVID-19 and subsequent lockdown measures in 2020 affected the mental health of adolescents in China. Using nationally-representative China Family Panel Studies (CFPS) data, we note that the overwhelming majority of students were subject to lockdown in 2020. Therefore, we employ a cohort difference-in-difference approach, comparing adolescents aged 10–15 years in 2018–2020 (treatment) with adolescents of the same age cohort in 2016–2018 (control). Our main finding is that the severity of depressive symptoms increased by 9–11%, on average, following the COVID outbreak and lockdown response. This estimate accounts for the typical decline in mental health experienced as students progress through school. Our results highlight two primary channels: economic hardship and rising intra-household conflict. However, the conflict channel appears independent of economic hardship and was likely driven by increased household exposure during lockdown, with adolescents from households experiencing pre-existing conflict suffering the greatest deterioration in mental health.

JEL codes: I12, I18, J12, J13, D19.
Keywords: CES-D, depression, depressive, child, partner, spouse, violence, IPV, exposure.

*Affiliation: Peking University HSBC Business School. Correspondence: lander@phbs.pku.edu.cn.
†Affiliation: Peking University HSBC Business School.

1 Introduction

The adverse effects of COVID-19 (henceforth, COVID) have been extensively studied since its emergence, which includes significant and widespread increases in depressive and anxiety symptoms among adolescents.[1] There are, however, relatively few longitudinal studies based in China, due to a lack of detailed nationally-representative individual survey data.[2] In this study, we utilise the 2016–2020 waves of the China Family Panel Studies (CFPS) to measure the causal effect of COVID (and subsequent policy responses) on the mental health of adolescents in mainland China.

A standard difference-in-difference (DiD) approach comparing affected to unaffected regions is infeasible in China, since virtually all adolescents experienced lockdown in 2020.[3] Therefore, we employ a cohort-based DiD strategy, comparing adolescents aged 10–15 in 2020 (i.e., post-COVID) with a similar-age cohort from two years prior.[4] Additionally, by observing adolescents for multiple waves in both the treated cohort (i.e., 2018–20) and untreated cohort (i.e., 2016–18), we can include individual fixed effects and thus control for unobservable time-invariant factors affecting each adolescent's mental health. This allows us to plausibly estimate COVID's short-term causal impact, assuming that adolescent mental health trends across these cohorts would have evolved similarly in the absence of COVID.

Our main findings reveal that COVID caused adolescents in China to experience a significant deterioration in mental health, corresponding to a 9–11% increase in the CES-D score (from the Center for Epidemiologic Studies Depression scale), on average. While a decline in adolescent mental health may not seem particularly surprising since it was observed in many countries after COVID, it stands in contrast with several recent studies in China: using longitudinal data from the Chengdu Positive Child Development (CPCD) Research project, Shi et al. (2022), Zhao et al. (2023), and Zeng et al. (2024) all find moderate improvements in adolescent mental health in 2020 following the COVID outbreak.

Additionally, we examine several channels through which COVID may affect mental health: economic hardship, intra-household conflict, physical exercise, and internet use/screen time. While we find evidence indicating each of these channels may be present in China, further analysis suggests that economic hardship (measured by a substantial fall in

[1] For recent comprehensive reviews of the literature on the impact of COVID on the mental health of children and adolescents, see, e.g., Deng et al. (2023), Miao et al. (2023), or Panchal et al. (2023).

[2] Several ad hoc surveys were administered shortly after the outbreak in 2020. They typically were narrow in scope, and targeted a single city/province; e.g., Chen et al. (2020), Xie et al. (2020), Liang et al. (2020).

[3] Such an approach has been used to measure the impact of lockdown policies in several recent studies, e.g., Serrano-Alarcón et al. (2022) and Colella et al. (2023).

[4] Guariso and Nyqvist (2023) also used a cohort-based approach to study the impact of school closure during COVID on the learning outcomes and mental well-being of primary-school students in India.

family income) and intra-household conflict (increasingly frequent arguments between children and their parents), were particularly harmful to adolescent mental health during COVID.

Our findings point toward lockdown as the likely cause underlying the most intense deteriorations in adolescent mental health in China during 2020. We arrive at this conclusion by first uncovering evidence that adolescents in households experiencing large income drops or increases in conflict (between children and parents) experience worsening mental health compared to adolescents experiencing the exact same 'shocks' several years prior to COVID. This is indicative of an amplification of these phenomena on mental health during COVID, and in both cases lockdowns plausibly intensify conflict and prohibit job search.

Additionally, we find that COVID caused adolescents in households with pre-existing conflict between family members (i.e., prior to COVID) to experience a relatively severe deterioration in mental health. On average, this entails an approximately 12.5–13.6% increase in CES-D20 score, which appears to result from these households experiencing a rise in arguments by around 25% per month, relative to a control group (also with pre-existing conflict).

Interestingly, we find that many households in China experienced either economic hardship or rising conflict, but rarely both. This stands in contrast to most related studies investigating the impact of COVID on intra-household conflict, particularly domestic violence, which invariably point towards economic hardship as the trigger for conflict.[5] In the absence of such a link, we reject common explanations for rising conflict based on bargaining power and backlash (Macmillan and Gartner, 1999; Aizer, 2010). Instead, we conclude that exposure (i.e., from lockdown/quarantine) is the most likely explanation for the rise in conflict in China (Dugan et al., 1999, 2003).[6]

Although the literature on the impact of COVID on mental health is extraordinarily vast, our study also contributes to this literature in several meaningful ways. Early studies completed shortly after the outbreak of COVID simply measured levels of depressive symptoms in the population, or subpopulations, and compared them to levels obtained in studies published prior to the COVID outbreak.[7] Since many of the studies in 2020 measured depression by issuing ad-hoc surveys, they were unable to adequately control for time-varying confounders, as pointed out, e.g., by Prati and Mancini (2021). Subsequently, a number of these surveys included follow-ups, enabling them to include individual fixed effects and control for time-

[5] See, e.g., Baranov et al. (2022) (Pakistan) and Bhalotra et al. (2024) (Chile).

[6] This finding is in line with Hsu and Henke (2021) (US), Agüero (2021) (Peru), and Ivandic et al. (2021) (UK), all of whom emphasise the role of exposure during COVID. We also note that our findings align with some results of Arenas-Arroyo et al. (2021), who finds evidence supporting exposure (from lockdowns) specifically for psychological abuse—which is the closest counterpart to our measure of conflict (i.e., arguments). In general, however, other forms of abuse (e.g., physical and sexual) only increased due to economic hardship.

[7] For an early review of such studies, see, e.g., Racine et al. (2021) or Ma et al. (2021).

invariant factors.[8] The most recent longitudinal studies in China, however, are confined to a particular city or province, e.g., Guangzhou (Wang et al., 2022), Chengdu (Zhao et al., 2023; Zeng et al., 2024), Shandong (Chen et al., 2022), Shanghai (Liu et al., 2024), and/or lack detailed individual data.

In contrast, our study utilises a robust cohort DiD design together with detailed nationally-representative individual-level panel data, which controls for both time-varying confounders and individual-level fixed effects. Thus, we are able to quantify the substantial mental health decline experienced by adolescents in China caused by COVID, as well as identify the increase in intra-household conflict as a key channel through which lockdown affected mental health. Finally, we show our results are robust to a variety of alternative specifications, including a 'doubly robust' DiD estimator, by Sant'Anna and Zhao (2020).

The paper proceeds as follows: Section 2 introduces the CFPS data, our sample selection criteria, and key variables of interest, Section 3 describes our empirical model and identification strategy, Section 4 reports our empirical findings, including a discussion of mechanisms and robustness, and Section 5 offers concluding remarks.

2 Data

2.1 China Family Panel Studies survey

This paper uses data from the China Family Panel Studies (CFPS) database, which is designed and constructed by the China Social Science Research Center of Peking University. The sample covers 25 provinces/municipalities/autonomous regions and is nationally representative. The survey focuses on the family relationships and economic activities of the participating households, as well as the education and health information of the participating individuals. It has collected and published six complete waves of data, up to 2020. The 2020 data were all collected from July to December in that year, and they reflect the basic conditions of Chinese society in the months following the outbreak of the COVID epidemic. This wave also contains several new questions specifically relating to COVID, therefore, it is the most detailed and appropriate Chinese survey data to study our research question.

To implement our empirical strategy, we require data from 2016, 2018, and 2020. The data from 2016 to 2018 are the most recent waves before the epidemic in the CFPS, and the characteristics of the student population as a whole do not change significantly during the relatively short time span from 2016 to 2020, which makes it easier to meet the prerequisite assumption of comparability across the treatment and control groups. Additionally, between

[8] For reviews of these studies, see Robinson et al. (2022), Madigan et al. (2023), Wolf and Schmitz (2024).

2016 and 2020, there are no major events or policy changes (besides the outbreak of COVID) that would clearly have an impact on the mental-health status of adolescents. Therefore, we are confident that the sample period covered by these waves enables us to identify the impact of the epidemic on adolescents' mental health.

2.2 Sample selection criteria

For these three waves of data (2016, 2018, and 2020), we first merged the individual self-answer questionnaire, the child proxy questionnaire, and the family questionnaire of the CFPS to construct a database containing both personal and family information. The combined database was then filtered as follows: (1) Only adolescents aged between 10–15 years were retained, because only this subsample provide both the self-answer information and child proxy information in the CFPS; (2) Non-student adolescents were excluded, i.e., those aged 10–15 who reported that they were not attending school;[9] (3) Observations with missing values for the variables (listed in Table 1) were excluded; (4) Retain individuals for whom the data were available for at least two consecutive years, i.e., only those observations for which data were available for 2016 and 2018, 2018 and 2020, or all three waves. This ensures we can include individual fixed effects in our main regression equation.

After processing, we obtained a total of 4,396 valid observations, from 1,767 unique students. In Section 3, we describe our empirical strategy and give a detailed breakdown of the distribution of students across waves.

2.3 Key variables

All of the variables we construct from the CFPS survey are listed in Table 1, and a detailed description of each can be found in Appendix Table A1. In this section, we focus on defining our measures of mental health, based on the information contained in the CFPS data, as well as providing a concise description of the other variables we use.

2.3.1 Mental-health measure

The main dependent variable in this paper is the score of the Center for Epidemiologic Studies Depression (CES-D) scale, designed by Radloff (1977).[10] It contains twenty questions that

[9]Non-student adolescents make up less than 1% of the overall sample. These adolescents should, according to compulsory education laws, be attending school in this age range.

[10]The CFPS utilises the Chinese-language version of the CES-D scale, which has been validated by, e.g., Cheung and Bagley (1998).

assess the respondent's level of depression by measuring the occurrence, frequency and intensity of symptoms of the past week. The scale assesses six areas: depressed mood, feelings of guilt and worthlessness, feelings of helplessness and hopelessness, psychomotor retardation, decreased appetite, and sleep disturbance. The frequency of the above negative symptoms in the past week is assigned one of four values: 0 for 'hardly ever' (less than one day), 1 for 'some of the time' (1–2 days), 2 for 'often' (3–4 days), or 3 for 'most of the time' (5–7 days). This yields a total score of 0–60, where a higher score indicates more severe depressive symptoms.[11]

The CFPS used this scale in all surveys waves from 2016 to 2020, however, the range of items/questions on the CES-D scale asked to respondents varied across waves. In 2016, 20% of respondents answered the full twenty questions (i.e., the 'long-form' CES-D20 scale) and the remaining 80% answered only eight representative questions (i.e., the 'short-form' CES-D8 scale). While, in 2018 and 2020, all respondents only answered the CES-D8 questions. Since the eight CES-D8 questions are a subset of the twenty CES-D20 questions, the CFPS used the results from the 2016 wave to determine a mapping from CES-D8 scores to CES-D20 scores (via 'equipercentile equating').[12] We primarily use the CES-D20 scale throughout the paper, however we report baseline estimates for both scales (in Table 2) to demonstrate that our main findings are not sensitive to the choice of scale.

2.3.2 Control variables

To control for relevant measurable time-varying factors that affect adolescent mental health status (e.g., own characteristics, school factors, and family factors), we include the following independent variables as controls in our empirical model.[13]

Individual-level controls. (i) Study-related variables: a dummy variable for whether the student is in a 'key' school, the number of study hours per day on weekdays, the number of study hours per day on weekends, the number of hours per week in extracurricular classes, a dummy variable for whether the student serves as a student leader, satisfaction with educational performance, pressure related to study, self-assessed excellence as a student, satisfaction with the school, satisfaction with the classroom teacher; (ii) Individual characteristics: a dummy variable for having any pocket money, a dummy variable for using a mobile device (e.g., phone or tablet) with internet access, a dummy variable for using a computer

[11]Note that four out of twenty questions in the scale measure the frequency of positive emotions. For these questions, a value is assigned in the reverse order to that of negative emotions in order to yield a consistent interpretation; i.e., a larger score is indicative of more (less) frequent negative (positive) emotions.

[12]For further details, see: http://www.isss.pku.edu.cn/cfps/docs/20201201085335172101.pdf (in Chinese).

[13]A detailed description of each variable is contained in Appendix Table A1. Additionally, several variables with implausible extreme values were winsorized at the 99$^{\text{th}}$ percentile (see Appendix B.1 for details).

Table 1: Summary statistics

	Mean	Std.Dev.	Min.	Max.	Obs.
Dependent variable					
CES-D20 score	9.880	6.090	2	48	4,396
CES-D8 score	3.888	3.088	0	22	4,396
Control variables					
Attend key school	0.210	0.407	0	1	4,396
Extracurricular class hours	3.024	7.394	0	42	4,396
Hours studying (weekdays)	8.372	2.991	0	20	4,396
Hours studying (weekends)	3.762	2.757	0	20	4,396
Student leader	0.340	0.474	0	1	4,396
Student performance	3.377	0.926	1	5	4,396
Study pressure	2.824	1.103	1	5	4,396
Excellence	3.144	0.832	1	5	4,396
Satisfaction (school)	4.127	0.972	1	5	4,396
Satisfaction (teacher)	4.348	0.975	1	5	4,396
Pocket money	0.774	0.419	0	1	4,396
Internet (mobile phone/tablet)	0.556	0.497	0	1	4,396
Internet (computer)	0.295	0.456	0	1	4,396
Sick (past month)	0.174	0.380	0	1	4,396
Sick (past year)	0.276	0.447	0	1	4,396
Out-of-pocket medical expenses (log)	3.792	2.769	0	12.324	4,396
Father lives at home	0.802	0.398	0	1	4,396
Mother lives at home	0.847	0.360	0	1	4,396
Family size	5.094	1.748	2	12	4,396
Mechanism variables					
Family income (log)	10.909	1.000	0	15.618	4,396
Physical exercise	0.646	0.478	0	1	4,396
Internet use (hours)	7.078	11.811	0	60	4,396
Conflict in the home	0.470	0.499	0	1	4,396
Total number of arguments	1.939	3.538	0	30	4,396

Source: CFPS (2016, 2018, 2020).

with internet access, a dummy variable for being ill in the past month, a dummy variable for being ill in the past year, and the log value of out-of-pocket medical expenses.

Family-level controls. A dummy variable for whether the father lives at home, a dummy variable for whether the mother lives at home, and family size.[14]

2.3.3 Mechanism variables

In Section 4.2, we investigate a number of channels/mechanisms through which COVID may potentially impact mental health. Since this set of variables are likely affected by COVID,

[14]Since it may be possible that these family-level control variables are affected by our treatment, COVID (e.g., a COVID death in the family or a parent being subject to COVID-related quarantine), in Appendix B.4 we demonstrate that our main findings are robust to excluding them as controls.

they are not included as controls. These include (i) a measure of the economic impact of COVID on the household (the log value of family income), (ii) a binary measure of physical exercise (in the past week), (iii) a measure of internet use/screen time (hours per week), and (iv) several measures of intra-household conflict (whether children had an argument with a parent or their parents had an argument in the past month, and the total number of either type of arguments in the past month).

3 Empirical strategy

The most direct approach to study the impact of COVID or quarantine/lockdown on adolescents' mental health would be to use a difference-in-difference (DiD) method, in which adolescents living in areas affected by the epidemic (and thus have quarantine/lockdown experience) are defined as the treatment group and compared to a control group of unaffected adolescents. This identification strategy, however, is not feasible using the CFPS data.

After COVID emerged in China, it rapidly spread to other provinces and cities within several months. This is reflected in the CFPS data, which reveals that over 90% of adolescents in our sample experienced quarantine/lockdown in 2020. Consequently, it is not possible to define a suitable control group in the 2020 CFPS wave. Therefore, we use CFPS data from before the epidemic period to construct a control group.

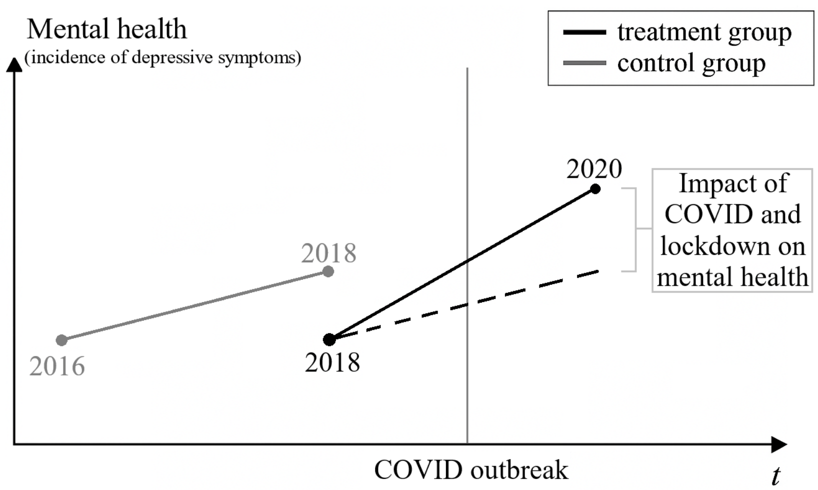
We formalise this identification strategy as follows: our **treatment group** (affected by COVID) is a cohort of students measured in the CFPS in both 2018 (pre-COVID) and 2020 (post-COVID), while our **control group** is a cohort of *similarly-aged* students measured in the CFPS in both 2016 and 2018 (i.e., both pre-COVID).[15] A stylised representation of this is shown in Figure 1.

Rather than control the year (i.e., compare some students in the 2020 to other students in 2020), we control the age of students (i.e., compare students in 2020 with similarly-aged students in 2018). This requires the average mental health level of students of similar ages (and hence at a similar grade in school) to not significantly vary across cohorts, in the absence of COVID. Specifically, for our sample, it requires that changes in mental health across grades (measured two years apart) are the same across cohorts (measured two years apart).

Intuitively, this assumption seems reasonable, since a two-year timespan is sufficiently short that it seems unlikely to produce systematically different cohorts due to, say, changes in technology, parenting methods, teaching methods, education policy, mental health policy, and

[15] A similar cohort-based identification strategy, in the context of COVID, was utilised by Guariso and Nyqvist (2023). Their focus was on measuring the impact of school closures on a cohort of primary-school children in India, both in terms of learning outcomes and psychological well-being.

Figure 1: Identification strategy



Note: The bold dashed line represents the counterfactual trend in mental health for the treatment group, in the absence of COVID.

so on.[16] Unfortunately, however, we are prevented from directly demonstrating the validity of our key identifying assumption due to a limitation in the CFPS survey design. That is, we cannot test whether the mental health of school-age children evolves in "parallel" over time (across cohorts two years apart), because we cannot compare the 2016–18 trend with the 2014–16 trend, since the CFPS uses a different measure of mental health in the 2014 wave compared with all subsequent waves.[17]

Instead, we show that these cohorts appear to be very similar for most key variables, *in levels*, in their respective pre-treatment periods in Appendix Table A2. This is in line with arguments made by Kahn-Lang and Lang (2020), who emphasise the importance of similarities in (pre-treatment) levels across groups, as well as noting that parallel pre-trends are neither necessary nor sufficient for parallel counterfactual trends. Nonetheless, any minor imbalance in covariates across cohorts (in levels) is addressed when we implement the doubly robust DiD estimator, in Section 4.1.4. There appears to be little evidence of selection (on observables) across cohorts, since it yields almost identical estimates to our baseline model.

The stylised trend in mental health for our control group (from 2016 to 2018) in Figure 1 is consistent with a variety of studies reporting positive time trends (i.e., a deterioration) in the mental health of school-age children in China between 2010–20 using different datasets (Tang et al., 2019; Li et al., 2019). The consensus is that the incidence of depressive symptoms

[16]This assumption will, however, become increasingly questionable over a longer timespan as time-varying factors (e.g., changing technology, teaching methods) differentially affect mental health of students across cohorts.

[17]The 2010 and 2014 waves of the CFPS use the Kessler (K6) scale, while the 2012, 2016, 2018, and 2020 waves all use the CES-D scale.

accelerates as age increases, particularly as children approach high school and ultimately face the daunting university entrance examination (‘*gaokao*’).

Based on the above strategy, the corresponding regression model for student $j = 1, \ldots, N$, in cohort $c = 0, 1$, at time period $t = 1, 2$, is:

$$MH_{j,c,t} = \alpha_0 + \delta(\mathcal{I}^{Treat}_{j,c} \times \mathcal{I}^{Post}_{t}) + \gamma_1 \mathcal{I}^{Treat}_{j,c} + \gamma_2 \mathcal{I}^{Post}_{t} + \beta X_{j,t} + \mu_j + \epsilon_{j,c,t}, \tag{1}$$

where $MH_{j,c,t}$ is the CES-D score for student j in cohort c at period t. $\mathcal{I}^{Treat}_{j,c}$ is an individual-specific indicator variable which equals 1 for students in the treated cohort $c = 1$ (i.e., 2018–20), and 0 for students in the control cohort $c = 0$ (i.e., 2016–18). $\mathcal{I}^{Post}_{t}$ is a period-specific indicator variable which equals 1 in the final period we observe each student ($t = 2$), and 0 in the first period ($t = 1$); specifically, in terms of years: for treated students it equals 1 in 2020 and equals 0 in 2018, and, for untreated students it equals 1 in 2018 and equals 0 in 2016.

Finally, $X_{j,t}$ includes all time-varying control variables, μ_j are individual FEs, and $\epsilon_{j,c,t}$ is the idiosyncratic error term. Note we do not include wave dummies here. The individual fixed effect here controls for factors that are unobservable and time-invariant for individuals, as well as ensuring that we take into account differences in depression levels across individuals.

This departs from the standard two-way fixed effects (TWFE) DiD setup, since treatment is cohort-specific in additional to individual-specific; hence it is not absorbed by the individual fixed effect. This is because some adolescents appear in both the treatment and control group, since they meet the age criteria to be included in both; e.g., a student aged 10–12 may appear in the control group in 2016–18, and that same student would be eligible to be in the treatment group aged 12–14 in 2018–20. If we were to restrict such a case, e.g., by only allowing these individuals to appear in either the treatment or control group, $\mathcal{I}^{Treat}_{j,c}$ would be absorbed by μ_j. Furthermore, in this case the setup is equivalent to TWFE, since $\mathcal{I}^{Treat}_{j,c} \times \mathcal{I}^{Post}_{t}$ becomes equivalent to a 2020 wave dummy, and $\mathcal{I}^{Post}_{t}$ to a 2018 wave dummy (with 2016 left as the reference group).[18]

To ensure we maintain a consistent age-distribution of students across our control and treatment groups, we must keep the (relatively young) students from the 2016–18 control cohort in the 2018–20 treated cohort (as relatively old students). This is demonstrated, by example and empirically, in Appendix B.2. We refer to these cases throughout the paper as ‘overlapping’ students. Using the specification in Equation (1) requires us to duplicate these overlapping students in the data (for the 2018 wave only), assigning one to the control group

[18] Specifically, if no student j appears in both the treatment and control group, we can define the binary treatment indicator as $\mathcal{I}^{Treat}_{j}$, and our regression becomes: $MH_{j,t} = \alpha + \delta(\mathcal{I}^{Treat}_{j} \times \mathcal{I}^{Post}_{t}) + \gamma \mathcal{I}^{Post}_{t} + \beta X_{j,t} + \mu_j + \epsilon_{j,t}$, where $\gamma \mathcal{I}^{Post}_{t}$ is equivalent to a period fixed effect.

and another to the treatment group.[19] This results in a sample of 1,767 unique students (965 treated cohort and 1,233 in the control cohort), including 431 that appear in both cohorts, yielding a total of 4,396 observations (2,466 control, 1,930 treated). In Section 4.1.4, we conduct several robustness tests on this assumption regarding overlapping students.

4 Results

4.1 Impact of COVID on adolescent mental health

Our main results are reported in Table 2, which contains the estimated coefficients from (1), with and without controlling for individual fixed effects. The coefficient of interest from (1) is δ, the coefficient on the interaction term $(\mathcal{I}_{j,c}^{Treat} \times \mathcal{I}_t^{Post})$, which represents the causal effect of COVID on adolescent mental health in China. It is reported in Table 2 under the row labelled Treat $\times$ Post. The estimates of γ_1 and γ_2 in (1), i.e., the coefficients on $\mathcal{I}_{j,c}^{Treat}$ and $\mathcal{I}_t^{Post}$, respectively, are reported in the rows labelled Treat and Post. All specifications include the control variables from Table 1, however their coefficient estimates are omitted for readability—the full table of estimates can be found in Appendix C.

All specifications have standard errors clustered by student (i.e., by each unique student in the CFPS) to avoid underestimating standard errors due to duplication of overlapping students. Finally, as a further robustness check regarding standard errors, we also consistently report wild-bootstrapped p-values (in brackets) for the causal effect of COVID.

Columns (1) and (2) present estimates with the CES-D20 scale as the dependent variable measuring mental health. These estimates show that the outbreak of COVID significantly increased the incidence of depressive symptoms among students (at a 5% level of significance). The magnitude of this effect is estimated to be 0.892 (column 1) without individual fixed effects, and 0.879 (column 2) when individual fixed effects are included—hence accounting for time-invariant unobservables marginally decreases the magnitude of the estimated effect. This effect is quite large, considering the mean CES-D20 score is 9.62 for students in the treated cohort before COVID (i.e., it is approximately a 9.1% increase).[20]

Columns (3) and (4) report results obtained with the CES-D8 scale. The effect remains positive and significant for both specifications. Note that the reduced magnitude of coefficient estimates, compared to columns (1) and (2), simply reflects the change in the range of the

[19] Alternatively, we could (equivalently) employ a two-step estimator, estimating the trend for each cohort separately, then taking the difference. This yields an identical estimate of the causal effect, since 'duplication' is essentially built into the two-step method.

[20] See Appendix Table A2 for summary statistics measured in the pre-treatment period.

Table 2: The effect of COVID on adolescent depressive symptoms

	CES-D20 score		CES-D8 score	
	(1)	(2)	(3)	(4)
Treat $\times$ Post	0.8916**	0.8793**	0.4308**	0.4241**
	(0.349)	(0.346)	(0.177)	(0.175)
	[0.009]	[0.011]	[0.010]	[0.015]
Post	0.3428	0.1345	0.1841*	0.0802
	(0.210)	(0.225)	(0.107)	(0.114)
Controls	Yes	Yes	Yes	Yes
Individual FE	No	Yes	No	Yes
R-squared	0.0912	0.5960	0.0913	0.5972
R-squared (within)	–	0.0472	–	0.0468
Observations	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). The dependent variable is listed at the top of the column, e.g., the CES-D20 scale for columns (1) and (2). The full table of coefficient estimates can be found in Appendix C Table C1. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

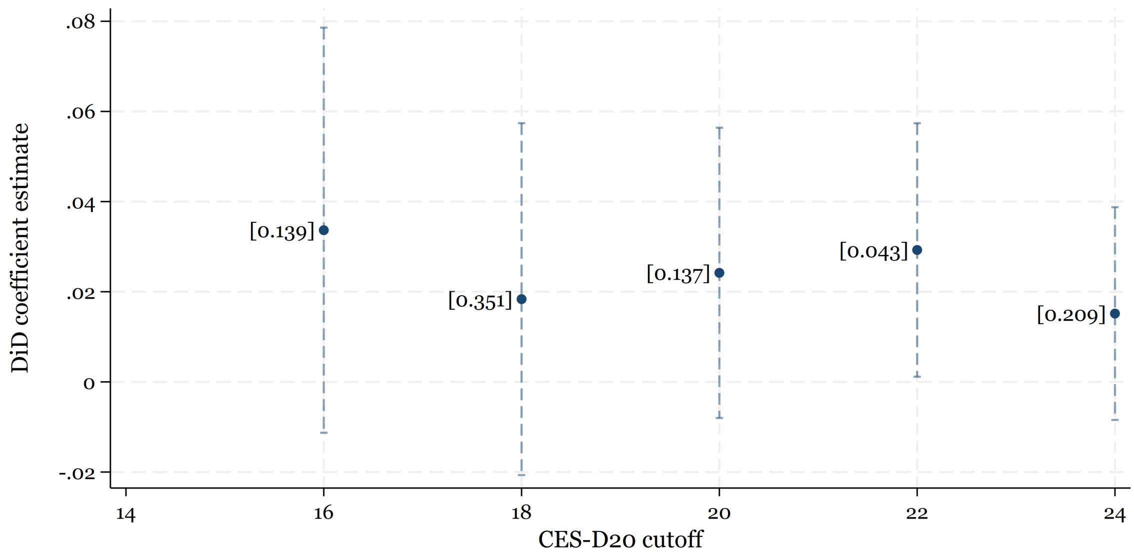
CES-D8 scale (i.e., from 0–24). Hence, given the mean CES-D8 score for students in the treated cohort before COVID is 3.76, this estimate represents a 11.3% increase.

An increasing trend in depressive symptoms by age/education level for adolescents in China has been well documented in several meta-analyses of past studies (prior to the COVID outbreak), e.g., Tang et al. (2019), Li et al. (2019). Our findings are consistent with this phenomenon, as indicated by the positive coefficient estimates for Post, which measures the trend in mental health over age/education level (common across cohorts). Interestingly, however, while the coefficient on Post is quite large in columns (1) and (3), both its magnitude and significance diminishes once individual fixed effects are included, in columns (2) and (4). This suggests prior studies may overstate the magnitude of this phenomenon by not accounting for child-specific time-invariant unobservables.

4.1.1 Indicator of depression

While a 9–11% CES-D score increase is indeed a measurable deterioration in average adolescent mental health, it is not apparent whether this magnitude is meaningful, from a psychometric perspective. Therefore, we define a binary indicator for relatively 'severe' (or elevated) depressive symptoms, which equals 1 if the CES-D20 score is above some cutoff (e.g., 20), and 0 otherwise, i.e., $\mathcal{I}(\text{CES-D20} \geq 20)$. This is constructed to convey specific information—a sufficiently high level of depressive symptoms, consistent with a diagnosis of depression—and the coefficient estimate on this indicator variable reflects the change in likelihood of having severe depressive symptoms caused by COVID. Since a variety of cutoff values are used in the psychology literature consistent with a potential diagnosis of depression, we choose to remain agnostic on this issue and proceed by considering a range of values, from 16–24.[21]

Figure 2: Coefficient estimates for different cutoff values of the binary CES-D20 measure



Note: The dashed line above/below each point estimate represents its 95% confidence interval, while the number in brackets beside it represents its bootstrapped *p*-value (10,000 replications).

The coefficient estimates are reported in Figure 2. The magnitude of the coefficient estimates invariably lie between 0.02–0.03, however, they are imprecisely estimated—only the cutoff of 22 is significant at the 5% level, which implies that COVID increased the likelihood of having severe depressive symptoms by 2.9%. Altogether, we interpret this set of results as indicating that, on average, COVID led to a significant increase in depressive symptoms (as measured by an increased CES-D score, in Table 2), but there is limited evidence that the magnitude of this increase would lead to a diagnosis of depression by health professionals.

4.1.2 Items on the CES-D scale

The CES-D scale measures the frequency and severity of symptoms and moods associated with depression across multiple dimensions by asking respondents several questions, e.g., eight for the CES-D8. We proceed by investigating how responses/scores for each question/item, rather than the aggregate CES-D score (as we reported in Table 2), were affected by COVID.

[21]While 16 is the 'traditional' cutoff for the CES-D20, following Weissman et al. (1977) and Comstock and Helsing (1977), Roberts et al. (1991) found that a cutoff as high as 24 maximises sensitivity and specificity when using the CES-D20 scale with adolescents.

The results, reported in Table 3, group these questions into similar categories as follows: columns (1) and (2) measure *mild negative emotions*, including the respondent reporting being in 'low spirits' or 'feeling sad', while columns (3) and (4) reflect *severe negative emotions*, including 'finding it difficult to do anything' or 'feeling they cannot continue with life'. The bottom half of the table features categories measuring *health behaviours*, and measures of *positive emotions*. Specifically, columns (5) and (6) measure the impact on sleep and interpersonal relationships, while columns (7) and (8) reflect positive emotions, including the frequency of respondents feeling 'happy emotions' or 'having a happy life', respectively.

Table 3: The effect of COVID on specific items from the CES-D scale

	Mild negative emotions		Severe negative emotions	
	(1) Low spirit	(2) Feel sad	(3) Everything difficult	(4) Cannot continue
Treat × Post	0.2633***	0.1130***	0.0764*	0.0145
	(0.042)	(0.040)	(0.044)	(0.026)
	[0.000]	[0.005]	[0.075]	[0.577]
Post	−0.0999***	0.0050	−0.0212	0.0052
	(0.028)	(0.026)	(0.027)	(0.016)
R-squared	0.5364	0.5325	0.5265	0.5101
R-squared (within)	0.0538	0.0268	0.0296	0.0109
	Behaviour		Positive emotions	
	(5) Poor sleep	(6) Relationships	(7) Happy mood	(8) Happy life
Treat × Post	0.0335	0.0248	−0.0609	−0.0406
	(0.045)	(0.039)	(0.051)	(0.048)
	[0.456]	[0.543]	[0.256]	[0.394]
Post	0.0596**	0.0362	0.0344	0.0609*
	(0.030)	(0.024)	(0.035)	(0.033)
R-squared	0.5185	0.5336	0.5200	0.5259
R-squared (within)	0.0242	0.0239	0.0214	0.0199
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Observations	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). The full table of coefficient estimates can be found in Appendix C Tables C2 and C3. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

The main insight from this exercise is that the decline in mental health among adolescents caused by COVID is largely attributed to a rise in mild negative emotions, particularly students reporting being in low spirits (coefficient estimate of 0.263, significant at the 1% level), and/or feeling sad (coefficient estimate of 0.113, significant at the 1% level). There is not consistently clear evidence of an increase in severe negative emotions, or adverse effects on sleep or interpersonal relations. There is, however, some evidence of a rise in more severe negative emotions, since the coefficient estimate on 'finding it difficult to do anything' is positive and significant at the 5% level. We also note that, although statistically insignificant, the signs of all other coefficient estimates are in line with our expectations (particularly negative coefficients on the measures of positive emotions).

Finally, an interesting result from this exercise relates to the causal effect of COVID on adolescent sleep quality, in column (5). Most prior studies across many countries, including China, find that COVID led to a rise in sleep disturbances for adolescents (Cai et al., 2024). Our result is not inconsistent with these findings: the positive coefficient on Post (0.060, significant at the 5% level) indicates that a comparison of students measured before and after COVID indeed reveals a rising incidence of (self-reported) poor sleep. Our empirical design, however, suggests that this may be a case of mistaken inference: on average, students in China experience increasingly poor sleep as they progress through school, even in the absence of COVID.[22]

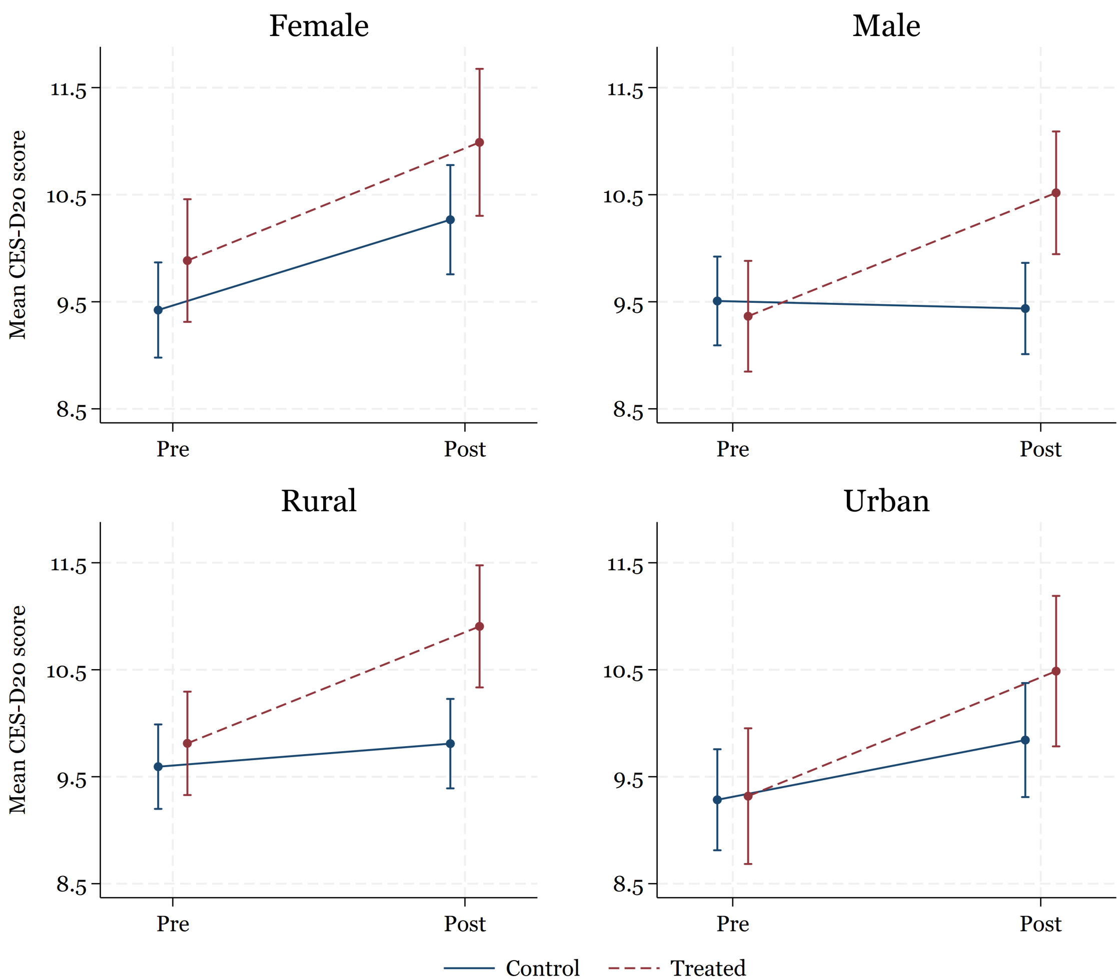
4.1.3 Analysis of differential impacts by gender and region

We find limited evidence of heterogeneity in the effect of COVID on adolescent mental health across typical policy-relevant dimensions, such as gender and region. Thus, we briefly report our findings and discuss the limitations on inference. The mean CES-D20 score and its corresponding 95% confidence interval for the treated and control cohorts, conditional on gender (i.e., boys vs girls) and region (i.e., urban vs rural), are depicted in Figure 3.

Regarding gender, Figure 3 reveals that the average level of depressive symptoms among girls in China is generally higher than boys, consistent with the overwhelming majority of prior studies across countries. Second, the figure shows us that both genders appear to have experience a similar rise (i.e., similar slope) in depressive symptoms at the onset of COVID (i.e., the 'treated' cohort, from 2018–20, in Figure 3). However, this does not provide us with a causal interpretation. Therefore, we report coefficient estimates from Equation (1), conditional on gender, in Table 4. The estimates contained in columns (1) and (2) reveal that, on average, boys experience a substantial deterioration in mental health (1.178, significant at the 1% level), while the effect on girls is not statistically significant.

[22]This is consistent with the main findings of Liang et al. (2021)—that senior high school students have a higher prevalence of sleep disturbances than those in junior high school—based on a systematic review of cross-sectional studies of sleep disturbance in Chinese adolescents.

Figure 3: Conditional mean CES-D20 score (with 95% CI), by gender/region and cohort



Note: CFPS data (2016, 2018, 2020). The top panel splits the sample by gender, while the bottom panel splits the sample by household registration type. 'Pre'/'Post' are 2016/18 for the control cohort, and 2018/20 for the treated cohort. The vertical line above/below each point estimate represents its 95% confidence interval.

Although the increase in mean CES-D20 score for the treated cohorts is similar across genders, this is clearly not the case for their control groups: Figure 3 reveals that, in the cohort prior to COVID, only girls experienced a decline in mental health, over time, in the absence of the treatment. This may provide insight into different findings across prior studies: without utilising a control group, we are effectively just comparing the slopes of the treated cohorts across genders, which Figure 3 reveals are indeed very similar.[23] While we observe a

[23]Existing studies have reached three different conclusions about gender differences in the impact of the pandemic on mental health status. Some scholars pointed out that the pandemic has a greater negative impact on the mental health status of women (Adams-Prassl et al., 2022; Butterworth et al., 2022). In contrast, Liang et al. (2020) found that men became more depressed after the epidemic, which they suggested may be related to men's

deterioration in mental health for both genders during COVID (i.e., comparing 2020 to 2018), it would be misleading to attribute the decline in mental health for girls to COVID, since our analysis reveals that, on average, girls experience a decline in mental health as they progress through school without COVID (according to the trend in the similarly-aged control cohort). We are, nonetheless, hesitant to conclude a differential effect across gender, since there is no measurable difference between these coefficient estimates (i.e., there is insufficient evidence to reject the null that these are the same).

Table 4: The effect of COVID on mental health, by gender and region

	Gender		Region	
	(1) Male	(2) Female	(3) Urban	(4) Rural
Treat × Post	1.1777***	0.5360	0.7935	1.0575**
	(0.454)	(0.541)	(0.573)	(0.438)
	[0.011]	[0.330]	[0.158]	[0.014]
Post	−0.3174	0.7339**	0.0989	0.1544
	(0.301)	(0.357)	(0.369)	(0.301)
Constant	9.9507***	8.5857***	11.5992***	8.4319***
	(1.943)	(2.180)	(2.425)	(1.886)
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.5948	0.6127	0.6188	0.5999
R-squared (within)	0.0513	0.0779	0.0738	0.0587
Observations	2,304	2,030	1,713	2,554

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

Regarding regional differences, in columns (3) and (4) of Table 4, we report coefficient estimates for Equation (1), conditional on urban and rural location, respectively. This is an important exercise, since it is reasonable to expect, *a priori*, that there may be potentially large regional differences in the mental-health response of adolescents to COVID across urban and rural areas, due to differences in the provision of medical services, policy responses, and economic factors.

The coefficient estimates reveal that, on average, the CES-D20 score of adolescents in rural areas increased by 1.058 (significant at the 5% level), but there was no statistically-significant change for students in urban areas (although the magnitude of the coefficient estimate is quite large, at 0.794). In fact, there is no measurable difference between these coefficient estimates,

coping styles and social roles. The remaining studies did not find gender differences in the effects of the epidemic (Ahmed et al., 2020; Xie et al., 2020).

which is consistent with Figure 3: while CES-D20 scores appear to be marginally higher in rural areas, the slopes are very similar across rural and urban areas, for both the treatment and control groups. Overall, we do not find compelling evidence indicating regional differences in terms of how COVID impacted the mental health of adolescents in China.

4.1.4 Alternative specifications

In this section, we briefly discuss a few important results obtained using alternative specifications to our baseline model. This includes (i) estimating cohort-specific fixed effects for overlapping students (i.e., students in both the treated and control group), and (ii) utilising a 'doubly robust' DiD estimator. In Appendix B.4, we present additional results to further demonstrate the robustness of our baseline estimates to other modelling assumptions.[24]

Overlapping students. As we described in Section 2, a subset of students appear in both the treatment and control group, since they meet the age criteria in all three waves of the CFPS. To address potential concerns that this subset of students is problematic, we allow these students to have a cohort-specific fixed effect (FE), rather than one per student (as in the baseline). This may seem conceptually unappealing if we only intend for the FE to absorb time-invariant unobservables, and hence should only have one FE per student in the study (as is the case in our baseline specification). However, allowing it to be an individual$\times$cohort FE allows it to further capture time-varying unobservables across waves for these overlapping students. This specification yields more precisely-estimated and marginally larger effect: 0.895, significant at the 1% level (Appendix Table B6, column 2).[25]

Doubly robust estimator. Although we have provided evidence that our treated and control cohorts are very similar across observables in the pre-treatment period *in levels* (Appendix Table A2), we concede that we simply cannot verify whether this extends to *trends*, due to the aforementioned data limitations in the CFPS. To improve upon this, we obtain estimates using a doubly robust DiD estimator, following Sant'Anna and Zhao (2020), which estimates the average treatment effect on the treated (ATT) by combining linear regression with propensity score matching. This method addresses potential imbalance in covariates and improves efficiency by modelling outcomes and selection simultaneously. It is 'doubly' robust in the sense that it yields a consistent ATT if either the outcome regression model for the control group is correctly specified or the propensity score model is correctly specified, but both conditions

[24] Our results are robust to including a commonly-used subjective measure of health, 'self-assessed' health, as a control (Table B9), as well as excluding several controls from our baseline specification to address concerns that they may be affected by COVID/lockdowns, e.g., extracurricular activities (Table B8).

[25] For further discussion on identification and interpretation of coefficient estimates with overlapping students, see Appendix B.3.

are not required. Compared to our baseline estimator, this approach requires a less strict 'conditional' parallel trend requirement (i.e., conditional on covariates).

The doubly-robust estimates, reported in Appendix Table B7, are a similar magnitude to our baseline estimates and estimated with similar precision (i.e., at the 5% level of significance) for both the CES-D20 and CES-D8 scale.[26] By including age, gender, and region in columns (2) to (4), respectively, we can infer how covariate imbalances along these dimensions affected our baseline estimates.[27] This exercise demonstrates, e.g., that our baseline estimate is marginally overestimated due to a gender imbalance (i.e., there are relatively more men in the control cohort than treated cohort, hence the aggregate decline in mental health in the control group is underestimated—see Figure 3).

Our preferred specification includes our full set of time-varying controls as well as gender, age, and region, and it yields an ATT of 0.918 (Table B7, column 5). Overall, there is little evidence of selection on observables—and, to the extent that it has an impact, the pre-treatment covariate imbalance observed in our sample appears to slightly attenuate our baseline estimate. Together with our baseline estimates, these additional results enable us to present a set of plausible estimates of the causal effect of COVID on the mental health of students: ranging from 0.879–0.918 for the CES-D20 (which corresponds to a 9.1–9.5% increase), and from 0.424–0.440 for the CES-D8 (i.e., an 11.3–11.7% increase).

4.2 Channels through which COVID affects adolescent mental health

We proceed with the analysis by investigating several potential mechanisms/channels behind the increase in the incidence of negative emotions among adolescents in the cohort affected by COVID. These include, (i) economic hardship, (ii) an increase in conflict in the home, (iii) a decline in physical activity, and (iv) increased internet use. The conditional mean of each mechanism (with 95% confidence intervals), by cohort, is depicted in Figure 4.[28]

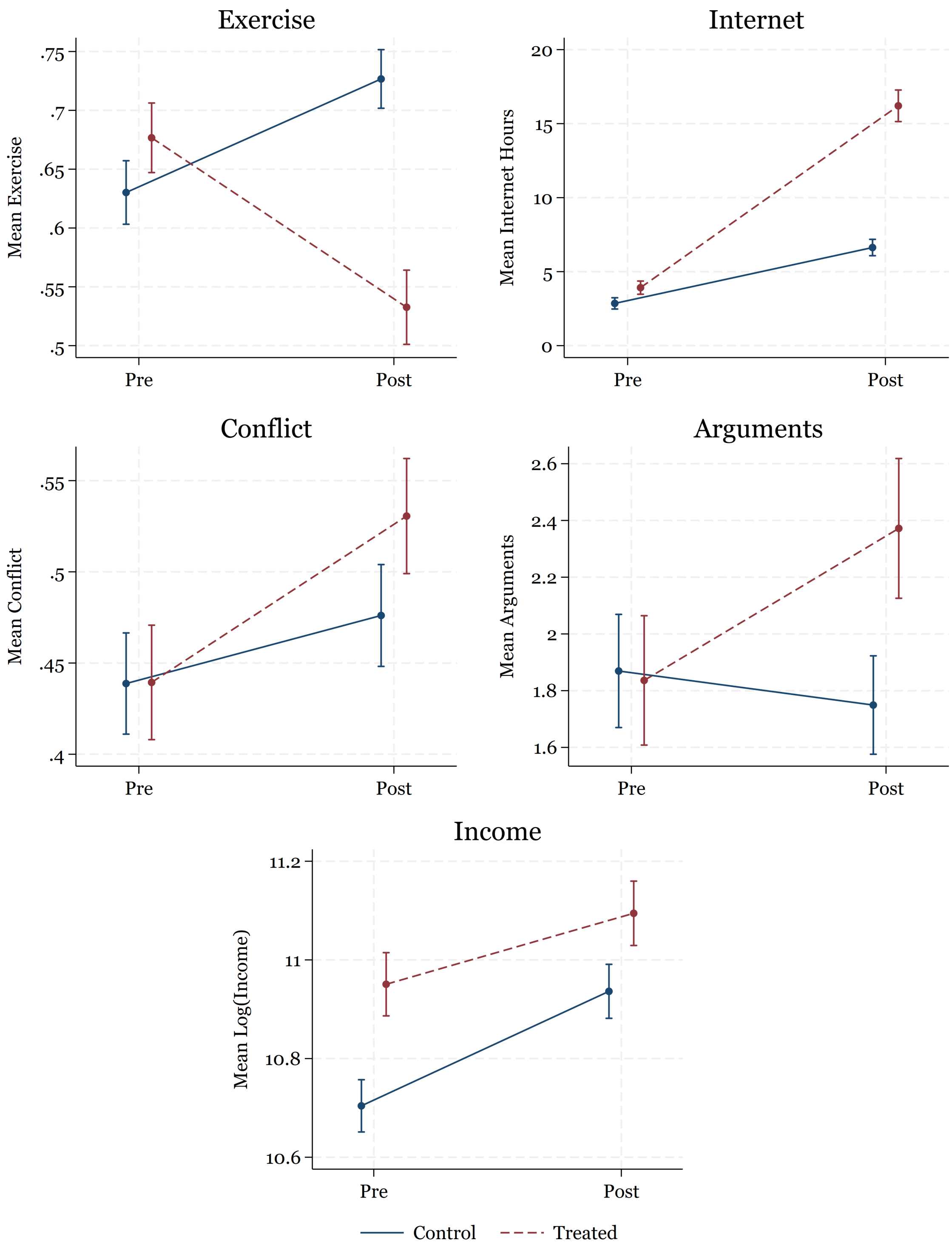
Figure 4 reveals several important points: First, the pre-treatment value of these mechanisms (in levels) was almost identical for adolescents in the treatment and control groups (with the exception of family income, which is rising over time). This echoes an important point made earlier, that we do not observe systematic differences between the control and treatment cohort along most observable dimensions in the CFPS data. Second, by comparing

[26] Since this estimator requires each individual to appear at most once in each time period, we cannot implement this with overlapping students appearing in both the treatment and control groups, as in our baseline specification. Therefore, we allow overlapping students to have a cohort-specific fixed effects, as above.

[27] The inverse probability weighting component of this estimator operates by re-weighting the control group, so that its overall distribution of covariates matches the distribution of the treated group.

[28] Throughout this section, we report changes in intra-household conflict at both the extensive margin (i.e., any conflict in the home) and intensive margin (i.e., the total number of arguments in the home).

Figure 4: Conditional mean of each mechanism/channel (with 95% CI), by cohort



Note: CFPS data (2016, 2018, 2020). Each panel refers to a mechanism variable listed in Table 1. 'Pre'/'Post' are 2016/18 for the control cohort, and 2018/20 for the treated cohort. The vertical line above/below each point estimate represents its 95% confidence interval.

slopes across the control and treatment cohorts, it appears that each of these mechanisms was impacted by COVID and the resulting change is consistent with each being a possible channel through which mental health deteriorated.

To precisely determine whether each mechanism contributes to the decline in adolescent mental health, we estimate the causal effect of COVID on each mechanism, using the linear regression model with fixed effects from Section 3; i.e., replacing the dependent variable in Equation (1) with a measure of each of the aforementioned mechanisms. As before, a significant coefficient estimate on Treat × Post from such a regression can be interpreted as the causal effect of COVID on the dependent variable. The estimates are reported in Table 5.

Table 5: The effect of COVID on the proposed mechanisms

	(1) Income	(2) Exercise	(3) Internet	(4) Conflict	(5) Arguments
Treat × Post	−0.1170**	−0.2431***	8.5346***	0.0606**	0.6789***
	(0.048)	(0.030)	(0.615)	(0.029)	(0.201)
	[0.014]	[0.000]	[0.000]	[0.040]	[0.000]
Post	0.2771***	0.0849***	1.7960***	0.0158	−0.3070**
	(0.031)	(0.019)	(0.329)	(0.020)	(0.125)
Controls	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.7334	0.5343	0.6531	0.5872	0.5809
R-squared (within)	0.0618	0.0437	0.4015	0.0274	0.0271
Observations	4,396	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). The full table of coefficient estimates can be found in Appendix C Table C4. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

Table 5 indeed confirms each of the proposed mechanisms/channels was active during COVID: on average, the growth rate of (log) family income declined, the likelihood of exercising declined by 24%, students spent 8.5 more hours on the internet (per week), the likelihood of conflict in the household increased by 6%, and the number of arguments (per month) in the household increased by 0.68 (with all of these variables significant at either the 1 or 5% levels).[29,30] Furthermore, the signs and magnitudes of these estimates indicate that any (or all) of these factors could theoretically explain why COVID adversely impacted the mental health of adolescents in China, since economic hardship, intra-household conflict,

[29] Although the negative coefficient estimate on (log) income appears to indicate that COVID reduced family incomes, it actually reflects that the average growth rate of income from 2018–20 (for the treated group) is less than that from 2016–18 (for the control group), due to the large positive significant coefficient estimate on Post.

[30] In Appendix Table B4 we report additional results for intra-household conflict using disaggregated measures: both the likelihood of conflict and arguments increased between parents as well as between parents and their children, and our findings are robust to transforming count data via logarithms, i.e., $\log(1 + \text{Arguments})$.

and internet use/screen time are negatively associated with mental health, while physical activity/exercise is positively associated with mental health (Liker and Elder Jr, 1983; Lempers et al., 1989; Taylor et al., 1985; Boers et al., 2019).

Although these mechanisms may separately (or jointly) explain the deterioration in mental health of adolescents in China, the relative contribution of each one cannot simply be inferred from the estimates reported in Table 5. However, attempting to disentangle these channels is a challenging task, particularly because they are not independent from each other. Physical exercise and internet use/screen time are substitutes for adolescents choosing to allocation their free time. Economic hardship can lead to intra-household conflict; e.g., Liker and Elder Jr (1983) and Lempers et al. (1989) documented the link between income constraints, marital stress, and conflict—between the parents, and also between parents and their children. Finally, all else equal, adolescents from relatively wealthy families are likely to have more opportunity to access the internet at home (e.g., via more devices).

Despite the difficulty disentangling these channels, by undertaking heterogeneity analysis with regards to each mechanism we can shed some light on their relative importance. We proceed along two paths: conditioning on *pre-treatment* status of each mechanism, and conditioning on *changes* in each mechanism throughout the sample period.

The former approach essentially seeks to evaluate whether groups of students, which are 'similar' in the pre-treatment period according to a given mechanism (e.g., adolescents from relatively low-income families) had measurably different mental-health responses to COVID. This exercise yields estimates with a causal interpretation: it is a 'local' average treatment effect, i.e., for students belonging to the group (e.g., low-income families).

The latter approach implements a triple-difference (DDD) estimator, conditioning on post-treatment outcomes, to evaluate whether groups of students that are 'similar' in their trend behaviour according to a given mechanism (e.g., adolescents from households that experienced a *drop* in family income). Since this approach conditions on post-treatment outcomes, which are likely impacted by COVID, we need to interpret these estimates with caution. It does not yield a causal interpretation; instead, it provides us with evidence of (conditional) heterogeneity across cohorts. It enables us to, e.g., determine whether average changes in mental health associated with a change in a particular mechanism (e.g., income drop) were larger during COVID, compared to similar changes in the same mechanism for the earlier cohort.

4.2.1 Heterogeneity by pre-treatment levels

We define groups based on pre-treatment status as follows: (i) for continuous variables (i.e income and internet use), we split the sample into those above/below the median level in their respective pre-treatment period; e.g., for income, we identify adolescents from households with relatively low (i.e., below-median) family income (where the median is calculated in 2016 for the control group, and 2018 for the treated group); (ii) for binary variables (i.e., conflict and exercise), we simply condition on either outcome; e.g., no intra-household conflict in the pre-treatment period. The coefficient estimates are reported in Table 6.

Family income/economic hardship. The adverse effect of COVID on mental health was clearly concentrated among relatively high-income households: the effect is small (0.162) and statistically insignificant for adolescents from below-median family income households, while it is large (1.746) and significant at the 1% level for adolescents from households with above-median family income.[31] These estimates are significantly different (p-value $= 0.020$).

While pre-treatment income is a powerful predictor of the types of households with adolescents experiencing a substantial deterioration of mental health during COVID, it does not appear to measure economic hardship. We don't find evidence that high-income households were more susceptible to economic hardship (e.g., large income drops).[32]

Instead, we find that high (pre) income coincides with large changes in other channels. On average, this group had a smaller reduction in the likelihood of exercise (-0.216 vs -0.273, both significant at the 1% level), but a larger increase in both internet use (9.41 vs 7.88, both significant at the 1% level) and the number of arguments (0.753 vs 0.614, significant at the 1% and 5% level, respectively). In each case, however, we cannot reject the null that these estimates are the same across groups. Nonetheless, given the stark differences in mental-health changes across groups, we take this as suggestive evidence of the relative importance of internet use and intra-household conflict.

Physical exercise. Physical exercise provides the least clear evidence of a difference between mental health changes across the two groups (by pre-treatment status). In this case, we are conditioning on students that did not regularly engage in physical exercise in the pre-treatment period (i.e., less than 1 time per week) vs those that did, respectively. The estimates are almost identical in magnitude, across groups, however only the estimate for those adolescents that previously exercised is precisely estimated (i.e., statistically significant). Moreover, we cannot reject the null that there is no difference in coefficient estimates across subgroups.

[31] This finding is robust to alternative definitions of median income; e.g., conditional on both year and location.

[32] Households with high-income were marginally more likely to experience an income drop, however, (i) this was similarly observed in both the treatment and control cohorts, and (ii) conditional on having income fall, both low- and high-income households experienced a similar drop (approximately 34% of pre-period income).

Table 6: The effect of COVID on depressive symptoms, by mechanism (pre-treatment)

	Income		Exercise	
	(1) Low	(2) High	(3) No exercise	(4) Prior exercise
Treat × Post	0.1640	1.7441***	0.9073	0.9006**
	(0.504)	(0.468)	(0.604)	(0.411)
	[0.753]	[0.000]	[0.138]	[0.022]
Post	0.5027	-0.3505	0.2928	0.0580
	(0.334)	(0.302)	(0.387)	(0.281)
Constant	10.6094***	8.2386***	8.0609***	10.7265***
	(2.160)	(1.882)	(2.524)	(1.705)
R-squared	0.6020	0.6227	0.6033	0.6248
R-squared (within)	0.0546	0.0641	0.0662	0.0529
Observations	2,204	2,192	1,536	2,860
	Internet		Conflict	
	(5) Low	(6) High	(7) No prior conflict	(8) Prior conflict
Treat × Post	0.6033	1.2367**	0.5714	1.3054**
	(0.457)	(0.509)	(0.422)	(0.557)
	[0.184]	[0.013]	[0.177]	[0.023]
Post	−0.3802	0.3137	0.6069**	−0.3714
	(0.355)	(0.332)	(0.285)	(0.359)
Constant	10.0304***	9.1539***	8.8544***	11.7308***
	(1.829)	(2.436)	(1.669)	(2.454)
R-squared	0.6014	0.6292	0.6177	0.6093
R-squared (within)	0.0373	0.0839	0.0535	0.0605
Observations	2,486	1,910	2,466	1,930
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes

Note: CFPS data (2016, 2018, 2020). The dependent variable in all specifications is the CES-D20 score. For each category (income, exercise, etc.), the sample is split into two mutually-exclusive groups based on pre-treatment values. The full table of coefficient estimates can be found in Appendix C Tables C5 and C6. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

While we cannot distinguish across groups in terms of the change in CES-D score, we can along other dimensions. We find evidence that those students engaging in prior exercise substituted their time from exercise (which dropped) to internet use/screen time, thus experiencing a greater rise in hours of internet use per week (9.179 vs 7.079, both significant at the 1% level). It is perhaps unsurprising that this relative rise in internet hours does not translate into more severe depressive symptoms for this group of adolescents, since it is only an additional 2 hours per week, on average. Nevertheless, considering this group experienced both a fall in exercise combined with a marginal rise in internet hours, on average, we find it informative that there appears to be little difference across these groups in terms of the average impact on mental health, and thus consider this as evidence against physical exercise.

Internet use/screen time. The adverse effect of COVID on mental health was concentrated among adolescents previously engaging in relatively high amounts of internet use/screen time. They experienced an average increase in CES-D20 score of 1.237 (significant at 5% level) compared with a 0.603 (insignificant) increase for those with relatively low internet usage. However, we cannot reject the null hypothesis that there is no difference in coefficient estimates across subgroups.

To investigate this further, we conducted additional heterogeneity analysis in terms of availability of devices in the home to access the internet (Appendix Table B3). Students engaging in prior use of either a mobile phone/tablet or computer to access the internet experienced a large rise in CES-D20 score of 1.056 (significant at the 5% level), compared to those without access to these devices (0.505, insignificant). Furthermore, when we condition on the subset of students that never use a mobile phone (throughout the entire sample period) vs those that do (at any point in the sample period), we find very stark differences.[33] Students that never use a phone have an insignificant change in depressive symptoms (-0.4), while those that do (at any point in the sample period) have a significant increase (1.134, significant at 1% level). Additionally, those that always have a phone throughout the entire sample period have an even larger increase (1.394, significant at 5% level).

Intra-household conflict. Finally, we investigate the impact of 'prior' conflict in the home by conditioning on adolescents with no intra-household conflict in the pre-treatment period vs conflict, respectively. The idea here is twofold: (i) if the rise in conflict is driven by economic hardship caused by COVID, then prior conflict should not predict mental health changes caused by COVID (insofar as prior conflict is not correlated with economic changes caused by COVID); (ii) prior conflict ought to be a good predictor of conflict in lockdown, i.e., forcing people who have conflict to spend more time together in close quarters should presumably lead to more (not less) conflict, all else equal. These results in Table 6 indicate that COVID predominately affected the mental health of children in families where conflict was already present in the home: 1.305 (significant at 5% level) compared with 0.571 (not significant).[34] However, again, we cannot reject that there is no difference across subgroups.

[33] We acknowledge that this latter exercise is potentially problematic if COVID affected mobile-phone/tablet take-up rates. However, the increase in internet access from mobile phones/tablets grows at an almost constant rate in both the control and treated cohorts: the proportion in each wave is 0.35 (2016), 0.57 (2018), 0.77 (2020). Thus, we do not find evidence that COVID affected mobile-phone access among adolescents.

[34] This pattern also holds when using disaggregated conflict measures (Appendix Table B5).

Interestingly, this connects with our results obtained from separating students according to their family-income levels: family income is generally higher in families with conflict occurring in the home (during the 'pre' period), compared to those with no conflict. Specifically, mean (log) family income was higher across all periods in the sample, as well as higher in the pre-treatment period (10.882 vs 10.743) and the post-treatment period (11.076 vs 10.930). Furthermore, this relationship also holds using other measures ranking income across groups, e.g., the median, 25%, and 10% percentile of the income distribution. Therefore, using family income as a measure of economic distress, our evidence suggests that households with no prior conflict were more likely to be subject to greater economic hardship (e.g., tighter budget constraints) than those with prior conflict—yet we observe a relatively substantial deterioration in mental health for children from relatively high-income households.

4.2.2 Heterogeneity by differences

We now consider the following changes to each mechanism: (i) an income drop of at least 50%, (ii) a drop in exercise (i.e., from exercising at least once per week to exercising less than once per week), (iii) a rise in hours of internet use per month, (iv) a rise in the number of arguments between children and parents per month (conditional on prior conflict in the household). The results are reported in Table 7, where 'Indicator' refers to the binary indicator constructed to measure the change in each mechanism. The DDD term is the interaction of DiD term with 'Indicator' (i.e., directly in the row below the DiD term, Treat $\times$ Post).

Family income/economic hardship. Our preferred (binary) measure for economic hardship is a drop in family income of at least 50% (Table 7, column 1).[35] The results indicate that economic hardship was associated with a marginally greater increase in depressive symptoms during the COVID period (2.029, significant at the 10% level), over and above the average COVID effect (0.699, significant at the 5% level). There was no significant association between income drops and depressive symptoms outside the COVID context, nor a period effect in the absence of COVID. These results suggest that a potentially heightened vulnerability among adolescents in households experiencing severe income shocks during the pandemic, though the evidence for heterogeneity is only marginally significant. Additionally, since the DDD term effectively separates households experiencing economic hardship from those that did not, we interpret the positive and significant DiD coefficient (0.699) as providing supportive evidence for the *other* channels during COVID (i.e., evidence of a post-COVID rise in depressive symptoms in households not experiencing economic hardship).

Physical exercise. Since our measure for exercise is binary (i.e., an indicator for exercising

[35] Approximately 8.7% of observations experienced such an income drop, so power may be limited.

Table 7: The effect of COVID on depressive symptoms, by changes in mechanisms

	(1) Income drop	(2) Exercise drop	(3) Internet rise	(4) Arguments rise
Treat × Post	0.6989**	0.9468**	0.0800	0.6352*
	(0.356)	(0.390)	(0.570)	(0.350)
	[0.051]	[0.015]	[0.888]	[0.070]
× Indicator	2.0287*	0.0186	1.0299	3.0179**
	(1.135)	(0.840)	(0.686)	(1.292)
	[0.076]	[0.983]	[0.138]	[0.022]
Post	0.1763	0.1853	−0.0988	0.1342
	(0.231)	(0.235)	(0.297)	(0.230)
× Indicator	−0.4089	−0.4013	0.6924	0.0276
	(0.703)	(0.662)	(0.438)	(0.796)
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.5972	0.5966	0.5981	0.5983
R-squared (within)	0.0500	0.0487	0.0522	0.0527
Observations	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). The dependent variable in all specifications is the CES-D20 score. The 'Indicator' variable in each specification is defined accordingly for the mechanism labelling each column: 'Income drop' equals one if family income dropped by at least 50%, 'Exercise drop' equals one if exercise reduced from at least once per week to less than once per week, 'Internet rise' equals one if a student increased their internet hours (per month), and 'Arguments rise' equals one if a student experienced a rise in the number of child-parent arguments (in the past month, conditional on having conflict in the 'pre' period). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD and DDD terms.

at least once per week), unfortunately we are prevented from investigating changes on the intensive margin. Therefore, in column (2) of Table 7, our measure of an exercise drop essentially reflects an extensive-margin adjustment (i.e., from exercising at least once per week to not). Nonetheless, this specification reveals that COVID was associated with a large increase in depressive symptoms (0.947, significant at the 5% level), for adolescents that did not stop exercising. Furthermore, since the DDD term is almost zero (0.019) and not significant, we find no evidence for differences in changes in depressive symptoms based on whether students continued exercising or stopped exercising. This appears to be compelling evidence against the physical-exercise mechanism.

Internet use/screen time. To study changes in internet usage, we construct a binary indicator for an increase in the hours of internet use (per month). The estimates in column (3) demonstrate that there is no statistically significant association between COVID and depressive symptoms for either group defined by internet hours change, nor was there evidence of a differential effect by change in internet hours. The absence of a significant interaction here is challenging to interpret due to a lack of precision: there may be meaningful subgroup differences, since the magnitude of the coefficient estimate is relatively large (1.030).

Intra-household conflict. Our preferred measure for changes in conflict over time is a binary variable indicating increases in the number of arguments between children and parents, conditional on pre-existing child-parent conflict in the household; i.e., it indicates intensive-margin changes for conflict between children and parents (Table 7, column 4).

A rise in child-parent conflict was associated with a significantly greater increase in depressive symptoms during COVID (3.018, significant at 5%), over and above the average COVID effect (0.635, significant at 10% level). It indicates that increases in intra-household conflict between children and parents uniquely intensified the negative mental-health impact of COVID among adolescents. While this is a stark finding, and relatively precisely estimated, we acknowledge that this measure reflects a small fraction of our sample (approximately 8%).

We do not find evidence of a similar effect for adolescents experiencing a rise in other types of intra-household conflict; i.e., parent-parent conflict (with or without prior conflict), or child-parent conflict (without prior conflict). Therefore, we conclude that the most vulnerable adolescents during lockdown were those with a history of conflict with their parents.

While the preceding analysis uncovered evidence that income drops and rising conflict may explain the rise in depressive symptoms for a small fraction of households; interestingly, we do not find evidence that the rise in conflict was a result of economic hardship, since there is little overlap between the households receiving these types of 'shocks' in either the control or treated cohorts. Specifically, 15.9% (15.5%) of households in the control (treated) cohort experienced either shock (i.e., an income drop, conflict rise, or both). Conditioning on this set of households, 98% (95%) received only one shock in the control (treated) cohort.

Additionally, we do not find evidence that households with pre-existing conflict were more likely to suffer economic hardship: estimating Equation (1) with the binary indicator for an income drop of at least 50% as the dependent variable, conditional on the subset of households with pre-existing conflict, yields extremely-low magnitude and insignificant estimates for the DiD term. This obtains for all measures of pre-existing conflict (i.e., aggregate, child-parent, and parent-parent conflict).

5 Concluding remarks

In this paper, we investigated the short-term causal effects of COVID on adolescent mental health in China using longitudinal data from the CFPS. Employing a rigorous cohort-DiD methodology, we identified a substantial deterioration in adolescent mental health, with

CES-D scores increasing by 9–11%, on average. This finding contrasts starkly with earlier regional studies, such as Shi et al. (2022), Zhao et al. (2023), and Zeng et al. (2024), which reported minimal or even positive mental-health trends during the same period in China.

A key feature of our study involved identifying several mechanisms potentially driving the observed mental-health decline: a rise in economic hardship (measured by changes in family income), increased conflict in the home (between children and parents, as well as between the parents), a decline in physical activity, and an increase in internet hours.

We provide novel evidence that intra-household conflict is a critical factor in China, since both the frequency and likelihood of conflict within households increased because of COVID. This is not an entirely new phenomenon, since many international studies have reported increases in domestic violence post COVID.[36] However, in the majority of these studies, the rise in conflict/violence has been attributed to economic hardship experienced by the family after the COVID outbreak—a well understood relationship, since Liker and Elder Jr (1983). In contrast, we did not find evidence of intra-household conflict being associated with economic hardship (before or after COVID) in China.

In the absence of such a link, we contend that the rise in arguments was most plausibly caused by exposure due to lockdown, since virtually all adolescents in China were subject to lockdown in 2020, and the adverse mental-health consequences were concentrated among families that already engaged in conflict prior to the pandemic. This aligns with several other studies arguing that exposure was the key mechanism underlying rises in intra-household conflict/violence during COVID, including Hsu and Henke (2021) (US), Agüero (2021) (Peru), and Ivandic et al. (2021) (UK).[37]

While we do not have such detailed measures of conflict (e.g., disaggregated by abuse type, as in Arenas-Arroyo et al. (2021)) as some other past studies, we are one of a few studies that measures the impact of the rise in conflict in the home on the mental health of adolescents. To the best of our knowledge, Baranov et al. (2022) (Pakistan) is the only other such study, and they attribute rising parental stress and domestic violence to economic hardship, rather than exposure.

Since our findings point toward exposure as the primary mechanism leading to increased intra-household conflict, this provides evidence to assist policymakers, both in terms of designing health interventions focusing on family counselling or conflict mediation during lockdowns, as well as informing cost-benefit analyses of crisis-induced lockdowns—even in the absence of economic hardship (e.g., when remote work is feasible).

[36] For a survey of conflict behaviour during COVID, see Chowdhury and Karmakar (2024).

[37] Prior to COVID, Chin (2012) also finds evidence supporting exposure *reduction* as the primarily factor influencing spousal violence in India.

References

ADAMS-PRASSL, A., T. BONEVA, M. GOLIN, AND C. RAUH (2022): "The impact of the coronavirus lockdown on mental health: evidence from the United States," *Economic Policy*, 37, 139–155.

AGÜERO, J. M. (2021): "COVID-19 and the rise of intimate partner violence," *World Development*, 137, 105217.

AHMED, M. Z., O. AHMED, Z. AIBAO, S. HANBIN, L. SIYU, AND A. AHMAD (2020): "Epidemic of COVID-19 in China and associated psychological problems," *Asian Journal of Psychiatry*, 51, 102092.

AIZER, A. (2010): "The gender wage gap and domestic violence," *American Economic Review*, 100, 1847–1859.

ARENAS-ARROYO, E., D. FERNANDEZ-KRANZ, AND N. NOLLENBERGER (2021): "Intimate partner violence under forced cohabitation and economic stress: Evidence from the COVID-19 pandemic," *Journal of Public Economics*, 194, 104350.

AU, N. AND D. W. JOHNSTON (2014): "Self-assessed health: what does it mean and what does it hide?" *Social Science & Medicine*, 121, 21–28.

BARANOV, V., P. GROSJEAN, F. J. KHAN, AND S. WALKER (2022): "The impact of COVID-Related economic shocks on household mental health in Pakistan," *Health Economics*, 31, 2208–2228.

BHALOTRA, S., E. BRITO, D. CLARKE, P. LARROULET, AND F. PINO (2024): "Dynamic impacts of lockdown on domestic violence: Evidence from multiple policy shifts in Chile," *Review of Economics and Statistics*, 1–29.

BOERS, E., M. H. AFZALI, N. NEWTON, AND P. CONROD (2019): "Association of screen time and depression in adolescence," *JAMA Pediatrics*, 173, 853–859.

BUTTERWORTH, P., S. SCHURER, T.-A. TRINH, E. VERA-TOSCANO, AND M. WOODEN (2022): "Effect of lockdown on mental health in Australia: evidence from a natural experiment analysing a longitudinal probability sample survey," *The Lancet Public Health*, 7, e427–e436.

CAI, H., P. CHEN, Y. JIN, Q. ZHANG, T. CHEUNG, C. H. NG, Y.-T. XIANG, AND Y. FENG (2024): "Prevalence of sleep disturbances in children and adolescents during COVID-19 pandemic: a meta-analysis and systematic review of epidemiological surveys," *Translational Psychiatry*, 14, 12.

CHEN, B., C. ZHAO, X. LI, AND J. LIU (2022): "COVID-19 pandemic and adolescent mental health in China: Micro evidence and socioeconomic mechanisms," *Frontiers in Psychology*, 13, 1041376.

CHEN, F., D. ZHENG, J. LIU, Y. GONG, Z. GUAN, AND D. LOU (2020): "Depression and anxiety among adolescents during COVID-19: A cross-sectional study," *Brain, Behavior, and Immunity*, 88, 36.

CHEUNG, C.-K. AND C. BAGLEY (1998): "Validating an American scale in Hong Kong: the center for epidemiological studies depression scale (CES-D)," *The Journal of Psychology*, 132, 169–186.

CHIN, Y.-M. (2012): "Male backlash, bargaining, or exposure reduction?: women's working status and physical spousal violence in India," *Journal of Population Economics*, 25, 175–200.

CHOWDHURY, S. M. AND S. KARMAKAR (2024): "The interrelationship between the COVID-19 pandemic and conflict behavior: A survey," *Journal of Economic Surveys*, 38, 1473–1489.

COLELLA, S., F. DUFOURT, V. A. HILDEBRAND, AND R. VIVÈS (2023): "Mental health effects of COVID-19 lockdowns: A Twitter-based analysis," *Economics & Human Biology*, 51, 101307.

Comstock, G. W. and K. J. Helsing (1977): "Symptoms of depression in two communities," *Psychological Medicine*, 6, 551–563.

Crossley, T. F. and S. Kennedy (2002): "The reliability of self-assessed health status," *Journal of Health Economics*, 21, 643–658.

Deng, J., F. Zhou, W. Hou, K. Heybati, S. Lohit, U. Abbas, Z. Silver, C. Y. Wong, O. Chang, E. Huang, et al. (2023): "Prevalence of mental health symptoms in children and adolescents during the COVID-19 pandemic: A meta-analysis," *Annals of the New York Academy of Sciences*, 1520, 53–73.

Doiron, D., D. G. Fiebig, M. Johar, and A. Suziedelyte (2015): "Does self-assessed health measure health?" *Applied Economics*, 47, 180–194.

Dugan, L., D. S. Nagin, and R. Rosenfeld (1999): "Explaining the decline in intimate partner homicide: The effects of changing domesticity, women's status, and domestic violence resources," *Homicide Studies*, 3, 187–214.

——— (2003): "Exposure reduction or retaliation? The effects of domestic violence resources on intimate-partner homicide," *Law & Society Review*, 37, 169–198.

Guariso, A. and M. B. Nyqvist (2023): "The impact of the COVID-19 pandemic on children's learning and wellbeing: Evidence from India," *Journal of Development Economics*, 164, 103133.

Hsu, L.-C. and A. Henke (2021): "COVID-19, staying at home, and domestic violence," *Review of Economics of the Household*, 19, 145–155.

Ivandic, R., T. Kirchmaier, and B. Linton (2021): "The role of exposure in domestic abuse victimization: Evidence from the COVID-19 lockdown," *Available at SSRN 3686873*.

Kahn-Lang, A. and K. Lang (2020): "The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications," *Journal of Business & Economic Statistics*, 38, 613–620.

Lempers, J. D., D. Clark-Lempers, and R. L. Simons (1989): "Economic hardship, parenting, and distress in adolescence," *Child Development*, 25–39.

Li, J.-Y., J. Li, J.-H. Liang, S. Qian, R.-X. Jia, Y.-Q. Wang, and Y. Xu (2019): "Depressive symptoms among children and adolescents in China: a systematic review and meta-analysis," *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 25, 7459.

Liang, L., H. Ren, R. Cao, Y. Hu, Z. Qin, C. Li, and S. Mei (2020): "The effect of COVID-19 on youth mental health," *Psychiatric Quarterly*, 91, 841–852.

Liang, M., L. Guo, J. Huo, and G. Zhou (2021): "Prevalence of sleep disturbances in Chinese adolescents: A systematic review and meta-analysis," *Plos One*, 16, e0247333.

Liker, J. K. and G. H. Elder Jr (1983): "Economic hardship and marital relations in the 1930s," *American Sociological Review*, 343–359.

Liu, Y., E. Zhang, H. Li, X. Ge, F. Hu, Y. Cai, and M. Xiang (2024): "Physical activity, recreational screen time, and depressive symptoms among Chinese children and adolescents: a three-wave cross-lagged study during the COVID-19 pandemic," *Child and Adolescent Psychiatry and Mental Health*, 18, 11.

MA, L., M. MAZIDI, K. LI, Y. LI, S. CHEN, R. KIRWAN, H. ZHOU, N. YAN, A. RAHMAN, W. WANG, ET AL. (2021): "Prevalence of mental health problems among children and adolescents during the COVID-19 pandemic: A systematic review and meta-analysis," *Journal of Affective Disorders*, 293, 78–89.

MACMILLAN, R. AND R. GARTNER (1999): "When she brings home the bacon: Labor-force participation and the risk of spousal violence against women," *Journal of Marriage and the Family*, 947–958.

MADIGAN, S., N. RACINE, T. VAILLANCOURT, D. J. KORCZAK, J. M. HEWITT, P. PADOR, J. L. PARK, B. A. MCARTHUR, C. HOLY, AND R. D. NEVILLE (2023): "Changes in depression and anxiety among children and adolescents from before to during the COVID-19 pandemic: a systematic review and meta-analysis," *JAMA Pediatrics*.

MIAO, R., C. LIU, J. ZHANG, AND H. JIN (2023): "Impact of the COVID-19 pandemic on the mental health of children and adolescents: a systematic review and meta-analysis of longitudinal studies," *Journal of Affective Disorders*, 340, 914–922.

PANCHAL, U., G. SALAZAR DE PABLO, M. FRANCO, C. MORENO, M. PARELLADA, C. ARANGO, AND P. FUSAR-POLI (2023): "The impact of COVID-19 lockdown on child and adolescent mental health: systematic review," *European Child & Adolescent Psychiatry*, 32, 1151–1177.

PRATI, G. AND A. D. MANCINI (2021): "The psychological impact of COVID-19 pandemic lockdowns: a review and meta-analysis of longitudinal studies and natural experiments," *Psychological Medicine*, 51, 201–211.

RACINE, N., B. A. MCARTHUR, J. E. COOKE, R. EIRICH, J. ZHU, AND S. MADIGAN (2021): "Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis," *JAMA Pediatrics*, 175, 1142–1150.

RADLOFF, L. S. (1977): "The CES-D scale: A self-report depression scale for research in the general population," *Applied Psychological Measurement*, 1, 385–401.

ROBERTS, R. E., P. M. LEWINSOHN, AND J. R. SEELEY (1991): "Screening for adolescent depression: A comparison of depression scales," *Journal of the American Academy of Child & Adolescent Psychiatry*, 30, 58–66.

ROBINSON, E., A. R. SUTIN, M. DALY, AND A. JONES (2022): "A systematic review and meta-analysis of longitudinal cohort studies comparing mental health before versus during the COVID-19 pandemic in 2020," *Journal of Affective Disorders*, 296, 567–576.

SANT'ANNA, P. H. AND J. ZHAO (2020): "Doubly robust difference-in-differences estimators," *Journal of Econometrics*, 219, 101–122.

SERRANO-ALARCÓN, M., A. KENTIKELENIS, M. MCKEE, AND D. STUCKLER (2022): "Impact of COVID-19 lockdowns on mental health: Evidence from a quasi-natural experiment in England and Scotland," *Health Economics*, 31, 284–296.

SHI, W., L. ZHAO, M. LIU, B. HONG, L. JIANG, AND P. JIA (2022): "Resilience and mental health: a longitudinal cohort study of Chinese adolescents before and during COVID-19," *Frontiers in Psychiatry*, 13, 948036.

TANG, X., S. TANG, Z. REN, AND D. F. K. WONG (2019): "Prevalence of depressive symptoms among adolescents in secondary school in mainland China: A systematic review and meta-analysis," *Journal of Affective Disorders*, 245, 498–507.

Taylor, C. B., J. F. Sallis, and R. Needle (1985): "The relation of physical activity and exercise to mental health." *Public Health Reports*, 100, 195.

Wang, D., J. Zhao, B. Ross, Z. Ma, J. Zhang, F. Fan, and X. Liu (2022): "Longitudinal trajectories of depression and anxiety among adolescents during COVID-19 lockdown in China," *Journal of Affective Disorders*, 299, 628–635.

Weissman, M. M., D. Sholomskas, M. Pottenger, B. A. Prusoff, and B. Z. Locke (1977): "Assessing depressive symptoms in five psychiatric populations: a validation study," *American Journal of Epidemiology*, 106, 203–214.

Wolf, K. and J. Schmitz (2024): "Scoping review: longitudinal effects of the COVID-19 pandemic on child and adolescent mental health," *European Child & Adolescent Psychiatry*, 33, 1257–1312.

Xie, X., Q. Xue, Y. Zhou, K. Zhu, Q. Liu, J. Zhang, and R. Song (2020): "Mental health status among children in home confinement during the coronavirus disease 2019 outbreak in Hubei Province, China," *JAMA Pediatrics*, 174, 898–900.

Zeng, Y., J. Song, Y. Zhang, X. Guo, X. Xu, L. Fan, L. Zhao, H. Song, and L. Jiang (2024): "Life changes and symptoms of depression and anxiety among Chinese children and adolescents before, during, and after the COVID-19 pandemic lockdown: a combination of cross-sectional, longitudinal, and clustering studies," *European Child & Adolescent Psychiatry*, 1–14.

Zhao, L., X. Li, Q. Yang, Y. Peng, L. Jiang, P. Jia, and W. Shi (2023): "The longitudinal association between internet addiction and depressive and anxiety symptoms among Chinese adolescents before and during the COVID-19 pandemic," *Frontiers in Public Health*, 10, 1096660.

APPENDIX

This appendix contains supplementary materials for *'Lockdown, Family Conflict, and Adolescent Mental Health in China'*. It has the following structure:

A. Data and summary statistics

B. Additional results and discussion

C. Full tables of coefficient estimates

Appendix A Data and summary statistics

Table A1: Variable names and definitions in the CFPS

Variable	Definition
Dependent variable	
CES-D20 score	Score calculated using the 'long-form' CES-D depression scale, ranges from 0–60.
CES-D8 score	Score calculated using the 'short-form' CES-D depression scale, ranges from 0–24.
Independent variables	
Treat	Dummy variable for the cohort affected by the COVID outbreak. 1= treatment group, i.e., individual i meets the age requirement and responds in both 2018 and 2020. 0= control group, i.e., individual i meets the age requirement and responds in both 2016 and 2018.
Post	Dummy variable for before/after period. 1= after event period, i.e. 2020 in treatment group and 2018 in control group. 0= before event period, i.e. 2018 in treatment group and 2016 in control group.
Treat $\times$ Post	Interaction term of the above two dummy variables. Used to measure the causal effect of COVID (and quarantine and lockdown measures) on dependent variable.
Control variables	
Attend key school	Dummy variable: 1 = student attends a 'key' school. (Note: refers to a highly-regarded class of schools)
Extracurricular class hours	Hours the student attended extracurricular classes every week during the recent month (that is not a summer/winter vacation).
Hours studying (weekdays)	Hours the student spends studying on workdays, in general.
Hours studying (weekends)	Hours the student spends studying on weekends, in general.
Student leader	Dummy variable: 1= student served as a leader of the class or school in the last semester.
Student performance	Self-rated degree of satisfaction with academic performance; from 'very dissatisfied' (1) to 'very satisfied' (5).
Study pressure	Self-rated degree of pressure on study; from 'no pressure' (1) to 'a lot of pressure' (5).

Table A1 (cont.): Variable names and definitions in the CFPS

Variable	**Definition**
Control variables (cont.)	
Excellence	Self-rated degree of excellence as a student; from 'very bad' (1) to 'very good' (5).
Satisfaction (school)	How satisfied the student is with the school; from 'very dissatisfied' (1) to 'very satisfied' (5).
Satisfaction (teacher)	How satisfied the student is with the teacher in charge of his/her class; from (1) 'very dissatisfied' to 'very satisfied' (5).
Pocket money	Dummy variable: 1 = receives pocket money.
Internet (mobile phone)	Dummy variable: 1= uses a mobile device, such as a phone or tablet, to access the internet.
Internet (computer)	Dummy variable: 1= uses a computer to access the internet.
Sick (past month)	Dummy variable: 1= sick in the past month.
Sick (past year)	Dummy variable: 1= visit a hospital or medical facility in the past 12 months due to illness.
OOP medical expenses (log)	The log value of the amount of money the family paid out of pocket (OOP) for this child's medical expenses in the past 12 months.
Father lives at home	Dummy variable: 1= father lives at home.
Mother lives at home	Dummy variable: 1= mother lives at home.
Family size	Number of family members.
Mechanism variables	
Family income (log)	The log value of total annual family income.
Physical exercise	Dummy variable: 1= exercises at least once a week.
Internet use (hours)	The total number of hours a student reports spending online each week.
Conflict in the home	Dummy variable: 1= at least one argument in the past month between either the child and parents or between the parents (observed by the child).
Total number of arguments	The sum of the number of arguments in the past month between (i) the child and parents and (ii) between the parents (observed by the child).

Note: CFPS data (2016, 2018, 2020).

Table A2: Descriptive statistics in the pre-treatment period

	Control group			**Treated group**		
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
Dependent variable						
CES-D20 score	9.469	5.417	1,233	9.618	6.087	965
CES-D8 score	3.680	2.758	1,233	3.755	3.088	965
Control variables						
Attend key school	0.193	0.395	1,233	0.225	0.418	965
Extracurricular class hours	1.593	5.222	1,233	3.004	6.804	965
Hours studying (weekday)	7.962	3.436	1,233	8.016	2.589	965
Hours studying (weekend)	3.410	2.701	1,233	3.551	2.681	965
Student leader	0.346	0.476	1,233	0.342	0.475	965
Student performance	3.402	0.961	1,233	3.396	0.953	965
Study pressure	2.796	1.124	1,233	2.776	1.163	965
Excellence	3.158	0.870	1,233	3.188	0.847	965
Satisfaction (school)	4.119	0.952	1,233	4.221	0.977	965
Satisfaction (teacher)	4.351	0.926	1,233	4.372	1.031	965
Pocket money	0.745	0.436	1,233	0.739	0.439	965
Internet (mobile phone/tablet)	0.348	0.477	1,233	0.476	0.500	965
Internet (computer)	0.297	0.457	1,233	0.247	0.431	965
Sick (past month)	0.194	0.395	1,233	0.176	0.381	965
Sick (past year)	0.278	0.448	1,233	0.302	0.459	965
OOP medical expenses (log)	3.911	2.664	1,233	4.044	2.757	965
Father lives at home	0.815	0.388	1,233	0.787	0.410	965
Mother lives at home	0.856	0.352	1,233	0.846	0.362	965
Family size	5.097	1.720	1,233	5.193	1.800	965
Mechanism variables						
Family income (log)	10.704	0.948	1,233	10.951	1.014	965
Physical exercise	0.630	0.483	1,233	0.677	0.468	965
Internet use (hours)	2.855	6.817	1,233	3.919	7.067	965
Conflict in the home	0.439	0.496	1,233	0.439	0.497	965
Total number of arguments	1.869	3.569	1,233	1.836	3.609	965

Note: CFPS data (2016, 2018, 2020). The pre-treatment period for the control group is 2016, and for the treated group it is 2018.

Appendix B Additional results and discussion

B.1 Outlier values

Several variables in the CFPS data appear to take implausibly large values for some students, even after repeated confirmation questions by the CFPS surveyors (designed specifically to confirm validity of such implausible responses). These variables are: Hours studying (weekdays), hours studying (weekends), family size, internet use, and total number of arguments. To address this, we winsorized these variables at the 99$^{\text{th}}$ percentile in our baseline model.

To investigate whether our results are sensitive to this choice, we obtained additional estimates where we: (1) do not winsorize (i.e., we do not adjust the data for these outliers); (2) winsorize each variable at the 95$^{\text{th}}$ percentile; (3) take logs of each of these variables (instead of keeping them in levels, as in our baseline). These results are reported in Table B1.

Table B1: The effect of COVID on mental health, considering outliers

	Winsorizing–99%		Winsorizing–95%		No winsorizing	
	(1) Levels	(2) Logs	(3) Levels	(4) Logs	(5) Levels	(6) Logs
Treat × Post	0.8793**	0.8593**	0.8791**	0.8450**	0.8734**	0.8619**
	(0.346)	(0.346)	(0.346)	(0.347)	(0.346)	(0.346)
	[0.009]	[0.014]	[0.011]	[0.019]	[0.011]	[0.013]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.5960	0.5962	0.5960	0.5963	0.5961	0.5962
R-squared (within)	0.0472	0.0477	0.0471	0.0480	0.0475	0.0478
Observations	4,396	4,396	4,396	4,396	4,396	4,396
Maximum value of winsorized/logged controls and mechanism variables (in levels):						
Extracurricular class hours	42		18		84	
Hours studying (weekday)	20		14		24	
Hours studying (weekend)	20		14		24	
Family size	12		9		15	
Internet use (hours)	60		35		133	
Total number of arguments	30		10		80	

Note: CFPS data (2016, 2018, 2020). The dependent variable in all specifications is the CES-D20 score. All control variables from Table 1 are included in each specification. Column (1) is our baseline specification. All columns labelled "Logs" take $\log(1 + \cdot)$ of all variables listed at the bottom of the table—except family size, which is simply log(family size), since its minimum sample value is 2. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator. Note that the mechanism variables listed at the bottom of the table (i.e., internet use and total number of arguments) are not included as controls—their maximum values are included here simply for reference.

The estimates in Table B1 demonstrate that these alternative specifications do not appear to impact our inference compared to our baseline estimate (0.8793, column 1). Specifically, whether accounting for outliers either through winsorizing (at either the 95th or 99th percentile) taking logs, or even leaving these variables at their originally-reported values in the data (albeit at implausible levels), our coefficient estimate is always estimated precisely (significant at 5% level), falling within the range of 0.84–0.88.

Finally, the bottom section of Table B1 sheds some light on why we chose to winsorize these variables at the 99th percentile in our baseline (i.e., preferred) specification. The maximum value of these variables is indeed implausibly large for some variables (e.g., 24 hours of study per day, 84 hours of extracurricular classes per week, 133 hours of internet use per week), and we deem winsorizing at the 95th percentile excessive for some variables (e.g., total number of arguments being reduced from 80 to 10 per month). Regardless, the results in this section suggest that this choice is not critically important for our inference.

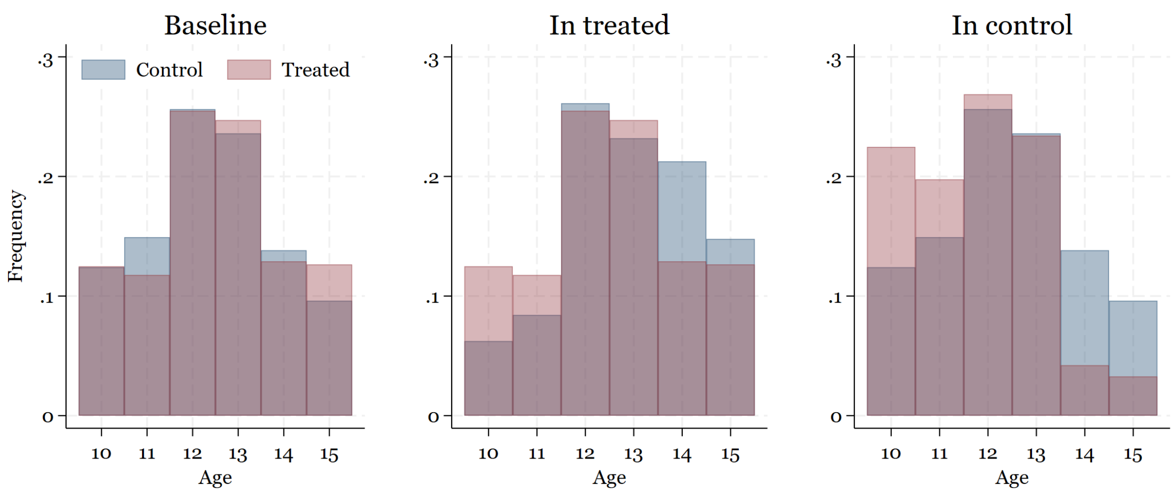
B.2 Age-distribution of students in the treated and control groups

In Section 3, we stated that it is necessary to include overlapping students in both the control and treated groups, otherwise the age-distribution will differ systematically across groups. We will proceed to demonstrate this both empirically, using our sample data, and by a stylised example.

B.2.1 Illustrating the age effect with our sample data

First, we illustrate this phenomenon empirically using the CFPS data. Figure B1 provides the empirical distribution of ages, across the range of ages in our sample (11–15), for our full sample (i.e., our 'baseline', which includes overlapping students in both the treatment and control group), and for the cases where they appear in only the treatment group ('In treated') or the control group ('In control').

Figure B1: Age distribution (empirical)



Note: CFPS data (2016, 2018, 2020). 'Baseline' includes overlapping students in both the treated and control groups, 'In treated' only includes overlapping students in the treatment group, and 'In control' only includes overlapping students in the control group.

Figure B1 reveals that excluding overlapping students from the control group will upwardly bias the age of students in the control group (see 'In treated'), while excluding them from the treated group will downwardly bias the age of students in the treated group (see 'In control'). On the other hand, duplicating overlapping students and thereby keeping them in each group results in a well-balanced age-distribution (see 'Baseline').

B.2.2 Illustrating the age effect by example

We now proceed to use a stylised example to demonstrate precisely why the outcome we observed using our sample data is a general result of not including overlapping students in both the treatment and control group.

To simplify matters we will consider a limited range of ages of students: from 11 to 13. To ensure we can include an individual fixed effect to account for time-invariant unobservable individual heterogeneity, our empirical strategy requires us observing students at least twice. Finally, to align with our dataset, we will assume we have 3 waves of data (however, to simply matters, we will assume each wave is observed one year apart).

Based on these few assumptions, in Table B2 we illustrate the effect of duplicating overlapping students vs only including them in either the treated or control group/cohort on the age-distribution in the control and treatment groups. For a given distribution of students aged 11–13 in each wave, the table shows which group they will fall into: (i) overlapping students, (ii) control group/cohort only, (iii) treated group/cohort only.

In Table B2 Panel A, we observe that, after duplication, the distribution of ages is balanced across the treatment and control group: we observe students aged 11–13 in both the treatment and control group. While, in Panel B we see that excluding overlapping students from the control group will upwardly bias the age of students in the control group. Conversely, in Panel C, we see that excluding overlapping students from the treated group will downwardly bias the age of students in the treated group.

Intuitively, this result arises since overlapping students are the youngest in earlier waves, and oldest students in later waves. Only by retaining these overlapping students in both groups can we avoid systematically distorting the age-distribution in either the treatment or control group.

Finally, in Section B.4.2, Table B6 columns (3) and (4) reports coefficient estimates when overlapping students are only in the treated group or only in the control group, respectively. Since, in both cases, coefficient estimates lie significantly below our baseline estimate, we can infer that there exists an age/grade effect for adolescents in China, consistent with prior studies; i.e., dropping relatively young students from the control group will, on average, increase the incidence of depressive symptoms in the control group, while dropping relatively old students from the treated group will, on average, decrease the incidence of depressive symptoms in the treated group—either way, the resulting effect on our DiD coefficient estimate is to downwardly bias it.

Table B2: Age-distribution (example)

	Age of students		
	Wave 1	Wave 2	Wave 3
Panel A: Overlapping students in both groups			
Representation of each group across waves:			
Overlapping	11	12	13
Control only	12	13	–
Treated only	–	11	12
Final age-distribution after duplication:			
Control (overlap)	11	12	–
Control only	12	13	–
Treated only	–	11	12
Treated (overlap)	–	12	13
Panel B: Overlapping students in treated group only			
Representation of each group across waves:			
Overlapping	–	12	13
Control only	12	13	–
Treated only	–	11	12
Final age-distribution after no duplication:			
Control only	12	13	
Treated only	–	11	12
Treated (overlap)	–	12	13
Panel C: Overlapping students in control group only			
Representation of each group across waves:			
Overlapping	11	12	–
Control only	12	13	–
Treated only	–	11	12
Final age-distribution after no duplication:			
Control (overlap)	11	12	–
Control only	12	13	–
Treated only	–	11	12

B.3 Identification

In this section, we will briefly demonstrate that our key parameter of interest is identified and explain why we do not report the coefficient estimate on Treat throughout the paper. We begin with a simple illustration of our empirical strategy, which departs from a standard cohort DiD due to overlapping individuals. For the purposes of exposition, we will simplify Equation (1) to just include the DiD terms. Thus, observed MH can be represented by:

$$MH_{j,c,t} = \alpha_0 + \delta(\mathcal{I}^{Treat}_{j,c} \times \mathcal{I}^{Post}_{t}) + \gamma_1 \mathcal{I}^{Treat}_{j,c} + \gamma_2 \mathcal{I}^{Post}_{t} + \epsilon_{j,c,t}, \tag{2}$$

for student $j = 1, \ldots, N$, in cohort $c = 0, 1$, at time period $t = 1, 2$, and where $E[\epsilon_{j,c,t}|c, t] = 0$.

Assumptions: (i) Constant (additive) causal effect of treatment: $E[MH_{1jct} - MH_{0jct}|c, t] = \delta_0 > 0$; (ii) Constant (group-specific) intercepts: $\mu^c_0 > 0$ (control group), $\mu^t_0 > 0$ (treated group); (iii) Constant (common) time trend: $\lambda_0 > 0$.

Under these assumptions, the parameters of Equation (2) are identified as follows:

- $\mathbb{E}[MH_{j,c,t}|c=1, t=1] = \alpha_0 = \mu^c_0 \Rightarrow \alpha_0 = \mu^c_0$ (i.e., control-group cohort intercept)
- $\mathbb{E}[MH_{j,c,t}|c=1, t=2] = \alpha_0 + \gamma_2 = \mu^c_0 + \lambda_0 \Rightarrow \gamma_2 = \lambda_0$ (i.e., time trend)
- $\mathbb{E}[MH_{j,c,t}|c=2, t=1] = \alpha_0 + \gamma_1 = \mu^t_0 \Rightarrow \gamma_1 = \mu^t_0 - \mu^c_0$ (i.e., cohort mean difference)
- $\mathbb{E}[MH_{j,c,t}|c=2, t=2] = \alpha_0 + \gamma_1 + \gamma_2 + \delta = \delta_0 \Rightarrow \delta = \delta_0$ (i.e., treatment effect).

The implication coming from the final line above can equivalently be obtained from evaluating population difference-in-differences for Equation (2). For example:

$$\begin{aligned}
&(\mathbb{E}[MH_{j,c,t}|c=2, t=2] - \mathbb{E}[MH_{j,c,t}|c=2, t=1]) - (\mathbb{E}[MH_{j,c,t}|c=1, t=2] - \mathbb{E}[MH_{j,c,t}|c=1, t=1]) \\
&\quad = [(\alpha_0 + \gamma_1 + \gamma_2 + \delta) - (\alpha_0 + \gamma_1)] - [(\alpha_0 + \gamma_2) - (\alpha_0)] \\
&\quad = [\gamma_2 + \delta] - [\gamma_2] \\
&\quad = \delta.
\end{aligned}$$

The constant time trend assumption, together with overlapping individuals, implies that $\mu^t_0 - \mu^c_0 = \lambda_0$ (i.e., that the cohort mean difference precisely equals the time trend). In terms of regression coefficients, $\gamma_1 = \gamma_2$. Note that this also means that differences in intercepts across cohorts is equivalent to differences in means across cohorts. So, while γ_1 is theoretically identified from within-variation for overlapping students, it is constrained to equal γ_2 and thus provides no additional information beyond reporting the coefficient on Post.[38]

Finally, in Equation (2), idiosyncratic (time-invariant) differences in means are captured by the individual fixed effects, and year fixed effects are omitted since they are not identified.

[38]This does not indicate, however, that Treat is a redundant variable in Equation (2). Indeed, Monte Carlo simulations reveal that excluding it in this setting induces omitted variable bias, upwardly biasing δ.

B.4 Robustness

B.4.1 Mechanisms

Table B3: The effect of COVID on mental health, by internet-access device

	Any technology		Mobile phone	
	(1) No prior use	(2) Prior use	(3) No prior use	(4) Prior use
Treat × Post	0.5051	1.0560**	0.5016	1.0195*
	(0.490)	(0.483)	(0.455)	(0.531)
	[0.327]	[0.031]	[0.268]	[0.043]
Post	−0.3191	0.3219	−0.2465	0.4976
	(0.364)	(0.331)	(0.345)	(0.381)
Constant	9.9437***	9.0669***	9.9285***	8.1740***
	(1.915)	(2.291)	(1.787)	(2.572)
R-squared	0.5954	0.6327	0.6022	0.6322
R-squared (within)	0.0341	0.0800	0.0337	0.0885
Observations	2,316	2,080	2,620	1,776
	Always use phone		Never use phone	
	(5) No	(6) Yes	(7) No	(8) Yes
Treat × Post	0.4281	1.3937**	1.1338***	−0.4140
	(0.430)	(0.585)	(0.383)	(0.773)
	[0.316]	[0.016]	[0.004]	[0.592]
Post	0.1603	0.3381	0.2245	0.0288
	(0.253)	(0.398)	(0.248)	(0.382)
Constant	8.9544***	11.8240***	10.5985***	7.5205***
	(1.732)	(2.648)	(1.686)	(2.597)
R-squared	0.5978	0.6371	0.6126	0.6156
R-squared (within)	0.0296	0.1036	0.0702	0.0405
Observations	2,896	1,500	3,384	1,012
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes

Note: CFPS data (2016, 2018, 2020). The dependent variable in all specifications is the CES-D20 score. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

Table B4: The effect of COVID on conflict in the home (disaggregated measures)

	Child-parent conflict			Parent-parent conflict		
	(1) Conflict	(2) Arguments	(3) log(Arg)	(4) Conflict	(5) Arguments	(6) log(Arg)
Treat × Post	0.0777***	0.3755***	0.1209***	0.0561**	0.3034***	0.0998***
	(0.028)	(0.135)	(0.038)	(0.027)	(0.111)	(0.034)
	[0.002]	[0.001]	[0.000]	[0.039]	[0.006]	[0.005]
Post	0.0222	−0.1092	0.0071	−0.0320*	−0.1978***	−0.0668***
	(0.019)	(0.086)	(0.025)	(0.018)	(0.073)	(0.023)
Constant	0.3959***	1.2546**	0.5004***	0.2402**	0.7326	0.2787**
	(0.103)	(0.516)	(0.144)	(0.101)	(0.461)	(0.133)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.5920	0.5691	0.6128	0.5592	0.5389	0.5753
R-squared (within)	0.0368	0.0267	0.0384	0.0144	0.0150	0.0169
Observations	4,396	4,396	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

Table B5: The effect of COVID on mental health, conditional on prior conflict

	Child-parent conflict		Parent-parent conflict	
	(1) No prior conflict	(2) Prior conflict	(3) No prior conflict	(4) Prior conflict
Treat × Post	0.6183	1.4237**	0.7119*	1.4679**
	(0.397)	(0.642)	(0.368)	(0.743)
	[0.105]	[0.027]	[0.049]	[0.043]
Post	0.3545	−0.3379	0.5498**	−0.8831*
	(0.269)	(0.407)	(0.249)	(0.480)
Constant	8.7506***	11.7900***	9.0294***	12.1407***
	(1.587)	(2.949)	(1.527)	(3.208)
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.5998	0.6290	0.6140	0.6076
R-squared (within)	0.0507	0.0647	0.0521	0.0681
Observations	2,882	1,514	3,200	1,196

Note: CFPS data (2016, 2018, 2020). The dependent variable in all specifications is the CES-D20 score. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

B.4.2 Overlapping students

In this section, we report estimates from a variant of Equation (1), which allows overlapping students to have a cohort-specific fixed effect, rather than one per student (as in the baseline). This estimate is reported in Table B6 column (2), while column (1) reports the baseline estimate for reference.

Table B6: Alternative classification of overlapping students

	(1) Baseline	(2) Recoded	(3) In treated	(4) In control
Treat × Post	0.8793**	0.8951***	0.7582**	0.6240
	(0.346)	(0.345)	(0.358)	(0.406)
	[0.011]	[0.011]	[0.034]	[0.130]
Post	0.1345	0.1206	0.2344	0.0556
	(0.225)	(0.225)	(0.277)	(0.231)
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.5960	0.6364	0.6438	0.6285
R-squared (within)	0.0472	0.0457	0.0513	0.0416
Observations	4,396	4,396	3,534	3,534
Students in treated group	965	965	965	534
Students in control group	1,233	1,233	802	1,233

Note: CFPS data (2016, 2018, 2020). The dependent variable in all specifications is the CES-D20 score. Treat is omitted in columns (2)–(4), as it is time-invariant for all units in these samples. Column (2) effectively classifies each student that appears in both the treatment and control group as a different student, so that they no longer share the same individual fixed effect (across the groups); i.e., this specification has student-cohort fixed effects, rather than a fixed effect for each student. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors for the DiD indicator.

Table B6 columns (3) and (4) reveal, using our model, the impact of restricting overlapping students to being present in either the treatment group or control group, respectively. We already demonstrated in Appendix B.2 that both of these cases will shift the age-distribution of students: including overlapping students only in the treated group means dropping some relatively young students from the control group, while including overlapping students only in the control group means dropping some relatively old students from the treated group.

The resulting reduction in the coefficient estimate is indicative of an age/education-level effect of mental health; i.e., in order to reduce the estimated impact of COVID on mental health, we must be either systematically dropping students with relatively good mental health from the control group (the former case), or dropping students with relatively poor mental health to the treated group (the latter case).

B.4.3 Alternative estimator

Table B7: The effect of COVID on mental health, using a doubly robust estimator

	(1) Baseline	(2) Age	(3) Gender	(4) Region	(5) All
Panel A: CES-D20 score					
ATT	0.8793**	0.8946**	0.8730**	0.9076**	0.9175**
	(0.385)	(0.384)	(0.385)	(0.383)	(0.383)
	[0.024]	[0.018]	[0.021]	[0.015]	[0.016]
Panel B: CES-D8 score					
ATT	0.4199**	0.4282**	0.4165**	0.4348**	0.4401**
	(0.194)	(0.194)	(0.194)	(0.193)	(0.193)
	[0.038]	[0.031]	[0.043]	[0.027]	[0.028]
Controls	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes
Observations	4,396	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). All estimates in this table were obtained using a 'doubly robust' DiD estimator (Sant'Anna and Zhao, 2020). In all columns of both panels, estimates were obtained using the sample where 'overlapping' individuals have cohort-student fixed effects, as in Table B6 column (2), i.e., 'Recoded'. Column (1) only includes all time-varying controls from our baseline specification. Columns (2) to (5) also include time-invariant variables (for determining propensity scores), e.g., column (2) includes age, column (5) includes age, gender, and region. Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild-bootstrap p-values (based on 10,000 replications) are in brackets directly below the standard errors.

B.4.4 Potentially problematic controls

To investigate whether a number of our control variables may be problematic, i.e., affected by COVID, we consider several alternative specifications below. In particular, we were concerned about household composition variables (i.e., co-residence with mother or father, and family size) as well as extracurricular class hours, since these could conceivably be affected by post-COVID lockdown policy responses.

In Table B8 Panel A, we investigate whether COVID appeared to impact these variables (by setting each candidate variable as the dependent variable). Then, in Panel B, we estimate the causal effect of COVID on mental health excluding each candidate variable as a control. The relevant candidate variable in each column is listed at the top and bottom row.

Table B8: Investigating potentially problematic controls

	(1) Extracurricular class hours	(2) Family size	(3) Father lives at home	(4) Mother lives at home
Panel A: Candidate 'bad' control as dependent variable				
Treat × Post	0.3218	0.1691***	0.0354**	−0.0359***
	(0.452)	(0.040)	(0.016)	(0.014)
Post	1.3451***	−0.1621***	−0.0103	0.0117
	(0.226)	(0.029)	(0.010)	(0.008)
Constant	0.6319	4.6144***	0.2663***	0.4573***
	(1.815)	(0.156)	(0.071)	(0.066)
R-squared	0.5558	0.9193	0.7904	0.8254
R-squared (within)	0.0415	0.0494	0.1819	0.1815
Panel B: CES-D20 score as dependent variable (excluding 'bad' control candidate)				
Treat × Post	0.8776**	0.9135***	0.8553**	0.8296**
	(0.346)	(0.346)	(0.346)	(0.345)
Post	0.2364	0.1902	0.2326	0.2417
	(0.206)	(0.209)	(0.209)	(0.209)
Constant	9.7289***	10.7669***	9.5741***	10.3072***
	(1.448)	(1.117)	(1.443)	(1.441)
R-squared	0.5958	0.5954	0.5955	0.5948
R-squared (within)	0.0466	0.0459	0.0460	0.0443
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Observations	4,396	4,396	4,396	4,396
Excluded control	Extracurricular class	Family size	Father lives at home	Mother lives at home

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.4.5 Subjective health measures

The CFPS survey includes a question on 'self-assessed health' (SAH) in each of its waves, however, we did not include it as one of our baseline controls. Due to its subjective nature, its causal connection to mental health may be unclear, possibly making it a poor control.[39] Nonetheless, in Table B9, we demonstrate that adding it to our list of controls (either directly or lagged) does not significantly affect our main conclusions.

In the CFPS, SAH is an ordinal variable, taking values from 1 to 5, which represent 'excellent', 'very good', 'good', 'fair', and 'poor' health, respectively. When we incorporate (i) SAH or (ii) lagged SAH, the coefficient on the DiD term (i.e., Treat × Post) with individual fixed effects from our baseline model becomes (i) 0.984, significant at 1% level, and (ii) 0.817, significant at 5% level, respectively.

Table B9: The effect of COVID on mental health, with additional health controls

	(1) Baseline	(2) SAH	(3) SAH (lagged)
Treat × Post	0.8793**	0.9841***	0.8169**
	(0.346)	(0.346)	(0.358)
Post	0.1345	0.0995	0.3255
	(0.225)	(0.224)	(0.286)
Constant	9.6532***	7.8486***	10.0267***
	(1.443)	(1.469)	(1.557)
Self-assessed health		0.8215***	
		(0.164)	
Self-assessed health (lagged)			−0.0189
			(0.034)
Controls	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
R-squared	0.5960	0.6019	0.5932
R-squared (within)	0.0472	0.0612	0.0457
Observations	4,396	4,396	4,028

Note: CFPS data (2016, 2018, 2020). SAH refers to self-assessed health, while SAH (lagged) is the corresponding one-wave lagged value. Column (1) is the baseline model (i.e., excluding SAH as a control). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[39] For a detailed investigation of objective vs subjective health measures, see, e.g., Crossley and Kennedy (2002), Au and Johnston (2014), and Doiron et al. (2015).

Appendix C Full tables of coefficient estimates

Table C1: The effect of COVID on adolescent depressive symptoms

	CES-D20 score		CES-D8 score	
	(1)	(2)	(3)	(4)
Treat × Post	0.8916**	0.8793**	0.4308**	0.4241**
	(0.349)	(0.346)	(0.177)	(0.175)
Treat	0.2202	0.2075	0.1123	0.1117
	(0.238)	(0.259)	(0.121)	(0.131)
Post	0.3428	0.1345	0.1841*	0.0802
	(0.210)	(0.225)	(0.107)	(0.114)
Attend key school	0.5972**	0.0537	0.3099**	0.0180
	(0.249)	(0.305)	(0.126)	(0.154)
Extracurricular class hours	-0.0080	0.0062	-0.0039	0.0032
	(0.014)	(0.018)	(0.007)	(0.009)
Hours studying (weekday)	-0.0394	0.0458	-0.0197	0.0222
	(0.037)	(0.044)	(0.019)	(0.023)
Hours studying (weekend)	-0.0238	0.0353	-0.0137	0.0179
	(0.035)	(0.045)	(0.018)	(0.023)
Student leader	−0.6747***	0.2126	−0.3496***	0.1040
	(0.214)	(0.294)	(0.109)	(0.149)
Student performance	−0.2565**	−0.1265	−0.1292**	−0.0653
	(0.125)	(0.155)	(0.063)	(0.078)
Study pressure	0.9301***	0.5920***	0.4699***	0.3019***
	(0.093)	(0.120)	(0.047)	(0.060)
Excellence	−0.6182***	−0.1504	−0.3173***	−0.0765
	(0.142)	(0.176)	(0.072)	(0.089)
Satisfaction (school)	−0.4034***	−0.3416**	−0.2052***	−0.1704**
	(0.123)	(0.152)	(0.063)	(0.078)
Satisfaction (teacher)	−0.5382***	−0.4138**	−0.2715***	−0.2117***
	(0.137)	(0.161)	(0.069)	(0.081)
Pocket money	−0.5125**	−0.6070*	−0.2540**	−0.3060*
	(0.246)	(0.321)	(0.124)	(0.163)
Internet (mobile phone)	−0.2057	0.3156	−0.1070	0.1537
	(0.216)	(0.286)	(0.110)	(0.145)
Internet (computer)	−0.2031	0.0726	−0.1047	0.0322
	(0.237)	(0.342)	(0.120)	(0.173)
Sick (past month)	0.1934	−0.1205	0.0996	−0.0476
	(0.287)	(0.381)	(0.146)	(0.193)
Sick (past year)	−0.5457**	−0.1512	−0.2838**	−0.0761
	(0.261)	(0.319)	(0.133)	(0.162)
Out-of-pocket medical expenses (log)	0.1205***	0.0600	0.0625***	0.0303
	(0.043)	(0.052)	(0.022)	(0.026)
Father lives at home	−0.3231	−0.5320	−0.1622	−0.2657
	(0.287)	(0.456)	(0.145)	(0.231)
Mother lives at home	0.0339	1.2553**	0.0047	0.6185**
	(0.328)	(0.537)	(0.166)	(0.271)
Family size	0.1609**	0.2288	0.0801**	0.1140
	(0.066)	(0.191)	(0.034)	(0.096)
Constant	13.7926***	9.6532***	5.8978***	3.8006***
	(0.972)	(1.443)	(0.491)	(0.732)
Individual FE	No	Yes	No	Yes
R-squared	0.0912	0.5960	0.0913	0.5972
R-squared (within)	–	0.0472	–	0.0468
Observations	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C2: The effect of COVID on specific items from the CES-D scale (I)

	(1) Low spirit	(2) Feel sad	(3) Everything difficult	(4) Cannot continue
Treat × Post	0.2633***	0.1130***	0.0764*	0.0145
	(0.042)	(0.040)	(0.044)	(0.026)
Treat	−0.1197***	0.0180	0.0154	0.0068
	(0.033)	(0.030)	(0.031)	(0.019)
Post	−0.0999***	0.0050	−0.0212	0.0052
	(0.028)	(0.026)	(0.027)	(0.016)
Attend key school	−0.0277	0.0427	0.0451	0.0412*
	(0.039)	(0.037)	(0.038)	(0.024)
Extracurricular class hours	0.0023	0.0028	0.0010	−0.0004
	(0.002)	(0.002)	(0.002)	(0.001)
Hours studying (weekday)	0.0131***	0.0035	0.0105**	0.0001
	(0.005)	(0.005)	(0.005)	(0.003)
Hours studying (weekend)	0.0097	−0.0007	−0.0036	−0.0000
	(0.006)	(0.005)	(0.006)	(0.003)
Student leader	0.0238	0.0139	0.0535	0.0312
	(0.036)	(0.031)	(0.036)	(0.021)
Student performance	−0.0038	−0.0031	−0.0068	0.0008
	(0.019)	(0.017)	(0.020)	(0.011)
Study pressure	0.0607***	0.0399***	0.0590***	0.0146
	(0.013)	(0.013)	(0.015)	(0.009)
Excellence	0.0164	−0.0205	−0.0447**	−0.0020
	(0.021)	(0.019)	(0.022)	(0.013)
Satisfaction (school)	−0.0205	−0.0088	−0.0411**	−0.0010
	(0.018)	(0.017)	(0.018)	(0.010)
Satisfaction (teacher)	−0.0366*	0.0032	−0.0223	−0.0168
	(0.019)	(0.018)	(0.019)	(0.011)
Pocket money	−0.0161	−0.0815**	−0.0132	−0.0006
	(0.040)	(0.038)	(0.043)	(0.024)
Internet (mobile phone)	0.0958***	0.0057	0.0464	0.0036
	(0.034)	(0.032)	(0.034)	(0.022)
Internet (computer)	0.0531	0.0476	0.0256	−0.0067
	(0.039)	(0.038)	(0.040)	(0.018)
Sick (past month)	0.0499	0.0033	−0.0337	−0.0341
	(0.044)	(0.040)	(0.049)	(0.026)
Sick (past year)	−0.0295	0.0131	−0.0316	0.0155
	(0.039)	(0.035)	(0.040)	(0.021)
OOP medical expenses (log)	−0.0061	0.0020	0.0065	−0.0027
	(0.007)	(0.006)	(0.007)	(0.004)
Father lives at home	−0.0244	0.0066	−0.0539	−0.0332
	(0.059)	(0.051)	(0.052)	(0.038)
Mother lives at home	0.1272*	0.0028	0.0446	0.0832*
	(0.068)	(0.056)	(0.066)	(0.049)
Family size	0.0342	0.0537***	0.0188	−0.0021
	(0.024)	(0.020)	(0.023)	(0.010)
Constant	1.1494***	1.0795***	1.5272***	1.0922***
	(0.168)	(0.151)	(0.175)	(0.101)
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.5364	0.5325	0.5265	0.5101
R-squared (within)	0.0538	0.0268	0.0296	0.0109
Observations	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C3: The effect of COVID on specific items from the CES-D scale (II)

	(5) Poor sleep	(6) Relationships	(7) Happy mood	(8) Happy life
Treat × Post	0.0335	0.0248	−0.0609	−0.0406
	(0.045)	(0.039)	(0.051)	(0.048)
Treat	0.0617*	0.0762***	0.0148	0.0385
	(0.034)	(0.029)	(0.041)	(0.039)
Post	0.0596**	0.0362	0.0344	0.0609*
	(0.030)	(0.024)	(0.035)	(0.033)
Attend key school	0.0551	0.0246	−0.0788*	−0.0843*
	(0.041)	(0.036)	(0.045)	(0.043)
Extracurricular class hours	0.0006	−0.0013	−0.0016	−0.0002
	(0.002)	(0.002)	(0.002)	(0.002)
Hours studying (weekday)	0.0098**	−0.0001	−0.0093	−0.0053
	(0.005)	(0.004)	(0.007)	(0.006)
Hours studying (weekend)	−0.0050	−0.0010	0.0099	0.0087
	(0.006)	(0.005)	(0.007)	(0.007)
Student leader	0.0266	0.0060	−0.0081	−0.0428
	(0.036)	(0.033)	(0.044)	(0.041)
Student performance	−0.0356*	0.0142	−0.0221	−0.0089
	(0.019)	(0.018)	(0.022)	(0.021)
Study pressure	0.0421***	0.0433***	0.0267	0.0157
	(0.016)	(0.014)	(0.017)	(0.016)
Excellence	−0.0455**	−0.0173	0.0175	0.0196
	(0.022)	(0.019)	(0.025)	(0.026)
Satisfaction (school)	−0.0053	−0.0232	−0.0288	−0.0418*
	(0.018)	(0.015)	(0.022)	(0.021)
Satisfaction (teacher)	0.0017	−0.0307*	−0.0732***	−0.0369*
	(0.018)	(0.016)	(0.022)	(0.022)
Pocket money	0.0132	−0.0611*	−0.0337	−0.1130**
	(0.042)	(0.036)	(0.049)	(0.047)
Internet (mobile phone)	−0.0297	0.0291	0.0321	−0.0293
	(0.038)	(0.031)	(0.042)	(0.040)
Internet (computer)	0.0127	0.0444	−0.0764	−0.0683
	(0.044)	(0.035)	(0.051)	(0.045)
Sick (past month)	0.0566	0.0253	−0.0668	−0.0482
	(0.047)	(0.041)	(0.055)	(0.050)
Sick (past year)	0.0272	−0.0125	−0.0571	−0.0014
	(0.039)	(0.034)	(0.046)	(0.043)
OOP medical expenses (log)	0.0086	0.0016	0.0158**	0.0044
	(0.007)	(0.006)	(0.008)	(0.007)
Father lives at home	−0.0763	−0.0067	−0.0495	−0.0283
	(0.059)	(0.051)	(0.069)	(0.062)
Mother lives at home	0.0559	0.0140	0.0947	0.1960**
	(0.064)	(0.064)	(0.082)	(0.077)
Family size	0.0008	0.0172	−0.0178	0.0092
	(0.020)	(0.017)	(0.026)	(0.028)
Constant	1.3720***	1.2880***	2.3509***	1.9413***
	(0.164)	(0.149)	(0.215)	(0.217)
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.5185	0.5336	0.5200	0.5259
R-squared (within)	0.0242	0.0239	0.0214	0.0199
Observations	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C4: The effect of COVID on the proposed mechanisms

	(1) Income	(2) Exercise	(3) Internet	(4) Conflict	(5) Arguments
Treat × Post	−0.1170**	−0.2431***	8.5346***	0.0606**	0.6789***
	(0.048)	(0.030)	(0.615)	(0.029)	(0.201)
Treat	0.2469***	0.0727***	2.5992***	−0.0191	−0.4667***
	(0.035)	(0.022)	(0.410)	(0.022)	(0.162)
Post	0.2771***	0.0849***	1.7960***	0.0158	−0.3070**
	(0.031)	(0.019)	(0.329)	(0.020)	(0.125)
Attend key school	−0.0046	0.0441	−0.6353	0.0692***	−0.0224
	(0.040)	(0.028)	(0.546)	(0.025)	(0.176)
Extracurricular class hours	−0.0013	0.0013	−0.0428	−0.0002	0.0046
	(0.002)	(0.001)	(0.031)	(0.001)	(0.009)
Hours studying (weekday)	−0.0005	0.0064*	0.0928	0.0025	0.0379
	(0.006)	(0.004)	(0.082)	(0.003)	(0.026)
Hours studying (weekend)	−0.0061	0.0060	−0.1488	−0.0013	−0.0319
	(0.007)	(0.004)	(0.093)	(0.004)	(0.030)
Student leader	−0.0104	−0.0332	0.2189	0.0349	0.0621
	(0.045)	(0.024)	(0.526)	(0.023)	(0.186)
Student performance	0.0246	0.0097	−0.3990	0.0146	−0.0171
	(0.023)	(0.012)	(0.245)	(0.012)	(0.092)
Study pressure	−0.0209	−0.0011	−0.3011	0.0337***	0.3536***
	(0.016)	(0.010)	(0.201)	(0.009)	(0.075)
Excellence	−0.0176	0.0195	0.1278	−0.0095	−0.0037
	(0.021)	(0.014)	(0.269)	(0.013)	(0.105)
Satisfaction (school)	0.0025	0.0002	−0.4712*	−0.0104	−0.1205
	(0.020)	(0.013)	(0.266)	(0.012)	(0.098)
Satisfaction (teacher)	−0.0155	−0.0062	0.1088	−0.0138	−0.0313
	(0.022)	(0.012)	(0.250)	(0.012)	(0.097)
Pocket money	0.0185	0.0319	−0.2909	−0.0035	−0.1074
	(0.042)	(0.030)	(0.590)	(0.029)	(0.182)
Internet (mobile phone)	−0.0660	0.0045	6.3618***	0.0446*	0.3302*
	(0.043)	(0.025)	(0.484)	(0.023)	(0.183)
Internet (computer)	0.0564*	0.0278	3.9697***	0.0278	0.1584
	(0.033)	(0.027)	(0.674)	(0.028)	(0.203)
Sick (past month)	−0.0764*	−0.0155	−0.1947	0.0092	0.0214
	(0.046)	(0.029)	(0.629)	(0.029)	(0.208)
Sick (past year)	−0.0477	−0.0243	−0.8125	0.0107	0.0130
	(0.040)	(0.025)	(0.549)	(0.025)	(0.181)
OOP medical expenses (log)	0.0073	0.0055	0.2007**	0.0006	−0.0494*
	(0.007)	(0.004)	(0.096)	(0.005)	(0.030)
Father lives at home	0.1271**	0.0468	−1.0001	0.0456	0.1964
	(0.055)	(0.039)	(0.768)	(0.036)	(0.204)
Mother lives at home	−0.0933	−0.0284	0.0901	−0.0242	−0.3004
	(0.066)	(0.050)	(1.011)	(0.046)	(0.281)
Family size	0.1051***	0.0087	−0.1388	−0.0042	−0.0457
	(0.024)	(0.014)	(0.267)	(0.014)	(0.079)
Constant	10.2578***	0.3770***	2.8532	0.3712***	1.9871**
	(0.193)	(0.116)	(2.237)	(0.110)	(0.806)
Individual FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.7334	0.5343	0.6531	0.5872	0.5809
R-squared (within)	0.0618	0.0437	0.4015	0.0274	0.0271
Observations	4,396	4,396	4,396	4,396	4,396

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C5: The effect of COVID on depressive symptoms, by mechanism (pre-treatment) (I)

	Income		Exercise	
	(1) Low	(2) High	(3) No exercise	(4) Prior exercise
Treat × Post	0.1640	1.7441***	0.9073	0.9006**
	(0.504)	(0.468)	(0.604)	(0.411)
Treat	0.5630	−0.4221	0.0054	−0.0927
	(0.416)	(0.365)	(0.500)	(0.344)
Post	0.5027	−0.3505	0.2928	0.0580
	(0.334)	(0.302)	(0.387)	(0.281)
Attend key school	0.2718	−0.2175	0.0190	−0.0753
	(0.464)	(0.389)	(0.575)	(0.353)
Extracurricular class hours	0.0078	0.0072	0.0017	0.0039
	(0.025)	(0.024)	(0.036)	(0.020)
Hours studying (weekday)	0.0411	0.0355	0.1307*	−0.0070
	(0.064)	(0.060)	(0.070)	(0.056)
Hours studying (weekend)	−0.0218	0.1096*	−0.0554	0.0853
	(0.063)	(0.065)	(0.079)	(0.055)
Student leader	0.7772*	−0.3385	1.2736***	−0.2306
	(0.409)	(0.418)	(0.475)	(0.365)
Student performance	−0.1757	−0.0988	−0.0498	−0.1805
	(0.207)	(0.221)	(0.236)	(0.197)
Study pressure	0.5697***	0.6699***	0.6368***	0.5630***
	(0.170)	(0.168)	(0.186)	(0.157)
Excellence	−0.0356	−0.2302	−0.1286	−0.1827
	(0.253)	(0.243)	(0.302)	(0.211)
Satisfaction (school)	−0.3156	−0.3721*	−0.3362	−0.2609
	(0.208)	(0.218)	(0.237)	(0.196)
Satisfaction (teacher)	−0.6088***	−0.1885	−0.2560	−0.5663***
	(0.222)	(0.233)	(0.258)	(0.208)
Pocket money	−0.8097*	−0.2184	−0.4450	−0.8266**
	(0.458)	(0.458)	(0.537)	(0.391)
Internet (mobile phone)	0.2148	0.5982	0.7585	0.0862
	(0.408)	(0.387)	(0.482)	(0.345)
Internet (computer)	0.1243	−0.0893	−0.7212	0.3089
	(0.530)	(0.435)	(0.644)	(0.396)
Sick (past month)	0.6783	−1.1465**	1.2287*	−0.8686**
	(0.540)	(0.523)	(0.681)	(0.437)
Sick (past year)	−0.2288	−0.1810	0.0877	−0.2135
	(0.451)	(0.436)	(0.574)	(0.380)
OOP medical expenses (log)	0.0490	0.0903	−0.0868	0.1170*
	(0.071)	(0.076)	(0.091)	(0.062)
Father lives at home	−1.0337	0.0300	−0.7318	−0.4885
	(0.631)	(0.663)	(0.705)	(0.597)
Mother lives at home	1.3997*	0.7576	0.8811	1.5173**
	(0.718)	(0.801)	(0.814)	(0.687)
Family size	0.2813	0.2504	0.3650	0.2210
	(0.312)	(0.241)	(0.319)	(0.228)
Constant	10.6094***	8.2386***	8.0609***	10.7265***
	(2.160)	(1.882)	(2.524)	(1.705)
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.6021	0.6227	0.6032	0.6248
R-squared (within)	0.0547	0.0640	0.0660	0.0530
Observations	2,204	2,192	1,536	2,860

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C6: The effect of COVID on depressive symptoms, by mechanism (pre-treatment) (II)

	Internet		Conflict	
	(5) Low	(6) High	(7) No prior conflict	(8) Prior conflict
Treat × Post	0.6033	1.2367**	0.5714	1.3054**
	(0.457)	(0.509)	(0.422)	(0.557)
Treat	−0.2195	0.3567	−0.0272	−0.3477
	(0.423)	(0.433)	(0.347)	(0.475)
Post	−0.3802	0.3137	0.6069**	−0.3714
	(0.355)	(0.332)	(0.285)	(0.359)
Attend key school	0.6370	−0.4953	0.2836	−0.3317
	(0.435)	(0.427)	(0.374)	(0.500)
Extracurricular class hours	0.0346	−0.0362	−0.0033	0.0060
	(0.026)	(0.024)	(0.018)	(0.034)
Hours studying (weekday)	0.0044	0.0706	0.0396	0.0258
	(0.060)	(0.066)	(0.050)	(0.078)
Hours studying (weekend)	0.0508	0.0380	−0.0067	0.0960
	(0.059)	(0.069)	(0.057)	(0.072)
Student leader	0.4775	−0.1456	0.2589	0.1424
	(0.376)	(0.451)	(0.348)	(0.509)
Student performance	0.0426	−0.4071*	−0.2530	0.0057
	(0.199)	(0.246)	(0.194)	(0.256)
Study pressure	0.5166***	0.6784***	0.3643**	0.8452***
	(0.155)	(0.189)	(0.148)	(0.195)
Excellence	−0.1653	−0.2218	0.0057	−0.3785
	(0.233)	(0.266)	(0.223)	(0.285)
Satisfaction (school)	−0.1614	−0.4980**	−0.4715**	−0.2480
	(0.195)	(0.241)	(0.184)	(0.246)
Satisfaction (teacher)	−0.3978**	−0.4202	−0.2585	−0.5823**
	(0.196)	(0.268)	(0.193)	(0.271)
Pocket money	−0.7878*	−0.5141	−0.0512	−1.0943**
	(0.432)	(0.485)	(0.425)	(0.469)
Internet (mobile phone)	0.7489*	0.5891	0.3146	0.2892
	(0.426)	(0.457)	(0.339)	(0.470)
Internet (computer)	0.4508	−0.0599	−0.4457	0.4184
	(0.549)	(0.471)	(0.396)	(0.582)
Sick (past month)	−0.2478	−0.1096	0.3656	−0.9028
	(0.500)	(0.579)	(0.469)	(0.603)
Sick (past year)	−0.0919	−0.3294	−0.1461	−0.3291
	(0.427)	(0.454)	(0.403)	(0.507)
OOP medical expenses (log)	0.0138	0.1252	0.0030	0.1639*
	(0.069)	(0.078)	(0.063)	(0.089)
Father lives at home	−0.6794	−0.4151	−1.0641*	−0.0085
	(0.516)	(0.815)	(0.564)	(0.741)
Mother lives at home	1.4639**	0.8142	1.6993**	0.5214
	(0.689)	(0.808)	(0.708)	(0.816)
Family size	0.1240	0.5260	0.3812*	0.0441
	(0.237)	(0.327)	(0.227)	(0.323)
Constant	10.0304***	9.1539***	8.8544***	11.7308***
	(1.829)	(2.436)	(1.669)	(2.454)
Individual FE	Yes	Yes	Yes	Yes
R-squared	0.6014	0.6292	0.6176	0.6093
R-squared (within)	0.0373	0.0839	0.0534	0.0606
Observations	2,486	1,910	2,466	1,930

Note: CFPS data (2016, 2018, 2020). Standard errors (clustered by student) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.