LUND UNIVERSITY
School of Economics and Management

Master's Programme in Economics

The Impact of Classroom Gender Composition on Mental Health of Adolescents: Evidence from the Czech Republic

by

Gustav Lundström & Hugo Morgado Azevedo

gustav.b.l.lundstrom@gmail.com | hugo.morgadoazevedo@gmail.com

Department of Economics

NEKP01

Master Essay II (15 credits)

June 2023

Supervisor: Jan Bietenbeck

Abstract

This essay investigates the impact of classroom gender composition on the mental health of adolescent students. Using data from the Health Behaviours of School-aged Children survey in the Czech Republic, we exploit idiosyncratic classroom variation in the gender composition within the same schools across different grades. We find that a higher proportion of female peers in the classroom has a detrimental effect on male students' mental health, whereas no comparable effects are observed among their female counterparts. Specifically, this effect is primarily driven by an increase in symptoms such as feeling low and nervous. Importantly, our study reveals non-linearity in the classroom gender composition effects, suggesting that a reduction of the number of girls in the classroom may not always be advantageous for boys, as its effectiveness depends on the initial gender composition.

Keywords: Education, Gender, Peer effects, Mental health, Fixed effects

Contents

List of Tables

1 Introduction

The determinants influencing mental health in adolescents persist as a subject that remains not fully understood. While economic research has traditionally directed its attention toward physical health,[1] an increasing number of recent studies have aimed to unravel the complexities of mental health (e.g. Eisenberg et al. 2009, Currie et al. 2010, Ding et al. 2009). Amidst the enduring COVID-19 crisis, marked by a concerning increase in mental distress among adolescents, the imperative role of mental health as a crucial aspect of overall well-being has become strikingly evident (OECD 2021). Notably, Currie & Stabile (2006) find that mental health conditions may exert a greater influence on later adult outcomes than physical health conditions.

Social interactions emerge as a pivotal factor influencing mental health, particularly among adolescents, as they navigate a distinct phase of brain development (Orben et al. 2020). Given that children allocate a large proportion of their time to school, the classroom environment occupies a central position wherein they actively engage in the formation of their social interactions. Consequently, extensive research has examined the impact of peer composition effects within the classroom (Henderson et al. 1978, Lazear 2001). The underlying concept of peer effects posits that students are not solely influenced by their own characteristics and circumstances, but also by the presence and actions of their peers. As an illustration, when individuals finds themselves in an environment where they are surrounded by peers with whom they are less likely to form close relationships, the quality of their social interactions may suffer as a result. Hill (2015) research reveals that boys tend to develop closer relationships with their male classmates, indicating that the gender composition within a classroom may influence students' mental health through this dynamic. Furthermore, gender disparities in students' habits and perceptions may also exert peer-influential effects within the classroom, thus affecting the individual student. Building on this rationale, Getik & Meier (2022) conducted a study examining the peer gender composition effects on mental health among adolescents.

In this paper, we build upon this literature and explore the effects of classroom gender composition on adolescent students in the Czech Republic. We use the Health Behaviours of School-aged Children (HBSC) study to gather self-reported health-related outcomes

[1]Theoretical perspective can be drawn from the works of Mushkin (1962); Becker (2009); and Grossman (1972), while applied research is explored in papers such as Cawley (2004); Cawley (2004); and Lindeboom et al. (2010).

and background characteristics of students. This survey encompasses students in the 6th, 8th, and 10th grades across five waves spanning from 2002 to 2018. Our treatment variable of interest employs the students' reported gender to determine the classroom proportion of female peers, excluding the student themselves. To assess mental health, we construct an index that measures psychological complaints among students, comprising of the following symptoms: feeling low, feeling irritable, feeling nervous, and experiencing sleeping difficulties.

HBSC utilizes schools as the primary sampling unit in Czechia, thereby providing data on classes from different grades within the same school. Inspired by the approach of Hoxby (2000), we exploit idiosyncratic classroom variation in the proportion of female students within the same schools across grades for identification. This method allows us to account for school and grade fixed effects to address any confounding factors shared among all students within the same school as well as those within the same grade. To assess the validity of our findings, we conduct a variety of balancing tests that support our identifying assumption that the classroom gender composition is exogenous, conditional on grade and school fixed effects

Our findings align with Getik & Meier (2022) and indicate that male students may experience a decline in their psychological well-being when surrounded by female peers, while female students do not seem to be similarly affected. Specifically, our investigation reveals that a 10 percentage point increase in the classroom proportion of female peers results in a significant rise of 0.0585 index points in the psychological complaints of male students, constituting a 1.2% deviation from the baseline. Notably, this effect appears to be particularly driven by the experience of symptoms such as feeling low and nervous.

We further investigate potential channels that may explain the adverse impact of class gender composition on boys' mental health, as well as the disparities in effects observed between boys and girls. We explore the influence of gender composition on the classroom environment, with a focus on mechanisms that we expect to be intertwined with mental health: physical activity, school environment, body perception, and drug use. Our results suggest that a reduction in physical activity levels among male students in the presence of more female peers may constitute one possible channel that could account for the observed increase in mental health complaints. Furthermore, while alterations in the school environment are unlikely to be the primary driver of the negative effect on boys'

mental health, they may offer a plausible explanation for the gender disparities in our estimates. However, body perception and drug use do not appear to be mechanisms driving the observed effects.

Finally, we perform a heterogeneity analysis to explore whether the influence of classroom gender composition varies based on the intensity of exposure. Interestingly, the findings show that the most pronounced negative effects on boys' mental health do not coincide with the highest level of proportion of female peers. This suggests that the removal of a female peer from the class may not always lead to positive outcomes on boys' mental health, depending on the initial proportion of female peers. Moreover, girls similarly exhibit non-linearity in the classroom gender composition effects.

This paper makes several notable contributions to the existing literature. Firstly, it offers novel insights into the determinants of mental health among adolescents, thereby expanding the understanding of a topic that has gained substantial prominence. Secondly, through our main analysis and exploration of potential mechanisms we expand on the peer effects literature by investigating the classroom gender composition effect on multiple outcomes. Lastly, we leverage the underutilized HBSC dataset and capitalize on the unique sampling procedure of the Czech Republic, in which schools serve as the primary sampling unit, to exploit idiosyncratic within-school variation.

The structure of the paper is as follows: Section 2 provides a comprehensive review of the relevant literature. Section 3 presents an institutional background and data. In Section 4, we outline the empirical approach, and Section 5 includes a set of balancing tests. The main findings are presented in Section 6, followed by an exploration of the potential mechanisms underlying these results in Section 7 Section 8 presents a policy-relevant heterogeneity analysis. Finally, Section 9 concludes the study.

2 Literature Review

The literature on the influence of peer effects on the mental health of adolescents is relatively limited but expanding rapidly. This literature review serves multiple purposes. Firstly, it aims to explore the concept of peer effects, encompassing both general peer influence and the specific impact of gender composition on the mental well-being of adolescents. Secondly, it seeks to identify the potential mechanisms through which the gender

composition of a classroom can affect mental health, shedding light on the underlying processes. Thirdly, it aims to situate our paper's contributions within the existing peer effects literature.

Peer effects are rooted in the idea that an individual's outcome is partly determined by peer composition, with the classroom being a relevant peer environment (Henderson et al. 1978, Lazear 2001). The ideal experiment to study classroom peer effects would involve randomly assigning students to school classes, thereby ensuring that a given peer composition remains exogenous to factors affecting the outcome of interest. However, the random assignment of students across schools and classes can be challenging to guarantee due to the inherent difficulty in completely eliminating selection bias. While randomized control trials offer a rigorous approach to implementing random assignment, they come with the drawback of costly enforcement. Thus, applied research has adopted different empirical strategies to address confounding factors when estimating peer effects. The prevailing method employed to control for the likely self-selection bias of students into schools and classrooms is to incorporate fixed effects at either the student or school level (Sacerdote 2011). In essence, these methods enable researchers to approximate the causal impact of group background characteristics on different individual outcomes (Angrist & Pischke 2009, p. 193).

One of the strands of the peer-effects literature examines the potential influence of gender composition within classrooms on students' outcomes. The gender of students offers a compelling avenue for exploring spillover effects among peers due to its inherent random nature. Drawing upon insights primarily from the psychological literature, studies have demonstrated that children tend to conform to their gender identity, a crucial factor in group socialization (Martin & Ruble 2004). In an applied and economic context, Hoxby (2000) made pioneering contributions by employing a fixed-effects model to leverage natural variations in gender compositions across cohorts within the same school. Her findings revealed that an increase in the proportion of girls in a cohort led to improvements in mean peer test scores for both genders.[2] Similar patterns have been observed in studies that rely on random assignment of students (Whitmore 2005).

The interest in evaluating peer compositional externalities on mental health is shared

[2]Using a similar methodological approach, Black et al. (2013) find consistent effects for females when investigating cohort variations in the proportion of females. However, male students are disadvantaged by an increase in the proportion of females peers.

interdisciplinarily. Within the non-economic literature, particularly in the fields of psychology and medicine, there exists a body of work exploring the impact of peers on mental health (see, Bearman & Moody 2004, Rosenquist et al. 2011). In the economics field, a recent paper by Giulietti et al. (2022) examined whether exposure to peer depression in adolescence affects one's own depression in adulthood using survey data of U.S. adolescents. Notably, their findings indicate that a higher proportion of same-gender peers experiencing depression amplifies the long-term depression for females, while no significant effect is observed for males. Considering the influence of peer depression exposure during adolescence on long-term mental well-being, it is reasonable to infer that the effects of peer depression can extend to the classmates' current mental health. Furthermore, given the well-documented evidence of a higher incidence of mental ill-health among females (Nolen-Hoeksema 2001, Kuehner 2003, Van de Velde et al. 2010), one might anticipate that a greater presence of females in the classroom would result in an increased exposure to peer depression, consequently leading to a potential rise in mental health complaints. Hence, investigating alterations in the classroom gender composition is an intriguing avenue to explore as it holds the potential to be a significant determinant of mental health among adolescents. Moreover, the existence of additional gender disparities among students may influence the classroom environment, further emphasizing the relevance of this peer effect.

Closely related to our work, Getik & Meier (2022) investigate the role of gender composition as a determinant of mental health. They leverage natural variations in gender compositions across cohorts within the same school to examine the association between gender composition and the mental health of adolescents using Swedish administrative data. Their findings reveal that a higher proportion of female peers in the classroom increases the prevalence of mental ill-health, particularly among boys.

In order to gain deeper insights into how the gender composition of classrooms can influence mental health outcomes in adolescents, we explore the existing literature that investigates diverse mechanisms shaping the classroom environment. The influence of the school environment on mental health is multifaceted, encompassing factors such as the quality of interactions and the emotional and physical dynamics among its members (Wang & Degol 2016).

Firstly, studies examining the role of bullying find that students report less fighting

when the share of females in the class increases (Lavy & Schlosser 2011). Moreover, the negative relationship between bullying and mental health is widely recognized in the medical literature (Eyuboglu et al. 2021). Considering these findings, it is plausible that a decrease in fighting behavior potentially associated with a higher proportion of females in the class could be one potential channel through which gender composition influences mental health among adolescents.

Secondly, in terms of gender differences in friendships, empirical evidence suggests that boys typically possess larger same-gender networks than girls (Ladd 1983, Anelli & Peri 2019). Furthermore, research indicates that boys tend to foster closer relationships with their male classmates (Hill 2015). Consequently, an increase in the proportion of female classmates can result in a reduction of potential friends for boys. Considering the positive impact of having more friends on mental health (Ho 2016), alterations in gender composition may have distinct implications for male and female students.

Thirdly, empirical studies conducted in the field of psychology have provided evidence regarding the impact of peer influence on an individual's subjective perception of body image (Michael et al. 2014). Notably, research findings suggest that gender may play a significant role in shaping these perceptions, as males often express a desire to increase muscle mass, whereas females tend to aspire to reduce overall body size (Ata et al. 2007). Considering the well-established positive association between body image and mental well-being (Holsen et al. 2001), it follows that changes in the gender composition of a classroom may have varying implications for students based on their gender.

Fourthly, a body of research has demonstrated the influential role of peer relationships in promoting physical activity among individuals (Salvy et al. 2008). Notably, studies focusing on adolescents consistently indicate gender differences in engagement with various forms of organized play, with boys exhibiting a higher participation rate compared to girls (Moller et al. 1992, Zarbatany et al. 2000). Consequently, an increase in male peers has the potential to elevate the level of physical activity of students. Given the documented positive correlation between physical activity and improved mental well-being (Bize et al. 2007), changes in the gender composition of a classroom may affect mental health among students.

Fifthly, empirical research in economics has provided evidence that juvenile risky behavior, such as teenage drinking and smoking, may be influenced by peer-group effects,

as presented by Gaviria & Raphael (2001). Such behaviors have been shown to be associated with the mental health of adolescents (Treur et al. 2021), thus representing a plausible mechanism through which gender composition affects students' mental health outcomes.

Relative to the papers reviewed above, our thesis offers both corroborating and novel insights into the existing literature. We present further evidence that aligns with the findings of Getik & Meier (2022), suggesting that increased exposure to female peers is associated with a higher prevalence of mental health issues among boys, while girls appear to be less affected. An important aspect of our study is the utilization of the seldom-used HBSC dataset, which allows us to explore the influence of peer effects on Czech adolescents. By leveraging this additional data source, we provide suggestive evidence that supports the generalizability of the observed effects to other institutional contexts. These findings provide a solid foundation for future research to substantiate and validate our conclusions. Additionally, this study delves extensively into the identification of mechanisms, particularly associated with our dataset, to explain the reasons underlying the gender disparity in the estimated effects of gender composition on mental health. Thus, it further contributes to the existing body of literature evaluating the effects of classroom gender composition. Finally, our research methodology deviates slightly from existing peer effects studies. In contrast to previous studies that have commonly relied on exploiting idiosyncratic cohort variation in the proportion of female peers within schools across cohorts, our approach differs due to the absence of longitudinal data on the same schools. Specifically, we leverage the availability of class-level data across different grades within schools from the HBSC survey to exploit class grade-by-grade variation. Consequently, our findings reside within a distinct niche of empirical strategies employed in studying peer effects, enhancing the richness and diversity of the existing literature.

3 Institutional Background and Data

This paper employs Czech data obtained from the Health Behaviours of School-aged Children survey; a World Health Organization collaborative cross-national study conducted every four years. Specifically, we analyze the survey data from 2002, 2006, 2010, 2014, and 2018 to investigate the mental health of 11-, 13-, and 15-year-old students in

representative samples of schools. The response rates for schools in the survey cycles ranged from 75% to 99%, while the response rate for adolescent respondents surpassed 80% (Sigmund et al. 2020).[3] Some regions and types of schools exhibited slightly lower participation rates in Czechia which creates minor inconsistencies in the composition of the actual sample. To address these issues and to enhance the sample's representativeness, we employ the HBSC survey weights. This variable was designed to adjust the actual proportions in relation to the target distribution by taking into account factors such as federal state, type of school, gender, and age (Moor et al. 2020).

The sampling procedure of HBSC in the Czech Republic employs schools as the primary sampling units,[4] randomly selecting one class from each grade in every school representing the three age categories previously mentioned (Cosma et al. 2021). This approach permits us to exploit within-school variation across grades to examine the effect of the proportion of female peers on students' mental health and well-being. This analysis is made possible due to our rich dataset, encompassing background information and mental health indicators for students across grades within the same school. The data, obtained from the HBSC, is self-reported by the students themselves. We pooled the data from all study years in our analysis, resulting in 30,865 school-aged Czech adolescents (observations). In the following section, we provide an institutional context, explicate the estimation sample, and present a summary of descriptive statistics and definitions of key variables.

3.1 Schooling in the Czech Republic

In the Czech Republic, students are legally required to attend school starting from the age of 6 and complete 9 years of compulsory education (European Commission 2023). As a result, all students included in our sample are still within the compulsory education age

[3] A valid concern arises regarding the non-response observed as it may be attributed to students with significant mental health issues, thereby rendering our sample non-representative of the overall student population. However, it is worth noting that empirical observations from the 2014 Czech study wave indicate that student non-response can be ascribed not only to illness but also to a range of diverse factors, including participation in academic competitions or sporting events (Zidkova et al. 2020). Therefore, we can reasonably conclude that the absence from the survey does not solely stem from mentally ill students, mitigating the concern of a skewed sample.

[4] Typically, the HBSC survey employs a stratified random cluster sampling method, using school classes as the primary sampling unit. However, in cases where this sampling frame is impractical, schools are then used as the primary sampling unit (Currie et al. 2014). One of those niche cases is the Czech Republic.

range, which ensures that our sample is representative of the targeted population. Basic education in the Czech Republic is divided into two stages, the primary stage (covering grades 1-5) and the lower secondary stage (covering grades 6-9). The majority of students attend a single regular nine-year elementary school. During the second stage, which is the final compulsory phase of elementary education, students are provided with the option to either continue their studies in their current school, or attend an eight-year gymnasium (*osmileté gymnázium*), or a six-year gymnasium (*šestileté gymnázium*). The latter option is relatively uncommon and represents the sole instance where 6th-grade students attend a different institution from their 8th and 10th-grade counterparts (Government of Czechia 2020). This specific feature facilitates the collection of data for a class in each grade within the same school, a particularity that we employ in our methodology to account for school and grade fixed effects. Moreover, using schools as the primary sampling unit in Czechia ensures that most schools provide information on students from each grade.

3.2 Estimation Sample

To ensure the effectiveness of our estimation strategy, which relies on variation among classes from different grades within the same school, we excluded classes where students from different grades were combined. These particular observations were predominantly derived from small-scale schools that may encounter resource limitations and do not represent the typical school in the Czech Republic. As a result, we removed 37 individuals from our sample, resulting in a final pooled sample of 30,828 students for our primary analysis.

3.3 Summary Statistics and Variable Definitions

The present study comprises a sample of nationally representative Czech adolescents who were enrolled in 1,671 classes across 572 schools, during the 2002 (n=5,012; 51.9% girls), 2006 (n=4,779; 49.5% girls), 2010 (n=4,391; 51.5% girls), 2014 (n=5,082; 52.4% girls), and 2018 (n=11,564; 49.7% girls) academic years. Notably, the final sampling period in 2018 consists of a greater number of observations in comparison to the preceding periods, implying that our findings may be primarily influenced by this year. An overwhelming proportion of the schools included in our sample, amounting to approximately 92%,

Table 1: Descriptive statistics.

	Boys (49.3%)			Girls (50.7%)		
	N	Mean	SD	*N*	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dependent Variables					
Psychological complaints index $[0, 16]$	14,581	4.87	3.61	15,231	6.21	3.91
Feeling low (= 1 if at least about every month)	14,682	.43	.49	15,320	.60	.49
Feeling irritated (= 1 if at least about every month)	14,749	.77	.42	15,365	.85	.36
Feeling nervous (= 1 if at least about every month)	14,766	.75	.43	15,376	.82	.38
Sleeping difficulties (= 1 if at least about every month)	14,722	.44	.50	15,349	.55	.50
School environment						
Do you like school (= 1 if positive feeling)	15,067	.66	.48	15,532	.74	.44
Feeling pressured by school (= 1 if at least some)	15,056	.31	.46	15,514	.34	.47
Students work together (= 1 if agree)	14,953	.59	.49	15,468	.50	.50
Students are helpful (= 1 if agree)	14,919	.49	.50	15,440	.48	.50
Students accept me (= 1 if agree)	14,921	.73	.44	15,449	.67	.47
Bullied by others (= 1 at least once)	14,862	.17	.38	15,355	.17	.37
Behavior change and body positivity						
Think about body (= 1 if about the right size)	15,047	.45	.50	15,541	.50	.50
Days physically active per week	15,118	4.48	2.06	15,568	4.03	1.99
Ever tried smoking	15,000	.36	.48	15,507	.34	.47
Ever been drunk	14,827	.29	.45	15,371	.24	.43
	B. Individual-level Characteristics					
Traditional Family (= 1 if living with both parents)	15,195	.71	.46	15,633	.68	.47
Family affluence scale (FAS II)	14,820	5.48	1.93	15,297	5.28	1.99
Age	15,169	13.39	1.66	15,605	13.39	1.65
	C. Class-level Characteristics					
Proportion of female peers	15,195	.50	.13	15,633	.51	.12
Proportion of above-mean FAS II students	15,195	.49	.54	15,633	.49	.56
Significant change in class size indicator	15,195	.05	.21	15,633	.04	.20
Age outlier in class indicator	15,195	.03	.17	15,633	.03	.17
Class size	15,195	19.58	4.62	15,633	19.76	4.54
School size	15,195	56.02	10.26	15,633	56.34	10.21

Notes: – Authors' tabulation of data supplied by the Health Behaviour in School-aged Children survey. The table provides descriptive statistics on key variables for two distinct groups within the estimation sample: male students ($n = 15,195$) and female students ($n = 15,633$). School size considers the number of students sampled in each school.

have three classes that are sampled. The majority of the remaining schools in our sample (around 7%) are known as six-year academies, which are schools that span two grades starting from the 8th grade. The average number of students sampled within school and class is approximately 56 and 20, respectively.

Psychological complaints index. Panel A in Table 1 lists the student outcomes analyzed in this study. Notably, the mental well-being of the students serves as the primary outcome. The utilization of the psychological complaints index derived from the HBSC data as the primary outcome variable has been widely recognized as a reliable non-clinical

measure in the psychological literature (Ravens-Sieberer et al. 2022, Cosma et al. 2020, Erhart et al. 2009, Haugland & Wold 2001). The HBSC Symptom Checklist was employed to assess subjective psychological complaints by asking participants to report the frequency of symptoms experienced over the previous six months. The responses were assessed on a 5-point Likert scale, ranging from zero ("about every day") to four ("rarely or never"). In total, the psychological complaints index includes four symptoms: feeling low, feeling irritable, feeling nervous, and experiencing sleeping difficulties. For the construction of the index, the scoring for each symptom was reversed, and the individual scores were summed. As a result, the index ranges from 0 to 16, with higher values indicating a greater frequency of complaints. Panel A presents the mean values and standard deviations of the psychological complaints index and its individual symptoms for male and female students separately. To analyze the individual symptoms, we transformed the categorical variables into binary indicators. A value of zero denotes rare or non-existent occurrences of the symptom in the preceding six months, while a value of one signifies that the symptom has recurred more than once per month. Female students report higher levels of psychological complaints index score as a whole and for each specific symptom. The mean psychological complaints index score for male students is approximately 4.9 out of 16, whereas female students report a score of approximately 6.2. Consistent with prior research examining gender differentials (Nolen-Hoeksema 2001, Kuehner 2003, Van de Velde et al. 2010), our findings align with the prevailing evidence indicating higher levels of psychological complaints among female adolescents.

Mechanisms. In regard to the mechanisms examined in the present study, the collected student responses from the HBSC survey are presented in Table 1, Panel A. To facilitate comprehension, we have transformed each question in the survey into a binary outcome variable. These measures have been constructed to capture diverse potential dynamic channels that may be affected by an increase in the proportion of female peers, while influencing the mental health of students. The initial set of variables, categorized as school environment outcomes, reveals that female students are more likely to exhibit a positive attitude toward school than their male counterparts (74% versus 66%). However, both genders report comparable levels of experiencing school-related pressure, instances of bullying, and receiving support from fellow students. Conversely, male students tend

to exhibit a greater likelihood of perceiving acceptance from their peers in comparison to female students (73% versus 67%), and also express a higher agreement regarding collaborative efforts among classmates (59% versus 50%). The second set of variables pertains to the examination of a student's body image and the alterations in their behavioral patterns. Notable gender disparities are observed, particularly in perceived body image, wherein males report lower satisfaction levels compared to females (45% versus 55%). Additionally, male students display higher levels of engagement in physical activity than their female peers. While both genders exhibit low levels of recreational behaviors, male students are slightly more likely to have experimented with smoking and drinking. The gender discrepancies observed across various mechanisms suggest that the classroom environment may be profoundly influenced by its gender composition. Considering the potential influence of classmates' perceptions and habits on individual students, the proportion of female peers within the classroom could potentially exert an influence on student mental well-being, should these mechanisms indeed demonstrate a correlation with mental health.

Individual-level controls. In Table 1, Panel B we present statistics of the individual-level characteristics of the students in our sample, which constitute the initial set of controls employed in our analysis. Firstly, to determine whether the students reside in a traditional family, we leverage the information provided by HBSC, which records data on whether the respondent lives with or without a mother or father in the main residence. Hence, the variable indicates a value of one if the individual lives with both parents and zero otherwise. Within our sample, we note that approximately 70 percent of the students come from traditional families, and this proportion remains consistent across both genders. Secondly, as a means of measuring an individual's socioeconomic status, SES, we employ the family affluence scale (FAS) as a background variable. In particular, we employ the second iteration of FAS, namely FAS II, which comprises four items specifically designed to capture indicators of family affluence. These items inquire about the ownership of a car or other motorized vehicle within the family, the presence of a personal bedroom for the respondent, the number of computers possessed by the family, and the frequency of family vacations outside the Czech Republic in the preceding year. By aggregating the numerical values corresponding to each survey question, we derive

a family affluence index ranging from zero to nine. A score of zero denotes the lowest level of family affluence, while a score of nine represents the highest level of affluence. The family affluence scores of male and female students are relatively comparable, with male students exhibiting an average score of approximately 5.5 out of 9, while female students report an average score of around 5.3. Furthermore, age, the final individual-level characteristic, was derived directly from the dataset. We note that the average age of female and male students is closely similar.

Classroom-level controls. Table 1, Panel C displays the class-level characteristics employed as controls in our analysis. Of particular interest is our treatment variable, the proportion of female peers, which is constructed at the classroom level excluding the student himself or herself. Notably, male and female students both experience an approximately equal proportion of female peers in the classroom on average. The first class-level control that we include in the analysis is a variable indicating the proportion of students in the class with an above-average level of FAS II (i.e., 5.4 out of 9), excluding the individual oneself. Given our expectation that socioeconomic status plays a significant role in self-selection, we incorporate an additional control by considering the class-level socioeconomic status. This is done to account for the possibility that parents of students may consider not only their own FAS level but also the FAS levels of their child's classmates when making educational decisions. While we anticipate that parental selection based on across-grade differences in schools is likely minimal, including this control variable allows us to mitigate any potential remaining confounding factors. Table 1 shows that this proportion does not differ significantly between genders on average. Moreover, we follow the approach outlined in Brenøe & Zölitz (2020) for the subsequent controls. We include an indicator variable to capture whether schools have undergone substantial fluctuations in class size across grades. Specifically, if the class size of the grade below deviates by at least 50 percent (either decreasing or increasing) compared to the grade above, the indicator variable is assigned a value of one; otherwise, it takes a value of zero. We note that around five percent of both male and female students are enrolled in classes that encounter this specific situation. Furthermore, we also construct an indicator variable to detect the presence of age outliers within the classroom. These outliers are defined as individuals whose age exceeds the standard age for their corresponding grade

level by two years. Male and female students experience a comparable occurrence of age outliers in their classes, approximately three percent for each gender. Finally, as our last control variable at the class level, we examine the class size, with our sample typically documenting an average of 20 students per class. However, it is worth noting that this figure exhibits minor variations between genders, potentially attributable to the presence of exclusive male classes, which tend to be smaller and consequently contribute to a lower average class size for the male cohorts.

4 Empirical Strategy

In this study, we evaluate the impact of classroom gender composition on students' mental health outcomes. An inherent challenge in estimating peer effects arises from the self-selection and sorting of students across schools. It is conceivable that the gender composition within a classroom is correlated with school or grade-level characteristics that may also affect students' mental health. For instance, parents of female students, who often experience more pronounced mental health concerns,[5] may select schools based on their perceived ability to provide appropriate support. This non-random assignment of students into schools would undermine the correct identification of the causal relationship between the outcome of interest and classroom gender composition, resulting in downward biased estimates for female students.

To address the issue of sorting, we use the HBSC cross-sectional data and exploit idiosyncratic classroom variation in the proportion of female students within the same schools across grades. The basic idea underlying this approach is to compare student outcomes across different grades, who encounter a similar school environment and possess comparable background characteristics, with the only distinguishing factor being the proportion of female students in their respective classrooms. The observed variations in this proportion are assumed to be driven exclusively by natural fluctuations, making it an appropriate variable to estimate peer effects. By adopting this methodology, we aim to account for a significant source of potential confounding factors that could bias the estimation of gender composition effects. Our approach is founded on the assumption that

[5]We observe a similar pattern in our sample, with female students reporting an average score of 6.21, while male students reporting an average score of 4.87 in the psychological complaints index. This gender disparity is consistent across all symptoms, suggesting that females experience more severe self-reported mental health issues.

parental decisions to enroll their children in a particular school are unlikely to be based on their knowledge of across-grade differences in schools. This research methodology has been previously employed by various studies investigating peer effects among students (e.g., Hoxby 2000, Brenøe & Zölitz 2020, Giulietti et al. 2022). However, a distinguishing feature of our study is that we utilize pooled data from the HBSC study waves and exploit variation in classes from different grades within schools for a given year. In contrast, other studies have made use of panel data and exploited changes in cohort composition within the same school across time. Hence, we exploit class grade-by-grade variation, while other studies use cohort year-by-year variation.

We present the main specification of our model below:

$$\begin{aligned} Y_{isg} = \beta_1 \text{Female}_i \times \text{ShareFemPeers}_{isg} + \beta_2 \text{Male}_i \times \text{ShareFemPeers}_{isg} \\ + \beta_3 \text{Female}_i + \gamma X_{isg} + \alpha_s + \delta_g + \varepsilon_{isg}, \end{aligned} \tag{1}$$

where Y_{isg} is the mental health outcome of student i attending school s in grade g. To investigate the influence of gender composition on student outcomes, we employ $\text{ShareFemPeers}_{isg}$ as our treatment variable of interest. This variable measures the proportion of female students in a given classroom, excluding the individual himself or herself. In the spirit of Brenøe & Zölitz (2020), we analyze how the proportion of female peers interacts with the gender of individual students to study heterogeneity by gender. Specifically, we denote an individual i as female if Female_i is equal to 1 (=0 otherwise), and as male if Male_i is equal to 1 (=0 otherwise).

Therefore, the coefficients of interest, β_1, and β_2, represent the extent to which the mental health outcomes of female and male students, respectively, are impacted by the gender composition of their classroom. A positive estimate would suggest that having a higher proportion of female peers in the classroom results in an increased incidence of mental health complaints among students. Additionally, β_3 provides insight into the gender disparity in mental health outcomes conditional on controls.

The terms α_s and δ_g represent school and grade fixed effects. The inclusion of school-fixed effects aims to address any confounding factors shared among all students within the same school, while grade-fixed effects account for confounding factors specific to a particular grade level.[6] X_{isg} is a vector comprising the three individual-level and the

[6]As demonstrated by Angrist (2014), the application of the fixed effects approach in estimating peer

four classroom-level characteristics shown in Table 1.[7] By including additional controls in our analysis, we seek to reduce the impact of any remaining confounding factors. Furthermore, it allows us to infer the potential influence of unobservable variables on our results by examining changes in the coefficients of interest.[8] Finally, we denote the error term by ε_{isg}. Throughout the regression analysis, we cluster the standard errors at the school level to allow for heteroskedasticity and incorporate survey weights to make the sample more representative of the population.

The specification in Equation 1 relies on the key identifying assumption that variations in classroom gender composition across grades are exogenous to factors affecting students' mental health, conditional on school-fixed effects, grade-fixed effects, and additional controls. This assumption implies the following strict exogeneity condition, $E\left[\varepsilon_{isg}|X_{isg}, \alpha_s, \delta_g\right] = 0$. In other words, to yield causal estimates from our model, no unobserved variable should satisfy all of the following three conditions:

1. it varies by grade and is school-specific;
2. it is correlated with both the classroom gender composition and the students' mental health outcomes; and
3. it is not accounted for in the set of controls for student- and classroom-level characteristics that we include in the model.

We find it challenging to conceive of any plausible omitted variable that would meet all three of these conditions. Still, to evaluate the credibility of the absence of such factors, a series of balancing tests will be performed in the ensuing section, testing whether changes in classroom gender composition display any systematic correlation with student background characteristics when accounting for school and grade fixed effects. Considering that our assumption holds, classrooms with different gender compositions should not differ in any other dimension when controlling for school and grade fixed effects. We also assess the degree to which fluctuations in the proportion of female peers conform to expected lev-

effects entails a potential concern for measurement error bias. The classical measurement error has been shown to bias estimates upward (Feld & Zölitz 2017). However, it seems reasonable to assume that the gender of students is correctly reported. Moreover, the high student response rates observed in the multiple survey waves instill confidence in the accurate and consistent measurement of our treatment variable.

[7] We include a control for the class size up to the third polynomial along with the other class-level controls

[8] Our approach adheres to the methodology of Altonji et al. (2005), which employs the level of selection on observed explanatory variables as an indicator for the level of selection on unobservable variables.

els of natural variability. Overall, the results from these randomization checks appear to provide support for our primary identifying assumption.

5 Balance Tests

To assess the validity of our identifying assumption that peer gender composition is quasi-randomly assigned conditioning on grade and school fixed effects, we regress individual-level characteristics on the proportion of female peers while accounting for the aforementioned fixed effects. Our fundamental identifying assumption would be compromised if, for example, the proportion of female peers exhibits systematic correlations with students' background characteristics.

The proportion of female peers and gender. In our initial balancing test, we examine whether the proportion of female classmates correlates with the gender of individual students. To address the mechanical relationship between one's own gender and the gender of randomly matched peers, we adopt the correction method proposed by Guryan et al. (2009). Specifically, this relationship arises from the fact that an individual cannot be paired with oneself, which implies that the pool of potential peers for female students is drawn from a group with a marginally lower mean proportion of female individuals in comparison to that of male students. To adjust for this, we control for the school-level leave-out mean of the proportion of females peers across classes within the school.

Table 2 presents our results with three successive specifications in which we include individual-level controls and subsequently class-level controls, in addition to grade and school fixed effects. The findings reveal that there is no persistent correlation between students' own gender and the proportion of female classmates. The estimated coefficients appear to be insignificantly different from zero. The incorporation of controls does not yield any significant impact on our estimations.

The proportion of female peers and student background characteristics. Although we do not find any association between the classroom proportion of female peers and students' own gender, it remains conceivable that our treatment variable may be related to

Table 2: Balancing Test of Gender.

	Gender (= 1 if Female)		
	(1)	(2)	(3)
Proportion of Female Peers	-.002	-.004	-.003
	(.017)	(.019)	(.019)
Observations	30,828	30,070	30,070
p-value for female coefficient	.904	.830	.855
Grade and school fixed effects	✓	✓	✓
Individual-level controls	–	✓	✓
Class-level controls	–	–	✓

Notes: – Standard errors clustered at the school level are in parentheses. Each column in this table is one separate regression indicated by the checkmarks. Utilizing the Guryan, Kroft, and Notowidigdo (2009) correction method, we regress the gender on the class proportion of female peers, excluding the student himself or herself (school-class leave-out mean) and the school leave-out mean (across all grades). In our models, we incorporate individual-level controls, including the individual's family structure (coded as 1 if both parents reside in the primary household and 0 otherwise), the socio-economic status of the individual's family measured through FAS II, and the age of the individual. The class-level controls we employ are defined as follows: indicators for whether any student in the class is two years older than the standard age for their grade level (i.e., 13.5 for grade 1, 15.5 for grade 2, and 17.5 for grade 3), indicators for significant changes in class size compared to the above grade (i.e., a change of 50% or greater), indicators for the proportion of students in the class who have a socio-economic status above the mean, and class size up to the third polynomial.

other background characteristics of the students. To explore this possibility, we conduct additional regression analyses where we include the following three individual-level variables as dependent variables: the family affluence scale index, a traditional family dummy variable, and the student's age. Specifically, we estimate two different specifications for each characteristic, with each specification interacting the proportion of female peers with the student's gender, and controlling for grade and school fixed effects. Additionally, in the second specification, we include class-level controls to further account for potential confounding factors. Our findings are presented in Table 3.

The estimated effects on traditional family and student age are negligible and statistically insignificant for both genders, which is in line with our underlying assumption that the proportion of female classmates is quasi-randomly assigned conditioning on grade and school fixed effects. Regarding the family affluence scale, the estimated coefficient for girls is negative but statistically insignificant and of a small magnitude, which is also consistent with our key assumption. However, the negative correlation between the classroom proportion of female peers and the family influence is twice as large for boys compared to girls. Furthermore, this estimate is significant at the 5-percent level. The inclusion of

Table 3: Complete Balancing Tests Including All Individual-Level Variables.

	Family Affluence Scale		Traditional Family		Age of Student	
	(1)	(2)	(3)	(4)	(5)	(6)
Male × proportion female peers	-.320**	-.372**	-.005	-.010	.017	.018
	(.143)	(.156)	(.036)	(.036)	(.030)	(.028)
Female × proportion female peers	-.168	-.203	.013	.008	-.013	-.009
	(.143)	(.152)	(.035)	(.035)	(.030)	(.028)
Gender (= 1 if Female)	-.284***	-.294***	-.033	-.034	-.079***	-.080***
	(.096)	(.097)	(.023)	(.023)	(.019)	(.019)
Observations	30,117	30,117	30,828	30,828	30,774	30,774
Mean dependent variable male	5.484	5.484	.705	.705	13.477	13.477
Mean dependent variable female	5.276	5.276	.681	.681	13.388	13.388
Grade and school fixed effects	✓	✓	✓	✓	✓	✓
Class-level controls	–	✓	–	✓	–	✓

Notes: – Standard errors clustered at the school level are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column in this table is one separate regression indicated by the checkmarks. The outcome variable Family Affluence Scale is an SES index that ranges from zero to nine, a score of zero denotes the lowest level of family affluence, whereas a score of nine indicates the highest level of affluence. For Traditional Family an individual is coded as 1 if both parents reside in the primary household and 0 otherwise; and the Age of Student is derived from the data set. The class-level controls we employ are defined as follows: indicators for whether any student in the class is two years older than the standard age for their grade level (i.e., 13.5 for grade 1, 15.5 for grade 2, and 17.5 for grade 3), indicators for significant changes in class size compared to the above grade (i.e., a change of 50% or greater), indicators for the proportion of students in the class who have a socio-economic status above the mean, and class size up to the third polynomial.

class-level controls does not alter our interpretation of these results. Although this seems inconsistent with our initial assumption since it suggests that parents with higher family affluence would select schools with lower proportions of female peers for their children, it is noteworthy that the estimated coefficient remains relatively small. Specifically, a 10% increase in the proportion of female peers is associated with a decrease of 0.037 in the family affluence index, which has a mean value of approximately 5.5 for boys. This effect size represents roughly 2% of the standard deviation observed for this variable reported in Table 1. Moreover, it is important to acknowledge the possibility of observing significant results due to false discovery rates. When conducting a considerable number of regression analyses and evaluating multiple hypotheses, it is expected that some coefficients will exhibit statistical significance. Specifically, we anticipate that 5% of coefficients will present statistical significance at the 5-percent level, solely by chance. Therefore, one

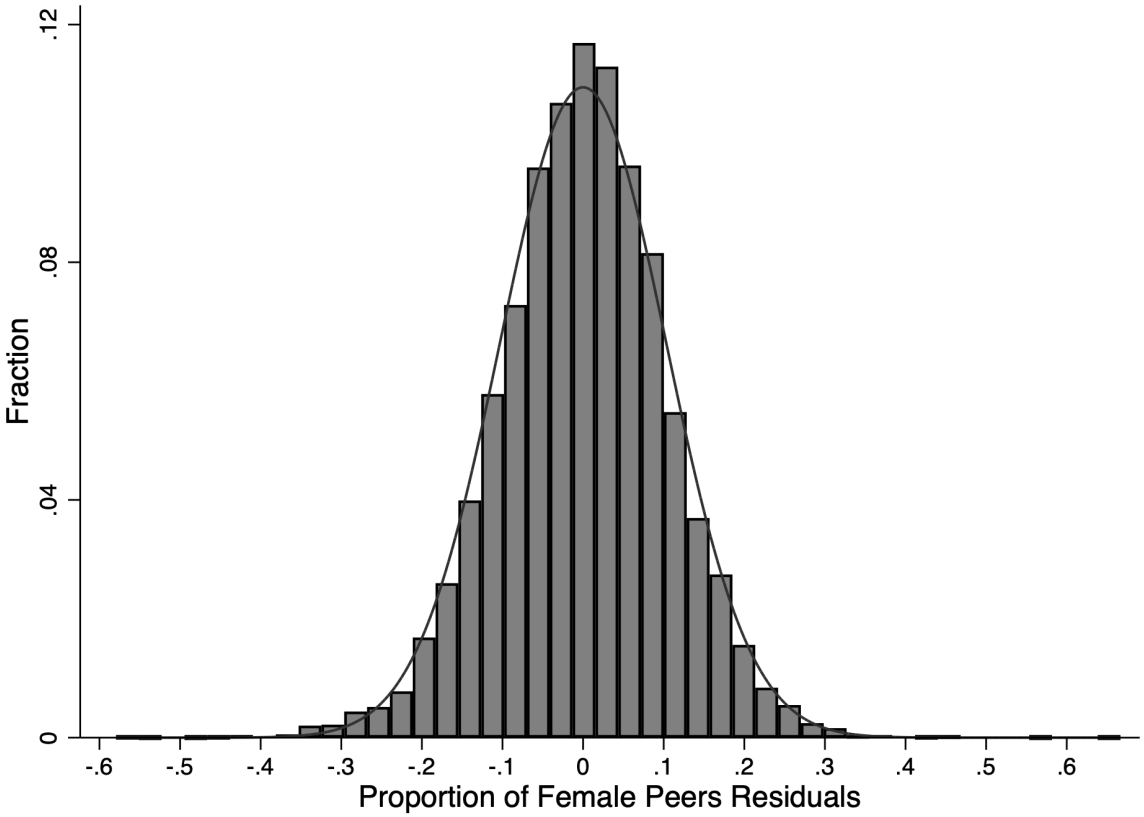

Figure 1: *Distribution of Residuals for the Proportion of Female Peers.* The figure presents a histogram depicting the distribution of residuals for the proportion of female peers in a sample of 1,671 classes across 572 schools. The overlaid curve illustrates a normal distribution that shares the same mean and standard deviation.

estimate should display statistical significance at the 5-percent significance level every 20 regressions. Given the negligible influence of class-level controls on our estimations, as well as the distinct estimates for males and females, we consider that we conducted a total of six regressions with one instance of 5-percent level significance. We believe that it remains plausible that this significance emerged from mere chance. Therefore, the results of these balance tests collectively lend support to the validity of our empirical approach.

Distribution and simulated distribution of the share of females. In order to further ensure the balance of our analysis, we conduct an examination into the extent to which the variation in the proportion of female peers can be treated as quasi-randomized, following the methodology outlined by Getik & Meier (2022), and Brenøe & Zölitz (2020). Specifically, we present in Figure 1 a histogram of the proportion of female peers at

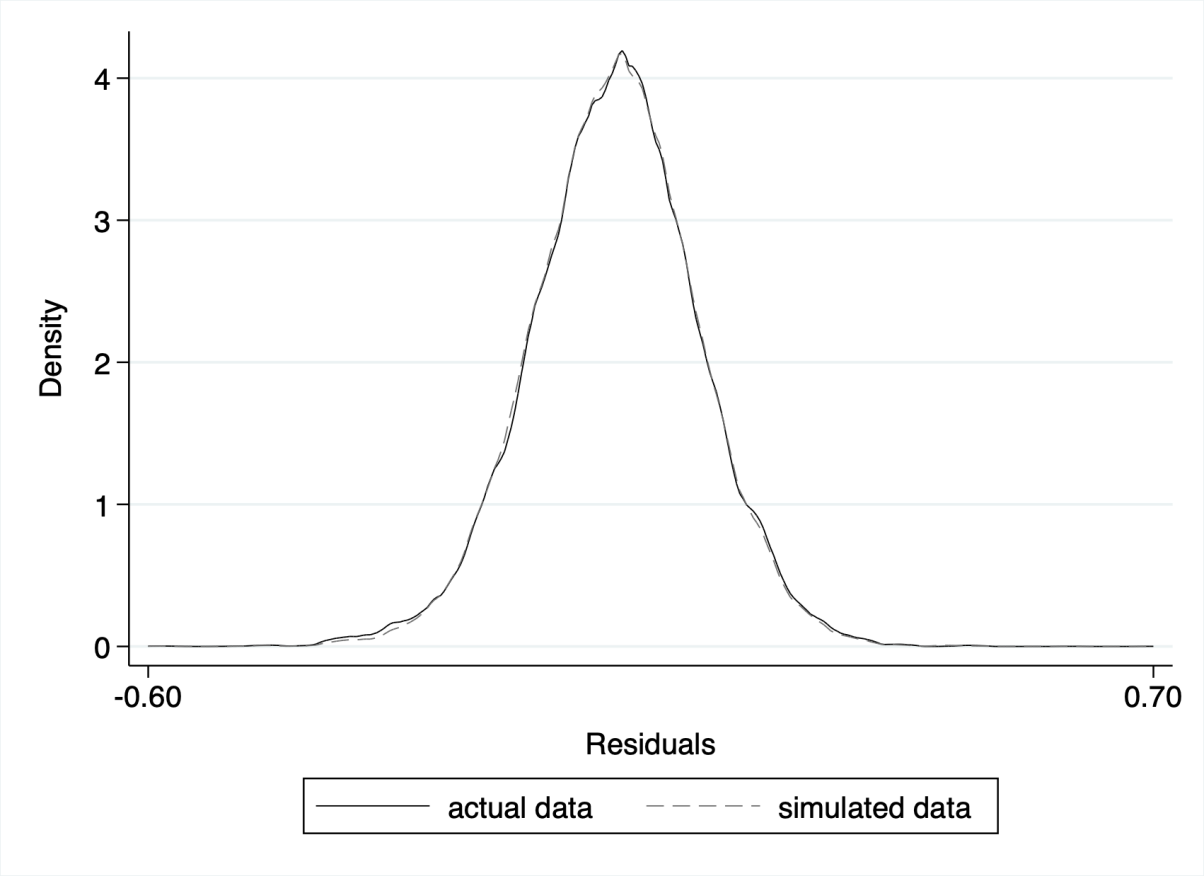

Figure 2: *Monte Carlo Simulation of the Share of Female Peers Across Classrooms Within Schools.* The depicted graph presents kernel density plots of residuals derived from regressions of the proportion of female peers on grade and school-fixed effects. The solid line pertains to the residuals obtained from a single regression using actual data, while the dashed line pertains to the residuals obtained from 1,000 regressions utilizing simulated data, where students were randomly allocated to classes within schools. The sample in both cases was confined to the number of classes per school and actual class size. The density computations were conducted through an Epanechnikov kernel using the optimal bandwidth of .011 in the actual data.

the class level, controlling for both grade and school fixed effects. The deviations from the mean in the proportion of female peers closely approximate a normal distribution, as evidenced by the overlaid normal curve with the same mean and variance. This finding lends support to the notion that the proportion of female peers may be considered as exogenous, conditional on grade and school fixed effects.

Monte Carlo simulation. In accordance with the methodology employed by Bietenbeck (2020), we seek to assess the extent to which the variation in the proportion of female peers in classrooms, as observed in our sample, can be regarded as equivalent to a

random allocation, given the prevailing data structure. To this end, we resort to a Monte Carlo simulation, wherein we randomly assign students to classes within their respective schools while keeping the total number of classes per school and the size of each class fixed. Subsequently, we estimate a regression model using the simulated data and regress the simulated proportion of female peers on school and grade fixed effects to obtain the corresponding residuals. By iterating this simulation 1,000 times and using the distribution of residuals collected in each regression we construct a kernel using the Epanechnikov function with an optimal bandwidth, minimizing the mean integrated squared error from the Gaussian distribution. Figure 2 displays the distribution of the residuals derived from the actual data alongside the simulation exercise, portrayed through two kernel density plots constructed with the aforementioned kernel density estimation. It is apparent that the two distributions are highly similar, lending support to the hypothesis that the proportion of female peers is randomly distributed across classes within schools in our sample, considering the underlying data structure.

In conclusion, the comprehensive set of balancing checks presented in this section offers robust support for our crucial identifying assumption. The results displayed in Table 2, Table 3, Figure 1 and Figure 2 indicate that accounting for grade and school fixed effects, the proportion of female peers may be considered exogenous, reinforcing our empirical strategy.

6 Main Results

Main analysis. Table 4 presents the estimates of the impact of gender peer effects on mental health. The first three columns report the estimates for the main outcome variable, the psychological complaints index. The subsequent columns (Columns 4 to 15) display the estimates for individual mental health symptoms that constitute the index across the various specifications.

To enhance the interpretability of the data, we have converted the four psychological symptoms we are examining – namely, feeling low, feeling irritable, feeling nervous, and experiencing sleeping difficulties – from categorical variables to binary ones. A value of zero represents rare or non-existent occurrences of the symptom in the past six months, whereas a value of one signifies that the symptom has recurred at least once a month. This

binary classification enables us to examine changes in the incidence of specific symptoms that persist over time, rather than scrutinizing increases in index values constructed by categorical variables, which may be challenging to interpret. We, therefore, employ a linear probability model to estimate peer effects, in line with the methodologies adopted by Giulietti et al. (2022) and Getik & Meier (2022). Although nonlinear models such as logit or probit offer a better fit, the marginal effects estimates differences between OLS and nonlinear models are of minimal significance (Angrist & Pischke 2009, p.107). In contrast, linear probability models possess notable advantages due to their standardized nature and ease of interpretability, rendering them more accessible for practical application. Furthermore, the inclusion of interaction terms in nonlinear models may pose challenges to their interpretability, and the utilization of fixed effects can further diminish the sample size (Beck 2018, Freedman 2008). Given these considerations, our decision to utilize linear probability models should not undermine the integrity of our findings, while facilitating a straightforward interpretation of the results.

To examine potential gender differences in the relationship between changes in the proportion of female peers and mental health, we introduce an interaction term between the student's gender and the proportion of female peers. To begin, we employ a basic model that regresses mental health on the proportion of female peers, as well as school and grade-fixed effects, as presented in Equation 1. We gradually introduce additional controls to each model specification starting with individual-level controls, followed by class-level controls. We also include, in Table 4, a test of the equality of coefficients between female and male students to compare if the effect varies across the two genders.

Overall, our findings suggest that male students experience moderately higher levels of psychological complaints when surrounded by a larger proportion of female peers. Our analysis reveals that this relationship is primarily driven by an increased propensity to feel low and nervous among male students. However, classroom gender composition does not appear to affect female students' psychological complaints.

An increase of 10 percentage points in the proportion of female peers in the classroom raises males' self-reported mental health complaints by 0.0585 index points, equivalent to a 1.2% increase from the baseline.[9] Although statistically significant at the 5-percent

[9]After running the regressions on samples restricted by survey years, we observed that the effect we found was primarily driven by the study waves in 2006 and 2010. This suggests that the negative influence of female peers on boys may not have remained consistent throughout the study period. Such variation in the effects could potentially be attributed to evolving dynamics of peer interactions across different

level, this estimate remains relatively modest in comparison to the mean of psychological complaints for male students. Incorporating individual and classroom-level controls into the model reduces the estimated coefficient, while the standard errors remain unaffected. Consequently, the resulting statistical significance level is attenuated to 10 percent. For female students, the estimates appear to be indistinguishable from zero across all specifications. While there seems to be a gender-based variation in the observed effects, our statistical analysis testing the equality of the coefficient between males and females reveals that the differences are not statistically significant at any conventional level.

Psychological complaints symptoms. With respect to mental health symptoms, we find that a 10 percentage point rise in the proportion of female peers in the classroom increases the prevalence of students persistently feeling low by 1.11 percentage points. This estimate is significant at the 1-percent level and represents a 2.6% deviation from the baseline. The effect of gender composition appears once more indistinguishable from zero for girls. Including additional controls does not substantially affect our estimates for either gender. Our hypothesis testing, assessing the equality of coefficients between male and female students, reveals a statistically significant disparity at the 5-percent level.

Additionally, the same change increases male (female) students' prevalence of being persistently nervous by 0.62 (-0.41) percentage points, a 0.8% (0.5%) change from the baseline. With the inclusion of controls, the estimates for both genders decrease slightly. This results in the male coefficient no longer attaining statistical significance at the 10-percent level, while the estimated coefficient for females achieves statistical significance at that same level. Despite the small magnitudes of the estimates, we observe a significant gender disparity in the effects, which can be attributed to the estimates having opposite signs. When it comes to irritability and sleeping difficulties among students, the impact of classroom gender composition appears to be of negligible magnitude across all specifications.

One potential caveat to our analysis of psychological symptoms is its sensitivity to the specification of variables. Our binary classification method relies solely on the variation between individuals who report persistent complaints from those who report no or rare complaints, without accounting for the intensity of their symptoms. To address this issue,

study waves, warranting further investigation and exploration in future research.

we employ the original categorical scale used in the index measure as an outcome variable in our regressions. Table A1 in the Appendix reports the estimates of the categorically scaled symptoms across all specifications. The gender disparity in the estimates of psychological complaints still appears to be mainly attributable to the variables of feeling low and nervous, although the discrepancy between the coefficients of those variables is less prominent. Moreover, irritability and sleeping difficulties now follow the same pattern with boys being comparatively more negatively affected than girls. Notably, the disparity in coefficients for irritability seems to have widened with this specification. One possible reason for this is that the variable has a relatively more uniform distribution across the categorically scaled responses, which allows for greater variation between the categories.

Taken together, our findings suggest that the mental well-being of boys is adversely affected by the proportion of female classmates, while girls appear to be less impacted. These results seem particularly noticeable in relation to symptoms of feeling low and nervous. Notably, our results align with the findings of Getik & Meier (2022), which similarly highlights that exposure to a greater number of female peers is associated with an increased incidence of mental health issues among boys, but not among girls. Although the inclusion of individual and classroom-level controls in our regression analysis generally results in a slight decrease in estimates for both genders, the small size of these changes suggests that any potential bias from remaining unobserved variables is likely to be negligible. Additionally, this anticipated decline in estimates for both genders is expected to maintain the existing gender gap observed in the effects.

Table 4: Impact of Classroom Gender Composition on Mental Health.

				Binary Components											
	Psycological Complaints Index [0, 16]			Feeling Low			Irritability			Nervosity			Sleep Difficulty		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Male × proportion female peers	.585**	.537*	.522*	.111***	.102***	.102***	.019	.017	.018	.062*	.051	.049	-.010	-.012	-.020
	(.292)	(.290)	(.290)	(.038)	(.039)	(.039)	(.032)	(.032)	(.032)	(.034)	(.034)	(.034)	(.040)	(.040)	(.040)
Female × proportion female peers	.009	.006	-.006	.012	.005	.005	-.004	-.001	-.001	-.041	-.047*	-.050*	.028	.024	.017
	(.326)	(.337)	(.338)	(.039)	(.041)	(.041)	(.029)	(.030)	(.030)	(.028)	(.028)	(.029)	(.039)	(.041)	(.041)
Gender (= 1 if Female)	1.623***	1.590***	1.590***	.217***	.215***	.215***	.089***	.084***	.084***	.119***	.116***	.116***	.091***	.092***	.092***
	(.207)	(.210)	(.209)	(.024)	(.025)	(.025)	(.020)	(.021)	(.021)	(.020)	(.020)	(.020)	(.027)	(.027)	(.027)
Observations	29,812	29,173	29,173	30,002	29,353	29,353	30,114	29,462	29,462	30,142	29,489	29,489	30,071	29,420	29,420
Mean dependent variable male	4.872	4.875	4.875	.428	.429	.429	.769	.770	.770	.754	.755	.755	.439	.438	.438
Mean dependent variable female	6.214	6.216	6.216	.595	.596	.596	.849	.850	.850	.821	.822	.822	.549	.550	.550
p-value from test for gender equality of proportion female peers	.147	.188	.199	.035	.045	.044	.547	.643	.632	.006	.011	.011	.454	.497	.482
Grade and school fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Individual-level controls	–	✓	✓	–	✓	✓	–	✓	✓	–	✓	✓	–	✓	✓
Class-level controls	–	–	✓	–	–	✓	–	–	✓	–	–	✓	–	–	✓

Notes: – Standard errors clustered at the school level are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column in this table is one separate regression indicated by the checkmarks. Psycological Complaints Index is an index ranging from 0 (lowest) to 16 (highest) derived from Feeling Low, Irritability, Nervosity, and Sleep Difficulty. The components were later made binary to indicate if an individual, during the last 6 months, has felt any psychological symptom more than about every month (equal 1) or never or rarely (equal 0). In our models, we incorporate individual-level controls, including the individual's family structure (coded as 1 if both parents reside in the primary household and 0 otherwise), the socio-economic status of the individual's family measured through FAS II and the age of the individual. The class-level controls we employ are defined as follows: indicators for whether any student in the class is two years older than the standard age for their grade level (i.e., 13.5 for grade 1, 15.5 for grade 2, and 17.5 for grade 3), indicators for significant changes in class size compared to the above grade (i.e., a change of 50% or greater), indicators for the proportion of students in the class who have a socio-economic status above the mean, and class size up to the third polynomial.

7 Mechanisms

In this section, we aim to delve into the underlying mechanisms that may explain why the proportion of female peers in a classroom has an impact on a student's mental well-being. Specifically, we seek to examine whether the gender composition of a classroom influences school environment factors, students' self-image, and students' behavioral change.

School environment. We focus on school environment dynamics that may be determinant drivers of students' mental well-being. Building upon previous research conducted by Hill (2015) and Ho (2016), we explore how variations in the proportion of female peers within the classroom can lead to changes in the school environment. Hill (2015) observed that male students tend to develop closer relationships with their male classmates, indicating the potential for positive social interactions. These interactions may have the potential to significantly impact mental health, particularly in alleviating symptoms such as feeling low. Furthermore, Ho (2016) sheds light on the influence of friendship networks on mental health. Specifically, his findings indicate that each additional friend is associated with a 0.018 standard deviation decrease in the prevalence of self-assessed depression symptoms.

To mitigate the risk of Type I errors associated with running a series of regression models and enhance statistical power, we adopt an index-based approach where a single regression is conducted to investigate the mechanisms related to the school environment. To develop the school environment index, we follow Field et al. (2016) by using the first component of principal component analysis. This index methodology effectively addresses the challenge of aggregating multiple outcomes into a single variable. Specifically, the procedure standardizes and assigns weights to the various responses related to school liking, students' enjoyment of being together, the helpfulness of students, students' feelings of acceptance, bullying, and school pressure. The resulting index uses the original four or five-scale categorical outcomes for school-environment dynamics. We reversed the scoring for bullying and school pressure, such that a higher score on the outcome variable reflects a positive school environment.

In our examination of the mechanisms, we adopt the same model as presented in Equation 1 and gradually incorporate individual and class-level controls. Consistent with the results part, we include an interaction term between gender and the proportion

Table 5: Gender Composition on School Environment Index.

	School Environment Index		
	(1)	(2)	(3)
Male × proportion female peers	-.072	-.057	-.061
	(.083)	(.083)	(.084)
Female × proportion female peers	.352***	.357***	.353***
	(.089)	(.089)	(.089)
Gender (= 1 if Female)	-.360***	-.346***	-.347***
	(.052)	(.052)	(.053)
Observations	29,718	29,194	29,194
p-value from test for gender equality of proportion female peers	<.001	<.001	<.001
Grade and school fixed effects	✓	✓	✓
Individual-level controls	–	✓	✓
Class-level controls	–	–	✓

Notes: – Standard errors clustered at the school level are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column in this table is one separate regression indicated by the checkmarks. The school-environment index is standardized to have a mean equal to zero and a standard deviation of 1.

of female peers to investigate any potential gender disparities in the effects. Additionally, we report the test for equality of coefficients between genders.

Table 5 presents estimates of how the proportion of female peers affects the school environment. All columns report the estimates for our school environment index, with each column including a different set of controls.

Our findings indicate that a 10 percentage point increase in the proportion of female classmates leads to a reduction of 0.0072 standard deviations in the school environment index for male students. However, that same change results in a rise of 0.0352 standard deviations in the school environment index for female students. The magnitude of this effect is particularly sizable for girls, with a coefficient that is statistically significant at the 1% level. The inclusion of controls does not substantially change our estimates, as presented in Columns 2 and 3. While our findings from this mechanism analysis cannot explain why gender compositions impact boys' mental health negatively, it can shed some light on why boys and girls were affected differently in their psychological complaints.

As evident from our test for gender equality between the coefficients, which shows a significant difference at the 1-percent level. While our analysis yielded results that are not in line with findings of Hill (2015) for male students, we found consistent evidence for female students. To be specific, an increase in the proportion of same-gender classmates has a positive impact on classroom behavior for girls, while no significant effect was observed for boys. This juxtaposition of findings provides suggestive evidence that the school environment serves as a channel through which changes in gender composition exert differential effects on mental health for the two genders.

To identify the specific components of the index leading to our results, we estimate the same regression for each component and present the results in Table A2 in the Appendix. Our findings suggest that the difference in estimated coefficients between the two genders is primarily influenced by students' enjoyment of school, the helpfulness of students, and students' feelings of acceptance.

Body image and Behavior change. Consistent with our approach to analyzing the impact of the school environment on mental well-being, we also investigate body image, physical activity, and drug use as possible mechanisms that may affect students' mental health.

Firstly, it is readily comprehensible that a student's self-perceived body image could exert a significant influence on their mental well-being. Furthermore, the classroom environment can serve as a platform where students' body perceptions may be influenced by their peers (Michael et al. 2014). Previous research investigating the internalization of body image by female and male students has indicated that females generally aspire for smaller overall body size, while males express a desire to increase muscle mass (Ata et al. 2007). These findings highlight the distinct ways in which both genders seek to modify their body image. Given the potential for peer influence, exposure to a higher proportion of female peers may impact students' body perceptions, providing a possible explanation for the influence of peer gender composition on mental health outcomes. We consider body image as a binary outcome variable, where it takes a value of one if perceived as the right size and zero otherwise. Our findings provide us with little evidence that changes in gender composition on students' mental health operate through its impact on self-perceived body image, as seen in Table 6.

Table 6: Gender Composition on Body Image and Behavioral Change.

	Think About Your Body			Physical Activity $[0, 7]$		
	(1)	(2)	(3)	(4)	(5)	(6)
Male × proportion female peers	-.024	-.026	-.020	-.536***	-.480***	-.490***
	(.036)	(.037)	(.037)	(.182)	(.180)	(.179)
Female × proportion female peers	-.005	-.016	-.013	-.170	-.165	-.173
	(.038)	(.038)	(.038)	(.174)	(.173)	(.174)
Gender (= 1 if Female)	.036	.035	.036	-.624***	-.592***	-.593***
	(.025)	(.025)	(.025)	(.111)	(.111)	(.111)
Observations	30,588	29,838	29,838	30,686	29,942	29,942
Mean dependent variable male	.454	.454	.454	4.478	4.484	4.484
Mean dependent variable female	.501	.500	.500	4.035	4.031	4.031
p-value from test for gender equality of proportion female peers	.685	.847	.872	.076	.126	.123
	Ever Smoked			Ever Drunk		
Male × proportion female peers	.044	.042	.033	.027	.027	.027
	(.039)	(.039)	(.040)	(.035)	(.035)	(.035)
Female × proportion female peers	.003	.011	.005	.032	.038	.037
	(.044)	(.043)	(.043)	(.037)	(.036)	(.037)
Gender (= 1 if Female)	-.001	-.004	-.006	-.045	-.040	-.040
	(.024)	(.024)	(.024)	(.023)	(.023)	(.023)
Observations	30,507	29,775	29,775	30,198	29,474	29,474
Mean dependent variable male	.365	.366	.366	.285	.285	.285
Mean dependent variable female	.344	.344	.344	.243	.243	.243
p-value from test for gender equality of proportion female peers	.337	.501	.547	.902	.804	.811
Grade and school fixed effects	✓	✓	✓	✓	✓	✓
Individual-level controls	–	✓	✓	–	✓	✓
Class-level controls	–	–	✓	–	–	✓

Notes: – Standard errors clustered at the school level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column in this table is one separate regression indicated by the checkmarks. For our outcome variable, each component was dichotomized to indicate agreement or disagreement with specific statements. In particular, for the Think About Your Body variable, we assigned a score of 1 if an individual reported that their body was "Too thin/too fat," and a score of 0 otherwise. The Physical Activity variable indicates whether an individual engaged in physical activity for at least 60 minutes in the last 7 days. The Ever Smoked variable was coded 1 if an individual reported having smoked at least once and 0 otherwise, while the Ever Drunk variable was coded 1 if an individual reported having been drunk at least once and 0 otherwise. In our models, we incorporate individual-level controls, including the individual's family structure (coded as 1 if both parents reside in the primary household and 0 otherwise), the socio-economic status of the individual's family, and the age of the individual. The class-level controls we employ are defined as follows: indicators for whether any student in the class is two years older than the standard age for their grade level (i.e., 13.5 for grade 1, 15.5 for grade 2, and 17.5 for grade 3), indicators for significant changes in class size compared to the above grade (i.e., a change of 50% or greater), indicators for the proportion of students in the class who have a socio-economic status above the mean, and class size up to the third polynomial.

Secondly, the impact of physical activity on psychological health is well documented (Bize et al. 2007). As female students typically engage in less physical activity than their male counterparts (Zarbatany et al. 2000),[10] an increase in the proportion of female peers in a class may potentially lead to a reduction in overall physical activity levels, if students are influenced by each other's habits. In fact, Salvy et al. (2008) finds that students are more likely to engage in intense physical activity when in the company of close friends. Given the tendency for students to develop closer relationships with same-sex classmates (Hill 2015), changes in gender composition may result in alterations in physical activity levels. Therefore, physical activity can serve as a compelling mechanism to explore its potential impact on mental well-being. The variable physical activity is defined as the number of days per week in which a student exercises for a duration of more than one hour. An increase of 10 percentage points in the proportion of female peers in the classroom decreases males' self-reported physical activity by 0.0536 days, a 1.2% change from the baseline. This estimate is significant at the 1-percent level across all specifications. Although female students also experience a decrease in physical activity with a higher proportion of female peers, the effect size is notably smaller and statistically non-significant. Once again, the inclusion of additional controls does not significantly alter our interpretation. Hence, we find some support for the notion that changes in gender composition have a negative impact on male students' mental health via its influence on students' physical activity. Moreover, we see notable disparities in the estimates between male and female students, although these differences do not reach statistical significance in all model specifications.

Thirdly, existing literature extensively documents the impact of drug use on the mental health of adolescents (Treur et al. 2021). Moreover, Gaviria & Raphael (2001) find suggestive evidence that both smoking and alcohol consumption display peer effects. Notably, they find a 10 percentage point rise in the proportion of smokers or drinkers within a peer group increases the probability of an individual engaging in smoking or alcohol consumption by 1.58% and 1.86%, respectively. Although no direct research has examined the influence of changes in the proportion of female peers on drug use, it is plausible to expect that such changes may contribute to mental ill-health. In examining drug use, we utilize data on students' smoking and drinking habits. Categorically, smoking is de-

[10]This pattern is consistent with our data, where female students, on average, exercise 4.035 days per week, while male students, on average exercise, 4.478 days per week

fined as a binary outcome equal to one if the student reports ever having smoked, and zero otherwise. Similarly, for drinking, we use a binary outcome variable that is equal to one if the student reports ever having been drunk, and zero otherwise. Table 6 presents our findings, which indicate that changes in gender composition have minimal impact on student's mental health through the mechanism of drug use.

After evaluating potential mechanisms, we have identified possible channels that may explain the negative effects of class gender composition on boys' mental health and its disparity with the effect observed on girls. Among these, we find that a decrease in physical activity for males with more female peers is one potential mechanism that could explain the increase in mental health complaints. While changes in the school environment could potentially explain the observed gender disparity in the estimates, it appears to be a less likely explanation for the negative effect on boys' mental health. This, however, relies on the assumption that changes in gender composition have an impact on student's mental health through these mechanisms. Furthermore, our analysis indicates that body perception and drug use are not likely to be the primary mechanisms driving the gender composition effects on mental health. Finally, we assess the impact of incorporating individual and classroom-level controls in our regression analysis, which generally results again in a slight downward adjustment of the estimated effects for both genders. Nonetheless, the modest magnitude of these changes implies that any potential bias stemming from unobserved variables should be small, thereby supporting the robustness of our findings.

8 Heterogeneity Analysis

In this study, we also explore any sources of heterogeneity in the impact of classroom gender composition by investigating whether the influence of female classmates varies based on the level of exposure. Following Giulietti et al. (2022), Table 7 presents the estimates of two non-linear peer-effects specifications. The first specification in Columns 1 and 3 employs an indicator variable to determine whether a student is subject to a proportion of female peers above the median. Columns 2 and 4 report the second specification where three indicator variables divide the classroom proportion of female peers into quartiles, the first quartile serving as the reference group. We consistently

Table 7: Heterogeneiry analysis: Psychological Complaints Index by Classroom Gender Composition Quartile.

Peer Proportion Quartile	Male		Female	
	(1)	(2)	(3)	(4)
Above median (.50)	.112		-.025	
	(.080)		(.085)	
2nd quartile (.43)		.129		-.096
		(.114)		(.112)
3rd quartile (.50)		.179		-.107
		(.119)		(.122)
4th quartile (.59)		.171		-.050
		(.115)		(.120)

Notes: – Standard errors clustered at the school level are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ This table presents the results of two regression models examining the relationship between the students' psychological complaints index and the proportion of female classmates. The first regression analysis (Columns 1 and 3) uses an indicator variable for whether an individual is exposed to an above-median proportion of female peers. The second regression analysis (Columns 2 and 4) employs three indicator variables to segment the proportion of female peers into quartiles, with the first quartile serving as the reference group. The quartile boundaries are reported in parentheses. To investigate the difference in effects between males and females, we interact the gender of students with the proportion of female peers. All outcomes are estimated using a specification that accounts for grade and school fixed effects, along with individual and classroom-level controls.

report separate estimates for males and females by interacting students' own gender with the proportion of female classmates. Finally, all specifications account for grade and school fixed effects, along with individual and classroom-level controls.

Our findings indicate that an above-median proportion of female peers is associated with a rise in the psychological complaints index score of male students by 0.112 compared to their counterparts who have a below-median proportion of female peers. This effect size corresponds to a 2.3% deviation from the baseline, although this estimate fails to reach statistical significance at conventional levels. For females, the effect seems negligible and indistinguishable from zero

According to the quartile dummies specification, the coefficients for male students exhibit a gradual and positive trend across the quartiles, except for a slight decrease between the 3rd and 4th quartiles. This suggests that boys' mental health outcomes are most negatively affected when they are in a class where the proportion of female classmates belongs to the third quartile, compared to the other quartiles. In particular, being in this situation is associated with a 0.179 increase in the psychological complaints

index score, compared to being in the first quartile. Interestingly, for female students, the opposite is observed as they experience the most positive impact on their mental health when the proportion of female classmates falls within the third quartile. In this instance, exposure to a third-quartile proportion of female classmates is associated with a 0.107 decrease in the psychological complaints index score, relative to the first quartile. However, it is noteworthy that none of these estimates are statistically significant at conventional levels.

In the context of policymaking aimed only at reducing the negative mental health outcomes associated with the peer environment, the findings of this study suggest that balancing class gender composition from the 4th quartile range (approximately 60% to 100% female peers) to the 3rd quartile range (approximately 50% to 60% female peers) may not be advisable if male students' mental health complaints are valued as important as those of females students. Instead, it may be more effective to place male students in a classroom with fewer female peers or even in a single-gender classroom to counteract any detrimental impact that female peers may have on their mental health. On the other hand, considering the existing vulnerability of female students to mental health issues, it may be advisable to balance class gender composition between the 2nd and the 3rd quartile range (approximately 40% to 60%). While male students appear to experience greater negative effects on their mental health compared to their female counterparts within this range, it may still be worthwhile to consider the potential positive impact on the mental health of female students, particularly given the prevalence of mental health challenges among this population. Nonetheless, it is imperative to recognize that the influence of classroom gender composition on student mental health is only one among several potential outcomes that may be impacted by peer effects. Therefore, a comprehensive approach should be adopted when assessing the efficacy and desirability of any policy intervention aimed at altering the classroom gender composition.

9 Conclusion

In this paper, we investigate the classroom gender composition effects on adolescents' mental health in the Czech Republic using HBSC data. Our results show that a higher proportion of female peers in the classroom may have a detrimental impact on the mental

health of male students, whereas no comparable effects are observed among female students. While the observed effect on male students is modest in magnitude, our findings align with the research conducted by Getik & Meier (2022), who suggest that gender peer effects exhibit a similar pattern in the context of Sweden. Moreover, our analysis reveals that the rise in psychological complaints among boys is primarily driven by an upsurge in symptoms such as feeling low and nervous.

Gender composition may impact the classroom environment profoundly, as there exist fundamental disparities in students' behaviors and perceptions based on gender. Our investigation delves into the various mechanisms through which gender composition can impact mental well-being. Specifically, our findings indicate that a decline in physical activity levels among boys, due to a higher proportion of girls in the classroom, may serve as a potential channel contributing to the increase in mental health complaints. This conclusion is supported by the well-documented benefits of physical activity on mental well-being. Furthermore, while changes in the students' perceived school environment may account for the gender disparities observed in our estimates, they are inadequate in explaining the rise in boys' mental health complaints. Finally, our research suggests that body perception and drug use do not play significant roles as driving factors behind the observed effects.

The model employed in this study is based on the key identifying assumption that the classroom proportion of female peers is exogenous to factors affecting students' mental health conditional on grade, and school fixed effects. Our balancing tests collectively lend support to this empirical strategy. We further evaluate concerns regarding confounding variables that may differ across grade levels within schools, by assessing how the gradual inclusion of controls influences our estimates. While the inclusion of controls generally leads to a slight decrease in our estimates, implying a potential upward bias, the negligible magnitude of these changes suggests that any bias from remaining unobserved variables is likely to be small.

In terms of policy implication, our results indicate that reducing the number of girls in the classroom may not always be recommended, as its effectiveness depends on the initial gender composition. In fact, our research reveals that male students experience the most adverse consequences when the proportion of female peers falls within the third quartile range (approximately 50% to 60% female peers), compared to the other quartiles. In-

stead, our results suggest that promoting single-gender classes or classrooms with a lower proportion of females may potentially be advantageous for boys' mental health. Regarding female students, although the effects are of lesser magnitude, our research indicates that the most positive impacts on their mental health occur when the gender composition lies within the 3rd quartile, compared to the other quartiles. Thus, a balanced proportion of the classroom may lead to beneficial results on girls' mental health. Nevertheless, there are other factors that could explain why this type of class is beneficial for the students. For instance, gender parity in the classroom may promote gender equality as it facilitates increased interaction between boys and girls, thereby yielding potential societal benefits. Additionally, Hoxby (2000) also demonstrated that the proportion of female peers can exert an influence on students' academic performance. Therefore, a nuanced perspective is needed when designing a policy targeting the classroom gender composition.

It is noteworthy to bear in mind that our results are situated within a specific context, as we examine Czech adolescents across various HBSC survey waves from 2002 to 2018. Nonetheless, the investigation conducted by Getik & Meier (2022) employing Swedish administrative data provides suggestive evidence that supports the generalizability of the observed effect in other institutional settings. It would be of interest for further research to assess the external validity of these conclusions.

Throughout our analysis, we employ self-reported outcomes on psychological well-being, rather than physician assessments. Consequently, our findings rely on a non-clinical assessment of mental health perceptions among adolescents, which may potentially deviate from an accurate representation of clinical mental health diagnoses. Moreover, our ability to examine long-term effects on mental health is compromised by our data, as we only observe students at a single point in time. To address these limitations, an alternative approach utilizing administrative data might gather long-term information on students' mental diagnoses and other outcomes, while also providing a larger sample size.

References

Altonji, J. G., Elder, T. E. & Taber, C. R. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of political economy* **113**(1), 151–184.

Anelli, M. & Peri, G. (2019), 'The effects of high school peers' gender on college major, college performance and income', *The Economic Journal* **129**(618), 553–602.

Angrist, J. D. (2014), 'The perils of peer effects', *Labour Economics* **30**, 98–108.

Angrist, J. D. & Pischke, J.-S. (2009), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Ata, R. N., Ludden, A. B. & Lally, M. M. (2007), 'The effects of gender and family, friend, and media influences on eating behaviors and body image during adolescence', *Journal of Youth and Adolescence* **36**, 1024–1037.

Bearman, P. S. & Moody, J. (2004), 'Suicide and friendships among american adolescents', *American journal of public health* **94**(1), 89–95.

Beck, N. (2018), 'Estimating grouped data models with a binary dependent variable and fixed effects: What are the issues', *arXiv preprint arXiv:1809.06505* .

Becker, G. S. (2009), *Human capital: A theoretical and empirical analysis, with special reference to education*, University of Chicago press.

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

Bize, R., Johnson, J. A. & Plotnikoff, R. C. (2007), 'Physical activity level and health-related quality of life in the general adult population: a systematic review', *Preventive medicine* **45**(6), 401–415.

Black, S. E., Devereux, P. J. & Salvanes, K. G. (2013), 'Under pressure? the effect of peers on outcomes of young adults', *Journal of Labor Economics* **31**(1), 119–153.

Brenøe, A. A. & Zölitz, U. (2020), 'Exposure to more female peers widens the gender gap in stem participation', *Journal of Labor Economics* **38**(4), 1009–1054.

Cawley, J. (2004), 'The impact of obesity on wages', *Journal of Human resources* **39**(2), 451–474.

Cosma, A., Költő, A., Badura, P., Winkler, P., Kalman, M. et al. (2021), 'Time trends in adolescent mental wellbeing in the czech republic between 2002 and 2018: gender, age and socioeconomic differences', *Central European Journal of Public Health* **29**(4), 271–

278.

Cosma, A., Stevens, G., Martin, G., Duinhof, E. L., Walsh, S. D., Garcia-Moya, I., Költő, A., Gobina, I., Canale, N., Catunda, C. et al. (2020), 'Cross-national time trends in adolescent mental well-being from 2002 to 2018 and the explanatory role of schoolwork pressure', *Journal of Adolescent Health* **66**(6), S50–S58.

Currie, C., Inchley, J., Molcho, M., Lenzi, M., Veselska, Z. & Wild, F. (2014), 'Health behaviour in school-aged children (hbsc) study protocol: background, methodology and mandatory items for the 2013/14 survey'.

Currie, J. & Stabile, M. (2006), 'Child mental health and human capital accumulation: the case of adhd', *Journal of health economics* **25**(6), 1094–1118.

Currie, J., Stabile, M., Manivong, P. & Roos, L. L. (2010), 'Child health and young adult outcomes', *Journal of Human resources* **45**(3), 517–548.

Ding, W., Lehrer, S. F., Rosenquist, J. N. & Audrain-McGovern, J. (2009), 'The impact of poor health on academic performance: New evidence using genetic markers', *Journal of health economics* **28**(3), 578–597.

Eisenberg, D., Golberstein, E. & Hunt, J. B. (2009), 'Mental health and academic success in college', *The BE Journal of Economic Analysis & Policy* **9**(1).

Erhart, M., Ottova, V., Gaspar, T., Jericek, H., Schnohr, C., Alikasifoglu, M., Morgan, A., Ravens-Sieberer, U. & Group, H. P. H. F. (2009), 'Measuring mental health and well-being of school-children in 15 european countries using the kidscreen-10 index', *International journal of public health* **54**, 160–166.

European Commission (2023), 'National education systems - czech republic'.
URL: *https://eurydice.eacea.ec.europa.eu/national-education-systems/czech-republic/overview*

Eyuboglu, M., Eyuboglu, D., Pala, S. C., Oktar, D., Demirtas, Z., Arslantas, D. & Unsal, A. (2021), 'Traditional school bullying and cyberbullying: Prevalence, the effect on mental health problems and self-harm behavior', *Psychiatry research* **297**, 113730.

Feld, J. & Zölitz, U. (2017), 'Understanding peer effects: On the nature, estimation, and channels of peer effects', *Journal of Labor Economics* **35**(2), 387–428.

Field, E., Jayachandran, S., Pande, R. & Rigol, N. (2016), 'Friendship at work: Can peer effects catalyze female entrepreneurship?', *American Economic Journal: Economic Policy* **8**(2), 125–153.

Freedman, D. A. (2008), 'On regression adjustments in experiments with several treatments'.

Gaviria, A. & Raphael, S. (2001), 'School-based peer effects and juvenile behavior', *Review of Economics and Statistics* **83**(2), 257–268.

Getik, D. & Meier, A. N. (2022), 'Peer gender and mental health', *Journal of Economic Behavior & Organization* **197**, 643–659.

Giulietti, C., Vlassopoulos, M. & Zenou, Y. (2022), 'Peers, gender, and long-term depression', *European Economic Review* **144**, 104084.

Government of Czechia (2020), 'Střední vzdělávání'.
URL: *https://portal.gov.cz/en/informace/stredni-vzdelavani-INF-95*

Grossman, M. (1972), 'On the concept of health capital and the demand for health', *Journal of Political economy* **80**(2), 223–255.

Guryan, J., Kroft, K. & Notowidigdo, M. J. (2009), 'Peer effects in the workplace: Evidence from random groupings in professional golf tournaments', *American Economic Journal: Applied Economics* **1**(4), 34–68.

Haugland, S. & Wold, B. (2001), 'Subjective health complaints in adolescence—reliability and validity of survey methods', *Journal of adolescence* **24**(5), 611–624.

Henderson, V., Mieszkowski, P. & Sauvageau, Y. (1978), 'Peer group effects and educational production functions', *Journal of Public Economics* **10**(1), 97–106.

Hill, A. J. (2015), 'The girl next door: The effect of opposite gender friends on high school achievement', *American Economic Journal: Applied Economics* **7**(3), 147–177.

Ho, C. Y. (2016), 'Better health with more friends: the role of social capital in producing health', *Health Economics* **25**(1), 91–100.

Holsen, I., Kraft, P. & Røysamb, E. (2001), 'The relationship between body image and depressed mood in adolescence: A 5-year longitudinal panel study', *Journal of health psychology* **6**(6), 613–627.

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

Kuehner, C. (2003), 'Gender differences in unipolar depression: an update of epidemiological findings and possible explanations', *Acta Psychiatrica Scandinavica* **108**(3), 163–174.

Ladd, G. W. (1983), 'Social networks of popular, average, and rejected children in school

settings', *Merrill-Palmer Quarterly (1982-)* pp. 283–307.

Lavy, V. & Schlosser, A. (2011), 'Mechanisms and impacts of gender peer effects at school', *American Economic Journal: Applied Economics* **3**(2), 1–33.
URL: *https://www.aeaweb.org/articles?id=10.1257/app.3.2.1*

Lazear, E. P. (2001), 'Educational production', *Technology* p. 777.

Lindeboom, M., Lundborg, P. & Van Der Klaauw, B. (2010), 'Assessing the impact of obesity on labor market outcomes', *Economics & Human Biology* **8**(3), 309–319.

Martin, C. L. & Ruble, D. (2004), 'Children's search for gender cues: Cognitive perspectives on gender development', *Current directions in psychological science* **13**(2), 67–70.

Michael, S. L., Wentzel, K., Elliott, M. N., Dittus, P. J., Kanouse, D. E., Wallander, J. L., Pasch, K. E., Franzini, L., Taylor, W. C., Qureshi, T. et al. (2014), 'Parental and peer factors associated with body image discrepancy among fifth-grade boys and girls', *Journal of youth and adolescence* **43**, 15–29.

Moller, L. C., Hymel, S. & Rubin, K. H. (1992), 'Sex typing in play and popularity in middle childhood', *Sex roles* **26**, 331–353.

Moor, I., Winter, K., Bilz, L., Bucksch, J., Finne, E., John, N., Kolip, P., Paulsen, L., Ravens-Sieberer, U., Schlattmann, M. et al. (2020), 'The 2017/18 health behaviour in school-aged children (hbsc) study–methodology of the world health organization's child and adolescent health study', *Journal of Health Monitoring* **5**(3), 88.

Mushkin, S. J. (1962), 'Health as an investment', *Journal of political economy* **70**(5, Part 2), 129–157.

Nolen-Hoeksema, S. (2001), 'Gender differences in depression', *Current directions in psychological science* **10**(5), 173–176.

OECD (2021), *A New Benchmark for Mental Health Systems*.
URL: *https://www.oecd-ilibrary.org/content/publication/4ed890f6-en*

Orben, A., Tomova, L. & Blakemore, S.-J. (2020), 'The effects of social deprivation on adolescent development and mental health', *The Lancet Child & Adolescent Health* **4**(8), 634–640.

Ravens-Sieberer, U., Erhart, M., Devine, J., Gilbert, M., Reiss, F., Barkmann, C., Siegel, N. A., Simon, A. M., Hurrelmann, K., Schlack, R. et al. (2022), 'Child and adolescent mental health during the covid-19 pandemic: results of the three-wave longitudinal copsy study', *Journal of Adolescent Health* **71**(5), 570–578.

Rosenquist, J. N., Fowler, J. H. & Christakis, N. A. (2011), 'Social network determinants of depression', *Molecular psychiatry* **16**(3), 273–281.

Sacerdote, B. (2011), Peer effects in education: How might they work, how big are they and how much do we know thus far?, *in* 'Handbook of the Economics of Education', Vol. 3, Elsevier, pp. 249–277.

Salvy, S.-J., Bowker, J. W., Roemmich, J. N., Romero, N., Kieffer, E., Paluch, R. & Epstein, L. H. (2008), 'Peer influence on children's physical activity: an experience sampling study', *Journal of pediatric psychology* **33**(1), 39–49.

Sigmund, E., Sigmundová, D., Badura, P., Voráčová, J., Vladimir, H., Hollein, T., Pavelka, J., Pžová, Z. & Kalman, M. (2020), 'Time-trends and correlates of obesity in czech adolescents in relation to family socioeconomic status over a 16-year study period (2002–2018)', *BMC Public Health* **20**, 1–12.

Treur, J. L., Munafò, M. R., Logtenberg, E., Wiers, R. W. & Verweij, K. J. (2021), 'Using mendelian randomization analysis to better understand the relationship between mental health and substance use: a systematic review', *Psychological medicine* **51**(10), 1593–1624.

Van de Velde, S., Bracke, P. & Levecque, K. (2010), 'Gender differences in depression in 23 european countries. cross-national variation in the gender gap in depression', *Social science & medicine* **71**(2), 305–313.

Wang, M.-T. & Degol, J. L. (2016), 'School climate: A review of the construct, measurement, and impact on student outcomes', *Educational psychology review* **28**(2), 315–352.

Whitmore, D. (2005), 'Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment', *American Economic Review* **95**(2), 199–203.

Zarbatany, L., McDougall, P. & Hymel, S. (2000), 'Gender-differentiated experience in the peer culture: links to intimacy in preadolescence', *Social Development* **9**(1), 62–79.

Zidkova, R., Glogar, P., Polackova Solcova, I., P. van Dijk, J., Kalman, M., Tavel, P. & Malinakova, K. (2020), 'Spirituality, religious attendance and health complaints in czech adolescents', *International journal of environmental research and public health* **17**(7), 2339.

A Appendix

Table A1: Robustness Test of the Psychological Complaints Symptoms.

	Feeling Low			Irritability		
	(1)	(2)	(3)	(4)	(5)	(6)
Male × proportion female peers	.203**	.181**	.183**	.121	.117	.133
	(.085)	(.086)	(.087)	(.092)	(.092)	(.092)
Female × proportion female peers	.073	.063	.063	-.009	.009	.022
	(.105)	(.107)	(.108)	(.105)	(.105)	(.106)
Gender (= 1 if Female)	.503***	.494***	.496***	.347***	.332***	.334***
	(.063)	(.064)	(.063)	(.062)	(.063)	(.063)
Observations	30,002	29,353	29,353	30,114	29,462	29,462
Mean dependent variable male	.773	.773	.773	1.522	1.524	1.524
Mean dependent variable female	1.212	1.212	1.212	1.809	1.810	1.810
p-value from test for gender equality of proportion female peers	.283	.338	.328	.273	.367	.353
	Nervosity			Sleep Difficulty		
Male × proportion female peers	.172*	.148	.142	.071	.063	.035
	(.102)	(.102)	(.103)	(.117)	(.118)	(.118)
Female × proportion female peers	-.054	-.063	-.066	-.003	-.012	-.035
	(.099)	(.103)	(.103)	(.119)	(.124)	(.123)
Gender (= 1 if Female)	.400***	.323***	.392***	.354***	.353***	.347***
	(.006)	(.068)	(.068)	(.081)	(.082)	(.082)
Observations	30,142	29,489	29,489	30,071	29,420	29,420
Mean dependent variable male	1.596	1.599	1.599	.975	.974	.974
Mean dependent variable female	1.884	1.885	1.885	1.293	1.294	1.294
p-value from test for gender equality of proportion female peers	.076	.105	.108	.628	.631	.652
Grade and school fixed effects	✓	✓	✓	✓	✓	✓
Individual-level controls	–	✓	✓	–	✓	✓
Class-level controls	–	–	✓	–	–	✓

Notes: – Standard errors clustered at the school level are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column in this table is one separate regression indicated by the checkmarks. In this table, we run the same regressions as in the main results but now keeping the original five-scaled categorical symptoms. In our models, we incorporate individual-level controls, including the individual's family structure (coded as 1 if both parents reside in the primary household and 0 otherwise), the socio-economic status of the individual's family measured through FAS II, and the age of the individual. The class-level controls we employ are defined as follows: indicators for whether any student in the class is two years older than the standard age for their grade level (i.e., 13.5 for grade 1, 15.5 for grade 2, and 17.5 for grade 3), indicators for significant changes in class size compared to the above grade (i.e., a change of 50% or greater), indicators for the proportion of students in the class who have a socio-economic status above the mean, and class size up to the third polynomial.

Table A2: Gender Composition on School and Student Dynamics.

	School Liking			Student Together			Student Helpful		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male × proportion female peers	-.065	-.069*	-.072*	-.003	.000	.000	.053	.058	.056
	(.041)	(.041)	(.041)	(.051)	(.052)	(.052)	(.049)	(.048)	(.049)
Female × proportion female peers	.039	.040	.036	.028	.026	.026	.142***	.133***	.130**
	(.043)	(.044)	(.044)	(.054)	(.054)	(.054)	(.050)	(.050)	(.050)
Gender (= 1 if Female)	.031	.027	.027	-.107***	-.102***	-.102***	-.054**	-.044*	-.043*
	(.025)	(.025)	(.025)	(.027)	(.027)	(.027)	(.026)	(.026)	(.026)
Observations	30,599	29,951	29,951	30,421	29,827	29,827	30,359	29,767	29,767
Mean dependent variable male	.655	.656	.656	.591	.592	.592	.486	.487	.487
Mean dependent variable female	.740	.740	.740	.500	.500	.500	.480	.480	.480
p-value from test for gender equality of proportion female peers	.030	.024	.025	.548	.610	.620	.073	.132	.138
	Student Accept			Been Bullied			School Pressure		
Male × proportion female peers	-.047	-.041	-.042	.030	.031	.028	.055	.059	.057
	(.036)	(.037)	(.037)	(.030)	(.030)	(.030)	(.041)	(.041)	(.041)
Female × proportion female peers	.173***	.174***	.173***	-.023	-.017	-.020	.056	.064	.062
	(.040)	(.040)	(.040)	(.033)	(.032)	(.032)	(.043)	(.043)	(.043)
Gender (= 1 if Female)	-.183***	-.178***	-.178***	.019	.017	.017	.028	.030	.029
	(.023)	(.023)	(.024)	(.020)	(.020)	(.020)	(.026)	(.026)	(.026)
Observations	30,370	29,783	29,783	30,217	29,655	29,655	30,570	29,959	29,959
Mean dependent variable male	.735	.736	.736	.175	.174	.174	.310	.309	.309
Mean dependent variable female	.667	.667	.667	.166	.165	.165	.340	.340	.340
p-value from test for gender equality of proportion female peers	<.001	<.001	<.001	.168	.206	.206	.978	.926	.912
Grade and school fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Individual-level controls	–	✓	✓	–	✓	✓	–	✓	✓
Class-level controls	–	–	✓	–	–	✓	–	–	✓

Notes: – Standard errors clustered at the school level are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column in this table is one separate regression indicated by the checkmarks. For our outcome variables, each component was dichotomized to indicate agreement or disagreement with specific statements. Specifically, for School Liking, a score of 1 was assigned if an individual reported liking school even "a bit/a lot," and a score of 0 was assigned otherwise. For Students Together, Student Helpful, and Students Accept, a score of 1 was assigned if an individual "agreed" with the statement, and 0 was assigned otherwise. Been Bullied was coded 1 if an individual reported having been bullied "more than once" and 0 otherwise, while School Pressure was coded 1 if an individual reported experiencing "some/a lot" of school pressure and 0 otherwise. In our models, we incorporate individual-level controls, including the individual's family structure (coded as 1 if both parents reside in the primary household and 0 otherwise), the socio-economic status of the individual's family, and the age of the individual. The class-level controls we employ are defined as follows: indicators for whether any student in the class is two years older than the standard age for their grade level (i.e., 13.5 for grade 1, 15.5 for grade 2, and 17.5 for grade 3), indicators for significant changes in class size compared to the above grade (i.e., a change of 50% or greater), indicators for the proportion of students in the class who have a socio-economic status above the mean, and class size up to the third polynomial.