The impact of bullying on educational performance in Ghana: A Bias-reducing Matching Approach

Authors:

Shahriar Kibriya
Department of Agricultural Economics
Texas A&M University
2124 TAMU, College Station, TX 77843-2124.
Email: shahriar.kibriya@gmail.com

Zhicheng Phil Xu
Department of Agricultural Economics
Texas A&M University
2124 TAMU, College Station, TX 77843-2124.
philxu@tamu.edu

Yu Zhang
Department of Agricultural Economics
Texas A&M University
2124 TAMU, College Station, TX 77843-2124.
Email: zhangyu523@tamu.edu

Selected Paper prepared for presentation for the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26-28.

The impact of bullying on educational performance in Ghana:
A Bias-reducing Matching Approach

Abstract:

School bullying is a serious problem in academic settings all over the world. Using data from Ghana through the Trends in International Mathematics and Science Study we show that bullying has a negative effect on academic performance. We also find evidence that female students are more affected by bullying. However, our analysis of the data reveals a female teacher in the classroom diminishes the negative effect of bullying on female students. The analysis uses a quasi-experimental propensity score matching and OLS methods followed by a series of robustness tests that validates the unconfoundedness and overlapping assumptions. The results of the study encourage policy makers to introduce gender sensitive anti-bullying program in academic settings.

1. Introduction and Literature Review

"If there's one goal of this conference, it's to dispel the myth that bullying is just a harmless rite of passage or an inevitable part of growing up."

US President Barack Obama, White House Anti Bullying Conference, 2011

Bullying is an important but often neglected issue that can hinder performance in school. It is a serious and worldwide phenomenon.[1] For instance, by analyzing data from a representative sample of 15686 U.S. students in grades 6 through 10, Nansel et al. (2001) show that almost 30% students of the sample reported moderate or frequent involvement in bullying. In another recent international survey conducted in 2011, consisting of more than 300000 students from 48 developed and developing countries, over 50% of these students reported that they experienced bullying at school; furthermore, 33% of the sample said that they were bullied "approximately weekly" (Mullis et al., 2012). Evidence of bulllying has also been found in Ghana, the United

[1] Olweus (1978) began to systematically study bullying at school in Scandinavia in the 1970s and proposed the definition of bullying that is widely accepted by the following researchers. According to Olweus (1993), a student is being bullied at school "when he or she is exposed, repeatedly and over time, to negative actions on the part of one or more other students." These negative actions include to attack or discomfort someone physically or verbally, spreading rumors, and intentionally excluding someone from a group.

Kingdom, Denmark, Italy, as well as other European countries (Ammermueller, 2012; Brown and Taylor, 2008; Dunne et al., 2013; Eriksen et al., 2012; Ponzo, 2013).

A large body of literature shows that education is essential for the private and social returns to human capital (see Card 1999 for an overview). Hence, examining effective ways to improve the quality of education remains a germane investigation. The literature increasingly focuses on the influences of individual, household, school and teacher characteristics on students' performances, such as the student's gender, school's quality, enrollment, location, and teacher's gender, experience, education level, etc. (Card and Krueger, 1992; Dearden et al., 2002; Ehrenberg and Brewer, 1994; Hanushek, 1986; Kukla-Acevedo, 2009). However, until recently few studies have addressed the effect of bullying on academic performance (Ponzo 2013). Furthermore, the gender specific effects of bullying on academic performance in developing nations have not being explored until now. With this study we attempt to contribute to the existing literature by studying the consequence of bullying on standardized tests differentiated by students' and teachers' gender in Ghana.

Bullying can impact academic performance in various ways. Victims of bullying are more likely to report feeling unhappy and lonely at school, and having fewer good friends (Boulton and Underwood, 1992). A victim of bullying is more likely to develop new psychosomatic and psychosocial problems compared with children who were not bullied (Kumpulainen et al., 2001; Fekkes et al., 2006), thereby an adverse effect on coping with loneliness, anxiety and depression in study and daily life. The evidence for the relationship between bullying and psychological problems has also been found in neuro-biological literature. For instance, Ouellet-Morin (2011) reports that physical maltreatment has long-lasting effects on the hypothalamic-pituitary-adrenal (HPA) reactivity that is associated with social, emotional, and behavioral problems. Hemphill et al. (2011) find that being bullied is highly correlated with binge drinking and depression. School avoidance and poor attendance also lead to poor academic performance. From a diverse sample of 5730 LGBT youths who had attended secondary schools in the United States, Kosciw et al. (2013) find that victims of bullying suffered from lower self-esteem and thereby lower academic performance. Exploiting the 2007 National Crime Victimization Survey's School Crime Supplement, Barrrett et al. (2012) suggest that fear of crime in school makes a student more willing to skip class and less aspired to pursue higher grades. The relationship between school violence and poor engagement has also been documented in Ripski and Gregory (2009) and Dunne (2013), etc.

Due to the prevalence and severity of school bullying observed around the world, recently there has been a rising academic interest to precisely quantify the causal effect of school bullying on academic achievement and lifetime earnings beyond school.[2] With a few exceptions[3] most research found direct associations between bullying and educational achievement. Most of the studies claim that bullying behavior leads to poor academic performance. Le et al. (2005) study a sample of twins chronologically in Australia and showed that childhood conduct disorder, such as seen in children who bully, can adversely affect an individual's academic attainment and their competency in the labor market. Brown and Taylor (2008) explore the same question by drawing a sample from the British National Child Development Study data. Their findings are consistent with those from Le et al.'s research. Furthermore, they showed that the effects from bullying outweigh the effects of class size, which has been a key determinant of educational attainment in the economic sciences literature (for example, see Card and Krueger, 1992; Dearden et al.,

[2] Experimental studies have also been designed to examine policy interventions of bullying in academic institutions. For instance, Persson and Svensson (2013) conducted a discrete choice experiment in Sweden to elicit the willingness to pay (WTP) to reduce school bullying.

[3] Such as Woods and Wolke (2004)

2002). Such empirical finding suggests the need for more studies to understand the causal linkage between bullying and academic performance. Ammermueller (2012) uses a much broader dataset, including 11 European countries, and with it analyzed the determinants of bullying and its effects on students' attainment. Similarly, it is found that being bullied has a significantly negative impact on present and future students' performance in both school and labor market. Notwithstanding significant correlation between bullying and attainment has been investigated in the above studies, the causal direction remains unclear. In other words, it is possible that a student has a lower academic performance because of being a victim, or the likelihood of a student being bullied is higher if he performs poorer. Ponzo (2013) attempted to solve the reverse causality problem by employing a non-parametric method (propensity score matching) for schools in Italy. She concluded that school bullying decreased student performances in both fourth and eighth grades.

This initiation of the study is related to Ponzo's (2013) basic framework and hypothesis followed by gender specific assignments as well as sophisticated robustness tests. Our aim is to better understand the gender specific effect of bullying in academic institutions of an African country, Ghana. However, it is challenging to draw causal inferences about the relationship between bullying and academic performance. Besides concerns of over selection and endogeneity bias, students' performance may also get affected by a heterogeneous learning environment both inside and outside school. Bullying is may be caused by other socio-economic indicators and a feedback loop may also exist towards bullying from academic achievement. So, employing only a linear regression analysis may under specify or over exemplify the effect of bullying. A randomized control trial would be ideal from an experimental design point of view, however very difficult to execute. For example: it would be ethically wrong to put a child in danger knowing that he/she is a victim of bullying. Hence, we decided to employ a quasi-experimental setting by employing both OLS estimation and Propensity Score Matching (PSM). PSM is used to reduce selection bias by equating groups based on these covariates. To estimate the average effect of bullying we compare the academic performance in standardized score (outcome variable) between the “bullied” (treatment) and “not bullied” (non-treatment). Following, we use matching technique to estimate the gender specific heterogeneous effects of bullying. Specifically, we estimate if being a female teacher or student in a classroom mitigates or enhances the average effect of bullying. To ensure the robustness of the estimated average effect we employ several robustness checks suggested by Imbens and Woolridge (2009). The tests are: graphical representation, matching quality estimator, and placebo regression. Employing this technique and associated robustness tests validates our findings and avoids several pitfalls. First, we do not require baseline information on students (Imbens and Woolridge 2009). Second, they ensure the comparison of the outcome variable is undertaken between students who have similar and overlapping characters (Dehejja and Wahba 2002). Third, the placebo regression validates the unconfoundedness (outcomes are independent to treatment) assumption of PSM. We use a rich dataset from the Trends in International Mathematics and Science Study (TIMSS) in 2011 that will be described in details in the next sections. The conclusions drawn from such analysis reveal the most vulnerable gender group and assist policy makers to design intervention programs that would alleviate the effects of bullying in enhancing academic performance.

The remaining parts of this paper proceed as follow. The next section describes the data used in our analysis. Section 3 presents the parametric and non-parametric research methodology. Section 4 shows the results and interpretations, with special attention paid to the heterogeneity analysis. In section 5, we check robustness via matching quality test and bias-corrected matching estimator with heteroskedasticity errors. Section 6 concludes.

2. Data Description and Summary Statistics

The International Association for the Evaluation of Educational Achievement (IEA) has conducted the Trends in International Mathematics and Science Study (TIMSS) in the past two decades. The TIMSS dataset is enriched by comprehensive background information related to students and their households, teachers, and schools. Students' achievements on math are reported with a scale of 0 to 1000; however their typical scores fall in the range from 300 to 700. In the following analysis, we use math score as the main measurement of students' performance. TIMSS sets four threshold scores as international benchmarks: Advanced International Benchmark (625), High International Benchmark (550), Intermediate International Benchmark (475), and Low International Benchmark (400).

Ghana is selected for our empirical analysis. Data is collected from the latest TIMSS survey, conducted in 2011. In Ghana, 7323 eighth grade students participated in 2011-TIMSS. All students and their associated schools were randomly chosen. The survey contains a set of questions regarding whether students suffer from school bullying in the background questionnaire. These questions were:

"During this year, how often were you made fun of or called names at school?"

"During this year, how often were you left out of games or activities by other students at school?"

"During this year, how often did someone spread lies about you at school?"

"During this year, how often was something stolen from you at school?"

"During this year, how often were you hit or hurt by other student(s) at school?"

"During this year, how often were you made to do things you didn't want to do by other students at school?"

Each respondent was asked to select one out of the following options: "once a week," "once or twice a month," "few times a year," or "never". Based on the answers collected from the respondents, the TIMSS dataset generates a derived variable indicating school bullying, that is, a student is graded as being "*bullied weekly*" if he/she at least experienced each of three of the six bullying behaviors "once or twice a month" and each of the other three "a few times a year".

Besides school bullying, we hypothesize that four primary factors affect students' achievements. The first of these factors is student's individual characteristics, such as student age and gender. The second factor is student's household characteristics, including parents' education level, as well as 5 indicators on home support for education. These indicators include computer possession, study desk, having their own room, internet accessibility, and number of books at home. The third factor is student's teacher characteristics, consisting of teacher's experience, gender, and education level. The fourth factor affecting students' academic performances is school characteristics, including school location, percentage of students coming from economically disadvantaged families, school enrollment, and the number of computers as a proxy of school facility.

Table 1 presents descriptive statistics for the variables discussed above for Ghana. It shows that the ratio of females to males is close to 1:1. The average age of eighth grade students is approximately 16 years. The statistics show that students from Ghana are almost 70 points lower from the Low International Benchmark (400). Furthermore, the female students score 24 points behind the male students. The primary choice variable, *bullied*, is also shown in Table 1. Bullying is found to be pervasive in Ghana: more than one half of students surveyed are bullied weekly. The likelihoods of being bullied are equal between the male and female students. Other variables of interest discussed above are also depicted in Table 1.

Table 2 demonstrates in detail about school bullying for Ghana. Instead of reporting one single derived variable of school violence, it reports the frequency of being affected by six different bullying behaviors. We find that the least prevalent activities of bullying are "being hurt by other students" and "forced to do things" (more than 50 percent students answered “*never*”). On the contrary, the most prevalent activities of bullying are "things stolen" and "being made fun of," generally more than 70 percent of the students have experienced such behaviors at least few times a year.

Table 1: Descriptive statistics for the main variables used

Variables		TIMSS 2011 Eighth grade	
		Mean	Std. dev.
Math Score		333.007	78.497
	Male	344.474	79.096
	Female	320.489	75.895
Bullied weekly		0.530	0.499
	Male	0.533	0.499
	Female	0.528	0.499
Student age		15.744	1.512
Female student		0.478	0.500
Parents' education level			
	university or above	0.106	0.307
	post-secondary	0.160	0.366
	upper secondary	0.221	0.415
	lower secondary	0.309	0.462
	primary or no school	0.204	0.403
Computer possession		0.250	0.433
Study desk		0.506	0.500
Own room		0.318	0.466
Internet at home		0.112	0.316
Books at home			
	0-10	0.401	0.490
	11-25	0.368	0.482
	26-100	0.139	0.346
	101-200	0.043	0.204
	>200	0.048	0.214
School location			

urban	0.178	0.382
suburban	0.166	0.372
Large town	0.167	0.373
small town or village	0.392	0.488
remote rural	0.098	0.297
Portion of students from disadvantage families		
0-10%	0.066	0.248
11-25%	0.112	0.315
26-50%	0.161	0.367
More than 50%	0.662	0.473
School enrollment	265.153	213.922
School computer		
1 computer for 1-2 students	0.443	0.497
1 computer for 3-5 students	0.118	0.323
1 computer for 6 or more students	0.290	0.454
no computers available	0.149	0.356
Years teacher has been teaching	8.266	6.557
Female teacher	0.121	0.326
Teacher education level		
upper secondary education	0.079	0.270
post-secondary non-tertiary level of education	0.450	0.498
short tertiary education	0.193	0.394
long tertiary education	0.274	0.446
university or higher	0.004	0.066
Observations	7323	

Source: TIMSS 2011.

Table 2: Descriptive statistics of the indicators for school bullying

Variables	Once a week		Once or twice a month		Few times a year		Never	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Being made fun of	0.429	0.495	0.154	0.361	0.112	0.315	0.304	0.460
Being left out of games	0.228	0.419	0.196	0.397	0.123	0.328	0.454	0.498
lies spread about me	0.170	0.376	0.164	0.371	0.181	0.385	0.485	0.500
Things stolen	0.271	0.445	0.223	0.416	0.206	0.404	0.300	0.458
Being hurt	0.172	0.377	0.156	0.363	0.140	0.347	0.532	0.499
Being forced to do things	0.190	0.393	0.140	0.347	0.128	0.334	0.542	0.498
Observations	7323							

Source: TIMSS 2011.

3. Research Methodology

We rely on parametric and non-parametric identification strategies.

We estimate the following model for student achievement using ordinary least squares (OLS):

$$Y_i = \beta_0 + \beta_1 bullied_i + \beta_2 X_i + \varepsilon_i \qquad (1)$$

where Y_i denotes the math score of student i , $bullied_i$ is a binary variable indicating whether or not the student has been a victim of school bullying, X_i is a vector of controls (including student, family, teacher, and school characteristics), ε_i is an error term capturing idiosyncratic shocks or unobserved characteristics. β_1 represents the measure of our major interest, i.e., the expected mean gap in academic performance between bullied students and non-bullied students.

However, OLS estimation may be biased due to endogeneity issues. Other models need to be used to overcome endogeneity problem and check the robustness of the estimations. In addition, since the data on bullying is observational survey data, OLS estimations may provide over-estimates of bullying impacts on academic performances.

Matching method is adopted to overcome the problems, for instance, grouping units based on a single variable (Dehejia and Wahba, 2002). Intuitively, matching estimator matches pairs of individuals with the characteristics from control and treatment groups. Accordingly, the pair of matched individuals is essentially similar in all aspects but being assigned into control or treatment group. Hence matching method makes the comparison between treatment and control group immune to selection bias. Nevertheless, problems arise when the number of covariates is high, deemed the curse of dimensionality in the literature. The method of propensity score matching (PSM) proposed by Rosenbaum and Rubin (1983) is suitable to solve this problem. PSM refers to the conditional probability (given a vector of covariates X) of being assigned to treatment. That is, propensity score takes into account the multi-dimensional covariates and compresses them into a single dimension, facilitating the matching process (Abadie and Imbens, 2009). Thus the key advantages of PSM are that by using a linear combination of covariates for a single score, it balances treatment and control groups on a large number of covariates without losing a large number of observations. Again, the pair-matched individuals in control and treatment group with the same propensity score are essentially comparable, since their only difference is whether being assigned in treatment or control group.

Formally, a propensity score is the probability of a unit (i.e., student in our research) being assigned to a particular treatment (i.e., being bullied) given a set of observed covariates. Propensity scores are used to reduce selection bias by equating groups based on these covariates. Suppose that we have a binary treatment T (T=1 if bullied, or 0 otherwise), an outcome Y (academic performance), and background variables X. The propensity score is defined as the conditional probability of treatment given background variables:

$$P(x) = \Pr(T = 1|X = x) \qquad (2)$$

Let Y(0) and Y(1) denote the potential outcomes under control and treatment, respectively. That is, Y(0) and Y(1) are expected academic performance of a student not being bullied and a student being bullied.

Then the treatment assignment is (conditionally) unconfounded if potential outcomes are independent of treatment conditional on background variables X. This can be written compactly as

$$Y(0), Y(1) \perp T|X \qquad (3)$$

where ⊥ denotes statistical independence. If unconfoundness holds, then

$$Y(0), Y(1) \perp T|P(X) \qquad (4)$$

In technical terms, we will obtain the Average Treatment Effect (ATE) as the mean difference in outcome between the treated and the control students, and the average treatment effect on the treated (ATT) that is the average effect from treatment for those who actually were treated.

In order to define formally the ATE, we define two potential outcomes: Y_{0i} is the value of the outcome variable for individual i if s/he is not treated, Y_{1i} is the value of the outcome variable for individual i if s/he is treated. The ATE is given by:

$$E(\mathrm{Y_{1i}} - \mathrm{Y_{0i}}) \qquad (5)$$

And ATT is given by

$$E[(\mathrm{Y_{1i}} - \mathrm{Y_{0i}})\ |\mathrm{T} = 1] \qquad (6)$$

This process will enable us to examine school bullying on students as a “treatment” and investigate the effect of violence on the treated group. Essentially, PSM method will allow us to compare two groups of students with similar characteristics with one group being a victim of school bullying. Intuitively, the effect of bullying can be identified as the treatment effect shown by the deviation in academic performance.

Several matching algorithms will be implemented, i.e., nearest neighbor, radius and kernel (Caliendo and Kopeining, 2008; Imbens, 2014). Nearest neighbor matching uses an algorithm that matches each bullied student with the non-bullied student with the closest propensity score. Nearest neighbor is applied with replacement, since a non-bullied student can be a best match for more than one bullied student. Since each bullied student is matched with only one non-bullied student, the number of matched students might be fewer for statistical efficiency. By using radius matching, we match each bullied student with all non-bullied students whose propensity score falls into predefined neighbor of the propensity score of the bullied student. We set the radius of the neighborhood is 0.005. Finally, we also apply Kernel matching, with which each bullied student is matched with a weighted average of all non-bullied students with weights declining with the distance between propensity scores of bullied and non-bullied students. In our analysis, we use Epanechnikov Kernel function where the bandwidth is 0.06, following Heckman et al. (1997).

For example, the treatment effect on the treated by NN estimator is

$$ATT^{NN} = \frac{1}{N^T} \sum_{w_i} \left[y_i^{obs} - \sum_{j \in C(i)_m} w_{ij} y_j^{obs} \right] \qquad (7)$$

where N^T is the number of observations in the treated group, N_i^C is the number of controls matched with treated observation i, w_{ij} is equal to $\frac{1}{N_i^C}$ if j is a control units of i, and zero otherwise, $w_j = \sum_i w_{ij}$. The other two matching algorithms are similar in principle but use different weighted average.

Besides the main effects of school bullying on students' academic performances discussed in the above section, we are also interested in the heterogeneous (differential) effects of bullying.

Specifically, we wonder if school bullying has a significant differential impact on performance of female and male students, of students from rural and urban areas, of students who attend school with different facilities, of students who attend school with more or less poor students, of students whose parents differ by education levels, of students who have access to female teachers, and of students who attend school with different enrollment size. The heterogeneous treatment effects are of great interests because it is possible that the relationship between academic achievements and the variable "*Bullied*" depends on the value of one or more other control variables. For example, it might be the case that being a victim of bullying at school, students who are female loses more points than those who are male.

To investigate the heterogeneous treatment effects of bullying on academic performance, we apply the PSM approach instead of OLS. The effects of bullying on students' performance are highly possible to be affected by more than one covariate. In such case, using OLS becomes cumbersome and could not guarantee the unbiased results. On the contrary, by separating the sample into two parts given any interactive variable (for example, student gender), we can apply PSM on each sub-sample, and compare the effects respectively. All other possible interactive variables related to bullying and student gender are ruled out in propensity score matching.

4. Empirical results

This section provides the main empirical results generated by OLS and PSM for Ghana. "*Bullied weekly,*" which is a binary variable, is used for our analysis. In all OLS specifications, standard errors are adjusted for school-level clustering and heteroskedasticity. In all matching specifications, we use a bootstrapping procedure to construct the standard errors for the average treatment effects of the treated.

Table 3: Impacts of being bullied weekly on math performance (OLS approach)

	(1)	(2)	(3)	(4)	(5)
Bullied	-18.144***	-20.440***	-19.630***	-18.634***	-15.447***
	(3.275)	(2.663)	(2.885)	(2.777)	(1.599)
Student age		-13.969***	-13.081***	-10.454***	-6.256***
		(1.455)	(1.378)	(1.115)	(0.647)
Student female		-28.785***	-27.910***	-28.048***	-27.223***
		(2.820)	(2.993)	(2.902)	(1.933)
Parents' highest education level					
University		39.646***	36.792***	22.432***	5.187
		(7.025)	(7.112)	(5.501)	(3.305)
Post-secondary		20.218***	15.516***	7.291	0.120
		(4.537)	(5.163)	(4.766)	(2.917)
Upper secondary		21.510***	17.320***	8.982*	5.819**
		(4.667)	(5.591)	(5.201)	(2.542)
Lower secondary		14.214***	11.557**	5.674	3.852
		(4.331)	(4.982)	(4.795)	(2.497)
Teacher female			-10.509	-13.724	
			(9.112)	(9.873)	
Teachers' experience			0.938	-0.657	
			(0.720)	(0.658)	
School location					
urban				18.929	
				(13.850)	
suburban				48.276***	
				(14.511)	
Large town				25.187*	
				(14.780)	
small town or village				11.541	
				(12.808)	
School enrollment				0.038	
				(0.021)	
Demographic controls	No	Yes	Yes	Yes	Yes
Teacher controls	No	Yes	Yes	Yes	No
School controls	No	Yes	No	Yes	No
School fixed effects	No	No	No	No	Yes
Observations	7323	5503	5002	4514	5503
R-squared	0.013	0.168	0.188	0.276	0.544

Note: "Demographic controls" include: 5 dummies for number of books at home, computer possession, study desk, own room, and internet accessibility. "Other teacher control" includes teacher's education level. "Other school controls" include: 4 dummies for portion of students coming from disadvantaged families, and 4 dummies for instructional computer accessibility. Standard errors are adjusted for school-level clustering and heteroskedasticity. *, **, *** indicates 10, 5, and 1 percent level statistical significance, respectively.

We first analyze the impacts of school bullying on students' academic achievements by OLS approach. As Table 3 shows, a variety of specifications are applied in the OLS analysis. Column (1) shows the simplest specification, nothing else is included in the model but "*Bullied*". In column (2) we add several variables to control for individual and household characteristics: student age, student gender, parents' highest education level, number of books at home, computer possession, study desk, own room, and internet accessibility. In column (3) we include additional variables to control for teacher characteristics: teacher' experience, gender, and education level. In column (4) we control for school characteristics as well: school enrollment, portion of students coming from disadvantaged families, school location, and number of school computers. In column (5) we use an alternative way to control for teacher and school characteristics: school fixed effects are employed instead of a set of control variables. Compared to the specification in column (4), specification in column (5) presents a more parsimonious model and includes the effects of potential unobserved school characteristics.

From column (1) to column (5), the R-square is increasing as more and more variables are added into regression. Although the coefficients of "*Bullied*" fluctuated from -15.4 to -20.4, they remain statistically significant at the 1 percent level across all five specifications. The magnitudes of the coefficients create a decreasing trend while the models become more comprehensive. This is reasonable because some control variables may be correlated with school bullying, leading to the overestimate of the impact on bullying in the less comprehensive specifications. Column (5) represents the most comprehensive specification, implying that the school fixed effects capture additional unobserved school characteristics that are correlated with the school bullying. Generally speaking, being bullied weekly at school is significantly negative related with eighth grade students' math performance. As shown in column (5), students being bullied scored approximately 15.5 points less than students not being bullied, which corresponds to a reduction of 20% standard deviation of sample mean score.

We also found that male and younger students in Ghana perform better. The other factors that significantly improve students' performance are parents' education level and school being located in suburban or large town. The rest of the controls seem not to be significantly correlated with students' performance.

Table 4 reports the results from the propensity score matching. The first step of this approach is to predict the propensity score, i.e., the probability of student being bullied conditional on pre-treatment control variables. For the sake of brevity, we do not report the logit estimation of propensity score. Three methods are employed to estimate average treatment effect on the treated (ATT): nearest neighbor matching, radius matching and Kernel matching. Nearest neighbor matching result suggests that students being bullied at school achieve 17.1 points lower than their non-bullied fellows in math scores. The similar results generated by radius matching and Kernel matching support the above finding. In fact, they provide even larger ATT estimates, about 18.5 and 18.3 points in reduction respectively. Table 3 and 4 reflect that the results are very robust across both parametric and nonparametric estimation approaches.

We now extend our analysis to the heterogeneous effects of school bullying. That is, the penalties for being a victim might be drastically different in terms of magnitude across gender groups.[4]

[4] We have also checked the heterogeneous effects across other categories: school located in urban or rural areas, parents' receiving post-secondary education or not, school comprised of low or high portion of poor students, and teachers with less or more experience. But we do not find significant differential effects on them.

Table 4: Impacts of being bullied weekly on math performance (matching approach)

Matching methods	8^{th} grade math scores
Nearest neighbor	-17.137***
	(3.643)
Number of treated	2357
Number of controls	2081
Radius/Caliper	-18.547***
	(2.025)
Number of treated	2341
Number of controls	2024
Epanechnikov Kernel	-18.300***
	(2.074)
Number of treated	2357
Number of controls	2081

Notes. Balancing property and common support are satisfied. Nearest neighbor is applied with replacement. Standard errors, estimated by 100 bootstrap replications, are reported in parentheses. *, **, *** indicates 10, 5, and 1 percent level statistical significance, respectively.

Panel A of Table 5 lists the effects of being bullied on female and male students separately. In Ghana, female students are more sensitive to school bullying compared with the male students. Specifically, if a girl is bullied, her math scores will be approximately 22 points less than an non-bullied girl. On the other hand, a victim boy suffered a drop of approximately 13 points in contrast to a non-bullied boy. While panel A illustrates the most vulnerable group of students, panel B of Table 5 displays a possible channel to alleviate the effects of bullying. Decomposing the sample by the gender of teachers, it is found that the presence of a female teacher may help. The different roles that female and male teachers play may be attributed to their distinct classroom management practices. Researchers have found that women teachers feel more comfortable to act nurturing, overtly affectionate, and empathic than their male counterparts; female teachers are also more responsive to school bullying, that is, they are more likely to help victims than males.[5] (Casey and Fuller, 1994; Martin and Ross, 2005; Hirdes, 2010)

[5] For example, female teachers determine situations to be more severe than the male teachers. They would talk to the students being bullied, find peer support, and involve other children more.

Table 5: Impacts of being bullied weekly by gender categories

	Female	Male
A. Students decomposed by gender		
	-22.314***	-12.973***
	(4.932)	(4.804)
Number of treated	1134	1215
Number of controls	974	1063
B. Teachers decomposed by gender		
	-2.491	-20.011***
	(8.369)	(3.810)
Number of treated	276	2077
Number of controls	221	1856

Notes. Balancing property and common support are satisfied. Nearest neighbor is applied with replacement. Standard errors, estimated by 100 bootstrap replications, are reported in parentheses. *, **, *** indicates 10, 5, and 1 percent level statistical significance, respectively.

5. Robustness check

This section evaluates the robustness of our estimates from three aspects. First, it assesses the overlap assumption by a graphical representation. Second, it measures the quality of matching, that is, whether the distribution of the covariates in both control and treatment groups are balanced. Third, a placebo regression approach is used to test the plausibility of the unconfoundedness assumption.

Overlap or common support is one the major assumptions in PSM, which ensures that students with the same propensity score have a positive probability of being the treated and untreated. Therefore, a straightforward method to test the overlap assumption is to plot the distribution of the propensity scores of the bullied and non-bullied students, and visually inspect whether the two distributions are overlapped. Fig. 1 compares the distributions of propensity score of the two groups. It is observed that the two distributions are considerably overlapped, suggesting that the overlap assumption holds.

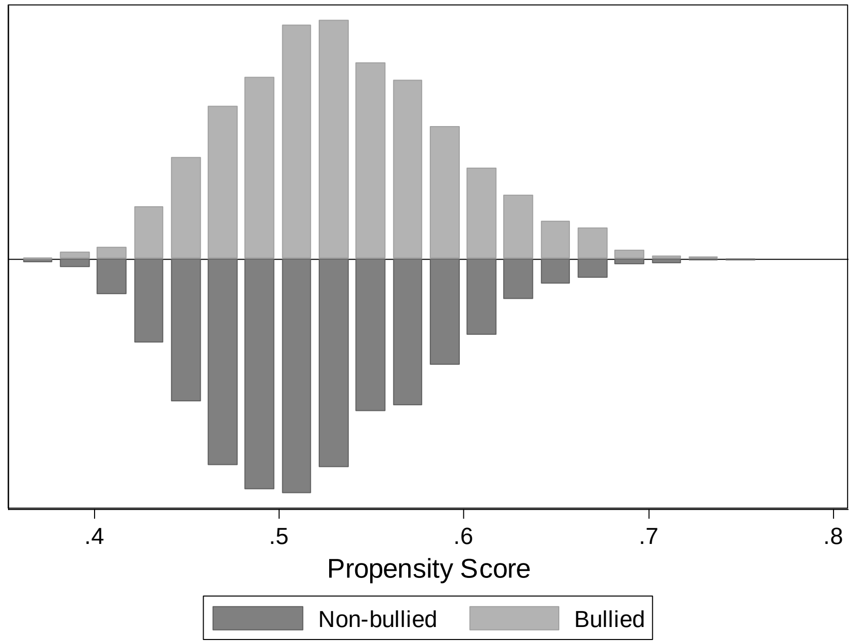

Fig. 1. Propensity score distributions by bully

Another concern rises over the quality of matching, since we do not condition on all covariates but on the propensity score. Before matching, the covariates mean differences between the treatment and control groups are expected to be relatively large. If a matching is successfully balanced, then there should be no statistically significant differences between covariate means of the treatment and control groups after matching. Following Diamond and Sekhon (2013), individual covariate balance is measured as the mean standardized difference in the empirical-QQ plot for each variable. An empirical-QQ plot compares the cumulative probability distribution of the treatment group with the control group. Larger mean standardized differences indicate better evidence for the conclusion that the individual covariate is not well balanced. Table 6 displays the results of match balancing test, showing the standard mean differences among covariates pre- and post-matching. Compared with the pre-matching case, the mean differences after matching become remarkably small and very close to zero. Table 6 suggests that the matching is good, since most of the variables have a reduction of at least 50%

in mean difference. *Student age* is the only exception that matching does not make any improvement. The reason is that all students in our study are in the eighth grade and their age range is relatively small.

Table 6: Matching quality: mean differences in covariates pre- and post-matching

Variable	**Mean Difference**		**Difference Reduction (%)**
	Before Match	After Match	
Home computer	0.018	0.006	66.67%
Study desk	0.028	0.001	96.43%
Own room	0.028	0.008	71.43%
Internet connection	0.017	0.003	82.35%
Home books	0.030	0.003	90.00%
Student gender	0.008	0.005	37.50%
Student age	0.008	0.008	0.00%
Parent education	0.017	0.006	64.71%
School location	0.015	0.008	46.67%
Portion of disadvantaged students	0.007	0.005	28.57%
School enrollment	0.010	0.005	50.00%
School computer availability	0.021	0.003	85.71%
Teacher's experience	0.016	0.003	81.25%
Teacher's education	0.014	0.002	85.71%
Teacher's gender	0.006	0.002	66.67%

The other major assumption of PSM is unconfoundedness, requiring that all covariates that relates with treatment and potential outcome are observed. A placebo regression is designed to assess the unconfoundedness assumption. Maintaining all right-hand variables used in the estimation of the propensity score, we insert a new dependent variable that is assumed to be exogenous with the treatment. If there are omitted variables correlated with the treatment, then the coefficient associated with *bullied* should be significantly different from zero. Otherwise, the unconfoundness assumption holds. We employ the birth date of each student as the pre-determined dependent variable, which can be treated as a variable randomly assigned to the students. Table 7 shows the results of the placebo regression. The null hypothesis that the coefficients of *bullied* cannot be rejected, indicating that omitted variables affecting the treatment do not exist.

Table 7: OLS results from the placebo regression test

	Coefficient	P value
Bullied	0.382	
Student age	-0.163	*
Student female	0.398	
Parents' education	-0.335	***
Teacher female	0.066	
Teachers' experience	-0.037	
Teacher's education	0.283	
School enrollment	0.001	
Portion of disadvantaged students	-0.283	**
School location	-0.271	*
Home books	-0.208	
Home computer	0.170	
Study desk	-0.197	
Own room	-0.009	
Internet connection	-0.399	
School computer availability	0.028	
Constant	18.686	***
Observations	3883	
R-squared	0.1	

Note: Standard errors are adjusted for school-level clustering and heteroskedasticity. *, **, *** indicates 10, 5, and 1 percent level statistical significance, respectively.

6. Concluding Remarks and policy implications:

The impact of school violence on students' health and psychological development has been well documented (Barrett et al., 2012; Dunne, 2013; Hazel, 2010; Hemphill et al., 2011; Kosciw et al., 2013; Ouellet-Morin et al., 2011; Ripski and Gregory, 2009). However, research on the consequence of school bullying on academic achievement in an African context was not addressed until recent times (Caputo, 2013; Perše et. al, 2011; Ponzo 2013; USAID, 2013). Quantitative evidence of such phenomena for developing countries has been absent. Our study has made an attempt to fill this gap in the literature. Using data on Ghanaian students from the TIMMS (2011) database we show that students' academic performance suffer due to school bullying. We find that female students suffer more than male students. However, our analysis reveals that female educators diminish the effect of bullying on female students. Several econometric and statistical techniques were used to account for potential biases. The overlapping assumption was validated by the quality of matching tests while unconfoundedness assumption was verified through placebo regressions. Furthermore, we estimated the pre and post match mean differences of covariates. These tests show no concern for the assumptions used, suggesting the causal interpretation and direction of bullying on academic performance convincing.

The finding highlights the importance of considering specific school development programs that address bullying. Our analysis establishes that bullying has a direct effect on academic performance and is not caused by other socio-economic determinants[6]. Our results also suggest that anti-bullying programs should have gender sensitive components. Female students were affected more by bullying; whereas a female teacher in the classroom can diminish this effect. Policy makers should consider promoting more female teachers in African schools and provide gender sensitivity trainings. We encourage further research of different educational environments, particularly those that utilize different approaches to students based on sex. For example, a systematic review of classrooms separated by sex is required to identify if the results found for students' sex are based on the educational system or other factors. Other factor may include different levels teacher of encouragement based on a student's sex, characteristics of

[6] We conducted a series of models to examine the interaction between different variables and find which effects are more significant in each case to consider for approaches.

teacher training, and curriculum. General questions about how teachers' sex interacts with student performance are yet to be answered. Future studies focused on multi-year studies designed to examine bullying that specifically focuses on teacher and school characters will shed more light on the this topic. Tracking cohorts of students will enhance our understanding of environmental factors and provide insight into changes over time for additional causal analysis.

References

Abadie, A., & Imbens, G. W. (2009). *Matching on the estimated propensity score* (No. w15301). National Bureau of Economic Research.

Ammermueller, A. (2012). Violence in European schools: A widespread phenomenon that matters for educational production. *Labour Economics*, *19*(6), 908-922.

Barrett, K. L., Jennings, W. G., & Lynch, M. J. (2012). The relation between youth fear and avoidance of crime in school and academic experiences. *Journal of School Violence*, *11*(1), 1-20.

Boulton, M. J., & Underwood, K. (1992). Bully/victim problems among middle school children. *British Journal of Educational Psychology*, *62*(1), 73-87.

Brown, S., & Taylor, K. (2008). Bullying, education and earnings: evidence from the National Child Development Study. *Economics of Education Review*, *27*(4), 387-401.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*,*22*(1), 31-72.

Caputo, A. (2013). Students' Perception of School Violence and Math Achievement in Middle Schools of Southern Italy. Educational Assessment, 18(4), 269-284.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, *3*, 1801-1863.

Card, D., & Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *The Journal of Political Economy*, *100*(1), 1-40.

Casey, R. J., & Fuller, L. L. (1994). Maternal regulation of children's emotions. *Journal of Nonverbal Behavior*, *18*(1), 57-89.

Chudgar, A., & Sankar, V. (2008). The relationship between teacher gender and student achievement: evidence from five Indian states. *Compare*, *38*(5), 627-642.

Dearden, L., Ferri, J., & Meghir, C. (2002). The effect of school quality on educational attainment and wages. *Review of Economics and Statistics*, *84*(1), 1-20.

Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association*, *94*(448), 1053-1062.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

Diamond, A., & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*, *95*(3), 932-945.

Dunne, M., Sabates, R., Bosumtwi-Sam, C., & Owusu, A. (2013). Peer relations, violence and school attendance: analyses of bullying in senior high schools in Ghana. *The Journal of Development Studies*, *49*(2), 285-300.

Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from< i> High School and Beyond</i>. *Economics of Education Review*, *13*(1), 1-17.

Eriksen, T. L. M., Nielsen, H. S., & Simonsen, M. (2012). *The Effects of Bullying in Elementary School* (No. 6718). The Institute for the Study of Labor.

Fekkes, M., Pijpers, F. I., Fredriks, A. M., Vogels, T., & Verloove-Vanhorick, S. P. (2006). Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms. *Pediatrics*, *117*(5), 1568-1574

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 1141-1177.

Hazel, C. (2010). Interactions between bullying and high-stakes testing at the elementary school level. *Journal of School Violence, 9*(4), 339-356.

Heckman, J., Ichimura, H. and Todd, P. E. (1997). Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4), 605–54.

Hemphill, S. A., Kotevski, A., Herrenkohl, T. I., Bond, L., Kim, M. J., Toumbourou, J. W., & Catalano, R. F. (2011). Longitudinal consequences of adolescent bullying perpetration and victimisation: A study of students in Victoria, Australia. *Criminal Behaviour and Mental Health*, *21*(2), 107-116.

Hirdes, C. L. (2010). Gender and Race of Teacher and Student: Are They Related to Teacher Responses to Incidents of School Bullying?.

Imbens, G. W. (2014). *Matching Methods in Practice: Three Examples* (No. w19959). National Bureau of Economic Research.

Kosciw, J. G., Palmer, N. A., Kull, R. M., & Greytak, E. A. (2013). The effect of negative school climate on academic outcomes for LGBT youth and the role of in-school supports. *Journal of School Violence*, *12*(1), 45-63.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review*, *28*(1), 49-57.

Kumpulainen, K., Räsänen, E., & Puura, K. (2001). Psychiatric disorders and the use of mental health services among children involved in bullying.*Aggressive Behavior*, *27*(2), 102-110.

Le, A. T., Miller, P. W., Heath, A. C., & Martin, N. (2005). Early childhood behaviours, schooling and labour market outcomes: Estimates from a sample of twins. *Economics of Education Review*, *24*(1), 1-17.

Martin, J., & Ross, H. (2005). Sibling aggression: Sex differences and parents' reactions. *International Journal of Behavioral Development*, *29*(2), 129-138.

Mullis, I. V., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 International Results in Reading*. International Association for the Evaluation of Educational Achievement. Amsterdam, The Netherlands.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, *285*(16), 2094-2100.

Olweus, D. (1993), Bullying at School: What We Know and What We Can Do. Blackwell Publishing.

Olweus, D. (1978). Aggression in the schools: Bullies and whipping boys. Washington, DC. Hemisphere Press.

Persson, M. & Svensson, M. (2013). The willingness to pay to reduce school bullying. *Economics of Education Review*, 35, 1-11.

Perše, T. V., Kozina, A., & Leban, T. R. (2011). Negative school factors and their influence on math and science achievement in TIMSS 2003. *Educational Studies*, 37(3), 265-276.

Ponzo, M. (2013). Does bullying reduce educational achievement? An evaluation using matching estimators. *Journal of Policy Modeling*, *35*(6), 1057-1078.

Ouellet-Morin, I., Odgers, C. L., Danese, A., Bowes, L., Shakoor, S., Papadopoulos, A. S., ... & Arseneault, L. (2011). Blunted cortisol responses to stress signal social and behavioral problems among maltreated/bullied 12-year-old children. *Biological psychiatry*, *70*(11), 1016-1023.

Ripski, M. B., & Gregory, A. (2009). Unfair, Unsafe, and Unwelcome: Do High School Students' Perceptions of Unfairness, Hostility, and Victimization in School Predict Engagement and Achievement?. *Journal of School Violence*, *8*(4), 355-375.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

United States Agency for International Development (2013, October). Literature Review on the Intersection of Safe Learning Environments and Educational Achievement.

Woods, S., & Wolke, D. (2004). Direct and relational bullying among primary school children and academic achievement. *Journal of School Psychology*, *42*(2), 135-155.