

NBER WORKING PAPER SERIES

BULLYING AMONG ADOLESCENTS:  
THE ROLE OF COGNITIVE AND NON-COGNITIVE SKILLS

Miguel Sarzosa  
Sergio Urzúa

Working Paper 21631  
<http://www.nber.org/papers/w21631>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
October 2015

We would like to thank Sebastian Galiani, John Ham, John Shea, and Tiago Pires for valuable comments on earlier versions of this paper. We are also indebted to Maria Fernanda Prada and Ricardo Espinoza for their comments on the computer codes used in this paper. In addition, we would also like to thank the seminar participants at the University of Maryland, University of North Carolina at Chapel Hill, LACEA meeting at Lima and George Washington University. This paper was prepared, in part, with support from a Grand Challenges Canada (GCC) Grant 0072-03, though only the authors, and not GCC or their employers, are responsible for the contents of this paper. Additionally, this research reported was supported by the National Institutes of Health under award number NICHD R01HD065436. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health nor of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Miguel Sarzosa and Sergio Urzúa. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Bullying among Adolescents: The Role of Cognitive and Non-Cognitive Skills  
Miguel Sarzosa and Sergio Urzúa  
NBER Working Paper No. 21631  
October 2015  
JEL No. C34,C38,I21,J24

**ABSTRACT**

Bullying is a behavioral phenomenon that has received increasing attention in recent times. This paper uses a structural model with latent skills and longitudinal information from Korean youths to identify the determinants and effects of bullying. We find that, unlike cognitive skills, non-cognitive skills significantly reduce the chances of being bullied during high school. We use the model to estimate average treatment effects of being bullied at age 15 on several outcomes measured at age 18. We show that bullying is very costly. It increases the probability of smoking as well as the likelihood of feeling sick, depressed, stressed and unsatisfied with life. It also reduces college enrollment and increases the dislike of school. We document that differences in non-cognitive and cognitive skill endowments palliate or exacerbate these consequences. Finally, we explore whether investing in non-cognitive skills could reduce the occurrence of bullying. Our findings indicate that the investment in skill development is key in any policy intended to fight the behavior.

Miguel Sarzosa  
Purdue University  
425 W. State Street  
Kenner Center  
Room 226  
West Lafayette , IN 47907  
msarzosa@purdue.edu

Sergio Urzúa  
Department of Economics  
University of Maryland  
3105 Tydings Hall #3115C  
College Park, MD 20742  
and NBER  
urzua@econ.umd.edu

# 1 Introduction

Bullying has been placed under the spotlight in many parts of the world in recent years. This behavioral phenomenon is by no means new, but it has increasingly become a predicament among young people. Frequent cases of suicide in school and college aged kids around the world keep reminding societies of the perils and immense costs that bullying victims—and communities in general—have to bear.

Psychologists have defined a bullying victim as a person that is repeatedly and intentionally exposed to injury or discomfort by others, with the harassment potentially triggered by violent contact, insulting, communicating private or inaccurate information and other unpleasant gestures like the exclusion from a group (Olweus, 1997). This explains why this behavioral phenomenon typically emerges in environments characterized by the imbalances of power and the needs for showing peer group status (Faris and Felmlee, 2011). Not surprisingly, schools are the perfect setting for bullying. The combination of peer pressure with the multidimensional heterogeneity of students, together with a sense of self-control still not fully developed, makes schools a petri dish for this improper behavior.

Bullying is not only widespread, but very costly.<sup>1</sup> According to official statistics, approximately 160,000 children miss school every day in the United States because of fear of being bullied (15% of those who do not show up to school every day); one of every ten students drops out or changes school because of bullying; homicide perpetrators were found to be twice as likely as homicide victims to have been bullied previously by their peers; bullying victims are between two to nine times more likely to consider suicide

---

<sup>1</sup>Anti-bullying campaigns and laws have been implemented in the US, Canada, UK, Germany, Scandinavia, Colombia and South Korea.

than non-victims.<sup>2</sup> In the United Kingdom, at least half of suicides among young people are related to bullying. Surprisingly, with few exceptions, the economic literature has mostly stayed away from research efforts to understand the bullying phenomenon.

Suicides and bullying also go hand in hand in the Republic of Korea (South Korea), the country we examine in this paper. Every day one school-aged Korean (ages 10 to 19) commits suicide, the country's largest cause of death among individuals between 15 and 24 years of age.<sup>3</sup> The fierce academic competition resulting from the high value the society gives to education, which in turn makes school grades and test scores extremely high stake events, has been identified as one of the reasons behind the phenomenon. In fact, South Korean households spend 0.8% of the GDP per year out of their pockets on education (more than twice the OECD average) and after-school academies or *hagwon* are increasingly popular.<sup>4</sup> The ultra-competitive high-pressure environment has fueled a climate of aggression that frequently evolves into physical and emotional violence, the underlying factor behind the nation's large number of students reporting to have been bullied every year (more than 77,000 in 2013). In fact, the problem of school violence is so prevalent that in an effort to curb this unwanted behavior in 2012 the South Korean government installed 100,000 closed circuit cameras in schools, and since 2013 private insurance companies offered bullying insurance policies.<sup>5</sup>

This paper assesses the determinants and medium-term consequences of being bullied.

---

<sup>2</sup>See [stopbullying.gov](http://stopbullying.gov).

<sup>3</sup>According to the World Health Organization, the overall suicide rate in South Korea is among highest in the world with 28.9 suicides per 100,000 people (2013).

<sup>4</sup>The degree of competition is such that there are *hagwons* exclusively dedicated to prepare students for the admission processes of the more prestigious *hagwons*. These investments are not remedial measures. They do not aim at helping less proficient individuals to keep up with their classmates. Instead, they are intended to make good students even better than their peers. Such are the incentives to study extra hours that the government had to prohibit instruction in *hagwon* after 10PM. See [www.economist.com/news/asia/21665029-korean-kids-pushy-parents-use-crammers-get-crammers-cr-me-de-la-cram](http://www.economist.com/news/asia/21665029-korean-kids-pushy-parents-use-crammers-get-crammers-cr-me-de-la-cram) and [www.economist.com/node/21541713](http://www.economist.com/node/21541713).

<sup>5</sup>[www.huffingtonpost.com/2014/02/07/south-korea-bullying-insurance\\_n\\_4746506.html](http://www.huffingtonpost.com/2014/02/07/south-korea-bullying-insurance_n_4746506.html).

Our empirical analysis is carried out using South Korean longitudinal information on teenagers, which allows us to examine the extent to which cognitive and non-cognitive skills can deter the occurrence of this unwanted behavior, and also how they might palliate or exacerbate its effects on several outcomes, including depression, life satisfaction, the incidence of smoking, drinking, health indicators and the ability to cope with stressful situations.<sup>6</sup>

Our conceptual framework is based on a structural model of endogenous bullying, multiple outcomes and latent skills. As we described below, the model is flexible enough to incorporate several desirable features. First, it recognizes that cognitive and non-cognitive measures observed by the researcher are only proxies of the true latent skills (Heckman et al., 2006a). Second, the model treats bullying as an endogenous behavior caused, to some extent, by individual’s skills and socio-economic characteristics. Third, flexible functional form assumptions used in the estimation of the latent skills guarantee the flexibility to recreate the patterns observed in the data. Fourth, the structural model allows us to simulate counterfactuals for individuals with different skill levels, which are then used to document heterogenous treatment effects of bullying on several outcomes. Finally, the model does not assume linearity in the estimation. In fact, simulations show that the estimated effects of skills on the outcomes analyzed are very non-linear.

The paper contributes to the literature in several ways. First, to the best of our knowledge, this is the first attempt to assess the determinants and consequences of bullying while dealing with the problems caused by the selection into becoming a victim or a

---

<sup>6</sup>In this paper we follow the literature and define *cognitive skills* as “all forms of knowing and awareness such as perceiving, conceiving, remembering, reasoning, judging, imagining, and problem solving” (APA, 2006), and *non-cognitive skills* as personality and motivational traits that determine the way individuals think, feel and behave (Borghans et al., 2008).

perpetrator of bullying, providing insights that can potentially motivate interventions to reduce their incidence. Second, we provide medium-term impacts of school victimization. In particular, the longitudinal information allows us to examine the transition from high school to adulthood, so we can identify the effects of early victimization on different outcomes observed during a decisive period of development. Third, we provide evidence on how cognitive and non-cognitive skills affect the likelihood of being bullied. Namely, we find that a one standard deviation increase in non-cognitive skills reduces the probability of being bullied by 37%. Fourth, we find that skills not only affect the probability of victimization, but they also palliate the consequences of bullying in subsequent years. In particular, we find that cognitive skills reduce the incidence of bad habits, such as drinking and smoking, proportionally more among bullying victims than among non-victims. Fifth, we quantify the effects bullying has on several behavioral outcomes. For instance, we find that being bullied at age 15 increases the incidence of sickness by 75%, the incidence of mental health issues by 50%, and raises stress levels caused by friendships by 20% of a standard deviation, all by age 18. We also find that there are differential effects of bullying victimization across skill levels. We document that bullying increases depression by 25% of a standard deviation in students from the bottom decile of the non-cognitive skill distribution, and reduces the likelihood of going to college by 10 to 18 percentage points in students that come from the lower half of the non-cognitive skill distribution. We also find that bullying increases the likelihood of smoking by 8.5 percentage points for students in the lowest decile of the cognitive skill distribution.

The paper is organized as follows. Section 2 puts our research in the context of the literature analyzing bullying. Section 3 describes our data. In Section 4, and following the existing literature, we present reduced form results of the impact of bullying on

different outcomes. Section 5 explains the empirical strategy we adopt in this paper. Section 6 presents and discusses our main results. Section 7 concludes.

## 2 Related Literature

Psychology and sociology literatures have been prolific in terms of descriptions of bullying as a social phenomenon. Findings from [Smith et al. \(2004\)](#), for example, show that bullying victims have fewer friends, are more likely to be absent from school, and do not like break times. This literature has also documented that younger kids are more likely to be bullied and that this phenomenon is more frequent among boys than among girls ([Boulton and Underwood, 1993](#); [Perry et al., 1988](#)). Interestingly, [Olweus \(1997\)](#) found that school and class size are not significant determinants of the likelihood of bullying occurrence. Furthermore, [Ouellet-Morin et al. \(2011\)](#) showed that bullying victims' brains have unhealthy cortisol reactions that make it difficult to cope with stressful situations, and according to psychological research, bullied children possess in general less favorable self-evaluation (self-esteem), and have a negative view of their situation ([Björkqvist et al., 1982](#); [Olweus, 1997](#)).

These analyses, although descriptive, provide critical guidelines for the specification of our empirical model. In particular, the characterization of the victims highlights the importance of controlling for skills, particularly non-cognitive dimensions, when analyzing the determinants and potential consequences of bullying.

But unlike in sociology and psychology, economic research on bullying is very scarce. At least two reasons explain this. First and foremost, the lack of information about bullying in both cross-sectional and longitudinal studies; and second, the fact that

selection into this behavioral phenomenon is complex and not random. In this context, the consequences of being bullied can be confounded by the intrinsic characteristics that made the person a victim or a perpetrator in the first place. And while few papers in economics have been able to analyze long longitudinal data, the efforts to control for selection have been more limited.

[Brown and Taylor \(2008\)](#), for example, estimate linear regression models using ordinary least squares (OLS) and ordered probits to examine the associations between bullying and educational attainment as well as wages in the United Kingdom. Their findings suggest that being bullied (and being a bully) is correlated with lower educational attainment and, as a result, with lower wages later in life. [Eriksen et al. \(2012\)](#), on the other hand, use detailed administrative information from Denmark and reduced form regressions (OLS and fixed-effect models) to document strong and robust correlations between bullying and 9<sup>th</sup> grade test scores, teenage pregnancy, use of psychopharmacological medication, height and weight at age 18. Although these are novel efforts, they only indirectly deal with the potential non-randomness of the bullying “treatment”.

By implementing an instrumental variable strategy to control for the endogeneity of bullying, [Eriksen et al. \(2014\)](#) try to deal with this criticism. Using the same data set as in [Eriksen et al. \(2012\)](#), they instrument the effect of teacher-parent reported bullying victimization on 9<sup>th</sup> grade GPA with the proportion of classroom peers whose parents have a criminal conviction. They confirm that bullying reduces GPA, although the estimates are clouded by the low variation of the instrument and by the possible district-neighborhood clustering of convicted parents that may relate with the endogenous sorting of students into schools/classrooms. These findings might also be subject to additional qualifications as the instrumental variable strategy does not necessarily identify causal effects when the selection into treatment (bullying) is driven by unob-

served heterogeneity (latent skills) (Angrist and Imbens, 1994; Heckman et al., 2006b). Our strategy does control for this.

Our work also relates to the literature evaluating the impact of conduct disorders during childhood and adolescence, as bullying can be conceptualized as one. Le et al. (2005), for example, bundle bullying with several other conduct disorders such as stealing, fighting, raping, damaging someone’s property on purpose and conning, among others. Using an Australian sample of twins, these authors control for the potential endogeneity arising from genetic and environmental factors. From the estimation of linear regression models, they find that conduct disorders were positively correlated with dropping-out from school and unemployed later in life. However, they are unable to deal with the endogeneity that might arise from intrinsic individual-level unobservable characteristics influencing both the conduct disorder and the outcome variables they assess.

Our approach is different from those in the existing literature in several ways. First, our empirical strategy incorporates the identification of unobserved heterogeneity in the form of cognitive and non-cognitive skills, showing that both dimensions are key determinants of the treatment and outcomes examined.<sup>7</sup> Second, even though our empirical strategy does not rely on exclusion restrictions in the instrumental variable sense—the structure imposed assumes that the outcomes and treatment equations are independent from each other once we control for observable characteristics and the latent skills (Heckman et al., 2006a)—we use a feature of the Korean schooling system, namely the random allocation of students to classrooms, as a source of exogenous variation affecting the probability of being victimized.<sup>8</sup> Third, we provide medium-

---

<sup>7</sup> The literature has shown that cognitive and non-cognitive skills are critical to the development of successful lives (see for example Murnane et al., 1995; Cawley et al., 2001; Heckman and Rubinstein, 2001; Duckworth and Seligman, 2005; Heckman et al., 2006a; Urzua, 2008).

<sup>8</sup>We provide further explanation on how we use the random allocation of students to classrooms in our empirical strategy in Section 6.1.1 and footnote 19.

term impacts of school victimization on behavioral outcomes. That is, we acknowledge that bullying affects the victims' lives beyond school, and consequently, we quantify its impacts on other future dimensions (e.g., health status, risky behaviors, social relations, life satisfaction and college attendance). Finally, unlike previous studies analyzing the behavioral phenomenon, our empirical strategy follows the recent literature on skill formation (Cunha and Heckman, 2008) and considers cognitive and non-cognitive skills not only as mechanisms that determine the chances of being bullied, but also as traits that palliate or exacerbate its negative effects.

### 3 Data

We use the Junior High School Panel (JHSP) of the Korean Youth Panel Survey (KYP). The KYP-JHSP is a longitudinal survey that started in 2003 sampling a group of second year junior high-school students (i.e., 14 year olds) who were then interviewed once a year until 2008, covering the transition from high-school into the beginning of their adult life. The data fully describes this process including rich longitudinal information on higher education choices from those that go to college and early employment choices for those not enrolling in college.

The sample consists of 12 regions including Seoul Metropolitan City. Children were sampled according to the proportion of the second year junior high-school students present in each region. The resulting panel consists of 3,449 youths and their parents or guardians (see descriptive statistics in Table 1). Subjects were consistently interviewed in six waves.<sup>9</sup> Each year, information was collected in two separate questionnaires: one

---

<sup>9</sup>As in any longitudinal survey, attrition is always an issue. By wave 2, 92% of the sample remained; by wave 3, 91% did so; by wave 4, 90%; and by wave 5, 86% remained in the sample.

for the teenager, and one for the parents or guardians.

As this is a sensitive age range regarding life-path choices, the KYP-JHSP provides a unique opportunity to understand the effects of cognitive and non-cognitive skills on multiple behaviors. It pays special interest in the life-path choices made by the surveyed population, inquiring not only about their decisions, but also about the environment surrounding their choices. For example, youths are often asked about their motives and the reasons that drive their decision-making process. Future goals and parental involvement in such choices are frequently elicited as well.

Besides inquiring about career planning and choices, the KYP-JHSP inquires about academic performance, student effort and participation in different kinds of private tutoring. The survey also asks about time allocation, leisure activities, social relations, attachment to friends and family, participation in deviant activity, and the number of times the respondent has been victimized in different settings. In addition, the survey performs a comprehensive battery of personality questions from which measures of self-esteem, self-stigmatization, self-reliance, aggressiveness, anger, self control and satisfaction with life can be constructed.

While the youngsters are often asked about the involvement of their parents in many aspects of their life, parents and guardians answer only a short questionnaire covering household composition and their education, occupation and income.

**The Construction of Non-Cognitive Measures.** The KYP-JHSP contains a comprehensive battery of measures related to personality traits. Among them, we select the scales of locus of control, irresponsibility and self-esteem. Locus of control relates to the extent to which a person believes her actions affect her destiny, as opposed to a person that believes that luck is more important than her own actions ([Rotter, 1966](#)). People

with internal locus of control face life with a positive attitude as they are more prone to believe that their future is in their hands (Tough, 2012). The irresponsibility measure relates negatively to perseverance and grit, that is, the ability to overcome obstacles and giving proportionally greater value to large future rewards over smaller immediate ones (Duckworth and Seligman, 2005). Finally, self-esteem provides a measure of self-worth.

The choice of these variables is backed by research in economics and psychology that shows that each of these personality traits are important determinants of future outcomes and the likelihood of victimization. For instance, works like Duckworth and Seligman (2005), Heckman et al. (2006a) and Urzua (2008) show that the unobserved heterogeneity captured by these measures are strong predictors of adult outcomes. In fact, our results presented in the Appendix A attest to that.<sup>10</sup> In the same vein, psychology literature like Björkqvist et al. (1982), Olweus (1997) and Smith and Brain (2000) provide evidence that shows that traits chosen correlate with school bullying victimization. People with external locus of control or a higher degree of perseverance may have a greater inclination to avoid/change a victimization situation. On the other hand, self-esteem may be proxying unobservable characteristics that make a school-age individual more popular and sociable, therefore affecting her chances of being bullied (Hodges and Perry, 1996).

It should be noted that most of the socio-emotional information in the KYP-JHSP is recorded in categories that group the reactions of the respondent in categories from “strongly agree” to “strongly disagree”. In consequence, and following common practice in the literature, we constructed socio-emotional skill measures by adding categorical answers of several questions regarding the same topic. This method incorporates some

---

<sup>10</sup>In footnote 18 we comment these results in detail.

degree of continuity in the scores, which is essential for our estimation procedure. The questions used can be found in Appendix B, and Table 2 show the descriptive statistics of the constructed measures.

**Academic Performance.** Although the KYP-JHSP has a rich battery of behavioral questions, it is somewhat limited regarding cognitive measures. Ideally, we would like to have measures closely linked to pure cognitive ability. However, the lack of such measures forces us to infer cognitive ability from grades and academic performance. In particular, we use the scores obtained in tests of i) math and science; ii) language (Korean) and social studies; and iii) overall school grades reported every year. See Table 3 for the descriptive statistics of these measures.

Importantly, previous literature has shown that academic performance is not orthogonal to non-cognitive skills (Borghans et al., 2011). In other words, the production function of academic test scores has to be modeled using both cognitive and socio-emotional skills as inputs. As will be shown in Section 5, our model takes into account this feature of the data and incorporates it into the estimation.

**Bullying.** In the KYP-JHSP, bullying is self-reported by the students. It refers to events where they have been severely teased or bantered, threatened, collectively harassed, severely beaten, or robbed. Therefore, it relates to the more direct and less subtle versions of bullying. The ones, according to psychological research, children tend to focus more on (Naylor et al., 2010). The reported incidence of bullying in the KYP-JHSP, presented in Table 4, is in line with the incidence reported in international studies for the ages studied (see Smith and Brain (2000), for a summary).

**Outcomes.** The outcomes we consider include depression, the likelihood of smoking, drinking, and college attendance, life satisfaction, self reported physical and mental

health, and standardized indexes of stress. All these are measured by age 18.

Specifically, we analyze an index of depressions which is constructed based on a battery of questions that assess its symptoms. Self reported physical health, on the other hand, measures whether the respondent considers she is in good health or not. The mental health outcome captures whether the respondent has been diagnosed to have psychological or mental problems. Regarding the stress measures, we use five outcome measures. The first four quantify how stressed the respondent gets regarding her friends, her parents, school and poverty. The last stress measure captures the stress levels associated with self-image. We also examine the impact of early bullying on college attendance by age 19. Descriptive statistics of the outcome variables are presented in Table 4.

## 4 Reduced Form Results

We would like to inquire about the effect of bullying at time  $t$ ,  $D_t$ , on the outcome of interest measured at time  $t + h$ ,  $Y_{t+h}$ . Following the literature (Brown and Taylor, 2008; Eriksen et al., 2012), we posit the following regression model:

$$Y_{t+h} = \mathbf{X}_Y \beta + \gamma D_t + e_{t+h} \tag{1}$$

where  $\mathbf{X}_Y$  is a matrix with all observable controls and  $e_{t+h}$  denotes the error term. In this paper we assume that bullying is an endogenous behavior driven by individual characteristics (latent skills) potentially correlated with the error term, so we would expect  $D_t \not\perp e_{t+h}$  (conditional on  $\mathbf{X}_Y$ ). This prevents the direct interpretation of the OLS estimator of  $\gamma$  as a causal effect, but the availability of proxies for skills can help

to mitigate the issue.

Let  $\mathbf{T}$  be a vector containing individual-level cognitive and non-cognitive test scores. In order to deter reverse causality, we assume these measures are taken before the bullying episode occurs, that is at time  $t - 1$ . Thus, the regression equation becomes:

$$Y_{t+h} = \mathbf{X}_Y\beta + \gamma D_t + \pi\mathbf{T}_{t-1} + \nu_{t+h} \quad (2)$$

where the identification of  $\gamma$  relies on the conditional mean independence assumption of the error term,  $\nu_{t+h}$ , with respect to the set of controls (including test scores).

Table 5 shows the results from the estimation of (1) and (2) using the KYP-JHSP data. The reduced form results suggest positive correlations between being bullied at 15 and depression, the likelihood of being sick, life satisfaction and feeling stressed at 18. The estimates, however, do not suggest significant correlations between being bullied and drinking or smoking (age 18).

But findings from regressions like (2) are problematic. On the one hand, even after controlling for test scores,  $\mathbf{T}_{t-1}$ , bullying  $D_t$  might be still endogenous. On the other hand, the literature has documented that components of  $\mathbf{T}_{t-1}$  should only be interpreted as proxies for skills (Heckman et al., 2006a). Therefore, as we confirm below, including  $\mathbf{T}_{t-1}$  in the set of controls might not be enough to control for the potential sources of biases arising from the endogeneity of bullying and the measurement error problems produced by test scores. Our empirical strategy deals with these issues.

## 5 Empirical Strategy

In this paper, we not only assume that underlying cognitive and non-cognitive skills are latent rather than observable, but also that they are determinants of—and the source of dependence among—outcomes, treatment (bullying) and test scores.

The core of the empirical strategy follows Heckman et al. (2006a), so it assumes a linear production function of test scores, whose inputs include both the individual observable characteristics and the latent skill endowments.<sup>11</sup> The insight provided by Kotlarski (1967) allows us to use observed test scores to identify the underlying distributions from which the realizations of latent endowments are drawn, facilitating the estimation of the complete structural model.

The model is estimated using maximum likelihood, incorporating mixture of Gaussian distributions to mimic the true underlying distributions of the latent skill endowments. This feature of the model grants the flexibility required to replicate the data.<sup>12</sup>

### 5.1 The General Setup

The structural models we implement can be described as a set of outcome equations linked by a factor structure with two underlying factors,  $\theta^A$  and  $\theta^B$ . In a general setup,

---

<sup>11</sup>In fact, variance decompositions of the test scores presented in Figures 1 show that latent endowments explain between 5 to 10 time more the variation of the scores than the observable characteristics. However, these figures also show that even after controlling for latent endowments more than half of the variation of the scores is still unexplained. These findings go in line with our argument against the use test scores as proxies for abilities in Section 4. The unexplained part of the variance of test scores will correlate with  $\nu_{t+h}$  in (2) biasing the results of the regressions. That is why we will rather identify the latent endowments from the test scores.

<sup>12</sup>The structural estimations presented in this paper were done using the `heterofactor` command for Stata. See Sarzosa and Urzua (2015).

suppose we face the following system for each individual in the sample:

$$\mathbf{Y} = \mathbf{X}_Y \beta^Y + \alpha^{Y,A} \theta^A + \alpha^{Y,B} \theta^B + \mathbf{e}^Y \quad (3)$$

where  $\mathbf{Y}$  is a  $M \times 1$  vector of outcome variables,  $\mathbf{X}_Y$  is a matrix with all observable controls for each outcome variable,  $\alpha^{Y,A}$  and  $\alpha^{Y,B}$  are vectors that contain the factor loadings for each underlying factor (i.e.,  $\theta^A$  and  $\theta^B$ ), and  $\mathbf{e}^Y$  is a vector of independent error terms (across individuals and outcomes) with distributions  $f_{e^Y m}(\cdot)$  for every  $m = 1, \dots, M$ . We assume that  $\mathbf{e}^Y \perp (\theta^A, \theta^B, \mathbf{X}_Y)$ . Furthermore, we assume the factors  $\theta^A$  and  $\theta^B$  follow the distributions  $f_{\theta^A}(\cdot)$  and  $f_{\theta^B}(\cdot)$ , respectively.

If  $M = 1$  then  $\mathbf{Y} = Y$  and (3) becomes

$$Y = \mathbf{X}_Y \beta^Y + \alpha^{Y,A} \theta^A + \alpha^{Y,B} \theta^B + e^Y \quad (4)$$

This general framework can be used to introduce a special case, relevant for this paper, in which there is a binary treatment  $D$  (e.g., being bullied or not) and a vector of subsequent factual and counterfactual outcomes, i.e., a model of potential outcomes inspired by the Roy model (Roy, 1951). In particular, let  $D$  be a dummy variable denoting whether or not an individual has a been victim of bullying and for the sake of exposition clarity, let's assume  $M = 1$  (e.g., the outcome of interest is depression at age 18). We posit the following model for  $D$ :

$$D = \mathbb{1} [\mathbf{X}_D \beta^{Y_D} + \alpha^{Y_D,A} \theta^A + \alpha^{Y_D,B} \theta^B + e^D > 0]$$

where  $\mathbb{1}[A]$  denotes an indicator function that takes a value of 1 if  $A$  is true, and  $\mathbf{X}_D$  represents a set of exogenous observables. Let  $Y_0, Y_1$  denote an outcome of interest

(e.g., the likelihood of depression) for those with  $D = 0$  and  $D = 1$  respectively (e.g., non victims and bullying victims). Then, the system of equations (5) will represent both potential outcomes and the choice equation. That is,  $\mathbf{Y} = [Y_1, Y_0, D]'$  where:

$$Y_1 = \mathbf{X}_Y \beta^{Y_1} + \alpha^{Y_1,A} \theta^A + \alpha^{Y_1,B} \theta^B + e^{Y_1} \quad \text{if } D = 1 \quad (5)$$

$$Y_0 = \mathbf{X}_Y \beta^{Y_0} + \alpha^{Y_0,A} \theta^A + \alpha^{Y_0,B} \theta^B + e^{Y_0} \quad \text{if } D = 0 \quad (6)$$

$$D = \mathbb{1} [\mathbf{X}_D \beta^{Y_D} + \alpha^{Y_D,A} \theta^A + \alpha^{Y_D,B} \theta^B + e^D > 0] \quad (7)$$

where  $\mathbf{X}_Y$  are a set of observable variables. Note that under our assumptions, although  $D$  is endogenous, once we control for unobserved heterogeneity  $(\theta^A, \theta^B)$ , the error terms  $(e^{Y_1}, e^{Y_0}, e^D)$  are jointly independent. Hence, this empirical strategy controls for the unobserved heterogeneity solving the problem of endogeneity (Heckman et al., 2006a). Furthermore, as it uses the underlying the latent skills instead of observed test scores, it avoids the measurement error problem affecting the reduced form results.

We complement our factor model with a measurement system of the form:

$$\mathbf{T} = \mathbf{X}_T \beta^T + \alpha^{\mathbf{T},\mathbf{A}} \theta^A + \alpha^{\mathbf{T},\mathbf{B}} \theta^B + \mathbf{e}^T \quad (8)$$

where  $\mathbf{T}$  is a  $L \times 1$  vector of measurements (e.g., test scores),  $\mathbf{X}_T$  is a matrix with all observable controls for each measurement, and  $\alpha^{\mathbf{T},\mathbf{A}}$  and  $\alpha^{\mathbf{T},\mathbf{B}}$ , are the loadings of the unobserved factors. Again, we assume that  $(\theta^A, \theta^B, \mathbf{X}_T) \perp \mathbf{e}^T$ , that all the elements of the  $L \times 1$  vector  $\mathbf{e}^T$  are mutually independent with associated distributions  $f_{e^h}(\cdot)$  for every  $h = 1, \dots, L$ . We follow Carneiro et al. (2003) and secure the identification of the factor loadings and underlying distributions (up to normalizations) imposing:  $\theta^A \perp \theta^B$  (independence) and  $L \geq 2k + 1$  (enough equations in the measurement system), where  $k$

is the number of factors. Therefore, the presence of two factors in (5) implies that there should be *at least* five measures in (8). Finally, we normalize one loading associated with each factor to one. This anchors the scale of each latent factor to a specific test score. Other loadings should be interpreted as relative to those used as numeraires.<sup>13</sup>

We estimate the model using a two-stage maximum likelihood estimation (MLE) procedure. The likelihood is

$$\mathcal{L} = \prod_{i=1}^N \int \int \left[ f_{e^1}(\mathbf{X}_{T_1i}, T_{1i}, \zeta^A, \zeta^B) \times \cdots \times f_{e^L}(\mathbf{X}_{T_Li}, T_{Li}, \zeta^A, \zeta^B) \right] dF_{\theta^A}(\zeta^A) dF_{\theta^B}(\zeta^B)$$

where we integrate over the distributions of the latent factors, obtaining  $\hat{\beta}^T, \hat{\alpha}^{T,A}, \hat{\alpha}^{T,B}$  and the parameters of the mixture of normal distributions describing the  $\hat{F}_{\theta^A}(\cdot)$  and  $\hat{F}_{\theta^B}(\cdot)$ .

Having identified the parameters characterizing  $F_{\theta^A}(\cdot)$  and  $F_{\theta^B}(\cdot)$  from (8), we can estimate the model described by equations (5), (6) and (7). The likelihood function in this case is

$$\mathcal{L} = \prod_{i=1}^N \int \int \left[ \begin{aligned} & [f_{e^{y_0}}(\mathbf{X}_{Y_0i}, Y_{0i}, \zeta^A, \zeta^B)]^{D_i=0} \times \Pr[D_i = 0 | \mathbf{X}_{Di}, \zeta^A, \zeta^B] \\ & \times [f_{e^{y_1}}(\mathbf{X}_{Y_1i}, Y_{1i}, \zeta^A, \zeta^B)]^{D_i=1} \times \Pr[D_i = 1 | \mathbf{X}_{Di}, \zeta^A, \zeta^B] \end{aligned} \right] d\hat{F}_{\theta^A}(\zeta^A) d\hat{F}_{\theta^B}(\zeta^B)$$

This procedure yields  $\hat{\beta}^Y, \hat{\alpha}^{Y,A}$  and  $\hat{\alpha}^{Y,B}$ .<sup>14</sup>

---

<sup>13</sup>The basic idea behind the identification argument follows Kotlarski (1967). More precisely, if there are three independent random variables  $e_{T_1}$ ,  $e_{T_2}$  and  $\theta$  and define  $T_1 = \theta + e_{T_1}$  and  $T_2 = \theta + e_{T_2}$ , the joint distribution of  $(T_1, T_2)$  determines the distributions of  $e_{T_1}$ ,  $e_{T_2}$  and  $\theta$ , up to one normalization. Note that, given that we have already identified all the loadings, we can write (8) in terms of  $T_\tau = \theta + e_{T_\tau}$  by dividing both sides by the loading. See more details in Carneiro et al. (2003).

<sup>14</sup>In this two-step procedure, we use a Limited Information Maximum Likelihood and correct the variance-covariance matrix of the second stage incorporating the estimated variance-covariance matrix and gradient of the first stage (Greene, 2000). Note that the two steps presented above can be joined

## 5.2 Measuring Skills and Joint Causality

The implementation of the two-factor model requires additional specificities due to the way the data has been collected. First, as explained in Section 3, cognitive skills need to be inferred from grades, which in turn are also affected by non-cognitive skills. As shown by Carneiro et al. (2003), this is easily accommodated by the model by allowing a more general factor loading matrix. To see this, let us write the measurement system (8) in a more compact matrix representation:

$$\mathbf{T} = \mathbf{X}_T \beta^T + \Lambda \Theta + \mathbf{e}^T \quad (9)$$

where  $\Theta = \begin{bmatrix} \theta^A & \theta^B \end{bmatrix}$  is a vector with the factors, and  $\Lambda$  is the factor loading matrix. Our empirical specification of the model assumes:

$$\Lambda = \begin{bmatrix} \alpha^{T_1,A} & \alpha^{T_1,B} \\ \alpha^{T_2,A} & \alpha^{T_2,B} \\ \alpha^{T_3,A} & \alpha^{T_3,B} \\ \alpha^{T_4,A} & \alpha^{T_4,B} \\ \alpha^{T_5,A} & \alpha^{T_5,B} \\ \alpha^{T_6,A} & \alpha^{T_6,B} \end{bmatrix} = \begin{bmatrix} \alpha^{T_1,A} & 0 \\ \alpha^{T_2,A} & 0 \\ 1 & 0 \\ \alpha^{T_4,A} & \alpha^{T_4,B} \\ \alpha^{T_5,A} & \alpha^{T_5,B} \\ \alpha^{T_6,A} & 1 \end{bmatrix}$$

This triangular factor loading matrix incorporates the fact that the KYP-JHSP provides “pure” measures of non-cognitive skills, while providing measures of cognitive ability in

---

and calculated in one likelihood of the form:

$$\mathcal{L} = \prod_{i=1}^N \int \int \left[ \begin{array}{l} [f_{e^{y_0}}(\mathbf{X}_{Y_0i}, Y_{0i}, \zeta^A, \zeta^B)]^{D_i=0} \times \Pr[D_i = 0 | \mathbf{X}_{D_i}, \zeta^A, \zeta^B] \\ \times [f_{e^{y_1}}(\mathbf{X}_{Y_1i}, Y_{1i}, \zeta^A, \zeta^B)]^{D_i=1} \times \Pr[D_i = 1 | \mathbf{X}_{D_i}, \zeta^A, \zeta^B] \\ \times f_{e^1}(\mathbf{X}_{T_1i}, T_{1i}, \zeta^A, \zeta^B) \times \dots \times f_{e^L}(\mathbf{X}_{T_Li}, T_{Li}, \zeta^A, \zeta^B) \end{array} \right] dF_{\theta^A}(\zeta^A) dF_{\theta^B}(\zeta^B)$$

the form of academic test scores, which likely reflect both cognitive and non-cognitive skills.

The second issue that arises from how the data was collected is the fact that skills are measured during the school year, therefore some kids may have been already exposed to the treatment prior to the skill measurement. This may cause a problem of joint causality, closely related to the effect of schooling at the time on the skill measures explored in [Hansen et al. \(2004\)](#). In their setting, the simultaneity issue comes from the fact that highly skilled people might attain higher education levels, but schooling, in turn, is believed to develop skills. Hence, when in presence of a high-skilled highly-educated person, econometricians have a hard time disentangling whether the person is highly-educated because she was highly skilled or she is highly skilled because she acquired more education. We follow the literature and impose the appropriate structure to control for the joint causality problem. In particular, the measurement system (9) needs to be modified to incorporate the probability of being treated at  $t = 1$ .

$$\mathbf{T} = \begin{cases} \mathbf{X}_T \beta_{D_1=1}^T + \Lambda_{D_1=1} \Theta' + \mathbf{e}_{D_1=1}^T & \text{if } D_1 = 1 \\ \mathbf{X}_T \beta_{D_1=0}^T + \Lambda_{D_1=0} \Theta' + \mathbf{e}_{D_1=0}^T & \text{if } D_1 = 0 \end{cases} \quad (10)$$

$$D_1 = \mathbb{1} [\mathbf{X}_{D_1} \beta^{Y_{D_1}} + \Lambda_{D_1} \Theta' + e^{D_1} > 0]$$

Notice that the matrix  $\Lambda$  is expanded to incorporate different loadings in the test score production functions depending on the state of bullying when test scores are collected,  $D_1$ . As a result of this, the loading normalization and test measurement

system structure required for identification become:

$$\Lambda = \begin{bmatrix} \Lambda_{D_1=1} \\ \Lambda_{D_1=0} \\ \Lambda_{D_1} \end{bmatrix}$$

where

$$\Lambda_{D_1=1} = \begin{bmatrix} \alpha_{D_1=1}^{T_1,A} & 0 \\ \alpha_{D_1=1}^{T_2,A} & 0 \\ 1 & 0 \\ \alpha_{D_1=1}^{T_4,A} & \alpha_{D_1=1}^{T_4,B} \\ \alpha_{D_1=1}^{T_5,A} & \alpha_{D_1=1}^{T_5,B} \\ \alpha_{D_1=1}^{T_6,A} & 1 \end{bmatrix}, \quad \Lambda_{D_1=0} = \begin{bmatrix} \alpha_{D_1=0}^{T_1,A} & 0 \\ \alpha_{D_1=0}^{T_2,A} & 0 \\ \alpha_{D_1=0}^{T_3,A} & 0 \\ \alpha_{D_1=0}^{T_4,A} & \alpha_{D_1=0}^{T_4,B} \\ \alpha_{D_1=0}^{T_5,A} & \alpha_{D_1=0}^{T_5,B} \\ \alpha_{D_1=0}^{T_6,A} & 1 \end{bmatrix}, \quad \Lambda_{D_1} = \begin{bmatrix} \alpha^{D_1,A} & \alpha^{D_1,B} \end{bmatrix}$$

## 6 Main Results

### 6.1 Measurement System

Table 6 displays the estimated parameters of the measurement system (10). The set of controls  $\mathbf{X}_T$  includes gender, family structure indicators, father's education attainment, monthly income per capita and the age stated in months starting from March 1989.<sup>15</sup>

As expected, cognitive and non-cognitive skills determine school grades as well as non-cognitive measures. In fact, our findings suggest that the loadings in  $\Lambda$  are large and

---

<sup>15</sup>All individuals in our sample were born within the same academic year, which goes from March to February.

statistically different from zero at the 1% level.<sup>16</sup> The estimates of  $\beta_{D_1=0}^T$  and  $\beta_{D_1=1}^T$ , on the other hand, imply that youths with wealthier and more educated parents tend to be more responsible, have higher levels of self-control and are more positive about themselves. These results are consistent with those in [Cunha et al. \(2006\)](#) and [Heckman and Masterov \(2007\)](#). Our estimates also suggest that family composition plays a big role in fostering desirable personality traits. Individuals with younger siblings and those who live with both parents tend to be more responsible. Interestingly, those who live with their mother have substantially higher levels of self-esteem than those who live only with their father. As with the non-cognitive measures, the cognitive scores are higher for individuals coming from wealthier and more educated parents, especially if the mother is present in the family. In addition, the presence of younger (older) siblings is associated with higher (lower) grades. Another notable finding, which is in line with [Borghans et al. \(2008\)](#), is that younger students are less responsible and have less self-control and self-esteem, even within the same year of age.

Using the estimates of equations (10) we are able to recreate the estimated distributions of initial non-cognitive and cognitive skills across the population evaluated at  $t = 1$ . These distributions are presented in [Figures 2a](#) and [2b](#).

### 6.1.1 The Bullying Model

We now analyze the determinants and impact of bullying at age 15 ( $t = 2$ ).<sup>17</sup> To do so we implement the model described by equations (5), (6) and (7). As outcomes, we use

---

<sup>16</sup>Figure 1 presents the variance decomposition of the test scores. It shows that the unobserved endowments represent a sizable proportion of the variance of the scores, being always more prominent than the variance captured by the observable controls.

<sup>17</sup>As explained in the previous section, latent skills were measured one survey wave before (i.e.,  $t = 1$ ).

the same measures examined in the context of our reduced form results (see section 4).<sup>18</sup> We are particularly interested in assessing the relation between skills, the problematic behavior and outcomes of interest.

The controls in the binary model describing bullying include age, gender, family structure variables, urban dummy, and father’s education. Moreover, in order to facilitate identification of the system of equations, we introduced an additional source of variation in the choice or treatment equation. For that, we use a very special institutional feature of South Korea: the fact that allocation of students to classrooms is random within school districts (Kang, 2007). Hence, the quality and characteristics of classmates a given student faces are exogenous, and affect each student’s probability of being bullied in a given classroom (Sarzos, 2015). In consequence, we introduce two of these classroom characteristics: the proportion of peers that report being bullies in the class and the proportion of peers in the classroom that come from a violent family.<sup>19</sup>

---

<sup>18</sup> Tables A.1 and A.2 in Appendix A present the estimated parameters associated with equation (4). That is, without the introduction of a treatment variable. It is presented as a benchmark to see how the outcome variables are affected overall by the observable and unobservable characteristics of the respondents. These results indicate that non-cognitive skills measured at age 14 are negatively associated with the likelihood of depression, the incidence of drinking and smoking, the likelihood of being sick, having mental health issues, or feeling stressed about friends and the economic situation at age 18. Furthermore, non-cognitive skills have a positive effect on the likelihood of having a positive perception of life. This is linked with the fact that while non-cognitive skills reduce the likelihood of depression, cognitive skills increase it. Just like happens with the stress variables. However, the reduction on the likelihood of depression is much larger than the increase in the likelihood of depression caused by cognitive skills. We find no effect of cognitive skills on the incidence of drinking alcohol, feeling sick or having mental health issues, while we find that cognitive skills are highly rewarded in the selection into college. Finally, our results indicate that both cognitive and non-cognitive skills reduce the incidence of smoking.

<sup>19</sup>We construct a measure of family violence based on the following questions: 1. I always get along well with brothers or sisters, 2. I frequently see parents verbally abuse each other, 3. I frequently see one of my parents beat the other one, 4. I am often verbally abused by parents, and 5. I am often severely beaten by parents. There are five possible answers to each of these five questions that range from very true to very untrue. We aggregated the answers and considered as students that come from violent families those who have an aggregate score of family violence above the mean. Then, we counted how many of these peers each student faced in her classroom as a proportion of the total number of students in the classroom. This variable is somewhat similar to the classroom proportion of incarcerated parents variable used as instrument by Eriksen et al. (2014) in that we relate household emotional trauma with violent behavior in school as in Carrell and Hoekstra (2010). However, we

The first variable affects the probability of being bullied it accounts for the fact that there are more suppliers of violence in the classroom.<sup>20</sup> The second variable seeks to capture the well established fact that youths with behavioral challenges are more likely to come from violent households (Carlson, 2000; Wolfe et al., 2003). Hence, randomly formed classrooms in which there are more students that come from violent families are more prone to witness violent behavior than classrooms with a lower concentration of students that come from violent families. Again, the incidence of victimization in the former type of classroom is expected to be higher because there are potentially more suppliers of violence.

Table 7 shows the results for three different specifications of the bullying equation (7) whereas Tables 8 and 9 do so for equations (5) and (6) (outcomes). The most salient finding of Table 7, regardless of the specification, is that while cognitive skills do not play a role in deterring or motivating the undesired behavior, non-cognitive skills are a very important determinants of the likelihood of the event. Our findings indicate that a one standard deviation increase in non-cognitive skills translates into a 4.16 percentage points reduction in the likelihood of being bullied. That is, one standard deviation increase in non-cognitive skills reduces by 37% the overall probability of being a victim of bullying. Notably, this significant effect remains unchanged throughout the three different specifications.<sup>21</sup> Table 7 also shows that the availability of suppliers

---

believe our measure is better in at least two ways. First, it results from reports of actual verbal and physical violence experienced by the students in their homes as opposed to a proxy of that, which parental incarceration is. Second, much of the domestic violence may happen within families that do not go to the extent of having a parent incarcerated. In that way, Eriksen et al. (2014) instrument is capturing extreme antisocial behavior, leaving out much of the relevant variation in domestic violence. Finally, it should be noted that unlike Eriksen et al. (2014) these variables (i.e., proportion of bullies and proportion of peers with violent families) are not crucial for our identification strategy which relies on the identification of unobserved heterogeneity.

<sup>20</sup>Importantly, there is always at least one bully and one victim per classroom in our data.

<sup>21</sup>Evidence of the sorting into treatment as a function of unobserved skills can be seen in Figure 3. Those selected as victims had a distribution of non-cognitive skills that lie to the left of that of non-victims.

of violence within each classroom matters. In fact, all else evaluated at the mean, to go from a classroom with a concentration of bullies in the 25<sup>th</sup> percentile to one in the 75<sup>th</sup> percentile increases the likelihood of being a victim by 2.26 percentage points. This represents an increase of 20% in the overall probability of being a victim of bullying. In the same vein, the marginal effect of increasing the concentration of peers in the classroom that come from violent families is positive and linearly increasing. For instance, although the marginal effect at the median is roughly zero, it reaches 9.15 percentage points at the classrooms in the 75<sup>th</sup> percentile of concentration of peers from violent families. Our finding also indicates that bullying is more prevalent among boys than among girls. This goes in line with psychological literature (Olweus, 1997; Wolke et al., 2001; Smith et al., 2004; Faris and Felmlee, 2011).

The results presented in Tables 8 and 9 indicate that skills have differential effects on the outcomes of interest depending on whether the person was involved in bullying or not.<sup>22</sup> These findings suggest that skills not only influence the likelihood of being involved in bullying, but they also play a role in dealing with the negative consequences after the bullying event has occurred. For instance, cognitive skills tend to deter drinking and smoking more among victims of bullying than among non-victims. In the same way, non-cognitive skills reduce stress more among victims than among non-victims. So regardless of whether bullying has large or small consequences on a particular dimension—which is the topic of Section 6.2, skill endowments help cope with these consequences in various ways depending on the outcome.

Although these findings are very informative, they do not inform about the effect of

---

<sup>22</sup>The results presented in Tables 8 and 9 and the subsequent simulations we obtained from a Roy model where the treatment equation followed specification (2) in Table 7. The results do not change if we use the remaining specifications for the treatment equations. For the sake of brevity, these results are not presented in the paper, but are available from the authors upon request.

bullying on later outcomes. We commit to this task in the next section.

## 6.2 The Effect of Bullying

The empirical strategy enables us to identify factual and counterfactual outcome levels as a function of the latent factors. This allows the estimation of the average treatment effect (ATE) and the treatment effect on the treated (TT) of being bullied conditional on each level of cognitive and non-cognitive skills. Formally, and using the notation introduced in 5.1, we estimate  $ATE(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C]$  and  $TT(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C, D = 1]$ . We also present versions of these parameters after integrating out cognitive and non-cognitive skills.

But before presenting the treatment effect results, we first document how the model replicates key features of the data. To this end, we use the results described in Section 6.1 to simulate different outcomes and compare them to the actual data.

While the data shows that 11.07% of the sample declares being bullied, our model predicts that 11.08% of the sample receives the “bullied treatment”. With respect to the other outcomes of interest, we compare the simulated and actual conditional means:  $E[Y_0 | D = 0]$  and  $E[Y_1 | D = 1]$ . Table 10 shows these data-model comparisons of the outcome variables. We see that our model is able to precisely recreate the observed average outcomes. These results do not only give us confidence about the capability of the model to fit the data, but also to generate counterfactuals:  $E[Y_1 | D = 0]$  and  $E[Y_0 | D = 1]$ .

### 6.2.1 ATE and TT of Being Bullied

Table 11 presents the unconditional ATE and TT estimates of being bullied at age 15 on future outcomes (age 18 and older). We document significant effects of victimization on feeling sick, mental health issues and stress caused by relationships with friends and parents, as well as in the total stress index. In particular, our results indicate that being bullied at age 15 causes the incidence of sickness to increase by about 75% three years later. In the same way, the incidence of mental health issues is increased by half among the treated due to victimization. Regarding the stress measures, we find that being bullied increases the stress caused by friendships by 20% of a standard deviation and the stress caused by the relationship with parents by 15% of a standard deviation. These findings contrast to the ones reported in the reduced form estimations, which ignore the endogeneity caused by the selection into treatment. For instance, while we find no overall effect of bullying on depression, life satisfaction and college attendance, the OLS estimates found an effect of -13.4%, -4.1 and -4.8 percentage points respectively.

Although we have calculated the overall ATE and TT, we can use our empirical strategy to inquire about these treatment parameters at different regions of the skills space. That is, we estimate treatment effects conditional on skill levels, with the intention of inquiring about subsets of teenager who face impacts even in the absence of impacts in the overall sample. These results are best presented in standard two-dimensional figures documenting the association between each skill (x-axis) and the outcome of interest (y-axis). However, for outcomes where the interaction between skills is important, we use three-dimensional graphs with cognitive and non-cognitive skills in the  $x$  and  $y$  axes and the treatment effect in the vertical axis.

Figures 4a to 15c show that there are in fact differential effects of victimization depend-

ing on the level of skills. In particular, these figures show that bullying causes people with low non-cognitive skills to face higher incidence of depression (Figure 4a), higher likelihood of feeling physically and mentally ill (Figures 6a and 7a), and higher levels of stress three years later (Figures 10a, 11a and 15a). Bullying also causes non-cognitive unskilled people to have less life satisfaction (Figure 8a) and reduce their likelihood of going to college (Figure 9a). Some of these effects on people with low non-cognitive skills are sizable. For instance, Figure 4a shows bullying increases the depression symptom index by about one fourth of a standard deviation among teenagers that belong to the first decile of the non-cognitive skill distribution. Figure 7a shows that victimization has a positive effect of about 0.8 percentage points in the likelihood of having mental health issues among those teenagers with levels of non-cognitive skills below the median. In the same vein, Figure 8a shows that bullying reduces likelihood of being satisfied with life by more than 20 percentage points for the students that belong to the first decile of the non-cognitive skill distribution.

Evidence from the psychological literature (e.g., [Smith et al. \(2004\)](#)) and the figures presented in the introduction, also suggest that bullying might affect schooling attainment, particularly by fostering a dislike for school that contributes to absenteeism and school drop out. By documenting the effect bullying has on college enrollment and stress caused by school, a measure that proxies a dislike for school and its related activities, we shed light on this idea.

Figure 9a shows that bullying is an important deterrent in tertiary education enrollment (by age 19). Teenagers that belong to the lower half of the non-cognitive skill distribution face a negative impact of bullying on college enrollment of the order of 10 to 18 percentage points. This is remarkable specially if we take into account that non-cognitive skills are not important in determining college enrollment (see Table A.2

in the appendix). However, bullying does have an impact among those with low non-cognitive skills. For them, the behavioral problem becomes an obstacle to higher education attainment. This finding also relates to the potential effect of victimization on the stress caused by school. Figure 10a indicates that although the overall ATE is not statistically different from zero, the impact of bullying on stress is significant for the students who possess non-cognitive skills below the median. In fact, the school stress is half of a standard deviation greater for the bullying victims than for the non-victims in the first decile of the non-cognitive distribution. All this evidence goes in line with the claim that bullying is a very harmful mechanism through which violence deters learning and schooling achievement, providing a channel through which the findings of Eriksen et al. (2014) on its effect on GPA materialize.

The effect bullying has on the stress caused by friendships is also greatly affected by the level of non-cognitive skill endowments. Figure 11a shows that this effect reaches half of a standard deviation among those in the bottom third of the non-cognitive skills distribution. Moreover, Figure 11b shows that this effect is exacerbated for cognitive skilled teenagers. In particular, the effect of bullying on a kid that comes from the first decile of the non-cognitive skills distribution and from the top decile of the cognitive skills distribution reaches 70% of a standard deviation in the stress caused by friendship scale. This contrasts to overall estimated ATE, which is 19.4% of a standard deviation.

Unlike previous outcomes the effect of bullying on smoking is mediated by cognitive instead of non-cognitive skills. Figure 12b documents significant effects of bullying on smoking for those who belong to the first decile of the cognitive skill distribution. For them, bullying increases the likelihood of smoking by about 8.5 percentage points. Furthermore, Figure 12c shows that among those in the first decile of the cognitive skill distribution, the effect is greater for those who lack non-cognitive skills. A similar

pattern emerges in the analysis of the perception of stress due to economic conditions. Figure 13c shows that bullying has a significant effect only for those cognitively skilled who, at the same time, lack non-cognitive skills. For them, the effect is greater than 20% of a standard deviation.

On the other hand, the effect of victimization on the stress caused by the relationship with the parents is particularly stable across the entire skills space, reaching approximately 15% of a standard deviation. However, Figure 14c shows that the effect is smaller for those that are skilled in both dimensions. In fact, the effect is no longer statistically different from zero for this population.

All these results attest to the fact that cognitive and non-cognitive skills not only affect bullying occurrence, but also, they mediate the extent to which these undesired behaviors affect subsequent outcomes.

### 6.3 Bullying and Investments in Skills

We have shown that skills are key determinants of bullying and its consequences. However, the findings are silent about the importance of skill investments. By modifying the stock of skills, they could reduce bullying occurrence.

To examine this issue, we re-estimated the bullying model (equation (7)), but this time including control variables proxying for investments. Formally, the model is augmented to include  $I_t$ , where  $I_t$  is a vector of skills' investment measures at time  $t$ . These measures include an index of parental control that measures whether the parents know where the youth is, who he is with and how long he will be there; an index of parental harmony that measures how much time the kid spends with their parents, whether the

child considers she is treated with affection by them, if she believes her parents treat each other well, and if her parents talk candidly and frequently with her as well as an index of parental abuse that measure whether the household is a violent setting. We also include two measures of school characteristics. The school quality measure is an index that aggregates measures of teacher responsiveness and learning conditions. The teacher responsiveness measure is based on the perceptions students have of their teacher, such as whether they think they can talk to their teacher openly and whether they would like to turn out to be like their teacher when they become adults. The learning conditions measure is based on the likelihood of students attending top institutions of higher education after graduating from that particular school, and whether students believe their school allows them to develop their talents and abilities. Finally, school environment is measured using information about robbery and criminal activity within or around the school and the presence of litter and garbage within the school or its surroundings.<sup>23</sup>

Our results are shown in the lower panel of Table 12. Column (1) of the table reproduces the original results just for comparison. Our findings show that the introduction of investment controls reduces the point estimate of the effect of non-cognitive skills on the likelihood of being bullied by only 25%. More importantly, however, the results show that less violence-prone parents and better schools significantly reduce the incidence of bullying. This suggests that the inertia caused by low skill levels, particularly non-cognitive traits, in previous periods on higher likelihoods of being involved in bullying can be reversed through the modification of tangible scenarios like the improvement of schools –including teachers– and diminishing aggressive behavior within households.

---

<sup>23</sup>School quality measures are coded in a reverse scale where high numbers mean less school quality.

## 7 Conclusions

This paper examines the determinants and consequences of bullying at age 15 on subsequent mental and physical health, risky behaviors, and schooling attainment. We base our analysis on the estimation of a structural model of endogenous bullying and counterfactual outcomes. That is, we explicitly model selection into the “bullying treatment”. In this framework, latent cognitive and non-cognitive skills are sources of unobserved heterogeneity. We estimate the model using longitudinal information from South Korea (KYP-JHSP).

Our findings show that non-cognitive skills significantly reduce the likelihood of being a victim of bullying. In particular, one standard deviation increase in non-cognitive skills reduces the probability of being bullied by 37%. In contrast to these results, we do not find significant effects of cognitive skills on bullying.

When analyzing the impact of bullying, we document that its victims have higher incidence of self-reported depression, sickness, mental health issues and stress, as well as less incidence of life satisfaction and college enrollment three years after the event. The magnitudes of the estimated treatment effects are by no means small, suggesting that bullying represents a heavy burden that needs to be carried for a long time.

Finally, consistent with the recent literature on skill formation, our findings suggest that investing in skills development is essential for any policy intended to fight bullying. We show that cognitive and non-cognitive skills are mediators that can palliate the effect of bullying on later outcomes. Consequently, investing in them might not only reduce the incidence of bullying, as there will be less people prone to be perpetrators and victims, but also significantly lessen its negative effects.

## References

- Angrist, J. and Imbens, G. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):447–475. [2](#)
- APA (2006). The american psychological association dictionary. *American Psychological Association Reference Books*. [6](#)
- Björkqvist, K., Ekman, K., and Lagerspetz, K. (1982). Bullies and victims: Their ego picture, ideal ego picture and normative ego picture. *Scandinavian Journal of Psychology*, 23(1):307–313. [2](#), [3](#)
- Borghans, L., Duckworth, A. L., Heckman, J. J., and Weel, B. T. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059. [6](#), [6.1](#)
- Borghans, L., Golsteyn, B. H. H., Heckman, J., and Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and Individual Differences*, 51(3):315–320. [3](#)
- Boulton, M. and Underwood, K. (1993). Bully/victim problems among middle school children. *European education*, 25(3):18–37. [2](#)
- Brown, S. and Taylor, K. (2008). Bullying, education and earnings: evidence from the national child development study. *Economics of Education Review*, 27(4):387–401. [2](#), [4](#)
- Carlson, B. E. (2000). Children Exposed to Intimate Partner Violence Research Findings and Implications for Intervention. *Trauma, Violence, & Abuse*, 1(4):321–342. [6.1.1](#)
- Carneiro, P., Hansen, K. T., and Heckman, J. (2003). Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review*, 44(2):361–422. [5.1](#), [13](#), [5.2](#)
- Carrell, S. E. and Hoekstra, M. L. (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone’s Kids. *American Economic Journal: Applied Economics*, 2(1):211–228. [19](#)
- Cawley, J., Heckman, J., and Vytlacil, E. (2001). Three observations on wages and measured cognitive ability. *Labour Economics*, 8(4):419–442. [7](#)
- Cunha, F., Heckman, J., Lochner, L., and Masterov, D. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1:697–812. [6.1](#)

- Cunha, F. and Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4):738–782. [2](#)
- Duckworth, A. and Seligman, M. (2005). Self-discipline outdoes iq in predicting academic performance of adolescents. *Psychological Science*, 16(12):939–944. [7](#), [3](#)
- Eriksen, T., Nielsen, H., and Simonsen, M. (2012). The effects of bullying in elementary school. *Economics Working Papers*, 16. [2](#), [4](#)
- Eriksen, T. L. M., Nielsen, H. S., and Simonsen, M. (2014). Bullying in elementary school. *Journal of Human Resources*, 49(4):839–71. [2](#), [19](#), [6.2.1](#)
- Faris, R. and Felmlee, D. (2011). Status struggles network centrality and gender segregation in same-and cross-gender aggression. *American Sociological Review*, 76(1):48–73. [1](#), [6.1.1](#)
- Greene, W. H. (2000). *Econometric analysis*. Prentice Hall, Upper Saddle River, New Jersey, 4th edition. [14](#)
- Hansen, K., Heckman, J., and Mullen, K. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, 121(1):39–98. [5.2](#)
- Heckman, J. and Masterov, D. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, 29(3):446–493. [6.1](#)
- Heckman, J. and Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *The American Economic Review*, 91(2):145–149. [7](#)
- Heckman, J., Stixrud, J., and Urzua, S. (2006a). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482. [1](#), [2](#), [7](#), [3](#), [4](#), [5](#), [5.1](#)
- Heckman, J. J., Urzua, S., and Vytlačil, E. (2006b). Understanding Instrumental Variables in Models with Essential Heterogeneity. *Review of Economics and Statistics*, 88(3):389–432. [2](#)
- Hodges, E. and Perry, D. (1996). Victims of peer abuse: An overview. *Relaiming Children and youth*, Spring:23–28. [3](#)
- Kang, C. (2007). Classroom peer effects and academic achievement: Quasi-randomization evidence from south korea. *Journal of Urban Economics*, 61(3):458–495. [6.1.1](#)
- Kotlarski, I. (1967). On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics*, 20(1):69–76. [5](#), [13](#)

- Le, A., Miller, P., Heath, A., and 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. [2](#)
- Murnane, R., Willett, J., and Levy, F. (1995). The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics*, 77(2):251–266. [7](#)
- Naylor, P., Cowie, H., Cossin, F., Bettencourt, R., and Lemme, F. (2010). Teachers’ and pupils’ definitions of bullying. *British Journal of Educational Psychology*, 76(3):553–576. [3](#)
- Olweus, D. (1997). Bully/victim problems in school: Facts and intervention. *European Journal of Psychology of Education*, 12(4):495–510. [1](#), [2](#), [3](#), [6.1.1](#)
- Ouellet-Morin, I., Odgers, C. L., Danese, A., Bowes, L., Shakoor, S., Papadopoulos, A. S., Caspi, A., Moffitt, T. E., and Arseneault, L. (2011). Blunted cortisol responses to stress signal social and behavioral problems among maltreated/bullied 12-year-old children. *BPS*, 70(11):1016–1023. [2](#)
- Perry, D., Kusel, S., and Perry, L. (1988). Victims of peer aggression. *Developmental psychology*, 24(6):807. [2](#)
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(Whole No. 609). [3](#)
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146. [5.1](#)
- Sarzosa, M. (2015). The Dynamic Consequences of Bullying on Skill Accumulation. [6.1.1](#)
- Sarzosa, M. and Urzua, S. (2015). Implementing factor models in stata: The heterofactor command. [12](#)
- Smith, P. and Brain, P. (2000). Bullying in schools: Lessons from two decades of research. *Aggressive behavior*, 26(1):1–9. [3](#)
- Smith, P., Talamelli, L., Cowie, H., Naylor, P., and Chauhan, P. (2004). Profiles of non-victims, escaped victims, continuing victims and new victims of school bullying. *British Journal of Educational Psychology*, 74(4):565–581. [2](#), [6.1.1](#), [6.2.1](#)
- Tough, P. (2012). *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character*. Houghton Mifflin Harcourt. [3](#)
- Urzua, S. (2008). Racial labor market gaps. *Journal of Human Resources*, 43(4):919. [7](#), [3](#)

- Wolfe, D. A., Crooks, C. V., Lee, V., McIntyre-Smith, A., and Jaffe, P. G. (2003). The Effects of Children's Exposure to Domestic Violence: A Meta-Analysis and Critique. *Clinical Child and Family Psychology Review*, 6(3):171–187. [6.1.1](#)
- Wolke, D., Woods, S., Stanford, K., and Schulz, H. (2001). Bullying and victimization of primary school children in england and germany: prevalence and school factors. *British Journal of Psychology*, 92(4):673–696. [6.1.1](#)

# Tables

Table 1: Descriptive Statistics

Total sample size	3,449		
Number of Females	1,724	Fathers Education:	
Proportion of urban households	78.55%	High-school	42.94%
Prop. of single-headed households	6%	4yr Coll. or above	36.56%
Median monthly income per-capita	1 mill won	Mothers Education:	
Prop. of Youths in College by 19	56.65%	High-school	56.31%
Incidence of smoking by 19*	19.08%	4yr Coll. or above	19.51%
Prop. of Single-child households	8.6%		

\*Incidence calculated as the proportion of people who has smoked at least once in the last year

Table 2: Descriptive Statistics: Non-Cognitive Measures

	Locus of Control		Irresponsibility		Self-esteem	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
All	10.679	2.142	8.288	2.403	-4.051	4.455
Males	10.835	2.182	8.310	2.397	-3.848	4.445
Females	10.524	2.091	8.267	2.409	-4.252	4.455
Attending College*	11.114	1.949	8.004	2.266	-2.913	4.103
Not Attending College*	11.166	2.007	8.124	2.347	-3.142	4.537

\* Sample limited to wave 6

Table 3: Descriptive Statistics: Test Scores

	Math and Science		Language and Social Studies		Class grade in last semester	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
All	0.115	1.043	-0.002	1.066	-0.137	1.074
Males	0.255	1.044	-0.141	1.081	-0.192	1.067
Females	-0.024	1.024	0.008	1.050	-0.081	1.079
Attending College*	0.236	1.007	0.101	1.011	0.027	1.015
Not Attending College*	-0.043	1.068	-0.138	1.119	-0.351	1.110

\* Sample limited to wave 6

Table 4: Descriptive Statistics: Bullying and Outcome Variables

Outcome	Mean	Std. Dev.	Min	Max
<i>Age 14</i>				
Bullied	0.2250	0.4176	0	1
<i>Age 15</i>				
Bullied	0.1119	0.3153	0	1
<i>Age 18</i>				
Depression	0	1	-2.084	3.093
Drinking	0.7213	0.4483	0	1
Smoking	0.1907	0.3929	0	1
Life Satisfaction	0.4772	0.4995	0	1
Sick	0.0735	0.2611	0	1
Mental Issues	0.0315	0.1749	0	1
Stressed Friends	0	1	-1.184	3.887
Stressed Parents	0	1	-1.858	2.831
Stressed School	0	1	-2.190	2.243
Stressed Poverty	0	1	-1.593	2.526
Stressed Total	0	1	-2.396	3.633
<i>Age 19</i>				
In College	0.690	0.463	0	1

Table 5: Reduced Form Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Depressed		Smoke		Drink		Sick	
Bullied	0.216*** (0.042)	0.134*** (0.042)	0.019 (0.014)	0.002 (0.014)	0.019 (0.021)	-0.002 (0.021)	0.042*** (0.012)	0.037*** (0.012)
Obs.	2,675	2,552	3,241	3,097	3,241	3,097	2,814	2,683
Observables	Y	Y	Y	Y	Y	Y	Y	Y
Test Scores	N	Y	N	Y	N	Y	N	Y
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Mental Health Prob		Life Satisfaction		In College		Stress: Friends	
Bullied	0.019** (0.009)	0.012 (0.009)	-0.062*** (0.021)	-0.041* (0.021)	-0.059*** (0.021)	-0.048** (0.022)	0.186*** (0.046)	0.144*** (0.047)
Obs.	2,814	2,683	3,241	3,097	2,681	2,558	2,806	2,676
Observables	Y	Y	Y	Y	Y	Y	Y	Y
Test Scores	N	Y	N	Y	N	Y	N	Y
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
	Stress: Parents		Stress: School		Stress: Poverty		Stress: Total	
Bullied	0.105** (0.045)	0.083* (0.046)	0.147*** (0.045)	0.141*** (0.045)	0.194*** (0.045)	0.145*** (0.046)	0.228*** (0.045)	0.183*** (0.046)
Obs.	2,806	2,676	2,806	2,676	2,806	2,676	2,806	2,676
Observables	Y	Y	Y	Y	Y	Y	Y	Y
Test Scores	N	Y	N	Y	N	Y	N	Y

Note: Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The observables controls used are: month of birth, gender, number of older siblings, number of younger siblings, natural log of household income per capita, urban, whether the kid lives with both parents, whether the kid lives only with her mother, father education: two-year college, father education: four-year college and, father education: graduate school. The tests scores used are: locus of control, irresponsibility, self-esteem, math and science, language and social studies and a class score.

Table 6: Identification of Skills at age 14

VARIABLES	Locus Control		Irresponsibility		Self-Esteem		Language - SocStudies		Math - Sci.		ClassScore
	Bullied	D = 1	D = 0	D = 1	D = 0	D = 1	D = 0	D = 1	D = 0	D = 1	
Age in Months	0.014* (0.008)	-0.014 (0.012)	-0.013** (0.006)	0.033*** (0.011)	0.016*** (0.006)	-0.006 (0.011)	-0.017*** (0.006)	-0.012 (0.008)	-0.013** (0.005)	-0.015*** (0.005)	-0.018*** (0.004)
Male	0.267*** (0.055)	0.222*** (0.084)	0.128*** (0.043)	-0.047 (0.082)	-0.050 (0.042)	0.333*** (0.083)	0.162*** (0.042)	-0.066 (0.062)	-0.001 (0.036)	0.295*** (0.034)	-0.083*** (0.029)
Oldersiblings	-0.043 (0.050)	0.047 (0.073)	0.026 (0.039)	-0.002 (0.071)	-0.021 (0.039)	0.023 (0.072)	-0.011 (0.039)	-0.063 (0.053)	-0.008 (0.033)	0.041 (0.031)	0.031 (0.027)
Youngsiblings	-0.080 (0.053)	-0.019 (0.078)	0.053 (0.040)	-0.061 (0.076)	-0.078* (0.040)	0.137* (0.077)	-0.006 (0.040)	0.104* (0.057)	0.108*** (0.034)	0.097*** (0.032)	0.102*** (0.027)
Lnmonthincpc	-0.081 (0.050)	0.080 (0.073)	0.127*** (0.039)	-0.180** (0.071)	-0.092** (0.039)	0.052 (0.072)	0.016 (0.039)	0.119** (0.054)	0.200*** (0.033)	0.188*** (0.031)	0.161*** (0.027)
Urban	0.022 (0.079)	0.030 (0.120)	0.193*** (0.062)	-0.046 (0.118)	-0.083 (0.061)	-0.026 (0.119)	0.056 (0.061)	0.278*** (0.088)	0.052 (0.052)	0.056 (0.050)	-0.006 (0.041)
Both Parents	-0.181 (0.149)	0.414** (0.211)	0.304** (0.126)	-0.391* (0.207)	-0.294** (0.125)	0.483** (0.209)	0.269** (0.125)	0.211 (0.156)	0.499*** (0.104)	0.464*** (0.096)	0.372*** (0.080)
OnlyMother	-0.126 (0.199)	0.796*** (0.283)	0.332** (0.167)	-0.565** (0.278)	-0.207 (0.166)	1.033*** (0.280)	0.226 (0.165)	0.058 (0.207)	0.511*** (0.137)	0.436*** (0.126)	0.181* (0.103)
FatherEd2yCol	0.032 (0.106)	0.386** (0.161)	0.047 (0.083)	-0.131 (0.157)	-0.132 (0.083)	0.244 (0.159)	-0.059 (0.082)	0.298** (0.119)	0.080 (0.073)	0.201*** (0.069)	0.182*** (0.060)
FatherEd4yCol	-0.041 (0.066)	0.115 (0.100)	0.165*** (0.050)	-0.130 (0.098)	-0.142*** (0.050)	0.081 (0.099)	0.126** (0.049)	0.344*** (0.073)	0.320*** (0.042)	0.179*** (0.040)	0.258*** (0.034)
FatherEducGS	0.186 (0.115)	0.132 (0.163)	0.357*** (0.094)	-0.064 (0.160)	-0.439*** (0.093)	0.135 (0.161)	0.217** (0.092)	0.585*** (0.120)	0.508*** (0.079)	0.327*** (0.073)	0.370*** (0.059)
Non-Cogn. Sk.	-0.076 (0.161)	1.331*** (0.176)	1.217*** (0.113)	-1.204*** (0.168)	-1.235*** (0.119)	1.123*** (0.167)	1 ■	1.172*** (0.150)	1.253*** (0.129)	1.327*** (0.132)	0.879*** (0.095)
Cognitive Skill	-0.052 (0.046)	-0.917** (0.360)	-1.077*** (0.214)	1.143*** (0.352)	0.737*** (0.213)	-1.228*** (0.355)	-0.254 (0.212)	0.631*** (0.269)	0.612*** (0.182)	0.583*** (0.169)	1 ■
Constant	-0.456* (0.258)	-0.917** (0.360)	-1.077*** (0.214)	1.143*** (0.352)	0.737*** (0.213)	-1.228*** (0.355)	-0.254 (0.212)	-0.981*** (0.269)	-1.451*** (0.182)	-1.480*** (0.169)	-1.022*** (0.153)

Observations

2,805

Note: Variable Oldersiblings corresponds to the number of older siblings the respondent has. Youngsiblings corresponds to the number of younger siblings the respondent has. Lnmonthincpc corresponds to the natural logarithm of the monthly income per capita. Both Parents takes the value of one if the respondent lives in a biparental household. OnlyMother takes the value of one if the respondent's father is absent from the household. FatherEd2yCol takes the value of one if the highest level of education attained by the respondent's father was a 2-year college degree. FatherEd4yCol takes the value of one if the highest level of education attained by the respondent's father was a 4-year college degree. FatherEducGS takes the value of one if the highest level of education attained by the respondent's father was graduate school. Columns headed as 1 collect the coefficients for those who were bullied at age 14. Columns headed as 0 collect the coefficients for those who were not bullied at age 14. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 7: Non-Cognitive and Cognitive Skills at age 14 and the Probability of Being Bullied at age 15

VARIABLES	(1)		(2)		(3)	
	Coeff.	Std. Err	Coeff.	Std. Err	Coeff.	Std. Err
Age in Months	0.0034	(0.010)	0.0021	(0.010)	0.0009	(0.010)
Male	0.3234***	(0.068)	0.2900***	(0.069)	0.3076***	(0.068)
Oldersiblings	-0.0452	(0.062)	-0.0397	(0.062)	-0.0457	(0.062)
Youngsiblings	-0.1153*	(0.067)	-0.1117*	(0.067)	-0.1159*	(0.067)
Lnmonthincpc	-0.0385	(0.061)	-0.0357	(0.061)	-0.0195	(0.061)
Urban	0.0469	(0.100)	0.0722	(0.100)	0.0789	(0.103)
Lives BothParents	-0.1580	(0.180)	-0.1654	(0.181)	-0.1503	(0.180)
Lives OnlyMother	0.0807	(0.236)	0.1298	(0.234)	0.1493	(0.234)
Fathereduc2yColl	0.0724	(0.129)	0.0786	(0.129)	0.0736	(0.129)
Fathereduc4yColl	-0.1046	(0.083)	-0.1100	(0.082)	-0.1218	(0.083)
FathereducGS	0.1479	(0.140)	0.1288	(0.140)	0.1303	(0.140)
% Peer Bullies			0.8107**	(0.340)		
% Peer Violent Fam					-3.8408**	(1.711)
% Peer Violent Fam <sup>2</sup>					4.8220**	(2.147)
Non-Cogn. Skills	-0.2810**	(0.132)	-0.2986**	(0.133)	-0.3004**	(0.133)
Cognitive Skills	0.0759	(0.058)	0.0784	(0.059)	0.0804	(0.059)
Constant	-1.0823***	(0.317)	-1.2904***	(0.328)	-0.4489	(0.443)
Observations	2,690		2,690		2,690	

*Note:* This Table presents the estimated coefficients of different specifications of the treatment equation  $D = \mathbb{1}[\mathbf{X}_D\beta^{Y_D} + \alpha^{Y_D,A}\theta^A + \alpha^{Y_D,B}\theta^B + e^D > 0]$ . Variable Oldersiblings corresponds to the number of older siblings the respondent has. Youngsiblings corresponds to the number of younger siblings the respondent has. Lnmonthincpc corresponds to the natural logarithm of the monthly income per capita. Variable Lives BothParents takes the value of one if the respondent lives in a biparental household. Variable Lives OnlyMother takes the value of one if the respondent's father is absent from the household. Fathereduc2yColl takes the value of one if the highest level of education attained by the respondent's father was a 2-year college degree. Fathereduc4yColl takes the value of one if the highest level of education attained by the respondent's father was a 4-year college degree. FathereducGS takes the value of one if the highest level of education attained by the respondent's father was graduate school. Variable % Peer Bullies corresponds to the (leave-one-out) proportion of peers that report being bullies in the respondent's class. The variable % Peer Violent Fam contains the (leave-one-out) proportion of peers in the respondent's classroom that come from a violent family, where a violent family is defined in footnote 19. % Peer Violent Fam<sup>2</sup> is the square of % Peer Violent Fam. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 8: Risky Behaviors and Mental and Physical Health Indicators: Outcome Equations (age 18) by Bullying Status  $D$  (age 15)

	(1) Depression		(2) Drink		(3) Smoke		(4) Life Satisfaction		(5) Sick		(6) Mental Health	
	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$
Age in Months	-0.000 (0.005)	0.009 (0.016)	-0.002 (0.003)	-0.000 (0.008)	0.003 (0.002)	-0.008 (0.006)	0.001 (0.003)	-0.001 (0.008)	0.000 (0.001)	0.001 (0.006)	0.001 (0.001)	0.007 (0.005)
Male	-0.208*** (0.039)	-0.394*** (0.111)	0.121*** (0.021)	0.274*** (0.060)	0.127*** (0.014)	0.143*** (0.044)	0.054*** (0.020)	0.108* (0.059)	-0.021** (0.011)	-0.035 (0.041)	0.007 (0.008)	0.023 (0.035)
Oldersiblings	-0.016 (0.037)	-0.050 (0.094)	0.031 (0.019)	-0.027 (0.050)	0.000 (0.013)	0.024 (0.036)	0.016 (0.019)	0.051 (0.049)	-0.004 (0.010)	-0.022 (0.034)	0.003 (0.008)	0.008 (0.029)
Youngsiblings	0.006 (0.038)	-0.072 (0.107)	-0.012 (0.020)	0.038 (0.058)	-0.021 (0.013)	0.033 (0.042)	0.007 (0.019)	0.028 (0.056)	-0.006 (0.010)	0.053 (0.040)	0.003 (0.008)	-0.028 (0.033)
Lnmonthincpc	-0.009 (0.037)	-0.094 (0.094)	-0.006 (0.019)	0.012 (0.050)	0.000 (0.013)	-0.040 (0.036)	0.028 (0.019)	0.087* (0.048)	-0.013 (0.010)	0.020 (0.034)	0.001 (0.008)	-0.024 (0.029)
Urban	-0.001 (0.057)	0.060 (0.162)	-0.036 (0.030)	0.004 (0.085)	-0.005 (0.020)	0.060 (0.063)	-0.046 (0.030)	-0.062 (0.083)	-0.013 (0.015)	-0.026 (0.061)	-0.016 (0.012)	-0.082 (0.051)
Both Parents	-0.337*** (0.118)	-0.400 (0.277)	-0.009 (0.060)	-0.054 (0.145)	0.021 (0.040)	0.000 (0.106)	0.214*** (0.060)	0.049 (0.143)	-0.025 (0.032)	-0.034 (0.098)	0.005 (0.024)	0.020 (0.083)
Only Mother	-0.081 (0.157)	-0.788** (0.355)	-0.038 (0.081)	0.044 (0.188)	0.094* (0.054)	0.046 (0.137)	0.196*** (0.080)	0.102 (0.184)	0.061 (0.042)	-0.041 (0.130)	0.045 (0.033)	-0.052 (0.110)
Fathereduc2yColl	-0.064 (0.076)	0.059 (0.192)	-0.030 (0.041)	-0.017 (0.107)	-0.014 (0.027)	0.162*** (0.078)	0.124*** (0.040)	-0.040 (0.106)	-0.015 (0.021)	-0.037 (0.072)	-0.004 (0.016)	-0.073 (0.061)
Fathereduc4yColl	-0.051 (0.047)	-0.051 (0.138)	-0.085*** (0.024)	-0.046 (0.073)	-0.040** (0.016)	0.079 (0.054)	0.014 (0.024)	0.018 (0.072)	0.012 (0.013)	0.023 (0.051)	-0.000 (0.010)	0.042 (0.043)
FathereducGS	-0.119 (0.087)	0.175 (0.223)	-0.047 (0.046)	-0.003 (0.116)	-0.045 (0.030)	0.097 (0.085)	0.166*** (0.045)	0.071 (0.114)	0.010 (0.023)	-0.083 (0.080)	0.000 (0.018)	0.083 (0.068)
Non-Cogn Skills	-0.425*** (0.070)	-0.729*** (0.192)	-0.108*** (0.038)	-0.219* (0.118)	-0.093*** (0.024)	-0.159* (0.084)	0.235*** (0.035)	0.477*** (0.098)	-0.043** (0.017)	-0.090 (0.068)	-0.022* (0.013)	-0.099* (0.059)
Cognitive Skills	0.108*** (0.033)	0.107 (0.089)	-0.005 (0.018)	-0.042 (0.052)	-0.034*** (0.012)	-0.138*** (0.037)	0.007 (0.017)	0.046 (0.044)	-0.001 (0.008)	0.011 (0.031)	0.008 (0.006)	0.016 (0.026)
Constant	0.541*** (0.196)	1.033** (0.471)	0.541*** (0.101)	0.352 (0.259)	0.053 (0.067)	0.193 (0.189)	0.159 (0.100)	0.047 (0.254)	0.164*** (0.052)	0.079 (0.176)	0.026 (0.040)	0.151 (0.149)
Observations	2,395	2,690	2,690	2,690	2,690	2,690	2,690	2,690	2,514	2,514	2,514	2,514

Note: This Table presents the estimated coefficients of the outcome equations  $Y_1 = \mathbf{X}_Y \beta^{Y_1} + \alpha^{Y_1, A} \theta^A + e^{Y_1}$  if  $D = 1$  and  $Y_0 = \mathbf{X}_Y \beta^{Y_0} + \alpha^{Y_0, A} \theta^A + e^{Y_0}$  if  $D = 0$ . "Depression" corresponds to a standardized index of depression symptoms. "Drink" takes the value of 1 if the respondent drank an alcoholic beverage at least once during the last year. "Smoke" takes the value of 1 if the respondent smoked a cigarette at least once during the last year. "Life Satisfaction" takes the value of 1 if the respondent reports being happy with the way she is leading her life. "Sick" takes the value of 1 if the respondent reports having felt physically ill during the last year. "Mental Health" takes the value of 1 if the respondent has been diagnosed with psychological or mental problems. Oldersiblings corresponds to the number of older siblings the respondent has. Youngsiblings corresponds to the number of younger siblings the respondent has. Lnmonthincpc corresponds to the natural logarithm of the monthly income per capita. Both Parents takes the value of one if the respondent lives in a biparental household. OnlyMother takes the value of one if the respondent's father is absent from the household. Fathereduc2yColl takes the value of one if the highest level of education attained by the respondent's father was a 2-year college degree. Fathereduc4yColl takes the value of one if the highest level of education attained by the respondent's father was a 4-year college degree. FathereducGS takes the value of one if the highest level of education attained by the respondent's father was a graduate school. Columns headed as 1 collect the coefficients for those who were bullied at age 15. Columns headed as 0 collect the coefficients for those who were not bullied at age 15. Standard errors in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 9: College Attendance and Stress Levels: Outcome Equations by Bullying Status  $D$  (age 15)

<i>Bullied</i>	(1) inCollege <sup>†</sup>		(2) Stress: Friends		(3) Stress: Parent		(4) Stress: School		(5) Stress: Total		(6) Stress: Poverty	
	$D = 1$	$D = 0$	$D = 1$	$D = 0$	$D = 1$	$D = 0$	$D = 1$	$D = 0$	$D = 1$	$D = 0$	$D = 1$	$D = 0$
	Age in Months	0.003 (0.003)	0.001 (0.009)	-0.003 (0.006)	-0.014 (0.020)	0.006 (0.006)	-0.026 (0.019)	0.003 (0.006)	-0.000 (0.018)	0.006 (0.006)	-0.005 (0.018)	0.006 (0.006)
male	-0.132*** (0.062)	-0.064 (0.062)	0.101** (0.042)	0.141 (0.141)	0.090** (0.042)	0.081 (0.132)	-0.192*** (0.041)	-0.048 (0.124)	-0.155*** (0.042)	-0.068 (0.129)	-0.256*** (0.042)	-0.127 (0.125)
Oldersiblings	-0.000 (0.018)	-0.049 (0.051)	-0.015 (0.039)	0.216* (0.117)	-0.033 (0.039)	0.066 (0.109)	0.056 (0.039)	0.129 (0.103)	-0.021 (0.039)	0.153 (0.107)	-0.033 (0.040)	0.069 (0.104)
Youngsiblings	0.003 (0.019)	-0.009 (0.059)	0.035 (0.040)	0.062 (0.134)	0.041 (0.040)	0.099 (0.126)	0.046 (0.039)	0.044 (0.120)	0.013 (0.040)	0.048 (0.123)	-0.024 (0.040)	-0.005 (0.119)
Lnmonthincpc	0.020 (0.018)	0.028 (0.055)	-0.005 (0.039)	0.023 (0.114)	0.107*** (0.039)	0.005 (0.107)	0.145*** (0.038)	0.067 (0.100)	0.036 (0.039)	0.035 (0.103)	-0.107*** (0.039)	-0.007 (0.101)
Urban	-0.022 (0.029)	0.064 (0.093)	-0.095 (0.060)	0.056 (0.205)	0.025 (0.061)	0.211 (0.192)	0.020 (0.059)	0.186 (0.180)	0.023 (0.060)	0.223 (0.186)	0.091 (0.061)	0.337* (0.182)
Both Parents	0.197*** (0.059)	0.163 (0.167)	-0.253*** (0.125)	-0.917*** (0.334)	-0.167 (0.126)	-0.445 (0.310)	0.202 (0.124)	-0.134 (0.292)	-0.097 (0.125)	-0.608** (0.304)	-0.090 (0.127)	-0.193 (0.294)
Only Mother	0.076 (0.079)	0.099 (0.207)	-0.260 (0.167)	-0.860* (0.444)	-0.095 (0.168)	-0.866** (0.412)	0.194 (0.165)	-0.551 (0.388)	-0.088 (0.167)	-0.874** (0.403)	-0.072 (0.170)	-0.365 (0.391)
Fathereduc2yColl	0.075* (0.040)	0.060 (0.114)	-0.002 (0.081)	-0.328 (0.247)	0.053 (0.082)	0.131 (0.229)	0.113 (0.080)	0.034 (0.217)	0.080 (0.081)	0.025 (0.227)	0.010 (0.083)	-0.042 (0.218)
Fathereduc4yColl	0.002 (0.023)	0.091 (0.078)	-0.054 (0.049)	-0.099 (0.173)	0.058 (0.050)	0.103 (0.162)	0.105** (0.049)	-0.053 (0.153)	-0.017 (0.050)	-0.155 (0.159)	-0.119** (0.050)	-0.345** (0.153)
FathereducGS	-0.050 (0.043)	0.043 (0.125)	-0.036 (0.092)	-0.513* (0.270)	0.011 (0.093)	0.010 (0.253)	0.022 (0.091)	-0.280 (0.239)	-0.170* (0.092)	-0.383 (0.250)	-0.374*** (0.093)	-0.426* (0.239)
Non-Cogn Skills	0.006 (0.036)	0.173 (0.122)	-0.251*** (0.073)	-0.829*** (0.252)	-0.079 (0.076)	-0.134 (0.237)	0.134* (0.075)	-0.445** (0.226)	-0.243*** (0.073)	-0.843*** (0.218)	-0.342*** (0.071)	-0.603*** (0.223)
Cognitive Skills	0.090*** (0.017)	0.057 (0.051)	0.137*** (0.035)	0.230** (0.107)	0.248*** (0.036)	0.190* (0.105)	0.349*** (0.035)	0.484*** (0.102)	0.273*** (0.035)	0.431*** (0.102)	0.056 (0.034)	0.187** (0.095)
Constant	0.470*** (0.098)	0.353 (0.282)	0.264 (0.207)	0.853 (0.600)	-0.484** (0.209)	0.415 (0.559)	-0.895*** (0.204)	-0.320 (0.525)	-0.083 (0.207)	0.394 (0.546)	0.627*** (0.211)	0.034 (0.529)
Observations	2,370	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508

*Note:* This Table presents the estimated coefficients of the outcome equations  $Y_1 = \mathbf{X}_Y \beta Y_1 + \alpha Y_1 A \theta A + e^{Y_1}$  if  $D = 1$  and  $Y_0 = \mathbf{X}_Y \beta Y_0 + \alpha Y_0 A \theta A + \alpha Y_0 B \theta B + e^{Y_0}$  if  $D = 0$ . Variable inCollege takes the value of 1 if the respondent attends college by age 19. The Stress variables are standardized indexes that collect stress symptoms triggered by different sources, namely friends, parents, school and poverty. Stress:Total aggregates the four triggers of stress. Oldersiblings corresponds to the number of older siblings the respondent has. Youngsiblings corresponds to the number of younger siblings the respondent has. Lnmonthincpc corresponds to the natural logarithm of the monthly income per capita. Both Parents takes the value of one if the respondent lives in a biparental household. OnlyMother takes the value of one if the respondent's father is absent from the household. Fathereduc2yColl takes the value of one if the highest level of education attained by the respondent's father was a 2-year college degree. Fathereduc4yColl takes the value of one if the highest level of education attained by the respondent's father was a 4-year college degree. FathereducGS takes the value of one if the highest level of education attained by the respondent's father was graduate school. Columns headed as 1 collect the coefficients for those who were bullied at age 15. Columns headed as 0 collect the coefficients for those who were not bullied at age 15. Standard errors in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

<sup>†</sup> College attendance is measured at age 19.

Table 10: Assessing the Fit of the Model

	Depression		Smoking		Sick		Life Satisfaction	
	Data	Model	Data	Model	Data	Model	Data	Model
$E[Y_0   D = 0]$	0.0561	0.0434	0.1307	0.1306	0.0639	0.0635	0.5201	0.5231
$E[Y_1   D = 1]$	0.1404	0.1218	0.1689	0.1839	0.1187	0.1181	0.4764	0.4767

	College		Mental Health		Stress: Friends		Stress: Parents	
	Data	Model	Data	Model	Data	Model	Data	Model
$E[Y_0   D = 0]$	0.7008	0.6899	0.0362	0.0352	-0.0538	-0.0665	-0.0181	-0.0353
$E[Y_1   D = 1]$	0.6398	0.6276	0.0827	0.0810	0.2928	0.2623	0.1587	0.1368

	Stress: School		Stress: Poverty		Stress: Total	
	Data	Model	Data	Model	Data	Model
$E[Y_0   D = 0]$	-0.0039	-0.0240	-0.0058	-0.0135	-0.0244	-0.0449
$E[Y_1   D = 1]$	0.1306	0.0644	0.1019	0.0772	0.2313	0.1796

*Note:* The mean simulated outcomes (i.e., Model) were calculated using 40,000 observations generated from the estimated model. The Data columns contain the outcomes' mean at age 18 obtained from the KYP-JHSP. The variable Depression corresponds to a standardized index of depression symptoms. Smoking takes the value of 1 if the respondent smoked a cigarette at least once during the last year. Sick takes the value of 1 if the respondent reports having felt physically ill during the last year. Life Satisfied takes the value of 1 if the respondent reports being happy with the way she is leading her life. Variable College takes the value of 1 if the respondent attends college by age 19. Mental Health takes the value of 1 if the respondent has been diagnosed with psychological or mental problems. The Stress variables are standardized indexes that collect stress symptoms triggered by different sources, namely friends, parents, school and poverty. Stress:Total aggregates the four triggers of stress.

Table 11: Treatment Effects: Outcomes at Age 18 of Being Bullied at Age 15

	Depression	Smoking	Drinking	Sick	Mental Health	Life Satisfaction
ATE	0.0416 (0.0755)	0.0183 (0.0261)	0.0001 (0.0347)	0.0552** (0.0244)	0.0298 (0.0210)	-0.0040 (0.0355)
TTE	0.0520 (0.0590)	0.0174 (0.0237)	0.0230 (0.0307)	0.0498** (0.0213)	0.0410** (0.0179)	-0.0268 (0.0311)
inCollege <sup>†</sup>	Stress:Friends	Stress:Parent	Stress:School	Stress:Poverty	Stress:Total	
ATE	-0.0257 (0.0337)	0.1939** (0.0924)	0.1508** (0.0752)	0.0513 (0.0818)	0.0242 (0.0783)	0.1467* (0.0814)
TTE	-0.0396 (0.0307)	0.2735*** (0.0725)	0.1450** (0.0684)	0.1422** (0.0635)	0.0825 (0.0646)	0.2191*** (0.0642)

Note: Standard errors in parenthesis. This Table presents the estimated treatment parameters.

$$ATE = \iint \mathbb{E} [Y_1 - Y_0 | \theta^{NC}, \theta^C] f(\theta^{NC}) f(\theta^C) d\theta^{NC} d\theta^C$$

and

$$TT = \iint \mathbb{E} [Y_1 - Y_0 | \theta^{NC}, \theta^C, D = 1] f(\theta^{NC}) f(\theta^C) d\theta^{NC} d\theta^C$$

The variable *Depress* corresponds to a standardized index of depression symptoms. Smoking takes the value of 1 if the respondent smoked a cigarette at least once during the last year. Drinking takes the value of 1 if the respondent drank an alcoholic beverage at least once during the last year. Sick takes the value of 1 if the respondent reports having felt physically ill during the last year. Mental Health takes the value of 1 if the respondent reports being diagnosed with psychological or mental problems. Satisfied takes the value of 1 if the respondent reports being happy with the way she is leading her life. The variable *inCollege* takes the value of 1 if the respondent attends college by age 19. The Stress variables are standardized indexes that collect stress symptoms triggered by different sources, namely friends, parents, school and poverty. Stress:Total aggregates the four triggers of stress. <sup>†</sup> College attendance is measured at age 19.

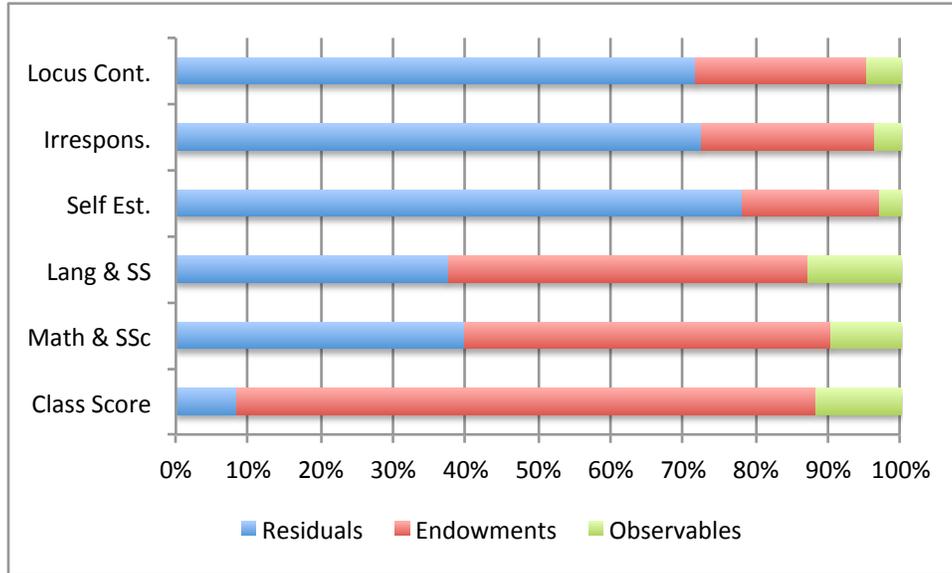
Table 12: The Model with Investment Controls

VARIABLES	Being Bullied			
	(1)		(2)	
	Coeff.	Std. Err.	Coeff.	Std. Err.
Non-Cognitive	-0.2986**	(0.133)	-0.2220	(0.137)
Cognitive	0.0784	(0.059)	0.0870	(0.069)
<i>Invest Cont at t</i>				
ParentControl			0.0299	(0.041)
ParentHarmony			-0.0281	(0.039)
ParentAbuse			0.1127**	(0.038)
SchQuality			0.0394***	(0.014)
SchEnviron			0.0185***	(0.007)
Observations	2,690		2,680	

*Note:* This Table presents the estimated coefficients of the treatment equation  $D = \mathbb{1} [\mathbf{X}_D \beta^{Y_D} + \alpha^{Y_D, A} \theta^A + \alpha^{Y_D, B} \theta^B + e^D > 0]$ . Column (1) is presented for reference and shows the results presented in Table 7. Column (2) add controls to the specification presented in Column (1). Namely ParentControl measures whether the parents know where the youth is, who he is with and how long he will be there, ParentHarmony measures how much time the respondent spends with their parents, whether she is treated with affection by them, if her parents treat each other well, and if her parents talk candidly and frequently to her, ParentAbuse measures whether the household is a violent setting, SchQuality measures teacher responsiveness and learning conditions (i.e., how likely are students to attend top institutions of higher education after graduating from that particular school, and whether students believe their school allows them to develop their talents and abilities), and SchEnviron is measured using information about robbery and criminal activity within or around the school and the presence of litter and garbage within the school or its surroundings. In both specifications we controlled for age in months, gender, rurality, the number of older and younger siblings the respondent has, the natural logarithm of the monthly income per capita, whether the respondent lives in a biparental household, whether the respondent's father is absent from the household., father's education, the (leave-one-out) proportion of peers that report being bullies in the respondent's class. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Figures

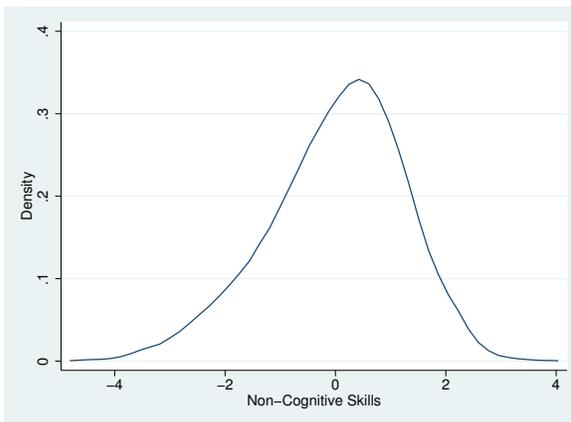
Figure 1: Variance Decomposition of Test Scores



*Note:* This Figure presents the proportion of the test scores variance that is explained by the observable controls and the endowments (latent cognitive and non-cognitive skills). The portion of the test score variance that remains unexplained is labeled Residuals.

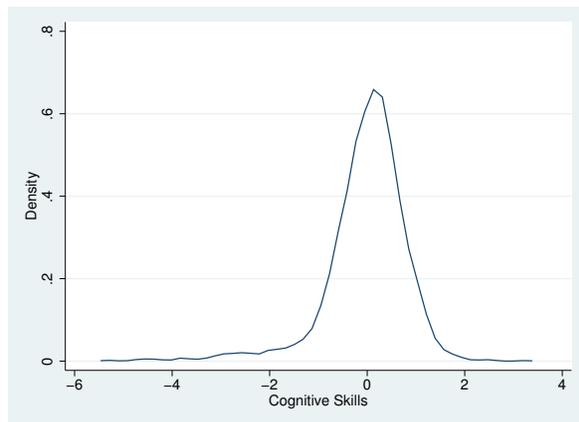
Figure 2: Distributions at  $t = 1$

(a) Non-Cognitive Skills Distribution at  $t = 1$



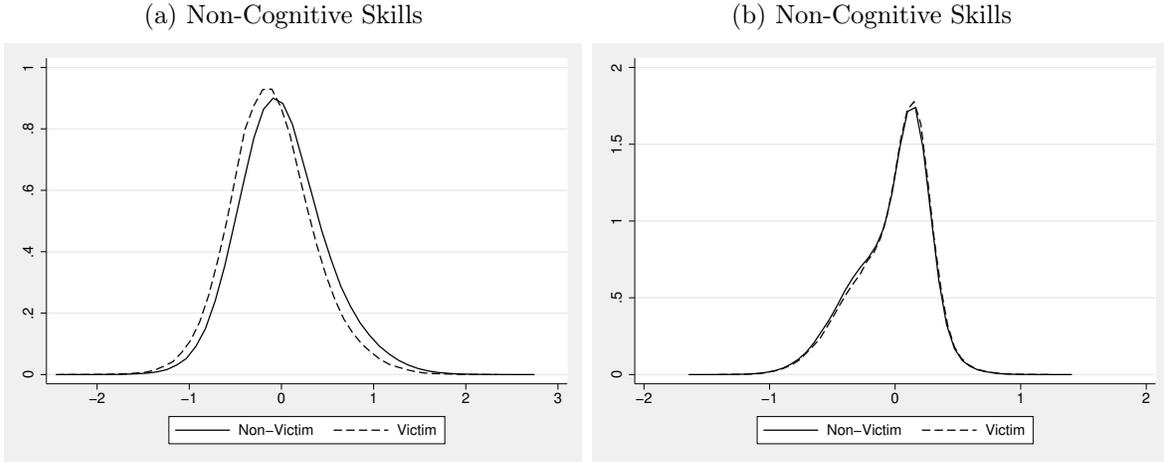
Mean	Std. Dev	Min	Max
-.0153126	.8338353	-5.359837	3.278021

(b) Cognitive Skills Distribution at Age 14



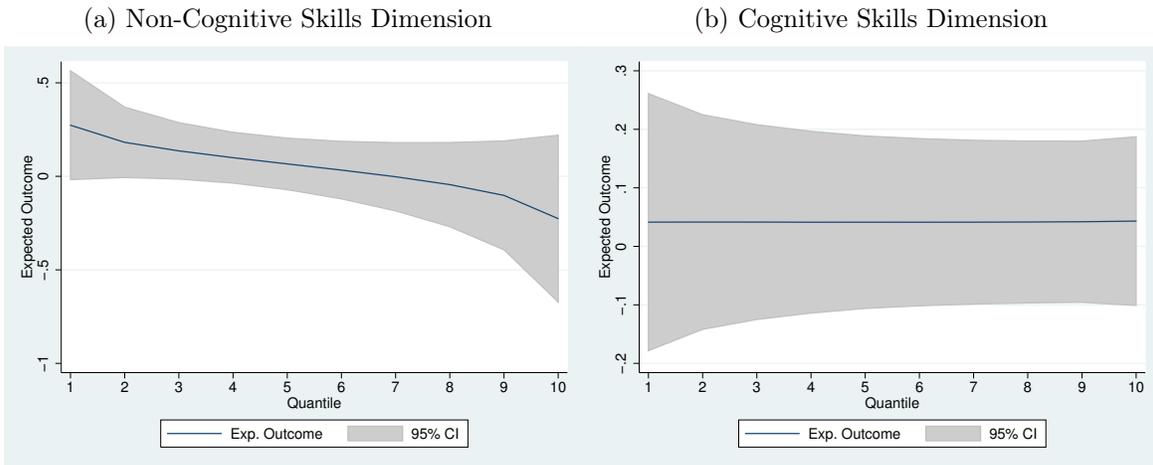
Mean	Std. Dev	Min	Max
.0111396	1.234135	-4.608246	3.834881

Figure 3: Skills Sorting into Being a Bullying Victim



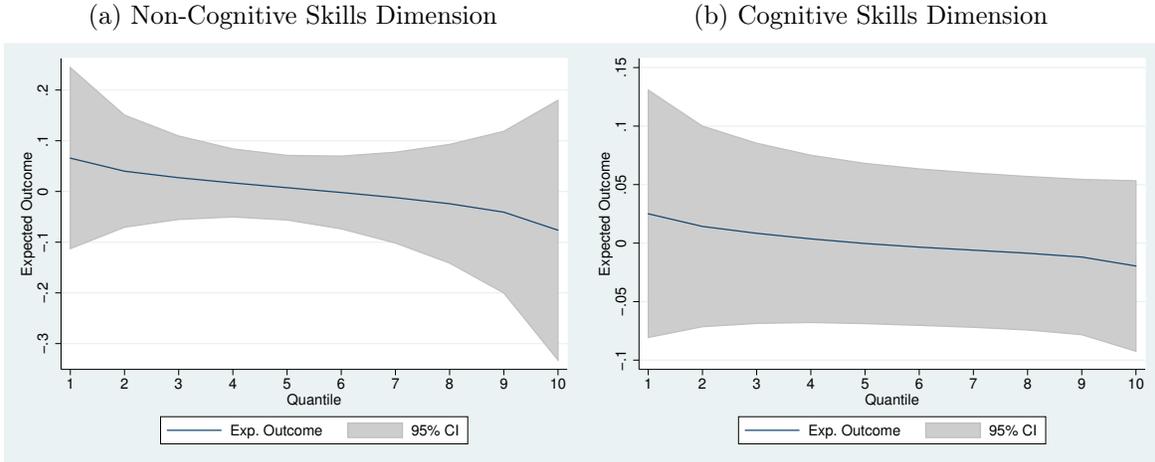
Note: Each panel in this Figure presents the distributions of unobserved abilities by victimization condition. The distributions are computed using simulated observations. The simulated data is generated using the estimates of the model.

Figure 4: ATE on Depression



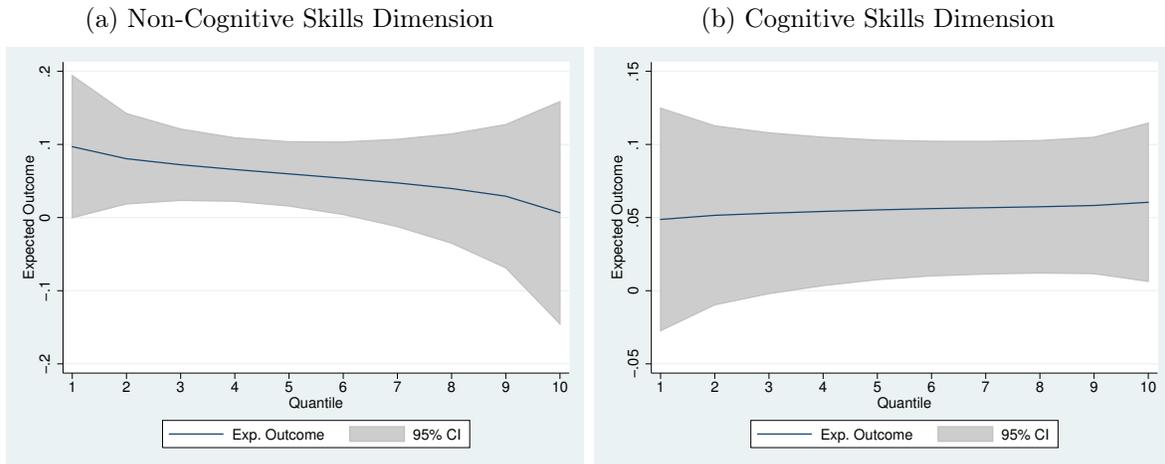
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. The depression variable is a standardized aggregated index of depression symptoms.

Figure 5: ATE on Drinking



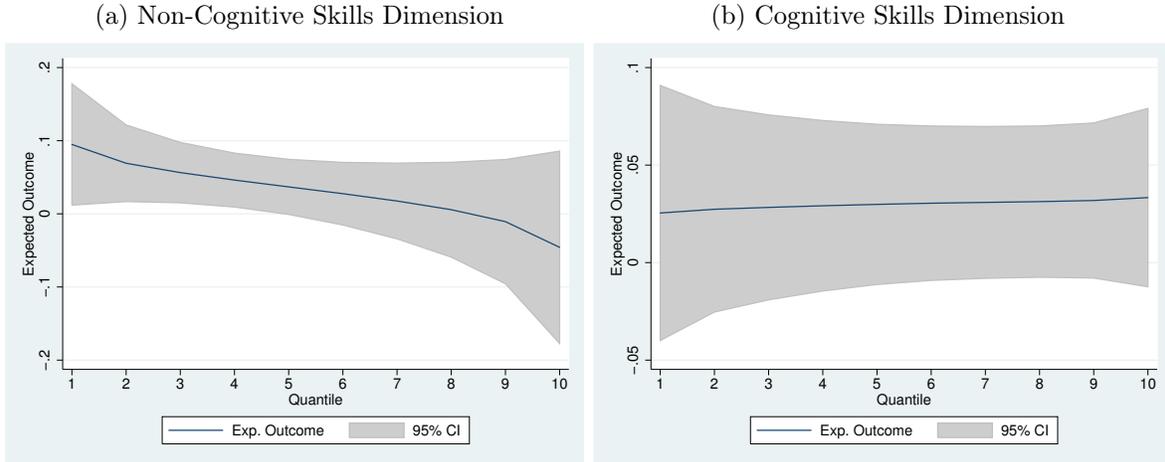
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. The variable Drinking takes the value of 1 if the respondent drank an alcoholic beverage at least once during the last year.

Figure 6: ATE on the Likelihood of Feeling Sick



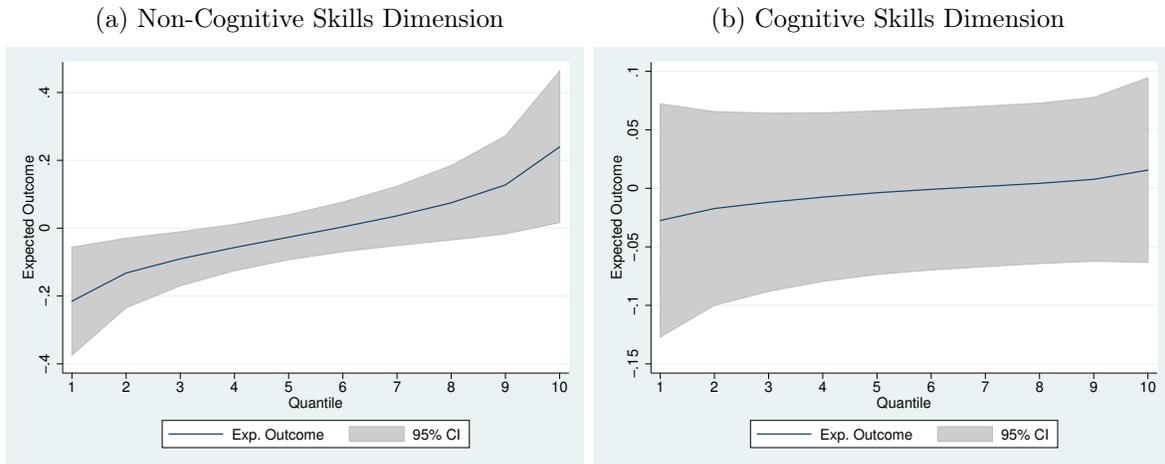
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. The variable Sick takes the value of 1 if the respondent reports having felt physically ill during the last year.

Figure 7: ATE on Mental Health Problems



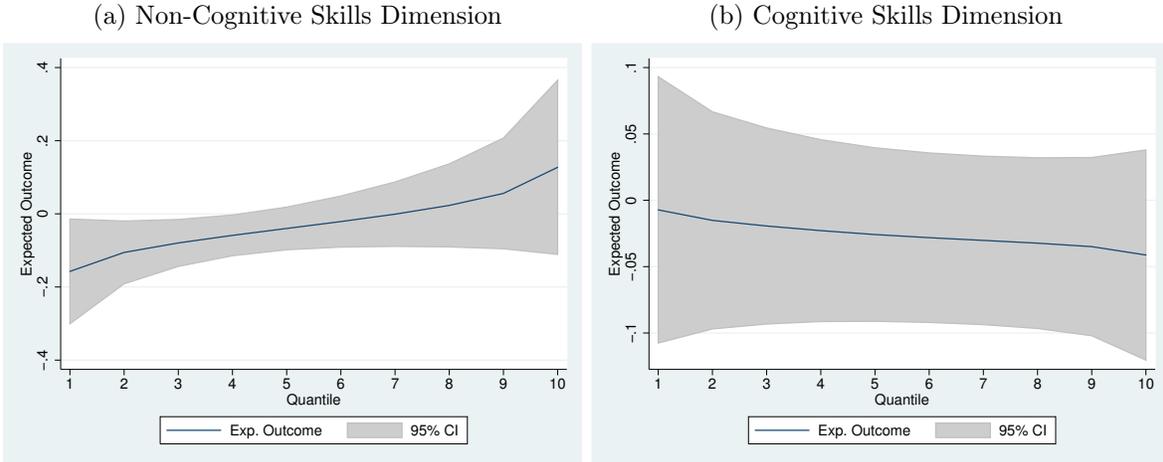
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The y-axes contain the deciles of each dimension of skills. The variable Mental Health Problems takes the value of 1 if the respondent has been diagnosed with psychological or mental problems.

Figure 8: ATE on Life Satisfaction



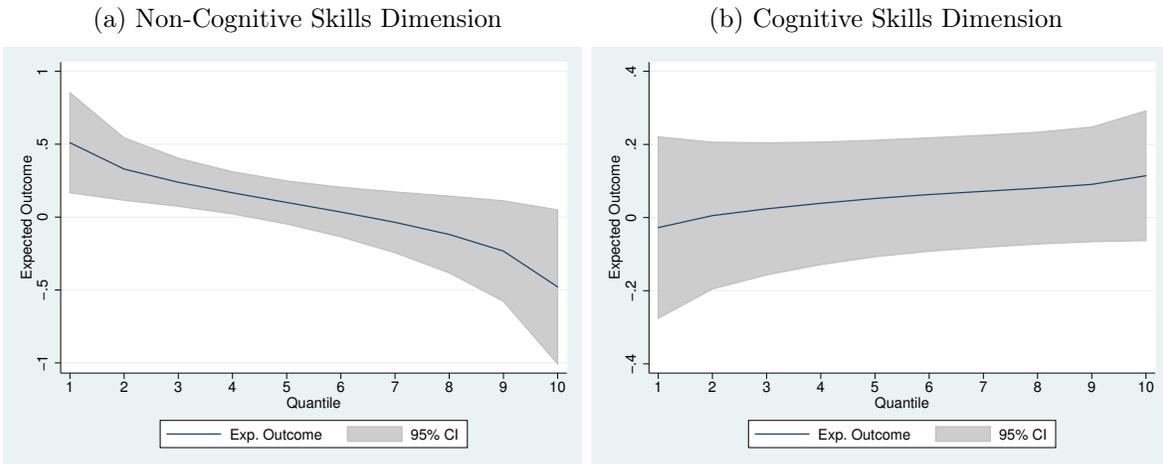
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The y-axes contain the deciles of each dimension of skills. Life Satisfaction takes the value of 1 if the respondent reports being happy with the way she is leading her life.

Figure 9: ATE on College Attendance



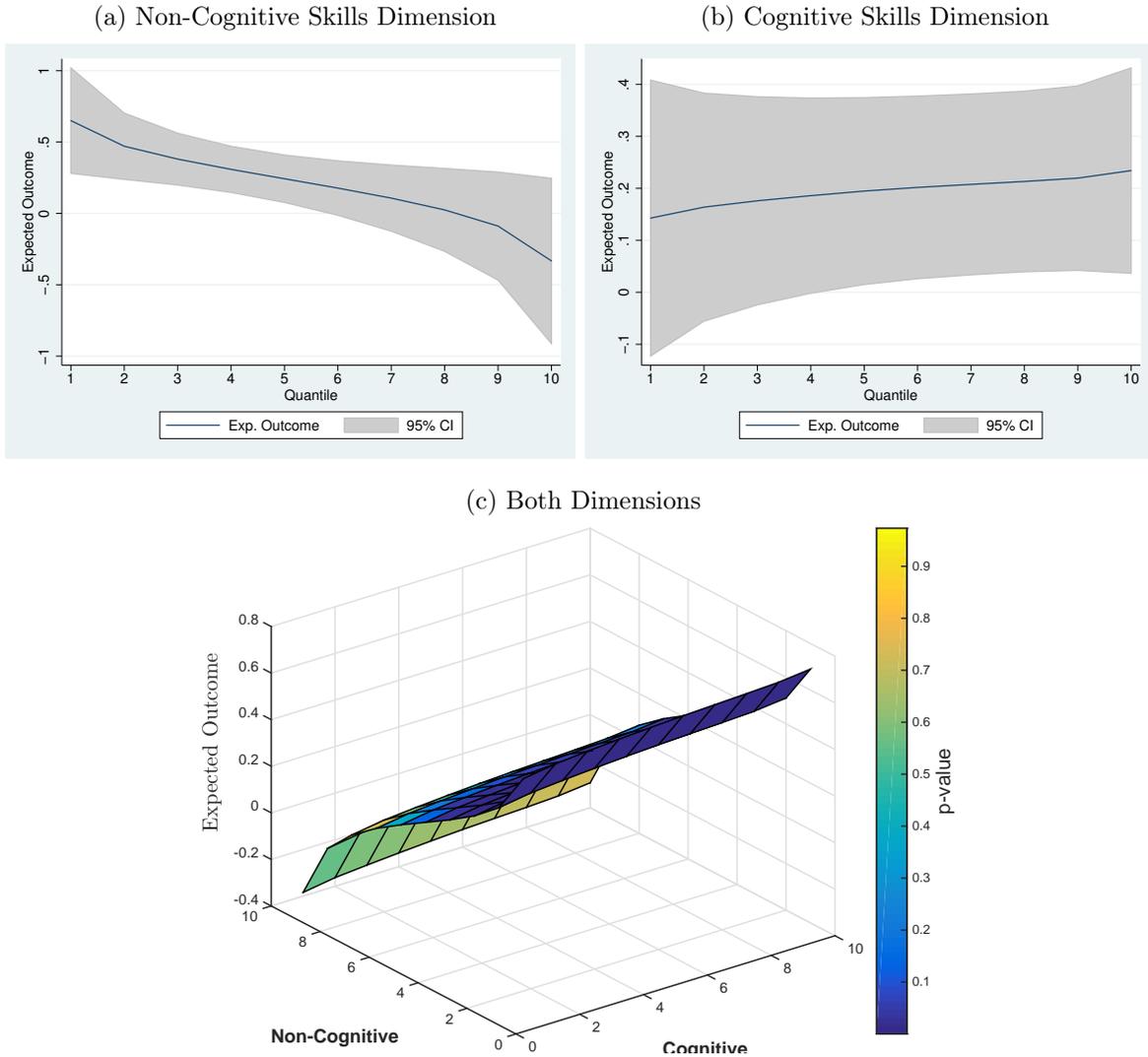
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The y-axes contain the deciles of each dimension of skills. College Attendance takes the value of 1 if the respondent attends college by age 19.

Figure 10: ATE on Stress: School



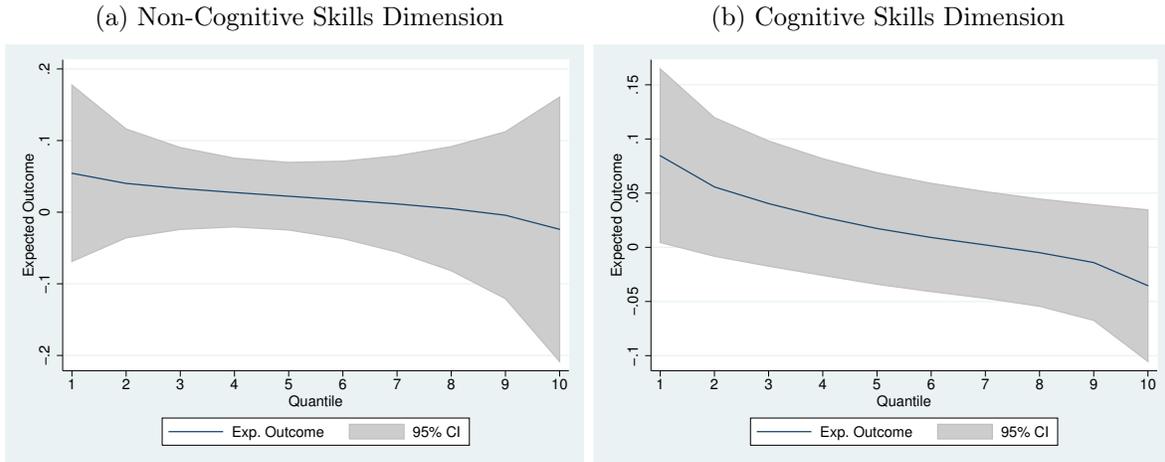
Note: Panels present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The y-axes contain the deciles of each dimension of skills. Stress: school is a variable that aggregates stress symptoms triggered by situations related with school.

Figure 11: ATE on Stress: Friends

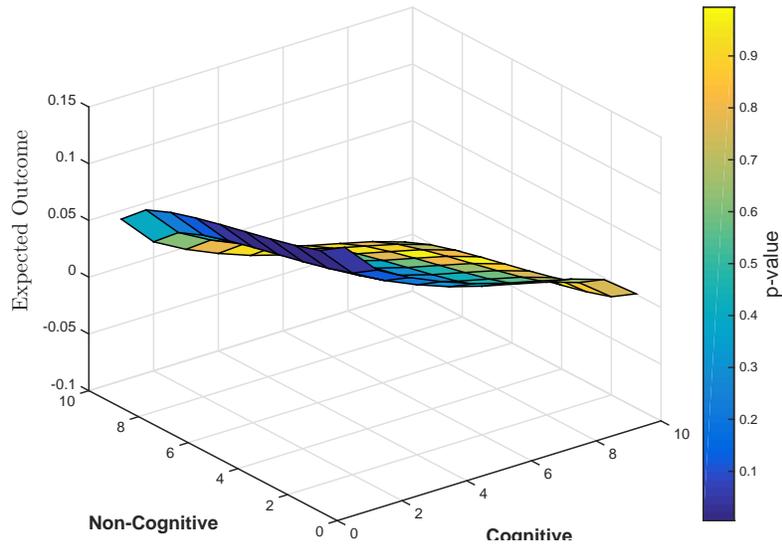


Note: Panels (a) and (b) present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. Panel (c) presents  $ATE(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C]$  in the  $z$ -axis and the deciles of cognitive and non-cognitive skills in the  $x$ -axis and  $y$ -axis. Stress: Friends is a variable that aggregates stress symptoms triggered by situations related with friends and social relations.

Figure 12: ATE on Smoking

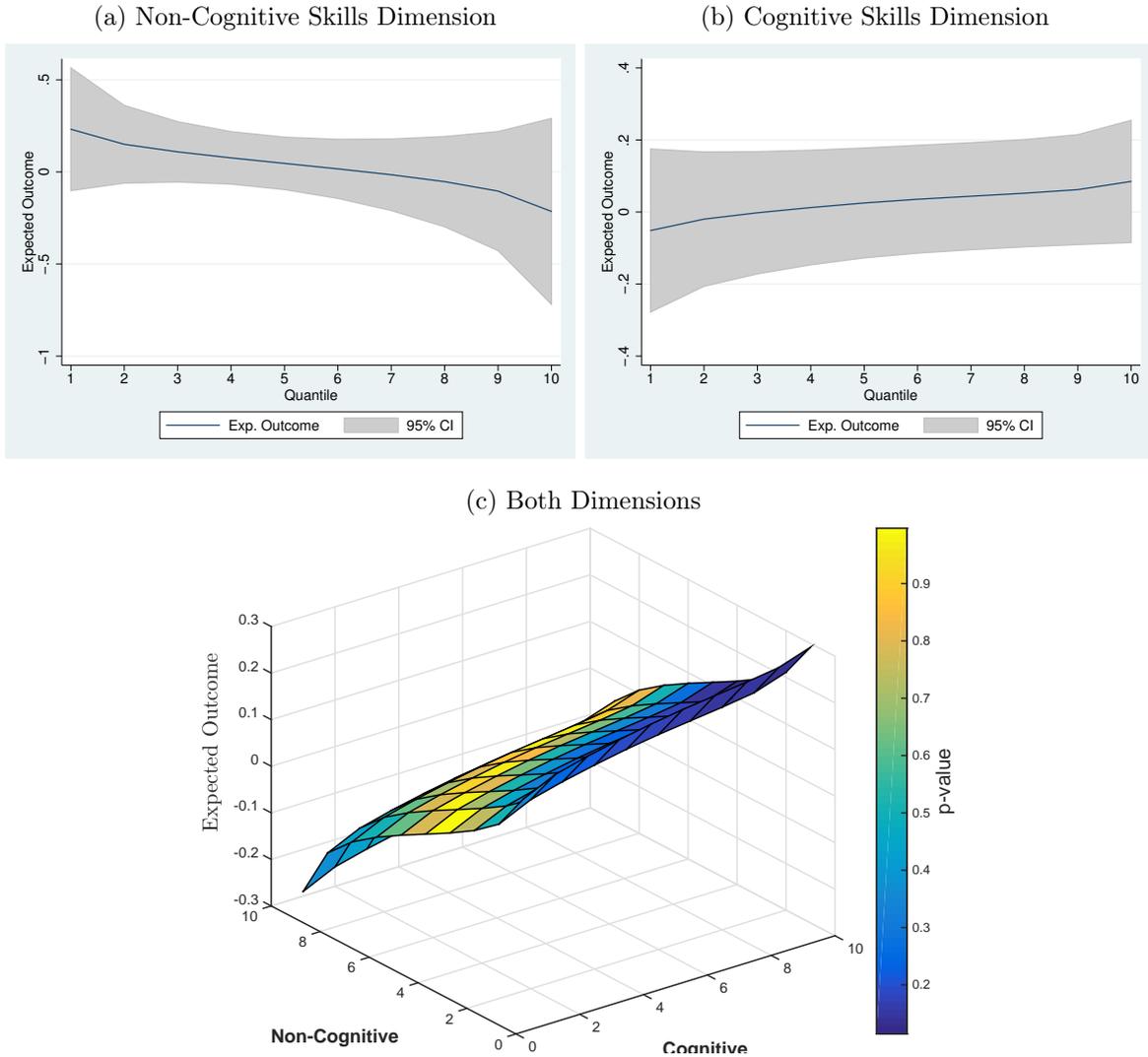


(c) Cognitive Skills Dimension



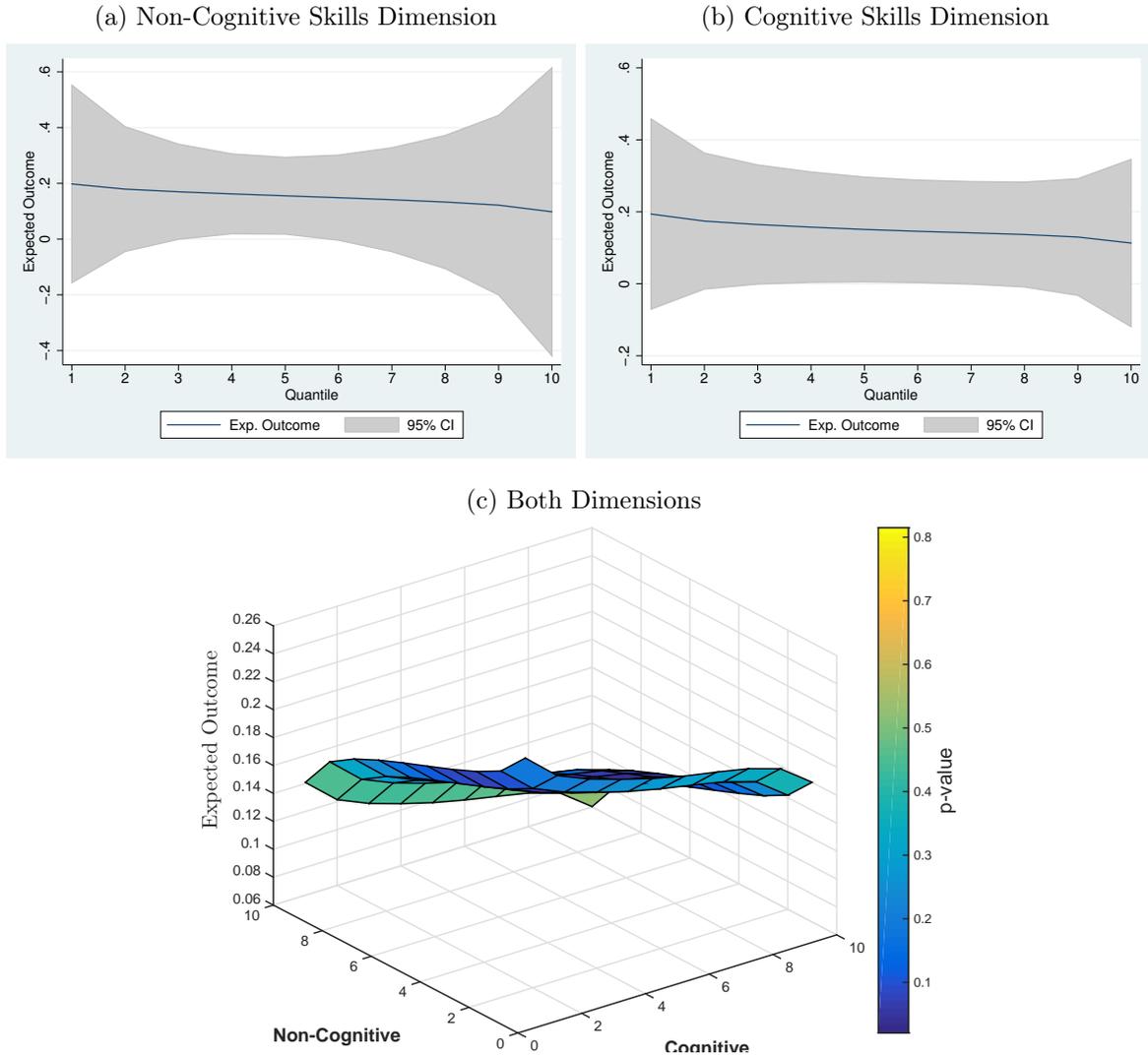
Note: Panels (a) and (b) present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The y-axes contain the deciles of each dimension of skills. Panel (c) presents  $ATE(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C]$  in the z-axis and the deciles of cognitive and non-cognitive skills in the x-axis and y-axis. Smoking takes the value of 1 if the respondent smoked a cigarette at least once during the last year.

Figure 13: ATE on Stress: Poverty



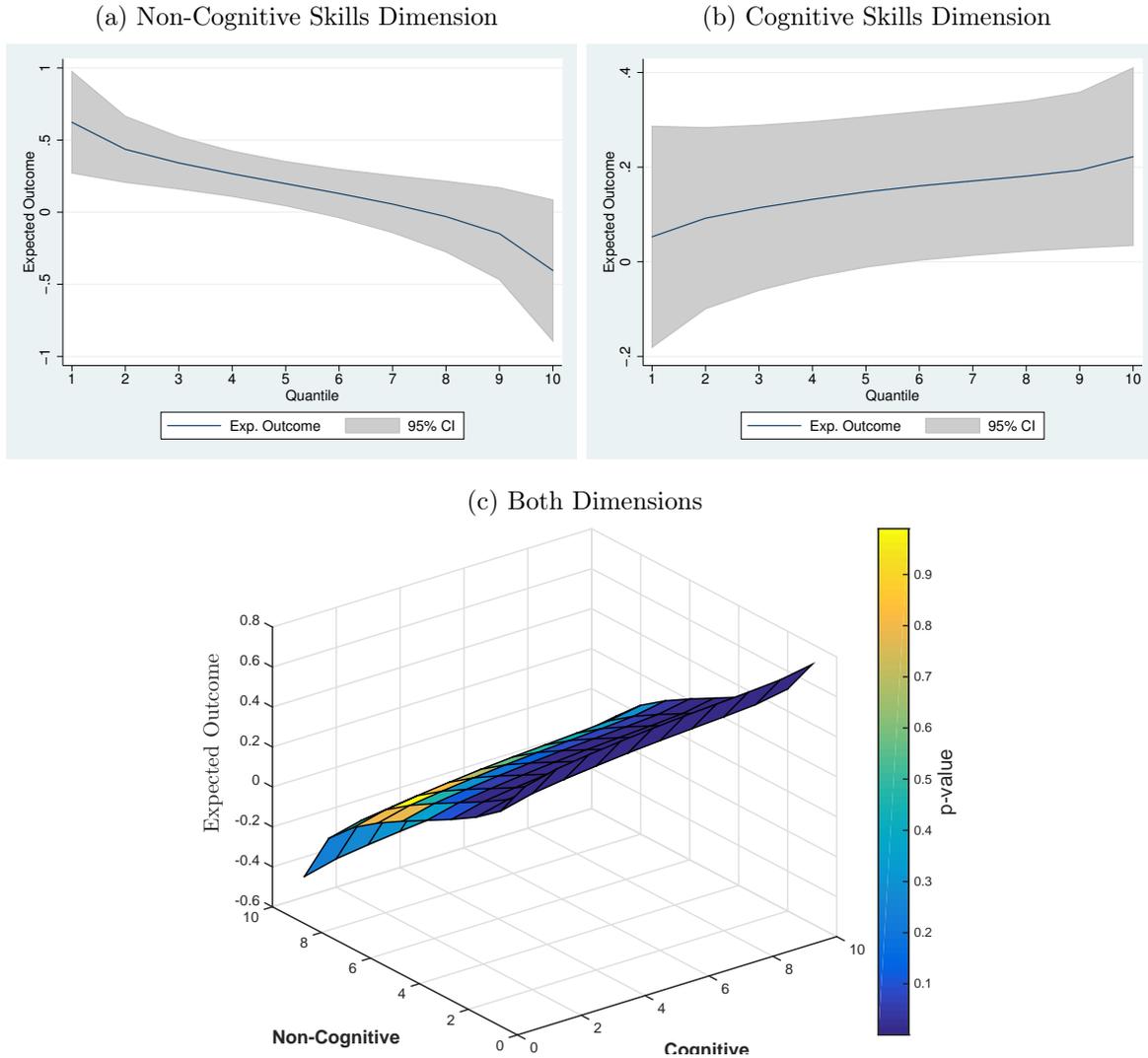
Note: Panels (a) and (b) present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. Panel (c) presents  $ATE(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C]$  in the  $z$ -axis and the deciles of cognitive and non-cognitive skills in the  $x$ -axis and  $y$ -axis. Stress: Poverty is a variable that aggregates stress symptoms triggered by situations related with economic difficulties.

Figure 14: ATE on Stress: Parents



Note: Panels (a) and (b) present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. Panel (c) presents  $ATE(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C]$  in the  $z$ -axis and the deciles of cognitive and non-cognitive skills in the  $x$ -axis and  $y$ -axis. Stress: Parents is a variable that aggregates stress symptoms triggered by the relation of the respondent with her parents.

Figure 15: ATE on Stress: Total



Note: Panels (a) and (b) present the  $ATE(\theta^S) = \mathbb{E}[Y_1 - Y_0 | \theta^S]$  for  $S = \{Non - Cognitive, Cognitive\}$ . The  $y$ -axes contain the deciles of each dimension of skills. Panel (c) presents  $ATE(\theta^{NC}, \theta^C) = \mathbb{E}[Y_1 - Y_0 | \theta^{NC}, \theta^C]$  in the  $z$ -axis and the deciles of cognitive and non-cognitive skills in the  $x$ -axis and  $y$ -axis. Stress: Total is a variable that aggregates stress symptoms triggered by situations related with friends, parents, school and poverty.

# Appendix

## A Estimations Without Treatment Effect Structure

Table A.1: Non-Cognitive and Cognitive Measures at Age 14 on Outcome Results at Age 18

VARIABLES	(1) Depre	(2) Drink	(3) Smoke	(4) Satisfied	(5) Sick	(6) Mental H.
Age in Months	0.001 (0.005)	-0.002 (0.003)	0.002 (0.002)	0.001 (0.003)	0.001 (0.001)	0.001 (0.001)
Male	-0.224*** (0.037)	0.137*** (0.019)	0.128*** (0.013)	0.059*** (0.019)	-0.018* (0.010)	0.012 (0.008)
Oldersiblings	-0.022 (0.035)	0.024 (0.018)	0.004 (0.012)	0.020 (0.018)	-0.008 (0.010)	0.003 (0.008)
Youngsiblings	-0.004 (0.035)	-0.007 (0.019)	-0.016 (0.012)	0.009 (0.018)	-0.002 (0.010)	-0.000 (0.008)
Lnmonthincpc	-0.020 (0.034)	-0.002 (0.018)	-0.003 (0.012)	0.034* (0.018)	-0.010 (0.010)	-0.002 (0.007)
Urban	0.008 (0.054)	-0.037 (0.028)	-0.002 (0.019)	-0.046 (0.028)	-0.014 (0.015)	-0.021* (0.012)
Both Parents	-0.360*** (0.109)	-0.019 (0.056)	0.017 (0.037)	0.193*** (0.055)	-0.032 (0.030)	0.005 (0.024)
OnlyMother	-0.205 (0.144)	-0.036 (0.074)	0.087* (0.050)	0.186** (0.074)	0.040 (0.041)	0.026 (0.032)
FatherEd2yCol	-0.056 (0.071)	-0.024 (0.038)	0.003 (0.026)	0.107*** (0.038)	-0.018 (0.020)	-0.011 (0.016)
FatherEd4yCol	-0.053 (0.044)	-0.082*** (0.023)	-0.031** (0.015)	0.017 (0.023)	0.012 (0.012)	0.003 (0.010)
FathereducGS	-0.083 (0.081)	-0.043 (0.043)	-0.029 (0.029)	0.154*** (0.042)	-0.002 (0.023)	0.012 (0.018)
Non-Cogn Sk.	-0.455*** (0.067)	-0.106*** (0.034)	-0.093*** (0.023)	0.256*** (0.033)	-0.053*** (0.017)	-0.031** (0.013)
Cognitive Skill	0.108*** (0.031)	-0.013 (0.016)	-0.046*** (0.011)	0.011 (0.016)	0.002 (0.008)	0.010 (0.006)
Constant	0.618*** (0.181)	0.524*** (0.094)	0.071 (0.063)	0.142 (0.094)	0.159*** (0.051)	0.044 (0.040)
Observations	2,395	2,690	2,690	2,690	2,514	2,514

*Note:* This Table presents the estimated coefficients of the outcome equations  $Y = \mathbf{X}_Y \beta^Y + \alpha^{Y,A} \theta^A + \alpha^{Y,B} \theta^B + e^Y$ . Depre corresponds to a standardized index of depression symptoms. Drink takes the value of 1 if the respondent drank an alcoholic beverage at least once during the last year. Smoke takes the value of 1 if the respondent smoked a cigarette at least once during the last year. Satisfied takes the value of 1 if the respondent reports being happy with the way she is leading her life. Sick takes the value of 1 if the respondent reports having felt physically ill during the last year. Mental H. takes the value of 1 if the respondent has been diagnosed with psychological or mental problems. See the descriptions of the controls used in Table A.2. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.2: Non-Cognitive and Cognitive Measures at Age 14 on Outcome Results at Age 18

VARIABLES	(1) inCollege <sup>†</sup>	(2) Stress: Friends	(3) Stress: Parent	(4) Stress: School	(5) Stress: Total	(6) Stress: Poverty
Age in Months	0.003 (0.003)	-0.003 (0.006)	0.003 (0.006)	0.003 (0.005)	0.005 (0.006)	0.006 (0.006)
Male	-0.128*** (0.019)	0.120*** (0.040)	0.097** (0.040)	-0.169*** (0.039)	-0.134*** (0.040)	-0.239*** (0.040)
Oldersiblings	-0.005 (0.017)	0.013 (0.037)	-0.022 (0.037)	0.065* (0.036)	-0.000 (0.037)	-0.019 (0.037)
Youngsiblings	0.001 (0.018)	0.042 (0.038)	0.044 (0.038)	0.049 (0.037)	0.018 (0.038)	-0.021 (0.038)
Lnmonthincpc	0.023 (0.018)	-0.004 (0.037)	0.092** (0.037)	0.132*** (0.036)	0.032 (0.037)	-0.094** (0.037)
Urban	-0.016 (0.028)	-0.063 (0.058)	0.048 (0.058)	0.041 (0.057)	0.054 (0.058)	0.119** (0.058)
Both Parents	0.192*** (0.056)	-0.366*** (0.118)	-0.214* (0.117)	0.150 (0.114)	-0.175 (0.116)	-0.102 (0.117)
OnlyMother	0.075 (0.073)	-0.353** (0.157)	-0.203 (0.156)	0.074 (0.152)	-0.208 (0.155)	-0.114 (0.156)
FatherEd2yColl	0.073* (0.038)	-0.046 (0.078)	0.058 (0.077)	0.097 (0.075)	0.064 (0.077)	-0.000 (0.077)
FatherEd4yColl	0.011 (0.022)	-0.063 (0.048)	0.061 (0.048)	0.091** (0.047)	-0.030 (0.048)	-0.139*** (0.048)
FathereducGS	-0.041 (0.041)	-0.089 (0.088)	0.016 (0.087)	-0.018 (0.085)	-0.197** (0.087)	-0.373*** (0.087)
Non-Cogn Sk.	0.028 (0.034)	-0.316*** (0.072)	-0.101 (0.073)	0.073 (0.073)	-0.310*** (0.071)	-0.369*** (0.068)
Cognitive Skill	0.085*** (0.016)	0.150*** (0.034)	0.247*** (0.034)	0.364*** (0.034)	0.291*** (0.033)	0.071** (0.032)
Constant	0.454*** (0.092)	0.362* (0.197)	-0.361* (0.195)	-0.793*** (0.191)	0.001 (0.195)	0.552*** (0.196)
Observations	2,370	2,508	2,508	2,508	2,508	2,508

*Note:* *Note:* This Table presents the estimated coefficients of the outcome equations  $Y = \mathbf{X}_Y \beta^Y + \alpha^{Y,A} \theta^A + \alpha^{Y,B} \theta^B + e^Y$ . Variable inCollege takes the value of 1 if the respondent attends college by age 19. The Stress variables are standardized indexes that collect stress symptoms triggered by different sources, namely friends, parents, school and poverty. Stress:Total aggregates the four triggers of stress. Oldersiblings corresponds to the number of older siblings the respondent has. Youngsiblings corresponds to the number of younger siblings the respondent has. Lnmonthincpc corresponds to the natural logarithm of the monthly income per capita. Both Parents takes the value of one if the respondent lives in a biparental household. OnlyMother takes the value of one if the respondent's father is absent from the household. Fathereduc2yColl takes the value of one if the highest level of education attained by the respondent's father was a 2-year college degree. Fathereduc4yColl takes the value of one if the highest level of education attained by the respondent's father was a 4-year college degree. FathereducGS takes the value of one if the highest level of education attained by the respondent's father was graduate school. Columns headed as 1 collect the coefficients for those who were bullied at age 15. Columns headed as 0 collect the coefficients for those who were not bullied at age 15. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

<sup>†</sup> College attendance is measured at age 19.

## B Information Used to Construct Non-cognitive score

In the case of locus of control, we added three questions:

1. I have confidence in my own decision
2. I believe that I can deal with my problems by myself
3. I am taking full responsibility of my own life

To create the self-esteem index we added:

1. I think that I have a good character
2. I think that I am a competent person
3. I think that I am a worthy person
4. Sometimes I think that I am a worthless person (the negative of)
5. Sometimes I think that I am a bad person (the negative of)
6. I generally feel that I am a failure in life (the negative of)
7. If I do something wrong, people around me will blame me much (the negative of)
8. If I do something wrong, I will be put to shame by people around me (the negative of)

Finally we created the irresponsibility index by adding:

1. I jump into exciting things even if I have to take an examination tomorrow
2. I abandon a task once it becomes hard and laborious to do
3. I am apt to enjoy risky activities