CENTRE FOR
SOCIAL SCIENCE RESEARCH

Social Surveys Unit

Exposure to violence and educational outcomes: evidence from Cape Town, South Africa

Duncan Pieterse

CSSR Working Paper No. 306

April 2012

Duncan Pieterse is a PhD student in the Department of Economics at the University of Cape Town and affiliated to AIDS and Society Research Unit (ASRU) at the University of Cape Town as a graduate student.

Exposure to violence and educational outcomes: evidence from Cape Town, South Africa

Abstract

We explore the relationship between exposure to violence during childhood perpetrated by adults inside the home and educational outcomes in the context of higher than average rates of violence in Cape Town, South Africa and the disproportionate exposure to violence of young South Africans (black and coloured youth in particular). We match official police murder statistics at the neighbourhood level to the Cape Area Panel Study to provide a unique descriptive analysis of violence in Cape Town and we determine the extent of selection bias using matching techniques. Using three measures of educational outcomes (numeracy and literacy test scores, dropout and high school exam results), we: (i) estimate kernel density functions of continuous educational outcomes measures by race and exposure to violence during childhood; (ii) remove constant differences in unobserved family and neighbourhood background that may bias the results by using sibling and neighbourhood fixed effect models; (iii) check the robustness of our sibling fixed effect regressions by including birth order effects. In the neighbourhood fixed effect regressions, the measures of exposure to violence are significant and have a large negative effect on educational outcomes (with the exception of literacy scores). In the sibling fixed effect regressions, the effect remains for two of the four measures of exposure to violence during childhood. The measure of exposure to emotional violence during childhood is least affected by selection bias and the only measure robust to the inclusion of birth order effects.

Introduction

The United Nations Study on Violence against Children defines violence against children as "the intentional use of physical force or power, threatened or actual, against a child, by an individual or group, that either results in or has a high likelihood of resulting in actual or potential harm to the child's health, survival, development or dignity" (OHCHR, 2006: 6). This definition highlights the adverse consequences of exposure to violence and incorporates the fact that violence can be threatened or actual. The distinction between threatened and actual violence allows us to differentiate direct from indirect exposure to violence. Direct exposure to violence refers to victimisation (including emotional violence such as threats of violence) while indirect exposure to violence is when an individual is a witness to an act of violence.[1] Violence against children can take place at home within the family, at school and in educational settings, in care and justice systems, in work settings or in the community. In this study we focus on the consequences of direct and indirect violence against children perpetrated by adults within the home.

Given the scale of violence in Cape Town, the disproportionate exposure to violence of young South Africans (black and coloured youth in particular) and established relationships between exposure to violence and a range of adverse outcomes, we are interested in the relationship between exposure to violence during childhood and educational outcomes during adolescence and adulthood.[2] We consider both intermediate educational outcomes (test scores) and final educational outcomes (exam results during the final year of high school and years of schooling). The relationship between exposure to violence during childhood and educational outcomes is based on the manner in which adverse experiences during childhood: (i) affect the formation of cognitive and non-cognitive skills; and (ii) induces a vulnerability to periods of acute stress.[3]

---

[1] There are broader definitions of violence that conceptualise violence along a continuum from direct physical assault to symbolic violence and normalised everyday violence (see Scheper-Hughes and Bourgois, 2004). Even broader definitions include the concept of structural violence, which refers to historically and institutionally embedded violence that constrain the individual agency of the poor such as exploitation, exclusion, inequality and injustice (see Galtung, 1969 and Farmer, 2004).

[2] These findings are from a separate paper that explores the experience of violence by young people in Cape Town.

[3] Cognitive skills are intelligence or mental ability where intelligence is defined as the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" (Neisser et al., 1996: 77). Non-cognitive skills include patience, self-control,

Evidence from neuroscience indicates that because childhood is a period of heightened sensitivity to positive and negative influences, adverse experiences or stress during early childhood can cause permanent changes to brain architecture and gene expression in a manner that influences the formation of cognitive and non-cognitive skills (see Knudsen et al., 2006). And, evidence from the epidemiology literature indicates that parents are the active ingredients of environmental influence during childhood (see Shonkoff, and Phillips, 2000). Taken together, this suggests that violence perpetrated by parents against children inside the home during childhood may impose a heavy developmental burden. The implications of these findings for the formation of cognitive and non-cognitive skills (and therefore educational outcomes) is formalised by the human skill formation model.[4]

The human skill formation model, which is developed in Cunha et al. (2006), is concerned with how skills are formed over the life cycle of an individual. In particular, it emphasises the role of parents in shaping both cognitive and non-cognitive skills through genetic endowments and pre- and post-natal environments (see Cunha et al., 2006, Cunha and Heckman, 2007 and 2009, and Cunha et al., 2010). In this model, each agent possesses a vector of cognitive and non-cognitive skills at each stage, which are used with different weights in different tasks. The process by which these cognitive and non-cognitive skills are produced is governed by a multi-stage technology, where all skills produced are a function of the inputs or investments at that stage.

Cunha and Heckman (2007) highlight two important features of the multi-stage technology: self-productivity and dynamic complementarity. Self-productivity has two elements: (i) skills produced during one stage persist into future periods; and (ii) non-cognitive skills may interact with cognitive skills to produce better outcomes (e.g. self-control and motivation may promote more vigorous acquisition of cognitive skills).[5] Dynamic complementarity refers to the idea that skills produced during one stage raise the productivity of investment at subsequent stages (therefore, investments in one period are more productive when there is a high level of capability in an earlier period). This model of skill formation consists of multiple stages of childhood where inputs at different

---

temperament, motivation and time-preference. See Almlund et al. (2011) for a more detailed discussion.

[4] An similar (less formal) theoretical model entitled the life-course perspective is reviewed by Furstenberg et al. (1987). According to the life-course perspective, which emphasises the role of family in shaping intellectual and social development of children, experiences early in life shape the behaviour and educational performance of adults (Morán et al., 2004).

[5] Almlund et al. (2011) provide detailed evidence of how cognitive skills and non-cognitive skills (what they refer to as personality traits) interact to influence academic performance.

stages are complements and where there is self-productivity of investment (Cunha and Heckman, 2007). Self-productivity and dynamic complementarity suggest that early adverse experiences may be especially detrimental to the human skill formation process because it may affect later stages of cognitive and non-cognitive skills development through the overall (accumulated) level of cognitive and non-cognitive skills and the efficiency with which cognitive and non-cognitive skills are produced.

It is well documented in the developmental psychology and epidemiology literature that exposure to violence during childhood is associated with various adverse outcomes. In one of the earliest studies, Bell and Jenkins (1991) showed that exposure to trauma, especially violence in the family, interferes with a child's normal development of trust and later exploratory behaviour, which lead to the development of autonomy. Several studies find an association between exposure to violence during childhood and anxiety, depression, attention-deficit or hyperactivity disorder and aggressive behaviour due to changes in brain functioning caused by elevated cortisol levels (see Famularo et al., 1992, Cooley-Quille et al., 1995, Schwab-Stone et al., 1995, Gorman-Smith and Tolan, 1998 and Guerra et al., 2003). Other studies have found an association between exposure to violence during childhood and increased risk of schizophrenia and substance abuse (see Felitti et al., 1998 and Kendler et al., 2000) as well as diabetes, heart disease, and immune disorders (see Felitti et al 1998 and Francis et al., 1999).

There are several studies from South Africa that confirm findings from the international literature. In one of the earliest studies from Cape Town, Ensink et al (1997) use a survey to determine exposure to community violence and questionnaires and clinical assessments to elicit post-traumatic stress disorder (PTSD) symptoms in a sample of 60 children aged 10-16 from Khayelitsha. The authors find that 95% had witnessed violent events, 56% had experienced violence themselves and 22% met criteria for PTSD. Using a sample of grade 10 adolescents in secondary schools in the Western Cape, Seedat et al. (2000) found a positive association between violent and multiple traumas and PTSD symptoms in adolescents. Barbarin et al. (2001) found that exposure to violence had a significant negative association with psychological functioning that was independent of gender and socioeconomic status and the main negative effects occurred in the domains of attention, aggression, and anxiety-depression. Using interviews with high school students in Cape Town, Ward et al. (2001) found positive associations between symptoms for PTSD and depression and exposure to several types of violence.

This review of the human skills formation perspective and the literature on the adverse consequences of exposure to violence during childhood, suggest the relationship between exposure to violence during childhood and educational outcomes is governed by two transmissions mechanisms. First, to the extent that the adverse mental and physical health outcomes associated with exposure to violence during childhood disrupt the formation of non-cognitive skills, exposure to violence during childhood may have an impact on educational outcomes. Given the association between exposure to violence during childhood and anxiety, depression and behavioural disorders later in life, it seems likely that exposure to violence during childhood affects the formation of non-cognitive skills such as patience, self-control, temperament, motivation and time-preference. To the extent that exposure to violence during childhood affects the formation of non-cognitive skills, it may affect educational outcomes through a direct impact on non-cognitive skills and may also have an indirect impact through the importance of these non-cognitive skills for the acquisition of cognitive skills. This is confirmed by evidence from the psychology literature that shows the importance of non-cognitive skills for schooling performance of children and adolescents (see Wolfe and Johnson, 1995 and Duckworth and Seligman, 2005). Heckman (2000), Carneiro and Heckman (2003) and Heckman et al. (2006) take this research further and show that both cognitive and non-cognitive skills matter for academic performance and educational attainment (what we broadly refer to as educational outcomes).

Second, the fact that exposure to violence during childhood is associated with a range of adverse physical and mental health outcomes later in life that are known to manifest or worsen in relation to acute or chronic life stress, suggests that adverse experiences during childhood may induce a vulnerability to the effects of stress during adolescence and adulthood (Heim and Nemeroff, 2001: 1024). To the extent that exposure to violence during childhood induces a vulnerability to stress during adolescence and adulthood, it may affect educational outcomes since academic performance is generally measured during periods of acute stress (such as tests or exams). Evidence from the neuroscience and developmental psychology literature shows that high levels of stress can impair academic performance due to the effect of elevated stress on memory loss (see Sauro et al., 2003, McEwen and Sapolsky, 1995 and Lupien et al., 2007). This view is consistent with social-psychological strain theory developed by Agnew (1992), which focuses on exposure to violence during childhood as a source of acute stress.

The extent to which exposure to violence during childhood affects the human skill formation process (and consequently educational outcomes) depends on how an individual responds to the stress associated with direct or indirect

violence. The transmission mechanisms that govern the relationship between exposure to violence and educational outcomes through its impact on cognitive and non-cognitive skills depend on: (i) individual characteristics (age, gender, individual personality or resilience and role within the family); (ii) the nature of the exposure to violence (type (severity), timing, frequency and length); and (iii) the presence of support systems (such as other family or school). Holt et al. (2008: 805) provide a summary of the literature on the differential impact of exposure to violence between boys and girls, Schwartz and Proctor (2000) show that the impact of exposure to violence related to being a victim differs significantly from that related to witnessing violence and Barbarin et al. (2001) provide evidence of the positive role of support systems (such as positive family relationships or school support).

The human skills formation process provides a theoretical framework within which to explore the relationship between exposure to violence and educational outcomes. In a separate descriptive paper we identified the main motivation for studying the consequences of exposure to violence, which is based on the relationship between violence and adverse welfare effects as well as its impact on equality of opportunity. The consequences of a disruption in the human skills formation process allow us to identify another motivation for studying the consequences of exposure to violence. To the extent that exposure to violence during childhood causes negative consequences for academic performance or educational attainment, it may inhibit the production benefits of education (lifetime earnings) as well as the non-production benefits of education for the young adult affected by exposure to violence.[6]

Furthermore, to the extent that exposure to violence causes an individual to drop out of school (due to poor grades related to a counterproductive learning environment, for example), it will affect the efficiency of the education system as well as labour productivity. Hanushek et al. (2006: 3) highlight two adverse consequences of dropout: (i) the investment-benefit perspective on school policy highlights potential lost productivity from premature school dropout (this should be clear from the discussion of the production of cognitive and non-cognitive skills); and (ii) the cost of education-efficiency perspective suggests that if the objective is to get a given number of students through some level of schooling, having students drop out earlier raises the cost of achieving that goal. Also, having disproportionately large numbers returning dropouts or grade repeaters in schools may distort normal instruction and raise the overall cost of education.

[6] Lochner (2011) provides an overview of the non-production benefits of education in three areas: (i) the effect of educational attainment and school quality on participation in crime; (ii) the effect of educational attainment on health and life expectancy; and (iii) the effect of education on citizenship, political participation and democracy.

In this paper we explore the relationship between exposure to violence during childhood and educational outcomes in the context of higher than average rates of violence in Cape Town, South Africa and the disproportionate exposure to violence of young South Africans (black and coloured youth in particular). The relationship between exposure to violence and educational outcomes has received some attention by economists.[7]

In one of the earliest studies, Grogger (1997) used High School and Beyond data from the United States to explore the impact of local violence on high school graduation. He used responses by high school principals to questions about: (i) fights among students; (ii) conflicts between students and teachers; and (iii) students bringing weapons to school to construct an index of school violence (Grogger, 1997: 662). His findings indicate that moderate levels of violence reduce the likelihood of graduation from high school by 5% and lower the probability that a student will attend college by 7% (Grogger, 1997: 659).

In a more recent study, Aizer (2008) used two measures of violence (the rate of hospitalisations for assault from California Hospital Discharge data to create measures of violence at the zipcode level and police data by reporting districts in Los Angeles) to determine the impact of violence on cognitive test scores. The study uses both family and neighbourhood fixed effects models to control for unobserved family and neighbourhood disadvantage. Once the author controls for unobserved underlying disadvantage, the impact of violence declines for some child outcomes, but it is still significant for others (having violent peers, for example).

Sharkey (2009) uses data from the Project on Human Development in Chicago Neighbourhoods merged with homicide data in Chicago from 1994 to 2002 and neighbourhood fixed effects to assess the impact of recent local homicides on cognitive assessments among children. He finds that African American children from neighbourhoods across Chicago interviewed within a week of a murder in their neighbourhood had achievement test scores one-half standard deviations lower than other children.

There is also some evidence from Latin America where many countries have similar levels of violence to South Africa. Using data from Chile and Nicaragua,

[7] The theoretical and empirical literature in economics have traditionally focused on two aspects of the economics of crime and violence: (i) the determinants of the decision to participate in crime (following from the pioneering theoretical work by Becker, 1968; see Freeman, 1999 for an empirical overview); and (ii) the adverse welfare effects of violent crime (including the economic costs of crime and the direct and indirect effects of violent crime).

Morrison and Orlando (1997) find that children who are exposed to domestic violence are more likely to have disciplinary problems at school and repeat grades and Knaul and Ramírez (2005) find a relationship between child abuse and educational attainment in Columbia and Mexico. The findings reviewed here are consistent with studies that document the negative educational shocks endured by children in the aftermath of violent conflicts (de Walque 2006; Akresh and de Walque, 2008; Blattman and Annan, 2010; and Shemyakina, 2011).

To our knowledge, similar studies have not been conducted for South Africa. My key contribution is to provide the most robust estimate of the relationship between exposure to violence during childhood and educational outcomes in the absence of exogenous variation in exposure to violence. Other contributions include the use of matching techniques to establish the extent of selection bias in the measures of exposure to violence and the inclusion of birth order effects and birth location effects. In addition, the neighbourhood fixed effect regressions performed here are based on a unique spatial definition of Cape Town neighbourhoods created by matching police precinct boundaries to the Cape Area Panel Study. Finally, I contribute to the literature on empirical evidence for the recently completed theoretical work on the human skills formation process. These contributions permit a better understanding of the long-term impacts of adverse childhood experience in the context of high rates of violence.

The remainder of this paper is organised as follows: first, we introduce a longitudinal study of the lives of young people in metropolitan Cape Town (the Cape Area Panel Study) and provide an overview of the four measures of exposure to violence during childhood, the three measures of educational outcomes and the two sibling samples. Second, we conduct a descriptive analysis, including an exploration of average educational outcomes by neighbourhood, examine the extent of "selection into exposure to violence bias" for each measure of exposure to violence during childhood and a non-parametric examination of the relationship between numeracy test scores and matric results on the one hand and exposure to violence during childhood on the other. Third, we introduce the empirical strategy, which is based on the reduced-form literature on the determinants of educational outcomes. Finally, we conduct neighbourhood and sibling fixed effect regressions to determine whether exposure to violence during childhood has an impact on educational outcomes during adolescence.

Data

The analysis in this paper relies on data from three sources: the Cape Area Panel Study (CAPS), the 2001 Census and the South African Police Service (SAPS). CAPS is a longitudinal study of the lives of young people in metropolitan Cape Town, South Africa.[8] Cape Town is the third largest city in South Africa (Statistics South Africa, 2011) and an ideal setting for our study due to its: (i) higher than average violent crime rates; (ii) large variation in violent crime across neighbourhoods; and (iii) substantial numbers of white, coloured, and black residents, which allow unique opportunities for the study of the post-apartheid experience of violence by young people from diverse backgrounds. Wave 1 of CAPS collected interviews from about 4,750 randomly selected young people aged 14-22 in August-December, 2002. Household, school, work, childbearing and sexual behaviour data were collected for each young adult in the sample.

The table below reports the weighted and unweighted sample characteristics for Wave 1. The sampling weights adjust for three elements of the CAPS sample design: (i) oversampling of black and white households to ensure racial distribution that matched the 1996 Census; (ii) differential sampling of households with and without young adults; and (iii) the addition of secondary households (backyard shacks) into the sample of screener households (Lam et al., 2008: 39). The sample weight, *weightyr*, adjusts for the three elements of sample design mentioned above as well as household and young adult non-response. Using *weightyr* creates a weighted distribution of 14-22 year-olds by population group that is within one percentage point of the population group distribution in Cape Town in the 1996 census (Lam et al., 2008).

There are slightly more females than males in Wave 1 of CAPS, the mean age is approximately 18 years, 3% of the young adults live in a household that receives a government grant, the mean household size is 5, 6% of the young adults in the sample indicate that their mother is deceased, 31% live in female-headed

[8] The Cape Area Panel Study Waves 1-2-3 were collected between 2002 and 2005 by the University of Cape Town and the University of Michigan, with funding provided by the U.S. National Institute for Child Health and Human Development and the Andrew W. Mellon Foundation. Wave 4 was collected in 2006 by the University of Cape Town, University of Michigan and Princeton University. Major funding for Wave 4 was provided by the National Institute on Aging through a grant to Princeton University, in addition to funding provided by NICHD through the University of Michigan. Additional information is available on the CAPS website: www.caps.uct.ac.za. In this study we also use data from Wave 5 of CAPS which has not yet been released publicly.

households and just over 20% of the sample speak English at home. The average young adult's mother has about 9 years of schooling, which is likely to be lower if the mothers for which education information are missing (11%) are less educated on average.

Table 1: Variable definitions and sample characteristics: individual and household characteristics (Wave 1)

Variable	Description	Wave 1	
		Unweighted	Weighted
male	0-1 dummy that equals 1 if the young adult is male	0.451 (0.498) *4746*	0.476 (0.008) *4746*
coloured	0-1 dummy that equals 1 if the young adult is coloured	0.422 (0.494) *4746*	0.526 (0.013) *4746*
black	0-1 dummy that equals 1 if the young adult is black	0.453 (0.498) *4746*	0.282 (0.010) *4746*
white	0-1 dummy that equals 1 if the young adult is white	0.125 (0.331) *4746*	0.187 (0.010) *4746*
age	Age of the young adult	17.875 (2.482) *4746*	17.923 (0.046) *4746*
log_hh_income	Log of combined household monthly income	7.696 (1.072) *3496*	7.939 (0.034) *3496*
grant	0-1 dummy that equals 1 if the household where the young adult resides receives a government grant	0.028 (0.164) *4712*	0.025 (0.003) *4712*
hh_size	Number of individuals in the young adult's household	5.431 (2.514) *4746*	5.307 (0.063) *4746*
mother_deceased	0-1 dummy that equals 1 if the young adult's mother is deceased	0.061 (0.239) *4,728*	0.053 (0.004) *4,728*
female_hh	0-1 dummy that equals 1 if the young adult lives in a household headed by a female	0.308 (0.462) *4,746*	0.292 (0.011) *4,746*
home_lang_english	0-1 dummy that equals 1 if English is the language spoken most often at home	0.206 (0.404) *4746*	0.287 (0.018) *4746*
mother_education	Years of schooling completed by the young adult's mother	8.845 (3.321) *4228*	9.334 (0.099) *4228*
mother_educ_missing	0-1 dummy that equals 1 if the years of schooling completed by the young adult's mother is missing	0.109 (0.312) *4746*	0.101 (0.006) *4746*

Source: Own calculations. CAPS Wave 1. The CAPS sample was weighted using *weightyr* as a probability weight. For the unweighted results the standard deviations are reported in brackets and for the weighted results the (linearised) standard errors are reported in brackets. The number of observations is given in italics.

The household measures included above are based on factors that have shown to have an adverse consequence for childhood development: (i) poverty (as measured by *grant* and *log_hh_income*; see Cunha et al., 2006); (ii) the lack of one loving and consistent adult (as measured by *mother_deceased* and *female_hh*; see Shonkoff, and Phillips, 2000); and socio-economic status (as measured by *mother_education* and *home_lang_english*; see Currie, 2009).[9]

Since *log_hh_income* is missing for a large part of the CAPS sample and is generally prone to measurement error, *grant* may be a better measure of household welfare. Household income is often measured with error due to recall bias (if an individual is unable to accurately recall their income) and reporting bias (intentional under- or over-reporting). Using the household income variable is particularly problematic if the households for which the income data are missing are systematically different in their exposure to violence during childhood compared to those for which income information is available.

The dummy that equals 1 if English is the language most often spoken at home is a proxy for class or social status under the assumption that, for coloured and black young adults in particular, speaking English at home is potentially an indication of a privileged background. Some studies find strong positive associations between English as a home language and positive educational and labour market outcomes (see Cornwell and Inder, 2008 and Casale and Posel, 2010).

The measures of exposure to violence against children are based on questions from Wave 1 that asked young people to reflect on their exposure to violence by adults at home when they were growing up (until the age of 14). We used their answers to construct four dummies that equal one if the young adult was exposed to violence during childhood (we characterise the period from infancy until the age of 14 as childhood).[10] There are two caveats to keep in mind when

[9] A measure of socio-economic status that is used in the empirical analysis, but not specified in the table above is area of residence (see Currie, 2009). We include neighbourhood dummies in all baseline regressions. We have a measure of household size (*hh_size*), but choose not to include this in the empirical analysis because evidence of the relationship between household size and child outcomes is mixed: some studies find no effect (see Angrist, Lavy and Scholsser, 2005) while others find a negative effect (Rosenzweig and Wolpin, 1980, Behrman et al., 1989, Black, Devereaux and Salvanjes, 2005 and Rosenzweig and Zhang, 2009).

[10] These measures were derived from questions were the young adults were asked to reflect on their exposure to violence childhood using a five-point scale: 1 (never), 2 (once or twice), 3 (sometimes), 4 (often) and 5 (very often). Each of these dummies was constructed as follows: 1 if the young adult indicated sometimes, often or very often and 0 if the young adult indicated never or once/twice. It can argued that the categorisation used to construct these

interpreting these reports of exposure to violence during childhood: (i) reporting bias exists if certain groups are systematically less or more inclined to talk about their experiences of violence or if more severe types of violence are more likely to be underrated on the five-point scale; and (ii) selection bias exists if the young adults that choose not to answer the questions on exposure to violence have systematically different experiences of violence to those who answer the questions (young adults were allowed to refuse to answer the question: the refusal rate was highest for *hit_hard* at 4% and less than 0.5% for the other measures).

The table below shows that exposure to violence during childhood is high and there is substantial variation across the four measures: 6% of the young adults report being hit hard during childhood, about 17% report being pushed and a fear of being hurt and 33% report being put down by adults. A higher proportion of coloured youth have been hit hard, pushed and put down by adults compared to other race groups. A small number of white youth in the sample have been hit hard (around 2%), compared to 5% for black youth and 8% for coloured youth. A substantially greater proportion of the youth have been pushed during childhood: 9% of white youth, 17% of black youth and 20% of coloured youth. Threatened violence (as measured by the fear of being hurt) is much greater amongst black youth (22%), compared to coloured (17%) and white (8%) youth.

The differences between youth for the *adult_put_down* variable (which can be seen as a measure of emotional violence) are less striking than the other measures: 36% of coloured youth have experienced emotional violence, compared to 32% of black youth and 23% of white youth. These sample means suggest that a greater number of coloured youth have experienced actual violence, compared to black youth, where the experience of violence is mostly threatened.

dummies is somewhat arbitrary, as a result, we re-estimate all regressions with dummies that reflect the full five-point scale.

Table 2: Variable definitions and sample characteristics: exposure to violence during childhood (unweighted)

Variable	Definition	Sample	Race			Gender	
			Black	Coloured	White	Male	Female
hit_hard	0-1 dummy that equals 1 if an adult, parent or stepparent living in the young adult's home often hit them so hard that they had marks or were injured	0.058	0.046	0.081	0.021	0.049	0.064
		(0.233)	(0.21)	(0.272)	(0.144)	(0.216)	(0.245)
		4,573	*2,094*	*1,909*	*570*	*2,055*	*2,518*
pushed	0-1 dummy that equals 1 if an adult, parent or stepparent living in the young adult's home often pushed, grabbed, slapped, or threw something at them	0.173	0.174	0.197	0.093	0.162	0.183
		(0.379)	(0.379)	(0.398)	(0.29)	(0.368)	(0.387)
		4,727	*2,141*	*1,992*	*594*	*2,131*	*2,596*
afraid_hurt	0-1 dummy that equals 1 if an adult, parent or stepparent living in the young adult's home often made them afraid that they might be physically hurt	0.179	0.215	0.171	0.076	0.165	0.19
		(0.383)	(0.411)	(0.377)	(0.265)	(0.371)	(0.393)
		4,739	*2,144*	*2,000*	*595*	*2,134*	*2,605*
adult_put_down	0-1 dummy that equals 1 if young adult was often verbally abused, insulted or put down by an adult, parent or stepparent living in their home	0.329	0.324	0.364	0.229	0.309	0.345
		(0.47)	(0.468)	(0.481)	(0.421)	(0.462)	(0.476)
		4,731	*2,143*	*1,994*	*594*	*2,130*	*2,601*
childhd_exposure	Score of exposure to violence during childhood that equals 0 if not exposed to any violence and 1 if exposed to all 4 types of violence during childhood	0.184	0.189	0.202	0.104	0.17	0.195
		(0.270)	(0.273)	(0.278)	(0.209)	(0.259)	(0.278)
		4,714	2,139	1,985	590	2,122	2,592

Source: Own calculations. CAPS Wave 1. Standard deviations in brackets and number of observations in italics.

The *childhd_exposure* measure, which is a score of total exposure to violence during childhood, is highest for coloured and black young adults (0.2 and 0.19) and much lower for white young adults (0.1).[11] These findings for coloured youth are consistent with a separate descriptive paper that shows higher than average exposure to violence experienced by coloured youth at school. For example, 18% of coloured young adults report teachers being threatened by students at their school (compared to about 4% for black and white young

[11] The *childhood_exposure* score is calculated to return missing if any of the dummies for the 4 types of exposure to violence during childhood is missing. This explains the slightly lower sample size for this variable.

adults), 11% of coloured young adults report that the teachers at their school are violent (compared to 3% of black and white young adults) and coloured youth report more than double the mean rates of bullying reported by black young adults.

Variation within the coloured and black groups is mostly large and similar for all the measures of exposure to violence during childhood. This is in contrast to white youth, where the standard deviations are lower for all measures, suggesting more similar experiences of childhood violence amongst white youth. The exposure to violence during childhood measures indicates that a higher proportion of female young adults have experienced actual and threatened violence compared to males.

These findings are consistent with evidence from South Africa's National Youth Victimisation Study: 27.1% of participants said they were often spanked at home as punishment for their wrongdoings (Leoschut and Burton, 2006: 30). In view of estimates from the National Youth Victimisation Study that 22% of youth witnessed family members intentionally hurting one another and in 40% of these cases weapons were used in the attacks, with 28% resulting in physical injuries (Leoschut and Burton, 2006: 31), we can assume that a significant number of the young adults in our sample were exposed to prenatal stress.[12]

We use three educational outcome measures in this study: (i) scores from numeracy and literacy tests administered to all young adults during Wave 1; (ii) educational attainment (dropout); and (iii) matric exam results.[13] The test scores are from Wave 1 of CAPS and the dropout and matric result measures were obtained from the full five-wave CAPS panel. Given the age profile of the young adults during Wave 1 of CAPS, Wave 5 (which has not yet been released publicly) provides the first occasion to measure dropout in CAPS.

[12] Prenatal stress has been associated with increased risk for major depression in adulthood (Hulshoff et al., 2000).

[13] All high school students write the matric exam at the end of Grade 12 and all students who obtained an aggregate mark of 40% of higher are known as matriculants. Details of how this variable was constructed are available in Appendix F.

Table 3: Variable definitions and sample characteristics: educational outcomes (unweighted)

Variable	Definition	Sample	Race			Gender	
			Black	Coloured	White	Male	Female
numeracy_score	Standardised numeracy test score	0	-0.438	0.069	1.342	0.052	-0.042
		(1)	(0.774)	(0.915)	(0.694)	(1.037)	(0.967)
		4,691	*2,120*	*1,981*	*590*	*2,115*	*2,576*
literacy_score	Standardised literacy test score	0	-0.468	0.231	0.906	-0.038	0.032
		(1)	(1.038)	(0.795)	(0.421)	(1.053)	(0.954)
		4,691	*2,120*	*1,981*	*590*	*2,115*	*2,576*
dropout	0-1 dummy that equals 1 if the young adult did not complete high school	0.469	0.545	0.505	0.094	0.499	0.446
		(0.499)	(0.498)	(0.5)	(0.292)	(0.5)	(0.497)
		3,757	*1,579*	*1,698*	*480*	*1,663*	*2,094*
matric_result	Average result obtained by the young adult for the final grade 12 (matric) exam	0.528	0.455	0.52	0.668	0.525	0.529
		(0.13)	(0.098)	(0.105)	(0.107)	(0.123)	(0.135)
		1,606	*611*	*645*	*350*	*654*	*952*

Source: Own calculations. These variables are constructed using all 5 waves of CAPS, with the exception of numeracy and literacy scores, which are available for Wave 1 only. Standard deviations in brackets and number of observations in italics.

The dropout and matric results measures are discussed in greater detail later, but there are two important caveats to keep in mind when interpreting these two measures of educational outcomes: (i) selection bias affects the matric results measure because we do not observe the matric results for those young adults who choose not to write the matric exam; and (ii) attrition bias exists if the young adults who drop out or write the matric exam after exiting the CAPS sample are systematically different in their exposure to violence during childhood compared to those who remain in the sample.

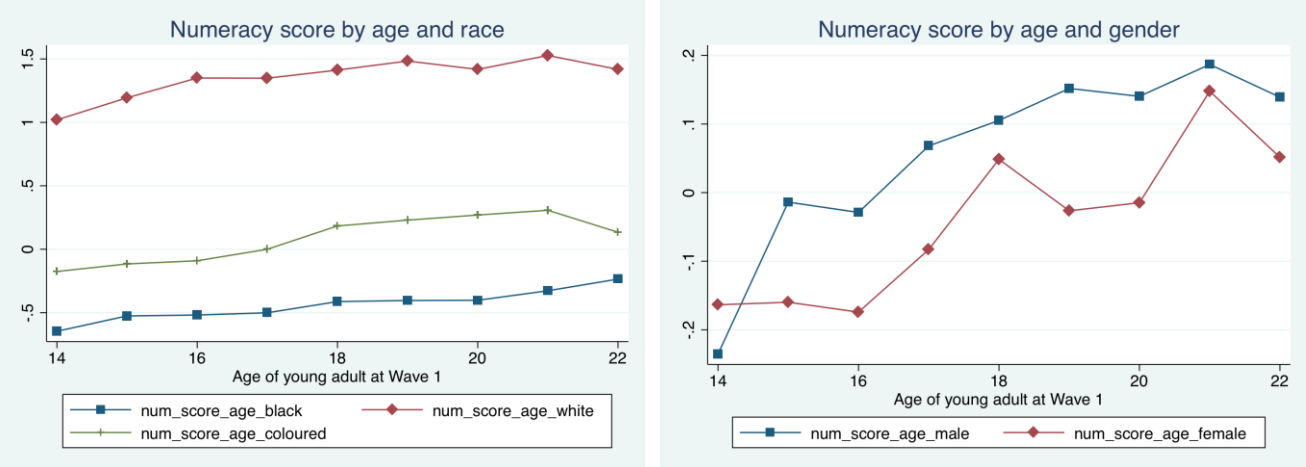
The first measure of educational outcomes is test scores. The same self-administered written literacy and numeracy test was administered to each young adult respondent in Wave 1 of CAPS. The tests could be taken in either English or Afrikaans (there was no Xhosa version, the home language of most black respondents). The test scores are standardised to zero mean and unit variance in the original dataset and are used unchanged here (numeracy test scores range between -1.7 and 2.3 and the literacy test scores range between -4.4 and 1.3).

The numeracy test scores indicate that the mean performance of black youth were about two standard deviations lower than their white counterparts and the mean performance of female youth were about one standard deviation lower than mean male performance. The average literacy test scores for coloured youth are much higher than their mean numeracy scores and mean female performance is about one standard deviation higher than mean male performance.[14] This is

[14] These results may represent differences in mean age across race and gender since all the young adults completed the same tests irrespective of age.

consistent with findings from the U.S. that girls tend to outperform boys on literacy tests in secondary school, while boys tend to outperform girls on numeracy tests (see, Downey and Yuan, 2005).

Figure 1: Numeracy test scores: by race and gender



Source: Own calculations. CAPS Wave 1.

The second measure of educational outcomes is dropout, which is a dummy that equals one if the young adult did not complete high school. This variable is constructed using educational attainment for all young adults in the sample based on their years of schooling from the full five-wave CAPS panel.[15] The means of the dropout variable show clear differences across race: black youth show the highest mean dropout rate (56%), followed by coloured (51%) and white (9%) youth.

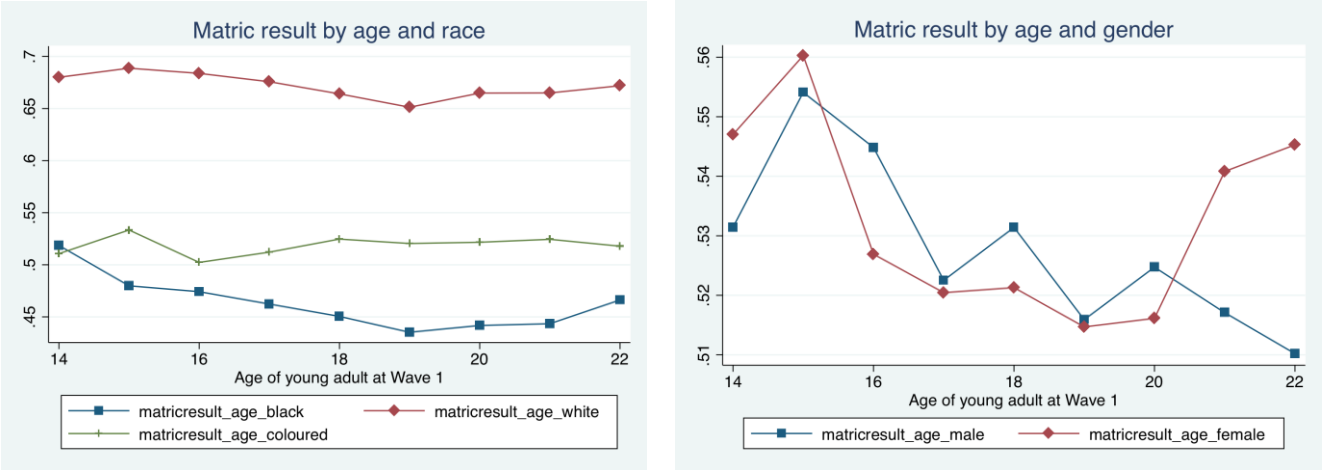
The third measure of educational outcomes is matric result. This measure is available only for those young adults that wrote the matric exam while part of the CAPS sample (it is possible that some young adults wrote the matric exam after they exited the CAPS sample).

The racial differences are striking: the mean matric performance of black youth is about 6 standard deviations lower than their coloured counterparts and the mean performance of coloured youth is about 14 standard deviations lower than

[15] Since the youngest individuals in the sample were 14 in Wave 1 (2002) and Wave 5 was completed in 2009, the youngest individuals in the sample are 20 years old by Wave 5. The dropout dummy equals 0 for individuals who completed high school between Waves 1 and 5 and equals 1 for all individuals who were dropouts by Wave 5. However, it is possible that some of the young adults who left the sample while in school dropped out of school after their attrition. As a result, the young adults for which the dropout measure is indicated as missing are a combination of individuals who are still in school in Wave 5 (and 20 years old), but may complete high school at some point and those who left the sample before completing high school.

their white counterparts. These racial disparities are consistent with extreme differences in school quality along racial lines before the end of apartheid (see Case and Deaton, 1999 and Case and Yogo, 1999) and after the end of apartheid (see Yamauchi, 2005).

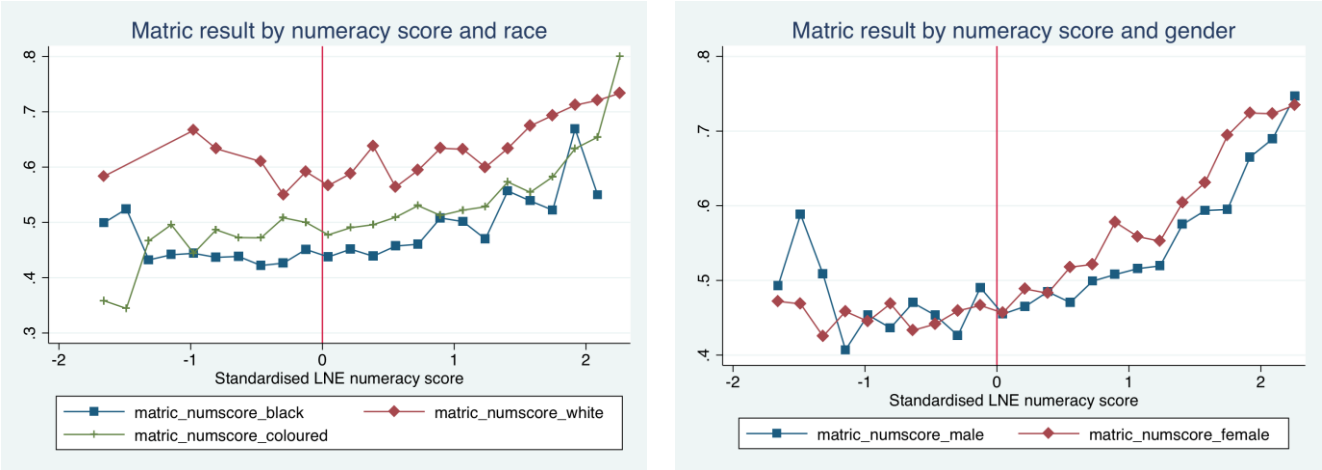
Figure 2: Matric results: by race and gender



Source: Own calculations. CAPS Panel.

There is a clear positive relationship between numeracy test scores and matric results for the young adults who wrote the matric exam: young adults who did well in the numeracy and literacy tests administered during Wave 1 also tend to do well in their matric exam. The slopes in the first graph indicate a much stronger positive association between numeracy test scores and matric results for coloured youth compared to black and white youth. Both graphs suggest that the positive association between numeracy test scores and matric results exists only for those young adults who obtained above average numeracy test scores.

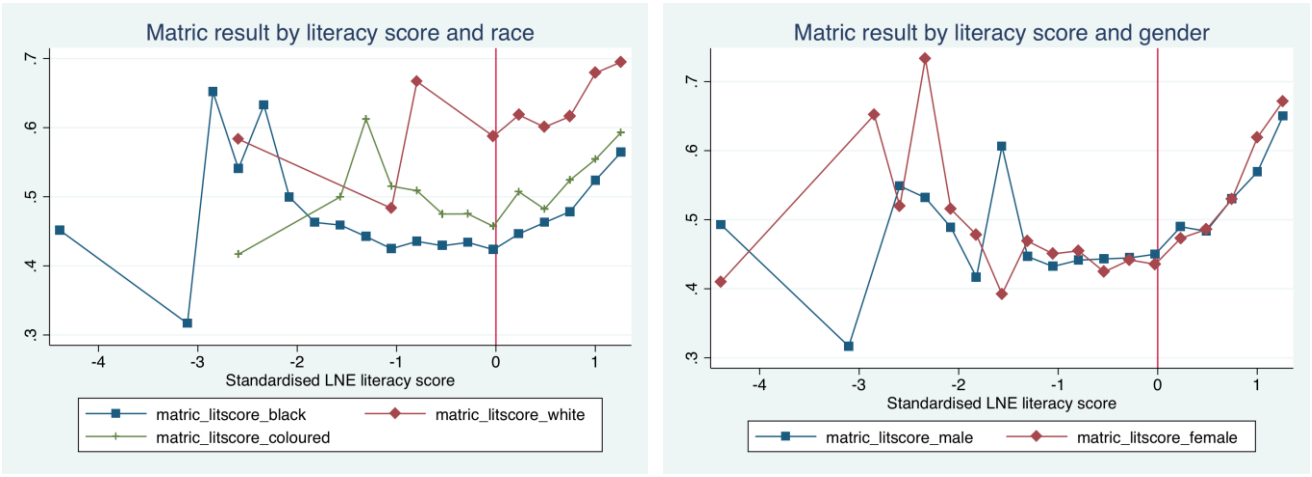
Figure 3: Numeracy test scores and matric results



Source: Own calculations. CAPS Panel.

These findings are consistent with the features of the multi-stage technology that governs human skills formation and the notion of critical periods for the development of cognitive and non-cognitive skills (see Heckman, 2012). The positive relationship between numeracy test scores and matric results (evident for those who obtained above average results for the numeracy tests) reflect the self-productivity of cognitive and non-cognitive skills: skills acquired in one period persist into future periods (the young adults who did very well in the numeracy tests also achieved very good matric results). The fact that, in general, those young adults who scored low on the numeracy tests were unable to achieve above average matric results is consistent with the concept of a "bottleneck" period: if skills at one stage of the life cycle are not formed at a sufficiently high level it is difficult to achieve excellence at the next stage (Cunha et al., 2006: 729).

Figure 4: Literacy test scores and matric results



Source: Own calculations. CAPS Panel.

The relationship between literacy test scores and matric results is similar, but more erratic below the mean (this is mainly because very few of the young adults who scored below the mean on the literacy test wrote the matric exam). Again, the positive relationship between literacy test scores and matric results is stronger for those young adults who obtained above average literacy test scores.

Wave 1 of CAPS contains 2,125 single young adults, 1,808 individuals in sibling pairs and 813 individuals in sibling trios; we create a matched pair sibling sample using sibling pairs (1,808) and a cluster sibling sample using sibling pairs and sibling trios (2,621 individuals).[16] The table below reports the

[16] This information is obtained from the *w1y_yatot2* variable, which gives the total number of young adults in the household, including only those who completed questionnaires. (Lam, Seekings and Sparks, 2007: 22). Combining this with *yaageorder* allows us to create: (i) 2 birth order dummies for the young adults in the matched pair sibling sample

variation within families for all measures of exposure to violence during childhood for both sibling samples.

Table 4: Variation in self-reported exposure to violence in CAPS sample and within families (unweighted)

Variation	Hit hard	Pushed	Afraid of being hurt	Adult put down
CAPS matched pair sibling sample				
Mean	0.067	0.181	0.178	0.329
Standard deviation	(0.25)	(0.386)	(0.383)	(0.47)
Number of observations	*116*	*327*	*322*	*594*
Percentage of households where young adults reported differences in exposure to violence	11%	22%	20%	31%
Standard deviation between families	(0.194)	(0.307)	(0.312)	(0.379)
Standard deviation within families	(0.163)	(0.233)	(0.222)	(0.279)
CAPS cluster sample				
Mean	0.062	0.183	0.187	0.339
Standard deviation	(0.241)	(0.386)	(0.39)	(0.474)
Number of observations	*156*	*477*	*491*	*887*
Standard deviation between families	(0.185)	(0.303)	(0.313)	(0.377)
Standard deviation within families	(0.162)	(0.243)	(0.234)	(0.287)

Source: Own calculations. CAPS Wave 1.

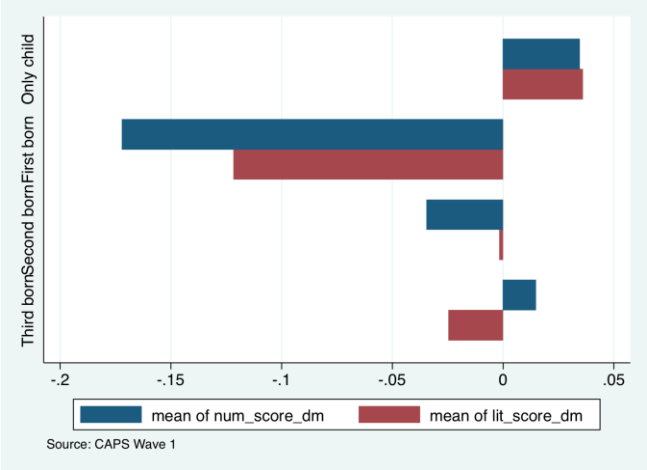
Within the CAPS matched pair sibling sample, the percentage of households where adults reported differences in exposure to violence is highest for *adult_put_down*, followed by *pushed* and *afraid_hurt* and is lowest for *hit_hard*. Variation between and within families is consistent with the variation reported for the CAPS sample: variation in exposure to violence during childhood between and within families is highest for *adult_put_down* and lowest for *hit_hard*. As expected, variation in exposure to violence is always lower within a family than between families. In other words, the difference in being exposed to violence during childhood is always higher for two young adults drawn at random from the sibling sample compared to two young adults from the same family.

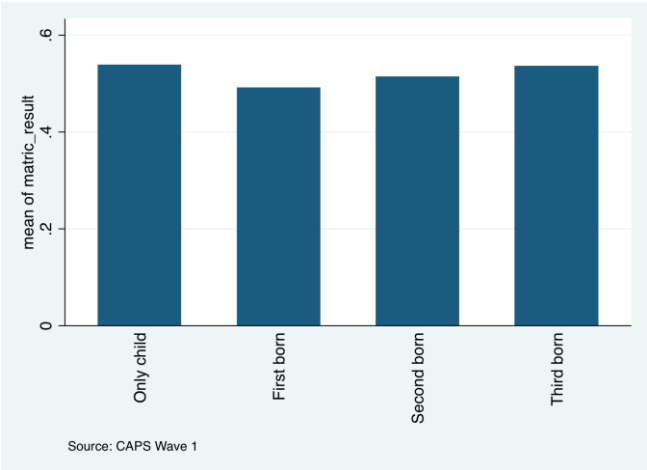
The variation in exposure to violence within families could be due to (i) observed individual heterogeneity (differences in gender, age, birth order or school environment); (ii) unobserved individual heterogeneity (it may be that one sibling has a greater propensity for behaviour deemed unacceptable by parents or siblings may have different interpretations or recollections of exposure to violence during childhood).

(*older_youngadult* and *younger_youngadult*); and (ii) 3 birth order dummies for the young adults in the cluster sibling sample (*oldest_youngadult*, *middle_youngadult* and *youngest_youngadult*).

The figures below show demeaned test scores and matric results by birth order for all the young adults in the sample.[17]

Figure 5: Test scores and matric results by birth order





Source: Own calculations. CAPS Wave 1.

Young adults from families where they are the only child generally perform better in numeracy and literacy test scores and matric results than young adults with siblings. For those young adults with siblings, being the oldest is associated with poorer academic performance compared to younger siblings. This is in contrast to a study by Hanushek (1992) that finds a U-shaped relationship between birth order and achievement: children in the earliest and latest birth positions significantly outperform those in the middle.[18]

Finally, given our emphasis on role of parental inputs in shaping the formation of cognitive and non-cognitive skills and therefore educational outcomes, we explore several measures of parental investment by race and gender below:

[17] Two variables were used to construct the 4 birth order categories used above (only child, first born, second born and third born): (i) *yaageorder* and (ii) *w1y_yatot1*. The *yaageorder* variable gives the number of the young adult in the household, including only those who completed questionnaires, and is ordered by age with *yaageorder*=1 for the youngest young adult with a completed questionnaire (Lam, Seekings and Sparks, 2007: 22). The *w1y_yatot1* variable gives the total number of young adults in the household, including those who did not complete questionnaires (Lam, Seekings and Sparks, 2007: 22).

[18] More recent studies by Black, Devereux and Salvanes (2005) and Black, Devereux and Salvanes (2007) found a relationship between birth order and educational attainment as well as birth order and IQ.

Table 5: Variable definitions and sample characteristics: parental inputs (unweighted)

Variable	Definition	Sample	Race			Gender	
			Black	Coloured	White	Male	Female
parent_inflsch	0-1 dummy that equals 1 if the young adult's biological parents were biggest influence on how well they performed in	0.753	0.758	0.73	0.812	0.756	0.751
		0.431	0.428	0.444	0.391	0.43	0.433
		4,736	*2,143*	*2,001*	*592*	*2,132*	*2,604*
parent_encrgoals	0-1 dummy that equals 1 if the young adult's biological parents gave the most encouragement towards achieving their personal goals during childhood	0.754	0.759	0.718	0.853	0.758	0.75
		0.431	0.428	0.45	0.354	0.429	0.433
		4,737	*2,145*	*1,999*	*593*	*2,134*	*2,603*
parent_helphmwrk	0-1 dummy that equals 1 if the young adult's parents helped with homework in the previous 12 months	0.207	0.128	0.222	0.445	0.222	0.195
		0.405	0.334	0.416	0.497	0.416	0.396
		4,746	*2,148*	*2,002*	*596*	*2,139*	*2,607*
parent_moneysch	0-1 dummy that equals 1 if the young adult's biological parents spent any money on their school fees, books or supplies in the previous 12 months	0.634	0.632	0.563	0.875	0.641	0.627
		0.482	0.482	0.496	0.331	0.48	0.484
		4,623	*2,062*	*1,970*	*591*	*2,076*	*2,547*
hh_ownbooks	0-1 dummy that equals 1 if the someone in the young adult's household owns more than 5 books	0.789	0.655	0.871	0.993	0.798	0.781
		0.408	0.475	0.335	0.082	0.402	0.413
		4,744	*2,147*	*2,002*	*595*	*2,138*	*2,606*

Source: Own calculations. CAPS Wave 1. Standard deviations in brackets and number of observations in italics.

The first two measures of parental inputs are based on questions that asked young adults to reflect on the adults who influenced them during childhood. About 75% of the youth in the sample say their biological parents were the biggest influence on their school performance and gave the most encouragement during their childhood. Coloured youth had slightly poorer outcomes than their black and white counterparts for both measures and the gender differences are negligible.

Just over 20% of young adults say their parents helped with their homework, 63% of the youth say their parents gave money for school and 79% of youth report that someone in their household owns more than 5 books. A small proportion of black youth report being helped in completing their homework by their parents (13%), compared to coloured (22%) and white (45%) youth. A smaller proportion of coloured youth (56%) report being given financial support for school, compared to black (63%) and white youths (88%). These measures

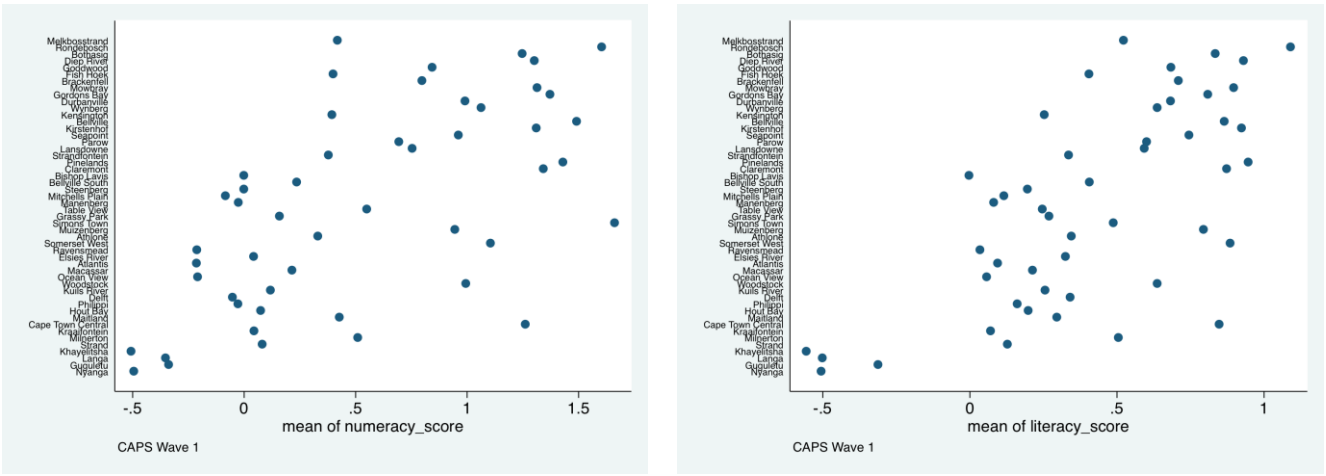
suggest that parents favour males in terms if their parental investment during childhood and adolescence.

The *hh_ownbooks* variables illustrates the disadvantage faced by black youth in terms of access to books at home: 66% of black youth report that someone in their household owns more than 5 books, compared to 87% of coloured youth and 99% of white youth.

Descriptive analysis

First, we explore test scores, dropout and matric results by neighbourhood (neighbourhoods are sorted by 2003 – 2004 murder rates, where neighbourhoods with the highest murder rates are at the bottom). The murder rates for Cape Town were obtained from the official crime statistics released by the South African Police Service (SAPS) through their website.[19] These murder statistics were matched to the CAPS dataset for each police precinct in Cape Town.

Figure 6: Test scores across CAPS neighbourhoods



Source: Own calculations. CAPS Wave 1.

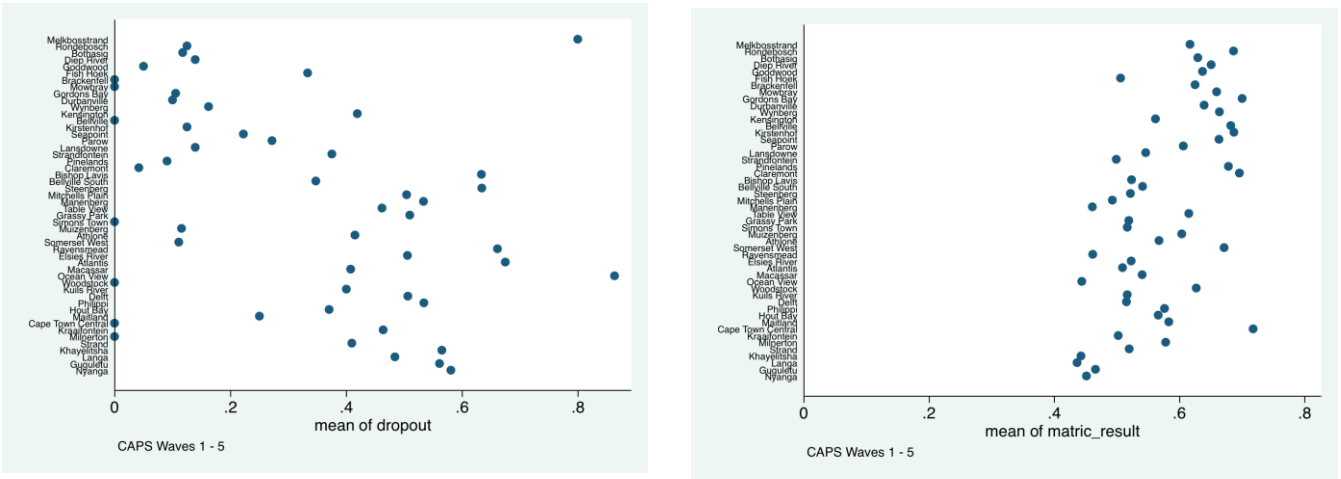
The graph in the first panel shows an inverse relationship between numeracy test scores and murder rate: in neighbourhoods where the murder rate is high (bottom left), numeracy scores are low and in neighbourhoods where the murder rate is low (top right), numeracy scores are high. The same relationship is shown for literacy test scores. In both cases the bottom four neighbourhoods are almost

[19] The official violent crime statistics were obtained from http://www.saps.gov.za/statistics/reports/crimestats/2011/crime_stats.htm. Appendix F contains details of how the murder statistics were matched to the CAPS dataset.

100% black and there is a clear difference between high murder neighbourhoods where a large proportion of the black youth reside and the other neighbourhoods in Cape Town.

The graphs below show dropout and matric results for each young adult against the neighbourhoods where they resided during Wave 1. In the first panel, there is a slight positive relationship between mean dropout at the neighbourhood level and murder rate: the neighbourhoods with higher murder rates (towards the bottom of the x-axis) have higher dropout rates. In the second panel, there is a clear inverse relationship between matric results and neighbourhood violence.

Figure 7: Dropout and matric results across CAPS neighbourhoods



Source: Own calculations. CAPS Waves 1 – 5.

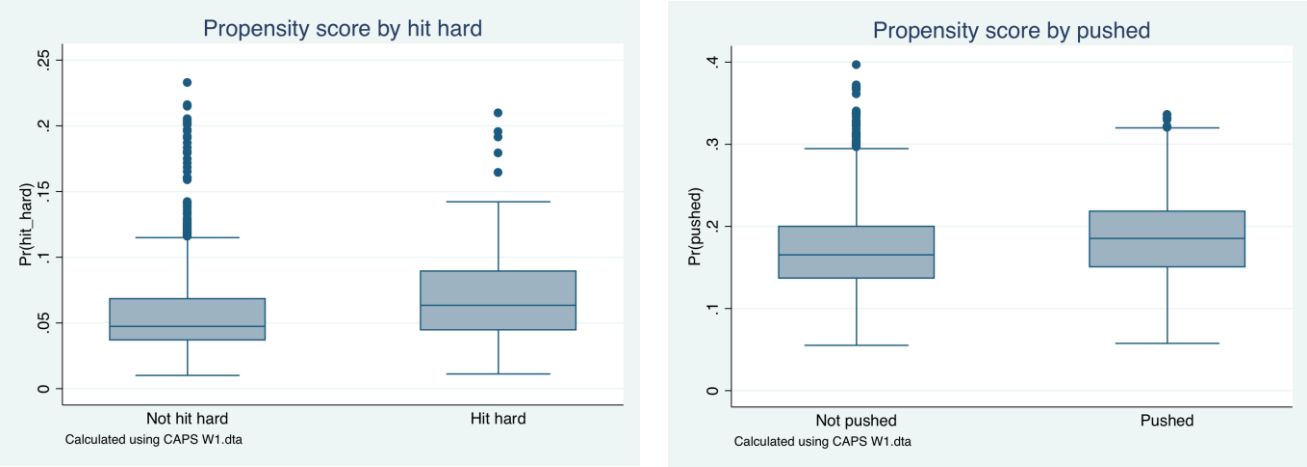
Second, we test the extent of "selection into exposure to violence bias" for each measure of exposure to violence during childhood. Preliminary findings from a separate descriptive chapter suggest that other forms of disadvantage, not simply exposure to violence, may be responsible for adverse consequences associated with exposure to violence. We can test the extent of "selection into exposure to violence bias" for each measure of exposure to violence during childhood using a graphical analysis associated with propensity score matching techniques.

We calculate the probability of exposure to violence during childhood conditional on the observables (all the variables in Table 1 except *white*, *mother_deceased* and *log_hh_income*). This is known as the propensity score index and since exposure to violence during childhood is binary, both the mean and the variance of exposure to violence during childhood conditional on the observables is determined by the propensity score index (which makes it a sufficient statistic for the relationship between exposure to violence during childhood and the explanatory variables used). In other words, the propensity score index reflects the "selection into exposure to violence bias" which is used

to compare how similar the exposed and non-exposed youth are in terms of their observable characteristics.[20]
The propensity score index is used to stratify the CAPS sample into 5 blocks.[21] In each of the blocks we want young adults that, based on their observables (as captured by the propensity score index), have similar probabilities of being exposed to violence during childhood, but some experienced violence and others did not.

A plot of the distribution of propensity scores by exposure to violence gives a sense of the similarity of observables for young adults that were exposed to violence and those that were not. If there is very little overlap between the distributions below (the box-plots are not in line with each other), it suggests that those who are exposed to violence and those who are not have very different observed characteristics and we are comparing "non-comparable" groups. Substantial overlap in the box-plots indicates very little "selection into exposure to violence" bias and means we have a good research design because the observables from Table 1 ensure sufficient overlap between treatment and control groups.

Figure 8: Distribution by treatment and control: hit hard and pushed



Source: Own calculations. CAPS Wave 1.

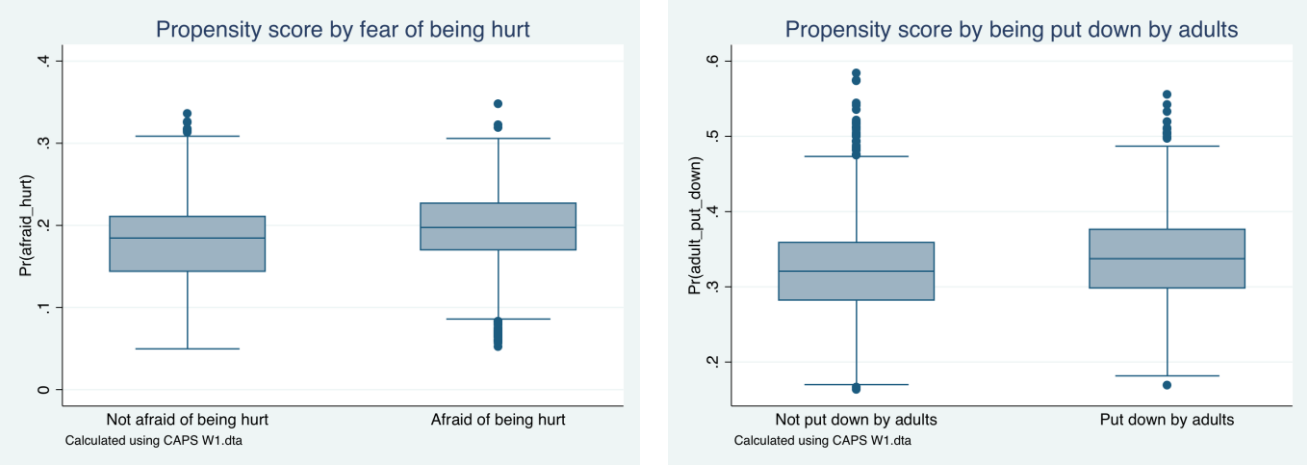
In the graphs above we illustrate the distributions of the propensity score by exposure to violence during childhood (exposed versus not exposed). This gives an indication of the "non-randomness" of exposure to violence in the CAPS

[20] The propensity score index is the predicted values from a regression of exposure to violence during childhood on all the observed characteristics of the young adults in the CAPS sample (the variables from Table 1 except *white*, *mother_deceased* and *log_hh_income*).
[21] The choice of 5 blocks is completely arbitrary and we ensure that our 5 blocks are "balanced" – that is, the sample means for each of the observed characteristics are equal for young adults that were exposed to violence during childhood and those that were not.

sample. Although there is some overlap for both *hit_hard* and *pushed*, there is evidence of more overlap for *pushed*. This suggests that for the *pushed* measure of exposure to violence during childhood, there is less selection into exposure to violence.

The findings are similar for *afraid_hurt* and *adult_put_down*: there is substantial overlap between young adults who were exposed to violence and those who were not, which means that the observables used ensure sufficient overlap between treatment and control groups for these measures.

Figure 9: Distribution by treatment and control: afraid of being hurt and adult put down

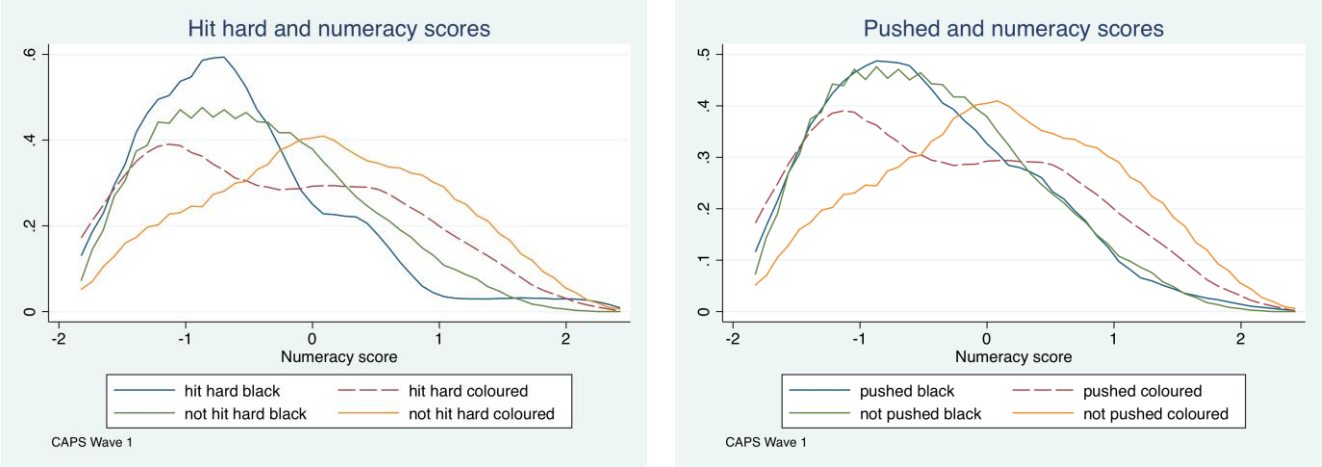


Source: Own calculations. CAPS Wave 1.

Finally, we examine the relationship between the continuous educational outcomes measures (numeracy test scores and matric result) and exposure to violence during childhood for coloured and black youth in Cape Town. We use kernel density functions, which is a non-parametric way of estimating the probability density function of test scores.[22]

[22] These kernel density functions are estimated in STATA using the *kdensity* command and the default settings were maintained (including the default epanechnikov kernel and optimal bandwidth procedure). Kernel density functions assume a lognormal distribution function, which is appropriate for the distribution of the normalised test scores.

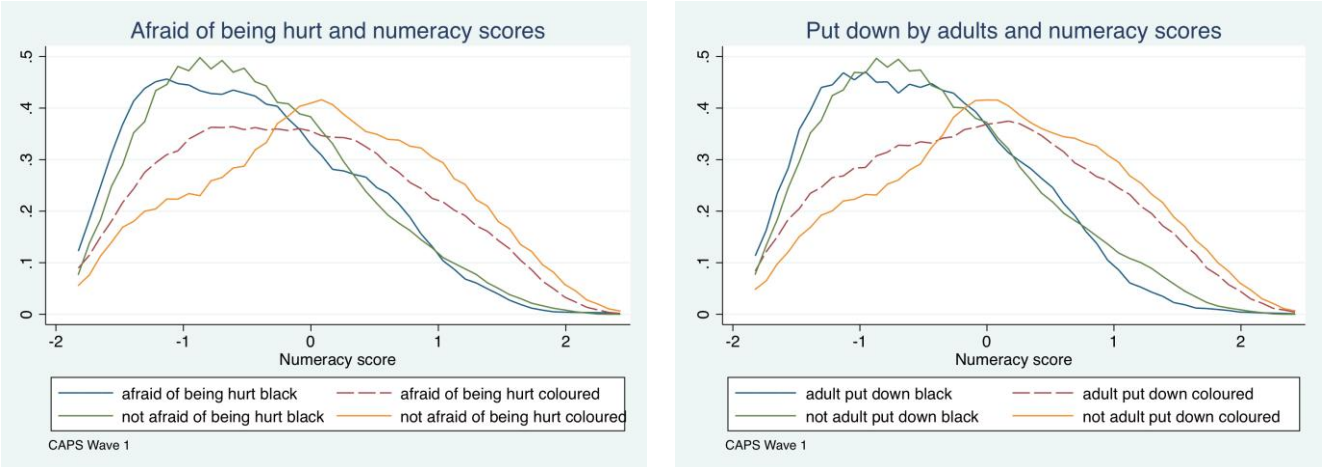
Figure 10: Hit hard and pushed during childhood and numeracy test scores



Source: Own calculations. CAPS Wave 1.

There is a difference in the numeracy test score distributions for coloured and black youth that were hit hard during childhood, but the difference is more striking for coloured youth (above left). There is a difference in the numeracy test score distributions of coloured youth that were pushed during childhood compared to those that were not (above right).

Figure 11: Afraid of being hurt and adult put down during childhood and numeracy test scores



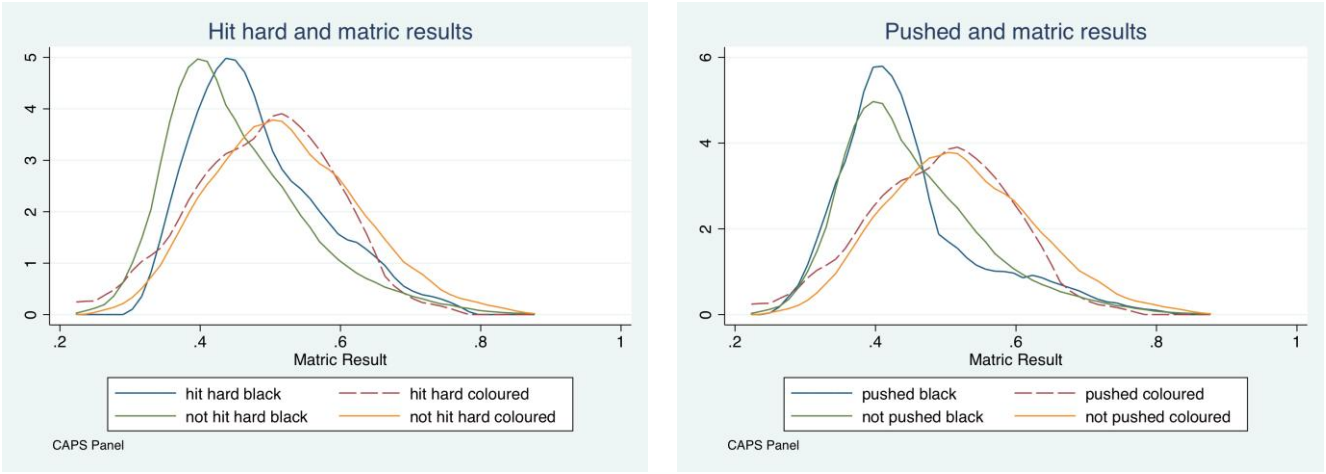
Source: Own calculations. CAPS Wave 1.

There is a difference in the numeracy test score distributions for coloured and black youth that were afraid of being hurt during childhood, but the difference is more striking for coloured youth (above left). There is a difference in numeracy test score distributions of coloured youth that were put down by adults during childhood compared to those that were not (above right). The same analysis was conducted for literacy test scores and there was no difference between the

literacy test score distributions for young adults that were exposed to violence during childhood using all four measures of exposure to violence. This finding is in contrast to results from Ludwig et al. (2010) who show a negative relationship between violent crime and children's test scores and find that the results are higher for literacy (reading) compared to numeracy (math) scores.

Second, we examine the relationship between matric results and exposure to violence during childhood for coloured and black youth in Cape Town. The graphs below indicate that matric results for black youth are much lower their coloured counterparts and the variation in matric results among black youth is much lower than for coloured youth.

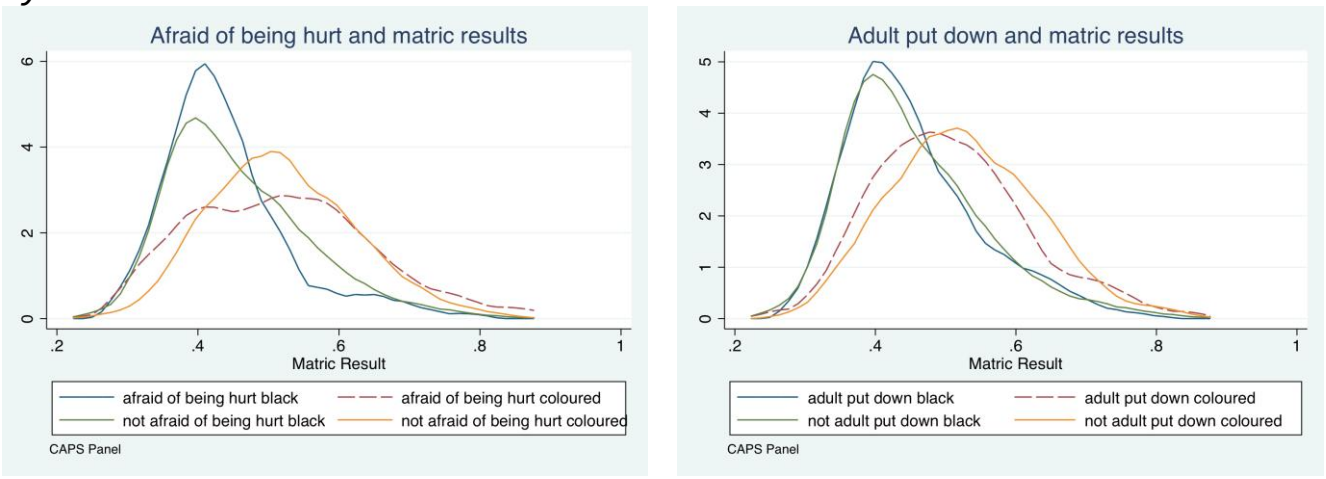
Figure 12: Hit hard and pushed during childhood and matric results: by race



Source: Own calculations. CAPS Panel.

The graph in the first panel illustrates an unexpected result: matric results are lower for black youth that were not hit hard in comparison to those that were. This finding contradicts earlier results that showed lower numeracy test scores for black youth that were hit hard during childhood. This finding, combined with Table 2 that showed only 4.6% of black youth in the CAPS sample were hit hard during childhood (and indicate significant variation within black youth), suggests that we should use the *hit_hard* measure with caution. The graph in the second panel shows no relationship between being pushed and matric results for black or coloured youth.

Figure 13: Fear of being hurt and put down by adults and matric results: by race



Source: Own calculations. CAPS Panel.

Although there is no clear relationship between the fear of being hurt during childhood and matric results for black or coloured youth, the matric results of coloured youth that were put down by adults during childhood are lower than those who were not (above right).

The graphs suggest an association between exposure to violence during childhood and educational outcomes that is stronger for coloured youth. In the regression analysis we test the robustness of these findings in a multivariate setting.

Empirical strategy

The empirical strategy used this paper is closely related to reduced-form literature on the determinants of educational outcomes (see Glewwe and Kremer, 2006 and Björklund and Salvanes, 2011). In this literature, educational outcomes are determined by individual characteristics, family background and neighbourhood characteristics:

$$\boldsymbol{Y} = \beta_0 + \beta_1 \boldsymbol{I} + \beta_2 \boldsymbol{F} + \beta_3 \boldsymbol{D}_N + \beta_4 \boldsymbol{V} + \boldsymbol{\varepsilon} \qquad (1)$$

Where $\boldsymbol{Y}$ is a measure of educational outcomes (test scores, matric results or dropout), $\boldsymbol{I}$ is a vector of individual characteristics (such as age and race), $\boldsymbol{F}$ is a vector of family characteristics (such as home language and household income), $\boldsymbol{D}_N$ is a vector of neighbourhood dummies and $\boldsymbol{V}$ is a measure of exposure to

violence during childhood.[23] The error term, $\boldsymbol{\varepsilon}$, can be decomposed into an individual specific error term that reflects unobserved individual characteristics, a family-specific error term that reflects unobserved family background such as parental competence, a neighbourhood-specific error term that reflects unobserved neighbourhood characteristics such as social cohesion and a random error term.

It is possible that collective efficacy, defined as social cohesion among neighbours combined with their willingness to intervene on behalf of the common good (see Sampson et al., 1999), is linked to reduced violence (if they report their neighbours who are perpetrating violence against their own children) and improved educational outcomes in a neighbourhood.[24] Therefore, we add a vector of neighbourhood dummies, $\boldsymbol{D_N}$, to equation 1 to control for unmeasured neighbourhood characteristics that may be correlated to exposure to violence during childhood and educational outcomes.

We can use equation 1 to obtain β_4, the association between exposure to violence and educational outcomes. However, the estimate of β_4 will be biased if: (i) young adults are selected into exposure to violence during childhood based on unobserved factors (omitted variables bias); (ii) households sort themselves into neighbourhoods in a way that affects both exposure to violence during childhood and educational outcomes (selection bias); and (iii) educational outcomes or exposure to violence during childhood is measured with error (measurement error bias).[25]

First, the estimate of β_4 will be biased if young adults are selected into exposure to violence during childhood based on unobserved factors; in other words, unobserved factors influence both educational outcomes and exposure to

[23] Glewwe and Kremer (2006: 14) include an additional variable ***EP*** (education policies that affect educational outcomes by changing what happens in the classroom) in their reduced form equation. They expect ***EP*** to interact with ***N*** (neighbourhood) to determine school quality.

[24] Each South African school has a School Governing Body (SGB) that governs the management of the school and is responsible for the school's finances. The SGB consists of the school principal, parents, teachers, senior students and other community members. For example, SGBs are able to raise additional funds (through deciding on higher fees or fundraising initiatives) that can be used to appoint SGB-paid teachers in addition to the government-paid teachers to reduce pupil-teacher ratios.

[25] The estimate of may be biased β_4 in other ways: (i) it is possible that academic performance or educational attainment affect how young adults reflect on their exposure to violence during childhood (simultaneity bias); and (ii) the relationship between exposure to violence and educational outcomes may not be linear and additive in which case a linear specification may be the incorrect functional form of the conditional expectation (misspecification bias).

violence during childhood. The unobserved factors that conceivably drive both exposure to violence during childhood and educational outcomes during adolescence may be at the individual level (delinquency) or household level (parenting style). For example, if a child has a greater propensity for delinquent behaviour that results in poor educational outcomes and exposure to violence by parents, then the estimate of β_4 would be biased unless selection into this behaviour is observed. Alternatively, if parents who practice corporal punishment when their children are young also fail to provide the necessary financial or other support during high school, then unobserved parental competence would bias the estimate of β_4.

Second, parents can influence the quality of a school (see SGB discussion in an earlier footnote) and all parents can choose from more than one school inside or outside their neighbourhood or can choose to move to a neighbourhood with better quality schools. In this case, because parents can affect school quality, selection bias is possible if neighbourhood sorting takes place in a way that affects educational outcomes in high school (through parental involvement, for example) and determines exposure to violence during childhood.

Third, the estimate of β_4 will be biased if educational outcomes or exposure to violence during childhood is measured with error. Measurement error in educational outcomes ($\boldsymbol{Y}$) would bias the estimate of β_4 if it were correlated with exposure to violence during childhood. For example, if the educational outcomes of young adults who were exposed to violence are systematically mismeasured compared to those that were not exposed to violence. Random measurement error in exposure to violence during childhood would lead to attenuation bias if the estimate of β_4 were reduced due to reporting error in exposure to violence (in other words, the true variation is reduced by the variation that is due to reporting bias).

Producing empirically valid estimates of the relationship between exposure to violence during childhood and educational outcomes is challenging because exposure to violence during childhood cannot be reproduced in an experimental setting and exposure to violence during childhood is not randomly distributed across households in Cape Town. Several strategies have been used throughout the economics literature to address the biases described above: (i) controls for observables; (ii) using area-level variables as instruments and exploiting variation between areas (see Evans, Oates and Schwab, 1992, Cutler and Glaeser, 1997, Card and Rothstein, 2007, Weinberg, 2004)[26]; (iii) individual,

[26] These studies do not always convincingly satisfy the exclusion criterion for the area-level variables used.

household, school and grade fixed effects (see Meghir and Rivkin, 2010); and (iv) social experiments or quasi experimental data (see Kling et al., 2007, Sacerdote, 2001 and Zimmerman 2003)[27].

To address unobserved differences between young adults due to fixed family background we conduct sibling fixed effect regressions. In the sibling fixed effect regressions we add a household dummy, $\mathbf{D}_F$, to equation 1 to control for fixed and unmeasured family background measures (such as parental style) that may be correlated to both exposure to violence during childhood and educational outcomes. The sibling fixed effect regressions rely on constant differences in exposure to violence during childhood between siblings in the same household.

The sibling fixed effect regressions are unable to address: (i) differences between siblings due to unobserved transitory family shocks (to a parent's employment status, for example) that affect both exposure to violence during childhood and educational outcomes later in life for one sibling and not the other; and (ii) individual behavioural differences between siblings that result in one sibling being exposed to violence during childhood in a way that affects their educational outcomes later in life.[28]

School fixed effects are not performed because they are problematic for four reasons: (i) school fixed effects are only useful if the exposure to violence took place during high school, which is unlikely since young adults are asked to reflect on their exposure to violence during childhood (unless individuals misunderstood the question); (ii) excluding all individuals that are the only representatives from their school in the CAPS sample would induce selection bias if these individuals share common background characteristics; (iii) for some schools we have very few observations and would therefore need to infer exposure to violence at the school on the basis of a few observations; and (iv) school fixed effects may accentuate the omitted variables bias problem if teachers sort young adults into classes on the basis of their exposure to violence during childhood. Also, if there is little variation in school quality between siblings, then the sibling fixed effect regressions help address some of the unobserved school heterogeneity.

[27] See Moffitt (2001) and Glewwe and Kremer (2006) for a critical review of the identification strategies employed by these studies.

[28] Sibling fixed effects are also unable to address unobserved differences in the school environment between siblings that result in one sibling being exposed to violence during childhood in a way that affects their educational outcomes later in life. This takes places if a teacher can influence whether one sibling is exposed to violence and not another (for example, by encouraging parents to practice corporal punishment with one delinquent child).

We use ordinary least squares (OLS) when the measure of educational outcomes if continuous (numeracy test scores and matric results) and a linear probability model (LPM) when the measure of educational outcomes is binary (such as dropout). The LPM is used for three reasons: (i) theoretical support for the use of LPMs when the outcome variable is binary is provided by Angrist and Pischke (2009); (ii) LPMs perform well in models that estimate fixed effects (see Angrist, 2001); and (iii) several related studies have used LPMs to estimate similar outcomes (see, for example, Aaronson, 1998). The trade-off associated with using LPMs is that the predicted probabilities are not always bound between 0 and 1.

Regression analysis

This section contains regressions of equation 1 using different measures of educational outcomes and exposure to violence during childhood. For reasons discussed in the previous section, these estimates cannot be used to draw conclusions about a causal relationship between educational outcomes and exposure to violence during childhood since there are likely to be other unobserved factors that affect exposure to violence during childhood and educational outcomes. All the regressions in this section use the unweighted CAPS sample.[29] All regressions report robust cluster corrected standard errors to correct for heteroskedasticity (due to correlation between the variance of unobserved factors at the neighbourhood level) and autocorrelation (due to correlation between unobserved factors within neighbourhoods).

First, we estimate neighbourhood fixed effect regressions of the relationship between exposure to violence during childhood and numeracy test scores based on equation 1 specified in the previous section.

Since these are OLS estimates, the coefficients measure the change in numeracy test scores with respect to each explanatory variable holding the other variables fixed. In the table below, being black is associated with a reduction in numeracy test scores of about 1.1 (more than one standard deviation) and being coloured is associated with a reduction in numeracy scores of about 0.8 relative to the omitted category of white. These are significant reductions given that numeracy test scores are normalised with mean zero and range between -1.7 and 2.3 for the CAPS sample. Living in a household that receives a government grant is

[29] Baseline regressions of the relationship between educational outcomes and exposure to violence during childhood are given in Appendix A and weighted regressions of equation 1 are given in Appendix B.

associated with a reduction in test scores of about 0.2, having English as a home language is associated with an increase of 0.4 in numeracy test scores and living in a female headed household is associated with a reduction of 0.07 in numeracy test scores. Living in a household where someone owns more than 5 books is associated with a 0.2 increase in numeracy test scores.

Table 6: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
age	0.0550***	0.0548***	0.0545***	0.0544***	0.0545***
	(0.00492)	(0.00477)	(0.00473)	(0.00468)	(0.00481)
male	0.0358	0.0368	0.0341	0.0321	0.0344
	(0.0415)	(0.0415)	(0.0396)	(0.0385)	(0.0401)
black	-1.069***	-1.070***	-1.062***	-1.063***	-1.055***
	(0.126)	(0.126)	(0.126)	(0.127)	(0.128)
coloured	-0.799***	-0.798***	-0.793***	-0.789***	-0.792***
	(0.0863)	(0.0865)	(0.0867)	(0.0867)	(0.0860)
grant	-0.182***	-0.188***	-0.202***	-0.204***	-0.185***
	(0.0566)	(0.0554)	(0.0560)	(0.0563)	(0.0573)
mother_education	0.0517***	0.0516***	0.0518***	0.0516***	0.0517***
	(0.00821)	(0.00815)	(0.00830)	(0.00820)	(0.00820)
home_lang_english	0.385***	0.387***	0.388***	0.385***	0.386***
	(0.0413)	(0.0417)	(0.0417)	(0.0417)	(0.0413)
female_hh	-0.0763**	-0.0749**	-0.0749**	-0.0761**	-0.0738**
	(0.0344)	(0.0336)	(0.0334)	(0.0341)	(0.0339)
hh_ownbooks	0.199***	0.202***	0.202***	0.199***	0.191***
	(0.0423)	(0.0410)	(0.0397)	(0.0413)	(0.0404)
Violence measure	-0.159***	-0.0654*	-0.0768**	-0.100***	-0.171***
	(0.0582)	(0.0339)	(0.0341)	(0.0230)	(0.0408)
Constant	-0.677***	-0.679***	-0.683***	-0.664***	-0.653***
	(0.186)	(0.177)	(0.179)	(0.174)	(0.179)
Observations	4,139	4,134	4,144	4,138	4,123
R-squared	0.430	0.431	0.429	0.431	0.432

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

All the measures of exposure to violence during childhood are significant and show a negative sign. The largest reduction in numeracy test scores is associated with *hit_hard* and the aggregate measure (*childhd_exposure*) while the smallest reduction in numeracy test scores is associated with *pushed* (0.07). The same regressions were performed for literacy test scores and the results were consistent with the kernel density functions in the descriptive analysis: with the exception of the aggregate measure (*childhd_exposure*), none of the exposure to violence during childhood measures was significant.

In Appendix C we repeat the above regressions, but instead of specifying the exposure to violence during childhood measures as binary, we create 5 dummies

from the full five-point scale. In all cases the omitted category is "never exposed to violence during childhood". These regressions show that the strongest association between exposure to violence during childhood and numeracy test scores occurs for those children that were exposed to violence often (relative to the omitted category of "never"): their numeracy test scores are lowered by between 0.22 (for *pushed* and *afraid_hurt*) and 0.45 (for *hit_hard*). For *adult_put_down*, the largest reduction in numeracy test scores is associated with children that were put down by adults very often, relative to the omitted category of never.

In Appendix D, we repeat the above regressions and add a measure of school quality. We are concerned that the neighbourhood fixed effect regressions conducted here are unable to account for differences between young people from the same neighbourhood who attend schools of different quality. The *school_quality* variable is a 0-1 dummy that equals 1 if the young adult attended a "previous disadvantaged" school.[30] The coefficients on the exposure to violence measures remain roughly the same and the school quality dummy is always significant and the negative effect is large. Attending a previously disadvantaged reduces numeracy test scores by 0.2 (20% of a standard deviation), which is a larger effect than exposure to violence and similar in magnitude to the welfare effect (as measured by whether the household receives a government grant) and similar in magnitude to whether someone in the house owns more than 5 books. However, the negative effect is still substantially smaller than the race effect and since these are neighbourhood fixed effect regressions, it suggests alot of heterogeneity within neighbourhoods between young adults from the same race that is not reduced once school quality is taken into account.

Second, we estimate LPM regressions of the relationship between exposure to violence during childhood and dropout based on equation 1. Since these are LPM estimates, the coefficients measure the change in the probability of dropout when one regressor changes, holding other factors fixed. These are neighbourhood fixed effect regressions since all specifications include neighbourhood dummies for 49 neighbourhoods in Cape Town based on the police precinct boundaries.

[30] A school is considered "previously disadvantaged" if it fell under the administration of a previously black department of education. In the CAPS sample, this includes the Department of Education and Training, Bophuthatswana Education Department, Ciskei Education Department, KwaZulu Department of Education and Culture, Lebowa Department of Education and Transkei Education Department. Since we are missing this information for about 422 young adults, we do not have the school quality measure for the full sample.

Table 7: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
male	0.130***	0.129***	0.131***	0.130***	0.130***
	(0.0222)	(0.0224)	(0.0222)	(0.0226)	(0.0224)
black	0.0451	0.0473	0.0346	0.0410	0.0419
	(0.0634)	(0.0622)	(0.0622)	(0.0646)	(0.0630)
coloured	0.164***	0.164***	0.157***	0.161***	0.157***
	(0.0426)	(0.0427)	(0.0429)	(0.0434)	(0.0448)
grant	0.236***	0.242***	0.238***	0.241***	0.238***
	(0.0184)	(0.0182)	(0.0175)	(0.0182)	(0.0178)
mother_education	-0.0274***	-0.0271***	-0.0274***	-0.0272***	-0.0271***
	(0.00225)	(0.00214)	(0.00223)	(0.00208)	(0.00211)
home_lang_english	-0.186***	-0.189***	-0.189***	-0.192***	-0.190***
	(0.0247)	(0.0242)	(0.0237)	(0.0239)	(0.0240)
female_hh	0.0179	0.0154	0.0150	0.0161	0.0146
	(0.0411)	(0.0408)	(0.0400)	(0.0405)	(0.0412)
hh_ownbooks	-0.147***	-0.151***	-0.146***	-0.152***	-0.144***
	(0.0253)	(0.0254)	(0.0264)	(0.0251)	(0.0266)
Violence measure	0.133***	0.0477***	0.0784***	0.0319*	0.115***
	(0.0384)	(0.0171)	(0.0244)	(0.0164)	(0.0298)
Constant	0.602***	0.605***	0.609***	0.616***	0.598***
	(0.0459)	(0.0443)	(0.0455)	(0.0457)	(0.0468)
Observations	3,088	3,083	3,093	3,087	3,075
R-squared	0.247	0.245	0.247	0.245	0.247

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

The coefficient for *male* means that, all other factors being equal, being male is associated with a 0.13 increase in the probability of dropout compared to females. Being coloured is associated with a 0.16 increase in the probability of dropout (compared to white, the omitted category). This is consistent with other studies on South Africa that find coloured youth show the earliest trend towards dropout before the completion of high school (Ministerial Committee on Learner Retention in the South African Schooling System, 2007). This is confirmed by Lam, Ardington and Leibbrandt (2006) who use data on 1,500 students from Cape Town who were enrolled in grades 8, 9, and 10 in 2002 to check whether they reached grade 11 and grade 12 and completed grade 12 by 2005. Their results indicate that coloured youth who were in grade 8 in 2002 were less likely than blacks to be enrolled in 2005, but those who were enrolled were more likely to have maintained normal progress through school (Lam, Ardington and Leibbrandt, 2006: 10).

Receiving a government grant and an increase in household size are associated with an increase in the probability of dropout. Having English as home language, living in a household where someone owns more than 5 books and

increases in mother's education are associated with a reduction in the probability of dropout.

According to these regressions, the factor that is associated with the biggest change in probability of dropout is whether the young adult lives in a household that receives a government grant. This is consistent with the idea that household income matters for educational attainment because of the costs associated with attending school and the pressure on young adults from poor households to drop out of school to take care of siblings or earn money to support the household. This finding is in contrast to the regressions on numeracy test scores where race indicators were associated with the largest reductions in numeracy test scores.

All the measures of exposure to violence during childhood are significant, show the expected positive signs and are consistent with the findings from the numeracy test scores. The largest increase in probability of dropout is associated with *hit_hard* and the smallest increase in probability of dropout is associated with *afraid_being_hurt*.

In Appendix C we repeat the dropout regressions with the full five-point scale of exposure to violence. The strongest association between exposure to violence during childhood and dropout occurs for those children that were exposed to violence often (relative to the omitted category of "never"): their probability of dropout increases by between 0.13 (for *pushed*) and 0.19 (for *hit_hard*). For *afraid_hurt* and *adult_put_down*, the largest increase in the probability of dropout is associated with children that that were exposed to violence sometimes, relative to the omitted category of never.

Third, we estimate neighbourhood fixed effect regressions of the relationship between exposure to violence during childhood and matric results based on equation 1. Being black is associated with a reduction in matric results of about 11% and being coloured is associated with a reduction in matric results of about 9% relative to the omitted category of white. These are significant reductions given that a change in 10% will always result in an aggregate symbol change (which could be the difference between obtaining university entrance or not). Living in a household that receives a government grant is associated with a reduction in matric results of 2%, having English as a home language is associated with an increase of 4% and living in female-headed household is associated with a reduction of 2%.

The *adult_put_down* and *childhd_exposure* measures are the only ones associated with a decrease in matric results. The reduction in matric results associated with *adult_put_down* is similar in magnitude to living in a female-

headed household. As mentioned before, these findings are subject to selection bias since we do not observe the matric results for young adults that were exposed to violence, but did not to write the matric exam.

Table 8: Exposure to violence during childhood (dependent variable: matric_result)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
male	-0.00938	-0.00917	-0.00964	-0.00989	-0.00935
	(0.00972)	(0.00961)	(0.00987)	(0.00974)	(0.00967)
black	-0.106***	-0.103***	-0.105***	-0.105***	-0.102***
	(0.0183)	(0.0179)	(0.0183)	(0.0180)	(0.0178)
coloured	-0.0887***	-0.0892***	-0.0882***	-0.0873***	-0.0886***
	(0.0205)	(0.0204)	(0.0205)	(0.0206)	(0.0205)
grant	-0.0214***	-0.0224***	-0.0214***	-0.0219***	-0.0223***
	(0.00769)	(0.00739)	(0.00760)	(0.00768)	(0.00737)
mother_education	0.00798***	0.00801***	0.00804***	0.00805***	0.00808***
	(0.00257)	(0.00255)	(0.00260)	(0.00256)	(0.00259)
home_lang_english	0.0362***	0.0364***	0.0363***	0.0354***	0.0361***
	(0.00834)	(0.00821)	(0.00819)	(0.00817)	(0.00828)
female_hh	-0.0155**	-0.0159**	-0.0157**	-0.0155**	-0.0154**
	(0.00706)	(0.00686)	(0.00716)	(0.00703)	(0.00685)
hh_ownbooks	0.000303	0.000187	3.53e-05	-0.000139	-0.000364
	(0.00935)	(0.00932)	(0.00927)	(0.00909)	(0.00927)
Violence measure	-0.00606	-0.00749	-0.00414	-0.0165**	-0.0210*
	(0.0153)	(0.00724)	(0.00750)	(0.00672)	(0.0114)
Constant	0.555***	0.556***	0.555***	0.556***	0.556***
	(0.0285)	(0.0284)	(0.0287)	(0.0285)	(0.0286)
Observations	1,331	1,329	1,332	1,331	1,326
R-squared	0.488	0.488	0.488	0.491	0.490

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

The final step in the regression analysis is to conduct sibling fixed effect regressions to control for fixed observed and unobserved household and family background characteristics that affect exposure to violence and educational outcomes. We conduct separate sibling fixed effect regressions for the matched pair sibling sample (shown in Appendix E) and the cluster sample. The cluster sample includes the 1,808 individuals that form the matched pair siblings as well as the 813 individuals that are part of the sibling trios in the CAPS sample.[31]

Once we control for fixed unobserved household and family background characteristics that affect exposure to violence and numeracy test scores, the effect disappears for two of the four measures of exposure to violence during

[31] When we apply sibling fixed effects to the cluster sample, it is equivalent to a regression of the intrapair difference in numeracy test scores on the intrapair difference in exposure to violence during childhood for the matched pairs and subtracting the mean from the family (cluster) to remove fixed family background effects for the sibling trios.

childhood. Being hit hard is associated with a reduction of 0.1 in numeracy test scores (compared to the baseline reduction of 0.34 and a reduction of 0.16 in the neighbourhood fixed effect regressions). Being put down by adults is associated with a 0.09 decline in numeracy test scores (less than one tenth of a standard deviation), which is a decline from 0.10 in the neighbourhood fixed effect regressions. The fact that the association between exposure to violence during childhood disappears for two measures, suggests that the bias in the neighbourhood fixed effect regressions are not due to measurement error if we expect the measurement error to be consistent for all measures of exposure to violence.

Table 9: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0301***	0.0298***	0.0293***	0.0293***
	(0.00675)	(0.00675)	(0.00625)	(0.00621)
male	0.0365	0.0381	0.0365	0.0340
	(0.0494)	(0.0482)	(0.0456)	(0.0442)
Violence measure	-0.101*	-0.0366	-0.0597	-0.0900**
	(0.0574)	(0.0362)	(0.0473)	(0.0355)
Constant	-0.589***	-0.583***	-0.570***	-0.551***
	(0.133)	(0.130)	(0.125)	(0.120)
Observations	2,584	2,581	2,587	2,583
R-squared	0.019	0.018	0.018	0.021
Number of households	1,171	1,172	1,172	1,172

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 cluster sibling sample. Robust and cluster corrected standard errors in brackets.

Below we repeat the sibling fixed effect regressions, but instead of specifying the exposure to violence during childhood measures as binary, we use the full five-point scale. In all cases the omitted category is "never exposed to violence during childhood". A fear of being hurt often is associated with a decline in numeracy test scores of 0.16 (relative to the omitted category of "never") and being put down by an adult often is associated with a reduction in numeracy test scores of 0.21.

Table 10: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0303***	0.0299***	0.0302***	0.0315***
	(0.00684)	(0.00658)	(0.00666)	(0.00691)
male	0.0378	0.0422	0.0397	0.0348
	(0.0485)	(0.0467)	(0.0465)	(0.0447)
_once	0.0484	0.0242	0.0674	0.0492
	(0.0742)	(0.0643)	(0.0580)	(0.0469)
_sometimes	-0.0871	-0.0235	-0.0308	-0.0475
	(0.0601)	(0.0409)	(0.0523)	(0.0383)
_often	-0.101	-0.0406	-0.157*	-0.214**
	(0.131)	(0.124)	(0.0927)	(0.0869)
_veryoften	-0.181	-0.108	-0.101	-0.200*
	(0.173)	(0.196)	(0.129)	(0.107)
Constant	-0.596***	-0.592***	-0.594***	-0.599***
	(0.131)	(0.131)	(0.134)	(0.136)
Observations	2,589	2,589	2,589	2,589
R-squared	0.020	0.019	0.021	0.027
Number of households	1,172	1,172	1,172	1,172

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 cluster sibling sample. Robust and cluster corrected standard errors in brackets.

As mentioned before, the sibling fixed effect regressions conducted above are unable to address: (i) differences between siblings due to unobserved transitory family shocks; and (ii) individual behavioural differences between siblings. If these unobserved transitory shocks or individual behavioural differences are related to birth order, then including birth order effects may account for some of these unobserved differences between siblings. We add dummies for birth order and interactions between the birth order dummies and exposure to violence measures to check the robustness of the sibling fixed effect regressions. If the two remaining significant measures of exposure to violence during childhood become insignificant, this suggests that unobserved differences between siblings may be driving these results.

The inclusion of birth order effects (shown in Appendix F) increases the coefficient on the *hit_hard* measure, but has no effect on the *adult_put_down* measure. The fact that the coefficient on the *adult_put_down* measure stays the same and earlier evidence that it may be the most robust measure of exposure to violence during childhood provides further evidence of a strong association between being verbally abused, insulted or put down by adults inside the home and numeracy test scores.

Table 11: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
male	0.0783***	0.0742***	0.0778***	0.0756***
	(0.0201)	(0.0212)	(0.0214)	(0.0210)
Violence measure	0.116**	0.0329	0.0796*	0.00239
	(0.0445)	(0.0302)	(0.0410)	(0.0369)
Constant	0.452***	0.454***	0.444***	0.459***
	(0.00762)	(0.0121)	(0.00786)	(0.0119)
Observations	2,139	2,136	2,141	2,137
R-squared	0.013	0.009	0.012	0.009
Number of households	1,087	1,086	1,087	1,085

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Above we repeat the sibling fixed effect regressions for the measure of educational attainment, dropout. In the neighbourhood fixed effect regressions, all the measures of exposure to violence during childhood were significantly associated with dropout. Once we control for fixed unobserved household and family background characteristics that affect exposure to violence and dropout, the effect disappears for two of the four measures of exposure to violence during childhood. Being hit hard is associated with a 0.12 increase in the probability of dropout (compared to the baseline increase of 0.21 and an increase of 0.13 in the neighbourhood fixed effect regressions). The fear of being hurt is associated with a 0.08 increase in the probability of dropout (compared to the baseline increase of 0.13 and an increase of 0.08 in the neighbourhood fixed effect regressions).

Below we repeat the sibling fixed effect regressions, but instead of specifying the exposure to violence during childhood measures as binary, we use the full five-point scale. In all cases the omitted category is "never exposed to violence during childhood". Being hit hard sometimes is associated with an increase in the probability of dropout of 0.12 (relative to the omitted category of "never") and the fear of being hurt sometimes is associated with an increase in the probability of dropout of 0.11.

Table 12: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
male	0.0791***	0.0721***	0.0761***	0.0712***
	(0.0203)	(0.0205)	(0.0203)	(0.0202)
_once	-0.0270	0.0348	-0.0338	0.0493
	(0.0611)	(0.0363)	(0.0473)	(0.0382)
_sometimes	0.115**	0.0531	0.108**	0.0401
	(0.0475)	(0.0396)	(0.0433)	(0.0350)
_often	0.0431	-0.0463	-0.0861	-0.0454
	(0.109)	(0.0701)	(0.111)	(0.0723)
_veryoften	0.290*	0.0523	0.00293	-0.0535
	(0.148)	(0.146)	(0.125)	(0.0592)
Constant	0.452***	0.450***	0.448***	0.447***
	(0.00864)	(0.0118)	(0.00825)	(0.0157)
Observations	2,143	2,143	2,143	2,143
R-squared	0.014	0.011	0.017	0.012
Number of households	1,087	1,087	1,087	1,087
***** p<0.01, ** p<0.05, * p<0.1**				

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

The inclusion of birth order effects leads to a further decrease in the coefficient on the *hit_hard* measure and increases in the coefficient on *afraid_hurt* (it also produces a significant coefficient for *pushed*). The fact that the coefficient on *hit_hard* decreases only slightly suggests a robust association between being hit to the point of injury and the probability of dropout.

Appendix A: Baseline regressions

Table 13: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhood_exposure*
Violence measure	-0.341***	-0.230***	-0.284***	-0.196***	-0.457***
	(0.0628)	(0.0672)	(0.0661)	(0.0655)	(0.113)
Constant	0.0162	0.0376	0.0476	0.0597	0.0792
	(0.130)	(0.137)	(0.132)	(0.145)	(0.145)
Observations	4,671	4,662	4,674	4,666	4,649
R-squared	0.006	0.008	0.012	0.009	0.015

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1. Robust and cluster corrected standard errors in brackets.

Table 14: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhood_exposure*
Violence measure	0.207***	0.120***	0.133***	0.0508***	0.203***
	(0.0332)	(0.0301)	(0.0279)	(0.0176)	(0.0346)
Constant	0.546***	0.537***	0.535***	0.542***	0.521***
	(0.0210)	(0.0240)	(0.0205)	(0.0211)	(0.0232)
Observations	3,148	3,141	3,149	3,145	3,133
R-squared	0.010	0.009	0.011	0.002	0.013

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets.

Table 15: Exposure to violence during childhood (dependent variable: matric_result)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhood_exposure*
Violence measures	-0.00910	-0.0186*	-0.0337***	-0.0315***	-0.0530***
	(0.0180)	(0.00989)	(0.0102)	(0.0116)	(0.0185)
Constant	0.527***	0.529***	0.531***	0.536***	0.535***
	(0.0186)	(0.0184)	(0.0185)	(0.0211)	(0.0200)
Observations	1,594	1,591	1,594	1,592	1,586
R-squared	0.000	0.002	0.008	0.012	0.010

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets.

Appendix B: Weighted regressions

Table 16: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
age	0.0575***	0.0573***	0.0570***	0.0568***	0.0569***
	(0.00583)	(0.00582)	(0.00580)	(0.00580)	(0.00581)
male	0.0717**	0.0719**	0.0684**	0.0669**	0.0700**
	(0.0289)	(0.0287)	(0.0286)	(0.0288)	(0.0287)
black	-1.050***	-1.050***	-1.043***	-1.045***	-1.035***
	(0.101)	(0.103)	(0.102)	(0.101)	(0.101)
coloured	-0.785***	-0.781***	-0.776***	-0.771***	-0.774***
	(0.0738)	(0.0738)	(0.0735)	(0.0735)	(0.0732)
grant	-0.179*	-0.184*	-0.203**	-0.204**	-0.176*
	(0.101)	(0.100)	(0.100)	(0.102)	(0.102)
mother_education	0.0614***	0.0614***	0.0617***	0.0616***	0.0615***
	(0.00636)	(0.00633)	(0.00634)	(0.00635)	(0.00633)
home_lang_english	0.371***	0.373***	0.373***	0.371***	0.372***
	(0.0477)	(0.0482)	(0.0477)	(0.0480)	(0.0483)
female_hh	-0.108***	-0.108***	-0.106***	-0.109***	-0.105***
	(0.0313)	(0.0314)	(0.0315)	(0.0315)	(0.0315)
hh_ownbooks	0.229***	0.229***	0.230***	0.229***	0.219***
	(0.0396)	(0.0400)	(0.0393)	(0.0389)	(0.0397)
Violence measure	-0.152***	-0.0742**	-0.0959***	-0.102***	-0.186***
	(0.0543)	(0.0367)	(0.0351)	(0.0291)	(0.0495)
Constant	-0.771***	-0.773***	-0.775***	-0.761***	-0.747***
	(0.180)	(0.178)	(0.177)	(0.178)	(0.177)
Observations	4,139	4,134	4,144	4,138	4,123
R-squared	0.469	0.470	0.469	0.470	0.472

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Table 17: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0571***	0.0570***	0.0573***	0.0572***
	(0.00583)	(0.00583)	(0.00581)	(0.00579)
male	0.0717**	0.0707**	0.0688**	0.0664**
	(0.0289)	(0.0288)	(0.0286)	(0.0288)
black	-1.052***	-1.056***	-1.051***	-1.050***
	(0.103)	(0.101)	(0.101)	(0.101)
coloured	-0.780***	-0.779***	-0.779***	-0.774***
	(0.0736)	(0.0739)	(0.0739)	(0.0739)
grant	-0.203**	-0.203**	-0.202**	-0.207**
	(0.0996)	(0.102)	(0.101)	(0.102)
mother_education	0.0609***	0.0608***	0.0608***	0.0609***
	(0.00637)	(0.00636)	(0.00635)	(0.00632)
home_lang_english	0.368***	0.372***	0.373***	0.370***
	(0.0477)	(0.0479)	(0.0481)	(0.0478)
female_hh	-0.109***	-0.106***	-0.106***	-0.107***
	(0.0311)	(0.0314)	(0.0313)	(0.0314)
hh_ownbooks	0.229***	0.229***	0.230***	0.229***
	(0.0399)	(0.0395)	(0.0399)	(0.0392)
_once	-0.0195	-0.0470	-0.0417	0.00662
	(0.0560)	(0.0398)	(0.0453)	(0.0397)
_sometimes	-0.0563	-0.0502	-0.0650	-0.0927***
	(0.0650)	(0.0386)	(0.0414)	(0.0336)
_often	-0.508***	-0.199**	-0.201**	-0.115
	(0.123)	(0.0853)	(0.0864)	(0.0700)
_veryoften	-0.285*	-0.189	-0.279**	-0.187**
	(0.160)	(0.136)	(0.108)	(0.0748)
Constant	-0.765***	-0.754***	-0.755***	-0.755***
	(0.180)	(0.177)	(0.178)	(0.178)
Observations	4,148	4,148	4,148	4,148
R-squared	0.470	0.469	0.469	0.470

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Table 18: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
male	0.125***	0.124***	0.127***	0.125***	0.124***
	(0.0187)	(0.0188)	(0.0185)	(0.0186)	(0.0186)
black	0.0230	0.0242	0.0147	0.0190	0.0204
	(0.0501)	(0.0496)	(0.0493)	(0.0504)	(0.0503)
coloured	0.139***	0.138***	0.133***	0.136***	0.131***
	(0.0330)	(0.0331)	(0.0326)	(0.0330)	(0.0338)
grant	0.233***	0.240***	0.236***	0.239***	0.235***
	(0.0293)	(0.0294)	(0.0292)	(0.0294)	(0.0296)
mother_education	-0.0290***	-0.0287***	-0.0292***	-0.0289***	-0.0288***
	(0.00349)	(0.00348)	(0.00346)	(0.00344)	(0.00344)
home_lang_english	-0.179***	-0.183***	-0.183***	-0.187***	-0.185***
	(0.0260)	(0.0260)	(0.0259)	(0.0256)	(0.0256)
female_hh	0.0344*	0.0322*	0.0315*	0.0336*	0.0322*
	(0.0184)	(0.0187)	(0.0184)	(0.0183)	(0.0183)
hh_ownbooks	-0.157***	-0.161***	-0.157***	-0.165***	-0.155***
	(0.0241)	(0.0238)	(0.0238)	(0.0233)	(0.0238)
Violence measure	0.142***	0.0566**	0.0691***	0.0234	0.111***
	(0.0318)	(0.0222)	(0.0230)	(0.0171)	(0.0300)
Constant	0.619***	0.622***	0.632***	0.638***	0.620***
	(0.101)	(0.0990)	(0.0979)	(0.0980)	(0.0994)
Observations	3,088	3,083	3,093	3,087	3,075
R-squared	0.291	0.290	0.291	0.289	0.292

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Table 19: Exposure to violence during childhood (dependent variable: matric_result)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
male	-0.0158***	-0.0159***	-0.0163***	-0.0161***	-0.0157***
	(0.00543)	(0.00543)	(0.00542)	(0.00532)	(0.00537)
black	-0.0878***	-0.0862***	-0.0879***	-0.0888***	-0.0855***
	(0.0166)	(0.0163)	(0.0167)	(0.0164)	(0.0161)
coloured	-0.0757***	-0.0756***	-0.0746***	-0.0738***	-0.0750***
	(0.0163)	(0.0162)	(0.0163)	(0.0160)	(0.0162)
grant	-0.0196**	-0.0220**	-0.0199**	-0.0202**	-0.0213**
	(0.00969)	(0.00931)	(0.00965)	(0.00974)	(0.00927)
mother_education	0.00887***	0.00887***	0.00897***	0.00895***	0.00894***
	(0.00140)	(0.00141)	(0.00141)	(0.00139)	(0.00140)
home_lang_english	0.0367***	0.0371***	0.0368***	0.0358***	0.0368***
	(0.00788)	(0.00787)	(0.00787)	(0.00784)	(0.00783)
female_hh	-0.0133**	-0.0145**	-0.0139**	-0.0136**	-0.0129**
	(0.00644)	(0.00646)	(0.00649)	(0.00636)	(0.00631)
hh_ownbooks	0.000362	0.000205	-0.000350	-0.000159	-0.000759
	(0.0112)	(0.0113)	(0.0113)	(0.0114)	(0.0113)
Violence measure	-0.0202	-0.00659	-0.00852	-0.0217***	-0.0300**
	(0.0151)	(0.00815)	(0.00841)	(0.00660)	(0.0123)
Constant	0.530***	0.530***	0.529***	0.531***	0.531***
	(0.0321)	(0.0319)	(0.0319)	(0.0321)	(0.0320)
Observations	1,331	1,329	1,332	1,331	1,326
R-squared	0.520	0.519	0.518	0.523	0.523

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Appendix C: Exposure to violence scale regressions

Table 20: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0548***	0.0545***	0.0547***	0.0548***
	(0.00469)	(0.00477)	(0.00486)	(0.00475)
male	0.0362	0.0351	0.0346	0.0319
	(0.0414)	(0.0400)	(0.0401)	(0.0394)
black	-1.072***	-1.076***	-1.071***	-1.069***
	(0.127)	(0.125)	(0.126)	(0.127)
coloured	-0.793***	-0.794***	-0.794***	-0.792***
	(0.0855)	(0.0860)	(0.0862)	(0.0863)
grant	-0.199***	-0.202***	-0.204***	-0.209***
	(0.0551)	(0.0549)	(0.0548)	(0.0561)
mother_education	0.0514***	0.0511***	0.0511***	0.0513***
	(0.00802)	(0.00808)	(0.00814)	(0.00826)
home_lang_english	0.382***	0.385***	0.388***	0.383***
	(0.0410)	(0.0415)	(0.0417)	(0.0417)
female_hh	-0.0760**	-0.0741**	-0.0749**	-0.0746**
	(0.0342)	(0.0335)	(0.0330)	(0.0341)
hh_ownbooks	0.199***	0.200***	0.200***	0.199***
	(0.0434)	(0.0419)	(0.0413)	(0.0426)
_once	-0.0246	-0.0372	-0.0489	0.0310
	(0.0444)	(0.0416)	(0.0442)	(0.0407)
_sometimes	-0.0863	-0.0390	-0.0490	-0.0818**
	(0.0804)	(0.0339)	(0.0339)	(0.0309)
_often	-0.446***	-0.223***	-0.238**	-0.124
	(0.127)	(0.0825)	(0.0891)	(0.0776)
_veryoften	-0.267	-0.149	-0.211*	-0.200**
	(0.227)	(0.138)	(0.122)	(0.0819)
Constant	-0.676***	-0.663***	-0.661***	-0.668***
	(0.181)	(0.179)	(0.182)	(0.183)
Observations	4,148	4,148	4,148	4,148
R-squared	0.431	0.430	0.430	0.431
***** p<0.01, ** p<0.05, * p<0.1**				

Source: Own calculations. CAPS Wave 1. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Table 21: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
male	0.131***	0.130***	0.131***	0.131***
	(0.0223)	(0.0225)	(0.0221)	(0.0224)
black	0.0437	0.0467	0.0364	0.0433
	(0.0619)	(0.0628)	(0.0621)	(0.0642)
coloured	0.165***	0.163***	0.158***	0.163***
	(0.0432)	(0.0424)	(0.0428)	(0.0425)
grant	0.237***	0.237***	0.236***	0.241***
	(0.0191)	(0.0190)	(0.0177)	(0.0187)
mother_education	-0.0271***	-0.0270***	-0.0272***	-0.0271***
	(0.00221)	(0.00215)	(0.00222)	(0.00220)
home_lang_english	-0.187***	-0.187***	-0.189***	-0.189***
	(0.0248)	(0.0248)	(0.0237)	(0.0237)
female_hh	0.0185	0.0155	0.0156	0.0164
	(0.0405)	(0.0396)	(0.0399)	(0.0401)
hh_ownbooks	-0.149***	-0.150***	-0.147***	-0.153***
	(0.0247)	(0.0247)	(0.0259)	(0.0249)
_once	-0.0856**	0.00176	-0.00412	-0.000470
	(0.0349)	(0.0223)	(0.0254)	(0.0211)
_sometimes	0.107***	0.0331*	0.0780***	0.0355**
	(0.0320)	(0.0175)	(0.0268)	(0.0141)
_often	0.189**	0.129**	0.116	0.0189
	(0.0879)	(0.0497)	(0.0913)	(0.0563)
_veryoften	0.171*	0.0726	0.0122	0.0297
	(0.0940)	(0.0821)	(0.0836)	(0.0564)
Constant	0.608***	0.603***	0.606***	0.613***
	(0.0462)	(0.0432)	(0.0450)	(0.0475)
Observations	3,096	3,096	3,096	3,096
R-squared	0.249	0.245	0.247	0.244

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Appendix D: Regressions with school quality dummy

Table 22: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*	*childhd_exposure*
age	0.0552***	0.0550***	0.0550***	0.0547***	0.0548***
	(0.00602)	(0.00582)	(0.00577)	(0.00574)	(0.00591)
male	0.0448	0.0469	0.0439	0.0415	0.0420
	(0.0445)	(0.0442)	(0.0424)	(0.0415)	(0.0427)
black	-0.993***	-0.990***	-0.978***	-0.978***	-0.978***
	(0.138)	(0.137)	(0.138)	(0.138)	(0.138)
coloured	-0.836***	-0.836***	-0.831***	-0.829***	-0.827***
	(0.0911)	(0.0910)	(0.0909)	(0.0913)	(0.0911)
grant	-0.173***	-0.181***	-0.196***	-0.199***	-0.178***
	(0.0571)	(0.0556)	(0.0568)	(0.0566)	(0.0572)
mother_education	0.0508***	0.0506***	0.0510***	0.0507***	0.0507***
	(0.00784)	(0.00780)	(0.00794)	(0.00783)	(0.00784)
home_lang_english	0.409***	0.413***	0.412***	0.410***	0.411***
	(0.0415)	(0.0418)	(0.0419)	(0.0421)	(0.0415)
female_hh	-0.0901**	-0.0883**	-0.0886**	-0.0903**	-0.0873**
	(0.0374)	(0.0365)	(0.0364)	(0.0372)	(0.0368)
hh_ownbooks	0.183***	0.185***	0.185***	0.182***	0.175***
	(0.0414)	(0.0399)	(0.0388)	(0.0405)	(0.0400)
school_quality	-0.198***	-0.198***	-0.198***	-0.199***	-0.200***
	(0.0542)	(0.0543)	(0.0550)	(0.0544)	(0.0538)
Violence measure	-0.182***	-0.0679**	-0.0784**	-0.101***	-0.181***
	(0.0592)	(0.0333)	(0.0307)	(0.0240)	(0.0411)
Constant	-0.645***	-0.651***	-0.657***	-0.634***	-0.623***
	(0.207)	(0.197)	(0.199)	(0.196)	(0.200)
Observations	3,799	3,793	3,802	3,796	3,784
R-squared	0.428	0.429	0.427	0.428	0.430

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. Own calculations. CAPS Wave 1. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Appendix E: Matched pair sample regressions

Table 23: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0322***	0.0322***	0.0326***	0.0321***
	(0.00977)	(0.00985)	(0.00943)	(0.00933)
male	0.0363	0.0394	0.0441	0.0403
	(0.0531)	(0.0512)	(0.0494)	(0.0471)
Violence measure	-0.0896	-0.00158	0.0290	-0.0823*
	(0.0832)	(0.0438)	(0.0546)	(0.0470)
Constant	-0.572***	-0.580***	-0.593***	-0.553***
	(0.189)	(0.188)	(0.183)	(0.176)
Observations	1,786	1,784	1,790	1,787
R-squared	0.022	0.021	0.022	0.024
Number of households	901	902	902	902

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 matched pair sibling sample. Robust and cluster corrected standard errors in brackets.

Table 24: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0325***	0.0322***	0.0331***	0.0335***
	(0.00953)	(0.00944)	(0.00919)	(0.00951)
male	0.0380	0.0455	0.0447	0.0371
	(0.0514)	(0.0500)	(0.0494)	(0.0468)
_once	0.0607	0.0505	0.0103	0.0618
	(0.0907)	(0.0784)	(0.0779)	(0.0667)
_sometimes	-0.0816	0.0109	0.0584	-0.0330
	(0.0780)	(0.0442)	(0.0621)	(0.0464)
_often	-0.0404	0.0686	-0.142*	-0.170
	(0.166)	(0.0838)	(0.0757)	(0.117)
_veryoften	-0.237	-0.102	0.0914	-0.192*
	(0.334)	(0.261)	(0.225)	(0.106)
Constant	-0.583***	-0.591***	-0.606***	-0.590***
	(0.180)	(0.185)	(0.183)	(0.188)
Observations	1,790	1,790	1,790	1,790
R-squared	0.024	0.023	0.024	0.028
Number of households	902	902	902	902

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 matched pair sibling sample. Robust and cluster corrected standard errors in brackets.

Table 25: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**
age	0.0239*	0.0236*	0.0247*	0.0234*
	(0.0135)	(0.0139)	(0.0135)	(0.0134)
male	0.0350	0.0397	0.0445	0.0409
	(0.0520)	(0.0514)	(0.0494)	(0.0476)
Violence measure	-0.145*	-0.0120	0.0219	-0.0983*
	(0.0831)	(0.0596)	(0.0612)	(0.0580)
younger_youngadult	-0.0422	-0.0399	-0.0343	-0.0471
	(0.0633)	(0.0628)	(0.0645)	(0.0725)
younger *** [Violence measure]**	0.118	0.0223	0.00792	0.0313
	(0.184)	(0.0796)	(0.0785)	(0.0768)
Constant	-0.402	-0.407	-0.435	-0.374
	(0.269)	(0.275)	(0.267)	(0.264)
Observations	1,786	1,784	1,790	1,787
R-squared	0.023	0.021	0.022	0.025
Number of households	901	902	902	902

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 matched pair sibling sample. Robust and cluster corrected standard errors in brackets.

Appendix F: Birth order effect regressions

Table 26: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	*hit_hard*	*pushed*	*afraid_hurt*	*adult_put_down*
age	0.0311***	0.0307***	0.0303***	0.0300***
	(0.00940)	(0.00933)	(0.00901)	(0.00890)
male	0.0372	0.0383	0.0364	0.0342
	(0.0507)	(0.0475)	(0.0459)	(0.0443)
Violence measure	-0.144*	-0.0123	-0.0317	-0.0909**
	(0.0757)	(0.0384)	(0.0449)	(0.0429)
youngest_youngadult	0.0120	0.0506	0.0460	0.00693
	(0.0732)	(0.0697)	(0.0679)	(0.0606)
middle_youngadult	-0.0109	0.0143	0.0373	0.0176
	(0.0594)	(0.0515)	(0.0510)	(0.0608)
youngest** * **[Violence measure]**	-0.0561	-0.188*	-0.151	0.0119
	(0.243)	(0.103)	(0.102)	(0.0629)
middle** * **[Violence measure]**	0.460**	0.00739	-0.0817	-0.000651
	(0.200)	(0.107)	(0.0958)	(0.144)
Constant	-0.606***	-0.607***	-0.596***	-0.566***
	(0.193)	(0.188)	(0.186)	(0.179)
Observations	2,584	2,581	2,587	2,583
R-squared	0.023	0.020	0.020	0.021
Number of households	1,171	1,172	1,172	1,172

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 cluster sibling sample. Robust and cluster corrected standard errors in brackets.

Table 27: Exposure to violence during childhood (dependent variable: numeracy_score)

VARIABLES	***hit_hard***	***pushed***	***afraid_hurt***	***adult_put_down***
age	0.0310***	0.0308***	0.0305***	0.0303***
	(0.00947)	(0.00949)	(0.00911)	(0.00906)
male	0.0360	0.0344	0.0328	0.0316
	(0.0514)	(0.0491)	(0.0472)	(0.0472)
_once	0.0441	0.0452	0.0559	0.0576
	(0.0726)	(0.0757)	(0.0743)	(0.0733)
_sometimes	-0.128*	-0.0875	-0.0771	-0.0825
	(0.0705)	(0.0642)	(0.0633)	(0.0628)
_often	-0.139	-0.107	-0.0938	-0.0944
	(0.141)	(0.134)	(0.132)	(0.135)
_veryoften	-0.269	-0.199	-0.190	-0.185
	(0.201)	(0.196)	(0.191)	(0.184)
youngest_youngadult	0.0132	0.0514	0.0518	0.0377
	(0.0725)	(0.0694)	(0.0637)	(0.0627)
middle_youngadult	-0.00974	0.0147	0.0427	0.0460
	(0.0596)	(0.0511)	(0.0504)	(0.0611)
youngest** **[Violence measure]	-0.0710	-0.197**	-0.174*	-0.0610
	(0.230)	(0.0924)	(0.0977)	(0.0602)
middle** **[Violence measure]	0.462**	0.0130	-0.0982	-0.0657
	(0.199)	(0.105)	(0.110)	(0.134)
Constant	-0.606***	-0.606***	-0.602***	-0.600***
	(0.191)	(0.190)	(0.181)	(0.181)
Observations	2,584	2,581	2,587	2,583
R-squared	0.024	0.022	0.021	0.020
Number of households	1,171	1,172	1,172	1,172

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Wave 1 cluster sibling sample. Robust and cluster corrected standard errors in brackets.

Table 28: Exposure to violence during childhood (dependent variable: dropout)

VARIABLES	***hit_hard***	***pushed***	***afraid_hurt***	***adult_put_down***
male	0.0782***	0.0736***	0.0769***	0.0750***
	(0.0206)	(0.0204)	(0.0209)	(0.0217)
Violence measure	0.108*	0.0538*	0.116**	0.0119
	(0.0615)	(0.0283)	(0.0571)	(0.0489)
youngest_youngadult	0.0751**	0.0900**	0.100***	0.0755**
	(0.0341)	(0.0367)	(0.0345)	(0.0372)
middle_youngadult	-0.0219	-0.000681	0.000667	-0.00803
	(0.0381)	(0.0348)	(0.0445)	(0.0481)
youngest *** [violence measure]**	-0.0199	-0.0895	-0.154*	-0.0154
	(0.234)	(0.118)	(0.0830)	(0.121)
middle *** [violence measure]**	0.150	-0.0910	-0.0899	-0.0301
	(0.166)	(0.0865)	(0.0824)	(0.0595)
Constant	0.446***	0.445***	0.432***	0.451***
	(0.0109)	(0.0126)	(0.0118)	(0.0135)
Observations	2,139	2,136	2,141	2,137
R-squared	0.019	0.016	0.021	0.014
Number of households	1,087	1,086	1,087	1,085

***** p<0.01, ** p<0.05, * p<0.1**

Source: Own calculations. CAPS Panel. Robust and cluster corrected standard errors in brackets. All regressions include neighbourhood dummies for 49 neighbourhoods in the CAPS sample.

Appendix G: Variable definitions

Matric results

The CAPS dataset contains the matric results for each individual in the sample that wrote the matric exam. The symbol and level were obtained for each subject written by the individual. We take the mean of the subjects written by the individual to construct the final matric result for every matric exam written by each individual (it is possible that an individual wrote the exam more than once). We used the following table to convert symbols into grades:

NSC Achievement level	NSC %	Senior Certificate Higher Grade	Senior Certificate Standard Grade	Matric grade in CAPS
7	80 – 99	A		85%
6	70 – 79	B	A	75%
5	60 – 69	C	B	65%
4	50 – 59	D	C	55%
3	40 – 49	E	D	45%
2	30 – 39	F	E	35%
1	0 - 29	G	F	25%

Matric results were constructed differently for individuals who took the exam until 2007 compared to those who took it after. Until 2007, the matric exam was known as the Senior Certificate and individuals who wrote the exam could choose to complete their subjects at higher grade or standard grade, which determined whether they would write the exam at higher or standard grade.

Differences between higher and standard grade varied across subjects, but generally it is accepted that a standard grade symbol is the equivalent to one level lower than the higher grade symbol (as shown above). To obtain the Senior Certificate, a learner had to obtain an aggregate of 720 marks in 6 subjects. If you met the requirement, but obtained 25% in one or two of the subjects, which represents failure in both standard and higher grades, the mark would be converted to lower grade and since 25% is a pass on lower grade, the student would qualify for a Senior Certificate.

As shown in the table above, our approach is based on the following principles: first, since we only know the symbol and not the exact grade, we assign each individual the midpoint of the range associated with the symbol they obtained

(with the exception of an A symbol obtained at higher grade, where we assign 85%). Second, when an individual recorded a symbol at the standard or lower grade level or the grade field was missing, we assigned the mid-point of the grade one level lower than the higher grade equivalent.

Finally, once we have assigned a mid-point equivalent for each subject, based on whether the subject was taken at the higher or standard grade level, we calculate the average matric result for the every matric exam written by each young adult in the sample.

Since 2008, the matric exam is known as the National Senior Certificate (NSC) and the higher, standard and lower grade distinctions were abolished. Under the NSC system, an individual has to take 7 subjects and obtain 30% to pass each subject. In this case the mid-point equivalent for each subject was assigned based on the symbol using the table above.

It is possible for an individual to write the exam more than once (if they fail the first time, for example), we choose to use results from the first matric exam written by each individual. Therefore, we create the matric result variable by assigning the earliest available matric result (starting with matric results from Wave 1) and continue until we have assigned the earliest available matric result to each individual using all five waves of the CAPS panel.

An alternative approach to the above may be to use the admissions system used by South African universities who calculate a score for each university applicant based on their matric results. This score is used to determine qualification for various academic programmes at the university. The challenge with using the admissions system is that it has changed over the years in response to changes in the matric curriculum or changes in the senior certificate system highlighted above so we would need a complete record of the admissions system used by a university to construct a similar score.

Neighbourhood violence

The CAPS dataset includes a community data file for Wave 1, which is based on the 2001 South Africa Census data tabulated at the "sub-place" level.

There are 683 "sub-places" in metropolitan Cape Town and 229 of these "sub-places were used as enumeration areas in CAPS. Each police precinct usually corresponds to several "sub-places" and police stations and "sub-places" were

linked on the basis of the police districts used by the SAPS. A customised map and list obtained from the City of Cape Town's Strategic Development Information and GIS Department was used to link police stations to sub-places in Cape Town. This allowed us to link murder statistics from the SAPS (available at www.saps.gov.za) and population counts from the Census 2001 to neighbourhoods in CAPS.

Since our primary concern is to explore the nature of violence in Cape Town, these police precincts were used to define the neighbourhoods in our sample. This process yields 50 neighbourhoods (or police precincts) based on 2001 police precinct boundaries. The same police precinct boundaries were used to define the population size for each neighbourhood based on the population size of each "sub-place" in that neighbourhood as given by the 2001 Census. The earliest publicly available SAPS murder statistics (April 2003 – March 2004) were linked to the neighbourhoods where the young adults reside for Wave 1 of CAPS which was completed at the end of 2002. Therefore, we have to assume that there were no large non-random changes in murder rates between 2002 and 2003. We exclude from our sample murder statistics from the police stations that were opened after Wave 1: Phillippi East from Nyanga, Lingelethu West and Harare from Khayelitsha, Kleinvlei from Kuilsrivier all in July 2004, Mfuleni from Kleinvlei in July 2005, Belhar from Delft in December 2007 and Lwandle from Strand in October 2008.

References

Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. Criminology, 30, 47-87.

Aizer, A. (2008). Neighborhood Violence and Urban Youth, NBER Working Paper Series, Vol. w13773 [online] Available: www.nber.org.

Akresh, R., and de Walque, D. (2008). Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide. World Bank Policy Research Working Paper No. 4606, World Bank, Washington, DC.

Almlund M, Duckworth A, Heckman JJ, Kautz T. 2011. Personality Psychology and Economics. In Handbook of the Economics of Education, Hanushek EA, Machin S, Wößmann L (eds.), vol. 4. Elsevier: Amsterdam; 1–181.

Angrist, J. (2001). Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice, Journal of Business and Economic Statistics 19(1), 2-16.

Angrist, J., Lavy, V., and Schlosser, A. (2005). New Evidence on the Causal Link Between the Quantity and Quality of Children, NBER Working Paper No. 11838.

Angrist, J., and Pischke, S. (2009). Mostly harmless econometrics: an empiricist's companion, Princeton University Press.

Atkinson, G., Mourato, S., and Healey, A. (2003). The Costs of Violent Crime, vol. 4(4): 79-94, Oxfordshire: World Economics, Economic & Financial Publishing.

Barbarin, O., Richter, L., and deWet, T. (2001). Exposure to violence, coping resources, and adjustment of South African children. American Journal of Orthopsychiatry, 7, 16–25.

Black, S., Devereux, P., and Salvanes, K. (2007). Older and Wiser? Birth Order and IQ of Young Men, NBER Working Papers 13237, National Bureau of Economic Research.

Black, S., Devereux, P., and Salvanes, K. (2005). The More the Merrier? The Effects of Family Size and Birth Order on Children's Education, Quarterly Journal of Economics, 669-700.

Becker, G. (1968). Crime and Punishment: An Economic Approach, The Journal of Political Economy 76: pp. 169–217.

Behrman, J., Pollak, R., and Taubman, P. (1982). Parental preferences and provision for progeny. Journal of Political Economy, 90(1): 52–73.

Bell, C., and Jenkins, E. (1991). Traumatic stress and children, Journal of Health Care for the Poor and Underserved, 2:175–85.

Björklund, A., and Salvanes, K. (2011). Education and Family Background: Mechanisms and Policies, in EA Hanushek, S Machin, and L Woessmann (eds.) Handbook of the Economics of Education, Volume 3, The Netherlands: North-Holland, pp. 201-247.

Blattman, C., and Annan, J. (2010). The Consequences of Child Soldiering. Review of Economics and Statistics 92(4): 882–98.

Burton, P. (2008). Merchants, Skollies and Stones: Experiences of School Violence in South Africa, Centre for Justice and Crime Prevention, Monograph Series, No. 4, Cape Town, April.

Card, D. and Rothstein, J. (2007). Racial segregation and the black-white test score gap, Journal of Public Economics 91, 2158-2184.

Carneiro, P., Heckman, J. (2003). Human capital policy, In: Heckman, J.J., Krueger, A.B., Friedman, B.M. (Eds.), Inequality in America: What Role for Human Capital Policies? MIT Press, Cambridge, MA.

Casale, D., and Posel, D. (2010). English language proficiency and earnings in a developing country: the case of South Africa, Paper prepared for the 5th IZA/World Bank Conference: Employment and Development.

Case, A., and Deaton, A. (1999). School Inputs and Educational Outcomes in South Africa, Quarterly Journal of Economics, Vol. 114, No. 3, pp. 1047-1084.

Case, A., and Yogo, M. (1999). Does School Quality Matter? Returns to Education and the Characteristics of Schools in South Africa, NBER Working Paper No. 7399.

Cooley-Quille, M., Turner, S., and Beidel, D. (1995). Emotional impact of children's exposure to community violence: A preliminary study, The American Academy of Child and Adolescent Psychiatry, 34:1362–68.

Cornwell, K., and Inder, B. (2008). Language and labour markets in South Africa, Journal of African Economies 17(3): 490-525.

Cullen, J., and Levitt, S. (1999). Crime, Urban Flight, and the Consequences for Cities, The Review of Economics & Statistics, 1999, vol. 81, issue 2, 159-169.
Cunha F., and Heckman J. (2007). The technology of skill formation. American Economic Review 97(2): 31–47.

Cunha F., and Heckman J. (2009). The economics and psychology of inequality and human development. Journal of the European Economic Association 7(2–3): 320–364. Presented as the Marshall Lecture, European Economics Association, Milan, Italy, August 29, 2008.

Cunha, F., Heckman, J., Lochner, L., and Masterov, D. (2006). Interpreting the Evidence on Life Cycle Skill Formation. In E. A. Hanushek and F. Welch (Eds.), Handbook of the Economics of Education. Amsterdam: North-Holland.

Cunha, F., Heckman, J., and Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. Econometrica 78(3): 883–931.

Currie, J. (2009). Healthy, wealthy and wise: Socioeconomic status, poor health in childhood and human capital development. Journal of Economic Literature, 47(7): 87–122.

Cutler, D., and Glaeser, E. (1997). Are ghettos good or bad?, Quarterly Journal of Economics 112, 827-872.

de Walque, D. (2006). The Socio-Demographic Legacy of the Khmer Rouge Period in Cambodia. Population Studies 60 (2): 223–31.

Downey, D., and Yuan, A. (2005). Sex differences in school performance during high school: puzzling patterns and possible explanations, The Sociological Quarterly 46, 299–321.

Duckworth, A., and Seligman, M. (2005). Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents. Psychological Science 16(12): 939-944.

Ensink, K., Robertson, B., Zissis, C., and Leger, P. (1997). Post-traumatic stress disorder in children exposed to violence, South African Medicine Journal, 87(11): 1526-30.

Evans, W., Oates, W. and Schwab, R. (1992). Measuring Peer Group Effects: A Model of Teenage Behavior, Journal of Political Economy 100 (5) 966-991.

Famularo, R., Kinscherff, R., and Fenton, T. (1992) Psychiatric diagnoses of maltreated children: Preliminary findings. Journal of American Academy of Child and Adolescent Psychiatry, 31:863-867.

Farmer P., (2004). An Anthropology of Structural Violence, Current Anthropology, Volume 45, Issue 3, pp. 305-325.

Felitti, V., Anda, R., Nordenberg, D., Williamson, D., Spitz, A., Edwards, V., et al (1998) Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults, The Adverse Childhood Experiences Study, American Journal of Preventative Medicine, 14: 245-258.

Francis, D., Caldji, C., Champagne, F., Plotsky, P., and Meaney, M. (1999) The role of corticotropin-releasing factor—norepinephrine systems in mediating the effects of early experience on the development of behavioral and endocrine responses to stress. Biological Psychiatry, 46: 1153-1166.

Freeman, R. (1999). The economics of crime, in O. Ashenfelter and D. Card (ed.), Handbook of Labor Economics, edition 1, volume 3, chapter 52, pages 3529-3571.

Furstenberg, F., Brooks-Gunn, J., and Morgan, P. (1987). Adolescent Mothers in Later Life. Cambridge University Press, New York.

Galtung, J. (1969). Violence, peace and peace research, Journal of Peace Research, 6(3): 171.

Glaeser, E., Sacerdote, B., and Scheinkman, J. (1996). Crime and social interactions, Quarterly Journal of Economics 111, 508-548.

Glaeser, E., and Saiz, A. (2004). The Rise of the Skilled City, Brookings-Wharton Papers on Urban Affairs 5: 47-94.

Glewwe, P., and Kremer, M. (2006). Schools, Teachers, and Education Outcomes in Developing Countries, Handbook of the Economics of Education, Elsevier.

Gorman-Smith, D., and Tolan, P (1998). The role of exposure to community violence and developmental problems among inner-city youth, Development and Psychopathology, 10:99–114.

Grogger, J. (1997). Local Violence and Educational Attainment, Journal of Human Resources 32(4): 659-682.

Guerra, N., Huesmann, R., and Spindler, A. (2003). Community violence exposure, social cognition and aggression among urban elementary school children, Child Development, Vol 73 (5): 1561-76.

Hanushek, E. (1992). The trade-off between child quantity and quality, The Journal of Political Economy, 100, 84–117.

Hanushek, E., Lavy, V., and Hitomi, K. (2006) Do Students Care About School Quality? Determinants of Dropout Behaviour in Developing Countries NBER Working Paper 12737.

Haveman, R. and Wolfe, B. (1995). The Determinants of Children Attainments: A Review of Methods and Findings, Journal of Economic Literature, 33(4): 1829-78.

Heckman, J. (2000). Policies to foster human capital, Research in Economics, 54 (1), 3–56.

Heckman, J. (2012). The Developmental Origins of Health, Health Economics, 21: 24–29.

Heckman, J., Stixrud, J., Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labormarket outcomes and social behavior, Journal of Labor Economics 24 (3), 411–482.

Heim, C., and Nemeroff, C. (2001). The Role of Childhood Trauma in the Neurobiology of Mood and Anxiety Disorders: Preclinical and Clinical Studies, Biological Psychiatry 49, 1023–1039.

Holt, S., Buckley, H., and Whelan, S. (2008). The impact of exposure to domestic violence on children and young people: A review of the literature, Child Abuse & Neglect, 32 (8): 797-810.

Hulshoff, P., Hoek, H., Susser, E., Brown, A., Dingemans, A., Schnack, H., et al (2000) Prenatal exposure to famine and brain morphology in schizophrenia, American Journal of Psychiatry, 157: 1170-1172.

Kendler, K., Bulik, C., Silberg, J., Hettema, J., Myers, J., and Prescott, C. (2000) Childhood sexual abuse and adult psychiatric and substance use disorders in women: An epidemiological and cotwin control analysis, Arch Gen Psychiatry, 57:953–959.

Kessler, R., Davis, C., and Kendler, K. (1997) Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey, Psychology Medicine, 27: 1101–1119.

Kling, J., Liebman, J., and Katz, L. (2007). Experimental Analysis of Neighborhood Effects, Econometrica 75 (1): 83–119.

Knaul, F., and Ramírez, M. (2005) Family Violence and Child Abuse in Latin America and the Caribbean: The Cases of Colombia and Mexico, Sustainable Development Department Technical Papers Series, Inter-American Development Bank, Washington, DC.

Knudsen E., Heckman J., Cameron J., and Shonkoff J. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proceedings of the National Academy of Sciences 103(27): 10155–10162.

Lam, D., Ardington, C., and Leibbrandt, M. (2006). Progress through School in Urban South Africa: Evidence from Panel Data, Draft prepared for submission to Eleventh Annual Conference of the African Econometric Society, Dakar, Senegal.

Lam, D., Seekings, J., and Sparks, M. (2006). User's Guide to the Cape Area Panel Study Integrated Waves 1-2-3 Data. The University of Cape Town.

Lam, D., Ardington, C., Branson, N., Case, A., Leibbrandt, M., Menendez, A., Seekings, J., and Sparks, M. (2008). The Cape Area Panel Study: A Very Short Introduction to the Integrated Waves 1-2-3-4 (2002-2006) Data [online] Available: www.caps.uct.ac.za.

Lochner, L. (2011). Nonproduction Benefits of Education: Crime, Health, and Good Citizenship, in E. Hanushek, S. Machin, and L. Woessmann (eds.), Handbook of the Economics of Education, Vol. 4, Amsterdam: Elsevier Science.

Ludwig, J., Burdick-Will, J., Sampson, R., Sanbonmatsu, L., and Sharkey, P. (2010). Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-experimental, and Observational Comparison, Prepared for the Brookings Institution Project on Social Inequality and Educational Disadvantage.

Lupien, S., Maheu, F., Tu, M., Fiocco, A., and Schramek, T. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. Brain Cognition, 65: 209–237.

McEwen, B., and Sapolsky, R. (1995) Stress and cognitive function, Current Opinion in Neurobiology, 5: 205–216.

Meghir, C., and Rivkin, S. (2011). Econometric methods for research in education in Handbook of the economics of education, edited by Eric A. Hanushek, Stephen Machin and Ludger Woessmann Elsevier.

Ministerial Committee on Learner Retention in the South African Schooling System (2007) Progress Report to the Minister of Education Mrs Gnm Pandor, M.P.

Moffitt, R. (2001). Policy interventions low-level equilibria, and social interactions, In: S. Durlauf and P. Young (Eds.), Social Dynamics, Cambridge, MA: MIT Press, pp. 45-82.

Morán, R., (ed) with Amartya K. Sen and Gro Harlem Brundtland. (2004). Escaping the Poverty Trap: Investing in Children in Latin America. Washington, DC: Inter-American Development Bank.

Morrison, A., and Orlando, M. (1999). Social and Economic Costs of Domestic Violence: Chile and Nicaragua in Morrison and Biehl (eds.), Too Close to Home: Domestic Violence in the Americas. Washington, D.C.: Inter-American Development Bank.

Neisser, U., Boodoo, G., Bouchard, T., Boykin, W., Brody, N., Ceci, S., Halpern, D., Loehlin, J., Perloff, R., Sternberg, R., and Urbina, S. (1996). Intelligence: Knowns and Unknowns, American Psychologist 51(2): 77-101.

OHCHR (2006). United Nations Study on Violence against Children, Pinheiro, P.S. (ed.), Office of the High Commissioner for Human Rights, Geneva.

Rosenzweig, M., and Wolpin, K. (1980). Testing the Quantity-Quality Fertility Model: the Use of Twins as a natural Experiment, Econometrica XLVIII: 227-240.

Rosenzweig, M., and Zhang, J. (2009). Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy. Review of Economic Studies. 76(3): 1149-1174.

Sacerdote, B. (2001). Peer effects with random assignment: Results from Dartmouth roomates, Quarterly Journal of Economics 116, 681-704.

Sampson, R., Morenoff, J., and Earls, F. (1999). Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children, American Sociological Review 64:633-60.

Sauro M., Jorgensen, R., and Pedlow, C. (2003) Stress, glucocorticoids, and memory: A meta-analytic review, Stress 6: 235–245.

Scheper-Hughes, N., and Bourgois, P. (2004). Introduction: making sense of violence, in Violence in War and Peace: An anthology, Edited by N. Scheper-Hughes and P. Bourgois, pp.1-27. Oxford: Blackwell.

Schwab-Stone, M., Ayers, T., Kasprow, W., Voyce C., Barone, C., Shriver, T., and Weissberg, R. (1995). No safe haven: A study of violence exposure in an urban community, Journal of the American Academy of Child and Adolescent Psychiatry 34: 1343-52.

Schwartz, D., and Proctor, L. (2000). Community Violence Exposure and Children's Social Adjustment in the School Peer Group: The Mediating Roles of Emotion Regulation and Social Cognition, Journal of Consulting and Clinical Psychology, 68, 670–683.

Seedat, S., van Nood, E., Vythlingum, B., Stein, D., and Kaminer, D. (2000). School survey of exposure to violence and posttraumatic stress symptoms in adolescents, Southern African Journal of Child and Adolescent Mental Health, 42, 38–44.

Sharkey, P. (2010). The acute effect of local homicides on children's cognitive performance, Working Paper, New York University.

Shemyakina, O. (2011). The Effect of Armed Conflict on Accumulation of Schooling: Results from Tajikistan. Journal of Development Economics, 95(2): 186-200.

Shonkoff, J., and Phillips, D. (2000). From Neurons to Neighborhoods: The Science of Early Childhood Development, National Academies Press, Washington, DC.

Statistics South Africa (2011). Statistics South Africa Interactive Data: SuperWeb [online] Available: http://interactive.statssa.gov.za. [Accessed: 21 September 2011].

Weinberg, B. (2004). Testing the spatial mismatch hypothesis using inter-city variations in industrial composition, Regional Science and Urban Economics 34, 505-532.

Wolfe, R., and Johnson, S. (1995). Personality as a Predictor of College Performance, Educational And Psychological Measurement, 55(2): 177-185.
Yamauchi, F. (2005). Race, equity, and public schools in post-Apartheid South Africa: Equal opportunity for all kids, Economics of Education Review, 24: 2132-33.

Zimmerman, D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment, Review of Economics and Statistics 9-23.