Households in Conflict Network
hicn.org

The Scar of Civil War
Exposure in (Early) Childhood and School Test Scores as a Teenager

Philip Verwimp[1]

HiCN Working Paper 429

April 2025

Abstract
This paper investigates the effect of civil war exposure in (early) childhood on school test scores as a teenager. It uses test scores from the Concours National in Burundi, a nationwide competitive exam taken at the end of primary school, consisting of four academic disciplines for the period 2010-2012. These data are combined with exposure to civil war at different stages in childhood. The paper finds that an average duration of war exposure from in utero to age 12 (4.3 years) increases the age at which the test is taken with 1.72 years and causes a drop in the test score of 5.5 points on average (which is about 5% of the average grade), of which 1.75 points can be attributed to the scarring effect of war exposure and 3.75 points to the cognitive effect. The effects vary according to the timing of the shocks in childhood and along the distribution of test scores. Boys suffer more from the scarring effect, obtaining significantly lower test scores than girls from taking the exam at a later age, whereas girls suffer more from the cognitive effect of war shocks, conditional on age-at-test. Girl's performance is more affected than boys for mathematics but not more for languages. The paper finds evidence for a sex-specific selection mechanism in utero.

Acknowledgements
The author thanks Edward Miguel, Tom Vogl and Joseph Cummins for insightful discussions at an early stage of this paper. I have benefitted from comments received from Janet Currie, Jean-Francois Maystadt and from seminar participants at i3h in Brussels, the Graduate Institute in Geneva, the 17th HiCN Annual Workshop in Göttingen, the 2024 IUSSP workshop on Conflict and Population in Southampton, the 2024 LEER Conference at KU Leuven and the 2025 Development Economics seminar series at KU Leuven. Only the author is responsible for the content of the paper and for potential errors.

[1] KU Leuven, Université libre de Bruxelles and University of Antwerp

1.Introduction

There is a large volume of research available on the importance of early childhood for human development (Almond et at 2018). Biomedical research has found that the first years of the life of a child are crucial for the child's growth trajectory and brain development. Serious illness such as malaria or diarrhea as well as nutritional deprivation or other adverse events impacting the health of the child will have life-long consequences. Key findings of this research are that nutritional deprivation in childhood leads to a later start of schooling, earlier and more drop out, more grade repetition and lower educational attainment all together (Alderman et al 2006; Shemyakina 2006; Verwimp and Van Bavel 2014) . There is also evidence of the impact of various forms of violence on academic achievement. War, terrorist attacks, school shootings, police violence to name but a few have been shown to affect educational achievement and test scores negatively (Brück et al 2019; Duque 2024)

In this paper I track the impact of exposure to civil war in (early) childhood on teenage cognitive development. I use the exposure to civil war at different stages in childhood to infer its effect on the test scores of Burundian children who completed primary school ten years after exposure. As educational achievement is an important driver of outcomes in adulthood such as wages, the paper contributes to the understanding of the long-term consequences of war exposure at the micro-level. The arguably exogenous nature of war exposure with respect to school outcomes, the spread of the war over time (birth cohorts) and across space (provinces, municipalities and schools), the disposal of the universe of test scores for all schools in the country and for multiple years as well as the availability of results for different academic subjects make the study unique. Leveraging the arguably exogenous nature of war exposure with respect to school outcomes and controlling for time and location fixed effects, allows us to infer the causal impact of war shocks on age-at-test and on test scores.

Currie and Vogl (2013) review the literature on early life health and adult outcomes in developing countries. They argue that the long-term effects of early-life shocks are larger in developing countries compared to industrialised ones, for the following reasons: (i) such shocks are more frequent in many developing countries suggesting that the lingering effects of early-life health problems may well be more important in developing countries; (ii) child health shocks are likely to interact. For example, a child who is malnourished may be less able to ward off, or to recover from, disease. Similarly, a mother who was malnourished may bear a child who is compromised in his or her ability to cope with health insults; and (iii) the long-run consequences of early-life health shocks depend on the availability and effectiveness of mitigation strategies. To the extent that parents in rich countries are better able to compensate for shocks, the long-term effects of poor health in childhood may be greater in developing countries.

Hence, the study of early-life health shocks is a window to the understanding of long-lasting, important human development issues. In order to investigate the impact of adverse shocks in early childhood, researchers have used anthropometric indicators of a child's health, in particular height-for-age. This is considered a long-term indicator for the general health condition of a child as it captures the accumulated nutritional history of the child. Bundervoet et al (2009) have found that children in war-affected areas of Burundi have lower height-for-age z-scores compared to children from non-affected areas. Akresh et al (2011) have found this too in Rwanda.

Several studies, including Schultz (2002) and Behrman et al. (2009a), use early-life conditions as instruments in estimating the effect of height on adult outcomes. Currie an Vogl (2013) argue that height is primarily a proxy for early-life conditions, and hence they see more justification for studying the reduced-form effects of the instruments directly. Height is useful as an independent variable precisely when we cannot observe all relevant determinants of childhood development. These authors add that for much the same reason, we cannot use adult height to distinguish the effects of in utero shocks from those of postnatal shocks. In addition, in case that the instruments used by the researcher affect the adult outcome directly, the exclusion restriction is violated.

That is the point of entry for this paper: we observe civil war exposure at different stages in childhood (*in utero, at birth, in early infancy, in late infancy and in primary*) and hence we do not want to or need to rely on height as a proxy for early-life conditions. Here lies *the first of five distinct contributions of the paper*: the observation of the direct effect of shocks, including their occurrence, intensity and duration at different moments in childhood on later life outcomes. This allows us to distinguish the effects of in utero shocks from postnatal shocks, and early-childhood shocks from late-childhood shocks, something only few contributions to the literature manage to do. The paper exploits the significant variation in the data over space and over time whereby children born in different provinces/municipalities and years are differentially exposed to war shocks.

In addition, whereas a large part of the literature has concentrated on the impact of adverse events on educational attainment, in particular years of schooling completed or grades attained, this paper uses test scores from nationwide high-stakes exams a measure of cognitive ability to infer the impact of early life exposure. *That constitutes the second contribution of this paper*. Whereas the below examples quantify, to a certain extent, the "extensive margin" of violence on educational attainment (years of schooling completed, grades repeated), this paper quantifies the impact of early childhood exposure to civil conflict on education and cognitive ability at the "intensive margin".

Shemyakina (2006) finds from her empirical work on violent conflict in Tajikistan, that girls suffer the greater loss in education compared to boys and she attributes this to concerns over safety and low returns to girls' education. In contrast, Akresh and de Walque (2008) find that male Rwandan children in non-poor households incur the strongest effect. Alderman, Hoddinott and Kinsey (2006) find that Zimbabwean children affected by the civil war in the 1970s completed less grades of schooling and/or started school later than those not affected by the shocks. Similar results are found by León (2011) for Peru; Angrist and Kugler (2008) and Rodriguez and Sanchez (2009) for Colombia; Chamarbagwala and Morán (2009) for Guatemala, de Walque (2006) for Cambodia. Evans and Miguel (2004) find that young children in rural Kenya are more likely to drop out of school after the parent's death and that effect is particularly strong for children who lost their mothers. While Kenya was not the scene of violent conflict during the observed period, the finding is relevant because violent conflict produces many orphans, which may have a similar effect on their schooling.

Boone et al (2013) have found that only few children in 4th and 6th grade of primary school can actually read a sentence or do simple calculus despite going to school already for several years in Guinea-Bissau. IAEEA studies from 2000 and 2003 find lower scores on mathematics and reading in developing countries compared to developed countries. Across Africa and in other parts of the developing world, schools are underperforming. They do not deliver what they are supposed to do: teach children how to read and write, teach language, geography, history, biology and so forth. Reasons for this are multiple: teacher absence, teacher lack of skills, overcrowded classrooms, absence of learning material, inadequate learning environment such as violence, drugs and so on. This relates to *the third contribution of the paper*: the analysis of the cognitive effect for different academic disciplines, to wit mathematics, French, environmental science and Kirundi (Burundi's national language) that together constitute the academic subjects tested in the Concours National, the nationwide exam at the end of primary school. As a child takes the tests at the same time, the age-at-test is the same for the four disciplines, we can detect potential heterogenous impact of exposure depending on academic subject.

The exposure to civil war has left a scar on the child akin to the effect of an unemployment spell on later labour market performance (Gregg and Tominey 2005; Mroz and Savage 2006). *The fourth contribution of the paper* lies in the fact that we are able to split this scar in its three constituting components: *a trajectory effect, a scarring effect and a cognitive effect*. The trajectory effect captures the observation that pupils who are exposed to civil war shocks take the exam at a later age. The effects of the adverse shocks can manifest themselves in delayed entry into school, grade repetition, absence from school to name but a few. The trajectory is also slowed down because the shocks can have local-level or school-level impact such as the destruction or closing of school buildings, teacher absence, road blocks, insecurity causing parents to keep children from school and so on.

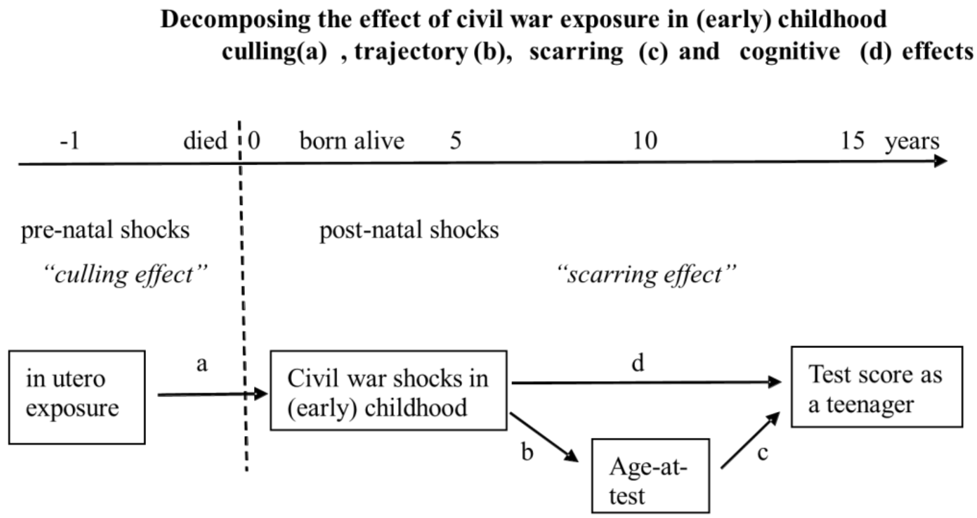
The scarring effect implies that pupils who take the exam at a later age obtain lower test scores, compatible with a depreciation of knowledge acquired in school over time. And the cognitive component captures the direct effect of civil war exposure on the test scores, conditional on age-at-test. The cognitive effect then is defined as the difference in the test score obtained in the exam, conditional on age-at-test, between an exposed and a non-exposed child. It captures inadequate brain development, concentration and learning problems, emotional dysfunctions such as anxieties or mistrust. The paper employs a *Conditional Mixed Process (CMP) estimator* as econometric method to estimate the three effects jointly, which to the best of my knowledge has not yet been performed in one unifying framework in this area of research.

The paper finds that an average duration of war exposure from in utero to age 12 (4.3 years) increases the age at which the test is taken with 1.72 years and causes a drop in the test score of 5.5 points on average (which is about 5% of the average grade), of which 1.75 points can be attributed to the scarring effect of war exposure and 3.75 points to the cognitive effect. For boys, the scarring effect is responsible for 38% of the drop and the cognitive effect for 62%, whereas for girls, the scarring part is only responsible for 18% of the drop and the cognitive effect for 82%. The impact is much larger in the capital Bujumbura and the effects also vary according to academic subject and the timing of the shock in childhood. It also varies along the distribution of test scores, and depends on the measure of exposure we employ, which I will document.

An additional finding is that in utero shocks have much less of an impact on the test scores of boys compared to girls and the paper investigates potential mechanisms that can explain that finding with reference to the literature on the culling effect. *Here lies the fifth contribution of the paper, to wit the finding of a sex-specific selection effect in utero*. Culling refers to a skewed male/female ratio at birth because more boys die in utero as a result of negative shocks compared to girls. Reason being that more boys are at the weaker end of the in-utero health distribution. Its consequence is that the surviving boys are on average stronger compared to the surviving girls. Stein, Susser, Saenger, and Marolla (1975) examined the relationship between prenatal exposure to famine and the conscription records of over 400,000 18-year-old men with famine exposure defined using birth date and place of birth. They did not find an effect on IQ scores. Combining data on adolescents aged 10–16 from the Adverse Childhood Experiences (ACE) project with the Malawi Longitudinal Study on Families and Health (N=1,559), Kämpfen et al (2022) show that girls whose households experienced two or more economic shocks in their year of birth have lower cognitive scores, which are measured using working memory, reading and mathematical tests. Girls also have lower educational attainment, conditional on age. These effects are gendered, as the authors do not observe similar effects among boys. Using historical data for Finland, Bruckner et al (2015) find that, consistent with culled cohorts, a one standard deviation decline in the annual cohort sex ratio precedes an 8% decrease in the risk of male infant mortality. Valente (2015) uses in utero exposure to the 1996-2006 Maoist insurgency in Nepal to assess culling and scarring effects. Her results show that exposed pregnancies are more likely to result in a miscarriage and in a female birth. Since exposed newborns do not seem to suffer of a poorer health, the author interprets her results as evidence of both scarring and culling effects operating. And Dagnelie et al (2018) in their work on the impact of violence on fetal health in the DRC find that in utero exposure increases fetal losses. More specifically, they find that conflict reduces the number of male live births, compatible with both scarring and culling effects. As they also find higher female mortality in infancy, they interpret their results as evidence of culling effect more then scarring: maternal exposure to conflict during pregnancy determined an increase of the health survival threshold, with the consequence of increasing the average health endowment of the newborns, particularly among boys, disproportionally populating the low end of the health distribution.

The remainder of the paper is structured as follows: section 2 presents the Conceptual Framework behind this study. Section 3 describes the features of Burundi's nationwide exam system at the end of primary school as well as the spread of the civil war over time and space. In Section 4 the data are described, including the construction of the exposure variables, together with a preliminary, graphic analysis and an outline of the identification strategy. Section 5 outlines our identification and econometric specification. Section 6 presents regression results for the estimation of age-at-test and test scores. Section 7 offers interpretation and discussion of the results. And section 8 presents detailed findings per academic discipline. Section 9 discusses a potential mechanism and section 10 concludes.

2. Conceptual Framework



We conceptualize the impact of civil war on test score using the four types of effects discussed in the introduction which occur between the moment of conception and the age at which the child takes the test. The first, (a) in the above graph, is *the culling effect*, referring to sex-specific selection as a results of the exposure to adverse shocks while in utero. As more girls than boys survive exposure in utero (Almond and Currie, 2011), this would mean that the surviving boys are on average stronger. The surviving children are then exposed to postnatal shocks in early and late infancy, resulting in *the trajectory effect* ((b) in the graph), meaning a higher age at which the test is taken compared to non-exposed children. This delay than leads to *the scarring effect*, which, narrowly defined, is the effect of age-at-test on the test score ((c) in the graph). In addition, there is also the direct effect of exposure on the test-scores, conditional on the age-at-test, which we define as *the cognitive effect* ((d) in the graph). In a broad definition, one could consider the overall effect of postnatal shocks (b+c+d) as the scarring effect, but in this paper I use the narrow definition (hence only effect (c)) as it allows us to separate the effect of postnatal shocks in its three constituting components.

3.Education and Civil War in Burundi

3.1 Burundi's National Test at the end of Primary School

Burundi's education system consists of 6 years primary, 3 years lower secondary, 3 years higher secondary and then higher education.[1] Instead of secondary education which prepares for higher education, pupils may also opt for technical or professional education. All pupils who want to continue their education have to participate in the Concours Nationale, an important test at the end of primary school. This test is administered by the Ministry of Education and all pupils have to take it on the same day. The test consists of four subjects: Mathematics, French, Kirundi (the national language) and Environmental Science. The latter a combination of geography and biology. The maximum score on the test is 200, with 70 points on the Maths, 80 points for the French test, 20 for Kirundi and 30 for Environmental Science.

Every year about 200,000 pupils take the test. Their performance in the test will determine their further school career and as such affect their opportunity to find a job and the level of income as an adult. It is considered a high-stakes exam in Burundi at the time. The 2% of pupils with the highest test scores, around 3500 every year, are allowed to continue into the Lycée, the secondary school tier that prepares for higher studies. Each year, a threshold is set that the pupil has to pass in order to be admitted to the Lycée. Depending on the difficulty of the exam of that year, that threshold varies to make sure that about 3500 pupils pass the threshold. In 2010, the threshold was set at 128/200, meaning that about 98% of the pupils had a lower score than this on the test. Until the late 1990's there was only one threshold. Since then, a second, lower threshold was introduced in line with the expansion of a second-tier secondary school system, called the communal colleges. In 2010, this second threshold was set at 81/200 and about 32% of the pupils obtained test scores between 81 and 128 and were thus allowed to enter a communal college. The remaining 66% who scored lower then 81/200 were not allowed to continue their education that year.[2]

[1] This was the case before the 2013 reform, the period relevant for this paper.
[2] The allocation of pupils to secondary school occurs at the national level. In a companion paper I have investigated the inequalities resulting from this: pupils from poor areas compete with pupils from rich areas such as parts of the capital Bujumbura for the same number of fixed seats in secondary school. I refer to Verwimp, P. (2023), Etno-regional Favouritism and the Political Economy of School Test Scores, Journal of Development Economics, October.

3.2 Burundi's Civil war: dispersion over time and space

As the exact timing and location of the civil war will play an important role for our identification strategy, we describe the evolution of the massacres and the civil war through time and space as follows:

- 1993 and 1994: massacres in many parts of the country but with different intensities
- End of 1994 to July 1996: Spread of civil war throughout the country.
- July 1996 to early 1999 : Return of Major Buyoya to power after a bloodless coup. Intense conflict in the capital but lower intensity in most provinces. International trade embargo in place.
- 1999-2003 : post-war era with signing of the Arusha Peace and Reconciliation Agreement in 2000. Sporadic violence continuing.

The early massacres were particularly intense in central and northern Burundi. Bundervoet (2009) estimates that in half of the provinces more than 7% of individuals interviewed in a large UNFPA survey lost their father in 1993. We present figures at the province level and sketch the evolution of the civil war based on Chrétien and Mukuri (2000). Fighting began in October 1994 in the northwestern provinces of Cibitoke, Bubanza, Bujumbura Rural and Ngozi. By early 1995, violence spread to the bordering Kayanza province, and by April 1995, massacres of civilians and confrontations between army and rebel forces happened in Karuzi, Bururi, Ruyigi and Muyinga. By late 1995, fighting took place in the central provinces of Gitega and Muramvya and the northern province of Kirundo. Map 1 depicts the situation at the end of 1995. By then, conflict had spread to almost all of the provinces of Burundi, with the exception of Cankuzo (in the east of the country) and Rutana and Makamba (in the south of the country). In July 1996, former president Buyoya seized power again in a bloodless coup d'état backed by the army. During late 1996 and early 1997, armed conflict continued in Kayanza, Muramvya, Kirundo and Gitega. Meanwhile in April 1997, the Arusha Peace talks between the principal conflict parties began. As of late 1997, insecurity increased again in Cibitoke, Bubanza and Bujumbura Rural, provinces which remained unsafe until 1999.

The various conflict accounts provide no definitive explanation for why the massacres and the civil war affected some provinces earlier than others. However, the conflict's spread was clearly influenced by the rebel base locations in the Democratic Republic of Congo's South Kivu region next to the borders of Cibitoke, Bubanza, and Bujumbura Rural, which explains why these provinces were first to experience war. The presence of the Kibira forest bordering these provinces could explain the subsequent spread of war to Kayanza and Ngozi, since rebels passed undetected through the forest. From these initial conflict provinces, the war spread to the rest of the country. Map 1 depicts the geographic spread of the civil war and Table 1 showing key figures for the duration and intensity of the violence.

¶

Table 1: Duration and Intensity of the Civil War

Province	Start and end of massacres and civil war	Duration of exposure in years at the province level		Province poverty headcount
		Max	average	in 1990
Bubanza	1995-2003	9	8.64	22.4
Bujumbura Rural	1995-2003	9	8.62	25.7
Bujumbura Capital	1996-1998	3	2.35	-
Bururi	1995-1996	2	1.5	37.7
Cankuzo	No	0	0	25.1
Cibitoke	1995-2003	9	8.63	19.6
Gitega (*)	1993-1997	5	3.02	35.2
Karuzi (*)	1993-1996	4	2.85	66.8
Kayanza (*)	1993-1996	4	2.64	44.9
Kirundo (*)	1993-1997	5	3.19	34.0
Makamba	1996-2003	8	7.82	-
Muramvya (*)	1993-1997	5	3.14	24.0
Muyinga (*)	1993-1996	4	2.57	27.8
Mwaro (*)	1993-1997	5	3.09	24.0
Ngozi (*)	1993-1996	4	2.78	42.5
Rutana	No	0	0	58.0
Ruyigi	1995-1996	2	1.68	41.0
Burundi	1993-2003	9	3,72	36.2

(*) these provinces were the scene of massacres in 1993 whereby more than 7% of respondents in a nationwide survey (the country-level median, UNFPA, 2002) lost their father, reason why we start the period of hostilities there in 1993.

4. Data, Definitions and Preliminary Observations

4.1 Data overview

We combine several datasets. The first dataset is an administrative record held by the Ministry of Education in Burundi which contains the results of the Concours National for the years 2010 to 2012. The Registry held by the Ministry of Education contains the family name and first name of the student, the province and municipality of residence, the year of birth, the gender of the child, the name of the school, the exam results for four academic disciplines and the ranking of each student. It does not contain household data or data on the characteristics of the school. In the three years under consideration, approx. 200,000 pupils took this test each year, reaching a total of approx. 600,000 pupils in the pooled dataset. Table 2 describes key variables in the data.[3]

We also have two datasets capturing exposure to civil war. The *first* of these is a self-constructed, historical dataset on the start, duration and end of hostilities in each of the Burundese provinces, as presented in Table 1. That dataset, which has been used elsewhere[4], determines for each province-year whether or not violent conflict was affecting the province in that year. There are 17 provinces in Burundi which were affected at different times during the decade-long war. Combining these historical data with the year of birth and residence data from the Registry above yields our first war exposure variable (see further details below). *The second war exposure data* come from the Uppsala Conflict Data Program (UCDP, https://ucdp.uu.se/). This program gives for the entire period of the civil war the number of fatalities at the municipality level (or commune as termed in Burundi). This is the administrative level below the province. We will use the number of fatalities at the municipality-year level indicated as "best" figure in the UCDP dataset as a measure of the severity of conflict. This represents our second violence variable.

[3] In the companion paper mentioned in footnote 2, I have documented irregularities in the grading of the exams in the province of Ngozi, reason why I perform robustness analysis in Appendix Table A1 where this province is excluded

[4] see for example, Verwimp and Van Bavel, 2014, *World Bank Economic Review*, vol.28, n2, pp.384-411, and Bundervoet et al, 2009, *Journal of Human Resources, 44 (2), 536-563.*

Table 2: Key indicators for the Concours Nationale 2010-2012

Year of the test Indicator	2010	2011	2012
Number of participants [(a)]	188,573	201,239	221,606
% participants are female	51	53	54
Average age at time of the test (stand.dev)	15.84 (1.6)	15.83 (1.7)	15.83 (1.7)
Average score (max is 200) (stand.dev)	70.5 (25.6)	63.0 (23.11)	125.7 (26.0)
% participants never exposed to civil war	8.14	23.5	43.3
Av.# of years of exposure during childhood	4.16 (2.7)	3.80 (2.8)	3.26 (2.9)
Av. # of fatalities during childhood [(b)]	120	118	111

(a) excluding pupils who enrolled for the test but did not show up (5%) ; (b) measured as the number of fatalities in the municipality of residence

4.2 Exposure variables

To examine the degree of exposure to civil war in childhood we utilize Table 1 in combination with the year of birth of the child to find the duration of exposure to conflict for each child. We define our two exposure variables as follows: (i) the number of years that the child was exposed in its province of residence during different stages in childhood and (ii) the number of war fatalities registered in the child's municipality at different stages in childhood.

Starting with the first variable, exposure means that fighting was taking place between warring parties in the province of residence at a specific time during childhood. In this exposure variable, the exact nature of the confrontation of the child with violence is not specified. It can involve the destruction or closure of schools, the absence of teachers, children kept at home, etc (Obura 2008). This war exposure variable is defined at the province-year level for the entire duration of the civil war. In the binary case, we code the exposure variable '1' if the child was exposed in the specific stages of his/her childhood and '0' if (s)he was not exposed. In the continuous case, we calculate the number of years of exposure in each stage of childhood. The degree of detail in our data allow us to distinguish between exposure to civil war at five stages in childhood: *while in utero, at birth, in early infancy (ages 1-3), late infancy (ages 4-6) and primary (ages 7-12).*

An example of exposure at different stages in childhood: a child born in 1998 and residing in the province of Makamba, will be exposed to civil war in his year of birth and in the first few years of live. As civil war was going on in the province from 1996 to 2003, the war-exposure dummy variable at the province-year level will have value 1 for the year *in utero* (1997), *the year of birth* (1998), the three years of *early infancy* (ages 1-3, 1999-2001), the first two years of *late infancy* (ages 4-6, 2002 and 2003) and no exposure thereafter (*in the age of primary*, 7-12). Calculating the duration of exposure during his entire childhood (in utero up to age 12 included) we arrive at 7 years out of 14.

Our definition of the exposure variable has *advantages and disadvantages*: <u>first</u>, as a province is a relatively large area, our variable may capture children that are affected less by the violence (eg. because the fighting occurs at the other end of the province) compared to other children (eg. residing in the fighting zone) who may be heavily affected. This measure does not capture the within-province variation in exposure. The solution to this is to provide a more fine-grained measure of exposure at the local level, which will be our second measure of exposure (see below). In defence of our exposure variable, *I want to highlight several particularities of the civil war in Burundi that give merit to a province-level measure of conflict exposure*: when violence occurs, some households leave the scene and migrate away. This can be for several days seeking temporary shelter, sometimes for weeks, months or even longer. The United Nations Population Fund (UNFPA) conducted a demographic household survey in 2001 and found over 50% of the rural population of Burundi had been displaced from their homes at least once between 1994 and 2000 due to the violence (UNFPA, 2002). Interestingly, the survey also documented that the overwhelming majority of households who are internally displaced remained in their own province (UNFPA, 2002, p.141).

Hence, under the condition of within-province displacement, using a province-level exposure variable is “migration-proof”. Meaning that households who sought refuge in a neighbouring municipality are coded as “exposed to violence” at the time of the violence, which they indeed are. This is particularly relevant in case the dataset does not contain information on migration, which is the case in our data on exam performance. Moreover, when violence occurs in a child’s province, but not in its village, it does not mean that the child is unaffected by it, even when the child is not displaced. Markets close, roads are blocked, armed rebels pass through the village, displaced households enter the area and so on. All of which occurred during the civil war in Burundi. Or, in the words of one seminar participant "*in some sense everyone in the province is exposed* ". Thus, with a province-level measure, residents who are further away from the fighting, possibly residing in a neighbouring municipality at that time, are also coded as "exposed to violence" at that time. Not modelling spatial spillovers, the province-level measure captures direct as well as indirect exposure of all residents from the same birth-cohort in a province and codes their exposure in the same way.

The alternative civil war variable that I propose is the number of fatalities in the municipality of residence of the child in each stage of childhood, calculated from the Uppsala Conflict Data Program. A child of age 12 for example taking the exam in 2010 in the commune of Mabanza will have the number of fatalities that occurred in Mabanza in the year preceding his year of birth (1997) as indicator of the child's exposure while in utero, and a similar allotment takes place for each subsequent period in the life of the child, similar to our first exposure variable described above. The number of fatalities are measured at the local level, making this is a more fine-grained indicator of exposure to violence. This measure also captures better the intensity of conflict compared to our first exposure variable. The limitations because of the absence of the month of birth however remain the same, reason why we employ the same three strategies to deal with that as mentioned above, in a robustness analysis.

Second, since the test-score dataset does not provide the place of birth nor does it feature migration variables, I have to assume that a child did not move to another province or municipality during its (early) childhood. In other words: I have to assume that the child spent his/her childhood in the area of the school where he/she participated in the Concours National. I acknowledge that this is a limitation of the study. However, for reasons specific to Burundi, this is not a far-fetched assumption and not a major limitation for two reasons : (i) every sector in Burundi (which is the administrative level below the municipality) has its own primary school. Hence it is not necessary to leave one's area of residence in order to send children to school. The practice of sending children away to boarding schools or to family members in another part of the country occurs for secondary schooling, but seldomly for primary; and (ii) when adults move residence for socio-economic reasons this mostly occurs at the time of marriage, when the new couple receives land to start their own household, or accept a job somewhere else, before they have children.

The later can be documented in a household panel of which the first wave was collected in 1998 in a nationally representative survey. In 2007, the Burundi Household Panel Survey reinterviewed a random selection of 1.000 households in 100 of the 390 villages sampled in 1998. The survey managed to track 872 parental households (with the same head or his spouse as in 1998) in their original location as well as 534 split-of households (newly formed households by a son or daughter who had left the parental household since 1998). From the 128 original households we did not manage to interview, 25 had migrated to another municipality (2.5%), were dissolved or had disappeared (2.5%), head and spouse had died (2%), or other unknown reason (2.5%). From the split-of group, 90% remained in the same municipality as their parents, 6% moved to another municipality within the same province and 4% moved out of the province. 86% of the first-born children in the newly formed households were born after the year in which the son/daughter left the parental household, 10% in the same year and 4% before (author's own calculations from the Burundi Household Panel Survey). Hence, these panel data show that moving to another province/municipality after household formation is a marginal phenomenon in Burundi in the first decade of the millennium.

And third, as we do not have the month of birth of a child, calculating *exposure in utero* based on the year of birth can yield an imprecise measure of exposure. If, for example the child is born in the months of October to December, then the *in utero* period is not the year preceding the year of birth, but this period is part of the year of birth. Whereas in the main analysis exposure in utero will be captured by inclusion of exposure in the year before birth, I employ three strategies as robustness to deal with the shortcoming of not having the month of birth[5]: (i) exclude the year before birth in the calculation of the length of exposure in childhood; (ii) exclude the year before birth as a separate shock in the regression analysis; and (iii) compare the impact of in utero exposure and after birth exposure between children that were exposed to only one of the

[5] For the exposure variables after birth (till age 12) the absence of the month of birth only has a small effect on the assignment of the periods of exposure, which encompass several years for each exposure variable. Hence the reason why I perform the robustness analysis only for *in utero exposure*.

two shocks. We do the latter to gauge the power of in utero versus after birth exposure. Each of these three strategies has its own limitations and neither can fully compensate the lack of data on the month of birth, but robustness analysis with all three can strengthen the findings from the main analysis. These are presented in Tables A2-4 in the Appendix.

4.3 Preliminary observations

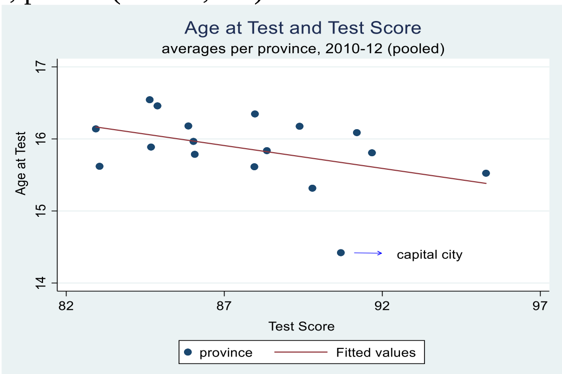
We start with the observation, in Table 3 and Figures 1 and 2 of provincial and municipality-level disparities in the age at which the national test is taken and the test score obtained. The average age at which pupils take the test is by far the lowest in the capital city Bujumbura.[6] And the difference between the province with the lowest average test score (Karuzi) and the highest one (Makamba) is more than 12 points, representing 14% of the country-level mean score. Plotting age-at-test and test scores we notice a negative relation: younger age-at-test correlates with higher test scores.

In table 4 we break down the data on age-at-test and test scores for children exposed to the civil war at least one year during their childhood versus children who were not exposed. We observe that the gap for age-at-test and for test scores between the exposed and non-exposed children is statistically significant at the 1% level, for any group we take (by gender, by residence, by poverty rate). This means, for instance, that poverty is unlikely to be the driving force behind the observed gap in age-at-test and test scores between exposed and non-exposed provinces, an issue to which we will return in section 5 when discussing identification. Nor is the gap a purely rural phenomenon. In fact, the gap is largest in the capital, an observation we will also return to later in the paper.

[6] Large variation exist between municipalities within the capital, reason why on average the capital as whole does not have the highest test score.

Table 3 and Figures 1 and 2 : Age at time of the national test and test score, 2010-2012, pooled (N=601,970)

Province	Age at test	Test score
		(max=200)
Bubanza	15.84	88.35
Buj. Rural	15.79	86.07
Buj. Cap.	14.42	90.69
Bururi	15.32	89.79
Cankuzo	16.17	89.38
Cibitoke	15.62	83.05
Gitega	16.18	85.86
Karuzi	16.14	82.94
Kayanza	15.89	84.68
Kirundo	16.54	84.64
Makamba	15.52	95.28
Muramvya	16.46	84.88
Muyinga	15.81	91.67
Mwaro	16.35	87.97
Ngozi	16.09	91.19
Rutana	15.61	87.95
Ruyigi	15.96	86.02
Burundi	15.83	87.55



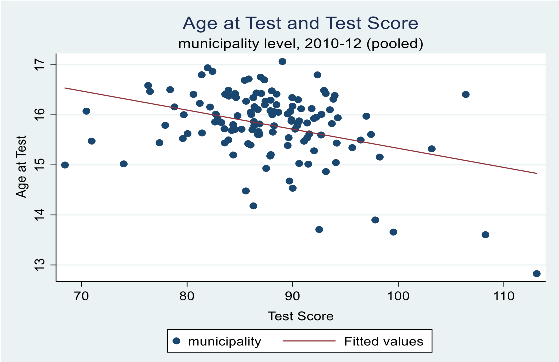


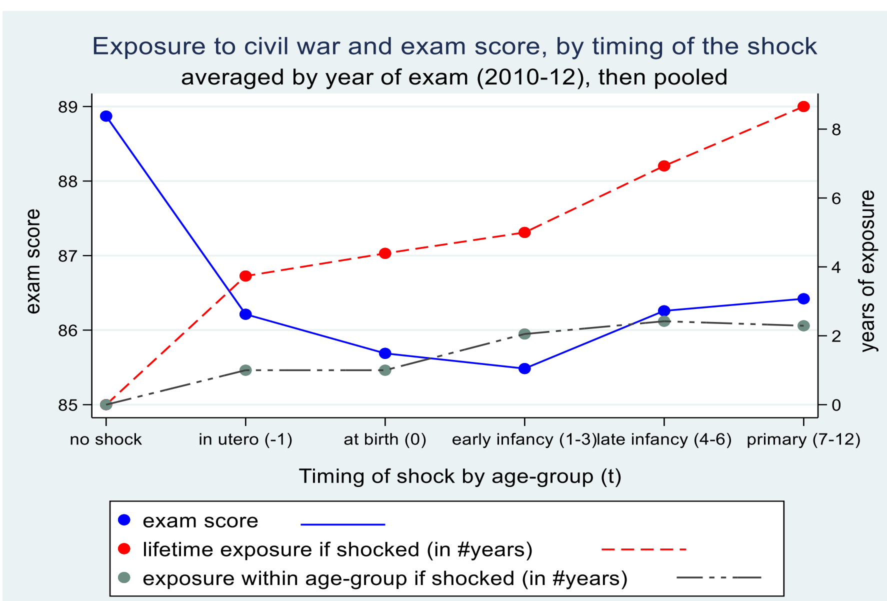
Table 4: Mean Age-at-test and Test Score for exposed and non-exposed children

	Mean Age at test				
	N	All	Not exposed	Exposed	Diff
Country	597,980	15.84	14.59	16.04	+1.45***
Boys	283,644	16,06	14.83	16,24	+1.41***
Girls	314,336	15,65	14.40	15.85	+1.46***
All but the capital	564,460	15.93	14.70	16.13	+1.42***
Capital	33,533	14.39	11.60	14.66	+3.06***
Poor Prov (a)	285,990	15.83	14.61	16.09	+1.48***
Nonpoor Prov (a)	278,470	16.03	14.86	16.16	+1.29***
	Mean Standarized Test score				
	N	All	Not exposed	Exposed	Diff
Country	597,980	0	0.110	-0.016	-0.126***
Boys	283,644	0.142	0.267	0.123	-0.144***
Girls	314,336	-0.127	-0.020	-0.144	-0.124***
All but the capital	564,460	-0.007	0.076	-0.020	-0.096***
Capital	33,533	0.127	1.022	0.042	-0.980***
Poor Prov (a)	285,990	0.022	0.046	0.017	-0.028***
Nonpoor Prov (a)	278,470	-0.037	0.129	-0.056	-0.185***

Note (a): poverty is defined using the headcount in 1990 (before the start of the war) which was calculated at 36.2% for the country. Provinces with a headcount above this % are considered poor, those below are considered nonpoor. Headcount poverty data for the capital were missing.

In Figure 3 we observe the red line (right axis) which gives the number of years that a child is exposed over its entire childhood - meaning from in utero to age 12 (thus a maximum of 14 years) - for children affected in different stages in childhood. Example: a child that is shocked at birth will be exposed on average 4 years during its childhood. And the green line (right axis) tells us the average duration of exposure for children exposed at a specific stage of childhood. The duration of exposure in utero and at birth is by definition 1 year as the coding for exposure is binary 0/1 in the year of birth and the year prior to the year of birth; in early infancy (1 to 3 years of age) as well as late infancy (4-6y) and primary (7-12y) the average duration of exposure is approximately 2 years. This means that exposed children receive most of their exposure before they reach primary school age. By the time the children in our sample reach that age, the civil war indeed is over in most provinces. We will use the timing of these shocks later in our analysis as we are interested in the impact of the shocks (and their duration) that occur at different stages in childhood.

Figure 3



We also observe the blue line (left axis) which gives the average test score for children affected by civil war at different stages in childhood, compared to non-affected children. The difference is on average approximately 3 points for a child who was exposed to a shock at any stage in childhood compared to a non-exposed child.[7]

In Figure 4 we use standarised test scores (which we will continue to use in the rest of the paper because they allow us to compare and pool data from different test years) as well as the number of fatalities a child encountered in its municipality of residence from the year before (s)he was born to the age of 12 (included), and we show the calculation by year of birth. The resulting graph has the shape of a whale, with the belly of the whale at its maximum for birth years 1995 and 1996, meaning the highest number of fatalities over the period under scrutiny together with the lowest test score attained in the Concours National. The curve of fatalities and test scores are each others mirror image: moving away from birth year 1995 to more recent or to older birth years, the number of fatalities over the period under scrutiny drops and test scores improve. Fatalities are the lowest and test scores the highest for the most recent birth years 1999 and 2000.

In Figure 5 we discuss the fatalities in more detail, showing their absolute number and relative importance in each stage of childhood (in utero, at birth, in early infancy, late infancy and primary) by year of birth. We observe that pre-war birth years (<1993) are only confronted with fatalities in later stages of childhood (late infancy and primary), whereas early-war (1993-1995) and mid-war birth years (1996-1998) confront fatalities in all stages of childhood and late-war birth years (1999-2000) confront fatalities mainly in early childhood stages.

[7] This does not take into account the variation in the level of difficulty in the annual tests. For that reason we will work with standarised scores in the remaining of the paper.

Figure 4

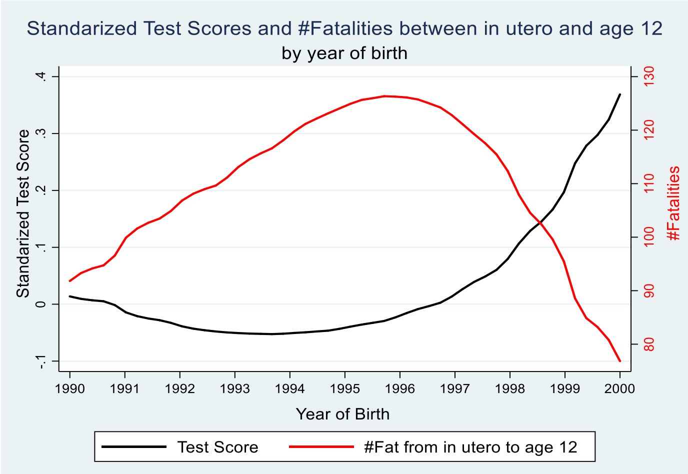
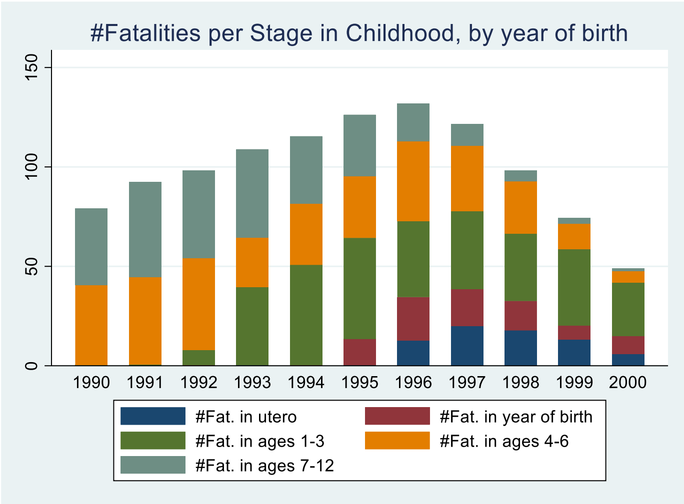


Figure 5



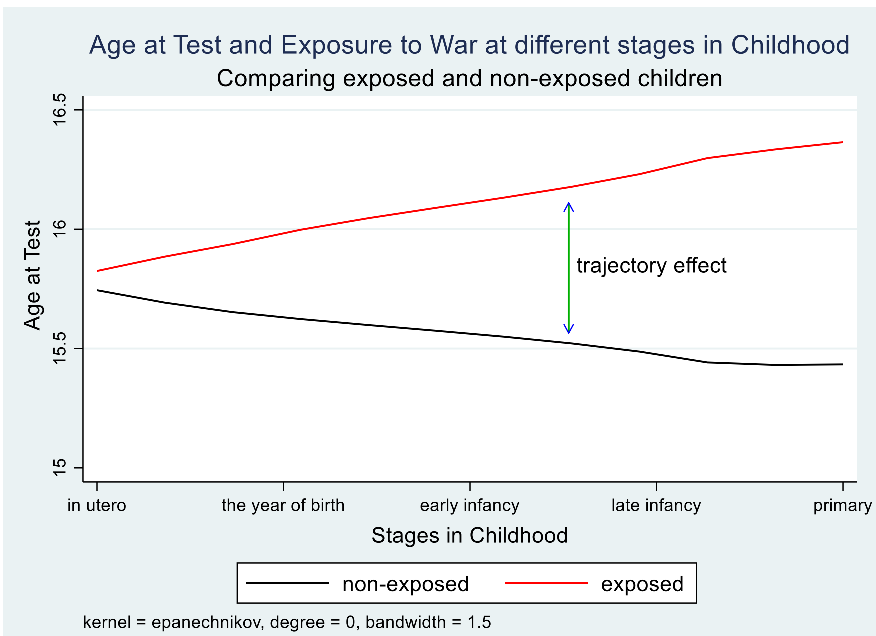
4.4 Defining the three components of the scar of civil war exposure

Naming the effect of exposure a 'scar' originates from the literature on unemployment where it refers to the long-lasting negative impact of an unemployment spell in future labour market outcomes, as if the person carries a 'scar' from such past adverse events. In a similar vein, children who are affected by a negative shock in (early) childhood bear the consequences of that later in their school career. The effect can be split in three parts: (i) *a trajectory effect*: the effect of exposure on the age at which the child takes the test; (ii) *a scarring effect*: the effect of age-at-test on the test score and (iii) *a cognitive effect*: the direct effect of war shocks on test scores, conditional on the age-at-test.

(i)The Trajectory Effect

Figure 6 shows that exposure to war leads to a higher age at which pupils take the exam, evidence of the delay in progress through the school system of an exposed child compared to a non-exposed, with more delay observed in the event of shocks occurring later in childhood (late infancy, primary) compared to early childhood (in utero, at birth, early infancy). For each stage in childhood, the mean age at test was computed for children who were exposed versus unexposed in that stage, not controlled for other factors, hence the unconditional mean. To be sure, the data sets do not have a variable indicating the year or the age at which the child entered primary school. Hence when we use the word ‘delay’ it covers both late entry into first grade as well as slower progress while in school, both resulting in older age-at-test.

Figure 6 : Trajectory effect: childhood shocks and the age-at-test



(ii) The Scarring Effect

Teenagers who take the exam at an older age obtain worse test scores, which we term the scarring effect. This effect is most outspoken in the capital city (Bujumbura) as Figure 7 shows. Only 3.300 of the 38.000 children in the capital were not exposed to the violence, and all of them took the exam between age 10-12. The average test score of these non-exposed urban children was +1 standard deviation away from the mean. Exposed urban children taking the test at an early age (11-15) have much higher grades than exposed children taking the test at a later age (16-20). The slope is the result of sharp differences between municipalities within the capital, of which some were hit hard by violence in the 1996-1998 period and others were not.

The scar also exists in the rural areas. The slope of the curve is less steep compared to urban areas (remark that the numbers on the y-axis in Figure 8 are a subset of those on the y-axis in Figure 7) . Exposure to civil war affects the slope, depending on the timing of the exposure, as we see in Figure 8: shocks early in infancy result in a steeper slope (=stronger scarring effects), whereas shocks later in childhood flatten the slope. The earlier the child was exposed, the steeper the slope. Above age 15 the slope remains steep for children affected early in childhood. This observation corresponds with the cumulative nature of the scarring effect: the earlier you are affected the stronger the accumulation of negative effects. Meaning that by the time a child who was exposed early in life takes the exam, (s)he will perform worse compared to a non-exposed child or a child exposed later in childhood.

Figures 7-8: the scarring effect

Figure 7: rural vs urban

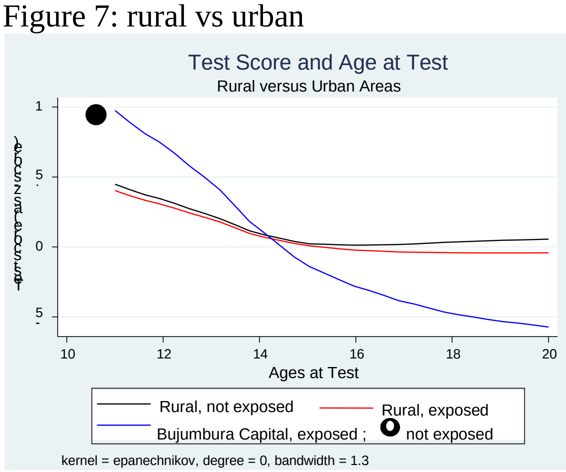
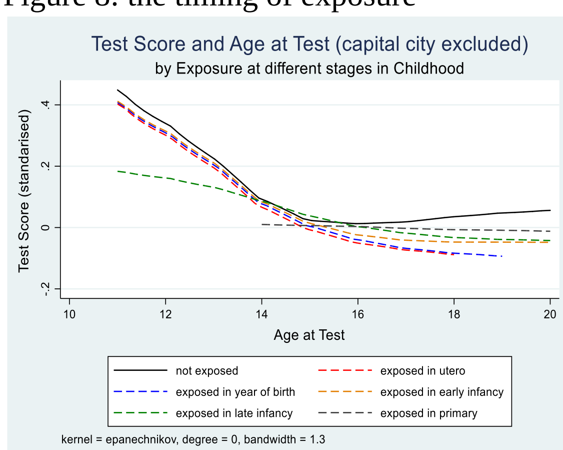


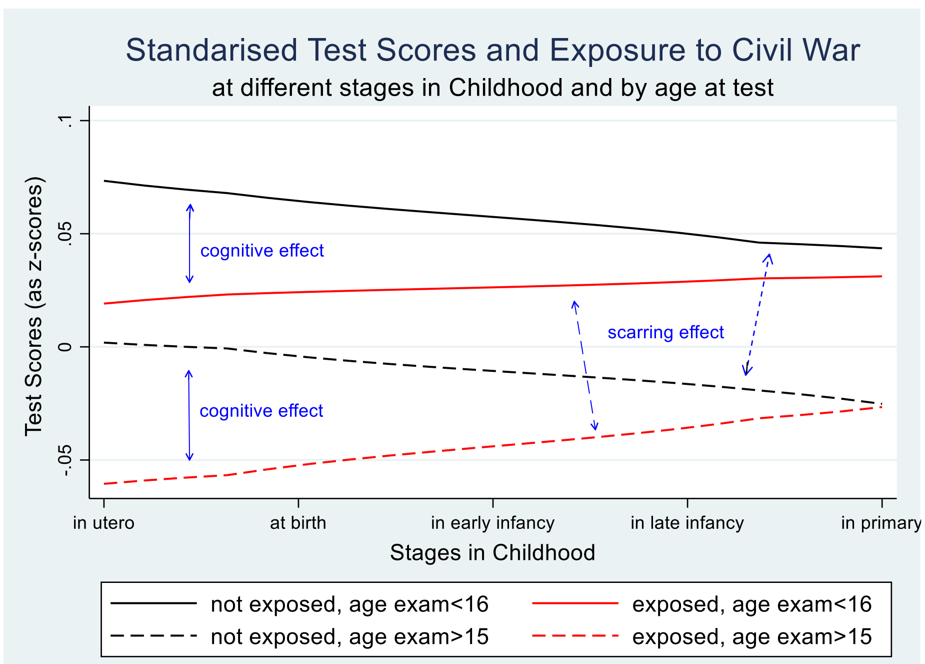
Figure 8: the timing of exposure



(iii)The Cognitive Effect

In Figure 9 we compare the test scores of older and younger children who are exposed at different stages in childhood with same-age children who are not exposed at that stage. We define the cognitive effect as the direct effect of war exposure on test scores, conditional on the age-at-test. We observe two issues, (i) once you control for the age-at-test - hence comparing the full black line with the full red line *or* the broken black line with the broken red line - the difference in test scores (i.e. the cognitive impact) between exposed an non-exposed children in largest in the event of in utero and early childhood exposure[8]; (ii) increased age-at-test, for exposed as well as non-exposed children (hence comparing the full and broken lines of the same colour) lowers the test scores, i.e. the scarring effect. Regression analysis will have to show the magnitude and statistical significance of these effects.

Figure 9: Scarring and Cognitive Effects



[8] A caveat here is of course that this binary relation does not take any of the fixed effects into account, which will be included latter in the regression analysis.

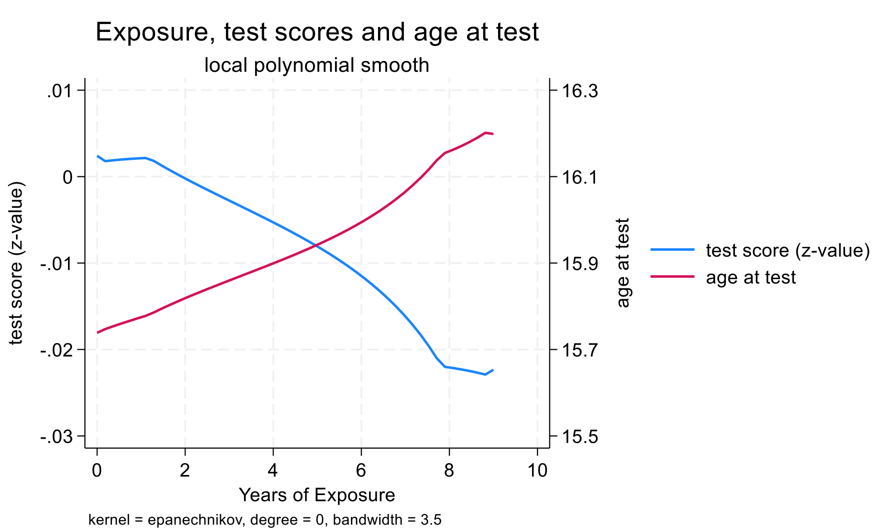
5. Identification and Econometric Specification

5.1 Identification and Potential Threats

Our empirical identification strategy can be illustrated by examining the nonparametric relationship between school test scores and age-at-test on the one hand and years of exposure to civil war on the other hand. In Figure 10 we estimate two kernel-weighted local polynomial regressions of, first, school test scores on years of exposure and, second, age-at-test on years of exposure, each time using an Epanechnikov kernel. The figure shows a considerable drop in average school test scores as well as an increase in average age-at-test for children who were exposed to additional years of civil war.

Our identification strategy exploits the variation of conflict over space (across provinces and municipalities) and over time (cohort of birth, the timing of the conflict and the year of the test). Even within a birth cohort not all children take the exam at the same age, which depends on their progression through the school system. The latter depends on their exposure to violence, which varies over provinces and municipalities. We distinguish between pre-war, early-war, mid-war and late-war birth cohorts with each birth cohort exposed to different levels of violence over the lifetime of a child. The presence of different birth cohorts guarantees that children are affected at different stages of their childhood with some children never affected.

Figure 10



Teenagers are between 10 and 25 when they take the Concours National with the bulk of pupils between ages 13 and 17 at the time of the test. We have test score results of three exam sessions, to wit the years 2010, 2011 and 2012. The variation over space comes from the geographic spread of the civil war, with provinces and municipalities (communes) affected differently throughout the course of the war. Some areas were never affected, some were severely affected for an extended period of time and some were affected for one or two years.

We argue that the duration and timing of exposure as well as the number of fatalities are not correlated with age-at-test and school test scores before the onset of the civil war. We take steps to show that these war events are plausibly exogenous to age-at-test and school test scores so that any effect we obtain from these war shocks can be interpreted as the causal impact of war on age-at-test and exam results. Since I do not have data on age-at-test and test scores dating from before the onset of the violence (1993), I refer to several analyses in previously published articles co-authored by the author that offer ample evidence for the near-exogeneity of violence with respect to pre-war household, local and province-level socio-economic, geographic and demographic characteristics. I am referring in particular to Bundervoet et al (2009); Voors et al (2012); Verwimp and Van Bavel (2014) as well as Mercier et al (2020).

First, reports by human rights organisations as well as scholarly articles (Human Rights Watch, 1998; Uvin, 1999; Krueger and Krueger 2007) demonstrate that violence in Burundi's civil war was largely indiscriminate because of the army's inability to identify rebels, but also by a desire for extermination, revenge, plundering and a perceived need to demonstrate power as part of the tactics of fear to control a population. This type of violence, Voors et al (2012, p.944) argue, "is near-exogenous to household characteristics and local economic conditions, hitting communities and civilians regardless of social status, education or income". Army and rebels arbitrary raided communities throughout the country. Using pre-war household, local and province-level demographic, economic and geographic characteristics collected from 1,000 households in 100 villages (mentioned above in section 4.2), Voors et al (2012) do not find correlations between these characteristics and subsequent occurrence of village-level violence or household-level victimisation during the civil war.[9]

In the same line of argument, Bundervoet et al (2009) find very few correlations between the timing of conflict onset (no, early or late) at the province level, the length of exposure to conflict at the individual level, and a range of household characteristics such as the level of education of the head of the household, his/her age, presence in the household or occupation using the same 1998 dataset. They use the same definition for the exposure to violence with the same data as I do here, which makes their analysis of exogeneity particularly applicable to the current paper.

Second, as poverty is a potential rival candidate for the impacts we are measuring on age-at-test and test scores, we want to make sure that we are not picking up this effect. In particular, if children in poor provinces go to school later and obtain worse test scores than children in nonpoor provinces and if it were the case that violence in the civil war was more intense or lasted longer in poor provinces, then we may be picking up the effect of poverty rather than the effect of violence. We respond to that potential threat with five pieces of evidence: (i) a higher % of children were exposed to massacres and civil war violence in non-poor provinces (89%) compared to poor provinces (80%) over the entire period (1993-2003), with the difference between the regions statistically significant at the 1% level; (ii) the eight provinces that were hit by an above median killing rate of 7% during the 1993 massacres had a higher poverty headcount (37.4% compared to 30.2%) but the difference is not statistically significant at the usual thresholds; (iii) when the civil war started at the end of 1994, the first affected provinces (Cibitoke, Bubanza and Bujumbura Rural) were among the richest in Burundi and were ranked first, second and fifth respectively in pre-war welfare rankings (Republic of Burundi, 2003); (iv) the five provinces that experiences civil war fighting early on (i.e. by January 1995) have lower headcount poverty (31.0%) compared to the nine provinces were the fighting started later (36.7%) and (v) Mercier et al (2020, p.830), using a household survey collected by UNFPA in 2002 with recall questions on pre-war (1993) assets, report that “average household-level wealth prior to violence exposure was not different between

[9] These 100 villages constitute a random subsample of the 390 villages visited during the 1998 Household Priority Survey for which village-level recall data on violence were collected.

subsequently exposed and non-exposed communes". Taken together, the evidence strongly suggests that there was no selection into violence based on pre-war household, local of province level characteristics that can be correlated with age-at-test or test scores. Children exposed to violence were not residing predominantly in poor provinces.

Another potential selection mechanism is participation in the national exam at the end of primary school. Children who were heavily affected by the civil war, most likely do not participate in the nationwide test. They may not even have attended school or dropped out from primary school before reaching grade 6. Indeed, as demonstrated in Verwimp (2012), low height-for-age z-scores at time (t-x) predict death and survival at time (t). While we do not study enrolment, grade attainment or drop out in this paper, it is clear that several instances of selection have already taken place before pupils participate in the test. The first is mortality. Infant and child mortality is high in Burundi. Children who die from illness, malnutrition or otherwise will never participate in the tests. Second is enrolment. Children who never enrol in school will not be able to study in primary school and participate in the test. Third is drop out. Pupils who drop out after a few years of primary school will not sit the test. And fourth is participation. Not all pupils who finish primary school participate in the test. In particular the children who do not want to continue their education do not take the test. These four reasons and in particular the first three can be affected and indeed amplified by the exposure to civil war. More infants and children die during the civil war, more pupils do not enrol and more pupils drop out. In this respect the results presented in this paper should be regarded as a lower bound of the effects of exposure to civil war on cognitive development. For papers that investigate enrolment and grade attainment in Burundi and Rwanda, I refer to Verwimp and Van Bavel (2014) and Bundervoet and Fransen (2018). We come back to other selection issues when we discuss potential mechanisms driving the results in this paper, in particular a gender-selection mechanism.

5.2 Econometric specification

Building on Figure 10 and previous figures, we estimate a specification that consists out of two parts. First the age-at-test specification, analysing the delay in progression through the school system, or in other words the detection of the trajectory effect, whereby

$$Age\text{-}at\text{-}test_{ijt} = a_1 + \beta_1 G_i + \beta_2 \sum Shock_{i,t-x} + \lambda_t + \mathfrak{z}_t + \Theta_j + v_{ijt} \qquad (1)$$

with Age-at-test the age of child *i* in school *j* at the time *t* of the test, G is the gender of the child and Shocks are a number of shocks that take place at different stages in childhood (*in utero, at birth, in early infancy, in late infancy, in primary*) represented by subscript *x*-years in the past. And λ the cohort of birth fixed effects, $\mathfrak{z}$ the year of exam fixed effects and Θ the location fixed effects. The latter can be at the level of the province, the municipality or the school. v_{ijt} is an idiosyncratic error term.

And the second part estimates the test score:

$$Test\ score_{ijt} = a_2 + \beta_3 Age\text{-}at\text{-}test_{ijt} + \beta_4 G_i + \beta_5 \sum Shock_{i,t-x} + \lambda_t + \mathfrak{z}_t + \Theta_j + Age\text{-}at\text{-}test_{ijt} * \Theta j + w_{ijt} \qquad (2)$$

with Test score the result of child *i* from school *j* taking the test at time *t* ; Age-at-test the age of the child *i* from school *j* at the time of the test and the other symbols as in equation (1). We also include a location specific time trend by interacting the province or municipality of residence dummy variables with age-at-test.

As both the test score and the age-at-test are determined by the civil war shocks, leaving out age-at-test from equation (2) would overestimate the effect of the civil war shocks on the test score. Hence, we bring the two equations together in *a conditional mixed process estimator* (CMP). CMP fits a large family of multi-equation, multi-level estimators which have a recursive structure, meaning the models have clearly defined stages. In our case exposure in childhood determines the progression made in the school system followed by a test at the end of primary school. "Mixed process" means that different equations can have different kinds of dependent variables, in our case age-at-test and test score. And a dependent variable in one equation can appear on the right side of another equation. The cohort of birth, year of test and location of residence fixed effects control for unobservables at the level of the birth cohort, the year of the test and the location.

One caveat to CMP is that we cannot include a fixed effect at the school level because the school variable is correlated with one of the regressors (age-at-test) and as a result the CMP model does not converge, reason why we perform the CMP regression only with a province and municipality fixed effect. Running the regressions separately as TWFE with school fixed effects (as we will show below) and comparing with municipality fixed effects in the CMP regression yields very similar results.

6. Regression Results

with exposure in years

We start the empirical work with the impact of civil war exposure on the age at which children take the test. The delay caused by the exposure slows down the progression of the pupil in the school system. In the first column of Table 5 we observe that a marginal increase of exposure during childhood (*from in utero to age 12*) increases the age-at-test by (0.405***) years. This effect is somewhat smaller for boys (0.376***, see column 2) and somewhat larger for girls (0.436*** see column 3). Column 1 however also reveals that girls take the test 0.1 years earlier. When we specify the shocks for different stages in childhood in Table 6, we find that exposure in each stage increases the age-at-test, an effect that is statistically significant in all stages, except for exposure in utero. The magnitudes of the effects are somewhat stronger for girls compared to boys. In addition, the magnitude of the effect of the late-childhood shocks on the age-at-test is larger compared to early-childhood shocks, demonstrating the widening gap in age-at-test between exposed and non-exposed children for shocks in later infancy compared to early infancy. This is the trajectory effect, as depicted in Figure 6.

Continuing with the second specification estimated in our CMP model, having the test score as dependent variable, we notice that increased age-at-test decreases the test score by (-0.041***) standard deviations on average for each extra year of exposure and that this effect is almost three times larger for boys (-0.059***) compared to girls

(-0.021***). This is the scarring effect of civil war. Hence, while, overall, girls take the test at a slightly earlier age, and exposure to civil war increases that age for girls more than for boys, this increase has a smaller adverse effect on girl's test scores compared to boys. This is confirmed in Table 6, where the effect of age-at-exam on test scores for boys is (-0.039**) and for girls it is (-0.10).

We also notice that the direct effect of civil war exposure on test scores (cognitive effect) conditional on age-at-test, is larger for girls (-0.041***) compared to boys (-0.036***). This finding is confirmed in Table 6 where we split the civil war shocks in different stages in childhood. The direct effect is larger (more negative) for girls in all stages except for exposure at birth. In Table 6 we also observe that the coefficient of the shock variables is not smaller for shocks in late infancy compared to early infancy, (columns 2, 4 and 6) something one may have expected based on Figure 9. In fact, only the coefficient of the shock occurring between ages 4 and 6 seems small and statistically insignificant. In section 7 below we will calculate the magnitude of the effects in points on the exam, but here we can already state that a decrease of 0.036 standard deviations as a result of one year of exposure means about a one point less in the exam.

Since we are using the Registry data for this estimation (N=600,000) we remind that we do not have information on the date or year of school entry nor on potential repetition of grades. Hence, we do not know if the later age at which the child takes the test is the result of later entry into school, grade repetition, teacher absence or the closing down or destruction of school buildings, or a possible combination of these reasons.

Table 5 : Civil War Exposure, Age-at-test and School Test Scores, 2010-2012 (pooled), CMP regressions *with years of exposure from in utero to age 12*

	Column 1		Column 2		Column 3	
Dependent variables: Age-at-test and Test Scores in the Concours National (as z-score) Regressors	Exposure variable is continuous (in years), All		Exposure variable is continuous (in years), Boys only		Exposure variable is continuous (in years), Girls only	
	Age-at-test	Test score	Age-at-test	Test score	Age-at-test	Test score
Expo in utero to 12	0.405*** (0.071)	-0.036*** (0.011)	0.376*** (0.071)	-0.036*** (0.013)	0.436*** (0.074)	-0.041*** (0.009)
Age-at-test		-0.041*** (0.003)		-0.059*** (0.011)		-0.021*** (0.003)
Gender of the child	-0.095*** (0.007)	-0.279*** (0.012)				
Constant	14.65*** (0.671)	1.203*** (0.064)	15.01*** (0.667)	1.543*** (0.098)	14.11*** (0.074)	0.625*** (0.052)
Year of test FE	Yes		Yes		Yes	
Cohort of birth FE	Yes		Yes		Yes	
Province FE	Yes		Yes		Yes	
Prov.spec. time tr.	No	Yes	No	Yes	No	Yes
N	597,979		283,644		314,336	

Note: standard errors clustered at the level of the shock (province) in all regressions.

Table 6 : Civil War Exposure, Age-at-test and School Test Scores, 2010-2012 (pooled), CMP regressions *with years of exposure in different stages of childhood*

Dependent variables: Age-at-test and Test Scores in the Concours National (as z-score) Regressors	Exposure variable is continuous (in years), different phases in childhood, all		Exposure variable is continuous (in years), different phases in childhood, boys only		Exposure variable is continuous (in years) , different phases in childhood, girls only	
	Age-at-test	Test score	Age-at-test	Test score	Age-at-test	Test score
Expo in utero	0.125 (0.160)	-0.035*** (0.008)	0.066 (0.173)	-0.028*** (0.011)	-0.018 (0.143)	-0.047*** (0.009)
Expo at birth	0.595*** (0.169)	-0.033** (0.014)	0.574*** (0.185)	-0.041** (0.016)	0.601*** (0.152)	-0.032** (0.015)
Expo age 1-3	0.487*** (0.147)	-0.039*** (0.015)	0.464*** (0.142)	-0.041** (0.018)	0.519*** (0.152)	-0.043*** (0.014)
Expo age 4-6	0.964*** (0.174)	-0.030 (0.020)	0.887*** (0.167)	-0.008 (0.022)	1.081*** (0.196)	-0.011 (0.026)
Expo ag 7-12	0.810*** (0.132)	-0.058** (0.025)	0.804*** (0.137)	-0.063** (0.025)	0.799*** (0.126)	-0.064** (0.029)
Age-at-test		-0.027 (0.018)		-0.039** (0.017)		-0.010 (0.021)
Gender of the child	-0.067*** (0.010)	-0.280*** (0.012)				
Constant	11.03*** (0.465)	0.914*** (0.215)	11.44*** (1.297)	1.202*** (0.224)	10.301*** (1.338)	0.393* (0.219)
Year of test FE	Yes		Yes		Yes	
Cohort of birth FE	Yes		Yes		Yes.	
Province FE	Yes		Yes		Yes	
Prov.spec. time tr.	No	Yes	No	Yes	No	Yes
N	597,979		283,643		314,336	

Note: standard errors clustered at the level of the shock (province) in all regressions.

with the Number of Fatalities

In Tables 7 and 8 we take the exact same CMP model, but we use our second measure of civil war exposure: the number of fatalities registered in one's municipality of residence. We use (log) Fatalities at the municipality level in each stage of childhood as a measure of the intensity of armed conflict in that location at that time. We obtain results that are very similar to those obtained with the years of exposure variable presented in Tables 5 and 6. In particular, the three core findings discussed above are reproduced in

Tables 7 and 8, strengthening the results: (i) fatalities in the entire period of childhood increase the age-at-test, in particular fatalities in late childhood (primary age)[10]; (ii) increased age-at-test lowers the test score substantially and the effect is statistically significant at the 1% level, with a magnitude of the coefficient for boys that is twice as large as that for girls, and (iii) the direct effect of the shocks on test scores, conditional on age-at-test, is negative and statically significant at the 1% and 5% level, in particular for girls and more so for fatalities in early-childhood compared to late childhood.

Table 7: Fatalities, Age-at-test and School Test Scores, 2010-2012 (pooled), CMP regressions with *Fatalities from in utero to age 12 as a one variable*

Dependent variables: Age-at-test and Test Scores in the Concours National (as z-score) Regressors	(log)Fatalities, all		(log)Fatalities, Boys only		(log)Fatalities, Girls only	
	Age-at-test	Test score	Age-at-test	Test score	Age-at-test	Test score
(log)Fatalities, from in utero to 12	0.234** (0.041)	-0.028** (0.014)	0.200** (0.038)	-0.018 (0.013)	0.264*** (0.046)	-0.040*** (0.015)
Age-at-test		-0.037*** (0.002)		-0.053*** (0.002)		-0.024*** (0.002)
Gender of the child	-0.097*** (0.004)	-0.279*** (0.008)				
Constant	17.18*** (0.225)	1.023*** (0.075)	17.43*** (.208)	1.272*** (0.081)	16.79*** (0.253)	0.561*** (0.090)
Year of test FE	Yes		Yes		Yes	
Cohort of birth FE	Yes		Yes		Yes	
Municipality FE	Yes		Yes		Yes	
Munic.spec time tr.	No	Yes	No	Yes	No	Yes
N	597,979		283,643		314,336	

Note: standard errors clustered at the level of the shock (municipality) in all regressions

[10] The sign of the coefficient for fatalities in utero and at birth in Table 8 is different compared to the exposure in years variable in Table 6, something we will discuss in the mechanism section.

Table 8: Fatalities, Age-at-test and School Test Scores, 2010-2012 (pooled), CMP regressions *with Fatalities in different stages in childhood*

Dependent variables: Age-at-test and Test Scores in the Concours National (as z-score)	(log)Fatalities, different phases in childhood, all		(log)Fatalities, different phases in childhood, boys only		(log)Fatalities, different phases in childhood, girls only	
Regressors	Age-at-test	Test score	Age-at-test	Test score	Age-at-test	Test score
(log)Fat. in utero	-0.061** (0.031)	-0.007* (0.003)	-0.062* (0.032)	-0.005 (0.004)	-0.057* (0.030)	-0.008** (0.004)
(log)Fat. at birth	-0.063* (0.032)	-0.011*** (0.004)	-0.079** (0.033)	-0.007* (0.004)	-0.049 (0.032)	-0.013*** (0.004)
(log)Fat. age 1-3	0.033 (0.030)	-0.012** (0.005)	0.045 (0.039)	-0.011** (0.005)	-0.020 (0.031)	-0.013** (0.006)
(log)Fat. age 4-6	0.020 (0.035)	-0.010* (0.005)	0.007 (0.043)	-0.007 (0.006)	0.035 (0.037)	-0.015** (0.006)
(log)Fat. age 7-12	0.235*** (0.004)	-0.004 (0.006)	0.215*** (0.045)	-0.003 (0.006)	0.260*** (0.047)	-0.007 (0.006)
Age-at-test		-0.048*** (0.004)		-0.061*** (0.005)		-0.036*** (0.005)
Gender of the child	-0.085*** (0.004)	-0.279*** (0.008)				
Constant	17.40*** (0.434)	1.154*** (0.110)	17.63*** (0.438)	1.392*** (0.123)	16.97*** (0.463)	0.692*** (0.116)
Year of test FE	Yes		Yes		Yes	
Cohort of birth FE	Yes		Yes		Yes	
Municipality FE	Yes		Yes		Yes	
Munic.spec time tr.	No	Yes	No	Yes	No	Yes
N	597,979		283,643		314,336	

Note: standard errors clustered at the level of the shock (municipality) in all regressions

Including School Fixed Effects

In the previous analyses we did not include school-level FE in our CMP estimation because the school variable is correlated with age-at-test as a result our estimation did not converge. Using a new STATA command, *reghdfe*, which is a linear regression allowing the absorption of multiple levels of fixed effects as well as multi-way clustering, we are able to include school FE in a TWFE estimation. It is a novel and robust algorithm to efficiently absorb the fixed effects (extending the work of Guimaraes and Portugal, 2010), making it very useful for the data we use because we have more than

3000 schools and 128 municipalities. One drawback is that *reghdfe* only estimates one equation, with the test score as independent variable. It does not simultaneously estimate a second equation with age-at-test as the independent variable. However, when using *redhdfe* to estimate the coefficient of age-at-test as well as the exposure variables in our regression of test scores yields very similar results as the ones reported above with CMP. Hence we are confident that the use of this new command produces reliable results in a TWFE regression of school test scores *with* school FE. We of course keep the caveat in mind that our unit of analysis remains the child, and not the school, because our data is a repeated cross-section of schools, not a panel of children.

Table 9 presents the results of the school FE effects regressions, for our two exposure variables and for girls and boys. As before, we include location-specific time trends, which are, as before, at the province level for the regression using exposure in years and at the municipality level for the regression using fatalities. The results for the exposure in years variable and the age-at-test variable in the school FE regression are very similar to the results presented earlier in columns 4 and 6 of Table 6 using province FE. And the results for the fatalities variable and age-at-test are very similar to the results presented in columns 4 and 6 of Table 8 using municipality FE. Seeing almost all of our previous results confirmed in the school FE regressions supports our claims that shocks in (early) infancy have detrimental effects on school test scores as a teenager.

As we cluster the standard errors at the level of the shock/treatment in all previous regressions, and given the debate on clustering in the literature, we also want to use other types of clustering to observe the difference. Abadie et al (2023) have demonstrated that in several instances clustered standard errors are unnecessarily large and robust standard errors are too small. They argue that the framework in which econometricians study clustering is motivated by a sampling mechanism in which researchers in a first stage select clusters at random from on infinite population and in a second stage select units randomly from the selected clusters (Abadie et al, 2023, p.5). In many applications in economics and social science however, researchers use data where they observe units from all the clusters they are interested in (as in the data used in this paper) and a framework based on randomly sampling a small fraction of a large population of clusters does not apply. In their new clustering framework the above authors explicitly include a design component so that the standard errors are not only computed based on the sampling procedure but also on the treatment assignment mechanism.

Following that discussion we provide in Table 9 estimation results with clustered standard errors as well as with robust standard errors. The magnitudes of the robust standard errors prove about ½ to 1/3 of the magnitudes of the clustered standard errors. This difference impacts in particular the results in column 3 (school FE regression of boys with fatalities as exposure indicator) where some coefficients turn statistically significant with robust standard errors but not with clustered standard errors. In the other three columns it does not make much of a difference.[11]

Table 9 : TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, exposure in years and Fatalities, *for different stages in childhood* and with clustered std.errors vs robust std.errors.

Dependent variable: Test Scores in the Concours National (as z-score) Regressors		Exposure in years in different phases in childhood		(log)Fatalities in different phases in childhood	
		Boys	Girls	Boys	Girls
Expo/Fat in utero	Coefficient Clust. Std.err Robust Std.err	-0.026 (0.009)** (0.006)***	-0.040 (0.008)*** (0.005)***	-0.002 (0.003) (0.002)	-0.008 (0.003)** (0.002)***
Expo/Fat at birth	Coefficient Clust. Std.err Robust Std.err	-0.039 (0.009)*** (0.006)***	-0.030 (0.012)** (0.005)***	-0.005 (0.004) (0.002)**	-0.010 (0.004)*** (0.002)***
Expo/Fat age 1-3	Coefficient Clust. Std.err Robust Std.err	-0.025 (0.012)** (0.004)***	-0.032 (0.008)*** (0.004)***	-0.009 (0.005)* (0.003)***	-0.011 (0.005)** (0.002)***
Expo/Fat age 4-6	Coefficient Clust. Std.err Robust Std.err	-0.017 (0.018) (0.007)**	-0.026 (0.021) (0.007)***	-0.005 (0.005) (0.003)*	-0.010 (0.005)** (0.002)***
Expo/Fat age 7-12	Coefficient Clust. Std.err Robust Std.err	-0.052 (0.019)** (0.007)***	-0.060 (0.023)** (0.007)***	-0.002 (0.006) (0.004)	-0.003 (0.006) (0.003)
Age-at-test	Coefficient Clust. Std.err Robust Std.err	-0.045 (0.016)** (0.008)***	-0.008 (0.018) (0.008)	-0.070 (0.004)*** (0.010)***	-0.050 (0.004)*** (0.012)***
Constant	Coefficient Clust. Std.err Robust Std.err	0.795 (0.079)*** (0.05)***	0.01 (0.08) (0.04)	0.94 (0.055)*** (0.048)***	0.219 (0.054)*** (0.043)***
Year of test FE		Yes	Yes	Yes	Yes
Cohort of birth FE		Yes	Yes	Yes	Yes
School FE		Yes	Yes	Yes	Yes
Loc. spec.time tr.		Yes	Yes	Yes	Yes
N		283,643	314,336	283,643	314,336

[11] As explained in footnote 3, we perform a robustness analysis of table 9 in Appendix (Table A1) without Ngozi province. And in Tables A2 and A3 we perform a robustness without taking the year before birth into account, as explained in section 4.2. And Table A4 compared children who are only exposed in utero to children who are only exposed after birth.

To improve on the precision of causal inference in cases where clustering is appropriate Abadie et al (2023) propose the *Causal Cluster Variance Estimator* (ccv) which explicitly includes a design component in the calculation of the standard errors. Implementing ccv in STATA we can only use one shock/treatment at a time, and it must be in binary form. Importantly, the standard errors can only be calculated when there is *variation in treatment within a cluster*. The latter is difficult to attain in our data because the years of exposure variable assigns the same treatment to all cohorts within the same province (=the cluster), and the number of fatalities indicator assigns the same treatment for all cohorts within the same municipality (=the cluster). Hence, the only way to arrive at within-cluster variation in treatment is to widen the cluster (i.e. take the province instead of the municipality) when using the fatalities variable as our exposure indicator. When doing so, just for the purpose of this robustness check, we only find within-cluster variation using the primary age shock (age 7-12). For all other shocks we do not find this variation. Table 10 compares the magnitudes of the three standard errors (clustered, robust and ccv) in a FE regression on test scores. The magnitude of the ccv standard errors situates itself halfway between the magnitudes of the clustered and robust standard errors.

Table 10: Magnitudes of the standard errors in a FE regression with Causal Cluster Variance

Boys					95% Confidence Interval	
Variable			z	P>|z|	Lower B	Upper B
At least one fatality in age 7 to 12	Coeffient	-0.121				
	ccv std.error	(0.014)***	-8.66	0.000	-0.149	-0.094
	clust.std.error	(0.032)***	-3.84	0.000	-0.183	-0.059
	robust st.error	(0.005)***	-26.35	0.000	-0.130	-0.112
Girls					95% Confidence Interval	
Variable			z	P>|z|	Lower B	Upper B
At least one fatality in age 7 to 12	Coeffient	-0.130				
	ccv std.error	(0.015)***	-8.41	0.000	-0.160	-0.100
	clust.std.error	(0.035)***	-3.76	0.000	-0.197	-0.062
	robust st.error	(0.004)***	-31.38	0.000	-0.138	-0.122

In summary, while our main analysis is performed with standard errors clustered at the level of the treatment, this may be too conservative as our data do not contain a sample of the population but indeed represent the universe of school test scores for all children in all schools. When we apply robust standard errors (in Table 9), some shocks in (early) childhood, in particular in column 3 (boys, fatalities) turn statistically significant. While this step may be not conservative enough, we also apply the ccv procedure proposed by Abadie et al (2023), which we can only test for the treatment/shock in the age 7 to 12 in binary from, and find that it is statistically significant at the 1% level for boys as well as girls (Table 10).

7. Interpretation and Discussion

Magnitude of the three components

As for the **trajectory effect** is concerned (we look here at Table 1, first column, with age-at-test as the dependent variable), we notice that one additional year of exposure from in utero to age 12 increases the age-at-test with 0.40 years. The average duration of exposure for the exposed in our data being 4.3 years this means that, on average, exposure to civil war increases the age-at-test by 1.72 years (in comparison: the maximum duration of exposure is 9 years, which increases the age-at-test by 3.6 years).

In order to calculate **the scarring effect** (defined as the effect of age-at-test on the test score), we are looking at the CMP regression in Table 1, column 2 and notice that the coefficient of the age-at-test variable is -0.041. We need to multiply this with the increase in the age-at-test as a result of the average duration of exposure, calculated as 1.72 years. That yields (-0.041*1.72=-0.070) which is expressed in standard deviations, as the dependent variable is the z-score of the grade obtained in the test. As one standard deviation in the grade equals approx. 25 points in each of the exam years, it means that the scarring effect of the increase in age-at-test resulting from the average duration of exposure to civil war is -0.070*25 points or -1.75 points.

Last but not least we calculate **the cognitive effect** of war exposure, which is the direct effect of exposure conditional on age-at-test. Again from Table 1, column 2 we learn that the coefficient of our exposure variable is -0.036, which, for an average length of 4.3 years of exposure this means -0.15 standard deviations. In terms of the grade obtained this equals -0.15*25 or -3.75 points. We perform the same calculations for boys and girls, also on the basis of the regression results presented in Table 1.

Hence, **in summary**, taking the entire sample in consideration, a child with an average duration of war exposure (4.3 years) during its entire childhood (in utero till age 12) will take the test on average 1.72 years later compared to a non-exposed child and will obtain on average 5.50 fewer points on the test, which equals approx. 5% of the average grade. The 'scar' of exposure is the sum of the scarring (-1.75 points) and the cognitive (-3.75 points) effect. As the Concours Nationale is a competitive exam to win a seat in secondary school, this reduction in points is enough to loose ones seat in secondary school. In the table 11 we present the magnitude of these effects for boys and girls separately as well as for poor and nonpoor provinces and for the capital Bujumbura. While the reduction of points as a results of civil war exposure is approx. 6.42 points for boys and 5.27 points for girls, we observe that the loss for boys resulting from the scarring effect counts for 38% of their overall decrease in the grade and the loss in terms of cognition counts for 62%. For girls these are respectively 18% and 82%, meaning that girls loose more in terms of cognition.

Summary Table 11: Magnitude of the scar of civil war exposure

Effect of years of civil war exposure	All	Boys	Girls	Poor Provinces	Nonpoor Provinces	Capital only
Y.of exposure (if exposed)	4.3	4.4	4.2	3.34	5.47	2.57
Trajectory Effect (in years)						
Marginal effect	0.40	0.38	0.44	0.48	0.33	1.35
Average effect	1.72	1.67	1.85	1.59	1.8	3.47
Scarring effect (in points)						
Marginal effect	-1.02	-1.47	-0.52	-0.45	-0.8	-2.75
Average effect	-1.75	-2.46	-0.97	-0.71	-1.44	-9.54
Cognitive effect (in points)						
Marginal effect	-0.90	-0.90	-1.02	-0.55	-0.45	-5.75
Average effect	-3.75	-3.96	-4.30	-1.84	-2.46	-15.81
Overall effect (Average, in points)	-5.50	-6.42	-5.27	-2.55	-3.86	-25.35
N	597,979	283,644	314,336	285,990	278,470	33,543

*Note: poor provinces are provinces with a headcount poverty above the country average, which was 36.2% in 1990. The capital city of Bujumbura is excluded here as no poverty headcount is available.

When we look at children in poor versus nonpoor provinces we observe that the nonpoor provinces are exposed to civil war substantially longer compared to the poor provinces, resulting in larger losses from the civil war, -3.86 points versus -2.55 points to be exact. It is the capital Bujumbura however that faces the strongest losses, with a reduction of no less than 25 points for children with the average length of exposure.

Quantile regression

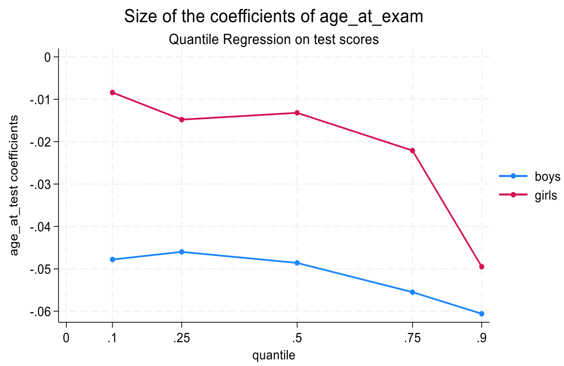
We also would like to know if the scarring effect and the cognitive effect of exposure to conflict varies over the distribution of test scores. To find out the magnitude of these effects, we run quantile regressions using the two conflict variables that we used earlier: exposure to conflict (as before measured in number of years of exposure to conflict - from in utero to age 12 - in the province of residence) and the number of fatalities – again from in utero to age 12 - in the municipality of residence).

In Figures 12 and 13, I am presenting the coefficients of age-at-test (scarring effect) and exposure to conflict (in #years from in utero to age 12) obtained in quantile regressions of test scores, for boys and girls. We notice that the size of the coefficients of age-at-test is stronger for boys (more negative) compared to girls, across the entire distribution of test scores and it is stronger for boys as well as girls in the top performing quantiles. In addition, the gender gap is the largest in the poorest performing quantiles.

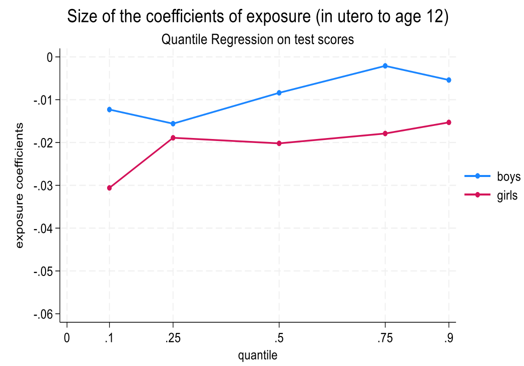
Considering the cognitive effect we see the opposite: the effect is smaller for boys compared to girls and it becomes smaller for both genders in the top performing quantiles. A similar picture emerges in Figure 14 and 15 where we present the findings using the number of fatalities.

Figure 12 and 13: The scarring and cognitive effects across the distribution of test scores, with years of exposure to conflict

Scarring Effect

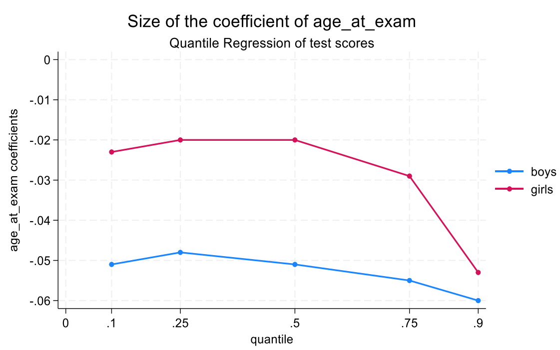


Cognitive Effect

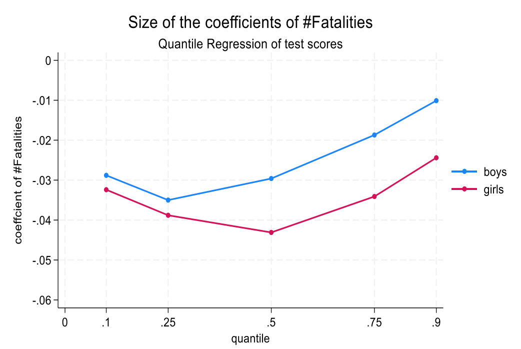


Figures 14 and 15: The scarring effect across the distribution of test scores, with #Fatalities

Scarring Effect



Cognitive Effect



Digging a little deeper into the cognitive effect, we want to find out if the timing of the shock in childhood makes a difference along the distribution of test scores. In Figures 16 and 17 we present the magnitude of the coefficients for both conflict variables at different stages in childhood (in utero, at birth, in early infancy, in late infancy and in primary age), just as we did in the main regression. I am presenting the coefficients and their level of statistical significance for the .25 and the .75 quantiles, for boys and girls. In case of the exposure to conflict variable, it seems that the effect of exposure *in utero* and *in primary* is larger (meaning more negative) compared to the three other stages in childhood, for all quantiles and for boys as well as girls. The shock at the other three stages in childhood (*at birth, in early infancy and in late infancy*) is statistically not significantly different from zero. We find that boys in the 0.75 quantile seem to incur the smallest cognitive impact from the conflict. Girls, across the distribution, suffer the largest cognitive impact, with coefficients below or at the same level of magnitude than the boys from the 0.25 quantile, a pattern that is confirmed (to a large degree) when using the #Fatalities.

Figure 16: Exposure in years at different stages in childhood

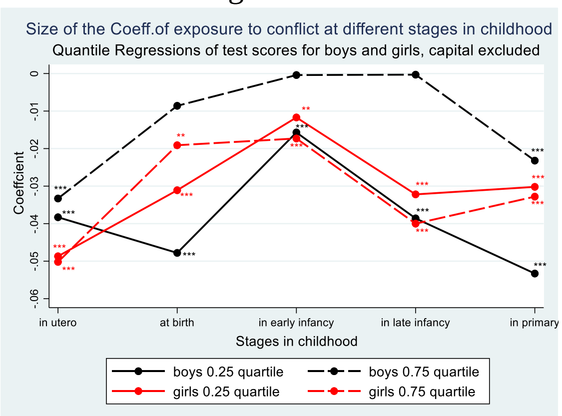
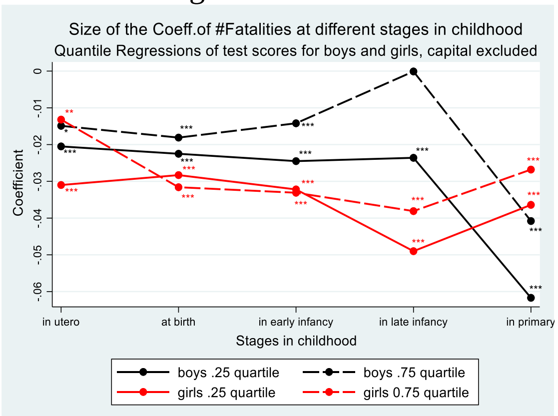


Figure 17: #Fatalities at different stages in childhood



8. Results per Academic Discipline

Next to the effects of the shocks at different stages in childhood along the distribution of test score results, we can also investigate the score for individual academic disciplines separately. Recall that the test score we have been using until now is a composite score of 4 tests including mathematics, French, environmental science and Kirundi, with the maximum number of points to obtain respectively 80,70, 30 and 20 for each test. In tables 12 and 13 we present the result for each discipline, using z-scores as before in each stage of childhood, for boys and girls. The estimation includes school fixed effects and location-specific time trends (province-level for our first exposure variable and municipality-level for the second). Standard errors are clustered a the location level. The TWFE specification is estimated with the *reghdfe* command in STATA (see above). Results are very similar to the CMP command.

In table 12, were we use the first exposure variable, we notice that test score results for boys are affected stronger (more negative) by the age at which they do the test (scarring effect) compared to girls for each academic discipline. This is in particular the case for mathematics. The direct effects of the childhood shocks, conditional on age (cognitive effect), proof to be worse for girls for mathematics, but not for French. The in utero shock is always worse for girls, and the birth shock worse for boys. We come back to this in the mechanism section.

Table 12:TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, exposure in years, *for different academic disciplines and stages in childhood*

Dependent variables: Test Scores for four disciplines (as z-score) Regressors	Mathematics		French		Environmental Science		Kirundi	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Expo in utero	-0.018* (.01)	-0.037*** (.007)	-0.019 (.012)	-0.027*** (.009)	-0.037*** (.007)	-0.042*** (.006)	-0.017** (.007)	-0.012 (.011)
Expo at birth	-0.024** (0.01)	-0.023** (.013)	-0.049*** (.011)	-0.029** (.011)	-0.034*** (.010)	-0.019** (.009)	-0.006 .(012)	-0.010 (.015)
Expo age 1-3	-0.012 (0.01)	-0.027*** (0.007)	-0.042*** (.013)	-0.031*** (.008)	-0.027** (.010)	-0.026*** (.008)	-0.012 (.015)	-0.001 (.0018)
Expo age 4-6	-0.005 (0.016)	-0.028* (0.02)	-0.03** (.012)	-0.021 (.013)	-0.027 (0.017)	-0.025 (.02)	-0.012 (0.022)	-0.011 (.031)
Expo ag 7-12	-0.033** (0.014)	-0.059** (0.02)	-0.061** (.021)	-0.046** (.021)	-0.048** (0.018)	-0.046** (.02)	-0.018 (.036)	-0.012 (.043)
Age-at-test	-0.047*** (0.013)	-0.007 (0.02)	-0.025 (.016)	-0.014 (.016)	-0.025 (.015)	-0.006 (.016)	-0.010 (.028)	0.000 .(034)
Constant	0.70*** (.077)	0.41 (0.078)	0.64*** (.067)	0.16** (.066)	0.575*** (.077)	-0.074 (0.082)	0.24* (.11)	0.008 (.14)
Y. of test FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coh.of b. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pr.sp. time tr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	283,463	314,179	283,500	314,199	283,458	314,165	283,463	314,169

In table 13, where we use the number of fatalities in each phase in childhood as our measure of conflict, we see an even more outspoken gender pattern: boys test scores are more affected through the age at which they take the exam (scarring), whereas girls suffer more the direct impact (cognitive), with mathematics as a clear example: no statistically significant impact on test scores for boys along the cognitive path. In fact, we only see direct impact here for boys for French, not for the other subjects.

Table 13: TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, number of Fatalities, *for different academic disciplines and stages in childhood*

Dependent variables: Test Scores for four disciplines (as z-score) Regressors	Mathematics		French		Environmental Science		Kirundi	
	Boys	Girls	Boys	girls	Boys	Girls	Boys	Girls
(log)Fat in utero	-0.004 (0.004)	-0.006* (.003)	-0.003 (.004)	-0.010*** (.003)	-0.000 (.0004)	-0.005 (.003)	-0.003 (.004)	-0.000 (.004)
(log)Fat at birth	-0.002 (0.003)	-0.007* (.006)	-0.009** (.009)	-0.014*** (.004)	-0.002 (.005)	-0.007** (.004)	-0.000 .(005)	-0.000 (.005)
(log)Fat age 1-3	-0.005 (0.004)	-0.009** (0.009)	-0.015*** (.005)	-0.014** (.005)	-0.005 (.005)	-0.002 (.005)	-0.003 (.006)	0.003 (.006)
(log)Fat age 4-6	-0.003 (0.004)	-0.012*** (0.046)	-0.013*** (.005)	-0.012** (.015)	-0.004 (0.005)	-0.002 (.005)	-0.006 (0.005)	0.005 (.0005)
(log)Fat age 7-12	-0.003 (0.005)	-0.0036 (0.0061)	-0.004 (.006)	-0.006 (.005)	-0.003 (0.006)	-0.002 (.005)	-0.002 (.008)	0.004 (.009)
Age-at-test	-0.063*** (0.004)	-0.058*** (0.004)	-0.06*** (.005)	-0.036*** (.004)	-0.043*** (.005)	-0.013*** (.004)	-0.014** (.005)	0.008 .(005)
Constant	0.797*** (.055)	0.169*** (0.05)	0.85*** (.054)	0.357*** (.051)	0.66*** (.059)	0.046 (0.117)	0.25*** (.061)	-0.005 (.065)
Y. of test FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coh.of b. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
M.sp.time tr	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	283,463	314,179	283,500	314,199	283,458	314,165	283,463	314,169

9. Mechanisms

Across our regressions we find that the test scores of girls are more negatively affected by in utero exposure compared to the test scores of boys. We observe that from the magnitude of the coefficient in tables 6 and 8 (0.28 vs 0.47 and 0.005 vs 0.008 respectively). Moreover, in table 8 the coefficient for in utero exposure for boys is not statistically significant at the usual thresholds, whereas for girls it is at 5%. This pattern is repeated when we use robust standard errors in table 9 and it is also repeated in the analysis per scientific discipline in tables 12 and 13.

One mechanism to explain this effect is sex-specific selection in utero or at birth. When more boys die in utero or at birth as a results of civil war exposure, the surviving boys would on average be stronger compared to the surviving girls. Such reduction could in turn be the result of two intertwined effects: (i) a “culling effect”, if sex-specific selection in utero is significant, surviving male children would be stronger since in utero shocks have more detrimental effects on boys than girls (Almond and Currie, 2011). Selection effects are likely to be particularly severe for large-scale shocks such as famines and civil wars (Neelsen and Stratmann, 2011; G rgens et al., 2012). This would mean that more girls than boys survive in utero exposure and that the surviving boys are on average stronger. Or (ii) an “in utero scarring effect” , resulting from a downwards shift of the entire foetal health distribution as a consequence of the negative shock, reducing the health endowment of each foetus in the population at the time of the shock. Since the low end of the health distribution is disproportionally populated by male foetuses, a reduction of the sex ratio follows. Even though both mechanisms result in a sex ratio reduction, they have a different impact on the health endowment of the surviving population: while the culling effect would generate a population of relatively more healthy individuals (notably among boys), the in utero scarring effect would imply the opposite.

To disentangle both mechanisms, we should see two types of evidence in the data: (i) the cohort of girls that survive in utero exposure to civil war shocks should be larger (meaning more numerous) than the cohort of boys exposed to the same shock, and (ii) in the regression analysis, in utero shocks should have less impact on the test scores of boys compared to girls, because the surviving boys are stronger. Weaker girls survive and are thus present in the data, whereas weaker boys do not survive in utero exposure and do not show up in the data.

Figures 18(a) and (b) offer evidence for the first type of effect: Figure 18(a) presents male/female infant death rations in the first 30 days after birth, calculated from the 2010 Burundi DHS. It shows that exposure in utero increases the male/female death ratio, supporting the culling effect. We refer to Sanders and Stoecker (2015) for using changes in the sex ratio of live births as a proxy for changes in fetal health. Figure 18(b) adds evidence to this: it depicts the % of female participants in the national test, thereby comparing children who are exposed and not exposed to civil war shocks at different phases in childhood. We observe that the sex composition of the cohort of participants in the test who are exposed to in utero shocks is very different compared to the sex composition of cohorts that are exposed later in childhood. In the first instance, girls are in the majority (56.5% and even 58.3% when we only take children affected by in utero shocks and not by shocks later in childhood), whereas in the second case girls are in the minority (47%). For non-exposed children it is exactly the opposite.

Figure 18a

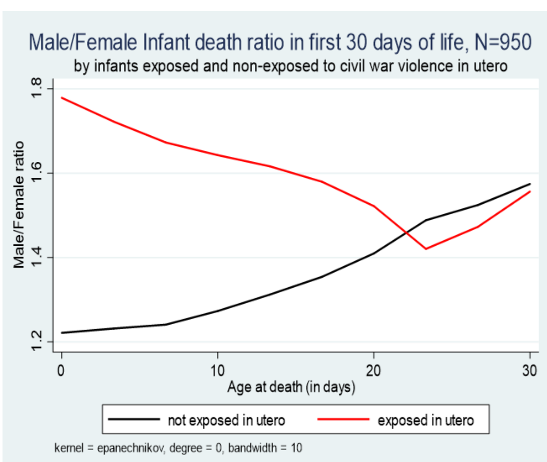
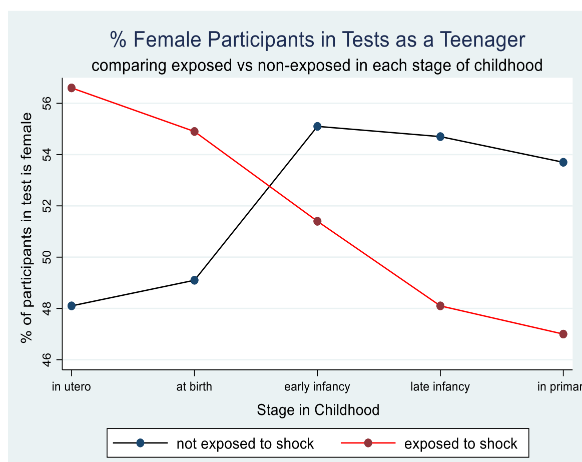


Figure 18b



And our regression analysis in many of the tables presented above offers evidence for the second mechanism: in utero shocks do not, or to a much lesser extent impact test scores of boys compared to girls. I conclude from this that the culling effect may dominate for boys and the ‘scar’ of civil war exposure is particularly felt by the surviving girls.

10. Conclusion

This paper uses the universe of test scores from a high-stakes nationwide exam for three consecutive years in Burundi to infer the impact of civil war exposure in (early) childhood on cognitive development as a teenager. The duration and spread of the civil war in Burundi offers an ideal setting for such investigation as different birth cohorts, residing in different provinces, are affected at different moments during childhood. We discussed the plausibly exogeneous spread of the civil war vis-à-vis the age-at-test and the test scores at length.

We use two different measures of exposure, we employ different estimation techniques, we look at various scientific disciplines, we perform quantile regressions and we discuss a potential mechanism. To the best of our knowledge, we are the first to conceptualize four distinct effects on war exposure in (early) childhood on test scores as a teenager in one framework: the culling effect, the trajectory effect, the scarring effect and the cognitive effect that, taken together, constitute the overall impact of civil war exposure on test scores. CMP allows us to test three of these effects simultaneously. As the CMP specification did not converge when we introduce school fixed effects, we complement our analysis with TWFE that arrives at very similar results as in the CMP estimation.

Our results show that exposure increases the age-at-test by on average 1.72 years (trajectory effect), which lead to less points at the test, in particular 5.5 points less of which 1.75 can be attributed to the scarring effect and 3.75 to the cognitive effect. Boys are more affected by the scarring effect compared to girls, whereas the latter are more affected by the cognitive effect. Robustness analysis does not change the main results. When we specify the different shocks in childhood separately, we find that boys are not, or to a lesser degree, affected by in utero shocks. We find evidence for sex-specific selection in utero whereby the weakest boys do not survive exposure when they are in utero. As a result, the surviving cohort of boys is relatively stronger, explaining the stronger cognitive impact of exposure in childhood on girls.

While the lay person as well as the scholarly community is well-aware of the detrimental effects of war on humans, the literature has quantified this impact mostly on the extensive margin. This paper quantifies the magnitude of the impact of civil war shocks in (early) childhood on the intensive margin. The impact on test scores and thereby on human capital formation 10 years later, when the children have become teenagers, is statistically significant and the magnitude of the effect is substantial. As such, the paper demonstrates the long-term impact of civil war for cohorts exposed in (early) childhood.

References

Abadie, A., S.Athey, G.W.Imbens, J.M.Wooldridge, 2023, When should we adjust standard errors for clustering ? , The Quaterly Journal of Economics, vol.138, issue 1, pp.1-35

Akresh, Richard; de Walque, Damien. 2008. Armed Conflict and Schooling : Evidence from the 1994 Rwandan Genocide. Policy Research Working Paper; No. 4606. World Bank.

Akresh, R., P.Verwimp and T.Bundervoet, 2011, Civil War and Child Stunting in Rwanda, Economic Development and Cultural Change,

Alderman, H., J Hoddinott, B Kinsey, 2006, Long term consequences of early childhood malnutrition, Oxford economic papers 58 (3), 450-474

Almond, D. and J. Currie, 2011, Killing me softly: The fetal origins hypothesis The Journal of Economic Perspectives, 25 (3), pp. 153-172

Almond, D., J. Currie, V. Duque, 2018, Childhood circumstances and adult outcomes: Act II, Journal of Economic Literature, 56 (4), pp. 1360-1446

Angrist, J.D., and Adriana D. Kugler, 2008, Rural Windfall or a New Resource Curse? Coca, Income, and Civil Conflict in Colombia , The Review of Economics and Statistics 90 (2): 191–215.

Behrman, J.R., John Hoddinott, John A. Maluccio and Reynaldo Martorell, 2009. "Brains versus Brawn: Labor Market Returns to Intellectual and Health Human Capital in a Poor Developing Country," Middlebury College Working Paper Series 0907, Middlebury College, Department of Economics.

Boone, P., Ila Fazzio, Kameshwari Jandhyala, Chitra Jayanty, Ganghadar Jayanty, Simon Johnson, Vimala Ramachandran, Ana Filipa Silva, Zhaoguo Zhan, 2014, The surprisingly dire situation of children's education in rural West Africa: results from the CREO study in Guinea-Bissau (Comprehensive Review of Education Outcomes), African Successes, Volume II: Human Capital, University of Chicago Press, Chicago, IL , pp. 255-280

Brück, T., Di Maio, M. and S.M.Miaari. 2019. Learning The Hard Way: The Effect of Violent Conflict on Student Academic Achievement, Journal of the European Economic Association, Vol.17 (5), pp.1502–1537

Bruckner, T.A., S. Helle, E.Bolund and V.Lummaa, 2015. Culled males, infant mortality and reproductive success in a pre-industrial Finnish population, Proceedings of the Royal Society B, January, Vol.282, Issue 1799

Bundervoet, T., P Verwimp, R Akresh, 2009, Health and Civil War in Rural Burundi, Journal of Human Resources 44 (2), 536-563

Bundervoet, T. 2009, Livestock, Land and Political Power: the 1993 Killings in Burundi, Journal of Peace Research, 46(3): 357-76

Bundervoet, T. and S.Fransen. 2018, The educational impact of shocks in utero: Evidence from Rwanda, Economics & Human Biology, Vol.29, May, pages 88-101

Chamarbagwala, R and H.E.,Morán, 2011, The human capital consequences of civil war: Evidence from Guatemala, Journal of Development Economics, Volume 94, Issue 1, Pages 41-61

Currie, J., and Tom Vogl, 2013, Early-life health and adult circumstance in developing countries, Annual. Review of Economics, 5 (1), 1-36

Dagnelie,O., G. De Luca and J.F.Maystadt, 2018, Violence, Selection and Infant Mortality in Congo, Journal of Health Economics,

De Walque, D., 2006, The socio-demographic legacy of the Khmer Rouge period in Cambodia, Population studies 60 (2), 223-231

Dickerson, A., S McIntosh, C Valente, 2015, Do the maths: An analysis of the gender gap in mathematics in Africa, Economics of Education Review 46, 1-22

Duque, V. 2024. Violence and Children's Education: Evidence From Administrative Data, Journal of Conflict Resolution, Vol. 68(5), pp. 903–937

Evans, D.K. and E Miguel, 2007, Orphans and schooling in Africa: A longitudinal analysis, Demography 44 (1), 35-57

Gregg, P., and E.Tominey, 2005, The wage scar from male youth unemployment Labour Economics 12 (4), 487-509

Guimaraes, P., P Portugal, B de Portugal, 2010, A simple feasible procedure to fit models with high-dimensional fixed effects, Stata Journal 4

Human Rights Watch, 1998, Proxy Targets: Civilians in the War in Burundi, New York

Kämpfen, F. , F Zahra, H.P. Kohler and R Kidman, 2022, The effects of negative economic shocks at birth on adolescents' cognitive outcomes and educational attainment in Malawi, SSM-Population Health 18, 101085

Krueger, R. and K. T. Krueger. 2007. "From Bloodshed to Hope in Burundi: Our Embassy Years During Genocide. Focus on American History Series." Tech. rep., Austin: University of Texas Press.

León, G., 2012. " Civil Conflict and Human Capital Accumulation: The Long-term Effects of Political Violence in Peru, The Journal of Human Resources 47 (4), 991-1022

Mercier, M. L.Ngenzebuke and P.Verwimp, 2020, Violence Exposure and poverty: Evidence from the Burundi Civil War, Journal of Comparative Economics

Mroz, T.A., and T.H. Savage, 2006, The long-term effects of youth unemployment Journal of Human Resources 41 (2), 259-293

Obura, A. 2008, Staying Power: struggling to reconstruct education in Burundi since 1993, UNESCO,

Republic of Burundi, 2003. Interim strategic framework for accelerating economic growth and reducing poverty. Poverty reduction strategy paper, Bujumbura

Rodríguez, C. and F.T. Sánchez, 2009, Armed Conflict Exposure, Human Capital Investments and Child Labor: Evidence from Colombia, Documentos CEDE 5400, Universidad de los Andes, Facultad de Economía, CEDE.

Sanders, N.J., and C. Stoecker, 2015, Where have all the young men gone? Using sex ratios to measure fetal death rates, Journal of Health Economics, 41 (2015), pp. 30-45

Shemyakina, O., 2011, The effect of armed conflict on accumulation of schooling: Results from Tajikistan, Journal of Development Economics 95 (2), 186-200

Stein, Z., Susser, M., Saenger, G., & Marolla, F. (1975). Famine and human development: The Dutch hunger winter of 1944-1945. Oxford University Press.

Schultz TP, 2002, Wage Gains Associated with Height as a Form of Health Human Capital, AEA Pap. Proc. 92 (2):349-53.

UNFPA, 2002, Enquête Démographique et de Santé, Bujumbura

Uvin, P., 1999, Ethnicity and Power in Burundi and Rwanda: Different Paths to Mass Violence, Comparative Politics 31(3): 253–71

Valente, C., 2015, Civil conflict, gender-specific fetal loss, and selection: a new test of the Trivers–Willard hypothesis, Journal of health economics 39, 31-50

Verwimp,P., 2012, Undernutrition, subsequent risk of Mortality and Civil War in Burundi, Economics and Human Biology, Vol. 10, Issue 3, pp. 221

Verwimp,P. and J.Van Bavel, 2014, Schooling, Violent Conflict and Gender in Burundi", World Bank Economic Review, vol.28, n2, pp.384-411

Verwimp, P, 2023, Ethno-regional Favouritism and the Political Economy of School Test Scores, Journal of Development Economics,

Voors, M., E.Nelissen, P.Verwimp, E.Bulte, Lensink, R. and D.Van Soest, 2012, Violent Conflict and Behavior ? Evidence from field experiments in Burundi", American Economic Review, 2012, April, Vol. 102, No. 2 pp. 941-64

Appendix

Robustness analysis without the inclusion of the province of Ngozi for reasons of potential manipulation of school exam results, see footnote 2. Results are very similar.

Table A1: TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, exposure in years and Fatalities, *for different stages in childhood* and with clustered std.errors, comparing results with and without the inclusion of the province of Ngozi.

Dependent variable: Test Scores in the Concours National (as z-score) Regressors		Exposure in years in different phases in childhood		(log)Fatalities in different phases in childhood	
		Boys	Girls	Boys	Girls
Expo/Fat in utero	With Ng. Without Ng.	-0.026 ** (0.009) -0.027 ** (0.009)	-0.040***(0.008) -0.037*** (0.008)	-0.002 (0.003) -0.001 (0.003)	-0.008** (0.003) -0.006** (0.003)
Expo/Fat at birth	With Ng. Without Ng.	-0.039***(0.009) -0.038*** (0.010)	-0.030** (0.012) -0.025** (0.011)	-0.005 (0.004) -0.003 (0.004)	-0.010***(0.004) -0.008*** (0.003)
Expo/Fat age 1-3	With Ng. Without Ng.	-0.025**(0.012) -0.021 (0.012)	-0.032*** (0.008) -0.033*** (0.008)	-0.009*(0.005) -0.007 (0.004)	-0.011** (0.005) -0.009* (0.004)
Expo/Fat age 4-6	With Ng. Without Ng.	-0.017 (0.018) -0.017 (0.017)	-0.026 (0.021) -0.027 (0.021)	-0.005 (0.005) -0.003 (0.005)	-0.010** (0.005) -0.009* (0.005)
Expo/Fat age 7-12	With Ng. Without Ng.	-0.052** (0.019) -0.055** (0.019)	-0.060** (0.023) -0.063** (0.024)	-0.002 (0.006) -0.003 (0.006)	-0.003 (0.006) -0.003 (0.006)
Age-at-test	With Ng. Without Ng.	-0.045** (0.016) -0.041** (0.016)	-0.008 (0.018) -0.005 (0.019)	-0.070*** (0.004) -0.068*** (0.004)	-0.050***(0.004) -0.048*** (0.004)
Constant	With Ng. Without Ng.	0.795*** (0.079) 0.805***(0.080)	0.01 (0.08) 0.01 (0.09)	0.94*** (0.055) 0.94*** (0.056)	0.219***(0.054) 0.204*** (0.053)
Year of test FE		Yes	Yes	Yes	Yes
Cohort of birth FE		Yes	Yes	Yes	Yes
School FE		Yes	Yes	Yes	Yes
Location sp.time tr.		Yes	Yes	Yes	Yes
N	With Ng. Without Ng.	283,643 267,236	314,336 298,075	283,643 267,236	314,336 298,075

Robustness analysis without the inclusion of exposure in the year before the year of birth, because we do not have the exact month of birth in the data. Using exposure and number of fatalities in entire childhood. Results are very similar.

Table A2: TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, exposure in years and Fatalities, *for exposure in entire childhood* and with clustered std.errors, comparing results with and without inclusion *of exposure* in *the year before the year of birth*

Dependent variable: Test Scores in the Concours National (as z-score) Regressors		Exposure in years				(log)Fatalities			
		including the year before the year of birth		excluding the year before the year of birth		including the year before the year of birth		excluding the year before the year of birth	
		boys	Girls	boys	girls	Boys	Girls	boys	Girls
Expo/Fat over entire childhood	Coeff. Clust.. Robust	-0.028 (0.009)** (0.003)***	-0.031 (0.007)*** (0.003)***	-0.022 (0.010)*** (0.004)***	-0.025 (0.006)*** (0.003)***	-0.014 (0.011) (0.005)***	-0.035 (0.013)*** (0.004)***	-0.012 (0.009) (0.004)***	-0.018 (0.010)** (0.004)***
Age-at-test	Coeff. Clust. Robust	-0.065 (0.004)*** (0.006)***	-0.267 (0.002)*** (0.007)***	-0.062 (0.004)*** (0.006)***	-0.024 (0.003)*** (0.007)***	-0.065 (0.014)** (0.010)***	-0.039 (0.003) (0.012)***	-0.065 (0.002)*** (0.010)***	-0.040 (0.003)*** (0.125)***
Constant	Coeff. Clust. Robust	0.865 (0.069)*** (0.048)***	0.057 (0.039)*** (0.042)	0.840 (0.066)*** (0.046)***	0.053 (0.037) (0.043)	0.923 (0.002)*** (0.045)***	0.195 (0.047)*** (0.042)***	0.917 (0.044)*** (0.045)***	0.161 (0.046)*** (0.041)***
Year.of the test FE		Yes		Yes		Yes		Yes	
Cohort of birth FE		Yes		Yes		Yes		Yes	
School FE		Yes		Yes		Yes		Yes	
Location spec. time tr.		Yes		Yes		Yes		Yes	
N		283,630	314,330	283,630	314,330	283,630	314,330	283,630	314,330

Robustness analysis without the inclusion of exposure in the year before the year of birth, because we do not have the exact month of birth in the data. Using exposure and number of fatalities in each phase of childhood seperately. Results are very similar.

Table A3: TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, exposure in years and Fatalities, *for different stages in childhood* and with clustered std.errors, comparing results with and without inclusion *of exposure in the year before the year of birth*

Dependent variable: Test Scores in the Concours National (as z-score) Regressors		Exposure in years in different phases in childhood				(log)Fatalities in different phases in childhood			
		Including exposure in the year before birth		Excluding exposure in the year before birth		Including exposure in the year before birth		Excluding exposure in the year before birth	
		B	g	b	g	B	G	b	g
Expo/Fat in utero	coeff	-0.026	-0.040			-0.002	-0.008		
	clus.	(0.009)**	(0.008)***			(0.003)	(0.003)**		
	robust	(0.006)***	(0.005)***			(0.002)	(0.002)***		
Expo/Fat at birth	coeff	-0.039	-0.030	-0.043	-0.036	-0.005	-0.010	-0.004	-0.007
	clus.	(0.009)***	(0.012)**	(0.009)***	(0.010)***	(0.004)	(0.004)***	(0.003)	(0.003)**
	robust	(0.006)***	(0.005)***	(0.006)***	(0.005)***	(0.002)**	(0.002)***	(0.002)**	(0.002)***
Expo/Fat age 1-3	coeff	-0.025	-0.032	-0.020	-0.024	-0.009	-0.011	-0.008	-0.007
	clus.	(0.012)**	(0.008)***	(0.010)*	(0.007)***	(0.005)*	(0.005)**	(0.004)*	(0.005)
	robust	(0.004)***	(0.004)***	(0.004)***	(0.004)***	(0.003)***	(0.002)***	(0.002)***	(0.002)***
Expo/Fat age 4-6	coeff	-0.017	-0.026	-0.005	-0.0004	-0.005	-0.010	-0.004	-0.008
	clus.	(0.018)	(0.021)	(0.016)	(0.023)	(0.005)	(0.005)**	(0.005)	(0.005)
	robust	(0.007)**	(0.007)***	(0.007)	(0.007)	(0.003)*	(0.002)***	(0.002)*	(0.003)***
Expo/Fat age 7-12	coeff	-0.052	-0.060	-0.05	-0.045	-0.002	-0.003	-0.001	-0.000
	clus.	(0.019)**	(0.023)**	(0.018)**	(0.024)*	(0.006)	(0.006)	(0.005)	(0.006)
	robust	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.004)	(0.003)	(0.004)	(0.003)
Age-at-test	coeff	-0.045	-0.008	-0.045	-0.014	-0.070	-0.050	-0.069***	-0.045
	clus.	(0.016)**	(0.018)	(0.015)***	(0.018)	(0.004)***	(0.004)***	(0.003)***	(0.003)***
	robust	(0.008)***	(0.008)	(0.008)***	(0.008)*	(0.010)***	(0.012)***	(0.010)***	(0.012)***
Const	coeff	0.795	0.01	0.773	0.009	0.94	0.219	0.935	0.186
	clus.	(0.079)***	(0.08)	(0.075)***	(0.067)	(0.055)***	(0.054)***	(0.051)***	(0.050)***
	robust	(0.05)***	(0.04)	(0.049)***	(0.045)	(0.048)***	(0.043)***	(0.047)***	(0.187)***
Year of test FE		Yes		Yes		Yes		Yes	
Cohort of birth FE		Yes		Yes		Yes		Yes	
School FE		Yes		Yes		Yes		Yes	
Location sp.time tr.		Yes		Yes		Yes		Yes	
N		283,630	314,330	283,630	314,330	283,630	314,330	283,630	314,330

Robustness analysis comparing children who are only exposed in utero with children who are only exposed after birth. When boys are only exposed in utero, it reduces the effects on school test results. In 15 of the 16 cells in Table A4 the results are in line with the culling effect described in the paper. Only in one cell (the first one, which is marginally significant) the result differs.

Table A4: TWFE regressions of school test scores (*reghdfe* command) with school fixed effects, exposure in years and Fatalities, comparing results for children who are only exposed in utero and children who are only exposed after birth

Dependent variable: Test Scores in the Concours National (as z-score) Regressors		Exposure in years				(log)Fatalities			
		Only exposed in utero		Only exposed after birth		Only exposed in utero		Only exposed after birth	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Expo/Fat	Coeff Clus Robust	-0.039 (0.021)* (0.018)**	-0.027 (0.019) (0.016)*	-0.025 (0.01)** (0.004)***	-0.037 (0.008)*** (0.004)***	-0.023 (0.025) -	-0.059 (0.028)** -	-0.012 (0.009) (0.005)***	-0.022 (0.012)* (0.004)***
Age-at-test	Coeff Clus Robust	-0.144 (0.010)*** (0.061)**	-0.152 (0.008)*** 0.047)***	-0.062 (0.004)*** (0.006)***	-0.017 (0.005)*** (0.007)***	0.47 (0.016)*** -	-1.45 (0.01)*** -	-0.065 (0.002)*** (0.010)***	-0.039 (0.003)*** (0.012)***
Constant	Coeff Clus Robust	0.794 (0.171)*** (0.164)***	0.217 (0.245) (0.165)	0.838 (0.062)*** (0.046)***	0.01 (0.042) (0.043)	0.755 (0.080)*** -	-0.273 (0.075)*** -	0.919 (0.004)*** (0.045)***	0.167 (0.046)*** (0.041)***
Year of test FE		Yes		Yes		Yes		Yes	
Cohort of birth FE		Yes		Yes		Yes		Yes	
School FE		Yes		Yes		Yes		Yes	
Location sp.time tr.		Yes		Yes		Yes		Yes	
N		57,607	73,754	262,749	285,042	68,729	75,247	281,659	311,663

Map 1: spread of civil war over provinces in Burundi