EC3320
2016-2017
Michael Spagat

Lecture 8

OP-ED COLUMNIST

The Pain of the G-8's Big Shrug

By NICHOLAS D. KRISTOF
Published: July 10, 2008

Genocide is regrettable, but don't lose perspective. It is simply one of many tragedies in the world today — and **a fairly modest one in terms of lives lost**.

Civil conflict in Congo has claimed more than 5.4 million lives over the last decade, according to careful mortality surveys by the International Rescue Committee. That's **at least 10 times the toll in Darfur,** but because Congo doesn't count as genocide — just as murderous chaos — no one has paid much attention to it.

OP-ED COLUMNIST

The **World Capital of Killing**

By NICHOLAS D. KRISTOF
Published: February 6, 2010

But so far the brutal war here in eastern Congo has not only **lasted longer than the Holocaust** but also appears to have claimed more lives. A peer- reviewed study put the Congo war's **death toll at 5.4 million** as of April 2007 and rising at 45,000 a month. That would leave the total today, after a dozen years, at **6.9 million.**

OP-ED COLUMNIST

Death by Gadget

By NICHOLAS D. KRISTOF

Published: June 26, 2010

A word of background: Eastern Congo is the site of the **most lethal conflict since World War II**, and is widely described as **the rape capital of the world.** The war had claimed **5.4 million deaths** as of April 2007, with the toll mounting by 45,000 a month, according to a study by the International Rescue Committee.

Number Mongering/Hyperbole

5.4 million, 5.4, million, 5.4 million, 5.4 million……

6.9 MILLION

If we just keep extrapolating 45,000 per month we're now up to around

8.6 million

There were a series of surveys done by the International Rescue Committee (IRC) measuring war deaths in the Democratic Republic of Congo (DRC).

Consider one that was published in the *Lancet* in 2006 covering the period January 2003 through April 2004. This one is of particular interest because it is the only one published in a prominent journal.

The main thing that the IRC estimated was a death rate which came out 2.1 deaths per 1,000 people per month with a 95% confidence interval of 1.6 to 2.6.

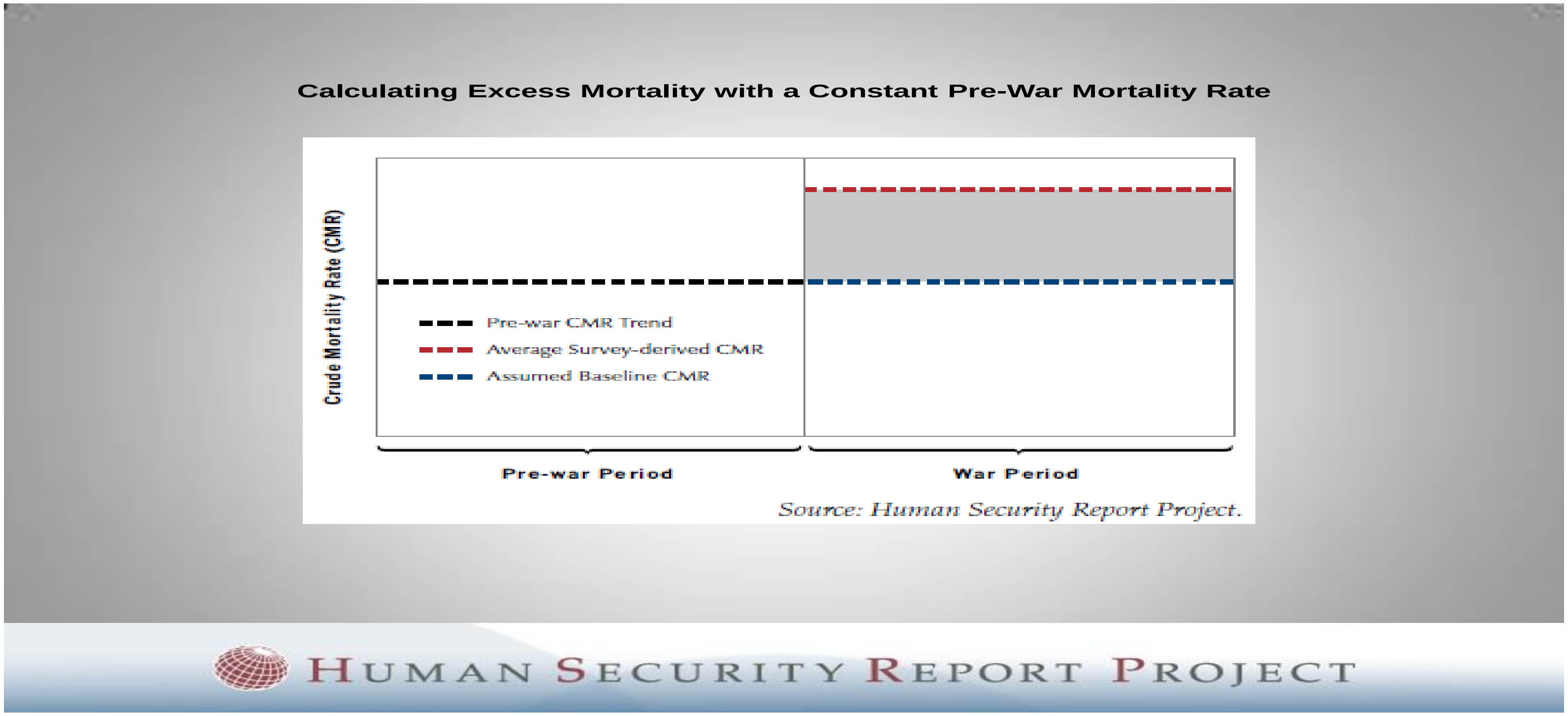
Notice that this estimate covers all deaths, not just violent deaths. This means that we definitely cannot attribute all these deaths to the war because, of course, anywhere in the world some people will always be dying. So an obvious question arises:

> How do they convert an estimated all-cause death rate into a number of deaths attributable to the war?

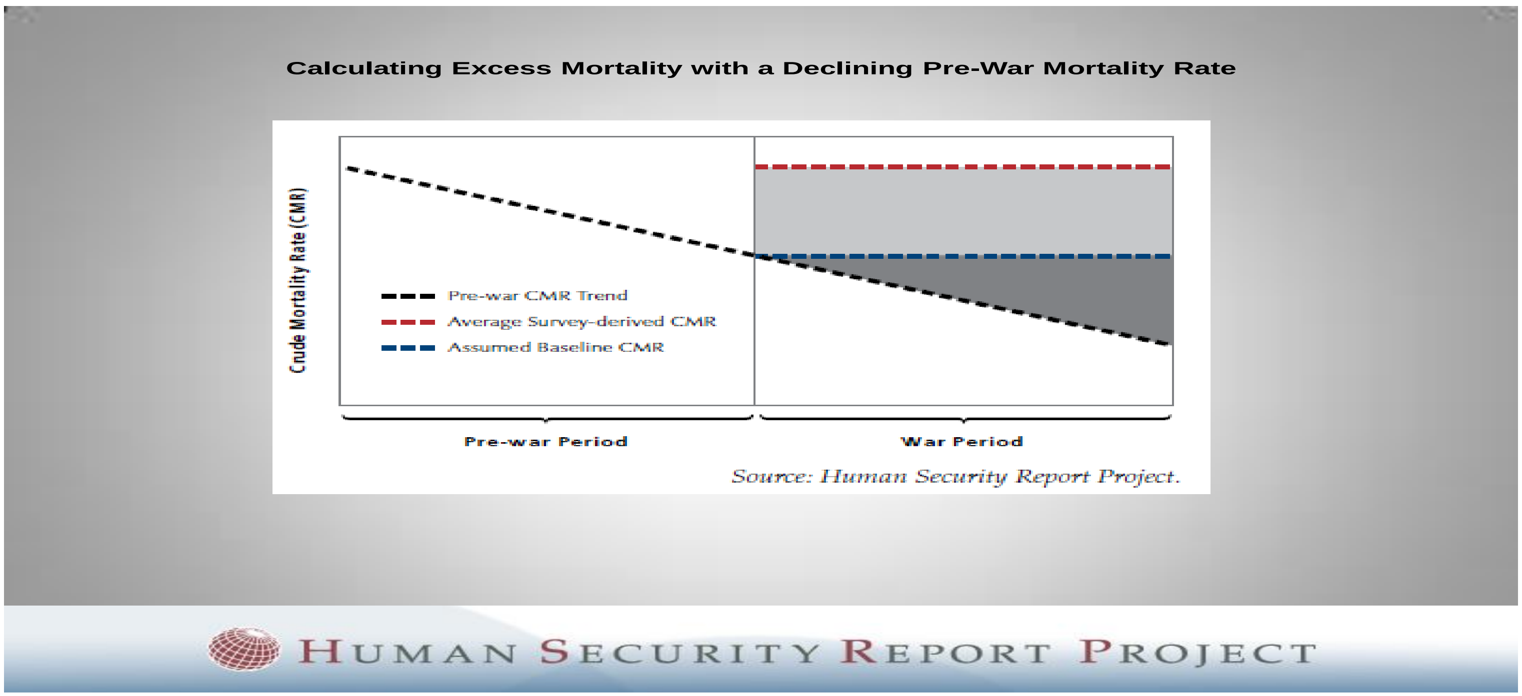
To do this we need to have some concept of an *excess death rate*, i.e., deaths above and beyond those that would have occurred if there had never been a war.

We should recognize that any such excess-death concept will necessarily include a fairly strong speculative component. This is because we will have to compare actual (estimated) deaths against theoretical deaths that we think would have occurred under a *counterfactual* situation. But we can't run history twice, once with and once without a war.

So to estimate excess deaths we need to somehow conjure up a baseline death rate. The typical approach is to take a pre-war mortality rate and assume that this would have carried forward if there hadn't been a war.



One weakness of this approach is that in most countries mortality rates tend to decline over time rather than staying constant. This suggests that a more appropriate counterfactual may be a declining trend rather than a flat line. If so, then the standard method for calculating excess deaths will tend to underestimate the "true" number of excess deaths.



But the IRC did not have a pre-war death rate to use as a baseline rate. So they had to conjure up a different baseline rate.

They made the strong assumption that if the DRC had not suffered a war then (with certainty) it would have experienced the death rate that was the average for this period for all of Sub-Saharan Africa. The IRC did this in the face of clear evidence that the DRC had long been toward the bottom of Sub-Saharan Africa across a wide range of economic indicators after suffering for decades at the hands of the notorious dictator Mobuta Sese-Seko.

The rate for Sub-Saharan Africa is measured (very poorly) at 1.5 deaths per 1,000 per month. This rate is taken as the baseline rate for the study.

We now have the information we need to calculate an excess death rate for the conflict in the DRC.

During-war rate - 2.1 per thousand per month

Baseline rate – 1.5 deaths per thousand per month

Excess death rate – 0.6 per thousand per month

It is as simple as that!

Do you think this is a high rate or a low rate?

Here is a simple heuristic to come to grips with this somewhat mysterious number.

There are around 9,000 RHUL students.

If 0.6 per thousand per month are dying that would mean 5-6 RHUL students dying every month. That would be a lot, wouldn't it?

One weakness with this heuristic is that the population of university students is young and healthy (at least compared to the rest of the population) so you would expect a particularly low death rate……but still 0.6 per thousand per month is massive.

Death rates are not very intuitive for most people. It is much easier for people to understand total numbers of deaths rather than rates of deaths. So the IRC calculates the number of deaths implied by the excess death rate. To do this we need an estimate for the total population of the DRC.

DRC Population - 63,700,000.   That is not super accurate but it is good enough.

0.6 deaths per thousand per month X 63,700 thousands = 38,220 deaths per month

The survey covers 15 months so this becomes **573,300 deaths**

That is a lot of deaths.

But what about a 95% confidence interval?

We can easily build one. Recall that the IRC estimate a death rate of 2.1, with a 95% confidence interval of 1.6 to 2.6. (Slide 4).

The IRC also treats the baseline rate as a certain 1.5.

So the 95% confidence interval on the excess death rate is 0.1 to 1.1 (per thousand per month).

This means that the 95% confidence interval on total excess deaths is 95,550 to 1,051,050

OK, let’s stop carrying everything out to a ridiculous number of digits. Basically the estimate is:

600,000 plus or minus 500,000

This is like saying that for tomorrow I am predicting the temperature will be between 3 degree centigrade and 33 degrees centigrade (18 plus or minus 15).

In short, this estimate is meaningless. By the way, you will not find the confidence interval anywhere in the published paper. No one would take the estimate seriously after seeing it.

Even this gargantuan confidence interval is based on two extremely dubious assumptions:

1. In the absence of war a reasonable estimate of what the DRC's death rate would have been is the Sub-Saharan average. Yet there is considerable evidence that this "baseline rate" is too low.

2. We take this counterfactual (baseline) death rate of 1.5 per 1,000 as a dead certainty, i.e, we do not place a confidence interval around the baseline itself.

If, instead, we use a baseline rate of 2.0 per 1,000 per month then the estimate becomes 100,000 with a confidence interval of -400,000 to 600,000. Translation: the IRC data is much too coarse to give us any useful handle on excess deaths in the DRC.

Excess Deaths in Iraq

Iraq also displays the phenomenon of a massive confidence interval for excess deaths.

The next slide give a table of excess deaths estimates I made with Stijn van Weezel in a paper that I have loaded onto the Moodle page based on the data from this paper. You can see that the confidence intervals surrounding the various estimates are extremely wide with confidence intervals for some of the estimates ranging even into negative territory.

Moreover, after you separate non-violent deaths from violent deaths you see that the bottoms of the confidence intervals for non-violent deaths dip far below 0.

Table A2: Estimates of excess deaths Iraq 2003-2011

	Replication	Re-analysis								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Excess deaths										
Central estimate	406,000	406,000	398,000	398,000	325,000	290,000	283,000	290,000	202,000	212,000
UI-lower	55,000	−56,000	52,000	−97,000	−29,000	−167,000	37,000	−123,000	−36,000	−197,000
UI-upper	687,000	721,000	724,000	747,000	575,000	654,000	550,000	679,000	441,000	576,000
Crude death rate	1.66	1.66	1.63	1.63	1.33	1.19	1.16	1.19	0.83	0.87
UI-lower	0.23	−0.23	0.21	−0.40	−0.12	−0.68	0.15	−0.50	−0.15	−0.81
UI-upper	2.81	2.95	2.96	3.05	2.35	2.67	2.25	2.78	1.80	2.35
P(excess deaths > 0)	0.987	0.954	0.982	0.949	0.963	0.912	0.99	0.938	0.955	0.889
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel B: Excess non-violent deaths										
Central estimate	160,000	160,000	184,000	184,000	92,000	115,000	102,000	122,000	35,000	54,000
UI-lower	−213,000	−307,000	−212,000	−345,000	−275,000	−398,000	−177,000	−339,000	−240,000	−401,000
UI-upper	409,000	474,000	489,000	547,000	332,000	467,000	366,000	489,000	264,000	417,000
Crude death rate	0.65	0.65	0.75	0.75	0.38	0.47	0.42	0.50	0.14	0.22
UI-lower	−0.87	−1.26	−0.87	−1.41	−1.12	−1.63	−0.72	−1.39	−0.98	−1.64
UI-upper	1.67	1.94	2.00	2.24	1.36	1.91	1.50	2.00	1.08	1.71
P(excess deaths > 0)	0.834	0.761	0.848	0.774	0.702	0.697	0.823	0.766	0.653	0.658
Reweighted data	–	–	Yes	Yes	–	Yes	–	Yes	–	Yes
Governorate resampling	–	Yes	–	Yes	–	Yes	–	Yes	–	Yes
Deaths with no certificate	Yes	Yes	Yes	Yes	–	–	Yes	Yes	–	–
Deaths with certificate not shown	Yes	Yes	Yes	Yes	Yes	Yes	–	–	–	–

N.B.− The uncertainty intervals (UI) are estimated with 1000 bootstrap replications. The lower bound is set at percentile 2.5 and the upper bound is set at percentile 97.5. Panel A gives estimates for excess violent plus non-violent deaths. Panel B gives estimates for just non-violent excess deaths. Reweighted data, central estimates account for stratification; Governorate resampling, uncertainty intervals account for stratification; Deaths with no certificates, includes reported deaths for which respondents admit to not having death certificates; Deaths with certificate not shown, includes reported deaths for which respondents fail to produce a death certificate they claim to have.

There are several things going on here that are difficult to separate:

1. We have a small sample. If the sample were bigger the confidence intervals would be narrower. It could be that in reality there are a substantial number of excess deaths but the sample size is just too small to demonstrate this fact.

2. As we already saw in the DRC case excess-death calculations are quite sensitive to small changes in the measured rates for either the baseline death rate or the during-war death rate.

3. Maybe there really are not very many excess, non-violent deaths. In other words, the problem might not be with the measurement system but, rather, the facts simply do not support the idea that there have been a large number of non-violent excess deaths.

Exploiting Regional Variation

Let’s look at the excess-death issue from a different angle. In particular, let’s take a regional perspective. In any war some regions will be more violent than other regions. (Recall the maps of Northern Ireland we saw in lecture 7).

War might cause indirect deaths as well as direct (violent) deaths. For example, war violence could damage water purification systems and then people may die because of impure water. Or war may cause people to flee their homes into a jungle where they then catch diseases and die. You are all creative people. You can think of your own possible mechanisms for how violence may translate into non-violent deaths.

Here is the point – If violence is causing non-violent deaths then we would expect that geographical areas that suffer relatively high levels of violence will also suffer relatively high levels of non-violent deaths. In other words, where violence spikes we also expect non-violent death rates to spike.

Actually, let's go a little slower on this key point.

Imagine you have two regions, call them A and B. A is violent during the war and B is not violent during the war.

Now imagine four non-violent death rates:

1. $A_{pre-war}$ - The non-violent death rate in region A just before the war starts
2. $A_{during-war}$ - The non-violent death rate in region A during the war
3. $B_{pre-war}$ - The non-violent death rate in region B just before the war starts
4. $B_{during-war}$ - The non-violent death rate in region B during the war

It is interesting to look at the change in the death rate, pre-war versus during war, in region A and the change in the death rate, pre-war versus during war, in region B:

1. $A_{during-war} - A_{pre-war}$

2. $B_{during-war} - B_{pre-war}$

These two differences are, essentially, excess non-violent death rates in the two regions. The observations at the bottom of slide 16 suggests that the difference number 1 should be bigger than difference number 2, i.e.,:

$$\left(A_{during-war} - A_{pre-war}\right) - \left(B_{during-war} - B_{pre-war}\right) > 0$$

This kind of calculation is known as a “difference in difference” calculation which it is, literally.

Notice that either of the two differences, $A_{during-war} - A_{pre-war}$ and $B_{during-war} - B_{pre-war}$, can potentially be negative. In fact, both differences can be negative while the difference in differences can still be positive.

That is, it can happen that death rates decrease after the war starts in both regions but they go down more in the non-violent region than they do in the violent one. In this case the difference in differences would be positive meaning that there would be evidence that the war is causing non-violent excess deaths.

Note that this sort of comparison resembles a controlled laboratory experiment. The peaceful region (B) is like a control group and the violent region (A) is like a treated group.

How can we translate this “diff-in-diff” thinking into a regression framework that can allow us to measure the size and statistical significance of violence on deaths rates or on other variables?

$$y_{hvt} = \alpha + \gamma D_t + \lambda D_v + \beta D_{vt} + \varepsilon_{hvt}$$

Where h denotes households, t refers to time periods (pre-war or during-war), v refers to regions (violent or non-violent), y is a number of non-violent deaths and ε is a random shock. D_t , D_v and D_{vt} are dummy variable, i.e, they are always 0 or 1. D_t is 1 during the war and 0 otherwise, D_v is 1 in violent regions and 0 otherwise and D_{vt} is 1 when the region is violent and it is the during-war period and 0 otherwise.

If β is positive then households in violent regions during the war tend to experience higher non-violent death rates than other households.

Before moving on I note that Stijn van Weezel and I do an estimate like this on the same Iraq data that we used for our excess deaths estimate. We got a statistically insignificant coefficient for β. I will not show the table here because there is an additional technical issue with this regression that goes beyond the scope of this course.

But the bottom line as far as the Iraq data is concerned is that non-violent death rates did not rise faster in the violent regions of Iraq than they rose in the non-violent regions of Iraq.

Next we use this “diff-in-diff” approach to study the impact of war on health.

We use “height-for-age” as our health indicator. Of course, it is nice to be tall but height is an important indicator for far more than cosmetic purposes. Being short is associated on average with poor nutrition while young which contributes to poor educational outcomes and further problems such as low life expectancy and unhealthy offspring.

Thus, we are interested in the height of children not because this is of great direct interest but because this gives us good information on how they will fare throughout their lives.

There has been considerable research on distributions of heights of children of all ages. The height-for-age measure is standardized to these observed distributions. A height—for-age score of -1.0, for example, means that the height is 1 standard deviation below the average for a child of that age.

Minouiu and Shemyakina apply the diff-in-diff to estimate the impact of the war in Côte d'Ivoire on height for age outcomes. Their baseline equation is:

$$(1)\ HAZ_{ijt} = \alpha_j + \delta_t + \lambda_{jt} + \beta_1(Conflict\ Region_j \times Cohort_t) + \beta_2(Conflict\ Region_j \times Cohort_t \times Female_i) + \varepsilon_{ijt},$$

where HAZ_{ijt} is the height-for-age z-score for child i in province j born during year t; α_j are the province-of-birth fixed effects, δ_t are birth-cohort fixed effects, λ_{jt} are province-specific trends in cohort health, and ε_{ijt} is a random, idiosyncratic error term. Indicator variables for female children and rural residence are included

This equation is quite close to the sort of diff-in-diff equation you have already seen. The main differences are that the left-hand-side variable is height-for-age and that includes a dummy variable for females.

The main coefficient of interest is β_1 which gives the effect on height suffered by children born in a conflict region during the conflict.

β_2 gives and additional effect that applies only to girls. However, it turns out that this effect is always statistically insignificant so you can pretty much ignore it.

Slide 26 gives the main table of results. Before I show the table let me make a few preliminary points.

The violence data comes from ACLED which is an important source of conflict data we are encountering for the first time.

The rest of the data come from the 2002 and 2008 household surveys known as the Côte d’Ivoire Household Living Standards Surveys.

The main conflict took place between 2002 and 2004 but some violence persisted until around March 2007. There is then a survey in 2008. Some people were in the conflict zone during the war but had left by the time the 2008 survey was done. Other people moved into the conflict zone after the conflict ended. This migration causes problems for the survey which the authors address by rerunning all their regressions on only the people who stayed in place throughout the whole period.

Table 3— Impact of Residence in Conflict Region on Child Health

	Full sample			Nonmigrant subsample		
	Baseline (1)	With controls (2)	With controls (3)	Baseline (4)	With controls (5)	With controls (6)
Conflict region × Born during conflict	−0.489**	−0.449**	−0.463*	−0.483**	−0.450**	−0.445*
	(0.210)	(0.199)	(0.227)	(0.201)	(0.201)	(0.224)
Conflict region × Born during conflict × Female	0.179	0.154	0.165	0.063	0.023	−0.005
	(0.117)	(0.129)	(0.125)	(0.179)	(0.198)	(0.196)
Female	0.217***	0.209***	0.220***	0.215***	0.204***	0.222***
	(0.061)	(0.060)	(0.063)	(0.070)	(0.066)	(0.074)
Rural household	−0.482***	−0.411***	−0.412***	−0.432***	−0.350***	−0.389**
	(0.095)	(0.091)	(0.096)	(0.126)	(0.120)	(0.136)
Child controls	No	Yes	Yes	No	Yes	Yes
Household head controls	No	Yes	No	No	Yes	No
Mother controls	No	No	Yes	No	No	Yes
p-value F-tests of zero joint effect of:						
Child ethnicity	—	0.001	0.064	—	0.006	0.064
Child religion	—	0.031	0.197	—	0.059	0.178
Household head characteristics	—	0.039	—	—	0.009	—
Mother characteristics	—	—	0.000	—	—	0.000
Observations	7,941	7,871	7,151	6,026	5,967	5,390
R^2	0.075	0.083	0.099	0.066	0.076	0.088

Notes: Robust standard errors in parentheses, clustered at the province level. The dependent variable is the height-for-age *z*-score. All regressions include province fixed effects, month-of-birth fixed effects, and province-specific time trends. Child controls include ethnicity (Akan, Northern Mandé, Southern Mandé, Voltaique, Krou, naturalized Ivorian, and non-Ivorian) and religion (Muslim, Christian, and other). Household head controls include age, education, gender, and marital status. Mother controls include age and education. Education is proxied by an indicator variable for French literacy. All estimates are weighted by inverse sampling probability.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Sources: 2002 and 2008 CIV HLSS, and Raleigh et al. (2010).

The key point of the table are the negative and statistically significant coefficients for children born during the conflict in the conflict zone.

These coefficients are pretty big as well. You lose about ½ a standard deviation in height if you had this misfortune.

There is weak evidence that maybe girls suffer slightly smaller height reductions than boys do but this isn't statistically significant.

Akresh et al. apply a similar approach to Nigeria using the Nigerian Demographic and Health Surveys from 2003 and 2008.

There are three main differences compared to Minoiu and Shemyakina.

1. Akresh et al. do not have data on conflict events. Instead, they use the fact that certain ethnicities bore the brunt of the violence in Nigeria and treat members of these ethnicities as potential conflict victims.

2. Akresh et al. calculate number of months of exposure to the conflict for each individual in the survey from the affected ethnic groups. So they can estimate what might be called "dosage effects".

3. Akresh et al. distinguish between different age groups.

The next slide gives their main table.

Table 1—Difference-in-Differences Estimates of the Impact of War on Stature

Duration of Exposure to War × Exposed Ethnicity

Dependent variable: adult height	(1)	(2)
Months exposure in utero × war ethnicity	−0.027	−0.047
	(0.030)	(0.038)
Months exposure at ages 0–3 × war ethnicity	−0.028***	−0.043**
	(0.009)	(0.021)
Months exposure at ages 4–6 × war ethnicity	−0.035**	−0.061*
	(0.014)	(0.034)
Months exposure at ages 7–12 × war ethnicity	−0.054***	−0.094*
	(0.011)	(0.050)
Months exposure at ages 13–16 × war ethnicity	−0.162***	−0.220***
	(0.033)	(0.076)
Months exposure in utero	−0.087	−0.081
	(0.059)	(0.060)
Months exposure at ages 0–3	−0.129***	−0.125***
	(0.034)	(0.034)
Months exposure at ages 4–6	−0.022	−0.015
	(0.070)	(0.071)
Months exposure at ages 7–12	−0.014	−0.003
	(0.083)	(0.083)
Months exposure at ages 13–16	−0.038	−0.022
	(0.138)	(0.137)
State fixed effects	Yes	Yes
Ethnicity fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Ethnicity time trends	No	Yes
Number of women	13,407	13,407

Notes: Robust standard errors in parentheses, clustered at the ethnicity × year level. In addition to the named fixed effects, regressions include a survey year dummy. The sample includes adult women born between 1954 and 1974. Ethnicity time trends are defined by an interaction between ethnicity category and birth year. War ethnicity equals 1 if the individual belongs to the Igbo ethnicity or to another southeastern minority. The mean height in the sample is 158.70 cm and the standard deviation is 7.06. The mean [standard deviation] of exposure duration in months (for the exposed ethnicities) at ages −1, 0–3, 4–6, 7–12, and 13–16 is 7.3 [2.4], 17.5 [9], 17.9 [8.6], 23 [9.9], and 20.6 [10.8], respectively.

*** Significant at the 10 percent level.

** Significant at the 5 percent level.

* Significant at the 1 percent level.

A rather surprising finding is that the biggest effects seem to be on adolescents.

Beyond that, there are statistically significant effects for all ages but not for fetuses ("in utero" in the table.)

These effects may look smaller than those of Minoiu and Shemyakina but these need to be multiplied by the number of months to get the full effects.

Summing up, it seems that there are substantial effects of war on the heights of children and the diff-in-diff method has been helpful in demonstrating these effects.