




Research Article

# Half a million excess deaths in the Iraq war: Terms and conditions may apply

Research and Politics  
October-December 2017: 1–8  
© The Author(s) 2017  
DOI: 10.1177/2053168017732642  
[journals.sagepub.com/home/rap](http://journals.sagepub.com/home/rap)  


Michael Spagat<sup>1</sup> and Stijn van Weezel<sup>2</sup>

## Abstract

Hagopian et al. (2013) published a headline-grabbing estimate for the Iraq war of half a million excess *deaths*, i.e. deaths that would not have happened without the war. We reanalyse the data from the University Collaborative Iraq Mortality Study and refute their dramatic claim. The Hagopian et al. (2013) estimate has four main defects: i) most importantly, it conflates non-violent deaths with violent ones; ii) it fails to account for the stratified sampling design of the UCIMS; iii) it fully includes all reported deaths regardless of death certificate backing, even when respondents say they have a death certificate but cannot produce one when prompted; iv) it adds approximately 100,000 speculative deaths not supported by data. Thus, we reject the 500,000 estimate. Indeed, we find that the UCIMS data cannot even support a claim that the number of non-violent excess deaths in the Iraq war has been greater than zero. We recommend future research to follow our methodological lead in two main directions; supplement traditional excess death estimates with excess death estimates for non-violent deaths alone, and use differences-in-differences estimates to uncover the relationship between violence and non-violent death rates.

## Keywords

Excess deaths, Iraq, differences-in-differences, bootstrapping, household survey data

## Introduction

Large excess death numbers draw big attention. Examples include the high profile estimates of excess deaths for the Democratic Republic of Congo (International Rescue Committee, 2007) and Darfur (Hagan and Palloni, 2006; Degomme and Guha-Sapir, 2010).<sup>1</sup> Such numbers are widely considered scientifically sound, of great interest to the general public and worthy of publication in top peer-reviewed journals, although they often generate controversy in both academic journals and the media (NBC News, 2008; Spagat et al., 2009; ScienceInsider, 2010).

The estimate of 500,000 excess deaths in the Iraq war, 2003–2011, by Hagopian et al. (2013) was a media event complete with a press release (University of Washington, 2013). In the present paper we reexamine the Hagopian et al. (2013) dramatic claim using exactly the same dataset as Hagopian et al. (2013) do, the University Collaborative Iraq Mortality Study (UCIMS). We find that the UCIMS data do not support the 500,000 claim. Most importantly, the estimate unravels as soon as we distinguish between violent deaths and non-violent deaths. Estimates for non-violent

excess deaths have extremely wide uncertainty intervals that stretch well below zero. Moreover, we employ a differences-in-differences (DiD) design that finds no local spillovers running from war violence to elevated non-violent death rates, reinforcing our null finding for non-violent excess deaths.<sup>2</sup> Three further corrections and improvements to the original analysis undermine the half-a-million claim even more. These are accounting for; i) stratification in the UCIMS sampling design; ii) weak death certificate backing for some deaths; and iii) the feeble basis for nearly 100,000 deaths which enter the estimate through rounding and confusion between deaths and excess deaths.<sup>3</sup>

<sup>1</sup>Department of Economics, Royal Holloway University of London, Egham, UK

<sup>2</sup>School of Economics, University College Dublin, Dublin, Ireland

### Corresponding author:

Michael Spagat, Department of Economics, Royal Holloway University of London, Egham, TW20 0EX, UK.  
Email: [m.spagat@rhul.ac.uk](mailto:m.spagat@rhul.ac.uk)



Our approach can serve as a model for future studies of excess war deaths. In particular, we recommend that future researchers make excess death estimates for non-violent deaths alone rather than following the established practice of conflating violent deaths with non-violent deaths in excess death estimates. We also advocate the use of DiD strategies to test whether war violence causes increases in non-violent death rates in areas where war violence is prevalent.

## Our analytical improvements

The UCIMS survey of deaths in the Iraq war is the most recent high-profile estimate of excess deaths in war. The data come from a nationally representative survey conducted in mid 2013 of individuals in 2,000 randomly selected households situated in 100 clusters across Iraq. The fieldwork for this study appears to be good and the data are freely available to a reasonable level of detail.<sup>4</sup> We are able to replicate the Hagopian et al. (2013) results almost exactly once we accept all their assumptions. So we turn now to several corrections and improvements to the original analysis.

### *Measuring excess war mortality*

War can cause both violent deaths, or direct battle deaths from the conflict, as well as non-violent deaths, arising indirectly from the consequences of the conflict. Yet violent deaths are easier to attribute to war than are non-violent deaths that can, at best, be only indirectly linked to war. Examples of non-violent war deaths can include hospital patients who die after not receiving crucial medical treatments due to war-caused power outages in hospitals or people who die of cholera after war-related damage to water treatment facilities. The number of non-violent deaths indirectly attributable to war might be substantial in some modern conflicts, although it is difficult to pin down these effects analytically.

The primary challenge for estimates of the number of non-violent deaths caused indirectly by a war is to avoid including non-violent deaths that would have occurred even without war. The main strategy used in the literature to tackle this separation problem is to define a counterfactual death rate that, one argues, would have occurred if war had been avoided. This counterfactual death rate, sometimes known as a *baseline* death rate, is subtracted from the death rate that actually occurs during the war. The difference between the actual death rate and the baseline death rate is commonly designated as an *excess* death rate. This excess death rate is then multiplied by the size of the at risk population to obtain a number of excess deaths.<sup>5</sup>

Hagopian et al. (2013) follow most literature in using the pre-war death rate as their baseline (Checchi and Roberts, 2005). The key assumption underpinning this choice is that the pre-war death rate would have carried

forward in time unaltered if war had never broken out. If this assumption is approximately true then it is reasonable to view the during-war death rate minus the baseline death rate as a war-caused excess death rate.

We stress that researchers generally conflate violent deaths with non-violent deaths when they make excess death calculations. Yet this conflation of two very different death types is fundamentally inconsistent with the whole point of invoking the excess deaths concept in the first place: to capture the possibility that war violence can lead indirectly to *non-violent* deaths. Mixing violent deaths into an excess death estimate confounds the channel running from violence to non-violent death with a distinct, and trivial, channel running from violence to violent deaths. Thus, a standard excess death estimate can create an illusion that a war is indirectly causing nonviolent deaths even for a war that causes exclusively violent deaths. Analogously, one might state that the average vegetarian consumes approximately one ton of vegetables plus meat per year. While technically correct this statement could create a false impression that the average vegetarian hypocritically consumes meat when, in fact, the ton of vegetables plus meat breaks down into one ton of vegetables and 0 tons of meat. Therefore, we argue that the best practice for evaluating the human cost of war should proceed along two tracks. First estimate direct violent deaths and then estimate excess *non-violent* deaths. Despite this advice, we maintain consistency with the existing literature by also estimating excess deaths in the standard manner that conflates violent deaths with non-violent ones.

### *Non-violent deaths versus violent deaths*

The excess deaths concept asserts that war violence leads indirectly to non-violent deaths. If so, then we should be able to link war violence with elevated non-violent death rates. We attempt to make this connection with a difference-in-difference (DiD) estimator (Angrist and Pischke, 2008), the mechanics of which we will explain in detail in section (Trying to link violence with non-violent death rates). Here we just provide some basic intuitions.

Our DiD estimator tests whether relatively large increases in non-violent deaths in particular locations are associated with relatively high violence levels in these locations. This approach detects spillovers from violence to non-violent deaths as long as these spillovers operate at a local level. If, for example, local bomb attacks cause substantial numbers of non-violent deaths by disrupting local water purification or hospital services, then the DiD estimator should detect these indirect effects of violence. On the other hand, the DiD approach will not detect indirect channels running from violence to non-violent deaths if these effects are diffused across space. A DiD estimator would not be useful for analysing the effects of violence that decreases national tax revenues, leading to national

**Table 1.** Relative weights by governorate.

Governorate	Weight
Thi Qar	2.75
Kirkuk	2.18
Najaf	2.02
Babylon	1.62
Muthanna	1.59
Kerbela	1.52
Salah Al Deen	1.42
Duhouk	1.25
Wasit	1.23
Maysan	0.98
Diala	0.98
Baghdad	0.94
Basra	0.90
Sulaimaniya	0.89
Ninevah	0.86
Erbil	0.63
Qadisiyya	0.61
Anbar	0.50

Each governorate's weight is its share of total population divided by its share of respondents in the survey.

budget cuts to health care that increase national death rates. Moreover, national or international policy may direct aid programme to violence-plagued local areas, offsetting real local spillovers in full or in part. Accounting for such contrary forces would be tricky. Aid would prevent excess deaths that would appear if aid were stopped and if would be difficult or impossible to measure these countervailing forces. Despite these caveats, many plausible spillovers will tend to operate at the local level and, therefore, the DiD estimator is a valuable tool for exposing such effects.

### Stratification

The UCIMS is a cluster survey stratified by governorate. Hagopian et al. (2013) account for clustering in their uncertainty intervals but, incorrectly, ignore the impact of stratification on both their central estimates and uncertainty intervals. This is a mistake because stratification can have big effects on both of these calculations (Working Group for Mortality Estimation in Emergencies, 2007; Lohr, 2009). Table 1 displays the UCIMS stratification scheme by governorate. Households in governorates with relative weights below 1 have inclusion probabilities higher than what they would have had under an equal probabilities scheme and vice versa. For example, a death in Anbar is 5.5 times more likely to appear in the sample than is a death in Thi-Qar ( $\frac{2.75}{0.50} = 5.5$ ). Therefore, in-sample deaths in Thi-Qar should represent 5.5 times more estimated deaths than in-sample deaths in Anbar do.<sup>6</sup>

A second stratification issue is that the procedures for estimating uncertainty intervals should incorporate the complexities of the sampling scheme, such as departures from equal inclusion probabilities. Hagopian et al. (2013) bootstrap (Shalizi, 2010) their uncertainty intervals but, incorrectly, ignore stratification in their computations. For example, the UCIMS sample contains 23 Baghdad clusters so each bootstrap resample should also contain 23 Baghdad clusters, each drawn with replacement from the 23 original Baghdad clusters. The Hagopian et al. (2013) resampling does not satisfy this requirement but we fix this problem in our bootstraps.

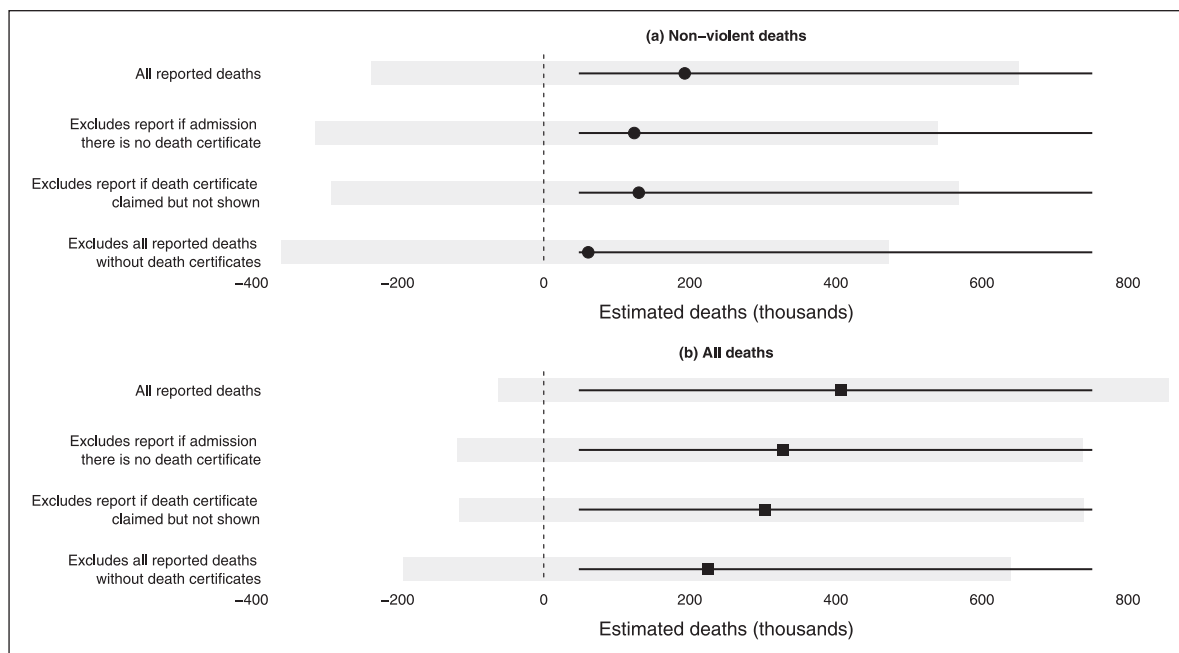
Note that all of the uncertainty intervals (UIs) in both our paper and in the original Hagopian et al. (2013) paper are obtained through bootstrapping, although the dataset has only 100 clusters which is not enough to be sanguine about using asymptotic theory (Flynn and Peters, 2004; Pons, 2007). Thus, the reader should regard all UIs in the paper as too narrow, a point which makes the Hagopian et al. (2013) claim of 500,000 excess deaths even less credible than the uncertainty intervals we publish in this paper suggest.

### Death certificates

The UCIMS dataset provides the outcome of each interviewer attempt to verify an in-sample death with a matching death certificate. There are 385 reported deaths in the dataset, of which interviewers record seeing death certificates for 284. However, 32 times respondents state that they do not possess a death certificate to document a reported death. In a further 66 cases, a respondent reports a death, claims to have a death certificate confirming the death, but fails to produce the supposed death certificate when prompted to do so in the interview.<sup>7</sup> We find this last type of reported death to be rather dubious; why would someone claim to have a death certificate he/she does not actually have? Nevertheless, many, perhaps even most, cases of reported deaths not backed by death certificates are real. People can lose or misplace a death certificate and in some cases, particularly during the first phase of the war, death certificates were never issued. Thus, it is difficult to quantify the uncertainty surrounding deaths not fully backed by death certificates although it is suboptimal to simply ignore this issue as Hagopian et al. (2013) do in their analysis. We address death certificate uncertainty by presenting four separate sets of estimates. These either i) accept all reported deaths, as in the original paper; ii) reject all reported deaths without a shown death certificate; iii) reject reported deaths for which respondents admit to not possessing a death certificate; or iv) reject reported deaths for which respondents say they have a death certificate but fail to produce one.

### From 400,000 to 500,000

There is a confusing conflict between the Hagopian et al. (2013) UCIMS-based central estimate of 406,000 excess deaths and the claim in the Hagopian et al. (2013) paper that



**Figure 1.** Excess Deaths Estimates and Uncertainty Intervals.

Central estimates and uncertainty intervals for excess deaths for non-violent deaths only in panel (a) and mixing together both violent and non-violent deaths in panel (b). The black squares and circles are central estimates. The grey bars give our 95% uncertainty intervals, each constructed from 1,000 bootstrap replications. The black lines give the uncertainty interval from the original study (48,000–751,000). All these calculations account for stratification by governorate. The numbers underlying this picture are in appendix Table A2.

they estimate 500,000 deaths.<sup>8</sup> The authors reach beyond their data to bridge the gap of nearly 100,000 deaths. Specifically, they argue that there is a war refugee population of approximately two million people not covered by their survey from which there were, they say, approximately 56,000 deaths. They then add 56,000 to 406,000 and round up to 500,000. There are two good reasons to reject this manoeuvre. First, 56,000 deaths in a refugee population of 2 million is barely above the baseline rate of 2.89 per 1,000 per year that Hagopian et al. (2013) use in their excess death estimates. Therefore, it is not clear that adding in refugees should increase the excess death estimate at all. Therefore, if anything, adding in refugees should decrease rather than increase the excess death estimate. Second, there can be no good basis to round from 462,000 all the way up to 500,000.

## In search of excess deaths

### Excess deaths estimates

Figure 1 presents a range of excess death calculations. In panel (a), we focus only on non-violent deaths whereas in panel (b), we revert to the standard approach of conflating violent deaths with non-violent deaths. Our primary and preferred estimates are the ones in panel (a) since, as argued in section (Measuring Excess War Mortality), the concept of excess deaths is fundamentally about war violence causing non-violent deaths, not about war violence causing violent deaths. We account for stratification and clustering in all of these results.

The top estimate in panel (a) shows that we cannot reject a hypothesis of negative non-violent excess deaths at a 10% level, even if we accept every reported death in the survey regardless of the quality of its death certificate backing. Of course, discounting reported deaths without death certificates drags the estimates further downwards. The central estimates for non-violent excess deaths in panel (a) are all well below 200,000 and the bottoms of these uncertainty intervals all drop below  $-300,000$ . Thus, the evidence in favor of even a positive number of non-violent excess deaths is exceedingly thin.

The estimates in panel (b) show that the UCIMS data do not support the Hagopian et al. (2013) claim of 500,000 excess deaths even when we follow Hagopian et al. (2013) in conflating violent deaths with non-violent ones in our excess deaths estimates. Although the stratification adjustment barely moves the central estimates, it does widen the uncertainty interval substantially, even pulling the bottom of the uncertainty interval below zero for the top estimate in the panel that accepts all deaths regardless of death certificate backing. With this estimate we can still reject a hypothesis of negative excess deaths at around a 5% level, but a hypothesis of 500,000 or more excess deaths, well above the central estimate, is roundly rejected. The other estimates include only deaths with various degrees of death certificate backing, lowering all the central estimates considerably. We estimate 210,000 excess deaths with the uncertainty interval swinging by more than 300,000 deaths in either direction when we



insist on death certificate backing for all deaths. This estimate is at an extreme, as noted in section 1, but the top estimate is also at an extreme although, perhaps, closer to the truth.

### Trying to link violence with non-violent death rates

The simplest DiD framework makes two binary divisions of the country along two separate dimensions: time and space. These divisions give us four groupings: a pre-war period and a during-war period crossed with a violent zone and a non-violent zone. The goal is to test whether living in the violent zone during the war is associated with relatively high non-violent death rates. We operationalise this idea with equation 1 in which  $h$  indexes households,  $v$  indexes zones (violent or non-violent), and  $t$  gives the time period (pre-war or during-war). We define the violent zones to be Anbar, Baghdad and Diala. Such a binary division is, of course, arbitrary although our violent governorates are all more than 28% more violent than Kirkuk, the next most violent governorate. Moreover, moving Kirkuk into the violent category does not change our results qualitatively.

$y_{hvt}$  is the count of non-violent deaths in household located in zone  $v$  during period  $t$ . Similarly,  $\epsilon_{hvt}$  is a random shock to household  $h$  located in zone  $v$  during period  $t$ . There are three dummy variables,  $D_t$ ,  $D_v$  and  $D_{vt}$  which are equal to 1, respectively, for households in the during-war period, for households in the violent zone and for households simultaneously in the the violent zone in the during-war period.<sup>9</sup>

$$y_{hvt} = \gamma D_t + \lambda D_v + \beta D_{vt} + \epsilon_{hvt} \quad (1)$$

The estimated parameters are  $\gamma$ ,  $\lambda$  and  $\beta$ .  $\gamma$  will be positive if there is a general tendency applying across both zones for households to suffer more non-violent deaths in the during-war period than they do in the the pre-war period.  $\lambda$  will be positive if, across both time periods, households tend to suffer more non-violent deaths in the violent zone than they do in the non-violent zone. The  $\beta$  coefficient is the key to the DiD approach. It gives the difference between the average difference in non-violent deaths, during-war minus pre-war, in the violent zone and the corresponding difference, during-war minus pre-war, in the non-violent zone. In other words,  $\beta$  gives the cross effect on non-violent deaths from being in the violent zone during the war after controlling for the effect of being in the violent zone and the effect of being in the during-war period. The estimates from this simple model offer no evidence that the key  $\beta$  coefficient might be positive (see appendix for Table A3).<sup>10</sup>

Our main DiD estimates use a generalization of the above framework that does not force locations into the binary categories of either violent or non-violent (equation 2).

$$y_{ht} = \beta_1 t + \beta_2 \text{ViolenceLevel}_h + \beta_3 (t * \text{ViolenceLevel}_{ht}) + \epsilon_{ht} \quad (2)$$

Table 2 gives the results from these regressions. Violence is measured as the number of fatalities per 100 inhabitants in the governorate containing household  $h$  ( $\text{ViolenceLevel}_{gov}$ ) except in column 6 where violence is measured at the cluster level ( $\text{ViolenceLevel}_{cluster}$ ). In one model we use a spatial lag of violence (column 5,  $\text{ViolenceLevel}_{neighbours}$ ) to account for cross-governorate violence spillovers. To facilitate easy comparisons between the models, we place all variables on a common scale by subtracting the mean and dividing by two times the standard deviation (Gelman, 2008).

The variable of interest is the product of the time period ( $\text{Period}$ ), coded 1 for the war period and 0 for the pre-war period, and the violence level. This estimated coefficient gives the relationship between local violence and local non-violent death counts during the war period, i.e. after controlling for the effect of living, both before and during the war, in a governorate (or cluster) that turns out to be violent during the war. These key estimates ( $\text{Period} * \text{ViolenceLevel}_{gov}$ ) are all 0, rounded to the first significant decimal point, or negative in the first five specifications. Specification 6 provides, at best, a tiny scrap of evidence for a positive impact of violence on non-violent deaths; but even in this specification we cannot reject a hypothesis of a 0 or negative effect at a standard significance level ( $P$ -value of 0.18). These results are consistent with our earlier null findings both for non-violent excess deaths and in the binary DiD.

### Comparisons of Counterfactuals

We have searched for possible indirect effects of violence on non-violent death rates using two different methods, based on two different counterfactual analyses. We now briefly consider the difference between the two counterfactuals.

The key assumption in the excess deaths counterfactual is that the pre-war death rate would have carried forward in time if war had never broken out, i.e. the pre-war mortality rate is what the during-war mortality rate would have been without the war. If this assumption is approximately true then it is reasonable to view the actual death rates minus the baseline death rate as excess deaths caused by war. A problem with this counterfactual is that there can be important changes that affect mortality rates that are separate from and unconnected with war onset but that, nevertheless, happen at roughly the same time as onset does. Moreover, the Human Security Report (HSR) made an important critique of the excess deaths concept while accepting its main premise of extrapolating pre-war death rates forward in time to create a baseline (Human Security Report Project, 2011). HSR argued that death rates tend to decrease over time so it recommended projecting forward a decreasing death-rate trend rather than a flat line frozen at a pre-war rate. Thus,

**Table 2.** Differences-in-differences regression models for the number of non-violent deaths per household.

Specification	Poisson	Poisson	NB	Quasi-Poisson	Poisson	Poisson
	(1)	(2)	(3)	(4)	(5)	(6)
Period	1.4*** (0.2)	1.5*** (0.2)	1.4*** (0.2)	1.5*** (0.2)	1.4*** (0.2)	1.5*** (0.2)
ViolenceLevel <sub>gov</sub>	0.1 (0.2)	0.4 (0.4)	0.1 (0.2)	0 (0.2)	0.1 (0.2)	
Period*ViolenceLevel <sub>gov</sub>	0 (0.4)	0 (0.4)	0 (0.4)	-0.3 (0.5)	0 (0.4)	
ViolenceLevel <sub>neighbours</sub>					0 (0.2)	
ViolenceLevel <sub>cluster</sub>						-0.1 (0.2)
Period*ViolenceLevel <sub>cluster</sub>						0.5 (0.4)
Intercept	2.62*** (0.10)	-2.9*** (0.3)	2.62*** (0.10)	2.6*** (0.1)	-2.62*** (0.10)	-2.62*** (0.10)
AIC	2024.4	2033.6	2019.2	-	2026.4	2022.7
AUC	0.647	0.682	0.648	0.643	0.650	0.653
Governorate FE	-	Yes	-	-	-	-

FE: fixed effects; NB: negative binomial; AIC: Akaike information criterion; AUC: area under curve.

Notes. Robust standard errors, clustered at cluster-level, are in parentheses where \*\*\*, \*\*, and \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.  $N = 3369$ . For column 6, the  $p$ -value for Period\*ViolenceLevel<sub>cluster</sub> is 0.18.

HSR would interpret deviations from pre-war trends, rather than deviations from pre-war levels, as evidence for excess deaths. The HSR modification is a step forward in principle. However, in practice it is quite difficult to accurately measure a pre-war death rate, let alone a trend. Hence, adopting the HSR recommendation would be likely to add considerable random noise into excess death calculations as a price for improving their theoretical foundations.<sup>11</sup>

The implied counterfactual for the binary DiD is that the trend in the non-violent zone, pre-war to during-war, gives the trend the violent zone would have experienced without the war. This approach potentially improves on the standard excess deaths approach by allowing for a common, but possibly nonzero, national mortality rate trend without war rather than mechanically applying a pre-war baseline mortality rate. Moreover, the binary DiD addresses the main HSR concern by baselining a trend, rather than just a fixed level. But the binary DiD has at least two important weaknesses. First, pre-war trends could be very different in the two zones, in which case a DiD regression might, spuriously, attribute differing during-war trends to violence. The standard practice in the literature of checking for common earlier trends will often be infeasible in war mortality studies because there will not be a long enough run of data to do so.<sup>12</sup> Second, this approach only picks up localized spillovers from violence to non-violent death rates. A DiD regression would give insignificant results if, for example, violence causes non-violent mortality rates to increase equally in both violent and non-violent zones.

## Conclusions

We use the same data as the original authors did and refute the claim (Hagopian et al., 2013) that the Iraq war caused half a million excess deaths. Our analysis is superior to the original work because we account for i) most importantly, the distinction between violent and non-violent deaths; ii) the stratified sampling design; iii) uncertainty over deaths not confirmed by death certificates; and iv) the baseless increase from 406,000 up to 500,000.

The war in Iraq has certainly led to several hundred thousand violent deaths and perhaps some non-violent excess deaths as well.<sup>13</sup> However, methodologically sound estimates based on the UCIMS data struggle to rule out even negative excess death estimates for violent plus non-violent deaths combined. This failure probably reflects data and methodological weaknesses as much as the underlying reality.

Our paper offers methodological lessons that extend beyond the Iraq application. First, excess death calculations should always include estimates for non-violent deaths alone. Second, researchers should supplement excess death estimates with DiD ones. These ideas have been helpful in our Iraq analysis and should bear further fruit in the future.

## Acknowledgements

We are very grateful for useful comments from Beth Daponte, Nils Petter Gleditsch, Joakim Kreutz, and Jon Pedersen, as well as the anonymous reviewers and the editor Kristian Skrede Gleditsch. We also thank Richard Podkolinski for programming assistance

and participants in the session Population and Health Impacts of Poverty and Armed Conflict at the 2015 International Studies Association (ISA) conference where an early version of this work was presented.

### Declaration of Conflicting Interest

The authors declare that there is no conflict of interest.

### Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Supplementary Material

The supplementary files are available at <http://journals.sagepub.com/doi/suppl/10.1177/2053168017732642>. The replication files are available at <http://thedata.harvard.edu/dvn/dv/researchandpolitics>.

### Notes

- Human Security Report Project (2011) debunked the Democratic Republic of Congo (DRC) estimate mainly for non-random sampling in its early rounds, for comparing during-war mortality rates with an inappropriately high baseline mortality rate, and for mortality rate estimates that are incompatible with other DRC mortality evidence. Guha-Sapir (2010) criticized the Hagan and Palloni (2006) estimate for excessive reliance on mortality rates in refugee camps, which might diverge substantially from in-country mortality rates
- There is a large potential pay off to exploiting geographical variation in violence levels in conflicts beyond the Iraq war since war intensity tends to vary greatly over space (Ward and Gleditsch, 2002; Buhaug and Gleditsch, 2008).
- We expand on point iii in the subsection 'From 400,000 to 500,000'.
- See the original paper for further methodological details.
- Typically, both the real and baseline death rates are estimated rather than calculated precisely, so we end up with an *estimated* excess death rate and an *estimated* number of excess deaths.
- There can be good reasons for stratification. For example, stratification can increase estimation precision in strata with high variation by taking relatively large samples in these strata, but it must be built into one's estimation scheme.
- Nothing is recorded in the death certificate field for three deaths
- The lead author of the study even characterized 500,000 as 'likely a low estimate' (KPLU, 2013; National Geographic News, 2013).
- To avoid confusion note that deaths are measured twice in each household, once for the pre-war and once for the during-war period whereas each household is either in the violent zone or in the non-violent zone during both periods.
- We use Poisson, Quasi-Poisson and Negative Binomial models to handle the count data and to account for possible overdispersion (Gelman and Hill, 2007).
- Indeed, the HSR suggestion is unworkable with the UCIMS data for the Iraq war because the pre-war measured downward trend is so steep in this dataset that the baseline quickly

plunges unrealistically all the way down to zero early in the during-war period (see appendix for Figure A2).

- This is the case with the Hagopian et al. (2013) study.
- See, for instance, the Iraq Body Count project for more information (<https://www.iraqbodycount.org/>).

### Carnegie Corporation of New York Grant

This publication was made possible (in part) by a grant from Carnegie Corporation of New York. The statements made and views expressed are solely the responsibility of the author.

### References

- Angrist JD and Pischke JS (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Buhaug H and Gleditsch K (2008) Contagion or Confusion? Why Conflicts Cluster in Space. *International Studies Quarterly* 52(2): 215–233.
- Cecchi F and Roberts L (2005) Interpreting and using mortality data in humanitarian emergencies: a primer for non-epidemiologists. *Humanitarian practice network, Overseas Development Institute*. No. 52, September 2005. Available at: <http://odihpn.org/wp-content/uploads/2005/09/network-paper052.pdf>
- Degomme O and Guha-Sapir D (2010) Patterns of mortality rates in Darfur conflict. *The Lancet* 375(9711): 294–300.
- Flynn T and Peters T (2004) Use of the bootstrap in analysing cost data from cluster randomised trials: Some simulation results. *BMC Health Services Research* 4: 33.
- Gelman A (2008) Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* 27(15): 2865–2873.
- Gelman A and Hill J (2007) *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.
- Hagan J and Palloni A (2006) Death in Darfur. *Science* 313(5793): 1578–1579.
- Hagopian A, Flaxman A, Takaro T, Esa al Shatari S, Rajaratnam J, Becker B, Levin-Rector A, Galway L, Hadi Al-Yasseri B, Weiss W, Murray C and Burnham G (2013) Mortality in Iraq Associated with the 2003–2011 War and Occupation: Findings from a National Cluster Sample Survey by the University Collaborative Iraq Mortality Study. *PLOS Medicine* 10(10): e1001533. Available at: <http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001533>
- Human Security Report Project (2011) *Human Security Report 2009/2010: The Causes of Peace and the Shrinking Costs of War*. Oxford: Oxford University Press.
- International Rescue Committee (2007) Mortality in the Democratic Republic of Congo: An ongoing crisis. Available at: <https://www.rescue.org/report/mortality-democratic-republic-congo-ongoing-crisis>
- KPLU (2013). Seattle researchers: Death toll for Iraq War likely near half million. Website KPLU 88.5 News for Seattle and the Northwest Available at: <http://knkx.org/post/seattle-researchers-death-toll-iraq-war-likely-near-half-million> (accessed 18 September 2017).
- Lohr S (2009) *Sampling: Design and Analysis*, 2nd edn. Pacific Grove, CA: Duxbury Press.

- National Geographic News (2013) Half-million Iraqis died in the war, new study says. National Geographic News. Available at: <http://news.nationalgeographic.com/news/2013/10/131015-iraq-war-deaths-survey-2013/> (accessed 18 September 2017).
- NBC News (2008) Death toll of 200,000 disputed in Darfur. NBC News. (accessed 19 May 2015. 18 March 2008. Available at: [http://www.nbcnews.com/id/23848444/ns/world\\_news-africa/t/death-toll-disputed-darfur/#.Wb\\_r4ciGOM8](http://www.nbcnews.com/id/23848444/ns/world_news-africa/t/death-toll-disputed-darfur/#.Wb_r4ciGOM8) (accessed 18 September 2017).
- Pons O (2007) Bootstrap of means under stratified sampling. *Electronic Journal of Statistics* 1: 381–391.
- ScienceInsider (2010). How Many Have Died Due to Congo's fighting? Scientists battle over how to estimate war-related deaths. Science. 21 January 2010. Available at: <http://www.sciencemag.org/news/2010/01/how-many-have-died-due-congos-fighting-scientists-battle-over-how-estimate-war-related> (accessed 18 September 2017).
- Shalizi C (2010) The bootstrap. Statisticians can reuse their data to quantify the uncertainty of complex models. *American Scientist* 98(3): 186–190.
- Spagat M, Mack A, Cooper T and Kreutz J (2009) Estimating war deaths: An arena of contestation. *Journal of Conflict Resolution* 53(6): 934–950.
- University of Washington (2013). Study: Nearly 500,000 perished in Iraq war. University of Washington. Bobbi Nodell 15 October 2013. Available at: <http://www.washington.edu/news/2013/10/15/study-nearly-500000-perished-in-iraq-war/> (accessed 18 September 2017).
- Ward MD and Gleditsch K (2002) Location, location, location: An MCMC approach to modeling the spatial context of war and peace. *Political Analysis* 10(3): 244–260.
- Working Group for Mortality Estimation in Emergencies (2007). Wanted: studies on mortality estimation methods for humanitarian emergencies, suggestions for future research. *Emerging Themes in Epidemiology* 4: 9.