Estimating Indirect War Deaths Directly

by

Michael Spagat

Department of Economics

Royal Holloway University of London

and

Stijn van Weezel

School of Economics

University College Dublin

Paper presented at ISA San Francisco, April, 2018

Abstract

There has been growing recent interest in measuring and accounting for non-violent deaths indirectly caused by war. However, the standard method for estimating such deaths is surprisingly indirect. It can be broken down as follows; 1) measure pre-war and during-war death rates for violent and non-violent deaths *combined*: 2) subtract the former from the latter to arrive at an "excess death rate": 3) separate the excess death rate into violent deaths and non-violent deaths indirectly caused by war. We suggest a more direct and parsimonious method; 1) measure pre-war and during-war non-violent death rates: 2) subtract the former from the latter to arrive at a non-"violent excess death rate." We show that the two methods are equivalent in a world where all these rates are measured perfectly but they diverge with imperfect measurement, at least under the standard paradigm of null hypothesis significance testing (NHST). We simulate measurement of these rates through sample surveys and find that our proposed (direct) method outperforms the standard (indirect) method when samples are small enough or when samples are large but the true non-violent excess death rate is small enough. We caution, however, that these simulations are not yet well calibrated to real data. We also show that the standard approach denies readers vital information on uncertainty about excess non-violent deaths. This suppression of uncertainty leads people to believe that estimated numbers are more precise than they really are.

1. Introduction

It is obvious that war violence leads directly to violent deaths. Many projects document these deaths one by one and/or grouped into the events in which they occurred. Other projects have used statistical methods (see this and this) to estimate the number of violent deaths in various conflicts. The methods and results in this field are big subjects that we will not address here.

It is equally obvious that war violence can lead indirectly to non-violent deaths and anyone can easily imagine possible transmission mechanisms. A water purification system could be sabotaged, causing the spread of waterborne diseases. A heart attack victim may die before an ambulance driver braves driving through a battle zone. Electricity outages may doom people who depend on life support machines. The ease with which we can construct such examples lays bare the sheer magnitude of the task of quantifying the scale of non-violent deaths indirectly caused by war.

We know of two distinct approaches for quantifying indirect non-violent war deaths. One is to investigate each non-violent death occurring during a war with the goal of determining whether we can convincingly trace it back to war violence. Such research can be illuminating although it faces two big challenges. First, it is very time consuming and resource intensive to seek ultimate explanations for every single death occurring during a war. So only very well financed projects can attempt such work. Second, even after intensive investigation it still will not be possible to make convincing binary classifications of deaths into either “caused by war violence” or “not caused by war violence”. Many deaths will fall somewhere in between these extremes with war perhaps playing some role but not a unique dispositive one. By no means do we dismiss the case-by-case approach but we will not discuss it further in this paper.[1]

The second approach is to compare during-war death rates with counterfactual death rates which, we assume, would have happened if there had never been a war. The most common version of this idea is to take the pre-war death rate as the counterfactual one. In this case the key assumption is that the pre-war death rate

---

[1] Lozado (2018) gives a good discussion of types of deaths indirectly caused by Hurricane María in Puerto Rico and the challenge confronting any project seeking to find all deaths indirectly attributable to the Hurricane.

would have carried forward unchanged if the war had never happened. The rest of our paper is focused on this version of the second approach.[2]

2. A Conceptual Problem - Indirect Measurement of Indirect Deaths

The standard method for estimating non-violent deaths indirectly caused by war is surprisingly indirect. It starts by unnecessarily mixing violent deaths in with non-violent deaths, even though the latter are the actual objects of interest. The violent deaths are then removed at a later stage.[3]

In particular, under the standard method we start by calculating two death rates: a during-war death rate and a pre-war death rate, both of which include both violent and non-violent deaths. The during-war death rate minus the pre-war death rate is called the “excess death rate.” Suppose that this rate is positive and that we accept the implied counterfactual that the pre-war rate would have continued forward without the war.[4] Then we can interpret the excess death rate as a rate of violent plus non-violent deaths caused by the war.

Note that we are now in a peculiar situation. We started by declaring an interest in non-violent deaths caused indirectly by war violence but we now have an estimate that mixes in violent deaths with the non-violent ones that we set out to understand. So we must now make an estimate of the number of violent deaths and subtract this from our excess death number. Only after following this circuitous path do we arrive at an estimate of non-violent excess deaths. If we make the further step of accepting the counterfactual assumption then we can view our estimate of non-violent excess deaths as giving non-violent deaths *caused* indirectly by war violence.

This is a surprisingly roundabout method for studying indirect deaths. The following analogy seems apt. We set out to determine whether introducing a polluting factory

[2] Jewell, Spagat and Jewell (forthcoming) gives an overview of this approach.

[3] See, for example, the Methods and Findings section of Hagoppian et al. (2013). It gives an excess deaths estimate for violent and non-violent deaths combined and then states that 60% of these are violent, implying that 40% are non-violent.

[4] It is important to maintain a critical attitude toward the counterfactual assumption because it might not hold in many real situations. Indeed, during war death rates can be lower than pre-war death rates in which case naively assuming that the counterfactual assumption must be true would force us to believe that a war is saving lives.

to a community has caused an increase in cancer deaths. So we measure the rate of cancer deaths plus heart disease deaths after the introduction of the factory and subtract off the rate of cancer deaths plus heart disease deaths before the introduction of the factory. We then deduct the part of this joint increase that is due to heart disease only to arrive at the increase in cancer deaths. This method may take us where we want to go but, at a minimum, it is unnecessarily complicated.

The following is a direct and intuitive method to estimate the number of non-violent excess deaths in a war. Start by calculating two death rates: a during-war *non-violent* death rate and a pre-war *non-violent* death rate. We call the during-war non-violent death rate minus the pre-war non-violent death rate the "excess non-violent death rate". If we then accept the counterfactual assumption we can interpret this number as our estimate of the rate of non-violent deaths caused indirectly by the war.

3. The Two Methods are Equivalent when Measurement is Perfect

In a world of perfect measurement the standard (indirect) method for estimating non-violent excess deaths is equivalent to our proposed direct method. We can see this clearly from the following example. In Table 1 "0" is the pre-war violent death rate, "B" is the during-war violent death rate, etc..

Table 1. An Example with Perfectly Measured Death Rates

	Violent Deaths	Non-Violent Deaths
Pre-war	0	C
During-war	B	D

The standard (indirect) method first calculates the excess death rate for violent plus non-violent deaths combined. This is (B + D) – (C + 0) = B + D – C. Next we remove during-war violent deaths by subtracting B. We arrive at D – C.

Our proposed direct method subtracts pre-war non-violent deaths, C, from during-war non-violent deaths, D. Again, we arrive at D – C.

4. The Two Methods can Diverge when Measurement is Imperfect

Everything in table 1 will be imperfectly measured in practice. Researchers normally quantify the uncertainty surrounding such estimates with uncertainty intervals (UI's). In addition, it is common practice to supplement UI's with null hypothesis significance tests (NHST). Of course, there are alternatives to NHST, such as Bayesian approaches, but NHST still dominates most practice.

Thinking in terms of NHST and non-violent deaths indirectly caused by war violence an obvious hypothesis one would like to reject is that excess non-violent deaths are negative. In most of the scientific world such a rejection would pass the research over a statistical significance bar, enabling a claim that war violence is associated with an increase in non-violent death rates (causing if you believe the counterfactual).

The NHST just described boils down to checking whether the bottom of a UI for non-violent excess deaths is greater than zero. In the terminology of Table 1 this means checking whether the bottom of the UI around D – C is greater than zero. However, all excess deaths estimates we know of perform this test on excess deaths, violent plus non-violent combined, rather than on non-violent excess deaths. In other words, the statistical significance test is applied to the UI around B + D – C rather than to a UI around D – C. The problem is that adding violent deaths in to an estimate before checking for statistical significance helps boost the estimate over the significance bar that must be cleared. Intuitively, this effect can be substantial if there are a large number of violent deaths.

Thus, we arrive at a problem that is located at the intersection of the standard (indirect) approach to estimating excess deaths and that stand method of NHST. In effect, the standard method allows researchers to bring in a stack of violent deaths just at the moment when it is time to clear the statistical significance bar.

We devote the remainder of the paper to comparing the two methods while maintaining a NHST perspective. It is possible that we could make good progress, even possibly restoring equivalence between the two methods, by abandoning the NHST paradigm. But we will not pursue such options here, leaving them to possible future work.

5. The Problem with the Standard Method is not just Theoretical

Hagopian et al. (2013), a survey of deaths in Iraq just prior to and during the war that began in 2003, falls directly into the trap we describe in section 3. This quote conveys the essence of their analysis:

> "From March 1, 2003, to June 30, 2011, the crude death rate in Iraq was 4.55 per 1,000 person-years (95% uncertainty interval 3.74–5.27), more than 0.5 times higher than the death rate during the 26-mo period preceding the war, resulting in approximately 405,000 (95% uncertainty interval 48,000–751,000) excess deaths attributable to the conflict….We estimate that more than 60% of excess deaths were directly attributable to violence, with the rest associated with the collapse of infrastructure and other indirect, but war-related, causes."

Note the following key points. First, the 405,000 estimate is off excess deaths, i.e., of violent and non-violent death combined. Second, the estimate of 405,000 excess deaths are then divided into 60% violent and, implicitly, 40% non-violent excess deaths. In other words we can infer an estimate of roughly 160,000 excess non-violent deaths from the information provided in the abstract. Hagopian et al. 2013 does not provide a UI for non-violent excess deaths but Spagat and van Weezel (2017) does: -210,000 to +410,000.[5]

This means that under any system that at least remotely resembles NHST we would not reject a hypothesis asserting that the number of non-violent excess deaths is less than 0. Indeed, this is not even a borderline case; the proposed test is already one-sided and we cannot even reject it at a 10% level (p = 0.17). Of course, NHST has its shortcomings and we would agree that the number of non-violent excess deaths in Iraq is probably positive. At the same time we think the information that this UI is 600,000 deaths wide and starts more than 200,000 deaths below 0 is very important and should engender great scepticism about any confident claim of 10's of thousands of non-violent deaths indirectly caused by the war. Hagopian et al. (2013) do not provide this information, in line with the practice from the standard (indirect) method.

[5] Spagat and van Weezel (2017) consider a number of scenarios and this is just one of several estimates with UI's provided but this one follows the Hagopian et al. (2013) methodology most closely. Hagopian et al. (2018) replied to this critique and Spagat and van Weezel (2018) made a rejoinder but the issue driving the present paper was not debated in this discussion.

From the point of view of the present paper here is the key point; violent deaths pull the Hagopian et al. (2013) estimate up and over the statistical significance bar. Thus, this example perfectly fits the scenario described in Section 3.

6. The Standard Approach Suppresses Uncertainty

There is a further problem with the standard (indirect) approach already alluded to in section 6. As noted above, the standard approach boils down to first making an excess-deaths estimate for violent and non-violent deaths combined and then splitting this estimate into violent and non-violent components. A subtle consequence of this two-step procedure is that a UI is calculated only for the combined excess-deaths estimate but not for the non-violent excess death component of this estimate.[6] This means that estimates of non-violent excess deaths are, d*e facto*, treated as certain.

Hagopian et al. (2013) again fits this pattern. They publish a UI for excess deaths but not for non-violent excess deaths. Instead, they just specify that the excess deaths estimate divides into 40% excess non-violent and 60% violent. People quoting the work are confronted with a choice of either quoting just the central estimates for excess non-violent and violent deaths or making their own UI calculations, a path that is onerous and not feasible for many readers.

The psychological effect of hearing, on the one hand, an estimate of 160,000 non-violent excess deaths versus hearing, on the other hand, an estimate of 160,000 non-violent excess deaths with a UI of -210,000 to +410,000 is large. Hence the suppression of uncertainty by the standard approach is a serious problem.

7. Simulations help us to Generate a Deeper Understanding

We have started doing simulations to further illuminate the problems created by the standard (indirect) approach to excess deaths estimation. We were also interested

[6] Sometimes UI's are published for a violent deaths estimate but not for the non-violent component of excess deaths.

in discovering whether the standard method may have some advantages we have not thought of.

Our simulations proceed as follows:

1. We randomly generate two million households with an average size of 5. These results are meant to represent a war zone, possibly a whole country at war.

2. We randomly assign deaths to some of the households. These are labelled as either pre-war or during-war and as either violent or non-violent. In all the simulations we have done so far the pre-war violent death rate is 0, the during-war violent death rate is 5 per 1,000 and the pre-war non-violent death rate is also 5 per 1,000.[7] We let the non-violent during-war death rate vary between 2.6 and 7.7 so that the true non-violent excess death rate varies between -2.4 and +2.7.

3. We simulate household surveys. For each value of the non-violent excess death rate we draw 1,000 samples of 10,000 households, 1,000 samples of 2,000 households and 1,000 samples of 100 households.

4. For each sample we estimate the non-violent excess death rate using both the standard (indirect) method and our proposed (direct) method.

5. In each sample we compute the error of each estimate. This step is important and tricky so consider the following breakdown;

a. Note that for every simulation we have created the two million households and, therefore, know all the relevant facts about them. In particular, we always know the true non-violent excess death rate. These non-violent excess death rates can only be estimated with error in our simulated surveys.

b. We apply a NHST perspective throughout. In particular, if the bottom of the UI on any estimate is below 0, i.e., if the statistical significance bar is not cleared, then we set the estimate for non-violent excess deaths equal to zero. This is a crude way to proceed but it is roughly in line with what normally happens in the scientific world. We plan to explore different scenarios in future work but this procedure seems like an appropriate baseline to start with.

c. Specifically, the estimates for each method are as described in section 3 when the statistical significance bar is cleared and 0 when the statistical significance bar is not cleared.

[7] We concentrate the violent deaths in just 300,000 of the households situated in 20 clusters/provinces. This fact does not matter for our simulations so far but this clustering is meant to pave the way to consider cluster surveys at a later date.

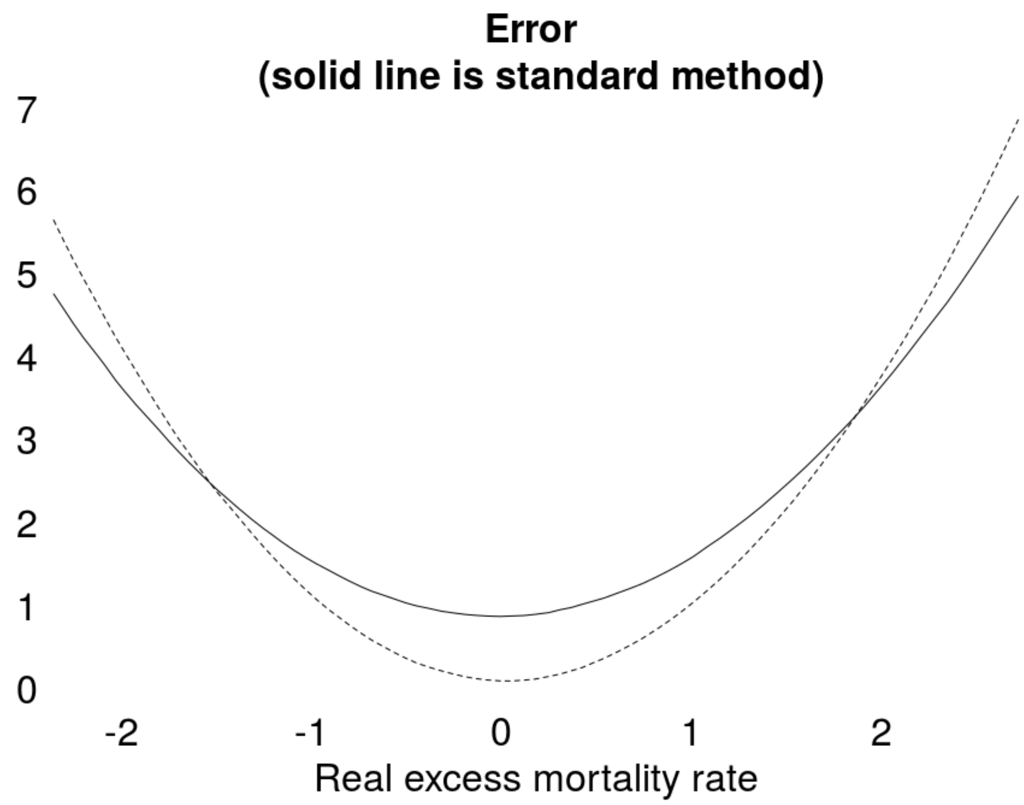
We make two important points before sharing the results of our early simulations. First, one should not take too seriously the specific numbers that emerge from these early results. The simulations are fairly complicated and we have not managed yet to calibrate the numbers to realistic values that might come up in practice. In particular, we are generating the deaths using a zero-inflated Poisson process for which the variance grows linearly with the mean. This means that when death rates are high the variances in deaths across households are also high, probably much higher than they are in real data. We believe this issue is distorting our simulation results at the moment. Still, we think that these simulations are already providing interesting results that we can build on in the future.

Second, some readers may be surprised by the thought of negative excess non-violent death rates but they should not be. This confusion probably arises from an inappropriately easy acceptance of the counterfactual assumption. If you simply assume that any difference between pre-war and during-war death rates is caused by war then a negative difference can only mean that a war is saving lives. Note that the Human Security Report examined many modern wars and found that national child mortality rates almost always decline over time, even during wars. Crucially, this is not because wars *cause* improvements in child health but, more likely, because improvements in child health have their own positive dynamics that are not completely derailed by war. Thus, it can easily happen that during-war rates can be lower than pre-war rates and it is important for us to consider such cases in our simulations.

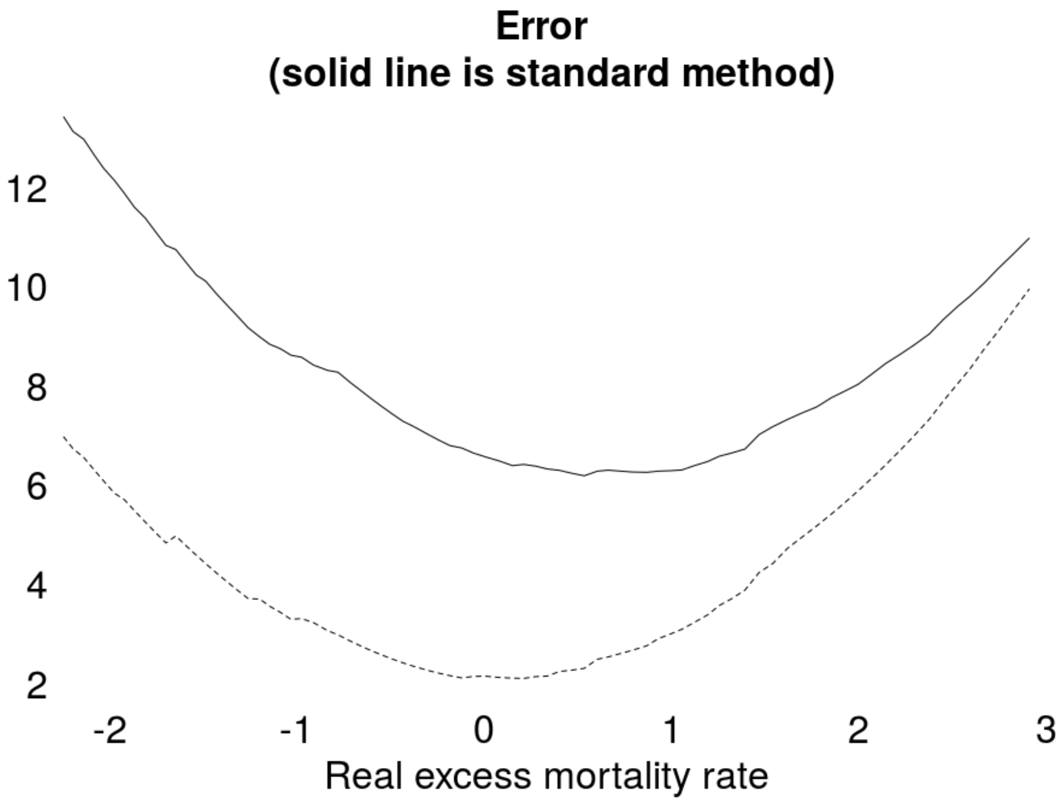
The next picture displays the results of the simulations for surveys of 10,000 households. Note that our surveys of 10,000 households are quite big, especially since they are not cluster surveys for which sample size is probably better understood as the number of cluster than as the number of households. The curves in the picture display the mean squared error for each method, averaging over the 1,000 runs and for each true value of the non-violent excess death rates.

We make the following observations about the picture. First, the direct method outperforms the standard method when the real excess non-violent death rate is 0 or close to 0. This makes sense. When the real rate is 0 then the direct rate is very likely to lead to a statistically insignificant result which will result in a 0 error. The indirect method, on the other hand, might get carried over the statistical significance bar by violent deaths. There is a similar phenomenon when the true non-violent excess death rates are near. However, now the difference in mean squared errors between the two techniques is smaller because the direct method frequently sets the estimates to 0 (statistical insignificance) when the true rates are slightly different from 0. In addition, the indirect method is penalized for variability; almost every estimate is statistically significant and that means that random overestimation and random underestimation cause errors that are avoided by the less variable direct method. However, the advantage flips to the indirect method when the real excess death rate grows sufficiently. This is because the continued tendency for the direct method to give statistically insignificant results, even some of the time, becomes a liability when the real rate is far enough from 0.

The next picture is constructed exactly like the last one but it is for samples of 2,000 households. It is qualitatively similar to the previous picture but the range of advantage for the direct method is now wider because, with smaller sample sizes, the estimates from the indirect method are more variable than they were with the smaller sample sizes.



The third picture is, again, showing the same things but now the sample sizes are 100. This may sound like an irrelevantly small sample size but, recall, that we are not yet simulating cluster samples. Thus, we can take our simple random samples of 100 as approximating what might happen in a cluster sample with 100 or fewer clusters. Cluster samples with 30 to 100 have been common in the literature (see this). The direct method performs much better in samples of this size at least for the class of simulations we have performed so far.



Finally, we would just like to reiterate that the actual numbers, e.g., the precise positions of the curves and where they cross should not be taken too seriously at this stage since our data generation process has not yet been calibrated to reality.

8. Conclusion

The direct method for estimating indirectly excess deaths always has the virtue that it quantifies uncertainty through an uncertainty interval attached to each estimate. The indirect method does not have this virtue.

Our simulations identify two broad classes of cases when the direct method outperforms the indirect method in the contest we run in the present paper. The first case is when estimation is imprecise as it will be with a small sample. The advantage of the direct method in very small samples appears to be strong and general, at least based on the range of simulations we have performed for this paper. The second broad class of cases seems to be when the true excess non-violent death rate is small, even in large samples. The larger the sample the smaller the true non-violent death rate needs to be for the direct method to have a smaller average mean squared error than the indirect one does. Of course, these results also suggest that there are situations (relatively large samples and large true excess non-violent death rates) when the indirect method outperforms the direct one according to our criteria.

We reiterate some caveats. Possibly better calibration of our simulated household with real data could change some of our findings. Also, we could consider other systems for interpreting and scoring results. In particular, a more Bayesian approach may change some of our findings based on NHST.

We will continue researching this subject and report back soon.

References

These are just built into links for now. We apologize for the inconvenience.