Extended Abstract

A Generalized Method for Reconstruction of War-Related Mortality

Dmitri A. Jdanov*, Domantas Jasilionis*, Dana A. Glei**,
and Vladimir M. Shkolnikov*

* *Max Planck Institute for Demographic Research, Rostock,*
** *University of California, Berkeley*

1. INTRODUCTION

War can have a substantial impact on mortality rates of a population, particularly among young adult males. Yet, official vital statistics data often exclude military deaths that occur abroad. Thus, estimates based on these data may vastly underestimate the true level of mortality during war periods. For example, among French males in 1915, e_0 was nearly 14 years greater for civilians compared with the overall population (40.8 vs. 27.0, respectively; HMD, 2007). For males aged 20-24, the mortality rate for the entire population was more than five times greater than that of the civilian population (9.6% vs. 1.8%; HMD, 2007). Similar figures for England and Wales also suggest that civilian mortality substantially underestimates the true mortality level during wartime.

Estimates of civilian mortality may also be biased by failure to account for population movements during wartime. Population estimates derived using the intercensal survival method do not account for the mobilization of military troops. Consequently, mortality rates based on these data are plagued by numerator-denominator bias because the numerator represents largely the civilian population whereas the denominator represents the total population.

In this paper, we propose a modeling strategy for estimation of mortality during war periods in which sufficiently detailed demographic data are not available. Our goal is to employ methods that enable us to estimate mortality surfaces from minimal input data while ensuring maximal comparability across time and countries. We follow a strategy which was successfully applied for estimating war-related and total mortality surfaces for England & Wales (Jdanov et al., 2005). This approach (with some modifications) has been tested for several other countries within the Human Mortality Database (HMD) project. However, estimating wartime mortality presents a special challenge because the availability, coverage, and completeness of data can vary considerably across countries. This paper is an attempt to further develop a methodology that can be used across various HMD countries affected by wartime mortality.

2. PREVIOUS APPROACHES

Over the years, various researchers have attempted to estimate total wartime mortality for particular countries. Several studies have attempted to estimate total deaths by country or evaluate total war-related losses. For every European country, aggregated estimates of battle, non-battle, and total human losses for both WWI and WWII were published by Urlanis (1971). In his work on *"Population changes in Europe since 1939"*, Frumkin (1951) provided data on demographics losses during WWII for each of 25 European countries. These works are based on the official sources or indirect estimations by other researchers and give only aggregated numbers of losses. More detail data can be found in country specific studies such as on England & Wales (Winter, 1976, 2003), France (Vallin, 1973), Italy (Glei et al., 2005) or Germany (Overmans, 2000).

3. DATA INPUTS

Cross-national comparability may be better served by a more general approach that requires fewer idiosyncratic assumptions. The model presented here is based on the same principles as the one used to reconstruct war losses for England and Wales (Jdanov et al., 2005) and can be adapted to other countries even when we lack detailed knowledge of historical events during wartime. Nonetheless, the availability of data differs across countries and the model must be modified accordingly. Furthermore, the modeling strategy still requires some assumptions, but they are more general in nature and could be applied in other contexts. While no single method for estimating mortality during wartime can be applied uniformly across countries, by following the same general principles we can produce estimates with greater consistency and comparability.

The necessary data inputs used in the model are summarized here:

1. The last pre-war and first post-war census counts by sex and age.
2. Civilian death counts by sex, age (here we can use standard HMD methods to get one-year age groups) and calendar year.
3. Total military deaths by sex and calendar year.

Other more detailed data can be used to specify the model and may result in more precise estimates, but such data are not required by the model.

4. MODEL

We use t_{begin} to refer to the year in which war began, t^* to denote the year in which war ended for the selected country, and t_{end} to indicate the last year of the war period. For modeling purposes, we divide the war period into the period of active war t_{begin} to t^* and the post-war period t^*+1 to t_{end}.

The first step is to estimate population size on January 1st of the year war began. Starting with the pre-war census count and civilian deaths, we use standard Human Mortality Database (HMD) methods (Wilmoth et al., 2005) to: i) redistribute counts of unknown age, ii) split deaths and population counts by birth cohort, and iii) calculate post-censal estimates $P(z,t_{\mathrm{begin}})$, representing the population size of cohort z on January 1st of year t_{begin}.

The aim of the second step is to distribute military deaths by cohort for each year of the war period:

$$\pi(z,t) = \frac{D^{\mathrm{m}}(z,t)}{\widetilde{\mathrm{D}}^{\mathrm{m}}(t)} \ ,$$

where $\widetilde{D}^{\mathrm{m}}(t)$ represents a known quantity: total military deaths for year t. With regard to notation, we use a tilde (~) above a variable to denote input data; other quantities remain unknown and must be estimated. The proportion $\pi(z,t)$ depends on the relative size of the cohort and the probability of military death:

$$\pi(z,t) = \frac{q^{\mathrm{m}}(z,t) \cdot P(z,t)}{\sum_{i=t-60}^{t-18} q^{\mathrm{m}}(i,t) \cdot P(i,t)} \ ,$$

where $q^{\mathrm{m}}(z,t)$ represents the probability of military death (among the total population) for cohort *z* during year *t*, and $P(z,t)$ is the size of the total population of cohort *z* on January 1st of year *t*. For simplicity, we assume the military population is restricted to ages 18-59 (i.e., cohorts $t-60$ to $t-18$). The probability $q^{\mathrm{m}}(z,t)$ is further defined as:

$$q^{\mathrm{m}}(z,t) = D^{\mathrm{m}}(z,t) / P(z,t) \ .$$

At this point, we know the population for the year the war began $P(z,t_{\mathrm{begin}})$ (obtained in Step 1), but $q^{\mathrm{m}}(z,t)$ and $D^{\mathrm{m}}(z,t)$ remain unknown. To estimate these quantities, we assume that $q^{\mathrm{m}}(z,t)$ follows a gamma p.d.f. with a maximum between ages 18 ($t-z=18$) and 24 (see example for France in Vallin, 1973, and for England & Wales in Winter, 1976). Thus, $q^{\mathrm{m}}(z,t)$ is described by parameterized function with parameters $a(t)$ and $b(t)$. We fit parameters $a(t)$ and $b(t)$ for $t = t_{\mathrm{begin}},...t_{\mathrm{end}}$ by minimizing differences between the input data and the resulting estimates of military deaths:

$$\sum_{t=t_{begin}}^{t_{end}} \left(\frac{\sum_{x=18}^{60} \hat{D}^{\mathrm{m}}(x,t)}{\widetilde{D}^{\mathrm{m}}(t)} \right)^2 + \sum_{t^* < t \le t_{end}} \left(\frac{\sum_x M(x,t)}{\sum_x P(x,t)^2} \right)^2 \rightarrow \min_{a(t_{\mathrm{begin}}),b(t_{\mathrm{begin}}),\ldots,a(t_{\mathrm{end}}),b(t_{\mathrm{end}})} ,$$

where $M(x,t)$ represents net migration.

Next, deaths for the total population (by age, birth cohort, and calendar year) are derived by summing deaths for civilian and military populations. The final step involves calculating all the estimates needed to produce complete life tables for the total population. As we did with pre-war census counts, we use standard HMD methods to split post-war census counts into birth cohorts. Then, using the pre- and post-war census counts and deaths for the total population, we apply the intercensal survival method to derive annual (January 1st) population estimates. The method is the same as that described in the HMD Methods Protocol (Wilmoth et al., 2005) except that we assume no migration (except mobilization and demobilization of military troops) during the war time. Finally, we estimate exposure-to-risk, death rates, and complete period life tables using standard HMD methods.

REFERENCES

Glei, D.A., S. Bruzzone, G. Caselli. (2005). Effects of war losses on mortality estimates for Italy: a first attempt. *Demographic Research*. 13(15):363-388. Available at: http://www.demographic-research.org/Volumes/Vol13/15/.

Human Mortality Database (HMD). 2007. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), Available at www.mortality.org. (accessed 17 April 2007).

Jdanov, D., Andreev, E., Jasilionis, D., Shkolnikov, V. (2005). Estimates of mortality and population changes in England and Wales over the two World Wars. *Demographic Research*. 13(16):389-414. Available at: http://www.demographic-research.org/Volumes/Vol13/16/13-16.pdf.

Vallin, J. 1973. *La mortalité par génération en France, depuis 1899; Travaux et Documents, Cahier no. 63.* Paris: Presses University de France. [In French]

Wilmoth, J. R., Andreev, K., Jdanov, D., and Glei, D.A. with the assistance of C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, and P. Vachon. 2005. "Methods Protocol for the Human Mortality Database.", Retrieved February 20, 2007 (http://www.mortality.org/Public/Docs/MethodsProtocol.pdf).

Winter, J.M. (1977). *Britain's "Lost Generation" of the First World War.* Population Studies, Vol.31, No.3, pp.449-466

Winter, J.M. (2003). *The Great War and the British People.* Palgrave Macmillan, 360 p.