Supplementary Information for

Estimates from 31 countries show the significant impact of COVID-19 excess mortality on the incidence of family bereavement

Mallika Snyder, Diego Alburez-Gutierrez, Iván Williams, and Emilio Zagheni

Mallika Snyder.
E-mail: mallikasnyder@berkeley.edu

This PDF file includes:

Supplementary text
SI References
Supporting Information Text

Methods.

Adjusting Simulation Inputs for Excess Mortality. In order to obtain simulation inputs, we first convert UNWPP 5-year mortality rates into monthly values assuming uniform mortality across the interval, except for the period from March 2020-June 2021. For this period, we employ weekly age and sex-disaggregated mortality data from the Human Mortality Database’s Short-term Mortality Fluctuations dataset (STMF), to adjust the monthly rates in line with two possible mortality scenarios: a COVID-19 case using rates from the 2020-2021 period, and a counterfactual case using an average of the 2016-2019 monthly rates.

Since SOCSIM employs monthly vital rates, we first convert HMD STMF weekly data to monthly data by dividing each week’s total by 7 and assigning each day thus generated to a month according to the ISO8601 standard, which is used by almost all of the countries in the STMF. We use this method to generate monthly exposures (from the STMF data) and death totals for 2016-2021. An average of 2016-2019 values forms our baseline of “expected” 2020-2021 mortality, and we use this to also calculate an annual average expected mortality rate by averaging across all monthly mortality rates. We then use this as the denominator of the ratio used in calculating the adjustment factors used in the two scenarios: in the COVID-19 case, the numerator is 2020-2021 monthly mortality rates, and in the counterfactual case, the numerator is the baseline 2016-2019 rate.

Since the original rates to be adjusted are annual averages which do not account for seasonality, this approach allows for us to consider by how much mortality would have been expected to change in the months from March 2020-June 2021 in the absence of COVID-19 excess mortality, and by how much it increased above annual expected values during the period as a result of COVID-19 excess mortality.

Additional Details on Fertility, Marriage, and Kinship in Our SOCSIM Microsimulations. Our approach follows closely that used by Alburez-Gutierrez et al. (2021) (1), which adopts a SOCSIM setup without marriage rates in which births are based on mortality data from England and Wales, Scotland, and Northern Ireland, and the United States of America.

Sample of Countries Considered. After running the simulations, as a robustness check we use simulation outputs to calculate excess mortality for the period by age group and compare this to equivalent estimates based on the STMF. In line with our input mortality adjustment factors, we calculate excess mortality as the average ratio across simulations of COVID-19 to counterfactual 65+- age-specific mortality rates for the March 2020-June 2021 period. There is a high correlation between input and output (simulated) excess mortality rates of .98 across the 31 countries considered, and there is at most a 3 percent difference between input and output estimates, despite differences in simulation sizes (simulated country populations range from around 12,000 to 120,000 in the middle of the period considered, with simulation size proportional, with some adjustments, to country population size) and in the exposures used in calculating both types of rates. This suggests that our simulations suitably account for excess mortality related to the pandemic in these countries. The sample of countries for which we provide estimates in the replication files are: Austria, Belgium, Bulgaria, Chile, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Israel, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Republic of Korea, Slovakia, Slovenia, Spain, Sweden, Switzerland, the United Kingdom (obtained by combining mortality data from England and Wales; Scotland, and Northern Ireland), and the United States of America.
Investigating the pre-1950 Demographic Stability Assumption. In this section, we consider the implications of assuming demographic stability before 1950. In particular, we examine the degree to which extrapolating the particular demographic characteristics of the 1950-1954 period into the past (including the baby boom) can bias the initial kinship structures for the countries we study.

One way to consider the impact of assuming demographic stability prior to 1950 on living kin is to think in terms of a particular kin relation, especially in terms of ancestors who experienced vital rates prior to the ego’s birth. In the case of mothers, the probability that an ego aged a in year y, from cohort c = y − a, had a living mother $M_1^y(a)$, can be approximated based on (5) with:

$$M_1^y(a) \approx p^{[c-k_c]}(c, a)$$  \[1\]

where $p^{[c-k_c]}(c, a)$ is the mother’s probability of surviving ages from age $y_a$ (the mean age of childbearing when ego was born), given cohort mortality risk for a person born in $[c-k_c]$.

Let us consider an ego born before 1950 who survives to subsequent years. This ego could be aged $a = 30$ in $y = 1960$, $c = 1930$, living 20 years prior to 1950 and 10 after. Formula 1 can be factorized with the year 1950 as pivot:

$$M_1^{1960}(30) \approx p^{(1930-k_{1930})}(k_{1930}, 20) \ast p^{1930-k_{1930}}(k_{1930} + 20, 10)$$  \[2\]

If rates (mortality and fertility) are assumed constant before 1950 (and female population structure by age too, which is relevant to the mean k), we should have a slightly different probability $M_1^{1960+}(30)$, given that all the rates before 1950 now are related to that last observed year:

$$M_1^{1960+}(30) \approx p^{1950}(k_{1950}, 20) \ast p^{1930-k_{1950}}(k_{1950} + 20, 10)$$  \[3\]

We can identify two main effects that drive the difference between $M_1^{1960+}(30)$ and $M_1^{1960}(30)$, given a direction of change for each one. Let us assume constant aging and a mortality decline. Before 1950 mothers would experience 1950 period rates instead of cohort ones, increasing mother’s survival during ego’s life. The second main effect goes in the opposite direction, given that the relation of maternal ages is $k_{1930} < k_{1950}$, survival chances decrease because of age (risk period starting in age $k_{1950}$ instead of $k_{1930}$) and a cohort effect (mortality schedule from $[1930 - k_{1950}]$ cohort instead $[1930 - k_{1930}]$). The net result will depend on the curvature of $L_y$, its change with time, and the slope of $k$ with time, not directly estimable without observed data or a parametric assumption on components. A similar approach can be made for older ancestors. For direct descendants both factors would also include cohort fertility rates.

This last case applies when ego is born before 1950 and is alive after. If she is born in 1950 or later, $y_t \geq 1950$, then this effect on living mothers would disappear. It also disappears for grandmothers if ego is born on $y_t \geq 1950 + k_{1950}$, and for great-grandmothers if ego is born in $y_t \geq 1950 + k_{1950} + k_{1950+1}$, and so on. Of course, if ego is born after 1950, all living descendants would not be affected by the change. In general terms, the bias in constant rates before 1950 in living ancestors older than mothers (when $i > 1$), will disappear for $M_i$ if ego is born at $y_{t+i} \geq y_t - k_{y_t-k_{1950}}$. In a stable population context, when there is a sudden change in rates (here in 1950), changing the regime itself, this series could be useful to ensure that ancestors until $i$ degree belong to the new and latest regime.

We carried out separate analyses to understand the effect of assuming demographic stability before 1950 for a specific population. For a numeric example we applied kin estimation for an ego aged 30 (an age where it is not unlikely to have a surviving grandmother in low mortality countries, and also a living child (1)) using vital rates from Sweden and locating ego in successive years from 1950 to 2015.

We ran two sets of analyses: ‘historical’ analyses used the estimated historical rates for the 1751-1950 period and ‘stable’ analyses assumed that the 1950 demographic rates applied throughout the 1751-1950 period. Age-specific mortality rates (starting from 1751) come from the Human Mortality Database and fertility rates (starting from 1891) come from the Human Fertility Database. In all cases, we assumed demographic stability for fertility in the 1751-1891 period. We combined these data with official demographic projections of fertility and mortality provided by Statistics Sweden. Sweden is the ideal case for this analysis given its long time series of high quality demographic rates. Crucially, Sweden also experienced a moderate increase in fertility after WWII (6). As such, the Swedish case would allow us to explore whether our stability assumption extrapolates the conditions of the post-WWII baby-boom into the past.

The results suggest that there is a general convergence in all the relatives types included in the study. The fact that older ancestors had later converged is consistent with the mentioned series $y_t$. In general terms the differences are null for all type of kin in year 2015 (last year of evaluation), when in 1950 we could see bigger differences, like 50% in cousins 9% in older sisters, -50% in grand-mothers and -80% in grand-daughters (\(\frac{\text{Historical}}{\text{Stable}} - 1\) * 100). In conclusion, at least in the Swedish case, no significant bias is expected in kin counts after 2015 because of assuming fixed rates before 1950.

Validating our Simulations with Formal Methods. To ensure that the simulated populations capture the kinship structure implied by the input demographic rates, we estimate a set of independent matrix kinship models with the same input data. We use an implementation of matrix kinship models (7) extended for populations with time-variant demographic rates, as implemented in the DemoKin R package (https://github.com/IvanWilli/DemoKin), to show that the estimates are consistent with the simulated kinship structures.

*The assumption of a concentrated fertility schedule around $k$ is necessary.
We report a comparison of estimates for living kin for a women in Sweden in July 2020 (the first wave of excess mortality), based on both microsimulation and formal methods. While the kin relations presented in the main text are for male and female kin, here we present only female kin in a single-sex population model so as to achieve comparability with the formal approach we present. The results suggest that there is considerable similarity between methods in terms of the results obtained, providing an important robustness check for our microsimulation-based results.

Kin counts are generally similar across the two methods. There are no significant differences for mothers, except for individuals aged 65+, for whom SOCSIM estimates .07 living mothers and the GKP approach estimates .02 mothers in both the COVID-19 and counterfactual scenarios. In general, for a given method, population-level kin counts are fairly similar across the two scenarios, reflecting the relative brevity of the mortality shock considered as well as the large number of individuals included in the simulations. Differences in the number of living grandmothers are all under 0.15, with the largest discrepancies in the 30-44 age interval, followed by the 15-29 age interval. This is likely associated with the increased rarity of having a living grandmother at older ages, leading to more statistical variability in the SOCSIM estimates (although even older individuals might be so unlikely to have a living grandmother as to make estimates of kin availability and associated differences relatively small). The expected number of daughters is higher in GKP-based estimates for most age groups by about 0.1 to 0.18. Consistent with our expectations, we tend to find the largest discrepancies for the relative types that are less common overall. Our simulations have a limited size (for example, across simulation runs, there were around 6,300 women aged 30-44 in Sweden who survived the COVID-19 mortality period from March 2020 to June 2021) and, as a result, there is bound to be more uncertainty around the prevalence of relatively rare kin ties.

**Comparison to Other Published Estimates.** As an additional robustness check, we compared our results to those from Hillis et al. (2021)’s modeled estimates of parental bereavement for children in which Hillis et al. (2021)’s estimates were calculated using excess mortality, and from Kidman et al. (2021)’s estimates for the United States. In general, our estimates are slightly lower, reflecting the lower age group endpoint (14 versus 17); most parentally bereaved children would likely be in the older portion of the age group. Hillis et al.’s estimate of 8,495 parentally bereaved children aged 0-17 in England and Wales alone by the end of April 2021 resembles our estimate of 7,567 children aged 0-14 for the United Kingdom as a whole for the same period, despite differences in the groups considered (8). Similarly, we estimate 2,298 orphans aged 0-14 in France, compared to the Hillis et al. (2021) estimate of 4,064 orphans aged 0-17. Some countries show greater variability: our results for Spain are higher than those estimated by Hillis et al. (2021) (3,905 orphans aged 0-14 versus 2,309 aged 0-17), likely reflecting differences in modeling strategies used. Existing published estimates for the United States vary considerably: Kidman et al. (2021) estimate 43,000 orphans aged 0-17 up to the end of January 2021 using excess mortality (9), while Hillis et al. (2021) estimate 104,884 for the same group up to the end of April 2021. Our results, when calculated up to the end of January 2021 (to match the time period Kidman et al. (2021) considered) are close to those of Kidman et al. (2021): We estimate a value of 36,716, which, given the shorter age interval of 0-14, is slightly lower, as we expected.

**References**

1. D Alburez-Gutierrez, C Mason, E Zagheni, The “Sandwich Generation” Revisited: Global Demographic Drivers of Care Time Demands. *Popul. Dev. Rev.* 47, 997–1023 (2021).
2. C Mason, Socsim Oversimplified. *University of California Demography Lab* (2016).
3. S Ruggles, Confessions of a Microsimulator. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 26, 161–169 (1993).
4. WD Manning, SL Brown, JB Stykes, Family Complexity among Children in the United States. *The ANNALS Am. Acad. Polit. Sci. Soc.* 654, 48–65 (2014).
5. N Keylitz, H Caswell, *Applied Mathematical Demography.* (Springer, New York), (2005).
6. J Van Bavel, DS Reher, The Baby Boom and Its Causes: What We Know and What We Need to Know. *Popul. Dev. Rev.* 39, 257–288 (2013).
7. H Caswell, The formal demography of kinship: A matrix formulation. *Demogr. Res.* 41, 679–712 (2019).
8. SD Hillis, et al., Global minimum estimates of children affected by COVID-19-associated orphanhood and deaths of caregivers: a modelling study. *The Lancet* 398, 391–402 (2021).
9. R Kidman, R Margolis, E Smith-Greenaway, AM Verdery, Estimates and Projections of COVID-19 and Parental Death in the US. *JAMA Pediatr.* 175, 745–746 (2021).