

Assessing the agreement between microsimulated and register-based kin counts in Sweden

Liliana P. Calderón-Bernal, Diego Alburez-Gutierrez, Martin Kolk & Emilio Zagheni

To cite this article: Liliana P. Calderón-Bernal, Diego Alburez-Gutierrez, Martin Kolk & Emilio Zagheni (30 Jul 2026): Assessing the agreement between microsimulated and register-based kin counts in Sweden, *Population Studies*, DOI: [10.1080/00324728.2026.2706778](https://doi.org/10.1080/00324728.2026.2706778)

To link to this article: <https://doi.org/10.1080/00324728.2026.2706778>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 30 Jul 2026.



[Submit your article to this journal](#)



Article views: 157



[View related articles](#)



[View Crossmark data](#)

Assessing the agreement between microsimulated and register-based kin counts in Sweden

Liliana P. Calderón-Bernal ^{1,2}, Diego Alburez-Gutierrez ^{1,3},
Martin Kolk ^{2,4,5} and Emilio Zagheni ¹

¹Max Planck Institute for Demographic Research, ²Stockholm University, ³Humboldt-Universität zu Berlin,
⁴Institute for Futures Studies, ⁵Åbo Akademi University

Estimating kinship networks is a data-intensive undertaking, usually conducted using empirical sources or demographic models. Although empirical data, such as population registers, provide observed individual-level information, they are limited by availability, truncation, and survivorship bias. Demographic models, including microsimulation, require less detailed data, but they often minimally account for population heterogeneity, family similarity, and multi-partner fertility. This study examines the agreement of kin counts (from grandparents to grandchildren) between SOCSIM microsimulation outputs and Swedish population registers for cohorts born between 1915 and 2017. The results show that microsimulation closely approximates mean kin numbers and reasonably reflects parity distributions. Microsimulation underestimates kin for recent cohorts not affected by truncation but provides a more accurate representation of kin for earlier cohorts with missing parent–child links. These findings highlight the utility of microsimulation for reconstructing kinship when only population-level rates are available, supporting its application in kinship analyses, both historical and projected.

Supplementary material for this article is available at: <https://doi.org/10.1080/00324728.2026.2706778>

Keywords: kinship; microsimulation; population registers; Sweden; demographic methods; family

[Submitted July 2025; Final version accepted March 2026]

Introduction

Kinship networks are a fundamental aspect of human life, surrounding individuals from birth to death. These networks extend beyond the nuclear family to include relatives linked by biological, legal, or normative ties. Despite their importance across the life course, as a source of support and intergenerational exchange, our knowledge of kinship networks remains limited, particularly from an international and comparative perspective. Few countries worldwide have high-quality longitudinal demographic data with the information necessary to study kinship, making the consideration of alternative methods for examining other contexts essential where such detailed data are not available.

In this study, we use the Swedish population registers as an empirical benchmark to assess the accuracy of kinship estimates derived from a relatively simple microsimulation model, which uses population-level demographic rates as inputs. By comparing these two approaches, we examine the potential and limitations of using demographic microsimulation to count kin. The term ‘kin counts’ refers to the mean number and the distribution of a given type of kin for individuals at each age within a birth cohort. While this analysis is based on the unique Swedish data context, the goal is to enhance our understanding of the accuracy of synthetic kinship networks, which could be used to study kinship in other settings where only demographic rates are available. This assessment contributes to the long-standing debate

on the usefulness of demographic microsimulation to study kinship, a subject that was first discussed in the 1990s (Ruggles 1993; Hammel and Wachter 1996; Wachter et al. 1997, 1998; Wolf 2001).

Given a population register of sufficient quality and longitudinal depth, it is possible to reconstruct the entire kinship network of a given individual (hereafter referred to as ‘ego’) based on data on their parents and children. Starting from ego’s parents and children, the researcher can move up and down generations to identify: ego’s parents’ parents (i.e. ego’s grandparents); ego’s parents’ children (i.e. ego’s siblings); ego’s children’s children (i.e. ego’s grandchildren); ego’s grandparents’ children (i.e. ego’s aunts and uncles); ego’s siblings’ children (i.e. ego’s nieces and nephews); ego’s aunts’ and uncles’ children (i.e. ego’s cousins); and so on. Kinship networks can be extended by including relatives through marriage. However, these reconstructions can be more challenging in practice due to limitations in the available data needed to link parents and children across multiple generations.

There are two common approaches to estimating kinship networks within a population. One approach is to derive them empirically using sources such as administrative registers, surveys, censuses, or ethnographic methods. The other approach uses modelling techniques, such as formal demographic methods (Caswell 2019, 2020, 2022, 2024; Caswell and Song 2021; Caswell et al. 2023; Caswell and de Vries 2025) or demographic microsimulation (Ruggles 1993; Zhao 2006; Zagheni 2015). Each approach has its advantages and limitations, which can result in differences in kinship estimates.

While empirical kinship counts offer detailed individual-level information, they are limited by factors such as data availability, coverage, truncation, and survivorship bias. For instance, kinship networks can be reconstructed from population registers using parent–child links (Kolk et al. 2023; De Bel et al. 2025). Surveys such as the Netherlands Kinship Panel Study (Dykstra et al. 2006) and the United States (US) Panel Study of Income Dynamics (Daw et al. 2016) can be used to estimate kinship, but missing links can result in incomplete networks. Censuses can provide insight into family relationships within households, both historically and in the present day. Ethnographic methods can also offer valuable qualitative insights into kinship within smaller groups or co-resident families. However, they do not allow for comprehensive population-level estimates. Conversely, analytical or microsimulation models can be run using population-level data but may not capture the full

complexity of real-world population dynamics. In such cases, the average number of kin, and sometimes kin distributions, can be estimated from mathematical models or by tracing kinship relationships for the synthetic individuals produced by microsimulation using a similar parent–child linking procedure to that already described for reconstructing kinship networks using registers.

As high-quality empirical data at the individual level are rare, demographic models of kinship using population-level data are a good alternative for estimating kinship networks. Specifically, demographic microsimulation is a computationally intensive, individual-based tool used to model individual demographic events that are consistent with macro-demographic rates, to gain insights into life-course transitions (Zagheni 2015). It allows the reconstruction of the demographic history and kinship networks that would have been observed at any given point in time during the simulation (Murphy 2010a). Among the most widely used microsimulation programmes, SOCSIM requires age-specific fertility rates and probabilities of death, whereas CAMSIM uses parity progression ratios, probabilities of death, mean ages at first marriage, and proportions never married. Therefore, the demographic events experienced by the synthetic individuals (and their timing) are determined by the input parameters and the assumptions underlying the microsimulation. However, the two simulators take different approaches. SOCSIM takes a population perspective and considers changes in demographic rates over time, whereas CAMSIM takes an ‘ego-centric’ perspective and is based on the assumption of demographic stability (Zhao 2006). SOCSIM’s population- and time-driven approach makes it ideal for estimating kinship networks over time, modelling groups of individuals (e.g. parents who have lost a sibling), and comparing these networks with empirical data over multiple generations.

Demographic microsimulation has been in use for several decades. However, since the 1990s there has been some debate about its ability to replicate the underlying dynamics of real populations and produce accurate kinship estimates. Notably, in ‘Confessions of a microsimulator’, Ruggles (1993) identified some issues with kinship microsimulation, particularly when correlations in demographic behaviour between family members were ignored. Subsequent external evaluations compared the accuracy of SOCSIM microsimulations with historical census data (Hammel and Wachter 1996) and survey results (Wachter et al. 1997, 1998), showing that the microsimulation outputs partially aligned with

the empirical data. However, these studies also identified factors that may limit the ability of microsimulations to describe the complexity of real populations fully. These analyses highlighted that discrepancies between microsimulated and empirical data were to be expected when simplifying assumptions were made. At the same time, such differences could have been taken as supporting Ruggles' concerns about the limitations of microsimulations, thereby hindering their wider implementation. Furthermore, Wolf (2001) cautioned that attempting to account for all sources of variability in the sampling process and for uncertainty in the parameters could yield meaningless results and that simple Monte Carlo replication under fixed inputs offered limited reassurance. Our study builds on this literature and contributes to the debate in two ways. First, it assesses the extent to which microsimulations can reasonably approximate observed empirical data when operating under simplifying assumptions. Second, it shows that certain limitations of register data for estimating kin can be addressed using 'fully tracked' microsimulations. The key potential sources of discrepancies between approaches are outlined next.

Some features of the data or assumptions underlying microsimulations may lead to differences between synthetic kinship networks and those based on empirical data. First, since microsimulations are the outcomes of the input data (and parameters), the only uncertainty in simulated kinship networks should arise from their stochasticity (Zagheni 2015). The demographic events experienced by the simulated individuals are randomly assigned and executed, often using a competing risks model, according to predetermined rates or probabilities (Ruggles 1993; Zhao 2006; Zagheni 2015). However, demographic data, typically available as population-level rates, may not fully reflect the heterogeneity of real populations. For instance, data by marital status or parity, which would allow for a more accurate kin distribution, are usually unavailable. Such a lack of data disaggregation may result in rather homogeneous synthetic populations.

Second, the resemblance in demographic behaviour between members of the same kinship group is often disregarded in microsimulation models. According to Ruggles (1993), ignoring these correlations can bias the mean number of kin downwards and result in a more homogeneous distribution than would be the case in real populations. In some simulators, this could be partially accounted for by adjusting some parameters (e.g. the heterogeneity and heritability of fertility) or by defining groups based on demographic resemblance within families.

However, this would require more detailed data, including family links that are rarely available and/or significant calibration of the input rates.

Third, partnership and fertility models must rely on simplifications of real-world dynamics. Microsimulation models can be closed (where partners are selected from simulated living individuals) or open (where partners with suitable characteristics are created when needed for marriage) (Zhao 2006; Zagheni 2015). Simulating childbearing within marriages or partnerships according to women's fertility schedules may not introduce bias into estimates for periods when marital fertility is the norm. However, when childbearing does not necessarily occur within marriage, accurately simulating fertility behaviour requires data disaggregated by age, parity, and marital status, and these are rarely available. Furthermore, to simulate realistic kinship networks, both the maternal and paternal lines would need to be captured. This would require a two-sex fertility and partnership model, demanding a substantial quantity of data and modifications to the common configuration of the simulator. When such detailed data are unavailable, microsimulation results may underestimate multi-partner and non-marital fertility and may produce less accurate estimates of men's fertility. These limitations may be most apparent when studying family relationships that arise from a separation followed by a new union.

In addition, certain characteristics of the empirical data may result in discrepancies with the microsimulated kinship networks. For instance, empirical kinship counts are often affected by left-truncation, whereby information is missing for individuals born close to the start of the registration system or data collection period. In particular, truncation of parent-child links in most administrative registers can limit the depth of complete kinship reconstruction, unlike in microsimulations, which track the entire population. Moreover, population registers often suffer from survivorship bias as they tend to count only individuals who have survived to a given date.

Administrative sources count the current population of a country, implying that both ancestral populations and recent arrivals are included. However, only long-settled populations have the deep kinship networks required for them to be counted using national data. The extended kinship networks of more recent arrivals cannot be fully traced using data from their destination country alone. Therefore, empirical kinship counts in national administrative data are limited to ancestral populations whose ancestors registered vital events. International migration can also truncate

administrative sources if vital events occur outside the country's borders and are thereby not recorded. Empirical kinship counts are therefore likely to differ from estimates derived from methods using population-level demographic rates as inputs (i.e. microsimulations). This discrepancy arises from the fact that population-level rates are calculated based on the total population, including, for example, births and deaths among short-term migrants. However, individuals with observable descendants within the country represent only a subset of the national population, as kinship ties within a country can be truncated by international migration.

Given the potential discrepancies between kinship networks derived from microsimulated and those derived from empirical data, it is scientifically valuable to compare the agreement between the two approaches systematically. Microsimulation, despite its necessary simplifications and assumptions, is a valuable approach for estimating kinship relationships when empirical individual-level data are unavailable. Using observed demographic rates, it produces fully tracked synthetic populations whose vital events and kinship networks can be traced over time, without the common limitations of empirical data (e.g. coverage gaps, truncation, and survivorship bias). Therefore, comparing kinship estimates derived from microsimulation with those derived from empirical sources can help us to understand where the two approaches align or differ and how model assumptions in the microsimulation or constraints in the empirical data contribute to any discrepancies. To our knowledge, few studies have directly compared microsimulated and empirical kin counts. Notably, Hammel and Wachter (1996) used a Slavonian historical census, and Wachter et al. (1997, 1998) used US national survey data. Other studies have contrasted microsimulated population structures with census enumerations, examining age distributions in Chinese censuses (Hammel et al. 1991) and the age and marital status distributions in British censuses, or compared the proportions of the population with different types of kin using a survey (Murphy 2010a, 2010b, 2011). However, no such comparisons have yet been conducted using a comprehensive kinship enumeration based on empirical data for an entire population. Our research contributes to filling this gap by providing a systematic assessment of agreement between microsimulated and register-based kin counts.

Among the empirical sources for counting kin, Sweden's long historical records of high-quality administrative population data provide an ideal basis for comparison. Here, we compare kin counts

derived from the SOCSIM demographic microsimulation programme (Hammel et al. 1976) using fertility and mortality data for Sweden with empirical kin counts estimated from the Swedish population registers, as provided in the 'Swedish kinship universe' by Kolk et al. (2023; see their Online Appendix 2).

With this systematic comparison, we aim to shed light on the accuracy of synthetic kinship networks derived only from fertility and mortality data. This will offer valuable insights for researchers seeking to use microsimulation to study kinship in contexts where more detailed demographic data are unavailable.

Data and methods

In this paper we assess the agreement between microsimulated and register-based kin counts by comparing the empirical estimates of kinship for a large number of kinship types obtained in a previous register-based study by Kolk et al. (2023) with the estimates of similar kin quantities obtained from a simulated population of Swedish individuals produced using the SOCSIM microsimulation programme. We started the analysis by running a demographic microsimulation of the Swedish population from 1751 to 2017 to obtain a complete register of all individuals ever alive during the simulated period, including their vital events and kinship relationships.

SOCSIM is an open-source microsimulation programme—originally developed at the University of California, Berkeley (Hammel et al. 1976) and written in C programming language—that has been used for decades in demographic research to study, among other things, kinship network availability (Murphy 2010a, 2010b, 2011; Verdery and Margolis 2017; Margolis and Verdery 2019; Alburez-Gutierrez et al. 2021) and loss (Zagheni 2011). The microsimulator requires as input an initial population file with individual information on date of birth and sex, and also monthly age-specific fertility rates and age-specific probabilities of death that apply over the simulated period to individuals of a given sex, 'group', and marital status. (SOCSIM offers the option to create different 'groups' within the simulated population and to define different rates for their members based on characteristics such as ethnicity, location, health status, wealth, and education. However, defining groups is a sensible option only when highly disaggregated data are available; otherwise, individuals are considered to belong to a single group.) If available, parity-specific fertility can be included using conditional age-specific rates by

birth order. If rates are available only by sex and age, the same set of sex- and age-specific demographic rates can be applied, regardless of marital status, group, or parity, using SOCSIM's default rules. In such cases, individuals are not categorized by group or marital status, and parity does not alter the stochastic exposure to sex- and age-specific rates. During the simulation, SOCSIM schedules and executes demographic events (births, marriages, and deaths) for each synthetic individual in the initial population and their descendants.

The operation of the microsimulator is described in Mason (2016) and summarized here. At the beginning of each simulation segment (i.e. when the demographic rates or societal constants change) or month, SOCSIM schedules an event for every synthetic individual to be executed in the future. Only one event can be scheduled for each individual at any one time. After the execution of an individual's event (except in the case of death) or a change in marital status or parity, a new event is scheduled for that individual. The next event to be scheduled is determined by creating random waiting times depending on the individual's sex-, age-, group-, and marital-status-specific rates. Once all potential events have randomly generated waiting times, the event with the shortest waiting time is selected and scheduled. Hence, the event competition follows a competing risks framework, wherein the probability of experiencing each event for which the individual of a given sex, age, group, and marital status is at risk is independent of all others. All scheduled events are sequentially executed in a random order. Then, SOCSIM increments the month and repeats the event execution. At the end of the simulation, SOCSIM writes an output population file with information about every synthetic individual who has ever lived and a marriage file containing information about each simulated marriage. The output population file contains one row for each individual, including their personal identification number (pid), sex, and dates of birth and death, along with the pids of some of their close kin and some other variables. Using the pids of the individuals and their mothers and fathers, it is possible to reconstruct each individual's kinship network, following the same approach as that used for the registers.

We ran one large simulation using the `rsocsim` package in R (Theile et al. 2023), with an initial population of 50,000 synthetic individuals and using age-specific rates obtained from the Human Fertility Collection (1751–1890), the Human Fertility Database (1891–2017), and the Human Mortality Database (1751–2017). The last two were retrieved via

the `HMDHFDplus` package in R (Riffe 2015). Although microsimulations involve stochasticity, fluctuations owing to stochastic variation diminish as the simulated population size increases, resulting in stable, near-deterministic outcomes. Conversely, smaller simulations show greater variability in mean numbers of kin and their distributions. Therefore, we relied on a single large-scale simulation to reduce such noise and yield more reliable estimates of kinship networks for comparison. It is worth clarifying that some earlier studies have used microsimulation to analyse smaller populations (e.g. Hammel et al. 1979), thereby demonstrating its value in such settings as well.

Due to data availability, we used age-specific fertility rates for all birth orders combined up to 1969 and conditional age-specific fertility rates by birth order from 1970 to 2017. Using parity-specific rates improved the accuracy of kinship distributions for the more recent cohorts compared with a simulation using combined rates over the entire period (results not shown). However, as parity-specific data are available only from 1970, this enhancement only partially affects the fertility of earlier cohorts.

Otherwise, our microsimulation set-up did not incorporate additional parameters for demographic heterogeneity or heritability, since such detailed information is lacking in the input data and alternative long-term data sources. By 'heterogeneity' we mean variations in the risks of experiencing an event, based on demographic characteristics. However, in our model, fertility and mortality rates are assumed to be the same for people with the same demographic characteristics. Although SOCSIM offers a 'heterogeneous fertility' option that models heterogeneity and maternal-line heritability, the default parameters of this option underestimate fertility, particularly during periods of high fertility, and would require significant rate calibration. Furthermore, while this option can mitigate the underestimation of certain collateral kin types (e.g. siblings, aunts, and nieces) to some extent, it also increases the occurrence of individuals with no kin and those with very large kin sets. This results in over-dispersion, which also requires calibration. Therefore, to ensure broader applicability, particularly in contexts lacking data for calibration or external comparison, we opted for a simplified simulation set-up without the 'heterogeneous fertility' option.

In line with earlier research using SOCSIM (see Hammel 2005), we initially ran the simulator for 100 years using the first available set of age-specific rates (for 1751) to produce a stable age structure. After this century of 'demographic stability', the

population tripled to 150,000 individuals alive at the beginning of 1751, who were then exposed to the corresponding annual rates for the period 1751–2017. At the end of the simulation (in 2017), the synthetic population reached about 1 million living individuals.

In the absence of accurate age-specific marriage rates by sex for the entire period, we used the ‘marriage after childbirth’ directive in `rsocsim` to establish lifelong partnerships and select living single partners whenever a previously single woman gave birth. Despite its name, this option corresponds to lifelong partnerships, rather than real marriages in the sociological sense. Following Alburez-Gutierrez et al. (2021), partners for each woman were selected from all living single men to minimize the squared difference between the observed distribution of ‘man’s age minus woman’s age’ and a normal distribution with a mean of 2.0 and a standard deviation of 3.0. In our model this lifelong partnership is formed when a previously single woman has her first child. Candidates are restricted to single men, who are selected based on the desired distribution of the age difference, with no consideration of additional characteristics. The partnership formed at the time of the first birth is used to assign paternity to subsequent births while both partners are alive. In terms of the event schedule and execution, individuals remain independent: women are exposed to fertility rates and both sexes to mortality rates. There is no joint couple behaviour. If a partner dies, a new partner can be assigned for subsequent births. Otherwise, the same partner is assigned as the father. Further information on this directive can be found in the supplemental materials of Alburez-Gutierrez et al. (2021).

It is worth mentioning that this partnership option is not intended to predict repartnering following a separation or a divorce, so this limits comparisons of multi-partner fertility, half-relatives, and stepfamilies. ‘Child-bearing partners’ refers to the number of distinct partners with whom an individual has at least one birth. In our set-up, this count is almost always ‘0’ or ‘1’, exceeding ‘1’ only after widowhood because it does not model separation or divorce. However, it provides a solution to the lack of empirical data on marriage, separation, divorce, and non-marital fertility over the whole period and produces an accurate estimate of women’s fertility overall. Moreover, using this option, which is less demanding in terms of data, for comparison with a ‘gold-standard’ source allows us to assess the accuracy of kinship networks simulated using only population-level fertility and mortality data. We deliberately chose a simplified microsimulation set-up (without heterogeneous fertility, marriage rates, or further data disaggregation) so that insights about its accuracy could

be applied to other settings where only population-level demographic data are available.

After checking the accuracy of the microsimulation output against the input age-specific demographic rates and summary measures (see Figure S1, supplementary material), we estimated the synthetic kinship networks following the approach outlined in the Swedish kinship universe (Kolk et al. 2023). Using the SOCSIM output, we traced the kinship networks (i.e. children, parents, siblings, grandchildren, grandparents, aunts and uncles, nieces and nephews, and cousins) of the synthetic individuals who were alive at the end of 2017, based on parent–child links, and we replicated the figures in Kolk et al. (2023). To facilitate comparison, we also reproduced the register-based figures using the population-level demographic data provided in Kolk et al. (2023; Online Appendix 2).

The empirical counts in this paper are based on the Swedish Total Population Register and the Swedish Multi-generation Register. These registers contain information on all individuals ever registered in Sweden since 1960, including their country of birth, sex, dates of birth and death, and names of their biological mother and father. The analytical sample comprises all individuals born in Sweden between 1915 and 2017 who were recorded in the registers and were living in Sweden at the end of 2017 ($N = 8,243,185$). For these reference individuals (used as the denominator), the kinship counts (numerator) include all relatives, whether alive or deceased in 2017. We applied the same restrictions to our microsimulation models to provide an account of kinship links in 2017 and to estimate the number of kin of each type in each age group (which corresponds to the cohort, as the period is fixed). To provide an assessment of the two approaches, we reproduced the same figures, used the same definitions of kin, and disaggregated in the same way as in Kolk et al. (2023).

The advantages and limitations of using Swedish register data to enumerate kinship networks are described in detail in Online Appendix 1 of Kolk et al. (2023). However, some of the data features should be acknowledged before comparing microsimulation estimates with those based on empirical data. Modern Swedish registers began with the introduction of a unique personal identity number (*personnummer*) in 1947–48, alongside the collection of data on parentage for all surviving children under the age of 16 living in Sweden. After 1947, data on children’s parents were obtained directly from birth records. The earliest cohort that can be linked to their parents and siblings is those born in 1932. Consequently, grandparents must have given birth

to their children (i.e. parents, aunts, and uncles) after 1932 for links between cousins to be traceable. The annual population registers were digitized only from 1968, marking the starting point for the availability of demographic event records. Another condition for inclusion in the Swedish register data, as discussed in Kolk et al. (2023), is survival until 1960, since the first digitized censuses took place in 1960 and 1965. Thus, despite their high quality, the Swedish registers are biased by left-truncation before 1932 and by survivorship until 1960.

Results

This study compares two sets of results concerning the number of kin in Sweden in 2017: one based on a newly created microsimulation model and the other based on previously published empirical results. The SOCSIM-based figures are generally consistent with the register-based figures from Kolk et al. (2023) in terms of mean numbers and distribution by type of kin. However, there are some differences due to the design and limitations of the microsimulation and/or due to incomplete information on parent–child relationships before 1932 in the registers. To systematically compare the average number and distribution of different types of kin derived from microsimulation and Swedish register data, all the figures are structured into three panels. Panel (a) shows the microsimulation-based estimates, and panel (b) shows the corresponding register-based estimates reproduced using the data from the Online Appendix 2 of Kolk et al. (2023). The figures are organized by birth cohort (equivalent to age measured in 2017) and illustrate the average number or distribution of a given type of kin at different ages for individuals alive in 2017. Panel (c) shows the difference between the two approaches, calculated as the microsimulation estimates minus the register-based estimates, indicating whether the microsimulation overestimates or underestimates the number or proportion of kin. A positive value means that the microsimulation estimates are higher and a negative value that the microsimulation estimates are lower than the register-based estimates. The replication and comparison of the figures from Kolk et al. (2023) not included in this section can be found in our supplementary material.

Average numbers of kin

The microsimulation closely aligns with the registers in estimating direct kinship ties, such as numbers of

living children, parents, and grandparents, with slight differences between cohorts. Figure 1 (panels (a) and (b)) shows that the average total number of children per woman is relatively consistent across sources, although it is lower in the microsimulation, especially for early cohorts. Panel (c) shows that this underestimation may be partly due to differences in the number of childbearing partners. While the microsimulation generally overestimates the number of children for women with only one childbearing partner (except for pre-1930 cohorts), it underestimates the number of children for women with two or three childbearing partners in almost all cohorts. This results from the partnership model used in the simulation, which is based on lifelong partnerships and considers repartnering only after the death of the previous partner. Therefore, as the simulation set-up is not intended to model multi-partner fertility, the distribution of children by childbearing partners may be biased, but the total number of children per woman remains comparable overall.

With regard to men, the number of children per man in the microsimulation is generally lower and shows greater variation between cohorts (see Figure S2, supplementary material). In addition to differences arising from multi-partner fertility, discrepancies between the sexes may be due to the SOCSIM fertility model, which is based on women's fertility schedules and does not permit direct incorporation of men's fertility rates. Therefore, men's fertility depends on their partnerships with women who are scheduled to have children.

For living grandchildren, the estimates are relatively consistent for women born after 1940, but the early cohorts have fewer children and then fewer grandchildren in the simulation (see Figure S3, supplementary material). However, the accuracy of the number of grandchildren per grandfather shows more variation across cohorts. As with fathers, the difference in accuracy between the sexes is probably due to the dependence of the SOCSIM fertility model on women's fertility schedules.

For living parents, Figure 2 suggests that the numbers are consistent overall, although the microsimulation estimates slightly more living fathers for some cohorts (1945–65) and fewer for others (1967–90), probably compensating for the unregistered parents in the register-based estimates. For living grandparents, Figure 3 shows that the mean number is similar up to the 1990 cohort but is higher in the microsimulation for more recent cohorts. The index population of the registers, that

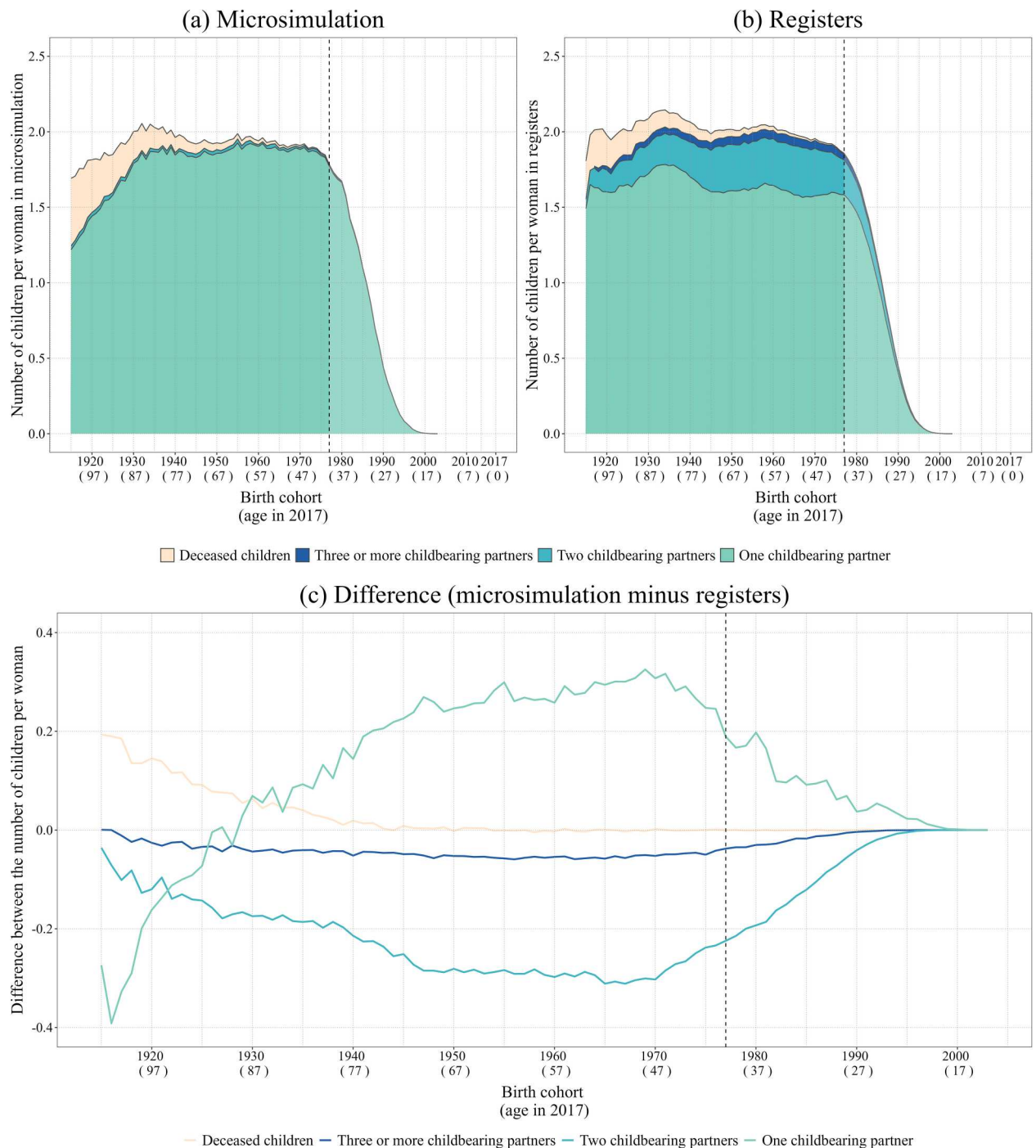


Figure 1 Average number of living and dead children per woman in 2017, by birth cohort and number of childbearing partners: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Notes: The vertical dashed lines distinguish the cohorts with complete and incomplete kin counts in the registers. The areas on the right-hand side represent cohorts for whom we may not be able to observe all their children (as some children will not yet be born in 2017). Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: For panel (a), see Data section. For panel (b), data are from Swedish registers and compiled in Kolk et al. (2023).

is, native-born Swedes, includes sizeable proportions of second-generation migrants whose grandparents are not resident in Sweden. This discrepancy

between a microsimulated population based on national vital rates and empirical kin counts based on actual populations reflects the specific migration

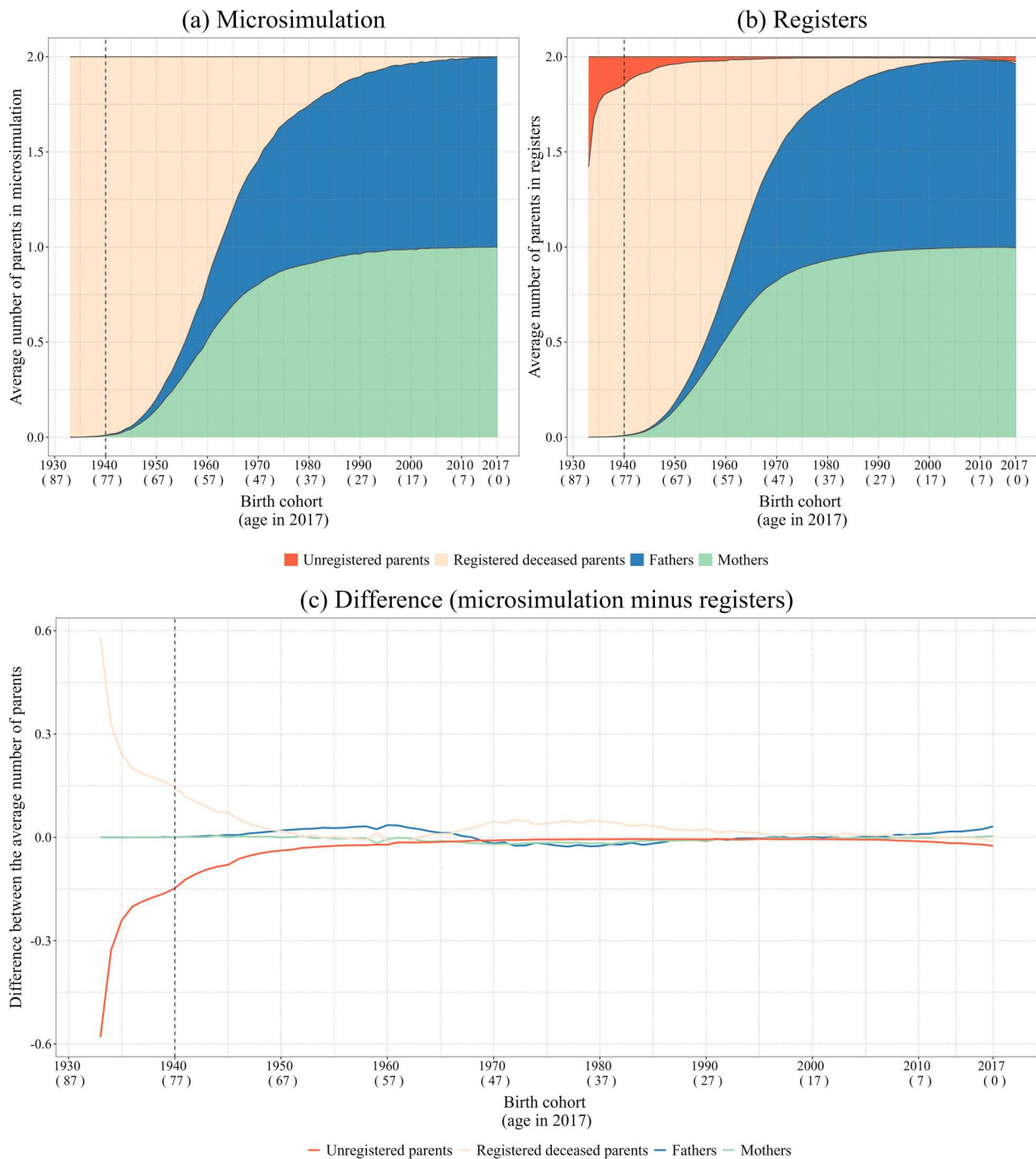


Figure 2 Average number of living, dead, and unregistered parents in 2017, by birth cohort: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Notes: The vertical dashed lines distinguish the cohorts with complete and incomplete kin counts in the registers. The areas on the left-hand side represent cohorts with incomplete coverage of parents due to missing parent–child links in the registers. Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: As for Figure 1.

histories of a given country. Many of the grandparents classed as unregistered are likely to be alive in other parts of the world.

For other types of living kin, such as siblings, nieces, nephews, aunts, uncles, and cousins, the microsimulation generally produces lower

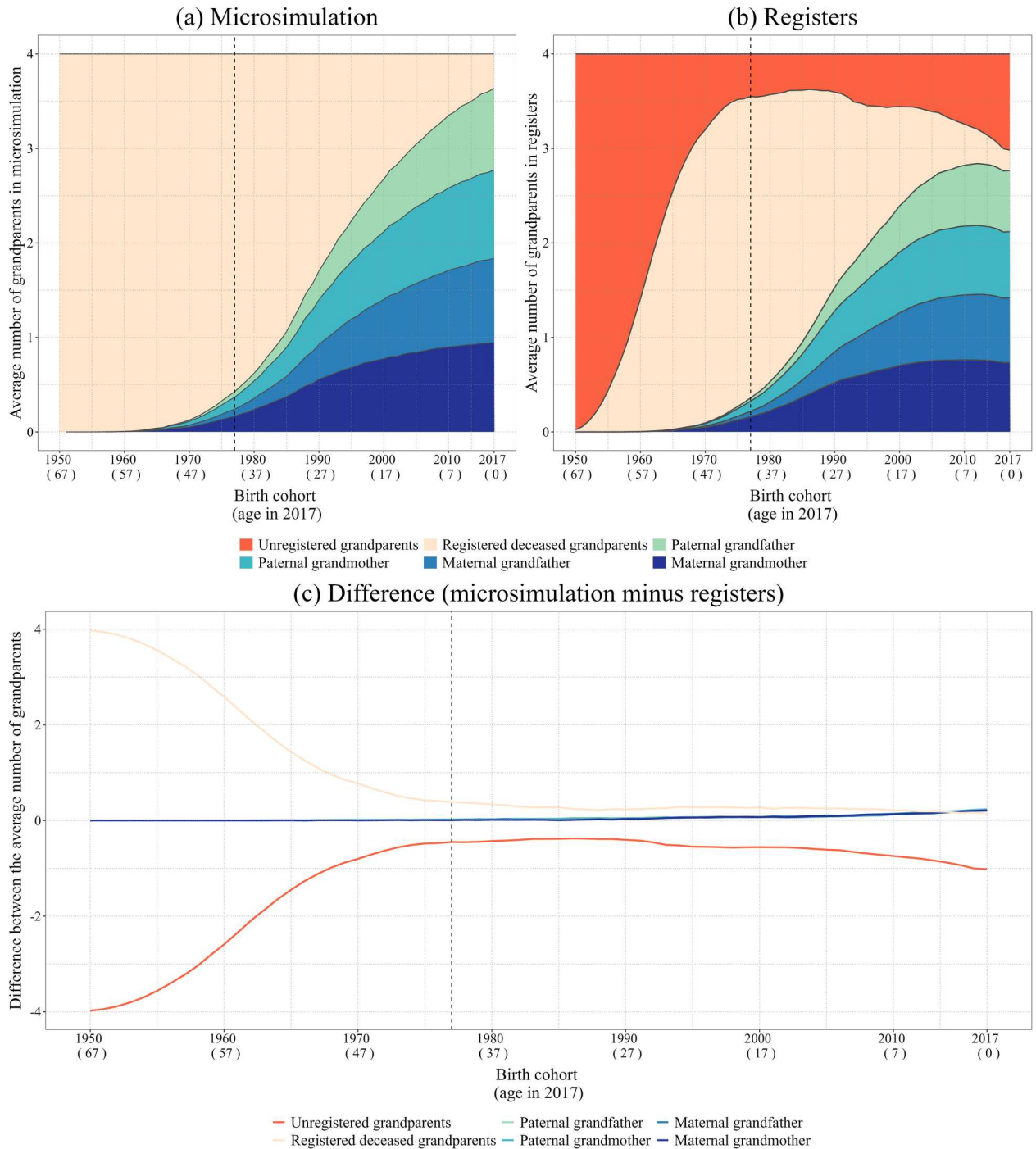


Figure 3 Average number of living, dead, and unregistered grandparents in 2017, by birth cohort and type of grandparent: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Notes: The vertical dashed lines distinguish the cohorts with complete and incomplete kin counts in the registers. The areas on the left-hand side represent the cohorts with incomplete coverage of grandparents due to missing parent–child links. Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: As for Figure 1.

numbers for the cohorts that, according to Kolk et al. (2023), have almost complete information in the registers (siblings for the 1940–2004 cohorts; nieces and nephews for the 1940–71 cohorts;

aunts and uncles for the post-1977 cohorts; and cousins for the 1977–98 cohorts), as well as for the most recent cohorts whose kin counts are incomplete due to ongoing fertility. However, it

produces higher numbers for earlier cohorts, which are affected by intergenerational parent–child truncation in the registers.

Looking first at siblings, [Figure 4](#) shows that the microsimulation produces lower numbers for the post-1940 cohorts. This undercount results from the differences between sources in estimating full and half-siblings. The microsimulation produces more full siblings for most cohorts, except for those born in 1933–53, and fewer (almost zero) half-siblings in all cohorts. This under-representation of half-siblings is expected, given that the partnership and fertility models we use in the simulation do not include marriage rates. As explained earlier, the microsimulation set-up is intentionally simple so that we can assess the accuracy of synthetic kinship networks derived only from fertility and mortality data.

The numbers of living nieces, nephews, aunts, uncles, and cousins follow a similar pattern, being lower in the microsimulation for the cohorts not affected by parent–child truncation, but higher for early cohorts. Specifically, the total number of living nieces and nephews is lower for people born after 1940 (see [Figure S4](#), supplementary material). This difference stems from the underestimation of half-siblings, leading to fewer nieces and nephews from half-siblings in the microsimulation. Although the microsimulation estimates more children of full siblings in most cohorts, this does not offset the lower number of nieces and nephews of half-siblings.

Similarly, the microsimulation estimates fewer living aunts and uncles for people born after 1965 (see [Figure S5](#), supplementary material) and fewer living cousins for people born after 1970 (see [Figure 5](#)). However, for earlier cohorts, missing information on grandparents in the registers (see unregistered grandparents in [Figure 3](#)) leads to missing aunts and uncles and then missing cousins in the empirical data. Therefore, while the microsimulation produces fewer horizontal kin for the cohorts with almost complete information in the registers, it provides higher estimates of horizontal kin for the early cohorts affected by parent–child truncation.

Differences in the number of living kin may also be due to different numbers of deaths, which also play a role in shaping kinship networks. The microsimulation produces similar numbers of deceased children and grandchildren as the registers for most cohorts but higher numbers for other types of kin for people born before 1970. The average number of deceased children is similar for women (see [Figure 1](#)) and men (see [Figure S2](#)) born after 1940, while the number of deceased grandchildren is

almost identical for both sexes in all cohorts (see [Figure S3](#)). In contrast, the numbers of deceased parents for most cohorts (see [Figure 2](#)) and of deceased grandparents for all cohorts (see [Figure 3](#)) are higher in the microsimulation. The larger numbers probably account for some individuals classified as unregistered in the empirical data. Given that the microsimulation tracks the entire population, any simulated parents or grandparents who are not alive are necessarily deceased, unlike in the Swedish registers, where the status of unregistered individuals remains uncertain. This may be due to migration since the Swedish data account only for people born and registered in Sweden.

Finally, the numbers of deceased aunts and uncles (see [Figure S5](#)) and deceased cousins (see [Figure 5](#)) are higher in the microsimulation for the early cohorts (up to +4.3 aunts and uncles and +2.0 cousins for the 1950 cohort), whereas the registers counted almost zero of these kin for the same cohorts. Missing information on deceased grandparents in the registers probably results in undercounting aunts, uncles, and cousins, which might be better estimated in the simulation.

To summarize the differences in the mean numbers of kin, [Figure 6](#) shows that the microsimulation produces similar numbers of *living* kin to the registers for the cohorts that are not affected by intergenerational parent–child truncation in the registers (i.e. post-1920, post-1940, or post-1977, depending on the kin) and also for the most recent cohorts that have not completed fertility (albeit slightly lower numbers for some kin). However, the discrepancies are more pronounced for early cohorts, where truncation and survivorship bias in the registers can lower kin counts.

Distributions of living kin

While the microsimulated and register-based counts are relatively consistent in terms of the mean numbers of living kin, the differences increase when examining the distribution of the numbers of living kin of a given type. For most cohorts and kin types, the microsimulation tends to over-represent individuals with the smallest family sizes while under-representing individuals with the largest family sizes. As shown in [Figure 7](#), the microsimulation estimates a larger proportion of women with only one child but a smaller proportion of women with two children, especially for the pre-1950 cohorts, which are less covered by parity-specific data. The overestimation of individuals with only

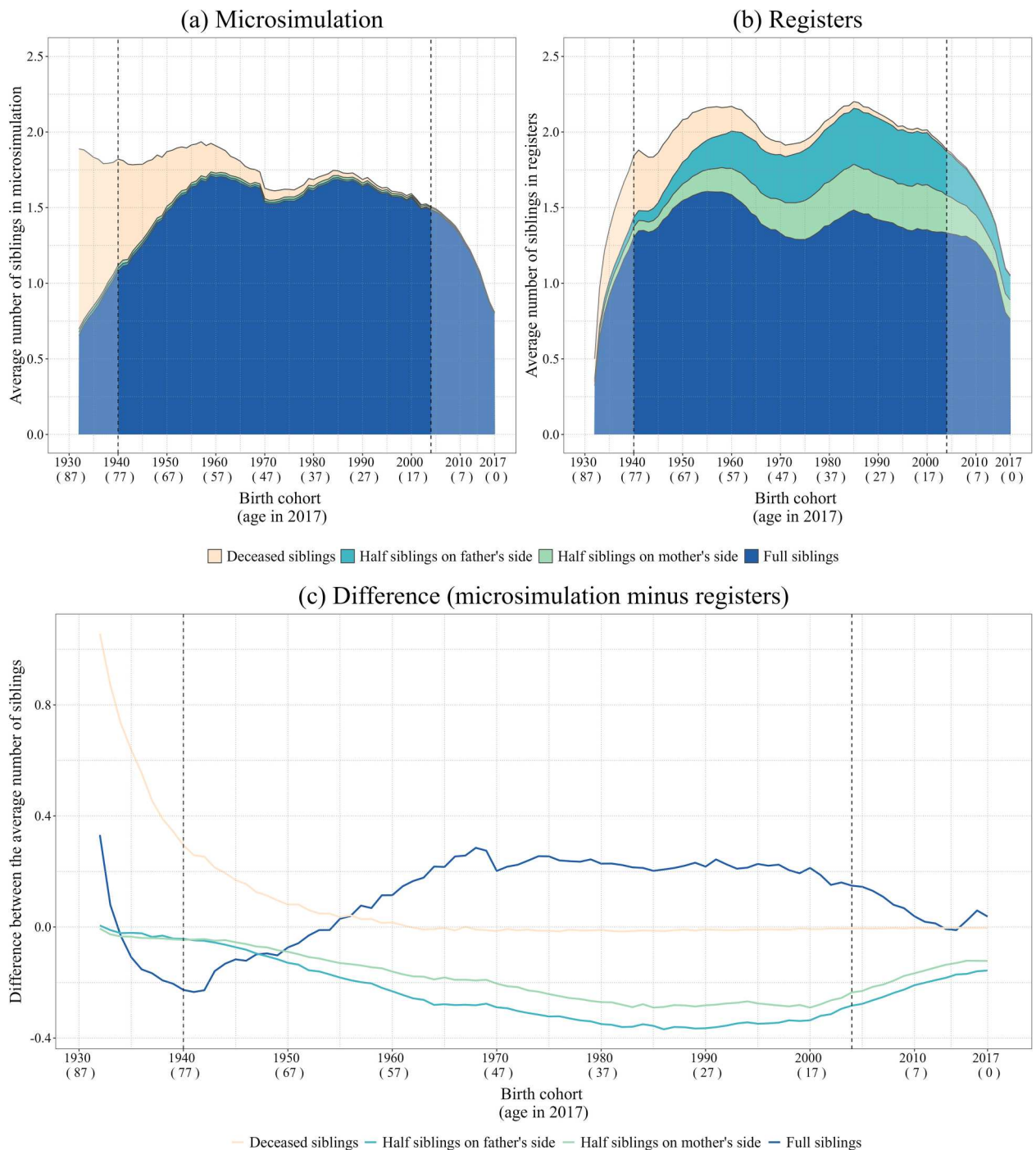


Figure 4 Average number of siblings by birth cohort and whether full or half-siblings: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Notes: The vertical dashed lines distinguish the cohorts with complete and incomplete kin counts in the registers. The areas on the right-hand side represent cohorts for whom we may not be able to observe all their siblings (as some will not yet be born in 2017). The areas on the left-hand side represent the cohorts with incomplete coverage due to missing parent–child links. Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: As for Figure 1.

one child and underestimation of those with two children also holds for most cohorts of men (see Figure S6, supplementary material). A similar pattern is

also observed for other types of kin. For most cohorts, the microsimulation over-represents individuals with one grandchild (see Figure S7,

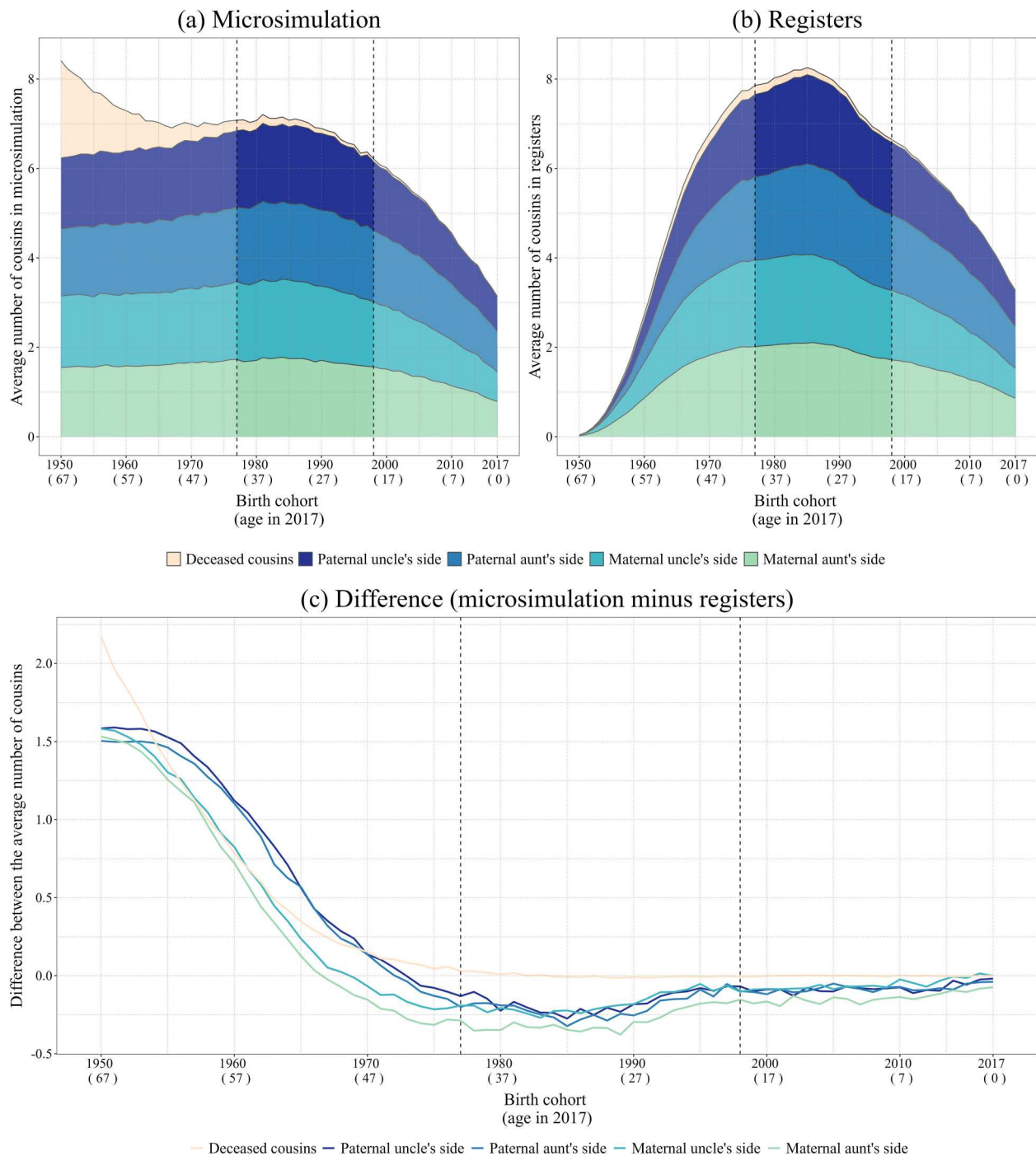


Figure 5 Average number of cousins, by birth cohort and type of cousin: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Notes: The vertical dashed lines distinguish the cohorts with complete and incomplete kin counts in the registers. The areas on the right-hand side represent cohorts for whom we may not be able to observe all their cousins (as some will not yet be born in 2017). The areas on the left-hand side represent the cohorts with incomplete coverage due to missing parent-child links in the registers. Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: As for Figure 1.

supplementary material) or one sibling (see Figure S8, supplementary material) but under-represents individuals with four grandchildren, those with four

or more siblings, and those with 11 or more cousins (see Figure S9, supplementary material). For early cohorts, the microsimulation also overestimates

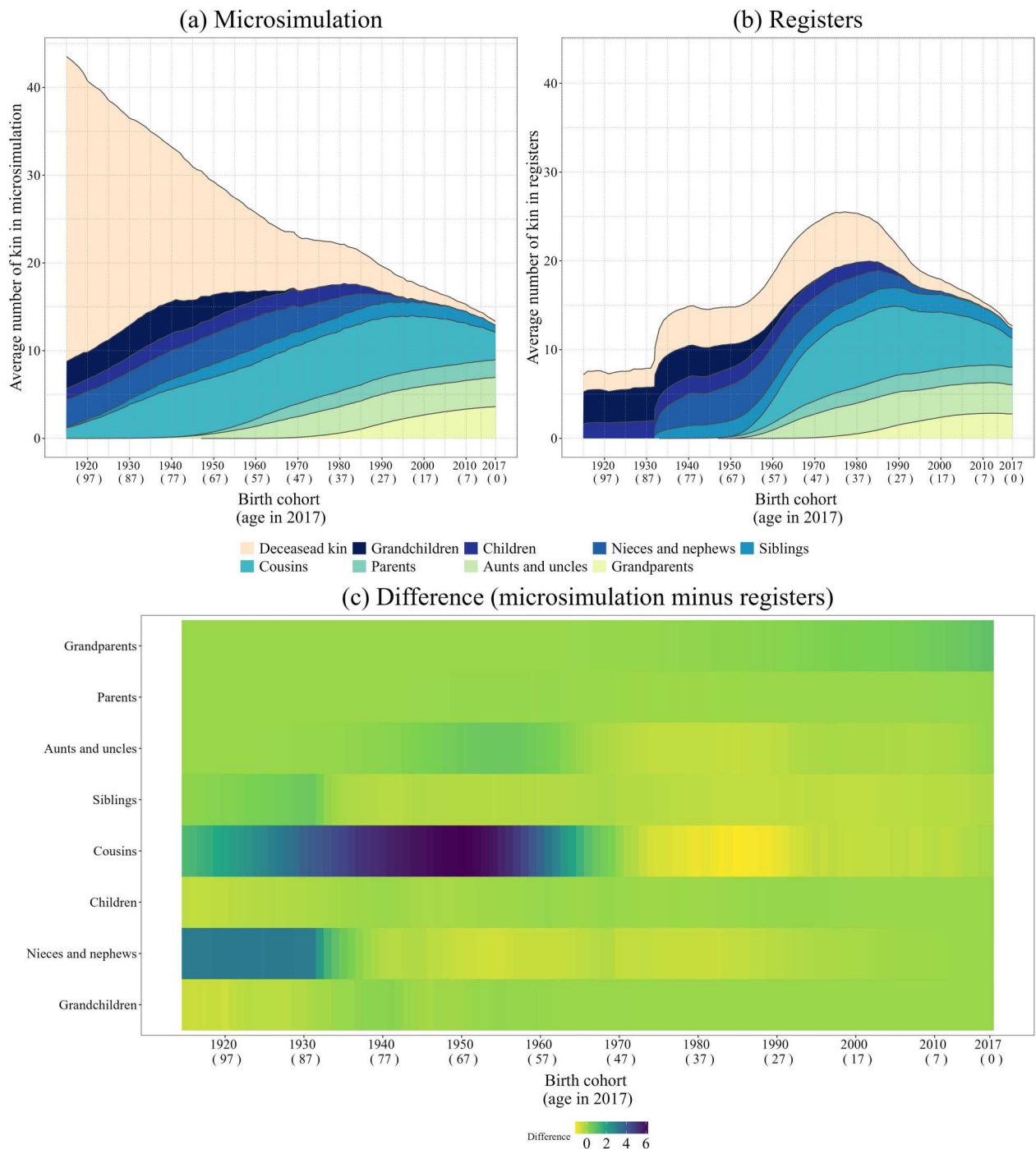


Figure 6 Average number of all types of kin in 2017, by birth cohort: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Note: These counts include children from all childbearing partners and of full and half-relatives. Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: As for Figure 1.

women and men without children or grandchildren but underestimates individuals without siblings or cousins, who are over-represented in the registers due to missing information.

Likewise, the microsimulation shows a more even distribution of the total number of living

kin across all simulated cohorts (see Figure S10, supplementary material). Therefore, fewer individuals from each simulated cohort have no kin or a large set of kin (i.e. 40+) compared with the empirical counts. The registers may lack information on some kin for individuals in the early

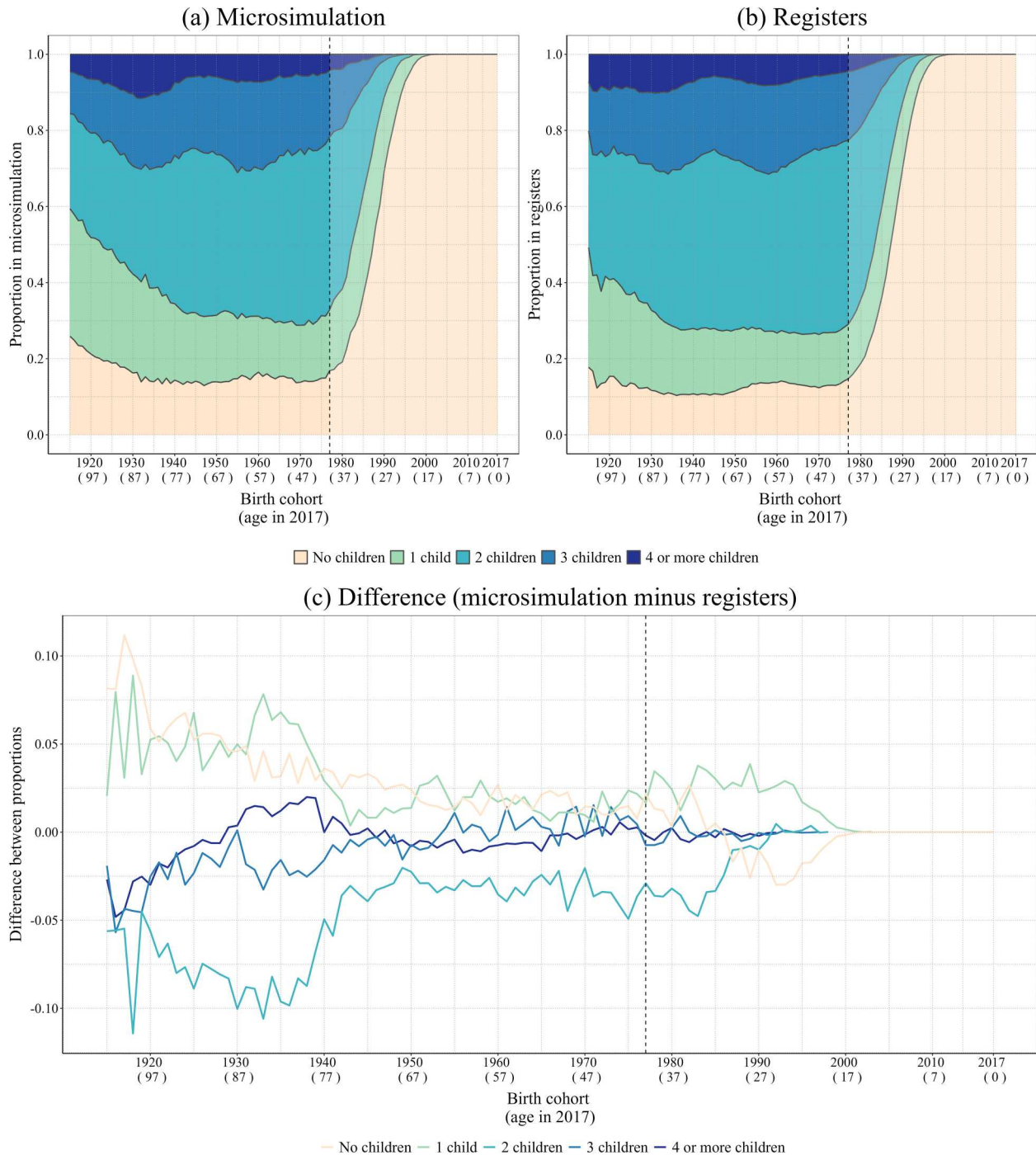


Figure 7 Proportional distribution of the number of living children per woman in 2017, by birth cohort: (a) estimated from a SOCSIM microsimulation; (b) estimated from the Swedish Registers; and (c) the difference between estimates (microsimulation minus registers)

Notes: The vertical dashed lines distinguish the cohorts with complete and incomplete kin counts in the registers. The areas on the right-hand side represent cohorts for whom we may not be able to observe all their children (as some will not yet be born in 2017). Negative values mean that microsimulation estimates are lower than register-based ones. This figure is best viewed online in colour.

Source: As for Figure 1.

cohorts, resulting in more individuals with no kin in those cohorts, but registers are more likely to include above-average family sizes than the microsimulation in nearly all cohorts.

Discussion

Estimating kinship networks can be challenging due to the limited availability of data on vital events and

family ties, particularly when considering relatives beyond the nuclear family or household. Two of the most common approaches to reconstructing these networks are the use of population registers and demographic microsimulation. Here, we compare kinship networks estimated using microsimulated and empirical data, for the contemporary Swedish population, to understand the main differences between the two sources and their effects on kin counts. We aim to shed light on the accuracy of synthetic kinship networks derived from a relatively simple microsimulation model by comparing them to a ‘gold-standard’ empirical source. The resulting insights should be valuable for future research using microsimulation to study kinship in contexts where only population-level data are accessible and a comparison with empirical data is not possible.

As shown in the Results section, the numbers of kin derived from the microsimulation and the registers are broadly similar, particularly regarding the mean number of living kin. However, the microsimulation produces slightly lower numbers of most types of kin for cohorts not affected by intergenerational parent–child truncation in the registers. This underestimation is partly due to the limited estimation of multi-partner fertility in the microsimulation, resulting in lower numbers of children from multiple child-bearing partners, half-siblings, and nieces and nephews of half-siblings.

According to Ruggles (1993), the underestimation of some kin may also result from ignoring correlations in demographic behaviour between family members. The SOCSIM microsimulator has an option to assign different fertility and mortality rates to specific groups. This could be used to replicate family clustering in demographic behaviour by creating groups of families with distinct fertility and mortality rates. However, simulating these correlations would require a more complex microsimulation set-up and substantial calibration of the input rates, because the original data are disaggregated only by sex and age. This would not align with this paper’s goal of using a simple microsimulation to assess the agreement between synthetic and empirical kinship networks.

As expected, the microsimulation produces less dispersion in kin counts across all kin types, thereby reflecting the limitations of the fertility model and the absence of family correlations in demographic behaviour. This was precisely the concern highlighted by Ruggles (1993). Empirical registers show a higher proportion of individuals with no kin, especially in earlier cohorts, probably due to missing data, whereas the simulation by

design includes all individuals. Conversely, large kin networks, which occur naturally in the real-world population but are less likely under the model’s simplified fertility assumptions, are under-represented in the simulation results.

Despite the undercount and under-dispersion of most types of kin in the microsimulation for the cohorts not affected by intergenerational truncation in the registers, a fully tracked synthetic population provides better accounting for some relatives of the early cohorts who are likely to be omitted from the registers due to missing information on parent–child links, conditioning on survival, or truncation of kinship ties due to migration. This applies to parents, siblings, nieces, nephews, and all types of deceased kin of people born before 1940 and also to grandparents, aunts, uncles, and cousins of people born before 1977. Moreover, microsimulation estimates can go further back in time where reliable register data are unavailable.

The results we present here have some limitations that we would like to acknowledge. Our results are based on simulated populations, which are not real-world populations and thus cannot fully reproduce the complexity of their dynamics and structures. Exact agreement between model-based kinship calculations and empirical measurements should not be expected (Keyfitz and Caswell 2005). First, we do not consider the familial transmission of fertility and mortality behaviour, as the input data are disaggregated only by sex and age. Thus, there is no pre-defined family clustering beyond the individual stochasticity resulting from the microsimulation. Second, compared with a simulation using fertility rates for all birth orders combined over the whole period (results not presented here), the accuracy of the results improves slightly on the inclusion of age-specific rates by birth order for 1970–2017. This suggests that using parity-specific data enables microsimulation to perform even better than when all birth orders are combined, although such detailed data are rarely available. Third, in the absence of data at higher levels of disaggregation (by age, parity, and marital status) for the whole period, our modelling of partnership and fertility may oversimplify real-world dynamics and the complexity of modern families. While these simplifications may limit the accurate estimation of non-marital fertility, men’s fertility, multi-partner fertility, and complex family structures involving half-siblings, estimates of women’s overall fertility remain reasonably accurate.

It is worth mentioning that some studies have used more detailed SOCSIM microsimulation set-ups,

whose results may be more finely tuned to specific contexts. In the US, Wachter (1997) used rates disaggregated by marital status and parity to examine the availability of kin among older people from 1990 to 2030, and most recent studies have used rates specific to marital status, stratified by race, to project the absence of kin among white and Black older adults from 2015 to 2060 (Verdery and Margolis 2017) and to analyse changes in the demography of grandparenthood for the 1880–1960 cohorts (Margolis and Verdery 2019). In Britain, age-specific fertility, mortality, and nuptiality rates (including divorce and cohabitation) have been used to analyse changes in the mean number of kin and the proportion of the population with specific kin, by age group for selected years (Murphy 2010a, 2010b) and birth cohorts (Murphy 2011). To model the HIV epidemic in Zimbabwe, SOCSIM's 'group' option was used to represent HIV status, with mortality rates distinguishing between HIV-positive and HIV-negative individuals (Zagheni 2011).

Apart from the simplifications mentioned, the demographic rates we use for our simulation consider the entire population, including migrants and their descendants. This can lead to discrepancies with empirical data such as population registers, which typically cover only the native-born population. According to Statistics Sweden (2026), international migration flows in Sweden were relatively small for most of the mid-nineteenth and twentieth centuries. Immigrants and emigrants typically accounted for less than 1 per cent of the total population each year, except during the late-nineteenth-century emigration wave. Therefore, using yearly demographic rates estimated for the total population, including migrants, is unlikely to distort kinship structures substantially for most of the studied period. However, deviations may arise in years of high migration. Thus, the effect of using vital rates with or without migrants is likely to be inconsequential. Conversely, there will always be a discrepancy between empirical kin counts (which reflect the fact that many inhabitants of countries with immigration do not have full data on kinship ties) and microsimulated approaches (which assume such networks are complete for everyone). A valuable area for future exploration would be to examine how migration affects kinship counts in synthetic populations. One approach would be to estimate demographic rates based on ancestral population groups in a given country and to use these as inputs for microsimulations to align the analysis populations better. Alternatively, future

research could estimate demographic rates that are accurate for different populations, disaggregated by country of origin, to simulate and count kin in both the destination and origin countries.

Matrix kinship models provide an alternative formal framework for analytically deriving the expected numbers and age distributions of kin from population-level demographic rates. These models share some of the characteristics, advantages, and limitations of microsimulation models. The initial time-invariant matrix formulation (Caswell 2019) has been expanded to incorporate 'stages' such as parity (Caswell 2020), time variation (Caswell and Song 2021), two sexes (Caswell 2022), causes of death (Caswell et al. 2023), stochasticity (Caswell 2024), and lifetime kin overlap (Caswell and de Vries 2025). All these factors can also be included in a microsimulation, provided the necessary data are available. Similarly to `rsocsim` (Theile et al. 2023), the `DemoKin` package (Williams et al. 2021) has operationalized the theoretical matrix models in R. Recent studies have used different implementations of these models to describe the demographic composition of kinship networks (Alburez-Gutierrez et al. 2023, 2024; Feng et al. 2024; Hünteler et al. 2024; Robles et al. 2024; Adhikari and Alburez-Gutierrez 2025; Matos-Moreno et al. 2025; van Damme et al. 2025) and to study the experience of kin loss (Schlüter et al. 2024; Acosta et al. 2026). Analytical and microsimulation approaches involve different trade-offs. Matrix models offer analytical transparency and computational efficiency but are limited in the inclusion of individual variability and interactions. Conversely, microsimulation produces demographic events at the individual level that are consistent with observed population-level rates, but it is more computationally intensive. A systematic comparison of SOCSIM microsimulation and matrix-model outputs is beyond the scope of this study but is the subject of ongoing work.

Despite the aforementioned limitations of the SOCSIM microsimulation model, it produces kinship networks that closely match those in high-quality population registers in terms of average kinship. It also provides reasonable estimates of the distribution of numbers of certain kin types. This highlights the utility of microsimulation for reconstructing kinship networks when only population-level demographic data are available, thereby providing indirect empirical support for previous studies that have used SOCSIM in other contexts. Furthermore, microsimulations can be used

to estimate kinship structures in historical periods that lack micro-level data and also to project future kinship dynamics based on demographic trends. By systematically comparing microsimulated networks with empirical data, this study strengthens the methodological basis of microsimulation in demographic research and encourages further investigation into kinship dynamics across diverse populations and contexts.

Notes and acknowledgments

- 1 Liliana P. Calderón-Bernal is based in the Department of Digital and Computational Demography, Max Planck Institute for Demographic Research (MPIDR), Rostock, Germany, and also the Demography Unit, Department of Sociology, Stockholm University, Sweden. Diego Alburez-Gutierrez is based in the Kinship Inequalities Research Group, MPIDR, and also the Humboldt-Universität zu Berlin. Martin Kolk is based in the Demography Unit, Department of Sociology, Stockholm University; the Institute for Futures Studies, Stockholm; and also Åbo Akademi University, Vasa, Finland. Emilio Zagheni is based in the Department of Digital and Computational Demography, MPIDR.
- 2 Please direct all correspondence to Liliana P. Calderón-Bernal, Max Planck Institute for Demographic Research, Konrad-Zuse-Str. 1, 18057, Rostock, Germany; or by E-mail: calderonbernal@demogr.mpg.de.
- 3 We thank Emma Pettersson for providing the R scripts developed for the Swedish kinship universe (Kolk et al. 2023), which enabled the reproduction of the figures in their paper and served as the basis for our replication using microsimulation. We thank Tom Theile for assistance with programming, especially with *rsocsim*. The manuscript was discussed at the Work-in-Progress Workshop of the Department of Digital and Computational Demography at MPIDR, where it benefited from several useful comments. We would like to express our sincere gratitude to Hal Caswell and an anonymous reviewer for their constructive and valuable comments on our manuscript. Liliana P. Calderón-Bernal gratefully acknowledges the resources provided by the International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS). Martin Kolk gratefully accepted funding from the Swedish Research Council (grant numbers 2025–00935; 2022–02361; 2022–02314).
- 4 Data and code availability statement: The code to retrieve the data, run the microsimulations, and reproduce the results is available online: https://github.com/liliana-calderon/SOCSIM_Registers.

Author contributions

CRedit: **Liliana P. Calderón-Bernal:** Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft; **Diego Alburez-Gutierrez:** Conceptualization, Methodology, Writing – review & editing; **Martin Kolk:** Conceptualization, Methodology, Writing – review & editing; **Emilio Zagheni:** Methodology, Writing – review & editing

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Liliana P. Calderón-Bernal  <http://orcid.org/0009-0002-7629-9718>

Diego Alburez-Gutierrez  <http://orcid.org/0000-0002-9823-5179>

Martin Kolk  <http://orcid.org/0000-0002-7175-4040>

Emilio Zagheni  <http://orcid.org/0000-0002-7660-8368>

References

- Acosta, E., D. Alburez-Gutierrez, M. Gargiulo, and C. Torres. 2026. Weaponizing kinship: A demographic analysis of bereavement in the Colombian conflict, *Population and Development Review*, 1–38. <https://doi.org/10.1111/padr.70048>
- Adhikari, S., and D. Alburez-Gutierrez. 2025. The future of grandparenthood in South Asia: The role of population aging and educational expansion, *MPIDR Working Papers* WP-2025-003. <https://doi.org/10.4054/MPIDR-WP-2025-003>
- Alburez-Gutierrez, D., E. Acosta, E. Zagheni, and N. E. Williams. 2024. The long-lasting effect of armed conflicts deaths on the living: Quantifying family bereavement, *Science Advances* 10(30): eado6951. <https://doi.org/10.1126/sciadv.ado6951>
- Alburez-Gutierrez, D., C. Mason, and E. Zagheni. 2021. The “Sandwich Generation” revisited: Global demographic drivers of care time demands, *Population and Development Review* 47(4): 997–1023. <https://doi.org/10.1111/padr.12436>
- Alburez-Gutierrez, D., I. Williams, and H. Caswell. 2023. Projections of human kinship for all countries, *Proceedings of the National Academy of Sciences* 120(52): e2315722120. <https://doi.org/10.1073/pnas.2315722120>

- Caswell, H. 2019. The formal demography of kinship: A matrix formulation, *Demographic Research* 41: 679–712. <https://doi.org/10.4054/DemRes.2019.41.24>
- Caswell, H. 2020. The formal demography of kinship II: Multistate models, parity, and sibship, *Demographic Research* 42: 1097–1146. <https://doi.org/10.4054/DemRes.2020.42.38>
- Caswell, H. 2022. The formal demography of kinship IV: Two-sex models and their approximations, *Demographic Research* 47(13): 359–396. <https://doi.org/10.4054/DemRes.2022.47.13>
- Caswell, H. 2024. The formal demography of kinship VI: Demographic stochasticity and variance in the kinship network, *Demographic Research* 51: 1201–1256. <https://doi.org/10.4054/DemRes.2024.51.39>
- Caswell, H., R. Margolis, and A. Verdery. 2023. The formal demography of kinship V: Kin loss, bereavement, and causes of death, *Demographic Research* 49: 1163–1200. <https://doi.org/10.4054/DemRes.2023.49.41>
- Caswell, H., and X. Song. 2021. The formal demography of kinship III: Kinship dynamics with time-varying demographic rates, *Demographic Research* 45: 517–546. <https://doi.org/10.4054/DemRes.2021.45.16>
- Caswell, H., and L. de Vries. 2025. The formal demography of kinship VII: Lifetime kin overlap within and across generations, *Demographic Research* 53: 123–174. <https://doi.org/10.4054/DemRes.2025.53.5>
- Daw, J., A. M. Verdery, and R. Margolis. 2016. Kin count(s): Educational and racial differences in extended kinship in the United States, *Population and Development Review* 42(3): 491–517. <https://doi.org/10.1111/j.1728-4457.2016.00150.x>
- De Bel, V., E. Bokányi, K. Hank, and T. Leopold. 2025. A parallel kinship universe? A replication of Kolk et al. (2023) with Dutch register data on kinship networks, *Demographic Research* 52: 915–938. <https://doi.org/10.4054/DemRes.2025.52.28>
- Dykstra, P. A., M. Kalmijn, T. C. M. Knijn, A. E. Komter, A. C. Liefbroer, and C. H. Mulder (eds). 2006. *Family Solidarity in the Netherlands*. Amsterdam: Dutch University Press.
- Feng, K., X. Song, and H. Caswell. 2024. Kinship and care: Racial disparities in potential dementia caregiving in the United States from 2000 to 2060, *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences* 79(Supplement_1): S32–S41. <https://doi.org/10.1093/gerona/glae106>
- Hammel, E. A. 2005. Demographic dynamics and kinship in anthropological populations, *Proceedings of the National Academy of Sciences* 102(6): 2248–2253. <https://doi.org/10.1073/pnas.0409762102>
- Hammel, E. A., D. W. Hutchinson, K. W. Wachter, R. T. Lundy, and R. Z. Deuel. 1976. *The SOCSIM Demographic-Sociological Microsimulation Program: Operating Manual*. Berkeley: Institute of International Studies. University of California.
- Hammel, E. A., C. Mason, K. W. Wachter, F. Wang, and H. Yang. 1991. Rapid population change and kinship: The effects of unstable demographic changes on Chinese kinship networks, 1750–2250, in *Consequences of Rapid Population Growth in Developing Countries*. New York: Taylor & Francis, pp. 243–271.
- Hammel, E. A., C. K. McDaniel, and K. W. Wachter. 1979. Demographic consequences of incest tabus: A microsimulation analysis, *Science* 205(4410): 972–977. <https://doi.org/10.1126/science.205.4410.972>
- Hammel, E. A., and K. W. Wachter. 1996. Evaluating the slavian census of 1698 part II: A microsimulation test and extension of the evidence, *European Journal of Population* 12(4): 295–326. <https://doi.org/10.1007/BF01796911>
- Human Fertility Collection. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at www.fertilitydata.org (data downloaded on 13 February 2024).
- Human Fertility Database. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at www.humanfertility.org (data downloaded on 11 April 2024).
- Human Mortality Database. Max Planck Institute for Demographic Research (Germany) and University of California, Berkeley (USA) and French Institute for Demographic Studies (France). Available at www.mortality.org (data downloaded on 13 February 2024).
- Hünteler, B., A. Polizzi, and A. A. van Raalte. 2024. Changing kinship structures in East and West Germany before and after reunification, *SocArXiv*, 6bf7k_v1. <https://doi.org/10.31235/osf.io/6bf7k>
- Keyfitz, N., and H. Caswell. 2005. *Applied Mathematical Demography* (Third Edition). New York: Springer.
- Kolk, M., L. Andersson, E. Pettersson, and S. Drefahl. 2023. The Swedish kinship universe: A demographic account of the number of children, parents, siblings, grandchildren, grandparents, aunts/uncles, nieces/nephews, and cousins using national population registers, *Demography* 60(5): 1359–1385. <https://doi.org/10.1215/00703370-10955240>
- Margolis, R., and A. M. Verdery. 2019. A cohort perspective on the demography of grandparenthood: Past, present, and future changes in race and sex disparities in the United States, *Demography* 56(4): 1495–1518. <https://doi.org/10.1007/s13524-019-00795-1>
- Mason, C. 2016. *Socsim Oversimplified*, Demography Lab, University of California. <https://lab.demog.berkeley.edu/socsim/CurrentDocs/socsimOversimplified.pdf>

- Matos-Moreno, A., D. Alburez-Gutierrez, I. Williams, A. M. Verdery, M. Fernández Soto, and A. Santos-Lozada. 2025. Kinship structures for left behind older adults in high outmigration contexts: Evidence from Puerto Rico, *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* 80(6): gbaf052. <https://doi.org/10.1093/geronb/gbaf052>
- Murphy, M. 2010a. Family and kinship networks in the context of ageing societies, in S. Tuljapurkar, N. Ogawa, and A. H. Gauthier (eds), *Aging in Advanced Industrial States*. Springer Netherlands, pp. 263–285. https://doi.org/10.1007/978-90-481-3553-0_11
- Murphy, M. 2010b. Changes in family and kinship networks consequent on the demographic transitions in England and Wales, *Continuity and Change* 25(1): 109–136. <https://doi.org/10.1017/S026841601000007X>
- Murphy, M. 2011. Long-term effects of the demographic transition on family and kinship networks in Britain, *Population and Development Review* 37: 55–80. <https://doi.org/10.1111/j.1728-4457.2011.00378.x>
- Riffe, T. 2015. *Reading Human Fertility Database and Human Mortality Database data into R*, MPIDR Technical Report TR TR-2015-004. Max Planck Institute for Demographic Research. https://www.demogr.mpg.de/en/publications_databases_6118/publications_1904/mpidr_technical_reports/reading_human_fertility_database_and_human_mortality_database_data_into_r_5438
- Robles, A., Martes-Camargo, P., and R. Rodríguez-Franco. 2024. ¿Quién cuida a la niña? Provisión de cuidados y estructura de parentesco en América Latina. [Who takes care of the girl? Care provision and kinship structure in Latin America], *Revista Brasileira de Estudos de População* 41. <https://doi.org/10.20947/S0102-3098a0274>
- Ruggles, S. 1993. Confessions of a microsimulator: Problems in modeling the demography of kinship, *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 26(4): 161–169. (world). <https://doi.org/10.1080/01615440.1993.9956352>
- Schlüter, B.-S., D. Alburez-Gutierrez, K. Bibbins-Domingo, M. J. Alexander, and M. V. Kiang. 2024. Youth experiencing parental death due to drug poisoning and firearm violence in the US, 1999–2020, *JAMA* 331(20): 1741–1747. <https://doi.org/10.1001/jama.2024.8391>
- Statistics Sweden. 2026. *Statistikdatabasen—Table—Population and Population Changes in Sweden by Sex. Year 1749–2024* [Dataset]. https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START_BE_BE0101_BE0101G/BefUtvKon1749/
- Theile, T., D. Alburez-Gutierrez, L. P. Calderón-Bernal, M. Snyder, and E. Zagheni. 2023. *rsocsim: SOCSIM with R* [Computer software], Max Planck Institute for Demographic Research. <https://github.com/MPIDR/rsocsim>
- van Damme, M., D. Alburez-Gutierrez, and A. F. Castro Torres. 2025. Research note: Estimating kinship size of older adults in Europe with models and surveys, *Demography* 62(3): 763–772. <https://doi.org/10.1215/00703370-11961236>
- Verdery, A. M., and R. Margolis. 2017. Projections of white and black older adults without living kin in the United States, 2015 to 2060, *Proceedings of the National Academy of Sciences* 114(42): 11109–11114. <https://doi.org/10.1073/pnas.1710341114>
- Wachter, K. W. 1997. Kinship resources for the elderly, *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 352(1363): 1811–1817. <https://doi.org/10.1098/rstb.1997.0166>
- Wachter, K. W., D. Blackwell, and E. A. Hammel. 1997. Testing the validity of kinship microsimulation, *Mathematical and Computer Modelling* 26(6): 89–104. [https://doi.org/10.1016/S0895-7177\(97\)00172-6](https://doi.org/10.1016/S0895-7177(97)00172-6)
- Wachter, K. W., D. Blackwell, and E. A. Hammel. 1998. *Testing the Validity of Kinship Microsimulation: An Update*. University of California. <https://citeseerx.ist.psu.edu/document?doi=b43eecd6d96e6b1aa9c81bbe78d8377d6901d8da&repid=rep1&type=pdf>
- Williams, I., D. Alburez-Gutierrez, X. Song, and H. Caswell. 2021. *DemoKin: An R Package to Implement Demographic Matrix Kinship Models* [Computer software]. <https://cran.r-project.org/web/packages/DemoKin/>
- Wolf, D. A. 2001. Note on statistical analysis and microsimulation for studying living arrangements and intergenerational transfers, in N. A. Apt and P. Lloyd-Sherlock (eds), *Population Bulletin of the United Nations. Special Issues Nos. 42/43, 2001: Living Arrangements of Older Persons: Critical Issues and Policy Responses*, UN, pp. 386–392. Technical Meeting on Population Ageing and Living Arrangements of Older Persons: Critical Issues and Policy Responses 2000, New York. <https://digitallibrary.un.org/record/464936>
- Zagheni, E. 2011. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe, *Population and Development Review* 37(4): 761–783. <https://doi.org/10.1111/j.1728-4457.2011.00456.x>
- Zagheni, E. 2015. Microsimulation in demographic research, in J. D. Wright (ed.), *International Encyclopedia of the Social & Behavioral Sciences* (Second Edition). Elsevier, pp. 343–346. <https://doi.org/10.1016/B978-0-08-097086-8.31018-2>
- Zhao, Z. 2006. Computer microsimulation and historical study of social structure: A comparative review of SOCSIM and CAMSIM, *Revista de Demografia Historica* XXIV(II): 59–88.