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

Martin Kolk, Linus Andersson, Emma Pettersson, and Sven Drefahl

ABSTRACT Given that surprisingly little is known about the demography of human kinship, we provide a demographic account of the kinship networks of individuals in Sweden in 2017 across sex and cohort between ages 0 and 102. We used administrative register data of the full population of Sweden to provide the first kinship enumeration for a complete population based on empirical data. We created ego-focused kinship networks of children, parents, siblings, grandchildren, grandparents, aunts and uncles, nieces and nephews, and cousins. We show the average number of kin of different types, the distribution of the number of kin, and changes in dispersion over time. A large share of all kin of an individual are horizontal kin, such as cousins. We observe the highest number of kin—on average, roughly 20—around age 35. We show differences between matrilineal and patrilineal kin and differences in the kinship structure arising from fertility with more than one childbearing partner, such as half-siblings. The results demonstrate substantial variability in kinship within a population. We discuss our findings in the context of other methods to estimate kinship.

KEYWORDS Kinship • Relatives • Extended family • Sweden • Genealogy

Introduction

Kin plays an important role in people’s lives. Family members, including those outside the household, are an important part of an individual’s social environment and often provide substantial informal care and support to each other. Grandparents often provide financial and emotional help and child-rearing assistance to younger generations. Grandchildren and adult children maintain vital care and social contact for the elderly. Individuals’ closest confidants often include their siblings. Networks of aunts, uncles, and cousins are among the few relationships that tend to last throughout a person’s life. Whereas social scientists of the twentieth century were relatively disinterested in kinship in developed countries, recent research has emphasized that individuals are far from decoupled from extended kin (Bengtson 2001; Furstenberg 2020; Mare 2011). Intergenerational kinship relations may have become more important as
increased longevity has extended the frequency of kinship over three or more generations (Bengtson 2001) and increasing dependency ratios have reshaped the conditions for public transfers and private transfers across generations (Lee and Mason 2011).

Despite evidence of the importance of kin, researchers know comparably little about the frequencies of different kin types in contemporary populations (Daw et al. 2016). Research on extended family members has focused primarily on the frequency of interactions with kin. The few current efforts to enumerate kin are drawn from survey samples, mathematical demography, and microsimulations.

This study provides the first enumeration of kinship networks using observed empirical data for all ages for an entire population. We sidestep several limitations evident in previous approaches in two ways: (1) by inferring kinship counts empirically from the total population of Sweden and (2) by linking individuals across five generations to construct a nearly full depiction of consanguineal relations by cohort for the population of Sweden alive in 2017. We present granular descriptions of biological kinship, including kin traced through full and half-siblings, children by partner order, and separate kin counts for the maternal and paternal sides. Several of these aspects (e.g., multipartner fertility) have been largely outside the reach of state-of-the-art simulation and mathematical models, although increasing sophistication in matrix models now allows two-sex kinship models (e.g., on two-sex kinship, see Caswell 2022). Similarly, most research has mainly predicted the average number of kin (but see Caswell 2020). In this study, we use both population averages and measures of spread and dispersion in total kin.

There is much to gain from descriptively enumerating kinship. The potential for an individual to receive support, be burdened, or otherwise be influenced by kin (Furstenberg 2020) is ultimately dictated by the availability of kin, which is highly variable in the population, as we will show. Given the importance of kinship structure in providing family care, kinship networks may also influence the overall demand for formal and informal care. Enumerating kin informs policies by clarifying the number of individuals without any living kin, which individuals can and cannot rely on grandparental care, and the distribution of working-age adults in an individual’s kinship network (Agree and Glaser 2009; Szebehely 2005). Moreover, our estimates provide data to validate simulation models of kinship.

Demographic changes affect the structure of kinship over time. Rapid increases in life expectancy over the twentieth century have extended the likelihood that people will interact with kin from older or younger generations (vertical kin), such as parents, grandparents, and grandchildren (Bengtson 2001; Murphy 2011), although changes in childbearing timing and quantity influence this likelihood. A population’s kinship network is the product of fertility, mortality, and other demographic behaviors as they change across periods and generations. Decreases in fertility mean that people have less collateral (horizontal) kin, such as siblings, cousins, aunts, and uncles, producing what has been described as a beanpole kinship structure: a network of kin with a “stem” reaching as high as four generations but where each member has few “branches” of generational contemporaries (Bengtson 2001; Uhlenberg 1996).

Demography drives kinship structure through more mechanisms than life expectancy and the average number of children, and kinship structures have many dimensions other than generational overlap (Murphy 2011). Changes in the mean age at childbearing can increase or decrease the number of generations subsequently alive.
Sex differences in fertility affect whether more kin exist on the mother’s or father’s side. For example, the likelihood of having two living cousins depends on the cousins’ longevity and the fertility of a second generation (i.e., aunts’/uncles’ fertility determines the number of cousins) and a third generation (i.e., grandparents’ fertility determines the number of aunts and uncles). The likelihood of having a paternal grandmother alive at age 50 is determined by sex-specific longevity and the generational overlap produced through the grandparent’s and parent’s sex-specific mean age at childbearing, and these factors will differ for matrilineal and patrilineal kin. Multipartner fertility affects the spacing of births, siblings, and cousins (Andersson 2020, 2021). In sum, kinship networks emerge from a diverse set of demographic behaviors across multiple generations. Parity-dependent birth spacing, multipartner fertility, assortative mating, and age differences between partners are particularly difficult to estimate from age-specific vital rates alone.

Previous Research on Kinship

Kinship Research in Anthropology, Demography, and Sociology

Early research on kinship occurred primarily within anthropology. Kinship was at the center of social science throughout the late nineteenth and early twentieth centuries (e.g., Davis and Warner 1937; Malinowski 1929/1982; Rivers 1914/1968), and it remained important well into the 1960s. Kinship research of the nineteenth century and the functionalist theory of the twentieth century saw the marginalization of extended kin as a key facet distinguishing modernity from premodern societal structures (Parsons 1943). Urbanization and wage labor have been argued to remove many of kinship’s prior functions in modern society (Giddens 1993; Parsons and Bales 1955; Schneider 1968). Family sociologists described the contemporary family as structurally isolated and linked the nuclear family to industrialization and development (e.g., Goode 1963; Palloni et al. 1996; Parsons 1943). This perspective prevailed in more recent theories on individualization and modernization (Giddens 1993; Habermas 1984/1985; Thornton 2005). Historical demographers have described the deep historical roots of nuclear families in northwestern Europe (Laslett 1983). Outside of evolutionary anthropology, contemporary anthropological studies on quantitative kinship composition and kinship systems remain rare (Furstenberg 2020; Johnson 2000; Mare 2011). However, there has been a recent increase in quantitative and analytical social science research connected to the rich anthropological/historical literature (e.g., Alburez-Gutierrez et al. 2022; Lansing et al. 2017; Schulz et al. 2019). Demography, gerontology, and sociology have studied dyadic kinship relations—such as intergenerational coresidence, geographical distance, and caregiving—and their effects on individual well-being for grandparenthood (e.g., Chudnovskaya and Kolk 2017; Margolis and Verdery 2019; Ruggles 1987; Ruggles and Brower 2003). In particular, research has found that quality and frequency of

---

1 Related, in matrix population models, the dynamics of kinship networks of a given individual can be modeled as the kin set of an individual as a “subsidized” population (Caswell 2019). Kin members are created via not only the index person’s childbearing but also the childbearing of future and past members.
contact influence senior citizens’ well-being (Brubaker 1990; Dykstra et al. 2006; Michielin and Mulder 2007; Tomassini et al. 2004).

In family sociology, most research on kin outside the nuclear family has focused on the role of kin for financial, emotional, and instrumental support or has tried to measure the intensity of affection and contact in the relationship (e.g., Albertini et al. 2007; Bengtson 2001; Heady and Kohli 2010; Rossi and Rossi 1990). Researchers have consistently shown a female bias in kin relationships, where women and maternal lines are described as having stronger ties (Di Leonardo 1987; Rossi and Rossi 1990; Schneider and Cottrell 1975; Schneider and Smith 1973; Young and Willmott 1957). The mother–daughter relationship has been described as the strongest tie in the Western kinship system (Hagestad 1986; Rossi and Rossi 1990).

Importantly, the structure of kinship networks matters for how these functions of kin members play out. The beanpole kinship structure’s growing vertical depth of kin characteristics has been hypothesized to strengthen bonds across generations (Bengtson 2001). Other consequences of changing kinship structures include the postponement into older ages of bequests and the experience of bereavement. Contemporary societies are likely in a historically unique situation in which the demographic availability of grandparent–grandchildren relationships is higher than at any other point in history and quite possibly higher than in the future because of changes in fertility timing and longevity (Wachter 1997).

Research on kinship’s influences on family members has rarely considered extended kin (Coall and Hartwig 2010; Kaptijn et al. 2010). However, growing evidence suggests that kin relationships outside the household are important predictors of several socioeconomic and demographic outcomes (Mare 2011; Pfeffer 2014). Some of this research has had a clear kinship focus by studying the associations among social status, fertility, and kin group size (Kolk and Hällsten 2017). Other work has emphasized the role of socialization and the effect of older kin on social mobility. Although the importance of kin relations for social stratification beyond the immediate family remains disputed (Engzell et al. 2020), this literature has brought extended kinship to the forefront. Recent work has traced increasingly larger kinship networks to examine how social background extending back many generations affects life chances (Collado et al. 2023; Hällsten and Kolk 2023; Modalsli 2023), focusing on great-grandparents and cousins many times removed.

Broadly, researchers have focused nearly all attention on what kin “do,” identifying the “effect” of kinship on various outcomes and contrasting ties and obligations across different kin types (Daw et al. 2016). Less research has examined who kin “are” or produced reliable demographic statistics on frequencies and distributions (Daw et al. 2016).

**Kinship in Sweden in a Comparative Perspective**

Kinship in contemporary Sweden is distinct from common ways of organizing kinship globally but is broadly similar to the northwestern European marriage and family formation patterns (Hajnal 1982; Reher 2005). This pattern is characterized by neolocal household formation, low coresidence with parents, late marriage and childbearing, and a large proportion of men and women remaining unmarried and
childless throughout their lives (Boholm 1983; Laslett 1977), in contrast to patrilineal and patrilocality kin systems, which are common across Eurasia. Kinship in Western and Northern Europe has been contrasted with the Mediterranean kinship pattern in which patrilineal relationships are pronounced in kinship and household formation (Reher 2005), with Eastern European kinship with its earlier marriage ages and complex households (Hajnal 1982), and with kinship in developed countries in East Asia (e.g., South Korea, Taiwan, and Japan), which retains strong patrilineal and patrilocality characteristics (Martin 1990). Kinship is defined differently in different cultural contexts, and the meaning itself of kinship differs culturally (Schneider 1968). For example, Swedish kinship terms distinguish patrilineal and matrilineal kin in contrast to English language kinship terms. In Figure S17 (shown in online appendix 1, along with all other figures and tables designated with an “S”), we provide a kinship diagram that precisely defines the kin relationships in our study and the Swedish language names for the kin relationships.

Intergenerational relationships in contemporary Nordic countries have been described as less intense (but more universal) than in many other European countries (Albertini et al. 2007). Sweden is also a global outlier with an unusually low share of intergenerational households (Ruggles 2009). Some scholars have described Swedish society as a contract between the individual and the state, where trust in public welfare relieves individuals from caring for dependent kin (Berggren and Trägårdh 2006). Yet, the notion of complete decoupling from extended kin is likely overstated. Examples that would suggest a more nuanced view of the continuing relevance of kinship in Sweden are the high shares of individuals maintaining multigenerational relationships and ties (Albertini et al. 2007) and the prevalence of family businesses (Sjögren 2006).

**Previous Research Quantifying Frequencies of Kinship**

Previous research has estimated the kinship structures of contemporary populations by using surveys, ethnographic methods, mathematical demography, or microsimulations. Specialized surveys connect information on survey respondents to their reported kinship networks. Survey data have been used to study kinship structure (Bonvalet and Lelièvre 2016; Dykstra and Komter 2006) and heterogeneity in kinship (Goldstein 1999; Goldstein and Warren 2000). Data sources include the Netherlands Kinship Panel Study; the Dutch OKiN; and, to a lesser extent, the GGS, the SHARE surveys of old individuals, the German PAIRFAM, and complex surveys, such as the U.S. PSID.

Dykstra and Komter (2006) used the Netherlands Kinship Panel Study to summarize kin types for 8,150 individuals aged 18–79 interviewed between 2002 and 2004 in the Netherlands. They examined a selection of kin similar to the one used in our study (from grandparents to grandchildren), finding that individuals in their sample had roughly 30 kin alive at different ages, and enumerated the mean number of different kin types. To our knowledge, this study is the one most similar to ours, although they did not focus in detail on different kin types, nor did they examine distributions or differences by sex, maternal and paternal kin, or family complexity. Using complex imputation methods combined with the rich PSID survey from the
United States, Daw and colleagues (2016) estimated different kin counts and identified socioeconomic and racial differences by kin type. Family surveys have also been used to study demographic traits of specific kin types, such as grandparents (Leopold and Skopek 2015).

Although surveys are undoubtedly useful for kinship research, they also have drawbacks. In addition to being costly, it is difficult to completely account for how mortality (and the resulting missing links) affects extended kin counts. Even large multi-actor, prospective surveys (Dykstra and Komter 2006; Kalmijn et al. 2018) have difficulty enumerating full kinship networks. In particular, non-response for one kin member undercuts entire branches of kin. The collected number of kinship networks also often falls below what is needed to achieve a statistically reliable age-stratified population because of the high financial and time costs of creating large surveys. Moreover, a strict genealogical definition of who is, for example, a cousin often differs from how people self-identify kinship (Finch 2007; Schneider 1968).

From the 1950s to the 1980s, some researchers asked respondents in interviews to recreate their kinship networks in ego-focused networks (Boholm 1983; Schneider and Cottrell 1975; Young and Willmott 1957). These data allowed researchers to examine whom individuals considered kin and how kinship was delimited. This approach is similar to ethnographic methods in anthropology that have been used to document kinship within small-scale societies, often aiming at the total reconstruction of kinship within smaller study areas.

Although complete empirical counts of kin (including nonresidential and deceased kin) are very sparse, data sources such as censuses provide insight into intergenerational family relationships within households. Census research extends far back into history (Ruggles and Brower 2003) and is available for a diverse set of countries (Ruggles and Heggeness 2008). We thus know more about coresident kin than noncoresident kin.

Available knowledge of kinship structures for entire populations is largely the result of microsimulations in which demographic rates have been used to create populations consisting of virtual individuals (Alburez-Gutierrez, Mason, and Zagheni 2021; Murphy 2011; Verdery 2015; Wachter 1997). This technique applies demographic rates to simulate a virtual population with intergenerational links. Microsimulations have been used to analyze the change in the availability of kin over the demographic transition (Murphy 2011) and to forecast the kinship of a future aging population (Verdery and Margolis 2017; Wachter 1997). The advantages of microsimulation include full kinship networks for simulated populations in the past, present, and future. Murphy (2011) estimated the average number of living kin by age and cohort using simulated British data, producing statistics on kinship similar to those in our study. Despite being state-of-the-art, this approach has limitations. First, although knowledge about kin relations is not a prerequisite,

---

2 For example, Daw et al. (2016) found that the number of cousins decreases substantially between individuals in their 20s and 30s and individuals in their 40s and 50s. This finding is likely due to the effect of the grandparents’ survival and probability of being observed in the survey on the identification of cousins through an individual’s grandparents. This decline by age is not found in microsimulation studies (e.g., Murphy 2011) and our results (discussed later).
the approach requires detailed demographic vital event data by age and sex for long
periods. Few countries have collected such data of adequate quality. Second, the
simulated approaches assume that demographic rates apply equally to all parts of
the population (or to subpopulations that then require rates). In simulations, each
individual has a stochastic rate of events, but the rate is usually identical for all
individuals of the same sex and age, although it can be expanded to other statuses
(e.g., race, socioeconomic status, and civil status) if such subpopulation rates are
available (Ruggles 1993; Verdery et al. 2020). The homogeneity assumption can be
partly relaxed in more complex microsimulation models (e.g., introducing intergen-
erational inheritance), but models require additional assumptions. Consequently,
microsimulation is most successful for estimating population-level averages but
may fail to account for some intricate complexities in kinship structure. Actual par-
ity distributions, mortality, birth spacing, and other demographic factors likely vary
across subpopulations, with high values concentrated in some families and groups
(Ruggles 1993). A different difficulty relates to accurately simulating partnering
processes that cannot be captured only by sex-specific demographic rates. There-
fore, simulations require validation from empirical data, as is often recognized (Post
et al. 1997; Ruggles 1993).

Finally, by applying demographic rates to derive kinship networks analytically,
mathematical demography forecasts the availability of kin who might, for example,
provide informal family care (Alburez-Gutierrez, Kolk, and Zagheni 2021; Himes
1992). Goodman et al. (1974) pioneered this approach in which age-specific rates
are used to derive corresponding kinship patterns. Matrix population models are a
more recent development within this approach (Caswell 2019, 2020). These models, like
microsimulation models, are based on demographic rates rather than empirical micro-
level data. Recent matrix approaches allow for time-varying rates and can estimate
kin frequencies for many kin (Caswell and Song 2021).

For both simulations and mathematical demographic models, some aspects of sta-
tistical variability can be included when estimating kinship (Dudel 2014), although
research has mostly focused on population averages. Other factors originating from
different demographic processes, such as multipartner fertility, are hard to model.
Analytical mathematical models or simulations will always be necessary for estimat-
ing future kinship structures (e.g., Verdery and Margolis 2017).

Descriptions of kinship based on administrative registers have advantages over
surveys, simulations, and analytical mathematical models because they are empirical
calculations based on real population data. Registers include individual-level data
to provide demographic information on entire populations. Biological, marital, or
household membership linkages can be used to construct kinship networks of full
populations. The availability of data for the entire population leaves few concerns
regarding selectivity and statistical power. Registers allow an analysis of heterogene-
ity across groups, such as socioeconomic groups.

Registers have not been used to calculate kin counts across an entire population
as we do in our study. However, researchers have used Nordic, Dutch, and Belgian
registers to study aspects of kinship (Kolk 2017; Lundholm and Malmberg 2009). A
future evolution of register-based kinship research may be to combine modern reg-
isters with historical parish records (cf. Kolk and Skirbekk 2022). In our study, we
introduce a straightforward approach to using registers to map kinship.
Data and Methods

We use the Swedish Total Population Register and the Swedish Multigeneration Register. These registers contain individual-level information on all persons ever registered in Sweden since 1960, including country of birth, sex, birth date, death date (if applicable), and biological mother and father. Our analytical population is all Swedish-born individuals observed in the population registers who were living in Sweden on December 31, 2017 \( (N=8,243,185) \). The analytical sample refers to egos (reference individuals, who serve as the denominator), and we include other kin regardless of whether they survived to 2017 (our numerator). We focus on individuals born in 1915 (aged 102 in 2017) or later; earlier cohorts had very few kin members, making our calculations increasingly unstable. A small number of surviving men occasionally affected our results for the 1915–1920 birth cohorts. For our index cohort, we therefore provide a cross-sectional snapshot of the Swedish biological kinship universe of men and women born in Sweden between 1915 and 2017 who were alive and residing in Sweden at the end of 2017. However, to calculate kinship, we also include all available information in Swedish administrative registers, and our kinship counts therefore include kin who were alive in 2017 and those who died in 2017 or earlier. When relevant, we show both living and deceased kin in our results.

We map the biological kinship network of each individual in the population, linking each person to their parents, children, siblings, grandparents, grandchildren, cousins, aunts, uncles, nieces, and nephews. Family complexity is also introduced in networks where parents or grandparents have multiple childbearing partners—in families with half- and full siblings or families where cousins may share only one grandparent. In these complex families, the observed siblings’ kin network may differ substantially. In Table S2, we show the kin relationship types, the genealogical distance from the ego, and the levels of kin complexity.

We also directly compute the average number of living or ever-registered kin and the distribution of kin across birth cohorts. All kinship links are formed based on available data on biological parenthood. Because social services conducted rigorous paternity investigations, the share of Swedish-born children with missing fathers was less than 1% in the second half of the twentieth century. For cohorts born in the 1930s and early 1940s, the information is of lower quality because digitized parent–child records were based on data on children residing in their parents’ households in 1947 (Kolk 2014). For the relatively small number of children born in same-sex unions, we observe biological parenthood as defined in registers (Kolk and Andersson 2020). Hence, for female same-sex couples, we observe the kinship of the birth mother.

Our study population for calculating kinship networks includes individuals who were digitized by Statistics Sweden: those ever residing in Sweden between 1960 and 2017. This has two important implications. We observed only individuals born or observed within this study period, and the study population is therefore conditioned on survival to early adulthood for those born early in the twentieth century. For example, when we study the 1990 cohort’s number of children, we observe the number of children up to age 27 in 2017 but not their presumably higher eventual number of children. Moreover, our registers are constrained by our observation of parent–child links only after 1932. Given this constraint, we do not observe an individual’s grandparent (or kin traced through the grandparent, such as aunts, uncles, and cousins) if
the grandparent was born before reliable parent–child links are available. Finally, no births outside Sweden were recorded (unless the person remigrated to Sweden with their children). For our study cohorts, Sweden was an immigration destination with quite limited emigration, so the effect of births outside Sweden is rather modest. Missing kin links have larger consequences for second-generation immigrants in our sample because it prevents us from observing information on grandparents. We discuss these issues, additional analyses, and the Swedish register data in much further detail in online appendix 1.

The oldest and youngest cohorts have an incomplete number of particular kin, albeit for distinct reasons. For the oldest cohorts, missing kin stem from pure data limitations: parental or grandparental linkages required to connect different kin types are missing. For the youngest cohorts, not-yet-born kin for individuals alive in 2017 cannot be observed. Full sets of siblings or cousins are acquired over time, and our younger cohorts will eventually have more kin than can be observed in a snapshot of kinship in a single year at an early age.

Vertical lines and shaded areas in our figures indicate our confidence that our cohorts include at least 95% of their eventual kin set. Lighter-shaded areas for younger cohorts highlight those cohorts that had not yet reached the age at which we could expect to observe a complete set of a certain kin type by 2017; lighter-shaded areas for older cohorts represent cohorts that did not have completed sets of kin because information on parental or grandparental linkages required to connect different kin types was limited. We describe this in further detail in online appendix 1. We also show the kin distribution using violin plots: the violin shape shows the entire kin distribution, and the box plot inside shows medians, quartiles, and 5th and 95th percentiles.

Illustrating the logic of constructing kinship from administrative data, Figure S1 shows the number of individuals in Sweden by birth cohort and our effective sample size. In online appendix 1, we explain how examining kinship from the perspective of individuals alive affects our findings. Our study population is Swedish-born individuals. However, over time, an increasing share of them have foreign-born parents. Thus, grandparents and kin connected with grandparents will be missing from the registers, which will be reflected in the results on grandparents for Swedish-born individuals (Figure 7 and Figure S2).

In online appendix 2, we provide a spreadsheet containing the (aggregated) data for the figures in our article to facilitate comparison between our results and other data sources.

**Results**

We present our results in figures showing (1) the average number of kin types and (2) the proportion of those with a certain number of kin (the distribution) for the population in Sweden in 2017, separately for men and women across birth cohorts. The figures are ordered by the generational distance from the reference index individuals, whom we will refer to as egos. We first focus on our (older) egos’ grandchildren and work our way toward descriptions of our (younger) egos’ grandparents. When comparing kin of the same generation (e.g., siblings and cousins), we first present the kin who are genealogically closer to the ego. Figures illustrating the distribution of total
How many grandchildren do people have in their kinship network? The top panels of Figure 1 show the average number of grandchildren for men and women across the 1915–2017 birth cohorts; the bottom panels show the numerical distribution of living grandchildren. In 2017, on average, men born between 1920 (aged 97) and 1940 (aged 77) had roughly 3.6 living grandchildren. One in five of these men had no living grandchildren, approximately half (47.2%) had between one and four living grandchildren, and the rest (32.6%) had five or more living grandchildren. The number of living grandchildren at the time of the study varied starkly among elderly men: one fifth had none, whereas one fifth had six or more grandchildren. The pattern for women is similar. Relative to men, however, women had a slightly higher average number of grandchildren (3.8) and a slightly lower share of women (17.4%) had no living grandchildren in 2017. The number of living grandchildren declined for
individuals born after 1940 because some or all of their grandchildren were not yet born. This decrease occurred among earlier cohorts for men, who tended to become grandparents at later ages than women. Approximately 69.9% of men born in 1960 (aged 58) had no living or observed grandchildren, compared with only 55.8% for women born in 1960. (Figure S8 shows the data in Figure 1 by the grandchild’s sex.)

Turning to the next generation, the top panels of Figure 2 show the average number of living children for men and women across the 1915–2017 birth cohorts, also broken down by the individual’s childbearing partners. The figure shows how much individuals with one, two, three, or more childbearing partners contributed to the average number of (living) children in that cohort. The bottom panels show the numerical distribution of living children across these cohorts. Cohort fertility was
quite stable for the 1915–1950 birth cohorts, at an average of roughly two children per individual, even though period fertility fluctuated strongly throughout the twentieth century. Examining final parity, we find an increasing two-child norm over time, although the broader pattern was stable. For women born before 1977, the proportion childless remained stable, at approximately 12.4%, on average. For men, this proportion increased across cohorts, and at least 21.5% of male cohorts born after 1955 had no children as of 2017. As a result, the average number of children for men born after 1955 declined across cohorts, whereas women across all cohorts had an average of roughly two children by 2017. For cohorts born after 1977, an increasingly large proportion of individuals had not yet completed their fertility by 2017. In addition, the average number of children from parents with two or more childbearing partners increased steadily across the 1915–1945 cohorts; for cohorts born after 1945, the average remained constant. Individuals with two partners contributed substantially to total childbearing, and individuals with three or more partners contributed little. Figure S9 shows fertility by parity and number of childbearing partners for the 1960–1970 cohorts.

Fig. 2 Average number of living and dead children (top two panels) and proportional distribution of the number of living children (bottom two panels), in 2017 by sex and birth cohort: Swedish-born individuals from the 1915–2017 birth cohorts who were alive at the end of 2017. The shaded areas on the right side of the dashed vertical lines indicate birth cohorts for which we are unlikely to observe the complete set of children because they were not yet born.

(continued)
Figure 3 shows the average number of nieces and nephews for the 1932–2017 birth cohorts (we trace parent–child links beginning with the 1932 cohort). For individuals born in 1955, the average number of nieces and nephews peaked at just over 4.13. For those born before 1940, a higher share of parental information is missing, giving us incomplete sibling information and thus a lack of information on nieces and nephews. For cohorts born after 1971, the number of nieces and nephews was not complete in 2017 because some of their siblings may not have completed their childbearing by 2017. Most individuals had nieces and nephews through their full siblings. However, the average number of nieces and nephews from half-siblings increased for later cohorts.

Next, we count kin who were the ego’s generational contemporaries. The top panel of Figure 4 shows the average number of siblings of egos across the 1932–2017 cohorts, and the bottom panel shows the numerical distribution of siblings. The
average number of siblings is around two (as cohort fertility was stable) for cohorts with almost complete information on siblings (1940–2004). Most of these siblings were full siblings, although the average number of half-siblings increased across cohorts, which explains the trends we previously observed for nieces and nephews. The bottom panel of the figure shows an increase in the proportion with one sibling beginning with the 1985 cohort. Again, for the cohorts outside the dashed lines, sibling information is incomplete because of a higher share of missing parental information for those born before 1940, and sibling sets are incomplete for those born after 2004. In Figure S10, we show the exact distribution of the number of half- and full siblings for the 1960–1970 cohorts. Figure S14 shows how widely spaced sibling sets are from the perspective of the 1980 and 1985 cohorts.

In Figure 5, we show the average number of cousins of egos for the cohorts born in 1950–2017 in the top panel and the distribution of cousins across the same cohorts in the bottom panel. Counting cousins requires that the index individual’s grandparents are in the registers so that the fertility of all aunts and uncles can be observed. Thus, we have close to complete information for the birth cohorts born between 1977 and 1998. These cohorts had between 6.58 and 8.09 living cousins, with those born in 1985 having the largest average number of cousins and the lowest share of individuals without any cousins (8.0%). Maternal cousins were about as frequent as paternal cousins. Much rarer than having no cousins was to have only one cousin, whereas more than 61% of those born in 1985 had six or more cousins. We found the number of cousins to be the type of kin that exhibited the largest variations across individuals, with more than 10% of those born in 1980 having more than 20 cousins on average (Figure S4). We show a more detailed version of the top panel of Figure 5.
Fig. 4 Average number of siblings by birth cohort and whether full or half-sibling (top panel) and proportional distribution of the number of siblings (half- or full) by birth cohort (bottom panel): Swedish-born individuals from the 1932–2017 birth cohorts who were alive at the end of 2017. The shaded areas on the right side of the dashed vertical lines indicate birth cohorts for which we are unlikely to observe the complete set of siblings because they were not yet born. The shaded areas on the left side of the dashed vertical lines indicate birth cohorts with incomplete coverage of siblings because of missing parent–child links.
Birth Cohort (age in 2017)

Average Number of Cousins

Deceased cousins  Paternal uncle's side  Paternal aunt's side
Maternal uncle's side  Maternal aunt's side

Fig. 5 Average number of cousins by birth cohort and by type of aunt or uncle (top panel) and proportional distribution of the number of cousins by birth cohort (bottom panel): Swedish-born individuals from the 1950–2017 birth cohorts who were alive at the end of 2017. The shaded areas on the right side of the dashed vertical lines indicate birth cohorts for which we are unlikely to observe the complete set of cousins because they were not yet born. The shaded areas on the left side of the dashed vertical lines indicate birth cohorts with incomplete coverage of cousins because of missing parent–child links.
in Figure S11. In Figures S15 and S16, we show how widely spaced cousin sets are (i.e., the birth years of cousins) from the perspective of the 1970, 1975, 1980, and 1985 cohorts.

Turning to generations younger than the ego, the top panel of Figure 6 shows the average number of mothers and fathers. For those born in 1960 (age 57), more than half had a living mother in 2017; for the 1980 cohort, more than 92% had a living mother. The number for fathers was slightly lower (86%). The figure also provides information on the share of parents who could be traced in the register. We had virtually complete information on parents for all cohorts born after 1950 and very good coverage for individuals born between 1935 and 1950, but an increasing share of missing information for older cohorts. We show the proportion with exactly one or two living parents in Figure S12.

The bottom panel of Figure 6 shows paternal and maternal aunts and uncles alive in 2017. There is little change across cohorts with nearly complete coverage—on average, 3.50 aunts and uncles were alive across cohorts. There were no substantial differences in the average numbers of aunts and uncles by maternal and paternal lineage nor between aunts and uncles. We here began from cohorts born in 1950 because of the necessity of having living grandparents. For cohorts born before 1977 (left of the dashed line), parental sibling sets were more likely to be incomplete because both parental and grandparental information is needed to establish the linkage to parental siblings. Missing grandparental information explains the slight dip for the most recent cohorts: an increasing share of younger persons in our data have foreign-born parents and grandparents. Figure S13 shows the data from the bottom panel of Figure 6 further broken down by half- and full aunts/uncles.

Figure 7 displays the average number of living grandparents of egos born between 1950 and 2017. Those born between 2005 and 2017 had, on average, 3.2 grandparents alive in 2017, and roughly 0.4 grandparents had been registered but died before 2017. In these cohorts, 0.8 (of 4) grandparents, on average, had never been registered in Sweden, most often because they had never lived in Sweden. The average number of living grandparents decreased rapidly for those born before 2005, and almost no grandparents were found for individuals born before 1965. Among older cohorts with nearly complete information on grandparents who had ever been registered in Sweden—such as those born in 1980 and thus aged 37 in 2017—we found that living grandmothers were substantially more frequent than living grandfathers, reflecting women’s greater longevity and tendency to marry older men.

We conclude by focusing on the sum of the kinship network we observed, pooling all kin types from our previous figures in Figure 8. The distributions of the total number of kin alive across the 1920–2010 cohorts are presented in the top panel. The upper tail of the distribution is long because some individuals had extensive kin networks. The figure is truncated at 70; 4,398 individuals across all cohorts in our study had more than 70 living kin in 2017. The median number of living kin was largest for the 1980 cohort (aged 37 in 2017), for whom we have high-quality coverage of both older and younger kin; the median number of living kin was smallest for the oldest cohorts born in 1920 and 1930 (ages 97 and 87). As seen in the bottom panel of Figure 8 (and in earlier figures), the small kinship size in the earliest cohorts is partly due to missing parent–child links in our data. In this panel, we show only living kin; similar data for all ever-observed kin are shown in Figure S6.
Fig. 6 Average number of living, dead, and unregistered parents, by birth cohort 1932–2017 (top panel), and average number of parent siblings by birth cohort 1950–2017 (bottom panel): Swedish-born individuals alive at the end of 2017. The shaded areas on the left side of the dashed vertical lines indicate birth cohorts with incomplete coverage of parents, aunts, and uncles because of missing parent–child links.
The dispersion of the number of kin is very wide. The interquartile range was greater than 10 for individuals born between 1960 and 1990. The distance between the 5th and 95th percentile is roughly 40 kin for the 1970 cohort. The large variation in the 1960 and 1970 cohorts reflects the greater influence of horizontal kin, such as siblings and cousins, among younger cohorts: these kin had more variation than vertical kin, such as children and particularly parents and grandparents. Still, the dispersion is substantial in older cohorts, and it is impossible to identify a standard kinship size among people in their 70s or 80s. A complete lack of kin is rare. The proportion of those with no close kin among the groups we identify was very small for all cohorts from 1970 onward (less than 2%); in the oldest cohort, 14.6% had no living kin.

The average total kin network size by kin type is shown in the bottom panel of Figure 8 for the 1915–2017 birth cohorts. For the earlier cohorts, we lack data because of missing parent–child links for some kin types (particularly cousins before the 1950s and siblings, nieces, and nephews before the 1930s). The average total number of kin alive in 2017 peaked for cohorts born around 1983 (age 35), who had an average of 20 living kin members.

The higher number of kin among individuals in midlife (approximately ages 27–52, or age cohorts 1965–1990) reflects that these members had both younger and older generations of kin alive at the same time (children and parents), and that they had more horizontal kin (e.g., cousins), whereas kin sets were not yet complete for younger cohorts. For the 1983 cohort (the year with the most thorough coverage), cousins
Fig. 8 Distribution of the total number of kin (top panel) and the average number of all types of kin by birth cohort 1915–2017 (bottom panel): Swedish-born individuals alive at the end of 2017 by birth cohort. In the top panel, box plots show quartiles and 5th and 95th percentiles; the top label indicates the percentage of the population with the exact median number of kin for that cohort (width of each curve). In the bottom panel, deceased kin includes all types of kin. The average number of kin for some birth cohorts is influenced by the lack of parent–child linkage for the oldest cohorts (see online appendix 1).
were the most common type of living kin; egos had an average of 8.03 cousins, who made up 40.3% of the full kinship network. The size of the remaining kin network was, in order of magnitude, aunts and uncles, followed by nieces and nephews, and then (more or less equally) parents, siblings, and children. However, if we had better data on cousins for earlier cohorts, the 1930–1960 cohorts’ total kinship networks would have been substantially larger. For younger cohorts, the average number of older kin increased slightly, whereas the number of younger kin decreased significantly. The inverse could be observed for cohorts born before the early 1980s. Note that we undercounted older cohorts’ horizontal kin because of missing parental (and grandparental) links before 1932. The lower number of kin in the 1940 cohort compared with the 1965 cohort reflects such undercounting.

We find that the size of an individual kin network is largely determined by the number of parental siblings and cousins and that these groups create large variations in kinship (Figure 8, top panel). Egos older than 25 have a notable number of deceased kin. The sum of living and deceased kin of younger cohorts is likely a reflection of their true kin size, whereas our registers do not cover some of the oldest kin of older egos (as discussed in online appendix 1). To conclude, most individuals of all ages have a substantial number of living kin in Sweden, but the kin type differs by the individual’s age. The number of kin also varies substantially but is not visible when examining averages.

Discussion

In our study, we conducted extensive empirical documentation of the demography of kinship, counting living kin for the entire population of Sweden in 2017 by year of birth. We documented the extent to which kin availability is shaped by the individual’s age and variation by kin type. Our results illustrate several properties of kinship demographics. First, kin network sizes are highly heterogeneous within birth cohorts. Kinship varies to an extent that is not trivial, self-evident, or previously analyzed in depth. For example, substantial minorities of elderly individuals had no grandchildren, whereas others had six or more grandchildren; cousin networks ranged from three or four individuals to dozens. The interquartile range of individuals’ total kinship size was as high as 10 in midlife. Because our results reveal substantial dispersion in kinship size across all cohorts, we conclude that it is hard to make generalizations about “typical” kinship. Using the statistical of average kin sizes to convey a typical kin network may be misleading. For example, although the mean number of cousins in a population is 8, close to 25% have more than 11 cousins, and 25% have 2 or fewer cousins. Interpreting an average of 8 cousins to mean that most people have 7–9 cousins is incorrect because dispersion is substantial. Dispersion for total kin size is even greater. We have also identified properties that are common denominators of contemporary kinship—that is, the aspects of kinship structure that appear to be shared by most people at a given age. For example, we numerically specified the dominance of horizontal kin (e.g., cousins) in adult kinship networks, and we detailed the younger-and-older kinship structure that characterizes kinship in midlife. We found the greatest variance for cousins and grandchildren, whose frequencies rely on the fertility behaviors of multiple generations.
Moreover, we found that the extent of complex relations usually not fully estimated in other accounts of kinship (e.g., half-siblings and parents of half-siblings) was nonnegligible and that the importance of kinship arising from multipartner fertility increased over the period we studied. Finally, by distinguishing between kin linked to the ego through paternal and maternal sides, we revealed the extent of numerical differences based on lineage across kin types. We found no lineage differences among nieces and nephews or aunts and uncles, and that the overall, contemporary kinship structure in Sweden is balanced across lineages, with the exception of more (living) maternal than paternal grandparents. Our results are specific to Sweden. However, because mortality and fertility in Sweden and many other high-income countries in the twentieth and twenty-first centuries are similar, many of our results can probably be generalized to demographically similar contexts. One exception is the relatively higher fertility in Sweden since the 1990s compared with some Southern European, Central European, and East Asian countries, which may produce more horizontal kin in higher fertility contexts, such as Sweden. Swedish fertility levels have been closer to those of the United States, France, and the United Kingdom. We can expect kinship patterns in developing countries to be entirely different and more affected by, for example, the more recent fertility transition, unlike the more stable twentieth-century Swedish mortality and fertility patterns.

In addition to presenting an empirical base for future research in kinship demography, our findings are relevant for research on the social significance of kinship. The distribution of living kin in the population represents a black box in the burgeoning research on the effects of kin (e.g., on educational outcomes, wealth, and emotional well-being), on what kin members do (e.g., frequency of interaction, provision of help), and the nature of the social relations ascribed to kin (e.g., perceived connectedness, closeness, and conflict). For example, transfers of care and resources across kin are all ultimately conditioned on the existence of a particular kin relation, which is determined by the demographic constraints given by the kinship structure (cf. Agree and Glaser 2009). Examining issues such as how much care a person can expect to receive from kin is conditioned on the individual’s kinship structure.

This study also demonstrated the usefulness of using administrative data to examine kinship demography. Although linked intergenerational administrative register data have been readily available for decades, this study was the first to use these data for comprehensively counting kin. By doing so, we obtained statistics (measures of dispersion and measures of kin linked through only one parent) that would be difficult to achieve with other approaches. We also demonstrated the limitations and pitfalls of using administrative data (e.g., the implications of truncated parent–child links), which is instructive for future research and important for highlighting areas where survey, register, and simulation approaches to kinship demography are preferable. In particular, we demonstrated the importance of delimiting plausible age and cohort spans for which register information may contain as-good-as-full coverage across generations of kin.

Our findings provide an empirical opportunity to validate kinship calculated from microsimulations and analytical approaches (Caswell 2020; Goodman et al. 1974; Ruggles 1993; Zagheni 2015), which is the only option for analyzing kin in many countries. Kin relationships of intact families can be reasonably well approximated by analytical models and microsimulation models: because birth intervals and age-specific
fertility follow predictable patterns, we expect analytical/simulation approaches to perform better for vertical kin relationships. However, union instability and multipartner fertility create very different demographic patterns (e.g., longer birth spacing and more variance in age at birth) that will diverge substantially from more stable family constellations. Our empirical data could help evaluate whether this divergence impacts kinship frequencies calculated from a rate-based approach. For example, our results could be compared with those of matrix approaches that either use only female rates or use male and female rates separately to estimate kinship (cf. Caswell 2022).

A complex factor not addressed in our study because of a lack of cohabitation data is fully accounting for in-law relationships through marriage or partnership, which is complicated by the lack of data on cohabitation in Sweden; including only married couples would have ignored a substantial share of union formation. This task remains important for future empirical research on kinship. Similarly, register data might be used to examine differences in kinship by socioeconomic status—differences that are hard to study with other methods of studying kinship. Another interesting dimension of kinship research possible with micro-level data (both administrative data and microsimulation) is studying pairwise attributes of kin (e.g., the prevalence of a kin type of a certain age) for egos of a given age.

To conclude, we argue that demographic characteristics of kinship have been neglected in empirical social research. We addressed this research gap by providing a full assessment of kinship links in contemporary Sweden—a complete accounting that has never been calculated for any national population. We identified important kin categories, including dimensions of family complexity, that are hard to assess with the analytical and simulated approaches that have been used to study this type of question. We hope that our work inspires future research quantifying kinship and serves as a reference point for the rich empirical research that examines how kin continues to play an active role in the lives of men and women in contemporary societies.

Acknowledgments We are grateful for the financial support from the Swedish Research Council for Health, Working Life, and Welfare (FORTE, grants 2016-07115 and 2016-07099), the Swedish Scientific Council (grants 2019-02552 and 2022-02314), the Academy of Finland (grant 321264), and the Global Challenges Foundation grant “Sustainable population in the time of climate change.” We acknowledge helpful comments from Diego Alburez-Gutiérrez and Tom Emery, and from two anonymous reviewers.

The study was approved by the Swedish Ethical Review Authority under the project “Ageing Well.” Access to underlying data requires approval of the Swedish Ethical Review Authority because the study used micro-level data of the complete Swedish population.

References


---

Martin Kolk (corresponding author)

martin.kolk@sociology.su.se

Kolk • Demography Unit, Department of Sociology, Stockholm University, Stockholm, Sweden; Institute for Future Studies, Stockholm, Sweden; Åbo Akademi, Vaasa, Finland; https://orcid.org/0000-0002-7175-4040

Andersson • Swedish Institute for Social Research, Stockholm University, Stockholm, Sweden; Department of Sociology, Turku University, Turku, Finland; https://orcid.org/0000-0002-0347-3802

Pettersson • Division of Insurance Medicine, Department of Clinical Neuroscience, Karolinska Institutet, Solna, Sweden; https://orcid.org/0000-0002-4776-9162

Drefahl • Demography Unit, Department of Sociology, Stockholm University, Stockholm, Sweden; https://orcid.org/0000-0003-4533-7558