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, Sven Drefahl

Stockholm
University

Demography Unit

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[1], Linus Andersson[2], Emma Pettersson[3], Sven Drefahl[4]

[1] Demography Unit, Department of Sociology, Stockholm University; Institute for Future Studies, Stockholm; Åbo Akademi, Finland.

[2] Swedish Institute for Social Research (SOFI), Stockholm University; Department of Sociolgoy; Turku University;

[3] Division of Insurance Medicine, Department of Clinical Neuroscience, Karolinska Institutet

[4] Demography Unit, Department of Sociology, Stockholm University

Abstract

We know surprisingly little about the demography of human kinship. In the current study, we give a demographic account of the kinship networks of individuals in 2018 in Sweden across sex and cohort. We used administrative register data of the entire Swedish population in order 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/uncles, nieces/nephews, and cousins. We show both the average number of kin of different types and the distribution of the number of kin as well as how dispersion has changed over time. We show trends for matrilineal and patrilineal kin and also show differences in the kinship structure arising from fertility with more than one childbearing partner, such as half siblings. The results demonstrate extensive variability as well as homogeneity in kinship structure. We discuss our findings in the context of other methods to estimate kinship.

Keywords: kinship, Sweden, extended family, grandparents, cousins, grandchildren

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 other support to each other. Grandparents often provide childrearing, financial, and emotional support to younger generations. Grandchildren and adult children maintain care and social contact vital for the wellbeing and sustenance of the elderly. Individuals' closest confidants often include their siblings. The network of aunts, uncles, and cousins maintains a constant presence, and these are among the few relationships which tend to last throughout a person's life. Whereas social scientists of the 20th century have had a relative disinterest in kinship in developed countries, more recent research has recognized that individuals are far from decoupled from extended kin (Bengtson 2001; Furstenberg 2020; Mare 2011). Intergeneration kinship relations have become more common as increased longevity extends the frequency of kinship over three or more generations, and as increasing dependency-ratios reshape the conditions for public transfers between cohorts and private transfers across age groups (Lee and Mason 2011).

Despite the abundance of evidence showing the importance of kin, researchers know comparably little about the actual frequencies of different types of kin in contemporary populations (Daw, Verdery, and Margolis 2016). Research on extended family members has focused primarily on the roles and behaviors of extended kin with regard to each other. The few current efforts to enumerate siblings, cousins, aunts/uncles, grandparents, and other kin are drawn from survey samples, mathematical demography, and micro-simulations. These approaches rely on strong assumptions, and the accuracy of estimated kinship networks derived from such methods is difficult to validate.

In this study, we provide the first-ever numeration of kinship networks using observed empirical data for all ages for an entire population. We have avoided several limitations of previous approaches by inferring kinship counts empirically from the total population of Sweden and 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 novel granular descriptions beyond the reach of current state of the art simulation and mathematical models, including kinship traced through full and half siblings, children by partner order, separate kin counts for the maternal and the paternal sides, and all genealogical types of kin. Whereas previous research was largely limited to predicting current population or sub-population averages, we also describe the variability and distribution of the number of total kin. Specifically, we compared the living siblings, cousins, aunts/uncles, grandparents, parents,

and grandchildren across birth cohorts of individuals alive in 2017. Our approach combines micro-level longitudinal data necessary to calculate kinship links across multiple generations with the advantages of coverage and accuracy from studying an entire population at one point in time.

There is much to gain from descriptively enumerating kinship. Researchers have demonstrated multiple and multifaceted ways in which kin members provide support, cause burdens, or otherwise shape everyday life (Furstenberg 2020). Yet, the potential for an individual to be influenced by kin-members is ultimately dictated by the availability of kin, which is, as we will show, highly variable in the population. Given the importance of kinship structure in the provision of family care, kinship networks may also influence the overall demand for formal and informal care. It informs policies by clearly showing the incidence of kinlessness, who can and cannot rely on grandparental care, and the distribution of working-age adults available in an individual's kinship network (Agree and Glaser 2009; Szebehely 2005). Moreover, our estimates will likely also provide opportunities to validate future simulation modeling of kinship structures and can be used as benchmarks to verify output from such models.

Our study is aimed to provide some general knowledge on the relatively uncharted field of kinship demography. Demographic changes over the past century have brought about changes to the structure of kinship that will continue in the future. Rapid increases in life expectancy over the 20th century have extended the likelihood that people will interact with vertical kin, such as parents, grandparents, and grandchildren (Bengtson 2001; Murphy 2011). The kinship network of a population is the aggregate 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, and aunts/uncles. This has been described as a beanpole kinship structure—a network of kin with a "stem" reaching as high as three or four generations but where each member has comparably few "branches" of generational contemporaries (Bengtson 2001; Uhlenberg 1996).

Demography drives kinship structure through more mechanisms than life expectancy and average number of children, and the kinship structures have many other dimensions than generational overlap (Murphy 2011). Changes in mean age at childbearing can also increase or decrease the number of generations subsequently alive. Sex differences in fertility influence the degree to which cousins, nieces, or nephews are found primarily on the side of the mother or on the side of the father, particularly if there are differences in parity distributions due to multi-partner fertility between men and women. As an example, the likelihood of anyone

having two living cousins depends on the cousins' longevity as well as the fertility of a second generation (i.e., fertility of aunts/uncles determines the number of cousins) and a third generation (i.e., fertility of grandparents determines the number of aunts/uncles). It is also based on when in life the second and third generation had their children The likelihood of having a paternal grandmother alive at about age 50 is decided by sex-specific longevity as well as the generational overlap produced through the sex-specific mean age childbearing of the grandparent and parent, showing that these factors will differ for matrilineal and patrilineal kin. Multi-partner fertility will lead to more widely spaced births, siblings, and cousins (Andersson 2020, 2021), which will affect the social relationships among horizontal kin. In sum, the shape of an individual's kinship network at a given period and age is the function of the demographic behaviors of multiple generations and pathways that are hard to capture with population averages of age-specific rates. Some aspects, such as parity-dependent birth spacing, multi-partner fertility, assortative mating, and age differences between partners, are particularly hard to estimate just from age-specific vital rates.

In the following sections, we give an overview of kinship in the social sciences and place the Swedish kinship system in this context. We then present and discuss previous attempts to enumerate kin and discuss the benefits and caveats of the different applied methods. Finally, we explain the data source, its limitations, and how we have measured kinship.

Kinship research in the social sciences

Early research on kinship occurred primarily within anthropology. Lewis Morgan considered changes in kinship to be a central dimension in the emergence of complex societies (Morgan 1997[1870]). Kinship continued to be at the center of social science throughout the late 19th and early 20th centuries (e.g. Malinowski 1982[1932]; Rivers 1968[1914]), and it remained central well into the 1960s. Anthropologists such as Murdock and Radcliff-Brown focused on examining kinship from a perspective where extended kin was described via successions of parent–child dyads and resulting lineages (such as clans or dynasties), whereas Levi-Strauss (1969[1949]) emphasized the role of marriage as a basis of alliances between individuals. Kinship research of the 19th century and functionalist theory of the 20th century emphasized concepts of modernization and saw marginalization of extended kin as a key facet that distinguished modernity from pre-modern societal structures. 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 (e.g. Goode 1963; Palloni, Hill and Aguirre 1996) described the contemporary family as 'structurally isolated' from other

domains in society and emphasized the centrality of the 'nuclear family.' This notion was reproduced in later reformulated individualization and modernization theories (Giddens 1993; Habermas 1985; Thornton 2005). Outside of evolutionary anthropology, contemporary anthropological studies on quantitative kinship composition and kinship systems remain comparatively rare (Furstenberg 2020; Johnson 2000; Mare 2011); however, there has been an increase in quantitative and analytical work in the social sciences connecting to the rich anthropological historical literature (e.g. Lansing et al. 2017; Schulz et al. 2019). In contrast, there is a sizable body of empirical literature in the areas of demography, gerontology, and sociology regarding dyadic kinship relations and their effects on individual wellbeing for grandparenthood and intergenerational co-residence (e.g. Margolis and Verdery 2019; Ruggles 1987; Ruggles and Brower 2003).

In family sociology, the majority of research on kin outside the nuclear family has typically focused on their role in instrumental support exchange across generations or tried to measure the intensity or affection in the relationship (e.g. Albertini, Kohli and Vogel 2007; Bengtson 2001; Heady and Kohli 2010; Rossi and Rossi 1990). Studies have focused on various forms of reciprocity and contact patterns, for example, between siblings (Hank and Steinbach 2018). Parent–child ties, especially the mother–daughter tie, have been described as particularly enduring and strong (Hagestad 1986). Even though kinship in most of Europe is mostly cognatic or bilateral (kinship is recognized along both maternal and paternal lines), researchers have consistently shown a female bias in kin relationships where women and maternal lines are described as "stronger" " (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).

Family sociologists and demographers have shown that grandparents provide substantial support through childrearing, especially for single parents. Studies in gerontology on the relationships between adult children and their aging parents suggest that quality and frequency of contact play an important role in the wellbeing of senior citizens (Brubaker 1990; Chudnovskaya and Kolk 2017; Dykstra et al. 2006; Michielin and Mulder 2007; Tomassini et al. 2004). The growing vertical depth of kin characteristics of the beanpole kinship structure has been hypothesized to strengthen bonds across generations (Bengtson 2001). More indisputable universal conditions of changing kinship structures include the postponement into older ages of bequests and the experience of bereavement. Some researchers have speculated that contemporary societies are in a historically unique situation in which the demographic

availability of grandparent–grandchildren relationships is higher than at any other point in history, and likely higher than in the future, due to changes in the timing of fertility and age at mortality (Wachter 1997).

Traditionally, research on kinship's influences on family members has focused on relationships within an individual's family of origin, relationships within their family of destination, or variations in non-nuclear family-type household compositions, thus ignoring more extended kin (Coall and Hartwig 2010; Kaptijn et al. 2010). However, there has recently been growing evidence that kin relationships outside the household are important predictors of a large number of socioeconomic and demographic outcomes, such as educational attainment, occupational attainment, and wealth the 2000s, social stratification research has expanded the focus from parent–child relations to include kin members outside the immediate nuclear family (Mare 2011; Pfeffer 2014). Some of this research has had a clear kinship focus by studying the association between social status, fertility, and lineage survival. Other work emphasized the roles of socializing and the intermittent effect living kin have on supporting upward (and preventing downward) social mobility. Although the importance of kin relations for social stratification beyond the immediate family remains disputed (Engzell, Mood and Jonsson 2020), this literature has brought extended kinship to the forefront. Recently, sociologists, economists, and demographers have been tracing larger and larger kinship networks to examine how the social background going back for many generations affects life chances (Colladoa, Ortuno-Ortın and Stuhler 2020; Hällsten and Kolk 2020; Modalsli 2016), focusing on great-grandparents and cousins many times removed.

A broad summary of the attention to kinship and family relationships outside the nuclear family illustrates that researchers have focused nearly all attention on what kin "do" and what the "effect" of kinship is on various outcomes as well as contrasting ties/obligations across different types of kin (Daw, Verdery and Margolis 2016). Only a very limited amount of research has examined who kin "are" or produced reliable demographic statistics on frequencies and distributions (Daw et al. 2016).

Kinship in Sweden

Kinship in contemporary Sweden is broadly similar to the northwestern European marriage and family formation patterns (Hajnal 1982; Reher 2005)(Hajnal 1982; Reher 2005). These are traditionally characterized by neolocal household formation (Laslett 1977), low co-residence with parents, late marriage and childbearing, and a large proportion of men and women remaining unmarried and childless throughout their lives (Boholm 1983). Kinship in

Western and Northern Europe have been contrasted with the Mediterranean kinship pattern in which patrilineal relationships are more pronounced in kinship and household formation (Reher 2005), Eastern European kinship with earlier marriage ages and more complex households (Hajnal 1982), and developed countries in East Asia (e.g., South Korea, Taiwan, and Japan) which retain strong patrilineal and patrilocal characteristics (Martin 1990).

Intergenerational relationships and contact in contemporary Nordic countries have been described as less intense (but more universal) than in many other European countries (Albertini et al. 2007). Some scholars have described the relationship between the individual and the state as one of trust in public institutions along with extensive welfare schemas that relieve individuals from caring obligations 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).

In the next section, we summarize the previous literature that has used various approaches to quantify kinship in contemporary populations using data aiming at representative and full pictures.

Previous research quantifying frequencies of kinship

Previous research has estimated 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 current kinship structures at a given time (Bonvalet and Lelièvre 2016; Dykstra and Komter 2006) and to examine issues of heterogeneity in kinship (Goldstein 1999; Goldstein and Warren 2000). Specially designed surveys are currently the most reliable data sources that both apply observed empirical data and aim to survey representative population samples. These include: the Netherlands Kinship Panel Study; the Parents and Children in the Netherlands (Ouders en Kinderen in Nederland); and to a lesser extent are the international Gender and Generation Survey (GGS); the SHARE surveys of old individuals; and complex surveys, such as the Panel Analysis of Intimate Relationships and Family Dynamics (PAIRFAM) in Germany and the Panel Study of Income Dynamics (PSID) in the United States.

The Netherlands Kinship Panel Study was used by Dykstra and Komter (2006) to give an overview of types of kin for 8,150 individuals aged 18–79 who were 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 around 30 kin alive at different ages (with a peak in mid-adult ages), and enumerated the mean number of different kin types. This is, to our knowledge, the most similar study to ours, though they did not focus in detail on different types of kin, nor did they examine distributions or differences by sex or, for example, maternal and paternal kin or family complexity. Using complex imputation methods combined with the rich PSID survey from the United States, Daw, Verdery, and Margolis (2016) estimates kin counts for different groups to enumerate different kin counts and gives socioeconomic and racial differences by kin type. Family surveys have also been used to study demographic traits of certain types of kin, such as grandparents (Leopold and Skopek 2015).

Although surveys are undoubtedly useful for kinship research, they also have several drawbacks. In addition to being costly, it is difficult with survey collection techniques to completely account for how mortality (and the consequential missing links) influences extended kin counts[1]. Survey data is also commonly collected for fixed numbers of kin (e.g., the oldest sibling or the youngest grandchild). With such ego-centric data, it is only possible to create a representative picture for the sampled individuals. Even with large multi-actor, prospective surveys (Dykstra and Komter 2006; Kalmijn et al. 2018), it is difficult to enumerate full kinship networks. In particular, partial non-responses or non-reports of one kin member often undercut an entire branch of kin; inflow and outflow to the kin pool through deaths and births are often unaccounted for because of attrition and because they occur after the last date of interview. The total number of constructed kinship networks also often falls below what is needed to achieve a statistically reliable age-stratified population as it is very expensive and time-consuming to create large surveys. When surveying kin outside the nuclear family, such as cousins, it seems that a strict genealogical definition of who is a cousin often differs from how people self-identify kinship (Finch 2007; Schneider 1968).

Between the 1950s and the 1980s, some researchers engaged in ambitious data collection efforts in which they asked respondents to recreate their own kinship networks in ego-focused networks (Boholm 1983; Schneider and Cottrell 1975; Young and Willmott 1957). These data allowed researchers to examine who individuals themselves thought were kin and

---

[1] For example, in Daw, Verdery, and Margolis (2016), the number of estimated cousins drops dramatically between individuals in their 20s and 30s and individuals in their 40s and 50s, which is likely because the identification of cousins through an individuals' grandparents is affected by the chance of the grandparents' survival and being observed in the survey(compare with microsimulation studies such as Murphy (2011) and our later results).

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. However, such heavily labor-intensive research data collection has been very uncommon in more recent research, and self-reported data are unlikely to give complete information on kinship for more distant kin in contemporary societies.

Although complete empirical counts of kin (also including both non-residential and deceased kin) are very sparse, the structure of data sources such as censuses means that we know much more about intergenerational family relationships within households. This research also extends far back into history (Ruggles and Brower 2003) and is available for a diverse set of countries (Ruggles and Heggeness 2008). Our knowledge of co-resident kin goes far beyond our knowledge about kinship in general, but as that is not the focus of our paper, we will not discuss this literature at length.

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; Wachter 1997; Verdery 2015). These virtual individuals are subjected to a stochastic set of possible events over long periods of time, which creates data of simulated kinship links between these virtual individuals that can then be used to examine how demographic availability of kin has changed over the demographic transition (Murphy 2011) or to forecast kinship of a future aging population (Wachter 1997; Verdery and Margolis 2017). This approach has many attractive features, including full information on kinship networks for simulated populations in the past, present, and future. Murphy (2011) used this approach to give a comprehensive overview of the average number of living kin by age and cohort using centuries of simulated UK data, producing statistics on kinship examine similar dimension as our study our results. Despite being 'state of the art', this approach has obvious limitations. One important drawback is that while knowledge on kin relations is not a necessary prerequisite, the approach still requires detailed demographic vital event data by age and sex for long historical periods of time. Only very few countries have collected such data in adequate quality. A second important drawback is that the simulated approaches assume that demographic rates apply equally to all parts of the population, meaning that although each individual has their own stochastic rate of events, that rate is identical for everyone of the same sex and age, and in some cases, this includes civil status and parity (Ruggles 1993). As a consequence, the approaches can be successfully used to estimate population level averages but fail to account for most of the intricate complexities

of how kinship networks systematically differ between individuals. For example, parity distributions, mortality, birth spacing, and other demographic factors likely vary across sub-populations, with high values concentrated in some families and groups (Ruggles 1993). A different difficulty relates to accurately matching partnering processes that cannot be captured only by demographic vital events of single individuals. Therefore, simulations require validation from empirical data, as is also often recognized (Post et al. 1997; Ruggles 1993).

Finally, demographic rates can be used to analytically derive information on kinship networks. Using mathematical demography, this approach has been used to forecast the availability of kin to, for example, provide informal family care (Alburez-Gutierrez, Kolk and Zagheni 2021; Himes 1992). Goodman, Keyfitz, and Pullum (1974) pioneered this approach in which age-specific rates are used to derive corresponding kinship patterns. A recent development of this approach is matrix population models (Caswell 2019, 2020). These approaches, like microsimulations models, are based on demographic rates rather than empirical micro-level data. Both simulations and mathematical demographic models focus on total population averages and do not provide information on heterogeneity in kinship size or kinship structure of a population. To estimate future kinship structures, some version of mathematical demographic tools and simulations will of course always be necessary (e.g. Verdery and Margolis 2017).

Descriptions of kinship based on the administrative register, as we introduce in this paper, have clear advantages over surveys, simulations, and analytical mathematical models due to being empirical calculations based on real population data. Registers use individual-level data to provide demographic information on entire populations. Depending on the characteristics of administrative data, biological, marital, or household membership linkages can be used to construct kinship networks of full populations. As data exist on the entire population rather than a sample, there are few concerns regarding selectivity and statistical power. Registers allow an analysis of heterogeneity, an important topic for kinship studies. Importantly, individual-level observational data of full populations allow researchers to reveal measures of variance rather than only population means.

Registers have not been used to calculate kin counts across an entire population in the way we have in this article, but some researchers have used Nordic, Dutch, and Belgian registers to study some aspects of kinship (Kolk 2017; Lundholm and Malmberg 2009). In our study, we introduce a straightforward approach to using registers in order to map kinship. We link multiple generations through birth registers to construct full kinship networks of parents, grandparents, full and half siblings, nieces and nephews, aunts and uncles, cousins, children,

and grandchildren. This combines the population-scale output of microsimulation with the detailed information of kin sometimes found in surveys while circumventing most causes of measurement error inherent in both those approaches.

Calculating kinship from Swedish register data

This study has made use of the Swedish Total Population Register and the Swedish Multigeneration Register. These registers contain individual-level information on all persons who have ever been registered in Sweden since 1960, including data on each individual's personal identification number, their country of birth, sex, birth date, death date (if applicable), and the personal identification numbers of their biological mother and father. The analytical population of interest in this study is all Swedish-born individuals observed in the population registers who were living in Sweden on December 31st, 2017 (N = 8,243,185). In our analyses, we focused on individuals born in 1915 and after (aged 102); earlier cohorts had very few members, and our calculations became increasingly unstable (for men, this also occasionally affected our results for cohorts between 1915 and 1920). For our index cohort, we therefore gave a cross-sectional snapshot of the complete Swedish kinship universe of men and women born in Sweden between 1915 and 2017 and who were alive and residing in Sweden in 2017. However, to calculate kinship, we also included all available information in Swedish administrative registers, and our kinship counts therefore include both kin who were alive in 2017 and those who had died in 2017 or before. When relevant, we showed both living and deceased kin in our results. Swedish digitized registers include individuals present in Sweden at any point after 1960. Later, we discuss the implications of this for our kin enumerations of older cohorts.

Provided that the information on parental identification numbers is complete, it is possible to map the entire 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. Different levels of family complexity are also introduced in networks where parents or grandparents have multiple child-bearing partners; for example, families with half and full siblings or families where cousins may only share one grandparent. In these complex families, the observed kin network of siblings may differ substantially. The table below presents the different types of kin relationships mapped in this study and shows genealogical distance from the reference individual, along with the different levels of kin complexity we address in the study.

Table 1: Genealogical distance of kin links and levels of family complexity

Kin type	Generation from ego	Nodes from ego	Genetic relatedness	Different types examined	1st decile ±	median ±	9th decile ±	At least one registered kin (share)[*]
grandchild	-2	2	1/4	through son/daughter	0	0	4	0.23
child	-1	1	1/2	child from 1st/2nd/3rd+ (childbearing) partner	0	1	3	0.52
sibling's child	-1	3	1/4	through full/half sibling brother/sister	0	1	6	0.52
sibling	0	2	1/2	full sibling/maternal half sibling/paternal half sibling	0	2	4	0.88
cousin	0	4	1/8	through maternal/paternal grandmother/ grandfather	0	3	12	0.60
parent	+1	1	1/2	mother/father	2	2	2	0.97
parent's sibling	+1	3	1/4	through mother's/father's full/half sibling	0	2	6	0.63
grandparent	+2	2	1/4	maternal/paternal grandmother/ grandfather	0	4	4	0.66

Notes: *For all cohorts in our dataset—for many kinship types, this is affected by data structure and kin availability. ± values refer to data used in the study, and refers to all cohorts.

Combining these kinship links with the information on birth and death dates from the registers, we could directly calculate the number of living or registered kin, of any type or complexity, for any individual in the population at any given time, though this study primarily focused on the number of living or ever-registered kin individuals had as of 2017. Using this data, the average number of living or ever-registered kin as well as the distribution of kin across birth cohorts could be computed directly. All kinship links are formed based on available data on biological parenthood and grandparenthood. Swedish government records have a very high

share of biological fatherhood observed, partly due to rigorous paternity investigations by the social services. The share of Swedish-born children with missing fathers is less than 1% in the second half of the 20th century. Only for cohorts born in the 1930s and early 1940s is the information worse, as digitized parent–child records were constructed from children residing in their parents' households in 1947 (Kolk 2014). For the relatively small number of children born in same-sex unions, we observed biological parenthood as defined in registers (Kolk and Andersson 2020). This means that for female same-sex couples, we observed the kinship of the birth mother.

Our study population used for calculating our kinship networks was conditional on the population having been digitized by Statistics Sweden, which means they must have been residing (and alive) in Sweden at some point between 1960 and 2017. This has two important implications. First, we only observed individuals who were born within our study period. Therefore, when we, for example, studied the number of children of the 1990 cohort, we observed the number of children up to age 27 in 2017, but not their presumably higher eventual number of children. This is an implication of creating a snapshot of kinship in 2017. Second, our registers were constrained by the fact that we only observed parent–child links after 1932. This means that we did not observe individuals' grandparents (or kin traced through them, such as aunts/uncles and cousins) if their grandparent was born before we had reliable parent–child links. That we only observed individuals who were alive in 1960 or later also involves conditioning on survival to early adulthood for those born early in the 20th century. As our data were based on Swedish register data, all events that occurred outside Sweden were not recorded (unless the person re-migrated to Sweden with their children). For our study cohorts, Sweden was an immigration destination with quite limited emigration, so this effect was rather modest, though it does mean we undercounted extended kinship to some extent if an intermediate generation moved outside Sweden. It had larger consequences for the second-generation immigrants in our sample, as it caused us to miss information on grandparents. We discuss these issues in further detail, including additional analyses and a more detailed description of Swedish register data, in supplemental text part 1.

In our results, we provided vertical lines and shaded areas to give the reader a visual cue for where we are certain that our cohorts included 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 a complete set of a certain type of kin by 2017; lighter shaded areas for older cohorts represent cohorts that did not have completed sets of kin because of limited information on parental or grandparental linkages required to connect different types

of kin. Whether a full set of kin was observable for a certain cohort depended on two factors. The first factor is that it takes a certain amount of time before a given kin set can be completed, and our younger cohorts will eventually have more kin. The second factor is the availability of parental or grandparental information, which is needed to link individuals and form kin networks. We have shown results of the distribution of kin using violin plots, where the violin shape shows the full distribution of kin and the box plot inside shows medians, quartiles, and 95% percentiles.

To better understand the logic of constructing kinship from administrative data, Figure 1 shows the number of individuals in Sweden by birth cohort. The age of individuals in these birth cohorts in 2017 is given in parenthesis. The red line shows the birth cohort size of all individuals ever recorded in the registers (i.e., present at any point after 1960), which includes individuals who have immigrated to Sweden (and often also emigrated). The yellow line represents cohort sizes of all Swedish-born individuals ever observed in the registers. We focused on Swedish-born individuals as we nearly always observed parents in national registers, which is necessary for most kin calculations (we could not trace kin who had never resided in Sweden). Fluctuations in cohort size among the Swedish-born individuals primarily mirror fertility trends, where the lower fertility during the Great Depression is visible, followed by a brief baby boom in the 1940s, as well as post-war fertility cycles, most marked in the 1980s and 2000s (low) and 1990s (high). Our analytical sample is represented by the green line, which additionally conditions the population on being alive at the end of 2017. By limiting our analytical sample to people alive in 2017, we maintained an increasingly select group of individuals with high longevity for cohorts born in 1940 and earlier. The blue line represents those individuals who were alive at the end of 2017 and had two Swedish-born parents. We did not observe anyone with two Swedish-born parents before 1932 because we lacked parent–child links, as discussed in the data section. For later-born cohorts, this mostly reflects that a large share of Swedish-born children had one or two immigrant parents. This will be reflected in results where we show data on grandparents for Swedish-born individuals (figure 9 and figure S1) .
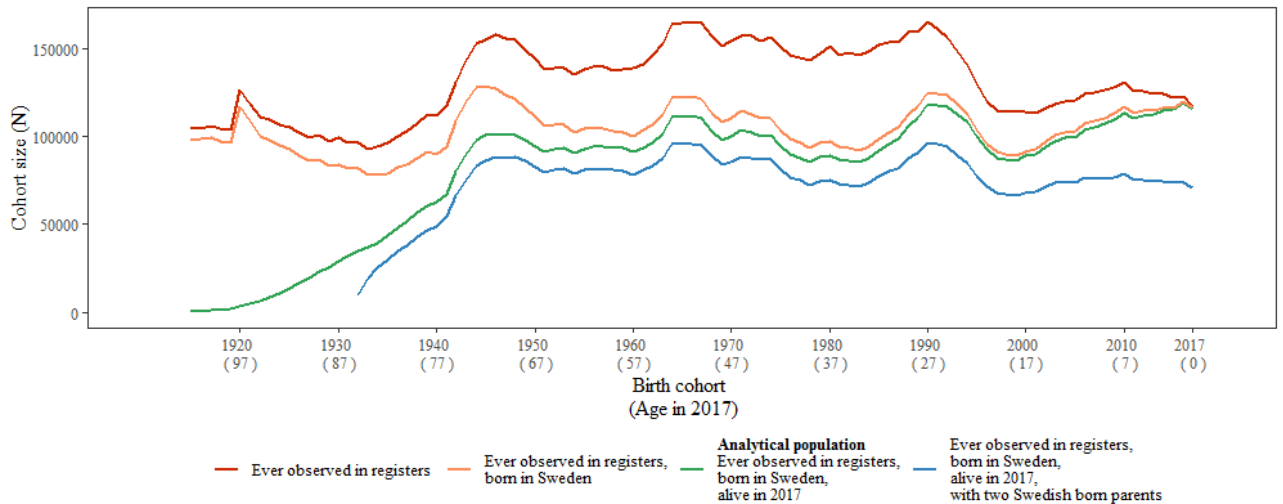
Figure 1: Population by cohort in Sweden using different definitions (1915–2017).

Results

Our results are presented via a series of figures showing (1) the average number of different kin types and (2) the proportion of those with a certain set of kin types among the living population in 2017. We did this separately for men and women and across all relevant birth cohorts. The figures were ordered by the generational distance from the reference index individuals, whom we will refer to as egos. We first focused on the grandchildren of our (old) egos and worked our way toward descriptions of the grandparents of our (young) egos. When comparing kin of the same generation (e.g., siblings and cousins), we first present the kin who are biologically closer to the ego. Figures illustrating the distribution of total kin are shown at the end of the section. Completed sets of kin were generally not observable across all birth cohorts for kin more than one generation away from the ego. The shaded areas in the plots show which cohorts we believe have incomplete kin sets. We also show the distribution of the number of kin in violin plots in supplemental figure S1–S5.
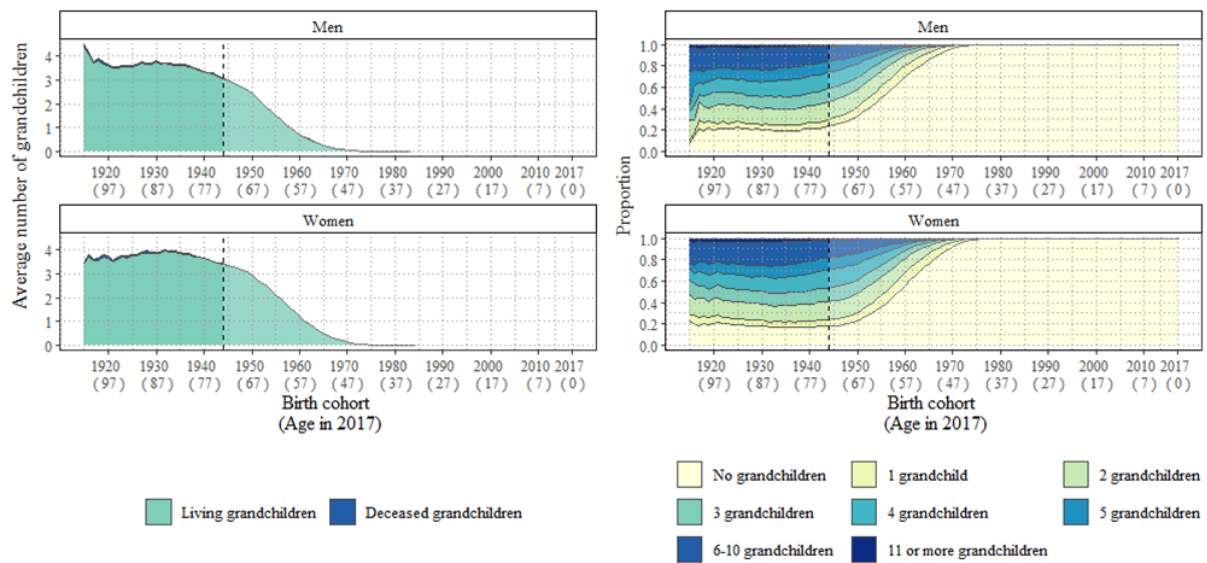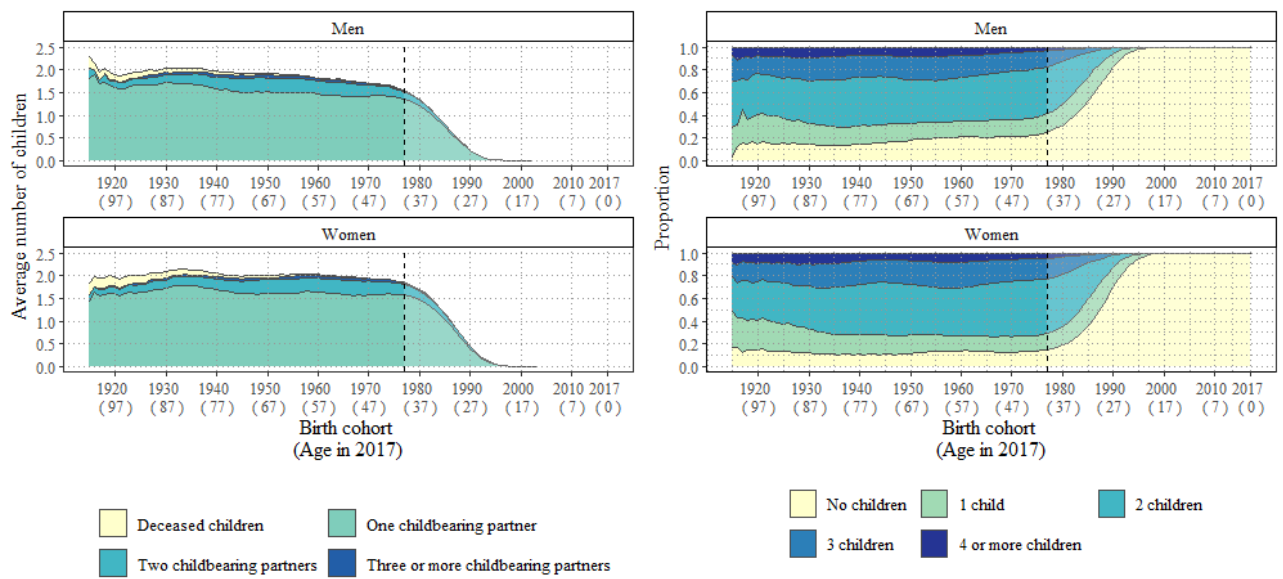
Figure 2a: Average number of living and dead grandchildren in 2017 by sex and birth cohort 1915–2017. Figure 2b: Proportional distribution of the number of living grandchildren in 2017 by sex and birth cohort 1915–2017. Swedish-born individuals alive in 2018.

How many grandchildren do people have in their kinship network? The left panel of Figure 2 shows the average number of grandchildren for men and women across the 1915–2017 birth cohorts. The right panel shows the numerical distribution of living grandchildren across the same cohorts. In 2017, on average, men born between 1920 (aged 97) and 1940 (aged 77) had 3.6 living grandchildren. One in five of these men (20.2%) had no living grandchildren, about half (47.2%) had between one and four living grandchildren, and the rest (32.6%) had five or more living grandchildren. Overall, there was stark variability among the elderly men in terms of having living grandchildren at the time of the study; one-fifth had none, whereas one-fifth had six or more grandchildren. The picture is similar for women born in this period, but they had a slightly higher average number of grandchildren compared to men (3.8), and a slightly lower share of women (17.4%) had no living grandchildren in 2017. Naturally, the number of living grandchildren declined for individuals born after 1940, as some or all of their grandchildren have yet to be born, though this decrease occurred more quickly for men, who tend to become grandparents at later ages than women. Approximately 69.9% of men born in 1960 (aged 58) had no living/observed grandchildren, whereas this percentage is only 55.8% for women born in 1960. In Figure S7, we show Figure 2 by sex of the grandchild.

Figure 3a: Average number of living and dead children in 2017 by sex and birth cohort 1915–2017. Figure 3b: Proportional distribution of the number children in 2017 by sex and birth cohort 1915–2017. Swedish-born individuals alive in 2018.

Turning to the next generation, the left panel of Figure 3 shows the average number of living children for men and women across the 1915–2017 birth cohorts, also broken down by childbearing partners of the individual. In other words, it shows how much individuals with one, two, three, or more childbearing partners contributed to the average number of (living) children in that cohort. The right panel shows the numerical distribution of living children across these cohorts. Cohort fertility was very stable in Sweden during across the 1915–1950 birth cohorts, at around two children per individual on average, though period fertility fluctuated strongly throughout the 20th century. When examining final parity, we found an increasing two-child norm over time, though the broader pattern was stable. For women born before 1977, the proportion of childless women remained stable, at around 12.4% on average. For men, this proportion increased steadily across cohorts, and for male cohorts born after 1955, at least 21.5% had no children as of 2017. As a result, the average number of children for men born after 1955 declined across cohorts, whereas we observed about two children on average for women across all cohorts with a complete set of children by 2017. For cohorts born after 1977, an increasingly large proportion of individuals had not yet completed their fertility by 2017. There was also a steady increase in the average number of children from parents with two or more childbearing partners across the 1915–1945 cohorts; for cohorts born after 1945, the average remained constant. Individuals with two partners contributed a substantial amount

18

to the total childbearing, whereas the contribution of individuals with three or more partners was very small. In Figure S8 we show fertility both by parity and number of childbearing partners for the 1960-1970 cohorts.
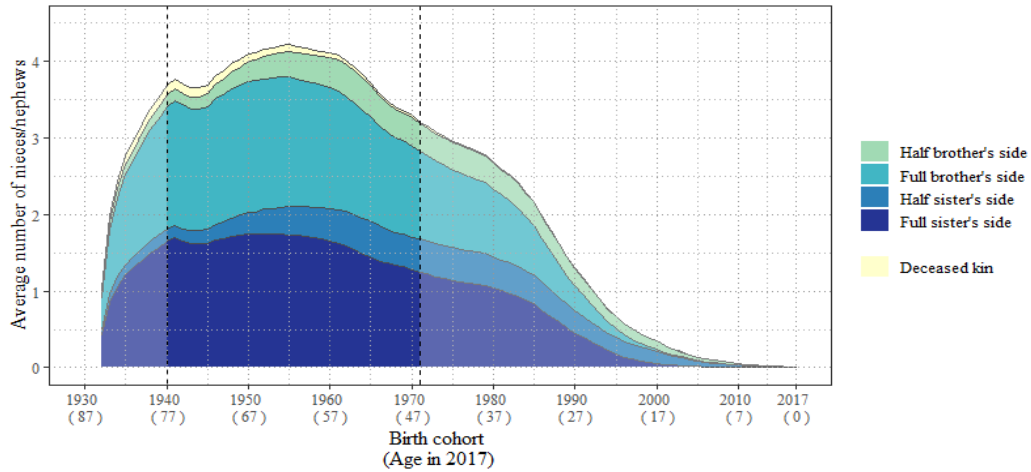

Figure 4: Average number of nieces and nephews by birth cohort and through full or half sister/brother, 1930–2017. Swedish-born individuals alive in 2018.

Figure 4 shows the average number of nieces and nephews of the ego across the 1932–2017 birth cohorts (1933 is when we began tracing parent–child links). We observed the peak of just over an average of four (4.13) nieces and nephews for individuals born in 1955. For those born before 1940, a higher share of parental information is missing. Therefore, we do not have complete sibling information, and consequently, we also lack information on nieces and nephews. For cohorts born after 1971, it is likely that the eventual full set of nieces and nephews was not yet complete in 2017, either because their own siblings were not yet born or because their siblings did not have a completed set of children by 2017. The majority of individuals had nieces and nephews through their full siblings. However, the average number of nieces and nephews from half siblings increased for later cohorts.
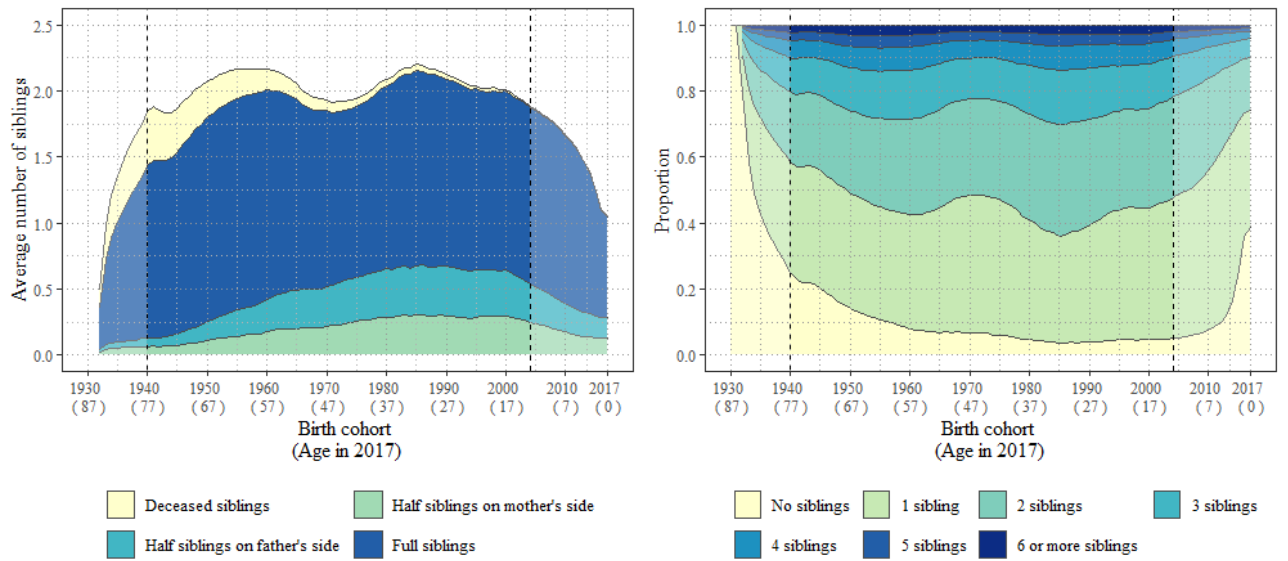
Figure 5a: Average number of siblings by birth cohort and whether full or half sibling, 1930–2017. Figure 5b: Proportional distribution of number of siblings (half or full) by birth cohort, 1930–2017. Swedish-born individuals alive in 2018.

Next, we enumerated kin who were the ego's generational contemporaries. The left panel of Figure 5 shows the average number of siblings of egos across the 1932–2017 cohorts, and the right panel shows the numerical distribution of siblings across the same cohorts. The average number of siblings oscillates around two (as cohort fertility was stable) for those with almost complete information on siblings (cohorts 1940–2004). The majority of these siblings were full siblings, though the average number of half siblings increased across cohorts, which is the reason for the trends we previously observed for nieces and nephews. Figure 5b most notably shows an increase from the 1985 cohorts onward in terms of the proportion of individuals with only one sibling. Again, for the cohorts outside the dashed lines, we observed incomplete information on siblings due to a higher share of missing parental information for those born before 1940 and for incomplete sibling sets for those born after 2004. In Figure S9 we show the distribution of number of children by exact number of siblings and whether half/full siblings for the 1960-1970 cohorts. In Figure S13 we show how widely spaced sibling sets are, from the perspective of the 1980 and 1985.
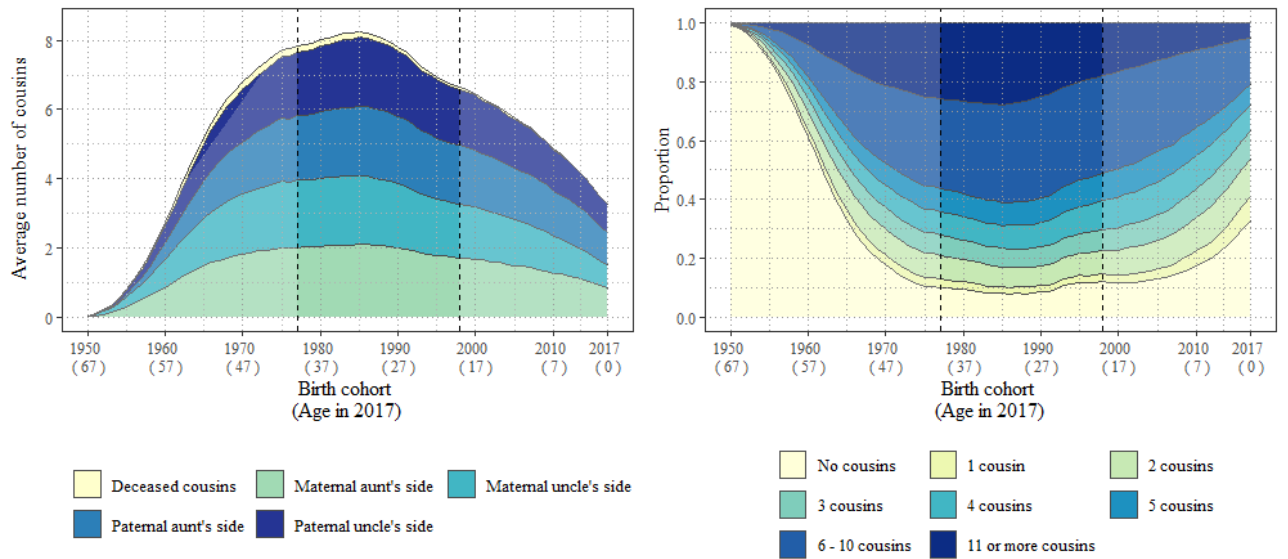
Figure 6a: Average number of cousins by birth cohort and by type of aunt/uncle, 1930–2017.
Figure 6b: Number of cousins by birth cohort, 1930–2017. Swedish-born individuals alive in
2018.

How many cousins are there in people's kinship networks? In Figure 6, we show the average number of cousins of egos for the cohorts 1950–2017 in the left panel and the numerical distribution of cousins across the same cohorts in the right panel. The enumeration of cousins had among the highest data requirements, as we needed to observe the grandparents of the index person in the registers while ensuring that the fertility of all parental siblings was completed. Thus, we have complete information for the 20 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 (8.01%). Maternal cousins were more or less exactly 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 among the largest variations across individuals, with more than 10% of those born in 1980 having more than 20 cousins on average (Figure S3). We show a more detailed version of Figure 6a, with more information on those sharing one or two grandparents in Figure S10. In Figure S14 and S15 we show how widely spaced cousin sets are, from the perspective of the 1970, 1975, 1980 and 1985 cohorts.
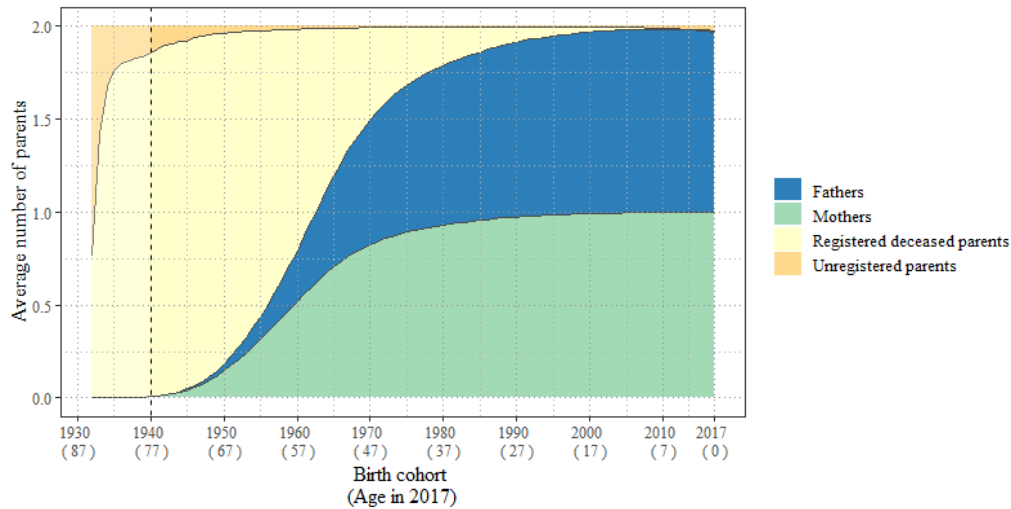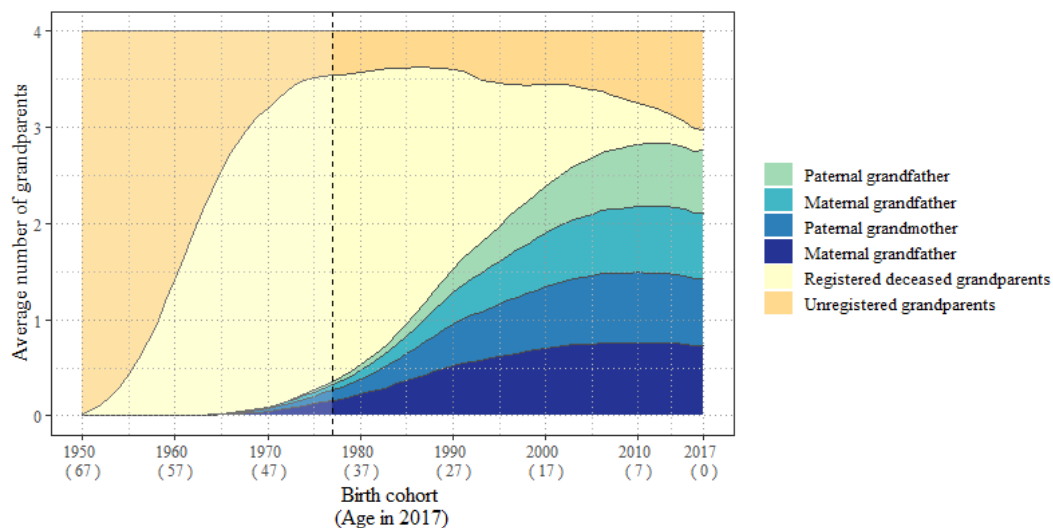
Figure 7: Average number of living, dead, and unregistered parents by birth cohort 1932–2017 for Swedish-born individuals alive in 2018.

Turning to generations preceding the ego, Figure 7 shows the average number of mothers and fathers alive for all cohorts of the ego born between 1932 and 2017. For those born after 1960 (age 57 or younger), more than half had a living mother in 2017; for the 1980 cohort or later, more than 92% had a living mother. The corresponding number for fathers were slightly lower (86%). The figure also provides information on the share of parents who could be traced in the register, as this was a necessary precondition for the reconstruction of kin not traced through the individual (kin other than children and grandchildren). 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 1 or 2 living parents in Figure S11
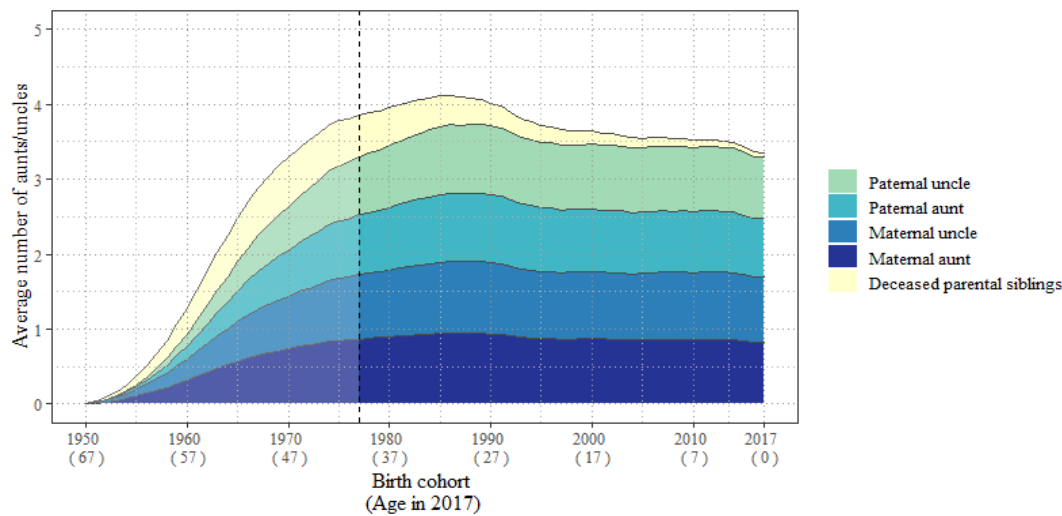
Figure 8: Average number of parent siblings by birth cohort 1950–2017 for Swedish-born individuals alive in 2018.

Figure 8 shows paternal and maternal aunts and uncles alive in 2017. We noted marked stability across cohorts with nearly complete coverage—on average, 3.50 aunts and uncles were alive, once again reflecting stable cohort fertility patterns in Sweden. 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 due to 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, as Sweden has had an increasing share of Swedish-born children born to foreign-born parents by cohort, for which we did not observe grandparents. In Figure S12 we show Figure 8 further broken down by half and full aunts/uncles.

Figure 9: Average number of living, dead, and unregistered grandparents by birth cohort 1950–2017 for Swedish-born individuals alive in 2018.

We finally show the average number of living grandparents of egos born between 1950 and 2017 in Figure 9. Those born between 2005 and 2017 had, on average, 3.2 grandparents alive in 2017, and about 0.41 grandparents on average had been registered but died earlier. For these cohorts, an average of 0.8 of their grandparents had never been registered in Sweden, presumably because they had never lived in Sweden. The average number of living grandparents decreased rapidly for those born before 2005; virtually no grandparents were found for individuals born before 1965. Among older cohorts with nearly complete information on grandparents who have ever been registered in Sweden—such as those born in 1980 and who thus were 37 in 2017—we found that living grandmothers were also substantially more frequent than living grandfathers, owing to the greater longevity of women.
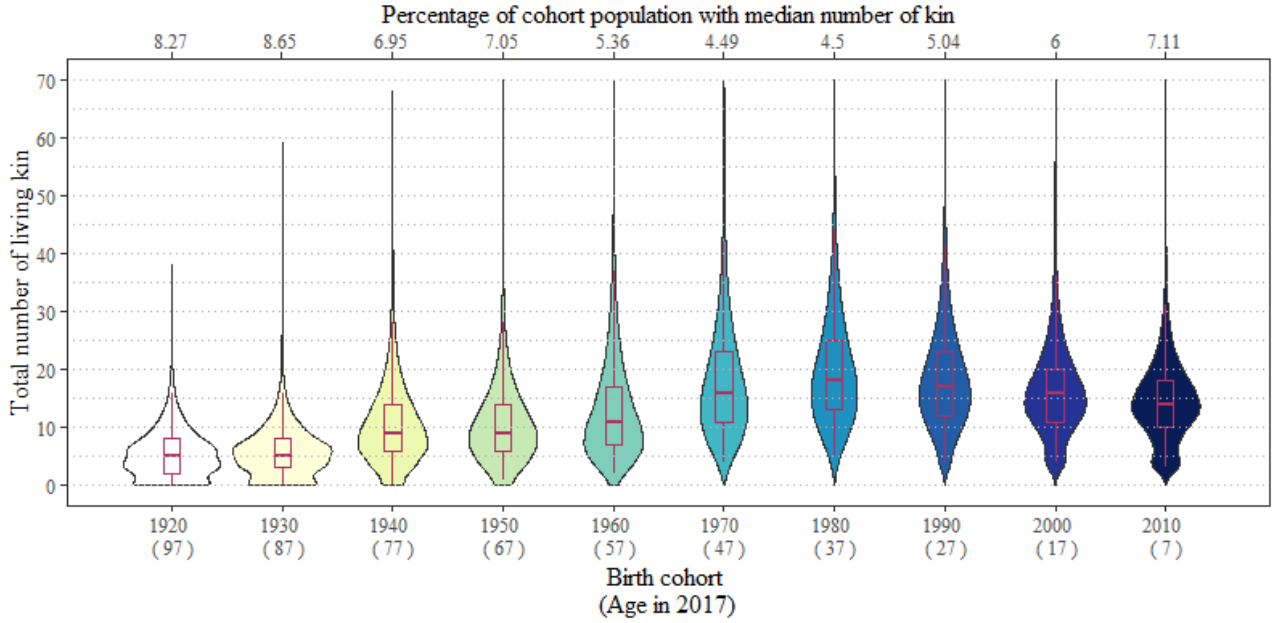
Figure 10: Distribution of the total number of kin for Swedish-born individuals alive in 2018 by birth cohort. Box plots show quartiles and 5th and 95th percentiles. The top shows the percentage of the population that had the exact median number of kin for that cohort (width of each curve).

After enumerating the extended kinship network of people alive in Sweden by type of kin, we conclude by focusing on the sum of the kinship network we observed. In this section, we have pooled all types of kin from our previous figures in Figure 10 and Figure 11. In Figure 10, we show the total number of kin, whereas in Figure 11, we also show the average number of kin by type. The distributions of the total number of kin alive across the 1920–2010 cohorts are given in Figure 10. The upper tail of the distribution is very long because some individuals had very large kin networks. The figure is truncated at 70; 4,398 individuals across all cohorts in our study had more than 70 living relatives in 2017. The median number of living kin was largest for the 1980 (age 37) cohort, where our coverage of both older and younger kin is of high-quality, and smallest for the oldest cohorts born in 1920 and 1930 (ages 97–87). As seen in Figure 11 and earlier figures, the small kinship size in the earliest cohorts is partly due to the missing parent-child links in our data.

The dispersion of the number of kin is very large. Individuals born between 1960 and 1990 had an interquartile range greater than 10, whereas the interquartile range of older cohorts was between 6 and 8. The distance between the 5th and 95th percentile is around 40 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, we found that the dispersion is also substantial in older cohorts, and it is not possible to identify a standard kinship size even among the elderly. Kinlessness is rare—the proportion of those who had no close living kin was very small for all cohorts from 1970 onward (less than 2%). Egos from the oldest cohort were those with the fewest kin, yet only 14.61% had no living relatives in this group.
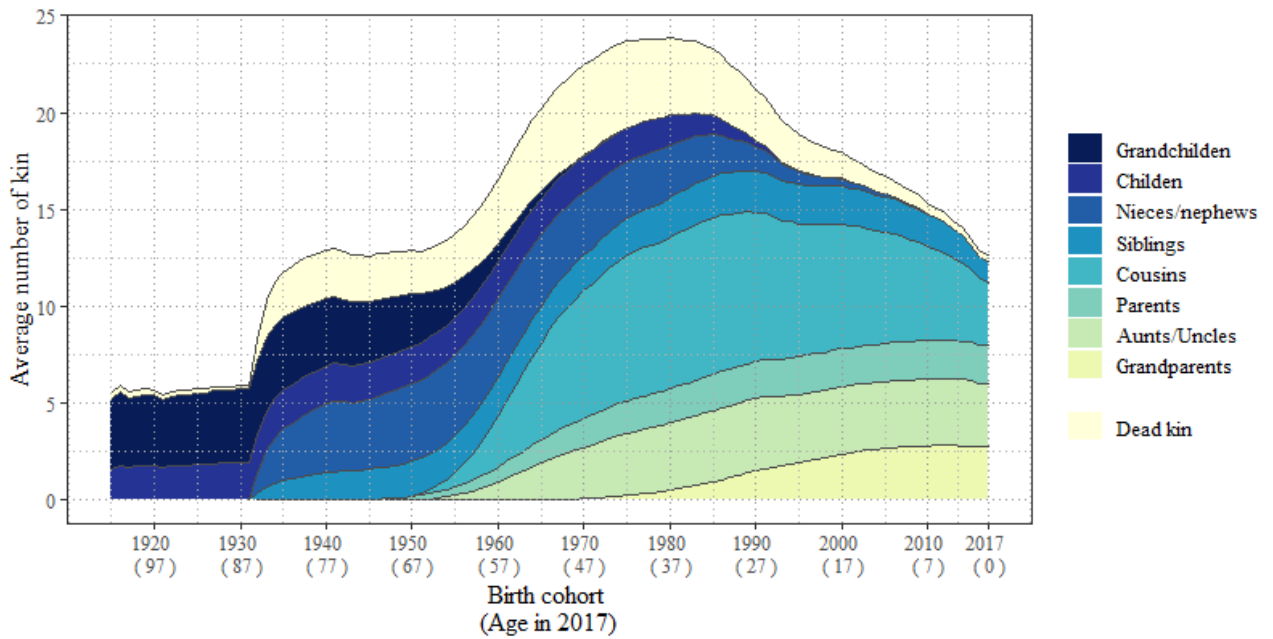

Figure 11: Average number of kin for all types of kin by birth cohort 1915–2017 for Swedish-born individuals alive in 2018. Dead kin includes all types of kin.

The average size of the total kin network by the type of kin is shown in Figure 11 for birth cohorts born from 1915–2017. As shown earlier, for the earlier cohorts, we lacked data due to missing parent–child links for some kin types (particularly cousins before the 1950s and siblings and nieces/nephews before the 1930s). The peak in the average total number of kin alive in 2017 was found to be for cohorts born around 1983 (age 35), who had, on average, 20 living kin members.

The higher number of kin among people in the middle of life (approximately age 28 to 53, or age cohorts 1965 to 1990) reflects that these members had both younger and older generations of kin alive at the same time (children as well as parents), and also that they had a larger number of horizontal kin (such as cousins, whereas sets were not yet completed for younger cohorts). For example, in the 1983 cohort (the year with the most complete coverage), the most common type of living relation was cousins; egos had, on average, 8.03 cousins, who made up 40.35% of the full kinship network. The size of the remaining kin network was, in

26

order of magnitude, aunts/uncles, followed by nieces/nephews and a more or less equally average number of parents, siblings, and children. However, if we had access to better data on cousins for earlier cohorts, the total kinship networks of cohorts born between 1930 and 1960 would likely have been substantially higher. For younger cohorts, the average number of older kin increased slightly while the number of younger kin decreased significantly. The inverse could be observed for cohorts born before the early 1980s. It is worth noting that we undercounted horizontal kin of older cohorts due to missing parental (and grandparental) links before 1932. The rapid decline in the average number of kin from cohorts born from 1965 to 1940 likely reflects such undercounting.

We found that a large part of the size of an individual kin network is determined by the number of parental siblings and their offspring (horizontal kin and horizontal kin of parents), and that this is also related to very large variations in kinship (Figure 10). We found a notable number of deceased kin at ages above 25; however, only for young cohorts is this a true reflection of their decreased kin for the truncation reasons discussed in our supplemental materials. To conclude, individuals of all ages have a substantial number of kin alive in Sweden on average, but the type of kin is a factor of the age of the individual. At the same time, such averages hide huge individual variations, as shown in Figure 10. In Figure 11 we only show living kin, we provide a similar figure with all ever-observed kin in Figure S6.

Discussion

In our study, we carried out extensive empirical documentation of the demography of kinship that helps fill a knowledge gap in the description of kin count and distributions of human. Our approach was to enumerate living kin by year of birth for the entire population of Sweden in 2017. We have shown the extent to which kin availability is shaped by the age of the individual and shown different estimates by family complexity. We have thus improved on previous attempts to estimate how common it is to have different categories of kin. Our results allowed us to establish several properties of kinship demographics. First, there is great heterogeneity in the sizes of kin networks within birth cohorts. We show substantive variability of kinship to an extent which is neither trivial, self-evident, nor previously analyzed in depth. For example, we have enumerated in detail the contrast between substantial minorities of elderly individuals with no grandchildren and those with six or more grandchildren; cousin networks ranged from three or four individuals to dozens. Come middle age, the interquartile range of individuals' total kinship size was as high as 10. Because we have demonstrated great dispersion in kinship size across all cohorts, we conclude that it is pernicious to make

generalizations about "typical" kinship. Complementary to this insight, we have pinpointed the properties that indeed are the "common denominators" of contemporary kinship; i.e., the aspects of kinship structure that appear to be shared by most people at a given age. For example, we have numerically specified the dominance of horizontal kin, such as cousins in adult kinship networks, and we showed in detail the younger-and-older kinship structure that characterizes kinship at mid-life. We found that kin ties that are structurally constrained in numbers (parents and grandparents) showed less variability than kin ties that are not (e.g., siblings). The greatest variance was found for cousins and grandchildren, which rely on the fertility behaviors of multiple generations (grandchildren and cousins).

Moreover, we have shown that the extent of relations usually not estimated in accounts of kinship, namely half siblings and parents of half siblings, was non-negligible. We also showed how the importance of kinship originating in multi-partner fertility increased over the period we studied. Our results indicate that the decision to include, exclude, or distinguish between full and half siblings is an important consideration in the study of kinship. Finally, by distinguishing between kin linked to the ego through paternal and maternal sides, we could show the extent to which and among which kin there were numerical differences based on lineage. Here, we conclude that such sex-based patterns were lacking among nieces, nephews, aunts, and uncles, and that overall, contemporary kinship structures in Sweden are not characterized by an overrepresentation of maternal or paternal kin, with the exception of a greater frequency of (living) grandmothers than grandfathers. Our results are specific to Sweden, but as mortality and fertility have shared broad similarities over the 20th century in many high-income countries, many of our results can probably be generalized to other demographically similar contexts. A partial exception may be the relatively higher fertility in Sweden compared to some southern European and central European countries, which may produce more horizontal kin. 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 fertility transition, unlike the more stable mortality and fertility patterns in Sweden.

In addition to presenting empirical numbers that can serve as a base for future kinship demography, our findings are relevant to the sociological literature on various aspects of 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 (on, for example, financial and emotional wellbeing), on what kin members do (frequency of interaction, provision of help), and the nature of the social relations ascribed to kin (degrees of 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 (Agree and Glaser 2009). Examining issues such as how much care a person can receive from kin is conditioned on the kinship structure of that individual.

Another contribution of this study is to demonstrate the usefulness and the limitations of using administrative data to conduct kinship demography. Although linked intergenerational administrative register data have been readily available for decades, this study presents the first attempt at providing a comprehensive approach to kin. A more granular and precise measure of kinship demography can be obtained by using an observed population rather than a sample estimate of the population. Moreover, many of the calculations that we performed (particularly measures of dispersion and measures of kin linked through only one parent) would be difficult to achieve with other approaches. We have also shown the limitations and pitfalls of using administrative data (such as implications of truncated parent–child links), which is instructive for future research and important to highlight tasks where survey, register, and simulation approaches to kinship demography are preferable. In particular, we have shown 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 are valuable as empirical evidence to validate other methods of calculating kinship, such as microsimulations and analytical approaches (Caswell 2020; Goodman et al. 1974; Ruggles 1993; Zagheni 2015), the only option for (most) other countries that lack administrative register data available over longer periods of the kind demonstrated in this study. Kin relationships of intact families can be reasonably well approximated by analytical models and microsimulation models, as birth intervals and age-specific fertility follow predictable patterns, and thus we expect analytical/simulation approaches to perform better for vertical kin relationships. However, union instability and multi-partner 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. Accurately measuring such factors is currently beyond contemporary approaches in micro-simulations and analytical modeling, and our descriptions are likely much more accurate than estimates from these methods regarding such particular kin relations. Similarly, the extent to which demographic rates are concentrated in families needs to be accounted for, as positive within-kin-group correlations in mortality and fertility will produce larger numbers of simultaneously small and large kin groups than predicted from models with uniform rates across members of population

subgroups. Similar mechanisms pertain to generational length and birth timing that vary strongly by, for example, education.

To conclude, we argue that demographic characteristics of kinship have been neglected in empirical social research. In the current study, we have addressed this by providing a full assessment of kinship links in contemporary Sweden, which had previously never been calculated for any national population. We show important categorizes of kin, including dimensions of family complexity, that are hard to assess with the analytical and simulated approaches which, until now, have been used to study this type of question. We hope that our work can serve as an inspiration for future research quantifying kinship as well as be a reference point for the rich empirical research that will be attempted to measure how kin continues to play an active role in the lives of men and women in high-income societies.

Acknowledgments

References

Agree, E.M.and K. Glaser. 2009. "Demography of informal caregiving." Pp. 647-668 in *International handbook of population aging*, edited by P. Uhlenberg. Dordrecht: Springer.

Albertini, M., M. Kohli, and C. Vogel. 2007. "Intergenerational transfers of time and money in European families: common patterns—different regimes?" *Journal of European Social Policy* 17(4):319-334.

Alburez-Gutierrez, D., M. Kolk, and E. Zagheni. 2021. "Women's experience of child death over the life course: A global demographic perspective." *Demography* forthcoming.

Alburez-Gutierrez, D., C. Mason, and E. Zagheni. 2021. "The "Sandwich Generation" revisited: global demographic drivers of care time demands." *Population and Development Review*.

Andersson, L. 2020. "Oh half-brother, where art thou? The boundaries of full-and half-sibling interaction." *Demographic Research* 43:431-460.

—. 2021. "Lifetime parenthood in the context of single-and multiple-partner fertility." *Advances in Life Course Research* 47:100355.

Bengtson, V.L. 2001. "Beyond the nuclear family: The increasing importance of multigenerational bonds." *Journal of Marriage and the Family* 63(1):1-16.

Berggren, H.and L. Trägårdh. 2006. *Är svensken människa?: gemenskap och oberoende i det moderna Sverige*. Stockholm: Norstedt.

Boholm, Å. 1983. *Swedish Kinship: an Exploration into Cultural Processes of Belonging and Continuity*. Gothenburg: Acta Universitatis Gothoburgensis.

Bonvalet, C.and É. Lelièvre. 2016. "Family Beyond Household and Kin: Life Event Histories and Entourage, a French Survey." in *INED Population Studies*. Paris: Springer.

Brubaker, T.H. 1990. "Families in later life: A burgeoning research area." *Journal of Marriage and the Family* 52(4):959-981.

Caswell, H. 2019. "The formal demography of kinship." *Demographic Research* 41:679-712.

—. 2020. "The formal demography of kinship. II. Multistate models, parity, and sibship." *bioRxiv*.

Chudnovskaya, M.and M. Kolk. 2017. "Educational expansion and intergenerational proximity in Sweden: developments in geographical distance between young adults and their parents, 1980-2007." *Population, Space and Place* 23(1): e1973.

Coall, D.A.and R. Hartwig. 2010. "Grandparental investment: Past, present, and future." *Behavioral and Brain Sciences* 33(1):1-58.

Colladoa, M.D., I. Ortuno-Ortın, and J. Stuhler. 2020. "Estimating Intergenerational and Assortative Processes in Extended Family Data." *online ms* https://janstuhler.files.wordpress.com/2020/10/collado-ortuno-and-stuhler-manucript_v37.pdf.

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.

Di Leonardo, M. 1987. "The female world of cards and holidays: Women, families, and the work of kinship." *Signs* 12(3):440-453.

Dykstra, P.A., M. Kalmijn, G. Knijn, A.E. Komter, A.C. Liefbroer, and C.H. Mulder. 2006. *Family solidarity in the Netherlands*. Amsterdam: Dutch University Press.

Dykstra, P.A.and A.E. Komter. 2006. "Structural characteristics of Dutch kin networks." Pp. 21-42 in *Family solidarity in the Netherlands*, edited by P.A. Dykstra, M. Kalmijn,

T.C.M. Knijn, A.E. Komter, A.C. Liefbroer, and C.H. Mulder. Amsterdam: Dutch University Press.

Engzell, P., C. Mood, and J.O. Jonsson. 2020. "It's All about the Parents: Inequality Transmission across Three Generations in Sweden." *Sociological Science* 7:242-267.

Finch, J. 2007. "Displaying families." *Sociology* 41(1):65-81.

Furstenberg, F.F. 2020. "Kinship reconsidered: Research on a neglected topic." *Journal of Marriage and Family* 82(1):364-382.

Giddens, A. 1993. *The transformation of intimacy: Sexuality, love and eroticism in modern societies*. Cambridge: Polity Press.

Goldstein, J.R. 1999. "Kinship networks that cross racial lines: the exception or the rule?" *Demography* 36(3):399-407.

Goldstein, J.R.and J.R. Warren. 2000. "Socioeconomic reach and heterogeneity in the extended family: contours and consequences." *Social Science Research* 29(3):382-404.

Goode, W.J. 1963. *World revolution and family patterns*. London: Free press of Glencoe.

Goodman, L.A., N. Keyfitz, and T.W. Pullum. 1974. "Family formation and the frequency of various kinship relationships." *Theoretical population biology* 5(1):1-27.

Habermas, J. 1985. *The Theory of Communicative Action: Reason and the rationalization of society*. Boston: Beacon Press.

Hagestad, G.O. 1986. "The family: Women and grandparents as kin-keepers." Pp. 141–160 in *Our Aging Society: Paradox and Promise*, edited by A.J. Pifer and L. Bronte. New York: WW Norton & Company.

Hajnal, J. 1982. "Two kinds of preindustrial household formation system." *Population and Development Review* 8(3):449-494.

Hank, K.and A. Steinbach. 2018. "Intergenerational solidarity and intragenerational relations between adult siblings." *Social Science Research* 76:55-64.

Heady, P.and M. Kohli. 2010. "Family, kinship and state in contemporary Europe." Chicago: The University of Chicago Press.

Himes, C.L. 1992. "Future caregivers: Projected family structures of older persons." *Journal of gerontology* 47(1):17-26.

Hällsten, M.and M. Kolk. 2020. "The shadow of peasant past: Seven generations of inequality persistence in Northern Sweden." *SocArXiv* https://doi.org/10.31235/osf.io/yjksz.

Johnson, C.L. 2000. "Perspectives on American kinship in the later 1990s." *Journal of Marriage and Family* 62(3):623-639.

Kalmijn, M., K. Ivanova, R. van Gaalen, S.G. de Leeuw, K. van Houdt, F. van Spijker, and M. Hornstra. 2018. "A multi-actor study of adult children and their parents in complex families: Design and content of the OKiN survey." *European sociological review* 34(4):452-470.

Kaptijn, R., F. Thomese, T.G. Van Tilburg, and A.C. Liefbroer. 2010. "How Grandparents Matter." *Human Nature* 21(4):393-405.

Kolk, M. 2014. "Multigenerational Processes in Demography." Stockholm University Demography Unit, Stockholm.

—. 2017. "A Life Course Analysis of Geographical Distance to Siblings, Parents and Grandparents in Sweden." *Population, Space and Place* 23(3):e2020.

Kolk, M.and G. Andersson. 2020. "Two Decades of Same-sex Marriage in Sweden. A Demographic Account of Developments in Marriage, Childbearing and Divorce." *Demography* 57(1):147-169.

Lansing, J.S., C. Abundo, G.S. Jacobs, E.G. Guillot, S. Thurner, S.S. Downey, L.Y. Chew, T. Bhattacharya, N.N. Chung, H. Sudoyo, and M.P. Cox. 2017. "Kinship structures

create persistent channels for language transmission." *Proceedings of the National Academy of Sciences* 114(49):12910.

Laslett, P. 1977. "Characteristics of the Western Family Considered Over Time." *Journal of Family History* 2(2):89-115.

Lee, R.and A. Mason. 2011. *Population aging and the generational economy : a global perspective*. Cheltenham: Edward Elgar.

Leopold, T.and J. Skopek. 2015. "The Demography of Grandparenthood: An International Profile." *Social Forces*:DOI: 10.1093/sf/sov1066.

Lévi-Strauss, C. 1969[1949]. *The elementary structures of kinship*. London: Eyre & Spottiswood.

Lundholm, E.and G. Malmberg. 2009. "Between elderly parents and grandchildren: Geographic proximity and trends in four-generation families." *Journal of Population Ageing* 2(3):121-137.

Malinowski, B. 1982[1932]. *The sexual life of savages in north-western Melanesia : an ethnographic account of courtship, marriage, and family life among the natives of the Trobriand Islands, British New Guinea*. London: Routledge.

Mare, R.D. 2011. "A multigenerational view of inequality." *Demography* 48(1):1-23.

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.

Martin, L.G. 1990. "Changing intergenerational family relations in East Asia." *The Annals of the American Academy of Political and Social Science*:102-114.

Michielin, F.and C. Mulder. 2007. "Geographical distances between adult children and their parents in the Netherlands." *Demographic Research* 17(22):655-678.

Modalsli, J. 2016. "Multigenerational persistence. Evidence from 146 years of administrative data." No 850.

Morgan, L.H. 1997[1870]. *Systems of consanguinity and affinity of the human family*. Lincoln, Neb.: Univ. of Nebraska Press.

Murphy, M. 2011. "Long Term Effects of the Demographic Transition on Family and Kinship Networks in Britain." *Population and Development Review* 37(supplement s1):55-80.

Palloni, A., K. Hill, and G.P. Aguirre. 1996. "Economic swings and demographic changes in the history of Latin America." *Population Studies* 50(1):105-132.

Parsons, T.and R.F. Bales. 1955. *Family, socialization and interaction process*. Glencoe, Ill: Free Press.

Pfeffer, F.T. 2014. "Multigenerational approaches to social mobility. A multifaceted research agenda." *Research in Social Stratification and Mobility* 35:1-12.

Post, W., F. Van Poppel, E. Van Imhoff, and E. Kruse. 1997. "Reconstructing the extended kin-network in the Netherlands with genealogical data: methods, problems, and results." *Population Studies* 51(3):263-278.

Reher, D. 2005. "Family ties in western Europe." Pp. 45-76 in *Strong Family and Low Fertility: A Paradox?*, edited by G. Dalla Zuanna and G.A. Michel: Springer.

Rivers, W.H.R. 1968[1914]. *Kinship and social organization : together with "The genealogical method of anthropological enquiry"*. London: Athlone.

Rossi, A.S.and P.H. Rossi. 1990. *Of Human Bonding: Parent-Child Relations Across the Life Course*. New York: Aldine de Gruyter.

Ruggles, S. 1987. *Prolonged connections: The rise of the extended family in nineteenth-century England and America*. Madison, Wisconsin: University of Wisconsin Press.

Ruggles, S. 1993. "Confessions of a microsimulator." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 26(4):161-169.

Ruggles, S.and S. Brower. 2003. "Measurement of household and family composition in the United States, 1850–2000." *Population and Development Review* 29(1):73-101.

Ruggles, S.and M. Heggeness. 2008. "Intergenerational coresidence in developing countries." *Population and Development Review* 34(2):253-281.

Schneider, D.M. 1968. *American kinship: A cultural account*. Chicago: University of Chicago Press.

Schneider, D.M.and C.B. Cottrell. 1975. *The American Kin Universe: A Genealogical Study*. Chicago: Department of Anthropology, University of Chicago.

Schneider, D.M.and R.T. Smith. 1973. *Class Differences and sex roles in American Kinship and family structure*. Engelwood Cliff, NJ: Prentice-hall.

Schulz, J.F., D. Bahrami-Rad, J.P. Beauchamp, and J. Henrich. 2019. "The Church, intensive kinship, and global psychological variation." *Science* 366(6466):eaau5141.

Sjögren, H. 2006. "Family capitalism within big business." *Scandinavian Economic History Review* 54(2):161-186.

Szebehely, M. 2005. "Anhörigas betalda och obetalda äldreomsorgsinsatser." *Statens offentliga utredningar* SOU 2005: 66:133-206.

Thornton, A. 2005. *Reading history sideways : the fallacy and enduring impact of the developmental paradigm on family life*. Chicago: University of Chicago Press.

Tomassini, C., S. Kalogirou, E. Grundy, T. Fokkema, P. Martikainen, M.B. Van Groenou, and A. Karisto. 2004. "Contacts between elderly parents and their children in four European countries: current patterns and future prospects." *European Journal of Ageing* 1(1):54-63.

Uhlenberg, P. 1996. "Mortality decline in the twentieth century and supply of kin over the life course." *The Gerontologist* 36(5):681-685.

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.

Verdery, A.M. 2015. "Links Between Demographic and Kinship Transitions." *Population and Development Review* 41(3):465-484.

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.

Young, M.D.and P. Willmott. 1957. *Family and Kinship in East London*. Glencoe, Ill.: Free Press.

Zagheni, E. 2015. "Microsimulation in demographic research." *International encyclopedia of social and behavioral sciences* 15:343-346.

Supplemental material – part 1

Evaluating kinship through Swedish register data – a description of constraints and data sources

In this supplemental text, we first describe Swedish register data in some detail, as it is important to understand the opportunities and limitations of numerating Swedish kinship. After giving a description of Swedish register data, we show how data availability affected kinship counts in some categories and motivated the cut-offs that were used in our study.

Swedish register data – data sources and quality issues

Sweden has had government administrative registers for a long period, having the longest series of vital events in the world. Modern Swedish registers originated with the introduction of a unique personal identity number (personnummer) in Sweden from 1947–1948. In connection with the introduction of this number, information was collected on parenthood for all (surviving) resident children under the age of 16. This information formed the basis of the parent–child links that served as the foundation for the type of kinship analysis carried out in this dissertation. After 1947, information on parents of children was collected directly from birth records and refers directly to biological parenthood. As information in 1947 was collected for children under the age of 16, 1932 is the earliest cohort that can be linked to their parents. This means that children born after 1932 can be linked to siblings born after 1932. Similarly, to estimate numbers of cousins, the grandparents of the cousins have to have had their children (parents, aunts, uncles) after 1932 for the link to be observed.

An additional limitation of contemporary Swedish registers is that they were digitized in the late 1960s. Yearly population registers dating back to 1968 were digitized and all subsequent demographic data in Sweden has been entered digitally. Therefore, 1968 is the earliest time for which event registers on demographic events are available. Additionally, Statistics Sweden digitized the census in 1960 and 1965. As a result, survival to 1960 is another condition for an individual to be included in contemporary Swedish register data. This also applied to our study, meaning that cohorts born in the early 1900s, for example, and those of their children who died before 1960, were (in most cases) not included in our study.

The Swedish multigenerational register (Statistics Sweden 2010) was created for research and is the source of the parent–child links in the study. It was constructed from the perspective of index individuals who had information on (biological) parents. Birth records are used to infer parenthood, and paternity for married parents is automatically assigned to the

husband, whereas for cohabiting couples, the male partner must fill in a form to claim paternity. In cases of missing/uncertain paternity, the civil service started very comprehensive searches for the missing father, and after 1960, missing paternity data is very rare (typically below 1%).

A consequence of the conditions outlined above, the target population of the register is individuals born after 1932 and alive in 1961. Various improvements in coverage have been made in order to make sure that data is of high quality for this target population, mostly adding information on people who died between 1947 and 1960. Overall, linkage rates are very high for individuals born in Sweden (Statistics Sweden 2010), particularly after the 1940s.

For our study, the most consequential limitation was the lack of parent–child linkages before 1932. Below we calculated the extent of such missingness in our data. The second most consequential limitation was conditioning on survival to 1960. In practice, this means that we did not capture child and infant mortality of boys and girls before that. To illustrate the extent of such missingness, for women born in 1920 survival to age 5 was 93%, and survival to age 45 was 87%. For other kin, except estimated childbearing of individuals born before 1950, the extent of this bias was very small (as calculating kinship often implicitly conditions on survival of intermediate generations).

Evaluating the extent of missing kinship in our study

As illustrated in our main manuscript, for some cohorts in our data, we could not enumerate their eventual total number of kin, as either a) the ego was so young that some individuals with a high certainty of being born in the future were not yet born, or b) some cohorts lacking parent–child links from around 1932 means that further back in time, our kinship networks were truncated. The first aspect is not a bias but a consequence of enumerating kinship at a given point. The second aspect is a direct bias introduced by data limitations that mostly affects estimates of older kin of people born earlier than 1977

To visually illustrate the second issue that many kin sets were not yet completed in 2020, we selected an 'ideal' cohort where we were certain we could observe all ever-born kin of that cohort for each kin type then calculate how many years it would take for 95% of their kin to be born, the results of which are shown in table S1 below. For example, 95% of children of the 1955 cohort were born within 40 years, in 1995 or before. Assuming it takes roughly 40 years for all cohorts to complete children sets implies that in 2017, we could not observe complete children sets for cohorts born in 1977 or later.

Table S1: Summary table for the year a reference cohort had a 95% complete kin set by type and the number of years required to complete such a set.

Kin type	Reference cohort (with a certainty of good coverage)	Cohort with an observed 95th percentile of a given kin type	Number of years to have a 95% complete set
grandchild	1930	2004	74
child	1955	1995	40
sibling's child	1955	2001	46
sibling	1985	1998	13
cousin	1985	2004	19

To illustrate the impact of cohorts where we might have lacked some kin due to missing parent links before 1932, we focused on the birth cohorts born in 1940 or after, for whom at least 95% of individuals had a mother observed in the registers. For types of kin that also require grandparental information—e.g., aunts, uncles, and cousins—we focused on those cohorts where at least 95% of individuals had two observed grandmothers, which are the cohorts born in 1977 or later.

Thus, for estimates of older kin, such as grandparents and cousins where we needed two intergenerational links, we used 1977 (and later) as an approximate cohort where our data did not suffer from intergenerational parent–child truncation. For generations with a distance of one, such as parents, we instead had 95% complete data from 1940 (and later) cohorts.

See also the supplemental figures S13–S15 that show the dispersion of kin for typical cohorts (1980 and 1985 for showing dispersion of birth years of siblings and half siblings, and 1980, 1985, 1990, and 1995 for cousins and half cousins). These figures give a good illustration of how many years are required to "complete" different sets of kinship.

Supplemental material - part 2

Supplementary figures

This part of the supplementary material contains additional descriptions of kin. These results have two purposes. First, the are used to assess the limits and validity of register data, as discussed in part1 and in the methodology section. See, for example, Figure S13 which explore the proportional distribution of birth spacing of siblings in order to assess how many living sibling kin are underestimated due to right-side censoring. The second purpose of these additional analyses is to provide a fuller catalogue of kinship demography and provide more detail to results in the main manuscript. For example, figure 10 in the main manuscript highlighted the exceptional variability in the total number of alive kin in the population, and in Figure S1-S5, we further show this across specific kin groups. Other figures, e.g. figure S14, show alive kin disaggregated across maternal and paternal lineage and whether kin are connected via one or two antecedents (e.g. maternal or paternal half-siblings, cousins connected by one or both grandparents). This is useful for researchers who need benchmark data, vital statistics, or input to simulation exercises, in order to describe kinship demography under family complexity. The authors can provide data in aggregate format upon request, to support such usage.
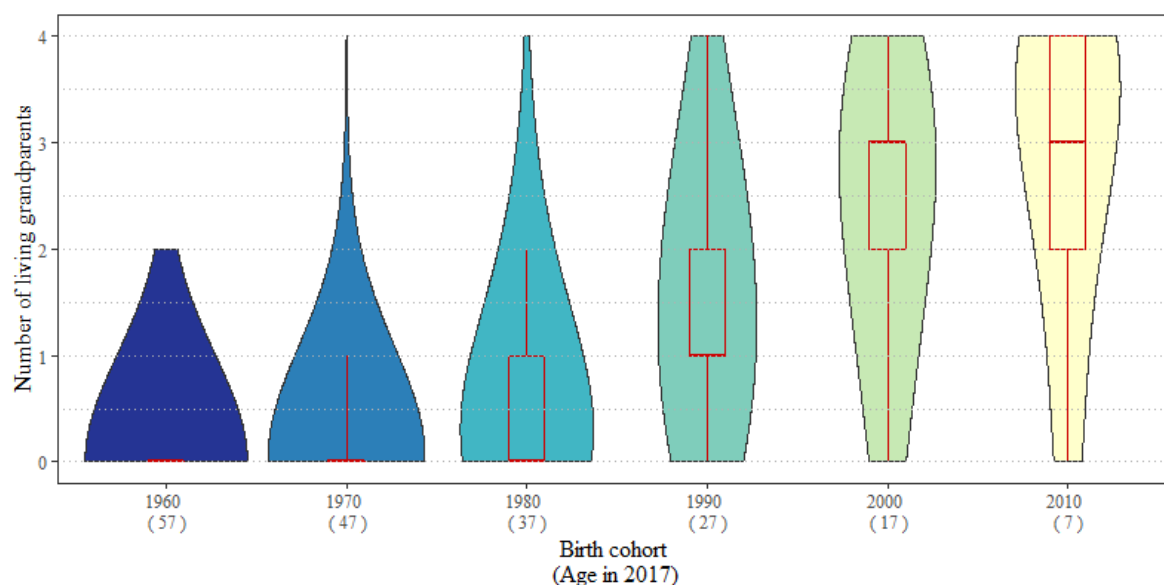
Figure S1: Distribution of the total number of living grandparents for Swedish-born individuals alive in 2018, birth cohorts 1960–2017. Box plots show quartiles and 5th and 95th percentiles.
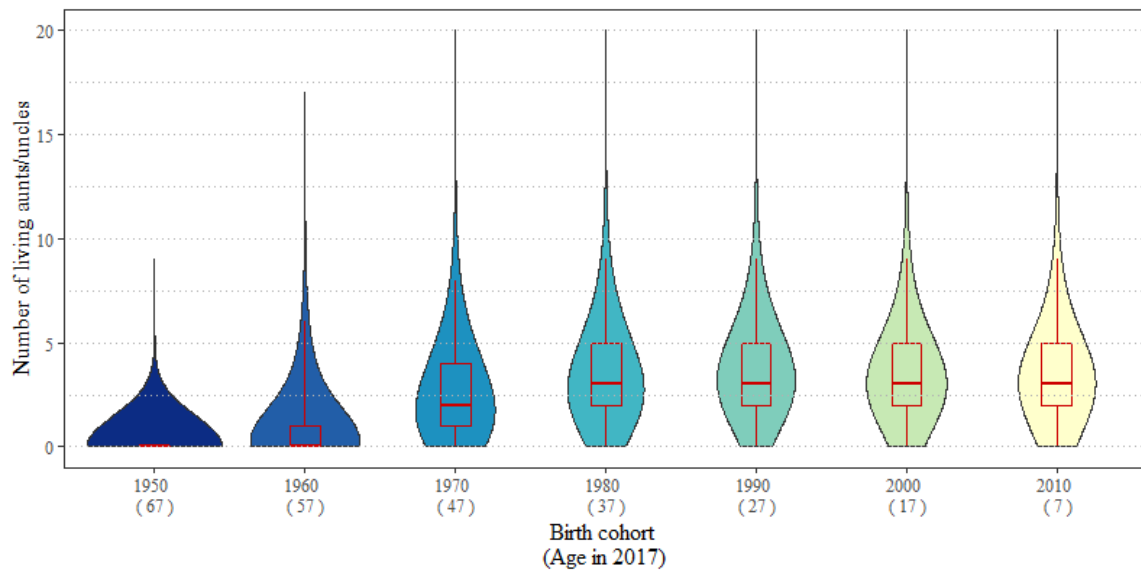
Figure S2: Distribution of the total number of living aunts and uncles for Swedish-born individuals alive in 2018, birth cohorts 1950–2017. Box plots show quartiles and 5th and 95th percentiles.
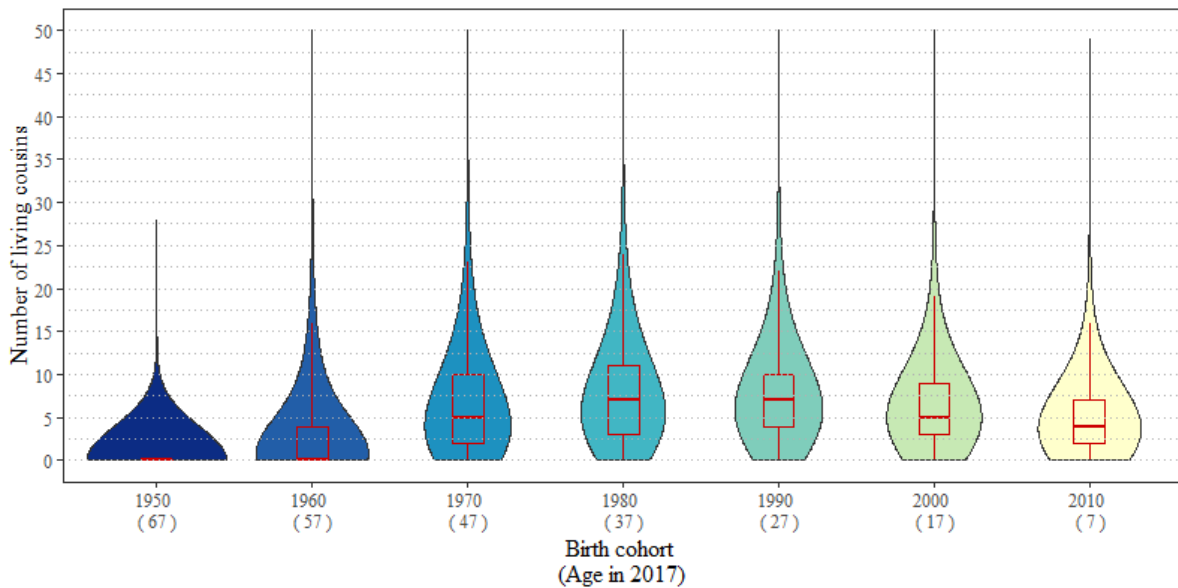
Figure S3: Distribution of the total number of living cousins for Swedish-born individuals alive in 2018, birth cohorts 1950–2010. Box plots show quartiles and 5th and 95th percentiles.
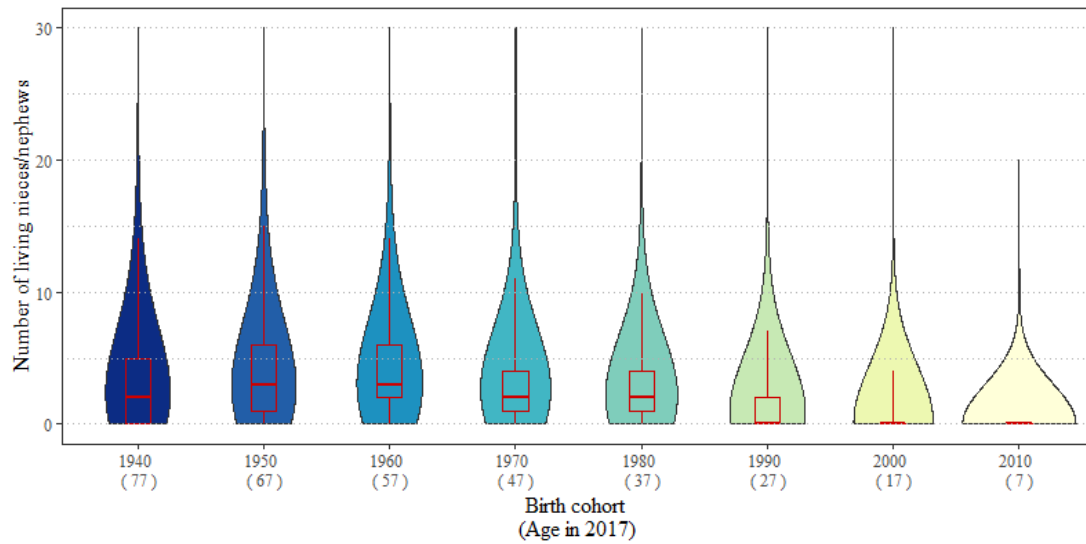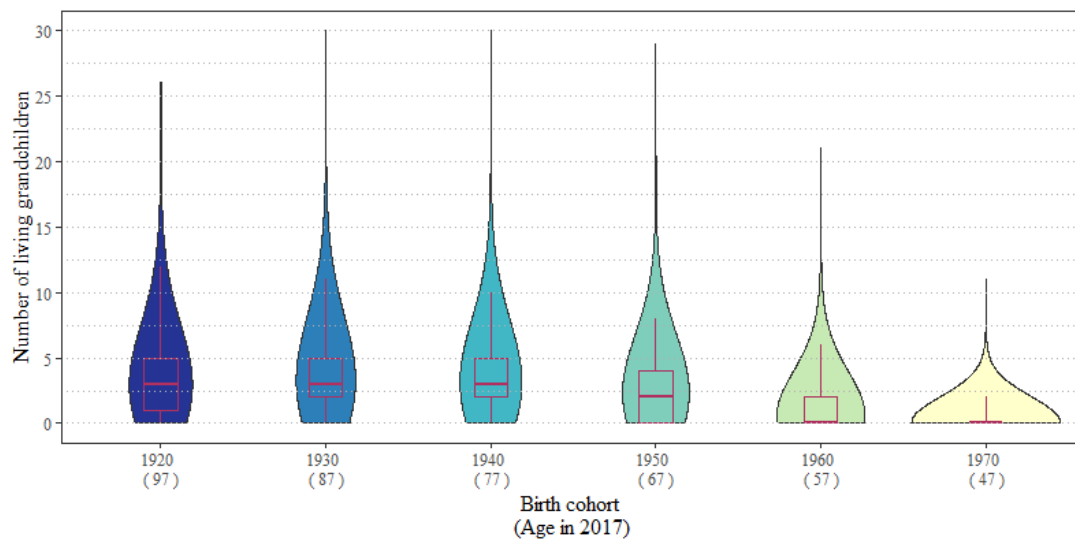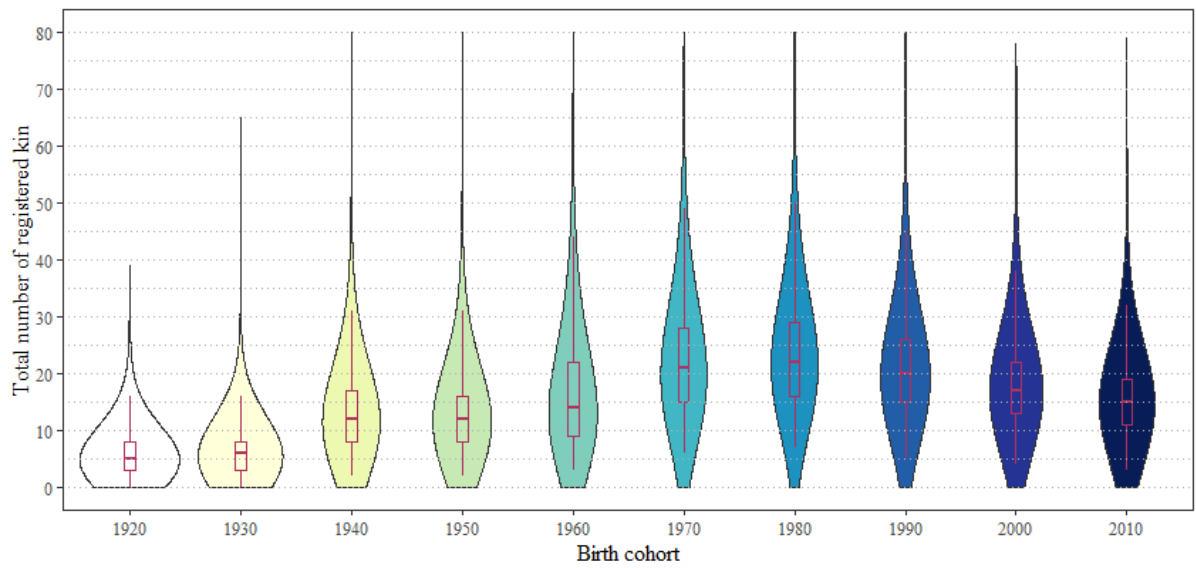
Figure S4: Distribution of the total number of living nieces and nephews for Swedish-born individuals alive in 2018, birth cohorts 1940–2010. Box plots show quartiles and 5th and 95th percentiles.

Figure S5: Distribution of the total number of living grandchildren for Swedish-born individuals alive in 2018, birth cohorts 1920–1970. Box plots show quartiles and 5th and 95th percentiles.

Figure S6: Distribution of the total number of ever-born registered kin for Swedish-born individuals alive in 2018, birth cohorts 1920–2010. Box plots show quartiles and 5ᵗʰ and 95ᵗʰ percentiles.
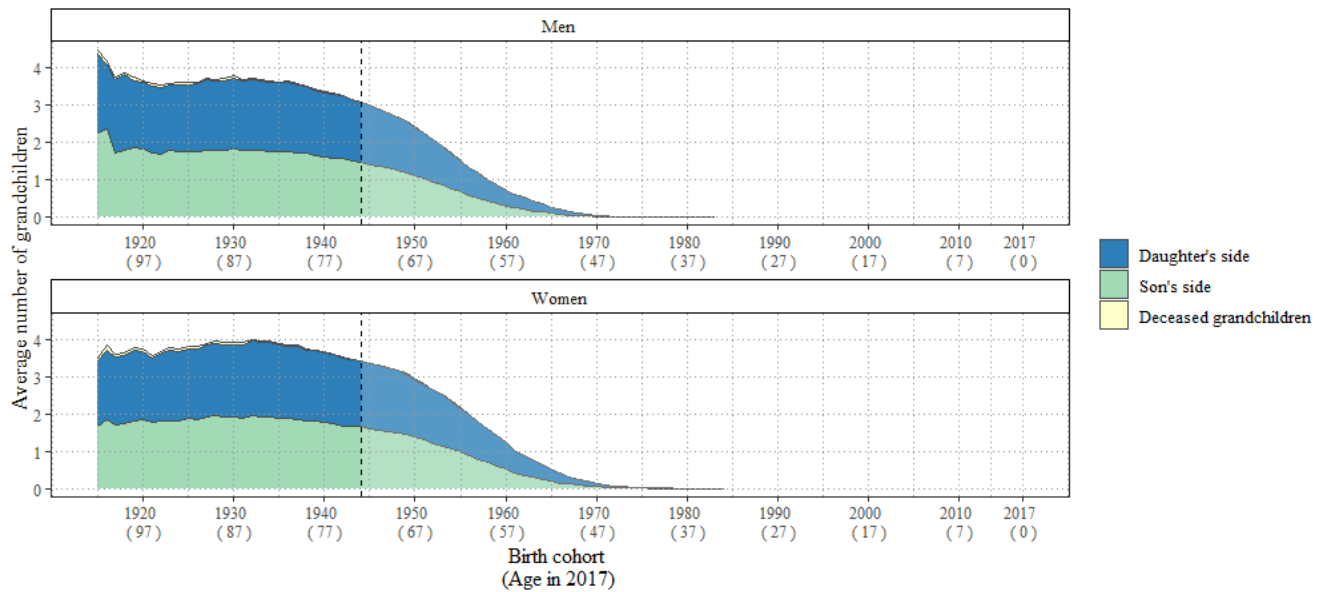
Figure S7: Average number of grandchildren from daughter's and son's side of Swedish born individuals alive in 2018 by, birth cohorts 1920–2017.

.
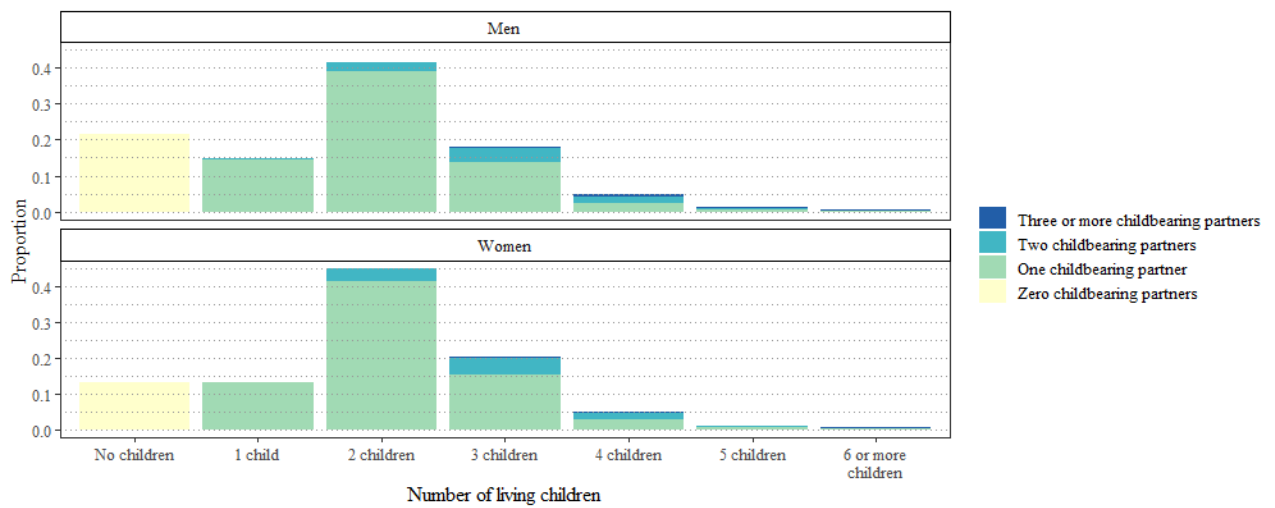
Figure S8: Proportional distribution of living children of Swedish born individuals alive in 2018, by number of children number of childbearing partners. Men and women born 1960-1970.
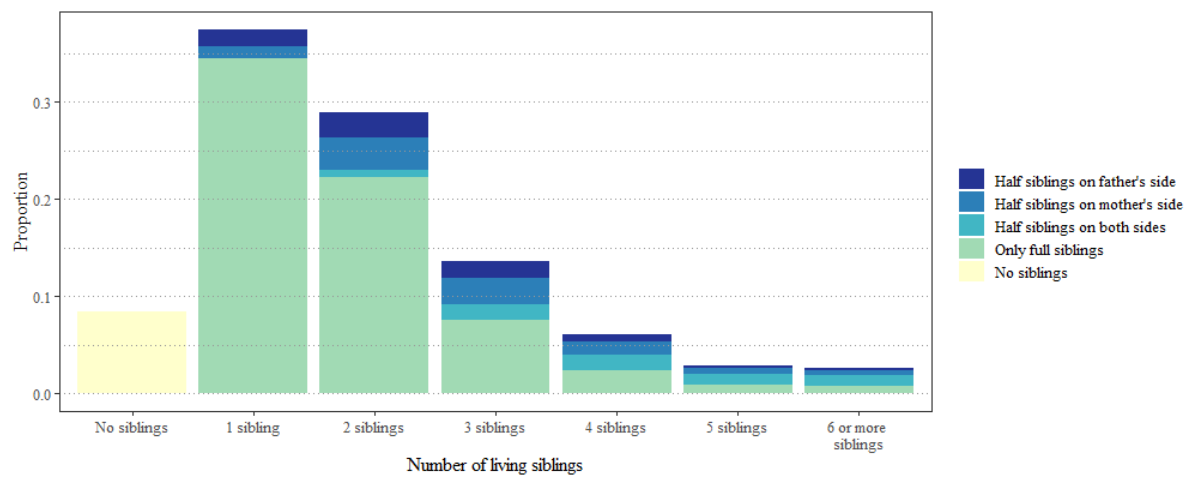
Figure S9: Proportional distribution of living siblings of Swedish born individuals alive in 2018, by maternal and paternal lineage. Men and women born 1960-1970
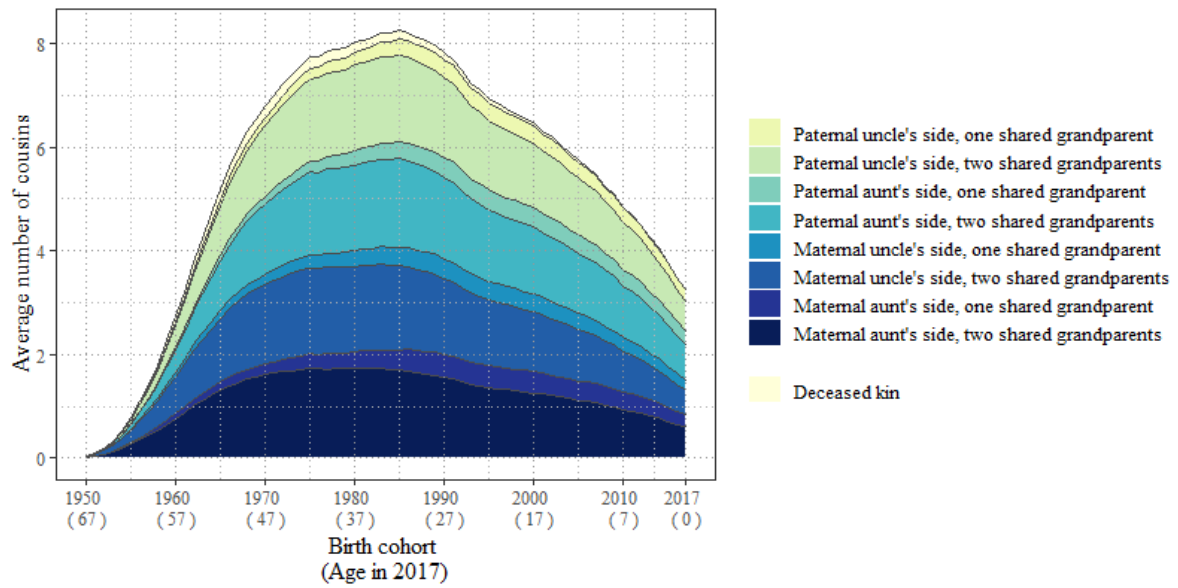
Figure S10: Average number of cousins of Swedish born individuals alive in 2018 by maternal and paternal lineage and by connection via one or two grandparents, birth cohorts 1950–2017.
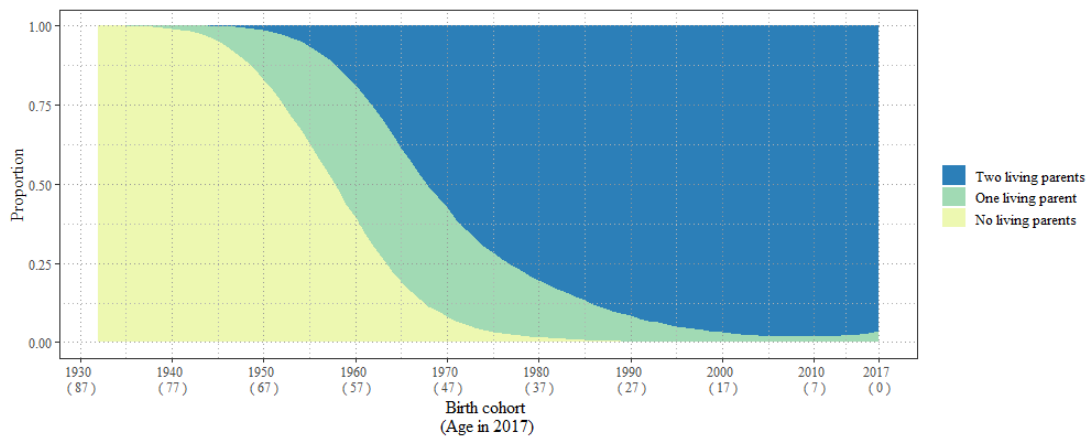
Figure S11: Proportion of Swedish born individuals alive in 2018 with none, one or two living parents, birth cohorts 1930–2017.
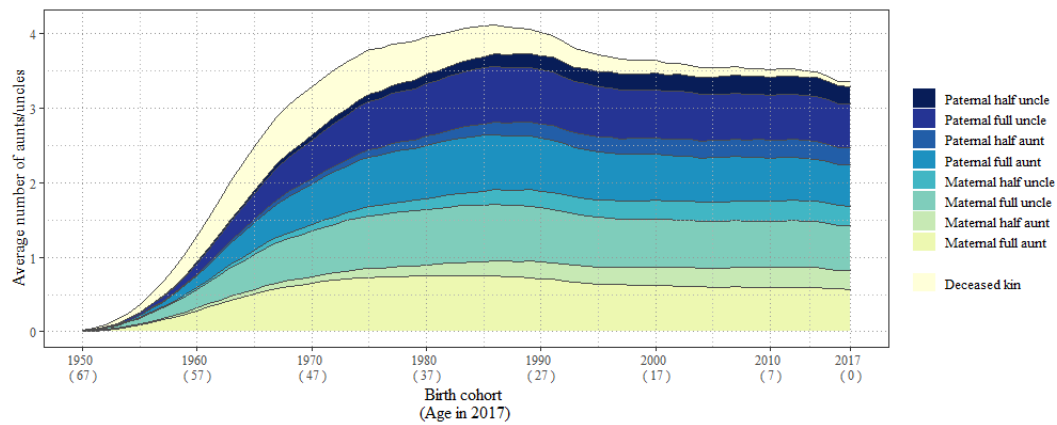
Figure S12: Average number of aunts and uncles of Swedish born individuals alive in 2018 by maternal and paternal lineage, birth cohorts 1950–2017.
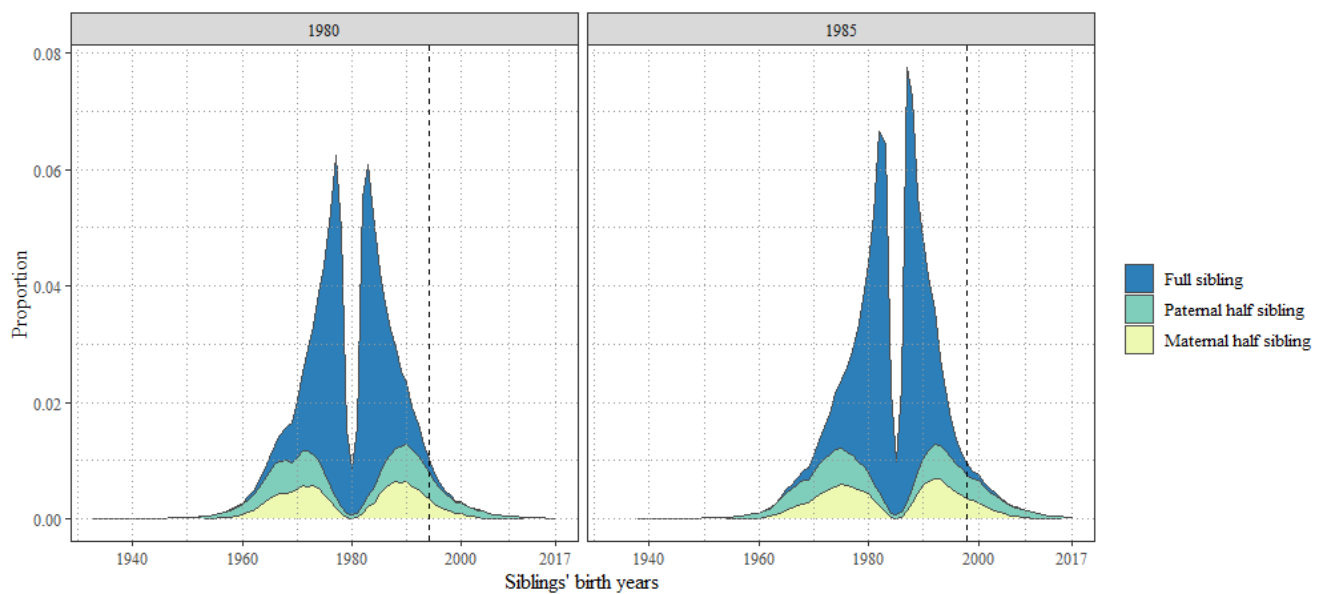

Figure S13: Proportional distribution of living siblings birth year, for Swedish born individuals alive in 2018 born 1980 (left) and 1985 (right). Full siblings and half siblings on mother's and father's side.
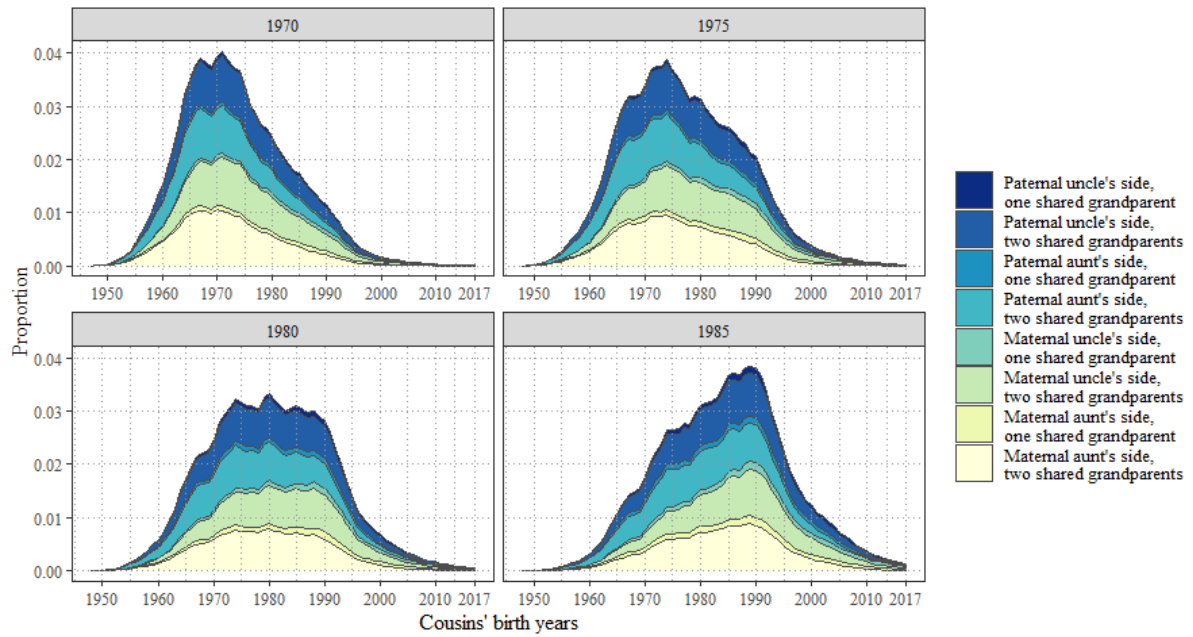
Figure S14: Proportional distribution of living cousins birth years, for Swedish born individuals alive in 2018 born 1970, 1975, 1980 and 1985, by maternal and paternal aunt/uncle and by connection to one or two grandparents.
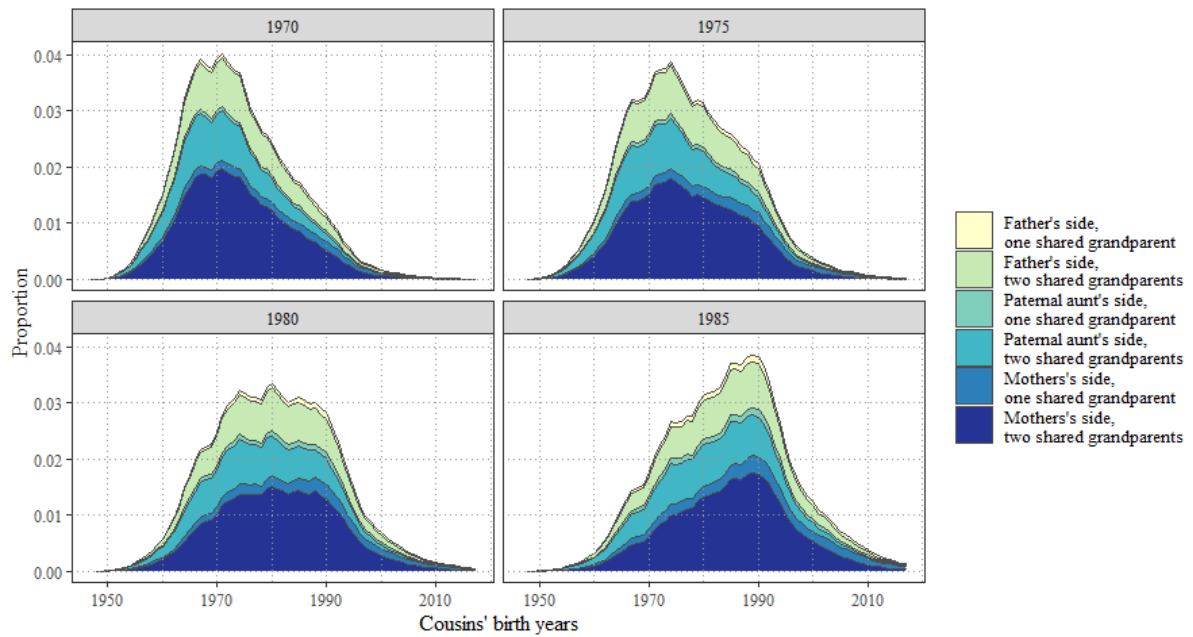

Figure S15: Proportional distribution of living cousins birth years, for Swedish born individuals alive in 2018 born 1970, 1975, 1980 and 1985, by maternal and paternal lineage and by connection to one or two grandparents.

Stockholm
University
Demography Unit