Generational Placement Trajectories in Norway: Combining Empirical and Simulated Data

Bettina Hünteler [1,2] & Diego Alburez-Gutierrez [1]

[1] Kinship Inequalities Research Group, Max Planck Institute for Demographic Research
[2] Institute of Sociology and Social Psychology, University of Cologne

Objective

In the current analysis, we combine empirical and synthetic micro-level data to project generational placement trajectories (GPT) in Norway for the 1953 birth cohort. GPT characterize the structural availability of intergenerational kin over the entire life course of individuals who are currently alive (Hünteler, 2022). Our use of microsimulation lets us overcome data limitations for their analysis, so that we can extend the characterization of life-course kinship structures to contexts, periods, and age ranges for which micro-level register data is not available. Through this, we pave the path towards obtaining results on a global level and from projections into the future.

Background

An increasing body of research is interested in the structural availability of (intergenerational) kin to individuals (Alburez-Gutierrez et al., 2022, 2023; Hünteler, 2022; Leopold & Skopek, 2015; Margolis & Verdery, 2019; Song & Mare, 2019), that is the number and kind of kin alive at a given age. Members of the intergenerational family (i.e., children, parents, and grandparents) are among the main providers and receivers of informal support of different kinds, ranging from emotional to instrumental (care). Family structures vary across individual life courses and between individuals, thus resulting in a different availability of support depending on the life stage and for specific subgroups of the population (Hünteler, 2022). Thus, focusing on average estimations of family structures does not capture the full picture of the support networks over the life course.

Previous research has introduced *generational placement trajectories* (GPT) as a framework to empirically examine such intergenerational life-course family structures (Hünteler, 2022). These trajectories depict a person's relative position within their intergenerational family and describe if, when, and for how long they are a child, parent, or grandparent. At the same time, GPT also define life-course intergenerational family structures because they highlight which family members are alive at a given time. The analysis of GPT requires extensive individual-level demographic data in which members of multiple generations[1] can be linked. Surveys can be used to study GPT, but they face limitations given lower data quality, reporting bias, a focus on a narrow range of kin, and selective participation. Administrative population registers can be used to overcome some of these limitations but they are only available for a few countries and periods. Moreover, both survey and register data are right-censored as of today, so it is unclear whether and when individuals will become grandparents in the future, resulting in an underestimation of the prevalence of and age at grandparenthood (Leopold & Skopek, 2015).

Given these data limitations, a separate stream of research has used demographic microsimulation to model kin availability for a range of kin and geographical contexts (Alburez-Gutierrez et al., 2023; Margolis & Verdery, 2019; Murphy, 2011). These models rely on macro-level input data on fertility and mortality that are available for most world countries. As a result, they can be used to study historical periods and also to project demographic trends into the future relying on the demographic projections that national and international statistical agencies routinely produce. The approach is limited because it ignores individual- and group-level heterogeneity and relies on a series of simplifying assumptions. Nevertheless, studies have consistently shown that model-based estimates of kinship are able to reproduce average patterns of kin availability. It is not known whether demographic microsimulation can be used to study the individual-level GPT that are the focus of this study.

Here, we integrate empirical, register-based, and synthetic, microsimulation-derived, data to examine typical kinship availability across individual life courses in Norway. We proceed in three steps: First, we will analyze GPT based on empirical register data from Norway using sequence and cluster analysis for the 1953 birth cohort. Second, we apply the same methods to a SOCSIM-generated synthetic population for the period in which both sources overlap (1953-2019). This corresponds to ages 0-67 of the 1953 birth cohort. Afterwards, we compare the results of the two data sources and discuss the degree to which the synthetic data captures the individual-level trajectories in the benchmark empirical data. As a third step, we use the microsimulations to project the GPT for the ages 68-100 of the 1953 birth cohort in order to provide an outlook on future structural availability of kin in Norway. Currently, we have

[1] The term 'generation' in this context refers to the genealogical position one occupies in relation to their direct line of ancestors or descendants.

conducted the first step (analyzing the registry data) and we are working on implementing, analyzing, and evaluating the demographic microsimulations.

Research Design

Sequence Analysis to Identify Generational Placement Trajectories

We apply sequence and cluster analysis to identify typical generational placement patterns. We consider six potential 'states,' ranging from not having any intergenerational kin alive, to having parents, children, and grandchildren alive at the same time (see the label of Figure 1 for a description of all the possible stages). Entry and exit into each of these states is determined by three life events ('transitions'): the death of the last remaining parent, the birth of a first child, and the birth of the first grandchild. Individuals can transition through any of these states on an annual basis and we generate one trajectory for each individual. For example, a given person born in 1953 was a *child only* from ages 0-25. At age 25, they had a child, so they were a *child and a parent* in the 26-40 age range. At age 40, both of their parents died, so that they were *parent only* from ages 40-70. At age 70, their child gave birth so that they were *parents and grandparents* over the 71-100 age range.

As a next step, we cluster the trajectories based on their similarity to one another. We use the timing-sensitive Chi-square distance measure (Studer & Ritschard, 2016), to estimate the pairwise similarities and the Partitioning Around Medoids cluster algorithm (Studer, 2013), to consecutively group the trajectories. This combination of methods has produced stable and reasonable cluster results in previous research (Hünteler, 2022; Hünteler et al., 2023). We use these clusters to characterize the typical GPT in the population.

Norwegian Register Data

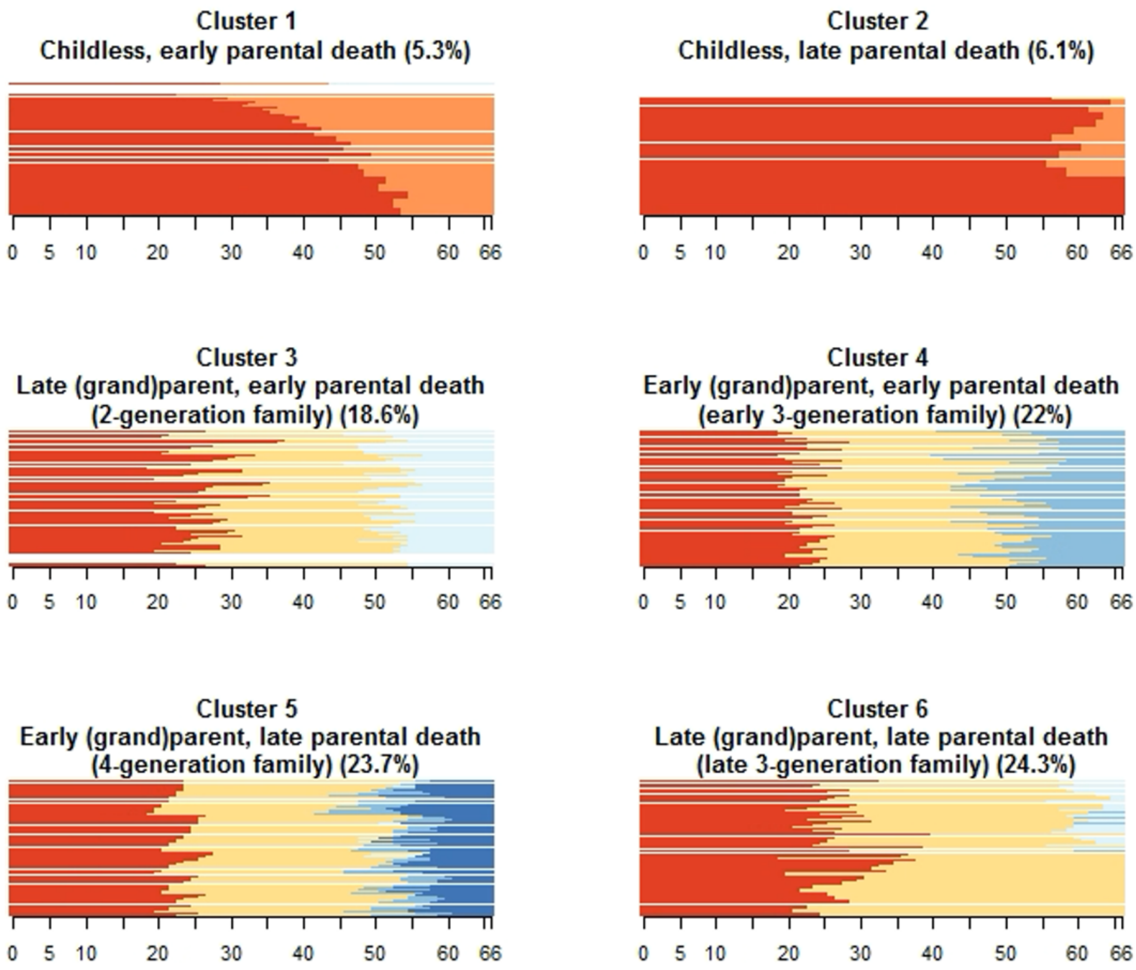
For the 'empirical' part of the analysis, we use data from the Norwegian population register. This contains information on individuals' registration status (alive, dead, emigrated) and demographic information (dates of birth and death, sex) for the population of Norway in a given year. The most recent data is available for 2019. Parents, children, and grandchildren can reliably be identified for individuals born in or after 1953. We use this birth cohort as the basis for our analyses as individuals born in 1953 have the longest-recorded life courses, allowing us to quantify generational placements between birth and age 67, observed between 1953 and 2019.

Demographic Microsimulation

Our projections of GPT relies on demographic kinship microsimulation. Concretely, we use the SOCSIM microsimulator as implemented in the R package *rsocsim* (Theile et al., 2023). SOCSIM takes macro-level age-specific fertility and mortality rates to generate synthetic micro-data with plausible generational structures. With the aim of generating a synthetic version of the Norwegian population registers, we use historical rates for the 1850-2019 period from the Human Mortality Database (HMD, 2023) and the Human Fertility Collection (HFC, 2023). In order to simulate the 2019-2051 Norwegian population, we use the probabilistic projections of demographic rates produced by the United Nations World Population Prospects. We follow the simulation setup of Alburez-Gutierrez and colleagues (2021), in which the simulator stochastically determines whether and when individuals experience specific demographic transitions, such as death or childbirth, on a monthly basis according to the input rates. The final output of the microsimulations is an individual-level rooster of all individuals that have ever lived in the population. The resulting synthetic data includes a column to identify the parents of each simulated individual. This allows us to identify parents, children and grandchildren, and determine the GPT for individuals born in 1953 in the age range 0-100. The microsimulation is based on stochastic models, so we run multiple microsimulations to account for this stochasticity and compare the different outputs. We assess the quality of the simulation by comparing the aggregated demographic rates based on the microsimulations to the input rates of the specific period. Finally, we identify typical life-course generational placement patterns based on the microsimulated individual data using the same methods applied to the register data.

Preliminary Results and Future Work

In a first step, following (Hünteler, 2022), we used the Norwegian microlevel register data for individuals born in 1953 and identified six typical groups of GPT, that were similar within but different from one another (Figure 1). In line with findings from Germany (Hünteler, 2022), two clusters comprised individuals who never became parents and lost their own parents either relatively earlier (5%) vs. later (6%) in their life course, that is at average ages 46 and 60, respectively. One cluster was characterized by a relatively early loss of parents at average age 45, paired with late parent- and grandparenthood at average ages 28 and 58. This patterning resulted in a two-generation family consisting of the sample individuals (G2) and their children (G3) for 19 years, on average (19%). For another cluster, the pattern was the opposite: the birth of children and grandchildren occurred relatively early (23 and 49, respectively) whereas parental death happened later, at age 62, on average. This brought about a four-generation family (22%) in which great-grandparenthood (G1-G4) existed for around 15 years. Lastly, two clusters were characterized by a pattern in which parental death occurred at a similar age as the birth of the first grandchild, so individuals transitioned from a three-generation family



Legend

Existing intergenerational family members					Corresponding generational placement
Individual only		G2			*None*
Individual and parent(s)	G1	G2			*Child*
Individual, parent(s) and child(ren)	G1	G2	G3		*Child & parent*
Individual and child(ren)		G2	G3		*Parent*
Individual, parent(s), child(ren) and grandchild(ren)	G1	G2	G3	G4	*Child, parent & grandparent*
Individual, child(ren) and grandchild(ren)		G2	G3	G4	*Parent & grandparent*

Figure 1 *N*=100 most frequent generational placement trajectories per cluster (Fasang & Liao, 2014)

consisting of their parents (G1), themselves (G2), and their children (G3) to a three-generation family with themselves (G2), their children (G3) and grandchildren (G4). The two clusters differed regarding the timing when they experienced this transition: 24% of all individuals experienced parental death and birth of grandchildren at around age 50, on average, (earlier 3-generation family), while for another 24% these transitions occurred ten years later (later 3-generation family). The results of this step will be used in the second step for comparing against the simulated GPT.

In the coming months, we will generate the synthetic register data to replicate the analyses outlined above. Afterwards, we will compare the output based on the two different data bases. This comparison will include: 1) the number of identified clusters, 2) the main features of the clusters through visual evaluation (state distribution plots, representative sequence plots, etc.), 3) specific cluster characteristics, such as the average age at the family transitions and their variation within each cluster, and 4) the prevalence of each of the typical clusters in the populations under study. Because microsimulation has yielded high-quality results in the analysis of kinship structures in the past, we expect that the GPT and patterns will not differ significantly between the two data sources. Differences that could arise may be related to the fact that heterogeneities in the simulated individual life courses only stem from the stochasticity underlying the microsimulation models and not from an explicit modeling thereof.

As a next step, we will extend the number of intergenerational family members accounted for to also include grandparents (G0) and project the life-course intergenerational family structure into the future to investigate how the GPT evolve over a 100-year lifespan.

Contribution

The current examination contributes to existing research on the analysis and projection of kinship structures by emphasizing their variation across and between life courses and, thus, moves beyond the investigation of average kin counts (Kolk et al., 2023). It provides a first test of the performance of applying the concept of GPT with sequence and cluster analyses to microsimulated data for this field. The study will provide novel evidence on the past, present, and future structural availability of kin in Norway. Beyond this empirical contribution, we expect that it will allow us to extend the analysis of GPT to countries in the global South that lack high-quality micro-data (from surveys or registers) but do have macro-level demographic rate information. This is a first step towards investigating patterns of GPT and their implications for structural kinship support on a global level.

References

Alburez-Gutierrez, D., Barban, N., Caswell, H., Kolk, M., Margolis, R., Smith-Greenaway, E., Song, X., Verdery, A. M., & Zagheni, E. (2022). *Kinship, demography, and inequality: Review and key areas for future development*. SocArXiv. https://doi.org/10.31235/osf.io/fk7x9

Alburez-Gutierrez, D., Mason, C., & Zagheni, E. (2021). The "sandwich generation" revisited: Global demographic drivers of care time demands. *Population and Development Review*, *47*(4), 997–1023. https://doi.org/10.1111/padr.12436

Alburez-Gutierrez, D., Williams, I., & Caswell, H. (2023). *Projections of human kinship for all countries*. SocArXiv. https://doi.org/10.31235/osf.io/hn3zm

Fasang, A., & Liao, T. (2014). Visualizing sequences in the social sciences: Relative frequency sequence plots. *Sociological Methods & Research*, *43*(4), 643–676. https://doi.org/10.1177/0049124113506563

HFC. (2023). *Human Fertility Collection* [dataset]. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). www.fertilitydata.org

HMD. (2023). *Human Mortality Database* [dataset]. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). www.mortality.org

Hünteler, B. (2022). A new perspective on the generational structures of families: Generational placements over the life course. *Advances in Life Course Research*, *51*, 100450. https://doi.org/10.1016/j.alcr.2021.100450

Hünteler, B., Nutz, T., & Wörn, J. (2023). *It all runs in the family? A life course perspective on intergenerational family positions and wealth accumulation*. SocArXiv. https://doi.org/10.31235/osf.io/hxuws

Kolk, M., Andersson, L., Pettersson, E., & Drefahl, S. (2023). The Swedish kinship universe: A demographic account of the number of children, parents, siblings, grandchildren, grandparents, aunts/uncles, nieces/nephews, and cousins using national population registers. *Demography*, 10955240. https://doi.org/10.1215/00703370-10955240

Leopold, T., & Skopek, J. (2015). The demography of grandparenthood: An international profile. *Social Forces*, *94*(2), 801–832. https://doi.org/10.1093/sf/sov066

Margolis, R., & Verdery, A. M. (2019). A cohort perspective on the demography of grandparenthood: Past, present, and future changes in race and sex disparities in the United States. *Demography*, *56*(4), 1495–1518. https://doi.org/10.1007/s13524-019-00795-1

Murphy, M. (2011). Long-term effects of the demographic transition on family and kinship networks in Britain. *Population and Development Review*, *37*, 55–80.

Song, X., & Mare, R. D. (2019). Shared lifetimes, multigenerational exposure, and educational mobility. *Demography*, *56*(3), 891–916. https://doi.org/10.1007/s13524-019-00772-8

Studer, M. (2013). WeightedCluster library manual: A practical guide to creating typologies of trajectories in the social sciences with R. *LIVES Working Papers*, *24*. https://doi.org/10.12682/lives.2296-1658.2013.24

Studer, M., & Ritschard, G. (2016). What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *179*(2), 481–511. https://doi.org/10.1111/rssa.12125

Theile, T., Alburez-Gutiérrez, D., Calderón Bernal, L. P., Snyder, M., & Zagheni, E. (2023). *rsocsim—Socsim microsimulation R package* [C++]. Max Planck Institute for Demographic Research; GitHub. https://github.com/MPIDR/rsocsim