PAA 2024, Columbus, Ohio
April 17-20, 2024

The intergenerational transmission of lifespan and longevity: evidence from the Swedish censuses 1880-1950

Isa Barraclough[1] isa.barraclough@ekh.lu.se

Ingrid van Dijk[1,2] ingrid.van_dijk@ekh.lu.se

1 Centre for Economic Demography, Department of Economic History, Lund University, Sweden

2 Centre for Family History and Historical Demography, Radboud University, Nijmegen, Netherlands

Short abstract (148/250)

Dramatic increases in life expectancy and declines in infant mortality have occurred in recent history, but these benefits have not been transferred equally to all. This study contributes to the existing literature on inequalities in mortality between families by investigating the familial association in lifespan and longevity and its historical variation in Sweden over two centuries 1820-2020. We use full-count census records linked to the Swedish death register. Preliminary analyses show that the correlation between lifespan of grandparents, parents and children is relatively stable across time despite large secular change in lifespan. Correlations are strongest between parents and children, and are stronger in same-sex relations, between grandmothers, mothers and daughters and grandfathers, fathers and sons. In planned work for this paper, we extend the analyses to same-generation relatives and to relatives who are not direct ancestors and use rank-rank associations to account for increases in lifespan across time.

1. Introduction and aim

Dramatic increases in life expectancy and declines in infant mortality have occurred in recent history, but these benefits have not been transferred equally to all those in society. There is a well-recognised SES gradient in mortality in most developed countries today, including Sweden (Mackenbach et al., 2017). Research investigating modern health disparities has revealed an increasingly complex picture of individual health and SES of individuals and their families of origin. We should therefore look beyond individual factors affecting health and consider the influence of an individuals' family network on their lifespan. In this paper, we provide evidence of the correlation in survival in broad family network using Swedish full-count census data. We are among the first to do so, and address long-term change in the role of the family in determining lifespan of individual family members.

The family has a powerful influence on health, providing resources to its members and generating differences in lifestyle through processes of socialization. Yet not much is known about changes over time in the role of the family in the production of lifespan and longevity. This study contributes to the existing literature by investigating the familial association in lifespan and longevity and its historical variation in Sweden over two centuries, using full-count representative census records. Our dataset consists of the 1880-1950 Swedish censuses with linked individual mortality information that allows us to identify mutigenerational vertical and horizontal family ties. Intergenerational elasticities, rank-rank slopes and sibling and cousin correlations are used to capture family associations in lifespan and longevity. Transmission of lifespan at the top and bottom percentiles, indicating advantageous and disadvantageous survival to old ages, will indicate whether there is segregation between the health of families at the top and low ends of survival, and importantly, how the relationship between relatives in mortality and lifespan has changed across time.

2. Background

Whilst literature documenting persistence across generations in outcomes such as education and income has a long history, much less is known about health transmission. Research examining patterns of intergenerational transmission (IGT) of lifespan has been fragmented across various disciplines and dates back over a century (Beeton and Pearson, 1900). Generally, this literature has confirmed the presence of a familial component to human longevity. However, there has been an overemphasis on direct relations between parents and children, to the neglect of broader kinship networks and multigenerational effects within the family (Mare, 2011). Also other relatives than parents may matter, as family resources are also transmitted over more than one generation (Hällsten & Kolk 2023; Knigge 2016). This study expands upon the previous two-generational and nuclear focus on families and includes grandparent-grandchild relations with plans to extend analysis to include wider kinship networks: siblings, cousins, aunts and uncles.

Previous studies on familial lifespan and longevity transmission have generally been unable to consider contextual and historical variation in familial health transmission over long periods of time until today. Our unique dataset allows us to explore whether intergenerational correlations in lifespan have changed over time, taking into account the economic, social and epidemiological developments. It is reasonable to expect that changes in the patterns of diseases, life expectancy and age dispersion at death might have influenced the familial contribution to individual life duration. On the one hand, increased access to health care and improved sanitary conditions might have allowed all individuals to reach older ages reducing the role of family endowments. Alternatively, the shift in main causes of death from infectious diseases to hereditary and lifestyle-related diseases related to family characteristics might have strengthened the familial association.

Finally, there is uncertainty over the role of SES in health inequalities over time, with some believing that these inequalities emerged only with the reduction in large epidemics and improvements in hygiene, nutrition and housing started to take effect. Debiasi and Dribe (2020) showed that the modern mortality gradient by SES has not always existed in Sweden and emerged in the second half of the twentieth century. The ability to track trends and patterns in the familial transmission of health over a long time period allows for greater understanding of the role of families in the emergence of the modern SES gradient.

3. Data and Results

We use a dataset comprising of the 1880, 1890, 1900, 1910, 1930 and 1950 Swedish censuses, containing variables at both the household and individual level including occupation, civil status, year of birth and residence location covering the whole country population. These data are distributed through IPUMS international (The Swedish National Archives, Umeå University, and the Minnesota Population Center). The Swedish censuses are a decadal summary of information from continuous parish registers kept by the clergy in which information for individuals within each household was kept. Due to the continuous collection of information, the data from the Swedish censuses is more reliable and under-enumeration of the population is less problematic than other comparable historical censuses.

Information from the census is linked to the Swedish Death Index (Sveriges Dödsbok) to obtain mortality information. The index contains all deaths occurring in Sweden between 1860 and 2016. The linking of these two datasets relies upon previous work using a probabilistic linkage approach of names and dates of birth. Due to the relatively high quality of both sources, this linkage process obtained a high linkage rate and low rate of false positives (Eriksson, 2015). Individuals can be followed across time between censuses, however in order to link parent-child relations we require them to be present in the same household at least once. The result is a dataset that covers a representative Swedish population over two centuries and allows for the identification of multigenerational and extended familial links with their respective mortality information.

Table 1. Number of observations per census year

Census year	Number of individuals with linked father	Number of individuals with linked mother	Number of individuals with at least one linked grandparent
1880	807,844	774,933	27,005
1890	566,522	537,498	147,414
1900	702,915	682,209	446,043
1910	762,529	748,098	699,876
1930[1]	392,638	394,252	409,933
1950	495,726	526,703	366,804
All years	3,728,173	3,663,693	2,097,075

Note: 1 Only one third of observations for 1930 census have been digitised

Our empirical specification relates the lifespan or longevity of an individual to the lifespan or longevity of one of their relatives or group of relatives. The outcome variable, lifespan is defined as the age at death with birth year controls included to account for the differences in life expectancy between generations. We restrict the sample to those born between 1860 and 1920. This is to remove excess of long-lived individuals in the 1880 census born before 1860, and ensure we have a death date for the last-born individuals of our sample. We focus on individuals who survive to at least age 25, consistent with prior literature in studies of intergenerational correlations of longevity, to study adult survival for those who have reached a childbearing age.

Preliminary results are shown in tables 2 and 3. Table 2 shows the magnitude of transmission of lifespan after age 25 for a pooled sample for all parent-child and grandparent-child relationships and tests the robustness of this relationship to the inclusion of controls for birth year. We find that the raw coefficient from a regression of the lifespan of a son on the lifespan of his father is 0.1, larger than that for a son and his mother at 0.08. For daughters, we see the daughter-mother lifespan coefficient is stronger at 0.09 than the daughter-father relationship at 0.08. There is a strong same-sex bias amongst the parent-child relationships, but this disappears when using the average of the parents' lifespan as a regressor with a coefficient of 0.12 for both sons and daughters. We obtain nearly identical coefficients when we include controls for parent and child birth year.

All grandparent-grandchild coefficients are of a magnitude smaller than parent-child relationships. As with the parent-child relationships, there are stronger same-sex links between grandparents and grandchildren, with the largest coefficients found for grandson-grandfather relationships. Coefficients for grandparents of the maternal line are higher for both grandsons and

granddaughters, although this maternal/paternal line difference is more pronounced for the latter. Controlling for grandparent and grandchild birth year raises the coefficients for all relationships. As we would expect, controlling for parental lifespan in the grandparent-grandchild regression lowers all coefficients suggesting part of grandparent lifespan transmission travels through the parents, however an independent effect of grandparents' lifespan on child lifespan remains.

Table 2: Linear regressions of parent and grandchild lifespan on child's lifespan, Sweden 1860-1920.

	(1)	**(2)**	**(3)**	**(4)**
Model	Lifespan (years) No Controls	(1) + Controls for Birth Year	(2) + Control for Parental Lifespan	# of Obs.[1]
Son-Father	0.095 (0.0011)	0.092 (0.0011)	-	1,330,125
Son-Mother	0.075 (0.0011)	0.074 (0.0011)	-	1,285,583
Son-Parents' Average	0.124 (0.0012)	0.122 (0.0012)	-	1,575,357
Daughter-Father	0.077 (0.0012)	0.073 (0.0012)	-	1,345,114
Daughter-Mother	0.088 (0.0011)	0.088 (0.0011)	-	1,301,060
Daughter-Parents' Average	0.119 (0.0013)	0.119 (0.0012)	-	1,596,474
Grandson-Paternal grandfather	0.025 (0.0021)	0.032 (0.0021)	0.026 (0.0021)	398,543
Grandson-Maternal grandfather	0.027 (0.0020)	0.031 (0.0020)	0.027 (0.0020)	422,140
Grandson-Paternal grandmother	0.011 (0.0020)	0.021 (0.0020)	0.020 (0.0022)	392,351
Grandson-Maternal grandmother	0.012 (0.0019)	0.017 (0.0019)	0.012 (0.0019)	406,555
Granddaughter-Paternal grandfather	0.005[2] (0.0022)	0.019 (0.0022)	0.015 (0.0022)	385,574
Granddaughter-Maternal grandfather	0.014 (0.0021)	0.020 (0.0021)	0.016 (0.0021)	408,767
Granddaughter-Paternal grandmother	0.005[2] (0.0021)	0.024 (0.0022)	0.022 (0.0024)	379,885
Granddaughter-Maternal grandmother	0.020 (0.0020)	0.025 (0.0020)	0.020 (0.0020)	393,784

Notes: Each cell represents the estimate regression coefficient in lifespan, log lifespan with or without controls for the familial relationship indicated in the row. Standard errors are displayed in parentheses.

1 Sample size for grandparent-grandchild regressions decreases marginally when including control for parental lifespan.

2 All coefficients are statistically significant at $p<0.001$ except for those indicated by [2] which are significant at $p<0.100$.

Table 3: Linear regressions of parent-child lifespan by birth cohort, Sweden 1860-1920.

	Birth cohort					
	(1)	(2)	(3)	(4)	(5)	(6)
	1860-1870	**1870-1880**	**1880-1890**	**1890-1900**	**1900-1910**	**1910-1920**
Son-Father	0.114 (0.0046)	0.097 (0.0031)	0.100 (0.0028)	0.086 (0.0024)	0.081 (0.0020)	0.086 (0.0032)
Son-Mother	0.100 (0.0044)	0.079 (0.0030)	0.077 (0.0027)	0.067 (0.0023)	0.065 (0.0019)	0.080 (0.0029)
Son-Parents' Average	0.147 (0.0047)	0.129 (0.0034)	0.125 (0.0029)	0.112 (0.0026)	0.120 (0.0022)	0.125 (0.0036)
Daughter-Father	0.085 (0.0044)	0.074 (0.0031)	0.075 (0.0027)	0.071 (0.0025)	0.064 (0.0022)	0.069 (0.0035)
Daughter-Mother	0.111 (0.0043)	0.095 (0.0029)	0.092 (0.0026)	0.084 (0.0023)	0.077 (0.0020)	0.072 (0.0032)
Daughter-Parents' Average	0.137 (0.0046)	0.122 (0.0033)	0.119 (0.0029)	0.115 (0.0026)	0.107 (0.0024)	0.109 (0.0038)

Notes: Each cell represents the estimate regression coefficient in lifespan with controls for parent and child birth year for the familial relationship indicated in the row and the birth cohort indicated in the column. Standard errors are displayed in parentheses. All coefficients are statistically significant at p<0.001.

Table 3 shows the evolution of the intergenerational transmission of longevity over time for six birth cohorts born between 1860 and 1920. For all parent-child relationships, the IGL is strongest for those born in the earliest birth cohort. There is a trend of increasingly smaller parent-child correlations for each successive cohort apart from the final 1910-1920 cohort where all coefficients increase except for mother-daughter. However, the relative strengths of the different parent-child relationships remain the same in the first and final birth cohorts, with stronger lifespan effects for sons than for daughters.

4. Conclusion and planned work

Preliminary analyses show evidence of consistent but relatively low intergenerational correlations in lifespan for Swedish families. Parent-child intergenerational correlations of lifespan are of a magnitude around 0.8-0.12 with grandparent-grandchild correlations lower around 0.01-0.03. We find an independent effect of the grandparent-grandchild relationships when controlling for parental lifespan, suggesting that grandparents are having a direct effect on the lifespan of their grandchildren or that there are indirect effects not captured in parental lifespan contributing to this relationship. The magnitude of our results is in line with recent work on a historical US population (Black et al., 2023).

Our results contribute to previous work that suggests the intergenerational persistence of lifespan appears to be substantially lower than the comparative persistence in income, education or wealth, which typically ranges between 0.3-0.45 in the literature (Chetty et al, 2014), potentially due to a stronger stochastic element to lifespan. Across all parent and grandparent relationships, our results point to familial lifespan transmission being strongest between family members of the same sex. This may be due to same sex individuals being exposed to the same risk factors to health due to similar occupations and lifestyles whilst some diseases and causes of death are sex-specific strengthening same-sex transmission. Alternatively, the influence of families in affecting healthy or unhealthy behaviours that affect lifespan may have gendered routes of transmission.

In the planned work for this paper, we will expand the analysis to include correlations in lifespan for extended and horizontal kinship relationships, including siblings, cousins, aunts, and uncles. Sibling and cousin correlations tend to be stronger than parent-child correlations as they live in more similar conditions due to closeness in age and account for unobserved factors at the family and community level not shared with parents. Analysis of same-generation relatives also enables better exploration of changing medical and social conditions and changes in life expectancy over time. We further plan to investigate lifespan persistence for individuals at the top and bottom of the lifespan distributions (Van den Berg et al., 2019). Following the social mobility literature, we will also explore alternate empirical specifications until today and e.g. use a rank-rank approach and estimates of period life expectancy by ancestral survival rank instead of the cohort approach presented here. Rank approaches correct for the secular improvement in lifespan across time and life expectancy approaches allow us to include currently living cohorts.

5. Selected references

Black, S., Duzett, N., Lleras-Muney, A., Pope, N., & Price, J. (2023) Intergenerational Correlations in Longevity. NBER Working Paper Series, no. w31034. Cambridge, Mass: National Bureau of Economic Research.

Debiasi, E., & Dribe, M. (2020). SES inequalities in cause-specific adult mortality: a study of the long-term trends using longitudinal individual data for Sweden (1813–2014). European Journal of Epidemiology, 35(11), 1043–1056.

Eriksson, B. (2015). Dynamic Decades: A micro perspective on late nineteenth century Sweden. [Doctoral Thesis, Department of Economic History].

Hällsten, M., & Kolk, M. (2023). The shadow of peasant past: Seven generations of inequality persistence in Northern Sweden. American Journal of Sociology, 128(6), 1716-1760.

Knigge, A. (2016). Beyond the parental generation: The influence of grandfathers and great-grandfathers on status attainment. Demography, 53(4), 1219-1244.

Mackenbach, J. P., Bopp, M., Deboosere, P., Kovacs, K., Leinsalu, M., Martikainen, P., Menvielle, G., Regidor, E., & de Gelder, R. (2017). Determinants of the magnitude of socioeconomic inequalities in mortality: A study of 17 European countries. Health & Place, 47, 44–53.

Mare, R. D. (2011). A Multigenerational View of Inequality. Demography, 48(1), 1–23.

Van den Berg, N., Rodríguez-Girondo, M., van Dijk, I. K., Mourits, R. J., Mandemakers, K., Janssens, A. A., ... & Slagboom, P. E. (2019). Longevity defined as top 10% survivors and beyond is transmitted as a quantitative genetic trait. Nature communications, 10(1), 35.