

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/187251>

Please be advised that this information was generated on 2018-04-11 and may be subject to change.

# Like Mother, Like Daughter. Intergenerational Transmission of Infant Mortality Clustering in Zeeland, the Netherlands, 1833-1912

Ingrid K. van Dijk & Kees Mandemakers

To cite this article: Van Dijk, I.K. & Mandemakers, K. (09-02-2018). Like Mother, Like Daughter. Intergenerational Transmission of Infant Mortality Clustering in Zeeland, the Netherlands, 1833-1912. *Historical Life Course Studies*, Online first, <http://hdl.handle.net/10622/23526343-2018-0003?locatt=view:master>

## HISTORICAL LIFE COURSE STUDIES

Intergenerational transmissions of infant  
mortality using the Intermediate Data  
Structure (IDS)

SPECIAL ISSUE  
2018

GUEST EDITORS

Luciana Quaranta  
Hilde L. Sommerseth



## MISSION STATEMENT

# HISTORICAL LIFE COURSE STUDIES

*Historical Life Course Studies* is the electronic journal of the *European Historical Population Samples Network* (EHPS-Net). The journal is the primary publishing outlet for research involved in the conversion of existing European and non-European large historical demographic databases into a common format, the Intermediate Data Structure, and for studies based on these databases. The journal publishes both methodological and substantive research articles.

### Methodological Articles

This section includes methodological articles that describe all forms of data handling involving large historical databases, including extensive descriptions of new or existing databases, syntax, algorithms and extraction programs. Authors are encouraged to share their syntaxes, applications and other forms of software presented in their article, if pertinent, on the EHPS-Net website.

### Research articles

This section includes substantive articles reporting the results of comparative longitudinal studies that are demographic and historical in nature, and that are based on micro-data from large historical databases.

*Historical Life Course Studies* is a no-fee double-blind, peer-reviewed open-access journal supported by the European Science Foundation (ESF, <http://www.esf.org>), the Scientific Research Network of Historical Demography (FWO Flanders, <http://www.historicaldemography.be>) and the International Institute of Social History Amsterdam (IISH, <http://socialhistory.org/>). Manuscripts are reviewed by the editors, members of the editorial and scientific boards, and by external reviewers. All journal content is freely available on the internet at <http://www.ehps-net.eu/journal>.

### Co-Editors-In-Chief:

Paul Puschmann (Radboud University & KU Leuven) & Luciana Quaranta (Lund University)  
[hislives@kuleuven.be](mailto:hislives@kuleuven.be)

**The European Science Foundation** (ESF) provides a platform for its Member Organisations to advance science and explore new directions for research at the European level. Established in 1974 as an independent non-governmental organisation, the ESF currently serves 78 Member Organisations across 30 countries. EHPS-Net is an ESF Research Networking Programme.



**The European Historical Population Samples Network** (EHPS-net) brings together scholars to create a common format for databases containing non-aggregated information on persons, families and households. The aim is to form an integrated and joint interface between many European and non-European databases to stimulate comparative research on the micro-level.

Visit: <http://www.ehps-net.eu>.



HISTORICAL LIFE COURSE STUDIES  
ONLINE FIRST (2018), published 09-02-2018

# Like Mother, Like Daughter. Intergenerational Transmission of Infant Mortality Clustering in Zeeland, the Netherlands, 1833-1912

Ingrid K. van Dijk  
Radboud University Nijmegen

Kees Mandemakers  
International Institute of Social History, Amsterdam & Erasmus University Rotterdam

## ABSTRACT

The burden of infant mortality is not shared equally by all families, but clusters in high risk families. As yet, it remains unclear why some families experience more infant deaths than other families. Earlier research has shown that the risk of early death among infants may at least partially be transmitted from grandmothers to mothers. In this paper, we focus on the intergenerational transmission of mortality clustering in the Netherlands in the province of Zeeland between 1833 and 1912, using LINKS Zeeland, a dataset containing family reconstitutions based on civil certificates of birth, marriage and death. We assess whether intergenerational transmission of mortality clustering occurred in Zeeland, and if so, whether it can be explained on the basis of the demographic characteristics of the families in which the infants were born. In addition, we explore the opportunities for comparative research using the Intermediate Data Structure (IDS). We find that mortality clustering is indeed transmitted from grandmothers to mothers, and that the socioeconomic status of the family, the survival of mothers and fathers, and the demographic characteristics of the family affected infant survival. However, they explain the heterogeneity in infant mortality at the level of the mother only partially.

**Keywords:** Mortality clustering, Infant mortality, Intermediate Data Structure, Biodemography, Early-life exposure

e-ISSN: 2352-6343  
PID article: <http://hdl.handle.net/10622/23526343-2018-0003?locatt=view:master>  
The article can be downloaded from [here](#).

© 2018, Ingrid K. van Dijk, Kees Mandemakers  
This open-access work is licensed under a [Creative Commons Attribution 4.0 International License](http://creativecommons.org/licenses/), which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit. See <http://creativecommons.org/licenses/>

## 1 INTRODUCTION

The burden of infant mortality is not shared equally by all families: some families lose a considerably larger share of children in infancy than others, a phenomenon known as early-life mortality clustering. In earlier research it has been shown that death clustering has an intergenerational component, as surviving children from high mortality families experience higher mortality among their own infants than individuals who grew up in low mortality families (Lindkvist & Broström, 2006; Vandezande, 2012). However, as yet, it is unknown how family histories of high infant mortality can be explained. The aim of this article is threefold. First, to determine whether intergenerational continuities between grandmothers and mothers in infant mortality occurred in the province of Zeeland, the Netherlands, in the 19<sup>th</sup> and early 20<sup>th</sup> century. Second, to disentangle the contribution of the socioeconomic status of the household in relation to mortality clustering among mothers vis-à-vis mechanisms of intergenerational transmission which work through exposure of mothers to sibling mortality and affecting their reproductive careers. Our third aim is methodological: we use the Intermediate Data Structure (IDS) for our research and reflect on opportunities for comparative research using this format.

Intergenerational transmission in increased mortality among infants has been found for several historical populations (Quaranta et al., 2017). This shows that death clustering is not limited within one generation but has roots in earlier generations. However, in earlier work explanations concerning intergenerational correlation within infant mortality have remained largely unaddressed, and as yet, it is unknown how these family histories of high mortality come about. Several explanations for this phenomenon are possible, which can roughly be divided into indirect and direct effects. On the one hand, grandmothers and mothers may share certain characteristics – including social and behavioral characteristics and genes – which are related to high infant mortality, causing high mortality in multiple generations (Vandezande, 2012). On the other hand, exposure to mortality in the family of origin itself may affect the likelihood that future mothers experience infant mortality (Nettle, Coall, & Dickins, 2011; Störmer & Lummaa, 2014).

In the past, data constraints often limited opportunities for conducting research into intergenerational similarities in infant mortality and survival. Here, we use LINKS, a large-scale historical database which contains up to seven generations of individuals experiencing at least one vital event in the province of Zeeland to reconstruct life courses of grandmothers and mothers and assess mortality hazards of infants in an intergenerational fashion. We aim to assess which factors play a role in the intergenerational transmission of high infant mortality from grandmothers to mothers.

## 2 THEORY

Almost 25 years ago, Monica Das Gupta (Das Gupta, 1990, 1997) recognized that child mortality clustered in a select group of families, despite high overall mortality rates in the Indian region that she was studying. Some families raised all their children to adulthood while others lost many children under the age of five. Although this variation was recognized in earlier research (for an overview, see Das Gupta, 1997 and Edvinsson & Janssens, 2012), Das Gupta's work put death clustering strongly on the research agenda. Differences in maternal ability were identified by Das Gupta as one of the prime causes of differences between families' infant mortality levels among their infants and young children. Over the last decades, it has been shown that early-life mortality does indeed cluster in a select group of high-risk families (Curtis, Diamond, & McDonald, 1993; Janssens, Messelink, & Need, 2010; Lynch & Greenhouse, 1994), including families in the province of Zeeland (Van Poppel, Bijwaard, Ekamper, & Mandemakers, 2012). These findings illustrate that some families face hardships that other families do not, and shift our understanding of early-life mortality from the level of the individual child to the level of the families in which they are embedded.

Earlier research has suggested that there are structural differences between families in which deaths cluster – high risk families – and families losing relatively few or no children. For instance, family size (Zaba & David, 1996), maternal ability (Das Gupta, 1990, 1997), maternal death (Pavard, Gagnon, Desjardins, & Heyer, 2005), remarriage of the mother (Edvinsson, Brändström, Rogers, & Broström, 2005), earlier stillbirths (Edvinsson et al, 2005) and Rh disease (Lundevaller & Edvinsson, 2012) have

been related to families' excess risk of early deaths (for an overview, see [Van Dijk \(in press\)](#)). In addition, it has been shown that increased mortality risk of infants is related between generations in the maternal line, which suggests that family histories of high mortality play a role in mortality clustering ([Vandezande & Matthijs, 2013](#); [Vandezande, Moreels, & Matthijs, 2010](#)). However, earlier work explanations for the intergenerational correlation in infant mortality have remained largely unaddressed. In this paper, we address direct and indirect mechanisms which may affect infant mortality in multiple generations, working through shared social characteristics between grandmothers and mothers or through exposure to mortality in the family of origin, and thus affecting life courses and infant survival.

With regard to the characteristics of the family especially the socioeconomic status of the family may be relevant ([Janssens & Pelzer, 2012](#)). Earlier research has shown that there existed a socioeconomic gradient in infant mortality in the Netherlands, with lower infant mortality among the middle and higher classes than the lower classes ([Van Poppel, Jonker, & Mandemakers, 2005](#)). It is hypothesized that this gradient in infant mortality has increased over the course of the 19<sup>th</sup> century, possibly due to healthier living environments for the upper classes ([Van Poppel et al., 2005](#)) and caused by a slow spread of knowledge about health and access to medical innovations from the upper classes to the lower classes ([Antonovsky & Bernstein, 1977](#)). Some evidence has been found for a relationship between socioeconomic status and death clustering, as variability in child mortality was larger between low-educated and low social class women in India ([Das Gupta, 1990; 1997](#)) and mortality clustering was lower among sharecroppers in 19<sup>th</sup>-century Italy than among landless laborers and non-rural workers ([Scalone, Agati, Angeli, & Donno, 2017](#)). The effect of socioeconomic status may not be limited to one generation, as the socioeconomic status of the grandparents may be reflected in the socioeconomic characteristics of the mother through a process of social reproduction ([Zijdeman, 2009](#)). In addition, socioeconomic conditions of the grandparent may affect health of the mother, including her reproductive health and caretaking capabilities.

Next to effects of socioeconomic status through social reproduction, exposure to high mortality itself may also directly affect future mothers, leading to increased mortality in several generations (see Figure 1). Several pathways may exist. The health of girls raised in high mortality households may be damaged by childhood adversity, including lack of nutrition and exposure to infectious disease in the parental home. Adults whose health is compromised by growing up in high mortality environments may live shorter and unhealthier lives ([Barker, Eriksson, Forsen, & Osmond, 2002](#); [Crimmins & Finch, 2006](#); [Lindeboom, Portrait, & Van den Berg, 2010](#)). As infants and young children are highly dependent on their parents, health and survival of parents is essential for survival of young-age offspring. Furthermore, childbearing capacities of women may be damaged, resulting in reproductive problems, stillbirths and child loss in adulthood ([Lalou, 1997](#)). Thus, exposure to high mortality may be related to high mortality among one's own offspring in adulthood, explaining why intergenerational continuities in high infant mortality between grandmothers and mothers may exist.

Alternatively, evolutionary biology emphasizes that exposure to childhood adversity affects life histories, including reproductive careers and infant mortality. Individuals who are exposed to childhood hardships, such as food shortages, or stressful events in childhood, including mortality of siblings, are expected to employ fast life strategies for reproduction, to increase the chances of successful reproduction in risky environments. These exposures are therefore related to a lower age at menarche, lower age at first marriage, lower age at first birth and shorter spaced births of children ([Nettle et al, 2011](#); [Störmer & Lummaa, 2014](#)). Furthermore, as part of this strategy individuals are expected to invest less resources in individual children and experience higher infant mortality ([Chisholm, 1993](#); [Störmer & Lummaa, 2014](#)). Through the reproductive strategy of the mother, grandmothers' experiences of infant mortality may be linked to survival of their grandchildren.

This paper is part of a larger project which investigates the extent to which intergenerational transmission of mortality clustering occurs in several historical populations in Europe (see also [Broström, Edvinsson, & Engberg \(in press\)](#), [Donrovich, Puschmann, & Matthijs \(in press\)](#), [Quaranta \(in press\)](#) and [Sommerseth \(in press\)](#)). The analyses are partially conducted using a common data base format and are based on a common script for the selection of cases, the construction of key variables, and the analysis of data (see [Quaranta, in press](#)), based on a release of the database LINKS Zeeland in the Intermediate Data Structure (from here: IDS). The IDS stores datasets originating from different sources and databases into a similar structure, which enhances opportunities for comparisons between datasets and therefore between regions and periods ([Alter & Mandemakers, 2014](#); [Alter, Mandemakers, & Gutmann, 2009](#)).

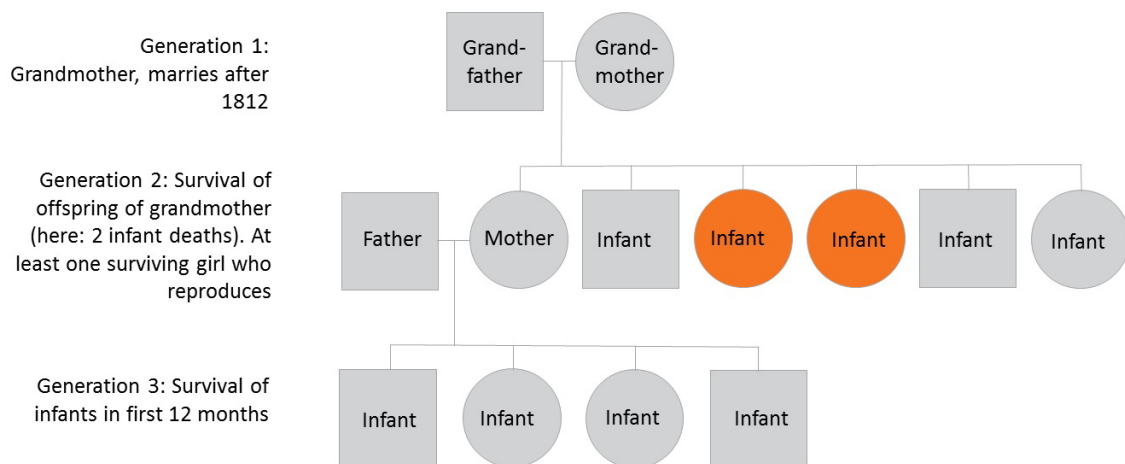
Results of common analyses conducted for several historical populations in Europe will be discussed in more depth elsewhere. However, most importantly, after selecting cases on similar criteria, constructing databases in a similar way, and using identical analyses, we find that there is indeed intergenerational transmission of mortality clustering in several historical populations. These effects are stronger for grandmothers who lost two or more children than for grandmothers who lost one child in infancy. Furthermore, the effects are not explained by demographic characteristics of the family. Because of the similar procedures for data selections, we can exclude that differences between populations are caused by differences in data construction between populations. In the conclusions of this paper, we will discuss our experienced in working with the IDS in the LINKS-Zeeland database.

In the current work, focus lies on intergenerational transmission of mortality clustering and its causes in the province of Zeeland. We contribute to the literature by including the role of socioeconomic status, fast life histories, and the compromised health of mothers in intergenerational family histories of high infant mortality. Now, we turn to the data construction.

### 3 DATA

The data analysis is based on LINKS, a database which contains vital events of the population of Zeeland, the Netherlands, between 1812 and 1912 (Mandemakers & Laan, 2017). The used data are released in the format of the Intermediate Data Structure (IDS) version 4, a common format intended for longitudinal databases concerning historical data (Alter & Mandemakers, 2014).

Figure 1 *Multi-generational influences on infant mortality*



Note: Infants who died before their first birthday are shown in orange.

#### 3.1 HISTORICAL CONTEXT

Zeeland is a largely Protestant province located in the south western coastal region of the Netherlands, and mainly consists of islands. In the 19<sup>th</sup> century, the province was a predominantly rural and agricultural province of which the economy was mainly dependent on market-oriented, capital intensive farming, focused on growing wheat, flax and madder. Although the majority of the population belonged to the unskilled agricultural worker class, wealthy farmers constituted a small but prosperous share of the population (Bras & Neven, 2007). At the end of the 19<sup>th</sup> century, the agricultural sector went through a prolonged crisis, which was especially felt among commercial farmers (Beekink, Boonstra, Engelen, & Knippenberg, 2003). Industrialization occurred relatively late (Blom & Lamberts, 2006), as it only took off in Zeeland after 1900 (Van Poppel et al., 2012).

Until water pipes were constructed in the province of Zeeland from 1910 onwards, in most municipalities rainwater was used as the main source of drinking water. Surface water was often unfit for human consumption due to salinization. Especially during warm summers, there could be a lack of drinking water. In addition, water in canals and cities was often unsafe and its use led to a high prevalence of

waterborne diseases. Malaria was endemic (Hoogerhuis, 2003), and the life prospects of infants and children were short, with a child mortality rate reaching more than 40% in some years (Hoogerhuis, 2003; Van Poppel & Mandemakers, 2002). At the time, the high infant mortality rate in Zeeland was blamed on women who did not breastfeed their children, as they were working outside the house, especially in agriculture (Hoogerhuis, 2003). In the province, the onset of the demographic transition occurred relatively late in comparison to other countries (Hofstee, 1981). Infant mortality began to decline after 1880, but remained high in an international perspective until the early decades of the 20<sup>th</sup> century (Van Poppel et al., 2005).

## 3.2 DATA SOURCES

LINKS stands for *LINKing System for historical family reconstruction*. The project aims at reconstructing all 19<sup>th</sup>- and early 20<sup>th</sup>-century families in the Netherlands, based on a digitized index of all civil certificates from this period called *WieWasWie (Who was Who)*. For over twenty years, numerous volunteers have been working to build the index, which contains not only the names of born, deceased and married persons, but also the names of their parents, places of birth, ages and partly their occupational titles. The database contains vital events of almost 2 million persons from the province of Zeeland.

Obligatory vital event registration was introduced in the Netherlands in 1812. The registers of birth, marriage and death become public with a delay of 100, 75 and 50 years, respectively. Indexes of the registers have been digitized completely for the province of Zeeland and have been released for the period 1812-1912 for births; 1812-1937 for marriages; and 1812-1962 for deaths. Unlike population registers, civil registers do not contain information about religion, addresses, and household composition. At the same time, the civil registers have a large scope, enabling identification of kin over a large area and in several generations. Using family reconstitutions based on civil registries, we can apply an intergenerational approach to demographic research questions. In this case, we can follow vital events from families occurring in Zeeland over time for up to five generations.

In the LINKS dataset, civil certificates of individuals were linked together based on first and last names of pairs of individuals. The linking process consisted of four steps. First, marriage certificates of children were linked to marriage certificates of parents, using the names and ages on the certificates. Second, death records were linked to birth records. Third, birth and death records were linked to the marriage certificate of the parents. Fourth, birth and death records were linked with marriage certificates of individuals. Small inconsistencies and double links – when certificates linked to several other certificates – were resolved by choosing for exact links or by not accepting established matches. To avoid false matches matching was limited within the timeframe on which events logically may occur. For example, matched marriage certificates belonging to women outside their fertile years but with matching parental names were not accepted. Matches could be missed for several reasons: if vital events occurred outside the province of Zeeland, if events occurred outside the time frame for which certificates have been indexed, if names were changed and if variation in spelling of names was too large between civil certificates. Persons who migrated between municipalities in Zeeland could be followed over time; however, individuals who left the province of Zeeland were lost from observation. Analyses have shown that family reconstructions are reliable with regard to family size, age at first birth, and other demographic indicators (Van den Berg et al., 2017).

An exact death date is known for almost half of our sample (47.6% of births). For an additional quarter of the sample (25.0% of births) other follow-up observations exist for vital events such as marriage or childbirth in Zeeland. For 27.3% of the births in the sample there is no follow-up observation. As our data consists of vital event certificates, the end of observation of marriages and deaths in 1937 and 1962 means that individuals who did not marry before 1937 or die before 1962 are not observed in the vital event registration after birth. In addition, individuals who out-migrate from Zeeland before the incidence of vital events are not found in the vital event indexes of Zeeland. Finally, for some individuals a link between their follow-up vital event certificates and birth certificates has not been established. For individuals without follow-up observation, we assume that they survive until their first birthday.

## 3.3 SELECTION OF CASES

To prepare the data for analysis and to conduct the analysis, a common approach is used for several papers. This approach is based on the condition that all data are structured in the same way, the

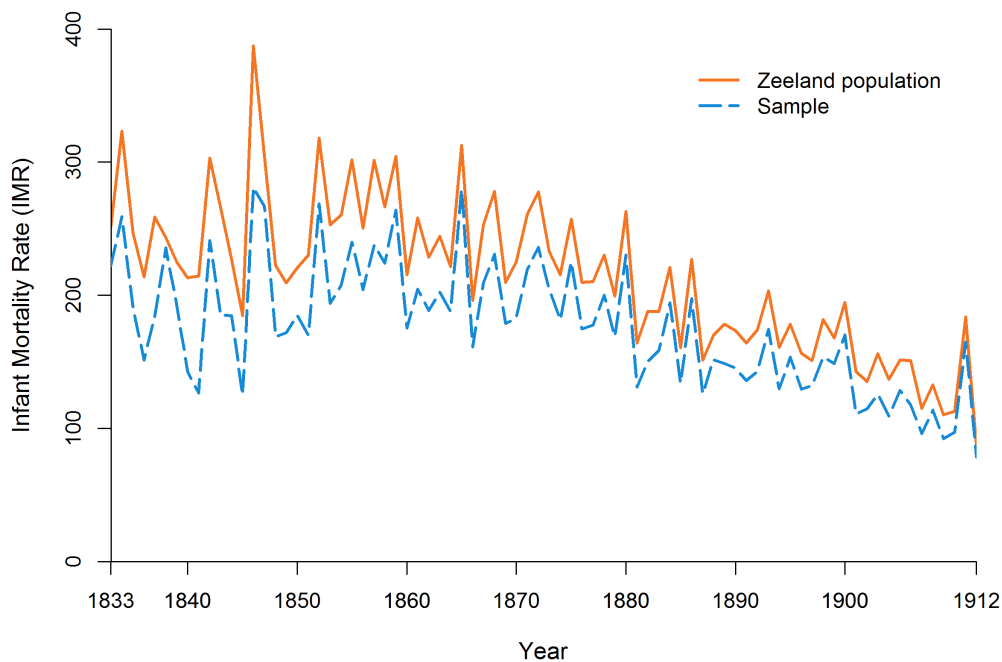


Intermediate Data Structure (Alter & Mandemakers, 2014). On the basis of this structure an extraction program, here developed by Quaranta (in press), selects cases based on similar criteria, builds variables for analysis in an identical fashion, and specifies the basic analyses. For an extended discussion of the data selection and variable construction, see Quaranta (in press).

The analysis focuses on the intergenerational transmission of mortality clustering, researching correlations in familial risk between grandmothers and mothers. Intergenerational transmission of mortality clustering encompasses a paradox. In extreme high-risk families, all children may die before they reach adulthood and/or are able to reproduce. To transmit high risk across generations, there must be a second generation that reproduces itself. Therefore, we conduct our analyses on a select group of families. First, there must be a second generation of mothers, which implies that at least one female child is born (i), who must survive childhood (ii) and reproduce (iii). In addition, grandmothers can only be selected if they were at risk to experience infant mortality and had a surviving, reproducing daughter which implies that they must have produced at least two children (iv).

For our analysis, we select singleton births in the province of Zeeland. The selected infants are the grandchildren of maternal grandmothers who marry in 1812 – the year that vital event registration became obligatory in the Netherlands – or later. The second generation is born from 1812 onwards, and the third generation (grandchildren) from 1833 onwards, which is therefore the starting point of our observations (see Figure 1). Observations of births end in 1912. We include 203,802 infants in the analysis, who are born to 43,582 mothers and 28,118 grandmothers. This means that for a considerable number of grandmothers several daughters are included in the analysis, which is taken into account in a robustness check (see Section 3.3). Lifeless reported infants are not included in the analysis. 4.3% of the births concern lifeless registered infants. These cases encompass both stillbirths and live births of infants who died before they were reported to the municipality, which had to be completed within 3 working days after birth (Vulsma, 1988). Estimates suggest that 71% of lifeless reported infants concern stillbirths and 29% concern live births of infants who died before they were reported to the municipality (Van Poppel et al., 2012). As it is unclear which children were born alive and for how long they lived after birth, we exclude lifeless registered children from the analysis, which will lead to a slight underestimation of infant deaths in the neonatal period and to conservative estimates of the intergenerational relation in infant mortality risk.

Figure 2 IMR Zeeland population and sample LINKS 1833 – 1912



Source: LINKS dataset Zeeland\_2017\_01 (IDS-version).

Figure 2 shows the population infant mortality rate (IMR), based on the indexes of all certificates of births and deaths in the province of Zeeland and the IMR of the sample on which this study is based. Although the IMR in the sample is lower than in the population, changes at the population level are followed rather well, and differences are related to the specific characteristics of the studied sample, due to our selection of infants which can be linked to their mothers and their maternal grandmothers. Children born in 1833 to known grandmothers are born to grandmothers who married in 1812 or later; thus, both their grandmothers and mothers were relatively young, increasing their own likelihood and their offspring's likelihood of survival. In addition, in the early period of our sample relatively strong deviations between the population and sample IMR can be seen; however, here, the number of infants in the sample is still very low, as these are the first children of the first children of the grandmothers. The low number of cases in these years results in relatively strong fluctuations of the sample IMR. Finally, we select infants originating from families who had been in Zeeland for at least two generations – as their grandmothers and mothers married and reproduced in Zeeland – suggesting relatively strong residential stability. These children may have had the added benefit of access to kin networks, which may have contributed positively to child survival through assistance in child-rearing.

### 3.4 VARIABLE CONSTRUCTION

In Table 1 we presented descriptive statistics for the included variables. Our analyses concern survival of the third generation, the infants, with analysis centered on survival between birth and the first birthday. 16.5% of the infants selected in the sample died before their first birthday. The primary explanatory variable is the number of infant deaths of the maternal grandmother, which is categorized as follows: zero deaths among her infants (28.0%), one death (27.0%), or 2 and more deaths (45.0%). We control for the number of births of the grandmother, as grandmothers with more births are more likely to have experienced an infant death. As grandmothers must produce a reproducing daughter and be at risk of experiencing at least one infant death, the minimum of observed births of the grandmother is two. The number of births has been categorized in four groups: 2 births (3.5%), 3 births (5.0%), 4 to 6 births (23.6%) and 7 or more births (67.8%). The birth order of the infant is included, as there usually exists a J-shaped relation between birth order and risk of infant death: earlier and later births have a higher risk of mortality. This variable is similarly categorized, with the addition of a category for first births (20.8%); second birth (17.5%); third birth (14.1%); fourth to sixth birth (27.3%); seventh or higher (19.8%).

Furthermore, we control for age of the mother, categorized into three groups: 15-24, 25-34 and 35-50 years old, as earlier research has shown that the risk of mortality differs between young mothers, older mothers, and average-aged mothers. To measure whether reproduction had started very early, in line with our hypotheses on fast life histories, we added a measure concerning whether mothers had their first child before the age of 20. In Zeeland, very early reproduction is rare: 1.3% of the infants have a mother who had her first child before the age of 20. In addition, we control for sex of the child, as there are substantial differences between boys and girls in survival through their first year of life. The infant mortality rate fluctuates over the course of the 19<sup>th</sup> century, slowly decreasing from 1880 onwards (Wolleswinkel-Van den Bosch, 1998) (see Figure 2). We include birth year of the child, which is centered on the mean birth year 1884.

The socioeconomic status of the grandfather and the father are measured using the HISCAM scale (Lambert, Zijdemans, Van Leeuwen, Maas, & Prandy, 2013). All codes were provided by the *HSN standardized, HISCO-coded and classified occupational titles, release 2013.01* (Mandemakers et al., 2013). The scale was divided into five categories because of possible non-linear relationships between occupational status and infant mortality. The categories were made in such a way that the group sizes were approximately equal, with the exception of the group for the highest statuses, and include the groups low (HISCAM score 41-46), low to average (47-52), average (53-59), average to high (60-69) and high (70-99) socioeconomic status. Socioeconomic status is missing for 849 fathers and 1432 grandfathers and is imputed with the mean. Socioeconomic status was measured using the highest known occupational title of the father and the grandfather. In a robustness check, we have included the lowest socioeconomic status of the grandfather and the father instead of the highest socioeconomic status (see Section 3.2). In addition, we include a measurement for farmer status of the father and the grandfather, which is based on the HISCLASS scheme code 8 (Van Leeuwen & Maas, 2011).

Finally, we include time-varying variables indicating death of the father or the mother through the first year of life of the child. Infants and young children are highly dependent on family members for

survival in their early life, and the death of a parent can therefore detrimentally affect the chances of survival of newborns and infants. We use the date of death of the parent to assess whether a parent died in the first year of life. If no date of death is known, we use the last date of observation of the parent. If the last observation is the birth of the child, we assume that the parent survives the first year of life of the infant.<sup>1</sup>

Table 1 *Descriptive statistics*

	Percentage	Mean	Min	Max
Infant deaths	16.6			
Grandmother child deaths		1.77	0	15
<i>Zero</i>	28.0			
<i>One</i>	27.0			
<i>Two or more</i>	45.0			
Grandmother births		8.55	2	27
<i>Two</i>	3.5			
<i>Three</i>	5.0			
<i>Four to six</i>	23.6			
<i>Seven or more</i>	67.8			
Females	48.7			
Birth order of the child		4.17	1	22
<i>First</i>	20.8			
<i>Second</i>	17.1			
<i>Third</i>	14.0			
<i>Fourth to sixth</i>	27.7			
<i>Seventh or higher</i>	20.4			
Age of the mother		30.5	15.4	50
15-24	20.1			
25-34	56.2			
35-50	23.8			
Mother has first child young (< 20)	1.3			
SES of the father		52	41	99
<i>Low</i>	38.8			
<i>Low to average</i>	28.1			
<i>Average</i>	19.6			
<i>Average to high</i>	9.9			
<i>High</i>	3.7			
Father is a farmer	17.1			
SES of the grandfather		50	41	99
<i>Low</i>	44.7			
<i>Low to average</i>	29.1			
<i>Average</i>	17.3			
<i>Average to high</i>	6.2			
<i>High</i>	2.6			
Grandfather is a farmer	19.6			
Death of the mother	0.89			
Death of the father	0.77			
N of infants	203,802			
N of mothers	43,582			
N of grandmothers	28,118			

Source: LINKS dataset Zeeland\_2017\_01 (IDS-version).

1 For some infants, observations on mothers (5.4%, N=11,052) or fathers (4.5%, N=9,248) end before the first birthday or death of the infant. Here, we assume that the parents did not die before the first birthday or death of the infant.

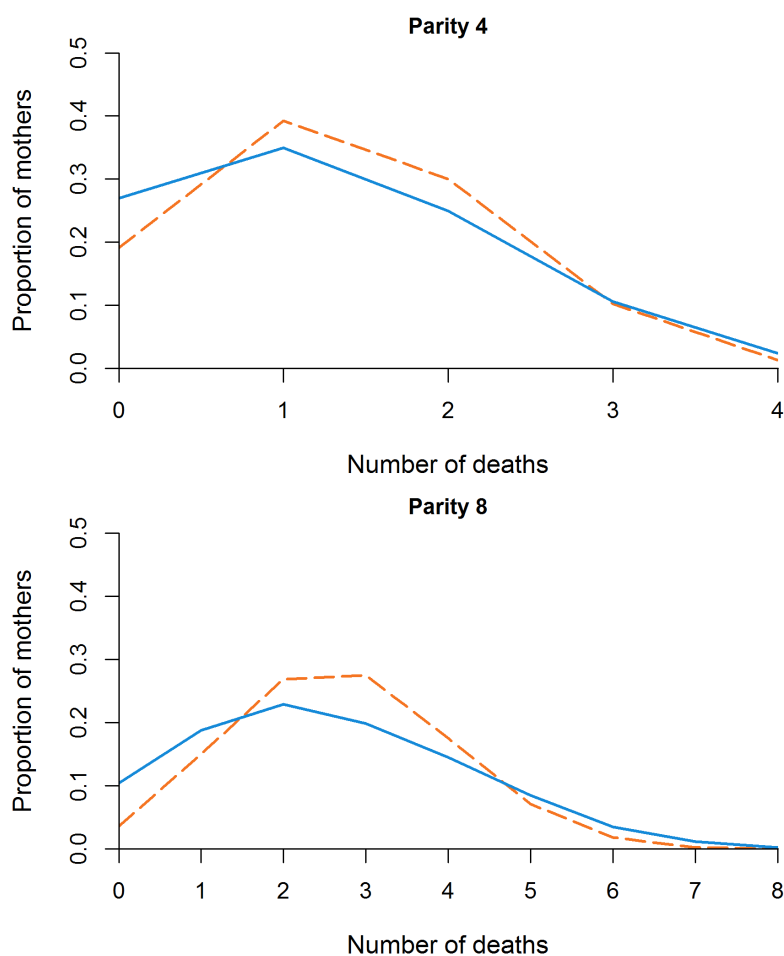
## 4 RESULTS

Analyses are divided into two parts. The first part is based on the code developed by Quaranta (in press) and conducted in Stata 14. Three survival models are estimated: 1) Cox regression models, 2) Weibull models with clustering on the mother, and 3) Weibull models with clustering on the grandmother. For a detailed discussion on methodology, see Quaranta (in press). The models are controlled for basic demographic variables, including age of the mother, the number of births of the grandmother, birth order of the child, and sex. The second part of the analysis is conducted in R 3.32 using the package *coxme*. We estimate frailty models with clustering at the mother level and stratify our analyses by the year of birth of the child. First, we discuss whether we find mortality clustering in families in the Zeeland population.

### 4.1 CLUSTERING OF INFANT MORTALITY

In measuring whether populations experience inequality in the number of early deaths between families, it is essential to take both family size and the binomial distribution of the probability of mortality into account. First, large families can be expected to lose more infants than small families. Furthermore, if early deaths follow a binomial probability distribution, the implication is that not all families will have the same realized mortality rate: there must be lucky large families that escape early death among their infants and unlucky small families who have to bury multiple infants. Even if one in five infants dies before their first birthday, chance predicts that more than 10% of large families with ten children would not lose a single infant. Thus, in assessing whether deaths cluster in the population we should take chance variations and the size of the family into account, and therefore start analysis at the level of the population.

Figure 3 Expected and observed number of families with given number of deaths, by parity



Source: LINKS dataset Zeeland\_2017\_01.

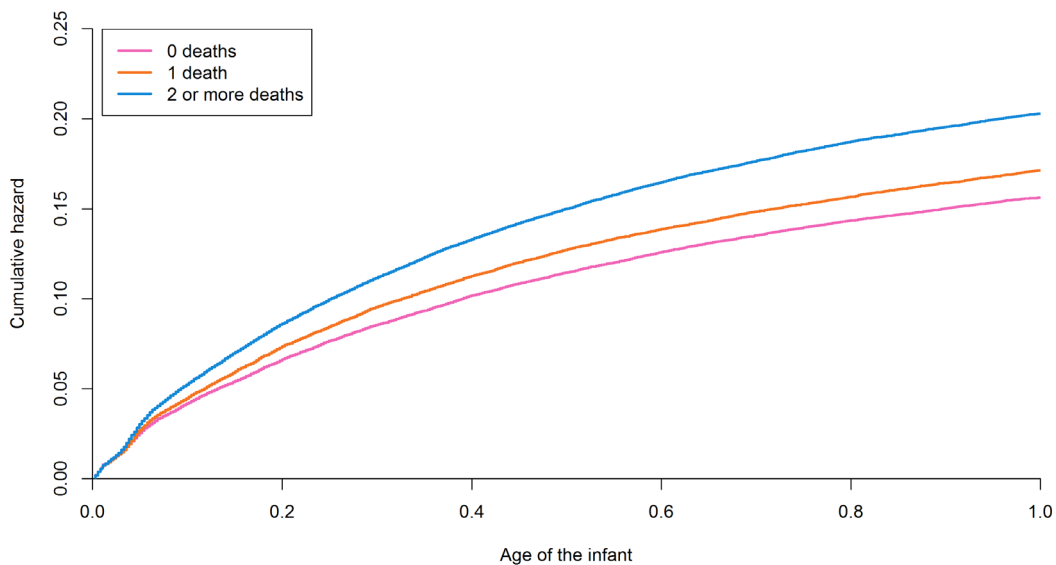
Note: Dashed orange line - expected distribution, filled blue line - observed distribution.

Figure 3 shows the expected and observed proportion of mothers experiencing a certain number of deaths. Here, we present figures for mothers with a total of 4 and 8 births. The expected number of deaths has been calculated using the infant mortality rates in the population; observed numbers are taken from the data. The dashed lines represent the expected number and the filled lines the observed number of mothers with a certain number of deaths. The figure shows that the number of mothers experiencing relatively moderate numbers of deaths is below what would be expected, while there are more mothers experiencing no deaths at all, or relatively high numbers of deaths in their families. These results are in line with earlier, similar explorations (see for instance Zaba & David, 1996), and illustrate that in the province of Zeeland deaths cluster in certain families with a relatively high mortality risk, while other families experience fewer deaths than would be expected or escape deaths among their infants altogether. In the next step, we turn to the data analysis.

## 4.2 MAIN RESULTS

Figure 4 shows the cumulative hazard, based on a Cox proportional hazards model, for infants by grandmothers' number of infant deaths: 0, 1, and 2 or more deaths. If the grandmother lost one infant, their grandchildren have a greater risk of death in infancy; if the grandmother lost two or more infants, this risk accumulates further. Estimated hazard ratios can be found in Table 2. All ratios are relatively stable across all three models estimated: Cox models without shared frailty and Weibull models with mother and grandmother shared frailty. The risk of infant death is 12% higher if the grandmother lost one infant. In case the grandmother lost two or more infants, this risk increases to a 34% higher risk of infant mortality. Thus, our main characteristic of interest, the number of infant deaths of the grandmother, is shown to be related to a higher likelihood of death among infants. The models control for demographic characteristics of the family, including the number of births of the grandmother, sex of the child, birth order, and age of the mother. In addition, we control for changes over time by including a measure for birth year (centered).

Figure 4 Cumulative hazard by number of infant deaths of the grandmother



Source: LINKS dataset Zeeland\_2017\_01 (IDS-version).

Furthermore, the models show that the sex of the child, his or her birth order and the age of the mother contribute to the risk of infant mortality: boys, later-born children and children from young mothers have a higher risk of dying in infancy. Without controlling for other characteristics of the family or child, the relative risk of infant mortality increases with 14% and 40% for infants with grandmothers who experienced one and two or more infant deaths versus grandmothers with zero infant deaths, respectively (results not shown). In addition, although a small share of frailty variance is explained, most variance between mothers in the likelihood that their infants die remains unexplained. Thus, although the demographic characteristics of the family do contribute to the risk of infant mortality they

only partially explain the association between the number of infant deaths of the grandmother and mortality among her grandchildren.

Differences in intergenerational transmission of mortality risk may exist within families. Certain infants within families may be more strongly affected than others as a result of gender preferences or biological differences. In most populations females have a survival advantage over males during their life course. Excess mortality of girls in comparison to boys is therefore often interpreted as the result of gender preferences (Bongaarts & Guilmoto, 2015; Curtis et al., 1993). We have tested whether there is a significant interaction between the sex of the infant and the effects of the intergenerational transmission of mortality risks. If intergenerational continuities in high infant mortality affect especially girls, possibly, the effect can be explained by preferences for boys over girls in multiple generations, leading to high mortality among girls due to infanticide or neglect. For the province of Zeeland we do not find evidence that there is a sex-specific process that might involve gender preferences (results available on request), contrasting some earlier work on survival of adolescent girls in Belgium (Devos, 2000) and 1 to 9 year old girls the Netherlands (Van Poppel, Schellekens, & Walhout, 2009) where especially girls from elite families experienced excess mortality.

Table 2 *Effects of grandmother's infant deaths on infant survival, Zeeland 1833-1912*

		Cox model		Weibull model, mother shared frailty		Weibull model, grandmother shared frailty	
		Hazard ratio	p-value	Hazard ratio	p-value	Hazard ratio	p-value
Infant deaths grandmother	Zero	<i>ref.</i>		<i>ref.</i>		<i>ref.</i>	
	One death	<b>1.12</b>	0.00	<b>1.12</b>	0.00	<b>1.13</b>	0.00
	Two or more deaths	<b>1.34</b>	0.00	<b>1.34</b>	0.00	<b>1.35</b>	0.00
Births of the grandmother	<i>Two births</i>	<i>ref.</i>		<i>ref.</i>		<i>ref.</i>	
	Three births	0.93	0.06	0.93	0.12	0.93	0.10
	Four to six births	<b>0.89</b>	0.00	<b>0.90</b>	0.00	<b>0.89</b>	0.00
	Seven or more births	<b>0.84</b>	0.00	<b>0.84</b>	0.00	<b>0.84</b>	0.00
Sex	Female	<i>ref.</i>		<i>ref.</i>		<i>ref.</i>	
	Male	<b>1.25</b>	0.00	<b>1.26</b>	0.00	<b>1.26</b>	0.00
Birth order	First	<i>ref.</i>		<i>ref.</i>		<i>ref.</i>	
	Second	<b>1.11</b>	0.00	<b>1.12</b>	0.00	<b>1.12</b>	0.00
	Third	<b>1.17</b>	0.00	<b>1.17</b>	0.00	<b>1.17</b>	0.00
	Fourth to sixth	<b>1.30</b>	0.00	<b>1.30</b>	0.00	<b>1.30</b>	0.00
	Seventh or higher	<b>1.67</b>	0.00	<b>1.60</b>	0.00	<b>1.62</b>	0.00
Child birth year, centered		<b>0.99</b>	0.00	<b>0.99</b>	0.00	<b>0.99</b>	0.00
Age of the mother	25-34	<i>ref.</i>		<i>ref.</i>		<i>ref.</i>	
	15-24	<b>1.09</b>	0.00	<b>1.09</b>	0.00	<b>1.09</b>	0.00
	35-50	0.99	0.37	1.02	0.20	1.01	0.43
Intercept				0.17		0.17	
Frailty variance				0.43		0.32	
N of infants		203,802		203,802		203,802	
N of infant deaths		33,763		33,763		33,763	

Source: LINKS dataset Zeeland\_2017\_01 (IDS-version).

### 4.3 THE ROLE OF SURVIVAL OF PARENTS AND SOCIOECONOMIC STATUS OF FATHER AND GRANDFATHER

In Table 3 results are presented for additional analyses using Cox regression models with mother shared frailty. In the first model, we analyze the effect of infant deaths of the grandmother controlling for her number of births and the gender of the infant. In the second model, we control for other demographic characteristics: the birth order of the child, the age of the mother at birth of the child. In addition to the earlier models (Table 2), we add the age of the mother at first birth. In the third model, we add socioeconomic status of the father and the maternal grandfather, farmer status of the father and grandfather, and death of the mother and the father in the first year of life of the infant. All models are stratified by year of birth of the child to account for changes in the likelihood of infant mortality for both grandmothers and infants over time, to ensure that these do not drive effects of the number of deaths of the grandmother on infant survival.

Our main results are in line with our earlier findings. A grandmother's infant deaths are significantly related to the survival chances of her grandchildren. In the second model, where we add control variables, the effect of infant deaths of the grandmother decreases slightly, but remains significant and sizeable (RR 1.12 and 1.34 for one and two or more infant deaths, respectively). The likelihood that an infant will die increases by birth order, while infants from average-aged mothers (25-35, reference category) are better off than those of younger mothers (15-24 year). Infants from relatively old mothers (35-50) do not significantly differ from those from average-aged mothers in their mortality risk, however. Since we controlled for the age group of the mother at the birth of the index child and the birth order of the child, we found that the age of the mother at her first birth affects the likelihood that her infants die: children born to a first time mother who is below the age of 20 are 19% more likely to die during their first year of life. This result is in line with theories of fast life histories and our expectations, but the strength of the effect is remarkable.

In the third model we add socioeconomic status of parents and grandparents. The effect of the number of grandmother infant deaths remain significant. We find that infants from the highest socioeconomic group have a better survival prospect than their peers from lower socioeconomic strata. In addition, infants whose mothers originate from highest-status households have better survival chances. Thus the effect of socioeconomic status can be tracked not only to the family of the infant itself, but also to the socioeconomic background of the mother. In addition, we find that infants whose fathers are farmers are better off than infants whose fathers have non-farming occupations. This effect cannot be found for farming occupations of grandparents.

Furthermore, in Model 3 we include survival of parents, as young children are highly dependent in their first year of life. We find that death of the mother during the first year of life is related to higher chances of infant death (RR: 1.49). On the other hand, we find that paternal death is not related to a higher chance of early-life mortality of infants; rather, the death of the father is related to lower mortality chances of infants (RR: 0.77). This result is in line with earlier research showing that maternal death is more detrimental for infant survival than paternal death (Beekink, Van Poppel, & Liefbroer, 1999; Breschi & Manfredini, 2002) and that death of parents can be compensated for, for instance, by the presence of other kin (Kok, Vandezande, & Mandemakers, 2011), remarriage (Andersson, Hogberg, & Åkerman, 1996) and out-migration (Breschi & Manfredini, 2002). Finally, by including socioeconomic status of fathers and maternal grandfathers and maternal and paternal survival in the model, very little frailty variance between mothers is explained. In addition, the main effect of the number of infant deaths of grandmothers on the survival of grandchildren is only slightly reduced and remains significant. Taken together, these results suggests that these characteristics play a minor role in multigenerational effects in infant mortality.

### 4.4 SENSITIVITY CHECKS

We conducted sensitivity checks for several assumptions with regard to data selection and variable construction. Results are available on request. Robustness checks include selections of infants whose grandmother survives until her 50<sup>th</sup> birthday, infants whose grandmother is observed until her 50<sup>th</sup> birthday or death, infants whose grandmothers' husband survives until her 50<sup>th</sup> birthday, and excluding infants with unknown dates of death or no observations after their birth from the analysis. No substantial differences were observed between our baseline models and the models with alternative case selections.

Table 3 *Effect of grandmother's infant deaths on infant survival, extended models, Zeeland 1833-1912*

		Cox model, mother shared frailty Model 1		Cox model, mother shared frailty Model 2		Cox model, mother shared frailty Model 3	
		Hazard ratio	p-value	Hazard ratio	p-value	Hazard ratio	p-value
Infant deaths grandmother	Zero infant deaths	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>
	One infant death	<b>1.12</b>	0.00	<b>1.12</b>	0.00	<b>1.11</b>	0.00
	Two or more infant deaths	<b>1.34</b>	0.00	<b>1.33</b>	0.00	<b>1.31</b>	0.00
Births of the grandmother	Two births	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>
	Three births	0.92	0.06	0.93	0.09	0.92	0.08
	Four to six births	<b>0.88</b>	0.00	<b>0.89</b>	0.00	<b>0.89</b>	0.00
	Seven or more births	<b>0.83</b>	0.00	<b>0.84</b>	0.00	<b>0.84</b>	0.00
Sex	Female	<i>ref</i>		<i>ref</i>			
	Male	<b>1.25</b>	0.00	<b>1.25</b>	0.00	<b>1.26</b>	0.00
Birth order	First			<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>
	Second			<b>1.11</b>	0.00	<b>1.12</b>	0.00
	Third			<b>1.16</b>	0.00	<b>1.16</b>	0.00
	Fourth to sixth			<b>1.27</b>	0.00	<b>1.27</b>	0.00
	Seventh or higher			<b>1.54</b>	0.00	<b>1.55</b>	0.00
Age of the mother	25-34			<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>
	15-24			<b>1.09</b>	0.00	<b>1.09</b>	0.00
	35-50			1.01	0.40	1.01	0.57
Mother early first birth			<b>1.19</b>	0.00	1.19	0.00	
SES of the father	Low					<i>ref</i>	<i>ref</i>
	Low to average					<b>1.01</b>	0.00
	Average					0.96	0.26
	Average to high					0.95	0.10
	High					<b>0.82</b>	0.00
	Father is a farmer					<b>0.80</b>	0.00
SES of the grandfather	Low					<i>ref</i>	<i>ref</i>
	Low to average					0.99	0.01
	Average					0.98	0.41
	Average to high					0.93	0.02
	High					<b>0.85</b>	0.00
Grandfather is a farmer					0.97	0.09	
Survival of the father					<b>0.77</b>	0.00	
Survival of the mother					<b>1.49</b>	0.00	
	Frailty variance	0.36		0.35		0.34	
	N of infants	203,802		203,802		203,802	
	N of infant deaths	33,763		33,763		33,763	

Source: LINKS dataset Zeeland\_2017\_01 (IDS-version).

Notes: Models are stratified by year of birth of the child.

Furthermore, we divided the analysis into three different time periods: 1833-1874, 1875-1894 and 1895-1912. Children born within each time range were analyzed separately. We find no substantial



differences between these analyses and our extended models. In addition, we have included the lowest socioeconomic status known in the data instead of the highest observed socioeconomic status. Here, we find some differences with the main results. With the lowest socioeconomic status included, we find that each paternal socioeconomic status group has lower infant mortality than the last. For the socioeconomic status of the grandparent or farmer status of father or grandfather, we do not find differences in the results. Our substantive conclusions do not change, however: a higher socioeconomic status of the father reduces the likelihood that infants die; however, this effect does not explain the relation between infant deaths of the grandmother and risk of death of the infant. In addition, it does not explain the frailty variance, or the differences between mothers in their chances of infant mortality.

Finally, instead of using a HISCAM scale of socioeconomic status we have included socioeconomic status using HISCLASS to investigate whether effects of socioeconomic status are class-specific beyond the difference between farmers and non-farmers. Here, we find that infants of higher professionals, lower clerical and sales personnel and farmers and fishermen are relatively well-off in comparison to elite groups, whereas children of lower-skilled farm workers have increased odds of dying in infancy. As for the social class of the maternal grandfather, we find that grandchildren of lower skilled workers and unskilled and lower skilled farm workers have a significantly higher risk of dying in infancy. However, these effects do not explain the relation between grandmothers' infant deaths and the risk of death of grandchildren. It appears that, comparable to the main results we presented, the main dichotomy exists between farmers and non-farmers, and higher classes versus lower classes. Thus, our substantive conclusions remain unaltered.

## 5 DISCUSSION AND CONCLUSION

In this paper, we have analyzed the intergenerational transmission of infant mortality risk between maternal grandmothers and mothers, using a large dataset from Zeeland, The Netherlands. The aim of this paper was threefold: first, to determine whether there existed intergenerational transmission of mortality risk in Zeeland; second, whether this intergenerational transmission can be explained by socioeconomic characteristics of the family, health of parents, and fast life histories; and third, to assess the advantages and potential disadvantages of conducting comparative research using the Intermediate Data Structure (IDS).

### 5.1 INTERGENERATIONAL TRANSMISSION OF MORTALITY RISK

In our analysis we have focused on intergenerational continuities between grandmothers and mothers in mortality among their infants. In line with earlier research, we have found that risk of infant deaths of grandmothers and mothers is correlated. This was shown in previous research on 19<sup>th</sup>-century Antwerp (Vandezande, 2012) and Skellefteå (Lindkvist & Broström, 2006). Analyses were conducted using Cox and Weibull survival models with shared frailty at the level of the mother or grandmother. All analyses led to similar conclusions, as did the sensitivity analyses: the risk of infant mortality is correlated between generations. Especially for grandmothers experiencing two or more deaths among their infants the risk of infant mortality of their grandchildren is higher. We found that these results hold after controlling for the number of births of the grandmother, birth order of the child, and several demographic characteristics of their families. Thus, we showed that there is a familial component in early-life mortality in 19<sup>th</sup> and early 20<sup>th</sup>-century Zeeland.

With regard to the mechanism explaining the relation between mothers and daughters in their experiences with infant mortality, we used Cox regression analyses to model the intergenerational correlation in infant mortality between grandmothers and mothers, incorporating several explanatory characteristics. We incorporated socioeconomic status of maternal grandfathers and fathers, including whether they were farmers or not; survival of parents during the first year of life of infants and early onset of reproductive careers.

In the current work we did not find evidence that survival of parents plays an important role in explaining intergenerational transmission of high infant mortality. We expected that survival of mothers could play an important mediating role, as mothers exposed to high mortality in early life may be unhealthier in adulthood (Crimmins & Finch, 2006) and maternal survival is important for survival

of infants and children. Although the death of mothers has a strong negative impact on the likelihood that their infants survive their first year of life, we found no evidence that parental survival explains the intergenerational relations between grandmothers and mothers in infant mortality.

We did find some preliminary evidence that fast life histories may play a role, as demographic characteristics of the family – including the birth order of the child and the age of the mother at the time of her first birth and at the time of birth of the infant – were found to partially explain the relation between mortality in the family of the mother and among her own children. At the same time, effects appear to be relatively minor: after including demographic characteristics, the effect of infant mortality of the grandmother remains strong and significant, and only a small share of the frailty variance between mothers is explained. Yet, lower investment in children, having more children and marrying earlier may all contribute to high-mortality patterns among women originating from high-mortality environments (Nettle et al., 2011; Störmer & Lummaa, 2014). In addition, sociocultural reasons could play a role: women from disadvantaged environments, characterized by early fertility and high infant mortality, may be socialized to pursue disadvantageous fertility careers, leading to high mortality in several generations (see, for an example, Laslett, 1980). In future research, more attention should be paid to the effects of mortality in the sibling set on mortality in the second generation, which may be mediated by marriage age and birth intervals of the mother.

Socioeconomic status affects mortality of infants – with lower infant mortality among farmers and in the highest status groups. Not only the socioeconomic status of the mother was found to affect her infant's survival, but the socioeconomic status of the grandfather mattered as well. However, socioeconomic status played only a small role in explaining mortality clustering among mothers. This result contrasts with work on late 20<sup>th</sup>-century Northeast Brazil, which has illustrated that in some regions socioeconomic status plays an important role in explaining concentration of early-life mortality in households (Sastry, 1997). For textile cities in 19<sup>th</sup>-century the Netherlands, however, no evidence was found that mothers from the working classes experienced higher mortality among their infants than mothers from other classes (Janssens & Pelzer, 2012). In addition, earlier historical work showed that the higher classes did not necessarily do better than the lower classes with regard to infant mortality, which may be explained by the use of wet nurses (Lalou, 1997). This contrasts to our findings, as we find a lower risk of infant mortality among the highest classes in Zeeland.

Preferences for boys over girls may lead to higher mortality among girls in several generations, at least partially due to neglect of girls. In the current research, such a preference would be indicated by intergenerational patterns of high infant mortality at the expense of survival of female infants. However, in contrast with earlier work (Devos, 2000; Van Poppel et al., 2009), we did not find evidence for family-patterns in sex-specific infant mortality, as boys and girls were similarly affected by high mortality among their grandmothers' infants.

In conclusion, we have emphasized the role of demographic characteristics in explaining infant mortality clustering in families, showing that there exists an intergenerational component to mortality clustering. However, questions about the causes of these intergenerational continuities remain. In future research, a further attempt should be made to disentangle the causes of these intergenerational continuities. While we have shown that socioeconomic status and demographic characteristics of the family cannot explain these similarities, other possible explanations include genetic factors and intergenerational stability in living conditions. Families who have lived in the same place for multiple generations may be subject to similar living conditions over time, resulting in intergenerational continuities in mortality among infants. On the other hand, if they migrate out of their region of origin, they may be subject to different mortality regimes, resulting in intergenerational change in the mortality among their children. Furthermore, in the current analysis we were unable to include infants born in families which had out-migrated from Zeeland. Possibly, out-migrants' infant mortality differs from stayers' mortality patterns, leading to possible under- or overestimations of intergenerational transmission of mortality clustering. Finally, future research should attempt to incorporate causes and timing of death, to assess whether intergenerational continuities extend to these characteristics as well.

## 5.2 USING IDS FOR A COMMON ANALYSIS

This paper is part of a project in which several research groups have used a common script for selection of cases, construction of variables and analysis, based on a common data format known as the Intermediate Data Structure (IDS). The IDS format provides a way to store datasets based on different sources into a similar structure, with the purpose of facilitating comparative research in historical demography. However, not all issues of comparative research can be solved by storing data in a similar format and constructing variables in similar ways. Each historical demographical data base has been constructed from different sources, with different purposes, and with varying requirements for inclusion of persons. Although the IDS structure ascertains that database structures are identical across datasets, which guarantees that the program for extraction of variables and analysis will run with every dataset, different characteristics of underlying databases may lead to problems in case selection, variable construction and analysis. Designing a common program for analysis is a complicated process, as approaches to data selection and variable construction relevant in one database may be less relevant in another, and necessary requirements specific to another database may be overlooked, possibly leading to problems downstream.

In the current work, analyses were based on three-generation information, including grandmothers, mothers, and children. For a precise selection of mothers and grandmothers with complete life course information, ideally, data with continuous observations from the date of birth or in-migration until the date of death or out-migration, such as data from population and parish registers, is used. With such data, decisions on sample selection can be based on the length of observation of individuals and their relatives, in this case concerning children, their mothers, and their grandmothers. However, for other types of datasets, including family reconstitutions based on vital event registration such as LINKS, no continuous observations are available and we cannot observe individuals and their relatives in a specific window of time, but we rely on more fragmented observations. As case selection based on the observation window is not possible, for such data, it is more difficult to select cases for data analysis.

A second problem arises when datasets contain persons who are related to others and satisfy the conditions for selection, but at the same time do not contain meaningful observations. For instance, in the LINKS database names of parents are given on certificates of birth, death and marriage. Thus, the database identifies these parents, and contains a record for them, but the case only contains relational information and a name. Therefore, to select a case in LINKS, observing a linked grandmother is insufficient: we have to know more about her than just her name. In several other commonly used historical demographical datasets similar problems could occur, for instance because of a sample-design to which incomplete life course data of significant others has been added, or because in multigenerational pedigrees ancestors are identified but not followed over time.

Thus, using the IDS format and common scripts for comparative research it is not always easy to decide which cases should be selected for analysis. To make the right selection for an analysis such as the one conducted in this paper, one has to ascertain that relevant characteristics are known for selected persons. To ease this problem, it would be useful if IDS structured databases would include a characteristic distinguishing relatively complete cases – key actors – from cases which do not contain meaningful observations. This indication may be based on the source of information about this person – the way the existence of this person has been verified – and the available data for this person in the database. Second, scripts could be developed which add these information to the IDS dataset, the so-called extended IDS (Quaranta, 2016). Third, extraction software for comparative research on IDS structured databases should be developed with care to ensure that the specific characteristics of each database are accounted for.

Whereas using a common script for data construction was not without problems, the advantages of using a common approach are also manifold. Within the current project intergenerational transmission of infant mortality between grandmothers and mothers has been approached in a comparable way between several historical populations. This has allowed new insights into the extent to which family histories of high mortality are relevant for survival of infants, providing valuable new insights into the determinants of infant survival in the past.

## FUNDING

This work was supported by the following grant: the Netherlands Organisation for Scientific Research (grant number 360-53-180).

## REFERENCES

- Alter, G., & Mandemakers, K. (2014). The Intermediate Data Structure (IDS) for longitudinal historical microdata, version 4. *Historical Life Course Studies*, 1, 1-26. <http://hdl.handle.net/10622/23526343-2014-0001?locatt=view:master>
- Alter, G., Mandemakers, K., & Gutmann, M. P. (2009). Defining and distributing longitudinal historical data in a general way through an intermediate structure. *Historical Social Research/Historische Sozialforschung*, 34(3), 78-114. doi:10.12759/hsr.34.2009.3.78-114
- Andersson, T., Hogberg, U., & Åkerman, S. (1996). Survival of orphans in 19th century Sweden the importance of remarriages. *Acta Paediatrica*, 85(8), 981-985. doi: 10.1111/j.1651-2227.1996.tb14198.x
- Antonovsky, A., & Bernstein, J. (1977). Social class and infant mortality. *Social Science & Medicine* (1967), 11(8-9), 453-470. doi: 10.1016/0037-7856(77)90022-1
- Barker, D. J., Eriksson, J. G., Forsen, T., & Osmond, C. (2002). Fetal origins of adult disease: strength of effects and biological basis. *International Journal of Epidemiology*, 31(6), 1235-1239. doi: 10.1093/ije/31.6.1235
- Beekink, E., Boonstra, O., Engelen, T., & Knippenberg, H. (2003). *Nederland in verandering: maatschappelijke ontwikkelingen in kaart gebracht*. Amsterdam: Aksant.
- Beekink, E., Van Poppel, F., & Liefbroer, A. C. (1999). Surviving the loss of the parent in a nineteenth-century Dutch provincial town. *Journal of Social History*, 32(3), 641-669. doi: 10.1353/jsh/32.3.641
- Blom, J. C., & Lamberts, E. (2006). *Geschiedenis van de Nederlanden*. Amsterdam: Bert Bakker.
- Bongaarts, J., & Guilimoto, C. Z. (2015). How many more missing women? Excess female mortality and prenatal sex selection, 1970–2050. *Population and Development Review*, 41(2), 241-269. doi: 10.1111/j.1728-4457.2015.00046.x
- Bras, H., & Neven, M. (2007). The effects of siblings on the migration of women in two rural areas of Belgium and the Netherlands, 1829–1940. *Population studies*, 61(1), 53-71. doi: 10.1080/00324720601048319
- Breschi, M., & Manfredini, M. (2002). Parental loss and kin networks: demographic repercussions in a rural Italian village. In R. Derosas & M. Oris (Eds.). *When dad died: Individuals and families coping with distress in past societies*. Bern: Peter Lang, 369-387.
- Broström, G., Edvinsson, S., & Engberg, E. (in press). Intergenerational transfers of infant mortality in 19th century northern Sweden. *Historical Life Course Studies*.
- Chisholm, J. S. (1993). Death, hope, and sex: Life-history theory and the development of reproductive strategies [and comments and reply]. *Current anthropology*, 34(1), 1-24. doi: 10.1086/204131
- Crimmins, E. M., & Finch, C. E. (2006). Infection, inflammation, height, and longevity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2), 498-503. doi: 10.1073/pnas.0501470103
- Curtis, S. L., Diamond, I., & McDonald, J. W. (1993). Birth interval and family effects on postneonatal mortality in Brazil. *Demography*, 30(1), 33-43. doi: 10.2307/2061861
- Das Gupta, M. (1990). Death clustering, mothers' education and the determinants of child mortality in rural Punjab, India. *Population studies*, 44(3), 489-505. doi: 10.1080/0032472031000144866
- Das Gupta, M. (1997). Socio-economic status and clustering of child deaths in rural Punjab. *Population studies*, 51(2), 191-202. doi: 10.1080/0032472031000149906
- Devos, I. (2000). Te jong om te sterven: de levenskansen van meisjes in België omstreeks 1900. *Tijdschrift voor sociale geschiedenis*, 1, 55-75.
- Donovich, R., Puschmann, P., & Matthijs, K. (in press). Mortality clustering in the family: Fast life history trajectories and the intergenerational transfer of infant death in late 19<sup>th</sup>- and early 20<sup>th</sup>-century Antwerp, Belgium. *Historical Life Course Studies*.

- Edvinsson, S., Brändström, A., Rogers, J., & Broström, G. (2005). High-risk families: The unequal distribution of infant mortality in nineteenth-century Sweden. *Population studies*, 59(3), 321-337. doi: [10.1080/00324720500223344](https://doi.org/10.1080/00324720500223344)
- Edvinsson, S., & Janssens, A. (2012). Clustering of deaths in families: Infant and child mortality in historical perspective. *Biodemography and social biology*, 58(2), 75-86. doi: [10.1080/19485565.2012.738575](https://doi.org/10.1080/19485565.2012.738575)
- Hofstee, E. W. (1981). *Korte demografische geschiedenis van Nederland van 1800 tot heden*. Haarlem: Fibula-Van Dishoeck.
- Hoogerhuis, O. W. (2003). *Baren op Beveland. Vruchtbaarheid en zuigelingensterfte in Goes en omliggende dorpen gedurende de 19e eeuw*. Wageningen: Wageningen Universiteit.
- Janssens, A., Messelink, M., & Need, A. (2010). Faulty genes or faulty parents? Gender, family and survival in early and late childhood in the Netherlands, 1860–1900. *The History of the Family*, 15(1), 91-108. doi: [10.1016/j.hisfam.2010.01.005](https://doi.org/10.1016/j.hisfam.2010.01.005)
- Janssens, A., & Pelzer, B. (2012). Did factory girls make bad mothers? Women's labor market experience, motherhood, and children's mortality risks in the past. *Biodemography and social biology*, 58(2), 133-148. doi: [10.1080/19485565.2012.720451](https://doi.org/10.1080/19485565.2012.720451)
- Kok, J., Vandezande, M., & Mandemakers, K. (2011). Household structure, resource allocation and child well-being. A comparison across family systems. *Tijdschrift voor Sociale en Economische Geschiedenis/ The Low Countries Journal of Social and Economic History*, 8(4), 76-101. doi: [10.18352/tseg.346](https://doi.org/10.18352/tseg.346)
- Lalou, R. (1997). Endogenous mortality in new France: At the crossroads of natural and social selection. In A. Bideau, B. Desjardins, & H. Pérez Brignoli (Eds.). *Infant and child mortality in the past*. Oxford: Clarendon Press.
- Lambert, P. S., Zijdeman, R. L., Van Leeuwen, M. H., Maas, I., & Prandy, K. (2013). The construction of HISCAM: A stratification scale based on social interactions for historical comparative research. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(2), 77-89. doi: [10.1080/01615440.2012.715569](https://doi.org/10.1080/01615440.2012.715569)
- Laslett, P. (1980). The bastardy prone sub-society. In P. Laslett, K. Oosterveen, & R. M. Smith (Eds.). *Bastardy and its comparative history: studies in the history of illegitimacy and marital nonconformism in Britain, France, Germany, Sweden, North America, Jamaica and Japan*. Cambridge: Harvard University Press, 217-246.
- Lindeboom, M., Portrait, F., & Van den Berg, G. J. (2010). Long-run effects on longevity of a nutritional shock early in life: The Dutch Potato famine of 1846–1847. *Journal of health economics*, 29(5), 617-629. doi: [10.1016/j.jhealeco.2010.06.001](https://doi.org/10.1016/j.jhealeco.2010.06.001)
- Lindkvist, M., & Broström, G. (2006). *Clustered infant mortality in an intergenerational perspective, Skellefteå 1831-1890*. Paper presented at the SSHA, Minneapolis.
- Lundevaller, E., & Edvinsson, S. (2012). The effect of Rh-negative disease on perinatal mortality: Some evidence from the Skellefteå region, Sweden, 1860-1900. *Biodemography and social biology*, 58(2), 116-132. doi: [10.1080/19485565.2012.720450](https://doi.org/10.1080/19485565.2012.720450)
- Lynch, K. A., & Greenhouse, J. B. (1994). Risk factors for infant mortality in nineteenth-century Sweden. *Population studies*, 48(1), 117-133. doi: [10.1080/0032472031000147506](https://doi.org/10.1080/0032472031000147506)
- Mandemakers, K., & Laan, F. (2017). *LINKS dataset Genes Germs and Resources. WieWasWie Zeeland. Civil Certificates. version 2017.01. IDS version. [Data file and code book]*.
- Mandemakers, K., Muurling, S., Maas, I., Van de Putte, B., Zijdeman, R. L., Lambert, P., Van Leeuwen, M.H.D., Van Poppel, F., & Miles, A. (2013). HSN standardized, HISCO-coded and classified occupational titles. *IISG Amsterdam*.
- Nettle, D., Coall, D. A., & Dickins, T. E. (2011). Early-life conditions and age at first pregnancy in British women. *Proceedings of the Royal Society B: Biological Sciences*, 278(1712), 1721-1727. doi: [10.1098/rspb.2010.1726](https://doi.org/10.1098/rspb.2010.1726)
- Pavard, S., Gagnon, A., Desjardins, B., & Heyer, E. (2005). Mother's death and child survival: the case of early Quebec. *Journal of biosocial science*, 37(2), 209-227. doi: [10.1017/S0021932004006571](https://doi.org/10.1017/S0021932004006571)
- Quaranta, L. (2016). STATA programs for using the Intermediate Data Structure (IDS) to construct files for statistical analysis. *Historical Life Course Studies*, 3, 1-19. <http://hdl.handle.net/10622/23526343-2016-0001?locatt=view:master>
- Quaranta, L., Broström, G., Van Dijk, I., Donrovich, R., Edvinsson, S., Engberg, E., Mandemakers, K., Matthijs, K., Puschmann, P., & Sommerseth, H. (2017). *Intergenerational transfers of infant mortality in historical contexts: a comparative study of five European populations*. Paper presented at the SSHA, Chicago.

- Quaranta, L. (in press). Program for studying intergenerational transmissions of infant mortality using the Intermediate Data Structure (IDS). *Historical Life Course Studies*.
- Quaranta, L. (in press). Intergenerational transfers in infant mortality in southern Sweden, 1740-1968. *Historical Life Course Studies*.
- Sastry, N. (1997). Family-level clustering of childhood mortality risk in Northeast Brazil. *Population studies*, 51(3), 245-261. doi: [10.1080/0032472031000150036](https://doi.org/10.1080/0032472031000150036)
- Scalone, F., Agati, P., Angeli, A., & Donno, A. (2017). Exploring unobserved heterogeneity in perinatal and neonatal mortality risks: The case of an Italian sharecropping community, 1900–39. *Population studies*, 71(1) 1-19. doi: [10.1080/00324728.2016.1254812](https://doi.org/10.1080/00324728.2016.1254812)
- Sommerseth, H. L. (in press). Intergenerational transfer in infant mortality in northern Norway during the 19<sup>th</sup>- and early 20<sup>th</sup>-century. *Historical Life Course Studies*.
- Störmer, C., & Lummaa, V. (2014). Increased mortality exposure within the family rather than individual mortality experiences triggers faster life-history strategies in historic human populations. *PLoS one*, 9(1). doi: [10.1371/journal.pone.0083633](https://doi.org/10.1371/journal.pone.0083633)
- Van Dijk, I.K. (in press). Early-life mortality clustering in families: A review of the literature. *Population Studies*.
- Van den Berg, N., Van Dijk, I., Mourits, R., Mandemakers, K., Slagboom, P. E., & Janssens, A. (2017). *Families in Comparison: An exploration of results of family reconstitutions using population registers and vital event registrations*. Paper presented at the IUSSP, Capetown.
- Van Leeuwen, M. H., & Maas, I. (2011). *HISCLASS: A historical international social class scheme*. Leuven: Universitaire Pers.
- Van Poppel, F., Bijwaard, G. E., Ekamper, P., & Mandemakers, K. (2012). Historical trends in the correlation of sibling deaths in infancy in the Netherlands. *Biodemography and social biology*, 58(2), 87-115. doi: [10.1080/19485565.2012.720448](https://doi.org/10.1080/19485565.2012.720448)
- Van Poppel, F., Jonker, M., & Mandemakers, K. (2005). Differential infant and child mortality in three Dutch regions, 1812–1909. *The Economic History Review*, 58(2), 272-309. doi: [10.1111/j.1468-0289.2005.00305.x](https://doi.org/10.1111/j.1468-0289.2005.00305.x)
- Van Poppel, F., & Mandemakers, K. (2002). Sociaal-economische verschillen in zuigelingen- en kindersterfte in Nederland, 1812-1912. *Bevolking en gezin*, 31(2), 5-40.
- Van Poppel, F., Schellekens, J., & Walhout, E. (2009). Oversterfte van jonge meisjes in Nederland in de negentiende en eerste helft twintigste eeuw. *Tijdschrift voor sociale en economische geschiedenis*, 6(4), 37-69. doi: <http://doi.org/10.18352/tseg.469>
- Vandezande, M. (2012). *Born to die. Death clustering and the intergenerational transmission of infant mortality, the Antwerp district, 1846-1905*. Leuven: Centrum voor Sociologisch Onderzoek.
- Vandezande, M., & Matthijs, K. (2013). Inherited dimensions of infant mortality. Detecting signs of disproportionate mortality risks in successive generations. *The History of the Family*, 18(2), 169-186. doi: [10.1080/1081602X.2012.755131](https://doi.org/10.1080/1081602X.2012.755131)
- Vandezande, M., Moreels, S., & Matthijs, K. (2010). Explaining death clustering. Intergenerational patterns in infant mortality, Antwerp, 1846-1905. *Working paper of the Scientific Research Community Historical Demography*, 13.
- Vulsma, R. (1988). *Burgerlijke stand en bevolkingsregister*. Den Haag: Centraal Bureau voor Genealogie.
- Wolleswinkel-Van den Bosch, J. H. (1998). *The epidemiological transition in the Netherlands*. Rotterdam: Erasmus University. [hdl.handle.net/1765/17663](https://hdl.handle.net/1765/17663)
- Zaba, B., & David, P. H. (1996). Fertility and the distribution of child mortality risk among women: an illustrative analysis. *Population studies*, 50(2), 263-278. doi: [10.1080/0032472031000149346](https://doi.org/10.1080/0032472031000149346)
- Zijdeman, R. L. (2009). Like my father before me: Intergenerational occupational status transfer during industrialization (Zeeland, 1811-1915). *Continuity and Change*, 24(3), 455-486. doi: [10.1017/S0268416009990166](https://doi.org/10.1017/S0268416009990166)