UNIVERSITÀ DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Statistiche
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE STATISTICHE
CICLO XXVII

CHILDREN SURVIVAL IN VENETO 1815-70
From the dark age to the dawn of change

Direttore della Scuola: Prof.ssa Monica Chiogna

Supervisore: Prof. Gianpiero Dalla Zuanna

Co-supervisori: Prof. Francesco Billari, Prof. Adriano Paggiaro

Dottorando: Leonardo Piccione

31/01/2016

Abstract

The infant mortality rate is universally used as a measure of population health, and it is particularly useful in contexts where limited resources require easily calculated measures. When studying underdeveloped countries or particular historical contexts, for example, an analysis of the dynamics driving to high levels of infant mortality is crucial in order to have a better picture of the society and the population. Since the death of a child is the last and worst outcome of a number of individual, familiar and community variables operating in the very initial part of life, a life course perspective is a good way to approach this issue.

In the thesis it is analyzed and discussed the effect of early life experiences in the short term, using an unpublished 19th century dataset collected in 46 parishes of the North-Eastern Italian region of Veneto. The first part of the study is devoted to the production of classical demographic tools such as life tables, while the second part is a collection of statistical analyses of increasing complexity aimed to increase the knowledge about the special case of Veneto. Large space is given to the debate about the existence of a mortality selection process, in favor of which the thesis provides a series of empirical proofs. Lastly, focussing on a subset of the dataset, the work discusses the role of the daily temperatures on the risk of dying in the first five years of life.

Sommario

Il tasso di mortalità infantile è universalmente utilizzato come indice del benessere di una popolazione, ed è particolarmente utile in contesti in cui risorse limitate richiedono misure calcolate in modi semplici. Nello studio di paesi sottosviluppati o in particolari contesti storici, ad esempio, l'analisi delle dinamiche che causano alti livelli di mortalità infantile è fondamentale per avere un quadro completo della società e della popolazione in analisi. Poiché la morte di un bambino è l'esito estremo di una serie di variabili individuali, familiari e comunitarie che operano nella primissima parte della vita, una prospettiva di 'corso di vita' è un buon modo per approcciare la questione.

Nella tesi viene analizzato e discusso l'effetto nel breve termine di esperienze vissute all'inizio della vita, usando un inedito dataset comprendente 46 parrocchie venete del periodo 1815-70. La prima parte dello studio è dedicata alla produzione di strumenti demografici classici (quali le tavole di mortalità), mentre la seconda parte presenta una serie di analisi statistiche di crescente complessità, orientate ad accrescere la nostra comprensione riguardo il caso specifico del Veneto. Ampio spazio è dedicato al dibattito riguardo l'esistenza di un processo di 'selezione per mortalità', in favore del quale la tesi fornisce diverse prove empiriche. Infine, concentrandosi su un sottocampione del dataset, il lavoro tratta il ruolo delle temperature giornaliere sulla probabilità di sopravvivenza nei primi cinque anni di vita.

Acknowledgements

It is very difficult to imagine my life without the morning when prof. Gianpiero Dalla Zuanna asked me if I wished to apply for a PhD position. Maybe he was a little bit unaware of the consequences, and so was I. But now I really want to thank him for that offer and for all that came next, from the academic support to the personal trust, from the scientific suggestions to the cycling discussions. Prof. Adriano Paggiaro is another person with whom I have communicated a lot, especially towards the end of the project: his push has been fundamental for the final sprint.

Padova has been the only constant in these four years of crazily changing variables. I wish to thank the city, the University and the Department of Statistics for having been really cozy places, and all the persons I have met in Padova for having been great travel companions. In particular, all my dear friends from my second house, the 'Residenza Cornaro', and, of course, my colleagues from the XXVII cycle: Gianluca, with his generosity and his thirst for IPAs and disputes; Vera, with her punctual curiosity and her human willingness; Marco, with his inspiring willpower; Paola, with her joyful pragmatism; Maereg, with is colorful composure; Valentina, with her lively restlessness. I am proud to have shared with them so many hours of the unique portion of our lives when we can become everything but we are still nothing. Actually, all of them have become something in the while; me, I am still searching (and it seems I am loving it pretty much: two time extensions, oh yeah).

Oxford has been a huge changing point not only in my PhD, but eventually in my entire existence. I want to thank prof. Francesco Billari and the Department of Sociology for the patience and the hospitality. Also, my months in the city of dreaming spires would have not been the same without the lovely Maria and Julie, and without the glorious Oxford University Basketball Club.

And now it comes to my old, tiny Corato, my feud and my refuge. My family is where I have always reloaded my batteries, it is the place where I can show my flaws without shame: I want to thank my parents and my brothers for bearing my worst. I can not forget the pillars represented by the people from 'Vivere In', with whom I have been walking for ages (and surely I will keep doing).

I am definitely forgetting someone and something, so I reserve the right to extend in pencil this list of significant places, persons and objects that I am going to mention here: thank to the Icelandic volcanoes, to the Alpe d'Huez, to all my beer-mates, to Nara's deers, to the Oregon Coast, to my rented bikes, to...

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Along with the multiple existing measures of life expectancy, the infant mortality rate is universally used as a measure of population health. Albeit its usage has been often criticized, it remains an important indicator of health for whole populations, and not only for a small part of it, especially in contexts where limited resources require easily calculated measures (Reidpath and Allotey 2003). Moreover, an infant mortality decline has usually been considered one of the decisive factors leading to the modern fertility decline and, consequently, to the demographic transition of a population (Van de Walle 1986; Haines 1997). These two considerations, highlighting the strong bond between lower infant mortality and higher development rate, make easier to understand the space devoted to infant mortality studies both in contemporary and historical demography.

Over the last few decades, in fact, several studies have sought to explain historical, individual and community differences in the risk of dying during the earliest stages of life. Thus far, however, two important patterns remain less than fully explained. First, although the risk of dying during the first year of life during the ancien regime was everywhere higher than 10%, considerable differences existed between populations living in adjacent regions, countries and parishes. Secondly, also the secular trends may have significantly varied among neighbouring populations. Although these issues are not easily disentangled, they are crucial to understand the historical determinants of infant mortality.

The demographic literature proposes a number of studies aimed to identify the most important determinants of infant mortality, the most influential of them focussing on the relations between medical, social and cultural factors (Schultz 1979; Palloni 1981;

Mosley and Chen 1984). The emerging analytical framework can be summarized by the following three large categories of socioeconomic determinants:

- **individual level variables**, e.g. mother's health and nutrition during pregnancy and lactation, personal characteristics, parental childcare;

- **household level variables**, essentially related to the family income (e.g. food, clothing, hygiene, sickness care);

- **community level variables**, e.g. ecological setting, political economy, health system.

It is particularly interesting - and increasingly adopted - to observe the interactions of these variables and their eventual outcome from a life course perspective. Actually, the spread of life course epidemiology has responded to the limitations of classic aetiological models (Kuh et al. 2003): the key point is that it exists a critical period at the very beginning of life (or even before birth), in which some particular exposures can permanently alter particular organ structures and metabolic functions (Braveman and Barclay 2009). Albeit this approach has been successfully adopted to explain some specific levels of morbidity and mortality during adult life, it may also be useful to reduce the temporal gap between the 'exposure' time and the 'follow-up' time, focussing then on the effects (the most tragic of them being death) of pre-natal and early-life experiences in a very short term (i.e. during infancy), rather than in a remote future.

In effect, this choice makes sense. Studying complex relations between biological and social variables over time is far from being straightforward. Issues arise both in the *causal inference* theoretical debate and from the *statistical analysis* point of view (De Stavola et al. 2006): temporal (and possibly causal) hierarchies among the exposures need to be taken into account. The risk structure can be represented in the 'simple accumulation of risk' form, in which exposures are independent (Fig. 1.1, model a) or clustered (b), or - more likely - in the more complex 'chain of risk' form, where each exposure has an effect both on a subsequent exposure and on the disease risk (c) or only on the subsequent exposure (d). In all the models, three types of variables must be distinguished: background variables, intermediate effects and outcomes. It is evident that the longer the time between background variables and outcomes, the greater the possibility of intervening intermediate effects. Long story short, reducing the time gap between background variables and outcomes, it is possible to avoid the troubles due to intervening intermediate effects.

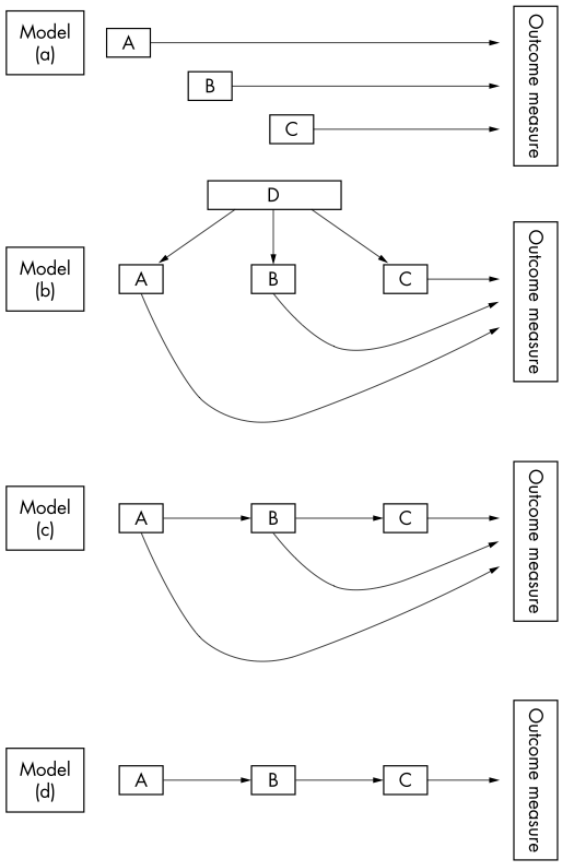


FIGURE 1.1: Life course causal models. Source: Kuh et al. 2003

Having highlighted that one of the key points for studying infant mortality is the identification of a early moment during which one or more 'negative experiences' take place, some issues remain still unsolved. Empirical evidences from the literature show that it is generally unclear which direction the influence of negative experiences during the critical periods will take (Hobcraft and Gilks 1984, quoted in Billari and Rosina 2000). It is possible that epidemic and famine experiences during early life debilitate a cohort that will later experience higher death rates than older or younger cohorts. On the other hand, others argue that high mortality experienced in early life selects out the frail members of the cohort, resulting in reduced subsequent mortality. With particular respect to the mortality selection issue, which can be really useful in order to understand these dynamics, the debate is still open.

Two final questions arise from a more practical point of view: which are, in detail, the variables that describe the personal characteristics and the context-related experiences in the most reliable way? How is it possible to collect them? In contexts like developing countries and historical studies, in fact, it is really rare to obtain anthropometric measures of child and mother in the prenatal and early life phases, that is in the critical periods. Measures like height, weight and placental size are golden information for this kind of analyses but, when missing, they need to be replaced by adequate - and easy to collect - proxy variables. The season of birth can be considered a good candidate for this role, as a proxy for prenatal and early-life environmental influences (Huntington 1938). Moore et al. (1999 and 2004), for example, showed that in rural Africa the season of birth strongly influences susceptibility to fatal infections in young adulthood. More specifically, it has been proved that the single month of birth is actually a factor in child survival, with different patterns according to age and geographic area of interest (Breschi and Livi Bacci 1997). But it is one of the challenges of historical demography to think of different, sharper and multidisciplinary solutions, both from the theoretical and the empirical point of view.

1.2 Main contributions of the thesis

In this thesis we use data drawn from 19th century parish archives in order to give an answer, with the historical and geographical peculiarities of our context of analysis, to some of the diverse questions arisen from the previous, general overview. Data come from the IMAV project (Infant Mortality in Asburgical Veneto) based at the University of Padua Department of Statistical Sciences. The aim of this project is to collect historical data that will aid in disentangling the main dynamics of the very high infant mortality that characterized the Veneto region during the Austrian Empire (1815-70). In that context, 40% of newborns died before their fifth birthday.

Across six chapters we provide a descriptive and speculative study of some relevant issues regarding infant mortality in Veneto during the 19th Century. These are, chapter by chapter, the main contributions of the thesis:

- In Chapter 2 the data collection is extensively described, with focus on data quality. Our dataset is made by unpublished nominative records from 46 parishes in the North-Eastern Italian region of Veneto and consists of more than 150,000 individual records, that is a remarkable dimension for the limited literature dedicated to the study of historical infant mortality in Veneto.

- In Chapter 3 we go deeper into the descriptive analyses of the data, presenting a number of classic demography tools such as general life tables and cohort-specific life tables: the first empirical evidences obtained give the directions for the following chapters. The final section of the chapter provides some confirms of previous studies about the influence of the month of birth, in the light of our original data resources.

- Chapter 4 is devoted to the debate about the direction of early-life negative experiences in the short term. Our results, obtained via Cox regression models, provide some initial proofs in favor of the mortality selection hypothesis. A new idea of 'context variable' is proposed, in order to summarize three early-life 'exposures' in an unique measure.

- In Chapter 5 we focus on the importance of community variables, since a population with a low mobility tends to bind its destiny to the characteristics of the place where it lives in. In effect, while the effects of individual factors have gained greater attention over time, not really much is known about the effects of community characteristics on health outcomes for children. Using a multilevel approach, we essentially confirm some findings from the 'mortality selection' chapter.

- Chapter 6 focuses on the Province of Padua, for which we own broader information. In particular, we study the influence of weather, that strongly influences the mortality levels of a population living in a specific geographical unit. With this aim, we include the daily temperatures in our life course frame, using discrete time models to measure the dangerous effect of - respectively - cold weather (in the very first part of life) and hot weather (after the first birthday). Several *robustness checks* in the last part of the chapter give strength to the results presented, that appear plausible and well grounded.

- In Chapter 7 we recap our most important findings and discuss about possible improvements to the study.

Chapter 2

Historical Context and Data Quality

2.1 Historical and Geographical Context

According to many authors, during the last decades of the Republic of Venice and at the time of the Napoleonic Empire, the area of our study was economically and socially underdeveloped. During the Habsburg period, an impulse was given to the region: the new administration accomplished new public works, rail and road transports, and reclamation of wetlands. However, this was not enough to improve significantly the social conditions and the life style of the local population: especially in rural areas, the severe taxation had very negative effects on a population that was completely devoted to agriculture. This one was seen as the only possibility to survive in an economically and socially static society.

In this context, the upper class proved to be tenacious in the defense of its outdated privileges, trying to prevent the social assault of the rising middle class. They were not interested in long term investments and completely indifferent to modernization. The ownership of lands - even if not expensive lands - was still a primary goal and this led to an extreme fractionation of both the properties and the land conduction.

Moreover, the land utilization was really backward: lack of irrigation, impracticability of the roads (and, as a result, difficulties in the commerce and in the market competitiveness), shortage of fertilization and deficiency of cattle and dairy farming, that were becoming an important propulsive factors in the bordering region of Lombardia, were only some of the severe problems that affected the agriculture of the region. In particular, the cultivation of corn had become almost a monoculture. Maize was everywhere, leading to a big shrinkage of woods and pastures and primarily to a dramatic pauperization of the diet: pellagra could easily spread and become one of the principal causes of death, especially after years of poor crops (Livi Bacci 1986)[1].

Peasants lived then in a psychological context of fear (they were afraid of nature), dissatisfaction (because of the unfavorable sharecropping agreement) and insecurity (the risk of getting into debts was very high for the small firms). Also, with the coming of factories and their revolutionary approach of domination on the nature, they were quickly losing the cult of tradition. They did not trust the public institutions and the myth of the patriarchal family was even fading. Unsurprisingly, the faith in God (or, better, the abandon to God) could not remain strictly intimate and individual, but turned to be a social necessity. The role of the parishes became central - not only geographically - in the daily life of rural towns.

The priests - who in general shared also the peasant origin with the churchgoers - accumulated a number of goods and functions well beyond the ritual sphere, gaining relevant administrative and economic roles. Apart from their employment as Civil Officers (that is actually fundamental for our research), they had an active role in the social life of the citizens: they spoke of education and politics (against the diffusion of socialism and secularization); they moved around the diocesan territory reporting the religious and social state of the people. From this point of view, the work of Gambasin (1973) is a milestone, describing in depth the relationships between priests and peasants in our context of analysis.

For example, we learn that 'In Vedelago, the peasants, miserable and sick, by far more numerous than artists and traders, emigrate. The landowners are scandalous and have a very bad conduct.' In some cases of true human disaster, the temptation to take refuge into superstition was real and the priests had to teach that 'God is the Lord of everything and Jesus has a redemptive power that frees from every demonic force. All the family is touched and delighted by what happens in the nature; they pray and invoke, they despair and damn.' The priests registered also 'a crack in the family ethics, due to a little care of the parents towards children'. This can be a key - albeit not the only one - to understand the exceptionally high levels of infant mortality in those situations (Gambasin 1973).

[1] Land organization was based on medium and large estates assigned through sharecropping, rent or a mix of the two. The landscape of the region was therefore characterized by large arable fields cultivated with cereals, flanked by rows of vines interspersed with elms, ash, walnut and cherry trees (the so-called aratorio arborato vitato, see Scarpa 1963). Between the 17th and the 19th century the cultivation of silkworms was introduced. It allowed the proliferation of small mills for silk processing mainly based on home work of women (Lazzarini 1981). Producing maize and wine to pay the high rents to the owners (the peasants did not usually own the land they cultivated) also meant limiting the surface devoted to meadow. The lack of sufficient resources to feed the animals continued to encourage the tendency to produce 'more polenta and less meat' (ivi, p. 88). About agrarian innovation and Malthusian dynamics in northern Italy during the Modern period, see Alfani 2013.

2.2 A Particular Type of Demographic Transition

Given the precarious life conditions described in the previous paragraph, it is not surprising to register that, from 1750 to 1840, the mortality of children in Veneto was extremely high, mainly due to mortality during the first month of life for children born during the winter (Fig. 2.1). However, during the eighty years that followed, mortality during the first year of life, and especially winter mortality during the first month, quickly declined, and in the first decade of the 20th century the region of Veneto was characterized by the lowest level of infant mortality of the Italian Kingdom (Dalla-Zuanna and Rosina 2011) - Fig. 2.2. This decline is quite surprising because - at least until the end of the 20th Century - in Veneto we did not observe special signs of improvement of food for the low class nor any significant modernization of the social structure.

These very particular trends of neonatal, infant and child mortality are the roots of the particular type of demographic transition of the Veneto region. For example, the enormous mass migration of the period 1880-1915 could not be understood without considering the great increase of the working-age population due to the decline in infant mortality in the second half of the 19th century. This increase, in fact, was not balanced by a parallel decline in fertility. The latter will drop below 5 children per woman only after the First World War.

Numerous scholars have tried to describe and interpret these particular demographic trends (Derosas 2009; Dalla-Zuanna and Rosina 2011). In particular, the Department of Statistics, University of Padua, has been working on the IMAV (Infant Mortality in Asburgical Veneto) project for several years. This project, with the help of researchers and students (including myself), aims to collect very detailed individual and territorial data for the Veneto region in the period 1816-70 which - as we shall see - is particularly fortunate on the side of available sources.

The goal is to systematize the IMAV data, with the aim to answer to some research questions that will be expressed in the different chapters, leading some demographic and statistical analysis of different levels of complexity, as stated in the introduction.

2.3 Individual Data

2.3.1 Collection

The aim of the IMAV project is to collect historical data that will help in disentangling the main dynamics of the very high infant mortality characterizing the region of Veneto

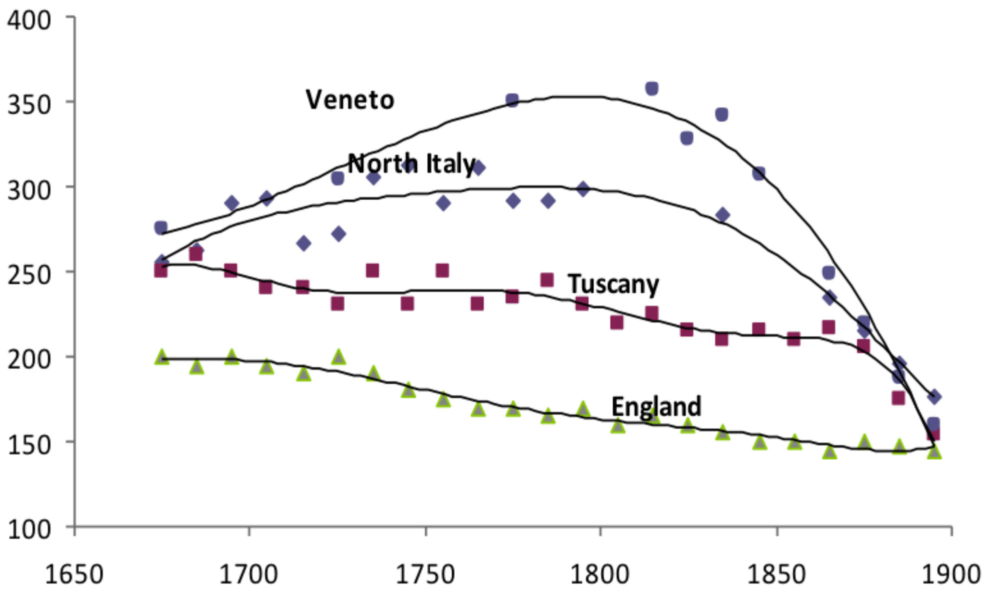


FIGURE 2.1: Secular trends in mortality rates during the first year of life in some areas of Italy and in England, 1675-1900. Source: Dalla-Zuanna and Rosina 2011.

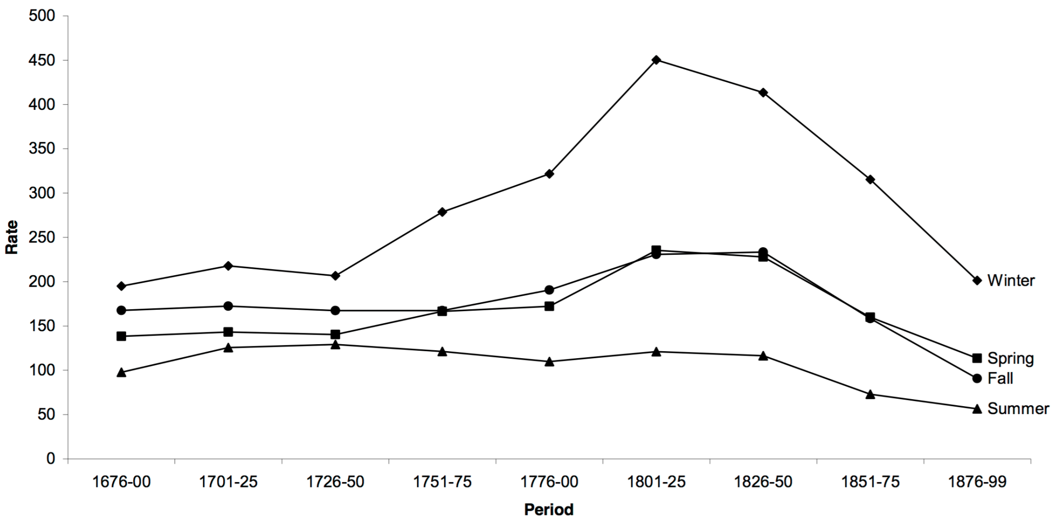


FIGURE 2.2: Probability of death (per thousand) during the first month of life by season of birth in the parish of Adria (province of Rovigo, South Veneto) in 1650-1900. Source: Dalla-Zuanna and Rosina 2011.

during the Austrian Empire (1816-66). Data have been collected from Civil Registers of parochial archives in the period 1816-70. During the Austrian Empire and in the first five years of the Kingdom of Italy, parish priests doubled as Civil Officers, paid by the State. They were carefully controlled by the provincial authorities, as testified by the frequent stamps on the pages of the pre-printed registers. All around the North-East of Italy (including Lombardy until 1860), these Civil Registers and/or their copies are preserved in the relevant Parish Archive and/or in the Diocesan Archive. They have not to be confused with the traditional Canonical Registers, that have been regularly written by the parish priests also during this period.

Church registers of baptisms from the mid-17th century on are widely available in the Catholic parishes of Veneto, while many burial registers can be found from the 18th century on, although in many parishes these registers date even further back. Unlike the civil registers, the church records were not pre-printed, and as such are much more difficult to read. The data recorded in the ecclesiastical acts of baptism are similar to those recorded in the civil acts of birth.

An important additional piece of information contained in the civil acts is the date of marriage, allowing us to identify children of the same parents without using the book of marriages, as would be necessary using only the church records. Moreover, the church records of baptisms do not - of course - report children who were not baptized (e.g., stillbirths), as was the case in the civil registers. The information contained in the civil registers of death are also similar to those observed in the ecclesiastical burial records, although in the former cause of death is recorded in a much more systematic - although not complete - way.

Back to our data, it is crucial to say that, during the first half of the 19th century, Veneto was divided into about 50 districts, that the Kingdom of Italy maintained as statistical territorial unit until the Census of 1911. In order to have a reliable statistical sample of each economic, social, and territorial context, the IMAV project collected data from at least one parish in each district of the Provinces of Venice, Padua, Treviso, and Vicenza, with a couple of districts of the province of Verona. Only for the district of Asiago (VI) it was not possible to collect data, as - for unknown reasons - in the parishes of this district the data collection stopped in 1848[2].

We focus on 46 parishes of the provinces of Padua, Venice, Treviso, Vicenza and Verona (Fig. 2.3). In Tab. 2.1, we indicate for each parish the dates of the first (and of

[2]None of the here considered parishes is in the mountain (the highest church is Valrovina, about 300 meters above the sea level). It is possible that in the communities of the Eastern Alps infant mortality, in the period considered here, was lower, because - paradoxically, given the low temperatures - poor families were better able to defend themselves from the cold. See - e.g. - http://www.forumeditrice.it/percorsi/forum-fvg/cultura-locale/sappada-plodn.

the last) certificate of birth and death that we collected. We can affirm that these 46 parishes represent an extensive picture of the rural environment in Veneto during the 19th century, excluding the Alpine area and the low plan between the Po and the Adige rivers.

In the period 1816-66, the registers of these 46 parishes include about 152 thousand births and 62 thousand deaths at ages 0-5 years[3].

Information from the registers included in the IMAV records is the following:

BIRTH REGISTERS

- date of birth
- given name of the child
- given and family name of the mother and father
- jobs of the mother and the father
- date of wedding
- date of the child's baptism.

DEATH REGISTERS

- date of death
- given name of the child
- given and family name of the mother and father
- jobs of the deceased (or father for young children)
- place of birth
- age at death
- cause of death.

[3]The data-entry process involved all the births in the period 1816-71, but obviously not all births in the last five years can be linked to the respective death for the entire age group 0-5. Therefore, in this chapter the performance indicators refer to the linkage of births until 1866 (or - for the parishes where the recording stops before, see table 2.1 - up to five years before the end of the survey). However, since the procedure linkage was performed for all the acts of birth and death, you can extend the analysis even at the last five years for the recognition of births, if we are content to study mortality from a very early age. For example, if you study the mortality in the first month, virtually the entire period 1816-70 can be analyzed.

Table 2.1: Reference dates for the birth and death certificates in the parishes.

	First Birth Act	**Last Birth Act**	**First Death Act**	**Last Death Act**
TV - Colle Umberto	1816	1867	1816	1869
TV - Bigolino	1816	1866	1816	1866
TV - S.Martino	1816	1870	1816	1873
TV - Selva	1816	1869	1816	1871
TV - Asolo	1832	1869	1832	1871
TV - Ormelle	1816	1866	1816	1869
TV - Piavon	1816	1866	1816	1870
TV - Monastier	1816	1870	1816	1870
TV - Vedelago	1816	1867	1816	1871
TV - S.Agnese (city)	1816	1866	1816	1870
TV - S.Lazzaro (city)	1816	1866	1816	1871
VE - S.Stino	1816	1867	1816	1870
VE - S.Michele	1817	1866	1818	1869
VE - Scorze'	1816	1870	1816	1871
VE - Carpenedo	1816	1870	1816	1871
VE - S.Pietro (city)	1816	1870	1816	1871
VE - S.Marco (city)	1816	1870	1816	1871
VE - Camponogara	1816	1866	1816	1872
VE - Chioggia	1816	1870	1816	1871
PD - Onara	1816	1870	1816	1870
PD - S.Giorgio	1816	1870	1816	1870
PD - Eremitani (city)	1817	1870	1817	1871
PD - S.Sofia (city)	1816	1870	1816	1871
PD - Chiesanuova	1816	1870	1816	1871
PD - Casalserugo	1818	1867	1818	1872
PD - Valnog./Faedo	1816	1870	1816	1874
PD - Pontelongo	1816	1866	1816	1867
PD - Pernumia	1816	1866	1816	1866
PD - Urbana	1817	1870	1817	1871
PD - Agna	1816	1871	1816	1871
VI - Valstagna	1816	1871	1816	1871
VI - Valrovina	1816	1871	1816	1871
VI - Thiene	1816	1870	1816	1871
VI - Nove	1816	1871	1816	1875
VI - S.Vito	1816	1871	1816	1871
VI - Lanze' e Setteca'	1816	1871	1816	1873
VI - Quargnenta	1816	1871	1816	1871
VI - Arzignano	1816	1871	1816	1871
VI - S.Pietro (city)	1816	1871	1816	1871
VI - Bosco e Nanto	1816	1871	1816	1871
VI - Sarego	1816	1871	1816	1871
VR - Vestenanova	1816	1871	1816	1871
VR - Caldiero	1816	1871	1816	1871

FIGURE 2.3: Location of the 43 parishes considered in the study.

2.3.2 Linkage

We have linked birth certificates to death certificates for children born in the same parish who (according to the priest notation) died at age 5 or younger, using as linkage keys name and family name of the child, name of the father, and name and family name of the mother. The linkage was facilitated by the availability of information on the children's age at death, as registered by priests; exact age was, however, ultimately calculated by matching the age at birth and death using Day Century Coding, or counting the number of days since January 1st, 1800 (Willekens 2013).

The linkage procedure has not been straightforward at all. Figure 2.4 helps us to disentangle the various steps. First of all, within the total number of children (cohort[4]) born

[4]The word cohort is commonly used in demography to identify people who experienced the same demographic event within the same time-span (e.g., children born in the year t, marriages celebrated in the year t). In this article, cohort is used in a broader sense, to identify children with identical characteristics at birth (e.g., newborns in the parish of Agna in the winter of years 1816-20; newborns in the parish of San Giorgio delle Pertiche in July of year 1832).

in a parish (group a) there are some children that clearly can not find in the death registers: they are children emigrated with their parents before their fifth birthday (group b) or children who simply survived after their fifth birthday (group c).

The children following the emigration of their parents can be a source of inaccuracy, since we do not know anything about them (migration was not reported in the registers): some of these emigrated children could have dead before their fifth birthday in a different parish, while we are 'forced' to consider them alive in the same parish where they were born. Anyway this possible underestimation of mortality is small in absolute number (the mobility of little children was relatively limited in our context of analysis) and is eventually balanced by a reasoned choice we made about death registers, which we are explaining soon.

In fact, the population of children dead in a parish before their fifth birthday (group d) is not homogeneous. It is quite easy to deal with the children dead in the same parish where they were born (e): their death act is linked to the relative birth act just through the linkage keys we described before. We call this procedure 'natural link' and this is of course the desirable situation.

But there is also a number of dead children for whom it was not possible to find the corresponding birth certificates (f). In this case we call it 'forced link', since we have generated the corresponding birth record: the procedure consisted in the identification of the records with incomplete information and in the consequent introduction of a number of 'dummy births'. More specifically, we indicated the date of birth starting from the age at death indicated by the priest on the death certificate. Using this system, we face the underestimation of mortality mentioned before.

We call no-linkage rate the proportion of 'not found' dead children over the total number of dead, once we exclude the children born in a different parish: in our notation, we have:

$$\textbf{NO LINKAGE RATE} = \frac{f}{d-g}$$

The 'no linkage rate' may depend on several factors: the accuracy of the parish priests in recording the events and the writing of the pastors themselves, errors in data collection (such as the interpretation of the spelling), but also the level of migratory exchanges of each parish. We will explore this phenomenon in the last part of this section.

This procedure of reconstructing the birth certificates is in line with that used by Schofield and Wrigley (1979) in reconstructing mortality in England during the 16th and 17th centuries through linking baptism and burial records. To avoid underestimating early mortality they also created 'dummy births', i.e., children found in the burial

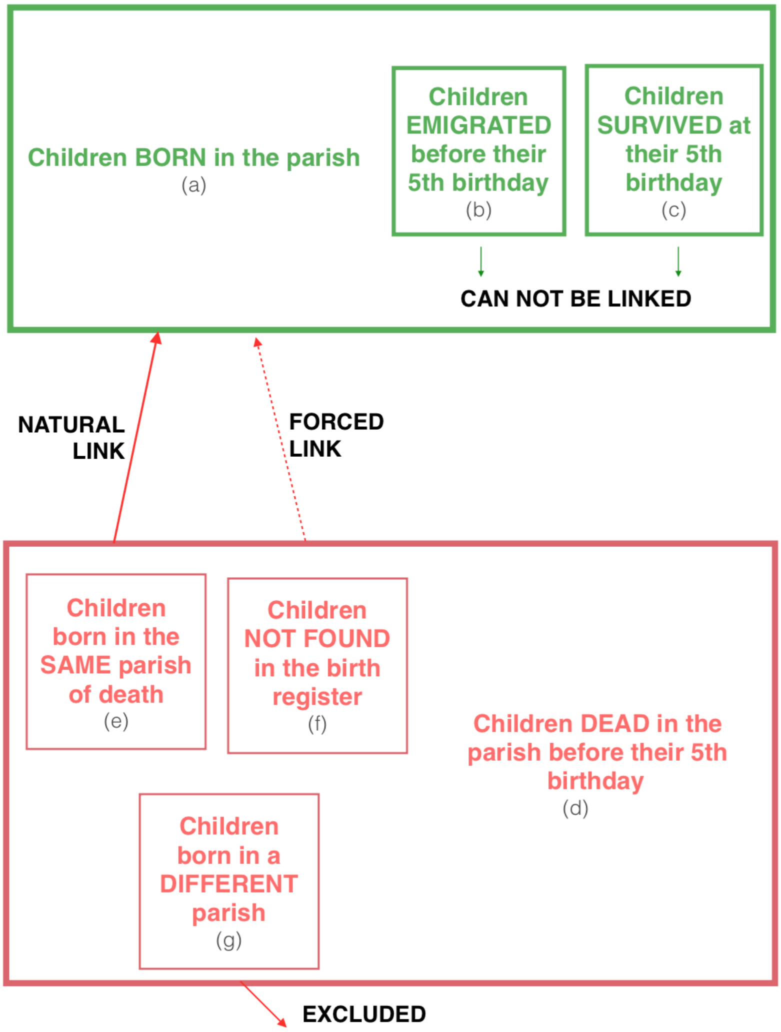


FIGURE 2.4: The procedure of data linkage between birth and death registers.

registers who died in the first days of life 'which cannot be linked to preceding baptisms, but clearly belong to a given family'.

It should be noted that in our approach the generated birth records refer only to children who died in the first few years of life, for whom we indicated both the place of birth and the parish where they died, and these two parishes coincide. Consequently, our procedure - while perhaps causing a slight increase in the number of births, due to a birth-record potentially already being present in the file and not recognized during the linkage procedure - allows us to avoid underestimating child mortality.

We should consider also another source of mortality inaccuracy, since it could also be underestimated due to systematic under-reporting of those children who died shortly after birth. This was observed in other Italian and European contexts in similar or slightly earlier periods (see e.g., Dalla-Zuanna et al. 2003; Dalla-Zuanna and Rossi 2010). However, in our parishes the deaths of these children were recorded with a good level of accuracy, as our data are consistent with those observed in high quality data situations.

Within our sample, 4.3% of the children are classified as having died within the first day of life; among this group, one half could be composed by stillbirths. According to table 2.2, the estimated proportion of stillbirths in our sample ranges from 40.76% in the period 1816-20 to 55.72% in the period 1841-45, with an average of 48.13%. While there is not a temporal pattern, important differences emerge at the territorial level, with the proportion of estimated stillbirths ranging from the mere 1% of S. Marco and Colle Umberto to the 100% of Casalserugo, Lanze'/Setteca', S. Pietro (VI) and Quargnenta. This does not mean, of course, that in some parishes there were almost zero stillbirths and in some other the totality of children dead during the first day of life were stillbirths: our estimates, in fact, can only be based on nameless children on both the birth record (when it exists) and the death record and on cases in which the date of baptism is not reported on the birth record, and these choices varied from priest to priest and from parish to parish.

As a result, both the stillbirth rate and mortality rate within the first day of life are around 2%-3%, compatible with that suggested by Woods (2009) for the pre-transition health period. In addition, a 5% probability of death in the first day of life in the period 1750-1810 was recorded for the sample of parishes used by the Cambridge Group in the reconstruction of the British population; as in our sample, stillbirths are included among the dead in the first day of life (Wrigley et al. 1997; Davenport 2010).

In his analytical reconstruction of the population history of three Bavarian villages, Knodel (1970) found 2% stillbirths during 1800-1849 while Oris et al. (2004) show that during the 18th and 19th centuries the proportion of stillbirths was higher in protestant than in catholic contexts, suggesting that in the latter, 'false live births' (i.e., stillbirths baptized as live births in order to allow the baptism to take place) were more diffused. As already stated, we cannot be sure of the 'true' values for stillbirths and real mortality rates in the first day of life; however, it is difficult to imagine dramatic underreporting in our data. More importantly, all the empirical results presented in the following chapters do not change if we exclude from the analyses all the 2,979 'suspected' stillbirths in the dataset, as we will highlight in one of the 'robustness checks' at the end of Chapter 6.

As showed in Table 2.3, 7.56% of all death acts were not linked to any birth act: for this proportion of dead certificates, 4,685 in total, we created the relative birth certificate. As reversal, more than 92% of the children born in the same parish who died at age 5 or younger were linked to their birth certificate, and the linkage performance is even greater for children who died before their third birthday: the wider the time span, the worse the linkage performance.

Table 2.2: Absolute number of stillbirths and proportion of stillbirths on the total of dead during the first day of life, according to ten five-years periods.

period	*stillbirths*	*firsrt day deaths*	*% of stillbirths*
1816-20	205	503	40.76
1821-25	246	549	44.81
1826-30	228	432	52.78
1831-35	327	642	50.93
1836-40	310	630	49.21
1841-45	336	603	55.72
1846-50	364	707	51.49
1851-55	326	700	46.57
1856-60	325	736	44.16
1861-65	312	687	45.41
Total	**2,979**	**6,189**	**48.13**

There are huge differences among the parishes, ranging from the 0.62% of Santa Sofia, a parish in center of Padua, to the 24.05% of San Marco, the most important parish in Venice's islands. In Table 2.4, we look closer at the nature of these 'not linked' children. Some of them (15.6% on average, with a maximum of 64% in the parish of Valrovina) are children who lived less than one day: they were included in the register of deaths, but not in that of births.

However, many unlinked death acts concern children dying at a 'late' age (1-4 years). Presumably a small part of them were not born in the parish of death, but their parish of birth was not indicated in the register. Other unlinked death acts are due to discrepancies between name/surnames, errors of the priests or in the data-entry.

2.4 Migrations and data quality

As described, our linkage procedure refers only to babies dead in the same parish where they were born. But migration may have triggered intense errors and potentially additional omissions. In fact, we can easily affirm that the variation of the no linkage rate among parishes is strictly related to the different levels of migrations among parishes, since we do not know the birth place of the 'not linked' dead children, and we can assume that some of them were not born in the same parish where they died.

For example, the parish of San Marco in central Venice shows a no linkage rate of 24%. This means that almost one quarter of the death certificates have not a linked birth certificate. We do not know much about these 479 children dead in San Marco, apart from the fact that 78% died when aged more than one month: it is likely that they were not born in San Marco (and this is the reason why we do not find them in the birth

Table 2.3: No linkage rate in the 46 parishes.

	Total Births	**Created Births**	**Deaths Before the 5th Birthday**	**No Linkage Rate %**
	(a)	**(f)**	**(d-g)**	**f/(d-g)**
TV - Colle Umberto	2,424	21	770	2.73
TV - Bigolino	2,016	30	866	3.46
TV - S.Martino	1,455	19	404	4.70
TV - Selva	2,891	57	1,182	4.82
TV - Asolo	1,811	83	728	11.40
TV - Ormelle	1,550	70	508	13.78
TV - Piavon	1,985	53	538	9.85
TV - Monastier	4,510	149	1,612	9.24
TV - Vedelago	1,662	39	844	4.62
TV - S.Agnese (city)	3,212	186	1,321	14.08
TV - S.Lazzaro (city)	1,292	29	460	6.30
VE - S.Stino	4,090	172	1,469	11.71
VE - S.Michele	1,985	39	861	4.53
VE - Scorze'	2,511	32	906	3.53
VE - Carpenedo	4,581	115	1,754	6.56
VE - S.Pietro (city)	16,444	399	7,022	5.68
VE - S.Marco (city)	5,571	479	1,992	24.05
VE - Camponogara	2,535	61	995	6.13
VE - Chioggia	4,091	202	1,400	14.43
PD - Onara	2,278	21	910	2.31
PD - S.Giorgio	2,707	145	1,130	12.83
PD - Eremitani (city)	3,965	280	2,062	13.58
PD - S.Sofia (city)	4,690	11	1,769	0.62
PD - Chiesanuova	3,201	63	1,289	4.89
PD - Casalserugo	2,354	15	1,096	1.37
PD - Valnog./Faedo	1,532	24	737	3.26
PD - Pontelongo	3,705	131	1,780	7.36
PD - Pernumia	4,609	136	2,070	6.57
PD - Urbana	2,202	17	973	1.75
PD - Agna	4,171	61	2,000	3.05
VI - Valstagna	5,508	96	1,904	5.04
VI - Valrovina	1,868	36	770	4.68
VI - Thiene	12,301	130	5,626	2.31
VI - Nove	3,758	159	1,771	8.98
VI - S.Vito	2,303	74	1,019	7.26
VI - Lanze' e Setteca'	1,531	108	726	14.88
VI - Quargnenta	1,521	12	557	2.15
VI - Arzignano	8,566	443	3,082	14.37
VI - S.Pietro (city)	2,602	185	1,164	15.89
VI - Bosco e Nanto	2,388	119	946	12.58
VI - Sarego	2,697	104	1,153	9.02
VR - Vestenanova	1,957	41	723	5.67
VR - Caldiero	3,180	39	1,058	3.69
TOTAL	**152,210**	**4,685**	**61,947**	**7.56**

Table 2.4: Distribution of the no linked records with respect to the age at death

	generated births	1st day	days 1-27	days 27+	1st day (%)	days 1-27 (%)	days 27+ (%)
TV - Colle Umberto	21	1	6	14	4.8	28.6	66.7
TV - Bigolino	30	3	7	20	10.0	23.3	66.7
TV - S.Martino	19	10	6	3	52.6	31.6	15.8
TV - Selva	57	5	19	33	8.8	33.3	57.9
TV - Asolo	83	30	24	29	36.1	28.9	34.9
TV - Ormelle	70	3	24	43	4.3	34.3	61.4
TV - Piavon	53	2	3	48	3.8	5.7	90.6
TV - Monastier	149	61	20	68	40.9	13.4	45.6
TV - Vedelago	39	2	11	26	5.1	28.2	66.7
TV - S.Agnese (city)	186	17	34	135	9.1	18.3	72.6
TV - S.Lazzaro (city)	29	10	6	13	34.5	20.7	44.8
VE - S.Stino	172	15	48	109	8.7	27.9	63.4
VE - S.Michele	39	0	3	36	0.0	7.7	92.3
VE - Scorze'	32	0	6	26	0.0	18.8	81.3
VE - Carpenedo	115	10	11	94	8.7	9.6	81.7
VE - S.Pietro (city)	399	4	3	392	1.0	0.8	98.2
VE - S.Marco (city)	479	77	28	374	16.1	5.8	78.1
VE - Camponogara	61	2	24	35	3.3	39.3	57.4
VE - Chioggia	202	10	13	179	5.0	6.4	88.6
PD - Onara	21	5	6	10	23.8	28.6	47.6
PD - S.Giorgio	145	20	14	111	13.8	9.7	76.6
PD - Eremitani (city)	280	19	38	223	6.8	13.6	79.6
PD - S.Sofia (city)	11	0	5	6	0.0	45.5	54.5
PD - Chiesanuova	63	0	1	62	0.0	1.6	98.4
PD - Casalserugo	15	4	11	0	26.7	73.3	0.0
PD - Valnog./Faedo	24	4	7	13	16.7	29.2	54.2
PD - Pontelongo	131	0	12	119	0.0	9.2	90.8
PD - Pernumia	136	12	38	86	8.8	27.9	63.2
PD - Urbana	17	0	9	8	0.0	52.9	47.1
PD - Agna	61	28	10	23	45.9	16.4	37.7
VI - Valstagna	96	38	13	45	39.6	13.5	46.9
VI - Valrovina	36	23	4	9	63.9	11.1	25.0
VI - Thiene	130	0	3	127	0.0	2.3	97.7
VI - Nove	159	68	28	63	42.8	17.6	39.6
VI - S.Vito	74	15	13	46	20.3	17.6	62.2
VI - Lanze' e Setteca'	108	20	20	68	18.5	18.5	63.0
VI - Quargnenta	12	0	2	10	0.0	16.7	83.3
VI - Arzignano	443	155	81	207	35.0	18.3	46.7
VI - S.Pietro (city)	185	39	25	121	21.1	13.5	65.4
VI - Bosco e Nanto	119	45	20	54	37.8	16.8	45.4
VI - Sarego	104	40	5	59	38.5	4.8	56.7
VR - Vestenanova	41	1	1	39	2.4	2.4	95.1
VR - Caldiero	39	0	1	38	0.0	2.6	97.4
TOTAL	**4,685**	**798**	**663**	**3,224**	**17.0**	**14.2**	**68.8**

registers) but they had moved earlier to Venice with their family. The intensity of the phenomenon in San Marco is likely due to the exceptionality of exchanges of people and goods that historically characterizes Venice, not adequately clued by the parish-priest in the death act.

But migrations took place everywhere, with much lower intensity. According to Berengo (1969), during the first half of the 18th century the population experimented little perturbations related to migrations: one third of them was limited to the same province, while they were practically irrelevant with respect to foreign countries. Much more stable were the seasonal migrations, that concerned the mountains and took place in the period April-October.

To better understand the extent of this possible underestimation we analyzed the deaths at age 0-5 years, born elsewhere, in the parish of San Vito di Leguzzano, a hilly town close to Vicenza. This is one of the parish that we analyzed early back in our collection, and we know it in depth. In our graphical representation (Fig 2.4), these records are indicated with letter *g* and they are excluded from the final version of the database.

In San Vito di Leguzzano, in our period of analysis, 2% of the total number of dead babies were born for sure elsewhere (we remember again that we do not know the birthplace for the children in group *f*, since in these cases the birthplace was not registered: some of them could have not been born in San Vito di Leguzzano). More than half of them were born in a parish located at less than 10 kilometers from San Vito di Leguzzano (cfr. Tab. 2.5) and we should note that, among the children coming from 'far', we have the three babies coming from the orphanage of Vicenza. On average a child dead in San Vito coming from a different place, was born at less than 15 kilometers from the parish.

In conclusion, if for families with very young children emigrations were similar in number to immigrations, it is possible that our measures of mortality are underestimated, and that these underestimations are greater where exchanges were more intense (as in San Marco) or lower where exchanges were less intense. This issue should not, however, greatly affect our research goals for two reasons: first, the introduction of the 'generated records' should restore almost completely the numeric balance; secondly, the underestimation should be similar, in relative terms, when and where early infant mortality was higher or lower.

TABLE 2.5: Town of origin of 19 children dead in San Vito di Leguzzano.

Origin	**n**	**distance (km)**
Magre', Monte Magre'	5	3
Malo, Monte di Malo	5	3
Marano	1	4
Isola Vicentina	1	8
Asiago	1	20
Vicenza	3	20
Castelfranco V.	1	43
Feltre	2	55
	19	**14.68**

Chapter 3

Differential Infant Mortality

3.1 Construction of the general life table

3.1.1 Giving a weight to each record

Before starting the procedure of construction of the life tables, which will help us in describing the patterns of survival in our context of analysis, we have to note that our database has not been selected according to a simple random sampling design. In our database, five provinces and thirty-four districts are represented by forty-six parishes. For each district we have at least one parish, and some parishes belong to the same district. Since the districts have different dimensions and the number of records registered in each parish does not reply those dimensions, some districts are over (under) represented, i.e. we have too many (too few) observations representing a small (big) district. This means we need to assign a specific weight to each record, in order to reproduce a sample that is representative of the whole population living in the thirty-four districts in 1815-70.

We also built sub-districts, in order to reduce the dimension of the biggest districts (coinciding with the cities in our sample, excluding Venice) and to take into account the specificity of the urban areas. For example, for the vast district of Padua (one third of the population of the province in both Italian censuses of 1857 and 1871) we have considered four parishes: two representing the city (Santa Sofia and Eremitani), one representing the hinterland (Chiesanuova) and one representing the countryside (Casalserugo). Something similar happened with the districts of Treviso (S. Lazzaro and S. Agnese representing the city, Monastier representing the country) and Vicenza

(S. Pietro representing the city, Lanze' and Setteca' representing the hinterland). Moreover, two other districts are represented by more than one parish: for example Oderzo (Treviso) is represented by the parishes of Ormelle and Piavon.

We calculated the weights in this way:

$$W_i = \frac{N_i/N}{B_i/B},$$

where i is the district or the sub-district, N_i is the population at the Italian census of 1871 (the first of good quality available) in the district or sub-district i, B_i is the number of children born in the parish(es) belonging to the district (or sub-district) i during 1816-70, N is the population at the census of 1871 in the 34 districts as a whole, B is the number of children born in the 46 parishes during 1816-70[1].These weights (reported in Table 3.1) have been assigned to each of the 152,210 birth records and, starting from this point, they will be considered in all the subsequent calculations we provide.

3.1.2 The general life table

The weighting procedure allows the life tables construction. The first is the most general one: in Table 3.2 we have reported $l(x)$ (the number of people surviving at age x)[2], $d(x)$ (the number of people who die aged x) and $q(x)$ (the probability of dying between age x and age $x + 1$).

As expected, this life table works very well in summarizing some relevant points:

- the first day of life was extremely lethal, with more than 41 children dying over 1,000 children born. Also, 10% of the total deaths registered in the first five years of life refer to the first day of life (see also Fig. 3.1);
- the first week (excluding the first day) and the first month (excluding the first week) continued to be difficult for the newborns, with a probability of dying of 96‰ and 97‰, respectively;

[1] For calculating the weights, the number of children born during the last years in Colle Umberto, Bigolino, Ormelle, Piavon, Vedelago, S. Agnese, S.Lazzaro, S. Stino, S. Michile, Camponogara, Casalserugo, Pernumia, Valnogaredo, Pontelongo and Urbana and the number of children born during in 1816-31 in Asolo have been estimated, since data for these parishes are not available for the complete period 1816-1870.

[2] Our sample is now statistically representative of the pool of the 34 districts. We have chosen $100,000$ (a very typical value) for $l(0)$, the radix of the table

Table 3.1: Data for calculating the sampling weights.

Parish	District	Bi	Ni	Wi
Agna	Conselve	4,171	25,754	0.697
Casalserugo	PD (country)	2,539	61,294	2.725
Chiesanuova	PD (hinterland)	3,201	21,500	0.758
Onara	Cittadella	2,278	32,528	1.611
Pernumia	Monselice	4,970	31,947	0.725
Pontelongo	Piove di Sacco	3,996	32,869	0.928
S.Giorgio	Camposamp.	2,707	37,331	1.556
S.Sofia, Eremitani	PD (city)	8,655	44,607	0.582
Urbana	Montagnana	2,202	32,508	1.666
Valnog./Faedo	Este	1,532	44,092	3.248
Asolo	Asolo	2,621	32,156	1.384
Bigolino	Valdobbiadene	2,174	23,374	1.213
Colle Umberto	Vittorio Veneto	2,564	43,235	1.903
Monastier	Treviso (country)	4,510	44,397	1.111
Ormelle, Piavon	Oderzo	3,813	46,922	1.389
S.Lazzaro, S.Agnese	TV (city)	4,857	28,291	0.657
S.Martino	Conegliano	1,455	45,280	3.512
Selva	Montebelluna	2,945	33,205	1.273
Vedelago	C.franco Veneto	1,758	31,239	2.006
Camponogara	Dolo	2,734	30,575	1.262
Carpenedo	Mestre	4,581	22,735	0.560
Chioggia	Chioggia	4,091	51,129	1.410
S.Marco, S.Pietro	VE (city)	22,015	128,901	0.661
S.Michele	S. Dona’	2,184	29,936	1.547
S.Stino	Portogruaro	4,326	35,440	0.925
Scorze’	Mirano	2,511	26,029	1.170
Arzignano	Arzignano	8,566	23,837	0.314
Bosco e Nanto	Barbarano	2,388	16,046	0.758
Lanze’ e Setteca’	VI (hinterland)	1,531	22,334	1.646
Nove	Marostica	3,758	32,532	0.977
Quargnenta	Valdagno	1,521	26,436	1.961
S.Pietro VI	VI (city)	2,602	37,686	1.635
S.Vito	Schio	2,303	45,060	2.208
Sarego	Lonigo	2,697	33,658	1.408
Thiene	Thiene	12,301	25,109	0.230
Valrovina, Valstagna	Bassano	7,376	47,342	0.724
Caldiero	S. Bonifacio	3,180	32,176	1.142
Vestenanova	Tregnago	1,957	18,998	1.096
		155,569	**1,378,488**	1.000

TABLE 3.2: Life table (weighted) for the total population. 46 parishes of Veneto, 1816-70

	l(x)	**d(x)**	**q(x)**
First Day	100,000	4,149	0.041
Days 2-6	95,851	9,211	0.096
Days 7-27	86,640	8,377	0.097
Months 1-2	78,263	3,232	0.041
Months 3-11	75,031	4,957	0.066
Year 1	70,074	5,543	0.079
Year 2	64,530	2,650	0.041
Year 3	61,881	1,447	0.023
Year 4	60,433	874	0.014
Days 0-6	100,000	13,360	0.134
Days 0-27	100,000	21,738	0.217
Year 0	100,000	29,928	0.299
Year 0-4	100,000	40,444	0.404

- between the fourth and the fifth birthday, $q(x)$ was 'only' 14‰ (the lowest in the table);

- the number of deaths registered in the four years following the first birthday are merely one fourth of the total death registered in the five-years period (see Fig. 3.1);

- in the pooled sample of 46 parishes of Veneto (1815-70) the initial population is reduced by more than 40% before the fifth birthday. It is a quite impressive proportion, especially if compared to other available life tables from the same period[3]. Remaining in Europe, among the countries with available data this proportion was equal to 38% in Austria; in Germany it was 33%; in the Netherlands 32%; in France 27%; in England and Wales 26%; in Sweden 25%. Outside Europe, in Canada 28% of newborns died before their fifth birthday, while in Japan only 23%. The level of infant mortality in this period was higher only in Russia (45%). For a comparison, in 2013 the highest children mortality was registered in Angola, where 17% of children died before their fifth birthday.

[3]These data come from the Human Life-Table Database, that is a collection of population life tables covering a multitude of countries and many years. Three institutions are jointly developing the HLD: the Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany, the Department of Demography at the University of California at Berkeley, USA and the Institut national d'etudes demographiques (INED) in Paris, France. Most of the HLD life tables are life tables for national populations, which have been officially published by national statistical offices. Some of the HLD life tables refer to certain regional or ethnic sub-populations within countries. Parts of the HLD life tables are non-official life tables produced by researchers.

The differences are even bigger if we look only at the first year of life: in our pooled sample of 46 parishes of Veneto (1816-70) the initial population is reduced by 30% before the first birthday. This is a level comparable to the Russian one (30.3%), while in the other European countries the proportion of newborns dying before their first birthday ranges from 14-15% (England and Wales, Netherlands and France) to 23-24% (Germany and Austria). Looking at the first month of life, in Veneto the mortality rate was 21.7%, while in the rest of Europe it peaked at 7.2% in Russia and was only 4.9% in the Netherlands [4] .

Before splitting the general table according to a series of variables that we are introducing soon, we have computed a life table restricted to the first month of life, i.e. we calculated the probabilities of dying in each of the first 30 days of life, for the pooled sample of 46 parishes. We have not reported the whole table, but Fig. 3.2 does the job well. After the peak in the first day, mortality drops in the second day to 15‰, and rises again up to 18-19‰ in days three and four. After a new, small peak at the end of the first week of life, the risk of dying starts its slow decrease until 1-2‰ (fourth week of life).

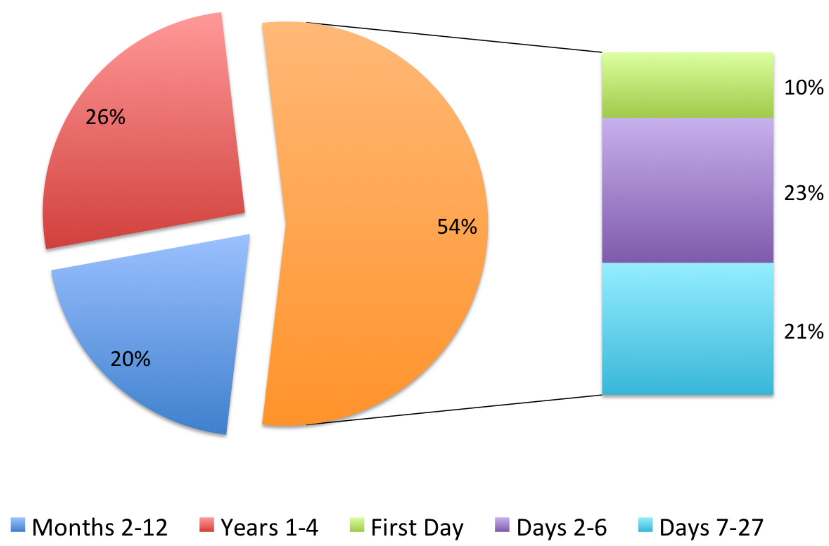


Figure 3.1: Pie graph of the number of deaths registered in the first five years of live in the pooled sample of 46 parishes. Veneto, 1816-70.

[4]The life tables are referred to the following years: Austria 1865-75; Canada 1841; France 1841; Germany 1871-81; Japan 1891; The Netherlands 1861-65; Russia 1896; Sweden 1816-40; England and Wales 1841.

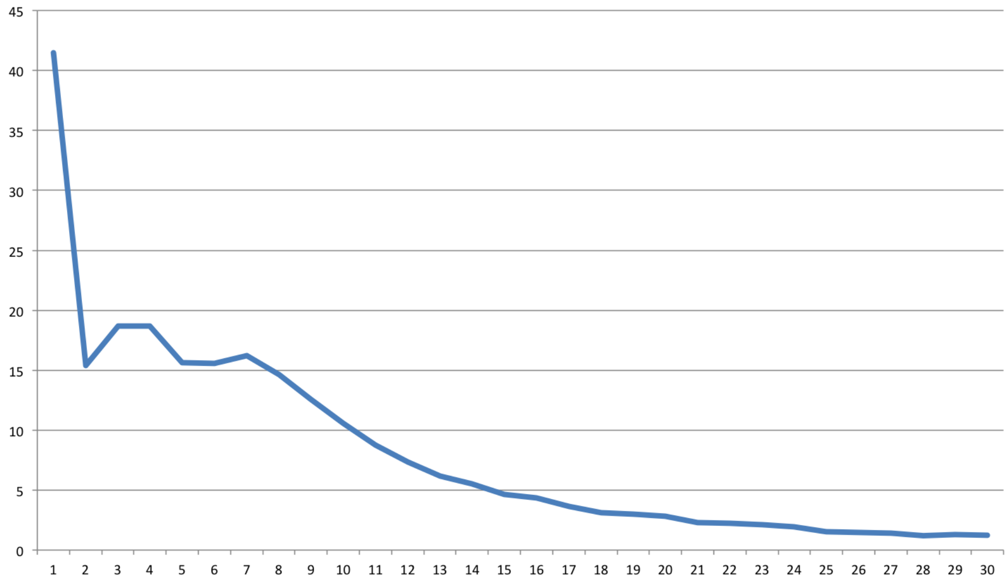


FIGURE 3.2: Probability of dying (per thousand) in each of the first 30 days of life, for the pooled sample of 46 parishes. Veneto, 1816-70

3.2 More specific life tables

In this paragraph we are using - for the first time in our dissertation - three individual characteristic of the newborns: sex, season of birth and father's job. At this point, anyway, our aim is not to provide a detailed differential analysis. We are presenting the mortality differences that are present in our database, according to some individual characteristics. In the future, an in-depth examination of these differences may lead to a broader interpretation.

Further in the chapter, moreover, we will analyze two other individual variables (parish of birth and month of birth) and we will give space to a description of the historical trend of mortality in our context of analysis. In Chapter 5, finally, we will introduce some community variables that are will require the switch to a multilevel approach.

3.2.1 Sex

The information about the sex of the newborn is always available and should be used in every model, as a control variable, when trying to explain the early mortality dynamics. We have 79,140 males (52% of the total) and 73,070 females (48% of the total) in our sample. The global sex ratio (M/F) at birth in our sample is 107, indicating that our data do not seem to be biased with respect to the proportion of male and female babies. However, there could be some locally-based gender biases in children registration, as significant territorial differences emerge in our sample, with the sex ratio at birth ranging from 99 to 118.

Males were exposed to a higher risk of dying in three out of the four time spans we have considered (see Tab. 3.3). In particular, their mortality is 16‰ higher than the one of females during the first week; the gap appears to be reduced in the following three weeks of life (7‰ higher) and almost disappears if we look at the months 1-11 (3‰ higher). In the four years between the first and the fifth birthday, the situation is reversed and the risk of dying is 4‰ higher for females. This last result, in particular, could be related to a sort of selection issue.

According to literature, mortality is higher for males at every age in Europe, where the baby girls were not discriminated from the child-care point of view, as it happened (and happens) in other areas of the planet. While lifestyle and behavioral factors contribute to sex differences in adult mortality, differences in infant mortality are more related to a combination of biological and environmental factors that disadvantage males (Waldron 1985, Pongou 2012). Actually, our results show a quite small male disadvantage if compared to six other European countries (Sweden, France, Denmark, England/Wales, Norway and Netherlands). In the pooled sample of our 46 parishes of Veneto, the male/female ratio of the probability of death before the first birthday is 1.072. In the six countries considered by Drevenstedt et al. (2008), the ratio ranged from 1.151 (in the Netherlands) to 1.208 (in England and Wales), in the period 1845-54. In Russia, using the data from the Human Life-Table Database, we calculated the ratio being 1.133, that is higher than in Veneto but lower than in the six European countries. Since in the six European countries the general level of infant mortality was much lower than the one observed in our sample and in Russia (cfr. Par. 2.1.2), we conclude that the male/female ratio actually increases when the infant mortality decreases, confirming a growing disadvantage for males related to the drop of infant mortality, as Drevenstedt et al. (2008) already proved with their work.

The male/female ratio reduces with age. Looking at the period between the first and the fifth birthday, for example, in England (where the first year ratio was the highest among the six European countries considered as reference) the ratio drops to 1.029 (that is the lowest among the six European countries considered as reference). This narrowing of the mortality gap between males and females takes place everywhere, from the Netherlands (1.059) to Russia (1.030). In the pooled sample of our 46 parishes of Veneto, the male/female ratio of the probability of death between the first and the fifth birthday is even inverted with respect to the first year of life: it is lower than 1 (0.974), indicating that now the risk of dying is slightly higher for females. As already stated, this phenomenon may partially be justified via mortality selection, but surely it deserves further explanations.

3.2.2 Season of birth

Previous studies (Kynast-Wolf et al. 2006; Lawlor 2004) showed that, in a context of analysis like ours, the season of birth may play a crucial role in the determination of the mortality level. In particular, to be born during the cold season was a very undesirable situation, with a high risk of dying within the first day (or first week, or first month) of life. In the longer term, hot months become lethal, because of the stronger incidence of gastro-enteric diseases. In Tab. 3.3 we see a very evident pattern: in the early phase of their lives, children born during wintertime were exposed to a much higher risk of dying with respect to the ones born in warmer seasons. Looking at the first week of life, the difference is huge: being born in winter caused a risk of dying 162‰ higher with respect to summer. The gap is reduced, but still big, if we look at the subsequent period (days 7-27): 141‰. Moreover, the season rank does not change among these two periods: after winter, the most lethal season is autumn, while spring and summer lay in the last two positions of the rank. Again, the situation appears to be reversed if we look at the longer term: in particular, children born in summertime had the highest risk of dying after their first birthday (years 1-4).

3.2.3 Social class

For the social class of each subject in the dataset, we are using a proxy variable. The priests used to register the jobs of both father and mother of the newborns in the birth acts. Our decision is to use the father's professional status, that is traditionally representative of the familiar wealth (especially in our context of analysis), to identify the social class of the child.

The first step is the classification of the different activities. In fact, our priests had not any sort of statistical know-how and sometimes they were really specific (e.g. 'brewery director', 'cheesemonger') or completely vague (e.g. 'rich') in the description of the job. In order to make our next analyses a little bit easier, in this chapter we group all the activities in four classes: two lower classes (farmers and craftsmen); an intermediate class (traders); an upper class (landowners, professionals, employees).

Unsurprisingly the four classes are not equally represented in our sample. In particular, 51% of the subjects in our sample were children of peasants and another 25% were sons of craftsmen. That is, three quarters of the sample is composed by children belonging to lower classes. Finally, 8% were sons of traders and another 14% could be considered part of a wealthy family.

There are significant differences according to the parishes. For example, in San Pietro (Venice) the proportion of peasants was only 0.5%, while in Valrovina it was more than 98%. In general (Fig. 3.3), the presence of peasants was unsurprisingly smaller in the parishes located in the cities (see S.Sofia and Eremitani for Padua, S.Agnese for Treviso, S.Pietro for Vicenza, S.Marco and S.Pietro for Venice), where more traders and landowners were living.

Tab. 3.3 represents a scenario in which children of peasants are exposed to the highest risk of dying in the short period, that is up to the first month of life. If we look at the long term (first year and years 1-4), the risk of dying turns to be higher for craftsmen and traders. Also, it is in the long term that the children living in wealthier families seem to take advantage of their situation: in the first year and in years 1-4, their risk of dying is from 17‰ to 59‰ lower if compared to the other three groups.

TABLE 3.3: Probability of dying (‰) according to four time spans and three individual characteristics (sex, season of birth, social class).

	Days 0-6	**Days 7-27**	**Months 1-11**	**Years 1-4**
Males	141	100	106	148
Females	125	93	103	152
Winter	231	176	97	144
Spring	100	71	89	157
Summer	69	35	110	158
Autumn	138	127	128	135
Peasants	137	103	102	133
Craftsmen	109	81	110	179
Traders	118	80	113	178
Landowners/Prof./Emp.	125	90	85	120

3.2.4 Parish of birth

As introduced by Fig. 3.3, we take into account the parish of birth of each children, since a number of characteristics are not homogenous among the different communities. We are now describing some macro-results emerging from a glance to the life tables of each parish. By the way, this is not a deep analysis of the peculiar effects of territorial and social features on infant mortality, since a multilevel approach to our problem will be presented in Chapter 5. Here we are only noting that (see Fig. 3.4) the differences in the infant mortality level can be huge.

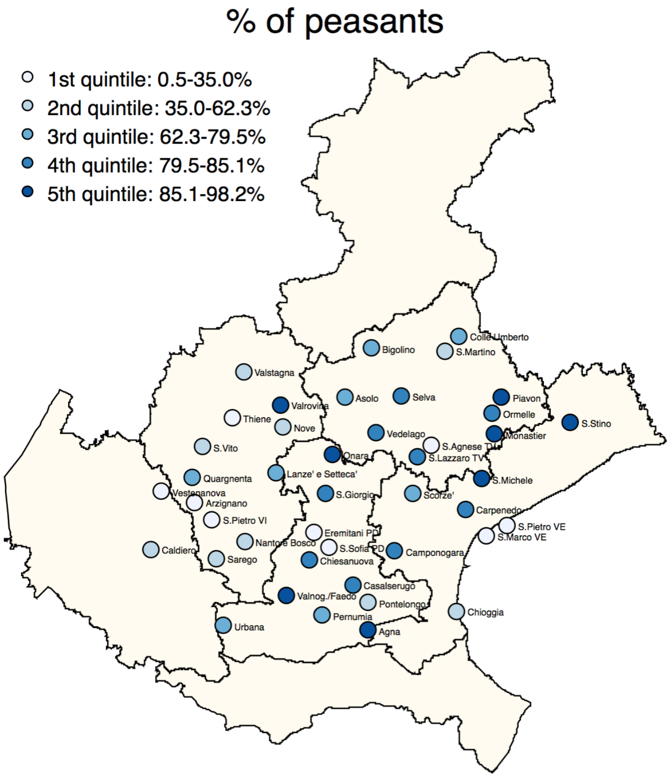


FIGURE 3.3: Proportion of children born in a peasants family in each of the 46 parishes in the database. Veneto, 1816-70

In particular, the province of Treviso (and especially its Eastern area) was the 'best' part of the region where to be born during the 19th century. In this area we have 5 parishes (Colle Umberto, S. Martino, Piavon, Ormelle and Monastier) belonging to the first quintile (mortality at years 0-4 between 273 and 346‰), with Piavon being the parish with the lowest rates in the whole sample. On the contrary, the province of Padua was the worst area with respect to this index. In the countryside, South of Padua, we have 4 parishes (Casalserugo, Pontelongo, Agna and Pernumia) belonging to the fifth quintile. In particular, in Pontelongo the level was 511‰.

The situation becomes more complex if we look at shorter spans of time. Specifically, we present here the mortality rates in four time spans: first week, first month (excluding the first week), first year (excluding the first month), years 1-4. It is evident that there is an inverse relation between the first two rates (Fig. 3.5 and Fig. 3.6) and the second two rates (Fig. 3.7 and Fig. 3.8): the parishes where the risk of dying was higher in the first week and in the first month are the same where the risk of dying in the two subsequent periods was lower, and viceversa. One good example are the two parishes located in Venice (San Marco and San Pietro), where the mortality was quite low in the first week (79‰ and 81‰, respectively) and in the first month (53‰ and 58‰, respectively), but in the subsequent periods it reached two of the highest registered levels in the region,

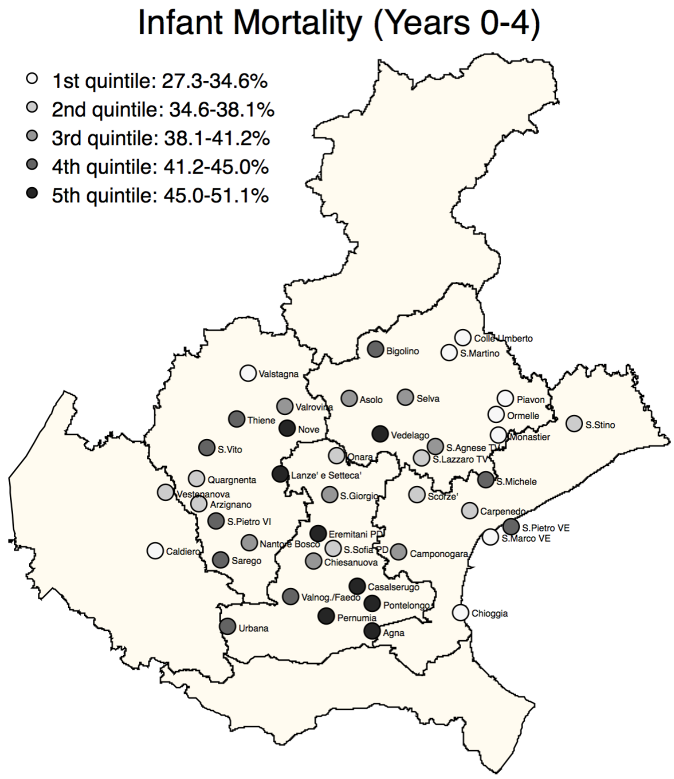


FIGURE 3.4: Child mortality rate (years 0-4) in each of the 46 parishes of the database. Veneto, 1816-70

TABLE 3.4: Correlation between mortality rates in four different time spans, calculated within the 46 parishes in the sample.

	days 0-6	**days 7-27**	**months 1-11**	**years 1-4**
days 0-6	1	0.783	-0.112	-0.214
days 7-27		1	0.103	-0.082
months 1-11			1	0.686
years 1-4				1

both in months 1-11 (106‰ and 140‰, respectively) and in years 1-4 (160‰ and 214‰, respectively). In Tab. 3.4 we show how the correlation between mortality in the first week and mortality in the subsequent periods turns negative starting from the second month, and becomes even stronger in the last period.

This evidence can be interpreted as a first clue of the existence of a sort of homeostatic mechanism related to the mortality in subsequent periods, which could be the result of a selection process (by death) of the young population. Anyway, we will say more about this topic further in this chapter and extensively in Chapter 4.

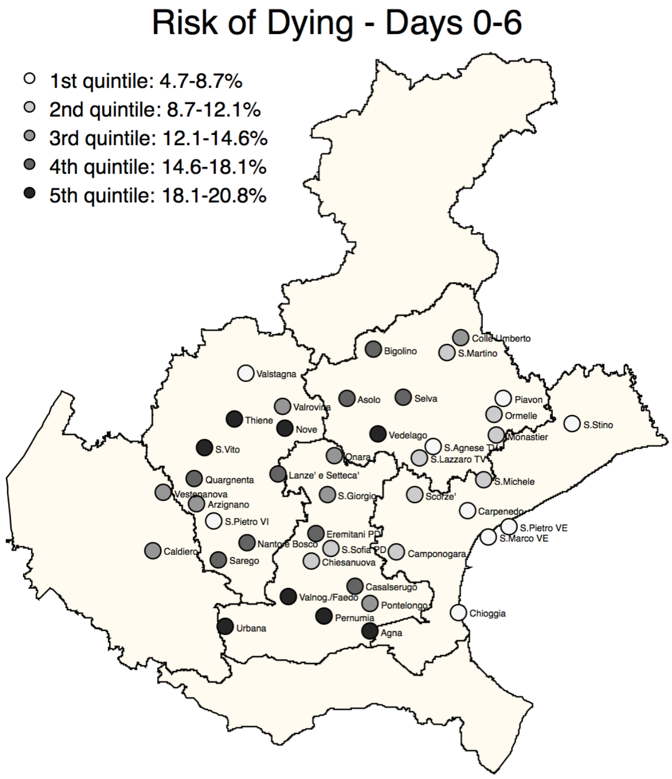


FIGURE 3.5: Child mortality rate (days 0-6) in each of the 46 parishes in the database. Veneto, 1816-70

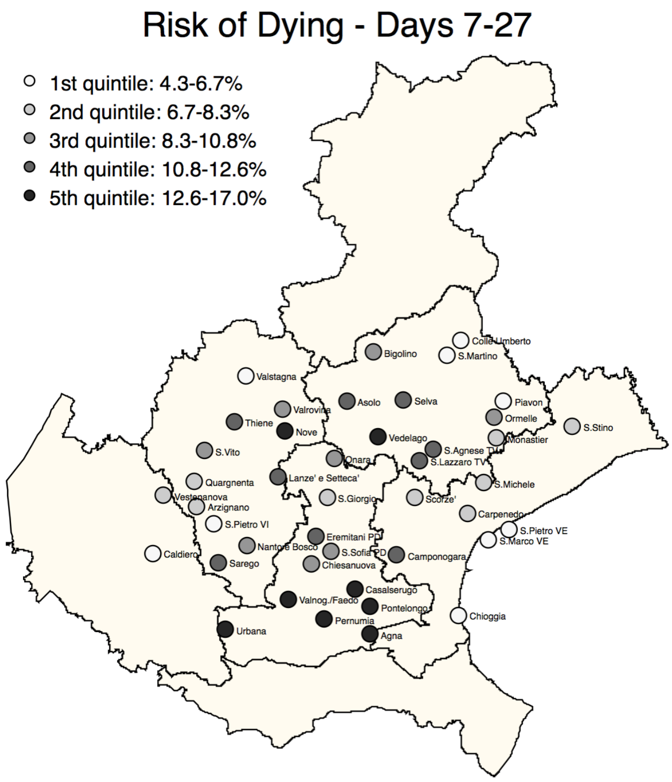


FIGURE 3.6: Child mortality rate (days 7-27) in each of the 46 parishes in the database. Veneto, 1816-70

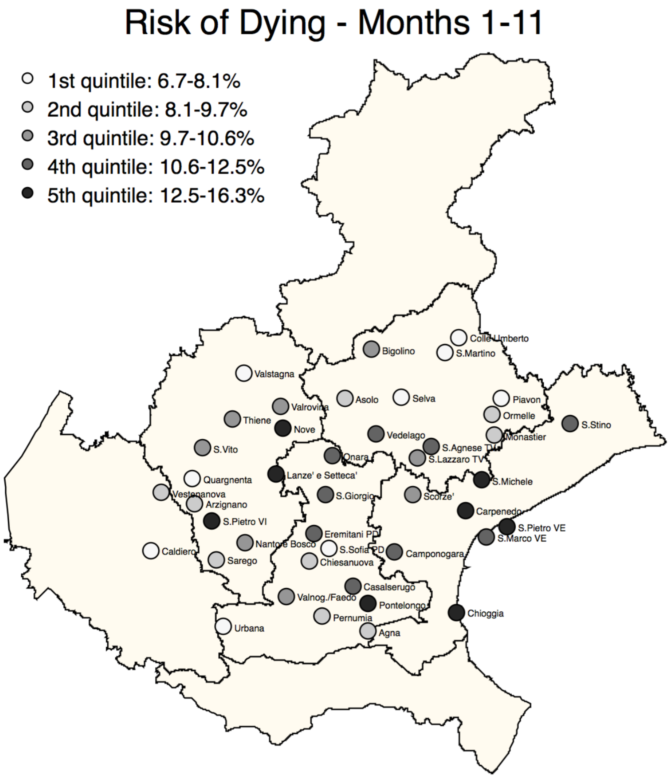


FIGURE 3.7: Child mortality rate (months 1-11) in each of the 46 parishes in the database. Veneto, 1816-70

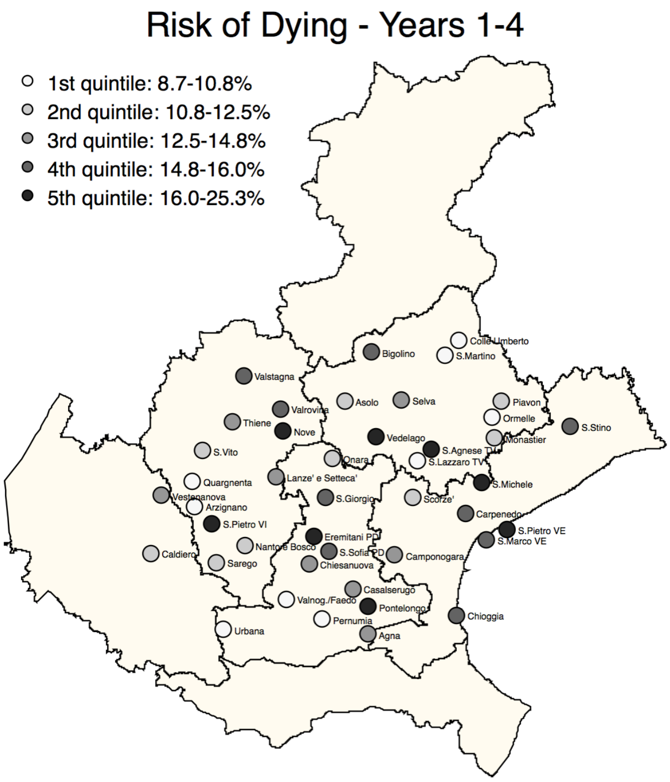


FIGURE 3.8: Child mortality rate (years 1-4) in each of the 46 parishes in the database. Veneto, 1816-70

3.2.5 Temporal trends

General life tables may hide the historical trends of mortality that likely characterizes a period of time more than 50 years long. Fig. 3.9 shows that the mortality rates in the first three ages importantly decrease starting from the mid-30s, that is twenty years after the first subjects in our database were born. The rates suddenly rise ten years later, likely due to the war wave of 1848 (and, more generally to the disorders connected to the First Italian War of Independence)[5], and then steadily drop again. In the mid-60s they are 35% lower than 50 years before.

The situation is partially different if we look at the mortality rate in the period between the first and the fifth birthday: we have a peak in 1826-30, then a decrease and after a new peak in the second half of the century (this peak is shifted with respect to the other three lines). At the end of the period under analysis, this rate is higher than it was at the beginning of the period.

This last evidence is not in line with the situation of a different region like England: according to McKeown (1962) and Woods (1992), the British neonatal mortality (i.e. mortality in the first month of life) had an almost constant level during the whole 19th century, starting its substantial decline only at the beginning of the next century, while the childhood mortality (i.e. mortality in age 1-4) already experimented a shrinkage of some 30% in the second half of the 19th century. This means that, inversely to what was happening in Veneto, in England the beneficiaries of the improvements of living conditions (and especially of child-care practices) were children aged one month and older. Anyway, at the beginning of our observational window neonatal mortality was much lower in England (see paragraph 3.1.2), and this is one of the reasons for the opposed trends registered in England and in Veneto. We will come back to this evidence in the Conclusions of the thesis.

This trend is confirmed if we look closer at what happens in the first month of life. Fig. 3.10 proves that the youngest cohorts (1835-49 and especially 1850-64) were exposed to a much lower risk of death than the oldest cohorts (1816-20). In particular, it is in the first seven day of life (starting from day 2) that we observe the hugest differences among the cohorts.

Finally, combining the analysis of the historical trends of mortality and the analysis of three individual variables and one community variable, we show here four additional results. According to Fig. 3.11, the decrease of mortality in the first week of life starts before for females than for males (Fig. 3.11a); looking at the first month excluding the first week (Fig. 3.11b) the situation is reversed - with males starting the decrease

[5]Also cholera outbreaks contributed to the mortality peak around 1848

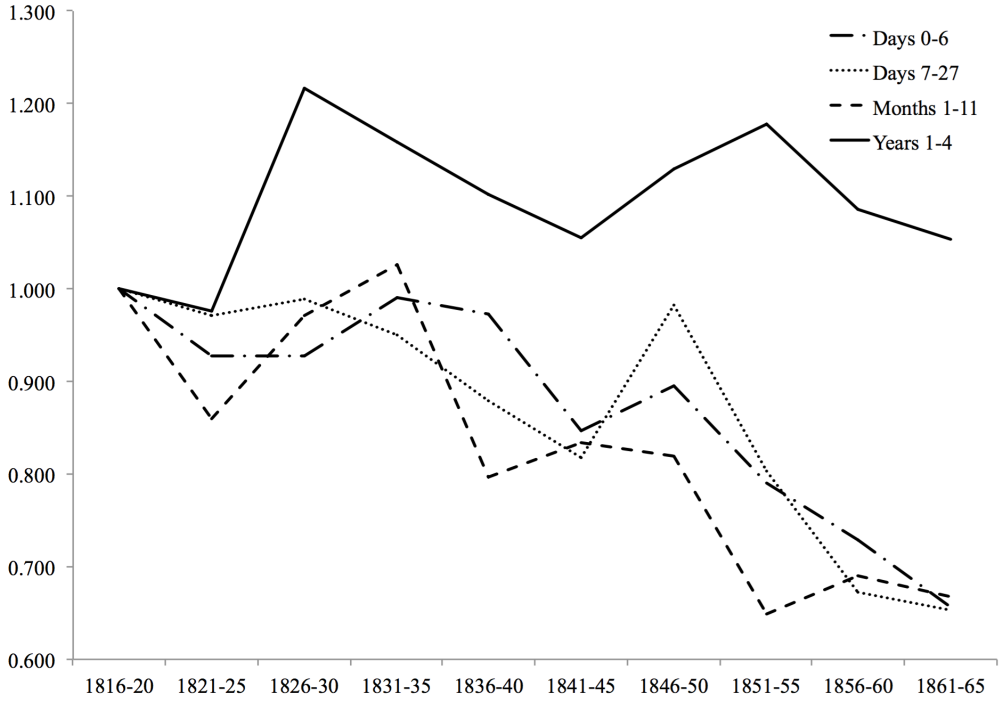


FIGURE 3.9: Historical trends of mortality in the first week, in the first month (excluding the first week), in the first year (excluding the first month) and in the first five years (excluding the first year) of life. On the vertical axis, the ratio with respect to the first period under analysis (1816-20). 46 parishes of Veneto, 1816-65.

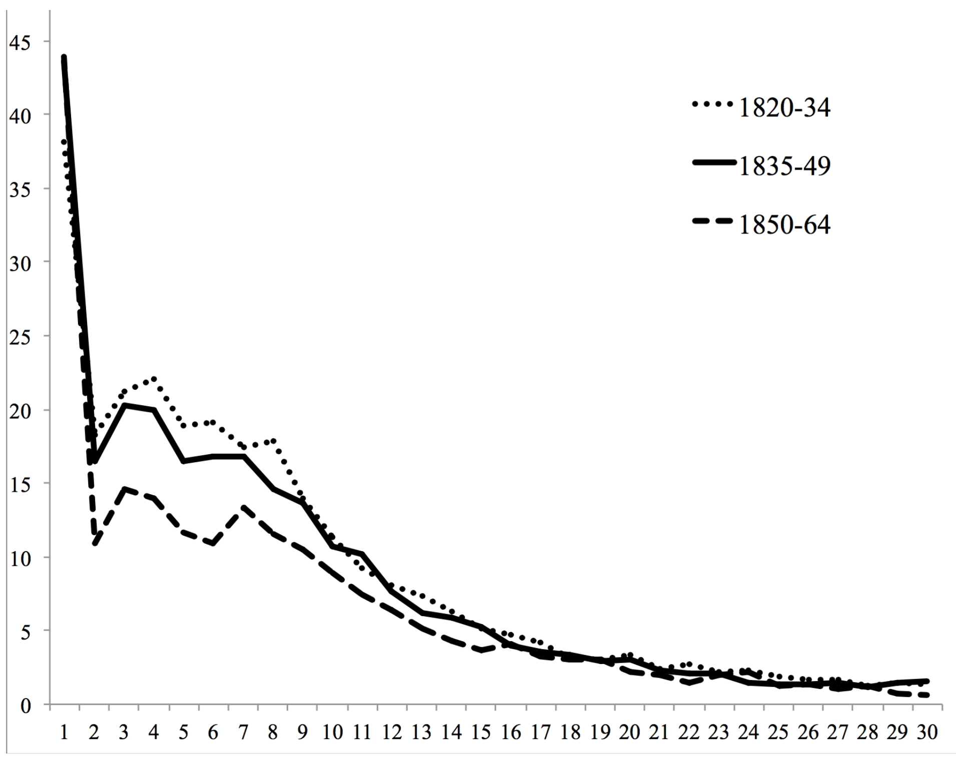


FIGURE 3.10: Historical trends of mortality in the first month of life, day by day, for three different periods of time. 46 parishes of Veneto, 1820-65.

before - while the trend is mixed in the months 1-11 (Fig. 3.11c). If in these three cases at the end of our observational window each group had reached a lower level than the beginning of the period (with the gains varying according to the group), this is not true looking at the four years following the first birthday (Fig. 3.11d): the level of mortality in 1861-65 is pretty much the same for males, and it is even higher for females. This pattern is repeated looking at the mortality trends of the following groups.

Comparing the first week of winter-borns and summer-borns (Fig. 3.12a), the former gain a bigger decrease of mortality at the end of the period (their initial level of mortality was much higher in absolute value); looking at the first month excluding the first week (Fig. 3.12b), the situation is reversed, with summer-borns first experiencing a mortality drop. Again, the trend is mixed in months 1-11 (Fig. 3.12c) and increasing in years 1-4 (Fig. 3.12d).

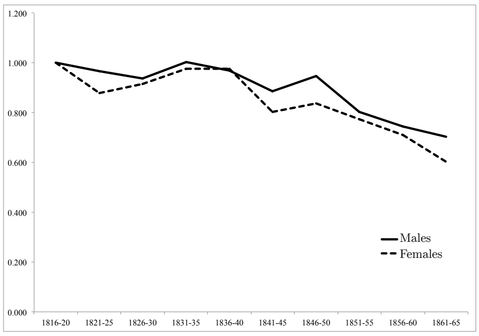
Children of non-peasants experience the mortality drop prior to children of peasants both in the first week and in the first month excluding the first week (Fig. 3.13a and Fig. 3.13b). The situation is inverted in months 1-11 (Fig. 3.13c), while in years 1-4 the mortality is quite constant for children of peasants and increasing for children of non-peasants (Fig. 3.13d).

The drop of first week and first month mortality starts before in the cities than in the country (Fig. 3.14a and Fig. 3.14b). The trend is less clear if we look at months 1-11 (Fig. 3.14c), while it is increasing (slightly in the country, decidedly in the cities) with respect to mortality in the years 1-4 of life (Fig. 3.14d).

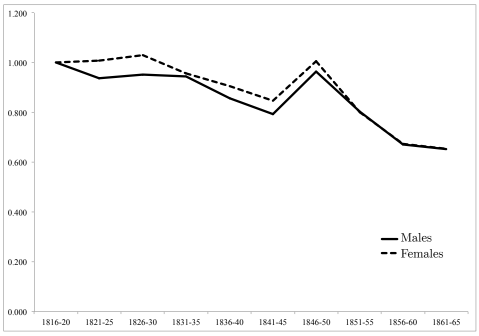
3.2.6 Death causes

In the death registers, in the last space of each row, usually the final cause of the child's death was described, according to a doctor's diagnosis. Albeit this information is not completely reliable, since we can presume that the medical knowledge was generally not really accurate, it could be useful to provide some more details about the general context. Anyway, since this kind of analyses would need a lot of specialistic efforts and a punctual dedicated study, the aim of this paragraph is simply to give a very general idea of what we are talking about.

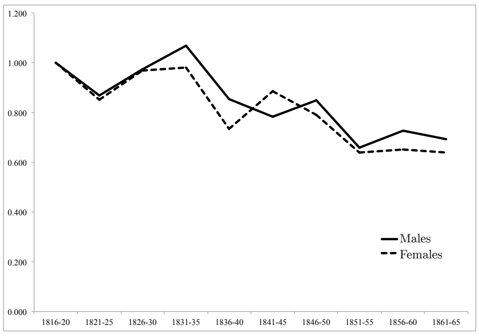
Out of the total 61,947 dead during the first 5 years of life, the cause of death has been registered for 60,783 children. The interpretation and transcription of the priests' judgments produced a list of more than 200 different causes of death, that - for the sole purpose of this overview - we have dimensionally reduced to 6 big groups. We have identified a class of respiratory diseases (47.2% of the total), a class of gastrointestinal



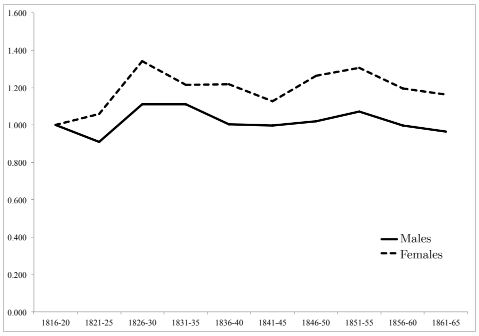
(A) 3.11a



(B) 3.11b

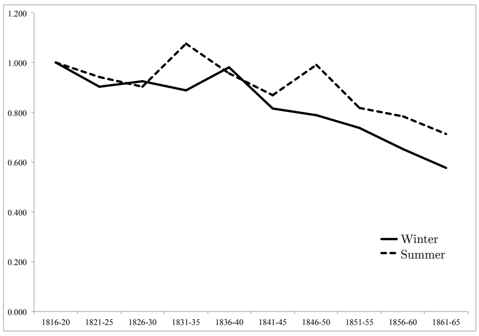


(C) 3.11c

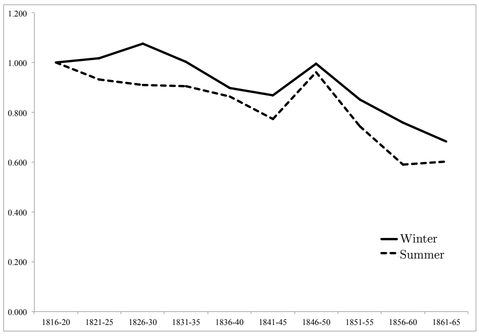


(D) 3.11d

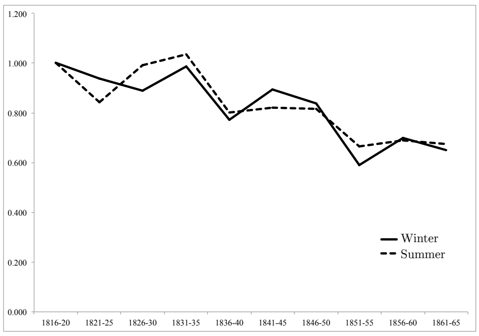
FIGURE 3.11: Historical trends of mortality in the first week (A), in the days 7-27 (B), in the months 1-11 (C) and in the years 1-4 (D) of life comparing males and females. On the vertical axis, the ratio with respect to the first period.



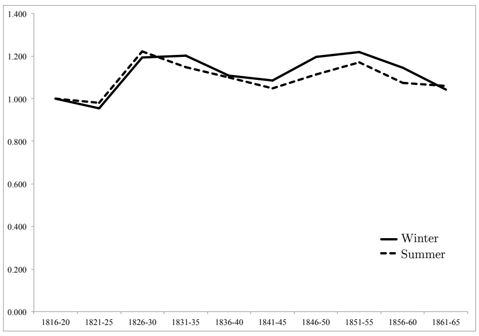
(A) 3.12a



(B) 3.12b

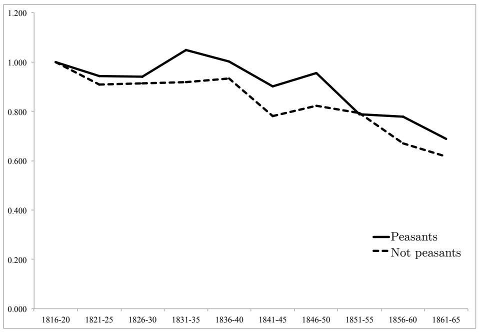


(C) 3.12c

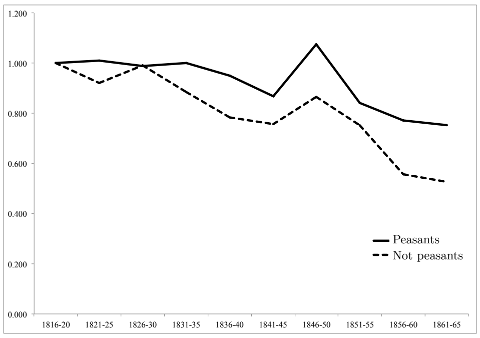


(D) 3.12d

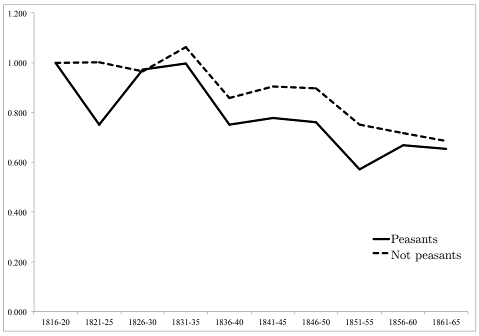
FIGURE 3.12: Historical trends of mortality in the first week (A), in the days 7-27 (B), in the months 1-11 (C) and in the years 1-4 (D) of life comparing children born in winter and children born in summer. On the vertical axis, the ratio with respect to the first period.



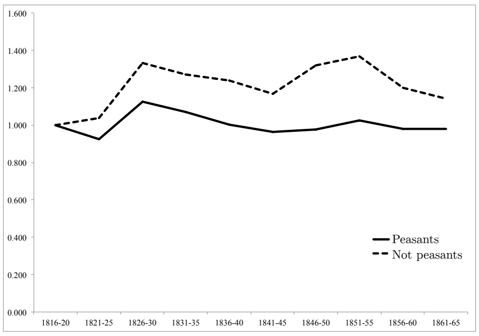
(A) 3.13a



(B) 3.13b

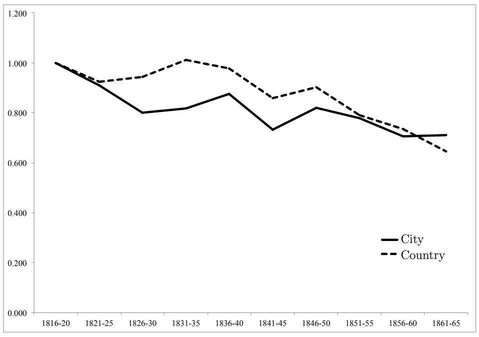


(C) 3.13c

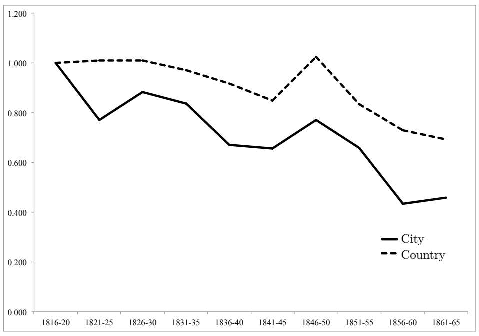


(D) 3.13d

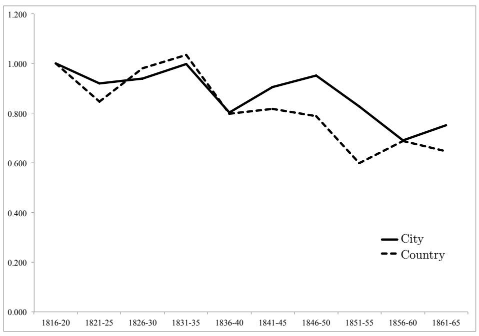
Figure 3.13: Historical trends of mortality in the first week (A), in the days 7-27 (B), in the months 1-11 (C) and in the years 1-4 (D) of life comparing peasants and non-peasants. On the vertical axis, the ratio with respect to the first period.



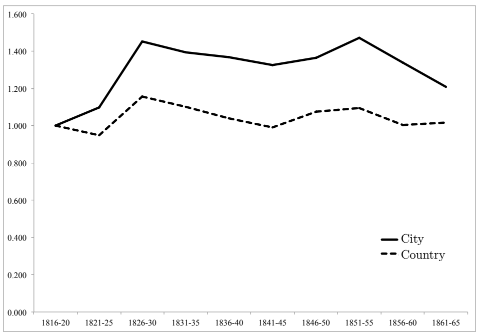
(A) 3.14a



(B) 3.14b



(C) 3.14c



(D) 3.14d

Figure 3.14: Historical trends of mortality in the first week (A), in the days 7-27 (B), in the months 1-11 (C) and in the years 1-4 (D) of life comparing cities and country. On the vertical axis, the ratio with respect to the first period.

diseases (19.7%), a class of birth-related diseases (15.7%), a class of vascular diseases (3.7%), a class of infectious diseases (3.0%) and a class of "other" diseases (10.7%).

As proven by Fig. 3.15, more than 35% of the deaths by respiratory diseases take place during the cold season, while only a small proportion (about 15%) take place during summer. On the other side if we look at the situation of gastrointestinal and infectious diseases, we see that those categories of illnesses are more lethal under warmer temperatures. Albeit coming from a coarse classification of the death causes, we will keep in mind this pattern in the upcoming section.

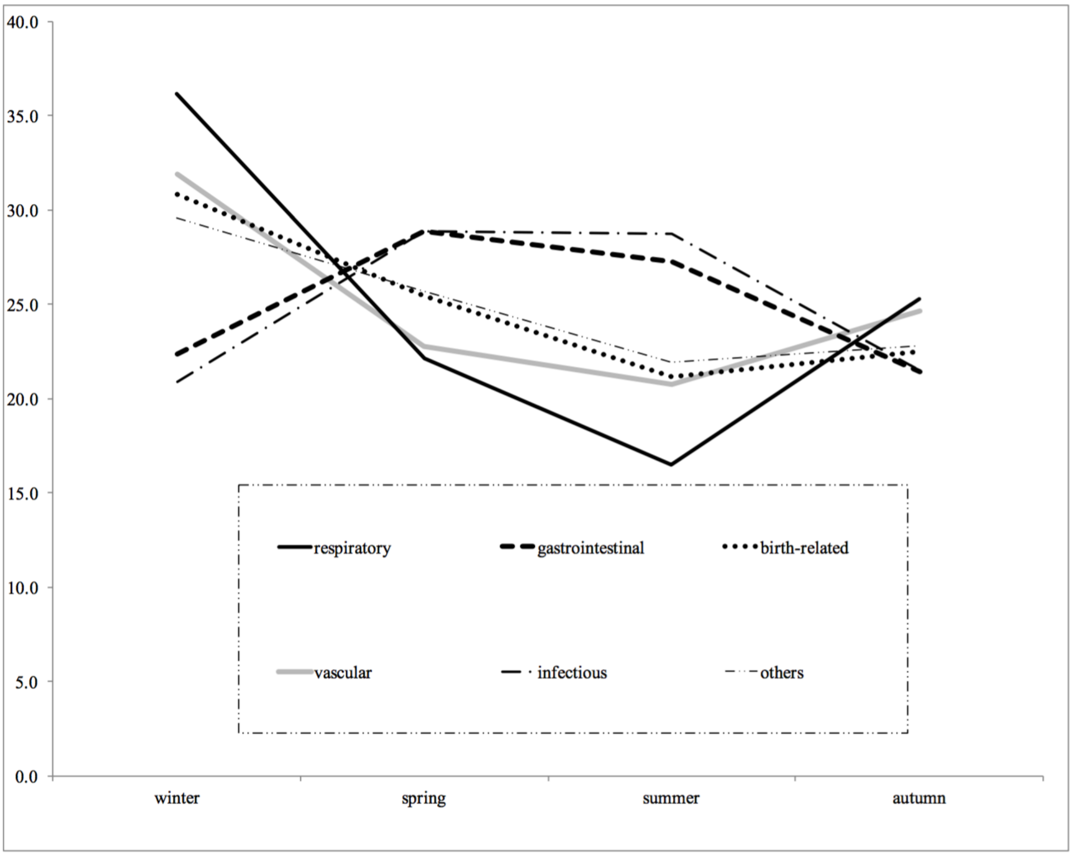


FIGURE 3.15: % of deaths happening in each season for six class of diseases. 46 parishes of Veneto, 1816-66.

The results presented up to this point in this chapter address our research in two directions. The first one is a focus on the geographical differences emerging among the parishes: why the levels of mortality are different from place to place? Why the temporal trends do not look geographically homogenous? What are the relations between the individual and the community characteristics influencing child and infant mortality?

On the other hand, it is also necessary to understand something more about the mortality selection process, that we have already invoked as a possible explanation for some results earlier in the chapter, stressing the regularity of the inverse relationship existing between mortality in the first month of life (and especially in the first week) and mortality in age 1-4 years. These two topics, that are likely related, will be discussed using statistical models in the next chapters, after the focus on the month of birth.

3.3 Focus on the Month of Birth

3.3.1 We are not astrologists

In February 2011, the Bombay High Court ruled that "astrology is a trusted science, and is being practiced for over 4000 years". Even after this judgment, a science-oriented mind is not starting to read daily horoscopes to discover what will happen to them next Friday. But the debate about the influence of interplanetary alignments at time of birth as explanation of one individual's health and character is not over. Nothing new under the sun, you could properly say. Masterpieces of the world literature like 'Romeo and Juliet' and 'King Lear' by William Shakespeare and 'Divine Comedy' by Dante Alighieri are literally driven by astrologically-based concepts. More recently, new attempts of showing the validity of some astrological bases have been registered. In 1955 the french psychologist Michel Gauquelin tried to show the existence of a statistical correlation between athletic eminence and the position of the planet Mars at time and place of birth, claiming that a significant number of sports champions were born just after the planet Mars rises or culminates. However, later research deleted the 'Mars effect' theory from the trusted scientific landscape.

Now, as demographers, we can not say a lot about interplanetary alignments and still less about divine grace. Also the study of a man's personality based on his birth time is something exotic for us. But we think there is a credible way demography can enter this debate. There is a scientific theory, increasingly accepted and confirmed by evidences, that links prenatal and early-life 'situations' to chronic conditions later in life. The thrifty phenotype hypothesis (also known as the Barker hypothesis), unsurprisingly, do not consider the position of stars at the birth time as a critical event for later life's diseases, but refers to more demographers-friendly information like weight at birth, mother's wealth, season of birth (as a social and meteorological proxy).

There is a wide literature proving the relationship between early-life and adult-life conditions. Much less explored is the effect of early-life events (i.e. first days of life) in a very short term (i.e. first months of life). As stated in Chapter 1, we aim to study the effect of pre-natal and very-early-life context on the immediately following life (i.e. infancy) with no time gap between causes and effects. We think this approach can be interesting not only for the lack of intermediate effects, but also because we are analyzing a context

(1800s in Veneto) in which more than 40% of children did not reach their 5th birthday. Moreover, it is a context in which the demographic transition was not started yet, since mortality and fertility were both very high. In this case, understanding the survival mechanism working in the very first phases of life could have more relevance rather than focussing on what happened to the few ones that were able to reach the adult age.

Kynast-Wolf et al. (De Stavola et al. 2006), with regards to Burkina Faso, observed an association between high infant mortality and being born at around the end of the rainy season. The authors explain this pattern as a malaria effect: children living in areas with high malaria transmission intensity are most susceptible for severe malaria during the second half of their infancy. In effect, the effect seems to be strong only within the first year of life. This result replicates the findings by Becher et al. (2004), where the season of birth was included among the major risk factors (along with death of the mother and twin birth) of infant and child mortality in the same area.

In order to go deeper in this kind of analysis, the first step was to describe the variables that we already have, which may help us in disentangling the very general situation represented by Table 3.2. We are going to use proxy variables for the five major risk groups for infant mortality:

- 'environmental health conditions' and 'nutritional status + infant feeding' can be represented by a combined use of geographical data and date of birth (or weather information, if available);
- 'socio-economic status' and 'use of health services' can be represented by the professional status of the child's father;
- 'fertility behavior' can be represented by the number of siblings (where available).

Summing up, if we wish to answer the questions about the effect of prenatal and early-life risky exposures on post-neonatal and infant mortality in our specific context (and if we wish to say something about horoscopes, too), we need to find a good proxy variable.

3.3.2 Reproducing some results

Breschi and Livi-Bacci (1997) examined the month of birth as a factor in child survival in Europe, during the mid-late 1800s. The existing relationship between month of birth and age seemed to operate in different ways in each of the five European countries under analysis (Belgium, Italy, Netherlands, Russia and Switzerland). In particular, the Italian model (1872-79) can be described by a U-shaped curve with first month mortality peaks in winter months (due to respiratory infections). The Belgian model is similar to the Italian one with smaller differences between winter and summer months. Switzerland and the Netherlands present a flat model, with reduced differences between winter and summer. This is an effect of the better care of children during the cold season. In Russia instead, mortality was higher for children born in the spring: it is likely that the participation of females in summer jobs would be linked to early weaning and lesser care. And it is known that gastroenteric diseases (due to contaminated food) are more lethal during the warm seasons.

Moreover, the authors presented a regional detail for some Italian regions (Savoy, Veneto, Sicily). Unsurprisingly, the highest level of infant mortality (q_{0-11}) was registered in Veneto, where the winter infant mortality was also the highest among all the five European countries under analysis (over 300 ‰). In this paragraph, we will reproduce some of the analyses conducted by Breschi and Livi-Bacci: we expect similar but even stronger evidences, since our data refer to a previous period (1816-71) and in the first half of that century the decrement of mortality was not started yet. Also, it is important to mention that here we are not using aggregated data, but a refined individual database.

3.3.3 The role of the month of birth

In paragraph 2.2.2 we showed that the season of birth played a major role in determining the risk of death for a newborn. That is, in our parishes it was not the same to be born in summer or in winter. Actually, we can be more precise and say that in our parishes it was not the same to be born in July or in January.

In Fig. 3.16 we present the monthly profile of infant mortality (first year of life), for the parishes from province of Padua. We selected just one province for simplicity of representation, but the results are analogous for the other provinces. All the parishes profiles are overlapped and they reproduce a U-shaped model with minimum in the summer months and high peaks of mortality in the winter, when the risk of dying was more than doubled. This confirm the results by Breschi and Livi-Bacci that, with data from late 1800s, identified the 'Italian model' as the U-shaped one, with respect to other models fitting different European countries situations. It is interesting to note that, even if the shape is almost the same for all the parishes, the absolute mortality levels (and in particular the winter peaks) are really different if we pass for example from Faedo (almost 700‰in December) to Santa Sofia (less than 40‰in December). We are disentangling these territorial differences later in our study.

Fig. 3.16 presented the profiles of the probability of dying during the first year of life according to the month of birth. But a whole year can be a very long period and it can

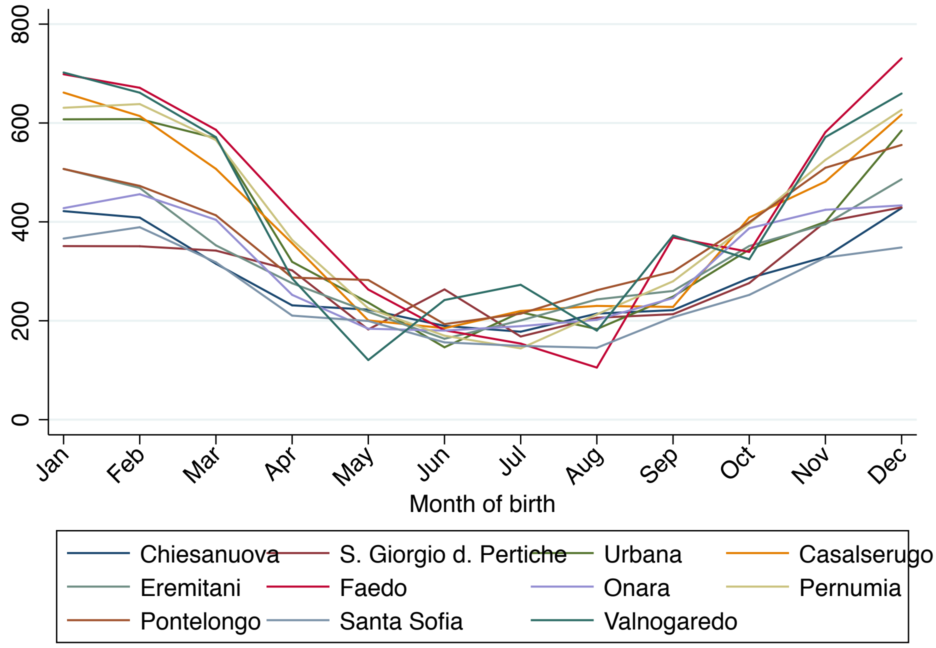


FIGURE 3.16: Probability of dying in the first three months of life by month of birth and parish. 11 parishes from the province of Padua (1816-70).

hide different dynamics. In effect, the situation is different if we decompose our time span (Fig. 3.17):

- the probability of dying during the first two days of life (blue line) is higher for children born in the winter months, but we can not appreciate big differences with the children born in the summer months. The first two days were always lethal. In other words, the profile is not properly U-shaped;
- the big difference between winter-born and summer-born children emerges when we look at the probability of dying during the first week (days 2-7 of life) and the first month (days 7-30 of life);
- the situation is reversed when looking at the yellow line (prob. of dying between days 30 and 365 of life). Spring and summer seem now to be the unluckiest seasons to be born in: the highest mortality is registered for children born in September, the lowest for children born in March. The next paragraph will help us to understand this 'new' situation, that is clearly related to the action of age and, to some extent, to the eventual selection occurred in the first month of life.

These results are comparable to those by Dalla Zuanna and Rosina (2011): "the most critical period is during the first week, omitting the first day."

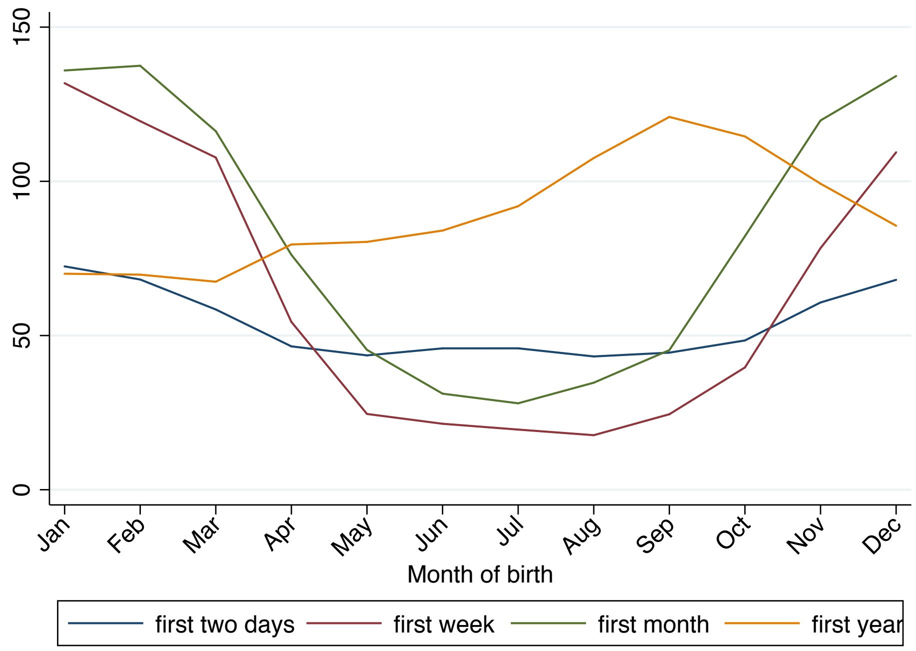


FIGURE 3.17: Probability of dying according to different time spans, by month of birth. 46 parishes from Veneto (1816-70).

3.3.4 Interaction between month of birth and age

The results from the previous paragraph confirm that there is a mechanism that relates month of birth (or season) and mortality in early life, as expected. This mechanism is not straightforward, as proved by Fig. 3.17 , since it involves the interaction of the seasonal effect of mortality per se (climate and socio-cultural behavior) and the age at which each child goes through a specific season. As suggested by Breschi and Livi-Bacci, there are three important links between season and mortality to take into account.

- **Winter-specific Effect**: in the first days/weeks of life children born in winter are at risk of respiratory infections. Inadequate protection and particularly cold winter increase the risk of death (as in Veneto, e.g.);
- **Summer-specific Effect**: risk of contracting infections of the digestive tract. This is a risk that varies according to the age at which the child goes through the summer. Moreover the maternal milk is a protective factor, granting a sort of immunity from gastro-enteric diseases;
- **Length of Breastfeeding**: for example, a child born in the summer season and weaned before his first birthday is subject to a higher risk of digestive infections during his second summer of life, with respect to a child born in the summer season and weaned after his first birthday.

Figures 3.18 and 3.19 help us in disentangling some of these issues and give us some confirms. Fig. 3.18 presents age-specific probabilities of dying for each month of age and for monthly cohorts (winter months, each color is a different month). Winter cohorts had a very high risk of death during their first month of life, and later the risk became low and not subject to particular variations. In this situation, we can see selection at work: the weakest individuals are 'eliminated' soon, because of adverse climate conditions. The strongest remain alive, and are subjected to a lower risk of dying, later [6].

On the other hand, fig. 3.19 presents the same probabilities for the summer cohorts. They had a lower mortality in the first month (if compared to winter cohorts) but a significant peak of mortality during their second summer - when they were 12, 13 or 14 months old - than during the first one, when they were 0,1 or 2 months old. In effect, the second summer was faced when the protective function of breastfeeding was over, as stated before.

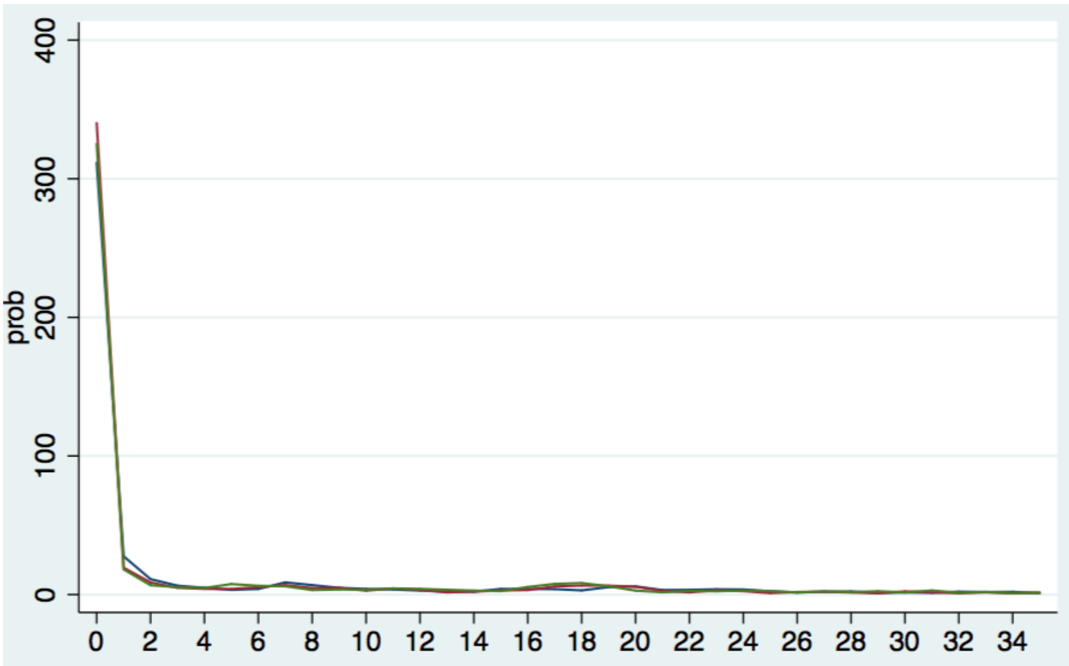


Figure 3.18: Probability of dying by age and month of birth (winter months; green line is december, red line is january, blue line is february). 46 parishes from Veneto (1816-70).

[6] The next chapter is entirely devoted to the analysis of the selection process.

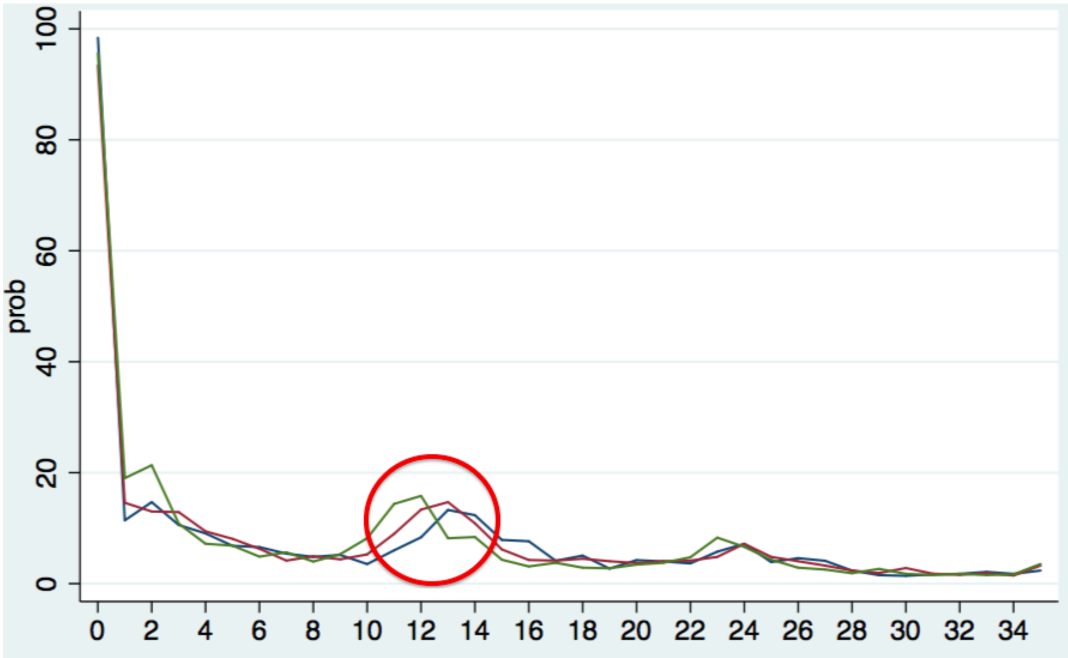


FIGURE 3.19: Probability of dying by age and month of birth (summer months; green line is june, red line is july, blue line is august). 46 parishes from Veneto (1816-710).

Chapter 4

Mortality Selection

In this chapter we resume the work done in the paper *Mortality selection in the first three months of life and survival in the following thirty-three months in rural Veneto (North-East Italy) from 1816 to 1835* (Piccione et al. 2014), adding new details and using a wider database, in order to obtain confirms and push forward the discussion about mortality selection.

4.1 What is mortality selection

The issue examined in the chapter can broadly be described as follows:

Many period-specific influences during life may affect mortality at a later stage, although it is often unclear which direction the resulting influence will take. For example, epidemic experiences or conditions of famine early in life may so debilitate a cohort that it subsequently experiences higher death rates relative to earlier or later cohorts. Others, however, argue that high mortality experienced in early life selects out the frail members of the cohort, resulting in reduced subsequent mortality. (Hobcraft and Gilks 1984, quoted in Billari and Rosina 2000).

While the two effects may coexist within the same situation, we can only measure the final balance within a population (Preston, Sill, and Drevenstedt 1998; Quaranta 2013). A number of papers attribute the influence of life conditions in infancy (and pregnancy) on mortality risks during adulthood or old age: while some authors find a prevalence of insult accumulation, others observe positive selection[1].

If the events experienced in utero and during the first years of life influence mortality risks many years later, the same dynamic could be present, and perhaps even stronger, between contiguous periods. We study the effect of mortality selection during the first three months of life (from here on 'early infant mortality') on mortality in the 33 months that followed (from here on 'late infant mortality').

We aim at understanding whether children who survived strong mortality risk during the first three months of life were 'positively selected' (i.e., they had a higher probability of surviving during the following 33 months) or if rather they experienced a so-called scar-effect, i.e., a process of insult accumulation prevailed such that they had a lower probability of surviving during the following 33 months [2].

We consider children born in Veneto (North-East Italy), using micro-data from the archives of seven parishes for the period 1816-1835, before the decline in mortality which characterized the demographic transition. The old demographic regime in Veneto provides an excellent context for studying this issue in that, as we will see below, there were considerable differences in early infant mortality, both across different areas of the region and especially across seasons. This intense variability makes it possible to identify, in detail, the possible influence of early infant mortality on mortality during the following 33 months.

The idea of examining the connections between early and late infant mortality builds on our previous research. Dalla-Zuanna and Rossi (2010), using aggregate data on the mortality of children within the 19 regions of the Austro-Hungarian Empire for the cohort born in 1851, show that in the Lander where mortality was high in the first three months of life there was relatively low mortality in the successive periods. Up until the 36th month of life the association between q_{0-2} and q_{3-35} (where the subscripts

[1] See e.g., Caselli and Capocaccia 1989; Kannisto, Christensen, and Vaupel 1997; Preston, Sill, and Drevenstedt 1998; Bengtsson and Lindstrom 2000, 2003; Doblhammer 2004; Shkolnikov 2012; Ekamper et al. 2013; Quaranta 2013. See also Bruckner, Catalano, and Smith 2013; Helgertz and Bengtsson 2013; Shen and Zeng 2013 (papers presented at session 249 *Pathways to health: direct and indirect effects of early life conditions on later health*, at the IUSSP Conference of Busan, South Korea, on August 2013).

[2] In this paper we define early infant mortality as mortality risk during the first 3 months of life (age 0-2 months), and late infant mortality as mortality risk from the 4th to the 36th month of life (age 3-35 months). The idea is to focus on two age intervals that differ with regard to the causes of death: mortality after the 3rd month of life should only be slightly affected by birth shocks or endogenous weakness, and mainly by exogenous causes (Bourgeois-Pichat 1951). Data on causes of death written in the death registers of our database (even if rough) partially support the choice of using the end of the 3rd month as a breakpoint between early and late infant mortality: during the 4th month of life, two causes of death strongly related to birth ('spasm' and 'asthenia') drop to 39% of deaths compared to 62% for children who died during the 1st month, 64% in the 2nd month, and 59% in the 3rd month. We do not consider children who died over the age of 3 because - as we described in Chapter 2 - the mechanism of linking birth and death records becomes more uncertain as age at death increases.

are months, as elsewhere in this article) was negative, albeit moderate (r = -0.35, see Figure 4.1). In the Habsburg Empire of the mid-1800s there was a kind of compensation between early infant mortality and mortality in the following months, what we describe as the prevalence of positive selection. In this work we are going to similarly assess whether such an effect exists within a smaller territory, studied with a much greater degree of detail thanks to the availability of individual data.

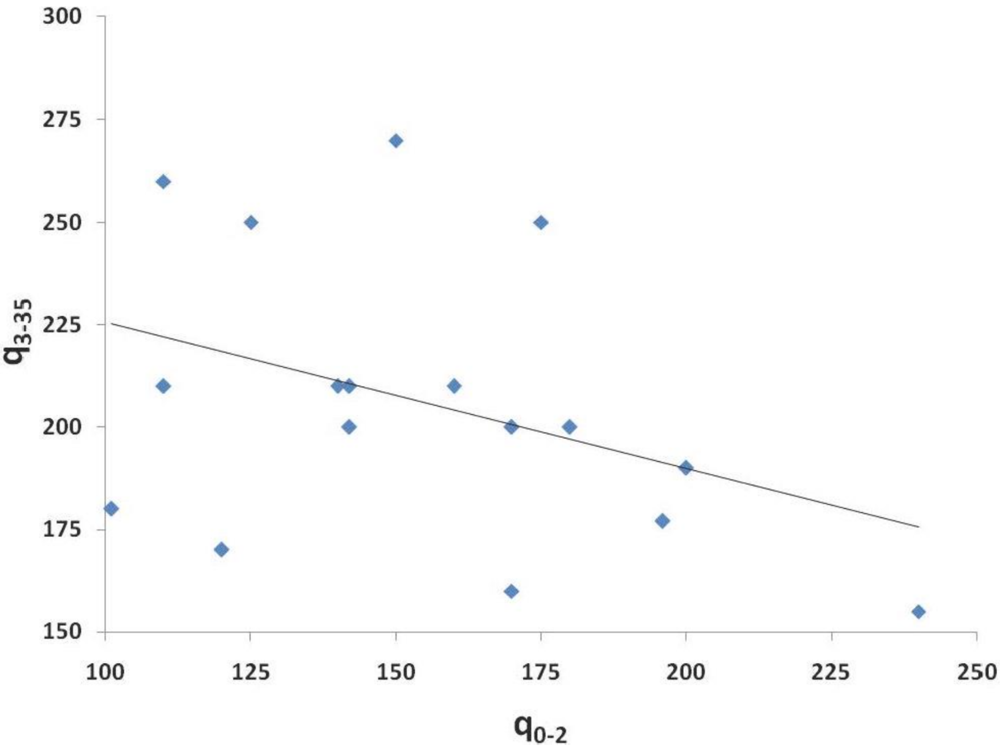


FIGURE 4.1: Probability of dying (per thousand) during months 0-2 and 3-35 in the nineteen Lander of the Austro-Hungarian Empire for the cohort born in 1851. Source: Dalla-Zuanna and Rossi 2010.

4.2 Theoretical background

The issue of selection has already attracted the attention of scholars from various fields of study. When thinking about selection it is impossible not to refer to Charles Darwin and his theory of evolution. In Darwin's theory, natural selection certainly plays a fundamental role. Yet a further reading of Darwin suggests that his idea of natural selection also included more than mortality. Mortality selection is indeed a specific case of natural selection; it is a directional positive selection that favors 'the fittest' (Endler 1986), i.e., people with the strongest probability of surviving. The selectionist argument, according to Peters (1991), is a logical truth and not a scientific theory per se. It is a very general concept that has been employed by different disciplines, from economics to epistemology. Demographers might very well benefit from knowledge gained by developments in sciences like biology and biomedicine. For example, Carey (1997) aimed to demonstrate the broad relevance of actuarial studies on Mediterranean fruit flies (medflies) to issues pertaining to humans. In turn, a number of researchers have used demographic selection arguments, with contrasting results, to explain the levelling-off of medfly mortality (Carey, Liedo, and Vaupel 1995).

In population studies the notion of selection is often evoked to explain the demographic behavior of migrant populations. For example, it has been suggested that low levels of mortality among migrants are the result of individuals who migrate being healthier and stronger compared to both those who did not emigrate and natives (the so-called 'healthy migrant effect', see e.g., McDonald and Kennedy 2004; Kennedy, McDonald, and Biddle 2006). A similar selection effect has been used to explain the marital and reproductive behavior of immigrant couples and women (Goldstein and Goldstein 1983; Chattopadhyay, White, and Debpuur 2006).

With regards to mortality analyses, however, things are quite different, in that selection is not determined by an external event, as with migration, but by the same processes of elimination that apply to death. Demographers most often employ a selectionist approach to theories on ageing. As a population ages it simply becomes increasingly positively 'selected', as individuals with higher risk of death die sooner. The result is a population formed mostly by individuals with relatively low death rates (Carey and Liedo 1995a; Vaupel and Carey 1993; Vaupel, Manton, and Stallard 1979). When focusing on the older age groups, this concept helps to explain deviation of the human mortality trajectory from the well-known Gompertz function (Manton and Stallard 1996). Due to positive selection, the increase in mortality rates at older ages is slower than the 8%-10% per year observed in middle age.

Both the demographic and biological approaches share at least one common point when employing the notion of selection. If two identically heterogeneous populations have different levels of mortality, then changes in the mortality trajectory related to selection will occur at younger ages in the population with higher mortality, ceteris paribus (Carey 1997). Another example of positive selection has been observed for Swedish cohorts born during the 20th century: the risk of ischemic heart disease mortality during cold periods drops among children who were in utero during a particularly cold period (Bruckner, Catalano, and Smith 2013). Positive selection has also been observed with regard to significant crises in mortality: the few available case studies on mortality before and after the plague epidemics that swept through Europe between the 14th and the 18th centuries verify this result. Following these plagues, people who survived epidemic outbreaks would have enjoyed higher survival rates than those before the plagues, although these results are somewhat controversial (Biraben 1975-76; Del Panta 1980; Billari and Rosina

1998, 2000; Alfani 2010). It is, in fact, difficult to discern whether the plagues swept away the weakest or if, alternatively, the smaller surviving populations (which may have been as much as 30%-50% less than before the plagues) relaxed Malthusian pressure, thanks to a higher ratio between resources and population. The lack of data, specifically on population during the centuries of plague in Europe, hinders further unravelling of the issue.

Other studies have shown the prevalence of the opposite effect: cohorts who survived serious mortality crises during the early years of life were also weaker later in life compared to cohorts slightly younger or older. For example, a process of insult accumulation occurred in Italy for male cohorts born in 1880-1899 who survived World War I: their mortality in adulthood and old age was slightly higher than that of their elder and younger 'brothers'; a consequence of diseases contracted during the war itself (Caselli and Capocaccia 1989; Caselli 1990). In southern Sweden during the 18th and 19th centuries, children who survived severe epidemics of smallpox or whooping cough, and severe nutritional stress, suffered higher levels of mortality as adults and in old age, particularly relative to airborne infectious diseases (Bengtsson and Lindstrom 2000, 2003). The same occurred with regard to Swedes born between 1912 and 1915 affected by the Spanish flu: exposure to the latter in the first years of life negatively affected both economic outcomes and health in old age. Health is also directly affected (not only indirectly via income and socio-economic position) (Helgertz and Bengtsson 2013; Quaranta 2013), and it has a long follow-up period, up to 70 years of age for cohorts born in Sweden during the 19th and the early 20th centuries.

Quaranta (2013) finds positive selection for male birth cohorts suffering from high infant mortality, from age one up to about age 30, after which the scar-effect is greater than the selection effect. An increase of 12% in all causes of death between the ages of 18-63 was observed in the Western Netherlands among people born in 1944-1945 whose mothers were affected by the Great Famine during the first half of gestation, compared with those whose mothers were not so affected. This difference is considerable, and is greater compared to individuals whose fathers were employed in manual vs. non-manual occupations (+8% for manual), and is independent of social class and education at age 18 (Ekamper et al. 2013).

A similar dynamic also holds true for the cohorts most severely affected by the Great Famine of the 1930s or by World War II in Ukraine (Shkolnikov 2012). Other studies do not, however, show adult mortality differences between cohorts affected or not affected by famine during childhood (Kannisto, Christensen, and Vaupel 1997). The scar-effect has also been observed in China, among individuals born in the first half of the 20th century. The total negative effect of adverse childhood conditions on current health is more prominent among the most elderly (aged 80+ in 2008-2009) compared to those who are slightly younger (aged 65-79, Shen and Zeng 2013).

4.3 An extremely high mortality

As already stated, in the period 1750-1840 the mortality of children in Veneto was extremely high, mainly due to mortality during the first month of life for children born during winter [3].

Extending the comparisons of Chapter 3, neonatal mortality in our context of analysis was even much higher than that observed at the end of the 20th century in the poorest countries of the world, where the probability of dying in the first week of life (including the first day) was never higher than 45‰ (Hill and Choi 2006), not far from that observed in some Indian contexts around 1970, but only for children with a weight at birth of 1500-2000 grams (Visaria 1988). However, levels similar to that of Veneto were found in other European pre-transitional contexts. For example, in the above cited family reconstruction of several Bavarian villages, Knodel (1988) finds that neonatal mortality (during the first month) ranged between 67‰ and 190‰; Goubert (1968) finds the level of neonatal mortality to be nearly 150‰ in some 18th century French rural villages, 50% higher than that observed in the white population of Quebec in the period 1621-1729 (Lalou 1997) [4].

During the eighty years that followed, mortality during the first year of life, and especially winter mortality during the first month, quickly declined (Dalla-Zuanna and Rosina 2011). While scholars have examined the demographic history of Veneto between the last stages of the ancien regime and the demographic transition (Pozzi 1991; Derosas 2002; Rosina and Zannini 2004), the patterns of high mortality during the first month of life remain far from explained. That said, there are clear indications that the influence of cold external temperatures was decisive (Derosas 2009; Dalla-Zuanna and Rosina 2011).

According to regression models fitted by Dalla-Zuanna and Rosina (2011), a decrease of 1 °C in the minimum external temperature during the winter corresponded to a 5% increase in the daily risk of death during the first month of life. The daily risk of

[3] Levels of infant mortality equal to 350‰, such as those recorded in the Veneto region in the century of 1750 to 1850, are very high, even if not exceptional in Europe before the health transition. For example, mortality during the first year of life was 486‰ (1862-70), 502‰ (1871-75), and 444‰ (1879-88) in the Bavarian district of Augsburg, and 428‰, 471‰, and 400‰ during the same three periods in the village of Anhausen, belonging to the same district, according to the family reconstruction by Knodel (1968). However, the pattern of Bavarian parishes is not in line with the Veneto pattern, as mortality during the first month does not vary according to season (Knodel 1988).

[4] For additional reviews and comparisons see also Oris, Derosas, and Breschi 2004; Pozzi and Barona 2012.

death during the third and fourth days of life varied from 80‰ to 130‰ to 220‰ if the minimum external temperature varied respectively from +5 °C to 0 °C to -5 °C.

4.4 Preliminary results and methodological choices

Figure 4.2 shows that the death probabilities q_{0-2} and q_{3-35} in the pool of the 46 parishes are relatively constant until 1840 and then decline in the years that follow. As our aim is to study selection before the decline in mortality we include in our analysis only the 62,637 children born in the period 1816-1835 [5].

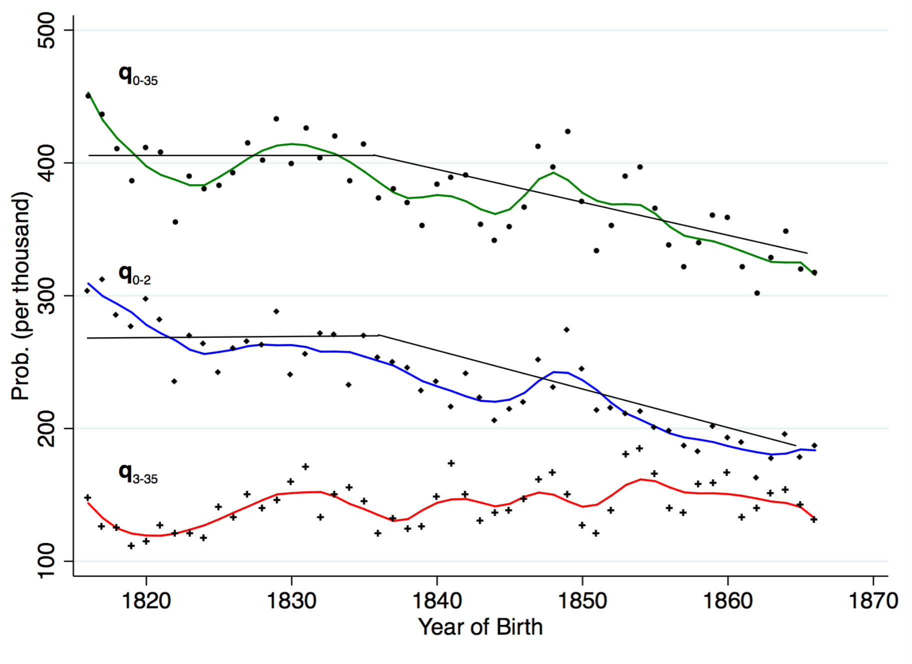


FIGURE 4.2: Annual probability of death at ages 0-2, 3-35, and 0-35 months (‰). Children born in 46 rural parishes in the region of Veneto in 1816-67.

The crucial point is to find the 'key covariate', or the variable that describes the mortality level during the first three months of life, that can in turn be used to explain the mortality risk of children who survived. Previous studies on the impact of temperature and climatic conditions on health reveal that winter mortality in Veneto was clearly higher than summer mortality, mainly during the early stages of life (Breschi, Derosas,

[5] In what follows we provide some additional information on data quality for the 62,637 children born in the seven parishes during 1816-1835. Sex is reported for all children but fifteen (which we exclude in our further analysis), and the M/F sex ratio at birth is 1.07, i.e., a value fully consistent with the sex ratio observed broadly for humankind. For only 4.7% of the children it was not possible to discern the job of the father (44.2% farmer, 26.9% craftsman or belonging to the working class, 8.4% merchant, 1.6% servant, 14.2% employee, civil servant, or land owner). Cause of death was reported in 94% of the death records: this proportion grows notably with age (91% for deaths at age 0-2 months, 98% for deaths at age 3-35 months). Moreover, causes are often confused with symptoms, as was common in death records dating from the 19th century.

and Manfredini 2000; Dalla-Zuanna and Rosina 2011). In our sample also, children born in the cold seasons were exposed to greater risk of death during the first month of life (see Tab. 4.1 and Fig. 4.3). Season of birth could thus be considered a proxy variable for early infant mortality risk.

Figure 4.4 shows the probabilities of death within 3-month periods for the ages 3-35 months, pooling the 46 parishes, according to season of birth (summer vs. winter). The shape of the two graphs is similar, dropping from month 3 to month 35 and peaking in the summer and autumn (mainly due to gastro-intestinal diseases). The general level of mortality during 3-35 months of age is, however, higher for children born in the summer, who survived a lower neonatal death risk compared to children born in the winter, who instead survived a higher neonatal death risk.

Table 4.1: Probability of dying in months 0-2 and in months 3-35 according to some individual characteristics: season of birth, sex, five-years period of birth, parish of birth.

	Months 0-2		**Months 3-35**		**Prob. of death**	
	Births	**Dead**	**Exposed**	**Dead**	**q0-2**	**q3-35**
	a	***b***	***c=a-b***	***d***	***b/a***	***d/c***
Season at birth						
Winter	16,056	6,845	9,211	1,643	426	178
Spring	17,466	3,412	14,054	2,663	195	189
Summer	15,410	2,336	13,074	2,682	152	205
Autumn	13,690	4,541	9,149	1,694	332	185
Sex						
Male	32,457	9,218	23,239	4,451	284	192
Female	30,165	7,916	22,249	4,231	262	190
Period						
1816-20	14,708	4,390	10,318	1,867	298	181
1821-25	16,690	4,396	12,294	2,120	263	172
1826-30	15,738	4,226	11,512	2,318	269	201
1831-35	15,486	4,122	11,364	2,377	266	209
Parish						
TV - Colle Umberto	931	186	745	84	200	113
TV - Bigolino	785	253	532	104	322	195
TV - S.Martino	653	127	526	62	194	118
TV - Selva	1,045	331	714	98	317	137
TV - Asolo	159	56	103	19	352	184
TV - Ormelle	582	169	413	45	290	109

Table 4.1: (continued)

	Months 0-2		Months 3-35		Prob. of death	
	Births	Dead	Exposed	Dead	q0-2	q3-35
	a	*b*	*c=a-b*	*d*	*b/a*	*d/c*
TV - Piavon	659	88	571	75	134	131
TV - Monastier	1,641	366	1,275	222	223	174
TV - Vedelago	627	244	383	73	389	191
TV - S.Agnese (city)	1,182	309	873	217	261	249
TV - S.Lazzaro (city)	441	121	320	64	274	200
VE - S.Stino	1,444	263	1,181	158	182	134
VE - S.Michele	543	140	403	94	258	233
VE - Scorze'	915	193	722	131	211	181
VE - Carpenedo	1,582	312	1,270	298	197	235
VE - S.Pietro (city)	5,984	1,316	4,668	1,087	220	233
VE - S.Marco (city)	2,024	421	1,603	289	208	180
VE - Camponogara	988	291	697	133	295	191
VE - Chioggia	1,639	235	1,404	243	143	173
PD - Onara	856	284	572	120	332	210
PD - S.Giorgio	1,133	302	831	244	267	294
PD - Eremitani (city)	1,462	441	1,021	314	302	308
PD - S.Sofia (city)	2,012	462	1,550	267	230	172
PD - Chiesanuova	1,182	296	886	218	250	246
PD - Casalserugo	851	288	563	112	338	199
PD - Valnog./Faedo	661	265	396	76	401	192
PD - Pontelongo	1,412	465	947	234	329	247
PD - Pernumia	1,901	794	1,107	151	418	136
PD - Urbana	804	317	487	70	394	144
PD - Agna	1,583	687	896	148	434	165
VI - Valstagna	2,208	461	1,747	265	209	152
VI - Valrovina	640	165	475	86	258	181
VI - Thiene	5,257	1,951	3,306	558	371	169
VI - Nove	1,639	542	1,097	220	331	201
VI - S.Vito	963	381	582	73	396	125
VI - Lanze' e Setteca'	605	229	376	81	379	215
VI - Quargnenta	551	154	397	60	279	151
VI - Arzignano	3,268	905	2,363	297	277	126
VI - S.Pietro (city)	5,985	1,317	4,668	1,087	220	233
VI - Bosco e Nanto	903	288	615	99	319	161

Table 4.1: (continued)

	Months 0-2		Months 3-35		Prob. of death	
	Births	**Dead**	**Exposed**	**Dead**	**q0-2**	**q3-35**
	a	***b***	***c=a-b***	***d***	***b/a***	***d/c***
VI - Sarego	1,024	349	675	123	341	182
VR - Vestenanova	766	160	606	115	209	190
VR - Caldiero	1,132	210	922	168	186	182
TOTAL	**62,622**	**17,134**	**45,488**	**8,682**	**274**	**191**

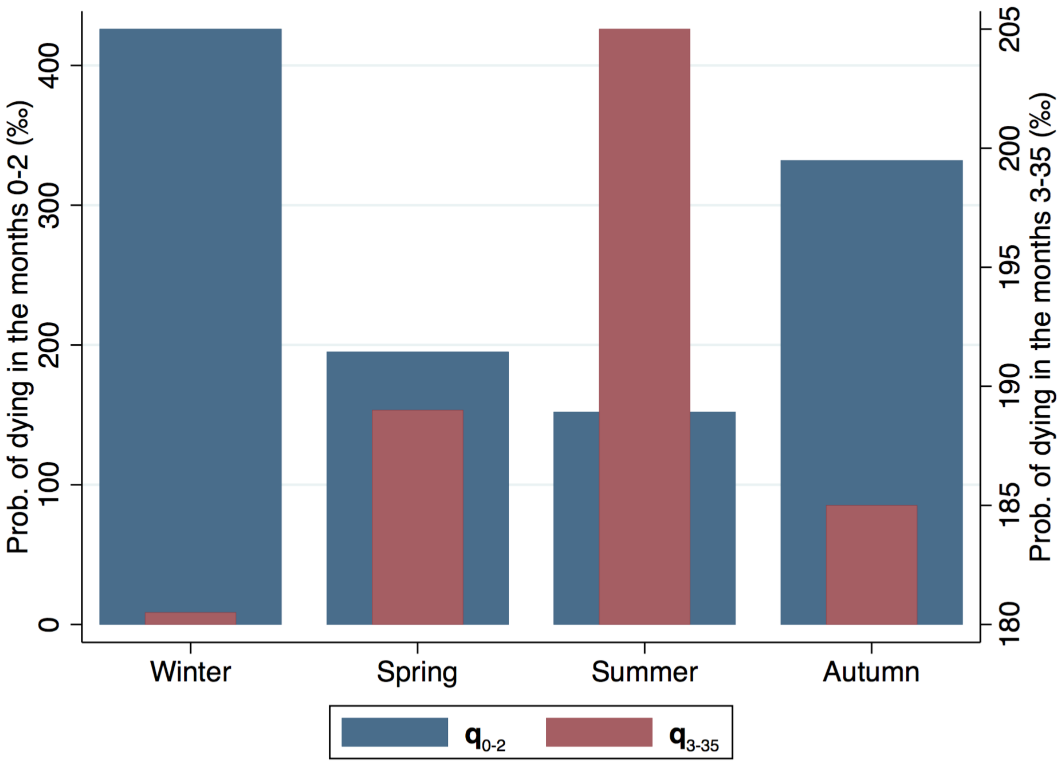


Figure 4.3: Early and late infant mortality (‰) by season of birth. Children born in 46 parishes in the region of Veneto in the period 1816-35.

While informative, the use of seasons is complicated by a scarcity of precise information on several aspects described in the literature as determinant in the definition of neonatal mortality in Veneto. A second option is thus the use of a more specific variable that combines season of birth with year of birth and parish of birth. These last two elements have in fact previously been shown to be relevant in the definition of mortality risk in Veneto (Dalla-Zuanna and Rosina 2011). We therefore decided to assign to each individual who survived the first three months the early infant mortality of his/her specific cohort. This designation was carried out in two different ways.

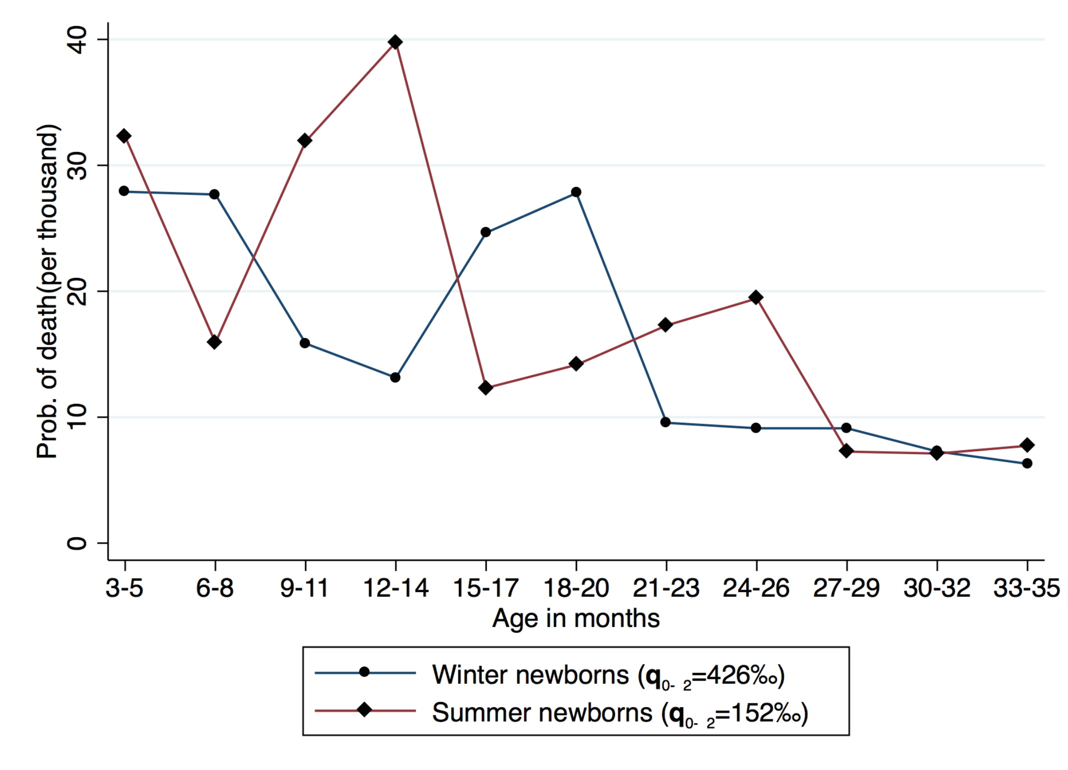


FIGURE 4.4: Risk of death within successive 3-month age groups for those aged 3-35 months children who survived the first three months of life, born during winter and summer in 46 parishes in the region Veneto in the period 1816-35.

First, we calculated 688 different levels of early infant mortality, provided by the combination of 4 seasons at birth, 4 groups of five-year intervals at birth and 43 parishes of birth. Rates of mortality in the first three months of life range from 822‰ (cohort of children born during the winter in the years 1816-1820 in the parish of Bigolino) to 54‰ (cohort of children born during the summer in the years 1831-1835 in the parish of Piavon). The scattergram between early and late infant mortality for these 688 cohorts is shown in Fig. 4.5. Finally, the 688 cohorts are pooled according to four increasing levels of early infant mortality, and Fig. 4.6 shows a clear bivariate inverse association between early and late infant mortality.

Using five-year intervals, parish and season may not be a precise enough measure to identify the specific environment a child experienced during his/her first three months of life. Early infant mortality could, for any number of reasons, vary from year to year or month to month, especially due to climate micro-oscillation. However, the number of children born in each year/season/parish is not large enough to calculate a stable measure of early infant mortality for smaller cohorts. To overcome this problem we estimated the 10,320 early infant mortality probabilities (12 months at birth, 20 years at birth, 43 parishes - or groups of parishes - of birth) using the following logit regression model:

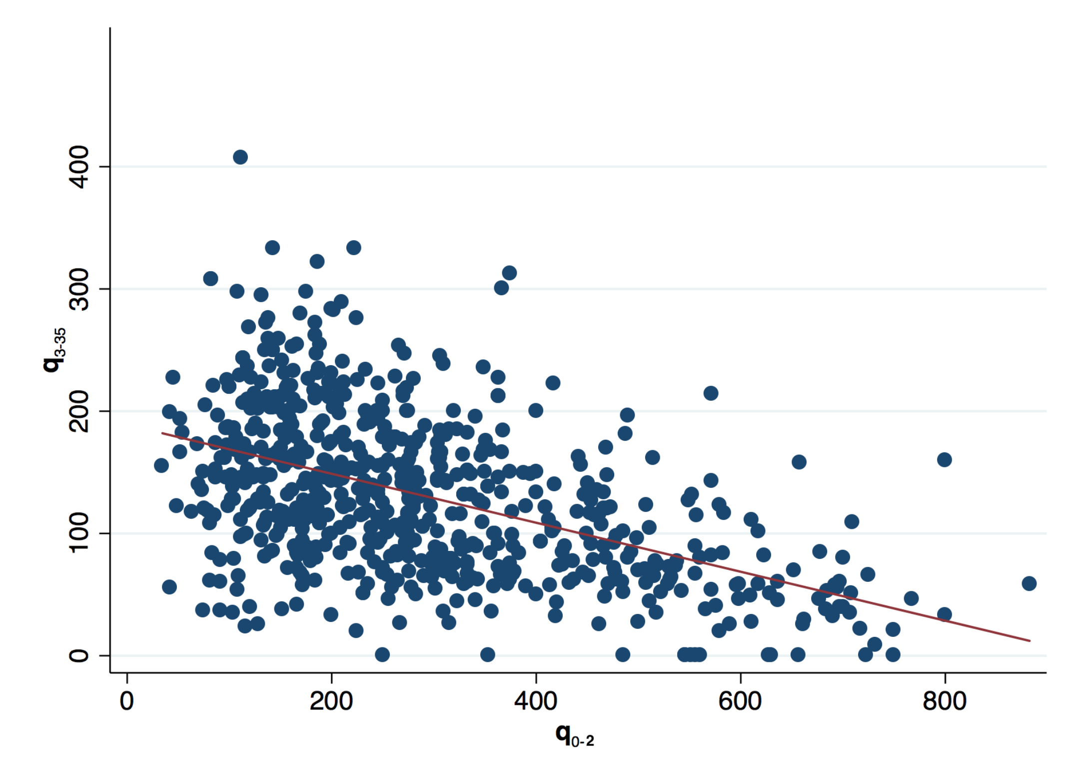


FIGURE 4.5: Early and late infant mortality for 688 cohorts provided by the combination of season, five-years period, and parish of birth. Children born in 46 parishes of the region of Veneto in the period 1816-35.

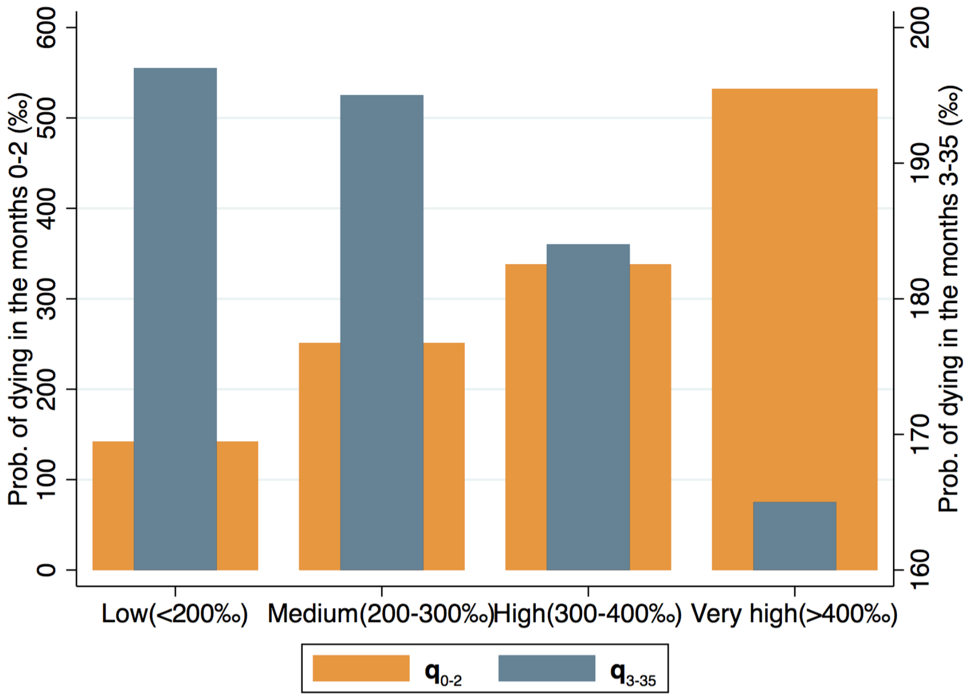


FIGURE 4.6: Early and late infant mortality (‰). 688 cohorts (43 communities x 4 time periods x 4 seasons) pooled according to the level of early infant mortality. Children born in 46 parishes of the region of Veneto during 1816-35.

$$q_{0-2} = \frac{\exp(\beta_0 + \beta_1 \dot{f}eb + \cdots + \beta_{11} dec + \beta_{12} 1817 + \cdots + \beta_{30} 1835 + \beta_{31} Col + \cdots + \beta_{72} Val)}{1 + \exp(\beta_0 + \beta_1 feb + \cdots + \beta_{11} dec + \beta_{12} 1817 + \cdots + \beta_{30} 1835 + \beta_{31} Col + \cdots + \beta_{72} Val)}$$

where *feb* and *dec* are months, *1817* and *1835* are years, *Col* and *Val* are the two parishes of Colle Umberto and Valrovina; the month of January, the year 1816, and the parish of Agna are the baseline.

The results range from 739‰ (cohort of children born in January of 1817 in the parish of Agna) to 42‰ (cohort of children born in July of 1822 in the parish of Piavon) and potentially provide a more detailed and complete picture compared to using only a 'rough' measure of early infant mortality for a five-year interval. We assigned the 'key covariates' to children who survived the first three months of life, together with several other variables from our dataset that may explain, at least partially, late infant mortality differences. In accordance with the literature - and given the preliminary results in Tab. 4.1 - we do not expect to observe large differences between the sexes in terms of late infant mortality. We similarly do not expect to see considerable differences in late infant mortality in the five-year intervals from 1816-1820 to 1831-1835, as the downward trend in late infant mortality had yet to begin (see Fig. 4.2). We expect, however, late infant mortality to vary according to parish of birth. Moreover, we expect late infant mortality to be higher during the summer, as suggested in Fig. 4.3.

Our aim is to verify whether children who survived a context of high early infant mortality had lower or higher mortality in the following 33 months, as suggested by Tab. 4.2. As children who died during the months 3-35 are censored, one way to assess this question is a semi-parametric Cox regression, which provides the monthly risk of death as a response variable, controlling for the elimination process (see e.g., Bruckner, Catalano, and Smith 2013; Ekamper et al. 2013). We model the risk of death between the 4th and the 36th months of life for individuals still surviving at the end of their 3rd month, using the above-described covariates as explanatory variables. Our explanatory variables do not vary over time, with the exception of month of life, which has instead been taken into consideration by transforming our database into a person-period format [6].

[6] Each individual has now a number of rows which correspond to the number of months lived. The last row corresponds to the month in which the subject died or is censored. We have associated each row with the relative season. For example, an individual born in May (season: spring) of 1816 who lived for 6 months has six rows in this new database format. The second row, in particular, indicates that the individual was alive in June (season: summer) of the same year. Notably, the last row indicates that the subject died in October (season: autumn) of 1816. In this way, it is possible to account for a possible effect of the transition from one season to another on the risk of death between the 4th and the 36th months of life.

TABLE 4.2: Probability of dying in months 0-2 and in months 3-35 according to early infant mortality (calculated and estimated).

	Months 0-2		**Months 3-35**		**Prob. of death**	
	Births	**Dead**	**Exposed**	**Dead**	**q0-2**	**q3-35**
	a	***b***	***c=a-b***	***d***	***b/a***	***d/c***
Early infant mortality (calculated)						
Low (≤ 200‰)	25,845	3,668	22,177	4,367	142	197
Medium (200-300‰)	14,938	3,754	11,184	2,186	251	195
High (300-400‰)	9,842	3,330	6,512	1,200	338	184
Very high (≥ 400‰)	11,997	6,382	5,615	929	532	165
Early infant mortality (estimated)						
Low (≤ 200‰)	25,107	3,574	21,533	4,360	142	202
Medium (200-300‰)	12,435	2,760	9,675	1,769	222	183
High (300-400‰)	12,496	4,088	8,408	1,602	327	191
Very high (≥ 400‰)	12,584	6,712	5,872	951	533	162
TOTAL	62,622	17,134	45,488	8,682	274	191

4.5 Models and Results

The covariates included in the first five models are simply indices of early infant mortality (see Table 4.3): season at birth (Model 1), calculated early infant mortality (Model 2), early infant mortality estimated by logit regression model (Model 3), season at birth and calculated early infant mortality (Model 4), and season at birth and estimated early infant mortality (Model 5). In Models 4 and 5 - when early infant mortality is included - the association between season at birth and late infant mortality disappears almost completely. In addition, as season at birth and early infant mortality are strongly interrelated, the latter seems to prevail in these two models. The goodness of fit for Models 2 and 3 is better than the goodness of fit for Model 1 (where season of birth is the only covariate). In light of these results, in the following models the season at birth variable is not included and we perform two parallel analyses using calculated early infant mortality and estimated early infant mortality as separate measures of selection during the first three months of life.

In order to determine whether selection during months 0-2 provides additional explanation of mortality at ages 3-35 months, we measure whether the statistical performance of a basic model (including only parish, time period, and sex) improves when early infant mortality is included as an explanatory variable.

In the basic model the risk of death between the 4th and the 36th months of life is strongly related to parish, whereas sex and period do not influence the risk (Tab. 4.4, Model 6). When early infant mortality is added the performance of the model improves: including the calculated early infant mortality (Model 7) the goodness of fit (log-likelihood) rises by 16 (Chi2=32, d.o.f.=3, p<0.001); including the estimated early infant mortality (Model 8) the goodness of fit (log-likelihood) rises by 21 (Chi2=42, d.o.f.=3, p<0.001). Summing up, the model improves when estimated early infant mortality is used. After including early infant mortality in the model the rank of parishes does not change, showing that interaction between the two variables is weak.

Model 9 introduces the transition between seasons as a time-varying covariate: when using winter as the baseline, the risk of dying during months 3-35 of life is significantly higher when there is a transition to a warmer season (4.4% higher in the transition to summer and 1.2% in the transition to spring, compared to 1.0% in the transition to autumn), confirming the results of Fig. 4.3, 4.4, and 4.6. This result is likely the outcome of the greater impact of gastro-enteric diseases during summer and autumn [7]. Model 10 and Model 11 reproduce the same analysis of Model 9 with respect to the subsequent cohorts: in particular, Model 10 refers to the children born in the period 1836-55, while Model 11 refers to the children born in the period 1856-66. Apart from some little differences in the magnitude of the coefficients, the results are comparable to those of Model 9.

[7] The cause of death is not always reported and is often confused with symptoms (see also footnote 5). However, by carefully examining death records at age 3-35 months, a cause of death likely identifiable as gastrointestinal disease is reported for about 20% of the deaths, with tremendous seasonal differences (87% of the deaths related to gastrointestinal diseases are referred to summer and autumn newborns). See also Chapter 5 for more details.

Table 4.3: Risk of dying (odds ratio) between the 4th and the 36th months of life by season at birth and early infant mortality, calculated or estimated. Cox model for children born in 46 Veneto parishes in the period 1816-35.

	MODEL 1		MODEL 2		MODEL 3		MODEL 4		MODEL 5	
	HR	**p-val**	**HR**	**p-val**	**HR**	**p-val**	**HR**	**p-val**	**HR**	**p-val**
Season at birth										
Winter	1	-					1	-	1	-
Spring	1.067	0.04					1.006	0.89	0.982	0.64
Summer	1.166	0					1.101	0.03	1.055	0.19
Autumn	1.048	0.18					1.019	0.61	1.008	0.83
Early infant mortality (calculated)										
Low ($\leq$ 200‰)			1	-			1	-		
Medium (200-300‰)			0.994	0.81			1.025	0.42		
High (300-400‰)			0.932	0.03			0.973	0.51		
Very High ($\geq$ 400‰)			0.831	0			0.871	0		
Early infant mortality (estimated)										
Low ($\leq$ 200‰)					1	-			1	-
Medium (200-300‰)					0.885	0			0.894	0
High (300-400‰)					0.940	0.04			0.957	0.23
Very High ($\geq$ 400‰)					0.786	0			0.801	0
LL	-91634		-91632		-91620		-91626		-91616	
LR	26.99		30.57		54.75		41.76		61.58	
Prob >Chisq.	0		0		0		0		0	

TABLE 4.4: Covariates of the risk of dying (odds ratio) between the 4th and the 36th months of life for children who survived the third month of life. Two complete models with early infant mortality calculated or estimated (7 and 8); three models with a time-varying covariate: cohort of 1816-35 (9), cohort of 1836-55 (10), cohort of 1855-65 (11).

	MODEL 6		**MODEL 7**		**MODEL 8**		**MODEL 9**		**MODEL 10**		**MODEL 11**	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Early infant mort.												
(calculated)												
Low (≤ 200‰)			1	-								
Medium (200-300‰)			0.945	0.039								
High (300-400‰)			0.883	0								
Very High (≥ 400‰)			0.825	0								
Early infant mort.												
(estimated)												
Low (≤ 200‰)					1	-	1	-	1	-	1	-
Medium (200-300‰)					0.919	0.005	0.918	0.004	0.928	0.002	0.876	0
High (300-400‰)					0.925	0.009	0.913	0.003	0.955	0.198	0.850	0.005
Very High (≥ 400‰)					0.785	0	0.778	0	0.857	0.001	0.756	0.057
Sex												
Male	1	-	1	-	1	-	1	-	1	-	1	-
Female	0.991	0.684	0.992	0.698	0.991	0.676	0.988	0.576	0.965	0.078	0.993	0.798

Table 4.4: (continued)

	MODEL 6		**MODEL 7**		**MODEL 8**		**MODEL 9**		**MODEL 10**		**MODEL 11**	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Period												
1816-20	1	-	1	-	1	-	1	-				
1821-25	0.942	0.060	0.922	0.012	0.925	0.014	0.918	0.007				
1826-30	1.119	0	1.101	0.002	1.103	0.002	1.092	0.005				
1831-35	1.161	0	1.141	0	1.143	0	1.140	0				
1836-40									1	-		
1841-45									1.078	0.012		
1846-50									1.079	0.011		
1851-55									1.000	0.994		
1856-60											1	-
1861-66											0.952	0.081
Parish												
PD - Agna	1	-	1	-	1	-	1	-	1	-	1	-
TV - Colle Umberto	0.667	0.003	0.614	0	0.601	0	0.598	0	0.973	0.840	0.860	0.361
TV - Bigolino	1.183	0.191	1.136	0.322	1.123	0.367	1.128	0.348	1.070	0.619	1.135	0.449
TV - S.Martino	0.605	0.002	0.552	0	0.546	0	0.629	0.002	1.399	0.033	0.618	0.039
TV - Selva	0.832	0.157	0.774	0.050	0.788	0.068	0.779	0.056	1.114	0.388	1.102	0.530
TV - Asolo	1.036	0.885	1.003	0.989	1.004	0.986	0.992	0.973	1.071	0.595	1.111	0.502
TV - Ormelle	0.606	0.004	0.574	0.001	0.578	0.002	0.600	0.003	1.065	0.678	0.686	0.064

Table 4.4: (continued)

	MODEL 6		**MODEL 7**		**MODEL 8**		**MODEL 9**		**MODEL 10**		**MODEL 11**	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
TV - Piavon	0.770	0.066	0.686	0.009	0.686	0.009	0.690	0.010	1.011	0.935	1.042	0.798
TV - Monastier	1.063	0.563	0.979	0.840	0.968	0.761	0.970	0.774	1.126	0.281	1.176	0.223
TV - Vedelago	1.160	0.304	1.116	0.449	1.143	0.354	1.168	0.277	1.161	0.294	1.362	0.093
TV - S.Agnese (city)	1.567	0	1.475	0	1.465	0	1.467	0	1.332	0.009	1.003	0.986
TV - S.Lazzaro (city)	1.223	0.183	1.158	0.333	1.152	0.350	1.183	0.264	1.392	0.053	0.671	0.064
VE - S.Stino	0.793	0.043	0.723	0.005	0.717	0.004	0.716	0.004	1.004	0.971	1.679	0
VE - S.Michele	1.464	0.004	1.384	0.014	1.376	0.016	1.374	0.016	1.253	0.051	1.659	0
VE - Scorze'	1.113	0.375	1.029	0.815	1.006	0.959	1.009	0.939	0.956	0.734	1.336	0.052
VE - Carpenedo	1.478	0	1.348	0.003	1.338	0.004	1.344	0.004	1.433	0.001	1.382	0.010
VE - S.Pietro (city)	1.468	0	1.363	0	1.332	0.001	1.331	0.001	1.294	0.004	1.970	0
VE - S.Marco (city)	1.139	0.196	1.049	0.639	1.032	0.761	1.007	0.947	1.118	0.262	1.215	0.122
VE - Camponogara	1.179	0.168	1.114	0.369	1.112	0.376	1.114	0.369	1.189	0.155	1.130	0.447
VE - Chioggia	1.051	0.635	0.943	0.580	0.942	0.572	0.943	0.577	1.222	0.055	1.470	0.003
PD - Onara	1.327	0.021	1.252	0.068	1.273	0.050	1.266	0.055	1.193	0.183	0.845	0.329
PD - S.Giorgio	1.924	0	1.824	0	1.816	0	1.835	0	1.193	0.146	0.940	0.711
PD - Eremitani (city)	2.006	0	1.917	0	1.918	0	1.919	0	1.233	0.038	1.782	0
PD - S.Sofia (city)	1.047	0.653	0.976	0.814	0.953	0.644	0.951	0.631	0.965	0.746	0.902	0.473
PD - Chiesanuova	1.570	0	1.474	0	1.465	0	1.443	0.001	0.958	0.713	0.848	0.285
PD - Casalserugo	1.206	0.138	1.155	0.253	1.161	0.236	1.176	0.197	1.438	0.005	0.732	0.119

Table 4.4: (continued)

	MODEL 6		**MODEL 7**		**MODEL 8**		**MODEL 9**		**MODEL 10**		**MODEL 11**	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
PD - Valnog./Faedo	1.175	0.255	1.152	0.317	1.156	0.308	1.171	0.264	1.075	0.676	0.611	0.085
PD - Pontelongo	1.546	0	1.499	0	1.484	0	1.476	0	1.221	0.060	1.251	0.116
PD - Pernumia	0.805	0.063	0.789	0.042	0.800	0.056	0.815	0.077	1.118	0.369	0.767	0.081
PD - Urbana	0.857	0.289	0.830	0.201	0.845	0.249	0.856	0.285	1.105	0.486	0.563	0.006
VI - Valstagna	0.904	0.325	0.838	0.089	0.816	0.050	0.814	0.048	1.074	0.483	1.202	0.150
VI - Valrovina	1.108	0.449	1.042	0.764	1.031	0.822	1.024	0.861	1.120	0.377	1.287	0.128
VI - Thiene	1.031	0.738	0.989	0.905	1.001	0.990	0.998	0.979	0.584	0	1.168	0.192
VI - Nove	1.245	0.040	1.180	0.123	1.186	0.110	1.186	0.111	1.144	0.227	0.943	0.710
VI - S.Vito	0.735	0.033	0.709	0.017	0.723	0.025	0.741	0.036	1.258	0.095	0.868	0.437
VI - Lanze' e Setteca'	1.341	0.034	1.323	0.044	1.318	0.047	1.329	0.040	1.425	0.008	0.972	0.894
VI - Quargnenta	0.923	0.599	0.873	0.374	0.862	0.335	0.861	0.332	1.034	0.839	0.913	0.632
VI - Arzignano	0.759	0.006	0.713	0.001	0.712	0.001	0.703	0.001	1.251	0.024	0.899	0.407
VI - S.Pietro (city)	1.468	0	1.367	0	1.332	0.001	1.331	0.001	1.294	0.004	1.970	0
VI - Bosco e Nanto	0.997	0.979	0.965	0.784	0.953	0.709	0.930	0.577	1.239	0.103	1.214	0.247
VI - Sarego	1.118	0.362	1.112	0.384	1.090	0.480	1.092	0.473	1.057	0.663	1.265	0.139
VR - Vestenanova	1.146	0.273	1.046	0.720	1.043	0.739	1.050	0.699	1.133	0.346	1.438	0.028
VR - Caldiero	1.117	0.327	1.015	0.897	1.008	0.947	1.013	0.910	0.830	0.131	0.610	0.003

Table 4.4: (continued)

	MODEL 6		**MODEL 7**		**MODEL 8**		**MODEL 9**		**MODEL 10**		**MODEL 11**	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Season transition												
(time varying)												
Winter							1	-	1	-	1	-
Spring							1.012	0	1.008	0	0.994	0.009
Summer							1.044	0	1.030	0	1.012	0
Autumn							1.010	0	1.002	0.191	0.992	0
LL	-91311		-91295		-91290		-91530		-85160		-56660	
LR	672.3		704.45		713.69		1466.6		994.57		832.97	
prob >Chisq.	0		0		0		0		0		0	

4.6 Summing up and giving new directions

The key result of this chapter is that mortality selection during months 0-2 may have played a significant role in determining differences in mortality during months 3-35 among cohorts born in 46 Veneto parishes in the period 1816-1835, i.e., before the start of the mortality decline which formed part of the demographic transition.

For children that survived a severe early infant mortality selection ($q_{0-2} \geq 400‰$, born mainly during winter), the hazard ratio of surviving the following 33 months was 20%-30% higher than that of children with relatively small early mortality selection ($q_{0-2} \leq 400‰$, born mainly during summer). This result suggests a sort of homeostatic mechanism, as mortality differences among the cohorts are lower for q_{0-35} than for both q_{0-2} and q_{3-35}: in considering the 688 cohorts obtained by crossing 4 seasons at birth, 4 five-year groupings at birth, and 43 parishes of birth, we observe that the variation coefficient (σ/μ) is 0.53 for q_{0-2}, 0.54 for q_{3-35}, but only 0.31 for q_{0-35}.

These results are limited to the context of our study (Veneto in the early decades of the 19th century), characterized by very strong mortality during the first week/month of life (excluding the first day), mainly for children born during the cold season. As we mentioned at the beginning of this chapter, while a scar-effect could coexist with positive selection, one predominates in different contexts and/or at different stages of life (even within the same cohort, as shown by Quaranta (2013) for Sweden). More generally, our results indicate the possibility that differences between cohorts in mortality during a given interval of age could also be due to differences in selection for death, for the same cohorts, during the age interval which immediately preceded.

Albeit this chapter provides several novel elements for the interpretation of differences in late infant mortality before the demographic transition, many questions need more in-depth analysis, which we will address through our database in the next two chapters. First, given that the predominating causes of death during the first few days of life were clearly quite different from those more common in the weeks that followed (Bourgeois-Pichat 1951), and using the fact that our database contains the exact day of birth and exact day of death, it is needed to consider narrower time spans, e.g. to measure whether the intensity of selection during the first week had an effect on the risk of dying during the subsequent periods. Also, the selection procedure must be tested across different social strata, using the job of the father as recorded on the birth and death records. Moreover, the multilevel structure of our dataset may be accounted for by suitable statistical models (see Chapter 5).

Finally, by taking advantage of the richer information about temperatures and families available for the province of Padua, it is possible to refine the models presented in this chapter (see Chapter 6).

Chapter 5

Multilevel Analysis

The data at our disposal are naturally organized in a hierarchic manner (see Fig. 5.1), and they are ideal to be analyzed via a multilevel approach. The general structure is quite simple: we can identify two level, one being the individual level (152,210 individuals in our sample) and the other one being the community level (46 parishes of different location and size).

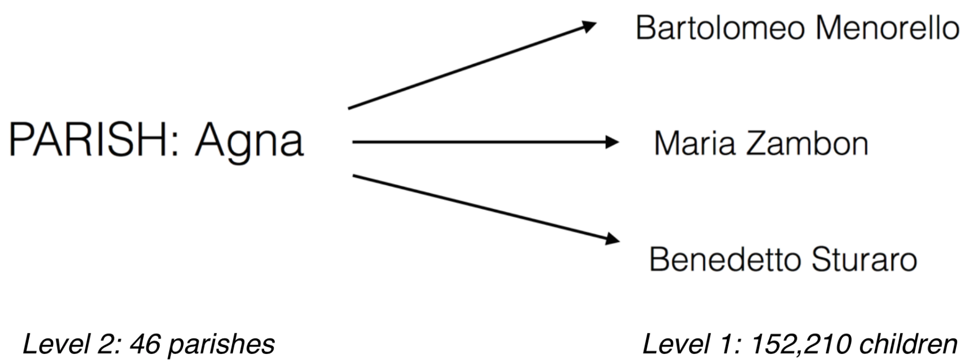


FIGURE 5.1: Hierarchical structure of the data for the whole sample.

The aim of this chapter is to verify the most important result of Chapter 4 (i.e. the existence of a mortality selection process) and to add some other details to the general dissertation, taking advantage of the strong multilevel structure of the data at our disposal.

In particular, in the first part of the chapter we will come back to the selection topic using a different temporal approach, while in the second part we will also introduce some territorial variables, i.e. some characteristic of the second-level units in our data's hierarchy.

5.1 Splitting the survival time and comparing fixed and random effects of parishes

As highlighted in paragraph 4.6, given the extremely different nature of the death causes operating at different ages, an 'exposure-to-selection' time of 3 months and a 'follow-up' time of 33 months are pretty long times, each of the two hiding some age-specific dynamics.

This is the reason why, using a slightly different approach, in models 12-15 (see Tab. 5.1), we split the children's first five years of life in four smaller periods (first week, first month excluding the first week, first year excluding the first month and years 1-4, as extensively done in the descriptive analyses of Chapter 3) and we focus on each of them. Also, we start using as explanatory variable not only an equivalent of the 'context variable' (i.e. the estimated early infant mortality) introduced in the previous chapter, but also the other, crucial individual variable at our disposal: the social class, derived from the father's job, as described in Chapter 3.

With respect to the estimated infant mortality in the birth context, we recall that in the models up to Model 11 it was referred to the first three months of life, as it was the probability of dying of the average child born in a particular combination of parish, year and month. Now - instead - it is referred to the only first week of life, i.e. we are trying to understand if the mortality selection process formerly discovered still takes place in a much narrower span of time. It is one of the other 'interesting steps' suggested in the Chapter 4's conclusions.

This means that, in Models 13-15, the early infant mortality is estimated according to the first week of life. In Model 12, on the other hand, we can not use the same context variable, since the model itself studies what happens in the first week of life: this is the reason why, in Model 12, the estimated early infant mortality is referred to the week prior to the birth, i.e. it is a proxy for the context that the children experimented during their last week in the womb. As a direct consequence, Models 13-15 are fully comparable among them, while Model 12 serves like reference.

Finally, in Models 12-15 we consider a parish fixed effect, like in all the models of Chapter 4. Since we are not directly interested in the specific effect of each parish, we have not reported the fixed effect coefficients in tab. 5.1. More interesting is the comparison between these models with fixed effects and similar model with random effects: in Tab. 5.2 we have run the equivalent of Cox random effect regressions, i.e. shared-frailty models. The underlying logic of these models is that some clusters of individuals (parishes) are intrinsically more or less prone to experiencing the event of interest than are others (Box-Steffensmeier et al.1999). With this purpose, in models 16-19 we have introduced a gamma-distributed variable that allows us to account for unobserved heterogeneity at the parish level: in fact, we already know that the parish of birth plays a determinant role in the survival process (i.e, there is a latent common group effect), even if until now we have omitted group covariates in our models.

The comparison of fixed effect and shared frailty models is common in epidemiology studies (e.g. Robertsson et al. 2006), in order to highlight pros and cons of the two approaches. Looking at our models, it seems that the coefficients of the four fixed effects models and the four random effects models are very similar, indicating that in our case it is important to take into account the parish level (since the likelihood of the multilevel models if always higher than the one of one-level models), but the fixed or random effects approaches do not make almost any difference. This means that the facts emerging from the 8 models are substantially the same, and we can interpret them in a pretty univocal way.

First, the mortality selection hypothesis seems to be confirmed. In fact, while the context experimented during the first week seems - unsurprisingly - a huge risk factor of dying in the first month of life (Model 17: for children survived to a context of very high first-week mortality, the risk is 8.6 times higher than children survived to a context of low first-week mortality), it becomes non-significant in the first year (Model 18) and turns to be a protective factor after the first birthday (Model 19: for children survived to a context of very high early infant mortality, the risk is 20% lower than children survived to a context of low early infant mortality). Compared to the coarser results of Chapter 4, this means that the selection produced by a high early infant mortality - or 'bad context' - seems to show its effect specifically after the first birthday, while in Chapter 4 they emerged more generally after the third month of life.

With respect to the other variable of interest, i.e. the social class, we have two bigger evidences: first, during the first week the risk of dying for children of peasants is not distinguishable from the risk for children of craftsmen and traders, and only the upper class babies have some advantage. Secondly, while the sign of the coefficients for the children of craftsmen and traders change over time (after the first birthday the risk for them is even higher than the one of the children of peasants), the children of landowners, professionals and employees always live in the best condition, unsurprisingly.

Focussing on the period during which selection showed its strongest effect, i.e. after the first birthday, we can think about verifying whether it was only the context of the first week of life to produce this kind of impact or, on the other hand, also the subsequent periods played some role. In particular, with the aim of expanding Model 19, we can

TABLE 5.1: Cox regressions modeling the risk of dying in four time spans, with parishes fixed effects. Odds ratio reported.

	MODEL 12		**MODEL 13**		**MODEL 14**		**MODEL 15**	
	First Week		***First Month***		***First Year***		***Years 1-4***	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Early infant mortality (estimated)								
Low (<200)	1	-	1	-	1	-	1	-
Medium (200-300)	1.727	0	1.875	0	1.007	0.801	0.953	0.045
High (300-400)	2.835	0	3.537	0	1.037	0.19	0.922	0.001
Very High (>400)	8.588	0	7.759	0	0.986	0.676	0.795	0
Sex								
Male	1	-	1	-	1	-	1	-
Female	0.95	0.005	0.943	0.002	0.948	0.004	1.028	0.085
Period								
1816-20	1	-	1	-	1	-	1	-
1821-25	0.965	0.347	1.017	0.675	0.842	0	0.987	0.748
1826-30	0.99	0.781	1.01	0.807	0.926	0.065	1.209	0
1831-35	0.924	0.041	0.906	0.018	0.996	0.93	1.176	0
1836-40	0.926	0.043	0.861	0	0.776	0	1.077	0.068
1841-45	0.885	0.002	0.921	0.052	0.827	0	1.091	0.029
1846-50	0.879	0.001	1.011	0.795	0.844	0	1.175	0
1851-55	0.775	0	0.936	0.118	0.691	0	1.291	0
1856-60	0.703	0	0.813	0	0.695	0	1.09	0.031
1861-65	0.743	0	0.844	0	0.669	0	1.043	0.293
1866-71	0.68	0	1.031	0.67	0.765	0	0.811	0.002
Social Class								
Peasants	1	-	1	-	1	-	1	-
Craftsmen	1.025	0.358	0.941	0.029	0.933	0.013	1.127	0
Traders	1.062	0.151	0.871	0.001	0.97	0.426	1.093	0.008
Land./Prof./Emp.	0.878	0	0.849	0	0.802	0	0.805	0

TABLE 5.2: Cox regressions modeling the risk of dying in four time spans, with shared frailty at the parish level. Odds ratio reported.

	MODEL 16		MODEL 17		MODEL 18		MODEL 19	
	First Week		***First Month***		***First Year***		***Years 1-4***	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Early infant mort. (estimated)								
Low (<200)	1	-	1	-	1	-	1	-
Medium (200-300)	1.723	0	1.864	0	1.007	0.805	0.954	0.047
High (300-400)	2.831	0	3.521	0	1.036	0.201	0.921	0.001
Very High (>400)	8.567	0	7.696	0	0.985	0.646	0.795	0
Sex								
Male	1	-	1	-	1	-	1	-
Female	0.95	0.005	0.943	0.002	0.948	0.004	1.028	0.086
Period								
1816-20	1	-	1	-	1	-	1	-
1821-25	0.966	0.354	1.017	0.674	0.843	0	0.987	0.749
1826-30	0.99	0.795	1.01	0.802	0.926	0.065	1.209	0
1831-35	0.925	0.044	0.906	0.018	0.997	0.936	1.176	0
1836-40	0.927	0.045	0.861	0	0.776	0	1.076	0.069
1841-45	0.885	0.002	0.921	0.051	0.827	0	1.091	0.03
1846-50	0.879	0.001	1.011	0.797	0.844	0	1.175	0
1851-55	0.775	0	0.936	0.117	0.691	0	1.291	0
1856-60	0.703	0	0.813	0	0.695	0	1.09	0.031
1861-65	0.743	0	0.844	0	0.669	0	1.042	0.296
1866-71	0.68	0	1.03	0.684	0.765	0	0.811	0.002
Social Class								
Peasants								
Craftsmen	1.022	0.407	0.942	0.03	0.935	0.016	1.13	0
Traders	1.056	0.193	0.872	0.001	0.972	0.465	1.097	0.005
Land./Prof./Emp.	0.878	0	0.85	0	0.805	0	0.807	0
Theta	0.096 (0.022)		0.123 (0.027)		0.124 (0.028)		0.187 (0.043)	

Table 5.3: Cox regression modeling the risk of dying in years 1-4, with shared frailty at the parish level and two different 'context' covariates. Odds ratio reported.

	MODEL 19b	
	Years 1-4	
	HR	p-val
Early infant mort. (first week)		
Low ($\leq$ 200‰)	1	-
Medium (200-300‰)	0.954	0.049
High (300-400‰)	0.922	0.001
Very High ($\geq$ 400‰)	0.791	0
Early infant mort. (months 1-11)		
Low ($\leq$ 200‰)	1	-
Medium (200-300‰)	0.952	0.059
High (300-400‰)	0.934	0.019
Very High ($\geq$ 400‰)	0.819	0

take into account two more contexts: the first month excluding the first week and the first year excluding the first month.

Given that first week and first month contexts are highly correlated (i.e. they are proxies of the same unobserved variable), in Model 19b we have kept only the first week context and we have added the first year context. In Tab. 5.3 we have reported only the coefficients of our key variables, while we have skipped all the controls (they do not vary from the ones in Model 19).

As result, the new added variable proves to be significant, showing a pattern fully comparable to the one of the first week context. Since these two contexts are not correlated, they must say something different: while the first week context captures the effect of a selection that is strictly linked to weather conditions and birth-related causes, the first year context may be a more direct proof of the existence of a 'late' selection, not limited to the temperature's action.

5.2 Describing the community variables

Having described the individual level variables in Chapter 3, we introduce here the community level variables - i.e. the ones that are shared by all the children belonging to the same parish - explaining the meaning and the nature of each variable.

5.2.1 Rurality

According to the OECD definition, a town (a parish, in our case) in which less than 15% of the population live in rural units is considered urban; a town in which a proportion of the population between 15 and 50% live in rural units is considered medially urban; a town in which more than 50% of the population live in rural units is considered rural.

Using data from the Italian Population Census of 1871 (the second after the unification, and the first available at the end of our temporal window), we have calculated the rurality rate for each parish. According to Fig. 5.2, the urban parishes coincide essentially with the cities or the biggest towns in our sample, while the rural communities are distributed in the hills and in the countryside along the entire region. In particular, the Southern area of the Province of Padua shows a high concentration of rural communities. Even if the urban parishes are numerically less numerous (9 out of 46) than the rural ones (23 out of 46), they represent the 35% of the individual records at our disposal, since they are more populated areas.

The distribution of rural communities is clearly related to the distribution of peasants families (cfr. Fig. 3.3), since the most rural parishes and the parishes with the highest proportion of peasants family are overlapped. Unsurprinsingly, the patterns of mortality are not really different, as well. According to Tab. 5.3, the risk of dying is lower in the urban communities if we look at the short period (first week and first month of life), while the contrary happens when we look at the subsequent periods (months 1-11 and years 1-4).

5.2.2 Altimetric zone

The Italian region of Veneto presents enormous geographical differences, as in a relatively small distance we pass from the Venice lagoon and the Po Valley to the hills (around Padua and Vicenza) and the big mountains (Dolomites). Anyway, the parishes in our sample are not representative of all of this altimetric variety, as we essentially have data coming from parishes located in the Valley. According to the definition of altimetric zone given by the Italian Institute of Statistics (ISTAT), 31 out of 46 parishes at our disposal (corresponding to 75% of the total population in our sample) belong to a plain zone; 13 parishes (namely Colle Umberto, Asolo, Bigolino, Selva del Montello, Faedo, Valnogaredo, Arzignano, Nanto, Bosco di Nanto, S. Pietro Vicentino, Quargnenta, Valrovina and S.Vito di Leguzzano), corresponding to one fifth of the sampled population, belong to hilly areas; only 2 parishes (Valstagna and Vestenanova) are considered part of a mountainous area. In our analyses, we will then consider the differences emerging between two big groups: plain versus hilly parishes.

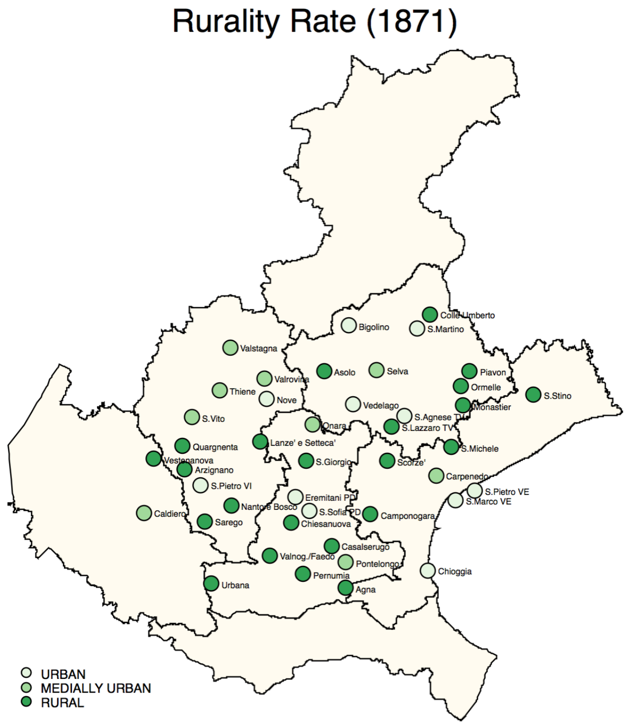


FIGURE 5.2: Distribution of urban and rural communities in 46 parishes of Veneto, 1815-70.

Looking at Tab. 5.3, the mortality in the first week of life is 10‰ higher in the hilly areas than in the plain. Starting from the subsequent period (first month excluding the first week) and also in the other two spans of time (first year excluding the first month and years from second to fifth) the situation is inverted in favor of the hilly parishes, where the risk of dying turns to be lower.

5.2.3 Illiteracy

A society with a large majority of people struggling for life and working hard in the fields for most of their lives is not supposed to be particularly educated. Our registers, anyway, do not provide information about education. In order to investigate the possible effect of education on early mortality, we must use data from external sources. Again, here we refer to an Italian Population Census (the one of 1881, in this case). In particular, we use the proportion of illiterate men born in 1812-16 and interviewed when they were 65-69 years old. As expected, the level of illiteracy was extremely high in our parishes: the average level is around 50%, ranging from 31.9% in the parishes located in the city of Venice to over 70% in the parish of Agna, in the Padua countryside (see Fig. 5.3).

Is the level of illiteracy associated to infant mortality? We can easily imagine that more educated people possessed some more concepts - even if basic ones - regarding the childcare in the first days/weeks of life. In effect, according to Tab. 5.3 the risk of dying int the first week and in the first month of life is definitely lower in the communities with lower percentages of illiterate men (this being a proxy of the level of education in the parish). As usual, there is a turning point after the first month of life and, in particular, in the years from second to fifth.

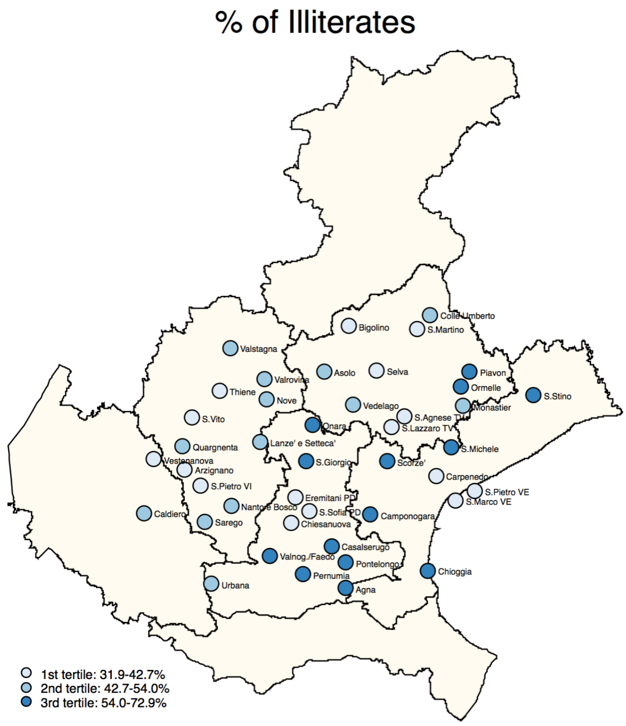


FIGURE 5.3: Proportion of illiterate men in each of the 46 parishes of our database. Cohort of 1812-16.

5.2.4 Religious affiliation

As already stated in Chapter 2, in the 19th century Veneto region the role played by priests in the communities was a preeminent one. Not only they had become civil servants, collecting data on births and deaths, but they still were very respected political and spiritual guides. In this context, the importance of religion (or, better, of the affiliation to religion) is not a matter of doubt, since it also influenced practical choices taking place in the first period of an individual's life (i.e. baptism) [1].

Using a document based on the pastoral visits of the bishops at the end of the 19th century in our parishes (Gambasin 1975), we have been able to estimate the proportion of people that regularly attended the religious functions. This is not, obviously, a variable registered and coded according to a proper definition of religious affiliation, since we are

[1] 1In this thesis, when we mention religion we refer to catholic religion, that is reported in the registers as the belief of 99% of the subjects under analysis.

Table 5.4: Probability of dying (‰) according to four time spans and four community variables (rurality, altimetric zone, illiteracy and religious affiliation)

	Days 0-6	***Days 7-28***	***Months 1-11***	***Years 1-4***
Urban	103	88	149	269
Med Urban	141	118	134	223
Rural	139	118	129	186
Hill	135	101	119	188
Plain	125	110	143	235
Low Illiteracy	117	96	146	247
Med Illiteracy	136	108	124	216
High Illiteracy	130	122	140	207
Low affiliation	76	64	152	256
Med affiliation	158	145	145	232
High affiliation	132	107	127	206

actually using the perception of the bishops about the 'proportion of subject that were used to regularly approach the sacrament of Penance'.

Unsurprisingly, more than half of the subjects in our dataset lived in communities where the affiliation index lied between 95 and 100% (i.e. almost the totality of the population regularly attended the sacrament of Penance); more than three quarters of the subjects lived in communities where the affiliation index was greater than 75%; less than one fifth of the subjects lived in communities there the affiliation index was between 58 and 75% (more precisely, as reported in Fig. 5.4, they all lived in the area of Venice, that then qualifies as the most secularized province in the region).

From Tab. 5.3, it emerges that the risk of dying was higher in the communities with a stronger religious affiliation. In particular, in the first week and in the first month of life the mortality was doubled in the parishes with an affiliation rate greater than 75% with respect to those with a lower rate. This result is coherent with the findings by Dalla-Zuanna et al. (2014) about the relationship between early baptism and early mortality. As usual, the situation is inverted if we look at the longer period, with the most secularized communities experiencing a higher mortality in the first year (excluding the first month) and in years 1-4.

Anyway, being all the 'low affiliation' parishes located in the same area (Venice), it is possible that the effect of the religious affiliation rate on mortality is spurious, hiding some other relations. This is one of the issues that we will try to clarify with the upcoming models.

FIGURE 5.4: Proportion of people regularly attending the religious functions in 46 parishes of Veneto, 1815-70.

5.3 Models at the community level

In this paragraph we switch from the individual approach that guides this research (until now we have only run regressions modeling the individual risk of dying) to a community approach, in order to say something more about the effect of community variables on infant mortality. To do so, in Models 20-23 (Tab. 5.5) the dependent variable is no more binomial (alive/dead), but it is a proportion indicating the mortality rate in each parish recorded in each of the four canonical time spans (the rates come from the life tables computed using our database and already mentioned in Chapter 3).

Having this kind of response variable, the effect of explanatory variables tends to be non-linear, so we rather model the relation between the mean proportion and the explanatory variables, using fractional logit models (Moeller 2013).

Before interpreting the results, three notes:

- there are 43 observations in each model, since the total number of parishes is 46 but 3 of them have been respectively merged to another one;

- as a proxy for the geographical proximity of the parishes, we use provincial dummies (the Province of Verona has been merged to the Province of Vicenza - the

Table 5.5: Fractional logit models modeling the mortality in four time spans according to five community variables.

	MODEL 20		**MODEL 21**		**MODEL 22**		**MODEL 23**	
	First Week		***First Month***		***First Year***		***Years 1-4***	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Province								
PD	1	-	1	-	1	-	1	-
TV	-0.295	0.055	-0.116	0.516	-0.078	0.442	-0.245	0.099
VE	-0.735	0	-0.647	0	0.15	0.119	0.066	0.554
VI+VR	-0.117	0.424	-0.215	0.21	0.087	0.402	-0.074	0.589
Rurality								
Urban	1	-	1	-	1	-	1	-
Med. Urban	-0.008	0.964	-0.028	0.867	-0.145	0.139	-0.244	0.013
Rural	0.036	0.818	-0.037	0.809	-0.19	0.021	-0.389	0
Illiteracy								
Low	1	-	1	-	1	-	1	-
Medium	0.16	0.224	0.12	0.393	-0.021	0.767	-0.009	0.906
High	0.005	0.97	0.126	0.363	0.118	0.15	-0.014	0.896
Rel. Affiliation								
<95%	1	-	1	-	1	-	1	-
>95%	-0.156	0.173	-0.188	0.128	0.001	0.986	0.031	0.749
Altitude								
Hill	1	-	1	-	1	-	1	-
Plain	-0.024	0.663	0.042	0.46	0.033	0.376	0.01	0.765
constant	-1.524	0	-2.137	0	-2.239	0	-1.547	0

closest one - since in our dataset it consisted of only 2 parishes);

- the 'religious affiliation' variable has been converted into a dummy having value 0 if the affiliation rate is lower than 95% and 1 if the rate is higher than 95%. This choice is justified by the fact that only 4 parishes out of 43 - albeit parishes with large populations - lied in the 'low affiliation' category that we can use in individual-level models.

First of all, there is a province effect: with respect to the reference group (Province of Padua), the parishes in the other three provinces were definitely a better place to live in during the first year of life: in particular, the province of Venice was the optimum. This protection vanishes (or even changes direction) in the subsequent periods.

Secondly, the rurality level of each parish plays a significant protective role after the first birthday, when living in a medially urban or rural community could be considered a real protective factor. This pattern has been also detected in countries like Belgium (Alter et al. 2005) and England (Williams et al. 1995), and - more generally - was very common across 19th century Europe.

Never significant is, on the contrary, the level of illiteracy of the community. Here the idea is that this kind of variable, that could be considered a proxy for the wealth and the social class of the family, does not work at the group level, being instead important at the individual level. In all the previous models, in fact, the father's job (that in some sense related to the illiteracy level) used to had a major impact on the risk of dying.

Also, in these models we do not find any religious affiliation effect that, as suspected, was essentially spurious and has likely been captured by the province dummies. Lastly, the altitude is not significant too.

5.4 Multilevel models with community variables

After the break of the previous paragraph, Models 24-27 (Tab. 5.6) use the same approach as Models 12-19. That is, we split again the first 5 years of life in four narrower time spans and for each of them we run a Cox regression with parish-shared gamma frailty in order to analyze the risk of dying in the period. The only difference is that now we consider in the models also the community variables described earlier.

We immedaitely note that evidences at the individual level (context, sex, period, social class) are essentially the same as the ones from Models 16-19. Thus, we focus here on interpreting the results regarding the community level, the only reported in Tab. 5.6:

- living in a rural community works as a protective factor, as already emerging from Models 20-23;

- living in a community with a high level of illiteracy (third tertile of the distribution of illiterated men born in the 1811-16 cohorts) is always a risk factor;

- living in a higher religious-affiliated community seems to be a risk factor in the first week of life, turning to be a protective condition afterwards;

- the altitude of the parish (plain or hill) presents alternate effects during time.

Table 5.6: Cox regressions modeling the risk of dying in four time spans, with shared frailty at the parish level and four community variables. Odds ratio reported.

	MODEL 24		**MODEL 25**		**MODEL 26**		**MODEL 27**	
	First Week		***First Month***		***First Year***		***Years 1-4***	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Rurality								
Urban	1	-	1	-	1	-	1	-
Med. Urban	0.932	0.633	0.872	0.402	0.939	0.707	0.862	0.459
Rural	0.977	0.87	0.855	0.318	0.905	0.536	0.731	0.106
Illiteracy								
Low	1	-	1	-	1	-	1	-
Medium	0.907	0.469	0.854	0.296	0.96	0.796	1.123	0.537
High	1.062	0.683	1.089	0.6	1.128	0.474	1.133	0.543
Rel. Affiliation								
Low	1	-	1	-	1	-	1	-
Medium	1.121	0.584	0.986	0.951	0.805	0.369	1.013	0.965
High	1.076	0.697	1.007	0.974	0.837	0.411	0.993	0.979
Altitude								
Hill	1	-	1	-	1	-	1	-
Plain	0.924	0.168	1.007	0.916	0.978	0.731	0.972	0.965
Theta	0.09		0.11		0.12		0.17	
	(-0.02)		(-0.02)		(-0.03)		(-0.04)	

Anyway - due to a matter of numerosity and to the fact that an individual variable like the social class already includes the characteristics described by community variables such as rurality and illiteracy - none of the coefficients related to the four community variables gains acceptable levels of statistical significance, leading to the conclusion that the inclusion of community variables does not add really much to the evidences coming from Models 12-19, the ones with just fixed or random parish effect.

Summing up, it is the combination of models at the community level (section 5.3) and multilevel models (section 5.4) to give us a more complete view about the patterns of infant mortality in 19th century Veneto and its interaction with territory. In particular, the community level allows us to introduce the importance of province, to obtain a confirm about the likely 'protecting' role of rurality rate, and to remove out all our conjectures about the possible effect of religious affiliation. On the other hand, from the multilevel models it emerges that there were two individual variables - i.e. context of birth and social class - strictly related to the risk of dying and with statistically significant coefficients; instead, we are not able to satisfactorily identify the effects of the community variables we control for.

Chapter 6

Further analyses: Daily Temperatures and Family Level

6.1 The Daily Temperatures

For a portion of our database corresponding to 9 parishes of the province of Padua (namely Chiesanuova, S. Giorgio delle Pertiche, Urbana, Eremitani, Onara, Pernumia, Pontelongo, S. Sofia, Valnog./Faedo) the information at our disposal is richer and gives us the chance for sharpening our analyses. For these 9 parishes, in particular, we can identify the familiar bonds between siblings and we can substitute the proxy variables we have used until now for the 'context of birth' with a more punctual and reliable one: the daily temperature.

6.1.1 The meteorological time series

In 2002, Camuffo published historical meteorological checked daily data for some European towns, derived from elaborations of early instrumental sources. Particular attention is given to the time series of Padua's temperatures: it is considered one of the most rich and interesting - due to the strong meteorological tradition of the University - and includes the registration of daily temperatures (minimum, maximum, and daily average) and atmospheric pressure, from 1725 to 1998.

In the following paragraphs, our idea is to assume that in the country parishes of the province (6 out of 9) the temperatures were very similar to those registered in the city center (3 out of 9). This assumption can be considered reliable since not only all the parishes were not too much geographically spread out in the province, but also their altitude is fully comparable. Also, we will use the daily mean temperature, in order to be able to reduce the impact of extremes and to evaluate both the effect of cold and hot temperatures [1].

Since - as showed along the previous chapters - the weather conditions were the most crucial variable in determining the risk of dying in early life, running statistical models that take into account this direct meteorological source can be very helpful.

6.1.2 Discrete time models

The availability of a daily weather measure introduces a possibility that we have not explored really much until now: an intensive use of a time-varying variable, in order to identify the effect of context changes over time. The perspective of building the core of this chapter around a time-varying structure lead us to the choice of a discrete time setting.

The first step to re-organize our data is to split each record in a number of rows equal to the number of days lived. Anyway, given the big amount of records and of days lived, the risk is that the total dimension of the database becomes too big to be properly analyzed. Also, it does not seem very useful to consider every single day lived - say - after the first birthday as a distinct observation.

The idea is that the records' splitting should be done differently according to the period under analysis: for example, it appears meaningful to split day by day during the first week of life, but after it is better to consider more separated splitting times (i.e. by weeks, by months, by trimesters). In particular, this means that our database is going to be divided in as many 'new' databases as the different splitting choices. Adding a fifth period (the sole first day of life) to the canonical four-periods of analysis (first week, first month excluding the first week, first year excluding the first month, and years 1-4), we can separately analyze these slices of early life:

- **First day**: this database includes all the children born in the 9 parishes of the Province of Padua in the period 1816-70. There is no splitting here, since there are no time-varying variables and each child has only one row (see Tab. 6.1): the interest here is in the effect of the temperature of the first day of life (*t1*) on the

[1] We know that in the very first part of life it is cold weather to be extremely dangerous for newborns while, after the first birthday, hot weather turns to be more lethal than cold. I.e. the role of minimum and maximum temperatures in the different periods of early life may require some more in-depth analyses. In this chapter we present the results of models in which we have used the daily average temperature, while we do not report the preliminary results of models with minimum and/or maximum temperatures. Anyway, the only remarkable evidence is that -unsurprisingly - the usage of minimum temperatures tends to only barely magnify the coefficients related to cold weather.

hazard of dying within the same day. The variable *dead* indicates if the subject died (1) or not (0).

Table 6.1: Four example records from first day database

id	day	t1	dead
11	3oct1827	16.5	0
189	7nov1829	7.4	0
433	13feb1844	1.6	1
456	9mar1856	11.2	0

- **Days 2-7**: it includes all the children survived to the first day of life. The splitting is done daily: this means that each child has a number of rows equal to the number of days lived. E.g.: a child survived until the end of the week (see Tab. 6.2) has six rows - one for each day - and the value of the variable *dead* is always 0. The daily temperature is time varying (*curr*), while the temperature of the first day of life is fixed in time (*t1*). Also the *age* of the child (in days) is - obviously - time varying.

Table 6.2: Six example records (same subject) from days 2-7 database

id	day	age	curr	t1	dead
75	7aug1819	2	20.3	20.5	0
75	8aug1819	3	23.5	20.5	0
75	9aug1819	4	22.9	20.5	0
75	10aug1819	5	21.1	20.5	0
75	11aug1819	6	23.0	20.5	0
75	12aug1819	7	22.7	20.5	0

- **Days 8-27**: it includes all the children survived to the first week of life. The splitting is done weekly: this means that each child has a number of rows equal to the number of weeks lived. E.g.: a child survived until the end of the month (see Tab. 6.3) has three rows - one for each week remaining in the month - and the value of the variable *dead* is always 0. The weekly average temperature is time varying (*curr*), while the temperature of the first day (*t1*) and the average temperature of the first week excluding the first day (*t2*) are fixed in time. The *age* is expressed in weeks.

Table 6.3: Three example records (same subject) from days 8-27 database

id	day	age	curr	t1	t2	dead
95	6nov1850	2	9.3	11.6	10.1	0
95	13nov1850	3	8.5	11.6	10.1	0
95	20nov1850	4	12.8	11.6	10.1	0

- **Months 1-11**: it includes all the children survived to the first month of life. The splitting is done monthly: this means that each child has a number of rows equal to the number of months lived. E.g.: a child survived until the fifth month (see Tab. 6.4) has four rows - one for each month after the first - and the value of the variable *dead* is 1 in the last row. The monthly average temperature is time varying (*curr*), while the temperature of the first day (*t1*), the average temperature of the first week excluding the first day (*t2*) and the average temperature of the first month excluding the first week (*t3*) are fixed in time. The *age* is expressed in months.

TABLE 6.4: Four example records (same subject) from months 1-11 database

id	**day**	**age**	**curr**	**t1**	**t2**	**t3**	**dead**
116	4mar1839	2	10.3	1.7	3.8	6.1	0
116	3apr1839	3	14.7	1.7	3.8	6.1	0
116	3may1839	4	16.9	1.7	3.8	6.1	0
116	2jun1839	5	19.1	1.7	3.8	6.1	1

- **Years 1-4**: it includes all the children survived to the first year of life. The splitting is done by trimesters: this means that each child has a number of rows equal to the number of trimesters lived. E.g.: a child survived until the end of the fifth year of life has sixteen rows - one for each trimester after the first birthday - and the value of the variable *dead* is always 0. The average temperature of the current trimester is time varying (*curr*), while the temperature of the first day (*t1*), the average temperature of the first week excluding the first day (*t2*), the average temperature of the first month excluding the first week (*t3*) and the average temperature of the first year excluding the first month (*t4*) are fixed in time. The *age* is expressed in trimesters (see Tab. 6.5).

TABLE 6.5: Four example records (same subject, died during the 8th trimester) from years 1-4 database

id	**day**	**age**	**curr**	**t1**	**t2**	**t3**	**t4**	**dead**
5	2oct1825	5	10.3	17.5	16.7	13.1	12.2	0
5	1jan1826	6	5.5	17.5	16.7	13.1	12.2	0
5	2apr1826	7	16.0	17.5	16.7	13.1	12.2	0
5	2jul1826	8	22.1	17.5	16.7	13.1	12.2	1

6.1.3 Models and Results

The previous paragraph has defined the starting point for five regression models (one for each of the new databases) whose aim is - as usual - to model the risk of dying in the corresponding span of time. The models are logit regressions in which the key covariates are the different temperatures (as defined before), and the controls are five individual variables (social class, sex, cohort, parish and age). Apart from the model running on the first day, the others contain two time-varying variables (current temperature and age of the child).

Tab. 6.6 shows the estimates regarding temperatures and age. Looking at the first day of life, we find that the the current temperature (the temperature of the day) has a big impact: the hazard is 1.7% lower for each extra Celsius degree. Nevertheless, if compared to the following models, this effect is smaller than - for example - what happens in the rest of the first week, indicating that the first day had its own dynamics and it is necessary to consider it separately: it was dramatically dangerous during wintertime, but still really dangerous during the rest of the year.

As just stated, the effect of the current temperature is bigger when we move forward in the first week: in days from 2nd to 7th, each extra Celsius degree corresponds to a 4.4% lower hazard. This effect has to be added to a similar one coming from the temperature of the first day of life, that - starting from this model - is considered as fixed in time: long story short, in the first week of life cold weather was particularly lethal.

In the first month (excluding the first week) the situation is not very different: we still have a strong effect of the current temperature, while the significant effect of the past temperature is now the one of the average temperature in days 2-7.

The picture suddenly changes in Model 31, the one referring to the first year excluding the first month. Here the current temperature has no more impact, while a higher temperature in day 1 turns to be a risk factor: the hazard is 3.3% higher for each extra Celsius degree experimented in day 1. Since we can not believe that the temperature of the first day of life continues to play a direct role one month - or even more - later, this is possibly the kind of compositional effect that we have called mortality selection. Nevertheless, this negative 'early life effect' is mitigated by the protective effect of the average temperature in days 7-27 (-4.4% per degree): as in Chapter 5, the period of life between the first and the twelfth month presents an unclear situation, in the sense that something is changing but two opposite effects seem to coexist.

Finally, in Model 32 we observe two crucial evidences. First, the current temperature is again significant, but in the other direction: after the first birthday, hot temperatures become more dangerous (for the bigger incidence of gastroenteric diseases), with each extra Celsius degree in the trimestral average temperature corresponding to a 3.4% higher risk of dying in the current trimester. Also, looking at the effects of past temperatures, there is the culmination of the 'effect inversion' that we have considered as a proof for the existence of selection along all the dissertation. In particular - like in model 31 - day 1 temperature is significant, with a +1.5% per degree hazard associated to children survived to a hotter first day of life. The other past temperatures seem not to be significant: there is not 'mitigation' effect (like in Model 31) nor a 'late selection' effect (like in Model 19b).

Looking at the duration dependence (that we can consider as an age effect), model 29 is the first one in which it can be evaluated: in the first week, it seems that, after the drop in the second day of life, from mid-week the hazard tends to rise towards the end of the week.

Model 31 (months 1-11) is the other one in which it is worth to look at the age effect, since it is not uniformly decreasing (older children are supposed to be exposed to a lower risk of dying). In particular, there is a rise of mortality in months 11 and 12, even if this is not very surprising: as introduced by the Focus at the end of Chapter 3, survival in the last months of the first year of life could be critical due to the end of breastfeeding protection.

In Tab. 6.7 we have reported the coefficients of the other individual covariates. The effect of social class seems to be not relevant in the first month of life, while in the last two periods (especially in years 1-4) sons of craftsmen and traders died more than sons of peasants, while for the sons of landowners the risk was lower across all the 5-years period (albeit not significant).

Females were exposed to a lower risk than males, apart from the last period.

The coefficients of the five-years period variable give us a confirm that in our observational window the mortality in the first day of life is the only one going against the general decreasing trend: coefficients are way bigger than 1 in Model 28, always smaller than 1 afterwards.

Finally, the parish fixed effects continue to be significant, confirming the importance of taking always into account the geographical dummies. Many coefficients show high variability due to small numbers of dead children, but the main evidence is that the inversion over time is no longer present. This confirms that we are controlling for selection through other channels.

Table 6.6: Logit regressions modeling the risk of dying in five time spans: temperatures and age. Odds ratio reported. In each *age* column, the row '1' is referred to the respective baseline.

	MODEL 28 Day 1		**MODEL 29 Days 2-7**		**MODEL 30 Days 8-27**		**MODEL 31 Months 1-11**		**MODEL 32 Years 1-4**	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
curr	0.983	0	0.956	0	0.949	0	0.999	0.73	1.034	0
t1			0.940	0	0.989	0.22	1.033	0	1.015	0.05
t2					0.965	0	1.009	0.49	0.994	0.58
t3							0.956	0	1.014	0.09
t4									1.001	0.98
Age										
1			1	-	1	-	1	-	1	-
2			1.163	0.05	0.369	0	0.535	0	0.981	0.79
3			1.241	0.01	0.187	0	0.370	0	0.825	0.01
4			1.167	0.05			0.308	0	0.621	0
5			1.187	0.03			0.218	0	0.604	0
6			1.306	0			0.228	0	0.417	0
7							0.201	0	0.340	0
8							0.216	0	0.283	0
9							0.196	0	0.282	0
10							0.255	0	0.229	0
11							0.352	0	0.245	0
12									0.215	0
13									0.200	0
14									0.114	0
15									0.121	0
16									0.147	0
Constant	0.060	0	0.055	0	0.271	0	0.047	0	0.011	0

Table 6.7: Logit regressions modeling the risk of dying in five time spans: social, class, sex, period and parish. Odds ratio reported.

	MODEL 28 Day 1		MODEL 29 Days 2-7		MODEL 30 Days 8-27		MODEL 31 Months 1-11		MODEL 32 Years 1-4	
	HR	p-val	HR	p-val	HR	p-val	HR	p-val	HR	p-val
Social										
Peasants	1	-	1	-	1	-	1	-	1	-
Crafts.	1.182	0.06	0.930	0.33	0.891	0.12	1.161	0.04	1.300	0
Traders	1.168	0.12	1.001	0.99	0.948	0.49	1.174	0.04	1.405	0
Land.	0.909	0.41	0.904	0.21	0.897	0.19	0.884	0.20	0.912	0.25
Sex										
Male	1	-	1	-	1	-	1	-	1	-
Female	0.729	0	0.919	0.06	0.923	0.08	0.920	0.08	1.025	0.55
Period										
1816-20	1	-	1	-	1	-	1	-	1	-
1821-25	1.223	0.22	0.842	0.07	1.002	0.98	0.726	0	0.972	0.76
1826-30	1.525	0.01	0.831	0.05	1.041	0.69	0.817	0.04	1.124	0.20
1831-35	1.959	0	0.932	0.46	0.996	0.97	1.099	0.32	1.128	0.21
1836-40	1.757	0	0.852	0.09	0.864	0.16	0.624	0.00	1.021	0.83
1841-45	1.567	0	0.841	0.07	0.856	0.14	0.725	0	0.875	0.16
1846-50	1.989	0	0.803	0.03	1.109	0.31	0.649	0	0.849	0.10
1851-55	2.078	0	0.575	0	0.790	0.03	0.516	0	0.986	0.89
1856-60	1.985	0	0.507	0	0.524	0	0.443	0	0.834	0.07
1861-65	1.557	0.01	0.609	0	0.700	0	0.598	0	0.767	0.01
1866-70	0.728	0.30	0.726	0.07	0.650	0.02	0.500	0	0.553	0
Parish										
Chiesan.	1	-	1	-	1	-	1	-	1	-
S.Giorgio	1.334	0.02	0.863	0.14	0.826	0.06	1.118	0.24	1.026	0.76
Urbana	0.789	0.10	2.548	0	1.924	0	0.975	0.82	0.742	0
Eremitani	1.451	0	1.014	0.89	1.162	0.11	1.097	0.33	1.478	0
Onara	1.056	0.68	1.118	0.26	1.080	0.44	1.263	0.02	0.846	0.07
Pernumia	0.896	0.41	0.248	0	0.176	0	0.075	0	0.007	0
Pontel.	0.820	0.11	1.336	0	1.530	0	1.353	0	1.180	0.03
S.Sofia	0.936	0.60	0.783	0.01	0.901	0.28	0.769	0.01	0.990	0.90
Valnog./F.	0.843	0.29	2.485	0	2.071	0	1.168	0.20	0.709	0

6.2 Robustness checks

6.2.1 Family level

In our path towards a better control of the context where a newborn lives his first period of life, the province of Padua allows us to add a crucial level in the multilevel approach driving our analyses, as it is obvious that children living in the same familiar context share a relevant bunch of biological and environmental characteristics. Albeit this is left for future research, the family identification gives us two more additional tools, one individual and one familiar, both interesting in our field of research: the parity, i.e. the birth order of each child, and family size, i.e. the number of children born in the same family.

The family identification process within the database is not straightforward and requires a huge amount of manual work: this is the reason why, at the moment, it has been done only for a subset of the whole sample. In particular, the process consists in grouping the children of the same parents and assigning them the same identification number, but the point is that sometimes minor differences in each name and family name's registration must be solved case by case, using control variables such as marriage date or baptism date.

At the end of the family identification process, we have been able to assign a family id to 26,312 out of 28,889 children (see Fig. 6.1) born in the 9 parishes in the period 1815-70, corresponding to 91% of the total records.

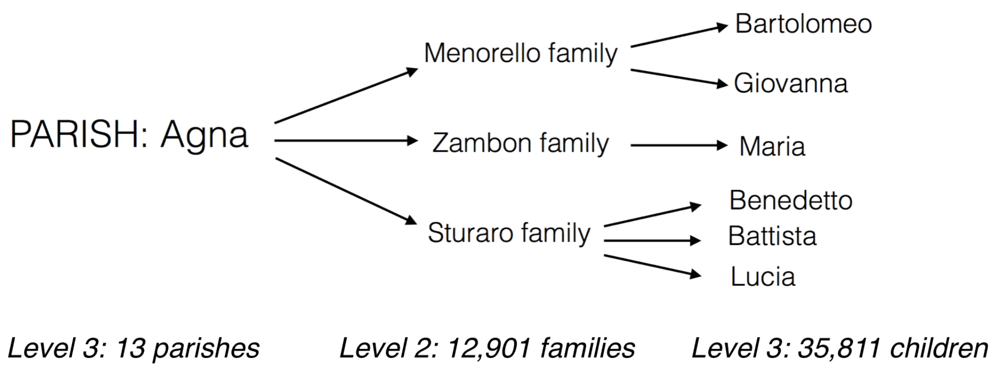


FIGURE 6.1: Hierarchical structure of the data for the parishes of the Padua province.

Within our 9-parishes sample we have been able to identify 9,784 families, corresponding to a total of 26,312 newborns (see Tab. 6.8). This means that the average number of children per family is a mere 2.69, much lower than the real averages of the time: the total fertility rate was 5 in 19th century Veneto (Caltabiano et al. 2015). In particular, it seems that we have a huge quantity of single-childed families (4,442, corresponding

Table 6.8: Total number of families and total number of children in the province of Padua's database after the family reconstruction process, according to the number of children per family.

n. of children	*n. of families*	*n. of children*	*% of children*
1	4,442	4,442	17
2	1,736	3,472	13
3	1,108	3,324	13
4	787	3,148	12
5	525	2,625	10
6	377	2,262	9
7	279	1,953	7
8	203	1,624	6
9	118	1,062	4.04
10	85	850	3.23
11	49	539	2.05
12	30	360	1.37
13	16	208	0.79
14	13	182	0.69
15	6	90	0.34
16	2	32	0.12
17	5	85	0.32
18	3	54	0.21
TOTAL	9,784	26,312	100

to 17% of the total children entries and to 45.4% of the total family entries), and this is a proof of the difficulties within the family reconstruction process: it is very likely that a large proportion of the 'only children' we have detected actually had one or more siblings, but they were not present in the registers (probably they were born in another parish, following a familiar migration[2], or - less likely - we failed in linking them).

Either way, in the two-level analysis we have run, we have excluded the one-childed families and we have kept all the families with at least two children, in order to identify - if existing - the effect of a shared familiar context. Our data can fulfill this purpose, since, albeit we can imagine that many of the children in the dataset had more siblings than the ones we found in the registers, we still can add in our models the crucial notion that some of the subjects must be considered very similar.

We have run several preliminary analyses including the family both as a fixed and random effect, and including the family indicator is always significant. Still, the main results about the temperature effects are essentially the same as those from Models 28-32, commented before.

[2]As stated in Chapter 2, the migrations - albeit existing - can be considered negligible with respect to the study of the early life, since it was unlikely that parents decided to move when their children were infant. When they moved to another parish, likely they did in the time between one child and another, thus making our family reconstruction more difficult .

This is a strong robustness check with respect to our previous results: the temperature effects (and so the mortality selection evidence) are the same also when we control for the familiar context, i.e. they work the same way between subjects that essentially differ only for their birth date and for the weather they experienced in the very first part of their lives.

6.2.2 Non-linear dependence from temperature

As a further check, we have hypothesized that the relationship between risk of dying and temperature (current and past) may not be linear, i.e. the effect of an extra Celsius degree may be stronger in a critical context (passing from 0 to 1 °C, for example) than in an almost neutral one (passing from 17 to 18 °C, for example).

To this purpose, we have divided all the temperature variables in 7 classes: less than 0 °C, from 0 to 5 °C, from 5 to 10 °C, from 10 to 15 °C, from 15 to 20 °C, from 20 to 25 °C and more than 25 °C, and we have re-run Models 28-32 using the so-created dummy variables.

The key point is that the main results of the chapter remain still unchanged, but this kind of relationship gives us some more details. In particular, looking at the first day and the first week of life we noticed that the effect of the current temperature is U-shaped, more than linear: the class 1 temperatures (less than 0 °C) is the one associated to the highest risk of dying; afterwards, the odds ratios drop and reach their minimum at class 5 temperatures (from 15 to 20 °C), and after grow again for the last two classes of temperatures, albeit not reaching the initial peak. This means that, in the very first days of life, cold temperatures were lethal, but also very hot weather was not desirable.

In the subsequent periods, the relationship between risk and temperature classes turns to be more linear: warmer temperatures are monotonically better than colder ones up to the first birthday, monotonically worse afterwards.

6.2.3 Different choices regarding the population of interest

It is also interesting to verify whether the results presented in the chapter change according to the theoretical choices we have done in Chapter 2 about the composition of the population of interest. In particular, the first choice regarded the inclusion in the dataset of the 'dummy births', i.e. the records we have created in order to reconstruct the life course of the unlinked dead children; the second choice regarded the inclusion in the dataset of the stillbirths or, better, the 'suspected' stillbirths.

Apart from a slight loss of statistical significance of the first day temperature coefficient in the model relative to years 1-4, the large majority of the substantial evidences remain unchanged, confirming that the discussion about the two choices regarding the population reconstruction are relevant from a theoretical point of view, but do not affect at all the empirical core of the study.

6.2.4 And the context?

As a logical consequence of all we have said until now, a final check is required as a proof of the progress we have achieved in this chapter. In fact, albeit we have obtained some additional outcomes of interest and consequently new results, the original aim of the daily temperature usage was the necessity to adopt a more reliable indicator for the early life context.

And actually the daily temperature proved to be better: with the inclusion of daily and past temperatures in the models, the effect of the old 'context' proxy we had used in Chapters 4 and 5 is no longer statistically significant. This result is coherent with the one found by Dalla Zuanna and Rosina (2011): in their study about the effects of minimum temperatures on the winter infant mortality within a single Veneto parish, they found that, with the inclusion of the minimum temperature in their models, the effect of the month of birth was no longer significant.

6.3 Summing up

The inclusion of the daily temperatures in our analyses provided strong confirms and new evidences. In particular, we have discovered that the 'early life conditions' (that in Chapters 4 and 5 we had tried to synthesize as a combination of parish, month and year of birth), can easily be described by the weather conditions experienced during the very first part of life. Albeit past temperatures do not influence directly the risk of dying in later life, their post-first-birthday statistically significant effect is a proof of the selection they produced early in life.

As along the dissertation we have progressively narrowed the 'exposure-to-selection' time, in this chapter we have tried to consider the shortest time we could, that is the sole first day of life. And we found that the strongest slice of selection actually appears to take place in the first day of life.

Among the other results, we had new confirms about two interesting evidences: the peak of mortality around the first birthday; the danger associated to hot temperatures rather than cold ones after the first birthday.

In the second part of the chapter we have introduced three different checks that, from different perspectives, give robustness to our results. All of them - especially the familiar one - have indicated several directions for further refinements.

Chapter 7

Conclusions

At the beginning of the Infant Mortality in Asburgical Veneto (IMAV) project, we had a few basic notions about the dramatic reality of infant mortality in the Veneto region during the 19th century. The fundamental keys of interests were essentially two: first, Veneto's situation was completely different not only with respect to other European countries, but also comparing it to neighbouring Italian regions; secondly, the levels of infant mortality in Veneto were as high as 400‰, i.e. one of the highest ever recorded levels in modernity.

We have tried to highlight the particularity of Veneto's case in the first two chapters of the thesis. In particular, it is the comparison with the English case to be quintessential, since at the end of the 19th century the two regions showed comparable levels on infant mortality, but this was the result of very different trajectories (see Fig. 2.1). In fact, if in England the neonatal mortality (i.e. the first-month mortality) had remained stable along the century while the childhood mortality (i.e. years 1-4 mortality) had considerably decreased, in Veneto the living conditions improvements experienced after the half of the century had caused an opposite pattern: a big neonatal mortality drop and, as a consequence, a relative childhood mortality peak (see Paragraph 3.2.5). The point is that the neonatal mortality levels at the beginning of the 19th century were extremely different in the two regions: the huge risk of dying in Veneto during the first month of life (explored in-depth along the Focus in Chapter 3) had to be the starting point to address the peculiarities of this study.

Apart from giving us a hint about *where* to investigate to better understand the problem (i.e. into the very first part of life), the unprecedented levels of infant mortality registered in 19th century Veneto work as a magnifying glass about *how* some early-life-related demographic dynamics operate: dealing with numbers as big as Veneto's ones, they appear more evident and easy-recognisable.

The second part of the thesis, then, has been spent looking for an adequate definition of 'critical period', that is a span of time during which the children under analysis experienced something crucial for the immediate continuation of their lives. Using different statistical methods, we have progressively shortened the length of this critical period, passing from the first 3 months of life (Chapter 4) to the first week of life (Chapter 5), ending with the sole first day of life (Chapter 6). The biggest evidence emerging from all the three choices is that, in the debate between *scar* and *selection* effects, our early-life analyses clearly go in the selection direction. The frail members - or the cohorts experiencing more problematic contexts of births - are selected out, so that the same cohorts turn out to survive more after having survived to the first days of life. Albeit adding details and statistical checks, the existence of a compositional effect (in the form of a homeostatic mechanism) we call *selection* has not been disproved, neither when - in the last chapter - we have limited the 'exposition-to-selection' period to a time as short as the sole first day of life.

Extending the basic and coarser selection results of Chapter 4, in Chapter 5 we have used a stronger two-levels approach, that we considered the one suiting best our database's structure. At the individual level, we found that the social class and the mortality levels experienced earlier in life (which we briefly called 'contexts') were the two variables more remarkably related to the risk of dying in periods following the first week of life. At the community level, albeit we have had confirms about the importance of including in the models a geographical (parish or province) indicator, we have discovered that only the *rurality rate* variable seemed to play a relevant role, the more rural communities being favored over the urban ones, especially after the first birthday.

Finally, in Chapter 6 we have moved to a discrete-time analysis, and we have substituted the 'context' variable with some measures of current and past temperatures, which were available - in a daily format - for the Province of Padua's subset of our data. The models have provided confirms both about the mortality selection issue and the big effect of weather on survival in the first stages of life: if cold weather was really lethal in the first part, hot weather turned to be at least dangerous after the first birthday (with the end of the breastfeeding protection).

In the future, one possible way to improve this evidence could be a model taking into account the additive effect of consecutive very cold (or very hot) days. Among the robustness checks we have presented in the last part of the chapter, the use of the family indicator suggests other future refinements: apart from a possible (albeit time-consuming) extension of the familiar reconstruction to the whole sample, it would be interesting to evaluate, for example, the parity effect, i.e. if the experiences of the previous siblings affected in some way the survival chances of the subsequent newborns.

Bibliography

[1] Alfani, G. (2013). Plague in seventeenth-century europe and the decline of italy: an epidemiological hypothesis. *European Review of Economic History*, 17(4):408–430.

[2] Alter, G. and Oris, M. (2005). Childhood conditions, migration, and mortality: Migrants and natives in 19th-century cities. *Social biology*, 52(3-4):178–191.

[3] Bacci, M. L. (2002). Storia minima della popolazione del mondo. *icon*, 39(051):256011.

[4] Bacci, M. L. (2005). *Storia minima della popolazione del mondo*. Il mulino.

[5] Barker, D. J., Osmond, C., Winter, P., Margetts, B., and Simmonds, S. (1989). Weight in infancy and death from ischaemic heart disease. *The Lancet*, 334(8663):577–580.

[6] Becher, H., Müller, O., Jahn, A., Gbangou, A., Kynast-Wolf, G., and Kouyaté, B. (2004). Risk factors of infant and child mortality in rural burkina faso. *Bulletin of the World Health Organization*, 82(4):265–273.

[7] Bengtsson, T. (2004a). *Life under pressure: Mortality and living standards in Europe and Asia, 1700-1900*. Mit Press.

[8] Bengtsson, T. (2004b). Mortality and social class in four scanian parishes, 1766-1865. *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700-1900*, pages 135–171.

[9] Bengtsson, T. and Lindström, M. (2000). Childhood misery and disease in later life: The effects on mortality in old age of hazards experienced in early life, southern sweden, 1760-1894. *Population Studies*, 54(3):263–277.

[10] Bengtsson, T. and Lindström, M. (2003). Airborne infectious diseases during infancy and mortality in later life in southern sweden, 1766–1894. *International Journal of Epidemiology*, 32(2):286–294.

[11] Billari, F. and Rosina, A. (2000). Analisi della mortalità alle età adulte e anziane: effetti di età, periodo e coorte. *Il sistema demografico alla fine delle grandi epidemie. Venezia, il Dogado e Chioggia fra Seicento e Settecento*, pages 213–230.

[12] Billari, F. C. and Rosina, A. (1998). Does cohort matter in pre-transitional mortality? analysis of adult mortality using an event history approach: the case of chioggia in the 17th century. *Genus*, pages 327–347.

[13] Boco, A. G. (2010). Individual and community level effects on child mortality: an analysis of 28 demographic and health surveys in sub-saharan africa.

[14] Bourgeois-Pichat, J. (1951). La mesure de la mortalité infantile. i. principes et méthodes. *Population (french edition)*, pages 233–248.

[15] Box-Steffensmeier, J. M. and Zorn, C. (1999). Modeling heterogeneity in duration models. In *Summer Meeting of the Political Methodology Society, July15-17*.

[16] Braveman, P. and Barclay, C. (2009). Health disparities beginning in childhood: a life-course perspective. *Pediatrics*, 124(Supplement 3):S163–S175.

[17] Breschi, M., Derosas, R., and Manfredini, M. (2000). 15 infant mortality in nineteenth-century italy: Interactions between ecology and society. *Population and Economy: From Hunger to Modern Economic Growth: From Hunger to Modern Economic Growth*, page 457.

[18] Breschi, M. et al. (1997). Month of birth as a factor in childrens survival.

[19] Breschi, M. and Pozzi, L. (2004). The determinants of infant and child mortality in past european populations. Forum.

[20] Bruckner, T. A., Van den Berg, G. J., Smith, K. R., and Catalano, R. A. (2014). Ambient temperature during gestation and cold-related adult mortality in a swedish cohort, 1915–2002. *Social Science & Medicine*, 119:191–197.

[21] Caltabiano, M. and Dalla-Zuanna, G. (2015). The delayed fertility transition in north-east italy. *European Journal of Population*, 31(1):21–49.

[22] Camuffo, D. (2002). History of the long series of daily air temperature in padova (1725–1998). In *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources*, pages 7–75. Springer.

[23] Caselli, G. (1990). The influence of cohort effects on differentials and trends in mortality.

[24] Corsini, C. A. and Viazzo, P. P. (1997). *The decline of infant and child mortality: the European experience, 1750-1990*. Martinus Nijhoff Publishers.

[25] Dalla-Zuanna, G., Minello, A., and Alfani, G. (2014). First clues of health transition. the parallel decline of early baptism and early mortality.

[26] Dalla-Zuanna, G. and Rosina, A. (2011). An analysis of extremely high nineteenth-century winter neonatal mortality in a local context of northeastern italy. *European Journal of Population/Revue européenne de Démographie*, 27(1):33–55.

[27] Dalla-Zuanna, G. and Rossi, F. (2010). Comparisons of infant mortality in the austrian empire länder using the tafeln (1851-54). *Demographic Research*, 22(26):813–862.

[28] D'Angelo, S., Dalla-Zuanna, G., and Rosina, A. (2003). Massimo risultato con il minimo sforzo. ricostruzione nominativa semiautomatica della sopravvivenza infantile, e analisi del regime demografico di alì nel xviii secolo. *Piccolo è bello. Approcci microanalitici alla ricerca storicodemografica*, pages 67–92.

[29] Davenport, R. (2010). The relationship between stillbirth and early neonatal mortality: evidence from eighteenth century london. *Cambridge Group for the History of Population and Social Structure*.

[30] Del Panta, L. (1980). *Le epidemie nella storia demografica italiana:(secoli XIV-XIX)*. Loescher.

[31] Derosas, R. (2002). Infant mortality broken into pieces: a case study on venice in mid-nineteenth century. *Società Italiana di Statistica, Atti della XLI Riunione Scientifica. Università di Milano-Bicocca*, pages 5–7.

[32] Derosas, R. (2009). The joint effect of maternal malnutrition and cold weather on neonatal mortality in nineteenth-century venice: An assessment of the hypothermia hypothesis. *Population studies*, 63(3):233–251.

[33] Doblhammer, G. (2004). *The late life legacy of very early life*. Springer.

[34] Doblhammer, G. and Vaupel, J. W. (2001). Lifespan depends on month of birth. *Proceedings of the National Academy of Sciences*, 98(5):2934–2939.

[35] Drevenstedt, G. L., Crimmins, E. M., Vasunilashorn, S., and Finch, C. E. (2008). The rise and fall of excess male infant mortality. *Proceedings of the National Academy of Sciences*, 105(13):5016–5021.

[36] Ekamper, P., van Poppel, F., Stein, A., and Lumey, L. (2013). Independent and additive association of prenatal famine exposure and intermediary life conditions with adult mortality between age 18–63 years. *Social Science & Medicine*.

[37] Endler, J. A. (1986). *Natural selection in the wild*. Number 21. Princeton University Press.

[38] Gagnon, A. (2012). Effect of birth season on longevity: Thrifty and hopeful phenotypes in historical québec. *American Journal of Human Biology*, 24(5):654–660.

[39] Gambasin, A. (1973). *Parroci e contadini nel Veneto alla fine dell'Ottocento*, volume 1. Edizioni di storia e letteratura.

[40] Gavrilov, L. A. and Gavrilova, N. S. (2011). Season of birth and exceptional longevity: comparative study of american centenarians, their siblings, and spouses. *Journal of aging research*, 2011.

[41] Glazier, I. A. (1969). L'agricoltura veneta dalla caduta della repubblica all'unità. by marino berengo.[studi e ricerche di storia economica italiana nell'età del risorgimento.](milan: Banca commerciale italiana. pp. 380.). *The American Historical Review*, 74(4):1321–1322.

[42] Goubert, P. (1968). Legitimate fecundity and infant mortality in france during the eighteenth century: A comparison. *Daedalus*, pages 593–603.

[43] Hill, K. and Choi, Y. (2006). Neonatal mortality in the developing world. *Demographic research*, 14(18):429–452.

[44] Hobcraft, J. and Gilks, W. (1981). Age period and cohort analysis effects in mortality studies. [Unpublished] 1981. Paper presented at International Union for the Scientific Study of Population Committee on Factors Affecting Mortality and the Length of Life Seminar on Methodology and Data Collection in Mortality Studies Dakar Senegal July 1981. Published in: Methodologies for the Collection and Analysis of Mortality Data. Proceedings of a Seminar in Dakar Senegal July 7-10 1981 [edited by] Jacques Vallin John H. Pollard and Larry Heligman. Liege Belgium Ordina 1984. p. 245-64.

[45] Huntington, E. (1938). Season of birth-its relation to human abilities.

[46] Kannisto, V., Christensen, K., and Vaupel, J. W. (1997). No increased mortality in later life for cohorts bom during famine. *American Journal of Epidemiology*, 145(11):987–994.

[47] Kennedy, S., McDonald, J. T., and Biddle, N. (2006). The healthy immigrant effect and immigrant selection: evidence from four countries.

[48] Knodel, J. (1970). Two and a half centuries of demographic history in a bavarian village. *Population studies*, 24(3):353–376.

[49] Knodel, J. E. (2002). *Demographic behavior in the past: A study of fourteen German village populations in the eighteenth and nineteenth centuries*, volume 6. Cambridge University Press.

[50] Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., and Power, C. (2003). Life course epidemiology. *Journal of epidemiology and community health*, 57(10):778.

[51] Kynast-Wolf, G., Hammer, G. P., Müller, O., Kouyaté, B., and Becher, H. (2006). Season of death and birth predict patterns of mortality in burkina faso. *International journal of epidemiology*, 35(2):427–435.

[52] Lalou, R. and LeGrand, T. K. (1997). Child mortality in the urban and rural sahel. *Population an English Selection*, pages 147–168.

[53] Lawlor, D. (2004). Commentary: The art and science of epidemiology: governed by the seasons? *International journal of epidemiology*, 33(1):144–146.

[54] Lazzarini, A. (1981). *Campagne venete ed emigrazione di massa, 1866-1900*, volume 3. Neri Pozza.

[55] Livi-Bacci, M. (1986). Fertility, nutrition, and pellagra: Italy during the vital revolution. *Journal of Interdisciplinary History*, pages 431–454.

[56] McDonald, J. T. and Kennedy, S. (2004). Insights into the ?healthy immigrant effect?: health status and health service use of immigrants to canada. *Social science & medicine*, 59(8):1613–1627.

[57] McKeown, T. and Record, R. (1962). Reasons for the decline of mortality in england and wales during the nineteenth century. *Population studies*, 16(2):94–122.

[58] Moeller, M. M. (2013). Methods for analyzing proportions.

[59] Moore, S. E., Cole, T. J., Collinson, A. C., Poskitt, E., McGregor, I. A., and Prentice, A. M. (1999). Prenatal or early postnatal events predict infectious deaths in young adulthood in rural africa. *International journal of epidemiology*, 28(6):1088–1095.

[60] Moore, S. E., Fulford, A. J., Streatfield, P. K., Persson, L. Å., and Prentice, A. M. (2004). Comparative analysis of patterns of survival by season of birth in rural bangladeshi and gambian populations. *International journal of epidemiology*, 33(1):137–143.

[61] Neupane, S. and Doku, D. T. (2014). Neonatal mortality in nepal: A multilevel analysis of a nationally representative. *Journal of Epidemiology and Global Health.*

[62] Oris, M., Derosas, R., Breschi, M., and Alter, G. (2004). *Infant and child mortality.*

[63] Piccione, L., Dalla Zuanna, G., and Minello, A. (2014). Mortality selection in the first three months of life and survival in the following thirty-three months in rural veneto (north-east italy) from 1816 to 1835. *Demographic research*.

[64] Pongou, R. (2013). Why is infant mortality higher in boys than in girls? a new hypothesis based on preconception environment and evidence from a large sample of twins. *Demography*, 50(2):421–444.

[65] Pozzi, L. and Barona, J. L. (2012). Vulnerable babies. late foetal, neonatal and infant mortality in europe (18th-20th centuries). In *Annales de démographie historique*, volume 123, pages 11–24. Belin.

[66] Preston, S. H. and Haines, M. R. (1991). Fatal years: child mortality in late nineteenth-century america.

[67] Quaranta, L. (2013). *Scarred for Life. How conditions in early life affect socioeconomic status, reproduction and mortality in Southern Sweden, 1813-1968*. PhD thesis, Lund University.

[68] Robertsson, O., Ranstam, J., and Lidgren, L. (2006). Variation in outcome and ranking of hospitals: an analysis from the swedish knee arthroplasty register. *Acta orthopaedica*, 77(3):487–493.

[69] Rosina, A. and Zannini, A. (2004). Mortalità infantile. *Il Veneto. Storia della popolazione dalla caduta di Venezia a oggi*, pages 177–194.

[70] Scarpa, G. and Luzzatto, G. (1963). *L'agricoltura del Veneto nella prima metà del XIX secolo: l'utilizzazione del suolo*. ILTE-Industria libraria tipografica editrice.

[71] Schofield, R. and Wrigley, E. (1979). Infant and child mortality in england in the late tudor and early stuart period. *Cambridge Monographs on the History of Medicine*, pages 61–95.

[72] Shen, K. and Zeng, Y. (2013). The direct and total effects of childhood conditions on current health among oldest-old are stronger than those among young-old.

[73] Simondon, K. B., Elguero, E., Marra, A., Diallo, A., Aaby, P., and Simondon, F. (2004). Season of birth is not associated with risk of early adult death in rural senegal. *International Journal of Epidemiology*, 33(1):130–136.

[74] Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.

[75] Visaria, L. (1988). Level trends and determinants of infant mortality in india.

[76] Waldron, I. (1985). What do we know about causes of sex differences in mortality? a review of the literature. *Population Bulletin of the United Nations*, (18):59.

[77] Willekens, F. (2013). Chronological objects in demographic research. *Demographic Research*, 28.

[78] Williams, N. and Galley, C. (1995). Urban-rural differentials in infant mortality in victorian england. *Population Studies*, 49(3):401–420.

[79] Woods, R. (1995). *The population of Britain in the nineteenth century*, volume 20. Cambridge University Press.

[80] Woods, R. (2009). *Death before birth: fetal health and mortality in historical perspective*. Oxford University Press.

[81] Wrigley, E. A. (1997). *English population history from family reconstitution 1580-1837*. Number 32. Cambridge University Press.

Leonardo Piccione

CURRICULUM VITAE

Personal Details

Date of Birth: February 10, 1987
Place of Birth: Corato (Bari), Italy
Nationality: Italian

Contact Information

University of Padova
Department of Statistics
via Cesare Battisti, 241-243
35121 Padova. Italy.

Tel. +39 049 827 4174
e-mail: piccione@stat.unipd.it

Current Position

Since January 2012; (expected completion: January 2016)
PhD Student in Statistical Sciences, University of Padova.
Thesis title: Children Survival in Veneto 1815-70. From the Dark Age to the Dawn of Change.
Supervisor: Prof. Gianpiero Dalla Zuanna
Co-supervisor: Prof. Francesco Billari
Co-supervisor: Prof. Adriano Paggiaro

Research interests

- Population Studies
- Historical Demography
- Life Course
- Social Statistics
- to discover

Education

October 2009 – October 2011
Master (*laurea magistrale*) degree in Statistics, Demography and Society .
University of Padua, Faculty of Statistics
Title of dissertation: "Neonatal and infant mortality in the Veneto region under the Austrian Empire. A nominative linkage and an analysis of mortality selection. " (in Italian)
Supervisor: Prof. Gianpiero Dalla Zuanna
Final mark: 110/110 cum laude

September 2006 – July 2009
Bachelor degree (*laurea triennale*) in Statistical and Economic Sciences .
University of Bari, Faculty of Economics
Title of dissertation: "The 'hot hand' phenomenon in basketball." (in Italian)
Supervisor: Prof. Alessio Pollice
Final mark: 110/110 cum laude.

Computer skills

- **Operating systems:** Mac OS - Windows
- **Statistical software:** Stata - R
- **Other software:** LaTeX- Microsoft Office - iWork

Language skills

Italian: native; **English:** good; **French:** basic.

Publications

Articles in journals
Piccione, L., Dalla Zuanna, G., Minello, A. (2014). Mortality selection in the first three months of life and survival in the following thirty-three months in rural Veneto (North-East Italy) from 1816 to 1835. *Demographic Research* **31**, 1199–1228.

Conference presentations

Piccione, L., Dalla Zuanna, G., Minello, A. (2013). The effect of neonatal selection on mortality in the following months. (poster) *XXVII IUSSP International Population Conference*, Busan, Korea, 26-31/08/2013.

Piccione, L. (2014). The role of the month of birth in a context of extremely high infant mortality. (invited) *Inaugural Conference of the European Society of Historical Demography*, Alghero, Italy, 25-27/09/2014.

Piccione, L., Dalla Zuanna, G., Minello, A. (2015). Mortality Selection in the First Three Months of Life and Survival in the following Thirty-Three Months in Rural Veneto (North-East Italy) from 1816 to 1835. (invited) *Population Association of America - 2015 Annual Meeing*, San Diego, USA, 30/04-2/05/2015.

Other Interests

Sports
Cinema
Social Media

References

Prof. Gianpiero Dalla Zuanna
Department of Statistical Sciences
University of Padua
Via Cesare Battisti, 241-243 — 35121 Padova
Phone: +39 049 827 4190
e-mail: gpdz@stat.unipd.it