# Nutrition early in life and longevity

Maarten Lindeboom, Malgorzata Poplawska, France Portrait,
Gerard van den Berg

This version: 25th March 2012


### Abstract

This paper investigates whether exposure to nutritional shocks in early life negatively affects mortality at older ages, using unique individual data on siblings born during and around the Dutch potato famine of 1846/47. The famine provides exogenous variation on the nutritional status of those exposed to the famine. Moreover, the sibling information allows us to control for other usually unobserved environmental factors at the family level that may affect later life health. For instance, family specific biological or socio-economic factors may mitigate or enforce the impact of nutritional shocks early in life and failing to control for this may bias estimates of the effect of the nutritional shock. We use a unique historical dataset that follows siblings in a family from the time of birth until death. This dataset is merged with data on regional food prices and calories available per capita in the nineteenth century in the Netherlands. Moreover, the calories data allows us to assess effects of the diet composition (proteins and proteins from animals intake) in utero and early life on longevity. We show that conditions during both utero and first five years of life are important for mortality patterns, with the latter having a higher impact, and they remain significant even if we correct for the family component. Next, we find evidence for long term effects of nutritional conditions during first years of life on mortality at age 40-60, especially of males, whereas negative nutrition conditions in utero seem to result in very high selection of survivors, which masks the effects on long term mortality. Finally, we find that children born in families of farmers and in families with low social status were particularly strongly affected by exposure to the Dutch potato famine during utero or early life.

# 1 Introduction

A life course framework is increasingly being used to study disease aetiology. This approach, aims at understanding how physical and social exposures during gestation, infancy, childhood, adulthood, and across generations influence health at older ages. Most epidemiologic, economic and demographic studies using a life course framework are successful in documenting associations between early life (nutritional) conditions and later life health (see e.g. Kuh and Ben-Shlomo, 2004; Gluckman et al., 2008). However, documented associations do not imply that early life (nutritional) conditions are causally related to health at older ages. Unobserved individual characteristics, like parental poverty or biological factors, may simultaneously influence early life (nutritional) conditions and later life health and mortality and this may frustrate the assessment of causality. To be able to detect causal effects one needs exogenous variation in the early life nutritional conditions and longitudinal data to relates these conditions to outcomes later in life.

Randomized experiments are often considered as the superior methodology for measuring causal effects. However, randomized experiments imply in the context of the current project that a randomly chosen population is, for example, restricted in access to food and that this population is followed until death. Such randomization is clearly not feasible. True experiments exist in the development economics literature, but for obvious reasons, the studies only examine the effects of minor variations in access to food, follow the individuals for relatively short periods of time, or are still ongoing (e.g. Martorell, Habicht, Rivera 1995). Experimental animal research provides convincing evidence of prenatal programming during adverse conditions, but these research results may not translate directly to humans (Cleal et al. 2007). Therefore, we must instead rely on non-experimental data and address possible selection issues. In view of this, an instrumental variable approach can be used that uses indicators of individual conditions early in life. These instruments should plausibly affect high-age morbidity or mortality and the only way that they can is by way of the individual early-life conditions. Typically, these are temporary characteristics of the macro environment into which the child is born. Such indicators do not give rise to endogeneity and simultaneity biases, because they are exogenous from the individual's point of view. If one observes an association between such an indicator and the health outcome later in life, then one can conclude that there is a causal effect of early-life conditions on that health outcome. This is the approach followed in this paper.

The paper analyzes effects of restrictions in food in utero and during childhood on longevity, conditional on having reached adolescence. We use the occurrence of the Potato famine of 1846-47, variations in food prices and in the availability of daily calories as exogenous indicators of exposure to adverse nutritional conditions in early life. The analyses are performed on individuals of the Historical Sample of the Netherlands (HSN) -release Long Term Mortality Effects of Potato crisis (LMP) 2012.01. For an extensive description of the HSN data, see Mandemakers, 2000.[1] The HSN-LMP individuals are born between January 1st, 1843 and January 1st, 1854 in the province of Zeeland in the Netherlands. These individual data are merged with data on the occurrence of the Potato famine, regional food prices and calories available per capita (Knibbe, 2007). We provide a detailed description of the data in Section 3.

The use of food price information and daily caloric intake is particularly useful in the current context. Like most other studies that use famines the start and end of the potato famine was not well defined. To be more precise, indeed, the food conditions were particularly worse at the peak of the famine (September 1846 – September 1847), but in general food conditions were bad between 1843 and 1854 and were also fluctuating a lot.[2] This makes it difficult to construct clearly defined groups that were exposed to under-nutrition (the treated) and those who are not (the controls) when one relies on the official period of the famine. With food price information and information on caloric intake a more continuous measure can be constructed that also takes into account the degree of exposure to adverse nutritional conditions early in life.

To our knowledge, information on availability of calories have not been used in the relevant literature yet. Our time series of caloric intake also distinguish between several types of nutrients, namely the total amount of proteins and the amount of proteins coming from

[1] The HSN database covers the full lifetimes of a representative sample of about 80,000 individuals born between 1812 and 1922 in the Netherlands. Besides information on birth, marriage and death of a given person, the HSN data also inform on parental and individual characteristics, like father's occupation.

[2] The severity of the nutritional restrictions in early life may help the identification of effects on long-term mortality.

animals. This allows us to examine whether besides the availability of calories also the diet composition (proteins and proteins from animals) in utero and early life is important for longevity.

Another distinguishing feature of our paper is that it uses data from the so called HSN-release LMP that also provides unique information on family relations between observations. From the beginning of 1850, the Dutch population registers collect (socio-economic) information at each successive family situation of all Dutch citizens (moves, changes in marital status, child births). The HSN-release LMP uses data from these population registers and gathers vital information on siblings of HSN respondents born in 1843-54 in Zeeland. All siblings had to be born between 1843 and 1854 as well, and had to survive until 1/1/1850. The sibling information allows us to control for family specific circumstances that are constant over time and that could enforce or mitigate the impact of the nutritional shock early in life. These family specific circumstances include socio-economic conditions as well as biologic or genetic conditions at the family level. For instance, for a rich family with strong networks it may be easier to have access to food in periods of food scarcity. Or, some families may have intrinsically healthier and stronger children than other families, which may mitigate the impact of a nutritional shock. Needless to say that such detailed information at the family level is generally not available in register data, nor in surveys. To our knowledge our paper is the first in this literature to exploit the sibling information to control for family specific circumstances.

We find that exposure to bad nutritional conditions in the first years of life results in higher mortality at ages 40-60, whereas being exposed to limited nutrition during utero results mostly in high selection of survivors. These two effects are particularly strong for males, confirming the theory that they are more vulnerable to nutritional shocks early in life than females. Moreover, we find that long run effects of limited access to food during early life are stronger for children of farmers than of fathers' with other occupation. Finally, the results are much stronger for children born in families with a low than with a high social status.

Section 2 describes in more detail the nutritional conditions in the Netherlands in 1830-60. Section 3 provides information on the data: the individual data on mortality and the macroeconomic indicators that we use to control for the nutritional conditions. Section 4 presents the results of the nonparametric tests for differences in life expectancy resulting from differences in nutrition levels during early life, proxied by calorie availability and potato prices. In Section 5 we analyze these effects by using parametric models, which allow us to correct for family effects, as well as to analyze the impact of nutrition in early life on mortality at different ages later in life.   In Section 6 we check sensitivity of our results, by investigating other proxies for food food availability: amounts of proteins and proteins from animals available and prices of wheat and rye.

# 2 Historical developments related to nutritional conditions in the period 1830-60

Compared to other European countries, the diet of the Dutch population at the beginning of the nineteenth century was varied and rich (Knibbe, 2007). In 1830-40, the amount of available daily calories reached on average 2,400 and the share of proteins from animals in the average diet was high compared to in other Westerns countries (see Figure 1 below). [3] In 1825-1844, the average Dutch diet consisted mostly of potatoes (15%), wheat (13%), rye (19%), dairy (30%) and meat (8%) (Knibbe, 2007).

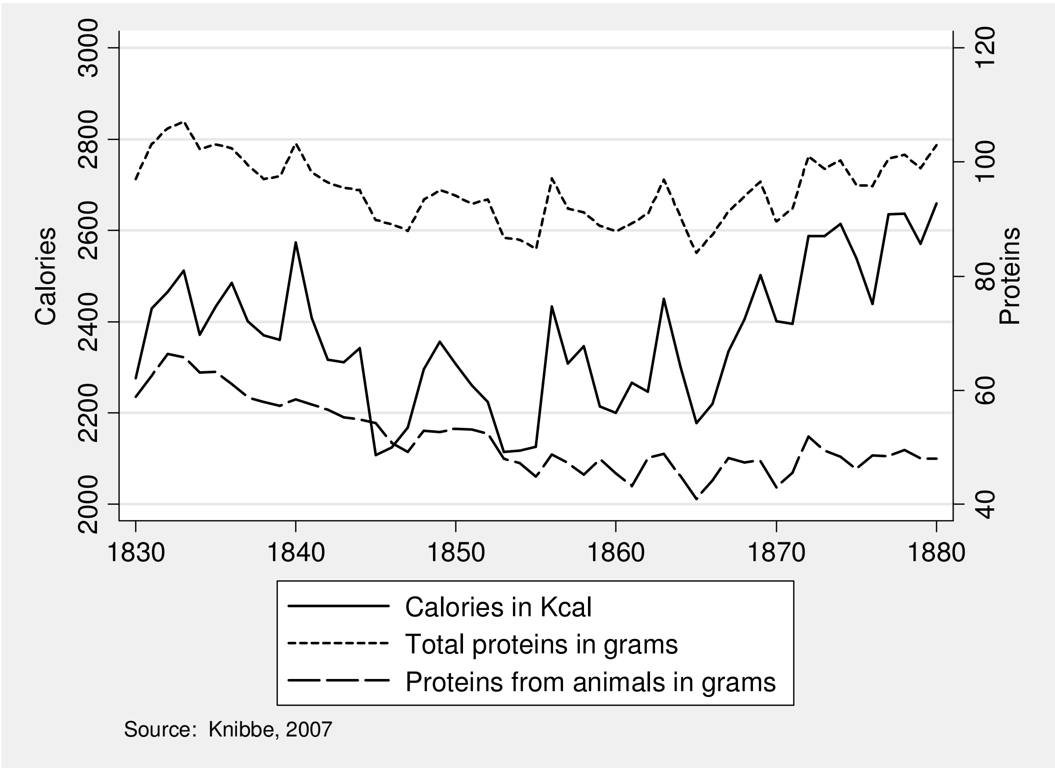


**Figure 1 Availability of daily calories and proteins per capita in the Netherlands per calendar year**

In the subsequent years, the Netherlands had to face several crop failures and a steady growth of its population. This led to a gradual deterioration of the nutritional conditions. Only after 1860 did the availability of food return to the levels of the beginning of the nineteenth century

[3] The share of proteins from animals in the Dutch diet equalled about 44% of the daily available calories, Knibbe, 2007.

and, from 1870, did massive imports of grains considerably improve the average Dutch diet (Knibbe, 2007).

Most importantly, in 1845, the until then unknown Potato blight devastated the potato crops of most European countries. Most of its victims were claimed in Ireland, but the blight also severely hit the Netherlands and the rest of Europe (Vanhaute, Paping, Ó Gráda, 2007). The disase was extremely devastating because no cure existed for quite a long time and because it affected all varieties of potatoes (Vanhaute, 2007).

In 1845, about 70% of the Dutch potato harvest was lost. The province of Zeeland was hit very hardly by the Potato blight as the average production of potatoes per hectare dropped from 158 hectolitres in 1842-44 to 8 hectolitres in 1845 (Paping and Tassenaar, 2007). The Dutch government enacted ad hoc relief measures that were successful in temporarily putting off a nutritional disaster. For instance, stocks from previous years were used to compensate for potato losses, financial assistance was provided to the most affected populations, measures to impede food exports and stimulate cereal and potato imports were taken. Unfortunately, in 1846-47, not only the potato but also the rye and wheat crops (three basic staples in the Dutch diet) partially failed. The potato crop and the rye harvests were seriously down in yields (namely a little bit less than half of the normal yield), and the wheat crop yield was about two-thirds of a normal one. The situation was aggravated by a long hot and wet summer (almost the perfect conditions for infectious diseases to spread) followed by a very cold winter in 1846-47. At the same time, there was a steady growth in dairy exports. Though the measures taken in 1845 were resumed, it was clear that the famine was imminent in the Netherlands. The most tragic period was between September 1846 and September 1847, as potatoes are usually harvested in September and early October. The years 1846 and 1847 are characterized by a sharp drop in the availability of calories (see Figure 1) and by high food prices (the prices of most food products more than doubled in the first half of 1847) (see Figure 2, for food prices in Zeeland). The potato harvests from 1847 until 1851 were also affected by the Potato blight, but to a lesser extent. In 1853-55, the calorie availability was also very low, when harvest failures coincided with high import prices caused by the Crimean War in 1853-56 (Knibbe, 2007). Only after 1856 did the potato yield and the nutritional conditions return to the levels of 1830. All these developments led to a gradual decrease of the share of proteins and of potatoes and to a gradual increase of the share of grains in the average Dutch diet.

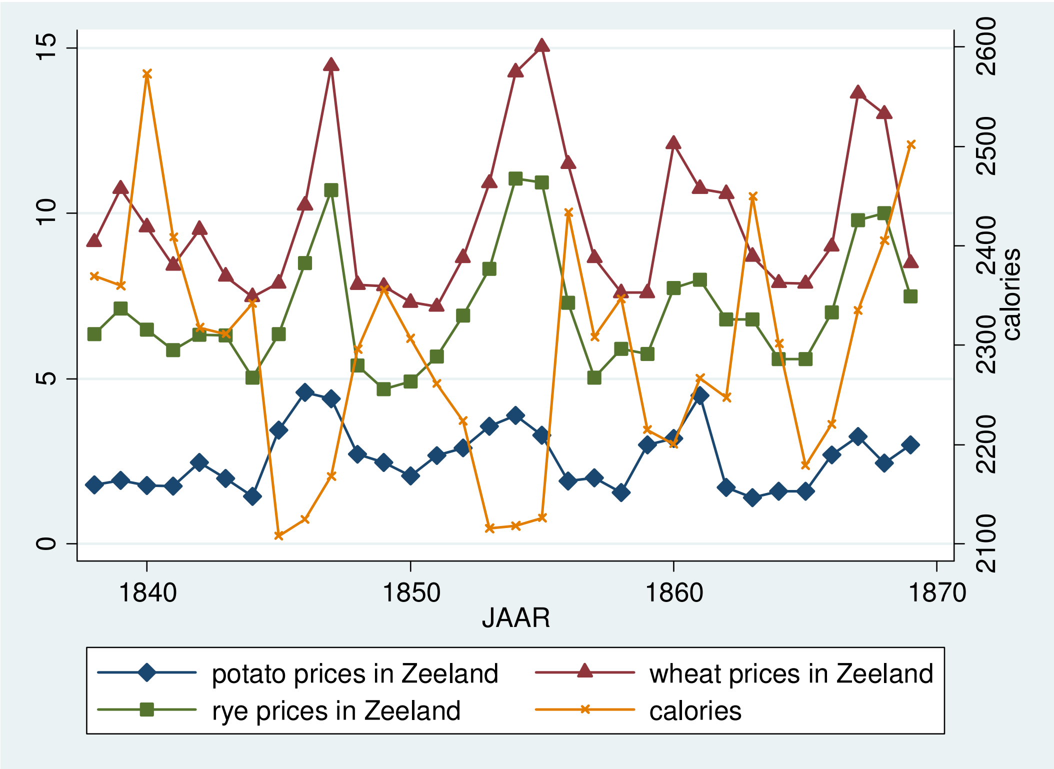


**Figure 2 Daily calorie availability per capita and average prices of potatoes, wheat and rye in Zeeland per calendar year**

The demographic consequences of the severe nutritional restrictions were tragic. Note that the province of Zeeland was disproportionally affected. This is partly explained by the fact that the province of Zeeland of the time was relatively poor (with high levels of unemployment and bad housing conditions) (Kort, 2001). The province was also highly rural: in 1840, only 24% of the Zeeland population was living in agglomerations above 5,000 inhabitants, compared to 35% in the Netherlands (De Meere, 1982) and, in 1849, 48% of the total labour force were working in the agricultural sector, compared to 36% in the Netherlands (Oomens and Den Bakker, 1994). As one may expect, the nutritional restrictions had its strongest immediate impact on the lower class, and especially on the rural lower class which was heavily dependent on agricultural produces not only for consumption but also as a source of income. Finally, the conditions of the water in the province of Zeeland were extremely bad. The ground water was mostly brackish, which made only the rain water suitable for consumption and led to periodic shortages in good quality drinking water (Kort, 2001).

In 1846-47, up to one third of the infants died before age one in the province of Zeeland (this compared to 20% in the Netherland at the time). This excessive infant mortality can be most probably ascribed to very low levels of breastfeeding, the concurrent relatively short birth intervals, the poor water conditions, and, to a lesser extent, the recurrence of malaria epidemics (Hoftee, 1983). Regarding to breastfeeding, the rural women had often to work hard on the land and were not able to suckle their babies. These were often fed with porridges and (contaminated) water, which was often fatal for weak babies.

The child mortality was also considerable, but much smaller than the infant mortality: about 15% of the children in Zeeland died at ages 1-5 in 1845-47 (check) (van Poppel and Mandemakers, 2002). Older individuals, especially in coastal area as in the province of Zeeland, also faced increased mortality rates.

There is also evidence of reduced fertility rates (by about 15% in September and October 1846 on average in the Netherlands, Paping and Tassenaar, 2007) and a slight increase in the number of emigrants, notably in 1847 and for the province of Zeeland (Fokker, 1877; Dominicus, 2011).

Of interest for our study, note that during the period we consider (1840-1860), there was no significant increase in life expectancy of Dutch people, the demographic transition occurred only after 1870 (ref[4]). Hence, we can directly compare life expectancies of cohorts born in 1840-60 but under different nutritional conditions, in order to analyze the impact of conditions in utero and during early life on longevity of individuals.

[4] ref

# 3 Data

## *3.1 Individual data*

Our main data come from the "*Historical Sample of the Netherlands (HSN). Dataset Long Term Mortality Effects of Potato Crisis (LMP), release 2012.01*" (HSN-release LMP 2012.01). The data includes information on 2200 individuals born between 1843 and 1854 in the province of Zeeland. The LMP dataset includes 436 original HSN respondents, who are followed from birth until death (377 have a known date of death), and 772 siblings of the HSN respondents. As mentioned in the introduction, information on siblings is only available from January 1st [1]1850. [5] Because the focus of the main data is on long-term effects of early life nutritional conditions and because of the very high levels of infant and child mortality in Zeeland, the original HSN sample was extended by 432 individuals. These individuals were randomly selected in the certificates of birth of Zeeland. Most important inclusion criteria were not to be a sibling of the original HSN respondents and having reached at least age 10. Similarly, only the siblings of the additional HSN respondents who had reached age 10 were included in the LMP data set (560 observations).[5] We end up with a sample of 2200 individuals and 868 families.

The data consists in 39.5% of the HSN respondents. Brothers and sisters of the main respondents account for 26.2% and 28.3%, respectively, of our sample. The remaining 5% of individuals are labelled as half-brothers or half-sisters of the main respondents. The siblings who were not born in Zeeland were excluded from the main analyses. The average number of individuals within one family equals 2.50 and 321 individuals have no sibling born in 1843-1854**.** Figure 3 shows the number of individuals in our sample by the year of birth. The exact age of death is known for 75% of our participants. We observe most of the dates of death of individuals who migrated to other parts of the Netherlands. For the remaining 25% we use censoring information, which we obtain from registered observations of various events in life, such as a moment of marriage. These observations are considered in the analysis as right-censored.

[5] Obviously, we will have to account for this in the statistical analyses.

**Table 1 Summary statistics**

| | All | Males | Females |
|---|---|---|---|
| **Total (#):** | 2200 | 1085 | 1115 |
| **% with date of death** | 74.9 | 74.7 | 75.2 |
| **Mean(age of death) (uncensored obs.)** | 49.8 | 50.2 | 49.5 |
| **% death above 40 (uncensored obs.)** | 62.8 | 64.1 | 61.4 |
| **Characteristics at birth*** | | | |
| **Mean HISCLASS (1:high -12:low)** | 9.03 | 9.16 | 8.9 |
| **# Fathers in low social class**** | 1,358 | 677 | 681 |
| **# Fathers in high social class**** | 557 | 266 | 311 |
| **% Fathers farmer** | 62.8 | 65 | 60.7 |
| **# Fathers farmer** | 719 | 330 | 389 |
| **# Fathers non farmer** | 1216 | 613 | 603 |
| **% HSN respondents born in rural area** | 81.8 | 81.6 | 82.2 |

** After exclusion of individuals with missing information on characteristics at birth*
*** Low social classes if HISCLASS >8 and high social classes if HISCLASS<8 (see Subsection 3.4)*

Figure 4 shows the histogram of ages of death of those observations for which we know the moment of death. It is important for the further analysis that for each age bracket we observe enough failures, hence we can compare mortality at different ages.

Table 1 presents summary statistics of most of the variables in the dataset we use. The average length of the observation spell in the whole sample, including censored observations, equals 43.1 years, while the average duration of life in the sample of uncensored observations equals 49.8 years. 49.3% of individuals in our sample are males and 62.8% of observations come from families where a father was a farmer. As much as 81.8% of observations in our sample were born in rural areas (which we define as a place with fewer than 5000 inhabitants).

In order to identify the social status of a given family we use information on father's occupation at birth, which is included in the HSN dataset. This information is available for 87.95% of the HSN individuals. By using the Historical International Standard Classification of Occupations (HISCO) codes (van Leeuwen, Maas, & Miles, 2002), we map father's occupation into one of 12 HISCLASS classes (van Leeuwen & Maas, A short note on HISCLASS, 2005). Each class clusters professions together with roughly the same workload, skill level and within the same economic sector. In this study, individuals with HISCLASS less than 8 are considered to have high social status, those with HISCLASS higher or equal to

8 are considered to have low social status. Note that low social status in many cases overlaps with father's occupation as a farmer (89.5% of individuals with low social status come from families of farmers). According to Table 1, 30% of individuals come from families with high social status.

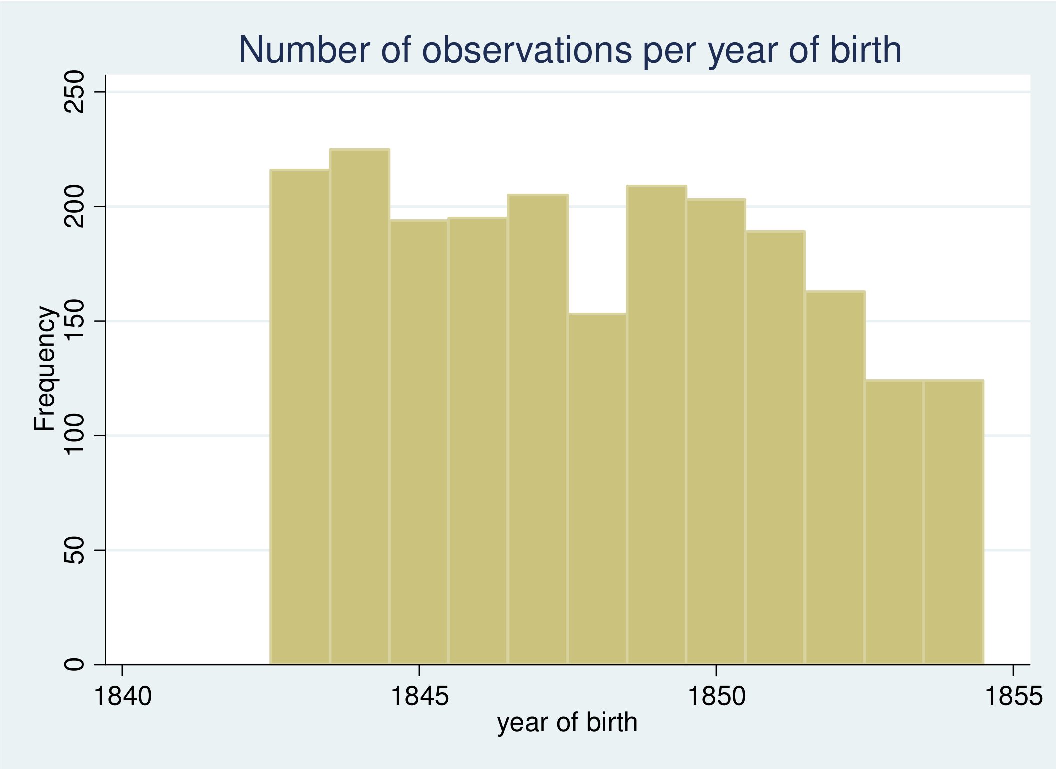


**Figure 3 Frequency of respondents by year of birth**

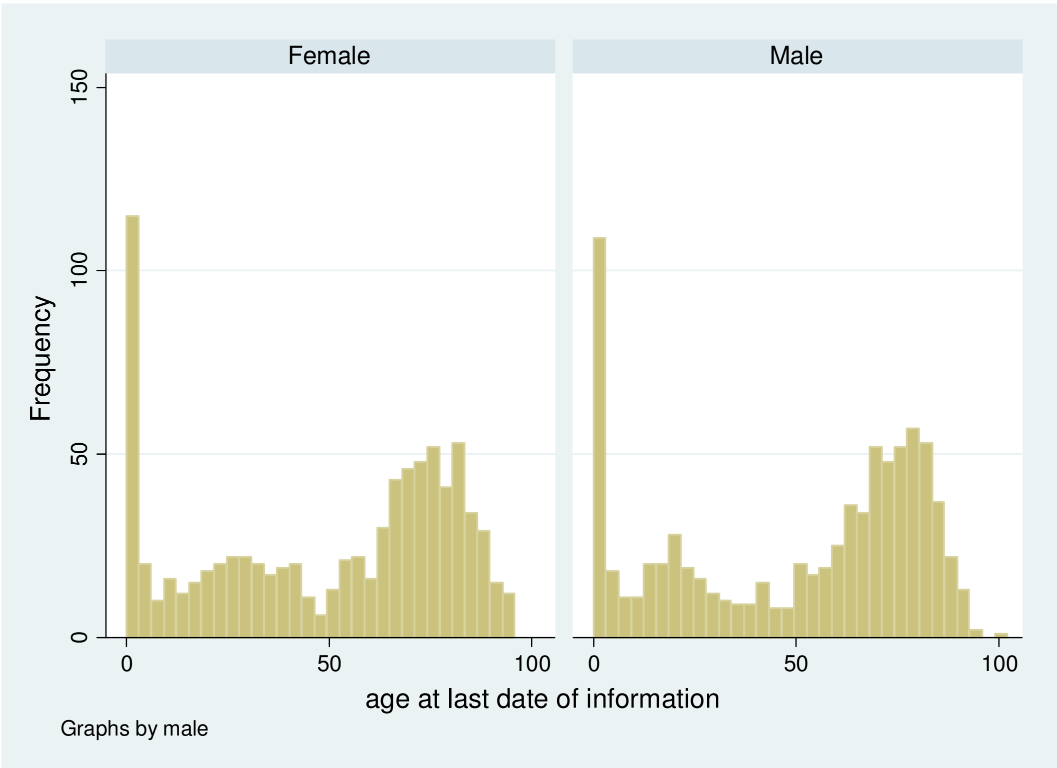


**Figure 4 Histograms of age of death of uncensored observations**

## *3.2 Macroeconomic data*

Besides information on the occurrence of the Potato famine, we merge the individual information with yearly average amount of calories, proteins and animal proteins available in the Netherlands (Knibbe, 2007) and with regional information on food prices in Zeeland (Riel). This is to proxy for nutritional conditions in utero and during early life. Average daily availability of calories seems to be a relatively precise measure of nutritional conditions in a given period and market food prices are commonly used indicators for access to food in historical studies (Almond 2002; Jacobs and Tassenaar 2004). However, note first that levels of available calories include food imports, but that they still may slightly differ from calories intakes. Indeed, in periods of shortage of food, individuals are likely to use stocks from previous years. Nevertheless, in the period under study, the stocks were most likely very low because of the recurrence of the nutritional shocks. Second, ideally we would like to have access to local or regional series. At the time of this study, these data were not available. Note however that, though the Netherlands witnessed in the first half of the nineteenth century substantial regional variations regarding food availability, transport considerably improved at that time, and, national food series could, as from the fifties, depict reasonably well the regional food availability (Knibbe, 2007).

In addition, note that the levels of calories available each year in the Netherlands, presented in Figure 1, are averages for males and females. Nowadays the recommended daily calorie intake for males equals 2,500 kcal, for females 2,000 kcal, and the official rate for minimum daily intake is 1,800 kcal. During the nineteenth century in Zeeland people did jobs that were on average physically very demanding and they were mostly working outside, often in stringent weather conditions, so they needed higher amounts of calories per day. Moreover, Zeeland was a poor region so the average daily consumption was most probably even lower than presented in Figure 1. Therefore, even though the amounts presented in Figure 1 seem to be relatively high, periods with average calorie availability below 2200 kcal refer to severe nutritional restrictions in Zeeland. According to Figure 1, between years 1840 and 1860 there were two periods with very low calorie availability: 1845-1847, the period of the potato famine, and 1853-1855, when harvest failures coincided with high import prices caused by the Crimean War (Knibbe, 2007).

As mentioned above, the calories series are calculated at the national level and may most likely overestimate access to food in Zeeland. Therefore, as another proxy for nutritional conditions we use a variable which is a measure calculated at the regional level: level of food prices in Zeeland. In our analysis we use data on prices of potatoes, rye and wheat. Figure 5 shows average daily calorie intake (at the country level) per year together with the level of potato, wheat and rye prices in Zeeland, expressed in guilders per mud (about 70 kilo). We can see that during years when food availability was low, prices were high and vice versa. Figure 6 shows that periods with high prices were usually also characterized by increased infant mortality in Zeeland. We use historical data on infant mortality from the report of Statistics Netherlands (Centraal Bureau voor de Statistiek,van der Bie & Smits, 2001).

We calculated correlation between prices of different food components between years 1839 and 1865 (this period covers utero and first years of life of all observations in our sample): it seems that prices of rye and wheat were highly correlated (correlation equals 0.93), while they were not too strongly correlated with potato prices (correlation with rye 0.67, and with wheat 0.54). Correlation of daily amount of calory intake and potato prices in the same period equals -0.69, whereas with rye and wheat equals, respectively, -0.50 and -0.37.

Figure 7 and Figure 8 show levels of potato prices and calorie availability together with the calculated values of the trend in years 1840-1870 (utero and early life periods in our sample) in Zeeland (to obtain trend components we have use Hodrick-Prescott filter with the smoothing parameter 1600). There seem to be almost no trend in the series of potato prices, therefore in the further analysis we use levels of potato prices. However, for the calorie availability data the trend component seems to be not constant, hence we decided to use deviations from the trend. In Section 5 we redo the main analysis using cyclical components of prices and calorie consumption levels.

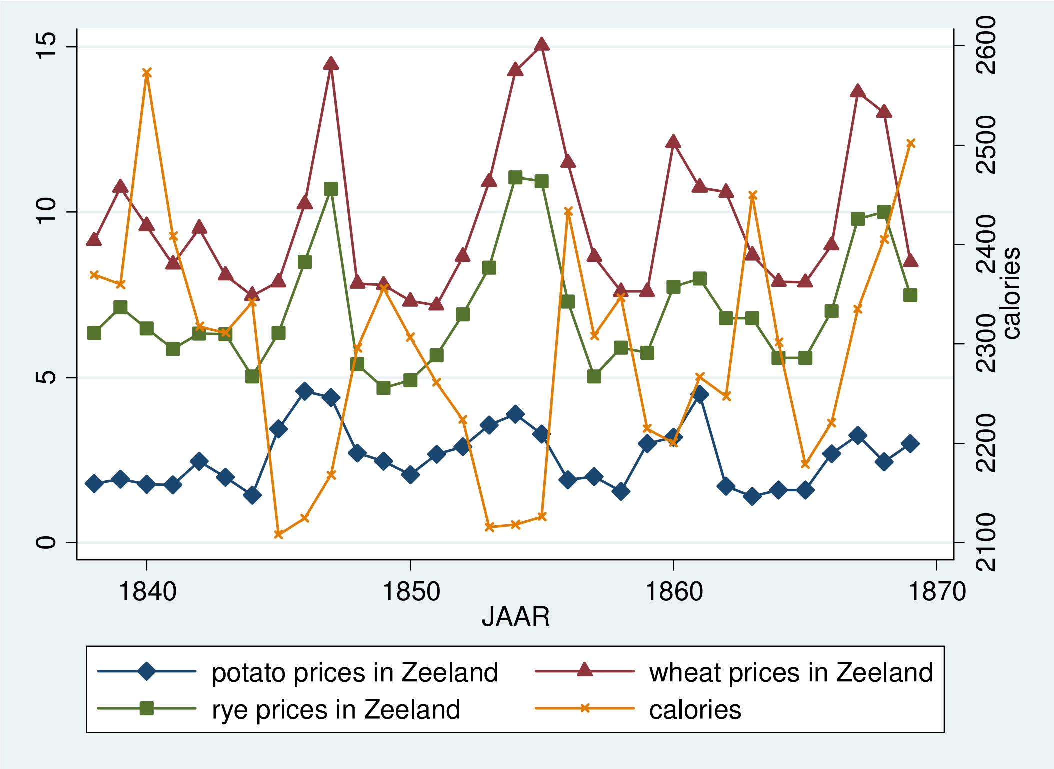


**Figure 5 Daily calorie availability per capita and average prices of potatoes, wheat and rye in Zeeland per year**

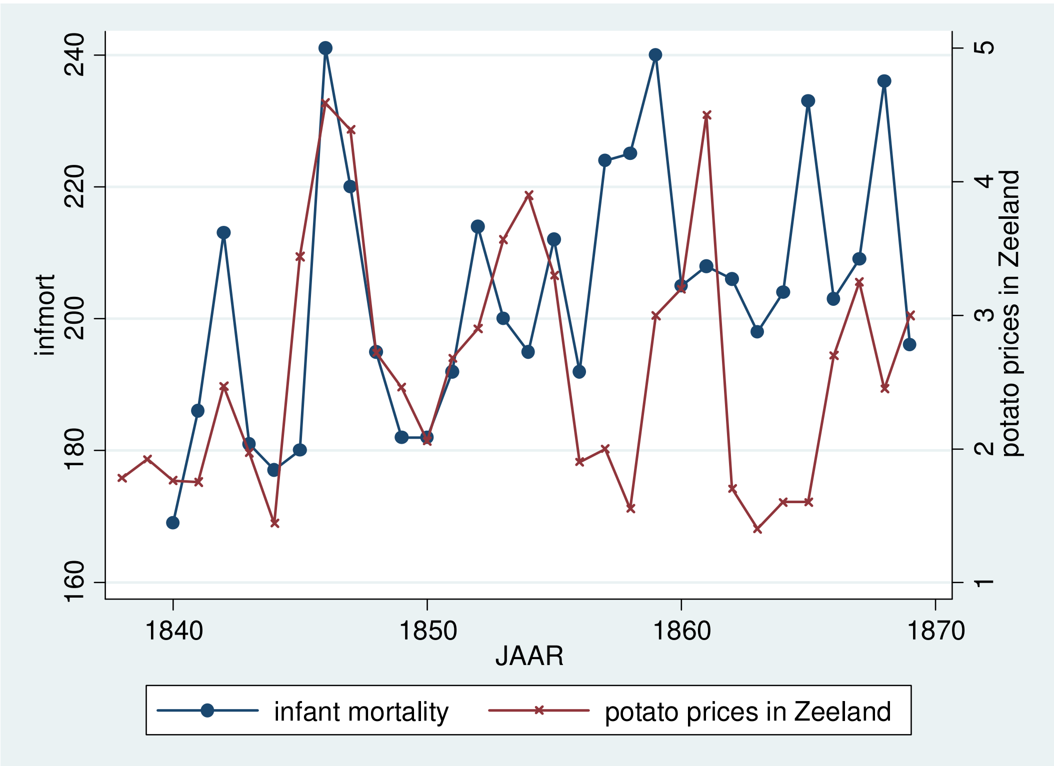


**Figure 6 Average yearly prices of potatoes and infant mortality in Zeeland**

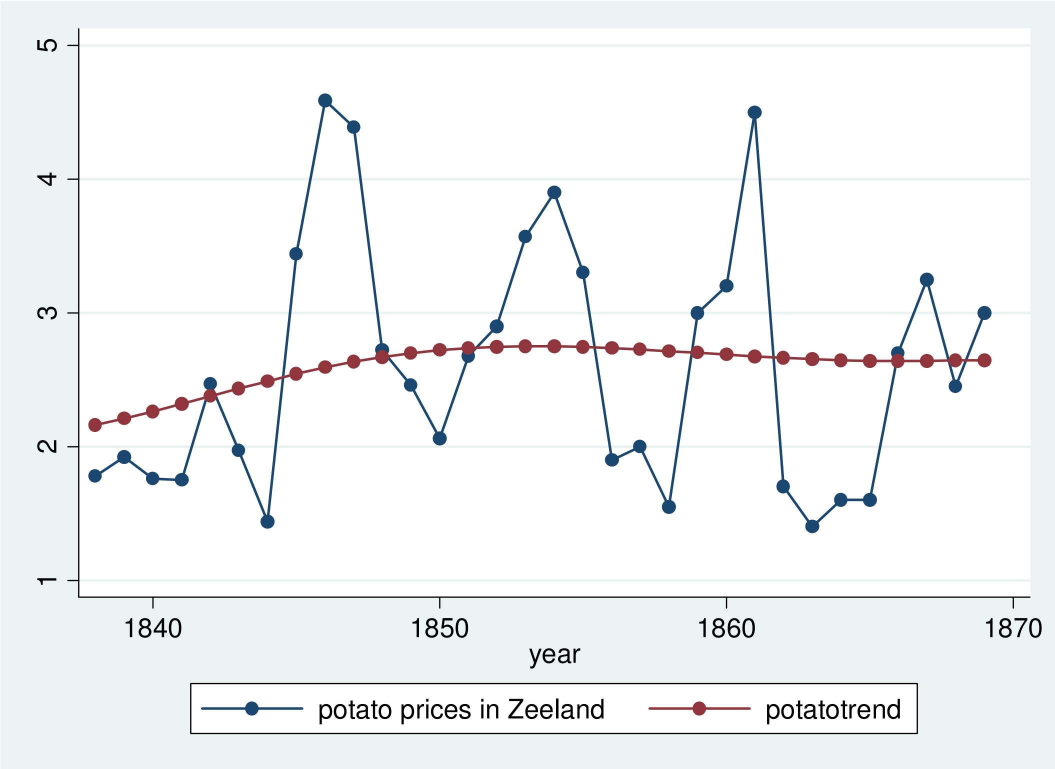


**Figure 7 Levels and cyclical components of potato prices in Zeeland**

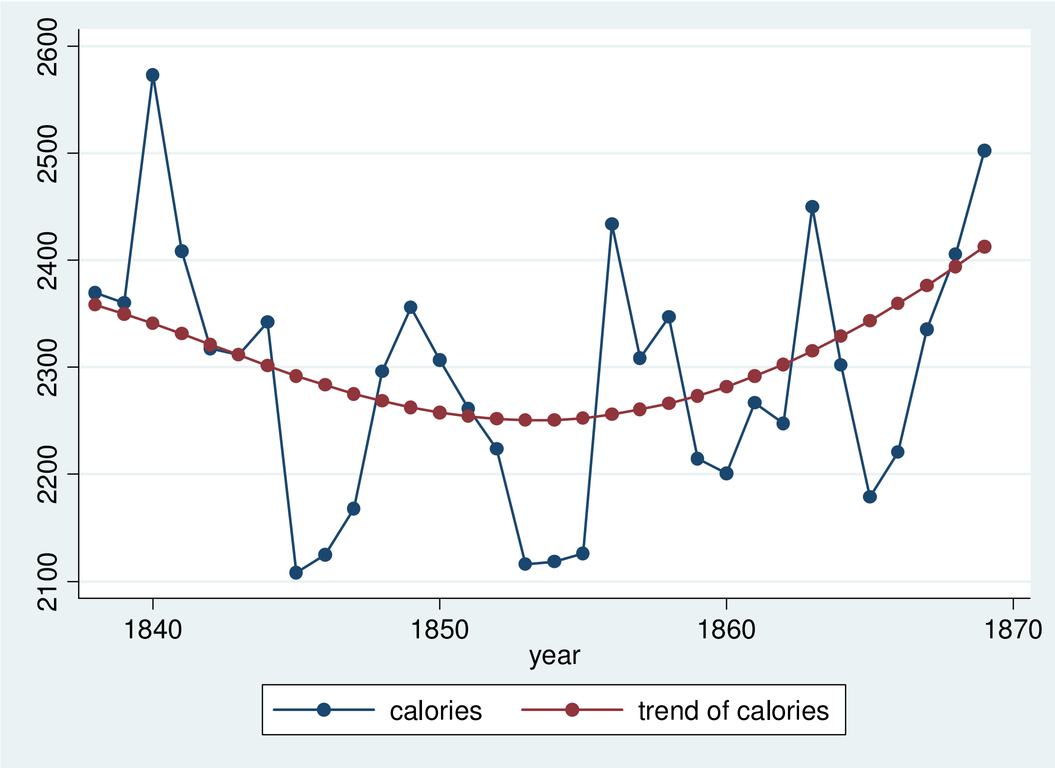


**Figure 8 Cyclical components vs levels of average daily number of calories available per year in the Netherlands**

# 4 Nonparametric analysis

We non parametrically compare lifetimes of individuals exposed to different nutritional conditions in utero and during first years of life. Due to important gender differences in survival patterns, we perform the analysis separately for males and females (ref Dorly[6]). We proxy nutritional conditions by the calorie availability, calculated by the average daily number of calories available per capita in the Netherlands, and by the level of potato prices in Zeeland, obtained from yearly averages of real food prices in Zeeland. In the analysis we consider cyclical components of calories available and levels of potato prices (see Section 3).

We consider exposure to nutritional conditions during two periods early in life: in utero and in the first 12 months of life, and between age 1 and 5. The first period refers to conditions during utero and around the moment of birth (with high infant mortality in the first year of life). We do not consider the two periods separately, because of high correlation (by construction) between measures of conditions in utero and the first year of life: we calculate calorie availability and level of potato prices in this period by taking a weighted average of yearly averages during the period starting from 9 months before birth until 12 months after the birth. If a person died before age 1, we calculate the weighted average between 9 months before birth until the moment of death. Similarly, we calculate weighted averages of prices/calorie availability between age 1 and 5 or until the moment of death, if it occurred before age 5 (if a person died before age 1, we specify these values as missing). This period refers to conditions during early childhood.

First, we want to compare survival of people exposed to good or bad conditions early in life, with low calorie availability and high prices referring to unfavourable nutritional conditions. To do this, we plot Kaplan Meier estimates of the survivor and the hazard function calculated separately for two groups of individuals who experienced different nutritional conditions early in life. For each exposure period of time we calculate KM estimates for people who experienced average calorie availability higher or lower than the 25$^{th}$ percentile of the values in the sample (we have calculated the threshold values as -122.35 for the exposure period until age 1 and -83.07 at age 1-5). Similarly, for the second measure – prices of potatoes, we compare KM estimates of those who experienced prices higher or lower than the 75$^{th}$

[6] ref

percentile of people in the sample (the threshold are 3.66 and 3.42, respectively). Because the graphs plotted using calorie availability and potato prices were very similar, we present only the ones using potato prices, taking into account local conditions in Zeeland.

According to Figure 9, the survival function estimate of females who experienced high prices of potatoes during the utero period and during the first 12 months of life lies below the one plotted for females born under more favourable conditions, which suggests higher mortality of females born under severe nutritional conditions. However, if we look at the plot of the hazard rate (Figure 10), it becomes clear that this relation is mostly due to very high mortality during first years of life of those who experienced high prices at the beginning of their lives. If we look at the KM estimates plotted for males, the survivor function of those born under bad conditions lies even above the one plotted for people born under better conditions, suggesting higher mortality of those born under lower prices. However, the plot of the hazard rate function suggests that this is mostly due to much lower mortality during age 20-30 of people who experienced bad conditions during utero and first 12 months of life. This can suggest that the severe nutritional shock experienced during utero resulted in very strong selection of those males who were actually born, therefore the male survivors are healthier than the comparable age groups born in a period with better food availability.

Figures 11 and 12 show KM estimates of the survivor function and the hazard rate function conditional on survival until age 1, depending on the level of potato prices between ages 1 and 5. We can see that being exposed to nutritional shock during age 1-5 resulted in very high mortality in the first years of life, but also in higher mortality around age 30 and 50-60 (males) and around age 40 (females).

Next, we perform log rank tests for equality of survival functions of individuals exposed to different nutritional conditions in utero and first 12 months of life and between age 1 and 5. The long rank test allows us to correct for the part of the unobserved heterogeneity related to family membership (including both genetic and environmental factors), by using a stratified version of the test. In this version of the test, we group together individuals born in the same family and we assume that all of them share the same value of the unobserved heterogeneity component. In order not to make the stratas too small, we run the stratified version of the test only for both genders together (otherwise we would have to consider men and women separately and would end up having very few observations for each family).

As before, we compare people who experienced calorie availability higher or lower than the 25th percentile or potato prices higher or lower than the 75th percentile. Moreover, now we also compare people born under very good and very bad conditions, by looking at those who experienced calorie availability in the top 25th percentile or prices in the bottom 25th percentile (the thresholds are: for calorie availability higher than 21.93 during utero and first 12 months of life and higher than 20.71 at age 1-5, for potato prices lower than 2.27 and 2.65, respectively), and we compare them with the bottom/top 25% percentile of calorie availability/prices.

Table 2 shows the p-values of the test (with small p-values suggesting significant differences between the survival functions of each subgroup). According to the table there are no significant differences in mortality patterns of people born under different calorie availability in utero and first 12 months of life. If we look at the potato prices, the stratified version of the test suggests that at the 5% significance level people born with 25% highest prices had different mortality patterns than those born under 25% lowest prices. If we consider the exposure period between age 1 and 5, there are significant differences between mortality patterns of people who experienced very high or very low calorie availability during this period. The result is significant only at 10% significance level for females, but at 5% significance level for males and for both genders together. However, it disappears if we run the stratified version of the test, correcting for family effects. If we use potato prices as an indicator for nutrition conditions at age 1-5, the test suggests that there are significant differences between the survival functions of those who exposed to prices higher or lower than top 25% during this period, and also of those who experienced very high or very low prices. This result is significant if we look at males and females, or at both genders together, or if we stratify the sample by the family membership.

Hence, the log rank test detects significant differences in mortality patterns mostly if we compare different nutritional conditions at ages 1-5. However, the test compares the distance between the survival functions, which is significant only if there is a high drop in survival at some age (which is visible only in Figure 10), and it does not really take into account differences in mortality at different ages. Therefore, in order to formally test long term differences in mortality patterns of people who experienced different nutritional conditions early in life, we need to use parametric models.

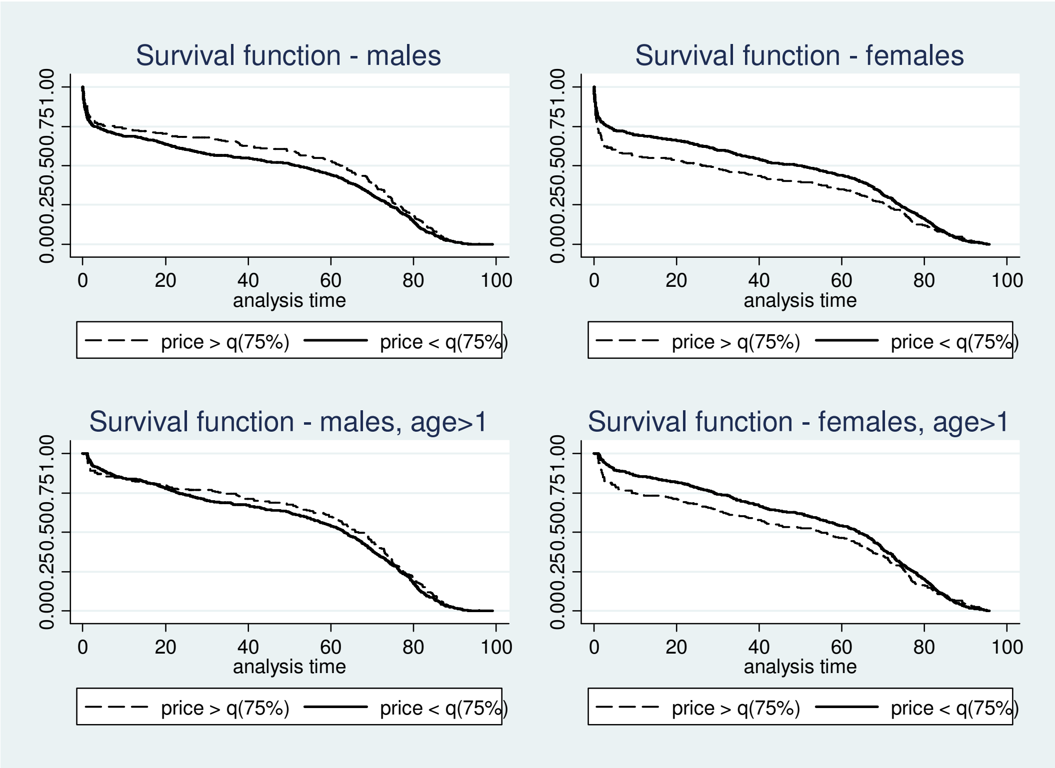


**Figure 7 Survival function estimates for all ages and conditional on survival until age 1, if the average level of potato price during utero and until age 1 was higher/lower than 75th percentile**

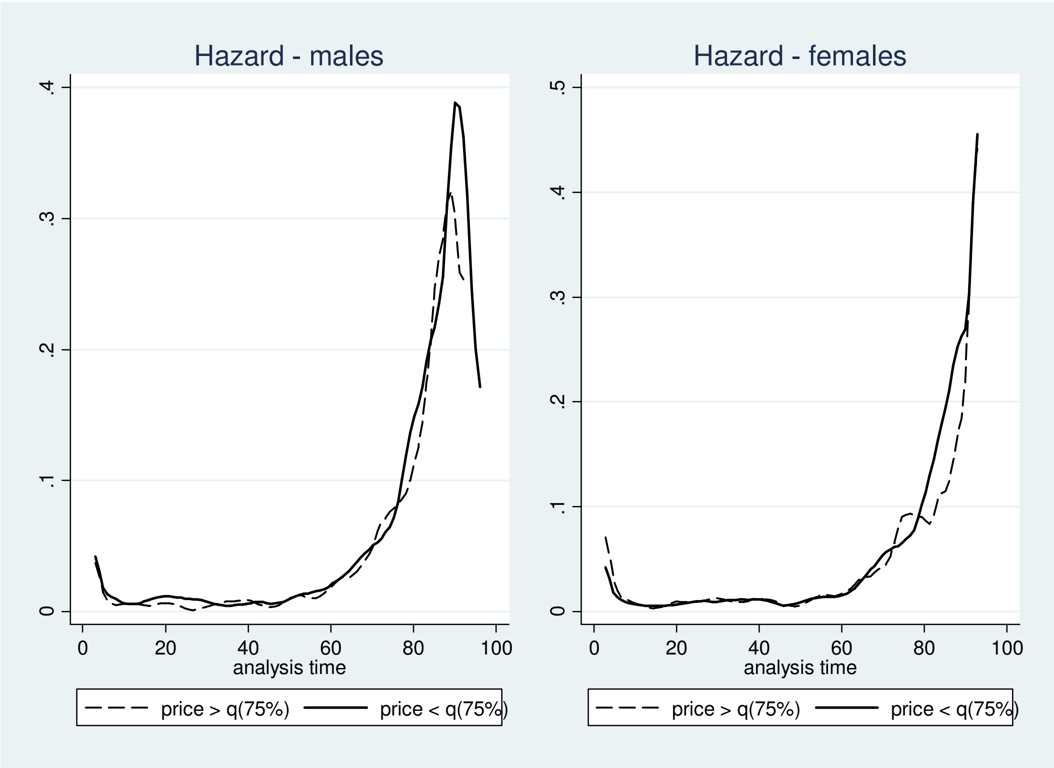


**Figure 8 Hazard rate estimates if the average level of potato price during utero and until age 1 was higher/lower than 75th percentile**

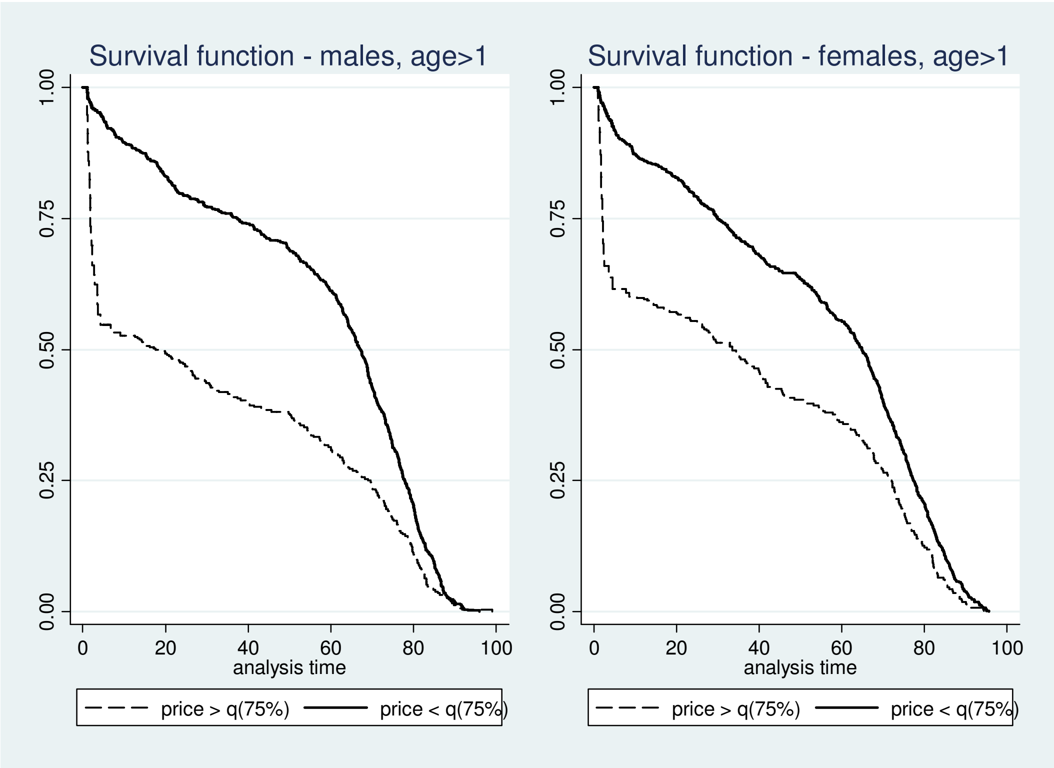


**Figure 9 Survival function estimate per level of potato price during age 1-5, plotted for conditional on survival until age 1**

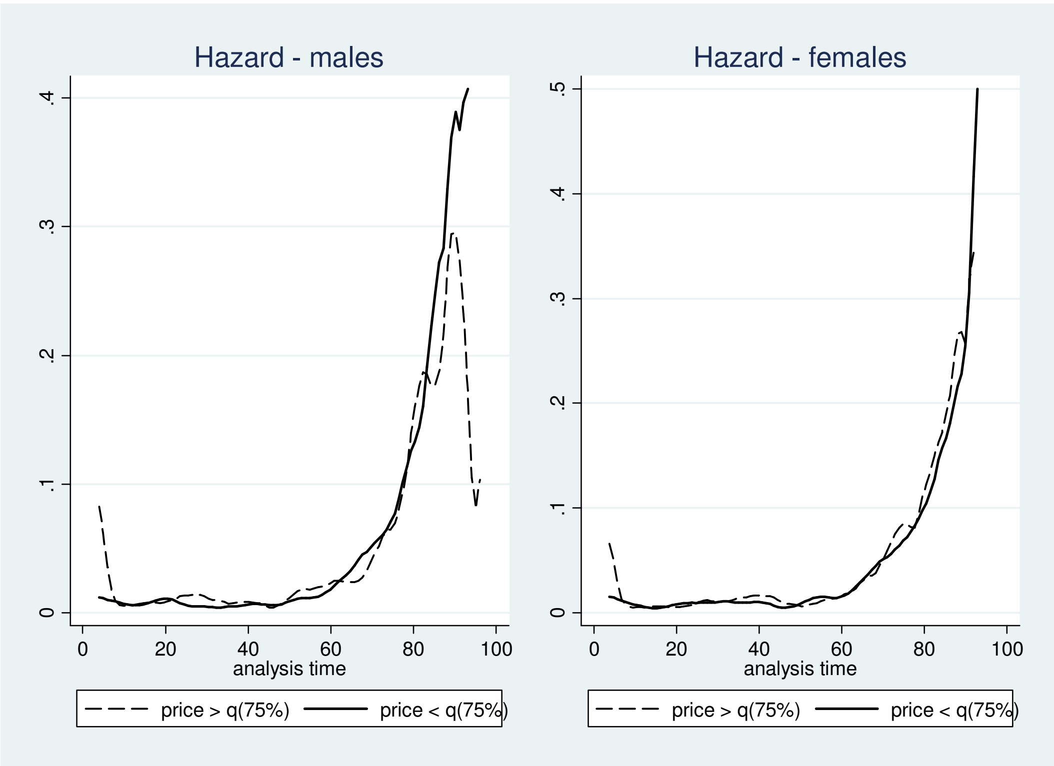


**Figure 10 Hazard rate per level of potato price during age 1-5, conditional on survival until age 1**

**Table 2 P-value of the log rank test for equality of survivor functions**

| | | male | female | both | strata |
|---|---|---|---|---|---|
| In utero and first 12 months of life | | | | | |
| Kcal | <25% or >25% | 0.2184 | 0.8935 | 0.3110 | 0.5699 |
| | <25% or >75% | 0.6046 | 0.8953 | 0.7301 | 0.7386 |
| Potato p | > or <75% | 0.0667 | 0.5138 | 0.3711 | 0.2016 |
| | <25% or >75% | 0.1906 | 0.3997 | 0.1906 | 0.0494 |
| Age 1-5 | | | | | |
| Kcal | <25% or >25% | 0.5209 | 0.5751 | 0.9728 | 0.5389 |
| | <25% or >75% | 0.0874 | 0.0298 | 0.0060 | 0.5050 |
| Potato p | > or <75% | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | <25% or >75% | 0.0000 | 0.0000 | 0.0000 | 0.0027 |

## 5 Parametric analysis

In this section we use parametric models to analyze effects of nutritional conditions early in life on survival over time. By using Cox proportional hazard models we examine how calorie intake or food prices during utero and early life affect individual hazard rate, so the probability of dying at different ages.

We analyze effects of nutritional conditions early in life, proxied by average calorie availability or potato prices during utero and first 12 months of life and during age 1-5 on mortality. In order to identify the effects for each gender without losing the number of observations, we interact each explanatory variable with a dummy for a given gender (using this approach, we have to assume the same baseline hazard for both genders). All regressions were run first on the whole sample, without correcting for family specific effects. Then we stratified our sample by family membership and the regressions were estimated by the stratified partial likelihood method (results in the second columns). Note that if we stratify the sample, then the individuals without any brothers or sisters (born in 1840-60) do not contribute to the partial likelihood function anymore. There are 321 such individuals. Hence, the actual number of observations used by regressions with stratification is actually lower by 321 than reported in the table and it is more difficult to obtain significant estimates. We have also tried to run the regressions controlling for the fact if an individual was born in a family of farmers, in a rural or urban area or if his family had a high or low social status, but these controls were all not significant and the other coefficients did not change much.

First, we analyze effects of early life conditions probability of dying at all ages. Hence, for now we do not distinguish between effects on mortality during early and later life. In this sense our parametric analysis measures the same effects as nonparametric (and it leads to similar conclusions). Table 3 contains the results of the regressions that analyze effects of average calorie availability or potato prices during utero and first 12 months of life and during age 1-5 on mortality. Because people who died before age 1 were not affected by conditions during 1-5 years after their birth, we interact the number of calories available/level of potato prices at ages 1-5 (for each gender) with the indicator, if a person survived until age 1. Recall that if a person died after age 1 but before age 5, we calculate the average over the period between age 1 and the moment of death of that person.

**Table 3 Total effects of nutrition conditions early in life on mortality**

| | | Calories available | | Potato prices | |
|---|---|---|---|---|---|
| | | | strata | | strata |
| males | utero-01 | -0.000291 | -0.00110 | -0.422*** | -0.348*** |
| | | (-0.57) | (-1.27) | (-10.11) | (-4.59) |
| | age 1-5 | -0.000587 | -0.00171 | -0.925*** | -0.651*** |
| | | (-0.79) | (-1.34) | (-17.87) | (-7.19) |
| | utero-01 | -0.000233 | -0.000185 | -0.422*** | -0.322*** |
| | | (-0.45) | (-0.22) | (-10.23) | (-4.45) |
| females | age 1-5 | -0.000225 | -0.000369 | -0.934*** | -0.677*** |
| | | (-0.30) | (-0.30) | (-18.35) | (-7.43) |
| | | | | | |
| N | | 2198 | 2198 | 2198 | 2198 |
| loglik | | -10426.0 | -874.5 | -10235.0 | -836.4 |

According to the coefficient estimates, calorie availability during early life does not have a significant effect on the mortality later in life. As far as the prices of potatoes are concerned, people (both males and females) who experienced high prices during utero and first 12 months of their life or at age 1-5, were characterized by increased mortality during their lives. Moreover, prices at ages 1-5 seem to have stronger effect on mortality than prices during utero and first 12 months of life.

Next, we analyze effects of early life conditions on mortality at different ages during life. As before, in order to proxy for the nutrition conditions early in life, we use average levels of potato prices and average deviations from the trend of calories available during utero and until age 1 and at ages 1-5. Table 4 and Table 5 present results of the estimation of regressions in which we allow nutrition conditions early in life to have different effects on mortality at ages 0-10, 10-20, 20-40, 40-60 and 60-80 (analyzing the conditions at age 1-5 we only consider mortality the effects on mortality at age 1-10). To do this, we use the episode splitting technique by splitting each life spell into yearly observations and we interact the average number of calories available or average potato prices observed during utero until age 1 or at age 1-5 with an indicator for a given age range. Moreover, we correct for current nutrition conditions (calculated for each year of life) by using the value of the deviation from the trend of calories available in a given year of life (contemporaneous effect). We only use current calorie availability (measured in deviations from the trend) instead of current levels of potato prices because of many missing values of potato prices for the years at the end of the nineteenth century. The information on yearly averages of calories availability is available until 1912, so until the last cohort in our sample turns 58. We first analyze effects of nutrition conditions in utero until age 1 and at age 1-5 separately, then we estimate regressions taking into account both effects. For each regression we first estimate the basic partial likelihood model, then we estimate a stratified version, using the family membership indicator. Finally, each variable is interacted with a gender specific dummy variable, so that we can analyze gender-specific effects without losing the number of observations in each regression and without making each strata smaller.

Table 4 presents the results of regressions using deviations from the trend of calories available during utero until age 1 and at ages 1-5 on mortality at different ages. According to the results of separate regressions for each exposure periods, females exposed to lower numbers of calories available during utero and first 12 months of life were characterized by increased mortality at ages 0-10, but not at later ages. For males this effect is significant only at 10% significance level, and only if we stratify the sample. Moreover, according to the same regression, males born under low calorie availability were characterized by lower mortality at age 20-40, which can be again a sign of strong selection of the people born during or shortly after the potato famine (however, this result is significant only at 10% significance level). According to the results of the regression using exposure at age 1-5, males and females who experienced low calorie availability at this age, were characterized by increased mortality at age 1-10 (for females this result disappears if we stratify the sample), but there were also no effects on mortality later in life. However, if we correct for exposure in both periods, it turns out that being exposed to low calorie availability in utero and during first 12 months of life or at age 1-5 resulted in increased mortality of males not only during the first 10 years of life, but also at ages 40-60. For females, there is some evidence that the exposure to low food availability during utero and first 12 months of life resulted in strong selection of children who were born (significant at 10% significance level positive effect of exposure to high calorie availability on mortality at age 10-20) and in higher mortality at age 20-40. However, both results disappear if we stratify the sample by the family membership. Finally, females exposed to low calorie availability during age 1-5 are characterized by increased mortality at age 40-60, but this result is also not significant if we stratify the sample.

The results seem to be even more clear if we consider variation in levels of potato prices during early life. According to Table 5, there is a strong evidence for selection of males who were born in a period of high potato prices: they are characterized by lower mortality at age

20-40. Being exposed to high prices during utero and the first 12 months of life results in increased mortality of both males and females at ages 0-10 (significant only if we correct for both exposure periods) and at ages 40-60 (but this result becomes not significant if we stratify the sample). Males and females exposed to high potato prices at ages 1-5 were characterized by higher mortality at ages 1-10, but also at ages 40-60 (this result is significant at 5% level if we do not and at 10% significance level if we do stratify the sample and it disappears if we do not correct for the early exposure period and stratify the sample). Hence there is a strong evidence for long term effects of bad nutritional conditions early in life.

Table 6 shows the comparison of the effect of nutritional conditions early in life for people born in the families of farmers and of non-farmers. We classify each person according to the occupation of his father. We present only the results of the regressions correcting for both exposure periods and using potato prices as an indication for nutritional conditions. According to the results, mortality of children of nonfarmers were affected by nutritional conditions early in life mostly at early ages, whereas children of farmers suffered from increased mortality also later in life, mostly at ages 40-60, as a result of exposure to higher potato prices at age 1-5. Moreover, males born during a period with high potato prices in families of farmers were characterized also by lower mortality at age 20-40 than those born in similar families during periods with lower prices (significant only if we correct for family membership). This result is not significant for children born in families were the father was not a farmer. There is also some evidence (at 10% significance level) of increased mortality at age 40-60 of females who were exposed to high potato prices during utero and first 12 months of life, but this result disappears if we stratify the sample. Finally, note that if we stratify the sample by family membership, than the effects of exposure to high prices in any of the two exposure periods on mortality of children of farmers during the first 10 years of life becomes not significant, while it remains significant for children of nonfarmers. This could suggest that families of farmers could compensate for temporary low food availability with home production, so that during periods with high food prices children would not die immediately. However, in the long term they still suffered from low food availability during early life, while children of nonfarmers did not.

Next, we analyze effects of nutrition during early life on longevity of people with different social statuses. Table 7 presents results of the estimation of Cox regressions on the subpopulations of people coming from high (HISCLASS class lower than 8) or low classes (HISCLASS between 8 and 12). According to the results, for people born in families with a high social status, high food prices during utero and first 12 months of life did not have a significant effect on mortality later in life. Moreover, being exposed to high prices at age 1-5 resulted in increased mortality at age 1-10, but not later in life. This could suggest that families with high social status could provide enough nutrition for pregnant women and newborn children even during periods of high food prices, whereas the nutrition of children at ages 1-5 was more affected by the level of prices. According to the results, children born in families with low social status under high food prices in any of the exposure periods were characterized by increased mortality at age 0-10, but this result becomes insignificant if we correct for family membership. One reason for this can be a high correlation of the social status with father's occupation as a farmer. Moreover, both males and females exposed to high prices at ages 1-5 were characterized by increased mortality at ages 40-60. Females who experienced high potato prices during utero until age 1 were also characterized by higher mortality at ages 40-60 (but higher only at 10% significance level if we stratify the sample), for males this result is significant only at 10% significance level if we do not stratify the sample, and disappears otherwise. Finally, males who experienced bad nutrition conditions during utero and first 12 months of life and who were born in families with low social status were characterized by positive selection, resulting in lower mortality at ages 20-40. To sum up, it is clear that children born in families with low social status were affected more by bad nutritional conditions during early life than people born in families with high social status.

****still to be added; comparison for rural/urban areas of birth

**Table 4 Effects of calorie availability early in life on mortality at different ages later in life**

| | | exposure in utero-01 | strata | exposure at age 1-5 | strata | exposure in utero-01 & 1-5 | strata |
|---|---|---|---|---|---|---|---|
| **males:** | | | | | | | |
| | contemp | -0.0001 | -0.0009 | 0.0003 | -0.0005 | 0.0013* | -0.0000 |
| | (calories) | (-0.20) | (-0.97) | -0.44 | (-0.56) | -1.86 | (-0.01) |
| utero-01 | 0-10 | -0.0007 | -0.0034* | | | -0.0045*** | -0.0059*** |
| | | (-0.64) | (-1.92) | | | (-3.68) | (-2.99) |
| utero-01 | 10-20 | -0.0009 | 0.0007 | | | 0.0000 | 0.0008 |
| | | (-0.56) | -0.34 | | | -0.02 | -0.31 |
| utero-01 | 20-40 | 0.0019 | 0.0033* | | | 0.0007 | 0.0026 |
| | | -1.5 | -1.69 | | | -0.45 | -1.11 |
| utero-01 | 40-60 | -0.0015 | -0.0019 | | | -0.0033** | -0.0048* |
| | | (-1.26) | (-0.98) | | | (-2.14) | (-1.94) |
| utero-01 | 60-80 | -0.0004 | -0.0040 | | | -0.0013 | -0.0070 |
| | | (-0.24) | (-1.21) | | | (-0.56) | (-1.59) |
| age 1-5 | 1-10 | | | -0.0121*** | -0.0069** | -0.0166*** | -0.0104*** |
| | | | | (-5.80) | (-2.22) | (-7.79) | (-3.10) |
| age 1-5 | 10-20 | | | -0.0001 | -0.0021 | 0.0010 | -0.0009 |
| | | | | (-0.03) | (-0.64) | -0.34 | (-0.25) |
| age 1-5 | 20-40 | | | -0.0017 | -0.0027 | -0.0025 | -0.0013 |
| | | | | (-0.95) | (-1.11) | (-1.10) | (-0.42) |
| age 1-5 | 40-60 | | | -0.0006 | -0.0038 | -0.0039* | -0.0073** |
| | | | | (-0.33) | (-1.26) | (-1.69) | (-1.96) |
| age 1-5 | 60-80 | | | 0.0002 | -0.0011 | -0.0017 | -0.0064 |
| | | | | -0.08 | (-0.19) | (-0.49) | (-0.81) |
| **females:** | | | | | | | |
| | contemp | 0.0006 | 0.0013 | 0.0002 | 0.0009 | 0.0013*** | 0.0000** |
| | (calories) | -0.99 | -1.36 | -0.37 | -0.93 | -2.67 | -2.01 |
| utero-01 | 0-10 | -0.0032*** | -0.0038** | | | -0.0071*** | -0.0060*** |
| | | (-3.11) | (-2.34) | | | (-5.96) | (-3.33) |
| utero-01 | 10-20 | 0.0023 | 0.0017 | | | 0.0038* | 0.0027 |
| | | -1.26 | -0.79 | | | -1.76 | -0.97 |
| utero-01 | 20-40 | -0.0018 | 0.0004 | | | -0.0032** | -0.0001 |
| | | (-1.57) | -0.27 | | | (-2.21) | (-0.07) |
| utero-01 | 40-60 | -0.0005 | 0.0021 | | | -0.0024 | -0.0004 |
| | | (-0.41) | -1.08 | | | (-1.45) | (-0.15) |
| utero-01 | 60-80 | -0.0005 | 0.0064 | | | -0.0014 | 0.0031 |
| | | (-0.27) | -1.51 | | | (-0.59) | -0.59 |
| age 1-5 | 1-10 | | | -0.0088*** | -0.0039 | -0.0146*** | -0.0091*** |
| | | | | (-3.77) | (-1.26) | (-6.19) | (-2.70) |
| age 1-5 | 10-20 | | | 0.0009 | -0.0001 | 0.0037 | 0.0021 |
| | | | | -0.39 | (-0.02) | -1.19 | -0.54 |
| age 1-5 | 20-40 | | | -0.0009 | -0.0014 | -0.0033 | -0.0011 |
| | | | | (-0.56) | (-0.57) | (-1.52) | (-0.37) |
| age 1-5 | 40-60 | | | -0.0011 | -0.0018 | -0.0039* | -0.0035 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | (-0.60) | (-0.67) | (-1.66) | (-0.99) |
| age 1-5 | 60-80 | | | 0.0001 | -0.0048 | -0.0012 | -0.0058 |
| | | | | -0.04 | (-0.82) | (-0.39) | (-0.75) |
| | N | 75005 | 75005 | 75005 | 75005 | 75005 | 75005 |
| | loglik | -6229.2 | -610.5 | -6218.5 | -614.3 | -6190.5 | -600 |

**Table 5 Effects of potato prices early in life on mortality at different ages later in life**

| | | exposure in utero-01 | | exposure at age 1-5 | | exposure in utero-01 & 1-5 | |
|---|---|---|---|---|---|---|---|
| | | | strata | | strata | | strata |
| males: | | | | | | | |
| | contemp | -0.0004 | -0.0011 | 0.0003 | -0.0008 | 0.0006 | -0.0005 |
| | (calories) | (-0.54) | (-1.25) | (0.48) | (-0.89) | (0.88) | (-0.46) |
| utero-01 | 0-10 | -0.0314 | 0.174 | | | 0.284*** | 0.405** |
| | | (-0.35) | (1.18) | | | (2.86) | (2.28) |
| utero-01 | 11-20 | 0.0337 | -0.0391 | | | 0.0150 | -0.107 |
| | | (0.29) | (-0.26) | | | (0.08) | (-0.41) |
| utero-01 | 20-40 | -0.181** | -0.318** | | | -0.168 | -0.560** |
| | | (-2.00) | (-2.37) | | | (-1.07) | (-2.36) |
| utero-01 | 40-60 | 0.0150 | -0.0691 | | | 0.361** | 0.230 |
| | | (0.17) | (-0.50) | | | (2.43) | (0.90) |
| utero-01 | 60-80 | 0.0854 | 0.0301 | | | 0.144 | 0.490 |
| | | (0.71) | (0.13) | | | (0.66) | (1.17) |
| age 1-5 | 1-10 | | | 1.217*** | 0.471 | 1.594*** | 0.896*** |
| | | | | (5.85) | (1.61) | (7.37) | (2.62) |
| age 1-5 | 11-20 | | | 0.0300 | 0.110 | -0.0865 | -0.209 |
| | | | | (0.14) | (0.41) | (-0.25) | (-0.48) |
| age 1-5 | 20-40 | | | 0.151 | 0.166 | 0.321 | -0.0684 |
| | | | | (0.96) | (0.80) | (1.22) | (-0.19) |
| age 1-5 | 40-60 | | | 0.285* | 0.487* | 0.821*** | 0.856* |
| | | | | (1.76) | (1.90) | (2.97) | (1.90) |
| age 1-5 | 60-80 | | | -0.113 | 0.637 | 0.0850 | 0.868 |
| | | | | (-0.43) | (1.19) | (0.18) | (0.90) |
| females: | | | | | | | |
| | contemp | 0.0002 | 0.0010 | 0.0003 | 0.0011 | 0.0012* | 0.0016 |
| | (calories) | (0.30) | (1.13) | (0.53) | (1.22) | (1.78) | (1.60) |
| utero-01 | 0-10 | 0.0387 | 0.179 | | | 0.393*** | 0.408** |
| | | (0.44) | (1.26) | | | (4.04) | (2.41) |
| utero-01 | 11-20 | -0.0666 | -0.151 | | | -0.113 | -0.255 |
| | | (-0.55) | (-0.95) | | | (-0.54) | (-0.96) |
| utero-01 | 20-40 | -0.0298 | -0.164 | | | 0.121 | -0.167 |
| | | (-0.35) | (-1.39) | | | (0.84) | (-0.81) |
| utero-01 | 40-60 | 0.0107 | -0.162 | | | 0.346** | 0.244 |
| | | (0.12) | (-1.12) | | | (2.29) | (0.94) |
| utero-01 | 60-80 | 0.0589 | -0.294 | | | 0.0699 | -0.131 |
| | | (0.48) | (-1.23) | | | (0.31) | (-0.29) |
| age 1-5 | 1-10 | | | 1.220*** | 0.479 | 1.511*** | 0.896** |
| | | | | (5.85) | (1.60) | (6.93) | (2.58) |
| age 1-5 | 11-20 | | | -0.0602 | 0.0170 | -0.0581 | -0.175 |
| | | | | (-0.28) | (0.06) | (-0.17) | (-0.40) |
| age 1-5 | 20-40 | | | 0.264* | 0.262 | 0.178 | -0.313 |
| | | | | (1.70) | (1.25) | (0.69) | (-0.90) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| age 1-5 | 40-60 | | | 0.285* | 0.379 | 0.832*** | 0.733* |
| | | | | (1.77) | (1.52) | (3.02) | (1.66) |
| age 1-5 | 60-80 | | | -0.112 | 0.386 | 0.133 | 1.184 |
| | | | | (-0.43) | (0.73) | (0.29) | (1.22) |
| | N | 75005 | 75005 | 75005 | 75005 | 75005 | 75005 |
| | loglik | -6229.6 | -610.9 | -6213.7 | -612.2 | -6198.6 | -603.5 |

**Table 6 Long run effects of nutrition conditions proxied by potato prices by father's occupation**

| | | Non farmers | | farmers | |
|---|---|---|---|---|---|
| | | | strata | | strata |
| males: | | | | | |
| | contemp | 0.0008 | -0.0013 | 0.0007 | 0.0000 |
| | (calories) | (0.63) | (-0.76) | (0.76) | (0.01) |
| utero-01 | 0-10 | 0.269 | 0.689** | 0.327** | 0.256 |
| | | (1.61) | (2.35) | (2.43) | (1.09) |
| utero-01 | 11-20 | -0.232 | -0.471 | 0.0514 | -0.00719 |
| | | (-0.66) | (-0.95) | (0.19) | (-0.02) |
| utero-01 | 20-40 | -0.262 | -0.393 | -0.175 | -0.577** |
| | | (-0.86) | (-0.79) | (-0.85) | (-2.01) |
| utero-01 | 40-60 | 0.271 | -0.398 | 0.319 | 0.413 |
| | | (0.95) | (-0.84) | (1.62) | (1.26) |
| utero-01 | 60-80 | -0.324 | -0.236 | 0.360 | 1.076 |
| | | (-0.84) | (-0.34) | (1.27) | (1.64) |
| age 1-5 | 1-10 | 1.451*** | 1.081** | 1.805*** | 0.811* |
| | | (4.23) | (2.00) | (6.42) | (1.72) |
| age 1-5 | 11-20 | -0.682 | -1.241 | 0.270 | -0.0244 |
| | | (-1.12) | (-1.52) | (0.56) | (-0.04) |
| age 1-5 | 20-40 | 0.0250 | -0.416 | 0.368 | 0.183 |
| | | (0.05) | (-0.57) | (1.06) | (0.40) |
| age 1-5 | 40-60 | 0.157 | -0.817 | 0.932** | 1.466** |
| | | (0.30) | (-0.86) | (2.54) | (2.42) |
| age 1-5 | 60-80 | -0.152 | 0.102 | 0.186 | 0.528 |
| | | (-0.19) | (0.07) | (0.30) | (0.38) |
| females: | | | | | |
| | contemp | 0.0011 | 0.0022 | 0.0018** | 0.0017 |
| | (calories) | (0.95) | (1.27) | (2.05) | (1.36) |
| utero-01 | 0-10 | 0.290* | 0.567** | 0.503*** | 0.357 |
| | | (1.77) | (2.06) | (3.85) | (1.56) |
| utero-01 | 11-20 | -0.495 | -0.767 | 0.0138 | -0.0953 |
| | | (-1.31) | (-1.46) | (0.05) | (-0.25) |
| utero-01 | 20-40 | -0.0969 | -0.210 | 0.151 | -0.0644 |
| | | (-0.37) | (-0.57) | (0.77) | (-0.23) |
| utero-01 | 40-60 | 0.0735 | -0.516 | 0.394* | 0.570 |
| | | (0.26) | (-1.02) | (1.92) | (1.60) |
| utero-01 | 60-80 | -0.208 | -0.363 | 0.165 | -0.392 |
| | | (-0.56) | (-0.55) | (0.55) | (-0.54) |
| age 1-5 | 1-10 | 1.464*** | 1.199** | 1.674*** | 0.740 |
| | | (4.26) | (2.19) | (5.86) | (1.53) |
| age 1-5 | 11-20 | -0.577 | -1.190 | 0.161 | -0.101 |
| | | (-0.97) | (-1.47) | (0.33) | (-0.17) |
| age 1-5 | 20-40 | 0.0388 | -0.366 | 0.155 | -0.223 |
| | | (0.08) | (-0.55) | (0.45) | (-0.51) |
| age 1-5 | 40-60 | 0.378 | -0.887 | 0.880** | 1.298** |

| | | | | | |
|---|---|---|---|---|---|
| | | (0.76) | (-1.03) | (2.42) | (2.24) |
| age 1-5 | 60-80 | -0.349 | 0.0620 | 0.350 | 1.484 |
| | | (-0.44) | (0.04) | (0.58) | (1.08) |
| | | | | | |
| | N | 23672 | 23672 | 43031 | 43031 |
| | loglik | -1636.6 | -199.5 | -3219.0 | -352.8 |

**Table 7 Long run effects of nutrition conditions proxied by potato prices by social status**

| | | High class | strata | Low class | strata |
|---|---|---|---|---|---|
| males: | | | | | |
| | contemp | 0.0008 | -0.0016 | 0.0007 | -0.0001 |
| | (calories) | (0.57) | (-0.77) | (0.76) | (-0.11) |
| utero-01 | 0-10 | 0.167 | 0.575* | 0.370*** | 0.311 |
| | | (0.94) | (1.83) | (2.85) | (1.37) |
| utero-01 | 11-20 | -0.194 | -0.430 | -0.00500 | -0.0875 |
| | | (-0.51) | (-0.78) | (-0.02) | (-0.26) |
| utero-01 | 20-40 | -0.153 | -0.00328 | -0.229 | -0.672** |
| | | (-0.41) | (-0.01) | (-1.19) | (-2.44) |
| utero-01 | 40-60 | 0.322 | -0.306 | 0.339* | 0.371 |
| | | (1.04) | (-0.61) | (1.76) | (1.15) |
| utero-01 | 60-80 | -0.419 | -0.282 | 0.368 | 0.984 |
| | | (-1.06) | (-0.38) | (1.31) | (1.58) |
| age 1-5 | 1-10 | 1.251*** | 1.392** | 1.860*** | 0.706 |
| | | (3.22) | (2.29) | (7.10) | (1.57) |
| age 1-5 | 11-20 | -0.514 | -0.844 | 0.0881 | -0.343 |
| | | (-0.77) | (-0.96) | (0.19) | (-0.59) |
| age 1-5 | 20-40 | 0.759 | 0.489 | 0.139 | -0.0457 |
| | | (1.22) | (0.55) | (0.43) | (-0.11) |
| age 1-5 | 40-60 | -0.306 | -1.855 | 1.020*** | 1.556*** |
| | | (-0.54) | (-1.60) | (2.85) | (2.62) |
| age 1-5 | 60-80 | -0.530 | -0.558 | 0.368 | 0.725 |
| | | (-0.65) | (-0.31) | (0.62) | (0.53) |
| females: | | | | | |
| | contemp | -0.0003 | 0.0009 | 0.0024*** | 0.0023* |
| | (calories) | (-0.20) | (0.48) | (2.79) | (1.87) |
| utero-01 | 0-10 | 0.111 | 0.398 | 0.577*** | 0.431* |
| | | (0.63) | (1.38) | (4.61) | (1.93) |
| utero-01 | 11-20 | -0.385 | -0.525 | -0.0989 | -0.289 |
| | | (-0.98) | (-0.96) | (-0.34) | (-0.79) |
| utero-01 | 20-40 | 0.111 | 0.199 | 0.0775 | -0.174 |
| | | (0.36) | (0.44) | (0.42) | (-0.68) |
| utero-01 | 40-60 | -0.114 | -0.875 | 0.451** | 0.582* |
| | | (-0.39) | (-1.49) | (2.25) | (1.67) |
| utero-01 | 60-80 | -0.251 | -0.522 | 0.161 | -0.435 |
| | | (-0.65) | (-0.75) | (0.55) | (-0.58) |
| age 1-5 | 1-10 | 1.324*** | 1.474** | 1.705*** | 0.654 |
| | | (3.45) | (2.45) | (6.35) | (1.41) |
| age 1-5 | 11-20 | -0.452 | -0.954 | 0.0159 | -0.340 |
| | | (-0.70) | (-1.07) | (0.03) | (-0.58) |
| age 1-5 | 20-40 | 0.700 | 0.549 | -0.0542 | -0.446 |
| | | (1.19) | (0.65) | (-0.17) | (-1.09) |

| | | | | | |
|---|---|---|---|---|---|
| age 1-5 | 40-60 | 0.247 | -1.377 | 0.902** | 1.317** |
| | | (0.48) | (-1.43) | (2.53) | (2.29) |
| age 1-5 | 60-80 | -0.777 | -0.335 | 0.528 | 1.498 |
| | | (-0.94) | (-0.18) | (0.88) | (1.12) |
| | | | | | |
| | N | 19359 | 19359 | 47344 | 47344 |
| | loglik | -1324.5 | -165.1 | -3558.0 | -383.4 |

# 6 Sensitivity analysis

In this section we check sensitivity of our results.

First, we perform a sensitivity check for potato prices: we reestimate the model correcting for conditions both during utero and early life using not levels, but cyclical components of potato prices. The results, presented in Table 8, are very similar to the results we obtained by using levels of prices.

Next, we reestimate our model examining effects of early life conditions on mortality at different ages, using prices of wheat and rye. We have plotted the graphs of the trend components of wheat and rye prices (we do not report them here) and there is a clear increasing trend component in them. Therefore, we base our analysis on cyclical components. Table 8, which reports the results of the estimation, confirms that there are strong effects of high food prices during utero and early life on increased mortality at early ages. Moreover, there is a clear evidence of higher mortality of males exposed to high prices at ages 1-5.

Next, we look at the composition of a diet, by reestimating the model using the amounts of proteins and proteins from animals available early in life as proxies for nutritional conditions. According to Table 9, again, there is a strong evidence for increased mortality at early ages of people exposed to low amount of proteins or proteins from animals during utero and first 12 months of life or during age 1-5. Moreover, the results suggest that males exposed to low amounts of proteins or proteins from animals during age 1-5 are characterized by higher mortality at ages 40-60 (the result is significant only if we correct for family membership).

[Decomposition of fixed effects]

**Table 8 Sensitivity analysis - cyclical deviations of prices of potatoes, rye and wheat**

| | | Potato prices | | Rye prices | | Wheat prices | |
|---|---|---|---|---|---|---|---|
| | | | strata | | strata | | strata |
| **males:** | | | | | | | |
| | contemp | 0.000723 | -0.00034 | 0.000391 | -0.000578 | 0.000162 | -0.000867 |
| | (calories) | (1.09) | (-0.34) | (0.62) | (-0.60) | (0.26) | (-0.90) |
| utero-01 | 0-10 | 0.221* | 0.534** | 0.0722 | 0.236*** | 0.0311 | 0.140** |
| | | (1.71) | (2.32) | (1.41) | (2.79) | (0.73) | (2.02) |
| utero-01 | 10-20 | 0.0856 | -0.0474 | -0.0248 | 0.00590 | -0.0133 | 0.0159 |
| | | (0.45) | (-0.18) | (-0.25) | (0.05) | (-0.15) | (0.15) |
| utero-01 | 20-40 | -0.234 | -0.475* | -0.0983 | -0.146 | -0.0948 | -0.127 |
| | | (-1.40) | (-1.83) | (-1.31) | (-1.35) | (-1.41) | (-1.36) |
| utero-01 | 40-60 | 0.319** | 0.204 | 0.00891 | 0.151 | -0.0212 | 0.0944 |
| | | (2.21) | (0.83) | (0.12) | (1.26) | (-0.33) | (0.89) |
| utero-01 | 60-80 | 0.148 | 0.578 | 0.242 | 0.407 | 0.228 | 0.325 |
| | | (0.67) | (1.44) | (1.55) | (1.34) | (1.55) | (1.21) |
| age 1-5 | 1-10 | 1.557*** | 0.933** | 0.443*** | 0.490*** | 0.257*** | 0.389*** |
| | | (6.93) | (2.34) | (4.72) | (3.47) | (3.43) | (3.53) |
| age 1-5 | 10-20 | 0.0499 | -0.0729 | -0.00919 | 0.0778 | 0.0112 | 0.0595 |
| | | (0.16) | (-0.18) | (-0.07) | (0.42) | (0.10) | (0.40) |
| age 1-5 | 20-40 | 0.134 | 0.00372 | -0.00773 | -0.0883 | -0.0362 | -0.118 |
| | | (0.55) | (0.01) | (-0.08) | (-0.61) | (-0.44) | (-0.99) |
| age 1-5 | 40-60 | 0.663*** | 0.771* | 0.0759 | 0.384** | 0.0684 | 0.323** |
| | | (2.71) | (1.87) | (0.75) | (2.38) | (0.83) | (2.44) |
| age 1-5 | 60-80 | 0.0849 | 0.944 | 0.246 | 0.314 | 0.232 | 0.236 |
| | | (0.21) | (1.12) | (1.14) | (0.63) | (1.10) | (0.55) |
| **females:** | | | | | | | |
| | contemp | 0.00137** | 0.00160 | 0.00107* | 0.00108 | 0.000799 | 0.000750 |
| | (calories) | (2.08) | (1.64) | (1.68) | (1.14) | (1.28) | (0.80) |
| utero-01 | 0-10 | 0.642*** | 0.460** | 0.175*** | 0.181** | 0.119*** | 0.121** |
| | | (5.75) | (2.24) | (3.75) | (2.57) | (3.12) | (2.06) |
| utero-01 | 10-20 | -0.213 | -0.332 | -0.159 | -0.142 | -0.113 | -0.125 |
| | | (-0.95) | (-1.13) | (-1.61) | (-1.13) | (-1.29) | (-1.14) |
| utero-01 | 20-40 | 0.209 | -0.0137 | 0.0772 | 0.0122 | 0.0644 | 0.0233 |
| | | (1.51) | (-0.07) | (1.09) | (0.14) | (1.09) | (0.31) |
| utero-01 | 40-60 | 0.267* | 0.0158 | -0.0700 | -0.123 | -0.0736 | -0.136 |
| | | (1.74) | (0.06) | (-0.92) | (-1.05) | (-1.10) | (-1.30) |
| utero-01 | 60-80 | 0.0630 | -0.518 | 0.109 | -0.495 | 0.108 | -0.523 |
| | | (0.27) | (-1.04) | (0.69) | (-1.22) | (0.72) | (-1.32) |
| age 1-5 | 1-10 | 1.377*** | 0.860* | 0.435*** | 0.433*** | 0.290*** | 0.320*** |
| | | (6.08) | (1.96) | (4.51) | (3.14) | (3.87) | (3.02) |
| age 1-5 | 10-20 | -0.240 | -0.304 | -0.268* | -0.145 | -0.235** | -0.137 |
| | | (-0.71) | (-0.67) | (-1.87) | (-0.77) | (-1.98) | (-0.88) |
| age 1-5 | 20-40 | 0.325 | -0.0447 | 0.0550 | -0.0267 | 0.0441 | -0.0397 |
| | | (1.41) | (-0.14) | (0.56) | (-0.20) | (0.55) | (-0.38) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| age 1-5 | 40-60 | 0.608** | 0.273 | -0.0553 | -0.0302 | -0.0658 | -0.0725 |
| | | (2.43) | (0.68) | (-0.53) | (-0.18) | (-0.77) | (-0.53) |
| age 1-5 | 60-80 | 0.0948 | 0.278 | 0.166 | 0.132 | 0.182 | -0.0190 |
| | | (0.23) | (0.30) | (0.79) | (0.27) | (0.88) | (-0.04) |
| | N | 75005 | 75005 | 75005 | 75005 | 75005 | 75005 |
| | loglik | -6195.0 | -605.8 | -6213.3 | -595.6 | -6218.9 | -596.7 |

**Table 9 Effects of diet composition on mortality at different ages**

| | | Proteins | | Proteins from animals | |
|---|---|---|---|---|---|
| | | | strata | | strata |
| **males:** | | | | | |
| | contemp | 0.00101 | -0.000324 | 0.000150 | -0.000848 |
| | (calories) | (1.53) | (-0.33) | (0.24) | (-0.89) |
| utero-01 | 0-10 | -0.0931*** | -0.107* | -0.0362 | -0.0388 |
| | | (-2.79) | (-1.94) | (-0.79) | (-0.50) |
| utero-01 | 10-20 | 0.00242 | 0.0281 | 0.0334 | 0.0483 |
| | | (0.04) | (0.41) | (0.38) | (0.43) |
| utero-01 | 20-40 | 0.0218 | 0.0381 | 0.0517 | 0.0349 |
| | | (0.50) | (0.60) | (0.80) | (0.38) |
| utero-01 | 40-60 | -0.0521 | -0.0720 | -0.00833 | -0.0300 |
| | | (-1.21) | (-1.07) | (-0.13) | (-0.28) |
| utero-01 | 60-80 | -0.0265 | -0.198 | -0.0372 | -0.296 |
| | | (-0.39) | (-1.61) | (-0.30) | (-1.37) |
| age 1-5 | 1-10 | -0.470*** | -0.292*** | -0.331*** | -0.366*** |
| | | (-7.44) | (-3.09) | (-3.28) | (-2.72) |
| age 1-5 | 10-20 | 0.0312 | 0.000137 | 0.0448 | 0.0458 |
| | | (0.39) | (0.00) | (0.39) | (0.32) |
| age 1-5 | 20-40 | -0.0752 | -0.0710 | -0.0394 | -0.0134 |
| | | (-1.23) | (-0.81) | (-0.46) | (-0.11) |
| age 1-5 | 40-60 | -0.0868 | -0.187* | -0.0832 | -0.307** |
| | | (-1.40) | (-1.85) | (-0.97) | (-2.18) |
| age 1-5 | 60-80 | -0.0228 | -0.0989 | 0.0189 | -0.0900 |
| | | (-0.26) | (-0.49) | (0.14) | (-0.32) |
| **females:** | | | | | |
| | contemp | 0.00144** | 0.00173* | 0.000516 | 0.000900 |
| | (calories) | (2.19) | (1.78) | (0.84) | (0.97) |
| utero-01 | 0-10 | -0.158*** | -0.119** | -0.130*** | -0.0753 |
| | | (-4.99) | (-2.44) | (-2.80) | (-1.07) |
| utero-01 | 10-20 | 0.0748 | 0.0701 | 0.0370 | 0.0647 |
| | | (1.29) | (0.93) | (0.43) | (0.58) |
| utero-01 | 20-40 | -0.0924** | -0.0179 | -0.0825 | -0.00743 |
| | | (-2.30) | (-0.35) | (-1.33) | (-0.09) |
| utero-01 | 40-60 | -0.0416 | 0.0505 | 0.00131 | 0.131 |
| | | (-0.94) | (0.76) | (0.02) | (1.32) |
| utero-01 | 60-80 | -0.0263 | 0.201 | -0.00588 | 0.452 |
| | | (-0.38) | (1.25) | (-0.05) | (1.35) |
| age 1-5 | 1-10 | -0.419*** | -0.262*** | -0.268*** | -0.310** |
| | | (-6.24) | (-2.65) | (-2.62) | (-2.34) |
| age 1-5 | 10-20 | 0.104 | 0.0568 | 0.155 | 0.0464 |
| | | (1.22) | (0.53) | (1.37) | (0.31) |
| age 1-5 | 20-40 | -0.0890 | -0.0157 | -0.0801 | 0.0437 |
| | | (-1.51) | (-0.19) | (-0.96) | (0.41) |

| | | | | | |
|---|---|---|---|---|---|
| age 1-5 | 40-60 | -0.0737 | -0.0212 | 0.0109 | 0.125 |
| | | (-1.16) | (-0.21) | (0.12) | (0.91) |
| age 1-5 | 60-80 | -0.0287 | -0.133 | -0.0431 | -0.150 |
| | | (-0.34) | (-0.61) | (-0.30) | (-0.41) |
| | N | 75005 | 75005 | 75005 | 75005 |
| | loglik | -6194.3 | -602.3 | -6223.7 | -605.1 |

# 7 Conclusions

## Bibliography


1. Dominicus, L (2011). Zeeuwen in America (in Dutch), de Buitenspelers.
2. Fokker, AP (1877) De volksvoeding in Zeeland (in Dutch). *Nederlands Tijdschrift voor Geneeskunde* 21:195-240.
3. Hofstee, EW (1978). De demografische ontwikkeling van Nederland in de eerste helft van de negentiende eeuw. Een historisch-demografische en sociologische studie (in Dutch). Van Loghum Slaterus, Deventer.
4. Knibbe, M. (2007) De hoofdelijke beschikbaarheid van voedsel en de levensstandaard in Nederland, 1807-1913 (in Dutch). *Tijdschrift voor sociale en economische geschiedenis* 4:71-107.
5. Kort, Albert Louis (2001), Geen cent te veel. Armoede en armenzorg op Zuid-Beveland 1850-1940. Rotterdam: Erasmus Universiteit
6. Lindeboom, M., Portrait, F., & Van den Berg, G. (2010). Long-run effects on longevity of nutritional shock early in life: the Dutch Potato Famine of 1846-1847. *Journal of Health Economics 29(5)* , 617-629.
7. Meere, J.M.M. de (1982) Economische ontwikkeling en levensstandaard in Nederland gedurende de eerste helft van de negentiende eeuw; Aspecten en Trends, 's-Gravenhage.
8. Oomens, C.A. and Den Bakker, G.P. (1994) 'De beroepsbevolking in Nederland 1849-1990', Supplement bij de sociaal economische maandstatistiek, part 2.
9. Paping, R, V Tassenaar (2007) The consequences of the Potato disease in the Netherlands 1845-1860: a regional approach, in C Ó Gráda, R Paping & E Vanhaute (eds) When the potato failed. Causes and effects of the 'last' European subsistence crisis, 1845-1850. Brepols Publishers, CORN publication series 9, Turnhout.
10. Poppel, F van, E Beekink (2001) De biometer in kaart gebracht: zuigelingen en totale sterftecijfers voor Nederlandse gemeenten 1812-1939, (in Dutch). *Gewina* 24:18-32.
11. Poppel, F van, K. Mandemakers (2002) Het geluk van borstvoeding; regionale en sociaal-economische verschillen in kindersterfte, 1812-1912 (in Dutch). *Index* 9 (10): 28-31.
12. Vanhaute, E (2007) So worthy an example to Ireland. The subsistence and industrial crisis of 1845-1850 in Flanders, in C Ó Gráda, R Paping & E Vanhaute (eds) When

the potato failed. Causes and effects of the 'last' European subsistence crisis, 1845-1850. Brepols Publishers, CORN publication series 9, Turnhout.

13. Vanhaute, E, R Paping, C Ó Gráda (2007) The European Subsistence Crisis of 1845–1850: a Comparative Perspective, in C Ó Gráda, R Paping & E Vanhaute (eds) When the potato failed. Causes and effects of the 'last' European subsistence crisis, 1845-1850. Brepols Publishers, CORN publication series 9, Turnhout.
14. van der Bie, R. J., & Smits, J. P. (2001). *Tweehonderd jaar statistiek in tijdreeksen 1800-1999.* Centraal Bureau voor de Statistiek.
15. van Leeuwen, M., & Maas, I. (2005). A short note on HISCLASS. *Unpublished* .
16. van Leeuwen, M., Maas, I., & Miles, A. (2002). *HISCO: Historical International Standard Classification of Occupations.* Leuven: Leuven University Press.