Inequality in Individual Mortality and Economic Conditions Earlier in Life

Gerard J. van den Berg
Maarten Lindeboom
Marta López

Inequality in Individual Mortality and Economic Conditions Earlier in Life

Gerard J. van den Berg
Free University Amsterdam, Princeton University,
IFAU Uppsala, Netspar, CEPR, IFS and IZA Bonn

Maarten Lindeboom
Free University Amsterdam, Netspar,
University of Bergen, Tinbergen Institute and IZA Bonn

Marta López
Free University Amsterdam

ABSTRACT

Inequality in Individual Mortality and Economic Conditions Earlier in Life[*]

We analyze the effect of being born in a recession on the mortality rate later in life in conjunction with social class. We use individual data records from Dutch registers of birth, marriage, and death certificates, covering the period 1815-2000, and we merge these with historical data on macro-economic outcomes and health indicators. We estimate duration models and inequality measures. The results indicate that being born in a recession increases the mortality rate later in life for most of the population. Lower social classes suffer disproportionally from being born in recessions. This exacerbates mortality inequality. This is not affected by social mobility: upward mobility does not vary much with the business cycle at birth. It turns out that the average long-run economic well-being of the family at birth, the transitory economic shocks at birth, and their interaction, are all relevant determinants of the mortality rate throughout the whole life.

Corresponding author:

Gerard J. van den Berg
Department of Economics
Free University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
E-mail: gjvdberg@xs4all.nl

---

1 Introduction

This paper studies the interaction between social class and the business cycle early in life as determinants of individual mortality later in life. In other words, we study whether the mortality rate later in life of individuals from lower backgrounds is more sensitive to being born in a recession than the mortality rate later in life of individuals from higher backgrounds.

There has recently been an increased interest in the effects of cyclical macro-economic conditions on health and mortality. Most of this literature deals with instantaneous effects of cyclical conditions, using data from relatively recent years (see e.g. Ruhm, 2000, and Dehejia and Lleras-Muney, 2004). Van den Berg, Lindeboom and Portrait (2006) show that being born in a recession causes a higher mortality rate later in life. This finding is consistent with the finding that poor living conditions and low socio-economic status early in life are associated with susceptibility to a range of health problems later in life (see e.g. Case, Fertig and Paxson, 2003, Hayward and Gorman, 2004, and references therein). It is also consistent with results in the recent demographic literature in which natural experiments are applied to test effects of nutrition and disease exposure in utero and during the first months of life on mortality later in life (see Doblhammer, 2004, for a survey).

It is interesting to know whether the long-run mortality effects of being born in a recession depend on the social class of the family. Mortality is often highly dependent on social class. Individuals with lower backgrounds have higher mortality rates. They may also suffer disproportionally from being born in a recession, for essentially two reasons. First, a recession may hit lower-class individuals harder, for example due to a lack of job security or due to an increased susceptibility to epidemics during recessions. Secondly, bad economic conditions at the individual level may have a stronger effect on mortality for lower-class individuals. In all these cases it may be sensible to target policy at infants born in low social-class families and born in recessions. Their mortality later in life could be significantly reduced if their conditions are improved upon, for example by monitoring their health shortly after birth and by providing food, housing, and health care. Indeed, the government may focus macroeconomic stabilization policies on sectors where lower-class workers are over-represented.

Our data set, called the Historical Sample of the Netherlands (HSN), covers around 14,000 individuals born in three large regions in the Netherlands in the period 1812-1912 and followed up to 2000. It has previously been used by e.g. Ganzeboom (2002), Van Poppel and Mandemakers (2003) and Van den Berg,

Lindeboom and Portrait (2006). The data set includes information taken from the standardized individual recordings of vital events (birth, marriage and death) kept by municipalities and provinces, and these also record the social class and occupation of the parents. We merge these with historical time-series data on macro-economic variables, notably GNP, and with external information on the incidence of epidemics. There are no reliable 19th century data on health expenditures and medical innovations, so one can not address their role in the causal chain from economic conditions to health to mortality.

We estimate duration models for individual mortality, where, for a given individual at a given moment in time, the mortality rate depends on current conditions, conditions earlier in life, social class, other characteristics, and, in particular, on the interaction between conditions early in life and social class. Note that recessions at birth can be regarded as unanticipated and exogenous at the individual level, so that the results provide insights into the long run effects of unexpected transitory macro shocks in economic conditions early in life. Similarly, we argue that social class can be regarded as a proxy for average long-run economic well-being of the family (or permanent income). It turns out that, in terms of this terminology, the interaction between permanent and transitory income is a major determinant of mortality. We demonstrate this by using the duration model estimates to decompose inequality in mortality across social classes into different components, including the long-run effect of social class on mortality, and the class-specific effect of early life cyclical conditions on mortality.

To the extent that our data and other empirical studies enable us, we examine whether social mobility during life mitigates adverse conditions early in life. In additional analyses, we study interactions with occupational indicators. After all, individuals in certain occupations may actually benefit from a recession. For example, if the recession goes along with high food prices then farmers may benefit.[1] We also examine how the size and the composition by social class of a birth year cohort changes with the cyclical indicator of the business cycle at birth. Furthermore, we develop and quantify inequality measures to capture the mortality variation across social classes due to variation in the effect of the business cycle early in life on mortality.

For current developing countries, which in certain aspects (like living and economic conditions) could be regarded as similar to the Netherlands in the period evaluated in the present paper, the existing literature has focused on inequalities

---

[1]See Johansson (2002) for an overview of evidence that child mortality depends on the amount of land owned and its interaction with socio-economic status.

in infant and child mortality by household socioeconomic status, since there are typically no long run data registers (see Sastry, 2004). In this sense, our paper aims to help to complement these studies by studying long run mortality effects.

The paper is organized as follows. Section 2 presents the data and discusses variables that we use in the analyses. We also provide descriptive analyses and some preliminary empirical results. Section 3 covers the duration analyses and provides the main results. Section 4 deals with the quantification of inequality, and Section 5 concludes.

2 The data

2.1 Individual records

The HSN data are derived from the registers of birth, marriage, and death certificates (see Mandemakers, 2000, for a general description in English). Currently, we have access to a cleansed sample of 13,718 individuals. This is a random sample of individuals born in the provinces of Utrecht, Friesland and Zeeland[2] between 1812 and 1922. The end of the observation window is December 31, 1999. The data provide the timing of vital life events as well as household and family characteristics that are recorded in the above registers, like the occupation of the father, whether or not the father was illiterate, gender, and geographic location.

The individual lifetime durations are observed in days. If the individual is still alive at the end of 1999 then we do not observe the date of death. We therefore restrict attention to individuals born before 1902. We observe dates of death of migrants out of the provinces. For some individuals in the sample born before 1902, the date of death is missing or has not been collected yet. The rate at which this occurs decreases quickly over time, from 21% for those born in 1812-1821 to around 6% for those born in later cohorts. Occasionally, marriage and/or childbirth dates of such individuals are observed, and in that case we right-censor the lifetime duration at the latest of these dates. Otherwise, the lifetime duration is right-censored at zero, and the individual is discarded from the data. Missing values of explanatory variables lead to an additional loss of individuals from the sample. The sample used below contains 8929 individuals. In Subsection 3.3

---

[2]At the time, the Netherlands had 11 provinces. Our three provinces were jointly rather representative of the Netherlands in terms of economic activity. For each province, a random generator was successively applied to the numbers that were attached to municipalities, years of birth, and rank of birth in the municipality and year. Patterns in aggregate mortality rates in our data are similar to those in national rates.

we examine to what extent the right-censored observations have systematically different characteristics, and we examine the sensitivity of the results with respect to the independence of the right-censoring.

When we select explanatory variables for individual mortality from these data, we restrict attention to characteristics that are realized at birth as opposed to later in life, for the reason that the latter may be endogenous or confounded. In particular, we do not include variables on the occurrence or timing of life events like marriage. The place of residence at birth is translated into a binary urbanization indicator (being 1 iff the individual is born in a city). Below we also consider urbanization indicators with more than two possible values.

2.2 Social class

The original data registers do not contain variables labelled as the social class of the individual or his/her family. We therefore have to derive social class indicators from the available variables. The aim is to obtain an indicator of the family's long-run economic well-being.[3] A family's long-run economic position may affect the sensitivity of mortality later in life to unexpected shocks like recessions early in life. In particular, if the long-run position is good then the family may be better insured against shocks, for example by way of accumulated assets, and the provision of nutrients to infants may then be better safeguarded in case of such shocks. It may also be easier to avoid epidemics.

To proceed, we operationalize social class by using the father's occupation at the birth of the individual of interest. For our purposes it is useful to have a social class indicator with a scale that is hierarchical, i.e. whose levels rank long-run economic well-being. After all, it is unfeasible to estimate models that include dummy variables for each possible occupation, and it is difficult to relate the corresponding estimates to long-run economic well-being.[4] We adopt Van Tulder (1962)'s mapping from occupation into a six-layer hierarchical social class scale. Mandemakers (1987) shows that this scale can be meaningfully applied to

---

[3]One could think of alternative class categorizations. However, in the Netherlands, at the time, there was no caste system and there were no segregated classes of priests, farmers, or rulers. The guild system had been abandoned before. Nobility was quantitatively negligible. Immigrants and Jews were relatively well-integrated, with the exception of gypsies, but these were also small in number. See Wintle (2000) for details.

[4]Moreover, using the non-hierachical dimension of the occupational status may lead to more endogeneity problems than using the hierarchical dimension, as individuals who choose a certain type of occupation may have systematically different degrees of care for their infants, whereas the hierarchical level may be more genetically determined.

nineteenth as well as twentieth century occupations. Using tax records from a small town, he also shows that there is a high degree of consistency across these centuries in terms of the ranking of a number of specific occupations in terms of income. As we shall see, the fact that the scale has only six levels makes it manageable with our data to include separate variables for the most frequently observed levels.

Table 1 provides the original general descriptions of the levels. For each level we give examples of occupations that were relatively common in the nineteenth century. In fact, Van Tulder (1962) originally designed the six social classes to capture different degrees of *prestige* of occupations in the Netherlands. Obviously, the prestige of an occupation does not necessarily correspond to the long-run average economic strength of the persons with this occupation. However, in his detailed survey of the Netherlands in the nineteenth century, Wintle (2000) concludes from a range of case studies that "income was the most important desideratum for social standing" and "the determining criterion for social status and class was *income* (and to a lesser extent wealth)" (where, presumably, income is supposed to reflect long-run average income). We take this as support for our use of the prestige of the occupation as a proxy for social class. This also fits in with the common practice to denote the six levels as lower lower class, higher lower class, lower middle class, higher middle class, lower higher class, and higher higher class (see e.g. Suurenbroek, 2001).

It is difficult to obtain more precise evidence on the relation between occupation and long-run average economic well-being. Our data do not contain income or wealth measures. In the nineteenth century, individual data on income or assets were not collected in a systematic fashion, and many tax records are not accessible. The census data collected throughout the nineteenth century are uninformative on income and assets. We analyzed a small number of historical datasets with information on individual occupation and indicators of income or assets. These are the personal tax registers for the cities of Haarlem in 1742 and Leiden in 1748, and a dataset containing personal taxable income for individuals in Leiden in 1900. All three have some disadvantages. The first two date from before our observation window. At the time, a fraction of individuals was exempt from paying direct taxes on the ground of being too poor, and these individuals are not represented in these datasets. The third dataset only includes individuals who (next to their main occupation) are member of a board of an institution, community, or charity, and these individuals are often relatively wealthy. Finally, the cities of Haarlem and Leiden are in the province of North-Holland which is

Table 1: Social class description

Social class	General description Examples
1	Unskilled laborers: diker, day laborer, dock worker
2	Semi-skilled laborers, low-level clerks: cow milker, beer brewer, farm laborer, gardener, florist, fisherman, woollen sorter, tailor, painter
3	Small old and new middle class, skilled laborers, small farmers and gardeners, clerks and low-level civil servants: potato farmer, barber, baker, shoemaker, smith, shopkeeper, mason, carpenter
4	Farmers and gardeners with average-sized firms, old and new middle class and medium-level civil servants: bailiff, corn dealer, merchant, innkeeper, miller
5	Higher staff, presidents of smaller firms, high-level civil servants, farmers and gardeners with large firms: factory manager, headmaster, infantry captain
6	Self-employed academics, teachers in secondary education, presidents of larger firms and top-level civil servants: auditor, lawyer, pharmacist, surgeon, professor

The general descriptions are based on Van Tulder (1962).

not in our main dataset.

In 1742 and 1748, personal taxes were intended to increase with consumption, current income, and wealth.[5] In Appendix 1 we show that individuals with occupations that score high on the Van Tulder prestige scale (i.e., that belong to a higher social class) also pay higher personal taxes (in 1742 and 1748) and have higher taxable income (in 1900). This clearly supports the sensibility of using our social class measure as a measure of long-run economic well-being of the family.

We also considered alternative mappings from occupation into social class, notably Treiman (1977)'s standardized international occupational prestige scale. This scale is useful for international comparisons but less useful for countries with a high percentage of the labor force engaged in agriculture. It turns out that in our data the occupational Van Tulder level is strongly positively related to the corresponding standardized Treiman level.[6]

Table 2 provides sample statistics for the sample of individuals by social class. The first row reports the total number of observations in the sample, followed by the percentage with censored lifetimes. For those with uncensored lifetimes, the subsequent rows give the sex ratios and the percentages of individuals whose father is not illiterate, as well as features of the lifetime distributions of men, women, and both genders.

In the sequel we denote the six social classes by SC1...SC6. Clearly, the number of individuals in SC5 and SC6 in our data is very small. We therefore merge them with those from SC4 to obtain a social class indicator with 4 categories instead of the original 6. As a sensitivity analysis, we also omit those in SC5 and SC6 from the analysis. In Subsection 2.5 we examine the size and distribution of the social classes of newborns over time.

The correlation between our social class measure and the mortality later in life will inevitably also capture mortality effects of the economic position later in life among the individuals with a given social class value at birth. After all, long-run economic well-being at birth and later in life are most likely positively correlated. Analogously, any mitigation of an early-life adverse macro shock does not necessarily only take place early in life, but could also continue later in life by way of compensatory actions by the individual him- or herself. One obvious route

---

[5]See in particular Diederiks, Noordam and Tjalsma (1985) for detailed descriptions of Leiden in 1748 and for analyses of the 1748 dataset.

[6]One could also construct a social class indicator by ranking occupations by income in the above-mentioned three historical datasets from Leiden and Haarlem. However, these are both relatively large cities, and rural occupations are rarely observed. Moreover, as explained, low-ranking occupations are underrepresented in these datasets.

Table 2: Summary statistics of the sample of individuals

| | Total | Social class | | | | | |
		1	2	3	4	5	6
Number of individuals	8929	3438	1400	2110	1715	175	91
% censored	19	17	19	21	21	20	24
Of the uncensored observations:							
% female	49	48	48	51	48	59	51
% father not illiterate	85	72	89	95	96	99	100
Lifetime duration	39.5	37.9	39.9	40.1	41.9	40.0	38.4
	[38.4]	[32.0]	[39.6]	[40.9]	[49.3]	[30.8]	[30.5]
	(36)	(36)	(35)	(36)	(35)	(35)	(35)
Lifetime duration	38.2	37.2	38.4	38.0	40.7	31.4	42.4
men	[33.0]	[28.9]	[33.1]	[29.8]	[46.7]	[19.3]	[42.5]
	(35)	(36)	(35)	(35)	(35)	(33)	(34)
Lifetime duration	40.9	38.7	41.6	42.0	43.3	45.9	34.5
women	[42.1]	[33.7]	[44.9]	[47.6]	[51.5]	[61.3]	[28.2]
	(36)	(36)	(35)	(37)	(35)	(36)	(36)

Explanatory note: for the durations we report averages, medians (between square brackets) and standard deviations (between parentheses).

is an individual transition to a higher social class by way of education, career, or marriage.

Unfortunately, our data do not not provide observation of social class at dates after birth, or observation of other informative outcomes, with the following exceptions. First, the data contain a subsample of 565 individuals for whom we observe the occupation (i.e. social class) of the individual at his/her death.[7] Secondly, the data contain a subsample of 655 men born in the province of Utrecht for whom we observe the occupation of the individual at his marriage date. In both subsamples, a comparison to the social class at birth provides an observation of social mobility. It turns out that in these data social mobility is rather low and that the amount of downward mobility is of similar magnitude as the amount of upward mobility, although total and upward mobility are larger among later cohorts. This is in agreement to the empirical literature on social mobility in the Netherlands in the nineteenth century, based on marriage certificates and census data (see in particular Van Dijk, Visser and Wolst, 1984, Boonstra and Mandemakers, 1995, and Ganzeboom, 2002). Social mobility patterns were stable throughout most of the nineteenth century. The earlier stages of industrialization towards the end of the century do substantially increase occupational mobility, but the corresponding class transitions are more often downward than upward. Only cohorts born after our observation window benefit on average from upward mobility due to industrialization. In Subsection 3.2 we examine whether social mobility in these subsamples depends on cyclical economic conditions in the birth year.

A third subsample of 575 women born in the province of Utrecht contains observations of the occupation of the woman at the date of marriage or (in the much more frequent case when she does not have an occupation) of her husband at that date. Upward mobility by way of marriage may have a different effect on the long run effect of early life conditions than upward mobility by way of career or education, which is why we discuss this in Subsection 3.3 when we report gender-specific model estimates.

---

[7]After the early stages of data collection it was decided to abandon the recording of this variable. Among individuals for whom we observe it, relatively few are born towards the end of the observation window. Also, most are male.
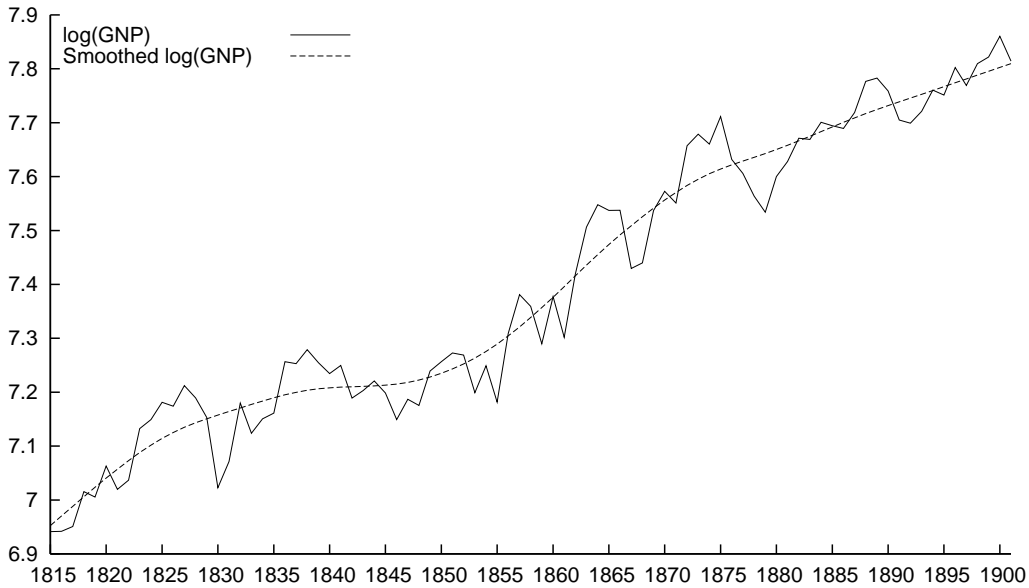
Figure 1: *log annual real per capita GNP.*

2.3 Data on macro-economic conditions, business cycles, and historical events

We merge the individual data records with external information. Most importantly, we use historical time-series data on annual GNP over the observation window. Our choice for GNP instead of obvious alternatives such as GDP is driven by the need for mutually consistent observations for as many years as possible. Since GNP is unavailable for the years 1812-1814, we discard these years, and so the paper will focus on individuals who were born in cohort years starting from 1815. Figure 1 plots the log annual real per capita GNP over the interval in which the sample members are born. Annual real per capita GNP is measured in 1,000 Euros with 1995 as a base year. Clearly, in addition to the upward trend, there are many cyclical fluctuations. Jacobs and Smits (2001) provide a detailed analysis of GDP movements in the Netherlands in the 19th century. Years with low and negative growth are observed more frequently than in the 20th century. The GDP fluctuations are strongly correlated to the business cycles in the UK and U.S.. We also consider other macro-economic indicators, like the inflation rate (see also Smits, Horlings and Van Zanden, 2000) and the share of agricultural production in annual GNP.

Ideally one would like to compare cohorts born in booms to those born in recessions with otherwise identical circumstances throughout life. This is unfeasible
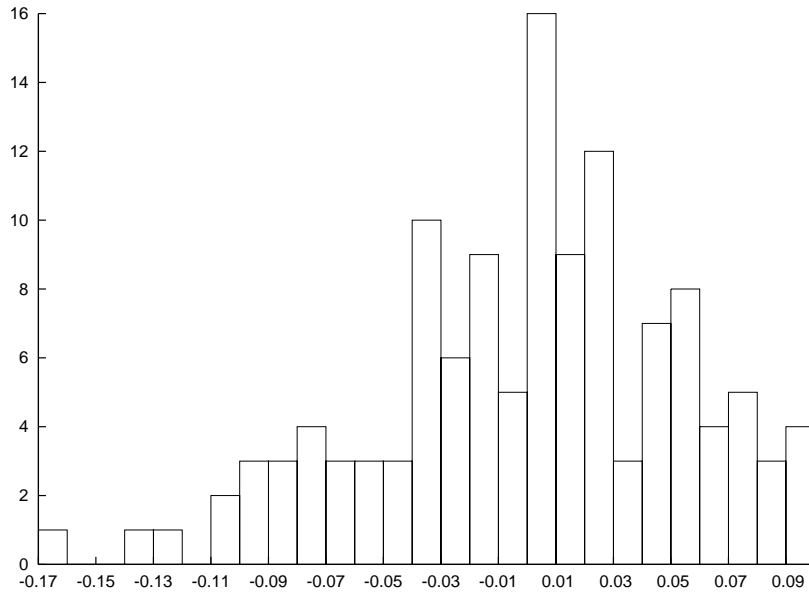
Figure 2: *The distribution of the cyclical indicator of GNP.*

due to the steady secular improvements in life conditions over time. In practice one may compare a cohort born in a boom to the cohort born in the subsequent recession, because the latter benefit from secular developments, so that a decrease of expected lifetimes can be attributed to the cyclical effect. More in general, one may relate the mortality rate to the state of the business cycle early in life. To proceed one therefore needs to assign a value of a cyclical indicator to each year. Most results below are based on a trend/cycle decomposition of log annual real per capita GNP using the Hodrick-Prescott filter with smoothing parameter 500. The values of the cyclical terms are very robust with respect to the actual decomposition method and smoothing parameter, and so are the resulting intervals within which the terms are positive or negative. We are therefore in the fortunate position that booms and recessions are clearly identifiable in the data. Moreover, the empirical analysis provides virtually identical results if other decomposition methods are used. Figure 1 displays the cycle and trend as functions of calendar time. Since the cyclical term (or indicator) will be our main explanatory variable, it is useful to know its distribution over time. Figure 2 provides a histogram where the cyclical term values are grouped in intervals of length 0.01. Below we occasionally round-off the value of the cyclical term to a binary outcome.

We also use external information on the incidence of epidemics and war because these cause pronounced spikes in the mortality rates. World War II (1940-1945) has been the only war and occupation on Dutch soil since Napoleon. It

included the famine of unprecedented severity of the winter of 1944/45, where mortality rates peaked because of malnutrition (Jewish genocide victims were excluded from the data). There are no reliable macro-economic statistics for the World War II period, so we represent it by a separate dummy variable.[8]

The price to be paid for the fact that the observation window is of unprecedented size concerns the absence of a number of variables that are often used in the mortality literature but that are unobserved in the 19th century records. Notably, we do not observe the individual's cause of death and aggregate amounts of health expenditures[9] and numbers of medical innovations.

2.4   Conditional lifetime durations and the cycle at birth

The sample is too small to be able to compare outcomes from two adjacent years with each other for each social class. To proceed, we aggregate rows of consecutive years into "boom" and "recession" periods. This enables us to make a non-parametric comparison of the average lifetimes in a cohort born in a single boom to those in the cohort born in the subsequent recession. Such a comparison is conservative in that it allows individuals in the recession to benefit from secular improvements. To prevent that the comparison is driven by the fact that lifetimes in a recession are smaller simply due to instantaneous effects of economic conditions on infant mortality, we condition on survival until age 2 or until age 5. The boom and recession should not include an epidemic, and we need both to last long in order to have reasonable sample sizes. We ignore the period after 1881 because the strong upward trends in GNP and lifetimes in that era may dominate the comparison.

Consider for example the boom of 1872-1876 and the subsequent recession of 1877-1881. Van den Berg, Lindeboom and Portrait (2006) showed that individuals born in this recession have significantly shorter mean conditional lifetimes than those born in the boom. It turns out that for most social classes the average conditional lifetimes are shorter among those born in the recession than among those born in the boom. However, the difference is only significantly negative for

---

[8]For overviews of developments in economic activity and demographic indicators in The Netherlands, see Petersen (1960), Mokyr (1974), Wolleswinkel-van den Bosch et al. (1998), De Vries (2000), and, in particular, Wintle (2000). See also Smits, Horlings and Van Zanden (2000) for aggregate time series on birth and mortality in the 19th century. The national population grew from 2.2 million in 1812 to 3.4 million in 1862 to 6.1 million in 1912.

[9]During most of the nineteenth century, government expenditure does not exceed 8% of GDP (see Wintle, 2000).

SC1. Apparently, even this approach suffers from the small sample size by social class.[10]

The main methodological disadvantage of simple comparisons of the cycle at birth and the average conditional lifetime is that they ignore periodicity of business cycles throughout the childhood years of an individual. Someone who is born in bad times is likely to experience good times during some childhood years, and vice versa, just because good and bad times succeed each other with an average frequency of a few years. If conditions at birth as well as during childhood affect mortality later in life, then the effect of the bad times at birth may be mitigated by the effect of the good times during childhood. To proceed, in the next section, we estimate duration models where the individual mortality rate is allowed to simultaneously depend on conditions at birth and on conditions during childhood. The estimation of these models exploits the variation in the timing of the stages of the business cycle across individuals, to disentangle the long-run effects of conditions at birth and during childhood. This leads to parameter estimates of the effect of cyclical conditions at birth on mortality later in life for a given social class and for given conditions during later childhood years. The duration analysis also controls for individual characteristics, albeit at the expense of model assumptions.

2.5 Birth cohort size and composition over the cycle

The main empirical analyses in the paper condition on social class, implying that exogenous variation in the number of births by social class does not affect the results. Nevertheless it is interesting to examine how the inflow into life and its composition across social class change with the business cycle, because such variation can reflect the extensive margin of health effects. Fertility may react to short-term economic stress, and this may in turn vary with social status. Also,

---

[10]Alternatively, one may use data from multiple cohorts simultaneously by performing a regression of conditional lifetimes on a binary boom/recession indicator at birth and its interaction with social class, including social class and dummy variables for being born in a particular "birth period" defined by a combination of a boom and the subsequent recession. Note that the latter dummy variables capture trends in average lifetimes across birth periods. It turns out that the effect of being born in the recession is typically significantly negative but that otherwise the coefficients of interest are insignificant with the expected signs but with values that are strongly dependent on details of the specification. Perhaps the loss of information due to (*i*) the discretization of time into "birth periods", booms, and recessions, and (*ii*) the focus on the mean of the lifetime distribution, is too large to be able to detect significant interaction effects.
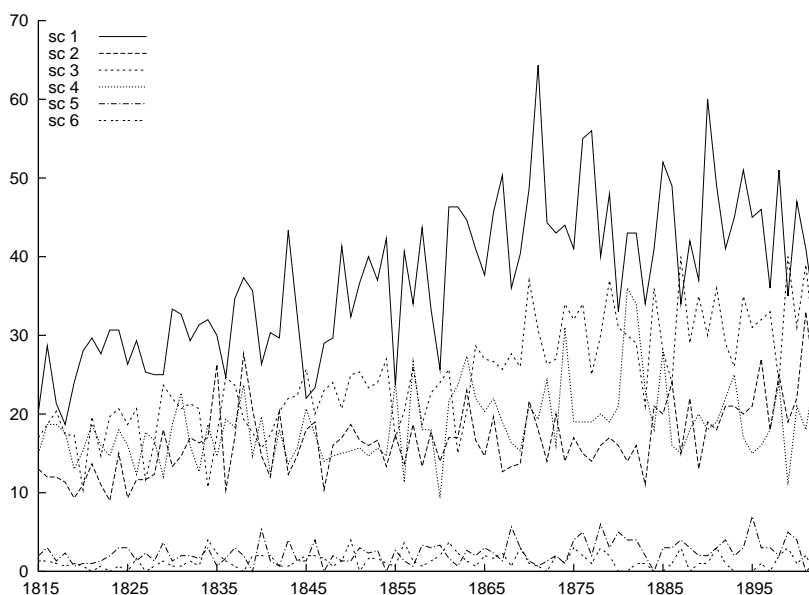
Figure 3: *Birth-year cohort size, over-all and by social class (standardized).*

individuals may postpone births upon perceiving a negative shock, so that the birth rate responds to economic fluctuations with a certain lag (see Bengtsson and Dribe, 2002, and references therein). On the other hand, it is observed in developing countries and modern societies that individuals with higher education or status more often have a child in a recession (Dehejia and Lleras-Muney, 2004), perhaps because the opportunity costs are lower.

Figure 3 displays the numbers of births by social class in our sample over the observation window. The numbers are adjusted for the fact that the annual sampling frequency for 1873–1901 differs from that for 1815–1872.[11] In Figure 4 the total annual inflow is normalized to 1, thus facilitating the view of how the composition of newborns in terms of social class changes over time. Towards the end of the nineteenth century there is a slight increase in the fractions of newborns from SC3, at the expense of the fractions from lower classes, but the effects are not overwhelming (see Wintle, 2000, for similar conclusions from the historical literature).

We also regress the (log) number of newborns per year on the cyclical indicator of the business cycle at birth, for the full sample, and also for each social class separately. We do not find significant effects of the cyclical indicator at birth on

---

[11]For Friesland and Zeeland for 1815–1872, the frequency is 0.0075, whereas for 1873–1901 it is 0.0050.
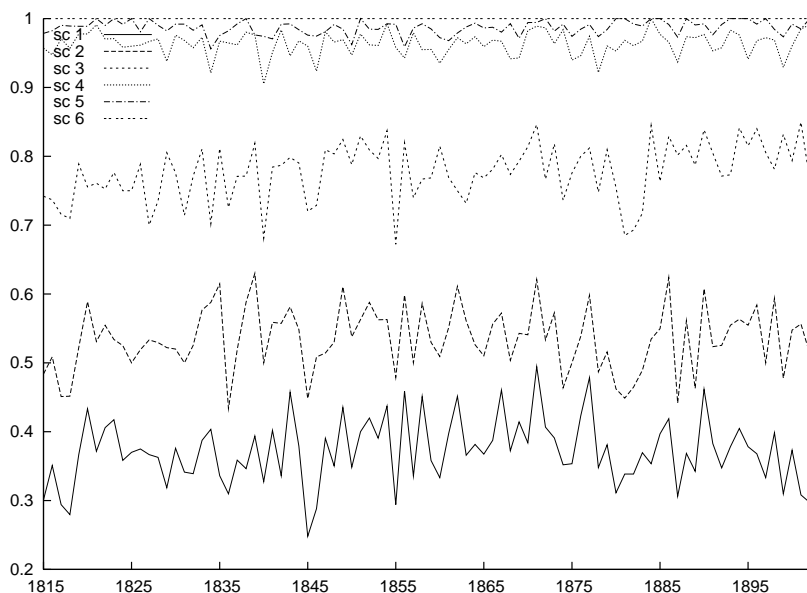
Figure 4: *Standardized composition of birth-year cohorts by social class.*

the over-all cohort size or the cohort size by social class. Other specifications, like those considering the fraction of births of each social class among newborns in a given year instead of the size, did not yield any significant results either. The same occurs when we fit a quadratic or cubic specification instead of a linear one, or when we use a lag of one year in the cyclical component. We conclude that there is no effect of the business cycle on the birth rate or on the composition of newborns by social class. By taking business cycle fluctuations as an indicator of unexpected economic shocks, it follows that such shocks do not affect the birth rate (over-all or by social class) in the Netherlands in the nineteenth century.

We expand on this by relating the fraction of newborns per year whose father has an occupation in the countryside to the cyclical indicator.[12] This relation turns out to be weak and insignificant, which confirms the conclusions of the previous paragraph.[13]

---

[12]This countryside occupation dummy equals 1 for about 16% of the sample. Its average by birth cohort decreases slowly over time, from about 20% in the 1810s until about 12% in the 1900s. The correlation with the social class indicator ranging from 1 to 4 is 0.40, and the correlation with the binary urban municipality indicator is -0.19.

[13]Of course, this does not rule out that the full distribution of occupations within a social class level displays trends and/or fluctuates over the cycle. Also, different occupations may be affected by the business cycle at different time lags. The data preclude an analysis of this.

3 Estimation of models for the individual mortality rate

3.1 Models for the individual mortality rate

The individual mortality rate is the natural starting point of the specification of the model, because of our interest in its dependence on conditions early in life. As our model specifications closely follow those in Van den Berg, Lindeboom and Portrait (2006), the present exposition can be brief. Age is measured in days, so we take it to be a continuous random variable. Let τ denote current calendar time. We express the mortality rate θ of an individual at a given point of time in terms of the prevailing age t, individual socio-economic and demographic background characteristics x, current macro-economic conditions $z(\tau)$, the trend components and cyclical indicators $z_{tr}(\tau-t+i)$ and $z_c(\tau-t+i)$ of macro-economic conditions earlier in life ($i \in \{0, ..., t-1\}$), and various interaction terms. The z variables are only measured once a year.

More specifically, we take $\log\theta(t|\tau, x, z)$ of an individual at age t at calendar year τ to be the sum of a number of terms (see also equation (1) below). The first of these, denoted by $\psi(t)$, captures the dependence on the prevailing age. This is a piecewise constant specification with 11 different intervals (0, 1, 2-6, 7-14, 15-34, 35-49, 50-59, 60-69, 70-79, 80-89, 90+), in consequence giving rise to 11 parameters. The second term is a linear parametric function of x. We restrict attention to characteristics at birth as opposed to later in life, for the reason that the latter may be endogenous or confounded. The third term is a linear parametric function of the prevailing $z(\tau)$. For $z(\tau)$ we take log annual real per capita GNP at t, as well as dummy variables for years with epidemics and for World War II. The latter also captures the fact that the GNP variable is missing for that period.

The fourth term is a linear parametric function of the cyclical indicators earlier in life. These parameters are the parameters of interest. As indicators we take the cyclical components $z_c(\tau)$ of the decomposition of log GNP from Section 2, at the birth year $\tau - t$, and also averaged over the years corresponding to age 1 up to 6. Of course, at $t = 0$ these are not allowed to affect θ. At $t = 1$, only the component in the previous year (i.e., $z_c(\tau - 1)$) is allowed to affect θ. At, for example, $t = 3$, θ depends on $z_c(\tau - 3)$ and on the average of $z_c(\tau - 2)$ and $z_c(\tau - 1)$.

The trend component $z_{tr}(\tau - t)$ of the decomposition from Subsection 2.3

at the birth year captures the secular long-run effects. The trend components in early childhood are typically very similar to this. Indeed, it is empirically difficult to disentangle the effects of these trend components of the conditions earlier in life from the effects of $z(\tau)$, due to multicollinearity. All of these variables are mostly increasing over time, and at the individual level the trend components of log GNP can be captured relatively well by the sum of current log GNP and an increasing function of age. We therefore omit $z_{tr}(\tau - t + i)$ from the model specification. Similar problems arise if we add calendar time polynomials to the specification.

This gives rise to

$$\log \theta(t|\tau, x, z, c) \;=\; \psi(t) + \beta' x + \alpha'_1 z(\tau) + \alpha_2(x)\, z_c(\tau - t)\, \mathrm{I}(t \geq 1) \;+$$

$$+ \alpha_3(x) \sum_{i=1}^{\min\{6,t-1\}} \frac{z_c(\tau - t + i)}{\min\{6, t-1\}}\, \mathrm{I}(t \geq 2),$$

where the dependence of α_2 and α_3 on x emphasizes that we allow these to depend on the social class, for example by way of $\alpha_j(x) = \gamma'_j x$. We also estimate more general specifications allowing for various other interaction effects between the mortality rate determinants. In particular, we fully interact the age dependence and current log GNP. Babies may suffer disproportionally from bad current conditions. If this is ignored then this may be picked up by the estimated coefficient of the cyclical indicator at birth, which may then be biased. The models are estimated by Maximum Likelihood.

As is well known, ignoring unobserved heterogeneity of mortality determinants across individuals may result in biased estimates of the duration model parameters (see Van den Berg, 2001, for an overview). Unobserved heterogeneity poses an additional problem if the current individual hazard rate is allowed to depend on the value of an explanatory variable at a point of time in the past but after the beginning of the spell; in our case the cyclical indicator during early childhood years (see e.g. Vaupel and Yashin, 1985). Basically, bad childhood years may give rise to selection of childhood survivors with favorable characteristics. So if unobserved heterogeneity is present but is not taken into account then the effect of the cycle during early childhood may be biased. However, the dynamic selection effect can be expected to generate a positive relation between the cycle during early childhood and observed mortality later in life. So, if we find a negative effect without taking account of unobserved heterogeneity, then the true effect is likely to be at least as negative. Of course, if the amount of unobserved heterogeneity

differs across social class at birth then the comparison of estimation results by social class at birth is hampered.

3.2   Estimation results

Table 3 presents the estimation results for the basic model specification, and Table 4 gives the results for the specification with interactions between age and current GNP. The estimates concern the mortality rate, so a positive value is associated with a shorter lifetime. The time unit is one year. To interpret the estimates for the cyclical indicators, notice that the reported marginal effects correspond to the hypothetical class with value zero. As an example, in Table 3, the effect of the cyclical indicator at birth for SC2 equals $-1.620 + 0.437 * 2$. Conversely, the marginal effect of social class at birth can be read off directly from the tables, because the cyclical indicators are mean-centered.

The most striking result is that for the lowest three classes (i.e., for most individuals) the cyclical indicator in the birth year has a negative effect on the mortality rate later in life. For individuals in social class 4, the effect is positive. However, notice that the effect for class 4 is very close to zero. We take this as evidence that among lower classes, at the individual level, economic conditions at birth have long lasting effects on mortality. Selection due to unobserved heterogeneity does not explain this result, because it would give a positive coefficient. Allowing for interaction effects between age and current macro-economic conditions gives a slightly smaller absolute value of the coefficient of the cyclical indicator in the birth year. The cyclical indicator at ages 1 up to 6 does not have a significant effect on mortality later in life for any social class.

The parameter estimates quantify the ceteris paribus effects of social class at birth and the business cycle indicator and their interactions, on the individual mortality rate later in life. If the cyclical indicator at birth has its average value of zero, then the mortality rate of SC4 individuals is approximately 10.5% lower than the rate of SC1 individuals. In a recession with a cyclical indicator of -0.05, the long-run mortality rate of SC4 newborns is approximately 16.5% lower than the rate of SC1 newborns. In a boom with a cyclical indicator of 0.05, the long-run mortality rate of SC4 newborns is approximately 4.6% lower than the rate of SC1 newborns. This implies that among SC1 newborns, the mortality rate of those born in a recession is 12% lower than of those born in a boom. For SC4 individuals the difference between being born in a boom or a recession is negligible. In Section 4 we translate these estimates into lifetime years gained or lost.

19

As noted in Subsection 2.2, individuals may mitigate effects of having been born in a recession by way of upward social mobility during the life course. This would not invalidate the empirical analysis, but it could affect the size and therefore the interpretation of the estimated net long-run effects. We shed some light on the importance of this by using the subsamples of 565 individuals for whom social class is observed at birth and at death and the subsample of 655 men for whom this is observed at birth and at the date of marriage. Specifically, we examine whether these individuals are more likely to be upwardly mobile if they are born in recessions. We define a recession as a year in which the cyclical indicator is negative. We have already seen that over-all mobility is low. In particular, there is virtually no upward mobility from SC4 to SC5 or SC6, whether the individual is born in a recession year or not. This suggests that upward social mobility does not explain the success of SC4 individuals in preventing or neutralizing any scars of having been born in a recession. It is more plausible that the actual economic well-being of SC4 families is simply not strongly affected by recessions, so that their newborns are not scarred in the first place. Across all social classes (and also among SC1 individuals only), upward social mobility is slightly lower among those born in a recession year. Specifically, the fraction of upwardly mobile individuals in the two subsamples is 3 and 2 percentage points lower among those born in a recession year than among those not born in a recession year. These numbers are not significantly different from zero. This suggests that social mobility is not systematically used to dampen negative long-run effects of birth in a recession.

In a related sensitivity analysis, we allow the marginal effect of social class and the interaction effect of social class and the cyclical indicator at birth to vary with age. This should reveal whether the importance of social class at birth decreases as age increases, and whether the negative effect of birth in a recession for those with a low social class diminishes as age increases. It turns out that these interactions with age are all insignificant.[14] The point estimates seem to suggest a slightly decreasing importance of social class at ages above 15. This suggests that perhaps individuals use other pathways than transitions to a higher social class in order to improve their health. Note however that coefficients interacted with age may be sensitive to misspecification of instantaneous (GNP) effects on the mortality rate, since older individuals are over-represented in later years. In any case, the result that the direction of social mobility does not depend on the cycle at birth and the result that social class has a mortality effect at all ages

---

[14]These estimation results are available upon request.

both confirm that long-run economic well-being of the family at birth is a relevant determinant of the mortality rate throughout life.

We briefly discuss some other parameter estimates. Current log annual real per capita GNP has a significantly negative estimated instantaneous effect. The estimated coefficient is the same whether we allow the long-run effect of the cycle at birth to vary with social class or not. A likelihood ratio test of the restriction that the effect is age-independent results in rejection. The estimated effect is largest for individuals aged between 7 and 34. This may reflect an instantaneous effect but may also reflect secular developments accumulated up to the current age. The model specification allows for the cholera epidemic in Utrecht in 1849, the smallpox epidemic in Utrecht in 1870 and in Friesland and Zeeland in 1871, and the influenza epidemic of 1918, as these are typically regarded to be the most severe in our observation window. We do not discuss these effects in detail, as they are not related to the focus of the paper.

The specification of the log mortality rate as a linear function of social class is restrictive, as the latter is an ordinal variable. To deal with this we replace the linear specification by a set of three dummy variables, for the marginal effect and also for the interactions with the cyclical indicators. The results are in Table 5. They mostly confirm the earlier results. The additive effects are monotonous, and the interaction effects are almost monotonous. Note that the effect of the business cycle at birth is significant only for the two lower social classes. The log likelihood values for estimated models in Tables 3 and 5 differ only by 2. A likelihood ratio test of the null hypothesis of a log linear specification (effectively reducing the number of parameters by 6, so that the test statistic asymptotically has a chi-square distribution with 6 degrees of freedom) therefore leads to acceptance of this specification.

3.3 Additional extensions and sensitivity analyses

We estimate a wide range of alternative model specifications. We start with specifications that are different with respect to the role of social class and its interaction with the cycle at birth. We then consider specifications that are different with respect to the cyclical indicators early in life, and we conclude with specifications differing with respect to other explanatory variables or their inclusion. For reasons of brevity we do not present estimates of all specifications and all parameters (these are available upon request).

Table 3: Parameter estimates of the baseline model for the individual mortality rate

variable	parameter estimate	standard error
Individual background characteristics		
female	-0.105	(0.024)*
social class	-0.034	(0.011)*
father not illiterate	-0.087	(0.034)*
born in urban area	0.083	(0.030)*
born in province Utrecht	0.240	(0.032)*
born in province Zeeland	0.297	(0.028)*
Conditions early in life		
cyclical indicator in birth year	-1.620	(0.511)*
idem × social class	0.437	(0.201)*
cyclical indicator at ages 1 up to 6	0.219	(1.046)
idem × social class	-0.105	(0.410)
Contemporaneous macro conditions		
current log(annual real per capita GNP)	-0.382	(0.033)*
1849 cholera in Utrecht	0.855	(0.198)*
1870/1 smallpox	0.486	(0.103)*
1918 influenza	-0.232	(0.178)
world war II (GNP missing)	-3.090	(0.282)*
Current age		
age 0	1.501	(0.243)*
age 1	0.195	(0.247)
age 2–6	-1.099	(0.246)*
age 7–14	-2.316	(0.252)*
age 15–34	-2.147	(0.252)*
age 35–49	-1.979	(0.262)*
age 59–59	-1.372	(0.268)*
age 60–69	-0.475	(0.271)
age 70–79	0.622	(0.278)*
age 80–89	1.613	(0.285)*
age 90+	2.414	(0.297)*

Explanatory note: An asterisk denotes significance at the 5% level. The province Friesland is the reference province.

Table 4: Parameter estimates of the model for the individual mortality rate with interactions between current GNP and age

variable	estimate	st.error	estimate	st.error
Individual background characteristics				
female	-0.107	(0.024)*		
social class	-0.035	(0.011)*		
father not illiterate	-0.080	(0.036)*		
born in urban area	0.080	(0.030)*		
born in province Utrecht	0.248	(0.031)*		
born in province Zeeland	0.300	(0.028)*		
Conditions early in life				
cyclical indicator in birth year	-1.586	(0.513)*		
idem × social class	0.433	(0.198)*		
cyclical indicator at ages 1 up to 6	0.280	(1.131)		
idem × social class	-0.091	(0.439)		
Contemporaneous macro conditions				
1849 cholera in Utrecht	0.756	(0.211)*		
1870/1 smallpox	0.495	(0.107)*		
1918 influenza	-0.141	(0.178)		
world war II (GNP missing)	-2.211	(0.312)*		
Current age, and interacted with current log(*annual real per capita GNP*)				
age 0	1.000	(0.687)	-0.315	(0.093)*
age 1	1.371	(1.387)	-0.541	(0.187)*
age 2–6	3.489	(1.260)*	-1.001	(0.170)*
age 7–14	5.387	(1.900)*	-1.414	(0.256)*
age 15–34	5.667	(1.093)*	-1.410	(0.144)*
age 35–49	-3.017	(0.389)*	-0.250	(0.050)*
age 50–59	-2.116	(0.347)*	-0.293	(0.044)*
age 60–69	-1.245	(0.328)*	-0.291	(0.040)*
age 70–79	-0.365	(0.314)	-0.265	(0.037)*
age 80–89	0.784	(0.317)*	-0.289	(0.036)*
age 90+	1.575	(0.512)*	-0.291	(0.056)*

Explanatory note: The first two columns contain the marginal effects of age while the last two columns give the coefficients of the interaction terms of age and GNP. An asterisk denotes significance at the 5% level. The province Friesland is the reference province.

Table 5: Parameter estimates of the model for the individual mortality rate with dummies for social class

variable	parameter estimate	standard error
Individual background characteristics		
female	-0.107	(0.024)*
social class 1	0.099	(0.034)*
social class 2	0.040	(0.041)
social class 3	0.015	(0.038)
father not illiterate	-0.083	(0.035)*
born in urban area	0.089	(0.030)*
born in province Utrecht	0.240	(0.032)*
born in province Zeeland	0.296	(0.028)*
Conditions early in life		
cyclical indicator in birth year × social class 1	-1.000	(0.374)*
idem × social class 2	-1.619	(0.625)*
idem × social class 3	-0.033	(0.511)
idem × social class 4	0.142	(0.516)
cyclical indicator at ages 1 up to 6 × social class 1	0.180	(0.765)
idem × social class 2	0.305	(1.305)
idem × social class 3	-0.858	(1.062)
idem × social class 4	0.238	(1.048)
Contemporaneous macro conditions		
current log(annual real per capita GNP)	-0.382	(0.033)*
1849 cholera in Utrecht	0.855	(0.198)*
1870/1 smallpox	0.487	(0.103)*
1918 influenza	-0.231	(0.178)
world war II (GNP missing)	-3.083	(0.282)*
Current age		
age 0	1.368	(0.243)*
age 1	0.062	(0.246)
age 2-6	-1.233	(0.246)*
age 7-14	-2.449	(0.252)*
age 15-34	-2.280	(0.251)*
age 35-49	-2.113	(0.262)*
age 50-59	-1.505	(0.268)*
age 60-69	-0.609	(0.271)*
age 70-79	0.488	(0.277)
age 80-89	1.480	(0.285)*
age 90+	2.281	(0.297)*

3.3.1 Countryside occupations

For families with work in the countryside, it may be easier to deal with negative shocks. For example, they may have better food access and better opportunities to avoid epidemics. We therefore add the countryside occupation dummy variable introduced in Subsection 2.5 as an individual background characteristic, and we estimate model versions with and without an interaction of this variable with the cyclical component at birth. Table 6 shows the relevant estimates. The effect of the countryside occupation variable and its interaction with the cycle at birth are both insignificant. The signs are as expected: if the father has a countryside occupation then the mortality rate is lower, and the effect of a negative economic shock is mitigated. The effects take away a small part of the effects of social class. Note that the fraction of individuals with a countryside occupation (16%) probably under-estimates the fraction of individuals with opportunities to mitigate adverse transitory shocks by way of access to land for food cultivation.

Table 6: Relevant parameter estimates of models for the individual mortality rate with countryside occupation dummy and its interaction with cyclical indicator at birth

variable	estimate	st.error
with countryside occupation × cycle at birth		
social class	-0.027	(0.012)*
born in urban area	0.072	(0.030)*
countryside occupation	-0.052	(0.037)
cyclical indicator at birth	-1.575	(0.516)*
idem × countryside occupation	0.473	(0.687)
idem × social class	0.383	(0.214)
without countryside occupation × cycle at birth		
social class	-0.027	(0.012)*
born in urban area	0.072	(0.030)*
countryside occupation	-0.051	(0.037)
cyclical indicator at birth	-1.631	(0.513)*
idem × social class	0.442	(0.202)*

By analogy to the modelling of countryside occupation effects, we therefore

incorporate the interaction between urbanization and the cycle at birth. Moreover, the effect of being born in an urban or rural area may not be the same for all social classes. Therefore, we estimate model versions that include interactions between the social class of the individual and the indicator of being born in an urban area, interactions between the cyclical indicator at birth and this urban indicator and also between these three variables described. None of these are significant.

3.3.2 Gender

We also estimate separate models for men and women. The parameter estimates are similar in magnitude and significance, in particular concerning the parameters of interest. In general, standard errors are somewhat larger, which is not surprising given the fact that the sample is cut in half. The main difference between the estimates by gender is that the marginal social class effects are larger in absolute value for women than for men. This suggests that women's mortality is more strongly driven by the average economic situation in the family into which they are born, whereas the sensitivity for a negative shock (recession) at birth is similar, as is the extent to which the sensitivity for such a shock is compensated for by social class.

We may relate this to the social mobility of women by marriage, as observed in the subsample of women for whom we observe the social class at birth and the social class of their husband. We find that marriage involves more often a transition to a higher social class than to a lower social class. Moreover, the fraction of upwardly mobile women is 4 percentage points higher among those born in a recession year than among those not born in a recession year. This number is marginally significant. By itself this is an interesting result, but apparently the upward mobility by marriage does not translate into a lower dependence of mortality on the social class at birth for women on average as compared to men, or into a lower dependence of mortality on the cycle at birth as compared to men. Perhaps the health gains from upward mobility are smaller if achieved by marriage than if achieved by education or one's own occupation. A full investigation should analyze mobility, marriage, and mortality in a joint framework, allowing for selectivity in marriage and mobility outcomes. This requires data that are more informative than at our disposal.

3.3.3 Social class grouping

As noted in Section 2, we merge social classes 4, 5 and 6 into one category. We also estimate models where the two highest groups are excluded from the sample. Table 7 gives the estimates, for a log-linear specification of the mortality rate as a function of social class, and for a specification where social class is captured by dummy variables. Clearly, the main results are insensitive to whether classes 5 and 6 are included.

Table 7: Relevant parameter estimates of the model for the individual mortality rate with individuals in social class 5 and 6 omitted

	estimate	st.error
with social class = 1,...,4		
social class	-0.035	(0.011)*
cyclical indicator at birth	-1.632	(0.517)*
idem × social class	0.445	(0.208)*
cyclical indicator at ages 1 up to 6	0.360	(1.056)
idem × social class	-0.192	(0.421)
with dummies for social class		
social class 1	0.103	(0.035)*
social class 2	0.045	(0.042)
social class 3	0.019	(0.039)
cyclical indicator at birth × social class 1	-1.000	(0.374)*
idem × social class 2	-1.615	(0.626)*
idem × social class 3	-0.034	(0.511)
idem × social class 4	0.178	(0.555)
cyclical indicator at ages 1 up to 6 × social class 1	0.201	(0.767)
idem × social class 2	0.318	(1.307)
idem × social class 3	-0.860	(1.064)
idem × social class 4	-0.015	(1.107)

3.3.4 Cyclical indicator

We now turn to specifications that are different with respect to the cyclical indicators early in life, First, instead of using the cyclical indicator at birth, we use the average cyclical indicator of the year previous to birth and the year of birth, so that in most cases the period of the pregnancy is also taken into account. This does not result in any substantial differences in the estimates. Secondly, recall that the average cyclical indicator at ages 1 up to 6 do not have any significant effects. To examine whether the estimates of the parameters of interest change if this average indicator is omitted, we estimate the model without it. Table 8 shows that the estimates are insensitive to this, whether social class is log-linear or captured by dummy variables. The average cyclical indicator for the age set $\{7, \ldots, \min\{14, t-1\}\}$ also turns out to have an insignificant effect.

Table 8: Relevant parameter estimates of models for the individual mortality rate with cyclical indicator at ages 1–6 excluded

	estimate	st.error
with social class = 1,…,4		
social class	-0.034	(0.011)*
cyclical indicator at birth	-1.624	(0.510)*
idem × social class	0.438	(0.201)*
with dummies for social class		
social class 1	0.098	(0.034)*
social class 2	0.040	(0.041)
social class 3	0.012	(0.037)
cyclical indicator at birth × social class 1	-1.003	(0.373)*
idem × social class 2	-1.623	(0.621)*
idem × social class 3	-0.023	(0.511)
idem × social class 4	0.141	(0.515)

3.3.5 Censoring

First, we estimate a probit model for whether an individual's lifetime is right-censored at zero or not. Most importantly, we find that the cycle at birth does not

have any significant effect. Other explanatory variables have significant effects. Notably, censoring is more likely in urbanized municipalities, among men, and among higher social classes. It is hard to interpret these results because censoring may have many causes, like non-compliance, the accidental destruction of certificates, and emigration abroad, where the latter two are quantitatively negligible at the national level. To complicate matters, note that right-censored observations necessarily have a positive duration, and indeed most likely survive childhood, whereas the uncensored include durations at zero and at low values. From such a point of view, one would expect opposite signs for the effect of an explanatory variable on the probability of censoring and the effect on the mortality rate in the full model, which is what we find for social class.

Next, we examine the sensitivity of the results with respect to the independence of the right-censoring. We take all durations that are censored at ages below 60 to be right-censored at age 60, and we re-estimate the model. It turns out that the effects of the cycle at birth and its interaction with social class are completely insensitive to this. The main differences are, not surprisingly, in the age coefficients and in the urbanization coefficient which is now less positive. We conclude that the main results are robust with respect to the way in which we deal with right-censoring.

3.3.6 Other interactions and specification issues

According to the estimation results in Tables 3 and 4, individuals born in Friesland have a lower mortality rate than those born in the other two provinces. This is consistent with the facts that in the period considered Zeeland and Utrecht had poorer sanitary conditions than Friesland and that in Friesland breastfeeding was much more common than in the other provinces (see Van Poppel and Mandemakers, 2003). It may be that being born in Zeeland or Utrecht does not affect mortality equally for individuals from different social classes. For this reason, we estimate a model where the interaction between the province of birth and the social class indicator is included. The effect is significant, implying a higher inequality of mortality across social classes in Zeeland and Utrecht.

The interaction between current log GNP and social class is not significant. Changes in the age intervals in the age dependence function do not affect the estimation results. In order to reduce the number of parameters to be estimated in the model, a sum of Chebyshev polynomials of the second kind[15] in t (pre-

---

[15]Such polynomials are mutually orthogonal in the observation window, ensuring absence of multicollinearity. See Abramowitz and Stegun (1970), for details, and Abbring, Van den Berg

vailing age) with unknown parameters can be considered as an alternative to the piecewise-constant specification for age. One may adopt a similar approach for current log(GNP). In both cases, the main results from the previous subsection remain the same, which gives evidence of their robustness.

The urbanization indicator used so far can be extended to a classification in three or more categories. Moreover, we may define the urbanization degree of the municipality by using the population density (since the spread of diseases may be faster among high concentrations of people) or the number of inhabitants of the town or the prevailing type of employment in the province the town belongs to (rural, industrial)[16]. We estimated several specifications. Estimates of the model with an urbanization variable with three levels are shown in Table 6. The other specifications lead to insignificant urbanization effect estimates.

Table 9: Relevant parameter estimates of the model for the individual mortality rate with urbanization degree at birth in 3 levels

	estimate	st.error
social class	-0.033	(0.011)*
urbanization degree	0.031	(0.016)
cyclical indicator at birth	-1.618	(0.510)*
idem × social class	0.435	(0.200)*
cyclical indicator at ages 1 up to 6	0.228	(1.047)
idem × social class	-0.114	(0.411)

4 Quantifying the effect of the business cycle early in life on the mortality variation across social classes

In this section we develop and quantify inequality measures to capture the mortality variation across social classes due to variation in the effect of the business

and Van Ours (2002) for an application in duration analysis.

[16]Data on this were obtained from StatLine, the central database of ''Statistics Netherlands'' (Centraal Bureau voor de Statistiek)

cycle early in life on mortality. This serves to translate the corresponding parameter estimates into lifetime years gained or lost in comparison to other social classes or to other early childhood conditions.

Consider, as a starting point, the mortality rate at a given point in calendar time τ. We could try to assess the effect of the business cycle early in life on this cross-sectional mortality rate, or we could try to decompose the cross-sectional variation in mortality into different terms, one of which captures the long-run effect of the business cycle early in life. One obvious way to proceed is to use the estimated model and compare the ensuing inequality to the counterfactual situation in which the interactions between social class and the cycle early in life are switched off. However, each cross-section contains different cohorts born at different stages of the business cycle, so we would consider some kind of average over the different stages of the cycle at birth. Clearly, cross-sectional measures are are not tailored to capture the long-run effects of events at birth. As an alternative, we make comparisons across birth cohorts, that is, use cohort-based measures.

For each birth cohort year τ, the variation across social classes of the mean (log) lifetime duration can be an indicator of social inequality in cohort τ. It will strongly reflect the cyclical indicator c of the cycle in the birth year, in the sense that in a cohort born in a recession this indicator of social inequality will be larger than in a cohort born in a boom. We take the estimated model instead of the raw data, because otherwise sampling variation distorts the analysis. The duration model estimates give

$$ \mathrm{E}(T|\tau, x, sc), $$

where we take τ to denote the birth year. Though T also depends on c and on the time-varying contemporaneous z, we omit the symbols for c and z here because c and z are given once we fix the birth year τ. In the calculations we use the actual time path of the contemporaneous effects $z(.)$ from τ onwards and we have to choose a value for c. By x we represent the personal characteristics excluding social class sc.

After estimating the distribution of individual characteristics X in the inflow, we take expectations over x and define:

$$ \mu(\tau, sc) \; := \; \mathrm{E}_X \, (\log) \, [\, \mathrm{E}(T|\tau, x, sc) \,]. $$

where the logarithm is optional. Alternatively, we could consider a "representative" individual with personal characteristics $x_i = \overline{x}$ equal to the average of x in

31

the sample, which gives

$$\mu^*(\tau, sc) := (\log) \, \mathrm{E}(T | \tau, \overline{x}, sc).$$

We can capture social inequality in birth cohort year τ by defining:

$$SIV(\tau) := \mathrm{var}_{sc|\tau} \, \mu(\tau, sc)$$

$$SIV^*(\tau) := \mathrm{var}_{sc|\tau} \, \mu^*(\tau, sc),$$

which requires estimation of the distribution of sc in the inflow at τ. Alternatively, we can compare the two extreme values of the social class indicator,

$$SID(\tau) := \mu(\tau, sc = 4) - \mu(\tau, sc = 1)$$

$$SID^*(\tau) := \mu^*(\tau, sc = 4) - \mu^*(\tau, sc = 1).$$

The latter measures do not take account of changes in the distribution of $sc|\tau$, and they are in line with the relative index of inequality as defined by Kunst and Mackenbach (1997). We can now compare the values of SID and SID^* across different τ at the peak b and at the trough r of a cycle, giving rise to our measure of interest:

$$DSID := \mu(\tau = b, sc = 4) - \mu(\tau = b, sc = 1) - [\mu(\tau = r, sc = 4) - \mu(\tau = r, sc = 1)].$$

This difference-in-differences captures the relative loss in years of life for low social class people due to the interaction between c and sc, i.e. due to the fact that the low social class people suffer more from being born in a recession.

In the applications we always compare two birth cohort years where one is a boom year and the other a year in the subsequent recession. Specifically, we compare 1824 to 1830, and 1825 to 1830, and 1841 to 1846. We use $\mu(\tau, sc)$ for all the calculations. The results are very close to those obtained when using $\mu^*(\tau, sc)$. Consider first the boom year 1824 and the recession year 1830. The difference $SID(1824)$ in expected lifetime between individuals from the highest and lowest social classes is approximately 5 years. This difference rises to 8 years for the recession year 1830 (i.e. $SID(1830) = 8$). As a result, $DSID = -3$. This means that low social class individuals lose 3 years extra lifetime compared to high social class individuals if they are born in a recession instead of a boom. Of course, the value of $DSID$ depends on the birth years considered. For other years, the lifetime path of GNP is different, as is the actual value of the cyclical indicator at birth. Comparing 1825 to 1830 gives rise to $DSID = -7$, and comparing 1841

to 1846 gives rise to $DSID = -1$. In general, for pronounced boom and recession years, the values of $DSID$ are somewhat dispersed around the value of 4 years.

We can remove instantaneous effects of infant mortality by conditioning on surviving into early childhood (until the age of 5) in all of the previous measures. In this case, all the expected lifetimes increase, and the value of $DSID$ is typically smaller in absolute value. For example, $SID(1830)$ decreases from 8 to 5 years, and as a result $DSID$ for the comparison of 1825 and 1830 changes from -7 to -4. If the interaction parameters are switched off (i.e. if we consider the counterfactual situation in which mortality later in life does not depend on the interaction of social class and the cycle early in life), the value of the $DSID$ reduces to zero. This is not surprising given that this measure is designed to capture the inequality due to these interactions.[17]

5   Conclusion

The high-age mortality of lower social classes suffers more from being born in recessions than the high-age mortality of middle and higher social classes. If one is born into middle and higher class families (with higher average income) then the long-run health effects of adverse transitory macro shocks at birth are mitigated. For the highest social classes, producing about 20% of newborns, such shocks do not affect the high-age mortality of their children at all. As a result, mortality inequality is higher than in a society in which the long-run effects of conditions early in life do not depend on social class.

---

[17]For completeness we also explain the computation of the cross-sectional measures discussed at the beginning of the section. We focus on the mortality rate $\theta_s(\tau, sc)$ of those alive at calender time τ with social class sc. This effectively involves the examination of elapsed durations in a stock sample of ongoing spells (Ridder, 1984). We assume that the inflow rate $q(\tau, x, sc)$ (i.e., basically, the number of people who are born at τ with characteristics x, sc) is multiplicative in x, so $q(\tau, x, sc) = q_0(\tau, sc)q_1(x)$. It follows that

$$\theta_s(\tau, sc) = \frac{\int_{p=0}^{\infty} \int_x q_0(\tau - p, sc) f(p|\tau - p, sc, x) dG(x) dp}{\int_{p=0}^{\infty} \int_x q_0(\tau - p, sc) \overline{F}(p|\tau - p, sc, x) dG(x) dp},$$

where G is the distribution of X in the inflow and f is the duration density of T, with c.d.f. F. To quantify this, the size of the inflow into life needs to be estimated, as a function of τ, sc and x. Since individuals in the sample considered were born in years 1815-1901, it would be useful to evaluate this mortality rate after 1890, where the ages of those alive that year range from newborns to at least 75 years. One may then compare the inequality in $\theta_s(\tau, sc)$ over sc to the inequality in the counterfactual situation in which interactions are absent.

The results are not explained by social mobility, in the sense that social mobility is rather low and the extent of upward mobility does not vary much with the business cycle at birth. Moreover, social class and its interaction with transitory macro shocks at birth have long-run mortality effects at all ages. All of this confirms that the economic well-being of the family at birth as well as transitory economic shocks at birth (and their interaction) are all relevant determinants of the mortality rate throughout life.

We also find that for those having been born into families with work in the countryside, it is easier to deal with negative macro shocks. The main difference between the estimates by gender is that social class effects are larger in absolute value for women than for men. This suggests that women's mortality is more strongly driven by the average economic situation in the family into which they are born, whereas the sensitivity for a negative shock (recession) at birth is similar, as is the extent to which the sensitivity for such a shock can be compensated for by social class.

The results have some implications for policy design and policy analysis, notably in developing countries. First, it may be sensible to monitor the health of babies and pregnant women among lower social classes in recessions. If the health status is deemed low, then these individuals can be helped by providing better food, housing, and health care. This may significantly reduce their mortality later in life. Secondly, projections of future health care and social security needs can be improved by taking long-run effects of early childhood business cycle conditions into account.

Some of the issues that this paper deals with could be further investigated with family tree data that provide lifetimes of members within families. Also, more comprehensive information about individual social mobility can be useful to provide insights into the differences in the results across gender. The information should be sufficiently rich to deal with selectivity of marriage and social mobility.

References

Abbring, J.H., G.J. van den Berg and J.C. van Ours (2002), The anatomy of unemployment dynamics, *European Economic Review* 46, 1785–1824.

Abramowitz, M. and I.A. Stegun (1970), *Handbook of Mathematical Functions; With Formulas, Graphs and Mathematical Tables*, Dover Publications, New York.

Bengtsson, T. and M. Dribe (2002), Fertility Response to Short-term Economic Stress: Deliberate Control or Reduced Fecundability?, Working paper, Lund University.

Boonstra, O.W.A. and C.A. Mandemakers (1995), Ieder is het kind zijner eigene werken; sociale stratificatie en mobiliteit in Nederland in the achttiende en negentiende eeuw, in: J. Dronkers, W.C. Ultee (eds.), *Verschuivende ongelijkheid in Nederland*, Van Gorcum, Assen.

Case, A., A. Fertig and C. Paxson (2003), From cradle to grave? The lasting impact of childhood health and circumstance, Working paper, NBER and Princeton University.

Dehejia, R. and A. Lleras-Muney (2004), Booms, busts and babies' health, *Quarterly Journal of Economics* 119, 1091–1130.

De Vries, J. (2000), Dutch economic growth in comparative historical perspective, 1500–2000, *De Economist* 148, 443–467.

Diederiks, H.A. (1985), Beroepsstructuur en sociale stratificatie in Leiden in het midden van de achttiende eeuw, in: H.A. Diederiks et al. (eds), *Armoede en sociale spanning: sociaal-historische studies over Leiden in de achtiende eeuw*, Verloren Publishers, Hilversum.

Diederiks, H.A., D.J. Noordam and H.D. Tjalsma (1985), *Armoede en sociale spanning: sociaal-historische studies over Leiden in de achtiende eeuw*, Verloren Publishers, Hilversum.

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

Ganzeboom, H.B.G. (2002), Long term trends in occupational status attainment in the Netherlands in the 19th and 20th century: evidence from marriage records and sibling data, Working paper, Utrecht University.

Hayward M.D. and B.K. Gorman (2004), The long arm of childhood: the influence of early-life social conditions on men's mortality, *Demography* 41, 87–107.

Jacobs, J. and J.P. Smits (2001), Business cycles in the Netherlands, 1815–1913, Working paper, Groningen University.

Johansson, K. (2002), Early life conditions and childhood mortality in Scania, Sweden, 1766–1894, Working paper, Lund University.

Mackenbach, J. and A. Kunst (1997), Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe, *Soc Sci Med* 44, 757–771.

Mandemakers, K. (1987), Aanzet tot een beroepsstratificatie voor Nederland rond 1900, Tiel 1884 en 1918, *Tijdschrift voor Sociale Geschiedenis* 13, 198–222.

Mandemakers, K. (2000), Historical sample of the Netherlands, in: P. Hall et al. (eds.), *Handbook of International Historical Microdata for Population Research*, Minnesota Population Center, Minneapolis.

Mokyr, J. (1974), The industrial revolution in the low countries in the first half of the nineteenth century: a comparative case study, *Journal of Economic History* 34, 365–391.

Petersen, W. (1960), The demographic transition in the Netherlands, *American Sociological Review* 25, 334–347.

Ridder, G. (1984), The distribution of single spell duration data, in: G.R. Neumann and N. Westergard-Nielsen (eds), *Studies in Labour Market Analysis*, Springer-Verlag, Berlin.

Ruhm, C.J. (2000), Are recessions good for your health?, *Quarterly Journal of Economics* 115, 617–650.

Sastry, N. (2004), Trends in socioeconomic inequalities in mortality in developing countries: the case of child survival in Sao Paulo, Brazil, *Demography* 41, 443–464.

Smits, J.P., E. Horlings and J.L. van Zanden (2000), Dutch GNP and its components, 1800–1913, Working paper, Groningen University.

Suurenbroek, F. (2001), Binnenlandse migratie naar en uit Amsterdam (1870–1890), Working Paper, CGM Amsterdam.

Treiman, D.J. (1977), *Occupational Prestige in Comparative Perspective*, Academic Press, New York.

Van den Berg, G.J. (2001), Duration models: specification, identification, and multiple durations, in: J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics, Volume V*, North-Holland, Amsterdam.

Van den Berg, G.J., M. Lindeboom and F. Portrait (2006), Economic conditions early in life and individual mortality, *American Economic Review* 96, 290–302.

Van Dijk, H., J. Visser and E. Wolst (1984), Regional differences in social mobility patterns in the Netherlands between 1830 and 1940, *Journal of Social History* 17, 435–452.

Van Poppel, F. and K. Mandemakers (2003), Differential infant and child mortality in three Dutch regions, 1812–1912, Working paper, Indiana University.

Van Tulder, J.J.M. (1962). *De beroepsmobiliteit in Nederland van 1919 tot 1954. Een sociaal-statistische studie*, Stenfert Kroese, Leiden.

Vaupel, J.W. and A.I. Yashin (1985), The deviant dynamics of death in heterogeneous populations, in: N.B. Tuma (ed.), *Sociological Methodology 1985*, Jossey-Bass, San Francisco.

Wintle, M. (2000), *An economic and social history of the Netherlands, 1800–1920: demographic, economic and social transition*, Cambridge University Press, Cambridge.

Wolleswinkel-van den Bosch, J., F. van Poppel, E. Tabeau and J. Mackenbach (1998), Mortality decline in the Netherlands during the period 1850–1992: a turning point analysis, *Social Science and Medicine* 47, 429–444.

Appendix

Appendix 1   External evidence on the relation between economic well-being and social class

The first two datasets we consider are from the personal tax registers for the cities of Haarlem (1742) and Leiden (1748). At the time, the tax was to be paid by the head of the household and was supposed to be a progressive function of the consumption and income level of the household (although there were deductibles). Note that total consumption can be seen as a proxy of average income and as such it reflects long-run economic well-being. Consumption was not directly observed but was estimated from the number of household members and their age, the number of servants, the possession of expensive property like horses and boats, and the income pattern. The income was sometimes not directly observed, in which case it could have been estimated from the occupation or from multi-year averages. For self-employed individuals, taxes were also based on stocks of production inputs and outputs. From a taxation point of view, the main difference between 1742 and 1748 is that in 1742 the tax rates were determined by certain private persons within the local community, whereas in 1748 this was done by officials. Individuals who were deemed too poor to pay taxes are not included in the data. This may lead to an underrepresentation of low social class individuals. In particular, it may lead to an underrepresentation of low income individuals within the lower social classes, which in turn may lead to an under-estimation of the effect of social class on taxes.

The third dataset contains personal taxable income for 1378 individuals in Leiden in 1900 who (next to their main occupation) were a member of a board of an institution, community, or charity. For each individual the taxable income is observed. The board-membership restriction leads to an overrepresentation of high-income individuals. Again, this may lead to an under-estimation of the effect of social class on income.

All three datasets contain detailed occupation classifications. We omit a discussion of the monetary measurement units, since we are primarily interested in the monotonicity and the sign of the relation between social class on the one hand and taxes or income on the other.

In Table 10 we use the individual records from 1742 to compare the social class (as a function of occupation) to the mean of the "income bracket" of the individual tax payer. The "income bracket" is a monetary interval that determines the tax amount; the higher the bracket the more taxes are to be paid. We only

38

use those individuals who are in the fifteen most common occupations in the 1742 dataset. In Table 11 we use the individual records from 1748 to compare the social class (as a function of occupation) to the tax amount of the individual tax payer. Again, we only use individuals in the most common occupations. High social class individuals are underrepresented in the tables because their occupations are not very common. From the tables we conclude that the relation between social class and the tax indicators is positive.

For the 1748 dataset, Diederiks (1985) shows that there are very high rank correlations between occupational averages of (i) taxation, (ii) whether one owns a house, and (iii) whether one is labelled as "poor" by the tax agency.

In Table 12 we use the individual records from 1900 to compare the social class (as a function of occupation) to the mean taxable income level of the individual. We use individuals in all occupations, excluding those with multiple occupations and those with a missing income level or occupation. We conclude that the relation between social class and taxable income is positive.

Table 10: Social class and the mean of the "income bracket", in Haarlem, 1742

Social class	Average mean "income bracket"	# individuals
1	650	1
2	718	20
3	767	218
4	1399	49
5	–	–
6	–	–

Table 11: Social class and the tax amount, in Leiden, 1748

Social class	Average tax amount	# individuals
1	3.5	7
2	6.5	114
3	11.3	541
4	33.3	34
5	–	–
6	33.6	23

Table 12: Social class and taxable income, in Leiden, 1900

Social class	Average taxable income	# individuals
1	550	3
2	2169	42
3	2228	133
4	2916	54
5	5604	130
6	8378	157