

Economic Conditions Early in Life and Individual Mortality

Gerard J. van den Berg

Free University Amsterdam, IFAU-Uppsala, CEPR, IZA, IFS, Tinbergen Institute

Maarten Lindeboom

Free University Amsterdam, Tinbergen Institute, IZA

France Portrait

Free University Amsterdam, Tinbergen Institute

January 24, 2005

Abstract

This paper analyzes the effect of economic conditions early in life on the individual mortality rate later in life, using business cycle conditions early in life as an exogenous indicator. We have individual data records from Dutch registers of birth, marriage, and death certificates, covering an observation window of unprecedented size (1812-2000). These are merged with historical data on macro-economic and health indicators. We correct for secular changes over time and other mortality determinants. We non-parametrically compare those born in a recession to those born in the preceding boom, and we estimate duration models where the individual's mortality rate depends on current conditions, conditions early in life, age, individual characteristics, including individual socio-economic indicators, and interaction terms. The results indicate a significant negative effect of economic conditions early in life on individual mortality rates at all ages.

Address for correspondence: Gerard J. van den Berg, Department of Economics, Free University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. E-mail: gberg@feweb.vu.nl.

Keywords: death, longevity, health, business cycle, recession, life expectancy, lifetimes, epidemics.

We thank the Editor, two anonymous referees, Anne Case, Angus Deaton, Gabriele Doblhammer, Jutta Gampe, Paul Gregg, Sören Hojgard, Jim Oeppen, Chris Ruhm, and participants at seminars at Princeton University and the Max Planck Institute for Demography (Rostock) and conferences on health and population economics and econometrics, in Lund, New York, and Stockholm, for helpful comments and suggestions. Márta Lopez provided excellent help with the empirical analyses. The "Historical Sample of the Netherlands Dataset release UZF.02" was kindly provided by the International Institute of Social History (IISG) in Amsterdam. A previous version of the paper was titled "Individual Mortality and Macro-Economic Conditions from Birth to Death".

1. Introduction

Socio-economic conditions during infancy and early childhood years may affect mortality later in life. It is well documented that poor living conditions early in life are associated with susceptibility to a wide range of health problems later in life (see e.g. Case, Fertig and Paxson, 2003, and references therein). In bad times, the provision of sufficient nutrients and good living conditions for children and pregnant women may be hampered, and exposure to disease or malnutrition early in life may lead to an increase in mortality. Such effects have important policy implications. If being born in bad times has a positive long-run effect on mortality then total lifetime earnings are reduced for those affected.¹ This would increase the benefits of policies that help young children and pregnant women in recessions, for example by way of enhanced provision of food, housing, and health care (see Almond, 2002, for a detailed discussion of policy implications).

The above-mentioned effects have been difficult to analyze empirically. First, data often lack a sufficiently long time span to cover full lifetimes of individuals. Secondly, individual socio-economic conditions during early childhood and health outcomes later in life may be jointly affected by unobserved heterogeneity, leading to simultaneity bias.

This paper analyzes the effects of economic conditions early in life on the individual mortality rate, by using micro data with an observation window of unprecedented size, and by using macro-economic conditions early in life as an “instrument” for the individual conditions. Macro-economic conditions do not give rise to the aforementioned simultaneity bias because they are exogenous at the individual level. We separate cyclical fluctuations from secular trends in macro-economic conditions, and to capture the individual conditions early in life we focus on the cyclical fluctuations early in life. We control for individual-specific background characteristics, contemporaneous conditions, and other individual mortality determinants.²

¹ Murphy and Topel (2003) demonstrate that the gains from mortality rate reductions have been enormous.

² In demography, natural experiments have recently been applied to test effects of nutrition and disease exposure in utero and during the first months of life on mortality later in life, using longitudinal data. These studies are primarily focused on medical explanations and not on economic conditions. For example, Doblhammer (2004) uses month of birth, and she surveys studies that use epidemics, wars, famines, or the rate of infant mortality in the cohort. The effect is often significant. Another related branch of literature focuses on the instantaneous effect of the business cycle on health and mortality (rather than the long-run effects we consider) in more

Our data set, called the Historical Sample of the Netherlands (HSN), covers around 14,000 individuals born in the Netherlands in the period 1812-1912 and followed up to 2000. It includes information taken from the standardized individual recordings of vital events (birth, marriage and death) kept by municipalities and provinces. 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. Also, the sample size precludes a detailed analysis of the effects of specific policy measures like the abolition of child labor.

The empirical analysis basically consists of two stages. First, we make non-parametric comparisons between the lifetimes of individuals, notably of those born in a boom and those born in the recession that directly follows the boom. Secondly, 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, and other characteristics. The main outcomes are in strong agreement. The state of the business cycle at birth affects mortality later in life. This supports the view that economic status early in life is a key determinant of health and mortality in adulthood.

The paper is organized as follows. Section 2 presents the data and discusses variables that we use in the analyses. Section 3 provides non-parametric analyses. Section 4 covers the estimation of the duration model, including sensitivity analyses, and it outlines policy implications of the results. 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

recent years. Brenner (1979) and Ruhm (2000) consider instantaneous effects among adults. Dehejia and Lleras-Muney (2004) consider instantaneous effects among children as well as selection effects on the composition of newborns. We return to their results below.

the provinces of Utrecht, Friesland and Zeeland³ 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 1903. We observe dates of death of migrants out of the provinces. For some individuals in the sample born before 1903, 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 9276 individuals.

We map the occupation of the father into a social class code ranging from 1 (lower lower) to 6 (upper upper).⁴ The place of residence at birth is translated into a binary urbanization indicator (being 1 iff the individual is born in a city). Table 1 gives summary statistics for the explanatory and endogenous variables.⁵ The lifetime statistics are for individuals whose date of death is observed.

We should note that the mortality rate as a function of age displays the familiar U-shaped curve. Among the uncensored observations, 25% died within a year. For ages up to 5, the mortality rate is highest among the cohorts born in the middle of the 19th century.

³ At the time, the Netherlands had 11 provinces. Our three provinces were jointly rather representative of the Netherlands in terms of economic activity. Aggregate mortality rates in our data display similar patterns as historical national averages.

⁴ Codes 1 to 6 represent occupations like: diker, day laborer, dock worker (1), cow milker, beer brewer, florist (2), potato farmer, barber (3), bailiff, corn dealer, miller (4), factory manager, headmaster, infantry captain (5), and auditor, lawyer, pharmacist, surgeon (6), respectively.

⁵ See e.g. 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.

2.2. A brief overview of historical developments related to mortality

Before we discuss the data on macro-economic conditions and historical events, we give a brief outline of the historical and economic developments related to mortality in the Netherlands in the era we consider. (Note however that the empirical analysis is not primarily concerned with secular developments but rather with cyclical fluctuations.)

The 19th and 20th century witnessed a transition from a rather poor agricultural society with high mortality rates to a rich society with a large service sector and low mortality rates. In the first half of the 19th century, the growth rate lags behind those of neighboring countries. Large-scale industrialization did not take off before the third quarter of the 19th century, which is relatively late (Mokyr, 1974, de Vries, 2000).

At 1830, large epidemics had been absent for around a century, but the period 1830-1875 contains a number of dramatic epidemics, exacerbated by the rise in urbanization.⁶ The cholera epidemics of 1848-49 and 1866-67 and the smallpox epidemic of 1870-72 each led to over 20,000 deaths nationwide (about 0.7% of the population). At the height of such epidemics, the national annual mortality rate was around 25% higher than otherwise. There were frequent smaller cholera epidemics. In towns and cities in particular, living standards were often poor, sanitary conditions were bad (de Swaan, 1988), and infectious and parasitic diseases were primary causes of death (see Woods, Watterson and Woodward, 1988, and Wolleswinkel-van den Bosch et al., 1998, on the strong negative health effects from early industrialization and urbanization in Britain and the Netherlands). After 1875, mortality rates dropped spectacularly. This is caused by better availability of food, the dissemination of nutritional, hygienic and medical knowledge among the population, and large improvements in public health provisions like sewage and water supply (Wolleswinkel-van den Bosch et al., 1998).

In the first half of the 20th century, mortality is further reduced. The improvements in socio-economic environment⁷ and the more intensified tuberculosis control are thought to be the most important determinants (see Wolleswinkel-van den Bosch et al., 1998). In the last decades of the 20th century, mortality rates reduce further, at the particular advantage of older age groups. This is due to improvements in medical technology and medical care (Cutler and Meara, 2001).

⁶ The single notable later epidemic is the influenza epidemic of 1918.

⁷ There was a period of rapid growth after the first world war, which ended with the depression of the 1930s (Van Zanden, 2000).

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.⁸ 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. Figure 1 plots the log annual real per capita GNP over the interval in which the sample members are born. 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 and the share of agricultural production in annual GNP (see also Smits, Horlings and van Zanden, 2000).

To capture the long-run effects of conditions early in life, we compare individuals born and grown up in good times to those born and grown up in bad times, following the line of thought that a high GNP goes along with better individual economic conditions for many individuals. However, for our purposes, good and bad times are not well represented simply by the level of GNP. An era with high GNP also leads to innovation and investment in hygiene and health care, which decreases mortality later in life for those born in this era. These can be labeled secular effects. The effects of cyclical movements in GNP, contrary to the effect of the trend value or current level of GNP, are not dominated by these secular developments. At the individual level, cyclical changes often involve unexpected income shocks. Indeed, the turbulence created by the cycle may itself lead to health problems (we return to this in Subsection 4.4).

Ideally one would like to compare cohorts born in booms to those born in recessions⁹ with otherwise identical circumstances throughout life. This is unfeasible due to the secular improvements 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

⁸ Details are on the Data Replication Appendix of the American Economic Review web site.

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. 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.

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 and numbers of medical innovations.

3. Descriptive and non-parametric analyses

We start with 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. To prevent that the latter are smaller simply due to instantaneous effects of economic conditions on infant mortality, we condition on survival until age 2 or until age 5. We require the boom and recession not to include an epidemic, and we require both to last at least three years, in order to have reasonable sample sizes. We also ignore the period after 1881 because the strong upward trends in GNP and lifetimes in that era may dominate the comparison.

⁹ With some abuse of language, we use the terms boom and recession to denote years in which current log GNP is above or below its trend level. See Subsection 4.1 for the formal definition used.

Consider the boom of 1872-1876 and the subsequent recession of 1877-1881. Among those born in the boom, the average of $T|T>2$ equals 66.0 years, whereas among those born in the recession it equals 62.5 (both standard deviations are 1.5 years). So, despite the secular trend that works to the advantage of those born later (i.e. in the recession), the latter have shorter conditional lifetimes. Formally, we perform a one-sided test of the null hypothesis of equal means against the alternative that the mean in the recession is smaller. The test statistic value equals 1.7, which exceeds the 5% critical value of the asymptotic standard normal distribution of the statistic under the null. For $T|T>5$ the results are similar. The averages equal 70.8 (1.2) and 67.5 (1.3), respectively, so the test statistic equals 1.9, which is again significantly different from zero. For other boom-recession combinations that we considered, the conditional average lifetimes are again smaller in the recession, but the tests do not result in rejection of the null hypothesis.

The above tests do not exploit all available data. To proceed, we perform a regression of (log) lifetimes on the binary boom (versus recession) indicator at birth, where we include 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. We need to restrict attention to birth years before 1896 because the last birth years in the data (1896-1902) constitute a boom. As above, we only consider individuals who survive beyond age 2 or age 5. We ignore individuals born in years with epidemics. We expect that the strong upward trend in lifetimes after 1881 affects the parameter estimate for the boom indicator at birth, so we interact it with a dummy for the period after 1881.

The regression results are in Table 2 for the case where we condition on $T>2$. The results for $T>5$ and for regressions of log lifetimes are qualitatively similar. The coefficient of the boom indicator at birth is significantly positive, implying that individuals born in recessions have higher long-run mortality rates than those born in better times. Note that, as above, the test is conservative to the extent that it does not take account of secular improvements within birth periods which necessarily benefit those born later (i.e. in the recession). The interaction coefficient of the boom indicator and the dummy for Birth Period VIII (which is the period after 1881) is as expected significantly negative.

We also perform this regression on the subsample of the three birth periods in which the business cycle fluctuates the most (as defined by the maximum amplitude of the cyclical component or by the average of this component in the boom minus the

average in the recession, leading in both cases to Birth Periods II, VI and VII). The results are also in Table 2.¹⁰ Clearly, the coefficient of the boom indicator at birth is larger and more significant than above, which lends more weight to our conclusions.¹¹

A disadvantage of the analyses so far 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 given conditions during later childhood years. The duration analysis also controls for individual characteristics and it incorporates infant mortality and secular trends within “birth periods”, albeit at the expense of functional-form model assumptions.

4. Duration analysis

4.1. 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 economic conditions. As age is measured in days, we take it to be a continuous random variable. Let the variable 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

¹⁰ We do not report the birth-period coefficients because they are virtually the same as for the baseline regression.

¹¹ Similarly to this, we detrended the time series of the average log conditional lifetime per birth year and regress it on the detrended log GNP. To reduce the noise due to measurement errors we time-aggregate the detrended terms by using averages over five-year intervals. The estimated regression coefficient is positive, and the correlation coefficient equals 0.33 if we consider $E(\log T|T>2)$ and 0.34 if we consider $E(\log T|T>5)$. Time-aggregation into smaller than five-year intervals leads to smaller but still positive correlation coefficients.

trend components and cyclical indicators $z_{tr}(\tau-t+i)$ and $z_c(\tau-t+i)$ of macro-economic conditions earlier in life ($i \in \{0, \dots, t-1\}$), and various interaction terms. The z variables are only measured once a year. In obvious notation,

$$\lim_{dt \downarrow 0} \frac{\Pr(t \leq T < t + dt \mid T \geq t, \tau, X = x)}{dt} = \theta(t, x, z(\tau), z_{tr}(\tau-t+i), z_c(\tau-t+i), \text{ with } i \in \{0, \dots, t-1\})$$

As our basic specification we take $\log \theta$ to be linear in the age dependence function of t , in x , $z(\tau)$, $z_{tr}(\tau-t)$, and, if $t \geq 1$, in $z_c(\tau-t)$ at birth and the average of $z_c(\tau-t+i)$ over i in the childhood age set $\{1, \dots, \min\{6, t-1\}\}$. This specification requires some discussion. First consider the cyclical indicators early in life. These are represented by the cyclical components of the decomposition of log GNP from Subsection 2.3, at the birth year, and averaged over the years corresponding to age 1 up to 6. At $t=0$ these do not affect θ . At $t=1$, only $z_c(\tau-1)$ affects θ . At $t=3$, $\log \theta$ depends on $z_c(\tau-3)$ and on the average of z_c in $\tau-2$ and $\tau-1$. To reduce disturbances due to measurement error in the values of the cyclical components we summarize them into two dummy variables.¹² Specifically, for the cycle in the birth year we use a binary boom/recession indicator, and for the cyclical conditions in the childhood age set we use a binary variable that captures whether the average cyclical component in this set is positive or not.

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

For the age dependence we adopt a piecewise constant specification with 10 different intervals (0, 1, 2-6, 7-14, 15-34, 35-59, 60-69, 70-79, 80-89, 90+). Concerning 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. For $z(\tau)$ we take log annual real per

capita GNP at τ , 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.

We also estimate more general specifications allowing for various interaction effects between the mortality rate determinants. We are particularly interested in interactions between t and $z_c(\tau-t)$, which arise if the mortality effect of economic conditions around birth varies over the course of a lifetime. By now there is substantial medical and epidemiological evidence that malnutrition of the mother during the final stages of pregnancy leads her child to have a particularly high incidence of health problems at ages over 50 (see e.g. Koupilová, 1997). We also fully interact the age dependence and current log GNP. Babies may suffer disproportionately 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.

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. In the empirical analysis we also estimate models in which an unobserved heterogeneity term is added to the expression for the log individual mortality rate. We assume that this term has a flexible discrete mass-point distribution in the population of newborns. In Subsection 4.3 we briefly address whether the composition of the population of newborns depends on macro-economic conditions.

¹² If a boom or recession defined this way consists of only one year then it is subsumed in the surrounding recession or boom. Details are on the Data Replication Appendix.

4.2. Estimation results

Table 3 presents the estimation results for the basic model specification, with and without interactions between t and $z(\tau)$. The estimates concern the mortality rate, so a positive value is associated with a shorter lifetime. The time unit is one year.

The most striking result is that the cyclical indicator in the birth year has a significantly negative effect on the mortality rate later in life. We take this as evidence that at the individual level, economic conditions at birth have long lasting effects on mortality. Selection due to unobserved heterogeneity cannot 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, but it remains significantly negative. It implies that the state of the cycle in the birth year gives rise to a difference of up to 7% in the mortality rate after the first year of life. As we shall see in Subsection 4.4, this is roughly comparable to the effect of 10 to 15 years of additional secular developments, so in that sense it is comparable to being born 10 to 15 years later in time. The cyclical indicator at ages 1 up to 6 does not have a significant effect on mortality later in life. One may argue that may be because of a negative causal effect and a positive selection effect. However, the result does not change when allowing for unobserved heterogeneity (see below).

Current log annual real per capita GNP has a significantly negative estimated instantaneous effect. Recall that this effect incorporates all secular effects from birth to the current age. The coefficient should therefore not be interpreted as an (average) individual income elasticity.¹³ The size of the effect varies with age. 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 15 and 34. Again, 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. The effects of other epidemics are insignificant and are omitted (see below for more details).

¹³ See also Pritchett and Summers (1996) for a discussion on this.

Women, individuals from a higher social class, and individuals from rural areas have lower mortality rates than their counterparts. The finding with respect to social class is consistent with the large literature on the health-income gradient.¹⁴ Individuals who are born in Friesland have a much lower mortality rate. This is a well known result (see e.g. van Poppel and Mandemakers, 2003) and is explained by the high prevalence of breastfeeding in Friesland and the bad quality of water in the other two provinces. The mortality rate is a standard U-shaped function of age. Note that this cannot be simply read off from the coefficients of the piecewise age dependence function, because the effects of the cyclical indicators are age-varying, and because of the interactions between t and $z(\tau)$.

The estimation results allowing for unobserved heterogeneity show that the variance of the unobserved heterogeneity term is significantly positive. However, the estimates of the other parameters are similar to those without. Sign, order of magnitude, and significance of the main coefficients are unaffected.

4.3. Sensitivity analyses

Rather than giving full sets of parameter estimates, we discuss the distinguishing features of each analysis, focusing on the parameters of interest. First, we re-estimate the models with samples in which the first 10-year or 20-year birth cohorts are fully omitted (recall that these have higher rates of unobserved death dates). The results are very similar to those above. We also estimate model specifications with alternative early-childhood age sets. For example, we consider cyclical indicators for the age sets $\{1, \dots, \min\{h, t-1\}\}$, with $h=1, \dots, 5$ as a determinant of the mortality rate at t . The estimates of the effects are similar to the reported estimate with $h=6$, suggesting that the even the actual state of the cycle between age 1 and 2 does not have important long-run effects on mortality. We also estimate specifications where the dummy cyclical indicators are replaced by the actual (average) values. This does not affect the conclusions either. Cyclical indicators for the age set $\{7, \dots, \min\{14, t-1\}\}$ and the cyclical indicator for the year prior to birth turn out to be insignificant. The estimation results are insensitive to changes in the age intervals in the age dependence function (for example, if one divides the age interval 35-59 into two

¹⁴ Using U.S. data from the mid 19th century, Steckel (1988) finds that individual wealth as measured by the value of real estate owned by the household head does not have a significant effect on the mortality of infants, children, or adult women. However, as Steckel (1988) notes, this wealth variable is imperfect, and, like our social class variable, it may be confounded. The household head's occupation level has a somewhat stronger effect on infant mortality.

intervals 35-49 and 50-59), except, not surprisingly, if one aggregates age intervals at low ages. If we include $z_{it}(\tau-t)$ in the model then the corresponding coefficient as well as the coefficient on log GNP are non-sensical, confirming that there is insufficient independent variation of these two variables in the data.

Next, we estimate a model in which the effect of the cyclical indicator at birth is allowed to vary with the age of the individual. Specifically, we interact it with an indicator of whether age exceeds 50. The interaction coefficient is insignificantly different from zero. This suggests that bad economic conditions during childhood do not have a weaker effect on the mortality rate at prime ages, although this should be taken with some caution, first of all because of the small number of observed deaths between ages 5 and 50, and secondly because the model specification at high ages is not very elaborate. Note that the absolute effect on mortality at high ages is larger simply because mortality itself is larger at high ages. The interaction of the effect of the cyclical indicator at ages 1 up to 6 with high age is not significant either. These results do not depend on whether we use the actual (averages of the) cyclical indicators or their rounded-off dummy variables.

Next, we consider other indicators of macro-economic conditions at birth, notably the inflation rate. Its coefficient is insignificant and close to zero. In fact, inflation seems to be unrelated to movements in real annual per capita GNP. The share of agricultural production in the annual GNP at birth has a positive coefficient, but this variable contains cyclical as well as trend components, so that its effect is hard to interpret.

We estimate separate models for men and women. Most parameter estimates are similar to each other and to those reported in Table 3. However, the cyclical indicator at birth is now insignificant for women, while it is still significant for men.¹⁵ This suggests that men suffer relatively heavily from adverse economic conditions at birth.¹⁶ Conversely, female mortality is significantly more dependent on their parents' social class.

¹⁵ This result is confirmed by non-parametric analyses by gender.

¹⁶ This is in agreement to some other recent evidence. The landmark medical study by Leon et al. (1998) on the relation between birth weight and mortality from heart disease finds stronger effects for men. Case and Paxson (2004) provide some evidence that men with certain health problems that might originate earlier in life are more likely to die than women with the same conditions. Also, in most countries examined, the long run mortality effects of the season of birth are smaller for women than for men (Doblhammer, 2004), although these differences are sometimes rather small.

We also estimate models with an additional dummy variable for 1866 in $z(\tau)$, because of the nationwide cholera epidemic in that year, but the corresponding coefficient is insignificantly positive and very small, so that it can be concluded that the epidemic did not have a sizeable effect on mortality.

Finally, we estimate models that include dummy variables for the season of birth. These turn out to have insignificant effects. At spring, summer, and fall, the estimated effects on the log mortality rate equal 0.02, 0.01, and -0.01 (standard errors 0.03).

Dehejia and Lleras-Muney (2004) show that in the late 20th century the composition of newborns varies with the stage of the cycle at birth. For example, in a recession in that era, among whites, more children are born with low-educated parents, i.e. with lower individual expected lifetimes. We check on this hypothesis in two ways. First, we examine how the number of newborns varies over the cycle, the argument being that a relation between these makes it more likely that the composition changes as well. Secondly, we examine whether the composition in terms of social class varies over the cycle. It turns out that in both cases we find no relation at all in our data. The slope coefficient of the regression of the log number of births in a birth year on the cyclical component of that year equals 0.13 (with the cyclical component and the log number of births being dispersed in $(-0.1, 0.1)$ and $(3.9, 5.1)$, respectively). The slope coefficient of the regression of average social class in a birth year on the cyclical component of that year equals 0.00. There is also no relation with the proportion of newborns from higher social classes.

To conclude this subsection, the main estimation results are qualitatively and quantitatively robust with respect to a wide range of assumptions.

4.4 Implications and policy

From the second set of results in Table 3, being born in a recession is associated with a 7% increase in the mortality rate after the first year of life. On average, around 25% of individuals die in their first year of life, and the total mean lifetime decreases by about 5%. This is in line with the non-parametric results. Using the results from Table 3 again, a 5% increase in the mean lifetime can be associated with a 12% increase in current GNP including the effects of all past secular developments. From Figure 1, it follows that this corresponds to being born on average 12 years later in time, that is, to benefiting from 12 years of additional secular developments.

It is possible that the state of the business cycle at birth has a smaller or different long-run effect in the 20th century. Dehejia and Lleras-Muney (2004) show that being born in a recession in the late 20th century improves the health of small children because parents' health-related behavior improves, for example because their opportunity cost of time reduces.¹⁷ This is an instantaneous effect, but it could suggest that long-run effects are different as well. However, recall that our primary interest is in the effect of individual economic conditions at birth on mortality later in life. In the 19th century, job loss or income loss could lead to starvation. There were no unemployment benefits and welfare programs to sufficiently safeguard household income. Childbearing was difficult to plan in advance. In this context, cyclical variation can be used to address our issue of primary interest. At the individual level, the presence of an effect of economic conditions at birth on mortality later in life does not depend on the effect of the cycle on these conditions.

It could also be argued that long-run effects of early childhood conditions in Western societies were more important in the 19th century than they are now because of the shift in the mortality spectrum from infectious diseases to chronic diseases. Our present data do not enable us to address this. However, the recent epidemiological literature using natural experiments demonstrates that cohort effects on health and mortality later in life are also significant in the 20th century (see e.g. Almond, 2002, and the survey in Doblhammer, 2004). Moreover, our results are confirmed by recent studies of the effects of observed individual socio-economic conditions on health outcomes like illness indicators later in life. These studies point towards childhood conditions as crucial determinants of health later in life (see e.g. Case, Fertig and Paxson, 2003).

The results indicate that from a policy point of view it is particularly useful to focus on children aged zero in bad economic conditions. The contemporaneous mortality of these children, as well as mortality later in their life, may be significantly reduced if their conditions are improved upon, for example by way of food, housing, and health provision.

Recessions in the 19th century may also have induced increased insecurity and stress (Brenner, 1979, shows that this effect is also present in the (mid) 20th century). This may also lead to higher long-run mortality rates for newborns. This is a long-run effect, from business cycles, by way of turbulence of the parents' economic conditions, to

¹⁷ This is in line with Ruhm (2000)'s finding that current mortality among adults themselves is

stress, to mortality. This effect may be relative in the sense that its magnitude depends less on the baseline income level than on the turbulence. The presence of such an effect provides a novel reason for the government to stabilize the macroeconomy.

5. Conclusions

Using data covering the period 1812-2000, our empirical analysis demonstrates a significant effect of the state of the business cycle at birth on the individual mortality rate later in life. On average, and keeping everything else constant, an individual born in a recession lives a few years less than an individual born in a boom. We take the significant effect of cyclical conditions at birth as evidence of a causal negative effect of individual economic conditions early in life on the mortality rate later in life. Note that the former is used to investigate the presence of the latter, but that the former effect is not necessary for the latter to be present. In this sense the relevance stretches beyond the birth cohorts in the data. Indeed, the results are in accordance to other studies on the significance of early life conditions. The results imply that from a policy point of view it is particularly useful to focus on children aged zero in bad economic conditions. The contemporaneous mortality of these children, as well as mortality later in their life, may be reduced if their conditions are improved upon, for example by way of food, housing, and health provision. To the extent that the cyclical turbulence creates stress among parents which in turn creates high mortality later in life among their infants, the results in this paper can be viewed as a justification for macroeconomic stabilization policy.

pro-cyclical in the late 20th century.

References

- Almond, D.V. (2002), Cohort differences in health: a duration analysis using the National Longitudinal Mortality Study, Working paper, University of Chicago.
- 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).
- Brenner M. (1979), Mortality and the national economy: a review and the experience of England and Wales, 1936-1976, *The Lancet* 8142, 568-573.
- Case, A., A. Fertig and C. Paxson (2003), From cradle to grave? The lasting impact of childhood health and circumstance, Working paper, Princeton University.
- Case, A. and C. Paxson (2004), Sex differences in morbidity and mortality, Working paper, Princeton University.
- Cutler, D.M. and E. Meara (2001), Changes in the age distribution of mortality over the 20th century, NBER Working paper 8556, Cambridge.
- Dehejia, R. and A. Lleras-Muney (2004), Booms, busts and babies' health, *Quarterly Journal of Economics* 119, 1091-1130.
- Doblhammer, G. (2004), *The late life legacy of very early life* (Springer, Berlin).
- Jacobs, J. and J.P. Smits (2001), Business cycles in the Netherlands, 1815-1913, Working paper, Groningen University.
- Koupilová, I. (1997), Fetal Growth, Social Factors and Circulatory Diseases, *Acta Universitatis Upsaliensis* 698, 7-60.
- Leon, D.A., H.O. Lithell, D. Vågerö, I. Koupilová, R. Mohsen, L. Berglund, U.B. Lithell and P.M. McKeigue (1998), Reduced fetal growth rate and increased risk of death from ischaemic heart disease: cohort study of 15000 Swedish men and women born 1915-29, *British Medical Journal* 317, 241-245.
- 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.
- Murphy, K.M. and R. Topel (2003), The value of health and longevity, Working paper, University of Chicago.
- van Poppel, F. and K. Mandemakers (2003), Differential infant and child mortality in three Dutch regions, 1812-1912, Working paper, Indiana University.
- Pritchett, L. and L.H. Summers (1996), Wealthier is healthier, *Journal of Human Resources* 31, 841-868.
- Ruhm C.J. (2000), Are recessions good for your health?, *Quarterly Journal of Economics* 115, 617-650.
- Smits, J.P., E. Horlings and J.L. van Zanden (2000), Dutch GNP and its components, 1800-1913, Working paper, Groningen University.
- Steckel, R.H. (1988), The health and mortality of women and children, 1850-1860, *Journal of Economic History* 48, 333-345.
- de Swaan, A. (1988), *In Care of the State; State Formation and Collectivization of Health Care, Education and Welfare in Europe and America in the Modern Era* (Oxford University Press, New York).
- 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).
- de Vries, J. (2000), Dutch Economic growth in comparative historical perspective, 1500-2000, *De Economist* 148, 443-467.
- 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.
- Woods, R., P.A. Watterson and J.H. Woodward, (1988), The causes of rapid infant decline in England and Wales, 1861-1921, *Population Studies* 42, 343-366.
- van Zanden, J.L. (2000), Post-war European economic development as an out of equilibrium growth path: the case of the Netherlands, *De Economist* 148, 539-555.

Table 1. Descriptive statistics

<i>variable</i>	<i>average</i>	<i>standard deviation</i>
Female (%)	49.8	
Social class (range 1-6):	2.29	1.3
Father not illiterate (%)	84	
Mother unmarried at birth (%)	2.0	
Share agriculture in GNP (%)	24.46	3.3
Log real per capita GNP at birth	7.41	0.26
Lifetime duration men: median=35.2	38.1	35.2
Lifetime duration women: median=41.2	40.8	35.9

Note: annual real per capita GNP is measured in 1,000 Euros with 1995 as base year. Reported moments and medians are based on uncensored lifetimes only.

Figure 1: log annual real per capita GNP

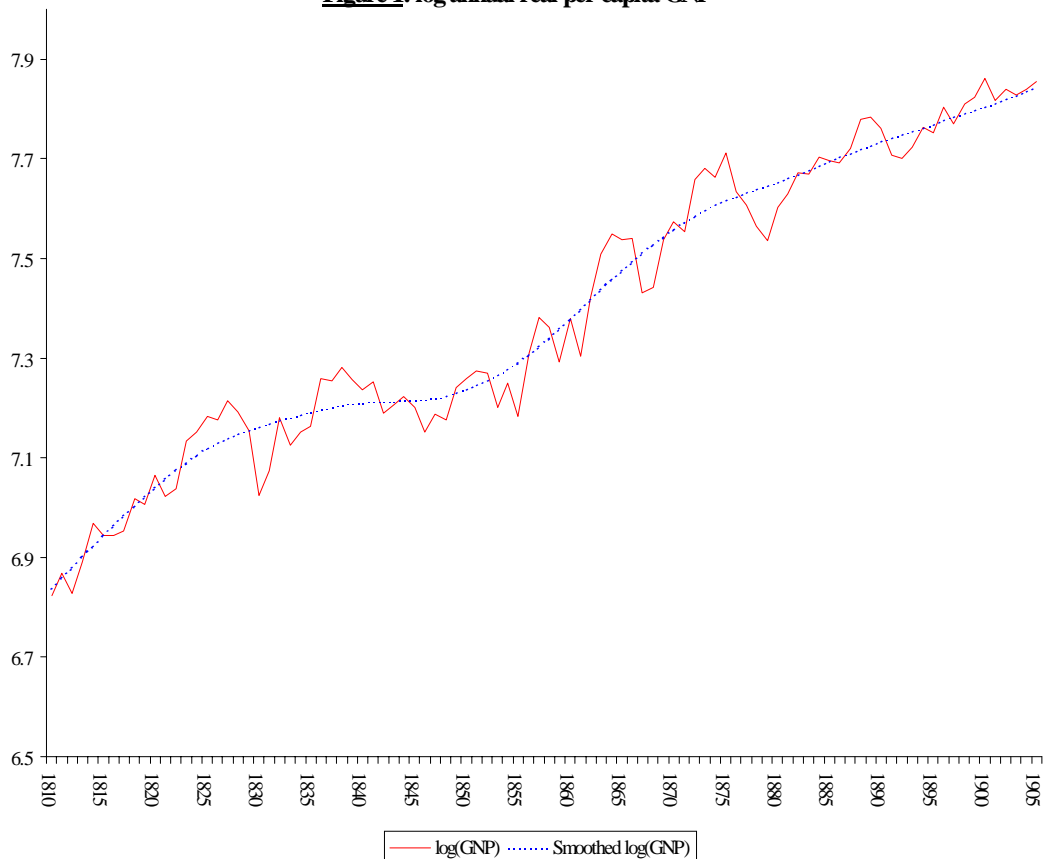


Table 2. Parameter estimates of regressions of conditional lifetimes on the binary boom/recession indicator at birth and the birth period dummy variables.

Variable	estimate	standard error
<i>Baseline regression:</i>		
Boom (instead of recession) at birth	1.58	0.95*
Birth Period I	48.92	1.29
Birth Period II	46.42	1.18
Birth Period III	45.24	1.17
Birth Period IV	46.82	1.57
Birth Period V	48.72	1.80
Birth Period VI	57.05	1.33
Birth Period VII	63.19	1.18
Birth Period VIII	68.68	1.40
Interaction Boom at birth × Birth Period VIII	-5.16	2.01
# of individuals	4774	
<i>Three birth periods with strongest cycle only:</i>		
Boom (instead of recession) at birth	2.36	1.33*
# of individuals	1885	

Notes: * denotes significance at the 5% level according to a one-sided test of a zero effect of the boom/recession indicator at birth. In both regressions, the outcome variable is the individual lifetime conditional on survival beyond age 2. The explanatory variables in the second regression include dummy variables for Birth Periods II, VI, and VII.

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

Variable	estimate	t-statistic	estimate	t-statistic
<i>Individual background characteristics</i>				
Female	-0.106	4.6	-0.107	4.6
Social class father at birth	-0.029	3.0	-0.030	3.0
Father is not illiterate	-0.072	2.2	-0.068	2.0
Mother was unmarried at the time of birth	0.074	0.9	0.077	0.9
Born in urban area	0.070	2.5	0.071	2.6
Born in province Utrecht*	0.662	20.8	0.662	20.6
Born in province Zeeland*	0.413	12.9	0.416	12.9
<i>Business cycle early in life</i>				
Boom (instead of recession) at birth	-0.086	3.1	-0.069	2.5
Cycle indicator for age 1 up to 6	0.002	0.1	0.005	0.2
<i>Contemporaneous macro conditions</i>				
1849 cholera in Utrecht	0.223	1.3	0.159	0.9
1870/1 smallpox	0.618	6.2	0.623	6.2
1918 influenza	-0.283	1.6	-0.191	1.1
World War II (GNP missing)	-2.629	9.9	-1.808	5.9
Current log(annual real per capita GNP)	-0.341	11.1		
idem at age 0			-0.251	2.8
idem at age 1			-0.399	2.1
idem at age 2-6			-0.837	4.7
idem at age 7-14			-1.274	4.7
idem at age 15-34			-1.395	9.2
idem at age 35-59			-0.230	5.5
idem at age 60-69			-0.245	6.2
idem at age 70-79			-0.239	6.6
idem at age 80-89			-0.260	7.2
idem at age 90+			-0.261	4.6
<i>Age</i>				
Age 0	1.006	4.4	0.339	0.5
Age 1	-0.255	1.1	0.166	0.1
Age 2-6	-1.535	6.7	2.131	1.6
Age 7-14	-2.749	11.6	4.201	2.1
Age 15-34	-2.589	11.0	5.409	4.7
Age 35-59	-2.236	9.1	-3.129	9.6
Age 60-69	-1.015	4.0	-1.822	5.7
Age 70-79	0.092	0.4	-0.782	2.6
Age 80-89	1.069	4.0	0.337	1.1
Age 90+	1.855	6.6	1.106	2.2
- log likelihood	28885		28837	
# of individuals	9276		9276	

Notes: Effects on log mortality rate reported. GNP-related variables are not mean-centered.

*: Province Friesland as reference.