Discussion Papers
in Economics

No. 01/2024

Transition to Motherhood: The Role of Health

Irina Simankova
FAU Erlangen-Nürnberg

Harald Tauchmann
FAU Erlangen-Nürnberg

ISSN 1867-6707

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Economics
https://www.iwf.rw.fau.de/discussion-papers/

Transition to Motherhood:

The Role of Health*

Irina Simankova

Universität Erlangen-Nürnberg

Harald Tauchmann

Universität Erlangen-Nürnberg

RWI – Leibniz Institut für Wirtschaftsforschung

CINCH – Health Economics Research Center

Tuesday 14th May, 2024

Preliminary working draft, please do not quote or cite without authors' permission.

Abstract

The age at which women become mothers for the first time is ever increasing in many industrialized countries. Therefore, fertility determinants that might deteriorate with age, such as health, and their effect on reproductive patterns, should be given more attention. We explore the effect of the subjective general health of women of reproductive age on the conditional probability of entering motherhood. Based on estimating linear discrete-time hazard models using survey data from Germany, we do not find a homogeneous health effect on the probability of having a first child. However, allowing effect heterogeneity over the span of reproductive age reveals that the role of health is ambiguous. While good health is associated with a lower probability of entering motherhood at the beginning of the reproductive phase, the opposite holds for the late reproductive phase. This pattern is robust to employing different estimation methods (parametric, non-parametric), conditioning on socio-economic characteristics, and taking unobserved individual-level heterogeneity into account.

JEL codes: C41, I19, J13
Keywords: motherhood, fertility, discrete-time survival analysis, instrumental variables estimation

*[*] **Address for correspondence:** Irina Simankova, Professur für Gesundheitsökonomie, Findelgasse 7/9, 90402 Nürnberg, Germany. Email: irina.simankova@fau.de. Phone: +49 (0) 911 / 5302 96291. **Acknowledgments:** We thank Franz Josef Zorzi, Maximilian Pöhnlein, and participants of annual meeting of the VfS Health Economics Study Group 2023, ZEW-FAU Health Economics Workshop 2023, and 36th BGPE Research Workshop for many valuable comments and suggestions. Excellent research assistance from Elena Yurkevich is gratefully acknowledged. Occasionally, the writing was enhanced with the assistance of AI applications DeepL, Grammarly and ChatGPT, specifically focusing on linguistic improvements rather than altering the context or factual content. **Conflict of Interests Declaration**: The authors have no conflicts of interest to declare.

1 Introduction

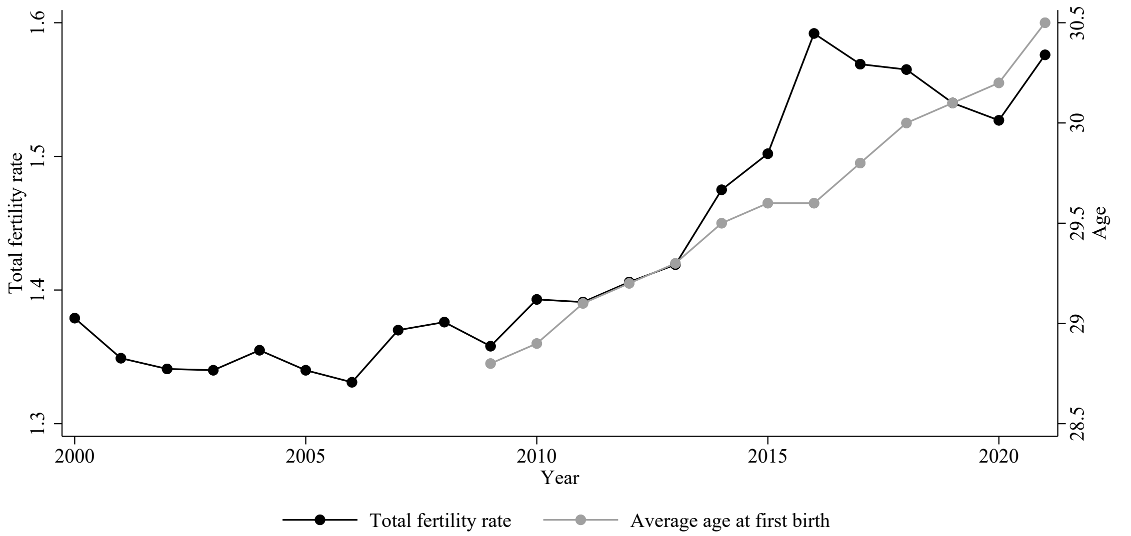
Since the mid-1970s, Germany – like many other industrialized countries – has been experiencing a transformation of reproductive behavior, resulting in fertility rates that persistently fall short of the replacement fertility rate of 2.1 (Van De Kaa, 1987). This pattern can be attributed to the invention of highly effective birth control measures and changes in women's preferences (Lesthaeghe & Surkyn, 1988). One facet of the change in reproductive behavior is an ever-advancing age at the birth of the first child. Delayed childbirth results in the deceleration of population growth and, consequently, changes in the demographic structure of society that might lead to severe economic consequences, e.g. challenges for the financing of the social security systems (Schleutker, 2014). Moreover, postponement of motherhood is associated with poor health outcomes for both mother and child, and a higher rate of mortality (Cnattingius, 1992). While the total fertility rate per woman has been slowly re-increasing in Germany since the beginning of the 2000s (Statistisches Bundesamt, 2023b), the average age of women at the first birth keeps rising (Statistisches Bundesamt, 2023c) (Figure 1). In other words, women tend to postpone motherhood and choose to remain temporarily childless for a longer period of time. This behavioral pattern may put them at an increased risk of becoming lifelong childless, possibly unwillingly. It is, nonetheless, crucial to better understand the interplay between the timing of motherhood and socio-economic conditions that may change over time, health in particular. For instance, a better understanding of the possible consequences of postponed motherhood is valuable for making decisions even at younger stages in life. Therefore, exploring the determinants and side-conditions of family planning is crucial, in particular with respect to those determinants that have heterogeneous effect over the life course.

Numerous studies have investigated factors influencing childbearing intentions and childbearing outcomes. Aspects that may play a role in this decision can be divided into the following categories: socio-economic (e.g. education), familial (e.g. family background and support), societal (e.g. maternity leave programs), and biological (e.g. infertility). The research in social and economic sciences has focussed on socio-economic factors, such as education (Fagbamigbe & Idemudia, 2016; Luc et al., 1993; Rindfuss & St. John, 1983), financial security (Kind & Kleibrink, 2013), and partnership stability (Kuhnt & Trappe, 2013). Somewhat surprisingly, health has attracted relatively little attention in this literature so far. As – besides education – another crucial element of human capital, individual health can be viewed as both a resource and a barrier to bearing and raising children (Syse et al., 2020). While some empirical studies consider health-related information in the analysis, they do not focus on its possible effect on the transition to motherhood (Benzies et al., 2006; Gordo, 2009; Tough et al., 2007). For instance, Gordo (2009) investigates the factors of postponement of motherhood in Germany and concludes that career costs are the primary drivers of delayed motherhood, whereas health does not seem to play a significant role.

Health may, however, become an increasingly more important determinant of women's fertility than it was hitherto considered in the existing literature. In developed countries, a major share of women postpone first childbirth to a – relative to their reproductive period – later age or even to its end. At an advanced age, women may become more susceptible to adverse health shocks, which may result in giving birth to a child even later or not having one at all. Thus, with advancing age, sub-optimal health may become a crucial aspect of entering motherhood, which so far has been insufficiently covered in the literature.

The medical literature has produced multiple papers that focus on the effect of specific diseases on fertility patterns, e.g. HIV (Chen et al., 2001), cancer (Cvancarova et al., 2009; Langeveld et al., 2002; Schover et al., 1999; Zebrack et al., 2004), mental disorder (Bhongade et al., 2015; McGrath et al., 1999), multiple sclerosis (Bonavita et al., 2021), obesity (Cheng & Ng, 2007; Frisco & Weden, 2013; Ramlau-Hansen et al., 2007). However, there is little evidence on the effect of general health on fertility (Alderotti & Trappolini, 2022; Holton et al., 2011; Mynarska & Wróblewska, 2017; Syse et al., 2020), i.e. on the possible effects of health that go beyond specific deficits that biologically affect reproduction. The effect of health can operate through non-biological channels as well. For instance, individuals in poor health might not feel sufficiently fit for parenting or might be afraid of being unable to afford children due to health-related wage deductions or job loss (Alderotti & Trappolini, 2022; Syse et al., 2020). Holton et al. (2011), for instance, question the common perception that childbearing decisions are primarily built on financial considerations and argue that poor health is a major obstacle to achieving family planning goals among Australian women. Mynarska and Wróblewska (2017) focus on the effect of health on childbearing decisions among women in Poland. They conclude that having health problems has a negative impact on motherhood plans. Alderotti and Trappolini (2022) investigate the role of health in fertility intentions among migrants in Italy. They conclude that poor health has a negative effect on fertility intentions. Syse et al. (2020) is, among the aforementioned studies, most closely related to our analysis as not self-reported fertility intentions but realized fertility is linked to general health. More specifically, Syse et al. (2020) study the influence of health on fertility among Norwegian women. They approximate health status with sickness absence and uptake of long-term benefits using nationwide registry data on women aged 16-45 from 2004-2018. Focusing on the question of whether fertility patterns have changed over the considered period, they find

Figure 1: Fertility patterns in Germany



Source: *Statistisches Bundesamt; own illustration.*

that receiving long-term benefits is associated with lower fertility, with this pattern becoming less clear in recent years. A positive association between fertility and women's health is, however, not found if sickness absence is used as an alternative measure of health.

Besides – with the exception of Syse et al. (2020) – not analyzing actual fertility, another shortcoming of these studies is that they do not take into account the deterioration of health capital with age and that its influence on reproductive behavior may not be constant over the life course. The contribution of our paper is to analyze the role of health in entering motherhood over the course of reproductive age among women. Thus, our research question is whether health plays a role in reproductive decisions and how this role differs at various stages of life. We focus on the decision to enter motherhood that is on the extensive margin of fertility. To address this research question, we utilize longitudinal population surveys from Germany and employ discrete-time survival models. We find that health is a significant determinant of entering motherhood, whose role changes remarkably over the life course of women. For women older than 30 years, our analyses yield a pronounced health gradient, with better health increasing the probability of entering motherhood. At younger ages, this pattern is reversed. Thus, we contribute to the existing literature by establishing the heterogeneity of the effect of health over the course of reproductive age.

The paper is structured as follows: in Section 2, we outline the data used for the analysis. Later on, in Section 3, the empirical strategy is described. Afterward, we present the estimation results and provide the output of robustness checks in Section 4. In Section 5, we discuss the results and limitations, and finally conclude.

2 Data

We combine two representative longitudinal annual surveys of the German population, namely the German Socio-Economic Panel v38 [SOEP] (Goebel et al., 2019, 2023; Wagner et al., 2007) and the Panel Analysis of Intimate Relationships and Family Dynamics v13 [PAIRFAM] (Brüderl et al., 2022). Though the SOEP is a general household survey while the PAIRFAM is tailored to questions regarding family formation and relationships, both surveys share the information that is key to our analysis: information on the timing of childbirth and subjective health in particular. These two surveys differ in many ways. Yet, from the perspective of our analysis, the sampling methodology is the core difference between them. The SOEP is an annual panel that aims to track all members of the initially[1] randomly sampled households as long as possible, while for the PAIRFAM-specific birth cohorts (1991-1993, 1981-1983, and 1971-1973) were sampled in 2008 and followed over time (Huinink et al., 2011). The two data sources, hence, substantially differ in terms of age composition. Nonetheless, the similarities across the surveys, in particular in terms of compatibility of dependent and explanatory variables, allow us to merge them in order to enlarge the sample. As our sample has quite a number of restrictions that lead to a significant decrease in the number of observations, we opt for combining two datasets[2].

Since our research question addresses a very specific group of individuals, we apply several inclusion criteria to restrict the estimation sample: (i) only women are included; (ii) childless by the first observed period; (iii) aged between 17 and 49 (reproductive age). The women are followed from the age of 17 or the age they entered the survey until (a) the birth of the first child, (b) reaching the end of reproductive age, or (c) dropping out from the sample[3].

The analysis data has a person-year panel structure, i.e. one individual contributes to the sample more than once. Only observations that have information on self-assessed health and selected control variables are included in the description and estimation samples. The sample size from the SOEP equals to 50,743 observations from 12,571 individuals, from the PAIRFAM 13,790 observations from 3,233 individuals, amounting in total to 64,533 observations from 15,804 individuals. Out of these 2,269 become mothers for the first time during the observation period.[4] As pregnancy can affect the health state, we redefine the event-time of becoming a mother as conception leading to the first birth, instead of the actual date of giving birth, by subtracting nine months from the date of birth of the first child. Moreover, in order to avoid reverse causality caused by changes

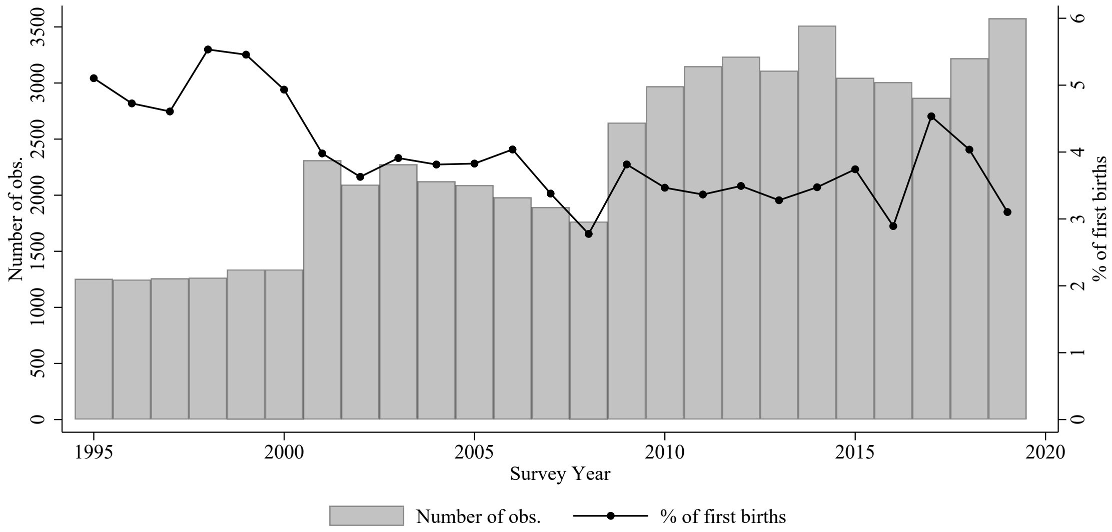
[1]The initial cross-section of households was sampled in 1984; panel attrition was addressed by drawing refreshment samples in later years. The PAIRFAM also includes refreshment samples.

[2]Similar strategy is employed in Schaubert (2015) and van der Vleuten et al. (2021).

[3]Missing waves are not considered a drop-out.

[4]Table A1.1 presents the distribution of length of the spell both in years (including gaps) and in observed waves.

Figure 2: Distribution of pregnancy year and number of observations by survey year



Notes: *The results are based on 58,582 observations. "% of first births" represents the share of women in the sample (in the respective year) who gave birth to their first child, i.e.* $N_t(\textit{gave birth to the first child})/N_t$. *Years 2021 and 2020 were excluded, because our definition of a birth (using information from future panel waves) renders them not comparable to earlier years.*
Source: *SOEP v38 (from 1994 - 2019) and PAIRFAM (2008-2019); own illustration.*

in health as a result of behavioral adjustments due to family planning, in the regression analyses we lag the covariates by one year relative to the event of interest.[5] We look at pregnancies that occurred place between 1995 and 2021 (Figure 2). As the PAIRFAM data is only available after 2008, we see an increase in observations in absolute terms in the last decade of the 21st century.

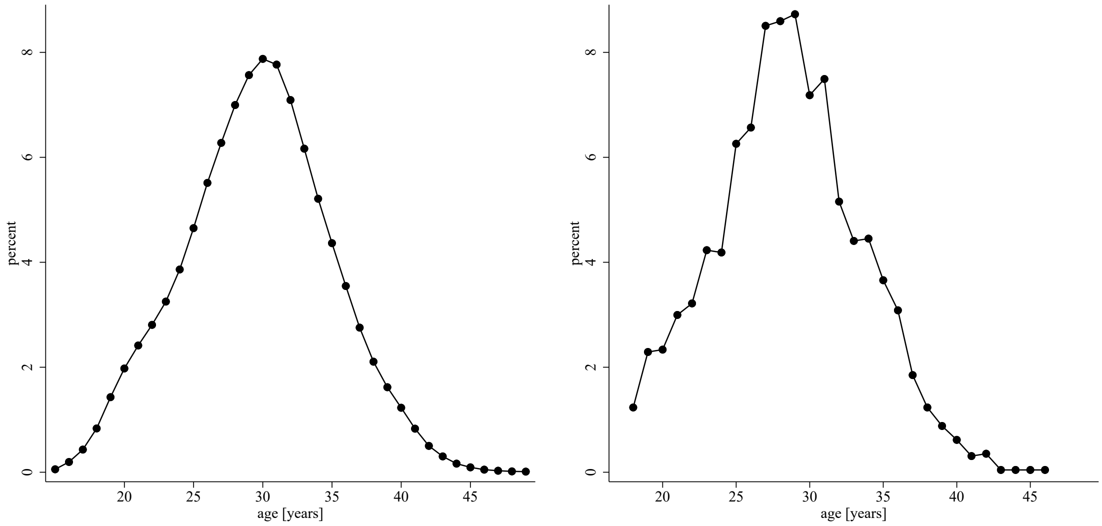
We compare estimated distribution of age at the first birth based on our sample with the closest population statistic that is available for Germany, namely the age distribution of mothers giving birth to the first child, considering the years 2009 to 2021 (Statistisches Bundesamt, 2023a) (Figure 3). The pattern found in the estimation sample is very similar to what is officially reported for the population, which provides some confidence in the data we use for further analyses. When comparing the two distributions, one has, however, to keep in mind that our conception-based definition of transition to motherhood shifts the location of the distribution in the estimation sample (Figure 3). [6] Moreover, the two distributions do not consider the same period of time. In any case, one recognizes the coherent pattern of the transition being most likely in the late 20s and early 30s. Teenage[7] birth is a rare event, yet also giving birth to the first child at the age of 40 or higher is very unlikely.

[5]Notwithstanding, we cannot rule out more indirect channels through which (past) health might be affected by fertility choices even among childless women, e.g. mental strain due to involuntarily remaining childless.

[6]The sample distribution peaks at the age of 29, while the peak value for the population statistics is 30. Yet, in our analysis, we look at the conception dates, which effectively means roughly a 1-year difference with the birth date.

[7]The SOEP includes only very limited information on underage individuals. Hence, our analysis cannot address childbearing at extremely young ages. This means that 'teenage motherhood', with respect to the survey data, means conceiving the first child at the age of 18 or 19.

Figure 3: Distribution of age at first birth (population vs. sample)

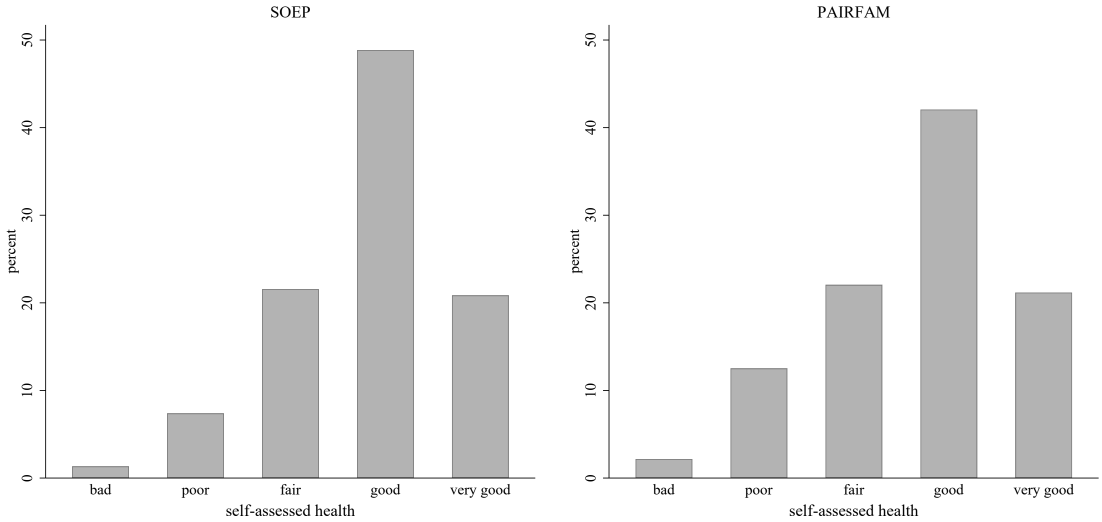


Notes: *The results of the right panel are based on 64,533 observations.*
Source: *Left panel - age at first birth in the population: Destatis (2009 - 2021); own illustration. Right panel - age at first birth in the sample: SOEP v38 (from 1994) and PAIRFAM v13, own illustration.*

The measure of health, which we want to relate to fertility, is self-assessed health [SAH], i.e. subjective health. This choice is partly data-driven since SAH – unlike more specific information on certain health deficits – is available in both the PAIRFAM and the SOEP data without gaps. Yet, we regard SAH as a well-suited health measure for the present analysis. From an economic perspective, the transition to motherhood is, first of all, a choice variable. This choice is likely to be influenced by health as it is perceived by the woman who makes this decision. One may argue that subjective health expectations are of greater importance for that decision than current health. Yet, expected health is not directly observable.[8] Nonetheless, just using contemporaneous SAH is consistent with a simple model of expectation formation, which assumes that health expectations are formed by projecting current self-perceived health into the future. The respective survey question in the SOEP is "How would you describe your current health?" with five answer options, namely bad, poor, fair, good, very good[9]. The question regarding self-reported health in the PAIRFAM is formulated with a reference to a specific time period, namely "How would you describe your health status during the past four weeks, generally speaking?", yet with the same answer choices. Nonetheless, the distribution of answers is highly comparable between the surveys (Figure 4), suggesting that the reference to the previous four weeks does not make the respondents think differently about their health. Overall, almost half of the individuals evaluated their health as good, about 20% each as fair and very good, and roughly 10% in total as poor and

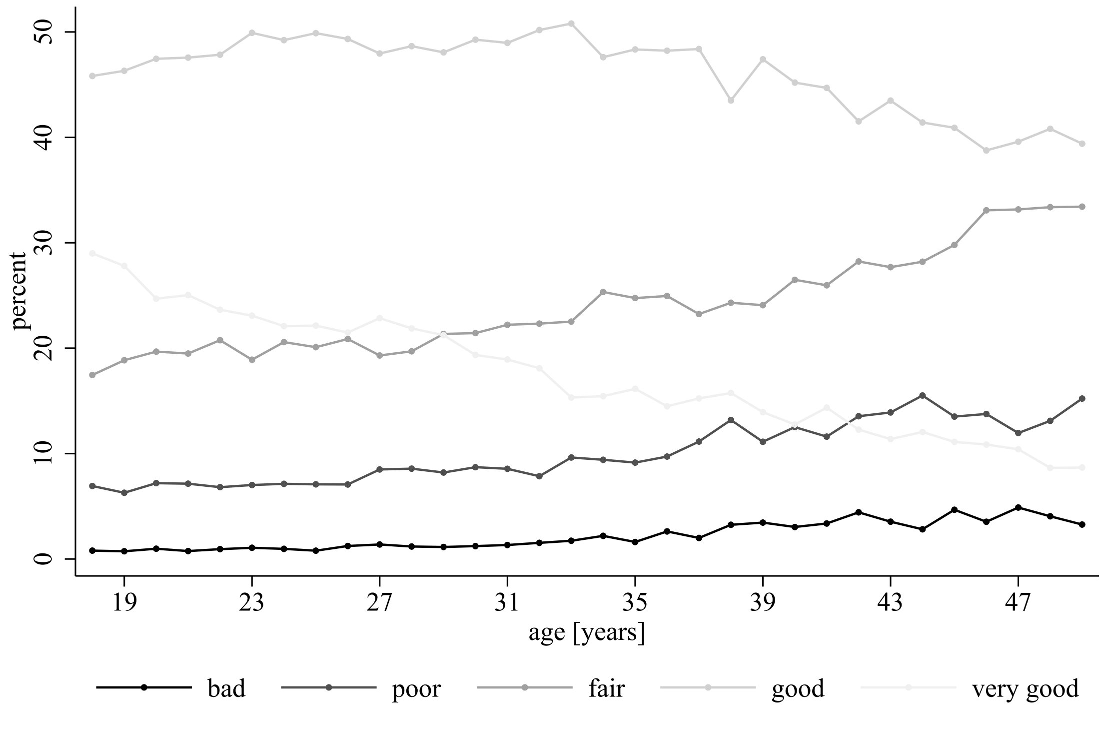
[8] The SOEP questionnaire only includes a question about being ‘worried about own health’.
[9] In the SOEP, this question has only been included since 1994. Thus, for our analysis, we use the survey data collected after 1994.

Figure 4: Distribution of self-assessed health



__Notes:__ The distributions of self-assessed health of respondents in the SOEP (left panel) and in the PAIRFAM (right panel) are presented separately. The results of the right panel are based on 50,743 observations, left panel – 13,790.
__Source:__ Lefr panel – SOEP v38 (from 1994), right panel – PAIRFAM v13; own illustration.

Figure 5: Distribution of self-assessed health by age



__Source:__ The distribution depicts the percentage of respondents in certain health state over different ages. The results are based on 64,533 observations.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

bad combined.

To describe the short-term and long-term variation in self-assessed health, we present the transition matrix of health status from period t to $t+1$ in Table 1 and the distribution of health status over different ages in Figure 5. There is some variation in health over the life course, mainly driven by a decrease in those in good health and an increase in those in fair health as age rises. The transition matrix provides evidence that most individuals whose health status changes switch to the nearest levels, while the extreme jumps, e.g. from very good to bad, are rarely observed. Nevertheless, Table 1 indicates that there is substantial within variation in health that can be exploited for identification. Except for women in good health, the share of observations for which no change in SAH is reported is smaller than 50%.

In the later stages of our analysis, we examine the interplay of fertility, health, and age conditionally on covariates. We, hence, consider several socio-economic controls that are available in the PAIRFAM and SOEP. In detail, these controls are: equalized household income [€ per month][10], highest achieved education at survey time [no/school drop-out, still in school, general elementary, middle vocational, vocational, higher education], employment [not employed, part-time, full-time], and migration background [no, indirect, direct]. In some specifications, we also condition on parental education[11], use alternative measures of education (years of education) and income (raw household income). The descriptive statistics for the control variables are presented in Table A1.2. Besides reporting these statistics for the full estimation sample, we also report them for women when entering motherhood. However, interpreting the descriptive statistics is not trivial, as due to the methodological design (described in Section 3), the individuals are dropped out once given the first birth. Thus, our analysis captures mostly a young population with an average age of 27.

3 Methods and Empirical Models

In this paper, we are interested in the role health plays in the transition to motherhood throughout the course of reproductive age. We take advantage of the longitudinal nature of the data and employ a survival analysis framework. That is, we aim to estimate the age-specific conditional probability of the transition to motherhood — in other words, the hazard – as a function of health. One major advantage of this modeling framework is that it straightforwardly accommodates imbalanced panels, right-censoring in particular (Jenkins, 1995). Right-censoring, i.e. losing track of

[10]Using a contemporaneous (yet lagged by one year) income measure is not ideal because it fails to capture income expectations, particularly for young individuals, that are crucial in family planning. Employing a moving average of income would possibly serve as a more suitable control in our case. However, due to the possibility of entering or exiting the survey at various ages, capturing income expectations within the same time frame for all individuals in the sample becomes impractical. Moreover, the possibility of dropping out of the sample due to entering motherhood relatively early in life makes this task even more challenging. Therefore, we opt to stick to the contemporaneous (yet, lagged by one year) equalized household income.

[11]More specifically, we condition on the woman's mother and father having obtained a high school certificate (Abitur).

Table 1: Transition matrix of self-assessed health

SAH_t	SAH_{t+1}					Total (%)
	bad	poor	fair	good	very good	
bad	30.75	31.42	21.12	13.50	3.21	100.00
poor	5.79	31.42	34.89	23.12	4.79	100.00
fair	1.34	13.14	42.42	37.10	6.01	100.00
good	0.47	4.68	17.82	61.55	15.47	100.00
very good	0.25	2.66	7.71	41.18	48.20	100.00

Notes: *The results are based on 64,533 observations. The values refer to the percentage of individuals that transitioned from a certain self-assessed health in period t to a certain self-assessed health in period $t+1$. For example, 30.75% of those that were in bad health remained in bad health in the next period, whereas 3.21% improved to very good.*

Source: *SOEP v38 (from 1994) and PAIRFAM v13; own calculation.*

an observational unit before the event of interest may have occurred, is a frequent issue in longitudinal survey data. Such data often suffers from sample attrition, e.g. because of individuals dropping out from the survey due to death, refusal to participate, or failure to follow up. Hazard models are, hence, more robust to selection issues than empirical models that directly use time to event – in our application, this means age at first birth – as the outcome variable.

Empirical survival analyses accommodate both continuous-time and discrete-time models. The fact that the SOEP and the PAIRFAM provide rather precise information regarding the date of birth suggests thinking of time as continuous and specifying a model in continuous time. However, the information on self-assessed health, as well as several control variables, is only collected on a yearly basis. Therefore, we think of time as a series of discrete periods and specify discrete-time hazard models (Singer & Willett, 1993).[12] This leads to, in technical terms, a rather simple empirical approach that coincides with estimating a (panel) binary outcome model. For women who remain lifelong childless, and also for women subject to right-censoring, the outcome is just coded as a sequence of zeros, while for women who are observed to become mothers, only the outcome for the period of the transition to motherhood is coded as one (Tutz & Schmid, 2016, pp. 52–53). Periods after that transition are not considered in the estimation sample since, in our context, the population at risk is childless women of reproductive age, and survival analyses only look at units at risk of experiencing the transition of interest. In other words, we regard motherhood (in technical terms, rather a pregnancy that results in the birth of the first child) as an absorbing state. Adopting the terminology used in the literature on survival analysis, surviving refers to staying at risk, i.e. remaining childless, while failing refers to becoming a mother. To be more specific about the specification of our discrete-time survival analysis, we use a linear prob-

[12]In the present application, thinking of time as discrete or as continuous makes probably little difference. The key explanatory variable SAH is observed on a yearly basis and exhibits substantial variation over time; see Table 1. Therefore, if following a continuous-time approach, one would probably heavily rely on episode splitting, which effectively leads to interpreting time as a sequence of one-year episodes.

ability model that specifies the probability of an event occurring conditional on surviving in the previous periods as a linear function of the explanatory variables.[13]

In the following, we use λ_{it} to denote the hazard rate, that is, the probability of a woman i becoming a mother at the age t, conditionally on not having made that transition at a younger age already. τ_i, hence, denotes the age of conception of the first child. Specifying the hazard as a linear function gives:

$$\lambda_{it} = \mathrm{P}(\tau_i = t | a_i, \boldsymbol{x_{it}}, \tau_i \geq t) = a_i + \boldsymbol{x_{it}\beta} \tag{1}$$

(cf. Farbmacher & Tauchmann, 2023). $\boldsymbol{x_{it}}$ denotes the vector of explanatory variables, including the key regressor *SAH*. $\boldsymbol{x_{it}}$ also includes some function of t. This renders the hazard duration dependent – i.e. in the present application, age-dependent – although there is no explicit reference to duration dependence in the notation used in (1). Since the average age at the transition to motherhood has increased throughout the considered period of time, we also add year fixed effect to the model to disentangle the effect of age from a possibly nonlinear, secular time trend. a_i denotes unobserved, individual (i.e. woman) level heterogeneity regarding the conditional childbearing probability. As an extreme case, a large negative a_i, which effectively places λ_{it} to zero, may just capture infertility.[14] Yet, a_i does not only capture somatic factors but also (time-invariant) preferences and socio-economic conditions that affect the probability of motherhood. From a methodological perspective, it is crucial whether or not a_i can be assumed to be unrelated to the variables in $\boldsymbol{x_{it}}$, *SAH* in particular. If uncorrelatedness holds, the effects of interest $\boldsymbol{\beta}$ can – thanks to the linear hazard specification – straightforwardly be estimated by ordinary least squares (OLS).[15] a_i and *SAH* might, however, be correlated for various reasons. Partnership quality, for instance, might be determined by unobserved and persistent factors, such as individual character traits. The quality of partnership is, however, likely to directly affect the decision about having a child (Berninger et al., 2011) and, at the same time, may also affect health outcomes (Robles et al., 2014). In one of our empirical models, we, hence, take possible correlations of a_i and $\boldsymbol{x_{it}}$ into account.

The focus of the empirical analysis is on the role health plays in the transition to motherhood and, in particular, on the interplay of health and age in that transition. This does not mean that we

[13]Though the linear probability is subject to critique (see e.g. Horrace & Oaxaca, 2006), using it is a common practice in applied econometrics, including applications that specify discrete-time hazard models (e.g. Corno et al., 2020; Currie & Neidell, 2005; Fernandes & Paunov, 2015). As a robustness check, we also report results from estimating a complementary log-log model, which has the appealing property of sharing its likelihood function with a discrete-time – Cox model like – proportional hazard model (Cameron & Trivedi, 2005, p. 600-603). In this alternative specification the hazard is specified as $\lambda_{it} = 1 - \exp\left(-\exp(a_i + \boldsymbol{x_{it}\beta})\right)$. The results the complementary log-log model yields hardly differ from its counterparts from using the linear probability model; see Appendix Figure A1.3.

[14]While the PAIRFAM questionnaire asks about infertility, there is no information on this issue in the SOEP.

[15]Consistent estimation, indeed, implies further (implicit) assumptions such as, censoring being at random and the linear specification approximating the true hazard sufficiently well.

claim an estimation of the causal effect of health on becoming a mother. We cannot claim causality because the empirical analysis is not based on an exogenous source of health variation. Since we are interested in health as a subjective concept, it is not even obvious what such exogenous variation would be.[16] Our analysis is, hence, descriptive in nature. Yet, we gradually increase the level of descriptive sophistication with the intention of cutting off more and more channels through which confounders may shape the association of SAH, age, and motherhood.

More specifically, we specify several models of different levels of flexibility to let the data speak about that interplay and to assess how robust the results are to the choice of model specification. In addition, we aim to figure out if some structure can be imposed on the empirical model that renders the analysis less data-hungry. We begin with (i) only allowing for self-assessed health to uniformly shift the level of hazard function. The next step is (ii) to introduce a saturated set of interactions of self-assessed health indicators and age dummies to enable the effect of health to be heterogeneous across the lifespan. Smoothing of the function can help to impose more structure on the estimates and, in consequence, ease economic interpretation since the fully saturated SAH-age specification appears rather rich, given the size of the estimation sample. To this end, we make use of a non-parametric estimator, (iii) namely kernel regression (Cattaneo & Jansson, 2018) that allows for data-driven modeling of the relationship between health and hazard function over the course of reproductive age. One drawback of such an approach is computationally expensive estimation, particularly with respect to estimating standard errors, which requires bootstrapping an already time-consuming estimation procedure. Another drawback is that including control variables and addressing the issue of unobserved individual heterogeneity is less straightforward in this estimation framework as compared to a conventional parametric linear regression. However, after the empirical visual comparison of the non-parametrically estimated hazard curves, we infer that it can be reasonably well approximated by (iv) a linear probability model with a third-order polynomial of age. To this simple empirical model, besides year indicators, (v) we add controls for several individual characteristics such as labor market participation, education, equalized household income, and migration background; see section Section 2 for a more detailed discussion of the control variables. We do not include marital or relationship status as it could cause a reverse causality problem even when letting the covariates enter the model as lagged values.

Finally, we take into account that unobserved heterogeneity a_i, may well be correlated with

[16] SAH is probably, affected by numerous exogenous factors such as severe diagnoses (Heller et al., 2009), accidents (Toft et al., 2010), daylight (Jin & Ziebarth, 2020), age, etc. Yet, establishing that these factors affect fertility only through the channel of self-perceived health appears virtually impossible.

self-assessed health and possibly the controls and hence may act as a source of bias.[17] The conventional linear fixed effects estimator (vi), which is seemingly an obvious choice in the considered setting, is however shown to be (heavily) biased when applied to outcome variables that indicate transitions into an absorbing state (Farbmacher & Tauchmann, 2023).[18] Therefore, we follow the approach suggested by Farbmacher and Tauchmann (2023) and (vii) employ – instead of fixed effects estimation – two-stage least-squares using first differences as instruments for the levels of the explanatory variables (see Cepec et al., 2022; Zhang & Axinn, 2021, for empirical applications of that method, based on an earlier working paper version). We report results from conventional fixed-effects estimation, together with several robustness checks (Section 4.2), only to illustrate that this classical method yields counter-intuitive results in the present application. For all other empirical approaches (i) to (v) we present results in Section 4.1. Yet, we consider model (v), that is, the conventional linear probability model with a third-order polynomial of age and controls as the preferred specification. We report (at the woman level) clustered standard errors for all the specifications. We provide an overview of all the models discussed above in Table 2.

Table 2: Overview of model specifications

	specification	results visualization
(i)	OLS with SAH dummies	Figure 6, right panel
(ii)	OLS with fully saturated set of interactions of SAH dummies and age dummies	Figure 7, panel: Fully saturated
(iii)	kernel regression with SAH dummies	Figure 7,panel: Non-parametric
(iv)	OLS with third-order polynomial of age interacted with SAH dummies	Figure 7, panel: Cubic
(v)	OLS with third-order polynomial of age interacted with SAH dummies, individual-level control variables included	Figure 8, panel: Cubic
(vi)	model (v) with individual-level fixed effects	Figure 8, panel: FE
(vii)	model (v) estimated by two-stage least-squares using first differences as instruments for the levels of the explanatory variables	Figure 8, panel: IV

Notes: *The Table presents the overview of specifications descried in Section 3 and corresponding estimation results.*

4 Results

In this section, we report the results from different estimation approaches discussed in Section 3. In doing this, we rely mainly on graphical representations of our findings in order to enhance

[17] It needs to be mentioned that a_i, even if uncorrelated with all explanatory variables in the population, may generate a bias in a non-repeated event setting. This bias is referred to as survivor bias and originates from selective survival, which renders a_i and x_{it} correlated in the estimation sample despite uncorrelatedness in the population; see Nicoletti and Rondinelli (2010) and Farbmacher and Tauchmann (2023) for more detailed discussions.

[18] This bias, which does not vanish asymptotically, originates from the within transformation rendering the transformed regressors a function of the outcome, i.e. the realized time at risk.

interpretability.[19] We begin with describing the estimation results based on different methodological approaches, in particular parametric and non-parametric. Afterward, we discuss the results of the preferred specification, which takes into account observed heterogeneity. Furthermore, we outline the results of estimation strategies aimed at resolving potential unobserved heterogeneity. Finally, we perform several robustness checks with respect to the definition of subjective health measure, sample selection criteria, and control variables.

4.1 Main Results

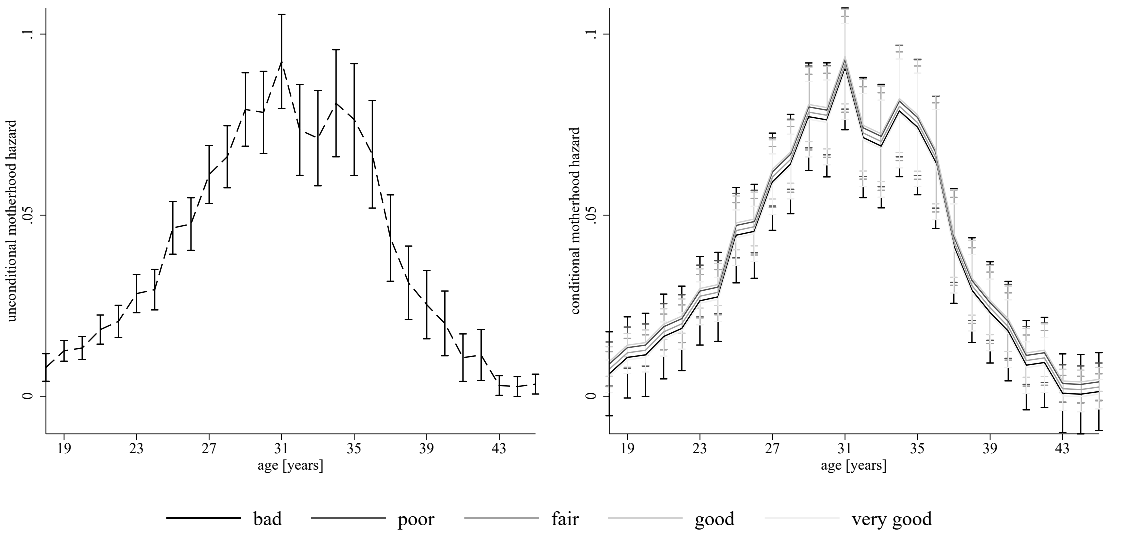
A descriptive estimate of the motherhood hazard is provided in Figure 6. The left panel depicts the estimated hazard function unconditionally on health, i.e. it originates from estimating equation (1) with – besides year dummies – age indicators as the only explanatory variables. The risk of transitioning to motherhood peaks at the age of 30, reaching almost 10%, and declines rapidly after age of 35. The right panel of Figure 6 reports results that are conditional on health in the sense that a set of SAH indicators is included in the regression model, allowing the different health states to shift the hazard curve (model (i) from Section 3). We observe that the individuals in better (with the exception of very good) health are at slightly greater risk of transitioning to motherhood than the ones with bad health (Figure 6, right panel). However, the test for joint statistical significance of the set of self-reported health dummies does not reject the null of coefficients being equal to zero (p-value = 0.245). These results, hence, seem to confirm earlier findings (Gordo, 2009) that health is not a significant determinant of fertility.

In the next step, we explore the role of health, allowing it to be heterogeneous over the course of reproductive age. By this, we take into account that the role of health in the timing of first childbirth might be more complex than just uniformly increasing or decreasing the hazard over the entire reproductive life. More specifically, we estimate a fully saturated model, i.e. include the set of age dummies interacted with self-assessed health dummies (Figure 7, Panel: Fully saturated[20]; model (ii) from Section 3). In this much richer model, the test for joint significance clearly rejects the null of the health coefficients jointly being zero (p-value = 0.000). This indicates that health plays some role in the decision about the transition to motherhood once allowing for a heterogeneous role over the age of women. However, interpreting the health-status specific hazard curves is not straightforward since the saturated model yields a large number of (individually) noisily estimated coefficients.

[19] To enhance readability, we present figures without confidence intervals in the main body of the text, and the ones with confidence intervals in the Appendix, see Figure A1.1 and Figure A1.2. Table with numeric results is presented in Table A1.3 and Table A1.5.

[20] See Figure A1.1 for confidence intervals.

Figure 6: Hazard function: without and with SAH as control



Notes: *95% confidence intervals are represented by vertical lines. The results are based on 64,533 observations. Year fixed effects are included. Right panel includes SAH dummies (model (i)).*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

To enhance the interpretability, we adopt a non-parametric approach to estimate the effect of health on motherhood hazard (model (iii) from Section 3). One of its major advantages is smoothing the hazard curves in a data-driven way without requiring assumptions regarding the functional form. More specifically, we employ non-parametric kernel regression. Unlike the saturated interaction model, the non-parametric approach (Figure 7, Panel: Non-parametric) yields a very salient pattern of health and age interaction regarding the transition to motherhood hazard. For younger ages, we see a gradient in such a way that the transition probability monotonically decreases in the health status. Yet, the pattern is fully reversed for women who have passed the age of about 28. For this age group, we see a gradient such that a transition to motherhood gets increasingly less likely for any less favorable health status. Only for women at the very end of their reproductive age, for whom the likelihood of becoming a mother is already minimal, Figure 7 (Panel: Non-parametric) does not indicate that health matters. That pattern does not conflict with what the fully saturated model yields. Yet, the lack of structure in the fully saturated model renders identifying such patterns much harder and by far less reliable.

The non-parametric approach provides a reference for choosing an appropriate specification in subsequent parametric regressions. Turning back to a more parametric estimation allows us to avoid the computational expense and large sample size requirements associated with computing standard errors using kernel regression and bootstrapping techniques. Furthermore, considering a substantial number of socio-economic controls, which may allow for judging whether the pattern found is just an artifact of confounders, turned out to be problematic with the non-parametric approach.[21]

According to Figure 7 (Panel: Non-parametric), we infer that the best suitable functional form for parametric modeling is a third-order polynomial of age interacted with self-assessed health which is presented in Figure 7 (Panel: Cubic; model (iv) from Section 3). This parametric yet flexible regression model qualitatively yields almost the same result as its non-parametric counterpart.[22] That is, the association of health and the transition probability is reversed over the course of reproductive life. This complex interplay of age and health is warranted by a statistical test of the joint significance of the interaction terms, which clearly rejects the null (p-value = 0.000).

The results discussed so far are purely descriptive. Unconsidered confounders may shape the association of childbearing, age, and health. For instance, the increased risk of teenage motherhood for women with poor health might just capture a low socio-economic status that affects either variable. We, hence, move towards an analysis that eliminates these sources of confoundedness by controlling for socio-economic characteristics.

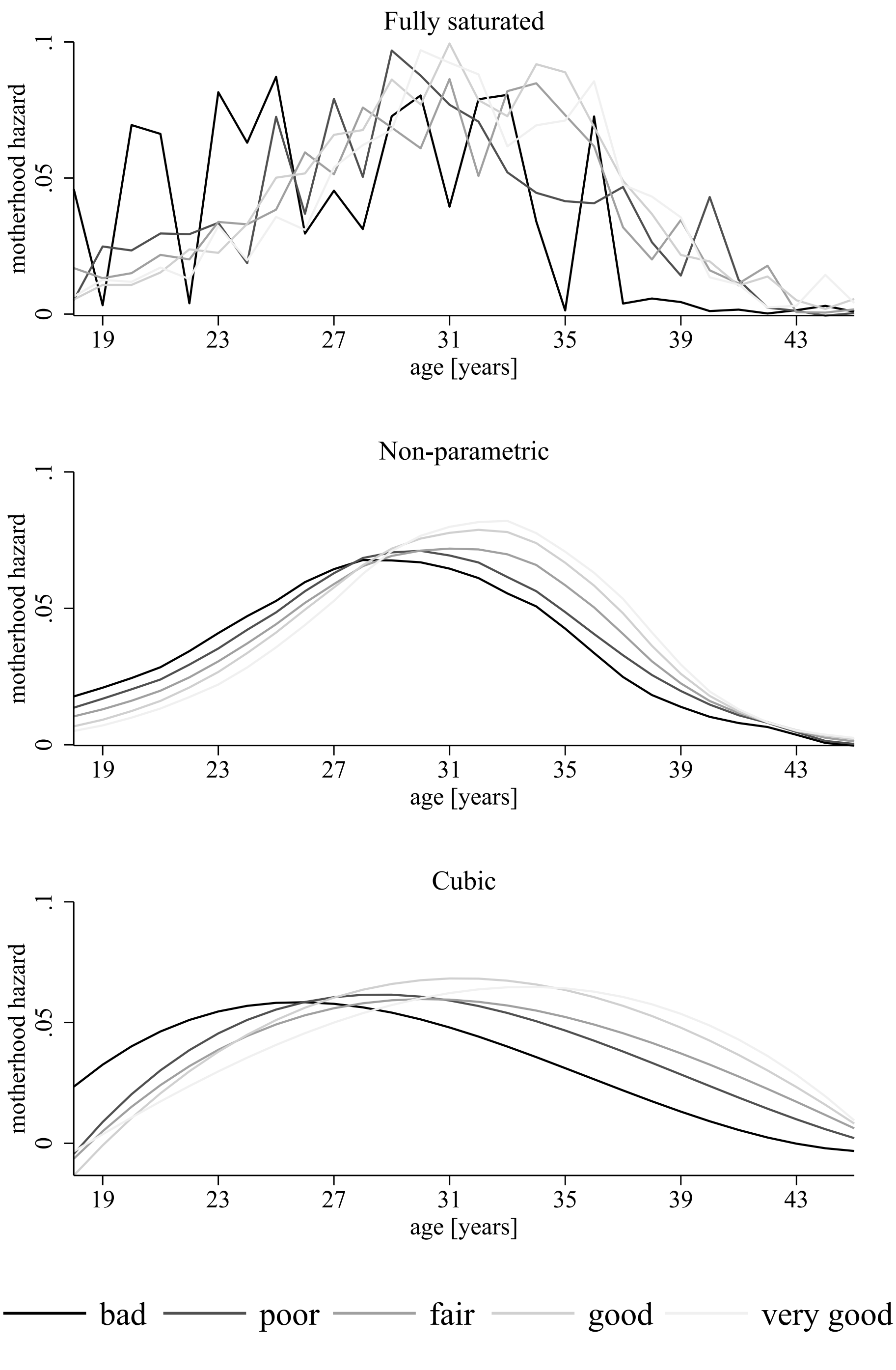
Figure 8[23] (Panel: Cubic; model (v) from Section 3) displays results for the specification that adds several controls to the third-order polynomial of age interacted with the five categories of SAH. We regard this specification as our preferred one. The socio-economic controls are highest achieved education, equalized household income, employment status, and migration background. Moreover, we include an indicator for observations from the SOEP to control for possible systematic differences in the two data sources. Year indicators are also included. Controlling for individual socio-economic characteristics has, however, little impact on the patterns of estimated health-specific hazard functions (compare Figure 8 to Figure 7, Panel: Cubic). The test for joint significance of the SAH age-polynomial interactions also yields the same results (p-value = 0.000). Our earlier key finding is, hence, robust to controlling for socio-economic characteristics. That is, the hazard curves are shifted to the left for worse health. This means that for women older than 30 years, the conditional probability of bearing the first child is much lower for women in bad or poor health in comparison to the ones in good or very good health. While the confidence intervals overlap for very young ages, we observe a distinct pattern for women in their early 20s (Figure A1.2). Figure A1.2 also suggests an even more distinct pattern at older ages. The former finding is particularly interesting as it suggests that the reverse health gradient at younger ages

[21]The more variables enter the model, the greater the number of observations needed to have sufficient 'cell size' in non-parametric modeling. We do not have enough data to increase dimensions by adding control variables.

[22]Several drawbacks of estimating a model using a non-parametric approach make us revert to a parametric one. More specifically, these disadvantages include the computational expense of obtaining standard errors and the inability to incorporate controls in the model to address the issue of confounding.

[23]See Figure A1.2 for confidence intervals.

Figure 7: Hazard function: estimated using different empirical approaches



Notes: *Control variables are not included, years dummies are included. The results are based on 64,533 observations. Panel: Fully saturated (model (ii)) – the set of all interactions of age and SAH is included. Panel: Non-parametric (model (iii)) – the model is estimated using kernel regression. Panel: Cubic (model (iv)) – the set of interactions of SAH with third-order polynomial of age is included. Negative values of the estimated hazard curves, which occur in the case of the parametric model, are primarily a scaling issue and an artifact of using a linear specification.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

cannot be (fully) attributed to socio-economic factors.

Quantifying the difference in conditional probabilities of becoming a mother results in rather large values in comparison to the share of women giving birth at a specific age, which is always less than 10% (see Figure 3). For example, at the age of 23, women in very good health are 2.78 percentage points less likely to transition to motherhood than women in bad health. At the age of 37, the same comparison yields that the former transition probability is higher by 2.78 percentage points.

Neglecting heterogeneity of the effect with respect to age would lead to the misleading conclusion that health is irrelevant in child-bearing decisions. While this erroneous result would be in line with Gordo (2009), it conflicts with the results of Holton et al. (2011), Mynarska and Wróblewska (2017), and Syse et al. (2020). However, allowing for age and health to interact in a complex way, we find that the role of health varies over the course of reproductive age and is of particular importance in older ages.

4.2 Robustness Checks

Our analysis does not tap a specific source of exogenous health variation. To nevertheless gain more confidence in the pattern we have found in the data – in the sense that it is not just an artifact of unconsidered confounders – in this subsection, we put additional effort into uncovering and possibly closing channels through which confounding factors might drive our earlier results.

4.2.1 Unobserved Heterogeneity

Some important determinants of fertility that are likely to be also related to general health are not observed in the data we use. As an extreme example, the SOEP lacks information regarding the physical ability to bear a child (infertility). Taking advantage of the panel structure of the data, we can adapt estimation approaches that can deal with such issues. As discussed in Section 3, we employ two competing methods for eliminating unobserved heterogeneity, namely inclusion of individual fixed effects (Figure 8, Panel: FE; model (vi)) and instrumenting levels of all explanatory variables with their first differences (Figure 8, Panel: IV; model (vii)). Naturally, neither approach allows for identifying coefficients of time-invariant regressors. Migration background is, hence, excluded from the set of control variables.

Conventional fixed effects estimation yields results that appear to make very little sense as the estimated motherhood hazard only increases (almost) monotonically in age, which is not at all plausible (Figure 8, Panel: FE). This corresponds to the argument of Farbmacher and Tauchmann

(2023) that conventional fixed effects estimation is heavily biased in a single-spell hazard model setting. On the contrary, the IV estimator yields results (Figure 8, Panel: IV) that are in terms of the point estimates comparable to their counterpart from the preferred specification and the results described previously. Yet, rather wide confidence bands indicate that instrumental variables estimation comes at the cost of reduced precision.[24] Nevertheless, the results of IV estimation do confirm the qualitative result of heterogeneous health effects with respect to age. The test of joint significance of the age-health interactions rejects the null at 5% significance level (p-value = 0.005). Finding that the qualitative pattern of results persists in the instrumental variables (IV) estimation is reassuring, as it suggests that this pattern is not merely an artifact of unobserved heterogeneity. It is also interesting that using an estimation method that primarily exploits within-woman health variation yields very similar results as the specification of reference does, which uses both within and between variation.

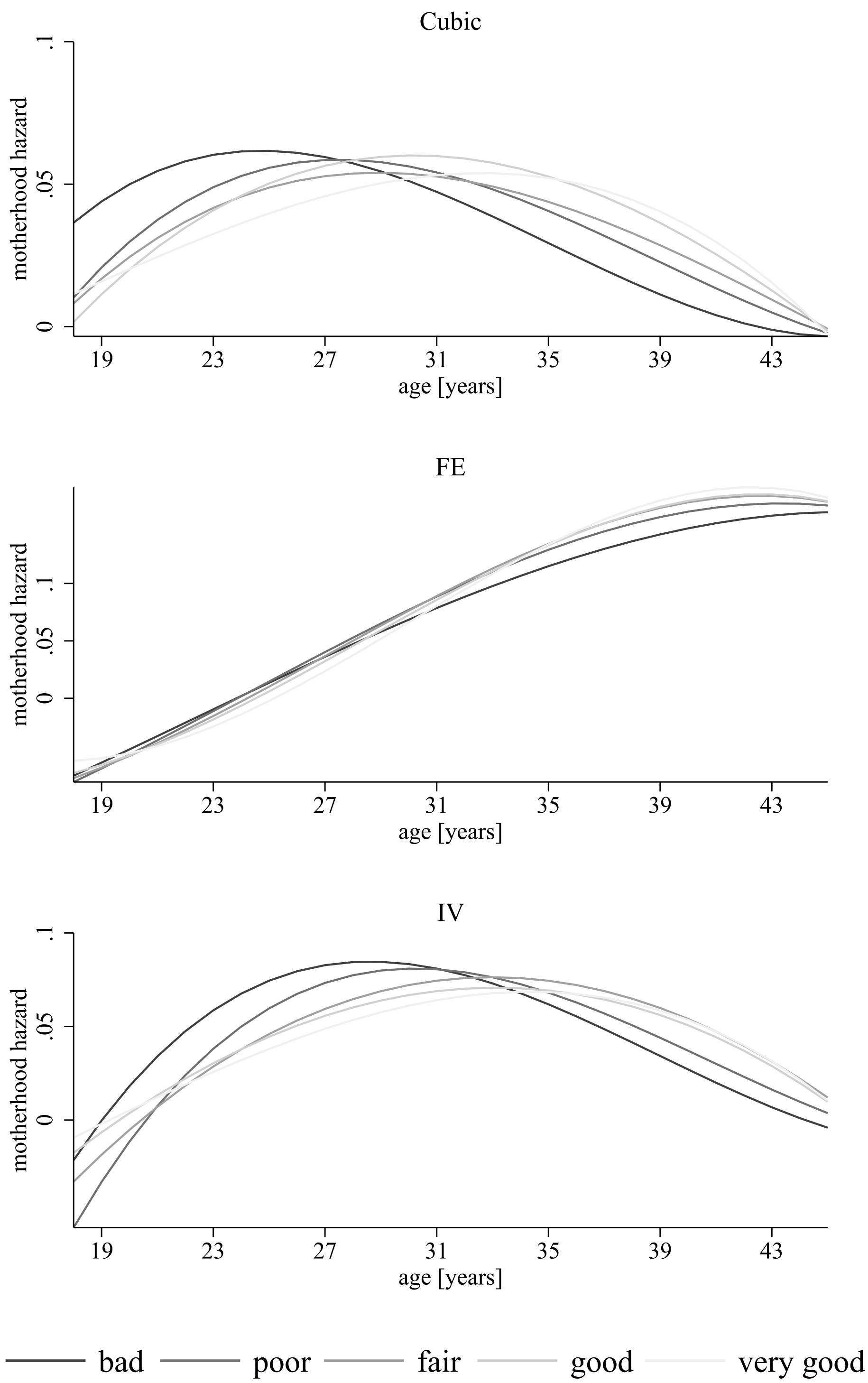
4.2.2 Functional Form

We choose our preferred parametric specification mainly relying on visual comparison of the results of parametric and non-parametric approaches, yet as a robustness check, we provide a formal comparison of the models with different degrees of complexity, namely linear, quadratic, cubic, fourth- and fifth-order polynomials (Table A1.4).[25] Conventional model fit metrics (AIC, BIC, R-squared adjusted) improve with the number of polynomial order. However, the interaction terms of self-assessed health and fifth-order polynomials are jointly not statistically different from zero; therefore, we exclude this model from consideration. Even though a model with a fourth-order polynomial is a better fit to the data in terms of AIC, BIC, and R-squared adjusted, it results in implausible curvature at the ends of the age range suggesting some over-parameterization by fourth-order polynomials (Figure A1.4). Nevertheless, the key pattern of a reversing role of health over the course of life is preserved. Yet, the hazard functions estimated by a model with cubic polynomial present the closest pattern to the results of non-parametric estimation and outperform linear and quadratic models, which reinforces us in regarding the cubic specification as the most suitable parametrization.

[24] According to the result of a Kleibergen-Paap under-identification test (Kleibergen & Paap, 2006), the null of general under-identification is nevertheless clearly rejected. This indicates that the IV estimation does not suffer from a weak-instruments issue.

[25] Models do not include control variables.

Figure 8: Hazard function: observed and unobserved heterogeneity



Notes: *Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included in all three panels. Panel Cubic (model (v)) is estimated using OLS; Panel FE (model (vi)) – using fixed effects; Panel IV (model (vii)) – using first differences as instruments for levels. The results are based on 64,533 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

4.2.3 Alternative Controls

We check the robustness of our results with respect to the measure of socio-economic status in three ways: (i) replace equalized income by raw net household income (Figure A1.5); (ii) drop indicators for highest achieved education and include years of education instead (Figure A1.6); (iii) add to the main specification indicators for mother and father having completed high school (Figure A1.7). The estimated hazard curves remain largely unchanged in either specification. This evidence provides us with even more confidence that the observed pattern is not an artifact of confounding factors. This, in particular, applies to the rather unexpected health gradient found for relatively young women. One could have argued that at a young age, one's own education and household income do not adequately capture socio-economic status. Consequently, the latter might generate a spurious negative correlation between health and fertility for young women. However, with parental education also being controlled, this argument appears rather weak.

4.2.4 Alternative Samples

The sampling design of the SOEP – unlike the PAIRFAM – is prone to fertility-related self-selection of women in the sample. If a woman joins an existing SOEP household or establishes a new one together with a SOEP respondent, she becomes a SOEP member as well. However, moving in with a partner is an important step of couple formation that can potentially lead to parenting. To address this possible issue, we perform another robustness check, in which we exclude these potentially self-selected individuals. We define a self-selected individual as a non-core SOEP member who is related to the head of the household as a partner in the first interviewed wave. Excluding 6,143 observations that satisfy our definition, we re-estimate the preferred model specification. Our results prove to be robust to the exclusion of potentially self-selected individuals. The pattern of the estimated hazard curves does not change (Figure A1.11), and health-specific heterogeneity of these curves is still warranted by statistical tests (p-value = 0.000).

In another robustness check concerning the sample selection criteria, we narrow down our sample to exclusively include women who do not have a migration background. This is done to eliminate the potential impact of variations in family-building cultures between migrants and native Germans. Our results are also robust to the exclusion of individuals with direct and indirect migration backgrounds (see Figure A1.12).

4.2.5 Redefining Self-assessed Health

We test the robustness of our results with respect to the definition of health measure in two ways. First, we aim at reducing the noise in the variation of SAH by grouping categories bad and poor into one as well as good and very good. The results of OLS estimation are presented in Figure A1.8. Secondly, we fix self-assessed health for each woman at the initially observed level. That is, we keep SAH constant tailored to the value that was reported in the first observed wave and, hence, only use between-women variation in health to estimate the hazard curves (Figure A1.9). The pattern remains consistent with our main results. That is, the pattern is robust to the source of variation used for the analysis, within, between or both. It suggests that women not only make their decisions based on the health expectation but also adjust their expectations with changes in health.

Lastly, we redefine health in terms of deviations from the best-observed health in the past. We aim to account for the dynamics of health development. Even though, in the long-term perspective, health inevitably deteriorates throughout life, short-term deviations from the past might still affect the decision to enter motherhood. We define a new variable equal to one if the health improved compared to the best-observed health in the past, minus one if health worsened, and zero if it did not change.[26] The size of the new estimation sample equals to 48,729 out of which 22,746 observations have a negative deviation, 4,693 have a positive deviation, and 21,290 exhibit no deviation, i.e. current health is as good as the the best past health status. Figure A1.10 indicates that women with positive health deviations are more likely to give birth in their early twenties. It is important to notice that women with the best possible health in the beginning, could not experience a positive health deviation, i.e. those women who could had, at best, good health at the beginning of the reproductive period. Between the ages of 23 and 37, women who undergo an improvement in health relative to the best health they experienced in the past are less likely to have a first child than women who experience a negative or no change in health. Intuitively, one could interpret this such that experiencing particularly good health later in life may lead to the formation of positive health expectations for the future and, as a result, to the postponement of motherhood. Although the results of this robustness check show a somewhat different pattern than all other specifications discussed so far, we do not consider this to be a lack of robustness. Using a measure of health that is defined relative to the past best health gives the health variables and the health-specific deviation in hazard completely different meaning.

[26]For example, if an individual reports $SAH_{2006} = good$, $SAH_{2007} = good$, $SAH_{2008} = very\ good$, $SAH_{2009} = good$, $SAH_{2010} = very\ good$; then the observation in 2006 is lost, and $deviation_{2007} = 0$, $deviation_{2008} = 1$, $deviation_{2009} = -1$, $deviation_{2010} = 0$.

5 Discussion and Conclusion

We empirically study the link between general health and the transition to motherhood among women in Germany. We take advantage of the longitudinal nature of population surveys and apply to them discrete-time survival analysis. Unlike the previous research that investigates the effect of health, assuming it to be homogeneous across women of different ages, we allow the health gradient in the transition probability to be heterogeneous with respect to age. Our analyses yield a remarkably robust pattern, in which better health is negatively associated with the likelihood of becoming a mother at a younger age, while this association is reversed later in reproductive life. Since the empirical analysis does not exploit a clean source of exogenous health variation, one cannot interpret the results in terms of precisely identified age-specific health effects. However, as the pattern of results is robust to controlling for several channels through which possible confounders may operate, the results appear to capture more than just the effects of third factors related to both health and fertility. One can, therefore, speculate about channels through which general health may directly affect fertility decisions and produce a pattern like the one we found. One such channel could be that women in poorer health have a preference for having their children early in life and do not want to wait until an age when health deficits are likely to become more severe, whereas women in better health are less cautious about delaying motherhood until a later age. In light of this way of interpreting the empirical results, the finding of qualitatively similar patterns, regardless of whether between- or within-women variation in health is used for estimation, suggests that such fertility intentions are adjusted to changes in health. For example, women who want to have children later in life may decide to have them earlier if they become concerned about their health relatively early. Another example would be that women who want to have children relatively late might not have them at all if their health deteriorates later in life.

Fertility decisions, though genuinely private matters are important for aging societies. A better understanding of what determines them is, hence, valuable for policymakers. Taking into account the deterioration of health with age opens up a new angle of the role of health. Given ever-increasing age at first birth, health as a fertility determinant should be given more attention in future research.

References

Alderotti, G., & Trappolini, E. (2022). Health status and fertility intentions among migrants. *International Migration*, *60*(4), 164–177. https://doi.org/10.1111/imig.12921

Benzies, K., Tough, S., Tofflemire, K., Frick, C., Faber, A., & Newburn-Cook, C. (2006). Factors Influencing Women's Decisions About Timing of Motherhood. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, *35*(5), 625–633. https://doi.org/10.1111/j.1552-6909.2006.00079.x

Berninger, I., Weiß, B., & Wagner, M. (2011). On the links between employment, partnership quality, and the intention to have a first child: The case of West Germany. *Demographic Research*, *24*, 579–610. Retrieved July 12, 2023, from https://www.jstor.org/stable/26349646

Bhongade, M. B., Prasad, S., Jiloha, R. C., Ray, P. C., Mohapatra, S., & Koner, B. C. (2015). Effect of psychological stress on fertility hormones and seminal quality in male partners of infertile couples. *Andrologia*, *47*(3), 336–342. https://doi.org/10.1111/and.12268

Bonavita, S., Lavorgna, L., Worton, H., Russell, S., & Jack, D. (2021). Family Planning Decision Making in People With Multiple Sclerosis. *Frontiers in Neurology*, *12*, 620772. https://doi.org/10.3389/fneur.2021.620772

Brüderl, J., Drobnič, S., Hank, K., Neyer, Franz. J., Walper, S., Alt, P., Borschel, E., Bozoyan, C., Garrett, M., Geissler, S., Gonzalez Avilés, T., Gröpler, N., Hajek, K., Herzig, M., Lenke, R., Lorenz, R., Lutz, K., Peter, T., Preetz, R., … Wetzel, M. (2022). *The German Family Panel (pairfam)Beziehungs- und Familienpanel (pairfam)* (Version 13.0.0). GESIS. https://doi.org/10.4232/PAIRFAM.5678.13.0.0

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications* (13th printing). Cambridge University Press.

Cattaneo, M. D., & Jansson, M. (2018). Kernel-Based Semiparametric Estimators: Small Bandwidth Asymptotics and Bootstrap Consistency. *Econometrica*, *86*(3), 955–995. Retrieved July 11, 2023, from https://www.jstor.org/stable/44955227

Cepec, J., Grajzl, P., & Mörec, B. (2022). Public cash and modes of firm exit. *Journal of Evolutionary Economics*, *32*(1), 247–298. https://doi.org/10.1007/s00191-021-00754-3

Chen, J. L., Phillips, K. A., Kanouse, D. E., Collins, R. L., & Miu, A. (2001). Fertility Desires and Intentions of HIV-Positive Men and Women. *Family Planning Perspectives*, *33*(4), 144. https://doi.org/10.2307/2673717

Cheng, J. Y. W., & Ng, E. M. L. (2007). Body mass index, physical activity and erectile dysfunction: An U-shaped relationship from population-based study. *International Journal of Obesity*, *31*(10), 1571–1578. https://doi.org/10.1038/sj.ijo.0803639

Cnattingius, S. (1992). Delayed Childbearing and Risk of Adverse Perinatal Outcome: A Population-Based Study. *JAMA*, *268*(7), 886. https://doi.org/10.1001/jama.1992.03490070068044

Corno, L., Hildebrandt, N., & Voena, A. (2020). Age of Marriage, Weather Shocks, and the Direction of Marriage Payments. *Econometrica*, *88*(3), 879–915. https://doi.org/10.3982/ECTA15505

Currie, J., & Neidell, M. (2005). Air pollution and infant health: What can we learn from California's recent experience? *Quarterly Journal of Economics*, *120*(3), 1003–1030. https://doi.org/10.1162/003355305774268219

Cvancarova, M., Samuelsen, S. O., Magelssen, H., & Fosså, S. D. (2009). Reproduction Rates After Cancer Treatment: Experience From the Norwegian Radium Hospital. *Journal of Clinical Oncology*, *27*(3), 334–343. https://doi.org/10.1200/JCO.2007.15.3130

Fagbamigbe, A. F., & Idemudia, E. S. (2016). Survival analysis and prognostic factors of timing of first childbirth among women in Nigeria. *BMC Pregnancy and Childbirth*, *16*(1), 102. https://doi.org/10.1186/s12884-016-0895-y

Farbmacher, H., & Tauchmann, H. (2023). Linear fixed-effects estimation with nonrepeated outcomes. *Econometric Reviews*, *42*(8), 635–654. https://doi.org/10.1080/07474938.2023.2224658

Fernandes, A. M., & Paunov, C. (2015). The Risks of Innovation: Are Innovating Firms Less Likely to Die? *The Review of Economics and Statistics*, *97*(3), 638–653. https://doi.org/10.1162/REST_a_00446

Frisco, M. L., & Weden, M. (2013). Early Adult Obesity and U.S. Women's Lifetime Childbearing Experiences: Early Adult Obesity and Lifetime Childbearing. *Journal of Marriage and Family*, *75*(4), 920–932. https://doi.org/10.1111/jomf.12049

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., & Schupp, J. (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*, *239*(2), 345–360. https://doi.org/10.1515/jbnst-2018-0022

Goebel, J., Grabka, M. M., Schröder, C., Zinn, S., Bartels, C., Franken, A., Gerike, M., Griese, F., Halbmeier, C., Kara, S., Krause, P., Liebau, E., Nebelin, J., Petrenz, M., Satilmis, S., Siegers, R., Steinhauer, H. W., Süttmann, F., Wenzig, K., ... Deutsches Institut Für Wirtschaftsforschung (DIW Berlin). (2023). *Socio-Economic Panel, data from 1984-2021, (SOEP-Core, v38, Remote Edition)Sozio-oekonomisches Panel, Daten der Jahre 1984-2021 (SOEP-Core, v38, Remote Edition)* (CSV,SPSS,Stata (bilingual),SPSS,RData; Version v38). CSV,SPSS,Stata (bilingual),SPSS,RData. SOEP Socio-Economic Panel Study. https://doi.org/10.5684/SOEP.CORE.V38R

Gordo, L. R. (2009). Why Are Women Delaying Motherhood in Germany? *Feminist Economics*, *15*(4), 57–75. https://doi.org/10.1080/13545700903153955

Heller, D. A., Ahern, F. M., Pringle, K. E., & Brown, T. V. (2009). Among older adults, the responsiveness of self-rated health to changes in Charlson comorbidity was moderated by age and baseline comorbidity. *Journal of Clinical Epidemiology*, *62*(2), 177–187. https://doi.org/10.1016/j.jclinepi.2008.05.009

Holton, S., Rowe, H., & Fisher, J. (2011). Women's Health and Their Childbearing Expectations and Outcomes: A Population-Based Survey from Victoria, Australia. *Women's Health Issues*, *21*(5), 366–373. https://doi.org/10.1016/j.whi.2011.02.006

Horrace, W. C., & Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, *90*(3), 321–327. https://doi.org/10.1016/j.econlet.2005.08.024

Huinink, J., Brüderl, J., Nauck, B., Walper, S., Castiglioni, L., & Feldhaus, M. (2011). Panel analysis of intimate relationships and family dynamics (pairfam): Conceptual framework and design. *Zeitschrift für Familienforschung*, *23*(1), 77–101.

Jenkins, S. P. (1995). Easy Estimation Methods for Discrete-Time Duration Models. *Oxford Bulletin of Economics and Statistics*, *57*(1), 129–136. https://doi.org/10.1111/j.1468-0084.1995.tb00031.x

Jin, L., & Ziebarth, N. R. (2020). Sleep, health, and human capital: Evidence from daylight saving time. *Journal of Economic Behavior & Organization*, *170*, 174–192. https://doi.org/10.1016/j.jebo.2019.12.003

Kind, M., & Kleibrink, J. (2013). Sooner or Later – Economic Insecurity and the Timing of First Birth. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2309908

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97–126. https://doi.org/10.1016/j.jeconom.2005.02.011

Kuhnt, A.-K., & Trappe, H. (2013, December). *Easier said than done: Childbearing intentions and their realization in a short term perspective* (0th ed., WP-2013-018). Max Planck Institute for Demographic Research. Rostock. https://doi.org/10.4054/MPIDR-WP-2013-018

Langeveld, N., Stam, H., Grootenhuis, M., & Last, B. (2002). Quality of life in young adult survivors of childhood cancer. *Supportive Care in Cancer*, *10*(8), 579–600. https://doi.org/10.1007/s00520-002-0388-6

Lesthaeghe, R., & Surkyn, J. (1988). Cultural Dynamics and Economic Theories of Fertility Change. *Population and Development Review*, *14*(1), 1. https://doi.org/10.2307/1972499

Luc, N., Thang, N. M., Swenson, I., & San, P. B. (1993). Selected determinants of fertility in vietnam: Age at marriage, marriage to first birth interval and age at first birth. *Journal of Biosocial Science*, *25*(3), 303–310. https://doi.org/10.1017/S0021932000020642

McGrath, J. J., Hearle, J., Jenner, L., Plant, K., Drummond, A., & Barkla, J. M. (1999). The fertility and fecundity of patients with psychoses. *Acta Psychiatrica Scandinavica*, *99*(6), 441–446. https://doi.org/10.1111/j.1600-0447.1999.tb00990.x

Mynarska, M., & Wróblewska, W. (2017). The health of women of reproductive age and their childbearing intentions. *Zdrowie Publiczne i Zarządzanie*, *15*(2). https://doi.org/10.4467/20842627OZ.17.014.6784

Nicoletti, C., & Rondinelli, C. (2010). The (mis)specification of discrete duration models with unobserved heterogeneity: A Monte Carlo study. *Journal of Econometrics*, *159*(1), 1–13. https://doi.org/10.1016/j.jeconom.2010.04.003

Ramlau-Hansen, C., Thulstrup, A., Nohr, E., Bonde, J., Sørensen, T., & Olsen, J. (2007). Subfecundity in overweight and obese couples. *Human Reproduction*, *22*(6), 1634–1637. https://doi.org/10.1093/humrep/dem035

Rindfuss, R. R., & St. John, C. (1983). Social Determinants of Age at First Birth. *Journal of Marriage and Family*, *45*(3), 553–565. https://doi.org/10.2307/351660

Robles, T. F., Slatcher, R. B., Trombello, J. M., & McGinn, M. M. (2014). Marital quality and health: A meta-analytic review. *Psychological Bulletin*, *140*(1), 140–187. https://doi.org/10.1037/a0031859

Schaubert, M. (2015). Career, Private Life, and the Well-Being among College-educated West German Women. Retrieved May 22, 2023, from https://www.econstor.eu/handle/10419/184885

Schleutker, E. (2014). Determinants of Childbearing: A Review of the Literature. *Zeitschrift für Soziologie*, *43*(3), 192–211. https://doi.org/10.1515/zfsoz-2014-0303

Schover, L. R., Rybicki, L. A., Martin, B. A., & Bringelsen, K. A. (1999). Having children after cancer: A pilot survey of survivors' attitudes and experiences. *Cancer*, *86*(4), 697–709. https://doi.org/10.1002/(SICI)1097-0142(19990815)86:4<697::AID-CNCR20>3.0.CO;2-J

Singer, J. D., & Willett, J. B. (1993, Summer). It's about Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events. *Journal of Educational Statistics*, *18*(2), 155. https://doi.org/10.2307/1165085

Statistisches Bundesamt. (2023a). *12612-0005: Lebendgeborene: Deutschland, Jahre, Alter der Mutter, Lebendgeburtenfolge* (Text). Text. Retrieved June 15, 2023, from https://www-genesis.destatis.de/genesis//online?operation=table&code=12612-0005&bypass=true&levelindex=1&levelid=1686820113270#abreadcrumb

Statistisches Bundesamt. (2023b). *12612-0009: Zusammengefasste Geburtenziffern (je Frau): Deutschland, Jahre, Altersgruppen*. Retrieved June 15, 2023, from https://www-genesis.destatis.de/genesis//online?operation=table&code=12612-0009&bypass=true&levelindex=0&levelid=1686820274133#abreadcrumb

Statistisches Bundesamt. (2023c). *Mothers' average age at childbirth Germany 2021*. Retrieved June 15, 2023, from https://www.statista.com/statistics/1087224/mothers-average-age-at-childbirths-germany/

Syse, A., Dommermuth, L., & Hart, R. K. (2020). Does health influence fertility? *Discussion Papers*.

Toft, A. M. H., Møller, H., & Laursen, B. (2010). The Years After an Injury: Long-Term Consequences of Injury on Self-Rated Health. *Journal of Trauma: Injury, Infection & Critical Care*, *69*(1), 26–30. https://doi.org/10.1097/TA.0b013e3181d3cbf2

Tough, S., Benzies, K., Fraser-Lee, N., & Newburn-Cook, C. (2007). Factors Influencing Childbearing Decisions and Knowledge of Perinatal Risks among Canadian Men and Women. *Maternal and Child Health Journal*, *11*(2), 189–198. https://doi.org/10.1007/s10995-006-0156-1

Tutz, G., & Schmid, M. (2016). *Modeling Discrete Time-to-Event Data* (1st ed. 2016). Springer International Publishing : Imprint: Springer. https://doi.org/10.1007/978-3-319-28158-2

Van De Kaa, D. J. (1987). Europe's second demographic transition. *Population Bulletin*, *42*(1), 1–59.

van der Vleuten, M., Jaspers, E., & van der Lippe, T. (2021). Same-Sex Couples' Division of Labor from a Cross-National Perspective. *Journal of GLBT Family Studies*, *17*(2), 150–167. https://doi.org/10.1080/1550428X.2020.1862012

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) - Evolution, Scope and Enhancements. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1028709

Zebrack, B. J., Casillas, J., Nohr, L., Adams, H., & Zeltzer, L. K. (2004). Fertility issues for young adult survivors of childhood cancer. *Psycho-Oncology*, *13*(10), 689–699. https://doi.org/10.1002/pon.784

Zhang, Y., & Axinn, W. G. (2021). Marital Experiences and Depression in an Arranged Marriage Setting. *American Journal of Sociology*, *126*(6), 1439–1486. https://doi.org/10.1086/714272

6 Appendix

Table A1.1: Statistics of observation period

Length of observed spell			Number of observed waves		
Number of years	Not mothers	Mothers	Number of waves	Not mothers	Mothers
1	4,290	466	1	4,290	466
2	2,295	328	2	2,635	359
3	1,377	259	3	1,474	268
4	1,139	225	4	1,114	232
5	742	189	5	748	203
6	634	155	6	597	159
7	513	130	7	490	127
8	479	116	8	449	109
9	379	109	9	374	95
10	423	79	10	334	75
11	353	60	11	301	51
12	319	44	12	228	39
13	112	36	13	93	28
14	84	23	14	73	16
15	67	18	15	51	16
16	61	9	16	57	5
17	54	7	17	43	8
18	38	7	18	32	7
19	44	4	19	38	5
20	28	3	20	34	0
21	37	2	21	25	1
22	19	0	22	19	0
23	21	0	23	19	0
24	5	0	24	2	0
25	5	0	25	7	0
26	4	0	26	4	0
27	13	0	27	4	0

Notes: *The Table presents the distribution of the years a woman is observed (length of observed spell (possibly including gaps)) as well as the number of waves grouped by ever becoming a mother. The results are based on 64,533 observations from 15,804 individuals. The length of spell is calculated as the latest year minus the first year of participation in the survey plus one.*

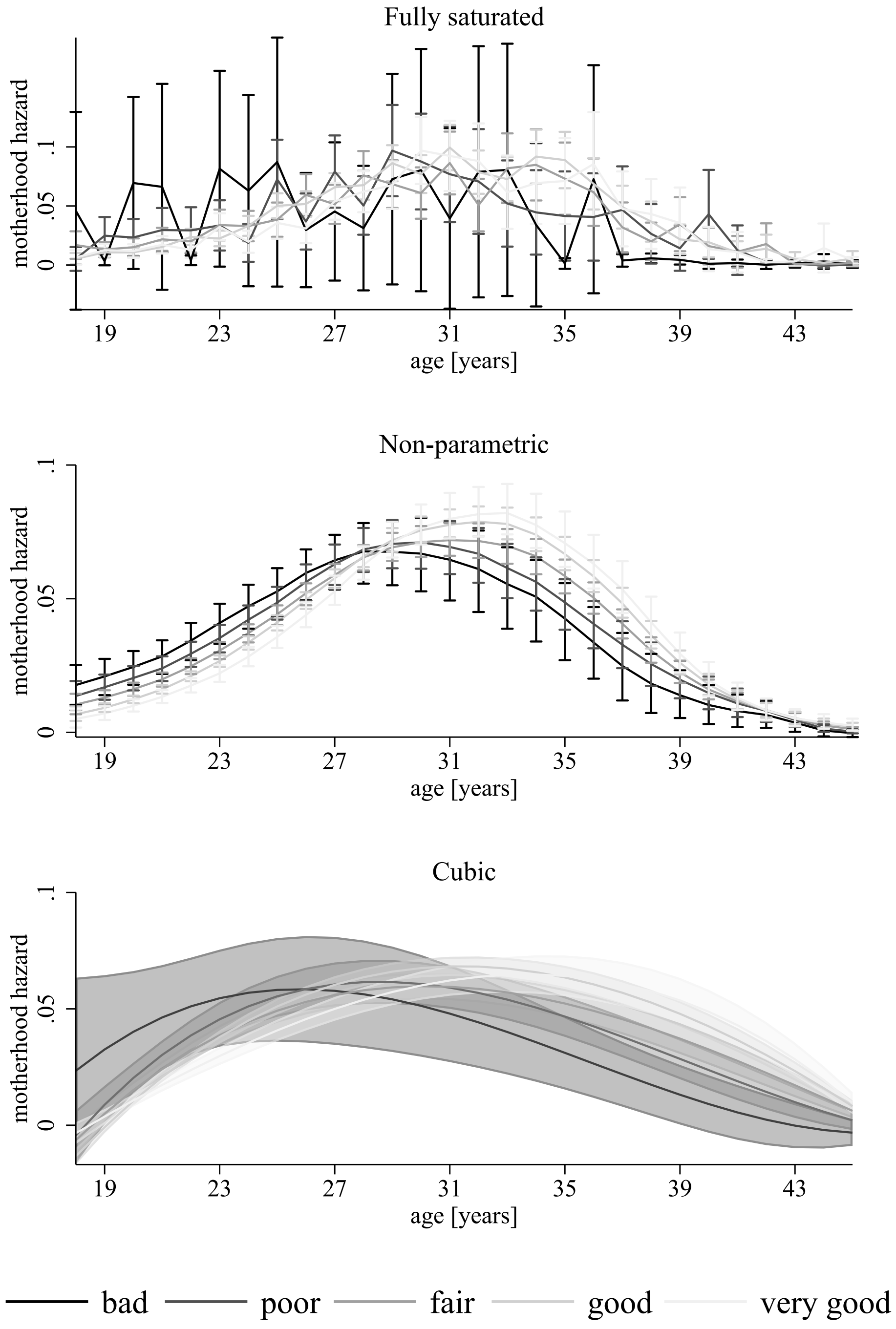
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own calculation.*

Table A1.2: Descriptive statistics of covariates

	Pooled sample		Birth = yes	
	Mean	SD	Mean	SD
Equivalized HH Income	1,950.25	1,525.71	2,049.30	1,309.61
Age	26.88	8.12	27.53	4.93
Education				
In school	0.16	0.37	0.04	0.19
No achieved education	0.02	0.13	0.02	0.13
General elementary	0.12	0.33	0.09	0.29
Middle vocational	0.38	0.49	0.38	0.49
Vocational (with Abitur)	0.09	0.29	0.12	0.33
Higher (vocational) education	0.22	0.42	0.35	0.48
Employment				
Not working	0.29	0.45	0.13	0.34
Full-time	0.45	0.50	0.66	0.48
Part-time	0.26	0.44	0.21	0.41
Migration background				
No migration background	0.76	0.43	0.76	0.43
Direct migration background	0.12	0.33	0.16	0.36
Indirect migration background	0.12	0.33	0.09	0.28
Number of obs.	64,533		2,269	
Number of ind.	15,804		2,269	

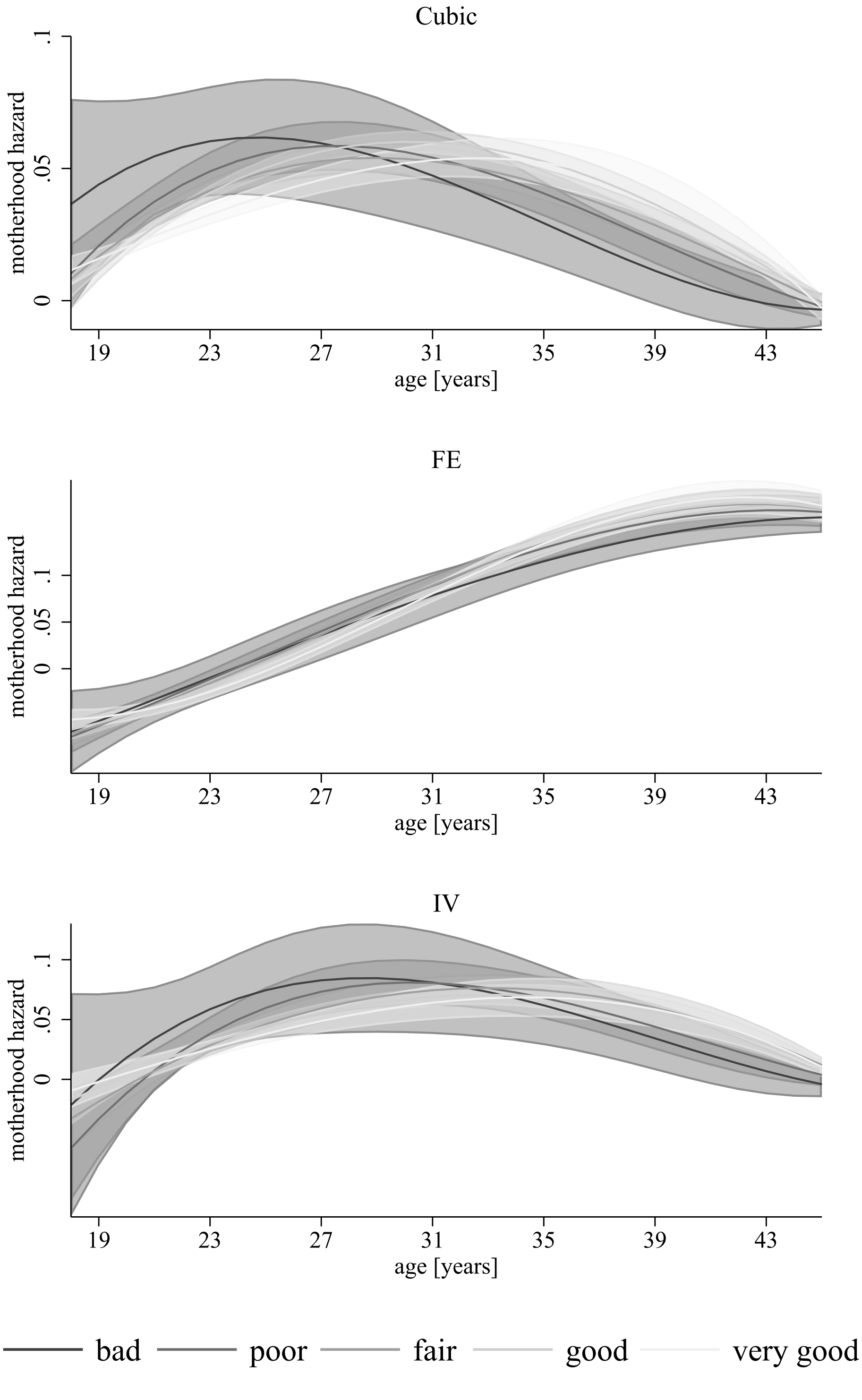
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own calculation.*

Figure A1.1: Hazard function with CI: estimated using different empirical approaches



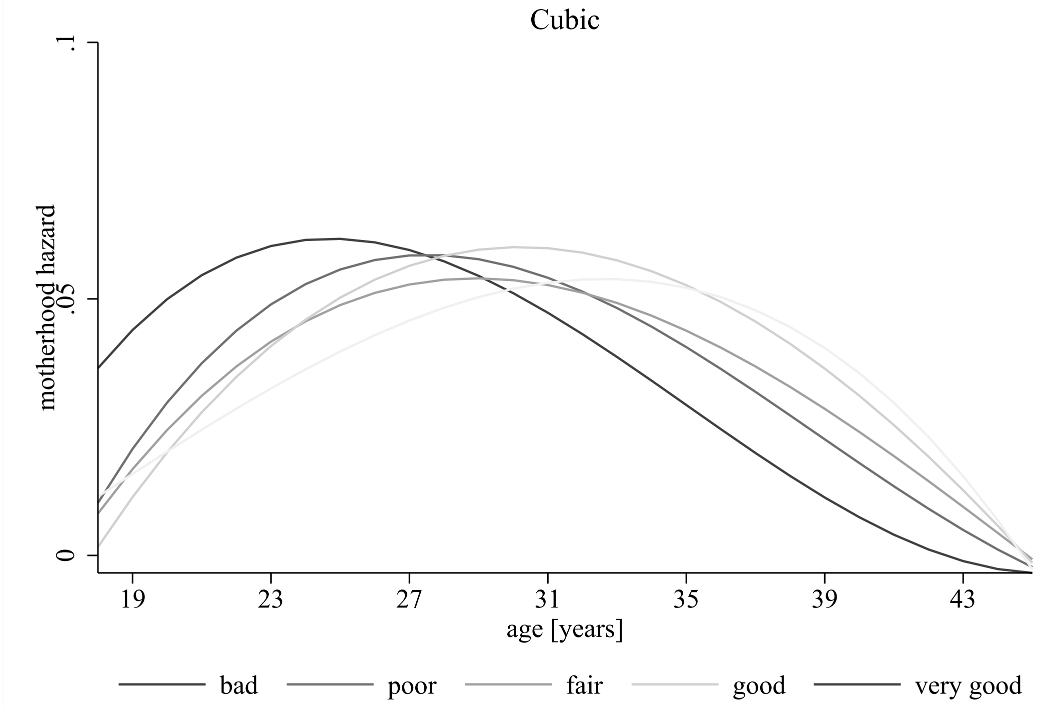
__Notes:__ 95% confidence intervals are presented. The CI for non-parametric estimation are computed non-parametrically and based on . bootstrap replications. The standard errors for cubic regression are calculated using delta-method. Control variables are not included; year dummies are included. The results are based on 64,533 observations. Panel: Fully saturated (model (ii)) – the set of all interactions of age and SAH is included. Panel: Non-parametric (model (iii)) – the model is estimated using kernel regression. Panel: Cubic (model (iv)) – the set of interactions of SAH with third-order polynomial of age is included.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

Figure A1.2: Hazard function with CI: observed and unobserved heterogeneity



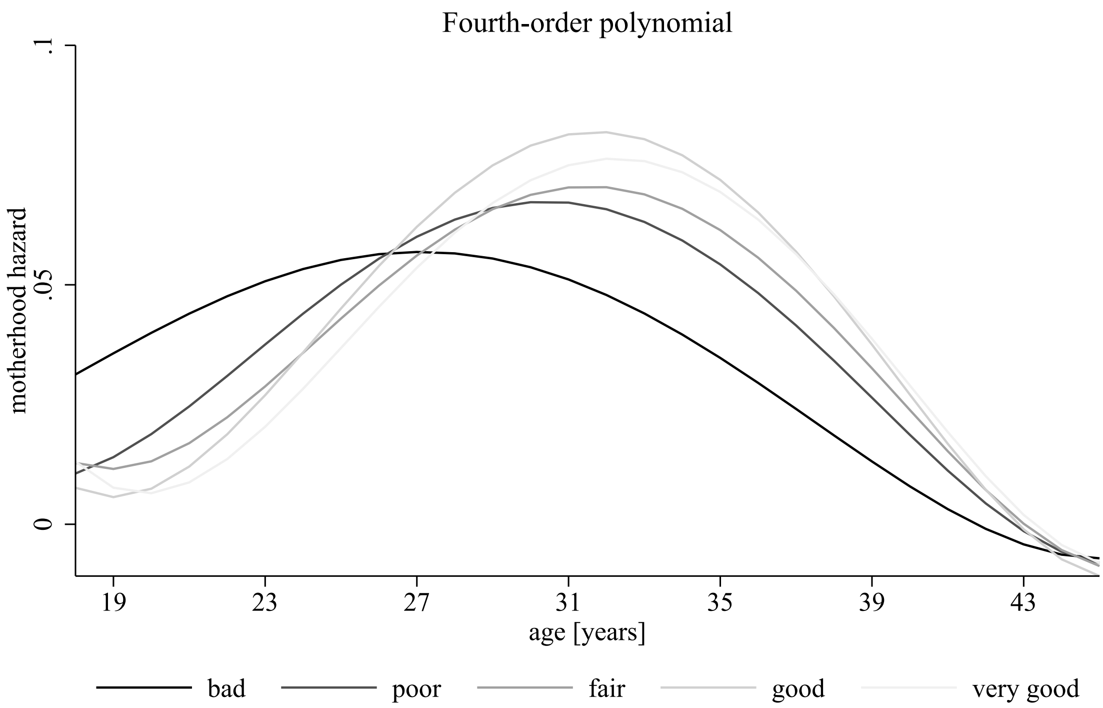
Notes: *95% confidence intervals are represented by filled areas. The standard errors are calculated using delta-method. Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included in all three panels. Panel Cubic (model (v)) is estimated using OLS; Panel FE (model (vi)) – using fixed effects; Panel IV (model (vii)) – using first differences as instruments for levels. The results are based on 64,533 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

Figure A1.3: Robustness check: complementary log-log



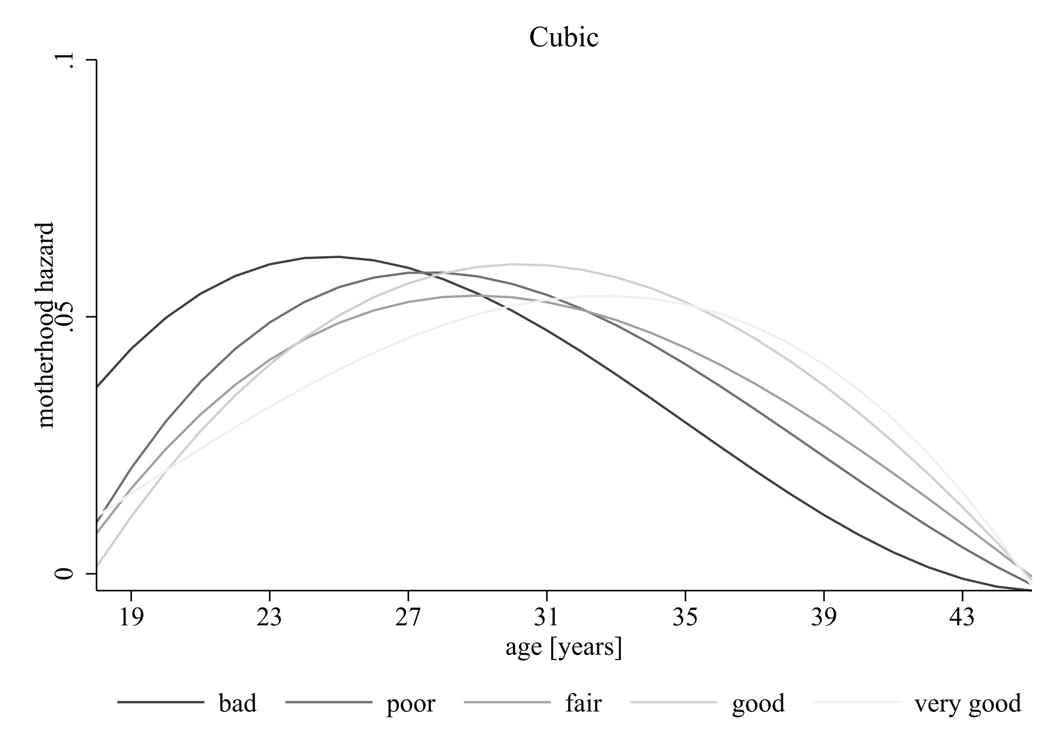
__Notes:__ The model is estimated using complementary log-log. Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to 0.000.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

Figure A1.4: Robustness check: fourth-order polynomial



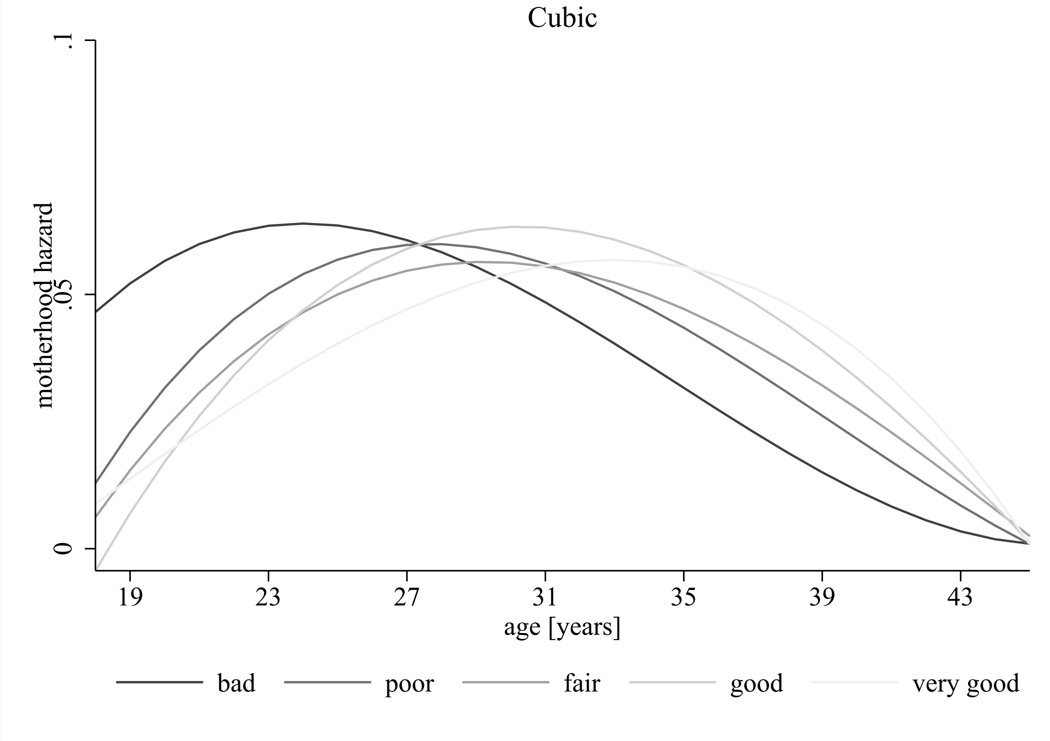
__Notes:__ Control variables are not included, years dummies are included. The results are based on 64,533 observations.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

Figure A1.5: Robustness check: other controls (net income)



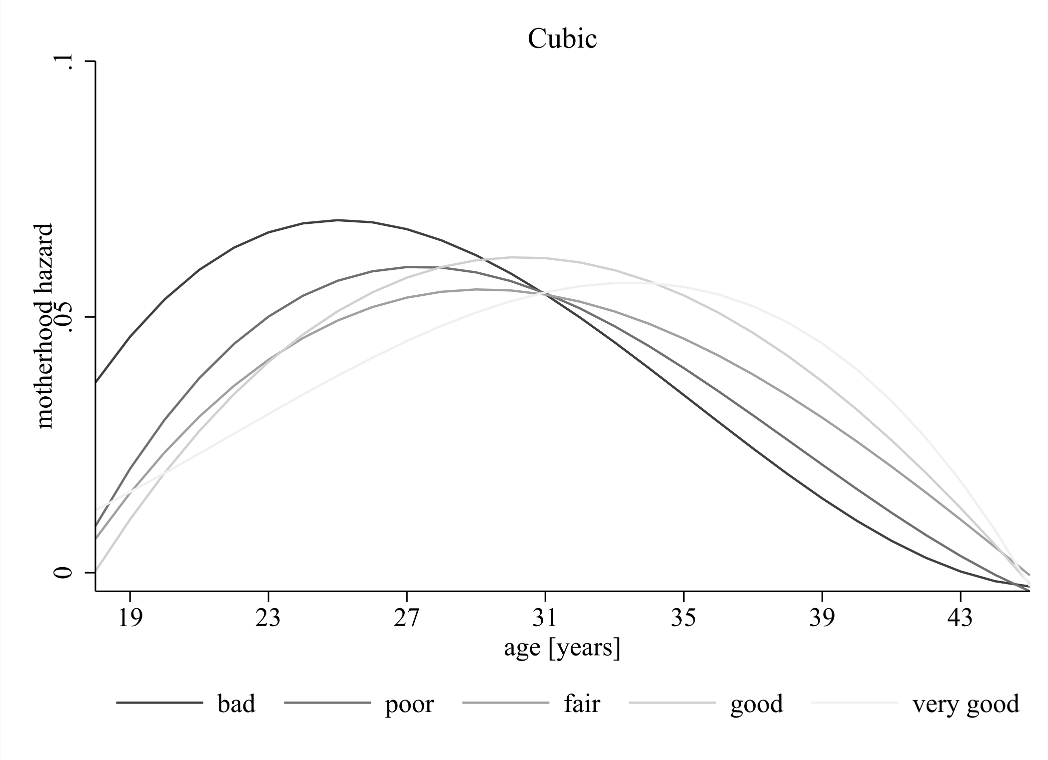
Notes: *Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to p-value = 0.000. The estimation is based on 64,533 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

Figure A1.6: Robustness check: other controls (years of education)



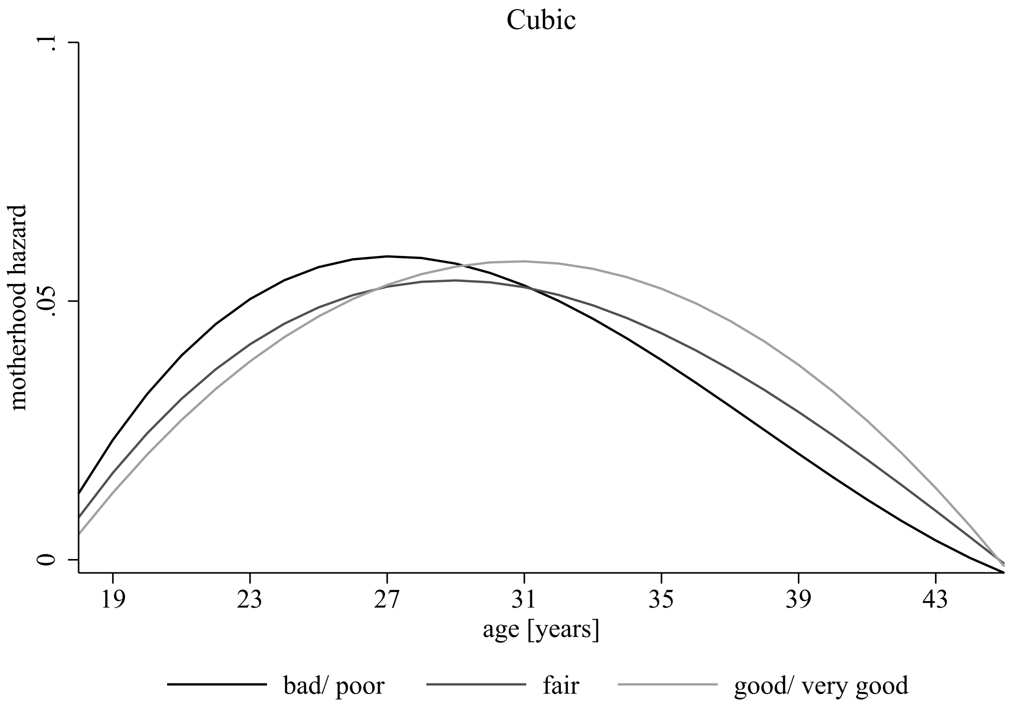
Notes: *Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to p-value = 0.000. The estimation is based on 57,646 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

Figure A1.7: Robustness check: other controls (additionally, indicator for whether parents completed high school)



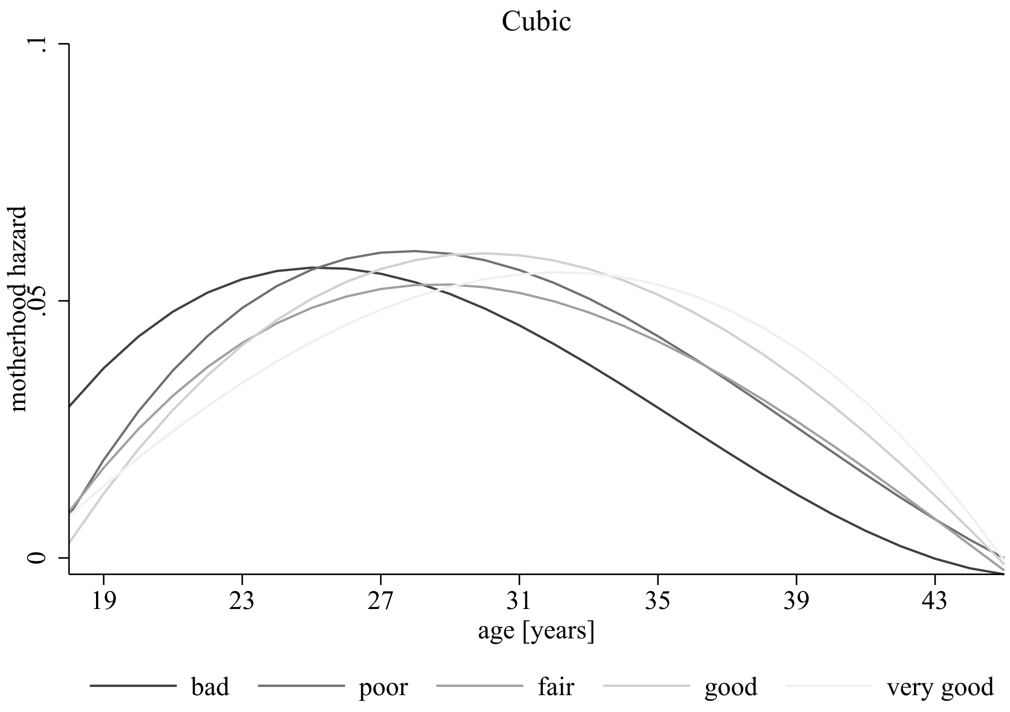
__Notes:__ Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to p-value = 0.000. The estimation is based on 57,621 observations.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

Figure A1.8: Robustness check: SAH in three categories



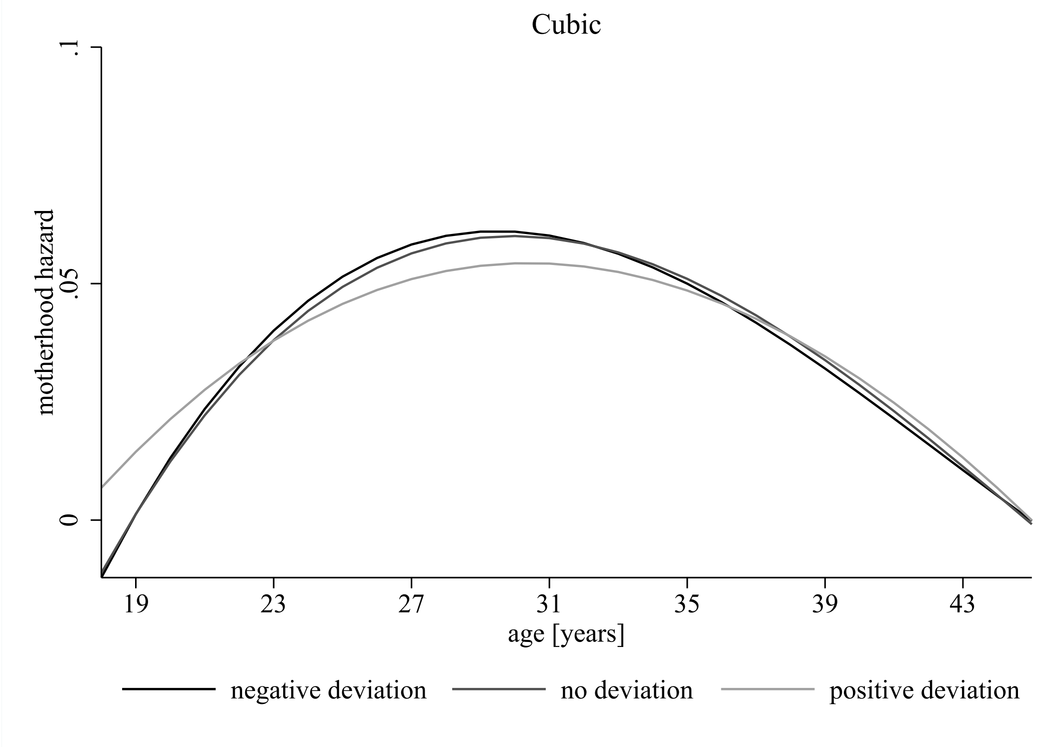
__Notes:__ Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to 0.000.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

Figure A1.9: Robustness check: SAH fixed at entry level



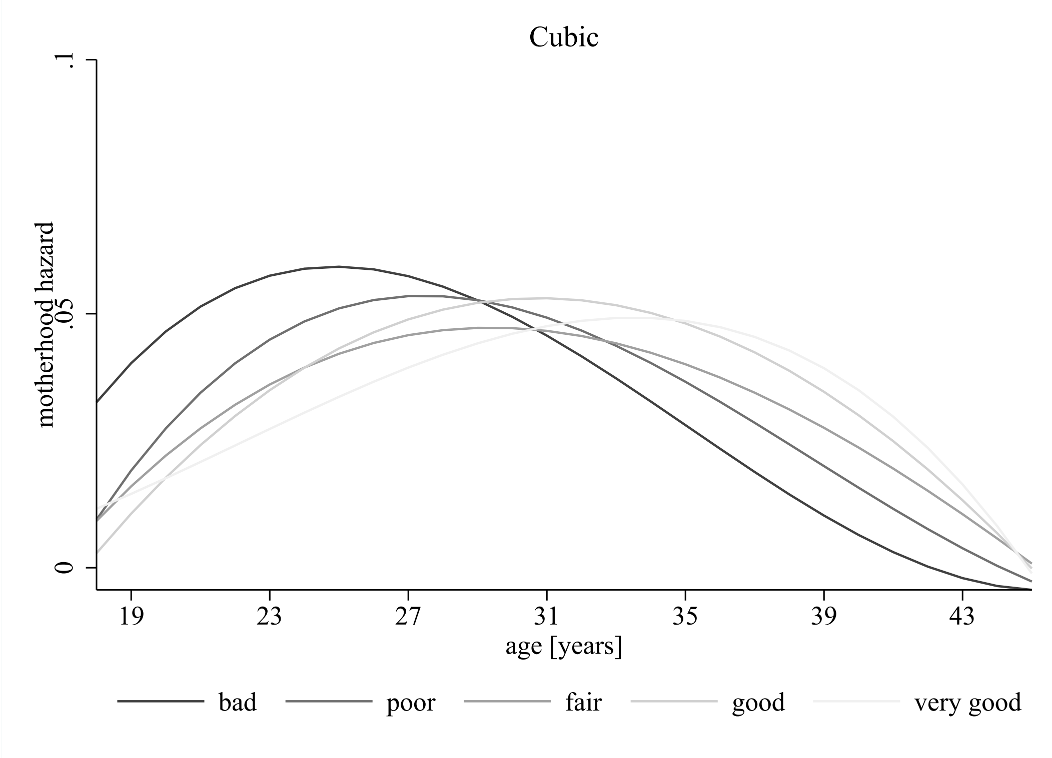
__Notes:__ Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to 0.000.
__Source:__ SOEP v38 (from 1994) and PAIRFAM v13; own illustration.

Figure A1.10: Robustness check: health deviations



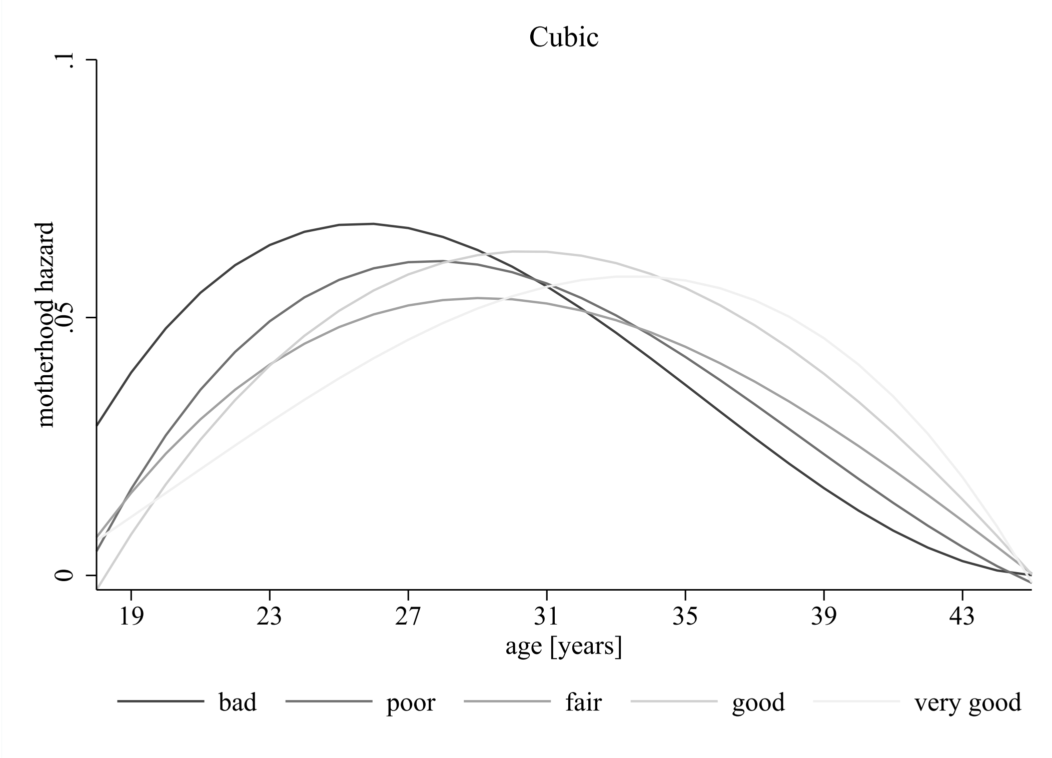
Notes: *Controls and year dummies are included. We define a positive deviation as an improvement and a negative deviation as worsening of health in comparison to the best observed health in the past. The set of interactions of deviation dummies with third-order polynomial of age is included. The p-value of test of joint significance of deviation indicators and age polynomials is equal to 0.039. The estimation is based on 48,729 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

Figure A1.11: Robustness check: self-selection into motherhood



Notes: *95% confidence intervals are represented by filled areas. Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included in all three panels. Potentially self-selected women are excluded. The results are based on 58,390 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

Figure A1.12: Robustness check: only women without migration background



Notes: *Controls and year dummies are included. The set of interactions of SAH with third-order polynomial of age is included. The p-value of test of joint significance of self-assessed health indicators and age polynomials is equal to p-value = 0.000. The estimation is based on 48,753 observations.*
Source: *SOEP v38 (from 1994) and PAIRFAM v13; own illustration.*

Table A1.3: Regression results (cubic specifications)

	OLS	OLS	FE	IV
Age	0.49	-0.13	-6.91***	0.00
	(0.79)	(0.81)	(1.14)	(.)
Age (squared)	2.21	3.33	27.95***	4.26***
	(2.67)	(2.69)	(3.80)	(0.66)
Age (cubic)	-5.83*	-6.37*	-31.18***	-8.27***
	(2.82)	(2.84)	(3.95)	(1.23)
Bad	-0.38	-0.41	-0.67*	-0.92
	(0.29)	(0.29)	(0.33)	(0.67)
Poor	-0.50***	-0.48***	-0.60***	-1.01**
	(0.13)	(0.13)	(0.17)	(0.36)
Fair	-0.34***	-0.33**	-0.42**	-0.38
	(0.10)	(0.10)	(0.13)	(0.24)
Good	-0.33***	-0.33***	-0.30**	-0.22
	(0.09)	(0.09)	(0.11)	(0.20)
Bad * age	4.94	5.12	7.06*	9.60
	(2.79)	(2.78)	(3.16)	(6.17)
Bad * age (squared)	-18.70*	-18.92*	-23.32*	-30.80
	(8.38)	(8.35)	(9.44)	(18.11)
Bad * age (cubic)	21.19**	21.20**	24.17**	31.00
	(8.06)	(8.02)	(9.03)	(17.10)
Poor * age	5.45***	5.22***	6.11***	9.83**
	(1.34)	(1.34)	(1.70)	(3.40)
Poor * age (squared)	-18.55***	-17.65***	-19.33***	-30.11**
	(4.29)	(4.27)	(5.40)	(10.32)
Poor * age (cubic)	19.68***	18.69***	19.35***	29.42**
	(4.36)	(4.34)	(5.46)	(10.06)
Fair * age	3.70***	3.56***	4.18**	3.55
	(1.06)	(1.05)	(1.37)	(2.28)
Fair * age (squared)	-12.56***	-12.03***	-12.95**	-10.40
	(3.46)	(3.45)	(4.45)	(7.04)
Fair * age (cubic)	13.34***	12.78***	12.72**	9.78
	(3.58)	(3.58)	(4.59)	(6.99)
Good * age	3.37***	3.33***	2.96*	2.17
	(0.93)	(0.93)	(1.17)	(1.93)
Good * age (squared)	-10.62***	-10.49***	-9.21*	-6.71
	(3.09)	(3.08)	(3.86)	(6.13)
Good * age (cubic)	10.61**	10.50**	9.09*	6.59
	(3.25)	(3.23)	(4.04)	(6.22)
Number of obs.	64,533	64,533	64,533	45,698
Joint sign. test	0.000	0.000	0.004	0.003
Controls	No	Yes	Yes	Yes

Notes: ** p<0.05, ** p<0.01, *** p<0.001. Individual-level clustered standard errors are in parentheses. Year fixed effects are included in the model. "Joint sign. test" refers to joint significant test of interactions of self-assessed health and age polynomials.*

Source: *SOEP v38 (from 1994) and PAIRFAM v13; own calculation.*

Table A1.4: Parametric regressions comparison

	Linear	Squared	Cubic	Fourth	Fifth
Age (linear)	0.15***	2.13***	0.49	-32.41***	-19.03
	(0.02)	(0.14)	(0.79)	(4.09)	(19.79)
Age (squared)		-3.27***	2.21	167.64***	77.29
		(0.21)	(2.67)	(20.33)	(132.47)
Age (cubic)			-5.83*	-362.30***	-65.94
			(2.82)	(43.11)	(430.38)
Age (fourth)				278.22***	-194.33
				(33.04)	(678.90)
Age (fifth)					293.36
					(416.51)
Number of obs.	64,533	64,533	64,533	64,533	64,533
Joint sign. test	0.000	0.000	0.000	0.031	0.316
adj. R-squared	0.005	0.020	0.021	0.024	0.024
AIC	-35,475	-36,492	-36,512	-36,719	-36,739
BIC	-35,149	-36,120	-36,095	-36,265	-36,340

Notes: ** p<0.05, ** p<0.01, *** p<0.001. Individual-level clustered standard errors are in parentheses. "Joint sign. test" refers to joint significant test of interactions of self-assessed health indicators and the highest polynomial in the model. Year fixed effects are included in the model.*

Source: *SOEP v38 (from 1994) and PAIRFAM v13; own calculation.*

Table A1.5: Regression results (dummies specification)

	Unconditional hazard	Conditional on SAH	Fully saturated
19	0.00	0.00	-0.04
	(0.00)	(0.00)	(0.04)
20	0.01*	0.01*	0.02
	(0.00)	(0.00)	(0.06)
21	0.01***	0.01***	0.02
	(0.00)	(0.00)	(0.06)
22	0.01***	0.01***	-0.04
	(0.00)	(0.00)	(0.04)
23	0.02***	0.02***	0.04
	(0.00)	(0.00)	(0.06)
24	0.02***	0.02***	0.02
	(0.00)	(0.00)	(0.06)
25	0.04***	0.04***	0.04
	(0.00)	(0.00)	(0.07)
26	0.04***	0.04***	-0.02
	(0.00)	(0.00)	(0.05)
27	0.05***	0.05***	-0.00
	(0.00)	(0.00)	(0.05)
28	0.06***	0.06***	-0.01
	(0.00)	(0.00)	(0.05)
29	0.07***	0.07***	0.03
	(0.01)	(0.01)	(0.06)
30	0.07***	0.07***	0.03
	(0.01)	(0.01)	(0.07)
31	0.08***	0.08***	-0.01
	(0.01)	(0.01)	(0.06)
32	0.07***	0.07***	0.03
	(0.01)	(0.01)	(0.07)
33	0.06***	0.06***	0.03
	(0.01)	(0.01)	(0.07)
34	0.07***	0.07***	-0.01
	(0.01)	(0.01)	(0.06)
35	0.07***	0.07***	-0.04
	(0.01)	(0.01)	(0.04)
36	0.06***	0.06***	0.03
	(0.01)	(0.01)	(0.07)
37	0.04***	0.04***	-0.04
	(0.01)	(0.01)	(0.04)
38	0.02***	0.02***	-0.04
	(0.01)	(0.01)	(0.04)
39	0.02***	0.02**	-0.04
	(0.01)	(0.01)	(0.04)
40	0.01*	0.01*	-0.04
	(0.00)	(0.00)	(0.04)
41	0.00	0.00	-0.04
	(0.00)	(0.00)	(0.04)
42	0.00	0.00	-0.05
	(0.00)	(0.00)	(0.04)
43	-0.00*	-0.01*	-0.04
	(0.00)	(0.00)	(0.04)
44	-0.01*	-0.01*	-0.04
	(0.00)	(0.00)	(0.04)
45	-0.00	-0.00*	-0.05

	(0.00)	(0.00)	(0.04)
46	-0.01*	-0.01*	-0.05
	(0.00)	(0.00)	(0.04)
47	-0.01*	-0.01**	-0.04
	(0.00)	(0.00)	(0.04)
48	-0.00*	-0.00*	-0.05
	(0.00)	(0.00)	(0.04)
49	-0.00*	-0.00*	-0.04
	(0.00)	(0.00)	(0.04)
Poor		0.00	-0.04
		(0.01)	(0.04)
Fair		0.00	-0.03
		(0.01)	(0.04)
Good		0.00	-0.04
		(0.01)	(0.04)
Very good		-0.00	-0.04
		(0.01)	(0.04)
Poor * 19			0.06
			(0.04)
Poor * 20			-0.01
			(0.06)
Poor * 21			0.00
			(0.06)
Poor * 22			0.07
			(0.04)
Poor * 23			-0.01
			(0.06)
Poor * 24			-0.00
			(0.06)
Poor * 25			0.03
			(0.07)
Poor * 26			0.05
			(0.05)
Poor * 27			0.07
			(0.05)
Poor * 28			0.06
			(0.05)
Poor * 29			0.06
			(0.07)
Poor * 30			0.05
			(0.07)
Poor * 31			0.08
			(0.06)
Poor * 32			0.03
			(0.07)
Poor * 33			0.01
			(0.07)
Poor * 34			0.05
			(0.06)
Poor * 35			0.08
			(0.05)
Poor * 36			0.01
			(0.07)
Poor * 37			0.08
			(0.05)
Poor * 38			0.06

	(0.04)
Poor * 39	0.05
	(0.04)
Poor * 40	0.08
	(0.05)
Poor * 41	0.05
	(0.04)
Poor * 42	0.04
	(0.04)
Poor * 43	0.04
	(0.04)
Poor * 44	0.04
	(0.04)
Poor * 45	0.04
	(0.04)
Poor * 46	0.04
	(0.04)
Poor * 47	0.04
	(0.04)
Poor * 48	0.04
	(0.04)
Poor * 49	0.04
	(0.04)
Fair * 19	0.04
	(0.04)
Fair * 20	-0.03
	(0.06)
Fair * 21	-0.02
	(0.06)
Fair * 22	0.05
	(0.04)
Fair * 23	-0.02
	(0.06)
Fair * 24	-0.00
	(0.06)
Fair * 25	-0.02
	(0.07)
Fair * 26	0.06
	(0.05)
Fair * 27	0.04
	(0.05)
Fair * 28	0.07
	(0.05)
Fair * 29	0.02
	(0.06)
Fair * 30	0.01
	(0.07)
Fair * 31	0.08
	(0.06)
Fair * 32	0.00
	(0.07)
Fair * 33	0.03
	(0.07)
Fair * 34	0.08
	(0.06)
Fair * 35	0.10*

	(0.05)
Fair * 36	0.02
	(0.07)
Fair * 37	0.06
	(0.04)
Fair * 38	0.04
	(0.04)
Fair * 39	0.06
	(0.04)
Fair * 40	0.04
	(0.04)
Fair * 41	0.04
	(0.04)
Fair * 42	0.05
	(0.04)
Fair * 43	0.03
	(0.04)
Fair * 44	0.03
	(0.04)
Fair * 45	0.03
	(0.04)
Fair * 46	0.03
	(0.04)
Fair * 47	0.03
	(0.04)
Fair * 48	0.03
	(0.04)
Fair * 49	0.03
	(0.04)
Good * 19	0.05
	(0.04)
Good * 20	-0.02
	(0.06)
Good * 21	-0.01
	(0.06)
Good * 22	0.06
	(0.04)
Good * 23	-0.02
	(0.06)
Good * 24	0.01
	(0.06)
Good * 25	0.00
	(0.07)
Good * 26	0.06
	(0.05)
Good * 27	0.06
	(0.05)
Good * 28	0.08
	(0.05)
Good * 29	0.05
	(0.06)
Good * 30	0.04
	(0.07)
Good * 31	0.10
	(0.06)
Good * 32	0.04

	(0.07)
Good * 33	0.03
	(0.07)
Good * 34	0.10
	(0.06)
Good * 35	0.13**
	(0.04)
Good * 36	0.04
	(0.07)
Good * 37	0.09
	(0.04)
Good * 38	0.07
	(0.04)
Good * 39	0.06
	(0.04)
Good * 40	0.06
	(0.04)
Good * 41	0.05
	(0.04)
Good * 42	0.05
	(0.04)
Good * 43	0.04
	(0.04)
Good * 44	0.04
	(0.04)
Good * 45	0.05
	(0.04)
Good * 46	0.05
	(0.04)
Good * 47	0.04
	(0.04)
Good * 48	0.04
	(0.04)
Good * 49	0.04
	(0.04)
Very good * 19	0.05
	(0.04)
Very good * 20	-0.02
	(0.06)
Very good * 21	-0.01
	(0.06)
Very good * 22	0.05
	(0.04)
Very good * 23	-0.01
	(0.06)
Very good * 24	-0.00
	(0.06)
Very good * 25	-0.01
	(0.07)
Very good * 26	0.04
	(0.05)
Very good * 27	0.05
	(0.05)
Very good * 28	0.07
	(0.05)
Very good * 29	0.03

			(0.06)
Very good * 30			0.06
			(0.07)
Very good * 31			0.09
			(0.06)
Very good * 32			0.05
			(0.07)
Very good * 33			0.02
			(0.07)
Very good * 34			0.07
			(0.06)
Very good * 35			0.11*
			(0.05)
Very good * 36			0.05
			(0.07)
Very good * 37			0.08
			(0.05)
Very good * 38			0.08
			(0.05)
Very good * 39			0.07
			(0.05)
Very good * 40			0.05
			(0.04)
Very good * 41			0.05
			(0.04)
Very good * 42			0.04
			(0.04)
Very good * 43			0.04
			(0.04)
Very good * 44			0.05
			(0.04)
Very good * 45			0.04
			(0.04)
Very good * 46			0.04
			(0.04)
Very good * 47			0.04
			(0.04)
Very good * 48			0.04
			(0.04)
Very good * 49			0.04
			(0.04)
Number of obs.	64,533	64,533	64,533
Joint sign. test		0.245	0.000

Notes: *$p<0.05$, ** $p<0.01$, *** $p<0.001$.*

Joint sign. test refers either to a test of joint significance of sah indicators or its full set of interactions with age.

Individual-level clustered standard errors are in parentheses. Year fixed effects are included.

Source: *SOEP v38 (from 1994) and PAIRFAM v13; own calculation.*