Raimond Maurer | Sehrish Usman

Dynamics of Life Course Family Transitions in Germany: Exploring Patterns, Process and Relationships

SAFE Working Paper No. 399 | September 2023

Leibniz Institute for Financial Research SAFE
Sustainable Architecture for Finance in Europe
info@safe-frankfurt.de | www.safe-frankfurt.de

Dynamics of Life Course Family Transitions in Germany:

Exploring Patterns, Process and Relationships

Raimond Maurer and Sehrish Usman

August 28, 2023

Raimond Maurer
Finance Department, Goethe- University Frankfurt and Leibniz-Institute for Financial Research (SAFE), Theodor-W.-Adorno-Platz 3, Frankfurt Germany
Email: maurer@finance.uni-frankfurt.de

Sehrish Usman
Finance Department, Goethe- University Frankfurt and Leibniz-Institute for Financial Research (SAFE), Theodor-W.-Adorno-Platz 3, Frankfurt Germany
Email: usman@finance.uni-frankfurt.de

Abstract

This paper explores dynamics of family life events in Germany using discrete time event history analysis based on SOEP data. We find that higher educational attainment, better income level, and marriage emerge as salient protective factors mitigating the risk of mortality; better education also reduces the likelihood of first marriage whereas, lower educational attainment, protracted period, and presence of children act as protective factors against divorce. Our key finding shows that disparity in mean life expectancies between individuals from low- and high-income brackets is observed to be 9 years among males and 6 years among females, thereby illustrating the mortality inequality attributed to income disparities. Our estimates show that West Germans have low risk of death, less likelihood of first marriage, and they have a high risk of divorce and remarriage compared to East Germans.

Keywords: Family dynamics, Life Events, Hazard estimation, Life course transitions

JEL classifications: C13, C34, J12

---

*We gratefully acknowledge research support from the Leibniz Institute for Financial Research (SAFE).

1 Introduction

People live in diverse living arrangements in Germany, and new living arrangements are shaping the demographics of the country. The aging population, low fertility and birth rates, changes in family composition and size, increase in non-traditional living relationships like cohabitation and delayed marriages are a few examples. The most important determinants of these changes include low mortality rates causing an increase in life expectancy, changes in societal and cultural factors, increase immigration, economic factors, and government policies. Furthermore, aggregate demographic changes are indeed influenced by personal choices and life course marital transitions of individuals, that are less talked about in literature. The preferences and decisions of individuals regarding their education, career, relationship, and family planning shape the current and future demographic landscape of the country for example, rise of non-marital cohabitation, increase in out-of-wedlock children and delayed marriages in Germany.

The aim of this paper is to explore dynamics of family life events which we broadly categorize into marriage, fertility, divorce, and mortality (self and spouse). This involves unfolding the patterns, process, and interactions within and between different family life events. We aim to predict the likelihood of experiencing these events across heterogeneous individuals using their demographic and economic characteristics. Building upon this, we aim to identify distinct distributions of these events among various subgroups. Specifically explore the following questions: Why certain individuals are at higher risk of experiencing certain events compared to others. What are the characteristics of these individuals that make them vulnerable to these events? To analyze these questions, we develop hazard models for each event of interest to predict the likelihood of an event occurring at each specific time, using a set of covariates. These models allow us to comprehend family dynamics by quantifying risks associated with these state transitions.

In pursuit of this objective first, we analyze different patterns of each life event using non-parametric analysis. Second, using empirical data we model each process to make predictions, third, we evaluate the interactions between these events to explore their interconnectedness. To accomplish this, we use German SOEP datasets (DIW) and adapt discrete time event history methodological framework. For this sort of data, event history models are preferred over standard statistical techniques as they can effectively handle censoring issues and time-varying covariates.

The findings of this paper provide researchers and policymakers with insights into actual dynamics of demographic and family transitions, thereby facilitating them to make informed policy decisions. This paper also contributes to discovering the social implications of various transitions such as consequences of low life expectancy for the people with low socioeconomic status. Understanding the relative vulnerabilities of different groups to specific events can help to uncover disparities and inequalities that may exist. By examining the factors which can affect the likelihood of such transitions, the decision makers can gain a comprehensive situation of the current policy and can evaluate the existing policies. This paper is not after establishing any causal analysis however, it serves as a foundation for normative analysis by pinpointing the empirical realities.

Our main results suggest that higher educational attainment, better income level, and marriage emerge as salient protective factors mitigating the risk of mortality. We find that discrepancy in mean life expectancies between low- and high-income segments amounts to 9 years for men and 6 years for women, illustrating the presence of mortality inequality arising from income disparities. The analysis of discrete hazard of first marital relationship shows that higher education goals decrease the likelihood of entering first marriage however, this probability changes also the distribution of age. Which means that as age increases, higher education increases the marriage hazard. Our divorce model predicts that lower educational attainment, protracted period, and presence of children act as protective factors against divorce. Our estimate shows that West Germans have low risk of death and first marriage, and they have a high risk of divorce and remarriage compared to East Germans.

The remainder of this study is structured as follows. Section 2.0 provides a review of the patterns of important life events and family dynamics, the major contributions, and implications of this paper. Section 3.0 presents the concepts and our econometrical framework - event history analysis, and section 4.0 covers data and variables for the modelling of family events including marriage, divorce, mortality, and fertility. Section 5.0 presents the results using parametric and non-parametric techniques. Section 6.0 discusses the main findings and their implications. Section 7.0 concludes the outcomes of this research and presents limitations of this research as well as scope for future research.

2 Literature Review

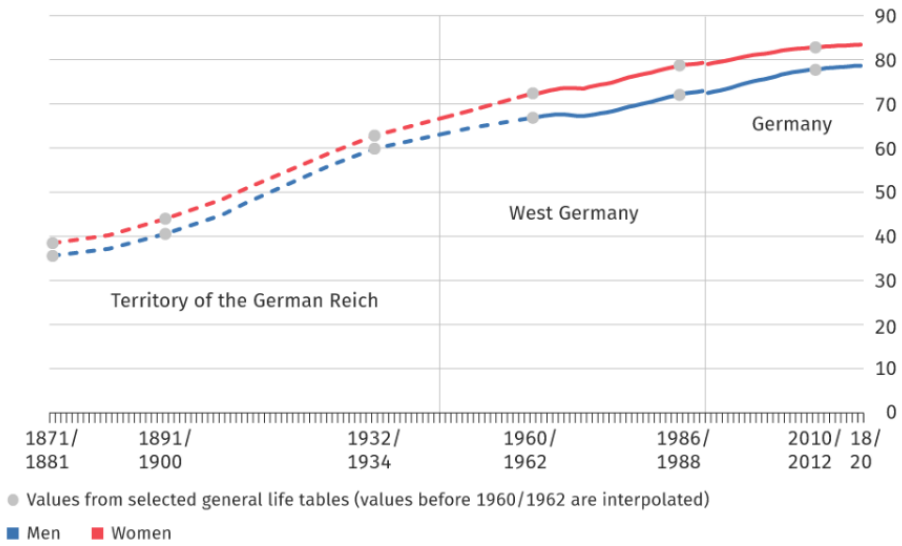
Life course marital transitions raise many interesting questions, such as why such transitions happen, how do they occur, what are their implications, and how to model them. Our study builds upon several streams of literature, which we will briefly introduce. First, we draw from demographic literature to understand the patterns of life course transitions within a population. Second, we explore the implications of life course transitions in various fields. Third, we employ the methodological framework of event history analysis to build hazard models for predicting these life course events.

2.1 Patterns of Life Course Events

Mortality Trends: Germany, like many other countries, has been facing a decline in mortality especially at higher ages. This process leads to an increase in the proportion of the elderly population that is of utmost consideration for social scientists, policymakers, insurance companies and social security institutions. Figure I shows the development of life expectancy at birth for men and women over different periods of German history. Overall, life expectancy is increasing for both genders, but women have higher expectancy compared to men. Average life expectancy at birth has increased for women from 82.3 years in 2005-2007 to 83.4 years in 2019-2021 and for men from 76.9 to 78.5 years for the same periods (Destatis Statistisches Bundesamt, 2022a). According to United Nations projections, the estimated average age of German population by 2100 would be 90.94 compared to 81.88 in 2023 (United Nations, 2023).

In general, mortality inequalities are quite visible and are seriously considered in literature e.g., the life tables of mortality distinguish mortality rates separately for men and women. Yet there are many other factors which contribute to mortality inequalities across different groups. Kibele, Klüsener & Scholz (2015) investigated regional mortality disparities in Germany and found substantial changes in disparities between southern and northern parts of Germany. While the regional mortality inequalities in Germany are shrinking over time especially between East and West Germany, socioeconomic driven mortality inequalities are becoming significant. This social gradient in longevity (Luy, Wegner-Siegmundt, Wiedemann & Spijker, 2015) has recently gained a lot of attention in Germany.

Figure I: Life Expectancy at Birth by Gender



Notes: **Error! Reference source not found.** shows average life expectancy (age) at birth in Germany (1871-2020). Red (blue) line shows average expectancy values for women (men). Broken line shows the interpolated values. The values are based on general life tables. Source: Destatis Statistisches Bunesamt (2022a)

In this paper, we are predicting the discrete hazard of mortality across various subgroups, with a special focus on the relationship between marital status and mortality. Marital status has received significant attention in the field of epidemiology. British physician Farr (1858) was the first to demonstrate lower mortality rates among married individuals (Oswald and Gardner, 2003). Oswald and Gardner explain that there are various factors contributing to this finding like marriage may offer protective benefits by reducing stress and stress-related illnesses, increasing material well-being, and ensuring timely and adequate medical treatment through the presence of a spouse. Gove (1973) argues that the psychological well-being associated with marriage contributes to lower mortality rates, while the isolation experienced by unmarried individuals results in higher mortality rates. Additionally, studies have shown that high social networks (House, Robbins & Metzner, 1982), selection effects, such as healthier individuals being more likely to marry (Hu and Goldman, 1990), and other unobserved characteristics of individuals may also contribute to the lower mortality observed among married couples. The changing mortality patterns and their association with sociodemographic characteristics of individuals are gaining particular interest (see Lampert & Kroll, 2014; Lampert, Hoebel & Kroll, 2019)

Marriage, Cohabitation and Divorce Trends: The family dynamics of European inhabitants are evolving over decades and Germany is one of the countries facing significant demographic and social changes. People live in diverse living arrangements in Germany and new living arrangements are shaping the demographics of Germany. Data from the Destatis Statistisches Bundesamt (2022b) show that married couples are the most common family form in 2022, with 5.87 million (69.6 percent of all families), but with a declining trend. In 2008, 6.1 million families (73 percent of all families) were still married. In comparison, the number of cohabiting couples and lone parents have increased. In 2022, cohabiting couples are at 1.01 million (12 percent of all families) compared to 694,000 (8 percent) in 2008 and lone parents are 1.6 million (19 percent) compared to 1.4 million (19 percent) in 2008. These statistics reflect the changing preferences of German households for family formation given the legal regulations for marital and non-marital couples.

We see a decline in divorces since 2012 and emotional factors have become important for a functioning partnership (Federal Ministry of Family Affairs, Senior Citizens, Women and Youth, 2022). There exist numerous factors that contribute to individuals choosing various forms of family formation (which are not addressed in this paper). Nevertheless, the key idea remains that these evolving preferences ultimately result in shifts within demographic structures over an extended period. The decline in number of marriages, increasing cohabitation and out-of-wedlock births in Germany, require accounting for changing dynamics of living forms and arrangements. It may help to understand better evaluation of distributional implications of various policies which erstwhile are disconnected with social realities. We aim to emphasize that we really need to understand family dynamics across subgroups because these subgroups may have different attitudes e.g., towards traditional marriage norms. After identifying these patterns, researchers and policy makers can better develop and evaluate policies for example, in the above context, promotion of diverse family structures and arrangements and promoting reforms that recognize and protect non-traditional forms of relationship.

2.2 Implications of Life Course Family Transitions

To gain an empirical understanding of demographic transitions, it is important to understand the dynamics of life events using positive analysis. The goal of performing a positive analysis is to describe how life course transitions occur and identify their key determinants in Germany. Life course transitions, in our context, can be broadly categorized into marriage, fertility, divorce, spouse death, and self-death. Few of these events are based on individuals' decisions such as marriage, divorce, birth of a child and few are unpredictable such as self or spouse death. These events have important implications in shaping society and policy considerations because understanding the varying vulnerabilities of different groups to specific events may uncover disparities and inequalities. Therefore, analyzing patterns and predictions of hazards of life events is an important area of research. Before exploring the novel literature in this area, we would like to discuss why it is important to understand hazards of different life events in the context of Germany.

Certain life course transitions lie beyond human control, such as the occurrence of death, which can profoundly alter a couple's dynamic, transforming them into a widow or widower. The precise measurement of an individual's probability of death is inherently unpredictable, However, certain personal characteristics can influence the likelihood of experiencing such an event. For instance, individuals who smoke have a higher probability of death compared to non-smokers. The increasing life expectancies because of advancements in technology, medicines etc. lead to changes in the age distribution of the population. Average life expectancy for men(women) at age 65 is 82.83(86.09) years in 2021 compared to 81.93(85.31) years in 2007 (Destatis Statistisches Bundesamt, 2022a). This can have important implications for social security systems, insurance companies and even at individual level.

These life course transitions, which are very personal events, can lead to significant demographic and structural changes. The most visible impact of a demographic shock is the change in the size and composition of households. For example, a birth or death in family can alter the size of the household and its composition, causing a change in age distribution or number of dependents in household. Furthermore, such transitions can affect the social norms and preferences within households. For instance, the birth of a child may shift the spending priorities of parents towards education and child-related activities.

Life course transitions are also associated with financial implications, and many theoretical and empirical studies especially relevant to financial decision making of households, account for these transitions. Research shows that life events can affect various dimensions of an individual's life, such as financial position, lifestyle, social/network structure, preferences/choices, and emotional/behavioral characteristics. Many studies have particularly focused on evaluating the impact of specific marital statuses on factors such as earnings, wealth, living standards, and labor supply. Marital status is an important economic and demographic variable because the propensity for marriage, its timing, and the duration of marriage have impacts on consumption, population growth rate, wage rates, fertility, migration, and mortality (Keeley, 1979). In a recent paper, Bonnet, Garbinti & Solaz (2021) present findings on the gender effects of divorce on living standards and labor supply using an administrative dataset on French households and employing a difference-in-difference approach. They found a decline in women's living standards relative to men after divorce and re-entry of divorced women into the labor market.

The importance of demographic transitions in understanding the consumption and savings decisions of households has been widely discussed by various authors using theoretical models. Attanasio, Banks and Maghir (1999) assumed a nonstationary income process and included the demographic effects in a lifecycle setting. They showed that demographic factors affect the dynamic optimization problem due to their direct influence on the marginal utility of consumption. Therefore, a same income shock can have varying impacts on households, depending on their size and compositions. Hong & Rios-Rull (2007) analyzed the role of marital transitions on the demand for life insurance. Scholz, Seshadri & Khitatrakun (2006) highlighted the impact of these transitions on wealth accumulation, using a life cycle model.

Family transitions can influence the optimal savings decisions and asset allocations portfolio choices of households. For example, marital transitions can cause changes in present and expected future earnings, sources of income, spending needs and wealth accumulations. Therefore, decisions such as how much to spend, save and where to invest are not only influenced by the variations in labor income but also by factors such as family status. Love (2010) investigated the importance of interactions between marital status and children on households' choices relevant to savings and portfolio decisions. His model was the first to discuss the effect of marital status and children on savings and portfolio allocation choices given an exogenous labor supply. He showed that households with children had less accumulated wealth and higher demand for life insurance. Hubener, Maurer & Mitchell (2016) contributed to the same direction and demonstrated the influence of uncertain family transitions on optimal household decisions regarding work, retirement, savings, portfolio allocation and life insurance in the U.S. In a recent paper by Altonji, Hynsjo & Vidangos (2021) investigated factors affecting the family income of men and women throughout their adult lives and found a larger impact of marital status on family income of women relative to men, who are more influenced by labor market shocks. In addition, there are many other papers in different streams of literature that investigate the role of family transitions in various contexts.

2.3 Modelling Life Course Family Transitions in Lifecycle Literature

In this section, we would explain how our statistical models can be used to calibrate family transitions in lifecycle consumption and portfolio literature - another contribution of our paper to household finance literature. In our previous section, we highlighted that marital transitions affect individuals in multiple ways by influencing their financial wealth, future expectations, labor supply, consumption expenditures, savings, and portfolio choices. Whether these transitions are decision-based or random moves, they appear to be another source of risk over the life cycle. However, these transitions are often overlooked in life cycle models due to the increasing complexity and curse of dimensionality associated with incorporating them. Nevertheless, to understand the consumption and portfolio choices of households belonging to different categories of marital status, it is important to comprehend the underlying process of marital transitions. This leads to another question of how to calibrate these transitions in theoretical lifecycle models.

In most of the previous studies the marital transition rate in the lifecycle model is calibrated based on the existing published research especially the econometric microsimulations models like Modeling Income in Near Term (MINT). This model from Urban Institute, is based on micro level datafile of actual and projected population of U.S. born between 1926 - 2070 (Smith et al., 2010). This model is used to evaluate the distributional consequences of social reforms by Social Security Administration (Panis & Lillard, 1999). In addition to the estimations of parameters in these models, projecting marital transitions is also an important step in dynamic microsimulations models because marriage histories can affect the beneficiary status, amount of government program benefits, labor supply, savings, and other behaviors of individuals (O'Harra & Sabelhaus, 2002). Therefore, the estimations based on dynamic microsimulations models are useful to project transitions and then address the questions relevant to distributional consequences of social security reform proposals and other policy issues. They provide the flexibility to make analysis using different datasets especially as the trends in demographic transitions change over time, and so do the policy implications.

The estimations coming from published dynamic microsimulations model are used by researchers to develop family models or "life history model" (Willekens, 2017). Love (2010) and Hubener, Maurer & Mitchell (2016) both have used estimations from MINT model to develop family process in the lifecycle models. Both papers assume the family process following Markovian property and calibrate the family process using MINT estimates. Most of the papers in this direction use first order Markov Chains Model to develop family transitions or Hidden Markov Models (Han, Liefbroer, & Elzinga, 2020). In the demographic literature, there is research that develops the effects of education, age, and other covariates on the probabilities of different real-world states like mortality and marriage (Case & Deaton, 2017), (Lampert & Kroll, 2014), (Kroh, Neiss, Kroll, & Lampert, 2012) and many more.

The estimations of life processes require special statistical models, survival analysis, because such processes are time and age dependent. In addition, there are many other reasons that make these methods suitable to predict transition rates/probabilities of life events. First, these methods perform time-to-event analysis, which means that in addition to the causal-effect analysis, the timing of the event is also of special interest like the age at death. Second, survival methods are special to deal with censoring issues unlike standard logistic and linear regression models. Censoring reflects the participants that do not encounter the event and by using standard models, the true time to event can be underestimated.

Therefore, our paper contributes to existing literature in various ways. First, we quantify the hazards of marital transitions for individuals with different demographic characteristics. Existing papers in this area are very fragmented, for example, linking marriage dissolution to mortality, marital status to health, income, or education to mortality. It is mainly because these topics belong to different areas of study like demography, sociology etc. We emphasize that in a lifecycle model, we should consider various potential marital states and estimate the hazards of moving from one state to another especially the models which are based on Germany.

Second, the findings of this paper are also important to make important policy evaluations for example, to understand the distribution of future marital statuses and survival probabilities of individuals to evaluate the distributional impact of the policy reforms. For example, in Germany, workers are encouraged to invest in tax-qualified Individual Retirement Accounts (IRAs) and the EET tax regime is implied on these accounts to incentivize savings. If the government decides to change the tax regime from EET to TEE, it will have different impact on individuals belonging to different marital statuses. In other words, the reactions of households with different marital statuses will be different. Therefore, understanding the distributional impact of such a tax reform in future, is only possible if the demographic distribution process of population is known.

This study does not make any claim regarding establishing the causality analysis and our primary objective of this research is to understand the family dynamics and develop statistical models that estimate family transition risk and then using these estimated parameters, develop a population model that can simulate future demographic projections of a heterogeneous population. Therefore, analysis here mostly covers the patterns of demographic transitions, the methodological frameworks model these transitions, estimation techniques and important implications.

3 Methodological Framework

3.1 Event History Analysis

Life is a series of events, and we are interested in explaining and predicting these events using event history analysis (survival analysis/duration models). We are using discrete time event history analysis to analyze our data. Discrete time event history models are useful when the process is continuous but measured observations are discrete and these processes[1] can be approximated using standard statistical tools like logistic regression. We have converted time-to-event data into a series of binary indicators to see if the event occurs or not. Event histories are ideal to investigate cause of events however, issues of censoring and time-varying variables create difficulties for standard statistical methods and lead to serious bias (Allison, 1982).

The timing of events is not always known to everyone, but for some individuals, we have precise information on actual timing of events. For others, the events can be right-censored, left-censored, or interval-censored. This variability in event knowledge calls for modeling techniques within survival analysis framework that can handle both complete and incomplete event information. Survival analysis requires data in an event-history data structure; therefore, understanding the theoretical setup is necessary to interpret the results. We assume that censoring is independent of the occurrence of an event because in our analysis, the individuals are not selectively withdrawn from the sample because of their likelihood to experience the event.

3.2 Analytical Framework and Technical Details

In survival models, basic analytical framework is based on time (e.g., age or survey date) and states of the individuals (Blossfeld, Rohwer & Schneider, 2019). All states of individuals belong to a state space and the time an individual spends in a particular state is known as spell/episode/waiting time/duration. For example, the time when an individual is born, he has a positive probability to die anytime therefore, *birth* is the point when the exposure time starts, and *death* is the event

[1] James S.Coleman (1981) defines event as a substantial process, in which a collection of units e.g. individuals, move among finite number of states. The transitions to these states may happen at any point in time a there are time constant/varying factors affecting this event (Blossfeld, Rohwer & Schneider, 2019).

when the time stops. The period between birth and death is denoted as survival time. These models can be further categorized into single episode or multiple episode models depending upon the number of origin and destination states. The model of mortality is a multi-state-single-episode model, which means individuals belonging to multiple states when enter the state of death do not revert. Therefore, the origin states are many, but the destination state is one. The model of marriage/divorce are the multi-state-multi-episode models (or repeated events) where the events may repeat. This interval can be further divided into many small intervals. The probability distribution of failure time can be specified by three ways including survivor function, probability density function and hazard function (Kalbfleisch & Prentice, 2002).

The detailed theoretical setup defined here is from Kalbfleisch & Prentice (2002) and Heeringa, West & Berguld (2010). Let us assume that failure time is denoted by T, which is a nonnegative random variable and takes the values $a_1 < a_2 < a_3 < \cdots$ and time of observation is denoted as t. Censoring time C is the time if the individual does not encounter the event. Survival time, T and censoring time C are assumed to be independent to each other after controlling for covariates. It is also one of the main assumptions of time-to-event data termed as non-informative censoring which means that individuals which do not encounter the event (censored) have the same probability of experiencing a subsequent event.

The function $f(a_i) = P(T = a_i)$ denotes the probability function of an event and is defined as

$$f(a_i) = \lambda_j \prod_{j=1}^{i-1}(1 - \lambda_j) \tag{1}$$

where, λ_j represents probability of an event e.g., the probability of dying at age 30, 65 or 90. The corresponding survivor function gives the probability that an event has not occurred prior to the end of observation period e.g., the probability that an individual does not die before the age 65 e.g., and can be represented as follows,

$$F(t) = P(T > t) \tag{2}$$

The survivor function is one minus cumulative distribution function that measures the probability that event occurs before the end of observation period, $P(T \leq t)$. e.g., the probability that an individual dies before the age 65.

$$F(t) = \prod_{j|a_j \leq t} (1 - \lambda_j) \tag{3}$$

Another important variable is the time origin, from which the spell/episode/time to event is measured. For example, in case of mortality, it is the time of birth of an individual and in medical sciences it can be the entry into the study. In survival analysis, different timescales can be used and after considering various strategies to choose appropriate time scale, suggested by Canchola, Stewart & Bernstein (2003), we have selected "age" of the participant as a timescale unit for non-parametric estimate. summarizes five ways to choose a timescale.

The aim now is to fit survival model where our covariates of interest change value over time. In general, a typical survival model can be mathematically expressed as a hazard function. Hazard function of failure time is the conditional probability that event will occur at time, t given it has not happened in previous period $t - 1$.

$$h_i(t) = h_0(t)e^{[\emptyset^T v_i + \propto y_i(t)]} \tag{4}$$

where, h_i represents the hazard (rate of an event). For example, in case of hazard of marriage, it implies hazard of individual i who marries ($w\ for\ (re)$wedding). The term, $h_0(t)$ represents baseline proportional hazard, $\emptyset^T v_i$ shows the effect of time variant covariates and $y_i(t)$ is the observed value of a covariate for the i^{th} participant. The most commonly used model to analyze such processes is the *Cox proportional hazard model* by Cox (1972), which is a semi-parametric model and uses a partial likelihood method and estimates a likelihood function at each event (Fisher & Lin, 1999; Canchola, et al. 2003).

In this paper we model all the hazard processes, the theoretical modeling framework is *Discrete Time Event-History Analysis*. All the processes like mortality are continuous processes, but all the datasets are in year format, which makes the discrete time model suitable for this analysis. Discrete-time hazard models can approximate the continuous-time hazard models using the standard statistical tools. The discrete-time hazard of an event can be defined as,

$$h(t) = \frac{f(t)}{S(t)} \quad (5)$$

where hazard of an event, $h(t)$ is the ratio of probability of event happen at time t, $f(t)$ divided by the survival probability $S(t)$ up to time t. The hazard conditional on covariates can be defined as follows.

$$h(t|x) = P(T = t|T \geq t; x) \quad (6)$$

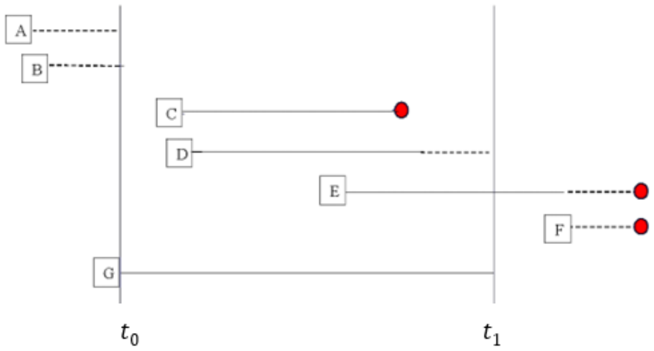
To analyze the model, the data is structured in person-year format, where each cross-section presents record per person per time unit at risk and person data ends if the person comes across the event or drops out of the study. Prior to estimation, defining the functional form of baseline hazard is important, which is done by creating the time-varying covariates that are the functions of survival time for each individual. For example, in the mortality process, the baseline hazard function is defined by assuming that hazard of death is piecewise linear in age.

3.3 Type of Censoring Issues in Data

In panel data the composition of core sample diminishes over time because of death, birth, migration, and other reasons which lead to panel attritions between waves. The SOEP data is right and left censored data. The right censoring occurs because participants drop out of the study without encountering the event (e.g., first marriage) and study is completed but all participants do not encounter the event. It is also left censored because all participants do not enter the study at the same time. Right censored data has the potential to introduce bias into the estimation of survival probabilities and survival analysis handles it very well (Kalbfleisch & Prentice, 2002). The frequent use of survival analysis in literature is based on the reason that survival analysis can handle the right censoring issue better as compared to standard statistical analysis.

In the event-history data the observations are often censored, and they may create a bias. Figure II shows the type of censoring issues in SOEP dataset. The x-axis shows the observation time window of the survey from t_0 to t_1. The solid line shows the participation period in the study and dotted line shows the period unknown. The red dots show the event has occurred e.g., the event of death. Suppose A, B, C, D, E, F and G are the participants of the survey. Participant A is fully censored on the left and not a part of the SOEP survey. B is partially censored on the left like the history of this individual before being a part of SOEP is unknown, but this person is right censored because he drops out of study before the end of survey. C does not face right censoring, but it is left censored because he enters the study later (e.g., refreshment samples in SOEP) and incurs the event of death. Participant D is a left and right censored as he enters late and leaves the study before incurring the event. E is left censored and dies at some point in the study, but its time of event is unknown. F is not a part of the study. Subject G is a part of the study, and it is still active in the yearly survey.

Figure II: Types of Censoring in SOEP Data



Notes: **Error! Reference source not found.** shows the types of censoring issues in SOEP datasets. X-axis shows the time period (t), letters A-G show the participants and potential participants of the study. The red dot shows the event has happened and lines without dots show no event. Solid line shows periods in which information of the participant is available while broken line shows the period in which information about participant is unknown. Diagram shows that the data is right and left censored which means that individuals have different entry and exit times in the survey, which leads to missing historical and future information about the participants.

3.4 Non-Parametric Analysis

For exploratory data analysis, we discuss nonparametric estimation methods to describe the characteristics of each event, without making any assumption about the distribution of the process. There are various methods to get survivor functions including Life Table Estimator, Nelson-Aalen Estimator and Kaplan-Meier Estimator (for details see, Cleves, 2008). The *Kaplan-Meier estimator* (or product limit) is mostly used for univariate analysis for categorical variables and provides a simple way to describe empirical distribution of survival data with respect to a specific factor, for example, education. Although the resulting survivor curve is not covariate-adjusted but it does highlight the importance of that variable in investigating the relationship with the event. The estimator is based on average risk set size per time interval to estimate a survival probability. For example, to calculate the survival probabilities for three time periods, the survivor function is estimated using the formulas below. The number of failed values show the count of individuals that have come across an event like (e.g., death).

$$\hat{F}(t_1) = \left[\frac{no.of\ sample\ values\ in\ t_1 - no.of\ failed\ values\ t_1}{no.of\ sample\ values\ in\ t_1}\right] \tag{7}$$

$$\hat{F}(t_2) = \hat{F}(t_1) * \left[\frac{no.of\ sample\ values\ in\ t_2 - no.of\ failed\ values\ t_2}{no.of\ sample\ values\ in\ t_2}\right] \tag{8}$$

$$\hat{F}(t_3) = \hat{F}(t_1) * \hat{F}(t_2) * \left[\frac{no.of\ sample\ values\ in\ t_3 - no.of\ failed\ values\ t_3}{no.of\ sample\ values\ in\ t_3}\right] \tag{9}$$

4 Data and Variables

4.1 Datasets

To estimate the parameters that can be used to calibrate the family model for Germany, data is accessed from Socio-Economic Panel (SOEP) by German Institute for Economic Research, (DIW) for birth cohorts 1882 to 2019 for the period 1984-2019. This is an annual household panel survey for Germany, in which households have been interviewed at multiple points in time since 1984. The advantage of using this dataset is that the participants have been surveyed repeatedly over time and information on marital histories, fertility histories and demographics of individuals is available.

To estimate the processes of marriage, divorce, and mortality, SOEP dataset provides a large number of observations for each subgroup. SOEP follow-up survey identifies the individuals who died and dropped out of the study. Therefore, through this data the relative risk of mortality (hazard ratios) can be estimated using survival analysis or time-to-event analysis. The marriage histories of individuals are also followed since they have joined the survey therefore, the models of marriage and divorce can also be estimated based on this dataset. We have merged various SOEP datasets to develop all our statistical models. For example, for mortality model and data containing information regarding the participants' identification-numbers and demographics including age, gender, and marital status, number of households, number of children, education and region (East or West Germany) is accessed from SOEP-PEQUIV data file. This data file is merged with SOEP-PPATHL data file to identify the death year of the participants, who have died. The data structure is in long-format where each observation is a person-year observation. Both PEQUIV and PPATHL datasets are individual tracking files and belong to Generated Data and Tracking Data categories respectively. For further details check the documentation for respective dataset (SOEP Companion, 2021).[2]

4.2 Sample Size

The models of mortality, marriage and divorce have different number of subjects and observations because of defined age brackets and variables of interest. Appendix (1A-5A) summarizes the distributional characteristics of each sample for each model of interest. The table shows that for modeling each event, we have sufficient number of observations and relevant events for each subgroup. Our mortality data is a panel dataset with a total of 101,912 subjects and total 6,901 events of deaths. Marriage data is based on retrospective histories of all adult SOEP-participants. For estimating hazard of first marriage, the dataset includes all adult singles (never married) and they get censored if the event of marriage does not happen. The sample includes 60,976 individuals with 41,063 first marriage events. For remarriage model, we have 18,621 subjects and a total of 7,958 events of remarriages are identified. For further details see Appendix (1A-5A).

4.3 Variables

The variable *age* is one of the most important predictors of life events because of biological reasons. In the context of life events, age also plays an important role in defining the behavior and attitude of individuals that may vary depending on the position of the individual in the lifecycle like a young adult, a middle-aged parent, or a retiree would have different opportunities, choices, preferences, and priorities. To control the cohort effect, cohorts are grouped based on the generations which are commonly defined in demographic and political studies. These age cohorts are further adjusted depending upon the model under consideration and the availability

[2] For respective documentations: $PEQUIV (Grabka, 2020), $PPATHL (SOEP Group 2019; DIW Berlin, 2022). Sometimes the documentation of exact version is not available therefore, previous versions can be referred to.

of number of observations in different age cohorts. SOEP dataset covers a vast diversity of individuals from various birth cohorts “Birth prior to 1915”, “First World War babies (1915-1935)”, “Second World War babies (1936-1945)”, “Baby Boomers (1946-1963)”, “Generation X (1964-1980)”, and “Generation Y (post 1980)”. We also control for time effects, for the models, which allow us to account for time changes.

The variable, *marital status* shows the family status of the including married, never married, widowed and divorced, and separated. Married couples include only the couples which are legally married and do not include the cohabitation couples.

Variable, *education* categorizes the individuals into three categories: *less than high school, high school and more than high school.* According to educational system in Germany as mentioned in the documentations available at SOEP Companion (2021), *less than High School* means intermediate or lower secondary school (Realschule and Hauptschule), *High School* means upper secondary school, certificate of aptitude for specialized short-course higher education, apprenticeship and specialized vocational school (Abitur, Fachhochschulreife, Lehre, Berufsfachschule) and *more than High School* covers school of health care, specialized college of higher education, post-secondary technical, college, technical university usually requiring practical training as part of the studies and civil service training (Schule des Gesundheitswesens, Fachhochschule and Universität).

Other important duration variables include the *duration of being unmarried* in case of divorce and *duration of marriage* in case of marriage. The demographic literature on models of marriage and divorce shows that duration effects are important in estimating the probabilities of divorce or (re)marriage. Therefore, the main idea of using the retrospective dataset is to have information on the durations of various such life events. The *outcome variable* is dichotomous which means that the event may have occurred or not. The event variables include Mortality (dead or alive), Marriage (married or unmarried), Birth and Divorce (legal marriage dissolution or no dissolution).

Moreover, the model assumes that the maximum number of children is 3, which means that all couples with three or more children are assumed to have three children. Information about kind of health insurance is not being controlled as models assumes a fair and equal health care system in Germany. Since all individuals are covered by the mandatory health care services in Germany therefore, all individuals have access to the same medical care without any additional financial burden.

4.4 Data Limitations

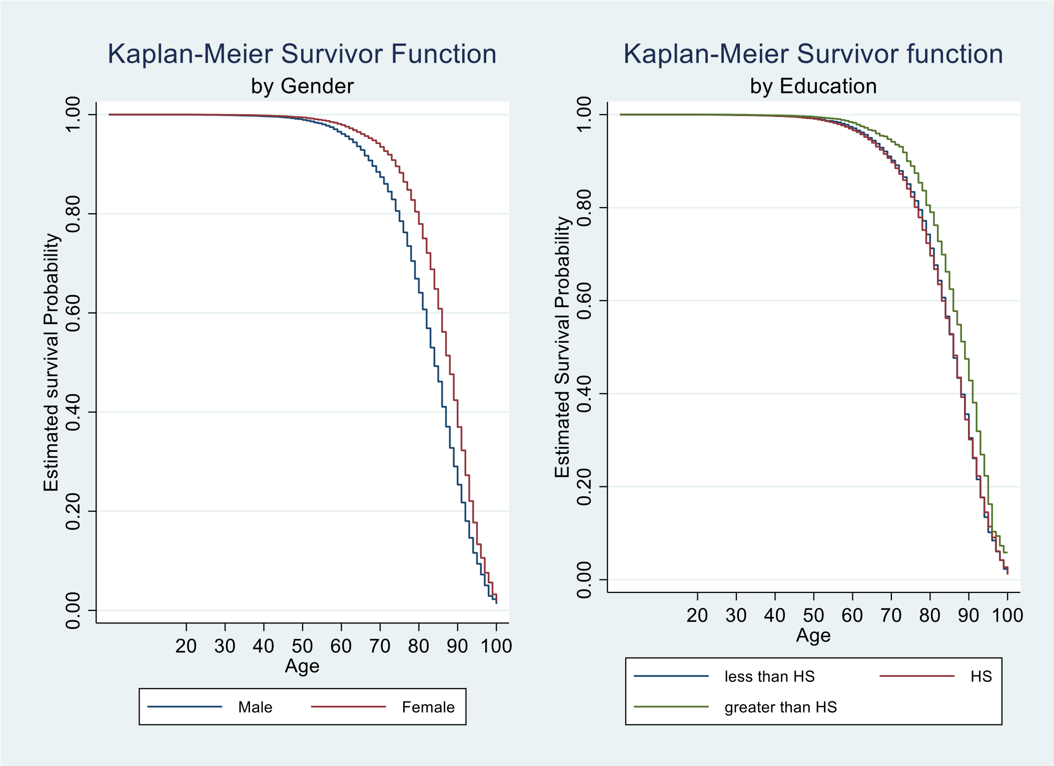
Information about the state of participants is available at the time of survey, which means that the course of events between the survey points remains unknown. Since many events can occur during survey intervals, we assume that only events happen within a year. For many respondents, there are issues of censoring. In the dataset of mortality model, individuals drop out of the study due to many reasons, leading to right censoring. Follow-up studies from SOEP attempt to investigate the reason of drop out; however, we don’t have reason of right-censoring for all the respondents therefore, we assume that an individual is alive as long as he/she is a part of survey. And the nature of event history analysis better handles such censoring and panel attrition problems. Finally, we are interested in the accurate approximation of hazard of an event and not after the disentangling the age, period and cohort effects therefore, we ignore specific problems confounding these factors.

5 Results

5.1 Model of Mortality

Figure III shows the Kaplan-Meier survivor functions by gender and educational groups. The curves by gender show differences in survival probabilities between men and women, in particular at higher age, men have higher hazard of mortality compared to women. We also see that proportional hazard of mortality varies across the age. The survivor curves by educational groups show the survivor probabilities for individuals with less than, equal to and more than a high school degree. The difference in survivor probabilities looks more significant for individuals with more than HS degree compared to other groups.

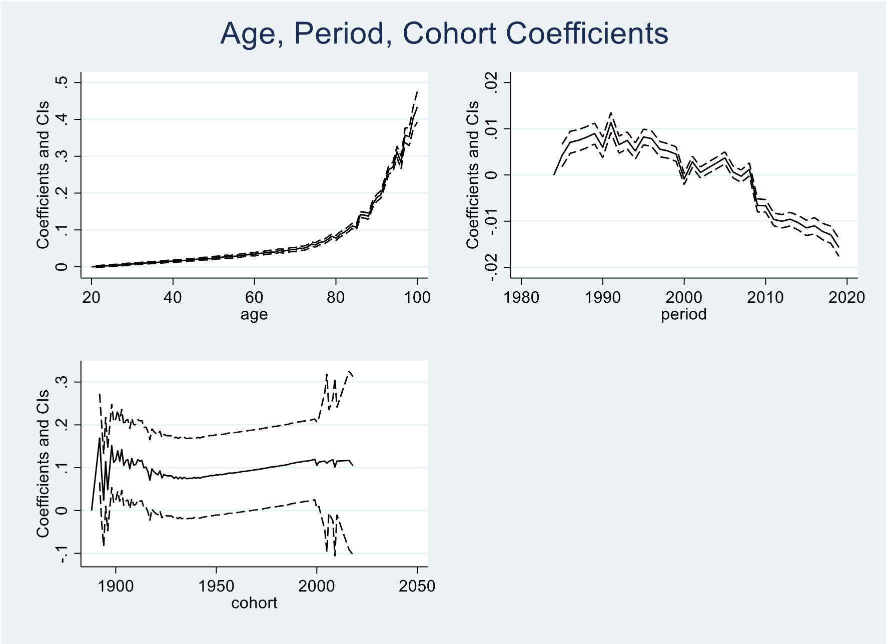
Figure III: Kaplan-Meier Survivor Functions by Gender and Education



Notes: Figure III shows Kaplan Meier Estimates for survivor functions by gender and education based on SOEP-v36 (1984-2019) for age 20 to 100. It shows higher but non-proportional hazard of mortality of men and less educated groups compared to women and higher educated group respectively. Source: Authors' calculation.

Next, we model mortality using discrete time logit regression framework and results are summarized in **Table *1***. We find that the expected hazard of mortality increases with age, advocating the law of nature. At higher ages, individuals may have higher risk of chronic illness, deteriorating immunity and declining physical strength. The period effects are captured by variable "time" that reflect the influence of social, economic, and environmental factors during the survey period. The results show that discrete hazard of mortality is decreasing between 1984 and 2019, which can be because of many factors like advancements in technology and medicine etc. The panel nature of our dataset allows us to capture the long-term effects of shared experiences and characteristics of groups of individuals from same birth cohorts. To disentangle the age, period, and cohort effects, we have run an analysis following computing program by O'Dea (2012) and method by Deaton & Paxson (1994). This method identifies the effects of coefficients by imposing restrictions that period effect is orthogonal to a trend that sums to zero. Figure IV shows that the effects of age, period and cohort are consistent with our main model.

Figure IV: Disentangling Age, Period and Cohort Effects



Notes: Figure IV distinguishes age, period, and cohort effects on mortality by estimating the coefficients of the linear age-period-cohort (APC) model, following user-written command by O'Dea (2012) and the parameters are identified using restriction proposed by Deaton & Paxson (1994). In APC model, dependent variable (discrete hazard of mortality) is a linear additive function of dummies of age, period and cohort. Figure (top-left) plots the coefficients and confidence intervals (CIs) of age and shows that coefficients are getting stronger with increasing age, reflecting increase in risk of mortality at higher ages. Figure (top-right) plots coefficients and CIs of each observed period (1984-2019) and a declining trend in coefficients values reflect a decrease in mortality risk over the time. Figure (bottom-left) plots coefficients and CIs of each cohort and we do not see any significant cohort effect on mortality.

Our regression results in Table-1 shows that education and income both significantly reduce the hazard of mortality. For both male and female participants in the survey, individuals with less than higher school education, have higher mortality hazard compared to individuals with higher school education. Participants with more than high school education have a lower hazard of mortality compared to those with higher level education. Better educated individuals may have better access to knowledge and health awareness, leading them to adapt to healthier health choices. Since higher income level leads to high socioeconomic status and better affordability of lifestyle, food, and health facilities therefore, it is also an important determinant of risk of mortality.

Marital status of both male and female is significantly related to mortality hazard. It is interesting that compared to married couples, people with all other marital statuses including single, divorced, widows and separated, have higher hazard of mortality. As literature shows, low mortality among married couples can be explained through the built-in-social support system, which provides companionship, emotional support, and assistance to each other. Compared to individuals without partners, couples enjoy better economic stability and protective effects like avoiding engaging in risky activities. Finally, for the given sample, individuals from West Germany have less hazard of mortality compared to East Germany. The male-female survival paradox across East and West Germany can be explained using historical, socioeconomic, etc. and have been widely discussed in the literature (Kühn, Dudel, Vogt, & Oksuzyan, 2019).

Following regression analysis, we have performed Wald tests to check if parameters of the fitted model of mortality are different from zero. We reject the null hypothesis both for education and marital status at 1 % significance level. Since our fitted logit model does not provide information about the size of effect, we have performed post-estimation analysis to analyze average marginal effects. Table-6A (Appendix) summarizes the average marginal effects by gender. The output indicates that the probability of mortality decreases by 0.0032 (0.002) percentage points for a female (male) with >HS education and increases probability by 0.0027 (0.001) percentage points for a female (male) with <HS education.

Table 1: Parameter Estimates of Mortality Hazard by Gender using Discrete Time Logit Model

Variables	Male	Female
Age Splines		
Age spline (20-65)	0.0999***	0.0884***
	(0.00501)	(0.00589)
Age spline (66-80)	0.0902***	0.114***
	(0.00596)	(0.00690)
Age spline (81-100)	0.0935***	0.141***
	(0.00821)	(0.00692)
time	-0.0273***	-0.0315***
	(0.00411)	(0.00433)
Cohort		
Cohort 1915-1935 (First World War)	0.0189	0.114
	(0.0979)	(0.0910)
Cohort 1936-1945 (Second World War)	-0.272	0.00209
	(0.141)	(0.148)
Cohort 1946-1963 (Baby Boomers)	-0.0909	0.214
	(0.186)	(0.197)
Cohort 1964-1980 (Gen X)	-0.238	0.0145
	(0.254)	(0.283)
Cohort > 1980 (Gen Y and onwards)	0.102	-0.149
	(0.371)	(0.490)
Education		
Education <HS	0.216***	0.109*
	(0.0549)	(0.0453)
Education >HS	-0.328***	-0.229**
	(0.0524)	(0.0747)
Marital Status		
Single	0.724***	0.487***
	(0.0763)	(0.0841)
Widow/er	0.271***	0.159**
	(0.0587)	(0.0560)
Divorced	0.652***	0.434***
	(0.0745)	(0.0799)
Separated	0.375**	0.184
	(0.123)	(0.179)
Region: East Germany	0.106*	0.124*
	(0.0451)	(0.0492)
Income	-0.642***	-0.623***
	(0.0620)	(0.0612)
Constant	-8.680***	-8.956***
	(0.354)	(0.407)
Observations	289,271	313,920
Log Pseudo Likelihood	-14838.489	-12463.687

Robust standard errors in parentheses brackets: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Notes: Table 1 shows the estimated model of mortality for male, and female based on SOEP-v36 (1984-2019) for age 20 to 100.The estimations are based on discrete time logit model using event-oriented observation design. Standard errors (in parentheses) are clustered at individual levels**.** Baseline groups: Education (HS degree); Marital Status (Married); Cohort (<1915); Region (West Germany)

5.2 Model of First Marriage and Remarriage

We define marriage as a formal, legal contractual union and does not encompass couples in living relationships (cohabitation). When estimating the model for marriage, we estimate the probabilities of first-time marriage and remarriage separately. Existing literature suggests that factors such as the presence of children and the duration of being unmarried after divorce, can influence the likelihood of remarriage. We drop the sample of same-sex partnerships and marriages that is only 0.10 percent and have high frequency of partner switching in the dataset. The high number of transitions within this subgroup could be attributed to various reasons, such as physical or psychological factors, which are beyond the scope of this paper.

Table-2 provides parameter estimates for male (Model 1 and Model 2) and female (Model 3 and Model 4) subgroups. Results show that getting older reduces the probability of marriage, however it is significantly higher at early ages. The hazard is also lower for the newer cohorts relative to the base group, possibly because marriage is getting less common in Germany. Increasing cohabitation in Germany, especially among the young cohorts, is substituting marriage. In Model 1, we do not find any effect of education on <HS group, however, after controlling for interaction between age and education, in Model 2, we find that male hazard of first marriage is highest for individuals with <HS and lowest for >HS relative to =HS. One explanation is that highly educated men, extended educational enrollment delays hazard of first marriage. For female, the probability of marriage decreases with education, and women who have less than higher school education are more likely to get married compared to other groups.

We add the interaction between age and education in Model 2 and Model 4 and we find that with increasing age, the likelihood of marriage increases for more educated people compared to less educated. One reason is that most of the less educated people already get married and drop out of the potential sample for first marriage at higher ages. Intuitively, it may be also harder for lower educated people to find a mate at relatively higher ages because of not-so-good occupation or financial situation. The pseudolikelihood of models has improved after adding the interaction term, which implies an improvement in model prediction. Furthermore, our results show that West Germans have less hazard of marriage compared to East Germans.

Table 2: Parameter Estimates of Hazard of First Marriage by Gender using Discrete Time Logit Model

Variables	Male		Female	
Age Splines	Model 1	Model 2	Model 3	Model 4
Age spline (18-25)	0.551***	0.549***	0.340***	0.342***
	(0.00605)	(0.00606)	(0.00344)	(0.00357)
Age spline (25-35)	0.0142***	0.00427	-0.0499***	-0.0595***
	(0.00261)	(0.00282)	(0.00290)	(0.00308)
Age spline (35-40)	-0.161***	-0.170***	-0.216***	-0.222***
	(0.00875)	(0.00883)	(0.0111)	(0.0112)
Age spline (40plus)	-0.00992**	-0.0243***	-0.00617	-0.0167***
	(0.00432)	(0.00499)	(0.00425)	(0.00510)
Education				
Less than High School	-0.0152	0.453***	0.294***	0.713***
	(0.0266)	(0.130)	(0.0226)	(0.0974)
More than High School	-0.262***	-1.593***	-0.525***	-1.946***
	(0.0166)	(0.0891)	(0.0169)	(0.0840)
Interaction: Age*<HS		-0.0160***		-0.0158***
		(0.00453)		(0.00381)
Interaction: Age*>HS		0.0443***		0.0508***
		(0.00308)		(0.00311)
Cohort				
Cohort 1915-1935 (First World War)	0.475***	0.446***	0.363***	0.358***
	(0.0569)	(0.0529)	(0.0526)	(0.0501)
Cohort 1936-1945 (Second World War)	0.511***	0.481***	0.797***	0.788***
	(0.0574)	(0.0536)	(0.0537)	(0.0513)
Cohort 1946-1963 (Baby Boomers)	0.181***	0.153***	0.638***	0.635***
	(0.0547)	(0.0509)	(0.0511)	(0.0488)
Cohort 1964-1980 (Gen X)	-0.322***	-0.348***	0.0229	0.0255
	(0.0545)	(0.0507)	(0.0502)	(0.0480)
Cohort > 1980 (Gen Y and onwards)	-0.897***	-0.937***	-0.391***	-0.387***
	(0.0650)	(0.0619)	(0.0559)	(0.0539)
Region: West Germany	-0.136***	-0.136***	-0.127***	-0.134***
	(0.0212)	(0.0213)	(0.0205)	(0.0203)
Constant	-15.89***	-15.78***	-10.47***	-10.48***
	(0.155)	(0.155)	(0.0935)	(0.0950)
Observations	423,867	423,867	373,880	373,880
Number of Events	20,501	20,501	20,562	20,562
Log Pseudo Likelihood	-75316.245	-75096.605	-83020.476	-82722.857

Robust Standard errors in parentheses brackets: * p<0.05, ** p<0.01, *** p<0.001
Notes: Table 2 shows the estimated model of first marriage for male, and female based on retrospective SOEP BIOMARSY for age 18 to 100. Standard errors are clustered at individual levels. Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany)

Following regression analysis, we have performed Wald test to check if parameters of the fitted model of education, cohort and region are different from zero. We reject the null hypothesis for all variables at 1 % significance level. Since our fitted logit model does not provide information on magnitude, we have performed post-estimation analysis to analyze average marginal effects. Table-7A (Appendix) summarizes the average marginal effects by gender. The output indicates that for female, probability of first marriage decreases by 0.022 percentage points >HS education and increases probability by 0.005 percentage points with <HS education.

For male, risk of marriage is lower for both <HS (0.00315 percentage points) and >HS (0.011 percentage points) relative to =HS education group.

Table-3 presents the estimations for discrete hazard estimates of high order marriages. Model 1 and Model 3 are estimated models for men and women respective without controlling for education and age interaction. However, Model 2 and Model 4 control interaction terms. We emphasize that remarriages are more common after the mid ages therefore, we have constructed age splines instead of polynomials. In all models, for both men and women, the likelihood of remarriages decreases significantly with age.

Education significantly affects the probability of marriage and individuals with <HS education has higher probability of marriage. However, when we control for age and education interaction, to determine if this relationship changes with age, we do not see any significant effects of education on remarriage for men. Our log pseudo likelihood estimates suggest that models without interactions are better predictive models. We also find that interaction is significant for group with >HS education female. This implies that women with more than HS education, have higher probability to get remarriage at any age level compared to the group with =HS education. Being from West Germany increases the probability of remarriage compared to being from the East region both for male and female.

Controlling the number of children, we found that presence of children in the households increases the hazard of remarriage. It might be possible that singles parents, either divorced or widowed, face increased financial obligations and time constraints and are more likely to consider remarrying, with the presence of children. We are also interested to investigate in the association between "how long an individual remains unmarried" after a divorce or death of spouse and likelihood of remarriage. We find a positive association between the increased duration of being unmarried and the likelihood of remarriage. For men, the likelihood of remarriage is higher right after the event while for female it is insignificant. After 7 years, there is still positive probability of remarrying however, average marginal effects (Table-8A - Appendix) shows that probability increases for female while decreases for male.

Table 3: Parameter Estimates of Hazard of Remarriage by Gender using Discrete Time Logit Model

Variables	Male		Female	
	Model 1	Model 2	Model 3	Model 4
Age spline (17-30)	-0.164***	-0.162***	-0.176***	-0.176***
	(0.0142)	(0.0142)	(0.00896)	(0.00905)
Age spline (30 onwards)	-0.0902***	-0.0896***	-0.118***	-0.116***
	(0.00255)	(0.00286)	(0.00290)	(0.00308)
Education				
Less than High School	-0.291***	-0.0874	-0.313***	-0.220
	(0.0918)	(0.254)	(0.0590)	(0.139)
More than High School	0.299***	0.332	0.232***	0.688***
	(0.0621)	(0.186)	(0.0541)	(0.177)
Interaction: Age*<HS		-0.0049		-0.0024
		(0.00627)		(0.0033)
Interaction: Age*>HS		-0.00085		-0.0114**
		(0.0044)		(0.0044)
Children				
No. of children = 1	0.311***	0.311***	0.155***	0.158***
	(0.0630)	(0.0629)	(0.0554)	(0.0554)
No. of children = 2	0.300***	0.300***	0.278***	0.282***
	(0.0684)	(0.0684)	(0.0578)	(0.0580)
No. of children = 3	0.514***	0.514***	0.391***	0.393***
	(0.0794)	(0.0795)	(0.0671)	(0.0672)
Cohort				
Cohort (Baby boomers)	-0.270***	-0.271***	0.276***	0.278***
	(0.0708)	(0.0710)	(0.0597)	(0.0598)
Cohort (Gen X)	-0.309***	-0.311***	0.234***	0.235***
	(0.0788)	(0.0790)	(0.0620)	(0.0620)
Cohort (Gen Y and onwards)	-0.0176	-0.0271	0.139	0.141
	(0.182)	(0.182)	(0.120)	(0.120)
Duration Splines				
Unmarried duration (0-2)	0.134***	0.134**	0.0373	0.0380
	(0.0486)	(0.0487)	(0.0484)	(0.0485)
Unmarried duration (2-7)	-0.00352	-0.00386	0.0542***	0.0548***
	(0.0116)	(0.0116)	(0.0105)	(0.0105)
Unmarried duration (7 onwards)	0.0976***	0.0978***	0.0909***	0.0903***
	(0.00382)	(0.00383)	(0.00339)	(0.00343)
Region: West Germany	0.462***	0.461***	0.330***	0.330***
	(0.0605)	(0.0604)	(0.0512)	(0.0512)
Constant	2.719***	2.673***	2.379***	2.343***
	(0.410)	(0.413)	(0.258)	(0.261)
Observations	58,602	58,602	130,963	130,963
Log Pseudo Likelihood	-12648.795	-12648.259	-14296.993	-14923.282

Robust Standard errors in parentheses brackets: * p<0.05, ** p<0.01, *** p<0.001
Notes:
Table 3 shows the estimated model of remarriage for male, and female based on retrospective SOEP BIOMARSY for age 17 to 100. Standard errors are clustered at individual levels. Baseline groups: Education (HS degree); Region (East Germany): Children (no kids); Cohort (Birthyear < 1946)

Following regression analysis, we have performed a Wald test to check if parameters of the fitted model of education, cohort, children, and region are different from zero. We reject the null hypothesis for all variables except education at 1 % significance level. Since our fitted logit model does not provide information on magnitude, we have performed post-estimation analysis to analyze average marginal effects. Table-8A (Appendix) summarizes the average marginal effects by gender. The output indicates that for >HS female, probability of remarriage after a divorce or widowhood, increases by 0.0075 percentage points. The probability of remarriage increases by 0.0186 percentage points for a male with >HS degree. We have found interesting average marginal effects for duration analysis. The probability of remarriage for a male decrease if a male remains "unmarried" for a longer time after losing a partner. Individuals, when remain independent and single for a longer period, get more comfortable in a solitary lifestyle and establish routines and hobbies that fit their own personal preferences. It becomes difficult for them to find a compatible partner compared to those who just move on and try to find a partner after the end of relationship. Moreover, the decreased social and cultural pressure after few years also reduce their want to find a partner. On the other hand, for female, the probability increases as the coefficient of average marginal effect gets higher for duration greater than seven years. For female, an event of divorce or widowhood may have a stronger emotional effect that reduces over time and they are more likely to consider remarrying.

5.3 Model of Divorce

Table-4 shows the results of likelihood of divorce for male and female. We see, for both male and female, there is a non-linear relationship between age and probability of divorce. We capture this non-linearity using age splines. Within the age-range of 17-30, the positive coefficient suggests that discrete hazard of divorce is higher in the early adulthood and as individuals' age, this hazard decreases, as the coefficient of age spline 30 onwards is negative. For female, we do not identify any significant effect of age on probability of getting divorced at early ages. In terms of education, both men and women with less than a high school degree exhibit significantly lower risk of divorce compared to those with a high school degree for men. This finding suggests that lower educational attainment acts as a protective factor against divorce for men. We identify regional differences in predicting the probability of divorce and sample in West Germany is at higher risk of divorce compared to East Germans. Cohort effects reflect interesting patterns that also match the empirical observation. For both genders, being part of Baby Boomers or born after 1963 significantly increases the likelihood of divorce compared to those born before 1946. This suggests that individuals in the younger cohorts experience higher risks of divorce, potentially due to societal changes, shifts in norms, reforms in divorce law[3] or other cohort-specific factors. We are also controlling for duration effects suggests that the longer the person remains in a relationship, the stronger relationship he/she builds. As the number of years in a marriage relationship increases, the probability of divorce decreases significantly. The presence of children significantly affects the prediction of divorce and children appear to reduce the risk of divorce. Moreover, couples with second or higher marriage have higher probability of divorce compared to those with first marriage.

Following regression analysis, we have performed a Wald test to check if parameters of the fitted model of education, cohort, children, and region are different from zero. We reject the null hypothesis for all variables except education at 1 % significance level. Since our fitted logit model does not provide information on magnitude, we have performed post-estimation analysis to analyze average marginal effects. Table-9A (Appendix) summarizes the average marginal effects by gender.

[3] Divorce reform 1976 – Unilateral divorce law in Germany (see details Müller-Freienfels, 1979)

Table 4: Parameter Estimates of Hazard of Divorce by Gender using Discrete Time Logit Model

Variables	Male	Female
Age Splines		
Age spline (17-30)	0.00594	0.0110
	(0.0144)	(0.00864)
Age spline (30 onwards)	-0.0347***	-0.0192***
	(0.00320)	(0.00322)
Education (Base: High School)		
Less than High School	-0.182***	-0.188***
	(0.0397)	(0.0349)
More than High School	-0.0670*	-0.0279
	(0.0402)	(0.0385)
Cohort (Base: birthyear < 1946)		
Cohort: Baby Boomers	0.783***	0.756***
	(0.0417)	(0.0408)
Cohort: >1963	0.776***	1.021***
	(0.0486)	(0.0445)
Duration Splines		
Marriage duration (0-2)	0.748***	0.869***
	(0.0492)	(0.0475)
Marriage duration (2-7)	0.0190	0.0161
	(0.0123)	(0.0105)
Marriage duration (7 onwards)	-0.0113***	-0.0332***
	(0.00397)	(0.00384)
Children (Base: No kids)		
No. of children = 1	-0.477***	-0.756***
	(0.160)	(0.151)
No. of children = 2	-1.297***	-1.049***
	(0.192)	(0.131)
No. of children = 3	-0.121	-0.382***
	(0.139)	(0.118)
Second or more marriages	1.483***	1.194***
	(0.0423)	(0.0391)
Region (Base: East)		
Region: West Germany	0.0883**	0.0472
	(0.0391)	(0.0362)
Constant	-6.809***	-7.050***
	(0.405)	(0.249)
Observations	468,615	519,753
Log Pseudo Likelihood	-18112.009	-23340.83

Robust Standard errors in parentheses brackets: * p<0.05, ** p<0.01, *** p<0.001
Notes: Table 4 shows the estimated model of divorce for male, and female based on retrospective SOEP BIOMARSY for age 18 to 100. Standard errors are clustered at individual levels. Baseline groups: Baseline groups: Education (HS degree); Cohort (<1946); Region (East Germany); Children (no kids)

5.4 Model of Fertility

Table-5 shows the results of likelihood of child by number of children. The results for likelihood of a child (for all columns) show that likelihood for a child is stronger for the first two age splines and it declines over age, mainly because of biological reasons. Women have significantly higher risk of having a child compared to male however, the negative coefficient of interaction of age and gender suggests that as women age, the likelihood of having a child is weaker for women compared to men. Education level negatively affects the likelihood of having a child, however, the negative coefficient of interaction between <HS and age predicts that as low qualified individuals age, their likelihood to have a child decreases. On the other hand, the positive coefficient of >HS and age suggests that as >HS graduates get older, their likelihood to have a child increases. We observe that sample from West Germany has less probability to have a child compared to East Germany. Our model also controls for cohort effects, and we find that the likelihood of having a third child has significantly reduced in younger cohorts compared to the base group. Another interesting result is that individuals who are either never-married cohabitants or previously married cohabitants have less likelihood to have children compared to married couples. Since our fitted logit model does not provide information on magnitude, we have performed post-estimation analysis to analyze average marginal effects. Table-10A (Appendix) shows the average marginal effects of likelihood of children. Following regression analysis, we have performed Wald test to check if parameters of the fitted model of education, cohort, children, and region are different from zero. We reject the null hypothesis for all variables except education at 1 % significance level.

Table 5: Parameter Estimates of Likelihood of Children using Discrete Time Logit Model

Variables	First Child	Second Child	Third Child
Age Splines			
Age spline (17-20)	0.409***	0.288***	0.0596
	(0.0157)	(0.0532)	(0.198)
Age spline (20-25)	0.115***	0.179***	0.174***
	(0.00434)	(0.00683)	(0.0170)
Age spline (25 onwards)	-0.0953***	-0.0706***	-0.0741***
	(0.00181)	(0.00190)	(0.00308)
Gender: female	1.771***	1.230***	1.182***
	(0.0559)	(0.0656)	(0.114)
Gender # age	-0.0371***	-0.0443***	-0.0435***
	(0.00212)	(0.00221)	(0.00349)
Education (Base: High School)			
Education <HS	1.619***	0.560***	0.523***
	(0.0780)	(0.0839)	(0.125)
Education >HS	-2.919***	-1.119***	-0.483***
	(0.0618)	(0.0773)	(0.140)
Interaction: Education <HS#Age	-0.0528***	-0.0120***	-0.00459
	(0.00312)	(0.00305)	(0.00407)
Interaction: Education >HS#Age	0.0985***	0.0453***	0.0247***
	(0.00224)	(0.00254)	(0.00412)
Region: West-Germany	-0.403***	0.190***	0.321***
	(0.0163)	(0.0171)	(0.0268)
Marital Status: Base (Currently married)			
Never-married Single	-2.002***	-0.927***	-0.105***
	(0.0143)	(0.0188)	(0.0372)
Previously married now Cohabiting	-1.310***	-0.861***	-0.000380
	(0.0501)	(0.0426)	(0.0514)
Cohort (Base: Birthyear < 1915)			
Cohort 1915-1935 (First World War)	0.380***	0.0850	-0.0928
	(0.0500)	(0.0639)	(0.0828)
Cohort 1936-1945 (Second World War)	0.621***	0.103	-0.547***
	(0.0498)	(0.0636)	(0.0832)
Cohort 1946-1963 (Baby Boomers)	0.977***	0.189***	-0.616***
	(0.0481)	(0.0621)	(0.0810)
Cohort 1964-1980 (Gen X)	1.275***	0.507***	0.0120
	(0.0478)	(0.0619)	(0.0805)
Cohort 1981-1989 (Gen Y)	1.206***	0.428***	-0.193**
	(0.0506)	(0.0660)	(0.0927)
Cohort > 1989 (Gen Z)	0.388***	0.478***	-0.273
	(0.0712)	(0.114)	(0.229)
Constant	-10.95***	-8.660***	-4.480
	(0.312)	(1.058)	(3.931)
Observations	822,341	288,687	256,405
Log Likelihood	-142423.75	-89041.064	-42867.459

Standard errors in parentheses brackets: * p<0.05, ** p<0.01, *** p<0.001
Notes: Table 5 shows the estimated model of fertility based on retrospective SOEP BIOBIRTH for age 18 to 100. Standard errors are clustered at individual levels. Baseline groups: Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany); Marital Status (Currently Married)

5.5 Simulations

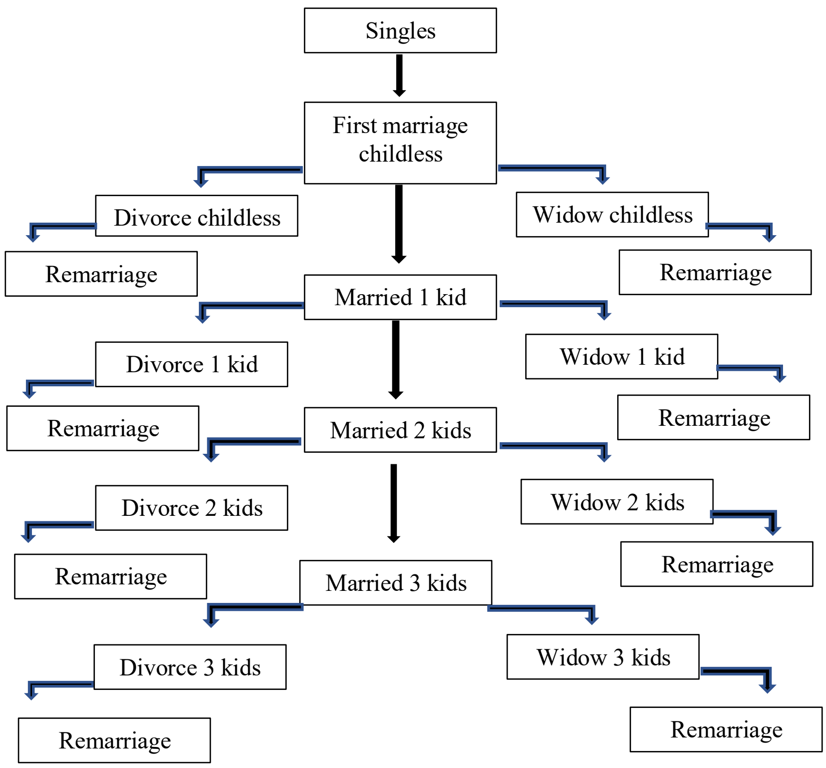
After estimating the models, we develop a population simulation model based on time-inhomogeneous first-order Markov chains. We calibrate our model using the estimated coefficients and run one million simulations. For our simulations we assume a population with 20 percent married population, 40 percent single men and 40 percent single women. We assume that the age difference between husband and wife is 2 years. For both cohorts, we assume that simulated population starts with age 20 and we assume a time period of 2019. For each simulation in each period, we estimate conditional transition probabilities with respect to origin state and make a random number draw. If the probability of transition is higher compared to random number, the individual changes the state. We repeat this for each year of age and that's how our simulated population evolves over the lifecycle. Our multistate transition states model has many states with and without children for both men and women.

Figure V shows the flowchart of possible marital transitions both for male and female. Only the state "Married" is not defined separately for any gender and just represents a legally married couple. Rectangular boxes show all the marital states whereas, the arrows represent possible transition from one state to another in the direction of arrows. All individuals are born singles and can get married at the age of 20. Once an individual gets married, he cannot get the status of single again. The transition from one state to another state happens because of the event. The event of death is the absorbing state, from which an individual cannot move to any other state.

There are different events that lead to state transitions and result in various family states. The states are as follows:

- First marriage and remarriage
- Birth of children
- Divorce or dissolution of marriage
- Death of spouse or Self-death

Figure V: State Transitions Model



Notes: Figure V shows the flowchart of possible marital transitions for male and female. Rectangular boxes show all the defined marital states whereas, the arrows represent possible transition from one state to another in the direction of arrows. For example, an individual is born single (never-married) and can move to various other states. The transition rate of moving from one state to another state varies across individuals and our regression models have estimated these transition probabilities. In our population model of time-inhomogeneous first order Markov chain, we have used this state transition model. It means that after defining an initial simulated population, the population evolves according to this state transition model each period with a transition probability that has been estimated using our regression coefficients. It is important to note that we can define and use any state transition model (with more states) and the presented model is just to show an implication of our results.

Figure VI (a) is for the baby boomers and Figure VI (b) is for Generation Y and onwards. For each birth cohort, we present the simulation results for three education categories. Simulation results show the distribution of population by marital status (average relative frequency distributions of population) over the life cycle.

Figure VI: Distribution of Simulated Population by Marital Status German versus U.S.

Figure VI (a): Baby Boomers (Birth year 1946-1963)

Education < HS group   Education = HS group   Education > HS group

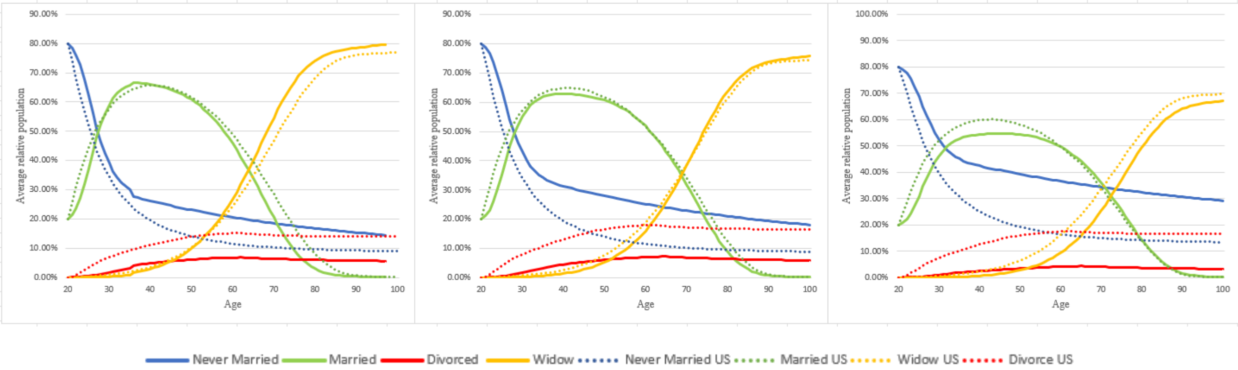
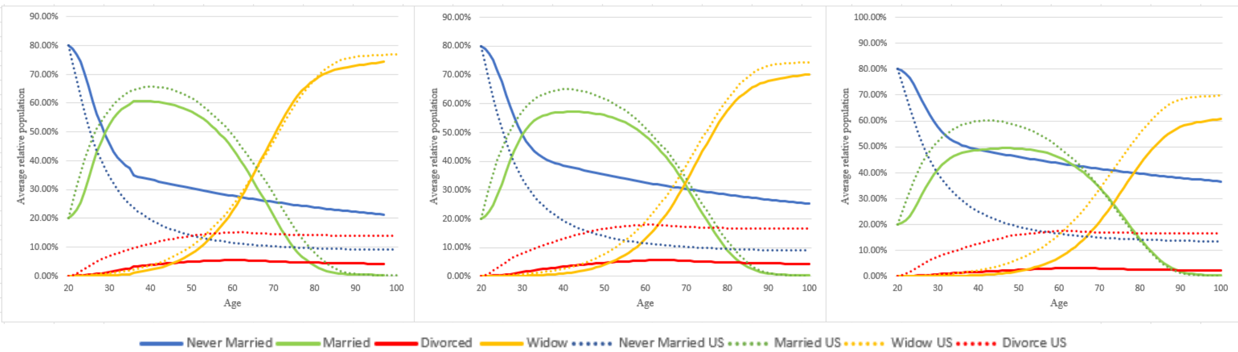


Figure VI (b): Generation Y (Birth year > 1980)

Education < HS group   Education = HS group   Education > HS group



Notes: Figures VI (a) (Baby Boomers: Cohort 1946-1963) and VI (b) (Generation Y: Cohort >1980) are based on 1 million simulations with the initial 80 percent population of single men and women and 20 percent married couples from age 20-100. Left graph (Population with <HS education), Middle graph (Population with =HS education), Right graph (Population with >HS education). The figures show the distribution (average relative population by percentage) for individuals with different marital status (singles/never married, married, divorced and widows). Solid line represents the marital states for Germany and dotted lines are based on MINT (Model 4) estimates (Smith, Favreault, Butrica & Issa, 2016).

In general, the distributions across <HS and =HS are quite similar, but we see notable differences for educational group >HS. The population of single individuals starts decreasing rapidly because of increasing transition rate to marriage. However, this transition rate is higher among <HS and =HS groups where, <HS group shows low transition rate from single to marriage. The initial marriage population (20 percent) rapidly increases, and more population get married, but we can see that for >HS the average population of married people is relatively lower compared to other two educational categories. As more individuals gets married, we see an increasing percentage of divorced individuals and an increasing percentage of widows at higher ages.

Our simulations of baby boomers and Generation Y captures a decreasing population of married individuals in the newer birth cohorts. By age 40 (baby boomers), the initial 80 percent single population has reduced to 28 percent (<HS), 31 percent (=HS) and 42 percent (>HS). For Generation Y, the initial 80 percent single population has reduced to 31 percent (<HS), 39 percent (=HS) and 49 percent (>HS). We see a decreasing population of people getting married in newer birth cohorts and the differences are widening among the higher educated people.

We have also run the simulations using MINT coefficients to compare the distributions across Germany and U.S. Our simulations show that transition to marriage across all education groups and ages is less in German simulated population compared to U.S. The proportion of divorced simulated population for Germany is far less compared to U.S., which supports the empirically observed crude divorce rates.[4] The state level data for Germany shows that crude divorce rate in Germany from 1970-2020 has increased from 1.3 to 1.7 and for U.S it has decreased from 3.5 to 2.3 people (Table-11A Appendix). However, the crude divorce rate is very high in U.S. compared to Germany. Moreover, the less proportion of divorced population in Germany does not only reflect less number of divorced but also driven by high estimated remarriage rates in German sample.

[4] The crude divorce rate (CDR), defined as the number of divorces during a given year per 1000 people (OECD, n.d.).

6 Discussion

We start our discussion with our model of morality and our results indicate the existence of considerable social differences in mortality for the given sample of Germany. This finding is in line with the existing literature which shows a significant relationship between socioeconomic factors and life expectancy Lampert & Kroll, 2014; Lampert, Hoebel & Kroll, 2019). To evaluate the variations in life expectancies for different levels of income groups, we have conducted a regression analysis with three income groups. These groups are based on net household income per month, including low income (<15001 Euro), middle income (>15000 Euro & <42001 Euro) and high income (>42000 Euro). In Table-6, we summarize our estimations of expected remaining life of a 65-year-old male/female belonging to a low, medium, or high level of income group.

Table 6: Average Remaining Life by Income Group

Estimated remaining life at age 65 (years) by gender and income			
Income Group	Low income	Middle income	High income
Male	5.34	8.96	14.46
Female	13.67	16.63	19.54

Note: This table shows the expected remaining life of a 65-year-old male/female belonging to a low, medium, or high level of income group. There are three income groups based on net household income per month, including low income (<15001 Euro), middle income (>15000 Euro & <42001 Euro) and high income (>42000 Euro). Table shows that on average, female and high income is associated with more years of remaining life.

Based on these estimations, the average life expectancy of a 65-year-old, is 81.75 years where, the mean life expectancies of a 65-year-old male, depending on low, medium, and high-income groups are 70, 74 and 79 respectively and for 65-year-old female, they are 79, 82 and 85 respectively. The difference of average life expectancies between low- and high-income groups is 9 years among men and 6 years among women, reflecting the mortality inequality caused by income. We see a clear gradient where high income is associated with increased longevity for both genders. The observed disparities caused by low socioeconomic status of individuals suggest that there is a dire need to evaluate the policies aimed at reducing income disparities and promoting fair economic opportunities and equal social welfare programs. Other than income and education, we have found that marital status of individuals also affect their mortality risks and life expectancy.

Moving to our next model on predictions of marriage which is built on SOEP retrospective dataset. Our findings show that marriage is facing a generational facelift as models predict that the likelihood of marriage of younger cohorts is declining. It reflects changing personal attitudes, experiences, social norms, legal reforms, and many other factors. The regression results identify the varying hazards of marriage across subgroups, based on age, gender, education and region. Our model predicts a negative relationship between education and first marriage which can be because of many reasons, consequences, and policy implications. A highly educated men or women is more ambitious for career goals, personal growth, financial stability, and prosperity, therefore, has less probability of getting into long-term relationship compared to other groups. Education may affect the attitude of individuals towards traditional family structures and these individuals are more likely to start with a cohabitation relationship rather than a binding long-term marriage relationship. Another reason can be educational assortative mating as individuals want to look for a similar partner.

The predicted decline in probability of marriage may have important consequences like changing family dynamics. For example, delayed age of first marriage, increase in cohabitation, alternative living arrangements, less number of children and also changing traditional gender roles. Our model shows that for the given sample, the average age differs across three educational groups, for women, the average ages at first marriages are 25 (<HS), 27(=HS) and 31(>HS). For men, the average ages at first marriages are 30 (<HS), 29(=HS) and 32(>HS). Although these averages are slightly below compared to country level statistical data[5], but we emphasize that our results highlight the heterogeneities that exists among these subgroups and understanding these differences is important to understand the needs of specific subgroups.

These changing patterns have important policy implications like introducing family-friendly policies under different living arrangements. Although social legislation has regularly treated cohabiting couples as mutually supportive unions, this status does not provide them with the right which they immediately get after making a marriage relationship (Ostner, 2001). For example, non-working

[5] According to Destatis Statistisches Bundesamt (2022b), the average ages at first marriage for men and women are 34.8 and 32.3 respectively.

cohabiting partners neither claim a part of their partner's pension nor exempt from health insurance contributions. In case of dissolution of relationship, they cannot claim distribution of wealth or financial support (like marital alimony) (Ostner, 2001). The death of cohabiting partner does not provide any compensations and claims for inheritance (Kreidler-Pleus, 2020). By pointing that out, we aim to emphasize that we really need to understand family dynamics across subgroups because these subgroups may have different attitudes e.g., towards traditional marriage norms. After identifying these patterns, researchers and policy makers can better develop and evaluate policies for example, in the above context, promotion of diverse family structures and arrangements and promoting reforms that recognize and protect non-traditional forms of relationship.

We now turn towards the limitations of our study. Determinants of family events encompass a wide range of factors other than personal characteristics like societal and cultural factors. We have tried to control such changes by controlling time which captures underlying shifts or trends in the data. However, marriage models are based on marital histories of survey participants which limits our ability to control time effects for retrospective data. Nevertheless, we have been able to control other important variables like duration of relationship histories which really matters in predicting such events. We also believe that there are many other unobserved determinants that can contribute towards such predictions and by not accounting for them, the results could be biased. However, our goal is to precisely predict the likelihoods of various life events and not to investigate causal analysis.

Our statistical findings are based on relationship histories of individuals since their birth, therefore, do not allow us to capture important economic variables like income and occupation. However, we are still able to explore many key determinants that can significantly predict the likelihoods of individuals coming across various family life events. These predictions can help us to explore family dynamics and can provide a background of recent demographic transitions in Germany. Furthermore, our event-oriented observation data may also have limitations like measurement errors. Since event histories strongly rely on memories of respondents therefore, respondents may find it hard to recall the exact years of events, leading to bias. Another limitation of this design is that we cannot control the behavioral aspect of individuals which may be key determinants of these events.

7 Conclusion

In general, family life events are endogenous events as they are based on personal choices, behaviors, and individuals' decisions. Yet they exhibit certain patterns that are predictable and can be captured using statistical models and accounting for social and cultural factors, economic and demographic characteristics. This paper delves into the dynamics of family life events in Germany by employing discrete time event history analysis. Our findings reveal crucial insights into the impact of various factors on demographic and family transitions, providing valuable information for informed policy decisions and understanding the social implications of these events. Our paper uncovers the importance of demographic factors including educational attainment, income levels, and marital status in shaping life outcomes and mortality risk. The discrete time survival models for mortality, marriage, divorce, and birth predict the likelihood of experiencing these events across heterogeneous individuals. We show the income disparities and educational differences contributing towards variations in life expectancies across individuals. These crucial insights unveil the impact of various factors on demographic and family transitions, provide valuable information for informed policy decisions and understanding the social implications of these events.

Bibliography

Allison, P. D. (1982). Discrete-Time Methods for the Analysis of Event Histories. Sociological Methodology, 13, 61–98. https://doi.org/10.2307/270718

Altonji, J. G., Hynsjo, D. M., & Vidangos, I. (2021). Marriage dynamics, earnings dynamics, and lifetime family income (No. w28400). National Bureau of Economic Research.

Attanasio, O. P., Banks, J., Meghir, C., & Weber, G. (1999). Humps and bumps in lifetime consumption. Journal of Business & Economic Statistics, 17(1), 22-35.

Blossfeld, H. P., Rohwer, G., & Schneider, T. (2019). Event history analysis with Stata. Routledge.

Bonnet, C., Garbinti, B., & Solaz, A. (2021). The flip side of marital specialization: the gendered effect of divorce on living standards and labor supply. Journal of Population Economics, 34, 515-573.

Canchola, A. J., Stewart, S. L., Bernstein, L., West, D. W., Ross, R. K., Deapen, D., & Horn-Ross, P. L. (2003). Cox regression using different timescales. Western Users of SAS Software. San Francisco, California.

Case, A., & Deaton, A. (2017). Mortality and morbidity in the 21st century. Brookings papers on economic activity, 2017, 397.

Cleves, M. (2008). An introduction to survival analysis using Stata. Stata press.

Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202.

Deaton, A. S., & Paxson, C. (1994). Saving, growth, and aging in Taiwan. In Studies in the Economics of Aging (pp. 331-362). University of Chicago Press.

Destatis Statistisches Bundesamt, (2022a). Average life expectancy (period life table): Germany, period of years, sex, completed age. (Available at: https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12621-0002&zeitscheiben=15&sachmerkmal=ALT577&sachschluessel=ALTVOLL000,ALTVOLL020,ALTVOLL040,ALTVOLL060,ALTVOLL065,ALTVOLL080&language=en#abreadcrumb)

Destatis Statistisches Bundesamt, (2022b). Families and family members with minor children living in the family.(Available at: https://www.destatis.de/EN/Themes/Society-Environment/Population/Households-Families/Tables/families.html)

DIW Berlin /SOEP (Ed.) (2022) : SOEP-Core v37 - PPATHL: Person-related meta-dataset, SOEP Survey Papers, No. 1187, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

Federal Ministry of Family Affairs, Senior Citizens, Women and Youth, (2022). Family Today. Data, Facts Trends Family Report - 2020, Berlin. (Available at: https://www.bmfsfj.de/resource/blob/190146/7313bd244bf3b7d05b334afc70ce7e1c/family-today-data-facts-trends-family-report-2020-data.pdf)

Fisher, L. D., & Lin, D. Y. (1999). Time-dependent covariates in the Cox proportional-hazards regression model. Annual review of public health, 20(1), 145-157.

Gove, W. R. (1973). Sex, Marital Status, and Mortality. American Journal of Sociology, 79(1), pp. 45-67.

Grabka, M. M. (2020). SOEP-core v35-codebook for the $ pequiv file 1984-2018: CNEF variables with extended income information for the SOEP (No. 772). SOEP Survey Papers. (Available at: https://www.econstor.eu/bitstream/10419/214236/1/1690462639.pdf)

Han, Y. S., Liefbroer, C. A. & Elzinga, H. C. (2020). Mechanisms of family formation: an application of Hidden Markov Models to a Life Course Process. Advances in Life Course Research.

Heeringa, S. G., West, B. T. & Berguld, P. A., (2010). Survival Analysis for Event History Survey Data. Dans: Apply Survey Data Analysis. Broken Sound Parkway NW: Taylor and Francis Group, LLC, pp. 303-329.

Hong, J. H. & Rios-Rull, J.-V. (2007). Social security, life insurance and annuities for families. Journal of Monetary Economics, 54, 118-140.

House, J. S., Robbins, C., & Metzner, H. L. (1982). The association of social relationships and activities with mortality: Prospective evidence from the Tecumseh Community Health Study. American journal of epidemiology, 116(1), 123-140.

Hubener, A., Maurer, R., & Mitchell, O. S. (2016). How family status and social security claiming options shape optimal life cycle portfolios. The Review of Financial Studies, 29(4), 937-978.

Hu, Y. Goldman. N. (1990). Mortality differentials by marital status: An international. Demography, 27(2), 233-250.

Kalbfleisch, J. D. & Prentice, R. L. (2002). The statistical analysis of failure time data. second edition. New Jersey: John Wiley & Sons.

Kibele, E. U., Klüsener, S., & Scholz, R. D. (2015). Regional mortality disparities in Germany: Long term dynamics and possible determinants. Kölner Z Sozpsychol, 67, 241-70. (doi: 10.1007/s11577-015-0329-2. PMID: 26412875; PMCID: PMC4577550).

Kreidler-Pleus, D. (2020). Family law in Germany: Overview. Thomas Reuters Practical Law. (Available at: https://ca.practicallaw.thomsonreuters.com/8-575-0676?transitionType=Default&contextData=(sc.Default)&firstPage=true)

Keeley, M. C. (1979). An analysis of the age pattern of first marriage. International Economic Review, 527-544.

Kroh, M., Neiss, H., Kroll, L., & Lampert, T. (2012). Menschen mit hohem Einkommen leben länger. DIW Wochenbericht, 79(38), 3-15.

Kühn, M., Dudel, C., Vogt, T., & Oksuzyan, A. (2019). Trends in gender differences in health at working ages among West and East Germans. SSM-Population Health, 7, 100326.

Lampert, T., & Kroll, L. E. (2014). Social differences in mortality and life expectancy. Robert Koch Institue.

Lampert, T., Hoebel, J., & Kroll, L. E. (2019). Social differences in mortality and life expectancy in Germany. Current situation and trends. Journal of health monitoring, 4(1), 3.

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. The Review of Financial Studies, 23(1), 385-432.

Luy, M., Wegner-Siegmundt, C., Wiedemann, A., & Spijker, J. (2015). Life expectancy by education, income and occupation in Germany: estimations using the longitudinal survival method. Comparative Population Studies, 40(4).

Müller-Freienfels, W. (1979). The Marriage Law Reform of 1976 in the Federal Republic of Germany. International & Comparative Law Quarterly, 28(2), 184-210

O'Dea, C. (2012). APC: Stata module to estimate age, period and cohort effects. Institute for Fiscal Studies.

OECD. (n.d.). OECD family database. (Available at: https://www.oecd.org/els/family/database.htm)

O'Harra, J. & Sabelhaus, J. (2002). Projecting Longitudinal Marriage Patterns for long-run policy analysis. Working Paper 14080, Congressional Budget Office.

Ostner, I. (2001). Cohabitation in Germany - Rules, Reality and Public Discourses. International Journal of Law, Policy and Family, 15, 88-101.

Oswald, A., & Gardner, J. (2003). Is it Money or Marriage that Keeps People Alive? In Royal Economic Society Annual Conference 2003 (No. 161). Royal Economic Society.

Panis, C. & Lillard, L. (1999). Final Report: Near Term Model Development, Santa Monica, California: RAND for the Social Security Administration, Office of Research, Evaluation and Statistics.

Scholz, J. K., Seshadri, A. & Khitatrakun, S. (2006). Are Americans Saving Optimally for Retirement? Journal of Political Economy, 114(4), 607-643.

Smith, K. E., Favreault, M. M., Butrica, B. A. & Issa, P. (2010). Modeling income in the near term version 6, Urban Institute. United States of America. (Available at: https://policycommons.net/artifacts/633658/modeling-income-in-the-near-term-version-6/1614973/)

Smith, K. E., Favreault, M. M., Ratcliffe, C. E., Butrica, B. A., Toder, E. J., & Bakija, J. (2007). Modeling income in the near term 5. (Available at SSRN 2206397.)

SOEP Group, (2019). SOEP-Core v34–PPATHL: Person-Related Meta-Dataset. SOEP Survey Papers 762: Series D–Variable Descriptions and Coding.

SOEP Companion, (2021). Working with SOEPhelp. (Available at: http://companion.soep.de/Working%20with%20SOEP%20Documentation/index.html)

United Nations, (2023). Germany Life Expectancy 1950-2023. Available at: https://www.macrotrends.net/countries/DEU/germany/life-expectancy

Willekens, F. (2017). Continuous-time microsimulation in longitudinal analysis. In New frontiers in microsimulation modelling (pp. 413-436). Routledge.

Appendix

Table 1A: Distributional Characteristics of Model of Mortality Sample

	Male	Female	Total Sample
Cohort			
Cohort(birthdate < 1915)	3,206 (0.93%)	6,190 (1.67%)	9,396 1.31%)
Cohort 1915-1935 (First World War)	36,600 (10.60%)	44,065 (11.87%)	9,396 (11.26%)
Cohort 1936-1945 (Second World War)	51,580 (14.94%)	49,926 (13.45%)	101,506 (14.17%)
Cohort 1946-1963 (Baby Boomers)	110,061 (31.89%)	114,630 (30.88%)	224,691 (31.37%)
Cohort 1964-1980 (Gen X)	103,325 (29.94%)	113,614 (30.61%)	216,939 (30.28%)
Cohort > 1980 (Gen Y and onwards)	40,380 (11.70%)	42,784 (11.53%)	83,164 (11.61%)
Education			
Less than High School	47,040 (13.63%)	75,090 (20.23%)	122,130 (17.05%)
High School	216,993 (62.87%)	223,578 (60.23%)	440,571 (61.50%)
More than High School	81,119 (23.50%)	72,541 (19.54%)	153,660 (21.45%)
Marital Status of Individual			
Married	224,322 (64.99%)	223,588 (60.23%)	447,910 (62.53%)
Single	82,709 (23.96%)	71,192 (19.18%)	153,901 (21.48%)
Widow	9,981 (2.89%)	4,473 (9.29%)	44,454 (6.21%)
Divorce	19,764 (5.73%)	32,351 (8.72%)	52,115 (7.27%)
Separate (not legally divorced)	8,376 (2.43%)	9,605 (2.59%)	17,981 (2.51%)
Region			
West-Germany	275,692 (79.88%)	295,444 (79.59%)	571,136 (79.73%)
East-Germany	69,460 (20.12%)	75,765 (20.41%)	145,225 (20.27%)
Age (Mean)	47.23	47.84	47.54
Age(Standard deviation)	16.67	17.15	16.92

Notes: Table 1A shows the distributional characteristics of the sample for modeling "mortality". Table summarizes the frequencies and percentages (in brackets) of each variable of interest for men, women and total sample. These variables include the cohort, education, marital status and region. In our sample for model of mortality, 60.23% of the women are married, 19.18% are singles, 9.29% are widows, 8,72% are divorced and 2.59% are separated but not legal divorced. Similarly, we can interpret the distributions of other variables in the given sample and across the two genders.

Table 2A: Distributional Characteristics of Model of First Marriage Sample

	Male	Female	Total sample
Cohort			
Cohort (birthdate < 1915)	6,452 (1.56%)	10,933 (3.09%)	17,385 (2.27%)
Cohort 1915-1935 (First World War)	39,024 (9.45%)	42,123 (11.90%)	81,147 (10.58%)
Cohort 1936-1945 (Second World War)	49,066 (11.88%)	34,514 (9.75%)	83,580 (10.90%)
Cohort 1946-1963 (Baby Boomers)	131,999 (31.95%)	94,188 (26.61%)	226,187 (29.49%)
Cohort 1964-1980 (Gen X)	145,649 (35.26%)	135,536 (38.30%)	281,185 (36.66%)
Cohort > 1980 (Gen Y and Onwards)	40,933 (9.91%)	36,610 (10.34%)	77,543 (10.11%)
Education			
Less than High School	51,906 (12.56%)	55,542 (15.69%)	107,448 (14.01%)
High School	240,349 (58.18%)	197,411 (55.78%)	437,760 (57.07%)
More Than High School	120,868 (29.26%)	100,951 (28.52%)	221,819 (28.92%)
Region			
West-Germany	339,794 (82.25%)	290,908 (82.20%)	630,702 (82.23%)
East-Germany	73,329 (17.75%)	62,996 (17.80%)	136,325 (17.77%)

Notes: Table 2A shows the distributional characteristics of the sample for modeling "first marriage". Table summarizes the frequencies and percentages (in brackets) of cohort, education and region for men, women and total sample. For example, in the given sample of men, 12.56% have <HS education, 58.18% have high school degree and 29.26% have >HS education. Similarly, we can interpret the distributions of other variables in the given sample and across the two genders.

Table 3A: Distributional Characteristics of Model of Remarriage Sample

	Male	Female	Total sample
Cohort			
Cohort (birthdate < 1946)	23,651 (39.5%)	71,361 (53.49%)	95,012 (49.16%)
Cohort (baby boomers)	25,837 (43.15%)	39,115 (29.32%)	64,952 (33.60%)
Cohort (gen x)	9,988 (16.68%)	21,575 (16.17%)	31,563 (16.33%)
Cohort (gen y and onwards)	397 (0.66%)	1,361 (1.02%)	1,758 (0.91%)
Education			
Less than High School	5,666 (9.54%)	35,535 (26.87%)	41,201 (21.50%)
High School	40,961 (68.98%)	78,902 (59.67%)	119,863 (62.56%)
More Than High School	12,756 (21.48%)	17,791 (13.45%)	30,547 (15.94%)
Region			
West-Germany	44,558 (74.42%)	100,789 (75.55%)	145,347 (75.20%)
East-Germany	15,315 (25.58%)	32,623 (24.45%)	47,938 (24.80%)

Notes: Table 3A shows the distributional characteristics of the sample for modeling "Remarriage". Table summarizes the frequencies and percentages (in brackets) of cohort, education and region for men, women and total sample. For example, in the given sample of men, 9.54% have <HS education, 68.98% have high school degree and 21.48% have >HS education. Similarly, we can interpret the distributions of other variables in the given sample and across the two genders.

Table 4A: Distributional Characteristics of Model of Divorce Sample

	Male	Female	Total sample
Cohort			
Cohort (birthdate < 1946)	238,555 (50.41%)	251,028 (47.67%)	489,583 (48.96%)
Cohort (baby boomers)	159,878 (33.78%)	174,967 (33.22%)	334,845 (33.49%)
Cohort (gen x onwards)	74,825 (15.81%)	100,626 (19.11%)	175,451 (17.55%)
Education			
Less than High School	209,583 (44.68%)	278,044 (53.43%)	487,627 (49.28%)
High School	130,968 (27,92%)	151,118 (29,04%)	282,086 (28,51%)
More Than High School	128,543 (27.40%)	9,230 (17.53%)	219,773 (22.21%)
Region			
West-Germany	377,346 (79.73%)	419,438 (79.65%)	796,784 (79.69%)
East-Germany	95,912 (20.27%)	107,183 (20.35%)	203,095 (20.31%)
Age (mean)	41,93	39,26	40,52
Age(standard deviation)	12,24	12,21	12,3
Duration (mean)	14,07	14,1	14,09
Duration(standard deviation)	11,66	11,73	11,69

Notes: Table 4A shows the distributional characteristics of the sample for modeling "Divorce". Table summarizes the frequencies and percentages (in brackets) of each variable of interest for men, women and total sample. These variables include the cohort, education and region. The mean and standard deviations for age and duration variable are presented in the last rows of the table. The table shows that mean duration of marriage in our sample is 14.07 years for males and 14.1 years for females. This is very close to the statistical value of 14.5 for 2021 average duration of marriage in Germany by Destatis Statistisches Bunesamt.

Table 5A: Distributional Characteristics of Model of Fertility Sample

	Male	Female	Total Sample
Cohort			
Cohort(birthdate < 1915)	10,546	17,829	28,375
	(1.52%)	(2.56%)	(2.04%)
Cohort 1915-1935 (First World War)	85,149	91,178	176,327
	(12.3%)	(13.1%)	(12.7%)
Cohort 1936-1945 (Second World War)	107,401	97,245	204,646
	(15.51%)	(13.97%)	(14.74%)
Cohort 1946-1963 (Baby Boomers)	231,615	222,070	453,685
	(33.45%)	(31.9%)	(32.67%)
Cohort 1964-1980 (Gen X)	204,846	217,730	422,576
	(29.59%)	(31.27%)	(30.43%)
Cohort > 1980 (Gen Y and onwards)	38,251	39,964	78,215
	(5.52%)	(5.74%)	(5.63%)
Education			
Less than High School	83,580	126,005	209,585
	(12.25%)	(18.33%)	(15.3%)
High School	415,181	407,599	822,780
	(60.85%)	(59.3%)	(60.07%)
More than High School	183,494	153,781	337,275
	(26.9%)	(22.37%)	(24.63%)
Children			
No kids	470,330	322,758	793,088
	(67.93%)	(46.36%)	(57.11%)
1 kid	100,398	161,673	262,071
	(14.5%)	(23.22%)	(18.87%)
2 kids	92,057	155,647	247,704
	(13.3%)	(22.36%))	(17.84%)
3 kids	29,606	56,129	85,735
	(4.28%)	(8.06%)	(6.17%)
Marital Status of Individual			
Married	378,088	314,629	692,717
Never Married Cohabitant	298,333	352,633	650,966
Previois Married Cohabitant	15,970	28,945	44,915
Region			
West-Germany	560,830	562,124	1,122,954
	(81%)	(80.74%)	(80.87%)
East-Germany	131,561	134,083	265,644
	(19%)	(19.26%)	(19.13%)
Age (Mean)	29.28	29.1	29.19
Age(Standard deviation)	7.78	7.72	7.75

Notes: Table 5A shows the distributional characteristics of the sample for modeling "fertility". Table summarizes the frequencies and percentages (in brackets) of each variable of interest for men, women and total sample. These variables include the cohort, education, marital status, number of children and region. For example, in the given sample of men, 67.93% have no child, 14.5% men have one child, 13.3% have 2 children and 4.28% have three children. We have capped the number of children till 3rd child. Similarly, we can interpret the distributions of other variables in the given sample and across the two genders.

Table 6A: Average Marginal Effects - Mortality Model

Dependent Variable: Likelihood/Hazard of Death		
Variables	Male	Female
Age spline (20-65)	0.000759***	0.00110***
	(5.20e-05)	(5.79e-05)
Age spline (66-80)	0.000976***	0.000992***
	(6.15e-05)	(6.73e-05)
Age spline (81-100)	0.00121***	0.00103***
	(6.09e-05)	(9.11e-05)
time	-0.000271***	-0.000300***
	(3.73e-05)	(4.55e-05)
Cohort 1915-1935 (First World War)	0.000958	0.000227
	(0.000723)	(0.00116)
Cohort 1936-1945 (Second World War)	1.67e-05	-0.00288*
	(0.00118)	(0.00162)
Cohort 1946-1963 (Baby Boomers)	0.00188	-0.00104
	(0.00174)	(0.00215)
Cohort 1964-1980 (Gen X)	0.000117	-0.00256
	(0.00228)	(0.00270)
Cohort > 1980 (Gen Y and onwards)	-0.00112	0.00128
	(0.00350)	(0.00474)
Education <HS	0.000968**	0.00269***
	(0.000406)	(0.000731)
Education >HS	-0.00176***	-0.00322***
	(0.000534)	(0.000472)
Single	0.00464***	0.00995***
	(0.000925)	(0.00136)
Widow/er	0.00131***	0.00303***
	(0.000465)	(0.000716)
Divorced	0.00404***	0.00868***
	(0.000847)	(0.00125)
Separated	0.00153	0.00439***
	(0.00161)	(0.00167)
Region (East Germany)	0.00110**	0.00120**
	(0.000448)	(0.000522)
Income	-0.00534***	-0.00706***
	(0.000534)	(0.000692)
Observations	313,920	289,271

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: Table 6A presents the average marginal effects (AMEs) after conducting a logit discrete-time regression to estimate the hazard of mortality for males and females. The AMEs indicate the average change in the probability of death for a one-unit change in each respective predictor variable, holding all other variables constant. The variables included in the analysis are Age splines (divided into three age groups: 20-65, 66-80 and 81-100), Education categories (< High School, =High School, > High School), Cohort categories (1915-1935, 1936-1945, 1946-1963, 1964-1980, and > 1980), and Region (West Germany). Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany)

Table 7A: Average Marginal Effects – Model of First Marriage

Dependent Variable: Likelihood/Hazard of First Marriage				
Variables	Male		Female	
	Model 1	Model 2	Model 3	Model 4
Age spline (18-25)	0.0253***	0.0252***	0.0199***	0.0199***
	(0.000314)	(0.000313)	(0.000236)	(0.000241)
Age spline (25-35)	0.000652***	0.000196	-0.00291***	-0.00347***
	(0.000120)	(0.000129)	(0.000169)	(0.000180)
Age spline (35-40)	-0.00737***	-0.00781***	-0.0126***	-0.0130***
	(0.000404)	(0.000408)	(0.000650)	(0.000659)
Age spline (40+)	-0.000455**	-0.00111***	-0.000361	-0.000977***
	(0.000198)	(0.000228)	(0.000248)	(0.000297)
Education <HS	-0.000737	-0.000328	0.0206***	0.0208***
	(0.00129)	(0.00119)	(0.00171)	(0.00160)
Education >HS	-0.0115***	-0.0115***	-0.0266***	-0.0269***
	(0.000715)	(0.000779)	(0.000821)	(0.000890)
Cohort 1915-1935 (FWW)	0.0262***	0.0249***	0.0194***	0.0191***
	(0.00276)	(0.00262)	(0.00256)	(0.00244)
Cohort 1936-1945 (SWW)	0.0287***	0.0273***	0.0514***	0.0506***
	(0.00282)	(0.00270)	(0.00297)	(0.00284)
Cohort 1946-1963	0.00882***	0.00754***	0.0384***	0.0381***
	(0.00250)	(0.00238)	(0.00252)	(0.00241)
Cohort 1964-1980 (Gen X)	-0.0127***	-0.0139***	0.00106	0.00118
	(0.00244)	(0.00231)	(0.00230)	(0.00220)
Cohort > 1980 (Gen Y and onwards)	-0.0281***	-0.0297***	-0.0152***	-0.0150***
	(0.00252)	(0.00240)	(0.00239)	(0.00230)
Region(West Germany)	-0.00645***	-0.00648***	-0.00769***	-0.00811***
	(0.00102)	(0.00105)	(0.00119)	(0.00127)
Observations	423,867	423,867	373,880	373,880

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: Table 7A presents the average marginal effects (AMEs) after conducting a logit discrete-time regression to estimate the hazard of first marriage for males and females. The AMEs indicate the average change in the probability of first marriage for a one-unit change in each respective predictor variable, holding all other variables constant. Model 1 and Model 2(with interaction terms) are the estimates for male and Model 3 and Model 4(with interactions) are the estimates for female. The variables included in the analysis are Age splines (divided into four age groups: 18-25, 25-35, 35-40, and 40+), Education categories (< High School, =High School, > High School), Cohort categories (1915-1935, 1936-1945, 1946-1963, 1964-1980, and > 1980), and Region (West Germany). Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany)

Table 8A: Average Marginal Effects – Model of Remarriage

Dependent Variable: Likelihood/Hazard of Remarriage				
Variables		Male		Female
	Model 1	Model 2	Model 3	Model 4
Age spline (18-30)	-0.00952***	-0.00944***	-0.00494***	-0.00493***
	(0.000820)	(0.000824)	(0.000252)	(0.000254)
Age (30+)	-0.00524***	-0.00521***	-0.00333***	-0.00326***
	(0.000168)	(0.00019)	(9.49e-05)	(0.00009)
Education <HS	-0.0150***	-0.0152***	-0.00795***	-0.00798***
	(0.00432)	(0.00431)	(0.00140)	(0.00140)
Education >HS	0.0186***	0.0186***	0.00729***	0.00753***
	(0.00413)	(0.00408)	(0.00180)	(0.00179)
Cohort (Baby boomers)	-0.0165***	-0.0166***	0.00749***	0.00753***
	(0.00453)	(0.00455)	(0.00158)	(0.00158)
Cohort (Gen X)	-0.0186***	-0.0188***	0.00622***	0.00625***
	(0.00489)	(0.00491)	(0.00162)	(0.00162)
Cohort (Gen Y and onwards)	-0.00118	-0.00180	0.00357	0.00360
	(0.0121)	(0.0120)	(0.00318)	(0.00319)
No. of children = 1	0.0176***	0.0176***	0.00395***	0.00403***
	(0.00370)	(0.00370)	(0.00141)	(0.00141)
No. of children = 2	0.0169***	0.0169***	0.00744***	0.00755***
	(0.00401)	(0.00401)	(0.00155)	(0.00155)
No. of children = 3	0.0313***	0.0313***	0.0110***	0.0110***
	(0.00534)	(0.00535)	(0.00196)	(0.00196)
Unmarried duration (0-2)	0.00778***	0.00780***	0.00105	0.00107
	(0.00283)	(0.00283)	(0.00136)	(0.00136)
Unmarried duration (1-7)	-0.000205	-0.000225	0.00152***	0.00154***
	(0.000674)	(0.000673)	(0.000295)	(0.000295)
Unmarried duration (7 onwards)	0.00568***	0.00568***	0.00255***	0.00254***
	(0.000254)	(0.000255)	(0.000104)	(0.000105)
Region,West-Germany	0.0248***	0.0247***	0.00870***	0.00868***
	(0.00299)	(0.00299)	(0.00127)	(0.00127)
Observations	58,602	58,602	130,963	130,963

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: Table 8A presents the average marginal effects (AMEs) after conducting a logit discrete-time regression to estimate the hazard of remarriage for males and females. The AMEs indicate the average change in the probability of remarriage for a one-unit change in each respective predictor variable, holding all other variables constant. Model 1 and Model 2(with interaction terms) are the estimates for male and Model 3 and Model 4(with interactions) are the estimates for female. The variables included in the analysis are Age splines (divided into two age groups: 18-30 and 30+), Education categories (< High School, =High School, > High School), Cohort categories (baby boomers, Gen X and Gen Y), Children (0, 1, 2 or 3 and more) and duration splines for those who are either divorced or widowed. Three duration splines include 0-2, 2-7 and 7 onwards. Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany)

Table 9A: Average Marginal Effects – Model of Divorce

Dependent Variable: Likelihood/Hazard of Divorce		
Variables	Male	Female
Age spline (18-30)	4.04e-05	8.94e-05
	(9.77e-05)	(7.03e-05)
Age spline (30 onwards)	-0.000236***	-0.000156***
	(2.18e-05)	(2.62e-05)
Less than High School	-0.00124***	-0.00152***
	(0.000271)	(0.000284)
More than High School	-0.000482*	-0.000243
	(0.000288)	(0.000335)
Region: West Germany	0.000615**	0.000389
	(0.000279)	(0.000303)
Cohort: Baby Boomers	0.00486***	0.00509***
	(0.000254)	(0.000267)
Cohort: >1963	0.00480***	0.00797***
	(0.000333)	(0.000375)
Marriage duration (0-2)	0.00509***	0.00707***
	(0.000344)	(0.000397)
Marriage duration (2-7)	0.000130	0.000131
	(8.33e-05)	(8.57e-05)
Marriage duration (7 onwards)	-7.68e-05***	-0.000270***
	(2.71e-05)	(3.16e-05)
No. of children = 1	-0.00268***	-0.00455***
	(0.000714)	(0.000624)
No. of children = 2	-0.00516***	-0.00558***
	(0.000397)	(0.000418)
No. of children = 3	-0.000799	-0.00271***
	(0.000872)	(0.000699)
Second or more marriages	0.0180***	0.0156***
	(0.000784)	(0.000742)
Observations	468,615	519,753

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: Table 9A presents the average marginal effects (AMEs) after conducting a logit discrete-time regression to estimate the hazard of divorce for males and females. The AMEs indicate the average change in the probability of divorce for a one-unit change in each respective predictor variable, holding all other variables constant. The variables included in the analysis are Age splines (divided into two age groups: 18-30 and 30+), Education categories (< High School, =High School, > High School), Cohort categories (baby boomers, Gen > 1963), Children (0, 1, 2 or 3 and more) and duration splines for capturing duration of marriage. Three duration splines include 0-2, 2-7 and 7 onwards. Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany)

Table 10A: Average Marginal Effects – Model of Fertility

Dependent Variable: Likelihood/Hazard of Birth			
Variables	First Child	Second Child	Third Child
Age spline (17-20)	0.0183***	0.0251***	0.00238
	(0.000704)	(0.00463)	(0.00789)
Age spline (20-25)	0.00514***	0.0156***	0.00694***
	(0.000195)	(0.000596)	(0.000683)
Age spline (25 onwards)	-0.00426***	-0.00614***	-0.00296***
	(8.12e-05)	(0.000163)	(0.000124)
Gender: female	0.0348***	-0.00754***	-0.00864***
	(0.000576)	(0.00116)	(0.000875)
age at the end of the period	-0.000278***	-0.00161***	-0.000903***
	(6.58e-05)	(0.000136)	(0.000103)
Education <HS	0.0132***	0.0181***	0.0155***
	(0.000901)	(0.00172)	(0.00117)
Education >HS	-0.00786***	0.0215***	0.0127***
	(0.000591)	(0.00147)	(0.00112)
Region: West-Germany	-0.0199***	0.0159***	0.0118***
	(0.000897)	(0.00137)	(0.000906)
Never-married Single	-0.128***	-0.0671***	-0.00405***
	(0.00132)	(0.00112)	(0.00138)
Previously married Single	-0.104***	-0.0638***	-1.53e-05
	(0.00254)	(0.00226)	(0.00207)
Cohort (Base: Birthyear < 1915)	0.00941***		
Cohort 1915-1935 (First World War)		0.00619	-0.00471
	(0.00111)	(0.00453)	(0.00434)
Cohort 1936-1945 (Second World War)	0.0171***	0.00758*	-0.0229***
	(0.00116)	(0.00451)	(0.00425)
Cohort 1946-1963 (Baby Boomers)	0.0316***	0.0144***	-0.0251***
	(0.00111)	(0.00440)	(0.00420)
Cohort 1964-1980 (Gen X)	0.0471***	0.0437***	0.000636
	(0.00113)	(0.00444)	(0.00425)
Cohort 1981-1989 (Gen Y)	0.0432***	0.0357***	-0.00940**
	(0.00145)	(0.00497)	(0.00469)
Cohort > 1989 (Gen Z)	0.00964***	0.0407***	-0.0128
	(0.00185)	(0.0106)	(0.00987)
Observations	822,341	288,687	256,405

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: Table 10A presents the average marginal effects (AMEs) after conducting a logit discrete-time regression to estimate the likelihood of first, second and third child The AMEs indicate the average change in the probability of a child for a one-unit change in each respective predictor variable, holding all other variables constant. The variables included in the analysis are Age splines (divided into three age groups: 17-20, 20-25 and 25+) Education categories (< High School, =High School, > High School), Cohort categories (1915-1935, 1936-1945, 1946-1963, 1964-1980, and > 1980), and Region (West Germany). Baseline groups: Education (HS degree); Cohort (<1915); Region (East Germany)

Table 11A: Crude divorce rate, 1970, 1995, 2019 and 2020 or latest available year

Country	1970	1995	2019		2020 (↘)
Latvia	4,6	3,1	3,1		2,7
Lithuania	2,2	2,8	3,1		2,7
Costa Rica			2,8		..
Cyprus	0,2	1,2	2,6		..
Denmark	1,9	2,5	1,8		2,7
Sweden	1,6	2,6	2,5		2,5
Finland	1,3	2,7	2,4		2,4
Luxembourg	0,6	1,8	3,1		2,3
United States	3,5	4,4	2,7		2,3
Chile			3,2		2,2
Korea	0,4	1,5	2,2		2,1
Czech Republic	2,2	3,0	2,3		2,0
Estonia	3,2	5,2	2,1		1,9
Switzerland	1,0	2,2	2,0		1,9
Australia	1,0	2,8	1,9		1,9
France		2,1	1,9	2016	..
Greece	0,4	1,0	1,8	2017	..
Israel	0,8	1,6	1,8		..
Iceland	1,2	1,8	1,6	2011	1,9
OECD-27 average	1,4	2,4	2,0		1,9
Belgium	0,7	3,5	2,0		1,8
Norway	0,9	2,4	1,9		1,8
EU-24 average			1,9		1,7
Portugal	0,1	1,2	2,0		1,7
Germany	1,3	2,1	1,8		1,7
Austria	1,4	2,3	1,8		1,7
United Kingdom	1,0	2,9	1,8		1,7
Netherlands	0,8	2,2	1,7		1,7
Spain		0,8	1,9		1,6
Turkey			1,9		1,6
Japan	0,9	1,6	1,7		1,6
New Zealand	1,1	2,6	1,7		1,5
Hungary	2,2	2,4	1,8		1,5
Slovak Republic	0,8	1,7	1,7		1,5
Poland	1,1	1,0	1,7		1,4
Bulgaria	1,2	1,3	1,6		1,3
Croatia	1,2	0,9	1,5		1,3
Romania	0,4	1,5	1,6		1,2
Italy		0,5	1,4		1,1
Slovenia	1,1	0,8	1,2		0,8
Mexico	0,6	0,4	1,3		0,7
Ireland			0,7	2017	..
Malta			0,7		0,5

Notes: Table 11A shows the crude divorce rate, 1970, 1995, 2019, 2020 or the latest year available for many countries around the world. The crude divorce rate (CDR), defined by OECD is the number of divorces during a given year per 1000 people. This measure of divorce is widely used in demography to compare the divorce rates across countries. The sources of above-mentioned crude rates for each country are available from OECD Family database.

Recent Issues

No. 398	Pantelis Karapanagiotis, Marius Liebald	Entity Matching with Similarity Encoding: A Supervised Learning Recommendation Framework for Linking (Big) Data
No. 397	Matteo Bagnara, Milad Goodarzi	Clustering-Based Sector Investing
No. 396	Nils Grevenbrock, Alexander Ludwig, Nawid Siassi	Homeownership Rates, Housing Policies, and Co-Residence Decisions
No. 395	Ruggero Jappelli, Loriana Pelizzon, Marti Subrahmanyam	Quantitative Easing, the Repo Market, and the Term Structure of Interest Rates
No. 394	Kevin Bauer, Oliver Hinz, Moritz von Zahn	Please Take Over: XAI, Delegation of Authority, and Domain Knowledge
No. 393	Michael Kosfeld, Zahra Sharafi	The Preference Survey Module: Evidence on Social Preferences from Tehran
No. 392	Christian Mücke	Bank Dividend Restrictions and Banks' Institutional Investors
No. 391	Carmelo Latino, Loriana Pelizzon, Max Riedel	How to Green the European Auto ABS Market? A Literature Survey
No. 390	Kamelia Kosekova, Angela Maddaloni, Melina Papoutsi, Fabiano Schivardi	Firm-Bank Relationships: A Cross-Country Comparison
No. 389	Stefan Goldbach, Philipp Harms, Axel Jochem, Volker Nitsch, Alfons J. Weichenrieder	Retained Earnings and Foreign Portfolio Ownership: Implications for the Current Account Debate
No. 388	Gill Segal, Ivan Shaliastovich	Uncertainty, Risk, and Capital Growth
No. 387	Michele Costola, Katia Vozian	Pricing Climate Transition Risk: Evidence from European Corporate CDS
No. 386	Alperen Afşin Gözlügöl, Wolf-Georg Ringe	Net-Zero Transition and Divestments of Carbon-Intensive Assets
No. 385	Angela Maddaloni	Liquidity Support and Distress Resilience in Bank-Affiliated Mutual Funds
No. 384	Julian Greth, Alperen Afşin Gözlügöl, Tobias Tröger	The Oscillating Domains of Public and Private Markets