MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Konrad-Zuse-Strasse 1 · D-18057 Rostock · Germany · Tel +49 (0) 3 81 20 81 - 0 · Fax +49 (0) 3 81 20 81 - 202 · www.demogr.mpg.de

MPIDR Working Paper WP 2024-023 I July 2024
https://doi.org/10.4054/MPIDR-WP-2024-023

Uncovering what matters: Family life course aspects and personal wealth in late working age

Nicole Kapelle
Carla Rowold I rowold@demogr.mpg.de

This working paper has been approved for release by: Christian Dudel (dudel@demogr.mpg.de),
Deputy Head of the Research Group Labor Demography.

Uncovering what matters: Family life course aspects and personal wealth in late working age

Nicole Kapelle, Department of Social Sciences, Humboldt-Universität zu Berlin and Nuffield College, University of Oxford [corresponding author]
Email: nicole.kapelle@hu-berlin.de
ORCID ID 0000-0001-5855-1153

Carla Rowold, Nuffield College, University of Oxford, Max Planck Institute for Demographic Research, Max Planck – University of Helsinki Center for Social Inequalities in Population Health, Rostock, Germany and Helsinki, Finland
Email: rowold@demogr.mpg.de
ORCID ID 0000-0003-0292-6096

Abstract

Background: Capturing the complexity of family life courses as predictors of later-life outcomes like wealth is challenging. Previous research has either (a) assessed a few selective but potentially irrelevant summary indicators, or (b) examined entire life course clusters without identifying specific important aspects within and between them.

Objective: To investigate which family life-course variables—encompassing variables that capture the order, duration, and timing of states and transitions—are key personal wealth predictors for Western Germans aged 50 to 59. And analyse the strength and direction of associations between relevant variables and personal wealth, and whether these differ by gender.

Methods: We used German Socio-economic panel study (SOEP) data and combined feature selection, sequence analysis tools, and regression techniques.

Results: We identified 23 family life-course variables as relevant predictors, with two—the time spent never-married, both without and with children—deemed most relevant. Most family life-course variables were negatively associated with personal wealth and characterised by

single parenthood, marital separation or early marital transitions with or without fertility transitions. The prevalence and significance of associations between these variables and personal wealth differed partly across genders. Results highlight the importance of previously concealed family life-course variables for wealth inequalities in late working age.

Contribution: We extend previous research on the nexus between family demography and wealth stratification by using a novel, data-driven approach that more effectively explores family life-course complexities by considering the 'entire' universe of variables that describe such life courses and identifying those life-course variables that are *relevant* wealth predictors.

Keywords: Family dynamics, Life course, Social stratification, Wealth, Feature Selection

Acknowledgements: This work was supported by the European Research Council under the European Union's Horizon 2020 research and innovation program under Grant Agreement 681546. Carla Rowold acknowledges financial support from the Economic and Social Research Council grant ES/P0000649/1, from Nuffield College, and was supported by grants to the Max Planck – University of Helsinki Center from the Max Planck Society, Jane and Aatos Erkko Foundation, Faculty of Social Sciences at the University of Helsinki, and Cities of Helsinki, Vantaa and Espoo. The views expressed herein are those of the authors and are not necessarily those of the funding bodies. The manuscript greatly benefited from discussions with participants of several conferences held in 2023, including the Population Association of America's (PAA) annual meeting (12th to 15th April 2023, New Orleans), the spring meeting of the Research Committee 28 on Social Stratification and Mobility (RC28) of the International Sociological Association (ISA) (24th to 26th May 2023, Paris), the annual conference of the European Consortium for Sociological Research (ECSR) (18th to 20th September 2023, Prague), and the annual conference of the Society for Longitudinal and Lifecourse Studies (SLLS) (9th to 11th October 2023, Munich). Additionally, we would like to thank participants of the research colloquium at the State Institute for Family Research at the University of Bamberg (ifb) for comments and suggestions that greatly improved the study.

Introduction

The relevance of sufficient private wealth to supplement or even replace public pensions has drastically increased over recent decades—even in countries with formerly monolithic, generous public pension systems—as a response to the financial pressures on the public pension systems generated by an ageing population. However, wealth inequalities in late working age are soaring in most OECD countries. Thus, rising personal responsibilities to secure economic living standards throughout the life course and particularly during older age cannot be met by a substantial proportion of the population. Understanding the drivers of wealth inequality is of utmost importance for policymakers and researchers alike. This understanding is crucial in reducing reliance on welfare and economic disparities, thereby maintaining social cohesion, and addressing the challenges associated with an ageing population.

Traditionally, debates and research on the drivers of wealth inequalities have focused on the role of social background and labour market position (Atkinson, 1971; Bernardi et al., 2018). More recently, a growing body of research has highlighted the relevance of the family as an important context for socio-economic stratification and wealth inequality (e.g., Halpern-Manners et al., 2015; Kapelle & Vidal, 2022; Ulker, 2008; Wilmoth & Koso, 2002; Zissimopoulos et al., 2015). Overall, such studies have identified a distinct set of life course aspects that are commonly associated with wealth advantages. These include *inter alia* continuous marriage that is entered at an average age and childbirth within wedlock with two or—depending on the cultural and historical context—a maximum of three children. Such aspects reflect a culturally and contextually idealised family life course. On the contrary, family life course patterns that break with this ‘ideal’—for instance, through a divorce or single parenthood—have often been linked to less favourable wealth outcomes.

Family life courses exhibit a high degree of complexity due to their variation among individuals. This variation encompasses the types of family events and transitions encountered throughout an individual's life, as well as the order, timing, and duration of these experiences. The analysis of the implications of this family complexity is becoming increasingly urgent, particularly when viewed through a gendered lens. Family complexity has increased, particularly for women (Van Winkle & Fasang, 2021). To analyse this complex association between family life courses and wealth, researchers have employed two strategies. First, family life courses were operationalised using a small range of summary indicators, such as age at first marriage, times married, the duration within first marriage, and the number of children, and these indicators were used to predict wealth levels (Halpern-Manners et al., 2015). Although research that followed this methodological tradition provided critical first impulses, the approach is highly selective and limited regarding the indicators and number of indicators that can be considered due to multicollinearity issues (Rowold et al., 2024). This means that the approach may not capture all *relevant* indicators of family life courses and provide an incomplete picture of the association of interest. Second, sequence analysis and cluster analysis were used to reduce the complex universe of family life course trajectories into a set of meaningful and distinct groups of life courses. Identified clusters were subsequently used within a regression framework to predict wealth in late working age (Kapelle & Vidal, 2022). While this approach offers a comprehensive, descriptive overview of the relationship between typical family life courses and wealth, it lacks specificity in identifying which particular life course aspects, within or across these groups, are most relevant for the accumulation of wealth and thus wealth levels in late working age.

Using a novel methodological approach, the present study addresses the shortcomings of previous research. Precisely, we (I) illustrate *which* family life course trajectory aspects out of a multitude of life course summary measures—focusing on both marital and fertility life course aspects—are most relevant for personal wealth at age 50 to 59 among cohorts of West Germans born between 1943 and 1967. We consider wealth at ages 50 to 59 because wealth penalties and advantages accumulate over the life course. As such, wealth inequalities should be particularly visible at pre-retirement age when wealth levels are expected to peak in anticipation of retirement (Alessie et al., 1997; Modigliani, 1988). Additionally, we (II) investigate *how* aspects identified as most relevant relate to personal wealth in late working age. Specifically, we assess the extent and direction in which “relevant” aspects predict personal wealth in late working age. Finally, we (III) examine whether the extent and direction of how family life course aspects are associated with personal wealth in late working age differ by gender.

To address our research objectives, we use data from the German Socio-Economic Panel Study (SOEP v38, 1984-2021; Goebel et al. (2019)) and apply a methodological approach that combines a feature selection algorithm and sequence analysis tools with regression techniques to identify wealth-relevant family life course aspects (i.e., variables—which are referred to as ‘features’ in machine learning—describing family life courses) and study the extent and direction to which they are associated with personal wealth of women and men in late working age (Bolano & Studer, 2020). In our analyses, we capture family life course aspects experienced between the ages of 15 to 50. Those aspects reflect the timing (i.e., at which age an event or transition took place), order (i.e., in which order events and transitions took place), and duration (i.e., how long individuals spent within a certain state) of family life course events and transitions. We focus on West Germany as an intriguing setting for our study because it has been marked by continuous cultural and institutional support for traditional family structures involving stable marriage and male breadwinner ideologies, despite widespread changes in women’s societal roles and endowments, as well as in relationship and fertility practices (Trappe et al., 2015). East Germany in contrast followed substantially different social policies regarding family structures, gender equality and wealth accumulation. Joining these two contexts in our study is thus not feasible.

Wealth accumulation and the life course

Wealth accumulation is a dynamic process that occurs through three main pathways (e.g., Keister & Moller, 2000; Killewald et al., 2017; Spilerman, 2000). First, *surplus income*, originating from labour earnings, social welfare, or assets like rent, interest, and dividends, may be saved or reinvested for exponential growth. Second, wealth may be obtained through financial transfers such as *inter vivos* transfers (i.e., transfers made during the grantor's life), inheritances, or other windfall profits. Finally, wealth may increase through capital appreciation, depending on the individual wealth portfolio and financial markets.

Wealth levels are expected to change over the life course as part of the dynamic process of wealth accumulation. Specifically, individuals commonly start with low or no wealth. Wealth then grows during working years before peaking before retirement. This may be considered the normative wealth accumulation pathway. Nevertheless, wealth levels are highly heterogeneous in late working age and financial preparedness for retirement varies drastically (Halpern-Manners et al., 2015; Hurd, 2002). These wealth inequalities can be understood as a result of age differentiation processes within a life course framework (Bernardi et al., 2019; Dannefer, 2003; O'Rand, 1996): Initial comparative advantages or disadvantages at an earlier age restrict or enhance future wealth levels and accumulation potentials. This path dependency means that individuals cumulate and compound disadvantages or advantages over the life course. Thus, wealth levels and accumulation potentials progressively differentiate between individuals as they age.

In line with path dependencies, the experience of certain life course events and transitions can generate enhancing or disruptive effects on an individual's wealth accumulation with potentially lasting and flow-on effects for future opportunities. Highly influential events or transitions are often denoted “turning points” (Abbott, 2001). However, decisive of an individual’s ability to accumulate wealth over their life course is not just whether enhancing or disruptive life course events and transitions occur at all, but also in which *order* (e.g., being married before childbirth) and at which *time* during the life course (e.g., getting married before the age of 20) they occur as well as the *duration* an individual stays within a certain state (e.g., the time spent as a single parent).

The relevance of family life courses for the accumulation of wealth

In the present study, we focus on family life course features—referring to the occurrence, timing and order of events and transitions as well as the duration within certain states—as structuring sources of opportunities or barriers to accumulate personal wealth until late working age. Parenthood and marital roles are known to affect labour income, consumption, investments, and wealth transfers which directly impact the process of wealth accumulation over the life course (e.g., Budig et al., 2012; Kapelle & Lersch, 2020; Leopold & Schneider, 2011; Lersch et al., 2017).

Linking marital status and wealth

Previous research has unequivocally found that married individuals exhibit higher wealth levels across the life course compared to never-married, divorced, and re-married individuals (e.g., Addo & Lichter, 2013; Halpern-Manners et al., 2015; Hao, 1996; Wilmoth & Koso, 2002). This marital wealth advantage has been attributed to a range of intertwined, wealth-enhancing contextual and normative factors. First, depending on the institutional context, married spouses may benefit from favourable tax rates next to other advantages such as joint insurances and pensions (Christl et al., 2023). Germany is one of the most prominent examples of marital tax benefits with its joint taxation and full income splitting (*Ehegattensplitting*) (Buslei &

Wrohlich, 2014). Second, marriage benefits are also visible in the acquisition of assets. Particularly homeownership—as one of the largest investments for the majority of households—is commonly more accessible to the married compared to non-married individuals, *inter alia*, because of high transaction costs (Thomas & Mulder, 2016). Third, social norms around marriage as a lifelong emotional and financial commitment encourage joint investments and resource integration (Knoll et al., 2012; Vogler et al., 2008), which improves economies of scale and makes investments more efficient and lucrative. Fourth, marriage entry has been shown to increase the likelihood of receiving intergenerational transfers, for instance, to support the couple with their start into a joint future (Leopold & Schneider, 2011). Finally, marriage is associated with a range of normative economic behaviours that couples may feel the need to achieve to merit marriage entry (Gibson-Davis et al., 2018). Marriage entry is therefore more likely with rising asset ownership and wealth accumulation potentials (e.g., higher levels of education or income)—particularly for men, but less so for women (Schneider, 2011; Xie et al., 2003). Thus, couples have been shown to postpone marriage until they have the economic means. On the contrary, early or very late marriage has been associated with economic disadvantage. Overall, these mentioned marriage benefits likely compound throughout the marriage as previous research has highlighted that the duration within (first) marriage is positively associated with wealth in older age (Kapelle & Lersch, 2020; Schmidt & Sevak, 2006; Zissimopoulos et al., 2015).

On the contrary, marital dissolution—in the form of separation and divorce—is detrimental to wealth in the short and long term (Addo & Lichter, 2013; Kapelle, 2022; Kapelle & Baxter, 2021; Ozawa & Lee, 2006). First, marital dissolution leads to a loss of previously discussed advantages. Second, as spouses split, they need to divide assets. Particularly larger assets such as the home are difficult to divide because spouses rarely have sufficient other assets to pay each other out or qualify for a mortgage by themselves. Thus, these assets are often liquified, which incurs substantial costs and can lead to a loss of wealth if assets have to be sold in unfavourable markets (Lersch & Vidal, 2014). Third, ex-spouses may continue to have financial ties that impede their wealth accumulation (e.g., spousal maintenance payments) with men more likely to support their ex-wife than vice versa due to a higher likelihood of children staying with mothers and women's lower economic resources. Finally, the likelihood of experiencing a marital dissolution is socially stratified with financially stressed couples more likely to dissolve (Dew, 2011; Dew et al., 2012; Eads & Tach, 2016).

In line with the previously discussed relevance of the timing, order and sequences of events, a range of aspects likely matter on how severe marital dissolution disrupts wealth accumulation until late working age. First, the age at which marital dissolution occurs may matter with earlier dissolution—when fewer assets are available and more time left until late working age for wealth accumulation—less wealth-detrimental than dissolutions at an older age (Zissimopoulos et al., 2015). Second, if divorce is followed by remarriage, some—but not all—divorce-related wealth penalties can be attenuated because certain marital advantages are re-established. Third, serial union dissolution leads to substantial wealth penalties due to the accumulation of disadvantages (Ulker, 2008; Wilmoth & Koso, 2002; Zissimopoulos et al., 2015).

Individuals may also abstain from marriage. As previously highlighted, marriage entry may be selective of economically more successful individuals. In turn, economically less well-off women and, more so, men with lower wealth accumulation potentials are more likely to stay single (Addo, 2014; Carlson et al., 2004; Gibson-Davis et al., 2005; Smock et al., 2005). While staying un-married results in a lack of all marital benefits, cohabiting individuals likely benefit from some although not all benefits—depending on the country context. However, considering the cohorts of interest and country context of the present study, long-term cohabitation was rather uncommon and only played a negligible role in the life courses of women and men (Le Goff, 2002).

Child-related wealth benefits and penalties

The anticipation or presence of children is associated with a range of wealth-enhancing mechanisms including increased overall savings incentives and changes in investment strategies including buying a family home and saving for children's education. It has also been shown to be associated with a higher likelihood of receiving intergenerational transfers (Leopold & Schneider, 2011; Lusardi et al., 2001). Despite some wealth advantages and institutional support for parents (e.g., child allowance), parenthood is also linked to a range of direct and indirect costs. Direct costs, disproportionately borne by women, include children's daily consumption expenses as well as costs for child care and education (Bradbury, 2011; Lanau, 2023). Indirect financial costs of childbirth particularly emerge for women as they take care-related career breaks, experience restrictions regarding employment opportunities, and reduced working hours (Budig & England, 2001), which translates into mothers' but not fathers' lower income and reduced wealth accumulation (Lersch et al., 2017). Particularly early childbirth has been linked to high indirect costs for women, for instance through effects on educational attainment or career entry (Gough & Noonan, 2013).

The interconnectedness of parenthood and marital status in the accumulation of wealth

How childbearing is linked to the accumulation of wealth is closely intertwined with the transition in and out of marriage or the absence of such a transition. Parenthood within marriage makes financial transfers from husband to wife more likely (Eickmeyer et al., 2019) and increases the ability to save jointly for children as married parents benefit from marital wealth premiums (Grinstein-Weiss et al., 2008). Particularly within the cohort and context understudy, parenthood within marriage was socially and institutionally encouraged while childbirth out-of-wedlock was stigmatised (Le Goff, 2002). As such, childbirth commonly followed marriage or parents married shortly after childbirth if they were partnered.

While parents who marry after childbirth can fully benefit from all marital advantages, lasting parenthood out-of-wedlock, which in the case of our cohort and context mostly refers to single-parenthood without a cohabiting partner, is accompanied by a lack of marital wealth premiums as well as reduced financial transfers, particularly between parents as previous research has illustrated that non-and under-payment of child-support or spousal alimony are common issues (Manning et al., 2003). Single parenthood can also be the result of marital dissolution, which is similarly accompanied by a loss of marital premiums and partner's support in addition to any other divorce-related wealth penalties. Both pathways into single parenthood are highly selective of economically less resourceful individuals (Upchurch et al., 2002). As children commonly stay with the mother (Walper et al., 2021), the costs of single parenthood are over-proportionally experienced by women affecting women's wealth accumulation. Costs of single parenthood may be reduced through (re)marriage. However, particularly children are a barrier and postponing factor for mothers but not fathers to (re)marry (Di Nallo, 2019).

Child-related costs and benefits for the accumulation of wealth are not only intertwined with marital status but also the number of children. Generally, consumption costs increase with the number of children. This is accompanied by a simultaneous increase in the barriers for parents particularly mothers to engage in the labour market. Despite some institutional support for parents, the threshold at what point child-related costs may outweigh benefits in turn differs by marital status with higher thresholds within marriage where costs can be covered jointly than outside of marriage (Zissimopoulos et al., 2015). Thus, while moderate fertility is socially and institutionally supported, high fertility, including associated child-related costs and social perceptions around it, may lead to penalties. Similarly, low fertility and childlessness are also regarded as a violation of social norms and values, which can in turn evoke adverse wealth-relevant repercussions (e.g., fewer intergenerational transfers, discrimination, etc.). Low fertility and childlessness as well as high fertility may also be socially stratified along wealth accumulation potentials.

The present study

Under the umbrella of the life course framework, a large body of research has sought to connect earlier life course aspects with later-life outcomes. Bernardi et al. (2019) emphasise that such research often conceptualises potentially relevant predictors as measures at a single point in time. This approach can be misleading as it overlooks the explanatory potential of other, potentially correlated predictors occurring earlier or later in life. Moreover, it often fails to fully capture all life course dimensions, including not only the occurrence but also the timing, order, and duration of events or transitions. At the same time, Bernardi et al. (2019) acknowledge the increasing difficulty, both theoretically and methodologically, of capturing a large universe of interconnected predictors referring to an extended life course timespan. As such, assessing which variables are truly relevant becomes progressively challenging. These challenges also impact the study of the association between family life courses and wealth in late working age. As illustrated in the previous sections, various family life course aspects might influence wealth accumulation.

Using a new methodological approach, the present study navigates the challenges of capturing life course complexity. Considering the ‘entire’ universe of variables that describe family life courses, we use a data-driven approach to identify those variables that are the most relevant wealth predictors. However, the present study does not seek to highlight causal links between specific family life course aspects and wealth. Instead, the aim is to provide a detailed description of the wealth-relevant features and explore the direction and magnitude of those features’ associations with respondents’ wealth in late working age. While we focus on family life courses, we acknowledge that these family life courses are closely interlinked with other life domains (e.g., work or educational trajectories). However, the interconnectedness of these domains is beyond the scope of the present study and should be considered an important avenue for future research.

Data and methods

Data and sample

The empirical analyses were based on longitudinal (prospective and retrospective) data from the German Socio-Economic Panel Study (SOEP v38, 1984-2021; Goebel et al. (2019)). The SOEP is a large and nationally representative study that tracks individuals living in eligible households annually since 1984. The dataset was suitable for our research purposes because it (i) collects information on a comprehensive set of wealth measures at the personal level in survey years 2002, 2007, 2012 and 2017, and (ii) contains detailed information on marital, and childbearing histories over entire respondents' life courses.

We selected respondents who were aged 50 to 59 at any time between 2002 and 2017, who lived in West Germany in 1989, and who provided complete retrospective marital and fertility histories from ages 15 to 50. Further, we restricted the sample to observations that had valid wealth information in at least one of the years in which wealth was measured (i.e., 2002, 2007, 2012, or 2017). For respondents who were captured more than once with a valid wealth wave between the ages of 50 to 59, we randomly selected one of the waves. This was necessary to reduce bias in the feature selection approach as selected features would have been more strongly driven by those respondents with several waves. We further excluded 181 individuals (women: 151, men: 30) who experienced widowhood before age 50 as the share of person-year-spells in widowhood was too small to be included. This led to an exclusion of 2.5 % of the sample. As men's retrospective fertility data have been collected less frequently within the

SOEP, the sample included fewer men than women. To prevent our results from being biased and driven by family states that are particularly important for women, we randomly selected the same number of women as men, establishing gender balance in our sample. This resulted in the deletion of 1,307 women. Robustness checks confirmed that women in the deleted group were similar to women in the sample group regarding family characteristics and other key measures. Our final sample consisted of 5,702 respondents with 2,851 women and the same number of men.

Personal net wealth

Our outcome variable of interest was a measure of personal net wealth, which was defined as the sum of all personally owned assets minus liabilities. Asset components in the SOEP include property assets, tangible and financial assets, private pensions, business assets and collectables, while liabilities refer to consumer credits or mortgage debts. For each household member aged 17 and older, SOEP personal wealth data have been collected in a three-step process: (1) a filter question is used to assess ownership of a certain wealth component; (2) the total market value of held wealth components is recorded; and (3) for jointly held wealth components, respondents are asked to provide the share they co-owned. Our outcome measure thus explicitly includes the personal share of any assets and liabilities that were owned with other individuals. As liabilities were subtracted from assets, respondents may hold negative net worth (around 5 % of the sample). We also included individuals with 0 net wealth (about 11 % of individuals in the sample). Figure A.2 in the appendix illustrates an overview of the distribution of personal wealth by gender and Table A.1 provides more detail on the wealth distribution by gender, showing wealth levels at the mean, median, 25th and 75th percentile.

We adjusted personal net wealth for inflation using the consumer price index of the German Statistical Office and top- and bottom-coded the extreme 0.1 % of reported wealth measures to reduce the influence of outliers. These adjustments were done on the entire SOEP sample and thus before the sample restriction. Although wealth—similar to income—is often transformed for analyses to account for the skewness of the data (Killewald et al., 2017), we used the original, absolute wealth distribution in Euros. This allowed us to better capture broader inequalities between respondents, which would be distorted through transformations. We used the imputed wealth variables provided by the SOEP team and took the mean value across the provided five imputed wealth sets (Grabka & Westermeier, 2015).

Family life courses

We generated yearly indicators capturing the succession of family states over time from age 15 to 50 (see Figure 1 for seven example sequences extracted from our sample). To this end, we used biographical information on respondents' marital and fertility histories that were collected retrospectively and prospectively within the SOEP. The life course states combined the two channels (partnership and fertility) into a total of 11 family states that could be experienced over the 46 years: (1) single and childless, (2) single with child(ren), (3) married and childless, (4) married with one child, (5) married with two children, (6) married with three children, (7) married with four or more children, (8) previously married (but currently unmarried) and childless, (9) previously married (but currently unmarried) with child(ren), (10) remarried and childless, and (11) remarried with child(ren). Note that our analysis did not differentiate whether children were residing with the respondent at any specific time. Instead, the 'child(ren)' indicator was based on whether respondents reported having a child or children.

>>>>>> FIGURE 1 ABOUT HERE <<<<<<

Due to methodological and theoretical reasons, the life course states and thus complexity that could be captured had to be restricted. First, we only disaggregated the number of children within marriage but distinguished only between childless individuals and parents for the other three categories to ensure sufficient cell sizes. Second, cohabitation episodes were not explicitly captured in our study but included under the categories ‘single’ or ‘previously married’. This limitation stems from the data collection approach of the SOEP, which emphasises retrospective partnership histories focusing on the formation and dissolution of marriages and does not collect detailed information on non-marital cohabitations. From a conceptual standpoint, we argue that the omission of cohabitation information is of minor concern for our study. For the cohort and context under study, individuals born between 1943 and 1967 in West Germany, cohabitation–especially long-term cohabitation and childbearing outside of marriage–was socially undesirable and discouraged (Le Goff, 2002). Consequently, such states are likely to be rare in our data. Reinforcing this perspective, Rowold et al. (2024) demonstrated that in a cohort of slightly older West Germans and Italians, less than one per cent of person-year spells between the ages of 18 to 65 were spent in non-married cohabitation.

Overview of methodological approach

In line with Bolano and Studer (2020), our empirical approach broadly followed three steps: First, we automatically extracted a wide set of features of the family life course trajectory.[1] Second, out of this large set we selected the features that were most relevant wealth predictors. To this end, we used a data-mining feature selection algorithm, the Boruta algorithm (see next section). Third, the selected, smaller set of family life course features (i.e., results from Boruta) were used as covariates in an Ordinary Least Square (OLS) regression model to estimate the direction and strength of each feature’s association with wealth in late working age.

For steps two and three of our approach, analyses were adjusted for a small set of three key covariates. First, we added a continuous measure of respondents’ age at the considered measurement of wealth ranging from age 50 to 59. Second, we added two dummy variables indicating whether respondents experienced a divorce or the death of their marital partner

[1] Note that ‘features’ is the terminology used within machine learning to refer to variables. We use the terms ‘features’ and ‘variables’ interchangeably.

between the age of 50 and their considered wealth observation. Both were rather rare occurrences. In our sample of 5’702 respondents, 207 experienced a divorce (women: 83, men: 124) and 70 the death of their partner (women: 53, men: 17).

The analytical steps one and two are discussed in more detail in the next section.

Describing the process of trajectory’s feature extraction and selection

Starting with an automatic extraction of family life course variables (i.e., features), we used the *WeightedCluster* R package by Studer (2013) and extracted 205 variables from our trajectories. These features covered the duration (I), ordering (II), and timing (III) of family life states.

Duration was measured as the sum of years spent in the different life course states. Using the sequences in Figure 1 as illustrative examples, individual 8, for instance, spent three years in the status ‘married and having one child’. *Ordering* (‘pattern’) captured whether and how often individuals experienced a period of the same state or sub-sequences of states. For example, individual 8 had one distinct period being ‘previously married and having children’, and individuals 74 and 76 had two periods in this state. The variable ‘Pattern PMC…RMC’ captures the subsequence of remarriage(s) following the state previously married and having children. For example, for individual 39 this subsequence occurred three times as the first period of being previously married is followed by two periods of being remarried and the second period of being previously married is followed by one period of being remarried. We operationalised *timing* as the age range in which a state started in five-year periods. For example, focusing on the start of ‘being married and having no children’, we can see that individual 28 experienced this between ages 25 to 29. Figure A.3 in the appendix provide more examples.

Next, the aim was to reduce the rather long list of 205 automatically extracted features to a parsimonious but theoretically relevant subset of features. To this end, we followed suggestions by Bolano and Studer (2020) and selected features that had a meaningful interpretation in relation to our association of interest, that were interpretable on their own, and that were theoretically relevant. We removed features that occurred for less than 2% of our sample and features that were highly correlated (r>.9). These restrictions resulted in a reduced set of 71 features.

To identify those family life-course features out of the list of 71 features that were most relevant for personal wealth in late working age, we applied the Boruta algorithm using the *Boruta* R package (Kursa & Rudnicki, 2010). Following Bolano and Studer (2020), Boruta was deemed appropriate because it considers various possible forms as well as potential interactions between features for assessing their importance for the outcome and it has recently been shown to perform best, especially for low-dimensional data sets (Degenhardt et al., 2019). The Boruta algorithm is a wrapping method based on a random forest (RF) approach. Essentially, an RF approach ranks all features according to their importance, but it usually does not determine a cut-off point distinguishing between important and unimportant features. Boruta provides such a cut-off point by generating shadow features which are randomised versions of the original features. The original features then compete against the shadow features and are only confirmed as being important for the outcome if they perform better in the random forest than the best-performing shadow feature. We provide a detailed description of the feature selection approach using the Boruta algorithm in the appendix.

Relevant family life course predictors of wealth in late working age

In total, 23 family life course features were selected as relevant predictors of personal net wealth in late working age by the Boruta algorithm. These features are displayed in Figure 2 ranked by their mean importance.[2] Additional summary descriptives (i.e., median, minimum

[2] The mean importance in the Boruta algorithm is the average significance of a feature in making accurate predictions, as measured by its contribution to reducing uncertainty or error in the model.

and maximum feature importance, means for continuous features and the prevalence for categorical features by gender) are provided in Table A.2 and Table A.3. As illustrated in Figure 2, features relating to the different life course dimensions—duration, order, and timing—were equally often selected as relevant features (duration: 8, order: 7, timing: 8). Additionally, features referring to being married (first time or higher order) or being un-married (never married single or previously married) were equally often selected, 12 and 11 times respectively.

Two features stand out as the most important wealth predictors according to the Boruta algorithm: the duration spent in the statuses ‘single without children’ or ‘single with child(ren)’. The high relevance of the duration spent within the status of ‘single with child(ren)’ may highlight the persistently emphasised economic disadvantages associated with single parenthood (Maldonado & Nieuwenhuis, 2015; Sierminska, 2018). The time spent within the status ‘single without children’ could be relevant for several reasons. It might be linked to a postponement of marriage and parenthood in favour of career progression and economic advantage (Amuedo-Dorantes & Kimmel, 2005; Uecker & Stokes, 2008). Conversely, time spent within this status could also reflect the selection of economically less successful individuals into singlehood (Gibson-Davis et al., 2005; Jalovaara & Fasang, 2017).

>>>>>> FIGURE 2 ABOUT HERE <<<<<<

Three other results are noteworthy when considering the features selected as relevant: First, several life course variables related to ‘previously married (but currently unmarried) with child(ren)’ are among the top 10 wealth predictors. Marital dissolution after having three children ('pattern..M2C...M3C...PMC.’) is the third most important family life course proxy for predicting wealth. The duration and the number of being previously married and having children are relatively important as well (‘duration.PMC’ and ‘pattern…PMC’). Second, and related to the first point, the features related to marital dissolution all pertained to the dissolution of marriages involving parents, while features of marital dissolution without children were not identified as relevant wealth predictors. This likely points to the high vulnerability of parents if their marriage breaks down. Third, concerning the timing of family events over the life course, the Boruta algorithm mostly identified periods at early ages as important predictors of personal wealth. This may suggest that family transitions at early or very early ages (up to age 24) lay the groundwork for wealth accumulation throughout the life course. For example, marrying or having already two children and being married between ages 20-24 ('Age20_24.Start_MNC', 'Age20_24.Start_M2C') are among the top-5 wealth predictors. In contrast, family transitions in mid to late adulthood appear to play a less substantial role with only a few selected features referring to transitions at or after the age of 30.

Overall, while all life course concepts are among the selected features, it is the duration within the statuses of 'single without children' and 'single with child(ren)' that were deemed the most important wealth predictors.

Association of relevant family life course features and wealth in late working age

While the feature selection approach indicated which family life course features are the most relevant predictors of personal wealth in late working age, it did not indicate the magnitude and direction of the association between relevant features and personal wealth. To further dive into this aspect and address our second research question, we moved to a regression approach. To this end, we ran a range of different regressions. We started with bivariate regressions (i.e., 23 regressions in total) and moved to stepwise multiple regressions that added up to eleven of the most relevant features into a single regression as explanatory variables for personal net wealth.

Figure 3 displays regression coefficients for the 23 bivariate associations. Note that we illustrate coefficients for duration features (in purple) on a different scale than coefficients for order and timing features (in green) due to the different variable types (i.e. continuous vs categorical). As can be seen in Figure 3, the majority of features are negatively associated with personal wealth in late working age. Negative associations are most often found for features relating to single parenthood, marital separation or comparatively early marital transitions with and without fertility transitions. Particularly strong negative associations can be found for features that refer to very early entry into parenthood (at ages 15-19, within or outside of marriage), the occurrence of marital separation with three children, the duration spent as an un-married parent after marital separation as well as experiencing this state more than once, and the duration spent as a single (never-married) parent.[3] Figure 3 also highlights five features that are significantly positively associated with personal net wealth in late working age. Specifically, these features capture the duration as a childless singlehood as well as being a married parent of three children and experiencing the state of being a married parent of one child, and relatively late transitions to the first marriage without and with children.

>>>>>> FIGURE 3 ABOUT HERE <<<<<<

As a next step, we moved to multiple regressions that added features in a stepwise manner. For ease of comparison between the models, we illustrate the results in Figure 4 (see Table A.4 for the full regression results). We started with the two most relevant features—the duration spent as a single parent and the duration spent single without children—according to the mean

[3] Note that ‘single’ may include a few cases of never-married, cohabitation. However, as previously highlighted never-married, cohabitation was a rather uncommon occurrence and mostly short-lived experience for the cohort and context under study.

importance of the feature selection analyses and moved on to the top five, top nine and, finally, top eleven features. We refrained from adding further features into the regression due to the decreasing importance of features and increasing multicollinearity issues.

Overall, results for these multiple regression models mirrored our bivariate results from Figure 3 and associations remained robust and in the expected direction. Some differences between the bivariate and the multiple regressions should be highlighted: the effect of the duration of single parenthood ('duration S1C') is slightly more negative in the multiple than the bivariate result. Effects of 'Pattern M2C-M3C-PMC' and 'Duration PMC' are less negative in the multiple regression. And finally, a range of features are no longer statistically significant ('Pattern PMC', 'Duration RMC', 'Age 15-19 Start M1C', 'Age 15-19 Start S1C').

Focusing on the differences between the stepwise multiple models, we can see that some dependencies become visible as we started adding additional features. For instance, the association of the duration as a childless single becomes less relevant and statistically non-significant once we start adding the timing variables referring to whether the transition to first marriage—at that point of the transition without children—takes place at age 20 to 24 and whether the birth of a second child within marriage takes place within this age bracket. Thus, not adding the timing of marriage or any other features, the association between the years spent as a childless single and wealth is ambiguous because longer singlehood likely also captures the positive association between late marriage and wealth—as can be seen in the bivariate associations (Figure 3).

Also, the effect of experiencing the pattern of being previously married with children is likely absorbed by the duration of this pattern. To assess this, we re-ran the regression of the top nine features dropping the duration indicator as a robustness check (results not shown but available upon request from authors). Indeed, this leads to the expected results of the pattern categories becoming negatively correlated with wealth. Specifically, experiencing the pattern of being previously married with children once or twice is associated with around 40’000€ and 165’000€ less personal wealth, respectively, once we no longer adjust for the duration in this status. This robustness check also reveals that the duration of remarried with children is no longer negative although still non-significant once we exclude the indicator for the duration within the status of previously married with children. All other variables remain stable across the full ‘top-9’ model and our robustness check.

>>>>>> FIGURE 4 ABOUT HERE <<<<<<

Exploring potential variations by gender

The average effects presented in the previous results section might differ importantly for women and men. We thus move to our third research aim: exploring the extent to which the magnitude and direction of the association between specific features and personal wealth in late working age differ across genders. To this end, we ran gender-specific regression models. Note that cell sizes were not sufficient for all variables once disaggregated by gender (see Table A.5). For instance, transitions to single parenthood and having the first child within marriage at ages 15 to 19 were extremely uncommon for men, while this pattern was more common among women. Thus, we did not include these two features in the regressions for men. Similarly, only a few men reported ‘being previously married but currently unmarried’ twice (‘Pattern_PMC’). We modified this variable for men to reflect experiencing this status at least once. Figure 5 shows results of gender-specific regression models for the top 9 and 11 features for women and the top 9 features for men (see Table A.6 for the full regression results).

A range of gender differences are noteworthy (Figure 5). First, early marriage of childless individuals (age 20-24) and early transitions to a second child within marriage (age 20-24) are substantially more negatively associated with wealth for men, hinting at disadvantage or selectivity associated with comparatively early family transitions for men. Second, the time spent as a childless single is negatively associated with men's wealth while we find no substantial effect for women. This is in line with the idea that economically less well-off men with lower wealth accumulation potentials are more likely to stay single (Addo, 2014; Carlson et al., 2004; Gibson-Davis et al., 2005). Third, the duration of being previously married with children is substantially more negatively associated with wealth for women, highlighting the lasting disadvantage of marital dissolution for wealth accumulation especially for mothers (Kapelle, 2022; Kapelle & Vidal, 2022).

For the remaining family life course features we found no substantial gender differences. This might seem surprising at least for some features. For example, the duration within single parenthood was similarly negative for both men's and women's wealth, adjusting for other features and our three main covariates. Overall, more women than men experienced this status and stayed on average longer within this status (Table A.3, Table A.5 and Table A.7) in line with previous research (Walper et al., 2021). However, men who experience this status might be more highly selective than women, which may explain the similarly negative effects for women and men for this feature.

Underlying gender differences might have also biased our feature selection model because features that are oppositely associated with wealth by gender might cancel each other out. As a result, gender-specific features may be concealed as relevant wealth predictors in our main model. As a supplementary analysis, we addressed this issue by running our feature selection approach separately for women and men. As a result, these analyses, however, focus on inequalities within each group rather than overall wealth inequality (see Table A.8 and Figure A.4 for the results).

In total, 13 and 14 features were deemed relevant in the separate models for women and men, respectively. Thus, a smaller set of features was selected as relevant in the gender-specific models compared to our main model. Those features that were selected, however, were mostly also confirmed within the main model. Compared to our main model where features from all life course dimensions were equally often selected as relevant, duration features were more often selected as relevant within both gender-specific models. The top five features within each gender group almost exclusively refer to the duration within certain states. Only among men, one pattern feature (‘Pattern..M1C’) was also selected within the top 5. Hence, as for the main model, two duration features were deemed as most important within each gender model: for women, this was the duration being previously married (‘Duration.PMC’) as well as the duration of single parenthood (‘Duration.S1C’). Overall features related to the life course state of being a previously married parent were more important wealth predictors for women compared to men. For men, the two most important features were the duration within childless singlehood (‘Duration.SNC’)—ranked the second most important feature in the main model—followed by the duration within remarriage with children (‘Duration.RMC’).

Thus, while some of the features that were selected in the main model appeared relevant for wealth inequalities within gender groups, the order of relevance differed between the main and gender-specific models. This also highlights that some features are more important to explain within-group inequalities but might be less important for overall inequalities.

Conclusion and discussion

In this study, we explored which family life course features (i.e., variables describing family life courses) are most relevant for personal net wealth at age 50 to 59 among cohorts of West Germans born between 1943 and 1967. We delved into the complexity of family life courses, considering not only the occurrence of transitions or events but also their timing, order, and duration. Furthermore, we examined the strength and direction of the associations between relevant features and personal wealth. This investigation aimed to determine whether selected features predict wealth positively or negatively and the magnitude of these effects. Finally, we also explored the extent to which gender stratifies these processes.

Our theoretical and empirical approach was informed by notions embedded within the life course framework. This framework has been used widely to predict individuals' later life outcomes. Due to empirical limitations, life course studies often focus on predictive factors at a single point in time and employ a limited set of predictors to represent entire life courses. However, the longer lives are studied, the more challenging it becomes to select a concise yet comprehensive set of predictors for later life outcomes. To navigate these challenges, the current study adopted a novel empirical approach. Using longitudinal (prospective and retrospective) data from the German Socio-Economic Panel Study, we first automatically extracted a broad set of family life course features. Subsequently, we employed the Boruta algorithm to identify features that are statistically relevant predictors of personal wealth. Finally, we applied a regression framework to assess the direction and strength of each selected feature's association with wealth in late working age. This approach allowed us to incorporate theoretical perspectives on the complexity of family life courses as predictors of wealth accumulation, thereby reflecting these concepts within our empirical methodology and understanding wealth levels in late working age.

Overall, our methodological approach identified 23 features that were deemed relevant predictors of wealth and that were differently associated with wealth in the overall sample population and to some degree by gender. A range of results are particularly noteworthy. *First*, features that were deemed relevant wealth predictors were diverse, highlighting the importance of being more aware of the complexity of family life courses when considering the link between life courses and later-life outcomes. This diversity was also reflected in the fact that all three life course dimensions—timing, order and duration of events and transitions—were selected as relevant.

Second, although all three dimensions appeared across the selected features, the duration spent within certain states was persistently identified as the most important. Thus, it is not only relevant whether and at what time adverse or beneficial transitions are experienced, but particularly how long individuals spend within these states.

Third, focusing on the timing of family transitions (i.e., age at the transition), we showed that comparatively (very) early life course events and transitions were deemed highly important wealth predictors for our cohort of interest. Regression results confirmed that these (very) early events and transitions were mostly negatively associated with wealth in late working age. This may hint at selection effects (i.e., economically less stable individuals transition earlier) but also at potential adverse effects for educational and career outcomes with early transitions inhibiting prolonged education and career advancements (Amuedo-Dorantes & Kimmel, 2005; Uecker & Stokes, 2008). Thus, policies that aim at reducing wealth inequalities may want to focus particularly on the causes and consequences of early family transitions. This may even become more relevant for more recent cohorts where norms around the “appropriate” age of family transitions have shifted and early family transitions may be deemed even more undesirable and selective.

Fourth, among the relevant ‘pattern’ features, it was patterns that ended with ‘being previously married with children’ that were most often selected. The duration spent within this state and whether the transition into this state took place at age 35 to 39 were also deemed relevant. As illustrated within the regressions, features characterised by being previously married (i.e., having experienced a separation or divorce) predicted wealth at a late working age negatively. Thus, and in line with previous research, our results highlight the tremendous negative effects of marital dissolution on wealth outcomes (Kapelle, 2022; Kapelle & Vidal, 2022). Considering persistently high divorce rates, discussions and interventions need to focus on how divorce wealth penalties—particularly for parents and more so for mothers—can be reduced and economic self-reliance after divorce can be strengthened.

Finally, some gender differences were prevalent. While early transitions were deemed particularly relevant, the definition of early differed between genders. For women, it was transitions at age 15 to 19 while it was transitions at age 20 to 24 for men. This is in line with observations that men tend to experience family transitions at slightly older ages than women (Ortega, 2014). Additionally, longer time spent as a childless single was identified as particularly adverse for men, hinting at selectivity in the marriage market for economically more successful men into marriage (Xie et al., 2003). Lastly, life course experiences related to marriage dissolutions of parents were more important wealth predictors for mothers, and partly also only negatively associated with wealth for them. This highlights the detrimental consequences particularly mothers face after divorces regarding their financial safety net.

Four notable limitations of the current study need to be highlighted. First, the SOEP respondents' cohabitation histories were not recorded retrospectively. Although we argue that the absence of cohabitation data in our sequences is not a significant concern for our study—given the social undesirability and discouragement of cohabitation within the studied cohort and context (Le Goff, 2002)—it should be noted that cohabitation may be a crucial factor in future research, particularly when examining more recent cohorts or more liberal contexts. Second, and connected to the previous point, although our study identifies wealth-relevant family life course features for cohorts of West Germans born between 1943 and 1967, it remains unclear whether the identified features are similarly important wealth predictors for later or earlier cohorts or other contexts. This is an important avenue for future research. Third, our methodological approach is, at least for the time being, limited in the number of life course channels it can consider. Although our focus was on the family as a context for stratification, it is important to acknowledge that family life courses are often closely intertwined with labour market trajectories. We were unable to explicitly explore this interconnectedness in this study due to methodological restrictions, but as methods advance further future research may want to consider this interconnectedness more thoroughly. Finally, survey data on wealth, including those used in our study, are subject to several limitations commonly found in previous research. Issues of misreporting or nonresponse are frequent due to the sensitivity and complexity of such data (e.g., Grabka & Westermeier, 2015; Riphahn & Serfling, 2005). Additionally, respondents are required to report their share of potentially jointly held wealth for the collection of personal wealth data. This process can be susceptible to errors, as the clarity of property rights may not be apparent to each individual, and their perceived ownership may not correspond with legal ownership (Joseph & Rowlingson, 2012). However, the SOEP data are uniquely valuable in providing wealth information fully at the personal level, coupled with detailed family histories. This has enabled us to examine gender differences more appropriately than would have been possible with other survey data.

Overall, this study provides a thorough description of (I) which family life course features are relevant for wealth levels in late working age and thus the accumulation of wealth over the life course for cohorts of West Germans born between 1943 and 1967, (II) the magnitude and direction in which these features contribute to explaining wealth in late working ages, and finally (III) the gendered nature of the association between relevant life course features and wealth. Understanding these aspects is crucial for the expanding body of literature on wealth inequality, family dynamics, and gender inequalities. While our study does not investigate causal relationships, our findings hint at potential causal connections and thus can contribute to the understanding of the potential causes and consequences of wealth inequalities. Our research results offer valuable insights for policy discussions on ways to mitigate rising wealth inequalities and to create an environment that enables individuals to build an economic safety net. Ensuring economic self-reliance throughout the life course independent of family lives, and particularly in older age, is of tremendous importance for current and upcoming generations in the context of an ageing population, rising economic inequalities, and increasing welfare expenditures.

References

Abbott, A. (2001). *Time matters: On theory and method*. Chicago , IL: University of Chicago Press.

Addo, F. R. (2014). Debt, cohabitation, and marriage in young adulthood. *Demography*, *51*(5), 1677-1701. https://doi.org/10.1007/s13524-014-0333-6

Addo, F. R., & Lichter, D. T. (2013). Marriage, marital history, and black–white wealth differentials among older women. *Journal of Marriage and Family*, *75*(2), 342-362. https://doi.org/10.1111/jomf.12007

Alessie, R., Lusardi, A., & Aldershof, T. (1997). Income and wealth over the life cycle: Evidence from panel data. *Review of Income and Wealth*, *43*(1), 1-32. https://doi.org/10.1111/j.1475-4991.1997.tb00198.x

Amuedo-Dorantes, C., & Kimmel, J. (2005). The motherhood wage Gap for women in the United States: The importance of college and fertility delay. *Review of Economics of the Household*, *3*(1), 17-48. https://doi.org/10.1007/s11150-004-0978-9

Atkinson, A. B. (1971). The distribution of wealth and the individual life-cycle. *Oxford Economic Papers*, *23*(2), 239-254.

Bernardi, F., Boertien, D., & Geven, K. (2018). Childhood family structure and the accumulation of wealth across the life course. *Journal of Marriage and Family*, *81*(1). https://doi.org/10.1111/jomf.12523

Bernardi, L., Huinink, J., & Settersten, R. A. (2019). The life course cube: A tool for studying lives. *Advances in Life Course Research*, *41*. https://doi.org/10.1016/j.alcr.2018.11.004

Bolano, D., & Studer, M. (2020). The link between previous life trajectories and a later life outcome: A feature selection approach. In LIVES Swiss Centre of Expertise in Life Course Research (Ed.), *LIVES Working paper Nr. 82*. Lausanne.

Bradbury, B. (2011). Child costs. In A. Ben-Arieh, F. Casas, I. Frønes, & J. E. Korbin (Eds.), *Handbook of child well-being* (pp. 1483–1507). Dordrecht, the Netherlands: Springer.

Budig, M. J., & England, P. (2001). The wage penalty for motherhood. *American Sociological Review*, *66*(2), 204-225. https://doi.org/10.2307/2657415

Budig, M. J., Misra, J., & Boeckmann, I. (2012). The motherhood penalty in cross-national perspective: The importance of work–family policies and cultural attitudes. *Social Politics: International Studies in Gender, State & Society*, *19*(2), 163-193. https://doi.org/10.1093/sp/jxs006

Buslei, H., & Wrohlich, K. (2014). Besteuerung von Paaren: Das Ehegattensplitting und seine Alternativen [Taxation of couples: The marital tax splitting and its alternatives]. In *DIW Roundup: Politik im Fokus, No. 21*. Berlin, Germany: DIW.

Carlson, M., McLanahan, S., & England, P. (2004). Union formation in fragile families. *Demography*, *41*(2), 237-261. https://doi.org/10.1353/dem.2004.0012

Christl, M., De Poli, S., & Ivaškaitė-Tamošiūnė, V. (2023). Does it pay to say" I do"? Marriage bonuses and penalties across the EU. *Journal of European Social Policy*, *33*(3), 317–336. https://doi.org/10.1177/09589287231159492

Dannefer, D. (2003). Cumulative advantage/disadvantage and the life course: Cross-fertilizing age and social science theory. *The Journals of Gerontology: Series B*, *58*(6), S327-S337. https://doi.org/10.1093/geronb/58.6.S327

Degenhardt, F., Seifert, S., & Szymczak, S. (2019). Evaluation of variable selection methods for random forests and omics data sets. *Briefings in Bioinformatics*, *20*(2), 492-503. https://doi.org/10.1093/bib/bbx124

Dew, J. (2011). The association between consumer debt and the likelihood of divorce. *Journal of Family and Economic Issues*, *32*(4), 554-565. https://doi.org/10.1007/s10834-011-9274-z

Dew, J., Britt, S., & Huston, S. J. (2012). Examining the relationship between financial issues and divorce. *Family Relations*, *61*(4), 615-628. https://doi.org/10.1111/j.1741-3729.2012.00715.x

Di Nallo, A. (2019). Gender gap in repartnering: The role of parental status and custodial arrangements. *Journal of Marriage and Family*, *81*(1), 59-78. https://doi.org/10.1111/jomf.12527

Eads, A., & Tach, L. (2016). Wealth and inequality in the stability of romantic relationships. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(6), 197-224. https://doi.org/10.7758/rsf.2016.2.6.10

Eickmeyer, K. J., Manning, W. D., & Brown, S. L. (2019). What's mine is ours? Income pooling in American families. *Journal of Marriage and Family*, *81*(4), 968-978. https://doi.org/10.1111/jomf.12565

Gibson-Davis, C. M., Edin, K., & McLanahan, S. (2005). High hopes but even higher expectations: The retreat from marriage among low-income couples. *Journal of Marriage and Family*, *67*(5), 1301-1312. https://doi.org/10.1111/j.1741-3737.2005.00218.x

Gibson-Davis, C. M., Gassman-Pines, A., & Lehrman, R. (2018). “His” and “Hers”: Meeting the economic bar to marriage. *Demography*, *55*(6), 2321-2343. https://doi.org/10.1007/s13524-018-0726-z

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

Gough, M., & Noonan, M. (2013). A review of the motherhood wage penalty in the United States. *Sociology Compass*, *7*(4), 328-342. https://doi.org/https://doi.org/10.1111/soc4.12031

Grabka, M. M., & Westermeier, C. (2015). *Editing and multiple imputation of item non-response in the wealth module of the German Socio-Economic Panel* (SOEP Survey Papers, Series C - Data Documentation, No. 272). Berlin, Germany: DIW. http://panel.gsoep.de/soep-docs/surveypapers/diw_ssp0272.pdf

Grinstein-Weiss, M., Hun Yeo, Y., Zhan, M., & Charles, P. (2008). Asset holding and net worth among households with children: Differences by household type. *Children and Youth Services Review*, *30*(1), 62-78. https://doi.org/10.1016/j.childyouth.2007.06.005

Halpern-Manners, A., Warren, J. R., Raymo, J. M., & Nicholson, D. A. (2015). The impact of work and family life histories on economic well-being at older ages. *Social Forces*, *93*(4), 1369-1396. https://doi.org/10.1093/sf/sov005

Hao, L. (1996). Family structure, private transfers, and the economic well-being of families with children. *Social Forces*, *75*(1), 269-292. https://doi.org/10.2307/2580765

Hurd, M. D. (2002). Portfolio holdings of the elderly. In L. Guiso, M. Haliassos, & T. Jappelli (Eds.), *Household portfolios* (pp. 431–472). Cambridge, MA, United States: MIT Press.

Jalovaara, M., & Fasang, A. E. (2017). From never partnered to serial cohabitors: Union trajectories to childlessness. *Demographic Research*, *36*, 1703+. https://doi.org/10.4054/DemRes.2017.36.55

Joseph, R., & Rowlingson, K. (2012). Her house, his pension? The division of assets among (ex-)couples and the role of policy. *Social Policy and Society*, *11*(1), 69-80. https://doi.org/10.1017/S147474641100042X

Kapelle, N. (2022). Time cannot heal all wounds: Wealth trajectories of divorcees and the married. *Journal of Marriage and Family*, *84*(2), 592-611. https://doi.org/10.1111/jomf.12824

Kapelle, N., & Baxter, J. (2021). Marital dissolution and personal wealth: Examining gendered trends across the dissolution process. *Journal of Marriage and Family*, *83*(1), 243-259. https://doi.org/10.1111/jomf.12707

Kapelle, N., & Lersch, P. M. (2020). The accumulation of wealth in marriage: Over-time change and within-couple inequalities. *European Sociological Review*, *36*(4), 580–593. https://doi.org/10.1093/esr/jcaa006

Kapelle, N., & Vidal, S. (2022). Heterogeneity in family life course patterns and intra-cohort wealth disparities in late working age. *European Journal of Population*, *38*(1), 59-92. https://doi.org/10.1007/s10680-021-09601-4

Keister, L. A., & Moller, S. (2000). Wealth inequality in the United States. *Annual Review of Sociology*, *26*, 63-81. https://doi.org/10.1146/annurev.soc.26.1.63

Killewald, A., Pfeffer, F. T., & Schachner, J. N. (2017). Wealth inequality and accumulation. *Annual Review of Sociology*, *43*(1), 379-404. https://doi.org/10.1146/annurev-soc-060116-053331

Knoll, M. A. Z., Tamborini, C. R., & Whitman, K. (2012). I do … want to save: Marriage and retirement savings in young households. *Journal of Marriage and Family*, *74*(1), 86-100. https://doi.org/10.1111/j.1741-3737.2011.00877.x

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, *36*(11), 1 - 13. https://doi.org/10.18637/jss.v036.i11

Lanau, A. (2023). First the children, then the employed: Deprivation and intra-household inequality in Europe. *Journal of Poverty*, *27*(4), 331-349. https://doi.org/10.1080/10875549.2022.2065561

Le Goff, J.-M. (2002). Cohabiting unions in France and West Germany: Transitions to first birth and first marriage. *Demographic Research*, *7*, 593-624. https://doi.org/10.4054/DemRes.2002.7.18

Leopold, T., & Schneider, T. (2011). Family events and the timing of intergenerational transfers. *Social Forces*, *90*(2), 595-616. https://doi.org/10.1093/sf/sor030

Lersch, P. M., Jacob, M., & Hank, K. (2017). Parenthood, gender, and personal wealth. *European Sociological Review*, *33*(3), 410-422. https://doi.org/10.1093/esr/jcx046

Lersch, P. M., & Vidal, S. (2014). Falling out of love and down the housing ladder: A longitudinal analysis of marital separation and home ownership. *European Sociological Review*, *30*(4), 512-524. https://doi.org/10.1093/esr/jcu055

Lusardi, A., Cossa, R., & Krupka, E. L. (2001). Savings of young parents. *The Journal of Human Resources*, *36*(4), 762-794. https://doi.org/10.2307/3069641

Maldonado, L. C., & Nieuwenhuis, R. (2015). Family policies and single parent poverty in 18 OECD countries, 1978–2008. *Community, Work & Family*, *18*(4), 395-415. https://doi.org/10.1080/13668803.2015.1080661

Manning, W. D., Stewart, S. D., & Smock, P. J. (2003). The complexity of fathers' parenting responsibilities and involvement with nonresident children. *Journal of Family Issues*, *24*(5), 645-667. https://doi.org/10.1177/0192513x03252573

Modigliani, F. (1988). The role of intergenerational transfers and life cycle saving in the accumulation of wealth. *The Journal of Economic Perspectives*, *2*(2), 15-40. https://doi.org/10.1257/jep.2.2.15

O'Rand, A. M. (1996). The precious and the precocious: Understanding cumulative disadvantage and cumulative advantage over the life course. *Gerontologist*, *36*(2), 230-238. https://doi.org/10.1093/geront/36.2.230

Ortega, J. A. (2014). A Characterization of world union patterns at the national and regional level. *Population Research and Policy Review*, *33*(2), 161-188. http://www.jstor.org/stable/43671594

Ozawa, M. N., & Lee, Y. (2006). The net worth of female-headed households: A comparison to other types of households. *Family Relations*, *55*(1), 132-145. https://doi.org/10.1111/j.1741-3729.2006.00362.x

Riphahn, R. T., & Serfling, O. (2005). Item non-response on income and wealth questions. *Empirical Economics*, *30*(2), 521-538. https://doi.org/10.1007/s00181-005-0247-7

Rowold, C., Struffolino, E., & Fasang, A. E. (2024). Life-Course-Sensitive Analysis of Group Inequalities: Combining Sequence Analysis With the Kitagawa–Oaxaca–Blinder Decomposition. *Sociological Methods & Research*, *Online first*. https://doi.org/10.1177/00491241231224226

Schmidt, L., & Sevak, P. (2006). Gender, marriage, and asset accumulation in the United States. *Feminist Economics*, *12*(1-2), 139-166. https://doi.org/10.1080/13545700500508445

Schneider, D. (2011). Wealth and the marital divide. *American Journal of Sociology*, *117*(2), 627-667. https://doi.org/10.1086/661594

Sierminska, E. (2018). The 'wealth-being' of single parents. In R. Nieuwenhuis & L. C. Maldonado (Eds.), *The triple bind of single-parent families: Resources, employment and policies to improve wellbeing* (1 ed., pp. 51-80). Bristol: Bristol University Press. https://doi.org/10.2307/j.ctt2204rvq.9

Smock, P. J., Manning, W. D., & Porter, M. (2005). "Everything's there except money": How money shapes decisions to marry among cohabitors. *Journal of Marriage and Family*, *67*(3), 680-696. https://doi.org/10.1111/j.1741-3737.2005.00162.x

Spilerman, S. (2000). Wealth and stratification processes. *Annual Review of Sociology*, *26*(1), 497-524. https://doi.org/10.1146/annurev.soc.26.1.497

Studer, M. (2013). WeightedCluster library manual: A practical guide to creating typologies of trajectories in the social sciences with R. In *LIVES Working Papers 24*. Lausanne, Switzerland: LIVES.

Thomas, M. J., & Mulder, C. H. (2016). Partnership patterns and homeownership: A cross-country comparison of Germany, the Netherlands and the United Kingdom. *Housing Studies*, *31*(8), 935-963. https://doi.org/10.1080/02673037.2016.1164832

Trappe, H., Pollmann-Schult, M., & Schmitt, C. (2015). The rise and decline of the male breadwinner model: Institutional underpinnings and future expectations. *European Sociological Review*, *31*(2), 230-242. https://doi.org/10.1093/esr/jcv015

Uecker, J. E., & Stokes, C. E. (2008). Early marriage in the United States. *Journal of Marriage and Family*, *70*(4), 835-846. https://doi.org/https://doi.org/10.1111/j.1741-3737.2008.00530.x

Ulker, A. (2008). Wealth holdings and portfolio allocation of the elderly: The role of marital history. *Journal of Family and Economic Issues*, *30*(1), 90-108. https://doi.org/10.1007/s10834-008-9139-2

Upchurch, D. M., Lillard, L. A., & Panis, C. W. A. (2002). Nonmarital childbearing: Influences of education, marriage, and fertility. *Demography*, *39*(2), 311-329. https://doi.org/10.1353/dem.2002.0020

Van Winkle, Z., & Fasang, A. E. (2021). The complexity of employment and family life courses across 20th century Europe. More evidence for larger cross-national differences but little change across 1916–1966 birth cohorts. *Demographic Research*, *44*, 775-810. https://www.jstor.org/stable/27032934

Vogler, C., Lyonette, C., & Wiggins, R. D. (2008). Money, power and spending decisions in intimate relationships. *The Sociological Review*, *56*(1), 117-143. https://doi.org/10.1111/j.1467-954X.2008.00779.x

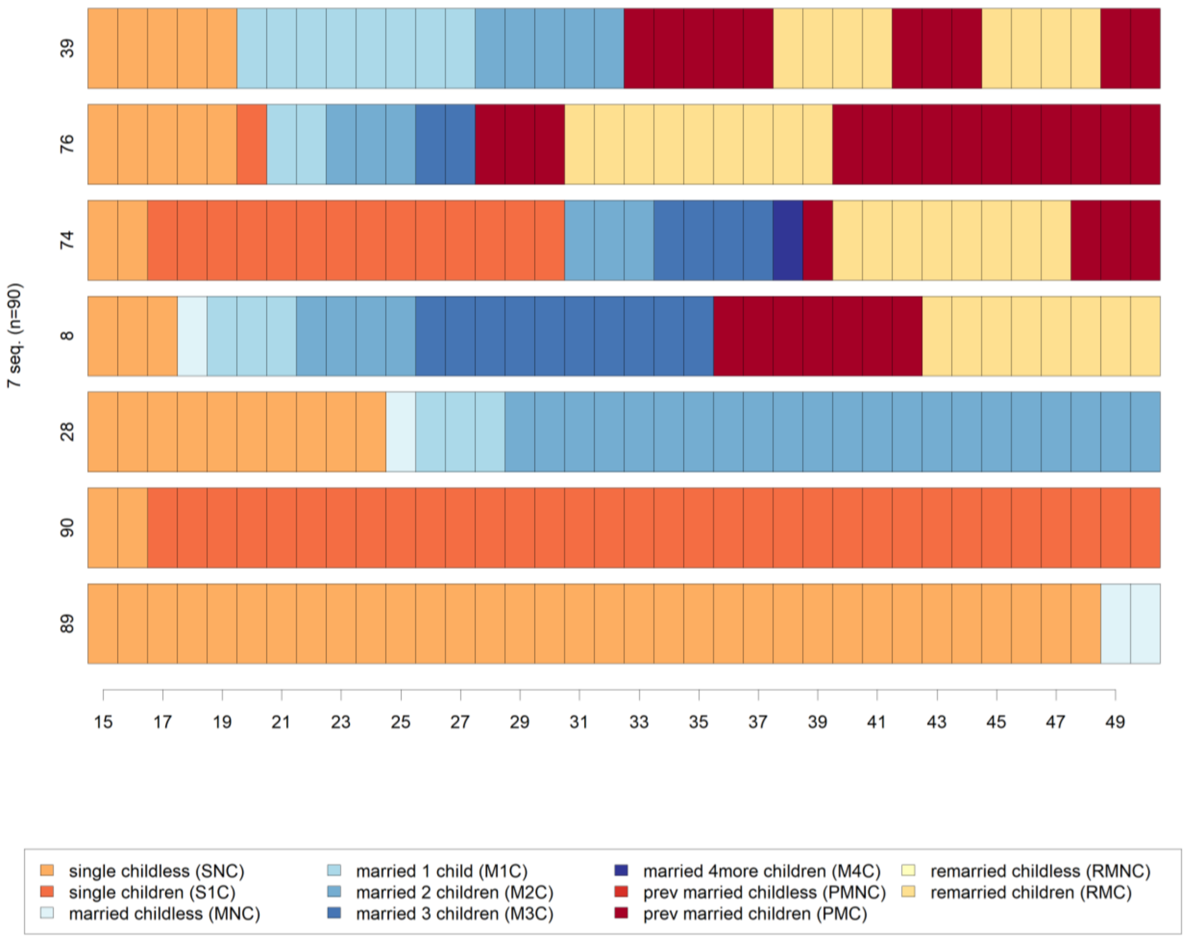
Walper, S., Entleitner-Phleps, C., & Langmeyer, A. N. (2021). Shared physical custody after parental separation: Evidence from Germany. In L. Bernardi & D. Mortelmans (Eds.), *Shared physical custody: Interdisciplinary insights in child custody arrangements* (pp. 285-308). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-68479-2_13

Wilmoth, J., & Koso, G. (2002). Does marital history matter? Marital status and wealth outcomes among preretirement adults. *Journal of Marriage and Family*, *64*(1), 254-268. https://doi.org/10.1111/j.1741-3737.2002.00254.x

Xie, Y., Raymo, J. M., Goyette, K., & Thornton, A. (2003). Economic potential and entry into marriage and cohabitation. *Demography*, *40*(2), 351-367. https://doi.org/10.1353/dem.2003.0019

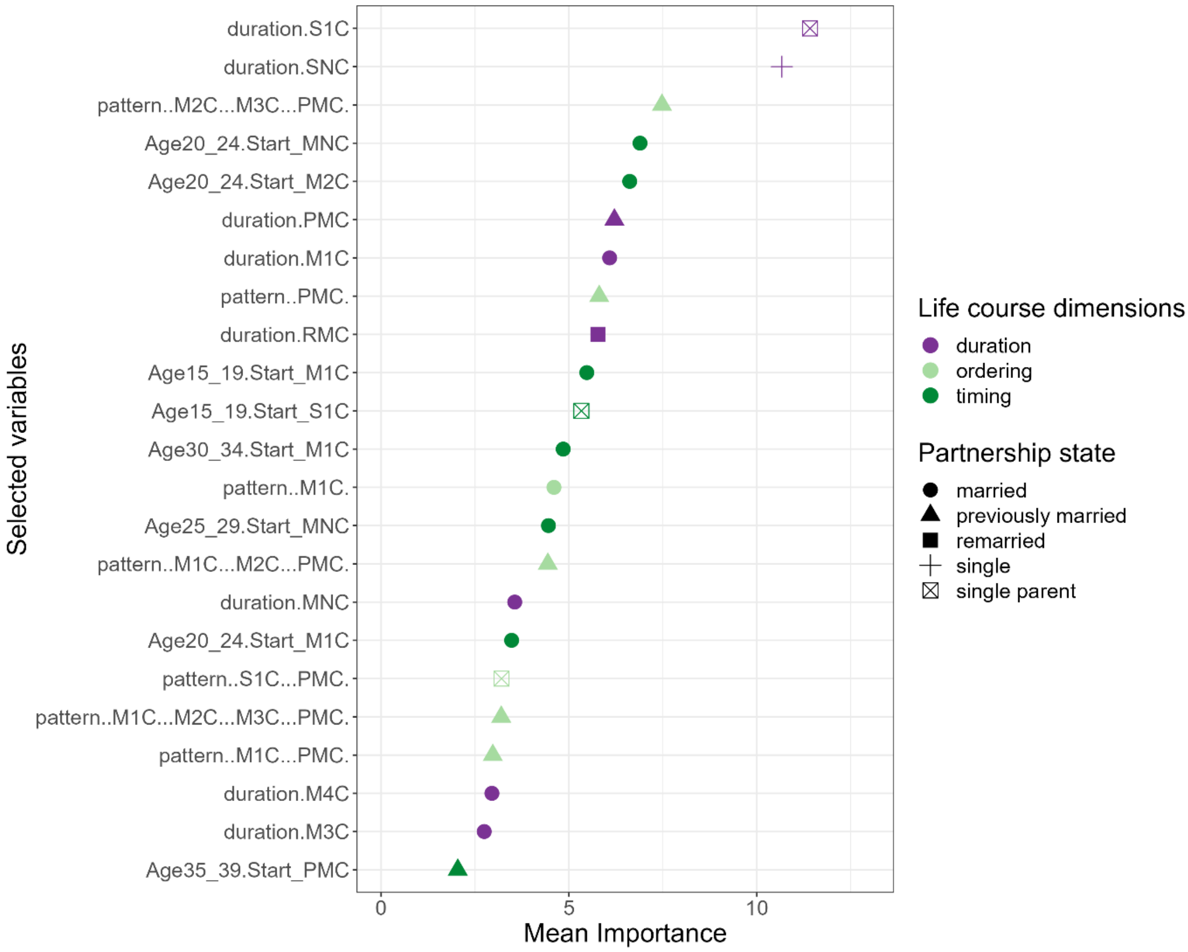
Zissimopoulos, J. M., Karney, B. R., & Rauer, A. J. (2015). Marriage and economic well being at older ages. *Review of Economics of the Household*, *13*(1), 1-35. https://doi.org/10.1007/s11150-013-9205-x

Figure 1 Example family sequences based on the analytical sample.



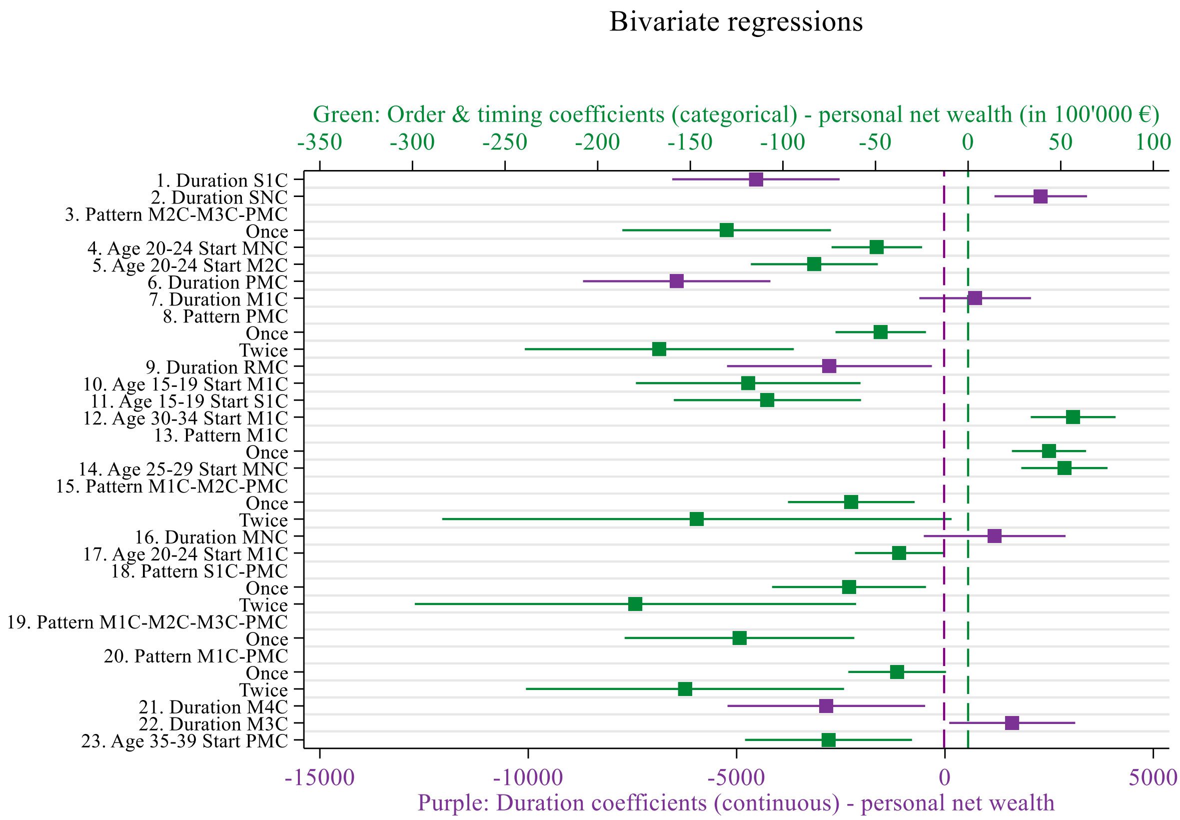
Notes: SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Figure 2 Features selected as relevant predictors of personal net wealth in late working age ranked by their mean importance according to the Boruta algorithm.



Notes: Feature importance is measured as z-scores. SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Figure 3 Bivariate regression results.

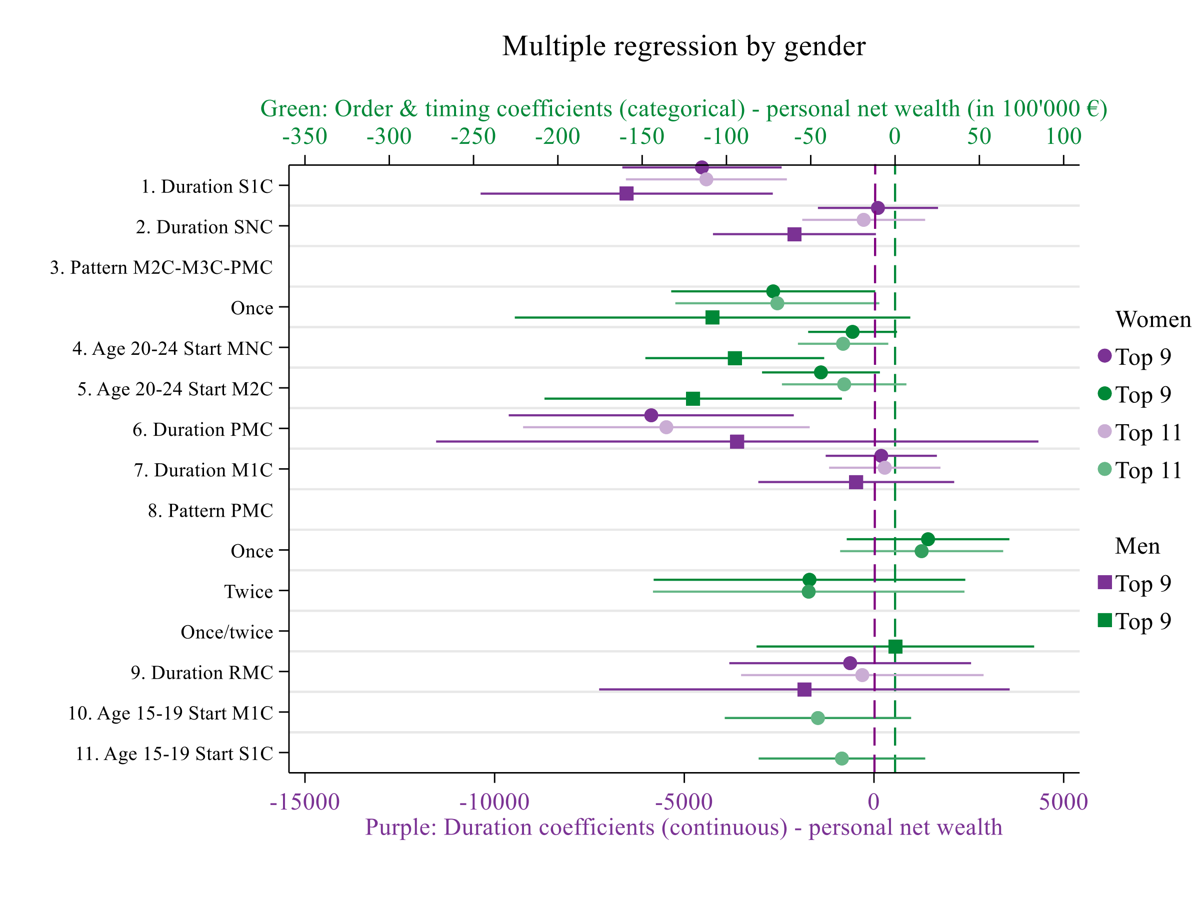


Notes: Whiskers indicate 95% confidence intervals. SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Figure 4 Stepwise multiple regression results.

Notes: Whiskers indicate 95% confidence intervals. SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Figure 5 Stepwise multiple regression results by gender.



Notes: Whiskers indicate 95% confidence intervals. SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

APPENDIX

Feature selection using the Boruta Algorithm

Feature selection is a critical process in the field of machine learning and data analysis. It involves automatically selecting a subset of ‘relevant’ features (i.e., variables) from the input data for use in model construction, tailored to the specific research question one aims to answer. The primary goal of feature selection is to improve the model's performance by eliminating unnecessary noise from the input data, resulting in a simpler, faster, more effective, and potentially more accurate model (Saeys et al., 2007).

There are two main approaches to feature selection. The ‘minimum optimum feature selection approach’ aims to find the minimal optimal subset of features that are sufficient for a model to predict the target variables. The ‘all relevant feature selection approach’ seeks to identify all relevant features contributing to the prediction of the target variable, offering insights into potential causal relationships for observed behaviours (Degenhardt et al., 2019).

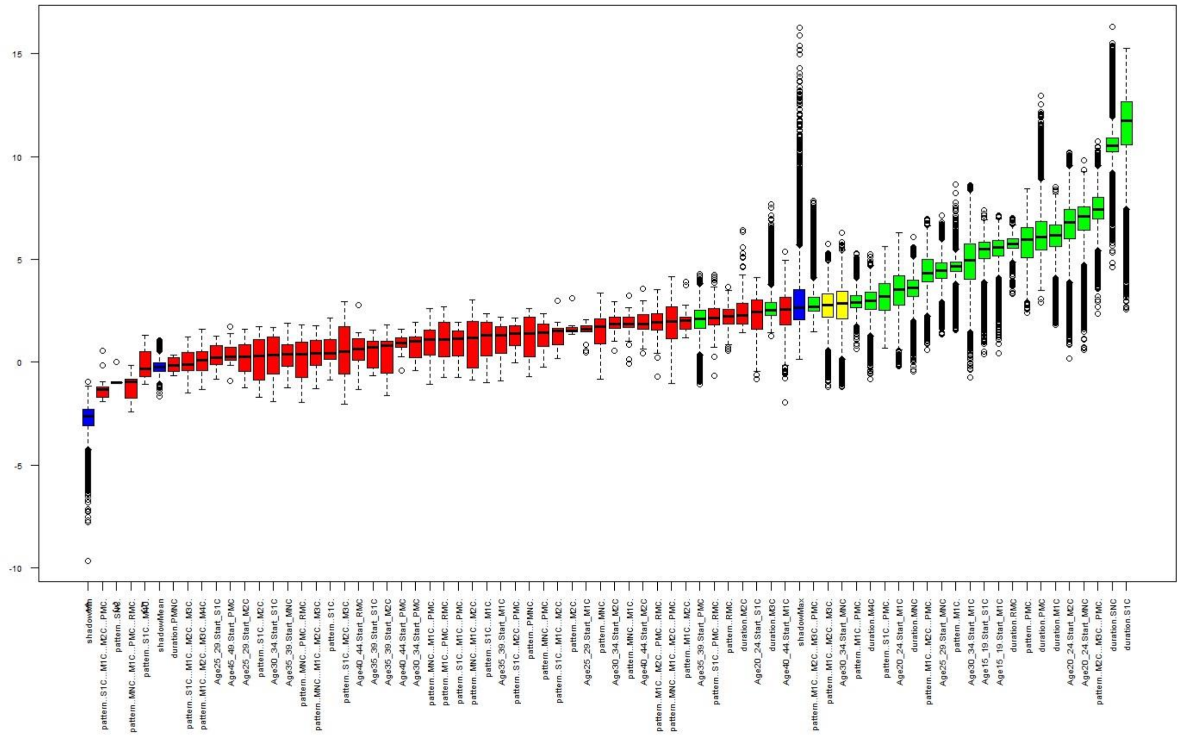
Three distinct types of feature selection methods can be distinguished: filter, wrapper, and embedded methods. Wrapper methods are particularly advantageous in their approach, as they directly analyse how subsets of variables perform within a specific predictive model, optimizing the selection based on actual model performance. This contrasts with filter methods, which independently assess features using statistical measures but do not account for model-specific interactions. While embedded methods like LASSO integrate feature selection into model training, wrapper methods offer a more focused evaluation, specifically optimising features in relation to the model's predictive accuracy, thus potentially yielding superior results in certain applications (Bolón-Canedo et al., 2014; Saeys et al., 2007).

In our study, we applied the Boruta algorithm, an 'all relevant feature' wrapper approach built around a random forest classification (Kursa et al., 2010; Kursa & Rudnicki, 2010). The Boruta algorithm aims to capture all relevant features in our dataset concerning our outcome variable—personal net wealth. This algorithm has a range of advantages—as already outlined in the main manuscript—that make it particularly suitable for our present study: its comprehensive approach ensures that no significant predictor is overlooked, its robustness against overfitting is crucial for handling our complex data, and its ability to deal with non-linear relationships and interactions is key given the complex and intertwined ways life course variables might be associated with wealth. Additionally, it has recently been shown to perform best among a set of selected algorithms, especially for low-dimensional data sets (Degenhardt et al., 2019).

The Boruta algorithm is a multi-step process (Kursa et al., 2010; Kursa & Rudnicki, 2010). First, it expands the dataset by adding 'shuffled' duplicates of all features, effectively doubling the dataset's number of features. These added features, known as shadow features, are randomised versions of the original features and have no meaningful relationship with the response variable. Second, a random forest classifier is trained on this expanded dataset, and feature importance is evaluated. A random forest classifier is a versatile machine-learning model that constructs multiple decision trees and aggregates their results for improved accuracy and control of overfitting. Essentially, the results from multiple feature importance evaluations are combined to decide on the relevance of a feature, reducing the likelihood of error. Feature importance is commonly assessed using Mean Decrease Accuracy, expressed in standardized Z-scores (i.e., the mean of accuracy loss divided by the standard deviation of accuracy loss). The Maximum Z-Score Among the Shadow Attributes (MZSA) is calculated and used to determine whether each original feature's Z-score exceeds the MZSA. Original features with Z-scores significantly exceeding those of the MZSA are deemed important and retained, while the rest are pruned. Finally, the algorithm iterates the second step, removing ‘unimportant’ features across different iterations, and stops either when all features are deemed significant or insignificant, or when a specific number of iterations is reached.

Our Boruta algorithm results are depicted in Figure A.1. The Figure presents boxplots of feature importance. The blue boxplots represent the minimum, average, and maximum Z-scores of the shadow features. Meanwhile, the red, yellow, and green boxplots correspond to the Z-scores of rejected, tentative and confirmed features, respectively.[4] A total of 23 features (depicted in green) are determined to have higher variable importance compared to the best-performing shadow feature, as indicated by Z-scores significantly exceeding those of the MZSA. These 23 features are considered in detail in the main manuscript.

Figure A.1 Boruta feature selection results sorted by median variable importance.



Notes: Feature importance is measured as z-scores.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

---

[4] Note that the boxplots are ordered based on their median Z-score, although decisions are based on whether features scored significantly higher or lower than the MZSA in a run.

Additional Tables

Table A.1 Summary statistics: Personal net wealth in 10’000 Euro

		Mean	Median	p25	p75	Min	Max	N
Personal	Total	19.30	10.95	1.55	23.47	-113.11	425.81	5702
net wealth	Men	23.09	12.80	2.91	27.65	-113.11	409.80	2851
	Women	15.52	9.10	0.88	19.91	-92.62	425.81	2851

Notes: Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Table A.2 Boruta feature selection results.

Features	Mean imp.	Min. imp.	Median imp.	Max. imp.	norm Hits
1. Duration.S1C	11.42	2.57	11.76	15.27	0.9942
2. Duration.SNC	10.67	4.61	10.55	16.31	0.9958
3. Pattern..M2C...M3C...PMC.	7.48	2.36	7.44	10.74	0.9805
4. Age20_24.Start_MNC	6.89	0.58	7.09	9.81	0.9622
5. Age20_24.Start_M2C	6.62	0.20	6.79	10.20	0.9525
6. Duration.PMC	6.21	2.90	6.11	12.95	0.9563
7. Duration.M1C	6.08	1.82	6.17	8.53	0.9506
8. Pattern..PMC.	5.80	2.41	5.97	8.45	0.9418
9. Duration.RMC	5.77	3.31	5.77	7.01	0.9519
10. Age15_19.Start_M1C	5.47	0.42	5.59	7.14	0.9294
11. Age15_19.Start_S1C	5.33	0.39	5.48	7.40	0.9177
12. Age30_34.Start_M1C	4.85	-0.73	4.95	8.60	0.8572
13. Pattern..M1C.	4.60	1.58	4.66	8.65	0.8849
14. Age25_29.Start_MNC	4.45	1.35	4.47	7.13	0.8690
15. Pattern..M1C...M2C...PMC.	4.44	0.62	4.34	6.99	0.8529
16. Duration.MNC	3.56	-0.46	3.61	6.08	0.7264
17. Age20_24.Start_M1C	3.47	-0.18	3.53	6.32	0.6667
18. Pattern..S1C...PMC.	3.20	0.70	3.18	5.64	0.6234
19. Pattern..M1C...M2C...M3C...PMC.	3.20	1.50	2.70	7.87	0.5773
20. Pattern..M1C...PMC.	2.97	0.64	2.91	5.30	0.5887
21. Duration.M4C	2.95	-0.84	3.01	5.23	0.5760
22. Duration.M3C	2.74	1.26	2.52	7.70	0.4900
23. Age35_39.Start_PMC	2.04	-1.08	2.12	4.30	0.2820

Notes: Mean, Min., Median and Max. refer to the mean, minimum, median and maximum importance of the estimated models. Hits refer to the share each feature had a higher importance than the MZSA. Feature importance is measured as z-scores.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Table A.3 Summary statistics of family life course features selected through the feature selection approach

Features			Men	Women
Order (prevalence)	Pattern_M2C_M3C_PMC	Not Such Pattern	98.11	96.77
		Once	1.89	3.23
	Pattern_PMC	Not Such Pattern	86.04	79.34
		Once	12.91	18.66
		At least twice	1.05	2.00
	Pattern_M1C	Not Such Pattern	28.17	25.96
		Once	71.83	74.04
	Pattern_M1C_M2C_PMC	Not Such Pattern	94.18	90.35
		Once	5.51	9.12
		Twice	0.32	0.53
	Pattern_S1C_PMC	Not Such Pattern	95.62	93.62
		Once	3.86	5.79
		At least twice	0.53	0.60
	Pattern_M1C_M2C_M3C_PMC	Not Such Pattern	98.35	97.44
		At least once	1.65	2.56
	Pattern_M1C_PMC	Not Such Pattern	88.60	82.95
		Once	10.56	15.71
		At least twice	0.84	1.33
Timing (prevalence)	Age20_24_Start_MNC	No	90.00	78.36
		Yes	10.00	21.64
	Age20_24_Start_M2C	No	96.81	88.64
		Yes	3.19	11.36
	Age15_19_Start_M1C	No	99.75	95.83
		Yes	0.25	4.17
	Age15_19_Start_S1C	No	98.81	94.77
		Yes	1.19	5.23
	Age30_34_Start_M1C	No	77.31	85.48
		Yes	22.69	14.52
	Age25_29_Start_MNC	No	79.80	84.71
		Yes	20.20	15.29
	Age20_24_Start_M1C	No	89.62	76.53
		Yes	10.38	23.47
	Age35_39_Start_PMC	No	96.60	95.26
		Yes	3.40	4.74
Duration in years (mean)	Duration S1C		1.19	1.56
	Duration SNC		15.63	11.58
	Duration PMC		0.96	1.85
	Duration M1C		4.30	5.06
	Duration RMC		0.88	1.00
	Duration MNC		2.61	2.75
	Duration M4C		0.79	1.02
	Duration M3C		2.49	2.78

Notes: Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Table A.4 Multiple regression: Adding features as explanatory variables in a stepwise fashion to predict personal net wealth in late working age.

	Top 2 B/(SE)	Top 5 B/(SE)	Top 9 B/(SE)	Top 11 B/(SE)
1. Duration: S1C	-4333.28*** (1024.71)	-5421.44*** (1038.08)	-5807.00*** (1056.62)	-5650.40*** (1067.21)
2. Duration: SNC	2174.96*** (566.45)	231.43 (640.82)	-431.24 (683.26)	-699.95 (692.29)
3. Pattern: M2C_M3C_PMC (ref.: no such pattern)				
Once		-112507.60*** (28911.72)	-88627.82** (31150.71)	-85295.99** (31167.65)
4. Timing: Aged 20-24 at start MNC		-59543.97*** (13553.43)	-62586.71*** (13590.72)	-67194.31*** (13706.87)
5. Timing: Aged 20-24 at start M2C		-81126.46*** (18683.61)	-83538.62*** (19103.23)	-65856.91** (20303.26)
6. Duration: PMC			-5204.85** (1982.15)	-4770.41* (1990.95)
7. Duration: M1C			-275.35 (726.41)	-189.48 (727.19)
8. Pattern: PMC (ref.: no such pattern)				
Once			9397.04 (23401.34)	5552.32 (23448.80)
Twice			-95141.04* (47652.59)	-95998.35* (47636.31)
9. Duration: RMC			-356.16 (1573.22)	-70.40 (1576.51)
10. Timing: Aged 15-19 at start M1C				-61817.43 (34098.96)
11. Timing: Aged 15-19 at start S1C				-40703.33 (28332.72)
N Individuals	5702	5702	5702	5702

Notes: SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). All regressions are accounted for respondent’s age, marital dissolution at or after age 50 and widowhood at or after age 50.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

* p<.05, ** p<.01, *** p<.001

Table A.5 Cell sizes for top-11 variables with continuous variables dichotomised by gender.

Features	Women	Men
1. Duration.S1C		
No	2130	2152
Yes	721	699
2. Duration.SNC		
No	5	2
Yes	2846	2849
3. Pattern..M2C...M3C...PMC		
Never	2759	2797
Once	92	54
4. Age20_24.Start_MNC		
No	2234	2566
Yes	617	285
5. Age20_24.Start_M2C		
No	2527	2760
Yes	324	91
6. Duration.PMC		
No	2262	2453
Yes	589	398
7. Duration.M1C		
No	740	803
Yes	2111	2048
8. Pattern..PMC		
Never	2262	2453
Once	532	368
Twice	57	30
9. Duration.RMC		
No	2609	2613
Yes	242	238
10. Age15_19.Start_M1C		
No	2732	2844
Yes	119	7
11. Age15_19.Start_S1C		
No	2702	2817
Yes	149	34

Notes: Duration variables were dichotomised in the way that 0 reflected no time spent within this state and 1 reflected any time spent within this state.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Table A.6 Multiple regression: Adding features as explanatory variables in a stepwise fashion to predict personal net wealth in late working age. Disaggregated analyses by gender

	Women: Top 2 B/(SE)	Men: Top 2 B/(SE)	Women: Top 5 B/(SE)	Men: Top 5 B/(SE)	Women: Top 9 B/(SE)	Men: Top 9 B/(SE)	Women: Top 11 B/(SE)
1. Duration_S1C	-3681.81***	-4974.96**	-4211.89***	-6209.91**	-4534.19***	-6518.84***	-4417.49***
	(1028.42)	(1921.22)	(1048.68)	(1934.45)	(1070.94)	(1964.23)	(1082.02)
2. Duration_SNC	1842.63**	373.19	692.66	-1613.20	104.79	-2097.30	-271.14
	(673.89)	(935.98)	(764.73)	(1023.36)	(807.69)	(1095.32)	(826.02)
3. Pattern_M2C_M3C_ PMC (ref.: no such pattern)							
Once			-96711.68***	-126656.26*	-72260.79*	-108246.46	-69810.71*
			(28565.79)	(55665.65)	(30864.76)	(59814.83)	(30867.30)
4. Age20_24_Start_MNC			-20407.50	-96198.25***	-25172.02	-95075.96***	-30823.46*
			(13399.57)	(26858.40)	(13441.16)	(27052.40)	(13666.34)
5. Age20_24_Start_M2C			-43417.20*	-115951.79**	-43950.02*	-119735.58**	-30174.49
			(17293.23)	(44370.60)	(17823.70)	(44989.10)	(18831.81)
6. Duration_PMC					-5870.77**	-3604.08	-5472.25**
					(1916.02)	(4048.85)	(1926.76)
7. Duration_M1C					192.79	-467.29	284.49
					(747.31)	(1317.04)	(748.51)
8. Pattern_PMC (ref.: no such pattern)							
Once					19569.25		15767.14
					(24604.75)		(24664.82)
Twice					-50749.22		-51210.87
					(47138.37)		(47116.35)
Once/twice						187.71	
						(41987.42)	
9. Duration_RMC					-627.52	-1833.83	-307.47
					(1624.87)	(2759.76)	(1630.18)
10. Age15_19_Start_ M1C							-45745.58
							(28189.85)

11. Age15_19_Start_S1C							-31504.40
							(25196.14)
N Individuals	2851	2851	2851	2851	2851	2851	2851

Notes: SNC = single and childless; S1C = single with child(ren); MNC = married and childless, M1C = married with one child, M2C = married with two children, M3C = married with three children, M4C = married with four or more children, PMNC = previously married and childless, PMC = previously married with child(ren), RMNC = remarried and childless, RMC = remarried with child(ren). All regressions are accounted for respondent's age, marital dissolution at or after age 50 and widowhood at or after age 50. Regressions for the top-11 features are only conducted for women because of the small cell sizes for men for features 10 and 11.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

* p<.05, ** p<.01, *** p<.001

Table A.7 Descriptive statistics for top-11 categorical variables by gender and over the pooled sample.

Features	Men mean/(SE)/min/max	Women mean/(SE)/min/max	Total mean/(SE)/min/max
1. Duration_S1C	1.19	1.56	1.37
	(3.92)	(4.88)	(4.43)
	0.00	0.00	0.00
	32.00	34.00	34.00
2. Duration_SNC	15.63	11.58	13.61
	(8.08)	(7.50)	(8.06)
	0.00	0.00	0.00
	36.00	36.00	36.00
3. Pattern..M2C...M3C...PMC			
Never	0.98	0.97	0.97
Once	0.02	0.03	0.03
4. Age20_24_Start_MNC	0.10	0.22	0.16
5. Age20_24_Start_M2C	0.03	0.11	0.07
6. Duration_PMC	0.96	1.85	1.41
	(3.08)	(4.63)	(3.95)
	0.00	0.00	0.00
	25.00	30.00	30.00
7. Duration_M1C	4.30	5.06	4.68
	(6.07)	(7.20)	(6.67)
	0.00	0.00	0.00
	31.00	34.00	34.00
8. Pattern..PMC			
Never	0.86	0.79	0.83
Once	0.13	0.19	0.16
Twice	0.01	0.02	0.02
9. Duration_RMC	0.88	1.00	0.94
	(3.37)	(3.87)	(3.63)
	0.00	0.00	0.00
	26.00	29.00	29.00
10. Age15_19_Start_M1C	0.00	0.04	0.02
11. Age15_19_Start_S1C	0.01	0.05	0.03
N	2851	2851	5702

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Table A.8 Boruta feature selection ranks for the pooled sample and by gender.

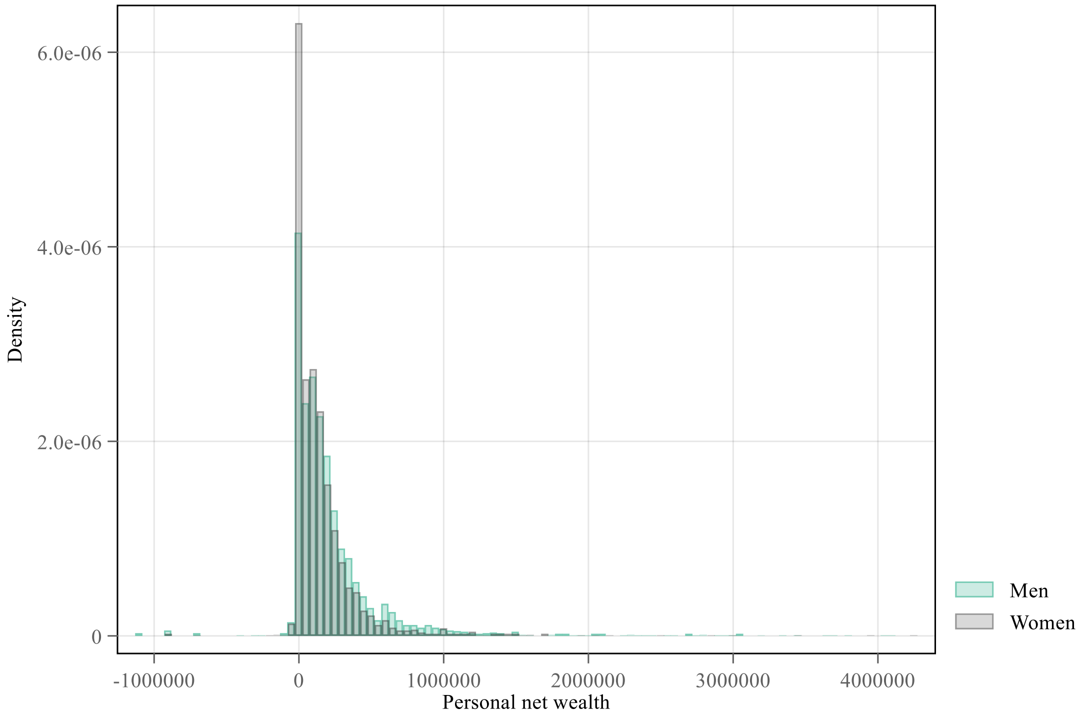
Features	rank Main	rank Men	rank Women
Duration.S1C	1*	9*	2*
Duration.SNC	2*	1*	3*
Pattern..M2C...M3C...PMC	3*	11*	11*
Age20_24.Start_MNC	4*	8*	23
Age20_24.Start_M2C	5*	15	15
Duration.PMC	6*	17	1*
Duration.M1C	7*	3*	5*
Pattern..PMC.	8*	16	7*
Duration.RMC	9*	2*	4*
Age15_19.Start_M1C	10*	25	12*
Age15_19.Start_S1C	11*	24	9*
Age30_34.Start_M1C	12*	18	20
Pattern..M1C	13*	4*	18
Age25_29.Start_MNC	14*	20	17
Pattern..M1C...M2C...PMC	15*	14	14
Duration.MNC	16*	5*	21
Age20_24.Start_M1C	17*	19	24
Pattern..S1C...PMC	18*	7*	13*
Pattern..M1C..M2C..M3C..PMC	19*	12*	16
Pattern..M1C...PMC	20*	22	10*
Duration.M4C	21*	21	22
Duration.M3C	22*	6*	6*
Duration.M2C	23 *(-)*	10*	8*
Age20_24.Start_S1C	24 *(-)*	13*	25
Age35_39.Start_PMC	25* *(23)*	23	19

Notes: Depicts all features that were either selected in the 'main' or gender-specific models. 'Main' refers to the main model for the whole sample, 'men' to the model restricted to men only and 'women' to the model restricted to women only. Ranked by mean importance in the main model measured as z-scores. *indicates features that were confirmed as important within the specific model. *(-)* refers to features that were not selected as important in the main model but were confirmed in the gender-specific models ('Age20_24.Start_S1C' and 'Age35_39.Start_PMC'). *(23)* depicts that feature 'Age35_39.Start_PMC' is ranked 23rd in the main model, when not considering features that were confirmed in the gender-specific model only.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

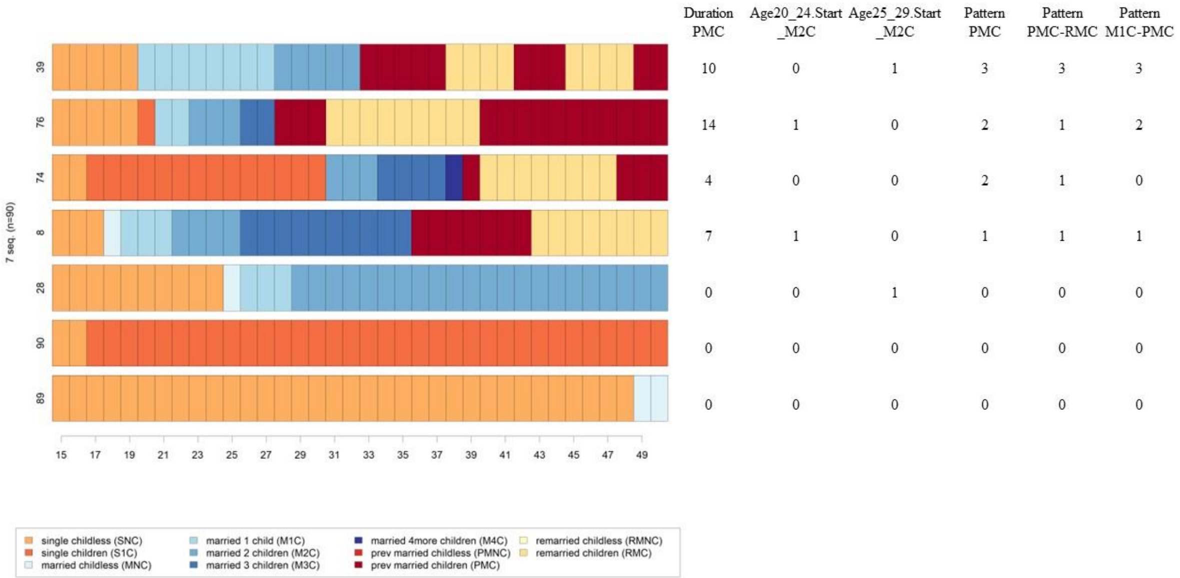
Additional Figures

Figure A.2 Density plot of personal net wealth by gender for the analytical sample



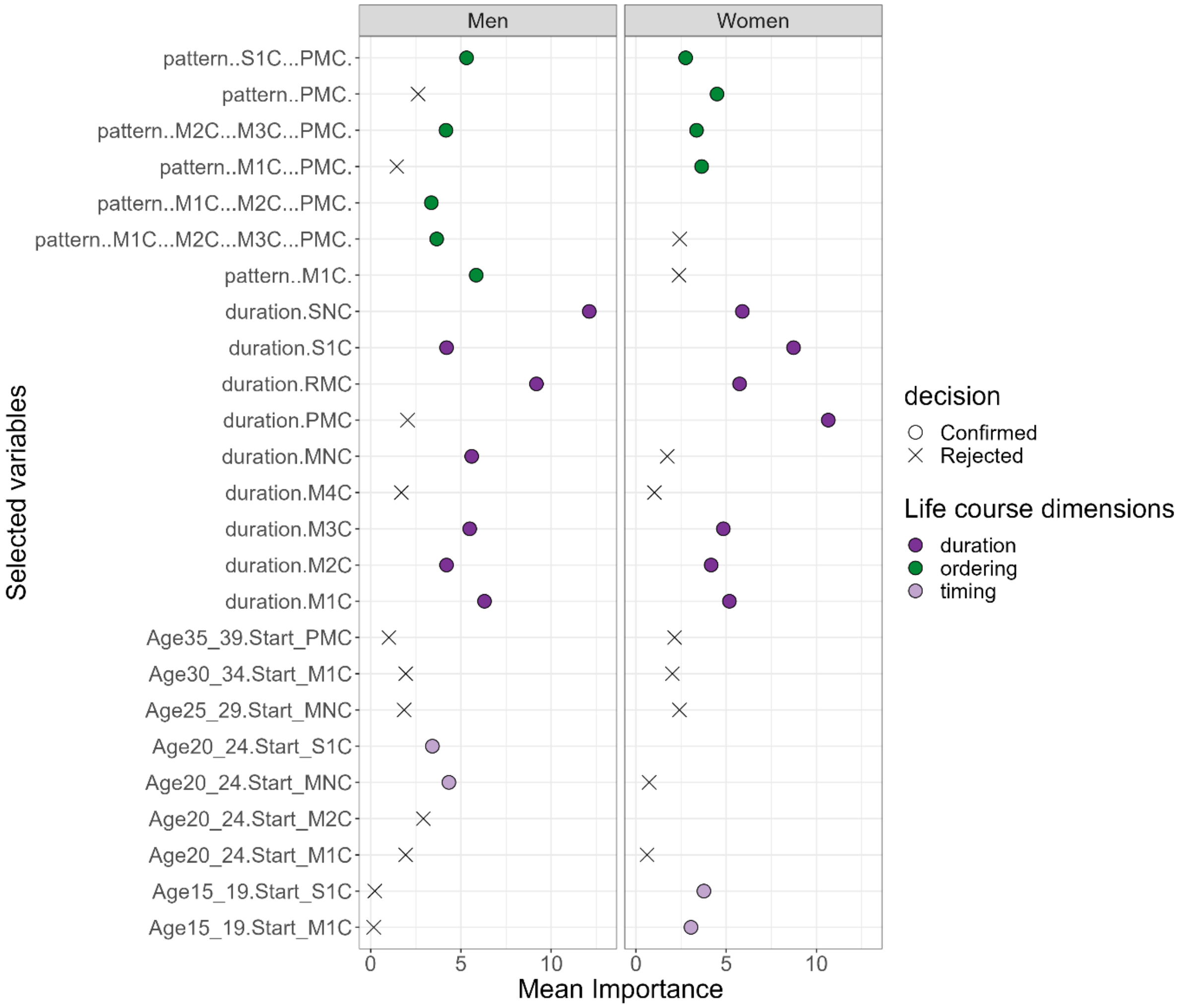
Notes: Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Figure A.3 Example trajectories and life course features



Notes: Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

Figure A.4 Boruta feature selection results considering women and men in separate samples.



Notes: Depicts the mean importance of all features that were confirmed important in the pooled model, in the model for wealth inequality among men only or among women only. Feature importance is measured as z-scores.

Data are from the Socio-Economic Panel Survey v38 (2002, 2007, 2012, 2017).

References

Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A., Benítez, J. M., & Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences*, *282*, 111-135. https://doi.org/https://doi.org/10.1016/j.ins.2014.05.042

Degenhardt, F., Seifert, S., & Szymczak, S. (2019). Evaluation of variable selection methods for random forests and omics data sets. *Briefings in Bioinformatics*, *20*(2), 492-503. https://doi.org/10.1093/bib/bbx124

Kursa, M. B., Jankowski, A., & Rudnicki, W. R. (2010). Boruta – A system for feature selection. *Fundamenta Informaticae*, *101*, 271-285. https://doi.org/10.3233/FI-2010-288

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, *36*(11), 1 - 13. https://doi.org/10.18637/jss.v036.i11

Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, *23*(19), 2507-2517. https://doi.org/10.1093/bioinformatics/btm344