Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Health Risks in Old Age

Implications for Household Savings and Insurance

Jeroen (Anne Jorn) van der Vaart

PhD 07/2024-001

Health Risks in Old Age
Implications for Household Savings and Insurance

Anne Jorn van der Vaart

Publisher: University of Groningen, Groningen, The Netherlands

Printer: Ipskamp Printing, Enschede, The Netherlands

Cover page: Rico van der Vaart

This document was prepared using LaTeX.

university of groningen

Health Risks in Old Age

Implications for Household Savings and Insurance

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. J.M.A. Scherpen
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Monday 1 July 2024 at 16:15 hours

by

Anne Jorn van der Vaart

born on 22 June 1996 in Dokkum

Supervisor
Prof. R.J.M. Alessie

Co-supervisors
Dr. M. Groneck
Dr. R. van Ooijen

Assessment Committee
Prof. V. Angelini
Prof. M.G. Knoef
Prof. S. Leth-Petersen

I wish that heaven had visiting hours...

Dedicated to my dearest father ♡

Acknowledgements

Together with my father, I visited the Bachelor's Open Day of the University of Groningen in Autumn 2013 to learn more about the Econometrics & Operations Research program. I cherish great memories of my father's support, especially now, some ten years later, having completed this wonderful program and even my PhD thesis. On the rough path to obtaining my PhD, I have been blessed with two of the greatest sources of wealth: love and support from others.

The supervision by Max, Raun, and Rob has been indispensable for my development as a researcher, but above all, as a person. We have known each other for about seven years now and much has happened in the meantime. I was first your research master student and thereafter your PhD candidate. The so many shared memories make me regularly realize: you have a special place in my heart.

Many other colleagues have also provided me with much guidance and support. Gerard, we have known each other for years, and we worked together successfully on the final chapter of my dissertation. Jochen & Viola, you have meant much to me both in the fore- and background, thereby opening doors for me that would otherwise remain closed. Agnieszka, what many may not know is that you supervised my bachelor's thesis and, together with Raun, advised me to pursue a research master's degree. Orest, under your guidance, I have been able to develop as a teacher, a role that brings me great satisfaction. I am very grateful to you and the many other colleagues that have guided me in the right direction.

In Spring 2022, I visited the University of Copenhagen. The stay taught me a lot about how other researchers work and what it is like to stand on my own feet. Søren, and everyone who is involved in the Center for Economic Behavior and Inequality (CEBI), your hospitality and help make me recommending anyone to visit CEBI and Denmark. Stefano, it was very enjoyable to spend time with you outside of work as well, it was the icing on the cake.

Research schools and -networks have played an important role in my career. The support from the back office, the research funding, the conferences, and the nice events organized by FEBRI have definitely improved my work. The same applies to the research network Netspar, which I have been part of for many years now. Especially, Marike and Lieke, I have often met you there, and greatly appreciate your endless help and encouragement. Lastly, I would like to thank the collaborations within the Centre for Public Health in Economics and Business (CPHEB) and Aletta's Talent Network (ATN). I experienced the board tasks and meetings to be a welcome broadening of my horizon.

A supportive environment is indispensable for bringing out the best in each other, day in and day out. My fellow PhD candidates, fellow research master students, and office mates have made me feel part of a group of friends. Also my friends outside of academia have supported me with more than just research. We have had great times throughout all the years we have known each other. These bonds continue to grow over time. Out of respect, I will not mention any of you by name. My friends know who I mean.

Someone recently told me that family means they are the first ones to call upon, no matter where you are or what the situation is. Dear Mum, Rico, Marianne, Ruud, and Amber, you make each day a little brighter. Your immense support and strength guide me through rough storms. The support by many other family members and my family-in-law is also invaluable to me.

Laura, my dear, with you I feel at home. Our love resides in small things but makes me move mountains. I love you very, very much.

Dear Dad, your last words "see you tonight, son" unexpectedly turned into a goodbye forever. It hurts me deeply that we can no longer share many milestones together, like my defense. But I still feel supported and advised by you every day, because in every tear hides the smile of a shared memory. See you someday, Dad ♡

Jeroen van der Vaart
Deventer, May 2024

Contents

Chapter 1

Introduction

1.1 General Motivation

Individuals are increasingly responsible for managing health and financial risks in old age. Due to population aging, public spending on pensions and long-term care (LTC) is expected to grow in OECD countries from 1.3% and 9.0% of GDP in 2018 to 2.3% and 10.3% in 2040 (OECD, 2021; de Biase and Dougherty, 2023). Substantial pension and LTC reforms have taken place to curb the cost. Many countries abolished pathways to early retirement and increased the statutory retirement age. Also, formal LTC is becoming less generous, often by encouraging individuals with LTC needs to live independently for longer in their own homes. For instance, by increasing co-payments for nursing home care and restricting access to nursing homes in case of lighter LTC needs. Consequently, the need to financially prepare for old age through self-insurance (e.g., through private savings or insurance) is expected to increase in the future. Also, informal care provision by family and relatives is expected to become more important.

When considering reforms of pensions and LTC and individuals' ability to self-insure old-age health and financial risks, it is essential to assess heterogeneity in risks and distributional consequences. It is well known that there is a large socioeconomic gradient in health and mortality, implying lower socioeconomic status groups to live shorter and to be unhealthier on average (Smith, 2007; Case and Deaton, 2017; Mackenbach et al., 2018). Besides, there is wide variation in the use of informal care: in the E.U. in 2016, 81% of the individuals needing LTC received informal care, whereas 19% did not receive informal care (European Commision, 2019). The socioeconomic gradient in health and mortality, and heterogeneity in informal care use may cause redistribution within pension schemes and public LTC insurance as some households benefit more than others. In private pension annuity and LTC insurance markets, the heterogeneities can lead to market inefficiencies as only above-average risks buy insurance.

Consequently, for the design of old-age (social) insurance, we need a good understanding of the heterogeneity and the dynamics of health and financial risks in old age. In this thesis, we examine mortality and LTC use as two of the most consequential uncertainties in old age. As an important mechanism to reduce formal LTC use, we consider whether the availability of informal care, income and savings affect LTC pathways: from no LTC to home-based care for lighter LTC needs to institutional care for more severe LTC needs (Chapter 2). We extend existing duration models to be able to estimate these transitions properly (Chapter 5). Turning to heterogeneities, we study the consequences of socioeconomic differences in LTC use and mortality for the design of LTC insurances and pension annuities (Chapter 3). Finally, we assess the consequences of inequalities in mortality and LTC for saving behavior and welfare of households (Chapter 4).

We analyze these questions in the context of the Netherlands, which provides a relevant setting for several reasons. First, there exist substantial socioeconomic differences in health and mortality (European Commision, 2021). Second, uncertainty related to pensions and LTC is limited (OECD, 2023; Bakx et al., 2023). Universal and comprehensive public LTC and generous pensions make precautionary saving for pension and health expenditures less relevant, as opposed to, e.g., the U.S, where it is indispensable to consider precautionary saving when studying health and financial risks in old age. Third, Statistics Netherlands (CBS) provides unique high-frequency administrative data on LTC use and mortality that can be linked on the individual and household level to tax and municipality registers containing socioeconomic and sociodemographic characteristics.

We now proceed with introducing each chapter of the thesis.

1.2 Motivation and Research Questions per Chapter

Insight into transitions across LTC arrangements (no LTC use, home-based care, and institutional care) is key for policy-makers aiming to reduce costly institutional care. Postponing entry into an institution may involve different policies and LTC needs than fostering a return from an institution to the home environment. Furthermore, the availability of informal care can vary across LTC arrangements and over time. With this in mind, Chapter 2 addresses the research questions:

2a. What is the duration of LTC use and the transition probability by type of LTC arrangement?

2b. What is the effect of disability type, availability of informal care, and economic resources on the transition probability across LTC arrangements?

In private insurance markets, the heterogeneity in risks contributes to a tendency to underinsure longevity risk and the risk of needing LTC, often referred to as the annuity puzzle and LTC insurance puzzle (for a review, see Lambregts and Schut, 2020). Adverse selection is one explanation for the limited market sizes, arising when those with above-average life expectancy more often buy annuities, and those with high expected LTC needs more often buy LTC insurance (cf. Finkelstein and Poterba, 2004). Another explanation for the low demand for LTC insurance is the availability of informal care (Mommaerts, 2024). Combining insurance when risks are negatively correlated has been proposed to reduce adverse selection (Murtaugh et al., 2001). Due to the well-documented wealth-health gradient, a negative correlation between LTC use and mortality risk might be present for socioeconomic groups. Despite their theoretical potential, old-age insurances that combine LTC insurance with annuities are still not very common, and their ability to cope with adverse selection is poorly understood.[1]

[1]The American Association for Long-Term Care Insurance highlights the favorable experience with LTC combination products over standalone LTC insurance; however, the number of policies sold remains limited, see: `https://www.aaltci.org/long-term-care-insurance/learning-center/ltcfacts-2019.php` and `https://www.aaltci.org/linked-benefit-faqs/` [*both retrieved on*: October 20th, 2023].

Chapter 3, therefore, addresses the following questions:

3a. How large are socioeconomic differences in LTC use and remaining life expectancy?

3b. What determines a combination of insurances that minimizes adverse selection?

3c. What is an optimal combination of annuity and LTC insurance?

While adverse selection might exist in private insurance markets, public programs warrant less redistribution when socioeconomic differences in health are present (Poterba, 2014; Auerbach et al., 2017). As the income-rich live longer than the income-poor, they receive more years of social security benefits.[2] In contrast, better health may induce lower LTC needs for the income-rich, implying fewer years of costly out-of-pocket LTC expenditures.[3] Consequently, there is growing consensus that reforms of old-age social insurance must account for both income *and* health disparities.

Despite being the workhorse model for studies on household's welfare (Low and Meghir, 2017), a life cycle model is rarely adopted in studies on the welfare or wealth effects of heterogeneous health in old age. Contrary to reduced-form models, life cycle models are structural models that directly link consumption and saving to obtaining utility. Beyond this, life cycle models are designed to analyze the contribution of heterogeneous risks, counterfactual policies, and saving motives to wealth accumulation. In the model, consumption and saving are the endogenous result of exposure to income and health risk, the available budget, preferences, and institutions. In this light, precautionary saving against uncertain future health (LTC) expenditures has proven relevant (De Nardi et al., 2010; Nakajima and Telyukova, 2024). Furthermore, ample literature shows that more affluent households hold a strong motive for saving for a bequest; the strength of this saving motive can be estimated with a life cycle model (De Nardi et al., 2010; Lockwood, 2018).

[2]For socioeconomic inequality in mortality, see, e.g., Deaton (2002); Smith (2007); Chetty et al. (2016).

[3]For socioeconomic inequality in LTC use, see, e.g., Goda et al. (2011b); Jones et al. (2018); Rodrigues et al. (2018); Tenand et al. (2020a).

Related, Chapter 4 aims to answer the following questions:

4a. How large are the distributional consequences of socioeconomic differences in LTC use and mortality?

4b. What underlying mechanisms drive the redistribution, in terms of saving preferences and elements of the pension and LTC system?

To answer the questions in this thesis, we make use of unique administrative data on the dates of LTC use and death. However, many administrative data sets are stock samples, implying a dynamic selection problem that has to be addressed. The data are typically observed from a given date onwards, and shorter spells ending before that date are left out, i.e., the sample is left-truncated. Consequently, subjects with favorable characteristics for long durations are over-sampled. This dynamic selection due to left truncation also happens to unobserved characteristics, so-called frailty. Ignoring dynamic selection due to left truncation can severely bias estimates of a duration model (van den Berg and Drepper, 2016). Related, we answer the following questions in Chapter 5:

5a. How can we account for dynamic selection due to left truncation when estimating LTC use and mortality risks?

5b. How large is estimation bias if dynamic selection due to left truncation is ignored?

Related to research question 5a., we develop a general estimation method and apply this method in Chapter 2 to 4.

1.3 Summary and Main Findings

In Chapter 2, we examine the durations of no LTC, home-based care, and institutional care *and* the transition probabilities between these care arrangements, given a need for care (a low or high physical or cognitive impairment). As a first step, we provide the empirical durations and transition probabilities, where the need for care is the only determinant that we study. We use unique administrative data that covers the history of

LTC needs and LTC use for the entire population of 65+ individuals in the Netherlands from 2009-2014. The data contain (single) spell data on home-based and institutional care use. The Dutch social insurance system featured universal and comprehensive LTC throughout this period, and co-payments were limited (Maarse and Jeurissen, 2016). As a second step, we study the impact of covariates on the availability of informal care and economic resources: having a partner or children (nearby), the health of the partner, income, and (financial) assets. To this end, we use the duration model from Chapter 5.

We find that the median duration of a home-based care spell with a low physical impairment is shorter than with a low cognitive impairment: two vs. five months. We also find that individuals with a low or high physical impairment are more likely to transition back to home-based care or no LTC. Turning to covariates, we find that having a healthy partner or children delays LTC entry and fosters a return from institutional to home-based care. However, while having a healthy partner delays institutional care use of physically impaired home-based care users, this, surprisingly, accelerates the use of institutional care for cognitively impaired individuals. Lastly, having more income, assets, or being a homeowner implies delayed LTC entry and a higher likelihood of returning to home-based care or no LTC use. Chapter 2 thus suggests substantial heterogeneity in the risk of using LTC.

Chapter 3 quantifies socioeconomic and socio-demographic differences in lifetime LTC use and mortality, and evaluates the implications for bundling LTC and annuity insurance. We extend the adverse selection model of Einav et al. (2010) to derive the combined insurance –a life care annuity– that minimizes adverse selection. In the model, heterogeneous types decide whether to be insured or not. Besides a negative correlation of risks, we derive two novel inputs for minimizing adverse selection with a life care annuity: (1) the mean duration in each of the two states 'no-LTC use' and 'LTC use', (2) the variance of the type-specific durations (reflecting heterogeneity). We then quantify differences in LTC use and mortality by gender, marital status, and lifetime income group. To this end, we use the same administrative data and duration models as in

Chapter 2. We use the estimation results to construct and evaluate the composition of optimal life care annuities.

We find substantial socioeconomic inequalities in LTC use and mortality. The difference in remaining life expectancy at age 65 between the bottom and the top lifetime household income quintile is 4.0 years for men and 2.3 years for women. Women in the bottom income quintile spend 1.7 more years in LTC than those in the top income quintile, while for men, this difference is 1.1 years. Hence, gender matters for the income gradient, which is stronger for men in terms of mortality, and for women in terms of LTC. Regarding informal care possibilities, being married reduces LTC duration by 22% for men and substantially flattens the socioeconomic gradient. At the same time, this is far less pronounced for women, potentially due to the high likelihood of outliving the spouse.

Following our theory and results, a life care annuity does not eliminate adverse selection if a uniform premium is offered. This is due to a gender effect that implies positively rather than negatively correlated risks: women live longer and use LTC longer. Group-specific premia instead yield large differences for the optimal insurance products over gender and marital status. Our results suggest that a life care annuity eliminates adverse selection for single men and women but less for married men and women due to unfavorable variances and correlations of the risks within these groups.

Chapter 4 quantifies the welfare implications of socioeconomic differences in LTC use and mortality. By adopting a life cycle model, we endogenize the consumption and saving decision of households. In the model, households draw utility from consuming, bequeathing, sharing a household, and living (remaining life expectancy). Importantly, using the model and evidence from Chapter 2 and 3, we allow LTC use and mortality risk to differ across gender, marital status and socioeconomic status (the lifetime household income quintiles). We estimate the parameters of the life cycle model using our unique administrative data, including tax-reported household assets. Next, we conduct counterfactual analysis to compute the additional welfare that higher socioeconomic status groups experience due to living longer and using less LTC. To this end, we endow each household with the health risks of the bottom lifetime income group and compute how much per-period consumption compensation each group requires to be as well off as with their true health risks (cf. De Nardi et al., 2024). Also, we study bequests and co-payments as drivers behind the welfare effects. To this end, we re-compute our welfare measure while removing bequest saving and LTC co-payments from the model one by one.

We report substantial distributional consequences of socioeconomic inequalities in LTC use and mortality. The welfare effect amounts to 23.4% additional consumption after age 65 for the households in the top lifetime income quintile compared to those in the bottom lifetime income quintile. We estimate a strong bequest saving motive for the income-rich, and consequently, their bequest saving motive explains 22.2 percentage points of their welfare gain. If we remove (abolish) co-payments, their welfare gain remains 21.8%, so only 1.6 percentage points are explained, implying that valuable bequests rather than co-payments explain the welfare gain.

In Chapter 5, we derive the likelihood-based estimator for duration models used in these chapters. We allow a frailty term (random effect) to be common among an arbitrary amount of left-truncated spells. For example, we frequently observe multiple LTC spells for the same individual: we assume individual frailty to be constant across these spells. In a Monte Carlo experiment, we show that ignoring the dynamic selection due to left truncation causes a substantial bias to time and covariate effects if frailty is spell-specific, but nullifies if frailty is shared among five spells. At the same time, the frailty variance is increasingly overestimated. Our user-written programs are available as `STATA` packages.

1.4 Policy Implications

The above results are relevant for the design of old-age social insurance and private insurance to accommodate heterogeneous LTC and mortality risks. We will now discuss policy implications that follow from our analyses.

Consider the care recipient's needs when facilitating informal caregiving

The results in Chapter 2 indicate that individuals with different types of disabilities have distinct LTC paths. For cognitively impaired individuals, institutional care use lasts longer and the partner is less effective in delaying the start of this state. Thus, family members of individuals with cognitive impairments have more prolonged exposure to burdensome caregiving, and their caregiving could be less effective in reducing the uptake of formal LTC use. With this in mind, our results stress a need for separate support programs for informal caregivers of cognitively and physically impaired LTC users. These programs include, for example, training and leave arrangements for informal caregivers, such as the 'Wet Arbeid en Zorg' (WaZo) in the Netherlands and the 'Family and Medical Leave Act' (FMLA) in the U.S.. While these programs exist in practice, they are not necessarily tailored to the disability type of a care recipient.

Develop separate care arrangements for private nursing homes

We show in Chapter 2 that individuals with more income or wealth postpone institutional care use and return home faster if they use institutional care. More affluent individuals thus live at home longer. In this light, our findings provide scope to opening more private residences and private nursing homes where individuals pay out of pocket for the accommodation while the government finances the care. This implies a shift away from publicly provided LTC to private providers, thus a possible demand for private insurance against nursing home costs.

Allow for flexible combinations of pensions and LTC insurance

Our analysis in Chapter 3 shows that adverse selection for stand-alone pension annuities and LTC insurance is reduced when combining the products, especially for single-person households. This reduction in adverse selection could be achieved with the life care annuity we propose, i.e., offering a top-up benefit when needing LTC. An attractive alternative would be to allow existing pension annuities or life insurance to pay for LTC cost, a so-called LTC rider. While the idea sounds intuitively appealing, the willingness to buy these products is very low in the U.S. (Chen et al., 2022a). With this in mind, governments could fiscally stimulate the use of LTC riders by counting this as a tax-deductible health expenditure.

Allow group-specific premia for a combined pension and LTC insurance

Our results in Chapter 3 reveal that a combined LTC insurance and annuity only effectively combats adverse selection if premia can depend on gender and marital status. However, in the European Union, the Court of Justice declared gender-specific premia invalid with European legislation and prohibited this practice in Europe in 2012. Gender-based pricing in insurance is still practice for many insurances and many states in the US, although the Affordable Care Act banned discrimination over gender for health insurance in 2014. Instead, allowing gender- and marital status-specific premia would lower adverse selection.

Consider health inequalities and redistribution when designing old-age social insurances

Our findings in Chapter 4 show that socioeconomic differences in mortality and LTC use lead to higher retirement income and lower co-payments for wealthier individuals. Governments could view this as an unintended income-regressive redistribution and might want to repair the welfare effects. Offering a lump sum payment at retirement could reduce the effect of health inequalities on retirement benefits. While lifetime annuity benefits depend on an individual's lifetime, the lump sum payment is instead based on the population life expectancy, and thus effectively the same for everyone. In line with more equal benefits, the Dutch government is currently discussing the plan to allow individuals to receive a maximum of 10% of their accrued second pillar pension in the form of a lump sum payment (Mehlkopf et al., 2019). Similarly, our findings speak for tying the statutory retirement age to career length because shorter-living (lower) lifetime income quintiles usually start working at younger ages. Consequently, the life expectancy in retirement is more homogenous across socioeconomic groups.

Increase taxation of bequests

Our results in Chapter 4 show that the more affluent leave larger bequests due to living longer (more pension) and using less LTC. If governments perceive this as an undesirable income-regressive redistribution, then higher taxes on bequests could introduce more actuarial fairness into the old-age (social) insurance system. It should, however, not be forgotten that this demands good coordination with the tax system of inter-vivos transfers, as inter-vivos can substitute bequests (Kopczuk, 2013).

1.5 Avenues for Future Research

Chapter 2 documents the importance of the availability of informal care and economic resources for realized transition paths in the LTC system. Future work can aim for a better understanding of the causal effects of these determinants by doing policy evaluation. A relevant question is, for example, to what extent do co-payments limit entry into home-based and institutional care? Furthermore, to what extent does mandating (some) informal care provision reduce the entry and duration of institutional care use? In recent years, the Dutch government implemented several reforms that could be leveraged to answer these questions, including more stringent eligibility criteria for institutional care and higher co-payment rates on assets. While other work already aimed to look at the cross-sectional effects of such policy reforms (for co-payments, see Tenand et al., 2023), studying transitions and persistence of using LTC remains an avenue for future research.

Policy evaluations are also a natural follow-up for our life cycle model that quantifies the welfare effect of health inequalities in Chapter 4. Our life cycle model can be used in an analysis that compares different setups of social security and public LTC insurance in terms of efficiency and equity, while considering health inequalities. In this respect, it would be interesting to compare two extreme systems that only consist of private or public LTC insurance, i.e., the U.S. welfare state versus the Northern European welfare state. Doing these more profound policy analyses would undoubtedly make the working age stage of the life cycle more relevant, i.e., when individuals decide to work, pay social contributions, accumulate savings, and buy insurance. Our current analyses focus more on the health inequalities that occur at a later age, so we pay less attention to the working age stage.

The life cycle model and the empirical results from Chapter 3 and 4 can also be used to compute the adverse selection costs of offering private LTC insurance and annuities. While adverse selection costs of single products were a topic in earlier work (see, e.g., Boyer et al., 2020), the adverse selection cost for a combined product has never been formally identified. Such analysis demands individual-level data on choices (trade-offs) between the different insurances and data on individual characteristics and subjective information; these data usually do not co-exist. It is worth mentioning that recently, using our results from Chapter 3 to price the products, a LISS panel survey was set up on hypothetical trade-offs between the products (de Bresser et al., 2022, 2023). The survey data can be merged with our administrative data, leaving an excellent laboratory setting to study the adverse selection cost of combined LTC and annuity insurance.

Another key insight from our study that requires further study is a strong bequest saving motive by Dutch households. A simple question could be: do individuals plan to leave a bequest, and if so, when do they start planning? One could also turn the question and take the recipient's perspective: do households anticipate bequests, and do their life cycle choices, including labor supply, housing, and informal care provision, differ because of anticipated bequests? Answering these typical questions is important for better coordination of taxing bequests and labor income, which are substitutes to some extent. At a higher level, this touches upon the extent to which intergenerational mobility of (human) wealth is present within the society and across multiple generations (Chetty et al., 2014; Lindahl et al., 2015), which we plan to study in future work.

Finally, in this thesis, we gain insights into the heterogeneity of life-course events occurring relatively late in life, such as widowhood, health decline, and eventual death. While these are substantive health risks late in life, these can also play on earlier in life. Therefore, we encourage future work to also focus on the many sources of uncertainty present earlier in life, e.g., divorce or disability risk. It would be interesting to see how these risks impact life cycle outcomes such as having adequate retirement income and savings at retirement.

CHAPTER 2

PATHWAYS IN LONG-TERM CARE AND ITS DETERMINANTS: EVIDENCE FROM DUTCH ADMINISTRATIVE DATA

This chapter is joint work with Rob Alessie, Max Groneck, and Raun van Ooijen. We thank Bertrand Achou, Julien Bergeot, Arjen Hussem, Marike Knoef, Kathleen McGarry, Torben Heien Nielsen, Arthur van Soest, Marianne Tenand, Bas Werker, and seminar participants at various seminars and conferences for helpful comments. We thank Netspar for financing data access through two theme grants: 'Uncertainty over the life cycle: implications for pensions and savings behavior' (2017-2020) and 'Health and labor market uncertainty over the life cycle: The impact on households' risk capacity and retirement income adequacy' (2022-2024).

2.1 Introduction

Confronted with an aging population, many OECD countries seek to provide adequate long-term care (LTC) while keeping the system sustainable. Policy-makers encourage cost-efficient alternatives to receiving care in an institution, such as home-based care. Two-third of care recipients use home-based care, while only making up one-third of LTC spending (OECD, 2019). Postponing an individual's transition into an institution and fostering a return home are two ways policy-makers can reduce institutional care use. Understanding the transitions into and out of LTC arrangements and their determinants, such as different needs for care, availability of informal care, and (non-)financial assets (de Meijer et al., 2011; Hiedemann et al., 2018; Diepstraten et al., 2020), is a necessary first step if policies seek to reduce public LTC costs.

While cross-sectional evidence on using LTC is fairly conclusive, the transition path between receiving home-based and institutional care and the role of informal care and economic resources is much less understood. For example, informal care substitutes personal care provided at home (van Houtven and Norton, 2004; Bonsang, 2009) but not necessarily in an institution (Charles and Sevak, 2005; Bergeot and Tenand, 2023). Is this because informal care postpones using home-based care, but not the transition from home-based care to a nursing home? Yet, identifying pathways in LTC conditional upon the need for care is challenging due to a lack of longitudinal datasets reporting both the need for and use of LTC arrangements at high frequency with the possibility to link individual and family members' characteristics.

In this chapter, we examine the determinants of the duration of no LTC, home-based care, and institutional care and transitions across the arrangements, given the need for care. We estimate a multi-state model on administrative data, covering the history of LTC needs and use of different formal LTC arrangements for all 65+ individuals in the Netherlands from 2009-2014. We link our unique data to personal records on family, socioeconomic and socio-demographic characteristics. We particularly focus on determinants that reflect availability of informal care and economic resources: having a partner or children (nearby), the health of the partner, income, and (financial) assets. To analyze whether the potential availability of informal care and economic resources accelerate or delay each of the transitions, we apply the novel estimation procedure for mixed proportional hazard models from Chapter 5. The estimated specification allows us to distinguish the covariate effects from other time-varying circumstances such as own health, duration dependence, and unobserved heterogeneity (frailty).

A main advantage of using administrative data is that they allow us to estimate the effects for subpopulations precisely. Specifically, we contrast users of LTC with a low or high physical or cognitive impairment. We are hereby motivated by a wide literature that documents large heterogeneity in the need for care (see, e.g. de Meijer et al., 2011). Moreover, Bonsang (2009) shows that severity of a care need is a factor that reduces informal care provision. The Dutch institutional setting offers an ideal opportunity to study them. A mandatory eligibility assessment grants access to the LTC arrangements for a particular hours per week (defining a low or high need here) based on a diagnosed type and severity of impairment. Universal and comprehensive public LTC insurance during the observational period allows us to largely abstract from dynamics that could be induced by private LTC insurance (see e.g., De Nardi et al., 2010).

We have four main findings. First, we find that physically and cognitively impaired individuals differ largely in their pathway through LTC arrangements, as measured by the duration and realized transition. More specifically, our results reveal that temporary use of LTC (cf. Einav et al., 2022) predominantly involves physically impaired individuals. The median time spent in home-based care is two months for individuals with low physical impairments, while it is about five months when having low cognitive impairments. Also, those with physical impairments more often transition to less specialized LTC arrangements, so from institutional care to home-based care or from home-based care to no LTC use. For instance, of those in home-based care, 37% with low physical impairments transition to no LTC, which is only 11% for those with low cognitive impairments.

Second, the effect of availability of informal care depends on the transition and potential source of informal care. While having a healthy partner *decreases* the hazard of institutional care use by 20% for home-based care users with a high physical impairment, we find that having a healthy partner *increases* the hazard of institutional care use by 19% for home-based care users with a high cognitive impairment (reference: singles). Third, the hazard of going from home-based care with a low physical impairment to no LTC use is 54% higher when having a healthy partner. Albeit weaker than the effect of the healthy partner, we find that having children delays the use of more specialized care and accelerates the use of less specialized care for all transitions. Lastly, turning to economic resources, we find that having more income, financial assets, or owning a house fosters a return to less specialized care and delays a transition into more specialized care.

We contribute to a broad literature on the need for LTC as a determinant for using LTC arrangements. The vast amount of work focuses on a single point in time (see e.g.: Portrait et al., 2000; Luppa et al., 2010; de Meijer et al., 2011; Sovinsky and Stern, 2016; Hiedemann et al., 2018; Duell et al., 2021). A multi-state framework is used by Dostie and Léger (2005) to examine transitions across living arrangements, which can be recurring but by definition exclude home-based care. Fuino and Wagner (2018) use the framework to study transitions into more specialized LTC arrangements, including home-based care but not allowing returns to less specialised care, i.e. states are not recurring. The distinctive aspects our model, which are crucial to explain observed heterogeneity in pathways, are including heterogeneity by need for care, recurrent transitions and home-based care.

Second, we provide evidence of the effects of informal care receipts on formal LTC use. While there seems to be consensus on informal care substituting several sorts of home-based care (Bolin et al., 2008; Kalwij et al., 2014; Barczyk and Kredler, 2018), the evidence on institutional care is mixed as some studies report substitutability with informal care (van Houtven and Norton, 2004; Charles and Sevak, 2005), whereas some do not (Bergeot and Tenand, 2023). Instead, by looking at the timing of a transition, we show that substitutability depends on the direction of a transition (i.e., to less or more specialized care) and the type of impairment. Particularly, we do not find evidence for substitution when going from home-based care to institutional care with a cognitive impairment. These individuals comprise a substantial part of the study population in Bergeot and Tenand (2023), so likely drive their findings.

Third, we contribute to the literature on income and assets as determinants for LTC use (e.g. Charles and Sevak, 2005; Luppa et al., 2010; Rouwendal and Thomese, 2013). Most of these results are mixed due to stringent eligibility criteria and the existence of private insurance in the U.S.. In turn, because eligibility is not means-tested and private insurance is virtually non-existent in the Netherlands, we attribute our estimated effects to selecting options other than a formal LTC arrangement, e.g. making adaptations to the home environment (Diepstraten et al., 2020), compensating informal caregivers (McGarry and Schoeni, 1997), or privately paying for LTC (De Nardi et al., 2010).

Knowing key determinants for pathways in LTC is essential for developing policies to encourage elderly to live independently at home for longer. Policy-makers can use our findings to equip high-risks groups with personalized support that anticipates their future need and use of LTC. For example, home adaptations for singles without children or informal care training for partners of cognitively impaired. Also, the findings can be used to make eligibility criterions stricter for low-risk groups or can be applied when having to prioritize risk groups on waiting lists for institutional care. Moreover, finding that more affluent live longer at home while receiving care, implies that it can be sensible to allow to receive publicly-covered care in a (private) facility of own choice, where individuals pay for accommodation out-of-pocket.

We proceed as follows. Section 2.2 describes the institutional context. Section 2.3 and 2.4 describe the data and provide descriptive statistics. Section 2.5 describes the empirical approach. Section 2.6 presents the results. Section 2.7 discusses and concludes.

2.2 Public LTC in the Netherlands 2009-2014

Public expenditures on LTC are 4.3% of GDP which is the highest among all OECD countries (OECD, 2015). The high expenditures go along with a comprehensive coverage (Colombo et al., 2011). While in most countries the out-of-pocket expenditures for formal LTC are substantial, in the Netherlands almost all formal LTC is paid for by social insurance contributions. Only 8% of the costs was financed by income- and asset-dependent co-payments in 2012 (Maarse and Jeurissen, 2016). Privately paid services were virtually absent until 2014 (Hussem et al., 2020). A large budgetary cut in 2015 mandated a shift from public institutional care to home-based and privately arranged care; we leave this outside the scope of our study.

Tailored to the type and severity of their health problem, individuals receive paid care at home or in specialized institutions. Home-based care includes social support (e.g. adult day care), personal care (e.g. washing and feeding) and nursing (e.g. wound dressing). About 30% of the LTC beneficiaries aged 65 and older live in an institution, where they receive a package of these services (Tenand et al., 2020a). A nursing home provides intensive care for elderly with severe cognitive or physical problems, e.g. following a stroke. Until 2013, care could also be rehabilitative, like Skilled Facility Nursing in the U.S. (Hackmann and Pohl, 2018). For elderly who cannot live independently but need less intensive care, residential care homes provide assisted living (Kok et al., 2015). The institutions have to adhere to strict guidelines on high quality of care (Bär et al., 2022).

To become eligible for publicly-provided care, the individual, or their family member or health care provider, has to apply at the government agency CIZ (Bakx et al., 2020). The eligibility decision is made by an assessor. The assessor collects information about functional limitations from current and prior applications, and might consult the applicant or health care provider (e.g. the GP) for additional information. Individual's income or wealth is not taken into account. Informal care reduces entitlements to LTC insofar as household members are capable of providing personal care. This applies only if the need is expected to last less than three months and home-based care is required;

but not if institutional or nursing care is required (Mot, 2010).

Following a standardized procedure with some discretionary power, the assessor then decides the type of care, amount of care, and length of time being eligible, or rejects the application. The entitlement is tailored to the main health problem that the assessor identifies: a physical impairment or disability, cognitive impairment, intellectual disability, psychiatric disorder, or sensory disability[1]. Beneficiaries of institutional care require round-the-clock and are granted access to one of the 52 default 'care severity' packages, specifying the care type (e.g. dementia care) and indicating the average number of hours they may use nursing care, personal care, and social support. Home-based care entitlements specify the hours of care for each of these LTC services apart. Care is not automatically granted; about 11% of the applications is rejected (CIZ, 2013). A new assessment occurs each five years or if the health situation changed.

Eligible elderly could opt for in-kind benefits or a lower personal budget used to pay for caregiving, which 4% did in 2012 (Algemene Rekenkamer, 2015). Also, they may convert an entitlement to institutional care into equivalent home-based care, implying lower co-payments. Wait times are limited: virtually all elderly can use their entitled care within the legal acceptable wait time of six weeks (CVZ, 2013). Bridged by using home-based care, elderly with a preferred institution may choose to wait longer.

A few relevant policy changes took effect in 2013. First, the co-pay rate on assets increased from 4% to 12%. However, in response individuals do not seem to have substituted institutional care by home-based care (Tenand et al., 2023). Further, the care severity packages involving only a few hours of residential care were no longer granted to new beneficiaries; they are granted home-based care instead. Lastly, the insurance of institutional rehabilitative care is privatized. Consequently, the number of publicly institutionalised elderly dropped by 24% between 2012 and 2014 (CBS, 2014).

Overall, the Dutch LTC system provided generous coverage until 2014, implying that pathways in LTC are primarily driven by an individual's health, personal circumstances and preferences.

[1] The classification is based on ICD-10, ICF and DSM-IV standards (CIZ, 2014).

2.3 Data

We use high-frequency administrative data on the individual's use of LTC arrangements, need for LTC, availability of informal care, and individual characteristics collected by Statistics Netherlands. We include all individuals aged 65 and over from 2009 to 2014 with information on their stays in institutional care, use of formal home-based care, date of death, eligibility assessment, family composition, and individual characteristics.

The data infrastructure at Statistics Netherlands allows us to link characteristics at the individual and family level across all registers and over time. We start with the Municipal Population Register, that reports basic socio-demographic characteristics on birth year, ethnicity and gender for all residents between 1995 and 2014. We only keep surviving residents on January 1^{st} 2009, because from then onwards the registry on the eligibility assessment is available. Also, we exclude individuals who migrated during this period.[2] Our target population comprises of 3,590,373 individuals aged 65 and over. Our studied window runs from 2009 to 2014.

2.3.1 Spell Data on LTC and Mortality

The key event in our analysis is the timing of a transition between no LTC use, home-based care, institutional care, and the transition into death. To compute the state duration and the subsequent transition, we use the dates on which the individuals used formal home-based or institutional care and passed away. We also observe the date and outcome of their eligibility assessment (see Table 2.1 for an overview of variables), which we will use to proxy for the need of care. We will study the timing of a transition conditional upon the need for care.

We must go through a few steps to construct the desired duration variables. Within a spell of no LTC use, we distinguish two sub-spells: someone has never used LTC before (a potential first-time user) or someone ever used LTC before.[3] The spell of 'a never

[2]We define being migrated as: the individual or partner does not have a living address in the Netherlands between 2004 and 2014.

[3]We observe use of LTC since 2004. Any previous use is assumed to have not occurred.

user of LTC' spell starts on the 65^{th} birthday and the spell of 'an ever user of LTC' starts on the date that LTC use stopped. The latter group is an relevant subgroup to look at because they possibly have health problems that are more difficult to cover with informal care. We keep all spells that are ongoing on January 1^{st} 2009 or start after that date.

Within a spell of home-based or institutional care use, we look at four sub-spells: having a low physical impairment, a high physical impairment, a low cognitive impairment, or a high cognitive impairment. Having a physical impairment includes having a physical disability. We only keep spells starting after January 1^{st} 2009, because for those we know the start date of the entitlement. The maximum observed spell duration is thus six years.

We define a low (high) impairment as an entitlement to hours of care below (above) the observed median (we report and discuss the cut-off values in Section 2.4). For example, a home-based care user with a physical impairment and who is entitled less than 5.5 hours of care (the median for this group), has a low physical impairment. If the individual is entitled to home-based care, we compute the hours of care as the sum of entitled personal and nursing care. To do so, we impute the absolute hours of personal and nursing care with the midpoint of each category.[4] For institutional care, we know the care severity package, which grants individuals access to particular hours of care. Appendix B.1 provides the mapping for each package into hours of care.

Home-based care use is reported every four weeks. We observe whether the individual used this care during that period, but not exactly when. To circumvent that we do not know the start and end date of use, we assume that the use occurred during the entire four weeks. To focus on individuals with health problems, we restrict home-based care users to users of personal or nursing care and exclude those who merely receive social support (see Appendix B.2 for a complete description of the LTC services). We observe in-kind care but not the minor share with a personal budget (4.4% of the users in 2012). The need for LTC and the use of institutional care are reported daily and unrestricted.

[4] The categorical values are not our choice, but applied by CIZ, see CIZ (2014) for further information.

Lastly, we close gaps of less than six weeks between two spells of home-based care, institutional care, or need for LTC. If home-based and institutional care use follow-up each other within six weeks, we assume that the first event continued during the gap. This smoothing aligns with the legally accepted waiting time of maximally six weeks between different LTC arrangements in the Netherlands ('de Treeknorm', see Hartmans et al., 2009).

2.3.2 Determinants of LTC use

Wherever we can, we stay close to the definitions in Duell et al. (2021). They use the same administrative data as us to study the link between used and entitled hours of care.

Availability of informal care. We construct two variables to capture the impact of the potential availability of informal care on the transitions: having a (healthy) partner and distance from the parent to the children. Partnered individuals are either married, have a partnership contract, or cohabit on a contractual basis. To account for health-dependent availability of informal care, we distinguish between a partner who uses no LTC, home-based care, or institutional care. Distance to the closest child proxies for the direct availability of informal care by offspring (cf. Bonsang, 2009). This distinguishes between a child living in the same municipality or a different municipality than the parent.

Economic resources. We also consider socioeconomic variables that could reflect the preference or financial possibility to choose a particular LTC arrangement. To this end, we construct decile groups on household income and financial assets and look at homeownership that could reflect a preference to age in one's own house (for evidence for institutional care, see Rouwendal and Thomese, 2013). Household income is measured pre-tax and comprises income from labor and capital, retirement income, and social insurance benefits of all household members. To make single and multi-person households comparable, we equivalize the household income by the OECD equivalence scale (OECD, 2011). Financial assets are the sum of checking and savings, stocks and bonds, and excludes entrepeneurial wealth and other real estate such as the own house.

Lastly, we include another list of covariates on health and socio-demographics to account for omitted variable bias: gender, ethnicity, region of residence, and health proxied with medication use. Unfortunately, the registry excludes medication prescribed in hospitals and nursing homes, and we, therefore, do not stratify our analyses by this variable.

Additional information about the data sets and variables is provided in Appendix B.1.

Table 2.1: Overview of Variables

Measurement frequency	
Use of formal home-based care	Every four weeks.
Use of institutional care	Daily.
Eligibility assessment	Daily.
Death	Daily.
Availability of informal care	Daily.
Individual characteristics	Yearly.
Eligibility assessment (entitlements)	
Main health problem	1: Physical impairment; 2: Physical disability; 3: Cognitive impairment; 4: Sensory disability; 5: Mental disorder; 6: Intellectual disability.
Hours per week of personal care at home*	1: 0-2; 2: 2-4; 3: 4-7; 4: 7-10; 5: 10-13; 6: 13-16; 7: 16-20; 8: 20-25; 9: 25+.
Hours per week of nursing care at home	1: 0-2; 2: 2-4; 3: 4-7; 4: 7-10; 5: 10-13; 6: 13-16; 7: 16-20; 8: 20-25; 9: 25+.
Care severity package for institutional care	52 possible packages: see Appendix B.1 for the entitled hours of care.
Availability of informal care	
Marital status/health	1: Single; 2: Partner without LTC; 3: Partner uses home-based care; 4: Partner uses institutional care.
Distance to the closest child	1: No children; 2: Lives in the same municipality; 3: Lives in a different municipality.
Economic resources**	
Equivalized household income***	Decile groups and levels in 2019 €s.
Household financial assets	Decile groups and levels in 2019 €s.
Homeowner	1: Yes; 0: No.

Notes: * For the category 25+ we know the exact amount of hours exceeding 25. The categorical values are not our choice, but applied by CIZ, see CIZ (2014). See Appendix B.2 for a complete description of the LTC services; ** Appendix B.1 provides the values of gender, ethnicity, region of residence, and medication use; *** Equivalized with the OECD equivalence scale (OECD, 2011).

2.4 Descriptive Statistics

Table 2.2 shows descriptive statistics by used LTC arrangement. To compare groups across LTC arrangements without age effects being present, we focus on the cross-section of 85-year-olds on January 1^{st} in 2009-2014.[5]

The first block shows that individuals without LTC use are more frequently men, married, homeowners, and have higher income and assets. The probability of having children does not vary substantially across arrangements. Another pattern worth highlighting is the correlation between the arrangements used by two partners. Respectively 88%, 37%, and 57% of the couples have both members using no LTC, home-based care, or institutional care. Lastly, we see that most home-based care users have a physical impairment as main health problem (81%). Instead, institutional care users more frequently have a cognitive impairment (40%).

To further understand the need for LTC, Table 2.3 provides the distribution of hours of care by used LTC arrangement and main health problem. For each arrangement, individuals with a physical impairment are entitled to fewer hours of care than those with a cognitive impairment. The median entitled hours of care is 5.5 for home-based care users with a physical impairment, while this is 13.25 when having a cognitive impairment. The medians serve as cut-offs to indicate a sub-spell of low or high need; below the median means a low need, and vice versa. For example, home-based care users with a physical impairment and entitlement below 5.5 hours have a 'low need'.

Part of the difference in hours of care comes from the entitled arrangement in home-based care. While 48% of the home-based care users with a cognitive impairment has an indication for institutional care, this is only 12% for those with a physical impairment (not shown). Given that there is a care severity package tailored to dementia which assigns to 19.25 hours of care, we already see a clustering at this value for home-based care users with a cognitive impairment. As a consequence, we cannot exclude the possibility that they wait for an open spot in their preferred institution.

[5] Some have an entitlement to LTC with an unknown start date as this entitlement starts before January 1^{st} 2009. We drop those observations once we apply our duration analysis.

Table 2.2: Observed Characteristics at Age 85 by Used LTC Arrangement

	No LTC use	**Home-based care**	**Institutional care**
Woman = 1	0.60	0.73	0.76
Partner = 1	0.42	0.27	0.20
Partner uses same arrangement = 1*	0.88	0.37	0.57
Has children = 1	0.85	0.85	0.81
Median equivalized household income**	21.0	19.6	18.3
Median household financial assets**	35.5	27.0	24.8
Homeowner = 1	0.40	0.29	0.17
Main health problem:			
Has physical impairment = 1***		0.81	0.52
Has cognitive impairment = 1		0.10	0.40
Has other problem = 1****		0.02	0.04
Has no entitlement = 1		0.07	0.04
Individuals (%):	207,819 (64)	63,083 (20)	52,580 (16)

Notes: * Conditional upon having a partner; ** 000s €; *** Physical impairment or disability; **** A sensory disability, intellectual disability or mental disorder; Appendix Table B.5 provides the numbers for the full population.

Table 2.3: Entitled Hours of Care at Age 85 by Used LTC arrangement and Health Problem

	Home-based care		Institutional care	
Impairment:	Physical	Cognitive	Physical	Cognitive
Hours of care per week:				
0-2	17	6	2	4
2-4	23	11	0	0
4-7	25	15	10	0
7-10	16	12	19	1
10-13	8	5	22	1
13-16	4	23	13	18
16-20	2	26	21	66
20-25	2	2	10	10
25+	1	0	4	0
$\sum$	100%	100%	100%	100%
Median*	5.5	13.25	11.5	19.25
N	51,001	6,572	27,395	21,275

Notes: * The median define a low and high need in this chapter. Low need: the entitled hours of care is below the reported median, and vice versa for a high need; Appendix Table B.6 provides the numbers for the full population.

2.5 Empirical Framework

The descriptive statistics focus on the use of LTC and the need for LTC in the cross-section, but we are rather interested in their longitudinal outcomes: for how long do individuals use an LTC arrangement with a particular need, and what is their subsequent state? We have to adopt a multi-state duration model to describe and answer how these outcomes relate to the potential availability of informal care and economic resources.

We are interested in the duration T and the next state of the different sub-spells from Section 2.3.1. To this end, we use a competing risks framework. Figure 2.1 provides a motivating example of home-based care users with a low physical impairment. Let i indicate the current state. Five mutually exclusive next states exist, i.e., competing risks. We closely follow the two-step counting procedure by Kalbfleisch and Prentice (1980) to estimate the underlying distributions. First, we compute the duration distribution $S_{ii}(t)$ of remaining in state i until time t. Next, to find out why some durations are shorter than others, we analyze the transition to state $j \in \mathcal{J} \backslash \{i\}$. These transition probabilities $S_{ij}(t)$ add up to $1 - S_{ii}(t)$, the complement of the duration distribution: a transition to another state j by time t. Lastly, and for parsimony, we estimate a mixed proportional hazard model to quantify the impact of informal care availability and other characteristics on the transitions.

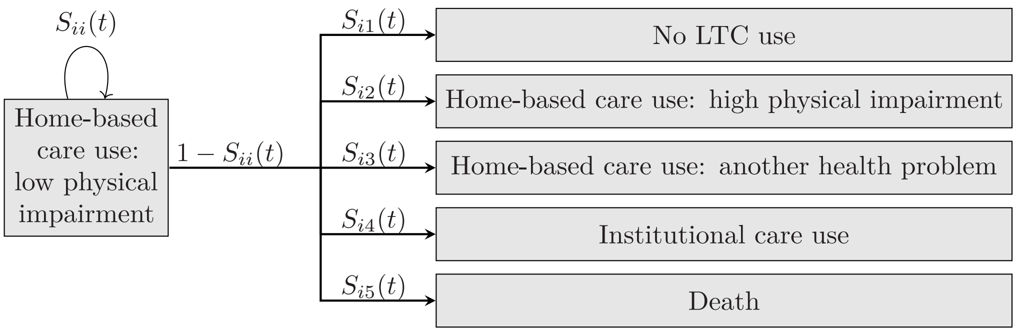


Figure 2.1: A Competing Risk Setting in Our Study

2.5.1 Duration Distribution and Transition Probabilities

We base estimation of $S_{ii}(t)$ on the conditional transition probability of leaving state i at time t: $\lambda_i(t)$. To compute $\lambda_i(t)$, we have data on the number of individuals in state i who can transition at time t, $N_i(t)$, and those who actually do so during the subsequent time interval dt, $D_i(t)$. Then:

$$\lambda_i(t) = \mathbb{P}(t \leq T < t + dt \mid T \geq t, \text{State} = i) = \frac{D_i(t)}{N_i(t)}, \tag{2.1}$$

which is the share of individuals at risk at time t who do transition.

In the data, individuals can have multiple spells of being in the same state, e.g. repeated home-based care use. For our analysis, this means that the same individual can be at risk multiple times at time t, each counting separately when measuring $N_i(t)$ and $D_i(t)$. We ignore possible dependence across spells in this descriptive analysis but include this in the parametric model of Section 2.5.2.

Intuitively, survival probability $S_{ii}(t)$ is now the same as the probability of not transiting in every preceding period, indicated by probability $1 - \lambda_i(t)$:

$$S_{ii}(t) = \prod_{k:t^{(k)} \leq t} \left(1 - \lambda_i(t^{(k)})\right), \tag{2.2}$$

where $t^{(k)}$ are all observed times in the data when a transition occurs. This is the non-parametric Kaplan and Meier (1958) estimator for survival functions, which is an asymptotic consistent estimator that accommodates left truncation of spells (sampling ongoing spells on 01/01/2009) and right censoring of spells (sampling spells that have not ended 01/01/2015). In Section 2.6, we will study the quartiles of this duration distribution, found by setting t such that $S_{ii}(t) = 0.5$, $S_{ii}(t) = 0.75$, and $S_{ii}(t) = 0.25$.

The second step decomposes the complement of duration distribution $S_{ii}(t)$, the unconditional transition probability $1 - S_{ii}(t)$, over the different destination states. Let $S_{ij}(t)$ be the probability to be transitioned from i to j by time t, with $S_{ii}(t) +$

$\sum_{j \in \mathcal{J} \setminus \{i\}} S_{ij}(t) = 1$. The size of $S_{ij}(t)$ helps us understand the likelihood of a particular transition.

To compute $S_{ij}(t)$, we use the survival distribution $S_{ii}(t)$ and conditional transition probability from state i to j at time t, $\lambda_{ij}(t)$. Like $\lambda_i(t)$, we identify $\lambda_{ij}(t)$ from the number of individuals in state i who can transition at time t, $N_i(t)$. However, we now take those who actually do so from state i to j during the subsequent time interval dt, $D_{ij}(t)$. Then, $\lambda_{ij}(t)$ is:

$$\lambda_{ij}(t) = \mathbb{P}(t \leq T < t + dt, \text{State} = j \mid T \geq t, \text{State} = i) = \frac{D_{ij}(t)}{N_i(t)}, \tag{2.3}$$

which is the share of individuals at risk at time t who do transition from i to j. Note the link with the overall transition probability (2.1): $\lambda_i(t) = \sum_{j \in \mathcal{J} \setminus \{i\}} \lambda_{ij}(t)$.

$S_{ij}(t)$ results from that the individual is in state i at some time point before t and then made the transition from i to j. The incidence of this transition at $t^{(k)}$ is $S_{ii}\left(t^{(k-1)}\right) \cdot \lambda_{ij}\left(t^{(k)}\right)$, the product of the probability of being at risk at time $t^{(k-1)}$ (2.2) and then subsequently making the transition (2.3). $S_{ii}(t)$ explicitly corrects here for the presence of competing risks: individuals could already have transitioned to any other state before time t, making a transition at t impossible anyway. The probability to transition between start time and time t is the sum of all incidence rates:

$$S_{ij}(t) = \sum_{k: t^{(k)} \leq t} \left(S_{ii}\left(t^{(k-1)}\right) \cdot \lambda_{ij}\left(t^{(k)}\right) \right). \tag{2.4}$$

If $t \to \infty$, then $S_{ii}(t) = 0$ and we can refer to $S_{ij}(t)$ as long-run transition probabilities. In our study, a finite t_{max} approximates the long run. $t_{max} = 40$ years for individuals who have never used LTC before. We assume they do not use LTC since age 65. $t_{max} = 40$ indicates the maximum observed age: 65 + 40 = 105. We take $t_{max} = 11$ years for current non-users of LTC who have ever used LTC before because we observe LTC use between 01/01/2004 and 31/12/2014. Lastly, $t_{max} = 6$ years for all other states because that is the duration of our data on the need for LTC.

2.5.2 Mixed Proportional Hazard Specification

Next, we adopt a parametric Gompertz model with a random effect to compute an interpretable estimate of the impact of the potential availability of informal care and economic resources on the transitions. The parsimonious parametric model overcomes that the estimations are yet non-parametric and meant to be descriptive. The random effect accounts for unobserved heterogeneity and dynamic selection (see Section 2.5.3).

In parametric hazard models, the conditional transition probability λ_{ij} is the commonly adopted outcome variable (cf. equation (2.3)). This probability is scaled by dt, so the outcome reflects transitions per time unit, i.e., a hazard rate. We assume a mixed proportional hazard rate specification for each transition i to j (van den Berg, 2001):

$$\lambda_{ij}(t|\boldsymbol{x}(t),\nu_{ij}) = \frac{\mathbb{P}(t \leq T < t+dt, j|T \geq t, \boldsymbol{x}(t), \nu_{ij}, i)}{dt} = \nu_{ij} \cdot \phi_{ij}(t) \cdot exp\left(\boldsymbol{x}(t)'\boldsymbol{\beta}_{ij}\right), \quad (2.5)$$

separating into an individual random effect $\nu_{ij} \sim \Gamma\left(\frac{1}{\sigma_{ij}^2}, \frac{1}{\sigma_{ij}^2}\right)$, duration effect $\phi_{ij}(t) = exp(\gamma_{ij} \cdot t)$ and a time-invariant effect $exp\left(\boldsymbol{x}(t)'\boldsymbol{\beta}_{ij}\right)$. $\boldsymbol{x}(t)$ is a vector with time-varying and time-invariant covariates whose impact $\boldsymbol{\beta}_{ij}$ we are interested in. We will report exponentiated coefficients $exp\left(\boldsymbol{\beta}_{ij}\right)$, meaning the estimates feature a hazard ratio interpretation. If the hazard ratio exceeds unit value, then an increase $x(t)$ accelerates a transition; it delays otherwise. Besides $\boldsymbol{\beta}_{ij}$, we also estimate the parameters γ_{ij} and σ_{ij}^2.

Covariate vector $\boldsymbol{x}(t)$ consists of all variables from Table 2.1. We include each category with a separate dummy. We proxy the potential availability of informal care with (the health) of the partner and having children. In a heterogeneity analysis, we will split the child effect by distance to the closest child and (the health) of the partner. Income- and asset decile proxy for the impact of economic resources and asset- and income-dependent co-payments to the transitions (Portrait et al., 2000). We also look at homeownership to see whether individuals stay in their own house. We control for LTC needs with our split by impairment type and level of need (low or high). Lastly, we include background variables on gender, ethnicity, drug uptake (a proxy for frailty), region of residence (to account for regional variation in the supply and use of LTC, see: Duell et al., 2017), age when entering the state, and year of observation (to account for policy changes).

A strength of the hazard model is its ability to encompass variables that vary continuously. The (health of the) partner, having children, distance to the children, and region of residence can change daily in the data. Income, assets, homeownership, observation year, and drug uptake vary yearly. Age upon entry, ethnicity, and gender are time-invariant.

The term $\phi_{ij}(t)$ accounts for the effect of time on making a transition, i.e., persistence or inertia of care arrangements (Hiedemann et al., 2018). For example, individuals might get emotionally attached to a state, or the cost and benefits of a transition vary over time (Dostie and Léger, 2005). We assume a Gompertz function: $\phi_{ij}(t) = exp(\gamma_{ij} \cdot t)$, $\gamma_{ij} < (>)\ 0$ indicating that time slows down (fastens) a transition, i.e., negative (positive) duration dependence. While the interpretation of $\widehat{\phi}_{ij}$ is mostly outside our scope, in a robustness check, we will see whether our functional choice for ϕ_{ij} affects estimates $\widehat{\boldsymbol{\beta}}_{ij}$.

The random effect –frailty– ν_{ij} controls for any remaining unobserved heterogeneity, such as the leniency of the randomly assigned assessor in granting access to LTC (Bakx et al., 2020). While the assessor is randomly assigned, this could still imply initial health differences leading to different duration lengths of LTC. In line with earlier work, we assume a gamma distribution for ν_{ij} because this well-proxies any frailty distribution for high t and thus reduces misspecification error (Abbring and van den Berg, 2007).

For identification, we have to assume that observed characteristics $\boldsymbol{x}(t)$ are exogenous with respect to unobserved characteristics ν_{ij} at any t: $\boldsymbol{x}(t) \perp \nu_{ij}$. The unobserved heterogeneity in our example is exogenous due to the random assignment of an assessor. Furthermore, our large battery of controls already reduces the role of unobserved heterogeneity upfront. Additionally, we have to assume no anticipation of future transitions with current characteristics $\boldsymbol{x}(t)$. For example, individuals foresee declining health and, therefore, already start to use home-based care. Given that a mandatory eligibility assessment is gatekeeper for home-based care, we can likely exclude such anticipation.

2.5.3 Model Estimation

We use the novel estimation procedure from Chapter 5 to estimate the mixed proportional hazard specifications. As they are present in our data, this procedure allows for time-varying covariates, right-censored and left-truncated spells, and individuals repeatedly experiencing spells of the same state. Before discussing these empirical challenges, we first discuss how we can tailor the estimation procedure of Chapter 5, meant to estimate a single risk, to the competing risks setting we operate in.

We take the example from Figure 2.1 and assume that the transition from i to $j=1$ occurred at time t. $\lambda_{i1}(t)$ cannot be used directly in a log-likelihood estimation because this is a rate and not a probability. Instead, we use the probability of going from i to $j=1$ by time t, $S_{i1}\left(t\right)$, in a log-likelihood estimation. This probability solely consists the hazard rates $\lambda_{i1}(s \mid \boldsymbol{x}(s), \nu_{i1}), ..., \lambda_{i5}(s \mid \boldsymbol{x}(s), \nu_{i5})$, $0 \leq s \leq t$. To see this, we adopt a latent failure time approach, defining S_{i1} as the joint probability (cf. Putter et al., 2007):

$$S_{i1}\left(t \mid \nu_{i1}, ., \nu_{i5}, \{\boldsymbol{x}(s)\}_{s=0}^{t}\right) = \mathbb{P}(T_{i1} \leq t, T_{i2} > t, ., T_{i5} > t \mid \nu_{i1}, ., \nu_{i5}, \{\boldsymbol{x}(s)\}_{s=0}^{t}), \quad (2.6)$$

where $\boldsymbol{x}(s)\}_{s=0}^{t}$ is the covariate path between start and t, and T_{ij} is the latent transition time of $i \rightarrow j$. For example, $\mathbb{P}(T_{i1} \leq t \mid \nu_{i1}, .., \nu_{i5}, \{\boldsymbol{x}(s)\}_{s=0}^{t})$ is the marginal probability that $i \rightarrow 1$ occurs before t if there would not exist competing risks. However, these exist, so S_{i1} had to incorporate that competing risks did not occur before t: $T_{ij} > t$ if $j > 1$.

Crucial for applicability of Chapter 5, we assume independent random effects across transition types, $\nu_{i1} \perp .. \perp \nu_{i5}$, implying that the joint probability (2.6) separates into marginal probabilities for each transition apart:

$$S_{i1}\left(t|\nu_{i1}, ., \nu_{i5}, \{\boldsymbol{x}(s)\}_{s=0}^{t}\right) = \mathbb{P}(T_{i1} \leq t|\nu_{i1}, \{\boldsymbol{x}(s)\}_{s=0}^{t}) \cdot \prod_{j=2}^{5} \mathbb{P}(T_{ij} > t|\nu_{ij}, \{\boldsymbol{x}(s)\}_{s=0}^{t})$$

$$= \left(1 - \exp\left(-\int_0^t \lambda_{i1}(\tau|\boldsymbol{x}(\tau), \nu_{i1})\right) d\tau\right) \cdot \prod_{j=2}^{5} \exp\left(-\int_0^t \lambda_{ij}(\tau|\boldsymbol{x}(\tau), \nu_{ij}) d\tau\right) \quad (2.7)$$

where T_{ij} only depends on ν_{ij} and not ν_{ik}, $k \neq j$ (see Appendix B.6 for further details). The independence assumption might seem arbitrary, but we already control for much observed variation and find very comparable estimates for a model without the random effect (see Section 2.6.3). Hence, assuming a more complex correlation structure would make interpretation harder while yielding similar estimates. Note that the final step in (2.7) defines each marginal probability in terms of underlying hazard (cf. Putter et al., 2007), implying the parameters of λ_{i1},..,λ_{i5} can be inferred from S_{i1} in the example.

Due to the separability of (2.7), we can apply the estimation procedure from Chapter 5: each transition has its own sub-log-likelihood, and the log-likelihood sums them. By optimizing the sub-log-likelihood, we find the parameters for the particular λ_{ij}. The sub-log-likelihood contribution for λ_{i1} would be the marginal density at T_{i1}, i.e. $\ln\left(\frac{\partial \mathbb{P}(T_{i1} \leq t \ \mid \ \nu_{i1}, \{\boldsymbol{x}(s)\}_{s=0}^{t})}{\partial T_{i1}}\right)$, because the latent transition time T_{i1} is observed completely. The other hazard rates involve right-censored latent times T_{ij} and have contribution $\ln\left(\mathbb{P}(T_{ij} > t \ \mid \ \nu_{ij}, \{\boldsymbol{x}(s)\}_{s=0}^{t})\right)$, representing survival beyond t.

We now turn to discuss the empirical challenges other than competing risks.

Right censoring - T_{ij} can also be right-censored at study end on 01/01/2015. A sub-log-likelihood contribution consists of a marginal survival probability then.

Left truncation and dynamic selection - Spells on no LTC use sampled on 01/01/2009 are left-truncated because of their ongoing duration $t_0 > 0$. So, we observe the marginal distribution conditional upon $T_{ij} > t_0$. Furthermore, observations at $t_0 > 0$ differ from the initial sample at $t = 0$ regarding frailty. Only observations with favorable frailty, i.e. low ν_{ij}, make it until t_0, yielding observed frailty distribution: $\Gamma(\nu_{ij} \mid \{\boldsymbol{x}(s)\}_{s=0}^{t}, T_{ij} > t_0)$. Ignoring the dynamic selection implies underestimated γ_{ij} and attenuated $\boldsymbol{\beta}_{ij}$ (van den Berg and Drepper, 2016). To account for left truncation and dynamic selection, we integrate each marginal distribution over its conditional frailty distribution:

$$\mathbb{P}(T_{ij} > t|\{\boldsymbol{x}(s)\}_{s=0}^{t}, T_{ij} > t_0) = \int_0^\infty \frac{\mathbb{P}(T_{ij} > t|\nu_{ij}, \{\boldsymbol{x}(s)\}_{s=0}^{t})}{\mathbb{P}(T_{ij} > t_0|\nu_{ij}, \{\boldsymbol{x}(s)\}_{s=0}^{t})} d\Gamma(\nu_{ij}|\{\boldsymbol{x}(s)\}_{s=0}^{t}), \quad (2.8)$$

where the denominator accounts for left truncation. In Section 2.6.3 we restrict estimation to $\sigma_{ij}^2 = 0$ to assess the impact of ignoring dynamic selection on our estimates.

Time-varying covariates - The probability in (2.8) takes into account the entire history of time-varying covariates $\{\boldsymbol{x}(s)\}_{s=0}^{t}$. While we ideally know the entire covariate path, this is $\{\boldsymbol{x}(s)\}_{s=t_0}^{t}$ for the left-truncated spells. We have to assume here that $\{\boldsymbol{x}(s)\}_{s=0}^{t} = \{\boldsymbol{x}(s)\}_{s=t_0}^{t}$, i.e. covariates are constant until t_0. Identification under milder assumptions has not been studied yet (see Chapter 5). Because we include many covariates and frailty, and left truncation only happens to spells of no LTC use, we believe the impact of missing information on $\boldsymbol{\beta}_{ij}$ is limited.

Repeated spells - Some individuals are repeatedly in a particular state, implying transitions may be correlated across spells. We assume the random effect ν_{ij} to be fixed, so transition $i \to j$ is positively correlated across individual's spells. Following Chapter 5, we slightly adapt the marginal distribution in (2.8) to account for repeated spells. Repeated spells and time-varying covariates yield parameter identification similar to a fixed-effects panel regression (van den Berg, 2001).

We refer to Chapter 5 and Appendix B.6 for further details on the complete sub-log-likelihoods and model identification.

2.6 Results

2.6.1 Duration of (no) LTC use and the Subsequent Transition

Table 2.4 shows the estimation results of the non-parametric model presented in Section 2.5.1. Each row indicates a currently used LTC arrangement and need for care. The first block of columns reports the duration of being in a particular state. We refer to a single spell and do not accumulate all individual's spells to have a lifetime duration. The second block is a transition matrix and reports the long-run transition probabilities. In addition, we show the probability of no transition, because our approximation of the long-run is finite due to the study period ending in 2015.

The top row reports the statistics for individuals who never used LTC before. We assume them to be without LTC since age 65. The median duration until a transition is 14.55 years, meaning a median age at transition of 79.55 years (65+14.55). Most individuals (71%) transition into home-based care rather than directly into institutional care (15%). The other 14% pass away without using LTC. So, most individuals use LTC at some point in life, and home-based care is likely to be the first LTC arrangement.

The second row provides the duration and transition from no LTC use if the individuals used LTC before. For this sizable group of 628,320 individuals, we find shorter spells than for those not using LTC since age 65 (17.5% of sampled individuals). The shorter durations follow from that the individuals with past LTC use are on average older and possibly frailer (they are on average 6.28 years older). In contrast, transition probabilities do not differ between the two groups.

Turning to the duration of using home-based and institutional care, we compare spells with a physical and cognitive impairment. The median and 75^{th} percentile spell duration are the longest for institutional care users with a high cognitive impairment. Their median duration is 1.41 years, while the 75^{th} percentile indicates that 25% have a spell longer than 3.14 years. In contrast, their counterparts in institutional care with a high physical impairment report a median and 75^{th} percentile of 0.20 years and 0.86 years, respectively. For a lower level of need and home-based care use, we report a similar pattern of longer spells when having cognitive impairment.

Instead of comparing by health problem, we can also compare the duration of home-based and institutional care. Given a low or high cognitive impairment, we report lower median durations in home-based care than in institutional care. For individuals with a physical impairment, we confirm this finding at the 75^{th} percentile but not at the median.

But which arrangement follows after the end of a spell? A first group transitions to less specialized care. 37% and 25% of the low and high physically impaired in home-based care have no LTC use as the next state. For institutional care users, we see a similar pattern to less specialized care: 25% and 23% of the institutional care users with a low or high physical impairment transition to home-based care. We also report substantial transition probabilities of going into more specialized care or death, i.e., the opposite direction. 48% and 34% of home-based and institutional care users with a low physical impairment transition to a higher level of need. Furthermore, 19% of the home-based care users with a high physical impairment go into institutional care.

In contrast, transitions to more specialized care or death more frequently occur if having a cognitive impairment. In home-based care, 49% with a low cognitive impairment transition to a high cognitive impairment. 25% of them get another main health problem, in particular a high physical impairment (not shown). Once having a high cognitive impairment, 72% go from home-based to institutional care. In institutional care, they either go from a low to a high need (62%) or die (81% with a high cognitive impairment).

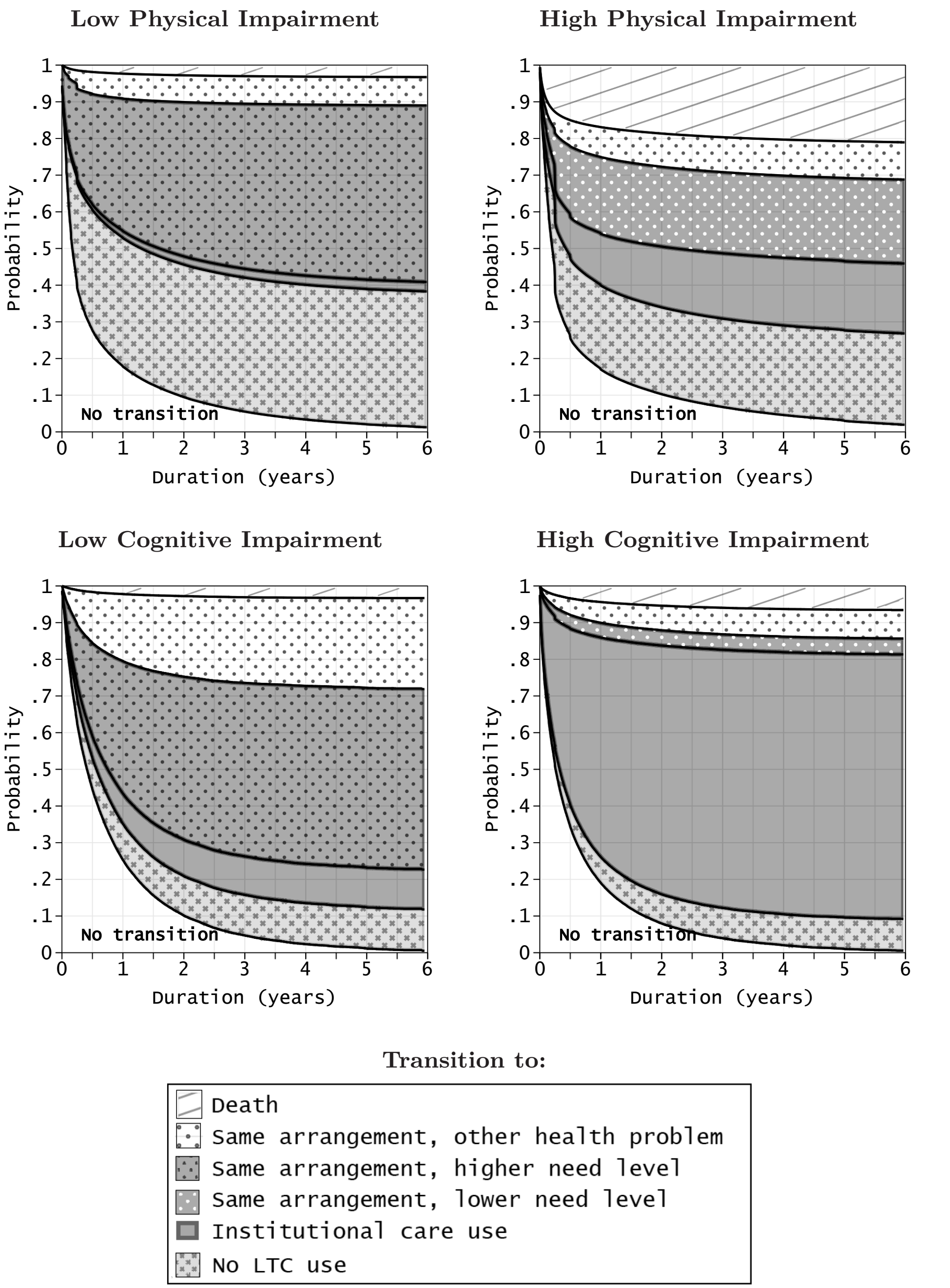
To highlight the timing of a transition, Figures 2.2 and 2.3 plot time against the transition probability of leaving home-based care or institutional care, respectively.. For this, we compare transition probabilities at $t = 0.5$ years and at $t = 6$, where the probabilities at $t = 6$ are also reported in Table 2.4. Generally, we observe that most transitions to less specialized care occur before time $t = 0.5$, whereas the transition to more specialized care or death also frequently occurs after $t = 0.5$. To exemplify this, consider those with a low physical impairment in home-based care. At $t = 0.5$, 33% have moved to no LTC use, while this is 37% at $t = 6$. On contrary, 30% has a high need at $t = 0.5$, increasing by 18 pp. to 48% at $t = 6$. This finding suggests that longer spells mainly end up in more specialized care or death.

Table 2.4: Spell Duration and Long-run Transition by LTC Arrangement and Need for LTC

	Spell duration (years)			**Long-run transition probability to next state:******								
	Median	25^{th} pct.	75^{th} pct.	No transition****	No LTC use	Home-based care use	Institutional care use	Same arrangement but lower need level	Same arrangement but higher need level	Same arrangement but other health problem	Death	Individuals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Current state:												
No LTC use: never used LTC before*												
All	14.55	8.96	19.54	0.00	X	0.71	0.15	X	X	X	0.14	2,888,760
No LTC use: ever used LTC before												
All	1.84	0.58	4.72	0.07	X	0.70	0.16	X	X	X	0.08	628,320
Home-based care use												
Physical impairment												
Low**	0.17	0.07	0.60	0.01	0.37	X	0.03	X	0.48	0.08	0.03	452,706
High	0.21	0.07	0.52	0.02	0.25	X	0.19	0.23	X	0.10	0.21	592,066
Institutional care use												
Physical impairment												
Low	0.23	0.08	1.63	0.04	0.13	0.25	X	X	0.34	0.09	0.16	109,304
High	0.20	0.07	0.86	0.04	0.14	0.23	X	0.05	X	0.11	0.44	245,412
Home-based care use												
Cognitive impairment												
Low	0.41	0.15	1.01	0.01	0.11	X	0.11	X	0.49	0.25	0.03	52,690
High	0.26	0.07	0.75	0.01	0.09	X	0.72	0.04	X	0.08	0.07	78,021
Institutional care use												
Cognitive impairment												
Low	0.92	0.23	2.09	0.01	0.04	0.07	X	X	0.62	0.08	0.17	34,203
High	1.41	0.38	3.14	0.07	0.04	0.04	X	0.01	X	0.03	0.81	137,902

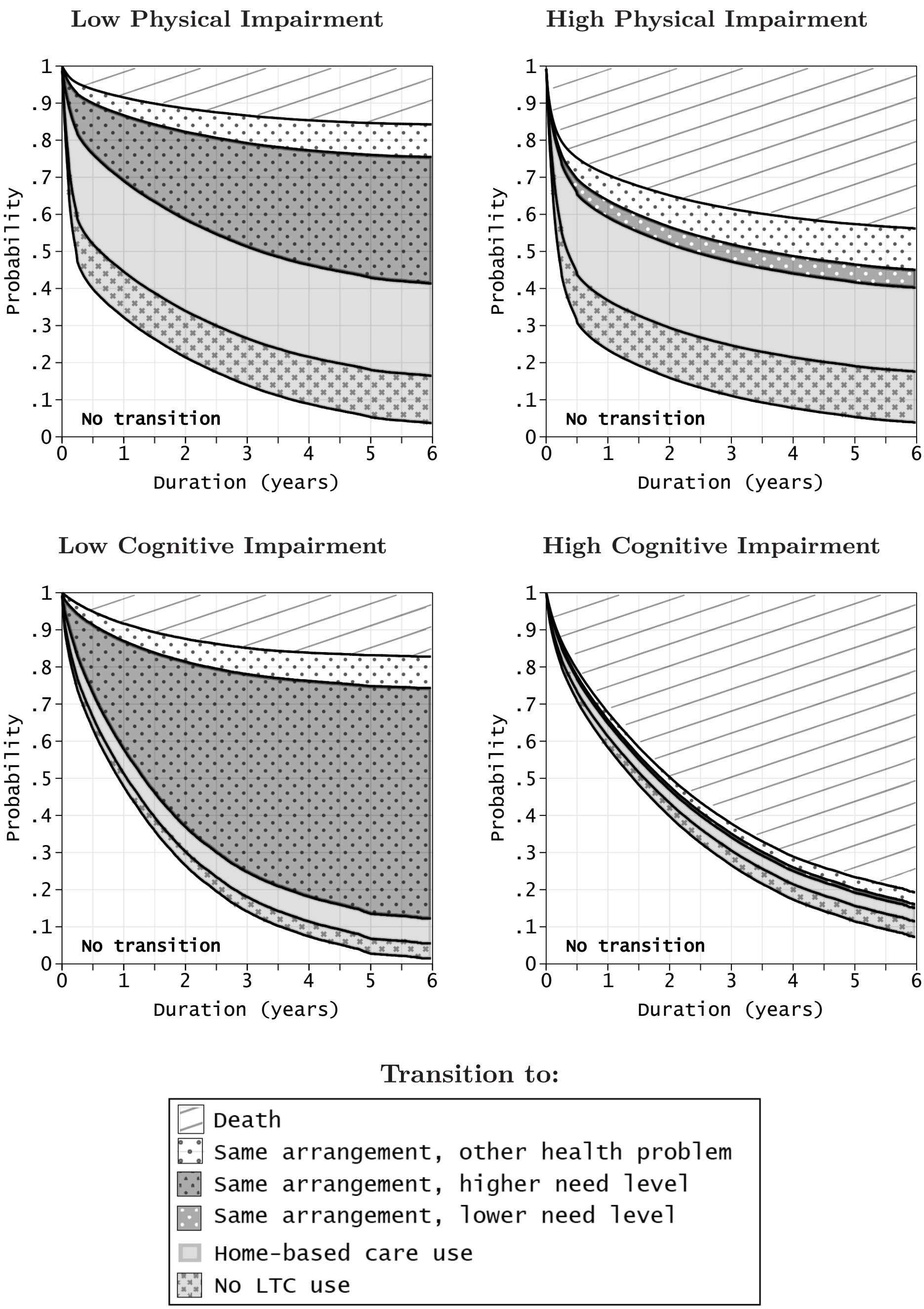
Notes: * We measure 'never use' as of age 65; ** Low need: the entitled hours of care is below the reported median in Table 2.3, and vice versa for a high need; *** These are the long-run transition probabilities from Section 2.5.1: $\lim_{t\to\infty} S_{ij}(t)$. Due to the study end, our approximation of the long-run is finite: $t_{max} = 40$ years if a non-user of LTC never used LTC before; $t_{max} = 11$ years if a non-user of LTC ever used LTC before; $t_{max} = 6$ years for the other spells; **** Some transitions occur after our long-run threshold t_{max} and thus remain unobserved.

Figure 2.2: Current Home-based Care Users: Transitions over Time by Arrangement and Need for LTC



Notes: The probabilities at time $t = 6$ are also reported in Table 2.4.

Figure 2.3: Current Institutional Care Users: Transitions over Time by Arrangement and Need for LTC



Notes: The probabilities at time $t = 6$ are also reported in Table 2.4.

2.6.2 Determinants of Transitions across LTC Arrangements

Table 2.5 shows the results for the transitions to more specialized LTC arrangements for the mixed proportional hazard model (cf. Section 2.5.2). If someone never used LTC before and has a healthy partner, the likelihood of going into home-based care is at any time 45% ($(1-0.55)\cdot 100\%$) lower when compared to being single (column 1). This delay of home-based care use suggests possible informal care provided by the healthy spouse. However, the effect shrinks to 21% if someone has ever used LTC before, possibly because the (past) health problems are harder to cover with informal care.

Having a healthy partner delays institutional care use for home-based care users with a high physical impairment. Surprisingly, it accelerates the use of institutional care for those with a low or high cognitive impairment (hazard ratios: 1.82 and 1.19). Cognitive impairments are different in that they could mean assistance with (cognitive) tasks throughout the entire day, putting a large care burden on the spouse. In turn, the partner might sooner alert the relevant authorities that institutional care is required. In addition, singles lack that 24-hour assistance and only singles who can (partially) take care of themselves use home-based care. As a consequence, the reference group contains relatively healthy singles.

The risk of using more specialized LTC is higher if the partner already uses LTC; the coefficients on 'partner in home-based care' and 'partner in institutional care' exceed unit value (reference: having a healthy partner). Moreover, the largest coefficients are found if the LTC arrangement at risk and partner's arrangement are the same, implying that couple members select into the same LTC arrangement. The individuals at risk possibly provided informal care themselves and thereby faced the inherent emotional and physical burden. In turn, their increased frailty exposes them to a higher risk on using LTC. Also, the partner might provide less informal care due to own health problems.

Albeit weaker than the effect of having a healthy partner, having children also delays the use of more specialized LTC. This suggests that children are a secondary source for informal care provision. To see how much of the effect is driven by access to informal care, we split the effect by distance to the children and partner status (see Appendix Tables B.7 and B.10). We test for the potential heterogeneity using a likelihood-ratio test statistic χ^2 reported in Table 2.5. For low and high physically impaired, having children has a stronger delaying effect if the closest child lives in the same municipality rather than in another municipality (further away). Also, for high physically impaired, the child effect is stronger for singles than for those with a healthy partner, suggesting that sources of informal care provision are substitutable. For the other transitions, i.e. different needs, we find no evidence for substitutability between child- and partner-provided care.

Lastly, we see that being a homeowner or having more assets and income delays transitions into more specialized care. The effect of being a home owner might in part reflect a preference for aging in one's own house. Furthermore, the impact of income and assets indicate that individuals might look for private alternatives for formal LTC.

Table 2.6 shows the estimation results for a transition into less specialized arrangements of LTC. In essence, the effects work in the same direction as in Table 2.5. While they delayed the use of more specialized care, we find that a return to less specialized care is accelerated when having a healthy partner, having children, being homeowner, having higher income, and having higher assets. Again, the availability of informal care, a preference for aging in one's own house, and the affordability of private care can explain these effects. Also, the signs on 'partner in home-based care' and 'partner in institutional care' indicate that individuals (start to) use the same LTC arrangement as their partner.

The findings further highlight that having children significantly impacts the transitions only if individuals have a low or high physical impairment. As these impairments possibly involve treatable symptoms and curative care, and children's caregiving skills can be sufficient for a parent to return to less specialized care. In turn, cognitive impairments probably require more skilled care that children less likely provide. Also, given the low incidence, we will not further discuss the cognitively impaired.

Particularly, the effect of having children is stronger for singles and if the closest child lives in another municipality (see Appendix Tables B.8 and B.11). We only see a significant and substantial distance effect for the transition from home-based care to no LTC use. From the preceding analysis we know that the risk of going from no LTC use to home-based care is higher if children live in another municipality. This could imply that those individuals are less frail when entering home-based care and thus more likely to return to no LTC use. For singles, the effect is stronger because the child is their primary source of informal care.

Rather than across LTC arrangements, Table 2.7 presents the transition within an arrangement of LTC, so from a low to a high need or vice versa.[6] Columns (1) and (3) suggest that a transition from a low to a high need in home-based care is delayed if having a healthy partner (hazard ratios 0.91 and 0.96). While this could indeed reflect informal care provision, we instead find that having a healthy partner accelerates a transition from a low to a high need in institutional care (columns 5 and 7). This contrasting finding may in part reflect selection into some arrangements (Table 2.5): partnered individual could be frailer in LTC because they initially postpone the use of more specialized arrangements. Consequently, the frail partnered individuals more frequently transition to a higher need within their arrangement.

In fact, we can explain all covariate effects in Table 2.7 with this selection due to observable characteristics. While we find evidence that having children delays LTC use and that this fosters the use of less specialized arrangements, a child effect is insignificant or reversed for transitions within arrangements. Similarly, the effect of income and homeownership is mostly opposed to what we find for transition across arrangements.

Lastly, we find that duration dependence is negative for almost all transitions from LTC use, implying fewer transition at longer durations, and positive for a transition from no LTC use. Hiedemann et al. (2018) also documents the negative duration dependence, which they call 'inertia'. The duration dependence of LTC users is, however, positive if they are high cognitively impaired home-based care users going into institutional care (column (6), Table 2.5), which we can explain by their required 24-hours supervision,

[6]While reported for completeness, we do not further discuss transitions (4), (6) and (8) because of low incidence.

making them increasingly likely to transition over time. The duration dependence of no LTC users who never used LTC reflects an age effect (see Section 2.6.1) and is positive (column (1), Table 2.5): older individuals are more likely to start using home-based care.

2.6.3 Sensitivity Analysis

To assess the robustness of our findings, we consider how estimates are affected by accounting for dynamic selection due to unobserved characteristics.[7] To test for the dynamic selection, we report for each hazard rate the estimated frailty variance σ^2_{ij} and accompanying likelihood-ratio test with H_0: $\sigma^2_{ij} = 0$. We report the model estimates restricting $\sigma^2_{ij} = 0$ in Appendix Tables B.16 to B.18.

Almost all frailty variances are significant, implying that there is unobserved heterogeneity and hence dynamic selection. For this, we compared the sub-log-likelihood of the unrestricted model ll_{ur} ($\sigma^2_{ij} \neq 0$) to that of the restricted model ll_r ($\sigma^2_{ij} = 0$) using likelihood-ratio statistic $2 \cdot (ll_{ur} - ll_r) \overset{a}{\sim} \chi^2_1$.[8] We chose to control for dynamic selection because its omission could lead to attenuated hazard ratios (i.e. closer to unit value) and underestimated duration dependence γ_{ij} (van den Berg and Drepper, 2016). However, the hazard ratios with restricted $\sigma^2_{ij} = 0$ are not substantially different (Appendix Tables B.16 to B.18). We do indeed report higher values of γ_{ij}, and duration dependence in columns (2) and (6) of Table 2.5 even becomes positive if controlling for frailty.

The similar hazard ratios could be explained by our rich set of covariates, making the role of frailty less important. This observation makes us confident that we properly account for dynamic selection and do not have to allow more flexible (correlated) frailty structures, making interpretation more complex but probably leaving results unchanged.

Another concern is assuming a baseline hazard function, which could affect hazard ratio estimates. To investigate this, we estimate a Cox proportional hazard model with $\sigma^2_{ij} = 0$, leaving the baseline hazard unspecified. We find similar results as before (Appendix Tables B.13 to B.15). Hence, our results are robust to the Gompertz assumption. We prefer our estimates because γ_{ij} allows us to study the sign of duration dependence.

[7] Note that in Section 2.6.2 we rather discuss selection due to observed characteristics.
[8] For completeness, we also report the log-likelihood but do not use this for hypothesis testing.

Table 2.5: Hazard Ratio Estimates for Transitions to a more specialized LTC Arrangement

From:	No LTC use Never used LTC before -	No LTC use Ever used LTC before -	Home-based care Physical impairment Low need	 High need	Home-based care Cognitive impairment Low need	 High need
To:	Home-based care		Institutional care		Institutional care	
	(1)	(2)	(3)	(4)	(5)	(6)
Partner (ref: single)						
Partner does not use LTC	0.55***	0.79***	1.00	0.80***	1.82***	1.19***
Partner (ref: partner does not use LTC)						
Partner uses home-based care	3.10***	1.98***	1.03	1.14***	1.16**	0.98
Partner uses institutional care	2.53***	1.47***	4.18***	3.05***	1.80***	3.00***
Children (ref: no children)						
Has children	0.91***	0.93***	0.86***	0.90***	0.76***	0.89***
Income (ref: lowest income decile)						
Highest income decile	0.72***	0.90***	0.98	0.92***	0.97	0.83***
Assets (ref: lowest asset decile)						
Highest asset decile	0.95***	0.91***	0.94	0.92***	1.02	0.83***
Homeowner (ref: renter)						
Homeowner	0.86***	0.87***	0.99	0.93***	0.91**	0.90***
Duration dependence						
γ_{ij}	0.11***	0.03***	-0.30**	-0.46***	-0.15***	0.19***
Individual-shared frailty[5]						
σ^2_{ij}	0.21***	0.43***	2.31***	0.43***	1.29***	1.22***
Individuals	2,888,623	628,283	452,678	592,030	52,687	78,012
Spells	2,888,623	867,987	635,791	829,767	62,357	86,727
Transition probability (%)	71	70	3	19	11	72
χ^2 (distance to the closest child)[1]	24.75***	10.25**	38.60***	136.06***	0.05	3.13
χ^2 (child effect × partner status)[2]	1.00	6.25	3.23	23.13***	0.99	0.54
Sub-log-likelihood (cf. (2.8))[3]	-1,854,960.00	-927,003.81	-54,628.71	-264,822.75	-15,846.55	-34,231.73
Log-likelihood[4]	-2,905,382.19	-1,449,380.42	-594,402.12	-1,090,061.55	-98,861.27	-91,798.37

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] Likelihood-ratio test (df=1) whether the child effect significantly differs by the distance to the closest child (lives in same or another municipality). Table B.7 in the appendix shows the unrestricted results. [2] Likelihood-ratio test (df=3) whether the child effect significantly differs by partner status (interaction between the partner status and child effect). Table B.10 in the appendix shows the unrestricted results. [3] The sub-log-likelihood is based on marginal survival probability (2.8). [4] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks. [5] The estimates for the restricted models ($\sigma^2_{ij} = 0$) are shown in Table B.16.

Table 2.6: Hazard Ratio Estimates for Transitions to a less specialized LTC Arrangement

From:	Home-based care Physical impairment		Home-based care Cognitive impairment		Institutional care Physical impairment		Institutional care Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	No LTC use		No LTC use		Home-based care		Home-based care	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	1.54***	1.25***	2.28***	1.41***	2.15***	1.33***	9.11***	3.32***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	0.48***	0.59***	0.61***	0.75***	1.28***	1.46***	1.45***	1.61***
Partner uses institutional care	0.58***	0.76***	0.56***	1.03	0.15***	0.38***	0.09***	0.30***
Children (ref: no children)								
Has children	1.14***	1.16***	0.93	1.03	1.12***	1.11***	1.15	0.98
Income (ref: lowest income decile)								
Highest income decile	1.48***	1.41***	1.39***	1.73***	1.72***	1.79***	1.85***	1.83***
Assets (ref: lowest asset decile)								
Highest asset decile	1.02	1.05***	0.83*	1.08	1.10**	1.01	1.01	0.82*
Homeowner (ref: renter)								
Homeowner	1.11***	1.13***	1.14***	1.12**	1.32***	1.24***	1.77***	1.36***
Duration dependence								
γ_{ij}	-1.68***	-1.24**	-0.84***	-0.59***	-4.75***	-3.29***	-4.93**	-6.06***
Individual-shared frailty[5]								
σ^2_{ij}	0.64***	0.89***	2.94***	2.39***	0.39***	0.20***	3.78***	3.98***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	37	25	11	9	13	14	4	4
χ^2 (distance to the closest child)[1]	161.81***	168.63***	11.08***	19.80***	3.18	9.22**	0.81	4.93*
χ^2 (child effect × partner status)[2]	13.22**	37.69***	4.24	3.49	0.96	4.46	0.10	2.13
Sub-log-likelihood (cf. (2.8))[3]	-157,466.30	-221,501.31	-15,214.30	-17,437.40	-22,885.14	-61,849.45	-4,101.64	-13,342.60
Log-likelihood[4]	-594,402.12	-1,090,061.55	-98,861.27	-91,798.37	-187,087.71	-333,576.20	-61,550.80	-198,680.01

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] Likelihood-ratio test (df=1) whether the child effect significantly differs by the distance to the closest child (lives in same or another municipality). Table B.8 in the appendix shows the unrestricted results. [2] Likelihood-ratio test (df=3) whether the child effect significantly differs by partner status (interaction between the partner status and child effect). Table B.11 in the appendix shows the unrestricted results. [3] The sub-log-likelihood is based on marginal survival probability (2.8). [4] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks. [5] The estimates for the restricted models ($\sigma^2_{ij} = 0$) are shown in Table B.17.

Table 2.7: Hazard Ratio Estimates for Transitions within LTC Arrangements

From:	Home-based care				Institutional care			
	Physical impairment		Cognitive impairment		Physical impairment		Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	High'need	Low need	High need	Low need	High need	Low need	High need	Low need
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	0.91***	0.94***	0.96*	1.75***	2.67***	0.55***	3.38***	0.57***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	1.04***	1.02	1.10***	0.98	1.20**	1.73***	1.18*	1.90***
Partner uses institutional care	1.42***	1.08***	1.36***	0.62**	0.23***	1.83***	0.29***	1.83***
Children (ref: no children)								
Has children	1.00	0.97***	1.02	0.87**	0.95*	0.96	1.01	0.99
Income (ref: lowest income decile)								
Highest income decile	1.14***	0.92***	1.12**	1.01	1.53***	1.12	1.22**	1.01
Assets (ref: lowest asset decile)								
Highest asset decile	0.96**	1.01	1.03	0.90	0.92*	0.91*	1.00	1.12
Homeowner (ref: renter)								
Homeowner	1.05***	0.99	0.97*	1.16***	1.15***	1.06*	1.12***	1.29***
Duration dependence								
γ_{ij}	-0.21***	-1.34***	-0.11***	-1.90***	0.07***	-1.45***	0.52***	-2.55***
Individual-shared frailty[5]								
σ^2_{ij}	1.28***	0.28***	0.34***	1.72***	1.47***	0.00	1.24***	0.50
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	48	23	49	4	34	5	62	1
χ^2 (distance to closest child)[1]	4.13*	6.19*	2.53	0.05	21.64***	1.83	1.00	1.30
χ^2 (child effect×partner status)[2]	4.81	12.06**	10.34*	8.63*	5.77	2.55	7.63	0.15
Sub-log-likelihood (cf. (2.8))[3]	-258,222.64	-267,997.66	-35,515.07	-10,499.89	-72,535.00	-38,217.66	-31,392.74	-6,834.48
Log-likelihood[4]	-594,402.12	-1,090,061.55	-98,861.27	-91,798.37	-187,087.71	-333,576.20	-61,550.80	-198,680.01

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] Likelihood-ratio test (df=1) whether the child effect significantly differs by the distance to the closest child (lives in same or another municipality). Table B.9 in the appendix shows the unrestricted results. [2] Likelihood-ratio test (df=3) whether the child effect significantly differs by partner status (interaction between the partner status and child effect). Table B.12 in the appendix shows the unrestricted results. [3] The sub-log-likelihood is based on marginal survival probability (2.8). [4] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks. [5] Estimates for the restricted models ($\sigma^2_{ij} = 0$) are shown in Table B.18.

2.7 Discussion and Conclusion

We provide evidence on pathways within the Dutch long-term care (LTC) system, given differences in the need for care, potential informal caregiving, and economic resources. We novelly study the incidence and timing of transitions across LTC arrangements, leading to insights that cannot be determined with cross-sectional (static) studies, i.e. when LTC arrangements and personal circumstances are not allowed to change. For the use of home-based or institutional care, we find that the duration is shorter if having a physical impairment; also, the use of LTC when having a physical impairment more often ends up in the use of less specialized LTC. We find that having a healthy partner *delays* institutional care use for home-based care users with a physical impairment, but, surprisingly, *accelerates* this for home-based care users with a cognitive impairment. Also, we find that having more income and financial assets or an own house *delays* the use of more specialized care and *accelerates* the use of less specialized care.

The ambiguous role of potential informal caregiving that we find is in line with prior research. Bonsang (2009) documents that informal care provision substitutes formal personal care at home, which we find for physically impaired. Our contrasting finding for cognitively impaired aligns with other work documenting that informal care provision not necessarily substitutes the use of more specialized institutional care (e.g., Bergeot and Tenand, 2023). We highlight that the initial delay of home-based care use by partnered individuals implies that they could be frailer when entering home-based care and thus subsequently could go faster into institutional care than singles.

Further, our findings highlight a possible emotional and physical burden of informal caregiving (for a review, see: Bom et al., 2019). First, as suggested by a weaker healthy partner effect, informal caregiving seems harder if the individual without LTC use has a past of LTC use. In addition, our findings reveal joint LTC use by partners, possibly implied by the detrimental consequences of informal care provision on physical and mental health outcomes. Lastly, because of their stronger effect on some transitions, we implicitly show that children of singles are exposed to a larger emotional and physical burden of informal care provision. These findings opt for policy tailored to specific groups of partnered individuals and children, e.g. by having specific care leave arrangements that unburden children of a single parent.

Our results should be seen as associations rather than as causal effects of informal care provision. In the analysis, we proxy for potential informal caregiving with marital status and having children. This effect does not mean that the family actually provides informal care because that is an individual decision. Instead, our estimates should be seen as an intention-to-treat (ITT) effect and are a lower bound to the effect of actual provision of informal care. Our large set of covariates and robustness checks make us confident that we find the mentioned ITT effects and no spurious effects. We encourage future research to shine more light on the causality and distinct channels behind our results, e.g. by looking at the impact on LTC transitions of having a medical professional in the family (Chen et al., 2022b).

Nevertheless, we can speculate on the broader implications of our findings on the determinants of LTC use. First, considering the high-risk groups of using LTC, it could be prudent to invest in informal care training for partners of cognitively impaired home-based care users or to invest in home adaptions for singles without children enabling them to age-in-place. Also, these individuals might be given priority on waiting lists for institutional care. Furthermore, our established risk factors advocate for a broader risk assessment of individuals not yet using LTC, e.g. by having algorithmic risk assessments or preventive home visits to those being aged 80+ years old, which already exist in some countries such as Denmark (Forebyggende hjemmebesøg, see: Vass et al., 2007). The outcomes of the assessment can prevent unnecessary institutional care spells, prevent other family members from being exposed to a period of stressful caregiving, and prevent the healthy partner from ending up in LTC.

Also, our findings show ways to update existing LTC arrangements. First, delayed institutional care use by homeowners provides scope to treat living in a public institution and caregiving as separate arrangements. As already is the case in the Netherlands, the elderly can then pay for living themselves, e.g. if wanting extra comfort in private institutions, while the caring component is paid for by the government. However, there should be enough places outside formal institutions where the care can be received, especially for cognitively impaired which we find to have a longer lead time. Lastly, more can be done to stimulate couples to age in place. A way to do so is to let them pay a double co-payment if they receive care within a formal institution and a single co-payment if they choose to receive this outside an institution. We encourage future work to shine further light on these incentives to age-in-place.

To summarize, we show distinct pathways across LTC arrangements and find its determinants, including having a cognitive or physical impairment, and availability of informal care. Also, the role of informal care differs given a physical or cognitive impairment. The findings advocate for a broad(er) risk assessment on the potential use of LTC arrangements. This way, more risk groups for using LTC are identified and policy can be developed that keeps public LTC systems viable in an era of aging.

Chapter 3

Combining Insurance Against Old-Age Risks to Accommodate Socioeconomic Differences in Long-Term Care Use and Mortality

This chapter is joint work with Rob Alessie, Max Groneck, and Raun van Ooijen. We thank Gerard van den Berg, Damiaan Chen, Franca Glenzer, Maarten Lindeboom, Joanna Tyrowicz, Jiayi Wen, Bram Wouterse, Nicolas Ziebarth, and seminar participants at various seminars and conferences for helpful comments. We thank Netspar for financing data access through two theme grants: 'Uncertainty over the life cycle: implications for pensions and savings behavior' (2017-2020) and 'Health and labor market uncertainty over the life cycle: The impact on households' risk capacity and retirement income adequacy' (2022-2024).

3.1 Introduction

Within the context of aging societies, the proper design of old-age insurance systems becomes increasingly salient. In private markets, a strong tendency to underinsure longevity risk and the risk of needing long-term care (LTC) has been empirically observed, often referred to as the annuity puzzle and LTC insurance puzzle, see Lambregts and Schut (2020) for a review. Adverse selection is one explanation for the limited market sizes, arising when those with above-average life expectancy more often buy annuities, and those with high expected long-term care needs more often buy LTC insurance.[1] Another explanation for the low demand for LTC insurance is the availability of informal care from the spouse or other family members (Mommaerts, 2024). To reduce adverse selection incentives, combining insurances to hedge long-term care- and mortality risks when they are negatively correlated has been proposed.[2] Despite its theoretical potential, old-age insurances that combine LTC insurance with annuities are still not very common, and its feasibility is poorly understood.[3]

This chapter quantifies socioeconomic and socio-demographic differences in long-term care use and mortality and evaluates the implications for combined insurance against these risks. We theoretically derive conditions for a combined insurance that minimizes adverse selection incentives. We then quantify differences in long-term care use and mortality employing a multi-state model using unique Dutch administrative data on exposure to formal long-term care use and mortality risk of over 3 million individuals aged 65 and above. Using these results we evaluate the factors that are important for a combined life care annuity.

[1] Cf. Finkelstein and Poterba (2004). However, the size of adverse selection problems in the LTC insurance market is subject to debates, cf. Brown and Finkelstein (2007), Brown and Finkelstein (2008), and Boyer et al. (2020), among others. Most notably, preference heterogeneity – low risks have a high preference for insurance – might even imply advantageous selection, cf. Finkelstein and McGarry (2006). We assume homogeneous preferences in this chapter and take the positive correlation of private information and insurance coverage as given.

[2] Murtaugh et al. (2001), Brown and Warshawsky (2013), Webb (2009), Solomon (2022),De Donder et al. (2022).

[3] The American Association for Long-Term Care Insurance highlights the favorable experience with LTC combination products over stand-alone LTC insurance; however, the number of policies sold remains limited, see: `https://www.aaltci.org/long-term-care-insurance/learning-center/ltcfacts-2019.php` and `https://www.aaltci.org/linked-benefit-faqs/` [*retrieved on*: October 20^{th}, 2023].

The point of departure for our study is the well-known socioeconomic gradient in longevity, according to which individuals with lower income die earlier than those with high incomes, cf. Deaton (2002). For long-term care use, low income individuals tend to be less healthy at older ages and require more LTC. Both of these socioeconomic gradients are significantly affected by gender and marital status. Importantly, the socioeconomic gradients and its heterogeneity greatly matter for the design of insurances. The implications apply to both private and public insurance systems.

For public social security, socioeconomic differences in mortality imply a redistribution of benefits from lower incomes, who die early, to higher incomes, who receive benefits for a more extended retirement period.[4] In private insurance markets, the implied differences in premium returns followed by inequalities in mortality can yield adverse selection problems. Pricing at average life expectancy would imply an actuarially unfair premium for higher income individuals, contributing to under-annuitization (Brown and Finkelstein, 2008). The picture is reversed for LTC insurance. Here, individuals with lower income tend to require more long-term care. This gradient imposes an opposite redistribution of benefits via LTC insurance from higher to lower income individuals in public insurance systems. In a private insurance market, medical underwriting and potentially low take-up rates of private LTC insurance might be the consequence (Braun et al., 2019). The negative correlation between longevity and long-term care needs implies that individuals with lower incomes are seen as lower risk types in the annuity market and higher risk types in the LTC insurance market, with the opposite for high income individuals. From a private insurance perspective, combining the two insurances to hedge these risks is appealing to reduce adverse selection problems.

In our study, to understand the implications of socioeconomic and socio-demographic differences in long-term care use and mortality for combined old-age insurance, we extend the standard adverse selection model of Einav et al. (2010). We formally derive an optimal

[4] For the US social security system, Groneck and Wallenius (2021) show that the (intended) progressivity turns regressive once the differences in life expectancy over socioeconomic status are considered. In Chapter 4, we quantify the welfare effects of social insurance programs stemming from inequalities in long-term care needs and mortality in a dynamic structural model.

combination of LTC insurance and pension annuity that minimizes adverse selection. More specifically, we obtain an expression showing how the optimal combination of insurance depends on three factors: (1) the mean duration in each of the two states 'no-long-term care use' and 'long-term care use', (2) the variances of money's worth for stand-alone LTC and annuity insurance over individual types, and (3) the correlation between the money's worth of the two insurances.

Next, we establish stylized facts on the socioeconomic gradients in longevity and LTC. We estimate the joint distribution of long-term care use and remaining life expectancy at age 65 by lifetime income, gender, and marital status. Exploiting rich administrative data provides us with sufficient observations also for the oldest-old, which is crucial for reliably estimating long-term care incidences. We develop a multi-state model and employ a recently developed method to estimate the underlying mixed proportional hazard rates (Chapter 5), incorporating frailty and allowing for time-varying covariates to capture the transition from being married to a single-person household.

We study the impact of these estimates for the design of insurances. We quantify adverse selection incentives for stand-alone pension annuity and LTC insurance measured as any deviation of the individual risk from actuarial fair pricing (i.e., non-zero premium returns). We then analyze the optimal combination of the insurances that minimizes adverse selection stemming from socioeconomic and socio-demographic inequalities. Our results allow us to understand the feasibility of combined old-age insurance for different socioeconomic groups and its determinants.

We have two main contributions. First, we establish new stylized facts simultaneously documenting a positive gradient in longevity and a negative gradient in long-term care use over lifetime income. We highlight to what extent informal care possibilities – proxied by having a spouse – affect these differences. Previous literature has studied this in isolation and focused on formal care only (cf. Kalwij et al. (2013) and Rodrigues et al. (2018), for example). Second, we theoretically and empirically study the optimal combination of insurances by determining the optimal specific benefit level for each future state of the world. Two additional factors – the relative duration and the heterogeneity in risks – are shown to be important for a combination of insurances, next to the the well-known negative correlation of risks. We further emphasize the need for group-specific premia to reduce adverse selection problems. Previous literature studied given insurance products (e.g. Murtaugh et al., 2001) and focused on the negative correlation between risks as the precondition for a successful combination, cf. Webb (2009), Solomon (2022).

We find substantial socioeconomic inequalities in long-term care use and mortality. The difference in remaining life expectancy at age 65 between the bottom and the top lifetime income quintile is 4.0 years for men and 2.3 years for women. Women in the bottom income quintile spend an additional 1.7 years in long-term care after age 65 than those in the top income quintile, while for men, this difference is 1.1 years. Hence, gender matters for the income gradient, which is stronger for men in terms of mortality, but stronger for women in terms of long-term care. Regarding informal care possibilities, proxied by having a spouse, being married reduces long-term care duration by 22% for men and it substantially flattens the socioeconomic gradient. At the same time, this is far less pronounced for women, potentially due to the high likelihood of outliving the spouse.

The implied consequences for valuing insurances show for LTC insurance a large positive premium return of +30 percent for the lowest income quintile and a negative premium return of –17% for the highest income. The gradient of the premium returns for annuities is reversed but flatter and ranges from –9% to +4% for the lowest and the highest income quintile.

Guided by our theory we determine the optimal combination of annuity and LTC insurance. Combining both insurances reveals that this is unfeasible with a uniform premium for everyone due to large gender-differences in long-term care use and mortality and a positive correlation of premium returns over gender. Group-specific premia yield large differences for the optimal insurance products over gender and marital status. Our results suggest that a life care annuity seems feasible for single men and women but less so for married men and women, due to unfavorable variances and correlations of the risks for these groups.

Our analysis is not limited to a combination of annuities and LTC insurance but holds more general for any bundling of insurances. Bundling risks in insurances is a widespread practice ranging from life-insurance with LTC-rider to home-car insurances, see Eling and Ghavibazoo (2019) and Solomon (2022) for further examples of combining insurances. Our results can help guiding the design of such bundled insurance products and inform about its feasibility to reduce adverse selection problems.

The remainder of the chapter proceeds as follows. The following Section 3.2 gives a brief literature review. Section 3.3 presents the theoretical model and Section 3.4 describes institutional details, the data, and the empirical approach. Section 3.5 presents the results, Section 3.6 discusses the main results and Section 3.7 concludes.

3.2 Literature

This chapter combines three related strands of literature studying (1) the causes of the annuity- and LTC insurance puzzles, (2) the potential to bundle insurances, and (3) the estimation of socioeconomic and socio-demographic differences in long-term care and mortality.

This chapter focuses on adverse selection and informal care possibilities as two factors affecting the low demand for annuities and LTC insurance. However, many other explanations for the so-called annuity- and LTC insurance puzzle have been put forward. Most notably, the risk for high out-of-pocket expenses for health-related expenses and bequest motives imply a tendency to hold sufficient liquid assets to prevent hitting the borrowing constraint, which implies low annuitization, cf. Lockwood (2018), Ameriks et al. (2018). Davidoff (2009) point to the importance of home equity, which can serve as a substitute for annuities and LTC insurance to some extent. Reichling and Smetters (2015) emphasize the role of correlated risks introduced via health shocks that simultaneously affect longevity and uninsured medical costs as a source for low valuation of annuitization. Pauly (1990) and Zweifel and Strüwe (1998), and more recently, Mommaerts (2024) and Coe et al. (2015) stress the importance of informal care availability for the low demand for private LTC insurance.

Most related to our approach are studies evaluating loads or the money's worth of insurance, which we will also apply in our analysis to evaluate adverse selection problems. Brown and Finkelstein (2007) find significant loads in the long-term care insurance market pointing to actuarial unfair pricing which varies by demographic and socioeconomic characteristics. Brown and Finkelstein (2008) determine large differences in the willingness to pay for insurance in a life cycle setting given the current government welfare system between insurances with these loads or without. Similarly, Mitchell et al. (1999) estimate the willingness to pay for actuarially fair pricing in the annuity market using a money's worth concept.

Theoretically, the extension of the standard adverse selection model to multiple risks to compare separate versus 'umbrella' contracts has been studied by Fluet and Pannequin (1997) focusing on the relationship between partial coverage and low-risk exposure under multiple risks, Gollier and Schlesinger (1995) analyzing the optimal structure of deductibles , and Picard (2020) studying optimal risk splitting in multidimensional screening models. Webb (2009) and Solomon (2022) investigate life care annuities directly. Webb (2009) sets up an adverse selection model in the presence of preference heterogeneity and unfair pricing, showing that the bundled product can be welfare-improving. Closely related to our theoretical model is Solomon (2022), who shows that the correlation structure and whether selection is adverse or advantageous are the key elements for the welfare effects of bundling. Solomon (2022) does not analyze an optimal combination of insurances, though.

Murtaugh et al. (2001) and Brown and Warshawsky (2013) have empirically studied the attractiveness of life care annuities relative to single products by determining how a combined product can be offered with a lower premium and less strict medical underwriting to attract more people. De Donder et al. (2022) show that a life care annuity can yield advantageous selection solely assuming differences in agent's risks.

This chapter also relates to the literature studying socioeconomic differences in mortality and long-term care. Pijoan-Mas and Ríos-Rull (2014) provide age-specific estimates for the negative relationship between mortality and socioeconomic status. Kalwij et al. (2013) also estimate longevity differences over income and gender using Dutch administrative data and report similar results to what we find. Similarly, a negative relationship between long-term care needs and long-term care use and socioeconomic status has been documented (Ilinca et al., 2017; Rodrigues et al., 2018; Garcia-Gomez et al., 2019; Tenand et al., 2020a). These findings align with the well-documented gender-health paradox, stating that women indeed do live longer but tend to be less healthy (Case and Paxson, 2005; Oksuzyan et al., 2008).

3.3 Adverse Selection Model with Multiple Risks

We extend the model of Einav et al. (2010) to describe how adverse selection for a stylized stand-alone annuity and LTC insurance can be reduced by a combined life care annuity and show that this insurance is welfare-increasing.[5] A precondition for this to work is a negative correlation between long-term care- and survival risk. We then use this simple framework to derive an optimal combination of the two insurances, allowing us to single out its determining components. We focus on comparing a world with single insurances to a world of a bundled product, which enables us to derive an optimal bundling in the sense that adverse selection problems are minimized. We abstract from multiple important aspects – such as screening, partial insurance, and the choice between stand-alone and bundled insurance – so that our simple model allows us to focus on optimally combining the two insurances and study the drivers of the optimal combination.

Suppose there is a continuum of individual types $\xi \in \Xi$ with distribution $G(\xi)$ who live for two periods. They differ by their probabilities of survival $s(\xi)$ and probability $q(\xi)$ to become in need of long-term care associated with costs of X. The probabilities are

[5]See also Einav and Finkelstein (2023), the 'self-indulgent' survey describing the recent studies using the Einav-model.

private information. Individuals receive utility U from consumption and are risk-averse with with $U' > 0, U'' < 0$. Lifetime utility[6] is:

$$\begin{aligned} V &= U(C_1) + s(\xi) \cdot \left\{\{1 - q(\xi)\} \cdot U(C_2^h) + q(\xi) \cdot U(C_2^l)\right\} \\ &= U(C_1) + \{s(\xi) - l(\xi)\} \cdot U(C_2^h) + l(\xi) \cdot U(C_2^l), \end{aligned} \tag{3.1}$$

where C_2^h is consumption when healthy and C_2^l is consumption when in need of long-term care at date $t = 2$, and $l(\xi) = s(\xi) \cdot q(\xi)$ is the unconditional probability of becoming in need of long-term care. In line with our later empirical results, we assume that individual types that live longer spend shorter time in long-term care, so that $\text{Corr}(s(\xi), l(\xi)) < 0$.

3.3.1 Stand-alone Annuity and LTC insurance

We first study two stand-alone contracts $k = \{A, L\}$ of an annuity A and a LTC insurance L. Individuals have initial wealth W_t in both periods $t = 1, 2$ where $W_1 > W_2$. In period 1 the agent can buy annuity insurance at premium P_A paying a benefit Υ in $t = 2$ in case of survival, and LTC insurance at premium P_L that covers long-term care costs X in the event of poor health at old age. Hence, the benefit is $B = \{\Upsilon, X\}$ under each insurance. There are no savings in the model so the budget constraint in period 1 is given by $C_1 = W_1 - P_A - P_L$. In period 2, the agent can consume $W_2 + \Upsilon$ in both states if insured. If uninsured, the agent has $C_2^h = W_2$ if surviving healthy and $C_2^l = W_2 - X$ if surviving in need of long-term care.

Rational individuals make a binary choice to buy insurance or stay uninsured, taking the other insurance as given.[7] Comparing the expected utility from being insured with the value from staying uninsured, we can derive the willingness to pay $\pi(\xi, k)$ (WTP) for an insurance for each type. With this, define aggregate demand $D_k(P_k)$ for insurance k as the mass of types whose willingness to pay exceeds the uniform price P_k for the

[6] Note, we assume homogeneous preferences implying that the only heterogeneity between households are the two risks.

[7] When studying one insurance, we assume that the respective other risk is fully insured so that we only have two groups: insured and uninsured agents. Solomon (2022) provides an extension where agents can decide to buy either insurance, both insurances or to stay uninsured. The main results are not affected by our simplifying assumption.

insurance product:[8]

$$D_k(P_k) = \int_{\Xi} \mathbb{1}(\pi(\xi, k) \geq P_k) dG(\xi). \tag{3.2}$$

Risk-neutral insurers have to cover only the costs $c(\xi, k)$ for each insured individual and compete in a Bertrand game over the price of the product. Firms cannot observe individual risk and have to price insurance based on an average risk type and cost AC_k.[9]

The distinguishing feature of the adverse selection model relative to the standard supply and demand model is that supply is not determined with an independent production technology. Instead, the average cost curve –the supply curve– AC_k is given by

$$AC_k(P_k) = \frac{1}{D_k(P_k)} \int_{\Xi} c(\xi, k) \cdot \mathbb{1}(\pi(\xi, k) \geq P_k) dG(\xi) = \mathbb{E}\left\{c(\xi, k) | \pi(\xi, k) \geq P_k\right\}, \tag{3.3}$$

which is determined by the types who choose to buy insurance.

The marginal cost curve in the market is given by $MC_k(P_k) = \mathbb{E}\left\{c(\xi, k) | \pi(\xi, k) = P_k\right\}$, and it is downward sloping so that marginal costs increase in price and decrease in quantity. This shape is generated by the fact that individuals with the highest willingness to pay for insurance are also those with the highest expected costs, but the type ξ is private information. Further, due to agents being risk-averse, the marginal cost curve locates below the demand curve.

Zero profit implies that the equilibrium insurance premium equals the average costs of the entire risk pool willing to buy the insurance at the given premium, so the firms' information problem implies welfare losses relative to a world with complete information.

Panel (a) in Figure 3.1 provides a stylized graphical representation of the welfare losses in a market for the stand-alone insurances.[10] In our example, the WTP curve is always

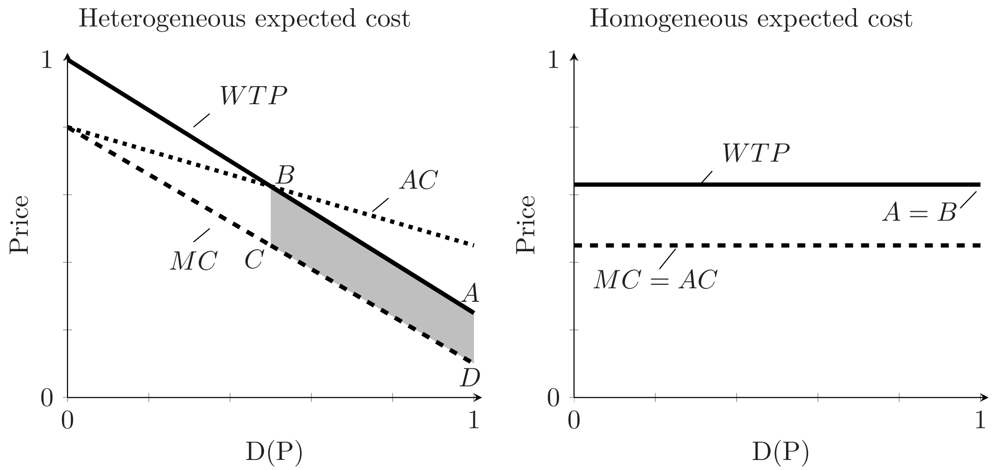
[8] See Appendix C.3 for the derivation of the demand- and WTP-curve for our two period model and accompanying comparative statics on the slope of the demand curve.

[9] Besides this adverse selection, we abstract from any other friction like, e.g., moral hazard. Firms are also not allowed to compete on the coverage features as in Rothschild and Stiglitz (1976) type models. Webb (2009) explicitly shows in this setup that bundling an annuity and a LTC insurance with negatively correlated risks for these states is a Pareto improvement.

[10] Linear demand and supply curves arise if the probabilities are uniformly distributed, which is assumed for the following analysis. Non-linearities, in contrast, arise from normally distributed probabilities.

above the MC curve due to the assumed risk-aversion, implying that agents always prefer being insured when pricing is at marginal costs. Due to asymmetric information, pricing occurs at average costs. Hence, the equilibrium price is in point B, where the willingness to pay of a new –lower-cost– individual no longer exceeds the average cost of the existing insurance pool. It is optimal for the marginal consumer to remain uninsured. The welfare loss due to asymmetric information is the deadweight loss $\overline{ABCD}$, which equals the sum of risk premia of uninsured individuals who are willing to pay a positive risk premium.

Figure 3.1: Adverse Selection Effects with Different Cost Patterns



The slope of the WTP curve is determined by the dispersion of types in the economy: a high willingness to pay for insurance implies a high underlying risk and vice versa. What happens if the heterogeneity in risk decreases? A lower dispersion in costs, $Var(c(\xi, k))$, flattens the WTP-, and the two cost curves. The willingness to pay across agents, as well as their costs, become more aligned. In effect, more agents would be insured (point B would move to the right), and the dead weight loss would decrease. Panel (b) Figure 3.1 depicts the extreme case without dispersion, $Var(c(\xi, k)) = 0$. With all individuals facing the same expected costs the demand- and supply curves become linear. Average costs are equal to marginal costs but below the WTP-curve due to the risk premium that agents are willing to pay. In that case, the first-best optimum of full insurance in point $A = B$ is possible for every risk-averse individual because asymmetric information no longer plays a role.

Of course, assuming risk-averse agents implies that agents would buy insurance even with a negative return on the insurance due to a positive risk premium that they are willing to pay. This means that the first best allocation is already achieved in this model before adverse selection is completely eliminated. In fact, reducing $Var(c(\xi, k))$ to the point where the WTP curve is above the AC-curve for all agents in Panel (a) of Figure 3.1 is enough to ensure full insurance. When studying an optimal comination of insurances, we will use the objective to minimize the variance in costs to get results that are independent of household preferences to simplify the analysis.

3.3.2 Optimal Combined Insurance

The Life Care Annuity In a combined insurance product, the life care annuity CA, agents can pay the premium P_{CA} that pays out the annuity Υ if the agent survives with probability $s(\xi) - l(\xi)$ and is healthy, and the payout is $(1 + \rho)\Upsilon$ if the agent survives but needs long-term care with probability $l(\xi)$.[11] The general idea is to hedge the two risks when $\text{Corr}(s(\xi), l(\xi)) < 0$ to attract a higher number of people choosing this insurance. Assume an individual with a high risk for annuities (high life expectancy), implying high costs, and simultaneously with a low risk of long-term care, implying low cost for LTC insurance. A second agent has low life expectancy and high long-term care risk, implying the reversed costs for the two insurances. The variation in the cost, $Var(c(\xi, k)) \neq 0$, implies adverse selection problems for stand-alone products. However, combining the two insurances hedges the risks and aligns the costs of these two agents. In the optimal outcome, the costs of the agents are equal, i.e., $c(\xi, k) = \bar{c}$ and $Var(c(\xi, k)) = 0$, which eliminates the adverse selection problem and makes the first best allocation feasible in our simple model so that everyone is insured, cf. Figure 3.1(b).

[11] We assume that stand-alone insurance is unavailable when studying combined insurance. This could be labeled as the 'managed competition' case – cf. Solomon (2022) – where a regulator, or market designer does not allow single insurance contracts.

Optimal Combination of Insurance We aim to determine a contract of a combined insurance that maximizes the fraction of insured agents by minimizing adverse selection. This can be achieved by the appropriate choice of ρ, which governs the relative size of the benefit in each state. Note that the benefit in case of LTC is then no longer restricted to be capped by the long-term care costs X, but we rather allow for arbitrary top-up values ρ, which might exceed the costs. The optimal size of the top-up ρ that minimizes adverse selection in this model is reached if expected individuals' cost are homogeneous for all types ξ:

$$\mathbb{E}\left(c(\xi, \mathrm{CA}, \rho)\right) - \overline{c} = 0. \tag{3.4}$$

Consider a simple example with only two types $\xi = (1, 2)$. To equalize benefits, the benefit level ρ needs to be chosen such that condition (3.4) is met for both types, implying that their costs are equal:[12]

$$c(1, CA, \rho) = c(2, CA, \rho) \implies s(1) \cdot \Upsilon + l(1) \cdot \rho \cdot \Upsilon = s(2) \cdot \Upsilon + l(2) \cdot \rho \cdot \Upsilon,$$

Solving for ρ gives:

$$\rho = \frac{s(2) - s(1)}{l(1) - l(2)}.$$

Obviously, $\rho > 0$ if $s(2) > s(1)$ and $l(1) > l(2)$, or vice versa: the time in long-term care $l(\xi)$ and remaining life expectancy $s(\xi)$ have to be negatively correlated to sustain a positive LTC insurance benefit.

If there are infinitely many types, there is no closed-solution possible, and we have to bring the average cost $\overline{c}$ as close as possible to individual expected cost. We do this by

[12] See Appendix C.3 for comparative static on how the slope of the demand curve changes if the correlation between $s(\xi)$ and $l(\xi)$ becomes more negative.

minimizing the squared difference of (3.4):[13,14]

$$\min_{\rho} \mathcal{F}(\rho) = \mathbb{E}\left\{\frac{s(\xi)+l(\xi)\cdot\rho}{\mathbb{E}(s(\xi))+\mathbb{E}(l(\xi))\cdot\rho}-1\right\}^2 = \mathbb{E}\left\{PR\left(\xi,\rho\right)^2\right\} = Var\left\{PR(\xi,\rho)\right\}$$

with

$$PR(\xi,\rho) = \frac{s(\xi)+l(\xi)\cdot\rho}{\mathbb{E}(s(\xi))+\mathbb{E}(l(\xi))\cdot\rho}-1. \tag{3.5}$$

In the multi-period framework, $s(\xi)$ represents the remaining life expectancy and $l(\xi)$, the unconditional remaining lifetime spent with long-term care needs. PR in Equation (3.5) is the *premium return* of the life care annuity, defined as the difference between the ratio of expected (present) value of benefits relative to its premium. An analogous concept of *money's worth* was suggested by Mitchell et al. (1999) and used by, e.g. Finkelstein and Poterba (2004) and Brown and Finkelstein (2007).[15] In the model discussed above, the expected value of benefits for each type is $s(\xi) + l(\xi) \cdot \rho$. The uniform premium in a competitive market is given by $P_{CA} = \mathbb{E}(s(\xi)) + \mathbb{E}(l(\xi)) \cdot \rho$ which is the denominator of Equation (3.5). A value of unity implies a premium return of zero: the pricing of the insurance is then actuarial fair with premia equal to the expected value of benefits. Our objective function aims to minimize the variance in premium returns, implying as little heterogeneity in marginal cost as possible. It is important to note that our objective is to minimize welfare loss due to adverse selection. We leave the explicit modeling of the choice of different insurance contracts for future research.

Deriving the first-order condition from the optimization problem (3.5) and solving for the optimal top-up ρ yields our main result:[16]

[13]Without affecting our main results, we divided the objective function by $\bar{c}$ so that we can express it in terms of premium returns to get a better intuition for the results.

[14]We here assume a real interest rate of zero so that the time value of money does not play a role in the model.

[15]The money's worth used by Finkelstein and Poterba (2004) is defined as the expected present discounted value of annuity payouts divided by the initial premium. In our terminology, this could be defined as the premium return plus one, equal to one if the benefits align with the premium.

[16]Appendix C.4 provides the derivation.

$$\rho^\star \quad = \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\frac{\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} - \text{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}, \frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}{\left\{\frac{\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}\right\}^{-1} - \text{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}, \frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}, \tag{3.6}$$

where $\frac{s(\xi)}{\mathbb{E}(s(\xi))}$ and $\frac{l(\xi)}{\mathbb{E}(l(\xi))}$ can be interpreted as the money's worth of the stand-alone insurances, i.e. one plus the premium return, for the stand-alone annuity-, and LTC insurance, respectively. The optimal size of benefit in long-term care relative to not in long-term care, $\rho^\star$, depends on three main elements: (1) the relative duration in each state, $\frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))}$, (2) the relative standard deviations of the money's worth for stand-alone LTC- and annuity insurance, $\frac{\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}$, and (3) the correlation of the money's worth of the two stand-alone insurances, $\text{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}, \frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}$.

The relative duration in each state, i.e., the expected life expectancy relative to the expected duration in need of long-term care, has a proportional impact on ρ^*. Assuming the relative standard deviations to be one and a perfectly negative correlation, the intuition is straightforward. Let life expectancy be two times higher than the time spend in need of long-term care so that $\frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} = 2$, then the level effect of optimal condition (3.6) implies that the top-up of LTC benefits must also be twice as high compared to the state when not in need of long-term care ($\rho^* = 2$) to eliminate the differences in premium returns.

The second factor is a measure for the heterogeneity in risks and can be interpreted as the heterogeneity in the money's worth in each stand-alone insurance. In effect, this measure is an indication which of the two insurances suffer from more severe adverse selection problems. In Section 3.3 we showed that decreasing heterogeneity in types also decreases the deadweight loss. A value above one implies that the heterogeneity in premium returns is larger for an annuity whereas this is reversed if this ratio is smaller one. The impact of this factor on the optimal combination of the two insurances is again straightforward. Assuming a relative duration of 1 and again a perfectly negative correlation, we also have a proportional effect on ρ^*. If the heterogeneity in premium returns is twice as large when in need of long-term care so that $\frac{\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} = 0.5$, then the implied top-up is given by $\rho^* = 0.5$. The intuition is that the optimal combination of the two insurances implies that a higher benefit should be granted in the state where heterogeneity in risks is lower. Finally, note, that taken both factors together assuming a perfectly negative correlation simply consist of the product of the two:

$$\rho^*_{\textbf{Corr}=-1} = \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}. \tag{3.7}$$

This equation shows that the two factors can offset each other: following the example above, if there is more heterogeneity in long-term care risk but the duration is shorter than surviving healthy, then it might be optimal to combine an annuity and a LTC insurance and pay out the same benefit in both states, i.e., $\rho^* = 2 \cdot 0.5 = 1.0$.

The third factor measures the correlation between the money's worth of the two stand-alone insurances. When the risks – and hence the premium returns of the stand-alone insurances – are not perfectly negatively correlated, the combination of the two insurances can only partially eliminate adverse selection incentives. The risks cannot be perfectly hedged and there are remaining differences in the premium returns in a combined insurance. Besides, a correlation between $[-1, 0]$ reinforces Equation (3.6)'s first two effects on the optimal top-up ρ^* in both directions. Hence, a lower correlation in absolute terms yields a positive (negative) effect on ρ^* if the ratio of the standard deviation (factor 2 in Equation (3.6)) is larger (less) than one.

Bringing the Model to the Data In our empirical section, we will elaborate on the quantitative importance of the effects described above for the different groups ξ, which we will specify as quintiles of lifetime income, gender, and marital status. Section 3.5 shows results for premium returns over lifetime income quintiles. We will label the slope of the line connecting the premium returns over income quintiles as *gradients*, referring to the well-known socioeconomic gradient in mortality discussed in the health-economics literature, see, e.g., Dow and Rehkopf (2010). Factor two in Equation (3.6) – the ratio of the standard deviations – is a measure for the sign and the steepness of these gradients. In contrast, the correlation can be seen from the shape of the premium returns over income and their relative (opposing) slopes: two linear and opposing slopes indicate a high negative correlation.

We further assume that agents can purchase insurance by paying a lump-sum premium P_k at (initial) age 65, priced at the average risk. We estimate the quantities $s(\xi)$ and $l(\xi)$ with our multi-state model. We discretize type distribution G, by taking the empirical probability of observing the type ξ at age 65. This also allows us to calculate the population's remaining life expectancy $\mathbb{E}(s(\xi))$ and the unconditional time spend in long-term care $\mathbb{E}(l(\xi))$.

3.4 Data and Empirical Approach

3.4.1 Institutional Context

The Netherlands has a universal and generous pension and long-term care system. The pension system consists of a tax-funded minimum social security benefit (first-pillar) that is paid from the statutory retirement age to each Dutch citizen with a required minimum time living in the country. This AOW (*Algemene Ouderdomswet*) pension is complemented with a (second-pilar) occupational defined benefit pension, which is mandatory (self-employed excluded) and based on lifetime earnings. The replacement rate is quite high, reaching around 70% of average lifetime earnings (Knoef et al., 2017).

The public long-term care system provides coverage for both formal long-term care at home and in a nursing home. Unlike the U.S., private LTC insurance and out-of-pocket expenditures are marginal, being less than 0.5% of total long-term care expenditures (Colombo et al., 2011). Everyone who lives in the Netherlands is insured and pays income-dependent premia. Total long-term care expenditures are 4.1% of GDP and among the highest of OECD countries (European Commision, 2015). Every request for long-term care is assessed by the Centre for Care Assessment (CIZ), taking into account the usual informal care that partners or other household members give to each other (Mot, 2010). Nursing home care is available for individuals with more severe conditions or a less supporting environment. However, individuals may also choose to receive personal care at home. When getting personal care at home, the partner is expected to provide the usual domestic and supportive care. Individuals are entitled to less personal care when the partner voluntarily provides personal care (Mot, 2010; Bakx et al., 2015). In 2015, a major long-term care reform has been implemented, reducing coverage and increasing co-payments. In the new system, only people who need care day and night are entitled to care in a nursing home. For people with lighter care needs, personal care at home is no longer publicly insured (Maarse and Jeurissen, 2016).

Overall, the Netherlands stands out from other OECD countries in old-age social insurance by providing an almost universal public long-term care scheme with generous coverage, which implies low out-of-pocket expenses so that adverse selection problems for using long-term care are arguably low. Eligibility rules depending on informal care availability also suggest low selection effects into long-term care. These institutional factors allow us to estimate arguably unbiased socioeconomic differences in long-term care use and mortality.

3.4.2 Data and Sample Selection

We use administrative data for the Netherlands containing detailed longitudinal information on formal long-term care use and mortality (exact date of death) for the entire population. Administrative data on formal nursing home care and home care is obtained from the Central Administration Office (CAK). These data cover all residents of the Netherlands aged 18 and older who have long-term care expenses that are covered by the public long-term system. Data on mortality is obtained from the causes of death registry. In addition, we use detailed income and assets data from tax registries to measure socioeconomic status. Demographic characteristics, including age, gender, and marital status are obtained from the municipality population register.

While the long-term care use data are available since 2004, we use them starting in

2006 when also assets data are available to determine socioeconomic status. Our study ends in 2014 before the major reforms of the long-term care system were implemented. We include retired individuals aged 65+ and their partners whose main source of income is pension income. We exclude individuals if they are not registered in the Netherlands for the entire sample period. Further, we exclude households remarrying or divorcing after age 65 (4.5%). We exclude a few households with negative income or assets and those with missing data (0.2%). This leaves us with a final sample of 3,219,297 individuals in 2,198,755 households.

3.4.3 Variables

Formal long-term care use is defined broadly, including institutionalized and home care. Institutionalized care comprises nursing home care and psychiatric or disabled care. For our sample, nursing home care covered about 93% of institutionalized care in 2006. Home care use is defined as receiving personal care, such as help with daily activities (ADL), and nursing care, such as wound dressing. We do not include domestic care. For institutionalized care, we measure each spell's starting and end date; for home care we measure the spells on a 4-week basis after 2008 and until 2008 as the first and last day of use in the year.[17] We excluded spells where home care was provided for less than one hour during the year.

For the covariates, marital status is defined as being in a couple (married, a registered partnership, or cohabiting) or a single-person household. Socioeconomic status is measured by average retirement income, which is the sum of personal gross income (deflated using CPI) – and for couples, its sum – and the annuity value of household financial assets. As our sample contains retired individuals only, average retirement income provides a good proxy for lifetime income. To compute the annuity value of household assets, we follow Knoef et al. (2016), see Appendix C.1 for details. Household financial assets are particularly important to include as a source of retirement income for former self-employed individuals. Retirement income is equivalized using OECD

[17] For 2008, each spell's start and end date is marked by the start and end of the calendar year.

scales to make couples and single-person households comparable regarding retirement income. Based on this measure, we construct lifetime income quintiles.

3.4.4 Multi-State Model

We use a multi-state model to estimate lifetime long-term care use and remaining life expectancy at age 65 for different groups $h \in \mathcal{H}$ (lifetime income quintile, gender, initial marital status at age 65). The model has three states (no long-term care use, long-term care use, and death) with transition rates $\lambda_k(t)$, and individuals can repeatedly visit the states (both in the model and data). To estimate the transition rates we apply a competing risk analysis, i.e. we take into account that only one of two possible transitions takes place, leaving the other transition unobserved. We assume the transition rates to be independent in terms of unobservable characteristics, so the transition rates can be separately estimated per state using a mixed proportional hazard (MPH) model (Hougaard, 2000; van den Berg, 2001):

$$\lambda_k(t, marstat_i(t); \nu_i^k, \gamma_k, \boldsymbol{\beta}_k) = \lambda_0(\boldsymbol{\gamma}_k, t) \cdot \phi(\boldsymbol{\beta}_k, marstat_i(t)) \cdot \nu_i^k, \tag{3.8}$$

where $\lambda_0(\boldsymbol{\gamma}_k, t) = \exp\{(\gamma_k + \gamma_{kh}) \cdot t\}$ is the baseline hazard capturing age-specific transition rates for each state and group, with t as the age-indicator. The parameter γ_{kh} captures the difference in the age-specific transition rates over groups. The advantage of using age as a time scale is that we abstract from unknown information regarding some individuals' beginning of the no-long-term care use or long-term care use spell. Otherwise, we should have imputed the starting dates or excluded these left-censored spells[18], which might result in biased estimates because of an initial conditions problem Heckman (1981). We assume a Gompertz functional form for the baseline hazard, which is a common specification for adult mortality in developed countries (see e.g. Missov et al., 2015).

The second term of the model, $\phi(\boldsymbol{\beta}_k, marstat_i(t)) = \exp\{\ \beta_k + \beta_{1kh} + \beta_{2kh} marstat_i(t)\}$, includes current marital status (for initially married couples) as a time-varying covariate to capture the transition from being married to a single-person household. Moreover,

[18]Contrary to left-truncated spells, left-censored spells have an unknown start date.

it captures differential mortality and differences in informal care possibilities between singles and couples. The parameter β_{1kh} measures the difference between initial singles and initially married individuals who have become single, and β_{2kh} picks up the additional impact of becoming single while currently married.

The third term of the MPH model $\nu_i^k \sim \Gamma\left(\frac{1}{\sigma_k^2}, \frac{1}{\sigma_k^2}\right)$ is an individual-specific random effect accounting for dynamic selection and other unobservable differences between individuals, for instance, factors explaining mortality among the oldest old and the mortality plateau (see e.g. Vaupel et al., 1998; Barbi et al., 2018). We assume this so-called frailty term to follow a Gamma distribution because it well describes observed heterogeneity over long durations (and, therefore, frailty in old age) (Abbring and van den Berg, 2007); moreover, unique parameter identification exists (Honoré, 1993). Individuals draw the random effect value at initial age 65. For tractability, the random effect is not shared over different states.

Estimating a mixed proportional hazard model with left truncation and frailty is computationally challenging because the left-truncated sample has a different frailty distribution. Allowing for time-varying covariates and repeated spells adds a layer of complexity. Because we assume independence across transitions, we follow the estimation technique from Chapter 5 addressing these challenges; see Appendix C.2 for more details and the maximum likelihood specification. Having estimates on the transition rates, we use a simulation model to determine long-term care use and remaining life expectancy at age 65 for different groups. As a starting point, we use the conditional distribution of our variables at age 65 (see Table C.1 in Appendix C.2). For the simulations, we extend the approach by Crowther and Lambert (2017) to allow for transitions from couples to single-person households. More specifically, for couples, we first simulate age profiles from age 65 until the end of life for both partners. Next, we re-simulate the remaining age profile for the surviving partner according to our simulation model. We simulate $N = 100,000$ households repeated 5,000 times to construct 95% confidence intervals; see Appendix C.2 for additional details.

3.5 Results

We show results on the simulated durations of long-term care and life expectancy over lifetime income and we highlight the importance of gender and marital status. We then show how these differences translate into the value of annuity- and LTC insurance and we finally present results for a life care annuity.[19]

3.5.1 Socioeconomic Differences in Long-term Care and Mortality

We find substantial gradients in long-term care use and remaining life expectancy over lifetime income. Table 3.1 shows that low-income individuals live shorter than high-income individuals but use more long-term care. On average, men and women in the bottom income quintile, respectively, live 4.0 and 2.3 years shorter than their high-income counterparts in the top income quintile (see last column). On contrary, low-income men and women spend 1.1 and 1.7 years longer in long-term care than their high-income counterparts. There are also gradients in the probability of ever using long-term care, ranging from 91% for women in the bottom income quintile to 86% in the top income quintile. Overall, the income gradient concerning life expectancy is steeper for men than for women while, reversely, the gradient for long-term care is steeper for women than for men.

To see the role of having a partner for these socioeconomic gradients, we turn to the difference for initially married versus initially single individuals. Marital status is an important factor influencing the transition into long-term care and mortality. We simulate the durations separately for individuals who married at age 65 and those single and compute the difference $\Delta(Married - Singles)$. The difference in life expectancy between initial married and singles is 2.5 years for men and 1.8 years for women. This survival advantage of being married is among others reported in Pijoan-Mas and Ríos-

[19]A robustness check confirmed a good match between simulated and empirical surival and long-term care use probabilities by age, marital status, lifetime income and gender. Results are available upon request.

Rull (2014). In addition, we find that single men spend 0.8 years more in long-term care than their married counterparts. We do not find a significant difference of long-term care use over marital status for women. This result suggests that women have fewer opportunities to get informal care from their spouse than men and tend to live longer.

The number in the last column corresponds to a differences-in-differences approach showing how being married or single affects the difference between the top and the bottom income quintile. Our results show that the gap in life expectancy between the bottom and top income group is 2.4 years smaller for married women than for single women. Essentially, this implies that being married flattens the income gradient of life expectancy for women: only for single women we observe a strong gradient over income while this is moderate for married women. This same number is only 0.4 years for men, implying that the gradient in life expectancy is only moderately flattened for married individuals. Similarly, the gap in long-term care use between the bottom and top income group is 1.6 years smaller for married men than for single men. This number is 0.1 years but insignificant for women. Again, this implies that the income gradient in long-term care use for married men is almost flat, whereas it is relatively strong for single men.

Turning to the socio-demographic difference, we find that women tend to live 3.9 years longer than men (21.9-18.0 years). In addition, women have a higher prevalence and longer duration of long-term care use than men: About 89 percent of women ever uses long-term care with an average duration of 5.1 years conditional upon use. In contrast, 77 percent of men use long-term care with an average duration of 3.1 years, amounting to 12% of their remaining lifetime, compared to 18% for women.

Table 3.1: Life Expectancy and Long-term Care Use by Lifetime Income Quintiles

(a) Men	All	Bottom	Second	Third	Fourth	Top	ΔTop - Bottom
LE at age 65 (years)	18.0 (17.9;18.1)	15.3 (15.0;15.5)	16.8 (16.6;17.0)	17.6 (17.5;17.8)	18.4 (18.2;18.6)	19.2 (19.1;19.4)	4.0 (3.7;4.2)
Δ(Married - Singles)	2.5 (2.3;2.7)	2.0 (1.6;2.5)	2.7 (2.2;3.1)	2.6 (2.2;3.0)	2.5 (2.1;2.9)	1.6 (1.2;2.0)	−0.4 (−1.0;0.1)
LTC (years)*	3.1 (3.0;3.1)	3.8 (3.7;4.0)	3.4 (3.3;3.6)	3.2 (3.1;3.2)	3.0 (2.9;3.0)	2.8 (2.7;2.8)	−1.1 (−1.2;−0.9)
Δ(Married - Singles)	−0.8 (−0.9;−0.7)	−1.8 (−2.0;−1.5)	−1.5 (−1.8;−1.3)	−0.7 (−0.9;−0.6)	−0.3 (−0.5;−0.2)	−0.2 (−0.4;−0.1)	1.6 (1.3;1.9)
Ever use LTC (%)	77 (76;78)	79 (78;80)	79 (78;80)	78 (77;79)	77 (76;78)	75 (75;76)	−3 (−5;−2)
(b) Women							
LE at age 65 (years)	21.9 (21.8;22.0)	20.1 (19.9;20.3)	21.8 (21.6;22.0)	22.0 (21.8;22.2)	22.2 (22.1;22.4)	22.3 (22.2;22.5)	2.3 (2.0;2.5)
Δ(Married - Singles)	1.8 (1.6;2.0)	3.0 (2.6;3.4)	2.0 (1.6;2.4)	1.4 (1.0;1.7)	1.2 (0.8;1.5)	0.6 (0.3;1.0)	−2.4 (−2.9;−1.9)
LTC (years)*	5.1 (5.1;5.2)	6.0 (5.9;6.2)	5.9 (5.8;6.0)	5.3 (5.2;5.4)	4.8 (4.7;4.9)	4.4 (4.3;4.4)	−1.7 (−1.8;−1.5)
Δ(Married - Singles)	0.1 (0.0;0.1)	0.4 (0.2;0.7)	0.4 (0.2;0.6)	0.3 (0.1;0.5)	0.4 (0.2;0.6)	0.4 (0.2;0.6)	−0.1 (−0.4;0.3)
Ever use LTC (%)	89 (88;89)	91 (91;92)	91 (91;92)	89 (89;90)	88 (88;89)	86 (85;87)	−5 (−6;−4)

Notes: These are population-averaged measures for the life cycle simulation of 100,000 households. We present the median estimates across 5,000 bootstrapped samples and the 2.5^{th} and 97.5^{th} percentiles between brackets. Sample sizes are reported in Table C.1 in Appendix C.2.

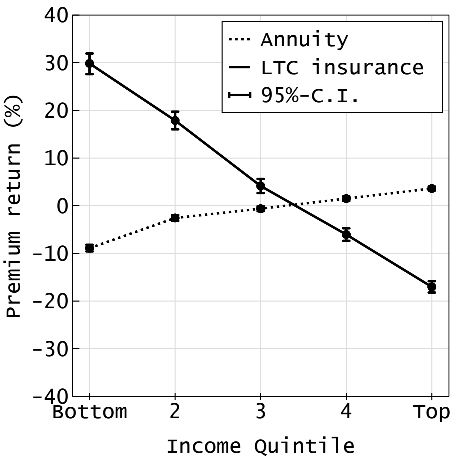
3.5.2 Premium Returns of Old-Age Insurances

Stand-Alone Contracts of Annuity and LTC Insurance

We translate the heterogeneity in long-term care use and life expectancy into a money's worth for the different insurances over subgroups according to Equation (3.5).

Uniform Premium We first study an annuity and a LTC insurance independently assuming a uniform premium for everyone for each insurance, implying that the total insured sample $\mathcal{H}$ comprises the whole population at age 65+. We focus on the income quintiles as our subgroups. The implied premium returns are depicted in Figure 3.2.

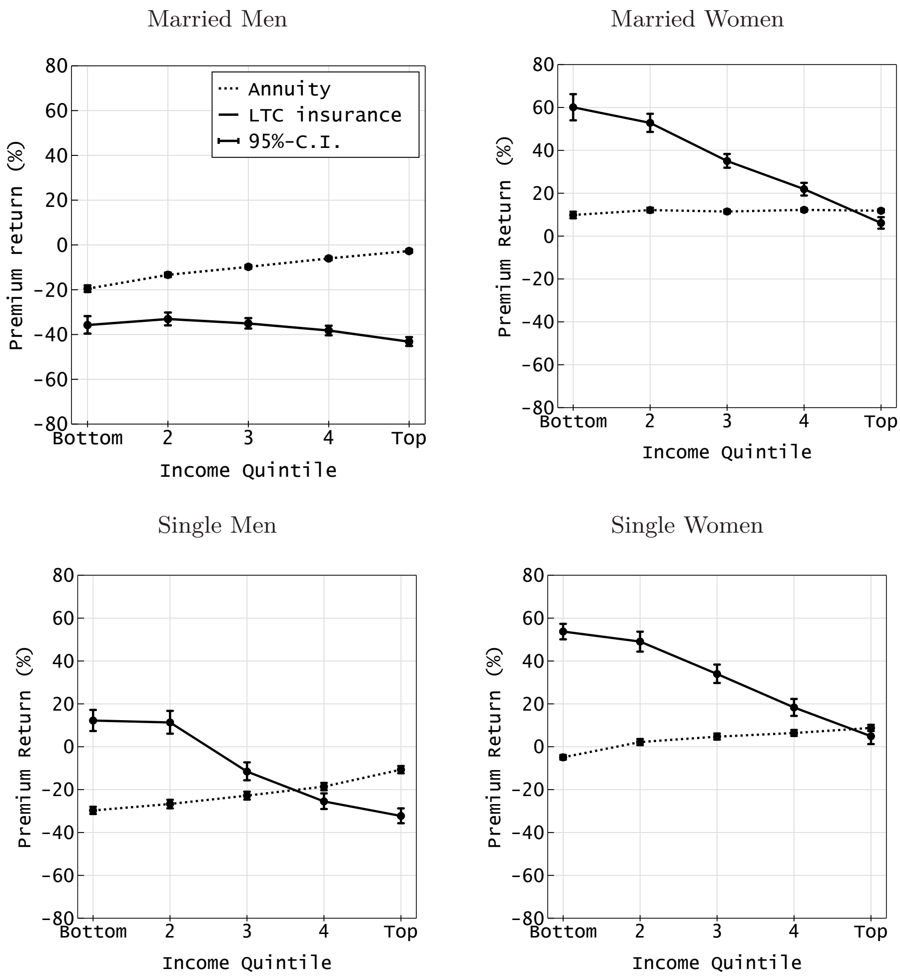
Figure 3.2: Premium Return with Uniform Premium



Notes: Population-averaged premium returns for the life cycle simulation of 100,000 individuals. Medians across 5,000 bootstrapped samples are shown. The underlying premium returns on the pension annuity, LTC insurance, and life care annuity are provided in Table C.5 in Appendix C.6.

As reflected by the steeper line, the results show that benefit inequality across income groups is larger for LTC insurance than for annuities. The premium return for the lowest income group is 29.9 percent, implying that a premium of one Euro yields an expected value of benefits of 1.299 Euro. On the other hand, the highest income groups lose 17.0 cents on every euro invested in the LTC insurance. On the contrary, for every euro invested in the annuity priced according to the average risk, households in the lowest income group receive only 91.1 cents. Households in the highest income groups have a positive return and earn 3.6 cents on top of every euro invested. The larger discrepancy in premium returns for LTC insurance makes this insurance product more prone to adverse selection by income groups than pension annuities in our case.

Figure 3.3: Premium Returns by Gender and Marital Status with Uniform Premium



Notes: These are population-averaged premium returns for the life cycle simulation of 100, 000 individuals. Medians across 5,000 bootstrapped samples are shown.

The socio-demographic differences of life expectancy and long-term care use over marital status and gender translate into heterogeneity in premium returns over these dimensions. There is also a negative correlation in risks for marital status as married individuals live longer but spend less time in long-term care - at least for men. Note, however, that the two risks are not negatively correlated over gender because women have a higher life expectancy and spend more time in long-term care.

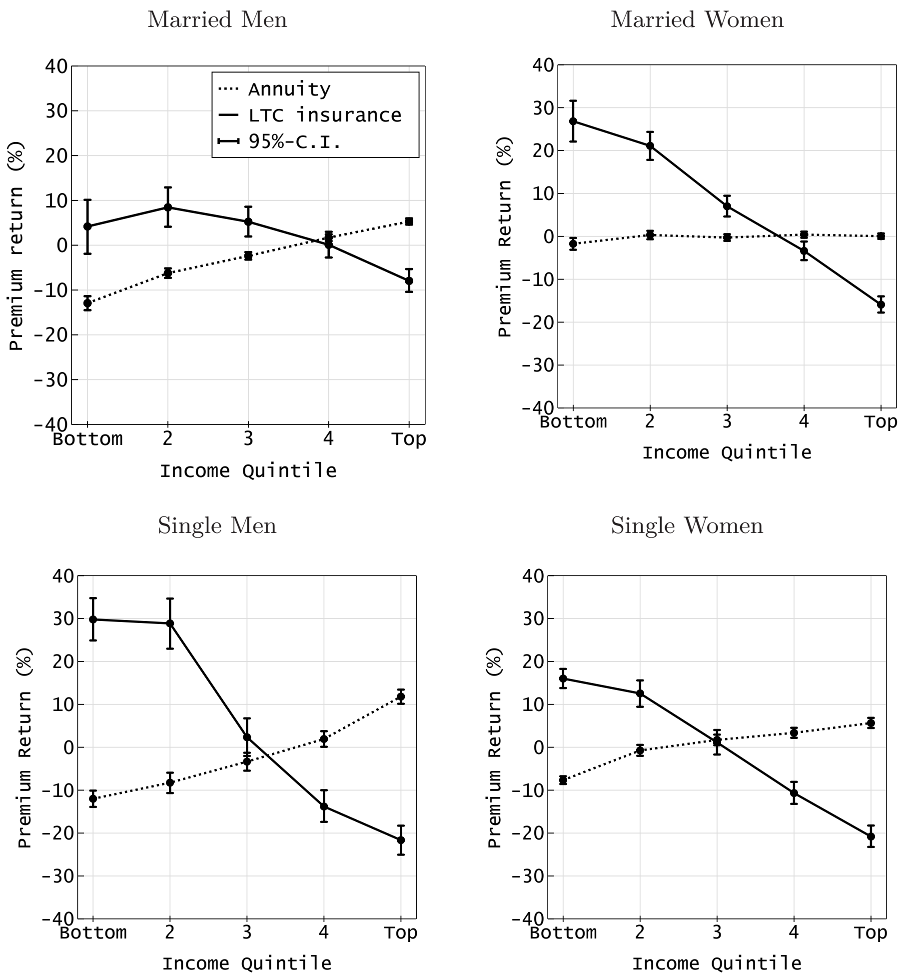
This is reflected in Figure 3.3 which shows the implied premium returns with uniform premium over marital status and gender. The large difference across panel (a)-(d) reveals strong level effects, particularly over gender. Married men have negative premium returns throughout the income distribution, whereas married women value both insurances. The reason for this outcome is simple: men die earlier and they use less long-term care. Insurances priced at the average risk are not valuable for this group.[20] The picture is similar for singles, except single men in the two lower income quintiles who enjoy positive returns of an LTC insurance.

Group-Specific Premia The large differences in premium returns for stand-alone pension annuities and LTC insurance can potentially lead to strong adverse selection effects based on marital status and gender. To prevent a potential unraveling of the insurance market, group-specific premia based on observables such as marital status and gender might reduce adverse selection problems.

Figure 3.4 shows the effect of marital-status-, and gender-specific premia on the premium returns. Compared to Figure 3.3, group-specific premia shift the lines closer to zero, while –unsurprisingly– the gradients over income still persist. Offering premia that may differ over gender and marital status, however, are able to eliminate the large level effects of the premium returns between these groups which decrease the adverse selection problem significantly. Figure 3.4 also shows large variations in the steepness and the shape of the gradients. For example, the income gradients in long-term care are particularly steep for married women and single men, while the shape of the gradient for

[20] For practical reasons, we assume that discrimination over lifetime income is not possible for insurance companies. This information is not only hard to obtain for insurances, it is also hard to imagine a regressive premium system where the income-poor need to pay higher premia than the income-rich to reduce differences in premium returns in the LTC insurance.

Figure 3.4: Premium Return over Gender and Marital Status with Group-Specific Premium



Notes: Population-averaged premium returns for the life cycle simulation of 100,000 individuals. Medians across 5,000 bootstrapped samples are shown. The underlying premium returns on the pension annuity and LTC insurance are provided in Table C.5 in Appendix C.6.

married men is more hump-shaped. The annuity gradient over income is stronger for men (both married and single) and almost non-existent for married women. These differences become important when analyzing the optimal combination of the two insurances which we turn to next.

A Life Care Annuity

As shown in Section 3.3, combined insurance can moderate welfare losses from adverse selection when the correlation between surviving and getting in need of long-term care is negative. In our setting, this is reflected by the reverse gradients of the premium-return lines depicted in Figures 3.2 to 3.4. However, at least with a uniform premium, we have a positive correlation of longevity and LTC risk over gender, which counteracts this negative correlation (cf. Figure 3.3).

We derive an optimal life care annuity according to Equation (3.6) and compare two cases assuming (i) a uniform premium over all observable groups (i.e., lifetime income, gender, marital status) and (ii) group-specific premia over gender and marital status where the optimal top-up ρ^* is found only over the remaining differences over lifetime income.

Table 3.2 shows the results of the optimal top-up of long-term care benefits and Table 3.3 presents standard deviations for stand-alone insurances and the life care annuity as a measure for the adverse selection problem with each of the three insurances.

Table 3.2: Optimal Life Care Annuity: ρ^* and Components

	Level effect	Heterogeneity in risk	Correlation between risks	Optimal LTC top-up	
	$\frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))}$	$\frac{\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}$	$\text{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}, \frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}$	$\rho^{\star}_{\text{Corr}=-1}$	$\rho^{\star}$
Uniform Premium	5.68	0.33	0.57	1.88	−0.56
Group-Specific Premium					
Married Men	8.50	1.26	−0.73	10.67	11.16
Married Women	5.03	0.04	−0.32	0.22	0.08
Single Men	5.25	0.42	−0.93	2.18	2.11
Single Women	4.41	0.35	−0.90	1.55	1.47

Notes: Median estimates across 5,000 bootstrapped samples. Optimal top-up and its components according to eq. (3.6) and (3.7).

We first turn to the case assuming a uniform premium paid by all individuals. The optimal top-up of LTC benefits needed to minimize the heterogeneity in premium returns across income groups is negative, $\rho^* = -0.56$, implying a lower benefit when needing long-term care, which is, of course, not a meaningful insurance. The result is coming from an overall strongly positive correlation between the two risks across the studied risk types. Recall our finding that women live longer and use more long-term care than men, implying a positive correlation of risks across gender. These large gender differences in longevity and long-term care use are stronger than the – negatively correlated – differences over income and marital status and induce an overall positive correlation between longevity and long-term care use. In addition, Table 3.3 shows that the standard deviation for the combined product is still very high, so the bundling does not reduce the adverse selection problem by much. Overall, this implies that a life care annuity with uniform premium does not work, so we now turn to group-specific premia.

First, note that with group-specific premia, all correlations turn negative, cf. column 3 in Table 3.2, which was already implied by the inverse gradients shown in Figure 3.4. To understand the different values of the optimal top-up, ρ^*, over these groups, let us decompose it into its components in the first three columns. The first column can be interpreted as the value of ρ^* if the heterogeneity in risk would be equal over states (implying a ratio of the standard deviations of one), and the correlation would be perfectly negative. Similarly, the second column would be the value of ρ^* if the duration would be equal for both states and the correlation -1. The value $\rho^\star_{\text{Corr}=-1}$ then is simply the product of the two, while the final column shows the sum of all three effects including the effect stemming from a non-perfectly negative correlation of the two risks.

Turning first to married men we observe that the duration in long-term care is rather short, so the optimal top-up would be 8.5 from the level effect alone. This can be seen from the optimality condition in Equation (3.6), prescribing a higher benefit to be paid in states with shorter duration. At the same time, the heterogeneity in longevity risk is larger, reinforcing the effect on the optimal top-up. If the heterogeneity effect is isolated, the optimal top-up would only be 1.26 because Equation (3.6) prescribes to put a higher weight on the less heterogenous state (needing long-term care in this case).

The combined effect in column 4 is actually quite close to the final optimal value of 11.16 because the correlation between the two risks is quite strongly negative (-0.73). The value of $\rho^* = 11.16$ implies that the benefit in the case of long-term care need to be more than 11 times larger than the annuity benefit, a high number that we put into perspective in the next section.

In stark contrast, married women have an optimal top-up of only 0.08 implying the optimal combination of insurances is close to a mere annuity. Two factors from the data drive this result: First, the heterogeneity in risk is very low for annuities compared to a strong one for LTC insurance implying the ratio to be 0.04; also when compared with the flat gradient for annuities and the strong gradient in long-term care insurance in panel (b) of Figure 3.4. In addition, the correlation is with -0.32 only moderately negative and a combination of the two insurances is not well-suited.

The picture is quite different for single individuals. Here, we find almost perfectly negative correlation between the risks as well as offsetting level effects and heterogeneity in risks yielding reasonable values for the optimal top-up between 2.11 for single men and 1.47 for single women.

Table 3.3: Standard Deviations of Premium Returns

	Annuity	LTC insurance	Life Care annuity
	$\text{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}$	$\text{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}$	$\text{SD}\left\{\frac{s(\xi)+\rho^*\cdot l(\xi)}{\mathbb{E}(s(\xi))+\rho^*\cdot\mathbb{E}(l(\xi))}\right\}$
Uniform Premium	11.47 (11.21;11.72)	34.64 (33.82;35.45)	11.00 (10.77;11.23)
Group-Specific Premium			
Married Men	7.57 (6.86;8.31)	6.04 (4.32;7.96)	2.42 (1.02;4.15)
Married Women	0.63 (0.31;0.98)	14.15 (12.88;15.37)	0.55 (0.22;0.93)
Single Men	8.74 (7.79;9.72)	21.06 (18.84;23.29)	2.22 (1.09;3.43)
Single Women	5.01 (4.46;5.55)	14.26 (13.0;15.55)	1.64 (0.97;2.34)

Notes: Values computed correspond to the objective function from eq. 3.5 and are multiplied with 100%. Median estimates across 5,000 bootstrapped samples and the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles between brackets.

The standard deviations for group-specific premium returns in Table 3.3 reveal the highest values for the stand-alone LTC insurances implying that adverse selection problems are most severe for this case. With group-specific premia, a life care annuity reduces these heterogeneities substantially always yielding lower standard deviations than in both of the stand-alone insurances.

Our results suggest that a life care annuity to hedge the two risks of longevity and long-term care is not quite possible for married men and women. The implied top-up of the benefit in the long-term care state is unreasonably high for married men and unreasonably low for married women. In contrast, a combined insurance is well-suited for single individuals.

3.6 Discussion

Our analysis points to a broader question of why, in practice, certain risks are covered under bundled policies while others are not. Examples for bundled insurances are not only life care annuities, but also life-insurances with a LTC rider, combined disability coverage, reverse mortgage, or home-car insurance, cf. Eling and Ghavibazoo (2019).

In our analysis, we shed further light on when and how to combine insurance products by disentangling the determinants of the risk structure when bundling is possible and what it depends on. To minimize the adverse selection problem, we show that it is not sufficient to only focus at the correlation between lifetime long-term care use and life expectancy, but rather also take into account the average size and variation of these correlated measures.

Our formula for ρ^* is easy to apply and to compare to other studies that report lifetime long-term care use and remaining life expectancy by socioeconomic group. For example, we can approximate a value of ρ^* using results from Ko (2022), Table 4, which documents longevity and long-term care needs over income deciles. Using these numbers yields a value of $\rho^* = 2.09$ for 60+ individuals in the U.S. ignoring the heterogeneity in gender and marital status.[21]

What can be learned from our analysis for the optimal top-up value ρ^* for the life

[21] In the computation we assumed an equal weight for each income group.

care annuity market in the U.S.? According to www.annuity.org, the monthly income stream paid when healthy in a typical life care annuity contract can be two to three times as large in the case of long-term care needs. This would imply a value of ρ of two or three for a typical life care annuity. According to our results in Table 3.2 these values are very close to the optimal top-up for single men and women, implying that for these groups the existing insurances in the U.S. would largely diminish adverse selection problems. However, the picture looks quite different for married men and married women: Men would require a benefit level 11 times higher than the annuity paid when not needing long-term care. This is not offered as a combined product and rather resembles a stand-alone LTC insurance. In contrast, a stand-alone annuity would rather fit for married women, which is implied by the value for ρ^* close to zero. Consequently, the current market for life care annuities does not seem to reduce adverse selection problems for married individuals.

Another important dimension that our study highlights are group-specific premia, in particular discrimination of premia over marital status and gender. In the U.S., discrimination over marital status are common practice by offering so-called 'couple discounts'. Solomon (2022) reports couple discounts for LTC insurance of around 25% compared to singles. Different premia also prevail for life care annuities, life insurance, and private annuities. Gender-based pricing in insurance is still practice for many insurances and many states in the US, although the Affordable Care Act banned discrimination over gender for health insurance in 2014. In the European Union, the Court of Justice declared gender-specific premia invalid with European legislation and prohibited this practice in Europe in 2012. For LTC and combined products, however, premia largely vary over gender and marital status, although couples tend to be insured jointly. According to American Association for Long-Term Care, premia for single women are around 50% higher than for men and per-capita also higher than for the combined premium for couples.[22]

[22] See: `https://www.aaltci.org/long-term-care-insurance/learning-center/ltcfacts-2022.php` [*retrieved on*: October 20^{th}, 2023]

We find sizable differences in the heterogeneity in risks over gender and marital status, which calls for the need to discriminate premia over these dimensions to tackle adverse selection problems adequately.

3.7 Conclusion

We quantify socioeconomic and socio-demographic differences in mortality and long-term care by estimating a flexible multi-state model on rich administrative data from the Netherlands. We use the estimated model to examine the adverse selection problems of stand-alone annuities and of LTC insurance for different groups. We further determine the optimal combination of these two products in a life care annuity that reduces the heterogeneity of premium returns across socioeconomic groups. We find a strong socioeconomic gradient in mortality and long-term care implying a negative correlation between the two risks and a large gender gradient in these two risks inducing a positive correlation. A third important factor influencing these differences is marital status indicating the importance of the availability of informal care by the spouse, particularly for men. A life care annuity aiming to minimize the heterogeneity of benefits between socioeconomic groups is not feasible with a uniform premium. Only with group-specific premia and then mostly for single individuals rather than for the married, a life care annuity can reduce adverse selection problems. Our results might provide an explanation for why the existing market for life care annuities in the U.S. is so small.

Chapter 4

Health Inequalities and the Progressivity of Old-Age Social Insurance Programs

This chapter is joint work with Max Groneck and Raun van Ooijen. We thank Rob Alessie, Lieke Beekers, Hugh Gravelle, Jawa Izza, Minjoon Lee, Raymond Montizaan, Patrick Moran, Søren Leth-Petersen, and seminar participants at various seminars and conferences for valuable contributions. We thank Netspar for financing data access through two theme grants: 'Uncertainty over the life cycle: implications for pensions and savings behavior' (2017-2020) and 'Health and labor market uncertainty over the life cycle: The impact on households' risk capacity and retirement income adequacy' (2022-2024).

4.1 Introduction

Health is strongly associated with socioeconomic status (Deaton, 2002; Chetty et al., 2016). This is a fundamental aspect of inequality in society with important implications for the progressivity of old-age social insurance programs, such as social security and public long-term care (LTC) insurance (Poterba, 2014; Auerbach et al., 2017). As the income-rich live longer than the income-poor, they receive more years of social security benefits (see, e.g.: Deaton, 2002; Smith, 2007; Chetty et al., 2016). In contrast, better health may induce lower LTC needs for the income-rich, implying fewer years of costly out-of-pocket LTC expenditures, such as co-payments for nursing home use (see, e.g.: Goda et al., 2011b; Jones et al., 2018; Rodrigues et al., 2018; Tenand et al., 2020a).

Health inequalities could imply an unintended income-regressive redistribution, raising two important questions: What is the size of the welfare gain for households with higher socioeconomic status due to expecting to live longer and to use LTC for a shorter time, and what mechanisms generate this gain? Such analysis requires a structural life cycle model that goes beyond conventional comparison of lifetime benefits and taxes (Goda et al., 2011a; Bosworth et al., 2016) because welfare consists non-monetary factors, including the utility of consumption, bequeathing, and living longer (Bernheim, 1987).

This chapter quantifies differences in the distribution of welfare that arise due to socioeconomic inequalities in health. Furthermore, we investigate the mechanisms behind the differences, particularly LTC co-payments and leaving bequests. Bequests are relevant as earlier research finds that wealthier households value these, and households can enlarge them when lifetime social insurance benefits are higher (De Nardi et al., 2010; Ameriks et al., 2011; Lockwood, 2018). We develop a life cycle model of singles and couples where households value consumption, bequeathing, and living longer and are exposed to uncertain income during working age, and uncertain LTC use and mortality after retirement. In the structural model, LTC use and survival risks differ exogenously by gender, marital status, and lifetime income quintiles to replicate the availability of informal care and the presence of socioeconomic differences in health. Our study focuses on the Netherlands which has generous and comprehensive public LTC provision (Bakx et al., 2023), including means-tested co-payments for nursing home care. We estimate the model using unique administrative data on income, assets, LTC needs, and mortality from 2006 to 2014. We use the estimated model to compute how much consumption compensation each lifetime income quintile would require to be indifferent to being exposed to the LTC use and mortality risk of the bottom lifetime income quintile (cf. De Nardi et al., 2024). To examine the impact of bequests and LTC co-payments, we remove them from the baseline model and re-compute the so-called consumption compensation equivalent.

Our approach and welfare measure are related to De Nardi et al. (2024). They use a structural life cycle model to quantify the lifetime cost of poor health for different initial health types. By assuming away the existence of poor health, they quantify the welfare cost of poor health for distinct health types. While we closely follow their approach, we conceptually differ as we shut down heterogeneity in poor health rather than the possibility of being in poor health.

Our estimation proceeds in two steps. First, we estimate income, LTC, and mortality risk processes and calibrate the risk aversion parameter. Second, we include these health and income risks in a structural life cycle model and estimate its key behavioral parameters: the subjective discount factor, consumption equivalence scale, the strength of the bequest motive, and the extent to which bequests are a luxury good. We estimate the parameters by matching simulated asset profiles to key aspects of the data, including asset holdings by marital status and lifetime income group. Also, we calibrate a parameter ensuring that households prefer living over death in utility terms (Hall and Jones, 2007). After that, we use the estimated model to make counterfactual predictions.

Aligned with studies from the U.S., our findings identify leaving bequests as an important channel for the income-rich to save: we find the marginal propensity to bequeath to be unit value for every euro above a consumption level of 40 thousand euros. This saving motive almost exclusively involves households in the top lifetime income quintile; hence, bequests are luxury goods. Turning to differences between singles and couples, the estimated equivalence scale of consumption is 1.15 and lower than usually documented in the literature (see e.g., De Nardi et al., 2021), implying Dutch households can save more due to stronger economies of scale. Lastly, the estimated subjective discount factor of 0.96 reveals a moderate preference for current consumption.

In a subsequent counterfactual analysis, we find that moving from the counterfactual of no health differences to the baseline where health differences exist, increases consumption by 23.4% for the top income quintile after age 65. In monetary terms this is a gain of 11.2%, driven mainly by more retirement benefits. Next, we assume away a preference for bequest saving and find that the welfare gain of 23.4% shrinks to 1.2% for the top lifetime income quintile. Hence, much of the welfare gain due to health inequalities stems from leaving larger bequests. Finally, if we remove co-payments, the welfare gain remains 21.8%, implying that valuable bequests rather than co-payments explain the welfare gain. For policy-makers, increased bequest taxes could thus be a way to alleviate welfare gains due to living longer and using less LTC.

This chapter contributes to several literature strands. A recently developed literature quantifies the lifetime cost of (self-reported) bad health (see De Nardi et al., 2024, and the references therein). We apply their approach to the large macro-oriented literature that characterizes the redistribution of old-age social insurance, programs including Medicaid (e.g., De Nardi et al., 2016; Braun et al., 2017), Medicare (e.g., McClellan and Skinner, 2006; Bhattacharya and Lakdawalla, 2006), social security (e.g., Goda et al., 2011a; Fehr et al., 2013; Groneck and Wallenius, 2021), and co-payments for LTC (e.g., Wouterse et al., 2021). Auerbach et al. (2017) advocates a more holistic accounting approach that includes all old-age social insurance programs to report progressivity. Closest to our study, Bagchi (2019) and Jones and Li (2023) use a structural life cycle model to study the interaction between heterogeneous mortality rates and social security benefit formula. We innovate this literature by examining the contribution of heterogenous LTC use and bequests to the redistribution of old-age social insurance.

This chapter also contributes to the quantitative-micro literature on retirees' saving behavior. The desire to leave a bequest has received considerable attention as a potential explanation for why more affluent households retain high levels of wealth in old age (De Nardi et al., 2010; Lockwood, 2018; Ameriks et al., 2020; Nakajima and Telyukova, 2024). However, the relative importance of saving for a bequest and precautionary saving varies depending on the estimation strategy and data. De Nardi et al. (2010) finds an insignificant bequest saving motive, arguably because savings in the U.S. are simultaneously used to pay for high out-of-pocket medical expenditures and to leave as a bequest (Dynan et al., 2004). Furthermore, the income-rich are under-represented in many surveys, including their Health Dynamics of the Oldest Old (AHEAD) data set. Lockwood (2018) instead finds a significant bequest saving motive by simultaneously fitting data on wealth and LTC insurance ownership. They argue that LTC insurance ownership acts as an exclusion restriction to separately identify a bequest saving motive. We add to this literature by using data from a country where the need for precautionary saving against out-of-pocket medical expenditures is low and where the income-rich are well-represented in the administrative data.

Besides, we link to the scarce literature that studies different saving behaviour by couples and singles within a life cycle model (e.g., De Nardi et al., 2021). Beside that couple member's can care about the welfare of a surviving partner, we capture the link between availability of informal care and formal LTC cost. It should be noted that for parsimony, we do not model the determinants of informal care; papers addressing such endogeneity stemming from altruistic and strategic informal care provision include Barczyk and Kredler (2018) and Ko (2022).

The chapter is organized as follows. Section 4.2 presents the socioeconomic differences in health. Section 4.3 describes the life cycle model. Section 4.4 provides the data and estimation procedure. Section 4.5 discusses the second-step estimation results. Section 4.6 performs the counterfactual health experiment. Section 4.7 discusses and concludes.

4.2 Socioeconomic Differences in LTC and Mortality

Before analyzing the welfare gain due to higher socioeconomic groups living longer and using less long-term care (LTC), it is crucial to examine how large these differences are. To quantify the differences, we use the same data and methods as in Chapter 3. In the analysis, we focus on 65+ individuals who are or were married at age 65. LTC use consists of institutional care use (Chapter 3 also included initial singles and home-based care use).[1] The sample contains 2,548,664 individuals and 1,487,109 households. See Appendix D.2.1 for a detailed description of the data and a summary of the estimation method.

Table 4.1 summarizes the remaining life expectancy (LE) and LTC use for men and women at age 65. We find opposite socioeconomic differences in LTC use and remaining life expectancy. Individuals within the top income quintile make less use of LTC but live longer. Men within the top income quintile live 3.6 years longer than their bottom income counterparts. For women this difference is 0.7 years. Men in the bottom income quintile use LTC for 0.1 years more years than their top income counterparts. For women this difference is 0.7 years. The difference in LTC use is larger for women and amounts to 26% of their average duration of using LTC. The larger difference for women can partially be explained by the fact that they often outlive their partner who often provides informal care.

[1] Home-based care use is not a separate state because its co-payments and, thus, redistributive effects are very limited in the Netherlands (Tenand et al., 2020b).

Table 4.1: Life Expectancy (LE) and Long-term Care Use (LTC) by Lifetime Income Quintiles

(a) Men	All	Bottom	Second	Third	Fourth	Top	ΔTop - Bottom
LE at age 65 (years)	18.9 (18.8;19.0)	16.4 (16.1;16.7)	17.6 (17.4;17.8)	18.4 (18.2;18.5)	19.2 (19.1;19.4)	20.0 (19.8;20.1)	3.6 (3.2;3.9)
LTC (years)*	1.9 (1.8;1.9)	1.8 (1.7;1.9)	2.0 (1.9;2.0)	2.0 (1.9;2.1)	1.9 (1.8;2.0)	1.7 (1.7;1.8)	−0.1 (−0.2;0.0)
Ever use LTC (%)	45 (44;45)	42 (40;43)	45 (44;46)	46 (45;47)	46 (45;47)	43 (42;44)	1 (0;3)
(b) Women							
LE at age 65 (years)	22.4 (22.3;22.4)	21.8 (21.5;22.0)	22.2 (22.0;22.4)	22.2 (22.1;22.4)	22.5 (22.3;22.6)	22.5 (22.4;22.7)	0.7 (0.4;1.1)
LTC (years)*	2.9 (2.9;2.9)	3.3 (3.1;3.4)	3.3 (3.2;3.4)	3.1 (3.0;3.1)	2.9 (2.8;2.9)	2.5 (2.5;2.6)	−0.7 (−0.9;−0.6)
Ever use LTC (%)	63 (62;63)	66 (65;68)	68 (67;69)	66 (65;67)	64 (63;64)	58 (57;58)	−9 (−10;−7)

Notes: The results are based on Dutch administrative records on individual and household marital status, gender, income, assets, institutional care use, and death between 2006 and 2014. The history of marital status dates back to 1995. Complete life histories on LTC use, deaths, and marital status are simulated according to the model of Chapter 3. The presented numbers are population-averaged measures for the life cycle simulation of 100,000 households. We present the median estimates across 1,000 bootstrapped samples and the 2.5^{th} and 97.5^{th} percentiles between brackets. Appendix D.2.2 provides the goodness-of-fit between the simulated and empirical survival probabilities and LTC use rates by age, lifetime income, and gender.

4.3 Life Cycle Model

We develop a life cycle model with uncertain LTC use and mortality to quantify the welfare gain of the higher lifetime income quintiles using less LTC and living longer. At every age $t \in \{25, 26, ...100\}$, a household maximizes lifetime utility by choosing total consumption expenditures c and savings a. The savings also determine the bequest that is left upon the death of the last household member. Households derive utility from consumption, leaving bequests, and being alive (independent of consumption). For tractability, we assume that household members are the same age such that a single age suffices to characterize the household.

A household has one of the following family statuses (f): a couple, single woman, or single man. Households enter the model as a couple initially and remain a couple until retirement at age 65, so there is no divorce or widowhood. Also, we assume no use of LTC before age 65 because of low likelihood.[2] After age 65, survival and use of LTC become uncertain, and couple households can become a single woman or single man household.

4.3.1 Preferences

The per-period CRRA utility functions of couples (C) and singles (S) are given by:

$$u^C(c) = 2 \cdot \frac{\left(\frac{c}{\eta}\right)^{1-\sigma}}{1-\sigma} + \bar{b}, \quad \text{and} \quad u^S(c) = \frac{c^{1-\sigma}}{1-\sigma} + \bar{b}, \quad \sigma \geq 0, \quad 1 \leq \eta \leq 2, \quad \bar{b} \geq 0,$$

where the parameter $\sigma \geq 0$ reflects the level of risk aversion.

Following the literature (De Nardi et al., 2021), we allow couples to benefit from economies of scale. Partners can pool their income and can consume goods jointly. η determines the extent to which households benefit from economies of scale. $\eta < 2$ features economies of scale: each couple member consumes $\frac{c}{\eta}$ units while this would be $\frac{c}{2} < \frac{c}{\eta}$ if they are single (Browning et al., 2013).

[2] At age 65, only 1% of the sample uses LTC.

Following De Nardi et al. (2024), we introduce scaling parameter $\bar{b} > 0$. This parameter is crucial when examining the welfare implications of altered life expectancies because households could attach value to the 'invisible' good of being alive that goes beyond consuming and bequeathing, e.g., the happiness of being alive.[3] In our model, risk-averse households would reach higher utility when being dead because $u^S < 0$ and $u^C < 0$ and utility from death is zero. We assume that utility from being alive is higher, thus calibrating a $\bar{b}$ yielding non-negative utility in any state when alive: $u^S \geq 0$ and $u^C \geq 0$.

The household derives utility $\mathcal{B}(a)$ from leaving bequest a. Following De Nardi (2004):

$$\mathcal{B}(a) = \frac{\phi}{1-\phi}^{\sigma} \cdot \frac{\left(\frac{\phi}{1-\phi} \cdot c_a + a\right)^{1-\sigma}}{1-\sigma} \quad \text{if } \phi \in (0,1),$$

$\mathcal{B}(a) = c_a^{-\sigma} \cdot a$ if $\phi = 1$ and $\mathcal{B}(a) = 0$ if $\phi = 0$, which De Nardi (2004) introduced to be consistent with wealth concentration among the wealthiest households in the U.S.. c_a is the consumption level below which households, under perfect certainty, will not leave a bequest (Lockwood, 2018). $c_a > 0$ implies bequests to be luxury goods. If households' wealth meets threshold c_a, ϕ is the share of excess wealth spent on a bequest: higher ϕ increases marginal utility from bequeathing relative to marginal utility from consuming.

4.3.2 Sources of Uncertainty

An important empirical artifact to be replicated is heterogeneity in asset holdings. A source for heterogeneity is uncertainty, forcing households to make precautionary savings (Carroll, 1997). We have uncertain health, family status, and income in our model.

Use of LTC and survival After age 65, exogenous health and family status shocks occur. The health of the husband and wife, h^m and h^f, evolve jointly and can differ between them ($h^m \neq h^f$). h^m and h^f take three values: a household member does not use public institutional care ($i = 1$), uses public institutional care ($i = 2$), or is dead

[3] In the literature, this parameter is used to compute the Value of a Statistical Life, i.e., the price that a population is willing to pay to prevent one certain death in the current period (see, e.g., Hall and Jones, 2007; St-Amour, 2022). This statistic is outside the scope of our study.

$(i = 3)$. LTC use induces co-payments (out-of-pocket expenditures) $m(y, a, h^m, h^f)$ that depend on income (y) and assets (a); these are paid to the government. We assume that LTC needs are homogenous across institutionalized individuals, so co-payments do not depend on the severity of the need for care.

We assume a Markovian process, so transition probabilities depend on the health and survival statuses of the preceding period: h_t^m and h_t^f. Survival status of a spouse controls for the potential availability of informal care. Furthermore, the health transition probability depends on lifetime income I and age t. Health transition probability π is:

$$\pi_{k,l}^{i,j}(t, I) = \mathbb{P}(h_{t+1}^m = k, h_{t+1}^f = l \mid h_t^m = i, h_t^f = j, t, I) \quad \text{with: } (i, j, k, l) \in \{1, 2, 3\}.$$

In particular, the death probability of the household is as follows:

$$\pi_{3,3}^{i,j}(t, I) = \mathbb{P}(h_{t+1}^m = 3, h_{t+1}^f = 3 \mid h_t^m = i, h_t^f = j, t, I) \quad \text{with: } (i, j) \in \{1, 2, 3\}.$$

Life cycle income: age 25 to 65 Exogenous income shocks happen during working life, reflecting the presence of labor supply shocks and health shocks. To save on the state space, we assume that these income shocks occur at the household level. Following the standard literature (Storesletten et al., 2004; French, 2005), household income dynamics follow an AR(1) process:

$$\begin{aligned} y_t &= \min(\widetilde{y}_t; \underline{y}) \qquad (4.1) \\ \widetilde{y}_t &= \alpha_t \cdot \exp(\theta) \cdot \exp(\eta_t) \cdot \exp(\epsilon_t) \\ \eta_t &= \rho \cdot \eta_{t-1} + u_t \\ \theta &\sim \mathcal{N}(0, \sigma_\theta^2); \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2); \quad u_t \sim \mathcal{N}(0, \sigma_u^2); \quad \eta_{24} = 0, \end{aligned}$$

where y_t is pre-tax household income, including income from labor, capital, and social insurance. α_t a deterministic age effect. θ is a fixed (labor) productivity effect. η_t is a persistent shock. ϵ_t is a transitory shock, in part reflecting transitory health shocks.

η_{24} is the initial level of the persistent income part. $\underline{y}$ is a government-provided income floor.

Life cycle income: age 65 and older Households receive retirement income $y_t = SS(f) + DB_{65}(f)$ consisting of a part independent of the income history, SS, and a part DB_{65}, whose defined benefit formula depends on the income history $\{y_s\}_{s=25}^{64}$. Income depends on the family structure because it becomes a smaller survivor benefit upon widowhood. Retirement income is stochastic due to random shocks in the income history until age 65 and time-varying family status.

4.3.3 The Government

The government provides income and LTC insurance after retirement by providing a first pillar pension and (partially) covering institutional care costs. Households pay mandatory for this insurance via dedicated taxes $\tau_{SS}(y)$ and $\tau_L(y, f, t)$. Moreover, co-payments $m(y, a, h^m, h^f)$ finance LTC use. Lastly, households pay a general income tax $\tau_G(y, f, t)$. We specify the functional forms of tax function τ in Appendix D.2.6. We specify m in Section 4.4.

Government revenues and costs in the model do not necessarily balance, which we ensure with additional lump-sum transfers Tr_{SS} and Tr_{LTC}. Appendix D.1.1 describes the procedure for how the government sets these transfer levels.

4.3.4 Optimization Problem

The timing is as follows: at the beginning of the period, households observe their state variables $\aleph$ that are relevant to their decision-making. The household obtains interest rate r on assets a, obtains income y, pays taxes τ and co-payments m, and makes the government-balancing transfers Tr. Then, based on state vector $\aleph$, households consume or save the remaining assets. Lastly, a survival and LTC use shock hits. If the final household member has died, any remaining assets go to the household's heirs (we assume households value their gross bequest and, therefore, ignore bequest taxes).

The state vector, $\aleph$, represents variables that are commonly observed by the household at the beginning of each period t:

$$\aleph_t^W = (a_t, \theta, \eta_t, \epsilon_t, \text{DB}_t, t)' \qquad \text{(if } t < 65)$$

$$\aleph_t^R = (a_t, \text{DB}_{65}, f_t, h_t^m, h_t^f, t)', \qquad \text{(if } t \geq 65)$$

where after age 65, retirement income replaces stochastic income, and family status f_t and health statuses h_t^m and h_t^f become uncertain. DB_t is the pension accrual until age t.

Note that all variables are known before deciding consumption c_t and next period's assets a_{t+1}, so we can recursively write the household's problem. Denote β the subjective discount factor. The household's value function at age t is:

$$V_t(\aleph_t^W) = \max_{c_t, a_{t+1}} \; u^C(c_t) + \beta \cdot \mathbb{E}[V_{t+1}(\aleph_{t+1}^W) \mid \boldsymbol{\aleph_t^W}]. \qquad \text{(if } t < 65)$$

$$V_t(\aleph_t^R) = \max_{c_t, a_{t+1}} \; u^f(c_t) + \beta \cdot \left(1 - \pi_{3,3}^{i,j}(t, I)\right) \cdot \mathbb{E}[V(\aleph_{t+1}^R) \mid \boldsymbol{\aleph_t^R}]$$
$$+ \beta \cdot \pi_{3,3}^{i,j}(t, I) \cdot \mathcal{B}(a_{t+1}), \qquad \text{(if } t \geq 65)$$

subject to a budget constraint and no-borrowing constraint, defining next period's assets:

$$a_{t+1} = (1 + r) \cdot a_t + y_t - \tau_G - \tau_{SS} - \tau_L - m_t - \text{Tr}_{SS} - \text{Tr}_{LTC} - c_t \geq 0.$$

The dynamic optimization problem after age 65 is different due to health uncertainty. A household survives into the next period with probability $1 - \pi_{3,3}^{i,j}(\cdot)$, and then faces the optimization problem again (V_{t+1}). With probability $\pi_{3,3}^{i,j}(\cdot)$, the household leaves a bequest with utility flow $\mathcal{B}(a_{t+1})$. Also co-payments for LTC use might occur ($m_t \neq 0$). We discuss the numerical implementation in Appendix D.1.2 to D.1.4.

As will be later important for our counterfactual analyses, health h_t^f and h_t^m impact the decision problem both via the utility function and budget constraint. The survival probabilities are lower when using LTC, implying that future consumption is more heavily discounted and households save less for future consumption. Health ambiguously affects the decision problem via co-payments. On the one hand, co-payments for LTC limit the available budget for consumption, inducing the household to precautionary save. On the other hand, a co-payment depends on assets and puts a penalty on saving.

4.4 Data and Estimation Procedure

We use administrative data from Statistics Netherlands that is available under restricted access. We can merge different data sets within the secured environment based on a unique individual and household identifier. Data come from multiple sources and registries: tax files (income and assets), municipal population registries (marital status, gender, birth year, and age), and a registry on institutional care use and deaths.

We use a two-step strategy similar to Gourinchas and Parker (2002) and De Nardi et al. (2010) to estimate the unknown parameters of our life cycle model. In the first step, we estimate the parameters for the health and income processes directly from the data, denoted by $\boldsymbol{\chi}$. Also, we tailor the pension and LTC use system to the Dutch setup 2006-2014. We fix the risk aversion and interest rate to $\sigma = 3$ and $r = 2\%$, values commonly used and found in life cycle studies (see, e.g., De Nardi et al., 2010).

Given the parameters from the first stage, we estimate the remaining parameters. To this end, we apply the method of simulated moments, i.e. we minimize the sum of squared differences between empirical and simulated moments of the asset distribution. The parameters to estimate are the subjective discount factor, bequest utility parameters, equivalence scale of consumption, and government-balancing transfers: $\boldsymbol{\delta} = (\beta, \phi, c_a, \eta, \mathrm{Tr}_{SS}, \mathrm{Tr}_{LTC})'$. After estimating all the parameters, we calibrate $\bar{b}$, i.e., the scaling parameter for the utility of surviving households.

4.4.1 First-Step Calibration and Estimation

Use of LTC and survival We estimate the health transition matrix using our simulated sample on household use of LTC from Section 4.2. We convert the life histories from continuous time to discrete time (an age period of one year), and compute transition probabilities accordingly. LTC use is assumed to be used throughout the entire age period and yearly costs the government €58,500 per user. The model is estimated using daily reported deaths, institutional care use, and marital status between 2006 and 2014. See Chapter 3 and Appendix D.2.1 for a detailed description of the data and a summary of the estimation method, including the computation of lifetime income (quintiles).

Co-payments for LTC use In the Netherlands, households make a co-payment to finance the use of LTC. The co-payment depends on the asset level a, household income y, and health statuses h^m and h^f. Households pay a low-rate or high-rate co-payment depending on the LTC used by the household members (h^m and h^f). The low-rate co-payment applies to couples with only one LTC user. The high-rate co-payment applies to singles and households with two LTC users:[4]

$$m(y,a,\cdot)\begin{cases}\max[1,900,\ \min[9,800,\ \ 0.125\cdot(y+0.04\cdot a)]] & \text{(low co-pay)}\\ \max[0\qquad,\ \min[27,000,\ \ 0.75\cdot(y_{AT}+0.04\cdot a-4,500)^{+}]] & \text{(high co-pay)}\end{cases}$$

The main difference between the two co-payment types stems from the cap on co-payments, €9,800 vs. €27,000, and the co-pay rate on income: 0.125 vs. 0.75. Also, note that contrary to low-rate co-payments, high-rate co-payments depend on income after taxes $y_{AT} = y - \tau_G - \tau_{SS} - \tau_L$. Lastly, 4% of the assets contribute to co-payments, implying endogenous co-payments in the model. In 2013, a policy change imposed an additional 8% of the assets to count for the co-payments. However, we stick to the 4% because that spans most of our sampling window (2006-2014).

Life cycle income: age 25 to 65 To estimate the income shock process (y_t), we use income data available for a representative sample of about 1% of the households (the IPO sample). In this sample, we have information on the distinct categories that

[4]We keep the formula simple for computational reasons, but the system is more complex in practice. Income and assets are measured with a two-year lag, implying we would have two additional state variables in our model. A low co-pay rate applies for the first four months of an institutional stay, which we cannot measure with our model specified at the year level instead. Also, there is an asset exemption of about €21,000 and €42,000 for singles and couples, but we follow Wouterse et al. (2021) and (partially) replace this with a general exemption of 0.75× €4,500 for the high co-payment.

comprise household income (the IPO sample), including taxes and private and public pension benefits. The data are available for a longer period than the data for the health processes: 2001-2014. A longer sampling window is important for estimating the persistence, i.e. longstanding effects, of income shocks.

We observe pre-tax income aggregated to the household level, including social transfers and pension income. This income definition also includes taxes for first pillar pension income and LTC provision, and a general income tax but excludes other dedicated taxes, e.g., for unemployment insurance. We only include the income of the household head and the partner (if applicable) and exclude the income of other household members. The variables are normalized to base year 2015 with the Consumer Price Index.

To abstract from early retirement decisions and schemes, we restrict our sample to couples whose oldest member is born after 1949 and whose primary income source is not retirement income. Further, we only include income above the government-provided safety net (welfare level): $y > \underline{y} = €15,600$ (2010-level). $\underline{y}$ is a government-provided income floor, equivalent to a consumption floor, as in, e.g., De Nardi et al. (2024).

We follow Storesletten et al. (2004) for the estimation of the income shock process. We estimate the age effect α_t and productivity effect θ_i by running a fixed effects regression of log income on age dummies (one for each $\log(\alpha_t)$) and a household fixed effect (θ_i):

$$\log(y_{it}) = \log(\alpha_t) + \theta_i + \eta_{it} + \epsilon_{it}, \tag{4.2a}$$

where i indexes a household and t the age of the oldest household member.

Ideally, our household-specific estimate $\widehat{\theta}_i$ excludes birth year effects. To wash out cohort effects, we run the following OLS regression of the predicted productivity effects on birth year dummies (cf. French, 2005; De Nardi et al., 2024):

$$\widehat{\theta}_i = \overline{c} + \overline{\theta}_c + \widetilde{\theta}_i, \ c \in \{1951, .., 1990\}, \tag{4.2b}$$

where $\overline{c}$ is the cohort effect of birth year 1950, $\overline{c} + \overline{\theta}_c$ is the cohort effect for birth years

1951-1990, and residual $\widetilde{\theta}_i$ is the household productivity effect excluding a cohort effect. We use $\widetilde{\theta}_i$ as the household-specific productivity effect.[5]

Next, we estimate the parameters of the income shock $\theta_i + \eta_{it} + \epsilon_{it}$: ρ, σ_θ, σ_u, and σ_ϵ. To this end, we construct the empirical auto-covariance matrix of the predicted residuals of $\tilde{\theta}_i + \eta_{it} + \epsilon_{it}$ from (4.2a) and (4.2b), and match them to the auto-covariances implied by equation (4.1). Appendix D.2.4 further explains the GMM procedure and shows the fit.

Table 4.2: Parameters of the AR(1) Income Process

Parameter:	ρ	σ_θ	σ_u	σ_ϵ
	0.966 (0.004)	0.184 (0.028)	0.131 (0.008)	0.166 (0.003)

Notes: Estimates for married households whose oldest member is younger than age 65 and born after 1949. Data from IPO 2001-2014: 77,118 households and 534,006 panel-year observations. Standard errors in parentheses.

Table 4.2 provides the results on the income shock. The estimated parameters align with results in the literature (Storesletten et al., 2004; Karahan and Ozkan, 2013; Blundell et al., 2015; Paz-Pardo and Galves, 2023). This also holds for the high income persistence $\rho = 0.966$ we estimate: income shocks have longstanding effects.

Life cycle income: after age 65 In the Netherlands, first pillar pension income is independent of income history $\{y_s\}_{s=25}^{64}$ but linked to minimum wage w. For couples, the benefit level is minimum wage ($SS = w$). For singles, the benefit level is 70% of the couple's benefit ($SS = 0.7w$). As minimum wage we take the 2010-value: $w = €18,240$.

A household in the model is also entitled to a second pillar pension benefit DB_{65}, which is linked to the history of income shocks $\{y_s\}_{s=25}^{64}$. In practice, the first and second pillar aim to replace about 75% of the average individual-earned income or obtained disability insurance income (Knoef et al., 2017).[6] We assume the same replacement rate and benefit formula at the household level. The second pillar pension income is

[5] Appendix D.2.3 shows the model estimates for the age profile $\{\overline{c} + \log(\alpha_t)\}_{t=25}^{64}$.

[6] We keep the formula simple for computational reasons, but the system is more complex in practice.

only accrued over the income y_t that exceeds $\frac{100}{75} \cdot SS$ because social security benefits are sufficient to replace the income below this level. The evolution of the second pillar pension benefit is:

$$\text{DB}_{t+1} = \text{DB}_t + \frac{1}{40} \cdot 0.75 \cdot \min\left(y_t - \frac{100}{75} \cdot SS \;\; ; \;\; 0\right) \qquad \text{if } t \;\leq\; 64,$$

where the factor $\frac{1}{40}$ makes sure we take a 40-year average of pre-tax household income.

Together, first and second pillar pensions compose income after retirement ($t \geq 65$):

$$y_t(\text{DB}_{65}, f) = y_t(\{y_s\}_{s=25}^{64}, f) = \begin{cases} w + \text{DB}_{65}, & \text{if} \quad f = \text{couple} \\ 0.7w + \text{rr}_w \cdot \text{DB}_{65}, & \text{if} \quad f = \text{single woman} \\ 0.7w + \text{rr}_m \cdot \text{DB}_{65}, & \text{if} \quad f = \text{single man.} \end{cases}$$

If a spouse dies, rr_w and rr_m convert a couple's pension benefit into a widow(er)'s pension benefit. Using the IPO data, we find $\text{rr}_m = 0.93$ (SE: 0.001) and $\text{rr}_f = 0.55$ (SE: 0.005). In line with our earlier work van der Vaart et al. (2020), we report $\text{rr}_m > \text{rr}_f$ implied by that men were the prime earner in the households and pension benefits mostly accrued to them. Appendix D.2.5 contains the estimation details.

A crucial variable in our model is lifetime income quintile I, which determines the health risks after retirement. We take DB_{65} as the model-equivalent level of lifetime income, which is exogenous because households do not decide on labor supply. Consequently, we can compute the quintiles of the distribution of DB_{65} without running the life cycle model. Next, when running the life cycle model, we use the quintiles and realization of DB_{65} to assign households a lifetime income quintile group.

Taxation We estimate the tax function $\tau_{SS}(y)$, $\tau_L(y, f, t)$ and $\tau_G(y, f, t)$ by regressing observed tax amounts in the IPO on household income according to a log-linear and sigmoid specification. We apply non-linear least squares estimation and estimate the functions separately for households below and above age 65 and for single and married households. Appendix D.2.6 reports the specifications and estimates.

Remaining calibrations Table 4.3 displays the remaining first-stage parameters.

Table 4.3: Other First-Step Parameters

	Symbol:	**Value:**	**Source:**
Relative risk aversion	σ	3	Several empirical studies[1]
Interest rate	r	0.02	The average interest rate on savings 2006-2014[2]
First pillar pension benefit	w	€18,240	2010-level
Yearly LTC cost per user (€)	LTC_{cost}	€58,500	van Ooijen et al. (2015)

Notes: [1] See estimates by Cagetti (2003); De Nardi et al. (2010); Lockwood (2018); [2] See DNB Statistics: https://www.dnb.nl/statistieken/dashboards/rente/ [*retrieved on*: August 7^{th}, 2023]

4.4.2 Second-Step Estimation

In this step, we apply the method of simulated moments (MSM) estimation to match asset moments in the administrative data with moments simulated with the life cycle model (see, e.g., De Nardi et al., 2010; Lockwood, 2018; De Nardi et al., 2021). Using our estimated first-stage parameter vector $\boldsymbol{\chi}$, we try to find preference vector $\boldsymbol{\delta} \in \boldsymbol{\Delta}$ that yields model-generated asset profiles that 'best match' observed asset profiles. We do the matching by applying standard generalized method-of-moments (GMM) techniques.

For the empirical moments, we use the same population and lifetime income quintiles that we used to compute the health process in Section 4.2, i.e., households whose members are aged older than 65 and were married at age 65. Following seminal work on the elderly's asset holdings (De Nardi et al., 2010; Ameriks et al., 2020; Nakajima and Telyukova, 2024), we take net worth as our measure of wealth. This is the total assets minus mortgages and other debt. Total assets are defined as the sum of the values of checking and savings accounts, risky assets (stocks and bonds), business wealth, the owner-occupied house, other real estate, and other assets such as cash-on-hand. The value of risky assets is normalized with the Amsterdam Exchange close index (AEX) on 31/12/2014, the owner-occupied house and other real estate with the house price index (base 2015), and debt and amounts deposited in checking and savings accounts with the Consumer Price Index (base 2015).

To prevent an overly complex model, we do not separately treat financial wealth and net housing wealth, i.e., the total value of real estate minus outstanding mortgage debt. The co-payments are, however, based on financial wealth and exclude housing wealth in practice. We assume that households liquidate their housing wealth (sell their house) once they enter a public care institution. Hence, net worth and financial assets coincide.

We base our estimator on the age profile of the median net worth of married and single individuals between ages 65 and 100 by lifetime income quintile I, implying $2 \times 36 \times 5 = 360$ moment conditions. We do not consider matching means (cf. De Nardi et al., 2010) because these empirical moments are sensitive to outliers, thereby driving estimation results. Furthermore, we restrict the analysis to matching the asset distribution after age 65 because our studied welfare effects primarily occur after this age.

However, similar to estimating the income processes before age 65, we must first deal with cohort effects to observed asset profiles. We similarly account for this as specifications (4.2a) and (4.2b) do for the income process. To stay as close as possible to the 1950 cohort for which we estimated the income process, we made the assets representative for a reference group of households born between 1945 and 1949. Appendix D.2.7 provides details about how we econometrically deal with the cohort problem of assets.[7]

We compute the moments also for our simulated sample and compare them with the data moments using the objective function:[8]

$$\sum_{k=1}^{K=360} \left[\left(M_k^d - M_k^s(\hat{\boldsymbol{\chi}}, \boldsymbol{\delta}) \right)^2 \right],$$

with $K = 360$ moments, and where M_k^d and M_k^s are the k-th data and simulated moment.

[7]The regressions involve the logarithm of assets, so we only keep non-negative assets. Furthermore, the regression is prone to outliers, so we drop assets above €2,500,000. We drop 0.9% of the households and 2.6% of the panel-year observations because of these restrictions.

[8]Instead of matching medians directly, existing work (e.g. Cagetti, 2003) looks at how many households in the observed population have assets below the simulated median, which is ideally 50%. This means that at each iteration, we would have to use our administrative data to determine how many individuals have assets below the group-specific simulated median, which is computationally expensive. That condition and our condition are equivalent at the true value $\boldsymbol{\delta}$ so we choose our current approach.

Our estimator $\hat{\boldsymbol{\delta}}$ minimizes this quadratic distance between the empirical and simulated data moments. We do not weight each moment with an asymptotically optimal weight matrix, implying we have less efficient estimates. Instead, efficient estimates would follow from taking the densities evaluated at the median as weights (Powell, 1994), but estimating these weights is computationally too expensive.[9]

The procedure can be summarized as follows. We first estimate asset profiles from the administrative data. Second, we estimate the unknown parameters for the first stage. Then, we take the first-stage calibrations $\hat{\boldsymbol{\chi}}$ and a given parameter value $\widetilde{\boldsymbol{\delta}}$ and run the life cycle model. We store the decision rules of the life cycle model. We know the steady-state distribution of individuals over the state variables and can compute the simulated asset moments from that (see Appendix D.1.4 for the computation of the distribution). Hereafter, the value of the objective function is computed. Lastly, we compute a new 'optimal' preference vector using a Gauss-Newton regression and repeat the procedure until parameter vectors of two consecutive iterations are arbitrarily close. See Appendix D.3.1 for the computation of the standard errors.

Lastly, we calibrate $\overline{b}$, a crucial parameter when examining the welfare implications of shortening and extending lifespans (cf. Hall and Jones, 2007). Our additive speciation implies that $\overline{b}$ does not depend on the consumption and saving decision, so we do not have to jointly estimate this parameter with the other preference parameters, but rather calibrate it conditionally upon them. We tailor the parameter to the group that has the lowest-per-period utility in our population: retired singles without private pensions ($\text{DB}_{65} = 0$). We set $\overline{b} = -\frac{\underline{c}^{1-\sigma}}{1-\sigma} = 0.3114$, where $\underline{c} = 0.7w = €12,768$ is their consumption level (in 0000s €) and implying this group has zero utility from consumption. In a similar spirit, De Nardi et al. (2024) used an estimated consumption floor to pinpoint $\overline{b}$. Because we tailor $\overline{b}$ to the lowest consumption level, our estimated welfare gains from living longer will be a lower bound to the true effect.

[9]We also tried inverse-variance weighting (cf. Altonji and Segal, 1996). However, this implied non-sensible estimates as there are extremely large weights for low compared to high lifetime income quintiles.

4.4.3 Model Identification

β is identified by the shape of the age profile on assets: higher β implies a stronger preference for future consumption and thus more saving. In addition, the Euler equation provides intuition for the identification of preference parameters on bequests ϕ and c_a, and equivalence scaling η in our model. To see how this works for η, suppose a simple model without bequests of a married household in period t, that will be not be married in period $t+1$ anymore. If the sole uncertainty is death, then the Euler equation implies the following consumption growth:

$$\begin{aligned}\log\left(\frac{c^S_{t+1}}{c^M_t}\right) &= \log\left(c^S_{t+1}\right) - \log\left(c^M_t\right) \\ &= -\log(\eta) + \frac{1}{\sigma}\cdot\left(\log(\beta) + \log(1-\pi_{3,3}) + \log(R) - \log(2)\right),\end{aligned}$$

where c^S_{t+1} is consumption when single, and c^M_t is consumption when married. Here, higher η (less economies of scaling) implies more consumption spending c^M_t when married, so lower savings when married. Hence, we identify η by comparing asset levels of married and single households of a given lifetime income quintile at two consecutive ages.

Also, the Euler equation shows a complication when having to estimate β and σ. Their joint effect on savings would be $\frac{1}{\sigma}\cdot\log(\beta)$, making it impossible to separately identify the two when studying a given asset level. Therefore, we follow Ameriks et al. (2011) and fix $\sigma = 3$, a value common in retirement-savings literature (e.g. De Nardi et al., 2010).

Lastly, to see how the bequest parameters are identified, we consider a single household that knows to die next period, does not subjectively discount utility from consumption c, and obtains utility from leaving a bequest a ($c_a > 0$ and $\phi \in (0,1)$). Assume that the household has cash-on-hand μ, then the decision problem is:

$$\max_{c,a}\ u^S(c) + \mathcal{B}(a) = \max_{c,a}\ \frac{c^{1-\sigma}}{1-\sigma} + \overline{b} + \frac{\phi}{1-\phi}^{\sigma}\cdot\frac{\left(\frac{\phi}{1-\phi}\cdot c_a + a\right)^{1-\sigma}}{1-\sigma},\ s.t.\ \ \mu = a + c.$$

The Euler equation with bequests is:

$$c^{-\sigma} = \frac{\phi}{1-\phi}^{\sigma} \cdot \left(\frac{\phi}{1-\phi} \cdot c_a + a \right)^{-\sigma}, \quad s.t. \;\; c = c_a + (1-\phi) \cdot \mu \text{ and } a = \phi \cdot (\mu - c_a),$$

so in optimum households equate the marginal utility from bequests and consumption.

Increasing c_a one-to-one increases consumption c, and one-to-one decreases the bequest size a. c_a is thus a terminal wealth level that must be met before a household intends to leave a bequest (bequests are luxury goods).[10] The likelihood of meeting this criterion is larger for higher lifetime income quintiles, from whose terminal assets we identify c_a. Furthermore, ϕ is the share of excess wealth they leave as a bequest. We identify ϕ by comparing the steepness of the asset profile for this group with $\mu > c_a$ compared to the groups with insufficient wealth $\mu \leq c_a$, i.e. groups with low lifetime income.

4.5 Second-Step Estimation Results

Figure 4.1 shows the empirical and simulated moments for our closest match. For exposition, we connect the moments with a line. Overall, we have a reasonable fit: we match the positive correlation between the level of assets and lifetime income quintile and the asset decumulation pattern after age 65. We also mimic the empirical artifact that households in the top income quintile die with substantial assets, i.e., leave a bequest. Our model is less capable of matching the low asset holdings for the bottom and second income quintile, which could be explained by that these groups contain relatively many hand-to-mouth consumers and have lower discount rates (Cherchye et al., 2023). However, introducing heterogenous preferences would make it less clear where a welfare redistribution stems from and abstracts from the standard in the retiree's saving literature that we stick to, i.e. a parsimonious model with homogenous preferences (De Nardi et al., 2010; Ameriks et al., 2020). Yet, the general picture of asset profiles seems to be reproduced by our MSM estimation, making us confident in using our estimated life cycle model for further inference.

[10]In the model, the survival probability is below unit value, so c_a refers to annuitized cash-on-hand rather than the level of cash-on-hand.

Figure 4.1: Empirical and Simulated Assets by Lifetime Income Quintile and Marital Status

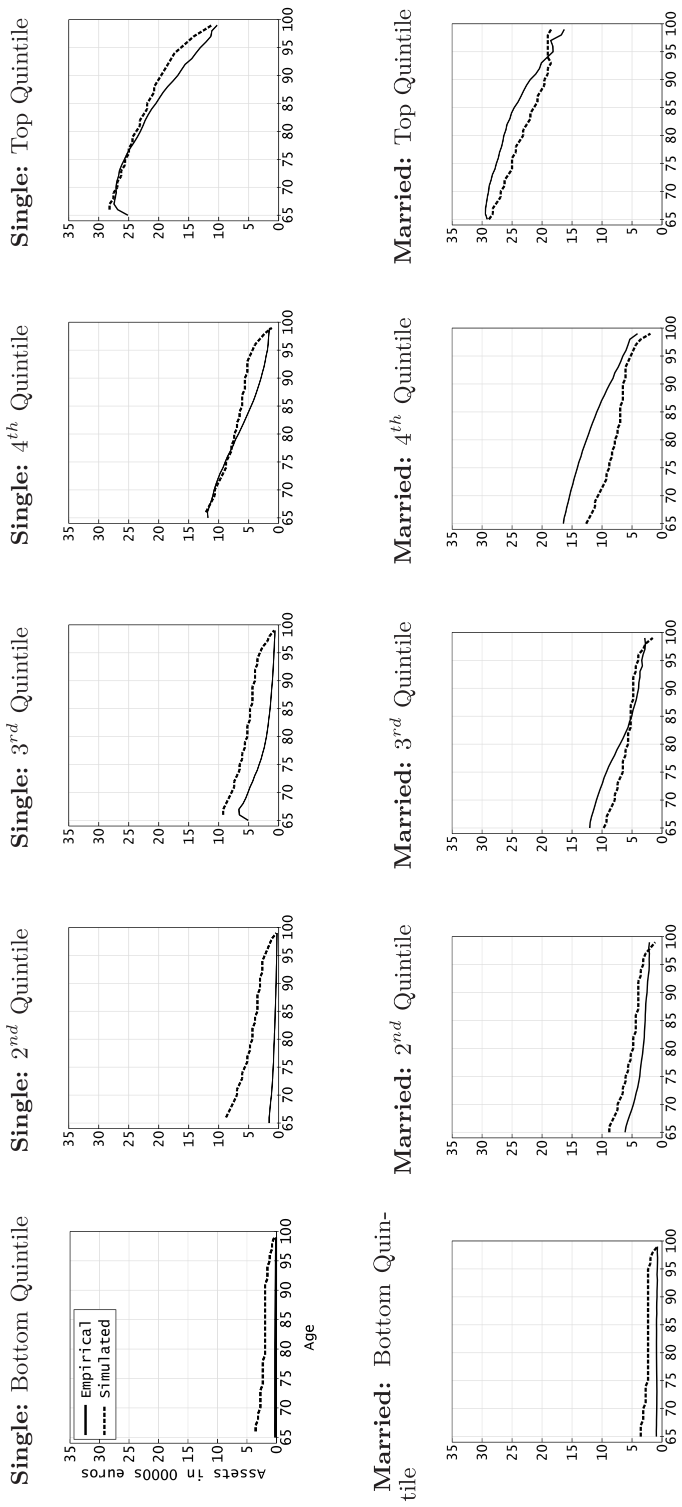


Table 4.4 presents the results of our preference parameter estimation. We estimate $\hat{\beta} = 0.960$, implying that households have a moderate preference for current over future consumption. The estimated bequest utility indicates a strong saving motive, where bequests are luxury goods ($\widehat{c_a} = 40,672 > 0$). We find the extreme case of $\widehat{\phi} = 1$, implying a linear bequest function, and all excess wealth is put into a bequest and not consumed. A high bequest propensity ($\phi > 0.88$) is common in the revealed preferences literature (De Nardi et al., 2010; Lockwood, 2018; De Nardi et al., 2024), while the stated preference literature finds lower values ($\widehat{\phi} > 0.48$, see, e.g. Ameriks et al., 2020). Our threshold consumption level $\widehat{c_a} = 40,672$ is close to De Nardi et al. (2010), who report $\widehat{c_a} = 34,000$, and slightly higher than other related studies (Lockwood, 2018; De Nardi et al., 2024).

Table 4.4: Estimated Structural Parameters

Discount factor	Bequest utility		Equivalence scale	Government transfer	
β	c_a	ϕ	η	Tr_{SS}	Tr_{LTC}
0.960	40,452	1.000	1.145	783.58	-433.05
(0.00002)	(1.03844)	(0.00054)	(0.00010)		

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors in parentheses. The data contain 1,471,858 households and 11,471,725 panel-year observations.

We find an equivalence rate of $\widehat{\eta} = 1.146$, which is lower than the commonly applied and estimated OECD-modified equivalence scale of 1.5 (for the life cycle model estimate, see, e.g., De Nardi et al., 2021). Lower equivalence scales are, however, also reported in the consumption-expenditure literature (see, e.g., Donaldson and Pendakur, 2004). Using the Euler equations from Section 4.4.3, our model predicts more savings than would be predicted if we take the OECD-modified equivalence scale $\eta = 1.5$. Hence, households in the Netherlands have relatively high economies of scale, implying they can save more.

The additional tax for singles to balance the government budget is $\widehat{\text{Tr}} = \widehat{\text{Tr}}_{\text{SS}} + \widehat{\text{Tr}}_{\text{LTC}} = 783.58 - 467.26 =$ €316.32 (for couples, this is double the amount). This consists of an additional tax to finance the first pillar pension ($\widehat{\text{Tr}}_{\text{SS}} > 0$) and a subsidy to finance the LTC system ($\widehat{\text{Tr}}_{\text{LTC}} < 0$). Given the low amounts we are talking about, we can think of these transfers reflecting measurement error due to calibration of the first-stage parameters.

4.6 Welfare Gain due to Lower LTC use and Mortality

In this section, we closely follow De Nardi et al. (2024) and use the estimated life cycle model to quantify the welfare gain arising from socioeconomic differences in LTC use and mortality (cf. Table 4.1). In the first step, we compute the monetary gain for any lifetime income quintile by counterfactually assigning them the health risks of the lowest lifetime income quintile. Besides, we evaluate the total welfare gain with a Willingness-To-Accept (WTA) metric that includes a non-monetary gain linked to reaching higher utility: the compensated consumption equivalence. As a final step, we utilize the unique feature of life cycle models that allows us to quantify the extent to which saving for a bequest and the existence of LTC co-payments contribute to the observed WTAs.

4.6.1 Counterfactual Analyses

At age 65, households draw an LTC use and mortality risk profile that depends on their lifetime income quintile, denote this baseline scenario by BS. We also have a counterfactual scenario, denoted by CF, where each household draws the health risks of the lowest lifetime income quintile, so health risks are homogenous. The counterfactual implies that higher lifetime income quintiles live shorter, so have lower lifetime retirement income, and have higher lifetime LTC use, so have higher lifetime co-payments for LTC. Furthermore, lifetime co-payments will be different under the counterfactual due to the endogeneity of assets. Lastly, lifetime government-balancing transfers will be different due to lower longevity and because we will re-calibrate $\widehat{\text{Tr}}_{\text{SS}}$ and $\widehat{\text{Tr}}_{\text{LTC}}$ to also match the government budget under the counterfactual.

We compute the net present value of retirement income net of co-payments and government-balancing transfers and take the difference between baseline and counterfactual scenarios as the monetary gain from heterogeneous health risks. We do this for each lifetime income quintile separately. In concordance with LTC use and mortality risk starting, we measure the net present value at age 65. For the two cases, denote with $y^{BS}(\aleph_t)$ and $y^{CF}(\aleph_t)$ the net incomes for a household aged $t \geq 65$ with state vector $\aleph_t$. Denote $\mathbb{E}_{65}(y^{BS}(\aleph_t))$ and $\mathbb{E}_{65}(y^{CF}(\aleph_t))$ their expected values measured when the household is 65. These expectations are unconditional upon survival after age $t \geq 65$ and thus include differential mortality. The difference $\mathbb{E}_{65}(y^{BS}(\aleph_t)) - \mathbb{E}_{65}(y^{CF}(\aleph_t))$ is the contribution of age t to the monetary gain, and the expected lifetime income gain is the sum of the age-specific gains:

$$\sum_{t=65}^{100} \frac{\mathbb{E}_{65}(y^{BS}(\aleph_t)) - \mathbb{E}_{65}(y^{CF}(\aleph_t))}{(1+r)^{t-65}}, \tag{Monetary gain}$$

where we deflate the income stream to age 65 with an interest rate of $r = 0.02$. Apart from this level estimate, we will decompose the monetary gain into parts stemming from pension income, LTC co-payments, and government transfers.

Because our counterfactual affects consumption decisions, and bequest decisions, and the utility of life expectancy, we follow De Nardi et al. (2024) and adopt the compensated consumption equivalence λ_c as a measure for the welfare gain. This measure is the minimum percentage points increase in counterfactual consumption that a household requires to prefer (accept) the 'worse' counterfactual over the baseline case, hence a Willingness-To-Accept (WTA).

Formally, the expected lifetime utility at age 65 in the baseline scenario, value function V_{65}^{BS}, is defined as:

$$\begin{aligned} V_{65}^{BS} &:= \sum_{t=65}^{100} \hat{\beta}^{t-65} \cdot \left\{ \mathbb{E}_{65}\left(u(c_t^{BS}(\aleph_t))\right) + \hat{\beta} \cdot \mathbb{E}_{65}\left(\mathcal{B}(a_{t+1}^{BS}(\aleph_t))\right) \right\} \\ &= \sum_{t=65}^{100} \hat{\beta}^{t-65} \cdot \left\{ \mathbb{E}_{65}\left((1+\mathbb{1}(\aleph_t)) \frac{\left(\frac{c_t^{BS}(\aleph_t)}{\hat{\eta}(\aleph_t)}\right)^{1-\sigma}}{1-\sigma} + \bar{b} \right) + \hat{\beta} \cdot \mathbb{E}_{65}\left(\hat{c_a}^{-\sigma} a_{t+1}^{BS}(\aleph_t)\right) \right\}, \end{aligned}$$

which is the sum of expected lifetime utility from consumption and bequeathing. c_t^{BS}

and a_{t+1}^{BS} are optimal consumption and a bequest at age t and $t+1$ for a household endowed with state vector $\aleph_t$. Note bequest utility is linear in assets because we estimate $\hat{\phi} = 1$.

Similarly, we determine the optimal consumption c_t^{CF}, bequests a_{t+1}^{CF}, and value function V^{CF} for the counterfactual case:

$$V^{CF}(\lambda_c) := \sum_{t=65}^{100} \hat{\beta}^{t-65} \cdot \left\{ \mathbb{E}_{65} \left((1 + \mathbb{1}(\aleph_t)) \cdot \frac{\left(\frac{(1+\lambda_c) \cdot c_t^{CF}(\aleph_t)}{\hat{\eta}(\aleph_t)} \right)^{1-\sigma}}{1-\sigma} + \bar{b} \right) \right.$$
$$\left. + \hat{\beta} \cdot \mathbb{E}_{65} \left(\hat{c_a}^{-\sigma} \cdot a_{t+1}^{CF}(\aleph_t) \right) \right\}.$$

We find WTA λ_c by solving: $V^{CF}(\lambda_c) = V_{65}^{BS}$. Without compensating ($\lambda_c = 0$), we expect less lifetime utility in the counterfactual scenario: $V^{CF}(0) < V_{65}^{BS}$. Because the utility is increasing in consumption, we have $\frac{\partial V^{CF}(\lambda_c)}{\partial \lambda_c} > 0$, and thus require $\lambda_c > 0$ to have $V^{CF}(\lambda_c) = V_{65}^{BS}$. $\lambda_c > 0$ represents the welfare gain: the closer this number is to zero, the smaller the welfare gain for the lifetime income quintile. Due to different deflation, we cannot directly compare λ_c to the monetary gain within a lifetime income quintile: monetary gains are obtained by using discount factor $\frac{1}{1+r}$, while λ_c is obtained by using discount factor $\hat{\beta} < \frac{1}{1+r}$.

In a final step, we look at the impact of LTC co-payments and saving for a bequest on the WTA. To this end, we one-by-one remove LTC co-payments and saving for a bequest ($\phi = 0$) for the baseline case and recompute optimal c_t^{BS} and a_{t+1}^{BS}, so V_{65}^{BS}. For the counterfactual, we keep c_t^{CF} and a_{t+1}^{CF} fixed and find λ_c that solves $V^{CF}(\lambda_c) = V_{65}^{BS}$.

4.6.2 Results

The first two lines in Table 4.5 show the average gain in lifetime income if health risks differ by lifetime income quintile, i.e., higher lifetime income quintiles use less LTC and live longer. The pecuniary gain per income group reveals a gradient favoring higher lifetime income quintiles. However, this result is incomplete because higher lifetime income quintiles by construction have higher absolute gains due to higher yearly income.

To back out level effects, the second row shows the gain relative to group-specific lifetime income under the counterfactual. The result confirms the gradient: the gain is −0.2% for the bottom income group and 11.0% for the top income group, a difference of 11.2pp.

When discussing welfare gains, we prefer the first-differenced estimate of 11.2pp., which accounts for the fact that bottom income groups, despite unchanged health risks, still loss or gain welfare under the counterfactual. The bottom income group namely loses 0.2% of lifetime income due a re-calibration of $\widehat{\text{Tr}}_{\text{SS}}$ and $\widehat{\text{Tr}}_{\text{LTC}}$. This gain is not directly linked to differences in health and is common to all income groups, and therefore we prefer the first-differenced estimate of 11.2pp.

With shares over 90%, we see that pension income is the largest contributor to the pecuniary gain for a lifetime income quintile. The role of LTC co-payments is non-negligible for the highest lifetime income quintile and explains 10.6% (€10,878) of their pecuniary gain. As a side-remark, the co-payments make up a small yet negative share for the second lifetime income quintile because their baseline LTC use is *higher* than under the counterfactual (see Table 4.1).

Table 4.5: Monetary and Welfare Gains Due To Socioeconomic Differences in LTC Use and Mortality: Levels and Decomposition

Lifetime income quintile	Bottom	Second	Third	Fourth	Top	ΔTop-Bottom
Monetary gain (€)	-917	14,075	26,864	49,812	102,474	103,391
Monetary gain[1] (%)	-0.2	3.2	5.2	8.0	11.0	11.2
Contribution to monetary gain[2] (%)						
Pension income	0.0	112.2	103.3	96.8	91.3	-
Co-payments	0.1	-2.8	2.3	6.8	10.6	-
Government transfers Tr_x	99.9	-9.4	-5.6	-3.5	-1.9	-
WTA[3]: $\lambda_c \times 100\%$	-0.2	0.8	2.9	7.9	23.2	23.4
No bequests ($\phi = 0$)	-0.5	0.1	1.7	4.9	0.7	1.2
No co-payments	2.4	3.6	5.6	10.4	24.2	21.8
No co-payments and bequests	2.3	3.4	5.1	8.7	3.3	1.0

Notes: [1] Expressed as a percentage of counterfactual lifetime income after age 65; [2] Gain of the particular income source in €s as a share of the monetary gain in €s (first row); [3] Willingness-To-Accept

The WTA confirms higher welfare gains for higher lifetime income quintiles, but what explains the gap of 23.4pp.? If we assume away saving for a bequest, the gap in welfare gain (WTA) between the top and bottom lifetime income quintiles shrinks from 23.4pp. to 1.2pp.. Hence, higher lifetime income quintiles benefit less from higher longevity and lower LTC use if they cannot save for a bequest. Their welfare gain dropped from 23.2% to 0.7% because they will not spend the additional lifetime income on their otherwise highly-valued bequests. On the contrary, lower lifetime income quintiles experience a much smaller drop in welfare gain because they value leaving bequests –luxury goods– much less. As a result, the difference in welfare gain between the top and bottom lifetime income quintile shrinks tremendously.

On the other hand, the gap between the top and bottom income groups remains a considerable 21.8pp. when we leave out LTC co-payments. Differences in co-payment duration are thus less important than a bequest to explain the excess welfare gain of the top lifetime income quintile. The gap remains large because LTC co-payments are a relatively small share of lifetime income gains (Row 4, Table 4.5). Moreover, higher lifetime income quintiles still receive the additional retirement income, which they can spend on –for them valuable– bequests.

Note that if we abolish LTC co-payments in the baseline scenario, any group experiences a welfare gain, which is good from a social planner's perspective. While their risks are not altered, the bottom lifetime income quintile has a welfare gain 2.4% because LTC co-payments are replaced by a higher transfer (tax) $\widehat{\text{Tr}}_{\text{LTC}}$ that is paid *unconditionally* upon LTC use. Lower socioeconomic groups can spend the otherwise co-paid resources on consumption, while the same is true for the higher socioeconomic groups who can additionally spend it on for them valuable bequests.

If we simultaneously assume away saving for bequests and LTC co-payments, we find welfare gains that are in between singling out only one of the two channels. In line with the evidence above, for lower lifetime income quintiles, the gain is closest to the case of singling out LTC co-payments only. In comparison, for higher lifetime income quintiles, the case is closer to singling out bequests only, as these are more valuable for them.

4.7 Discussion and Conclusion

We evaluate the welfare gain that Dutch households with higher lifetime income experience due to using less long-term care (LTC) and living longer. To this end, we estimated a life cycle model on singles and couples' consumption and saving behavior, including idiosyncratic risks on income, LTC use, and mortality. We calibrated the model to match Dutch administrative data on asset holdings from 2006-2014. Using the estimated model, we conducted three counterfactual experiments to quantify and explore possible channels of the welfare gain: (1) assign each household the LTC use and mortality risk of the bottom lifetime income quintile, (2) additionally remove a preference for bequest saving, and (3) replace co-payments for LTC with a fixed tax that is paid unconditionally upon using LTC.

Our findings highlight a sizeable excess welfare gain of 23.4pp. higher consumption for the highest lifetime income quintile if their health follows the true process rather than the counterfactual one. The large welfare gain for the top lifetime income quintile can almost exclusively be attributed to their preference for leaving bequests: the welfare gain reduces to 1.2pp. if households would not hold a preference for bequest saving. Our ranking exercise shows that LTC co-payments are less important when explaining the excess welfare gain: the gap remains 21.8pp..

The estimated welfare effects should be interpreted as a lower bound estimate because we calibrate the utility of remaining life-expectancy $\bar{b}$ at the lower end of possible values. This seems a sensible choice as Hall and Jones (2007) show that lower values of $\bar{b}$ better match healtcare expenditures in the U.S.. Yet it must be said that, in line with Hall and Jones (2007) and our own computations (available upon request), the estimated welfare effects are sensitive to higher choices of $\bar{b}$.

In line with our findings, earlier work emphasized that modelling bequest saving is crucial for understanding the asset holdings of the income- and asset-rich (De Nardi et al., 2010). However, earlier work is primarily conducted in the U.S., where public LTC provision is less generous: savings data alone need not separately identify precautionary and bequest motives because wealthier households simultaneously save assets for both uses (Dynan et al., 2004). Our study is one of the first attempts to estimate the bequest saving motive in a country where precautionary saving is less important, and thus saving data alone could suffice. From our findings, we conclude that the estimated preference for bequest saving seems consistent across countries and estimation strategies.

For policy design, we can conclude that higher taxes on bequests could be a way to introduce more actuarial fairness into the system of old-age social insurances. Along the spectrum of other possible policy interventions, having social security benefits tailored to the career length is another way to increase actuarial fairness because the working life of shorter-living (lower) lifetime income quintiles usually starts at younger ages.

While our findings opt for those kinds of policies, we keep the quantitative importance of these alternative policy proposals and their interaction with heterogenous mortality *and* LTC use for future work (see Bagchi (2019) for an example involving differential mortality only). In that case, there should be paid more attention to the working age stage than we do, because reforming the system makes precautionary saving and tax contributions increasingly relevant. Also, when assessing different retirement policies, we might have to extend the life cycle model with endogenous health and labor supply decisions and health-dependent utility (cf. French, 2005; Finkelstein et al., 2009). In our specification, we pursue parsimony and thus treat health as exogenous and do not model labor supply explicitly. That does not mean we entirely ignore these variables; the income risk reflects them and, therefore lifetime income status at age 65.

Besides, future research could estimate the effects of changing the LTC insurance system besides the co-payments. In doing so, we can assess whether our studied welfare gains are larger in a system with exclusively private LTC insurance or a mix of public and private LTC insurance. Lastly, future research can include other behavioral frictions that likely matter for evaluating of the impact of bequests. The typical frictions that one can think of are taxes on bequests, taxes on capital gains, and trade-offs between leaving bequests and inter-vivos transfers, which are not part of our model.

Chapter 5

Estimating Left-truncated Shared Frailty Models

This chapter is joint work with Gerard J. van den Berg from the University of Groningen, University Medical Center Groningen, IFAU Uppsala, and IZA. The authors thank Rob Alessie for helpful comments when developing the programming software. The syntax is available upon request.

5.1 Introduction

Hazard rate estimation is widely applied in research on event times in various social, behavioral, and medical sciences disciplines.[1] Hazard rate models apply when subjects experience single or repeated events, e.g., the time between subsequent COVID-19 infections, or where events involve a cluster, e.g., the times until the other household members are infected. Correlations across hazard rates due to common characteristics, e.g. being vaccinated, should be accounted for when doing statistical inference (van den Berg, 2001). However, part of the common characteristics is often unobserved, and ignoring the unobserved heterogeneity leads to biased inference. The way to account for the bias is to augment the hazard rate specification with a random effect ν, called frailty, that is common across grouped data of size J: a shared frailty model.[2]

Due to a sampled subpopulation, hazard rate models can suffer from sample selection problems akin to truncated regression models. While inflow samples comprise a random draw of the population at inflow into the state of interest, selective left-truncated samples arise if subjects are only drawn when sufficiently long in the state of interest, e.g., in population data, where exits before a particular date are typically not observed. Left-truncated subjects have favorable characteristics (low ν) for a high event time. Several empirical studies ignored the dynamic selection due to left truncation, implying underestimated time effects and covariate impacts suffering from attenuation bias, as shown by van den Berg and Drepper (2016) for a case with two shared spells ($J = 2$) and time-invariant covariates. In turn, accounting for time-varying covariates requires observing the entire covariate history or making identifying assumptions, especially on the unobserved part of the covariate history (Lancaster, 1990).

In this chapter, we build upon van den Berg and Drepper (2016), and analyze the bias if dynamic selection due to left truncation is ignored and frailty is shared among more than two spells ($J > 2$). First, we derive the theoretical conditional likelihood

[1] See Wang et al., 2019, for an overview of applications. Examples include the Bass diffusion model (marketing, Gopinath et al., 2014), COVID-19 mortality (epidemiology, Baqui et al., 2020), and labor market transitions (economics, Chodorow-Reich and Coglianese, 2021), among others.

[2] See Hougaard (2000) for an extensive description of shared frailty models.

specifications $l_{\mathtt{A}}$, accounting for the dynamic selection implying unbiased estimates, and $l_{\mathtt{B}}$, ignoring the dynamic selection implying possibly biased estimates. While van den Berg and Drepper (2016) show that $l_{\mathtt{B}}$ provides unbiased results if the dynamic selection is *absent*, we derive a sufficient condition for $l_{\mathtt{B}}$ to also imply unbiased results if the dynamic selection is *present* but ignored when estimating. Subsequently, we quantify the bias in parameter estimates with a Monte Carlo experiment employing several data-generating processes, thereby tuning the number of shared spells and truncation rates. In addition to left truncation, our likelihoods allow for time-varying covariates, right censoring, an arbitrary number of shared spells, and group-specific time effects. We implement both specifications in publicly available `STATA` software packages.[3]

We find that the dynamic selection can be ignored if the entry times in the observed sample are zero. While this nests the case of an inflow sample, this setting also occurs if only subjects with an entry threshold of zero are sampled. At this point, it is worth highlighting that we look at a natural extension of the left truncation of single spells, which restricts a subject only to be sampled if all its spells exceed a threshold. Ever increasing the number of shared spells within a subject approaches the case of observed entry times of zero because, ceteris paribus, more entry thresholds have to be met, which is more likely if these thresholds are lower. This novel sufficient condition is good news for researchers who ignore(d) dynamic selection due to left truncation in their estimation.

Our Monte Carlo experiment provides supportive evidence by revealing an attenuation bias to a time effect and covariate impact of 54% and 47% if there is one spell per subject and a truncation rate of 0.5, declining to 2% and 3% if there are five spells per subject. The biases are also smaller at lower truncation rates. At the same time, we find that the frailty variance is underestimated by 59% for one spell per subject changing into an overestimation of 27% for five spells per subject. The surprising bias at a higher number spells is likely coming from the thresholds not being exactly zero while the observed frailty distribution being highly selective.

[3] The programs are available upon request.

In addition to these methodological results contributing to the presented literature, our user-written `STATA` commands complement and contribute to the set of `STATA`- and `R`-commands available to estimate shared frailty models. The most frequently used package to estimate frailty models in `STATA` is `streg` (Gutierrez, 2002) with option `shared()`, but in the presence of dynamic selection due to left truncation, this requires frailty clusters to have a size of at most one (van den Berg and Drepper, 2016), as is explained by `STATA` if the routine is called upon in such cases. As an extension, van den Berg and Drepper (2016) offer `STATA`-code for clusters up to size two but not beyond, and they do not allow for time-varying covariates. In `R`, the packages `parfm`, `frailtyEM`, and `frailtypack` (Munda et al., 2012; Balan and Putter, 2019; Rondeau et al., 2022) extend frailty to be shared within a cluster of arbitrary size. However, none of these packages can properly deal with time-varying covariates, left truncation and group-specific time effects simultaneously, which we do allow for in our code.[4]

In the remainder of this introduction section, we discuss some of the empirical challenges that estimating a shared frailty model is exposed to. Then, the rest of the article unfolds as follows. Section 5.2 derives the analytical likelihoods. In particular, Section 5.2.4 provides the likelihood that accounts for dynamic selection due to left truncation. Section 5.2.5 provides the Monte Carlo experiment. Section 5.3 discusses and concludes.

5.1.1 Mixed Proportional Hazard Model with Shared Frailty

The Mixed Proportional Hazard model (MPH) is a commonly adopted approach to model event time T with frailty, introduced in economics and demography by Lancaster (1979), Nickell (1979) and Vaupel et al. (1979). MPHs fully characterize the distribution of T. A MPH is as follows:

$$\lambda(t|\boldsymbol{x}(t)) := \lim_{dt \to 0} \frac{\mathbb{P}\left(T \in [t, t+dt) \mid T \geq t, \boldsymbol{x}(t)\right)}{dt} = \nu \cdot \lambda_0(t) \cdot \exp(\boldsymbol{x}(t)'\boldsymbol{\beta}), \text{ with } \nu \sim G(\nu),$$

[4] See Balan and Putter (2019) and Gorfine and Zucker (2023) for recent overviews of `R` packages that have the option to estimate shared frailty models.

where λ is the number of transitions per time unit at time t. The proportional rate splits up into a random effect ν, a time effect $\lambda_0(t)$, and a covariate effect $\exp(\boldsymbol{x}(t)'\boldsymbol{\beta})$. The random effect ν is the unobserved frailty term that can be shared within a group (cluster) or across spells of a subject and follows some parametric distribution G. $\boldsymbol{x}(t)$ denotes a time-dependent vector with impact $\boldsymbol{\beta}$. For inference, the parameter vector $\boldsymbol{\beta}$ and, to a lesser extent, baseline hazard function λ_0 are most important.[5] Without further reference, we assume interest in the shape of λ_0, implying that the functional form λ_0 is also assumed and parameterized. The distribution of T is fully parameterized by λ_0, G and $\boldsymbol{\beta}$, allowing us to estimate the unknown model parameters with full information log-likelihood.[6] The estimation is exposed to several empirical challenges that we will discuss now.

Time-varying covariates $\boldsymbol{x}(t)$ provide a challenge that perhaps received less attention than the other challenges addressed by our estimation procedure. Eventually, the distribution of event times depends on the aggregated risk up to time t: $\int_0^t \lambda(s \mid \boldsymbol{x}(s))ds$ (Lancaster, 1990). This sum consists of the risk that was experienced at each time period, and explicitly depends on the different covariate paths of $\{\boldsymbol{x}(s)\}_0^t$ that subjects went through. Contrary to hazard models without frailty, this aggregated risk is an explicit part of the estimation procedure for shared frailty models.

Inference on hazard rates greatly benefits from time-varying covariates when meeting some conditions. Notably, the time path $\boldsymbol{x}(t)$ has to be exogenous to t and ν. Also, the hazard may not depend on the future path of $\boldsymbol{x}(t)$ if that path is unobserved for the researcher (Kalbfleisch and Prentice, 1980; van den Berg, 2001). Lastly, to have no initial conditions problem, we have to assume that we observe the full covariate path $\{\boldsymbol{x}(s)\}_0^t$ instead of only the partial path $\{\boldsymbol{x}(s)\}_{s>0}^t$ (Lancaster, 1990). If these assumptions are satisfied, then time-varying covariates aid identification of λ_0 and $\boldsymbol{\beta}$ (Honoré, 1993). For now, we assume that the data meet these assumptions.

[5] Occasionally frailty is of prime interest, such as studies on extreme-age plateaus in human mortality (Barbi et al., 2018).

[6] If the baseline hazard λ_0 is not of interest, then λ_0 can be unspecified and limited information likelihood methods can be applied (cf. Cox (1972); Ridder and Tunali (1999)).

Dynamic selection is addressed by including random effect ν. Dynamic selection implies that as time ticks, the subpopulation of survivors changes in terms of observed and unobserved characteristics, i.e. subjects with low ν remain longer in the state of interest. If ν is omitted from the specification, then the estimate of $\lambda_0(t)$ reflects that low ν remain longer. In general, the estimated slope of λ_0 is downward biased. Moreover, estimates of covariate impacts suffer from attenuation bias (Lancaster, 1990). We adopt the often-used gamma distribution $\nu \sim \Gamma\left(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}\right)$, with $\mathbb{E}(\nu) = 1$ and $\text{Var}(\nu) = \sigma^2$, because of its good approximation for general G at high durations and mathematical tractability in the likelihood function (Abbring and van den Berg, 2007).

Left truncation of events can exacerbate the survivorship bias of parameter estimates due to dynamic selection (van den Berg and Drepper, 2016). Left-truncated subjects are namely sampled with characteristics that favor high durations, including low ν, because they have durations of at least t_0 for some $t_0 > 0$. To correct for the bias due to left-truncated subjects, estimation procedures have to condition on survival up to t_0 rather than assume that the subjects at t_0 are a random draw from the underlying distribution of spell lengths.

Right-censored event times are a last source of bias to hazard rate estimation that we allow for. Right censoring occurs if the event time is cut off at some date before the actual event takes place, implying an unrepresentative sample of too short event times. To this end, hazard rate estimation procedures adjust the likelihood contribution of censored subjects to the likelihood of not having experienced the event before the censored time; instead of that this is the date at which subjects experience the event.

5.2 Likelihood Function for Shared Frailty Models with Left-truncated Data

5.2.1 Notation

We draw subjects from a population. Each subject, in turn, provides a sample of J spells that are spent in some state of interest. The spell length is measured as the duration since the start time $t = 0$. However, a subject is only observed and included in the data if each of the J spells satisfies the following constraint: the random length T_j of spell j exceeds a certain amount of time $t_{0j} \geq 0$, with $j = 1, \ldots, J$. Apart from this observability constraint, we take the spells to be drawn from the inflow into the state of interest.[7] Let t_j be the observed end time, with $t_j > t_{0j}$. We introduce binary variable d_j that reflects whether we observe the true spell end ($d_j = 1$) or a right-censored spell ($d_j = 0$) at t_j. Right-censored spells have an observed end time t_j that lies before the actual end time, for example, because the study ends before the event occurs.

Following Lancaster (1990) we use the following notation for covariates:

$$\begin{aligned}
\boldsymbol{x}_j(t) &:= \{\text{covariate values of spell } j \text{ at time } t\};\\
\boldsymbol{X}_j(t) &:= \{\text{covariate path of spell } j \text{ between 0 and } t\};\\
\boldsymbol{X}_j &:= \boldsymbol{X}_j(\infty),
\end{aligned}$$

and if we consider all spells together, we have:

$$\begin{aligned}
\boldsymbol{x}(\boldsymbol{t}) &:= \{\text{covariate values of all spells at time } \boldsymbol{t}\};\\
\boldsymbol{X}(\boldsymbol{t}) &:= \{\text{covariate path of all spells between } \boldsymbol{0} \text{ and } \boldsymbol{t}\};\\
\boldsymbol{X} &:= \boldsymbol{X}(\infty),
\end{aligned}$$

where $\boldsymbol{t}$ is the collection of all observed times for the spells.

We allow the time effect λ_0 in our specifications to depend on time-invariant covariates. To this end, we introduce the covariate vector $\widetilde{\boldsymbol{x}}_j$, which can be the same as $\boldsymbol{x}_j(t)$. The two vectors are already different if $\boldsymbol{x}_j(t)$ contains at least one time-dependent covariate.

[7] This is the natural generalization of the concept of left truncation to the setting with multiple spells per subject. In practice, other observation schemes may apply. For example, one may sample subjects without restriction and subsequently only observe those $K \leq J$ spells for which the length t satisfies $t \geq t_{0j}$. In that case, if the researcher knows J and $\boldsymbol{t}_0$, it follows that the $J - K$ spells not meeting this constraint are known to be left-censored at t_{0j}, which is informative on the model parameters of interest. Alternatively, one may sample subjects without restriction and continue sampling spells satisfying $t \geq t_{0j}$ until J of those are observed. In that case, spell lengths are left-truncated, but the distribution of unobservables is as in the population of subjects. This constellation does not generalize the concept of left truncation from a single-spell setting.

5.2.2 Hazard Rate and Survival Probabilities

Integrated Hazard Rates

We are interested in the distribution function of event end time T, say F. Note that $F(t) = \mathbb{P}(T < t)$ with $F(0) = 0$. Inherently, $S(t) = \mathbb{P}(T \geq t) = 1 - F(t)$ is the survival function. The survival function closely relates to hazard rate $\lambda(t)$, which is the number of transitions per time unit:

$$\begin{aligned}\lambda(t) &:= \lim_{dt \to 0} \frac{\#\text{Realized transitions between } t \text{ and } t + dt \mid T \geq t}{dt} \\ &= \lim_{dt \to 0} \frac{\frac{\mathbb{P}(t \leq T < t+dt)}{dt}}{\mathbb{P}(T \geq t)} = \frac{f(t)}{S(t)}.\end{aligned}$$

The hazard rate λ intimately links back to the survival distribution S:

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} = -\frac{\frac{\partial 1 - F(t)}{\partial t}}{1 - F(t)} = -\frac{\partial \ln(1 - F(t))}{\partial t} \rightarrow$$
$$-\int_0^t \lambda(\tau) d\tau = \ln(1 - F(t)) - \ln(1 - F(0)) = \ln(S(t)).$$

So:

$$\mathbb{P}(T \geq t) = S(t) = \exp(-m(t)), \quad \text{with: } m(t) = \int_0^t \lambda(\tau) d\tau,$$

implying that the survival probability solely depends on integrated hazard m, i.e. the total hazard of having experienced the event before time t. Furthermore, a survival probability S is fully characterized by the chosen hazard rate specification λ.

The Weibull and Gompertz are two common choices for the shape of hazard rate λ. These specifications have duration dependence parameter γ that we will estimate. In Table 5.1 we provide the hazard rate, integrated hazard, and survival distribution in which these particular choices result. In this chapter, we restrict ourselves to these two choices.

For inference, we must specify the hazard rate in the presence of covariates. To do so,

Table 5.1: Characteristics of the Weibull and Gompertz Distribution (cf. Bender et al., 2005)

	Distribution	
	Weibull	Gompertz
Hazard rate $\lambda(t)$	$\gamma \cdot t^{\gamma-1}$	$\exp(\gamma t)$
Integrated hazard $\int_0^t \lambda(\tau)d\tau$	t^γ	$\frac{1}{\gamma} \cdot (\exp(\gamma t) - 1)$
Inverse integrated hazard	$t^{\frac{1}{\gamma}}$	$\frac{1}{\gamma} \cdot \ln(\gamma t + 1)$
Survival function $S = \exp\left(-\int_0^t \lambda(\tau)d\tau\right)$	$\exp(-t^\gamma)$	$\exp\left(\frac{1}{\gamma} \cdot (1 - \exp(\gamma t))\right)$
γ	$(0, \infty)$	$(-\infty, \infty)\setminus\{0\}$

we follow Lancaster (1979) and use a mixed proportional hazard rate (MPH) specification. Effectively, the hazard rate λ now depends on time t_j, observed characteristics $\boldsymbol{x}_j(t_j)$ and $\widetilde{\boldsymbol{x}}_j$, and the unobserved term ν. A MPH model is as follows:

$$\lambda(t_j, \boldsymbol{x}_j(t_j), \widetilde{\boldsymbol{x}}_j, \nu) = \nu \cdot \phi(\boldsymbol{x}_j(t_j)) \cdot \lambda_0(t_j, \widetilde{\boldsymbol{x}}_j).$$

This specification features a part with a time-invariant impact, namely the proportional hazard ϕ; a part containing the duration dependence or time effect, the baseline hazard λ_0; and a part containing random noise, the frailty term ν. For the proceeding, we assume functional forms of λ_0 to restrict to the shapes denoted in Table 5.1.

The derivations of the integrated hazard m and survival probability S are non-standard because the integration should take into account that $\boldsymbol{x}_j(t)$ also varies over time. Specifically, we have to take the entire covariate path $\boldsymbol{X}_j$ into account when integrating. Integrated hazard will thus be a function of $\boldsymbol{X}_j$:

$$\begin{aligned} m(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j, \nu) &= \int_0^{t_j} \lambda(\tau, \boldsymbol{x}_j(\tau), \widetilde{\boldsymbol{x}}_j, \nu)d\tau \\ &= \nu \cdot \int_0^{t_j} \phi(\boldsymbol{x}_j(\tau)) \cdot \lambda_0(\tau, \widetilde{\boldsymbol{x}}_j)d\tau \\ &= \nu \cdot M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j), \end{aligned} \tag{5.1}$$

where integrated hazard splits into an unobserved part ν and observed part M.

Conditional Survival Probabilities

The conditional survival probability of spell j becomes:

$$\begin{aligned}\mathbb{P}(T_j \geq t_j \mid \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j, \nu) = S(t_j \mid \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j, \nu) &= \exp(-m(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j, \nu)) \\ &= \exp(-\nu \cdot M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)).\end{aligned}$$

Most important for further derivations is that this spell-specific probability contains a distinct parts linked to unobserved heterogeneity, ν, and observed heterogeneity, $M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)$.

Because frailty is shared across spells, we base our likelihood estimation on the joint survival probability. To construct the joint conditional survival probability for $\boldsymbol{T} = (T_1, .., T_J)$, we assume that their realisations $\boldsymbol{t} = (t_1, .., t_J)$ are independently distributed conditional upon frailty. Given the earlier structure of the spell-specific conditional survival probability, the joint survival probability depends on the collection of covariate paths $\boldsymbol{X}$, covariate vectors $\widetilde{\boldsymbol{x}}$, and frailty term ν. Consequently, the joint survival function $S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \nu)$ is:

$$\begin{aligned}\mathbb{P}(\boldsymbol{T} \geq \boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \nu) = S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \nu) = \prod_{j=1}^{J} S(t_j \mid \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j, \nu)) &= \prod_{j=1}^{J} \exp(-\nu \cdot M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)) \\ &= \exp\left(-\nu \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)\right), \end{aligned} \tag{5.2}$$

which is the product of independent survival probabilities.

Unconditional Survival Probability

The conditional survival probabilities cannot be readily applied in likelihood estimation because they depend on the unobserved term ν. To obtain a completely defined likelihood, we first have to derive the unconditional survival probability with an assumed distribution G for ν. For this derivation, we follow van den Berg and Drepper (2016) (note that they do not consider time-varying covariates).

For ease of derivation, we define the distribution of ν in general terms. Let ν have a distribution that belongs to the exponential family. The probability density function of ν, denoted by $g(\nu)$, then takes the following general form:

$$g(\nu) = \nu^{\zeta} \cdot \exp(-\eta \cdot \nu) \cdot \widetilde{m}(\nu) \cdot \widetilde{\phi}(\zeta, \eta)^{-1}, \quad \text{with: } \widetilde{\phi} > 0.$$

This family of frailty distributions has a unique Laplace transform, which is sufficient for identifying the duration dependence λ_0 and the time-invariant effect ϕ (Elbers and Ridder, 1982; Honoré, 1993). The Laplace transform can be derived as follows:

$$\int_0^\infty dG(\nu) = 1 \rightarrow \int_0^\infty g(\nu)d\nu = 1 \rightarrow \int_0^\infty \nu^{\zeta} \cdot \exp(-\nu \cdot \eta) \cdot \widetilde{m}(\nu) \cdot \widetilde{\phi}(\zeta, \eta)^{-1} d\nu = 1$$
$$\rightarrow \widetilde{\phi}(\zeta, \eta) = \int_0^\infty \nu^{\zeta} \cdot \exp(-\nu \cdot \eta) \cdot \widetilde{m}(\nu) d\nu.$$

Using the Laplace transform, the unconditional survival probability $S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}})$ is:

$$\begin{aligned} S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}}) &= \int_0^\infty S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \nu) dG(\nu) \\ &= \int_0^\infty \exp(-\nu \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)) \cdot g(\nu) d\nu \\ &= \int_0^\infty \exp(-\nu \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)) \cdot \nu^{\zeta} \cdot \exp(-\nu \cdot \eta) \cdot \widetilde{m}(\nu) \cdot \widetilde{\phi}(\zeta, \eta)^{-1} d\nu \\ &= \frac{\widetilde{\phi}(\zeta, \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + \eta)}{\widetilde{\phi}(\zeta, \eta)}, \end{aligned}$$

where we applied the Laplace transform in the last step.

We further assume that ν is gamma-distributed. For this to hold, take $\widetilde{\phi}(\zeta, \eta) = \eta^{-(\zeta+1)} \cdot \Gamma(\zeta + 1)$ and $\widetilde{m}(\nu) = 1$, with $\zeta = \frac{1}{\sigma^2} - 1$ and $\eta = \frac{1}{\sigma^2}$. The assumed gamma distribution is useful because of its mathematical tractability and good approximation for any conditional frailty distribution at long durations (Abbring and van den Berg, 2007). With these assumptions we have $\mathbb{E}(\nu) = \frac{1}{\frac{1}{\sigma^2}} \cdot \frac{1}{\sigma^2} = 1$ and $\mathrm{Var}(\nu) = \sigma^2 \geq 0$. Using the gamma distribution, we obtain the following joint unconditional survival probability:

$$
\begin{aligned}
S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}}) &= \frac{\left(\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + \eta\right)^{-(\zeta+1)} \cdot \Gamma(\zeta+1)}{\eta^{-(\zeta+1)} \cdot \Gamma(\zeta+1)} \\
&= \left(\frac{\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + \eta}{\eta}\right)^{-(\zeta+1)} \\
&= \left(\frac{\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + \frac{1}{\sigma^2}}{\frac{1}{\sigma^2}}\right)^{-(\frac{1}{\sigma^2}-1+1)} \\
&= \left(\sigma^2 \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + 1\right)^{-\frac{1}{\sigma^2}} = \mathcal{L}\left(\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)\right),
\end{aligned}
$$

where $\mathcal{L}$ maps integrated hazard M into the joint unconditional survival probability.

This survival function nests a survival function without frailty, i.e. $\sigma^2 = 0$. Note:

$$
\begin{aligned}
\lim_{\sigma^2 \to 0} \left(\sigma^2 \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + 1\right)^{-\frac{1}{\sigma^2}} &= \left(\lim_{\sigma^2 \to 0} \left(\sigma^2 \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + 1\right)^{\frac{1}{\sigma^2}}\right)^{-1} \\
&= \exp(-M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)),
\end{aligned}
$$

where the second step is the definition of an exponential number (see Simon and Blume, 1994). We find the same result using (5.2) if we impose $\nu = \mathbb{E}(\nu) = 1$.

5.2.3 Analytical Log-Likelihood

We base the log-likelihood on survival probability $S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}})$. For the construction of the likelihood on actual observations, we have information on exact survival times t_j with event indicator $D_j = 1$, and information on censored survival times t_j with event indicator $D_j = 0$. Intuitively, censored times enter the likelihood function as survival probability $S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}})$, because they made it till that particular time but also beyond. Uncensored times enter the likelihood by taking the partial derivative of $1 - S(\boldsymbol{t} \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}})$ w.r.t. t_j, i.e. the density evaluated at t_j. Let $\overline{D} = \sum_{j=1}^{J} D_j$ be the number uncensored survival times. Let $\mathcal{L}^{(\overline{D})}$ denote the $\overline{D}^{th}$−derivative of $\mathcal{L}$. Then, the term that forms the likelihood is obtained from the chain rule on differentiation (cf. 7.6 in Hougaard, 2000):

$$(-1)^{\overline{D}} \cdot \prod_{j=1}^{J} \left\{ \frac{\partial M(t, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)}{\partial t}\bigg|_{t=t_j} \right\}^{D_j} \cdot \mathcal{L}^{(\overline{D})} \left(\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) \right) =$$

$$(-1)^{\overline{D}} \cdot \prod_{j=1}^{J} (\phi(\boldsymbol{x}_j(t_j)) \cdot \lambda_0(t_j, \widetilde{\boldsymbol{x}}_j))^{D_j} \cdot \mathcal{L}^{(\overline{D})} \left(\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) \right)$$

where censored times only contribute to the second part of the multiplication. We used expression (5.1) to arrive at the right-hand side. Note that the likelihood contribution is not spell-specific, but defined at the subject level at which frailty is shared. As a consequence, we take higher-order partial derivatives if multiple event times are uncensored.

The likelihood function does account for right censoring, but not for left truncation yet. We consequently have to condition the likelihood on survival up to time t_{0j}. We use the joint survival probability $S(\boldsymbol{t}_0 \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}})$ until the left truncation times $\boldsymbol{t}_0$ for this. Then, the log-likelihood contribution for a subject consisting of J spells with shared frailty:

$$l(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}) = \ln \left\{ \frac{(-1)^{\overline{D}} \cdot \prod_{j=1}^{J} (\phi(\boldsymbol{x}_j(t_j)) \cdot \lambda_0(t_j, \widetilde{\boldsymbol{x}}_j))^{D_j} \cdot \mathcal{L}^{(\overline{D})} \left(\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) \right)}{S(\boldsymbol{t}_0 \mid \boldsymbol{X}, \widetilde{\boldsymbol{x}})} \right\}$$

$$= \ln \left\{ \frac{(-1)^{\overline{D}} \cdot \prod_{j=1}^{J} (\phi(\boldsymbol{x}_j(t_j)) \cdot \lambda_0(t_j, \widetilde{\boldsymbol{x}}_j))^{D_j} \cdot \mathcal{L}^{(\overline{D})} \left(\sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) \right)}{\mathcal{L} \left(\sum_{j=1}^{J} M(t_{0j}, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) \right)} \right\}$$

where the function $\mathcal{L}^{(\overline{D})}$ is given by:

$$\mathcal{L}^{(\overline{D})}(y) = \frac{\partial \mathcal{L}(y)}{\partial y^{\overline{D}}} = (-1)^{\overline{D}} \cdot (\sigma^2 \cdot y + 1)^{-\frac{1}{\sigma^2} - \overline{D}} \cdot \prod_{q=0}^{(\overline{D}-1)_+} (q\sigma^2 + 1), \text{ with } y = \sum_{j=1}^{J} M(t_{0j}, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j)$$

and whose complete derivation we provide in Appendix E.1.

Using the expression for $\mathcal{L}^{(\overline{D})}$, the log-likelihood contribution becomes:

Analytical Log-Likelihood Contribution for a Subject with J Spells and Shared Frailty ν:

$$l(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}) = \sum_{q=0}^{(\overline{D}-1)_+} \ln(q\sigma^2 + 1) + \sum_{j=1}^{J} D_j \cdot \ln(\phi(\boldsymbol{x}_j(t_j))) + \sum_{j=1}^{J} D_j \cdot \ln(\lambda_0(t_j, \widetilde{\boldsymbol{x}}_j))$$
$$-(\frac{1}{\sigma^2} + \overline{D}) \cdot \ln\left(\sigma^2 \cdot \sum_{j=1}^{J} M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + 1\right) + \frac{1}{\sigma^2} \cdot \ln\left(\sigma^2 \cdot \sum_{j=1}^{J} M(t_{0j}, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) + 1\right) \quad (5.3)$$

Notice that identification warrants that the values of the time-varying covariates are observed before t_0, which might for example be the case in retrospective studies. Henceforth, we assume that these values are constant until and including the truncation point t_{0j}. A formal identification analysis of the model with left truncation is beyond the scope of this chapter. However, it is clear that also requires an assumption on the baseline hazard λ_0 as a function of t on $[0, t_{0j})$. In particular, one may assume that λ_0 is constant until a point in time that exceeds the values of t_{0j} for all j for every subject.

5.2.4 Method A: Accounting for Dynamic Selection due to Left Truncation

In this part, we convert the analytical log-likelihood (5.3) into its counterpart with parameterized functions. We aim to estimate the unknown parameters. For estimation, we have developed a user-written command in `STATA`, which we discuss below and refer as 'Method A'. This method accommodates dynamic selection due to left truncation. In Section 5.2.5, we contrast Method A to its counterpart ignoring the dynamic selection, i.e. 'Method B'.

Model Parameterization

We parameterize covariate impacts ϕ and duration dependence λ_0 to infer them. To this end, we assume ϕ has an unknown parameter vector $\boldsymbol{\beta}$ and λ_0 and an unknown parameter vector $\boldsymbol{\gamma}$ that we will estimate. In addition, we have to assume a functional form of λ_0. Table 5.1 provided the different choices that we can make for λ_0.

Mixed proportional hazard assumes a single-index and log-linear structure for ϕ:

$$\phi(\boldsymbol{x}_j(t), \boldsymbol{\beta}) = \exp(\theta_{j1}(t)), \text{ with: } \theta_{j1}(t) = \boldsymbol{x}_j(t)'\boldsymbol{\beta},$$

where the covariate impacts are exponentiated to yield positive hazard at any time t. Due to this restriction, the impact of $\boldsymbol{x}_j(t)$ on hazard ϕ has a relative risk interpretation.

Similarly, we restrict how duration dependence depends on its determinants:

$$\lambda_0(t, \widetilde{\boldsymbol{x}}_j, \boldsymbol{\gamma}) = \lambda_0(t, \theta_{j2}) \text{ with: } \theta_{j2} = \widetilde{\boldsymbol{x}}_j'\boldsymbol{\gamma}.$$

Notice that some normalizations may be required, especially if θ_{j1} and θ_{j2} share covariates.

Lastly, to ensure that the variance is positive when estimating, we parameterize it as:

$$\sigma^2 = \exp(\theta_3).$$

The log-linearity and single-index assumption are not only crucial for the interpretation of the results, they also make the log-likelihood computations tremendously easier. Usually, log-likelihood optimization routines namely require score functions and the Fisher information matrix. This Jacobian and Hessian of the log-likelihood function can be numerically approximated or analytically computed, where numerical approximations are notoriously slow and inefficient if extensive data sets are used. In turn, if these derivatives can be calculated analytically we can save much computational time and have better estimation accuracy. The linear form restrictions allow us to calculate the derivatives in a straightforward way analytically. We can just do scalar differentiation for $\theta_{j1}(t), \theta_{j2}$, and θ_3, and subsequently apply the chain rule of optimization to reach at the Jacobian and Hessian analyzed in $\boldsymbol{\beta}, \boldsymbol{\gamma}$, and σ^2. `STATA` has further optimized the implementation of the chain rule via the built-in routines `mlvecsum`, `mlmatsum` and `mlmatbysum` (Gould et al., 2010), which we therefore also apply.

Time-varying Covariates

Based on the above parameterizations, we may program the log-likelihood contributions. The first three terms in (5.3) are a scalar sum, which we can therefore easily compute. The last two terms are more complex because they involve hazard M, integrated over the entire covariate path. We have to convert this integral into a computable quantity.

For tractability, we split up the integral function M into distinct S_j domains $t \in \{t_j^{(s-1)}, t_j^{(s)}\}$ on which the covariate vector $\boldsymbol{x}_j(t)$ is constant (comparable to piecewise constant specifications for λ_0 as a function of t). In such a domain, the time-invariant impact ϕ leaves the integral and we only integrate over λ_0, which is a simple integral problem:

$$
\begin{aligned}
M(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) &= \int_0^{t_j} \phi(\boldsymbol{x}_j(\tau)) \cdot \lambda_0(\tau, \widetilde{\boldsymbol{x}}_j) d\tau \\
&= \sum_{s=1}^{S_j} \int_{t_j^{(s-1)}}^{t_j^{(s)}} \phi\left(\boldsymbol{x}_j(\tau)\right) \cdot \lambda_0(\tau, \widetilde{\boldsymbol{x}}_j) d\tau \\
&= \sum_{s=1}^{S_j} \int_{t_j^{(s-1)}}^{t_j^{(s)}} \phi\left(\boldsymbol{x}_j(t_j^{(s-1)})\right) \cdot \lambda_0(\tau, \widetilde{\boldsymbol{x}}_j) d\tau \\
&= \sum_{s=1}^{S_j} \phi\left(\boldsymbol{x}_j(t_j^{(s-1)})\right) \cdot \int_{t_j^{(s-1)}}^{t_j^{(s)}} \lambda_0(\tau, \widetilde{\boldsymbol{x}}_j) d\tau \\
&= \sum_{s=1}^{S_j} \phi\left(\boldsymbol{x}_j(t_j^{(s-1)})\right) \cdot \left(\int_0^{t_j^{(s)}} \lambda_0\left(\tau, \widetilde{\boldsymbol{x}}_j\right) d\tau - \int_0^{t_j^{(s-1)}} \lambda_0\left(\tau, \widetilde{\boldsymbol{x}}_j\right) d\tau \right) \\
&= \sum_{s=1}^{S_j} M\left(t_j^{(s)}, \boldsymbol{x}_j(t_j^{(s-1)}), \widetilde{\boldsymbol{x}}_j\right) - M\left(t_j^{(s-1)}, \boldsymbol{x}_j(t_j^{(s-1)}), \widetilde{\boldsymbol{x}}_j\right), \\
&\text{with: } t_j^{(0)} = 0; \;\; t_j^{(1)} = t_{0j}; \;\; t_j^{(S_j)} = t_j, \text{ and:} \\
M(t_{0j}, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j) &= \int_0^{t_{0j}} \phi(\boldsymbol{x}_j(\tau)) \cdot \lambda_0\left(\tau, \widetilde{\boldsymbol{x}}_j\right) d\tau = \int_0^{t^{(1)}} \phi\left(\boldsymbol{x}_j(t^{(0)})\right) \cdot \lambda_0(\tau, \widetilde{\boldsymbol{x}}_j) d\tau \\
&= M\left(t^{(1)}, \boldsymbol{x}_j(t^{(0)}), \widetilde{\boldsymbol{x}}_j\right),
\end{aligned}
$$

since we assume covariates are fixed until the left truncation time t_{0j}.

Programmed Log-Likelihood

Combining the results above, we reach the following log-likelihood function that we program in `STATA` using `MATA`, i.e. Method A:

$$
\begin{aligned}
l_{\mathtt{A}}\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}}\right) = & \sum_{q=0}^{\left(\overline{D}-1\right)_+} \ln\left(q\sigma^2+1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\boldsymbol{x}_j\left(t_j^{\left(S_j-1\right)}\right)\right)\right) \\
& - \left(\frac{1}{\sigma^2}+\overline{D}\right)\cdot \ln\left(\sigma^2 \cdot \sum_{j=1}^{J}\sum_{s=1}^{S_j}\left\{M\left(t_j^{(s)},\boldsymbol{x}_j\left(t_j^{(s-1)}\right),\widetilde{\boldsymbol{x}}_j\right) - M\left(t_j^{(s-1)},\boldsymbol{x}_j\left(t_j^{(s-1)}\right),\widetilde{\boldsymbol{x}}_j\right)\right\}+1\right) \\
& + \frac{1}{\sigma^2}\cdot \ln\left(\sigma^2\cdot\sum_{j=1}^{J} M\left(t^{(1)},\boldsymbol{x}_j\left(t^{(0)}\right),\widetilde{\boldsymbol{x}}_j\right)+1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{\left(S_j\right)},\widetilde{\boldsymbol{x}}_j\right)\right)
\end{aligned}
$$

with: $t_j^{(0)}=0$; $t_j^{(1)}=t_{0j}$; $t_j^{\left(S_j\right)}=t_j$.

If we explicitly consider the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we get :

Programmed Log-Likelihood Contribution for a Subject with J Spells and Shared Frailty:

$$
\begin{aligned}
l_{\mathtt{A}}\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right) = & \sum_{q=0}^{\left(\overline{D}-1\right)_+} \ln\left(q\cdot\exp\left(\theta_3\right)+1\right) \\
& + \sum_{j=1}^{J} D_j\cdot\ln\left(\phi\left(\theta_{j1}\left(t_j^{\left(S_j-1\right)}\right)\right)\right) + \sum_{j=1}^{J} D_j\cdot\ln\left(\lambda_0\left(t_j^{\left(S_j\right)},\theta_{j2}\right)\right) - \left(\frac{1}{\exp\left(\theta_3\right)}+\overline{D}\right)\cdot \\
& \ln\left(\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}\left\{M\left(t_j^{(s)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right) - M\left(t_j^{(s-1)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)\right\}+1\right) \\
& + \frac{1}{\exp\left(\theta_3\right)}\cdot\ln\left(\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J} M\left(t^{(1)},\theta_{j1}\left(t^{(0)}\right),\theta_{j2}\right)+1\right)
\end{aligned}
$$

with: $t_j^{(0)}=0$; $t_j^{(1)}=t_{0j}$; $t_j^{\left(S_j\right)}=t_j$, $\theta_{j1}\left(t\right)=\boldsymbol{x}_j\left(t\right)'\boldsymbol{\beta}$, $\theta_{j2}=\widetilde{\boldsymbol{x}}_j'\boldsymbol{\gamma}$, and $\theta_3=\ln\left(\sigma^2\right)$,

which is the log-likelihood that we programmed in `STATA` and to which we can apply the chain rule. Appendix E.2 provides the score function and Fisher information matrix related to the programmed log-likelihood.

The estimation procedure is suitable to estimate the shared frailty model in the presence of time-varying covariates, right censoring and left truncation. Furthermore, the time effect λ_0 is allowed to depend on model covariates $\widetilde{\boldsymbol{x}}$. The estimation program is available upon request.

5.2.5 Method B: Ignoring Dynamic Selection due to Left Truncation

Method A differs from the `STATA`-build-in option to estimate shared frailty models with left truncation: `streg` with the option `shared()`, which we will refer to as 'Method B'. This is simply because Method B is not intended or advised to be used in such settings. As outlined in van den Berg and Drepper (2016), nevertheless applying this method via input `forceshared` would impose that the observed frailty distribution of sampled subjects is the same for untruncated and left-truncated subjects, i.e. $G(\nu)$. However, to be in our sample, the left-truncated subjects must not have experienced any event before $\boldsymbol{t}_0$. A relatively low frailty ν increases the likelihood of meeting thresholds $\boldsymbol{t}_0$. Thus, left-truncated subjects are sampled from the distribution $G(\nu \mid \boldsymbol{T} \geq \boldsymbol{t}_0)$. Only if the strong assumption of absence of unobserved heterogeneity is made, the distributions $G(\nu \mid \boldsymbol{T} \geq \boldsymbol{t}_0)$ and $G(\nu)$ are identical under this truncation scheme. Otherwise the distribution of ν in the left-truncated sample differs from that in the untruncated sample and we have $\mathbb{E}(\nu \mid \boldsymbol{T} \geq \boldsymbol{t}_0) < \mathbb{E}(\nu)$. This dependence is not taken into account if Method B is incorrectly applied, and consequently duration dependence –a time effect– is underestimated because it compensates for the omitted dynamic selection in the model. Also, the magnitude of covariate impacts attenuates towards zero (van den Berg and Drepper, 2016). The size of the bias increases with more severe left truncation.

To further elucidate the difference between the applicability of the two methods, consider a simple example of a subject with multiple spells but no covariates. Assume that all observed end times are right-censored, so conditional survival probabilities enter the log-likelihood. The spells are left-truncated and meet truncation thresholds $\boldsymbol{t}_0$. Method B takes the log-likelihood contribution of the subject as (Gutierrez, 2002):

$$l_{\mathtt{B}}(\boldsymbol{t}, \boldsymbol{t}_0) = \int_0^\infty \mathbb{P}(\boldsymbol{T} \geq \boldsymbol{t} \mid \nu, \boldsymbol{T} \geq \boldsymbol{t}_0)\; dG(\nu)$$

where the integral over unconditional frailty distribution G implies that the frailty distribution is assumed to be independent from truncation times $\boldsymbol{t}_0$. However, in general dynamic selection takes place and only subjects with favorable low ν reach truncated start times $\boldsymbol{t}_0$. van den Berg and Drepper (2016) show that our log-likelihood (Method A) is correct for the chosen truncation scheme and integrates over the conditional frailty distribution instead:

$$l_{\mathtt{A}}(\boldsymbol{t}, \boldsymbol{t}_0) = \int_0^\infty \mathbb{P}(\boldsymbol{T} \geq \boldsymbol{t} \mid \nu, \boldsymbol{T} \geq \boldsymbol{t}_0) \, dG(\nu \mid \boldsymbol{T} \geq \boldsymbol{t}_0)$$

Using Method B effectively means that conditioning correction in dG is left out, i.e. the last term in (5.3). Method B only deals with left truncation by adapting the integrated hazard on observed end times in (5.3) (the sum starts from $s = 2$ instead of from $s = 1$):

$$\begin{aligned} l_{\mathtt{B}}(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = & \sum_{q=0}^{(\overline{D}-1)_+} \ln\left(q \cdot \exp(\theta_3) + 1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\theta_{j1}\left(t_j^{(S_j-1)}\right)\right)\right) \\ & + \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)\right) - \left(\frac{1}{\exp(\theta_3)} + \overline{D}\right) \cdot \\ & \ln\left(\exp(\theta_3) \cdot \sum_{j=1}^{J} \sum_{s=2}^{S_j} \left\{ M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\right\} + 1\right), \\ & \text{with: } t_j^{(0)} = 0; \;\; t_j^{(1)} = t_{0j}; \;\; t_j^{(S_j)} = t_j, \;\; \theta_{j1}(t) = \boldsymbol{x}_j(t)' \boldsymbol{\beta}, \;\; \theta_{j2} = \gamma, \;\; \text{and} \;\; \theta_3 = \ln\left(\sigma^2\right). \end{aligned}$$

In addition to coping with dynamic selection due to left truncation, note that our Method A is more flexible by allowing observed covariates $\widetilde{\boldsymbol{x}}$ and duration dependence $\boldsymbol{\gamma}$ to interact, i.e. $\theta_{j2} = \widetilde{\boldsymbol{x}}_j' \boldsymbol{\gamma}$, whereas Method B does not: $\theta_{j2} = \gamma$. For Method B, we have an alternative user-written version allowing $\theta_{j2} = \widetilde{\boldsymbol{x}}_j' \boldsymbol{\gamma}$.

$l_{\mathtt{A}}$ and $l_{\mathtt{B}}$ are asymptotically the same (i.e. when sufficiently many subjects are sampled) if the starting times $\boldsymbol{t}_0$ approach zero in the *observed* population. Then, we do not need the corrective probability in our log-likelihood (5.3) for survival up to $\boldsymbol{t}_0$ because this approaches unit value for each sampled subject (see Appendix E.3 for the proof). This not only nests the special case of no left truncation $t_j^{(0)} = 0$, but also left truncation schemes imposing that only spells with very low $t_j^{(0)} \to 0$ are sampled, e.g. when there are many spells and all must meet the threshold. The subject's sampling likelihood is namely higher if the drawn $t_j^{(0)}$'s are smaller. Lastly, $l_{\mathtt{A}}$ and $l_{\mathtt{B}}$ are asymptotically the same if there is no frailty, i.e. when $\sigma^2 = 0$.

Closely following van den Berg and Drepper (2016), we perform a Monte Carlo experiment to compare the estimation outcomes of using Method A to using Method B in settings with up to five spells of shared frailty and different degrees of left truncation. We proceed in three steps. First, we generate data according to a mixed proportional hazard rate with shared frailty and apply the truncation scheme. Next, to have a better understanding of its implications for estimation outcomes, we discuss the simulated observed frailty distribution $G(\nu \mid \boldsymbol{T} \geq \boldsymbol{t}_0)$ for different truncation rates and number of shared spells. Lastly, we estimate hazard rate parameters on the simulated data and compare outcomes across the estimation methods.

We consider a data generating process of random duration T_{ij} according to a mixed proportional hazard rate with shared frailty ν:

$$\lambda(t_{ij} \mid \gamma, \beta, \nu_i, x_{ij}) = \nu_i \cdot \lambda_0(t_{ij} \mid \gamma) \cdot \exp(\beta x_{ij}), \text{ with: } \nu_i \sim \Gamma_d\left(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}\right),$$

where $i \in \{1, .., N\}$ indicates the subject and $j \in \{1, .., J\}$ the spell number of the subject. For subject i we draw J random event times; we denote the random event times and their realisations with the vectors $\boldsymbol{T}_i = (T_{i1}, .., T_{iJ})$ and $\boldsymbol{t}_i = (t_{i1}, .., t_{iJ})$. The frailty term ν_i is shared across all J spells of the subject here. The parameters β, γ, σ^2 are of interest, and we aim provide estimates $\widehat{\beta}$, $\widehat{\gamma}$, and $\widehat{\sigma^2}$. The data is generated in three steps.

Step 1: Set model parameters and draw characteristics. For every case we assume $\beta = \gamma = \sigma^2 = 1$. Subjects draw characteristic $x_{ij} \sim \mathcal{N}(0, 1)$, which is independent across their spells. Also, subjects draw frailty $\nu_i \sim \Gamma_d(1, 1)$ which is shared across their spells. To see the relation between the number of shared spells within a subject and the bias, we one-by-one consider the case where frailty is shared across $J = 1, 2, 3, 4$ and 5 spells.

Furthermore, we separately analyze the cases where baseline hazard λ_0 is specified as Weibull and Gompertz hazard.

Step 2: Simulate end times t_{ij} based on the parameter constellations of Step 1. To this end, we draw value t_{ij} from the conditional survival distribution $S(t_{ij} \mid \nu_i, x_{ij})$. For this, consider the fundamental relationship:

$$\begin{aligned}
\mathbb{P}(T_{ij} \geq t_{ij} \mid \nu_i, x_{ij}) = S(t_{ij} \mid \nu_i, x_{ij}) &= \exp\left(-\int_0^{t_{ij}} \lambda(\tau \mid \gamma, \beta, \nu_i, x_{ij}) d\tau\right) \\
&= \exp\left(-\nu_i \cdot \exp(\beta x_{ij}) \cdot \int_0^{t_{ij}} \lambda_0(\tau \mid \gamma) d\tau\right) \\
&\rightarrow -\frac{1}{\nu_i \cdot \exp(\beta x_{ij})} \cdot \ln(S_{ij}) = \int_0^{t_{ij}} \lambda_0(\tau \mid \gamma)) d\tau.
\end{aligned}$$

We can simulate t_{ij} by drawing $S_{ij} \sim \mathcal{U}(0,1)$, and subsequently solving the last equation for t_{ij}. Solving the equation requires the inverse of $\int_0^{t_{ij}} \lambda_0(\tau \mid \gamma) d\tau$ reported in Table 5.1.

Step 3: Apply the left truncation scheme. We consider different truncation rates c running from 0 to 0.95 with increments of 0.05. For each t_{ij} the subject independently draws a truncation time $t_{0ij} \sim \mathcal{U}(b(c))$, and $\boldsymbol{t}_{0i}$ is the collection of the subject's threshold values. A subject is left-truncated and entirely dropped from the data if one of the subject's spells does not meet the threshold, i.e. $t_{ij} \leq t_{0ij}$ for some $j \in \{1, .., J\}$. The parameter $b(c)$ is tuned so that it guarantees truncation rate c. For each truncation rate we keep the number observed subjects constant at $N = 5,000$, because otherwise higher truncation rates mechanically lead to small sample sizes and lower accuracy. We repeat this procedure 100 times, so for each truncation rate, number of shared spells, and baseline hazard type, we generate 100 data sets containing 5,000 subjects.

Before estimating $\widehat{\beta}, \widehat{\gamma}, \widehat{\sigma^2}$ from the simulated data, we highlight in Figure 5.1 the observed threshold values t_{0ij} and the properties of our observed frailty distribution $G(\nu_i \mid \boldsymbol{t}_i \geq \boldsymbol{t}_{0i})$. Each graph plots the truncation rate against the median observed t_{0ij} and the expected observed frailty $\mathbb{E}(\nu_i \mid \boldsymbol{t}_i \geq \boldsymbol{t}_{0i})$. Each line is the connection of twenty scatters, with a separate scatter for each truncation rate. Each scatter represents the median estimate across the 100 simulated data sets.

Panels A and B show how the median value of t_{0ij} in the observed population varies with the truncation rate. Obviously, this median is zero at a truncation rate of zero, because we assume no truncation and thus no thresholds. Given the number of shared spells, a higher truncation rate implies a higher median threshold by construction. Instead taking the truncation rate as given, we see that the median threshold is lower if there are more shared spells; ceteris paribus, the subject also meets the additional thresholds if these thresholds are lower. Taken together, we expect that Method A and B produce more similar results if truncation rates are low or when there are more shared spells (the log-likelihoods $l_{\mathtt{A}}$ and $l_{\mathtt{B}}$ are then becoming asymptotically equivalent as $t_{0ij} \to 0$, see Section 5.2.5). When there are fewer spells or higher truncation rates we expect the estimation output by the two estimation methods to be more dissimilar.

Panels C and D reflect how mean observed frailty, measuring the degree of dynamic selection, varies with the truncation rate. To explain the patterns, we take the expected frailty in the underlying population $\mathbb{E}(\nu_i) = 1$ as reference point. Indeed, at a truncation rate of zero, i.e. no dynamic selection, we have that the observed mean $\mathbb{E}(\nu_i \mid \boldsymbol{t}_i \geq \boldsymbol{t}_{0i}) = 1$ is the same as in the underlying population. However, $\mathbb{E}(\nu_i \mid \boldsymbol{t}_i \geq \boldsymbol{t}_{0i})$ is lower at higher truncation rates so when truncation is more severe. The dynamic selection becomes stronger at higher truncation rates because a lower ν makes meeting the higher truncation thresholds more likely. Focussing on differences at a given truncation rate, we see that the dynamic selection is stronger if there are more shared spells: ceteris paribus, a lower ν makes it more likely to also meet the thresholds for the additional shared spells.

We now turn to comparing estimation outcomes $\widehat{\beta}$, $\widehat{\gamma}$, and $\widehat{\sigma^2}$ of using Method A, i.e. allowing for dynamic selection due left truncation and unobserved heterogeneity, to using Method B, i.e. not allowing for dynamic selection due left truncation and unobserved heterogeneity. Additionally, we introduce Method C which restricts estimation to $\sigma^2 = 0$ (the conventional `STATA`-command `streg` without the option `shared()`). The latter estimation is used to see how severe the bias from using Method B is compared to conventional biases that arise if hazard models are misspecified with an excluded frailty term. For the estimation, we assume that the researcher knows that the time-invariant impact does not feature a constant and that the functional form of the baseline hazard λ_0 is correct. We provide the median parameter estimates across the 100 simulated data sets in Figures 2 and 3. Each line is again a connection of twenty scatters. The second column in Figure 3 is a replication of van den Berg and Drepper (2016).

We see that Method A reports the underlying parameter values for any truncation rate, number of shared spells and baseline hazard. To this end, we had to compare the estimated parameters values $\widehat{\beta}$, $\widehat{\gamma}$, and $\widehat{\sigma^2}$ to their underlying true values: $\widehat{\beta} = \widehat{\gamma} = \widehat{\sigma^2} = 1$. This is unsurprising as the used log-likelihood specification $l_{\mathtt{A}}$ is tailored to the underlying data generating process in this simulation. Only at the very high truncation rate of 0.95, the parameter estimates deviate somewhat from the true parameter values.

We do see large deviations in $\widehat{\beta}$ and duration dependence $\widehat{\gamma}$ from their true values when we apply Method B. Figure 2 reveals that if the truncation rate is 0.5 (i.e. we leave out 50% of the subjects because one or more of their end times are below the truncation threshold) then our implied estimates are $\widehat{\beta} = 0.46$ and $\widehat{\gamma} = 0.53$ for a single-spell Weibull model. This suggests a substantial attenuation bias towards zero of 54% ($(1 - 0.46) \times 100\%$) and 47% if we use Method B. We also see that the bias increases if we have heavier truncation, reflected by a higher truncation rate. As pointed out in Section 2.5, the attenuation bias increases with the truncation rate because we observe a more selective sample of ν. The bias is comparable for the Gompertz case (see Figure 3).

The bias in parameter estimates $\widehat{\beta}$ and $\widehat{\gamma}$ remains substantial and negative if we consider shared frailty across more than one spell, but the size of the bias becomes smaller. E.g. at a truncation rate of 50% and five shared spells, the attenuation bias for $\widehat{\beta}$ and $\widehat{\gamma}$ with Weibull hazard declines from 54% and 47% for single spells to 2% and 3% for five spells. The decline in bias is a direct result of that $\boldsymbol{t}_{0i} \to 0$ (Figure 5.1) making $l_{\mathtt{B}}$ more similar to the unbiased method $l_{\mathtt{A}}$ (see Section 5.2.5).

Contrary to the bias in $\widehat{\beta}$ and $\widehat{\gamma}$, the bias in $\widehat{\sigma}^2$ can take different signs and explode if we consider more spells and heavier truncation schemes. Even though $\boldsymbol{t}_{0i} \to 0$, the thresholds are still positive and frailty distribution highly selective; the estimate $\widehat{\sigma}^2$ is sensitive for the positive small thresholds in combination with the strong selection.

Even though we obtained seemingly unbiased estimates $\widehat{\beta}$ and $\widehat{\gamma}$ at many spells, we consider the bias in $\widehat{\sigma}^2$ undesirable because this is sometimes as well a parameter of interest: take for example studies on mortality plateaus in old age where it is crucial to distinguish between frailty and age effects (see e.g. Barbi et al., 2018).

To put the bias from Method B into perspective, we also provide the parameter estimates that follow from erroneously specifying a hazard model without frailty, i.e. Method C. In all cases, we see that the bias from a wrongly specified hazard rate is a lower bound to the bias from applying Method B. Hence, with the current parameter constellation and truncation scheme it is more useful to specify a model with frailty.

5.3 Discussion and Conclusion

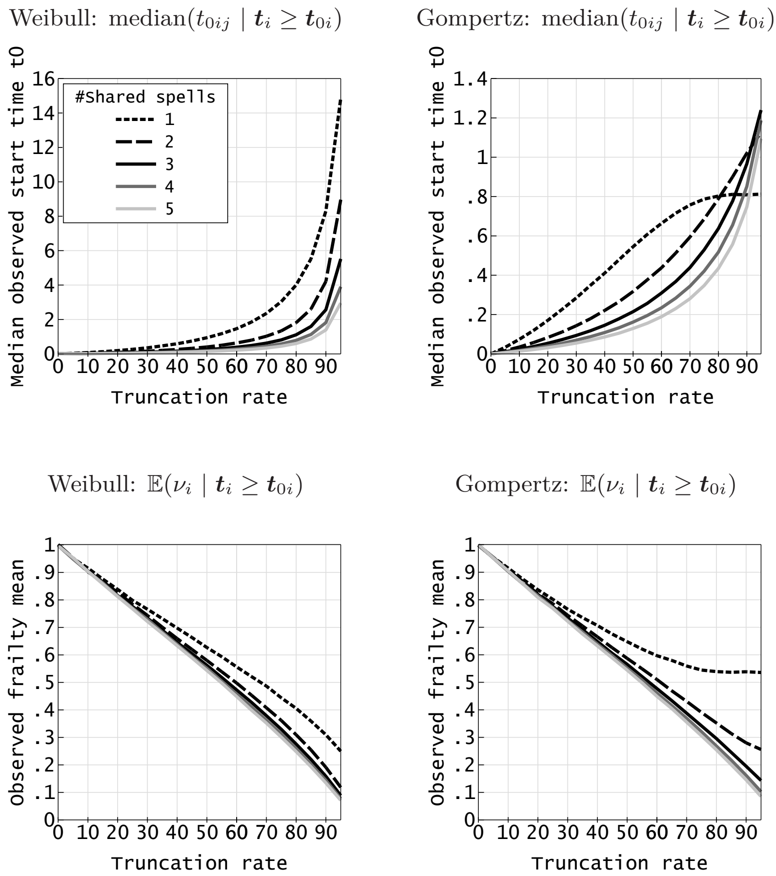
We developed a general estimation procedure to estimate shared frailty models with left truncation. Using a Monte Carlo experiment, we find that duration and covariate effects are downwards biased. At the same time, the frailty variance can be underestimated or overestimated if the interplay between left truncation and shared frailty is ignored. Our procedure resolves all biases, but the biases in time effects and covariate impacts also nullify if entry thresholds are equal to zero, a truncation rate is low, or the number of spells within a subject increases. Our estimation procedure is more versatile than other approaches because we can simultaneously allow for left truncation, right censoring, time-varying covariates, and an arbitrary amount of shared spells.

Caution remains warranted when interpreting our Monte Carlo experiment results because the biases strongly depend on our chosen truncation scheme and parameter values. We consider the natural extension of a left-truncated single spell per subject, i.e., all spells within a subject must meet a threshold in order to be in our sample. For instance, sampling families only if all members are alive on a particular date. Alternative schemes could involve repeated spells within a subject, where the subject is sampled only if its first spell meets a threshold. The size of the biases and the signs, particularly frailty variance, may differ under the different truncation schemes.

Our estimation procedure distinguishes from earlier programmed work in `R` and `STATA` particularly by that we simultaneously allow for time-varying covariates and left truncation (van den Berg and Drepper, 2016; Balan and Putter, 2019; Gorfine and Zucker, 2023). In the Monte Carlo experiment, we, however, only included time-invariant covariates for expository purposes. The direction of time-varying covariates and left truncation is an interesting avenue for future research for this Monte Carlo experiment and parameter identification. In this chapter, we have to assume that covariates are constant until the left truncation point; we encourage new work to find milder assumptions that yield parameter identification in a shared frailty model with time-varying covariates and left truncation (Lancaster, 1990; Honoré, 1993). Our programs can support a Monte Carlo experiment on such identification as well.

Researchers should be aware of the implications of ignoring dynamic selection due to left truncation. Even more so, because the `STATA` package `streg, shared()` assumes no dynamic selection due to left truncation, and several studies applied this estimation technique (for examples, see: van den Berg and Drepper, 2016). The good news of our research is that there are settings where dynamic selection due to left truncation is present but can be ignored to have unbiased estimates, in particular studies with many spells or low truncation rates.

Figure 5.1: Observed Left Truncation Thresholds and Observed Frailty by Left Truncation rate



Notes: Each line is the connection of twenty scatter points, each for a truncation rate between 0 to 0.95 (increments of 0.05). Each scatter is the median across 100 simulated data sets.

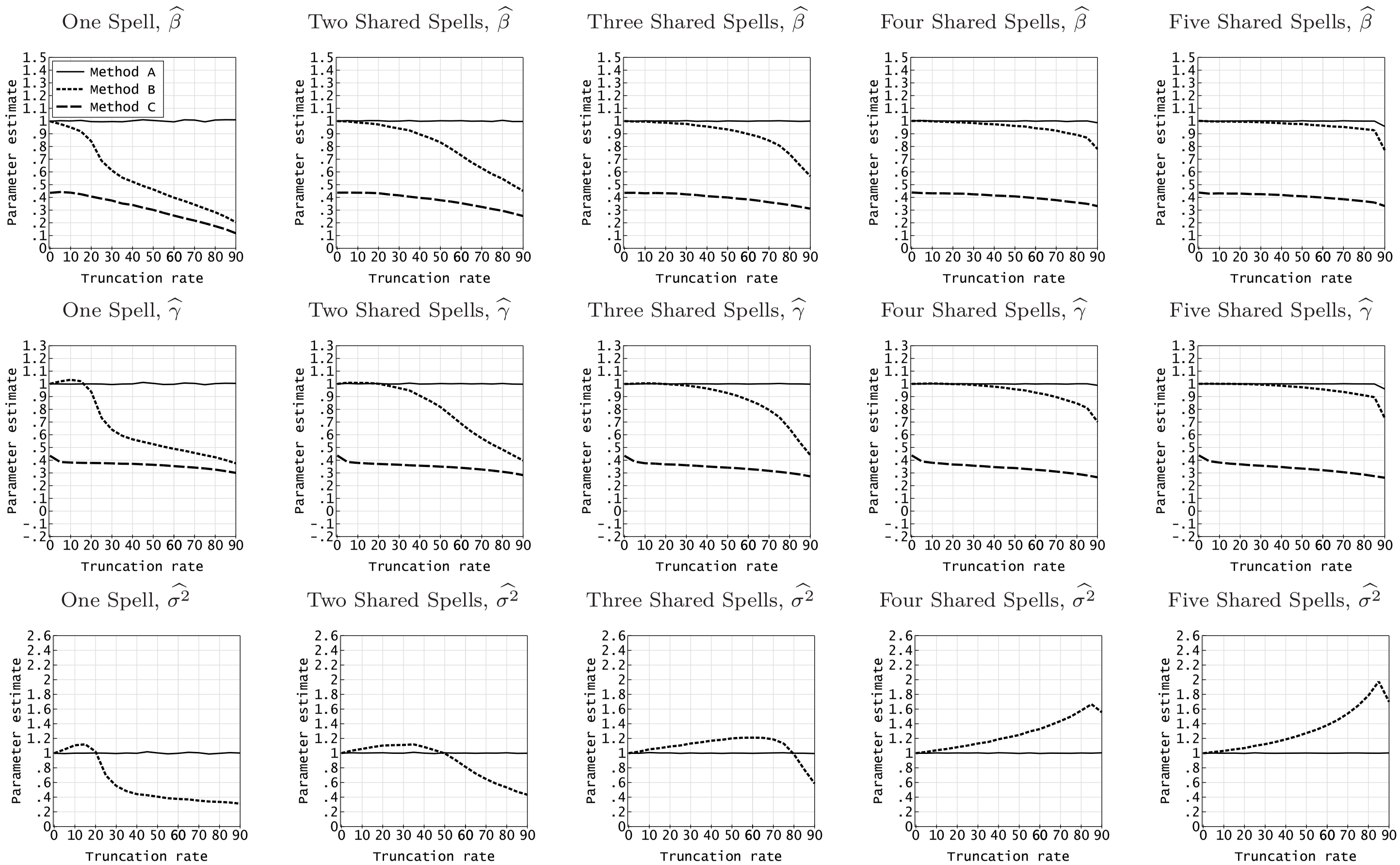


Figure 2: Parameter estimates for a Weibull specification with $\gamma = 1$, $\beta = 1$ and $\sigma^2 = 1$. Each line is the connection of twenty scatter points, each for a truncation rate between 0 to 0.95 (increments of 0.05). Each scatter is the median across 100 simulated data sets. Method A: log-likelihood according to l_A in Section 5.2.4. Method B: log-likelihood according to l_b in Section 5.2.5. Method C: Method A and B but restricting $\sigma^2 = 0$.

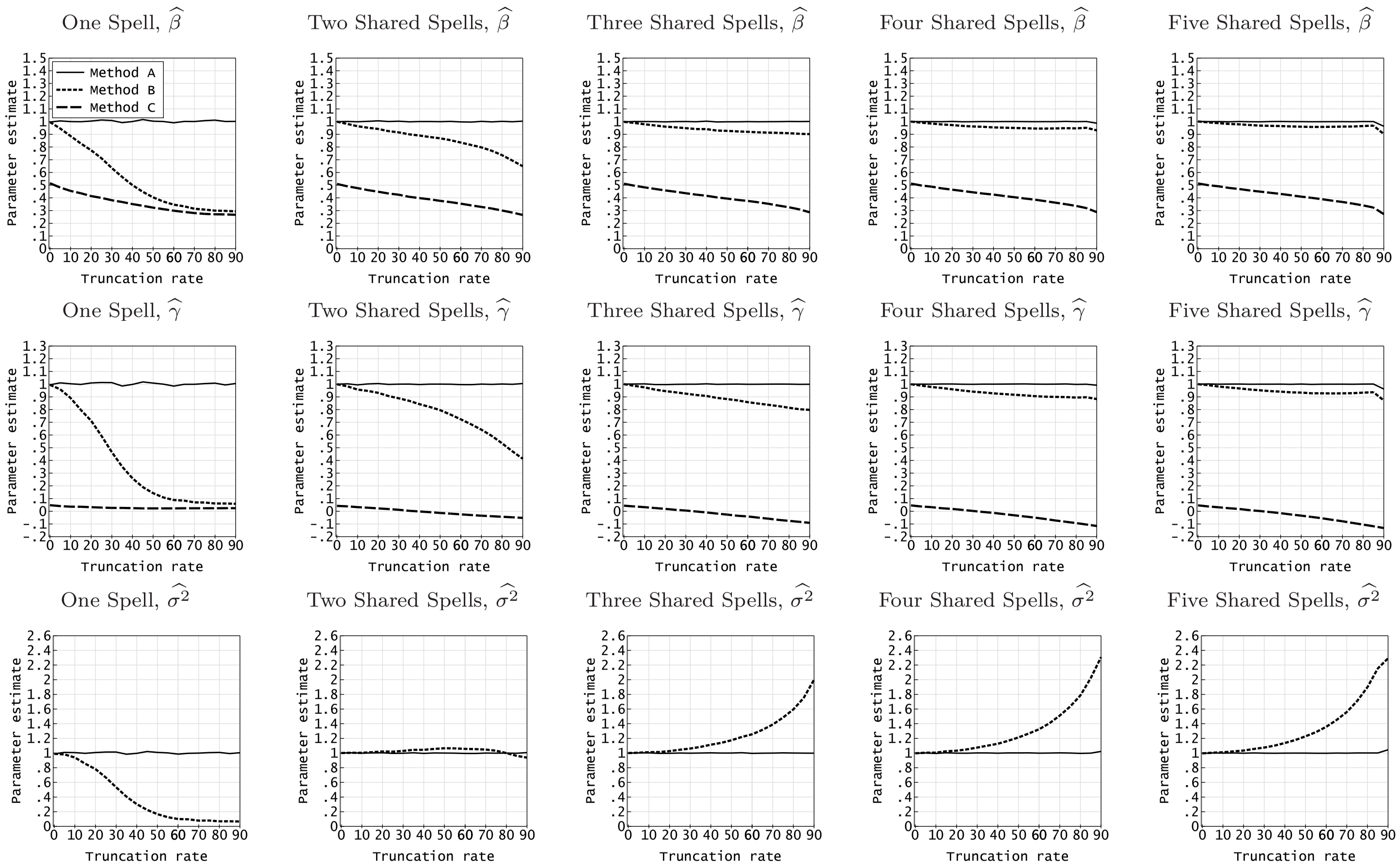


Figure 3: Parameter estimates for a Gompertz specification with $\gamma = 1$, $\beta = 1$ and $\sigma^2 = 1$. Each line is the connection of twenty scatter points, each for a truncation rate between 0 to 0.95 (increments of 0.05). Each scatter is the median across 100 simulated data sets. Method A: log-likelihood according to l_A in Section 5.2.4. Method B: log-likelihood according to l_B in Section 5.2.5. Method C: Method A and B but restricting $\sigma^2 = 0$.

Appendices

Appendix B: Chapter 2

B.1 Data

Table B.1: Variables Not Reported in Table 2.1

Variables:	Categorical values:
Woman	1: Yes; 0: No.
Prescribed medication for: (cf. Van Ooijen et al., 2015)	1: High blood cholesterol; 2: High blood pressure; 3: Coronary and peripheral vascular disease; 4: Cardiac diseases 5: Respiratory illness, asthma; 6: Diabetes; 7: Rheumatic Disease; 8: Gout; 9: Osteoporosis; 10: Peptic acid; 11: Cancer; 12: Cataract or glaucoma; 13: Epilepsy; 14: Thyroid disorder; 15: Pain; 16: Alzheimer; 17: Parkinson; 18: Sleep problems; 19: Depression; 20: Anxiety; 21: Psychotic illness; 22: Other: non-chronic
Ethnicity	1: Indigenous; 2: Moroccan; 3: Turkish; 4: Surinam; 5: Former Dutch Antilles incl. Aruba; 6: Other non-Western countries; 7: Other Western countries; 8: Unknown
Region of residence: (31 Regions with their own LTC procurement office that matches LTC supply and demand for the regional population)	1: Amstelland en de Meerland; 2: Amsterdam; 3: Apeldoorn/Zuthpen; 4: Arnhem; 5: Drenthe; 6: Flevoland; 7: Friesland; 8: Groningen; 9: Haaglanden; 10: Kennemerland; 11: Midden-Brabant; 12: Midden-Holland; 13: Midden-IJssel; 14: Nijmegen; 15: Noord- en Midden-Limburg; 16: Noord-Holland-Noord; 17: Noordoost-Brabant; 18: Rotterdam; 19: 't Gooi 20: Twente; 21: Utrecht; 22: Waardenland; 23: West-Brabant; 24: Westland; 25: Zaanstreek/Waterland; 26: Zeeland; 27: Zuid-Holland-Noord; 28: Zuid-Hollandse Eilanden; 29: Zuid-Limburg; 30: Zuidoost-Brabant; 31: Zwolle

Table B.2: Overview of Used Registers

Name	Waves	Content	Used for variable on:
Zorgzvtab	2004-2008	Dates of use of formal home-based care	Past LTC use
Gebzzvtab	2009-2014	Dates of use of formal home-based care	Spell on home-based care use; Partner's health; Past LTC use
Zorgmvtab	2004-2014	Dates of use of institutional care	Spell on institutional care use; Partner's health; Past LTC use
Indicawbztab	2009-2014	LTC eligibility assessment	Spell on the need for LTC
Gbaoverlijdentab	2021	Date of death*	Time till death; Children alive
Gbapersoontab	2021	Residents within the Netherlands or those who live abroad and hold a relation with the Dutch government*	Initial sample; Gender; Ethnicity; Birth date (age)
Kindoudertab	2021	Child-parent link*	Children
Gbaverbintenispartnerbus	2019	Married couples and registered partners*	Marital status
Gbahuishoudensbus	2020	Household members and relations among them	Cohabiting children or partner
Gbaadresobjectbus	2021	Individual's address*	Migration; Region of residence; Distance to the children
Vslgwbtab	2021	Municipality where the address is located*	Region of residence; Distance to the children
Integraal huishoudens inkomen	2006-2013	Household income and homeownership	Pre-tax equivalized household income (decile); Homeownership
Vehtab	2006-2013	Household's assets	Household financial assets (decile)
Medicijntab	2006-2013	Prescribed drugs with ATC codes	Medication use

* Longitudinal file that daily reports the variables between 1995 and 2021.

Table B.3: Care Severity Packages for Institutional Care

Type of care	Severity level	LTC services	Hours per week	Client profile
Nursing and caring	1	All but NC	3.0-5.0	Until 2013. Residential home: 'light' user.
Nursing and caring	2	All	5.5-7.5	Until 2013. Residential home: 'average' user or recovering from hospitalization
Nursing and caring	3	All	9.5-11.5	Until 2014. Nursing home: light physical impairments. Residential home: complex physical impairments.
Nursing and caring	4	All	11.0-13.5	Nursing home: starting cognitive impairments Nursing home: average physical impairments
Nursing and caring	5	All	16.5-20.0	Nursing home: intensive dementia care
Nursing and caring	6	All	16.5-20.0	Nursing home: severe physical impairments such as Parkinson's disease, chronic heart failure, etc.
Nursing and caring	7	All	20.0-24.5	More complex impairments than level 6.
Nursing and caring	8	All	24.5-29.5	Last stage of a physical impairment.
Nursing and caring	9a; 9b	All	18.0-22.0	Until 2013: Rehabilitation in a nursing home
Nursing and caring	10	All	26.5-32.5	Palliative care
Physically disabled	1	All	14.0-17.0	Clients with permanent physical impairments that do not improve with treatment.
Physically disabled	2	All	17.5-21.5	
Physically disabled	3	All	18.0-22.0	
Physically disabled	4	All	19.5-24.0	
Physically disabled	5	All	24.0-29.0	
Physically disabled	6	All	25.5-31.5	
Physically disabled	7	All	27.0-33.5	
Mild intellectually disabled	1	All but NC	12.5-15.5	Focus on sheltered living (no nursing).
Mild intellectually disabled	2	All but NC	17.0-21.0	
Mild intellectually disabled	3	All but NC	22.0-27.0	

AS: function 'Social support'; PC: function 'Personal Care'; NC: function 'Nursing Care'; TR: function 'Treatment'. An indication for institutional care grants access to a type of care (sector), LTC services, and an number of hours per week that the client can access the services: a care severity package. Found on: `https://wetten.overheid.nl/BWBR0014855/2014-02-06#Bijlage2`.

Table B.4: Care Severity Packages for Institutional Care (continued)

Type of care	Severity level	LTC services	Hours per week	Client profile
Care for mild intellectually disabled	4	All but NC	27.0-33.0	
Care for mild intellectually disabled	5	All but NC	27.0-33.0	
Intellectually disabled	1	All	10.0-12.0	
Intellectually disabled	2	All	12.5-15.0	
Intellectually disabled	3	All	15.0-18.5	
Intellectually disabled	4	All	17.0-21.0	
Intellectually disabled	5	All	22.0-27.0	
Intellectually disabled	6	All	21.0-26.0	
Intellectually disabled	7	All	20.5-37.0	
Mental health care	1B; 1C	All; PC & AS	7.5-9.5; 6.5-8.5	B: Treatment; C: Social support
Mental health care	2B; 2C	All; PC & AS	12.0-14.5; 11.0-13.5	
Mental health care	3B; 3C	All; PC & AS	13.5-16.5; 12.5-15.0	
Mental health care	4B; 4C	All; PC & AS & NC	16.0-19.5; 15.0-18.5	
Mental health care	5B; 5C	All; PC & AS & NC	17.0-21.0; 16.5-20.0	
Mental health care	6B; 6C	All; PC & AS & NC	22.5-27.5;20.5-25.5	
Mental health care	7B	All; PC & AS	32.5-39.5	
Sensory disabled: hearing	1	PC & AS & TR	17.5-21.5	Severe hearing problems
Sensory disabled: hearing	2	All	34.0-42.0	
Sensory disabled: hearing	3	All	40.5-49.5	
Sensory disabled: hearing	4	PC & AS & TR	25.0-31.0	
Sensory disabled: sight	1	PC & AS	11.0-13.5	Severe sight problems or blindness
Sensory disabled: sight	2	PC & AS	15.0-18.0	
Sensory disabled: sight	3	All	19.0-23.5	
Sensory disabled: sight	4	All	26.0-31.5	
Sensory disabled: sight	5	All	28.5-35.0	

AS: function 'Social support'; PC: function 'Personal Care'; NC: function 'Nursing Care'; TR: function 'Treatment'. An indication for institutional care grants access to a type of care (sector), LTC services, and an number of hours per week that the client can access the services: a care severity package. Found on: `https://wetten.overheid.nl/BWBR0014855/2014-02-06#Bijlage2`.

B.2 LTC Assessment

B.2.1 Description of LTC Services

- **Personal care**: involves (simple) medication uptake and help with activities of daily living (ADLs), including bathing, dressing, ambulating, toileting (including continence), feeding;
- **Nursing care**: more specialized care such as needle injection and catheterization;
- **Treatment**: aims at rehabilitation or prevents worsening of the limitation (we consider this to be nursing care);
- **Social support** includes daycare in groups or personal assistance, e.g., help with organizing the household and doing administration (Mot, 2010).

B.2.2 Entitlement to Home-based Care Services

The entitlements are in hours per week per LTC service. The classification is as follows:

- **Personal care**: 0-1.9 hours; 2-3.9 hours; 4-6.9 hours; 7-9.9 hours; 10-12.9 hours; 13-15.9 hours; 16-19.9 hours; 20-24.9 hours; 25 hours or more;
- **Nursing care**: 0-1.9 hours; 2-3.9 hours; 4-6.9 hours; 7-9.9 hours; 10-12.9 hours; 13-15.9 hours; 16-19.9 hours; 20-24.9 hours; 25 hours or more;
- **Treatment (individual)**: no hours allocation;
- **Treatment (group)**: 1 day part, 2 day parts; 3 day parts; 4 day parts; 5 day parts; 6 day parts; 7 day parts; 8 day parts; 9 day parts. A day part is 4 hours;
- **Social support (individual)** 0-1.9 hours; 2-3.9 hours; 4-6.9 hours; 7-9.9 hours; 10-12.9 hours; 13-15.9 hours; 16-19.9 hours; 20-24.9 hours; 25 hours or more;
- **Social support (group)** 1 day part, 2 day parts; 3 day parts; 4 day parts; 5 day parts; 6 day parts; 7 day parts; 8 day parts; 9 day parts. A day part is 4 hours.

B.3 Descriptive Statistics

Table B.5: Observed Characteristics by Used LTC Arrangement (all)

	No LTC use	**Home-based care**	**Institutional care**
Woman = 1	0.54	0.71	0.73
Partner = 1	0.68	0.30	0.19
Partner uses same arrangement = 1*	0.97	0.29	0.44
Has children = 1	0.87	0.84	0.76
Median equivalized household income**	22.9	19.6	17.8
Median household financial assets**	30.6	23.2	23.2
Homeowner = 1	0.56	0.30	0.16
Age	73.3	82.1	84.1
Main health problem:			
Has physical impairment = 1**		0.80	0.51
Has cognitive impairment = 1		0.09	0.36
Has other problem = 1***		0.03	0.09
Has no entitlement = 1		0.08	0.04
Individuals (%):	13,598,785 (87)	1,115,609 (7)	963,410 (6)

* Conditional upon having a partner; ** 000s€ *** Physical impairment or disability; **** A sensory disability, intellectual disability or mental disorder

Table B.6: Entitled Hours of Care by Used LTC Arrangement and Health Problem (all)

	Home-based care		Institutional care	
Impairment:	Physical	Cognitive	Physical	Cognitive
Hours of care per week:				
0-2	17	6	2	4
2-4	23	11	0	0
4-7	25	16	9	0
7-10	16	12	17	0
10-13	8	6	22	1
13-16	4	21	13	16
16-20	3	25	22	67
20-25	3	2	12	12
25+	2	0	5	0
$\sum$	100%	100%	100%	100%
Median	5.5	11.5	13.25	19.25
N	897,806	98,201	491,449	348,990

B.4 Mixed Proportional Hazard Model: Distance to the Closest Child

Table B.7: Hazard Ratio Estimates for Transitions to a More Specialized LTC Arrangement

From:	No LTC use Never used LTC before	No LTC use Ever used LTC before	Home-based care Physical impairment		Home-based care Cognitive impairment	
	-	-	Low need	High need	Low need	High need
To:	Home-based care		Institutional care		Institutional care	
	(1)	(2)	(3)	(4)	(5)	(6)
Partner (ref: single)						
Partner does not use LTC	0.55***	0.79***	1.00	0.80***	1.82***	1.19***
Partner (ref: partner does not use LTC)						
Partner uses home-based care	3.10***	1.98***	1.03	1.14***	1.16**	0.98
Partner uses institutional care	2.53***	1.47***	4.17***	3.05***	1.80***	3.00***
Distance to the closest child (ref: no children)						
Lives in the same municipality	0.91***	0.93****	0.82***	0.88***	0.76***	0.88***
Lives in another municipality	0.90***	0.94***	0.94*	0.96***	0.77***	0.91***
Income (ref: lowest income decile)						
Highest income decile	0.73***	0.90***	0.96	0.90***	0.96	0.82***
Assets (ref: lowest asset decile)						
Highest asset decile	0.95***	0.91***	0.94	0.92***	1.02	0.83***
Homeowner (ref: renter)						
Homeowner	0.86***	0.87***	0.99	0.93***	0.91**	0.90***
Duration dependence						
γ_{ij}	0.11***	0.03***	-0.30**	-0.46***	-0.15***	0.19***
Individual-shared frailty						
σ^2_{ij}	0.21***	0.43***	2.31***	0.43***	1.29***	1.22***
Individuals	2,888,623	628,283	452,678	592,030	52,687	78,012
Spells	2,888,623	867,987	635,791	829,767	62,357	86,727
Transition probability (%)	71	70	3	19	11	72
Sub-log-likelihood (cf. (2.8))	-1,854,947.63	-926,998.69	-54,609.41	-264,754.72	-15,846.52	-34,265.42
Log-likelihood	-2,905,300.63	-1,449,334.09	-594,298.36	-1,089,868.88	-98,854.21	-91,784.15

Notes: The estimates are hazard ratios exp(β_{ij}) cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test).

Table B.8: Hazard Ratio Estimates for Transitions to a Less Specialized LTC Arrangement

From:	Home-based care Physical impairment		Home-based care Cognitive impairment		Institutional care Physical impairment		Institutional care Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	No LTC use		No LTC use		Home-based care		Home-based care	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	1.54***	1.25***	2.28***	1.41***	2.15***	1.33***	9.13***	3.33***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	0.48***	0.59***	0.61***	0.75***	1.28***	1.46***	1.45***	1.61***
Partner uses institutional care	0.58***	0.75***	0.56***	1.03	0.15***	0.38***	0.09***	0.30***
Distance to the closest child (ref: no children)								
Lives in the same municipality	1.11***	1.13***	0.90*	0.98	1.13***	1.12***	1.17	1.00
Lives in another municipality	1.20***	1.24***	1.02	1.14**	1.10***	1.09***	1.11	0.92
Income (ref: lowest income decile)								
Highest income decile	1.46***	1.39***	1.36***	1.69***	1.72***	1.79***	1.86***	1.85***
Assets (ref: lowest asset decile)								
Highest asset decile	1.01	1.05***	0.82*	1.08	1.10**	1.01	1.01	0.82*
Homeowner (ref: renter)								
Homeowner	1.11***	1.13***	1.14***	1.11**	1.32***	1.24***	1.77***	1.37***
Duration dependence								
γ_{ij}	-1.68***	-1.24***	-0.84***	-0.59***	-4.75***	-3.28***	-4.93**	-6.06***
Individual-shared frailty								
σ^2_{ij}	0.63***	0.89***	2.93***	2.39***	0.39***	0.20***	3.79***	4.00***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	37	25	11	9	13	14	4	4
Sub-log-likelihood (cf. (2.8))[1]	-157,385.39	-221,417.00	-15,208.75	-17,427.50	-22,883.54	-61,844.84	-4,101.24	-13,340.14
Log-likelihood[2]	-594,298.36	-1,089,868.88	-98,854.21	-91,784.15	-187,080.90	-333,568.80	-61,546.94	-198,666.52

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8). [2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

Table B.9: Hazard Ratio Estimates for Transitions Within a LTC Arrangement

From:	Home-based care				Institutional care			
Impairment:	Physical impairment		Cognitive impairment		Physical impairment		Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	High‘need	Low need	High need	Low need	High need	Low need	High need	Low need
Partner (ref: single)								
Partner does not use LTC	0.91***	0.94***	0.96*	1.75***	2.67***	0.55***	3.38***	0.57***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	1.04***	1.02	1.10***	0.98	1.20**	1.73***	1.18*	1.90***
Partner uses institutional care	1.42***	1.08***	1.36***	0.62**	0.23***	1.83***	0.29***	1.83***
Distance to the closest child (ref: no children)								
Lives in the same municipality	1.00	0.96***	1.01	0.87*	0.95*	0.95	1.02	0.97
Lives in another municipality	1.01	0.98**	1.04	0.86*	0.95*	0.98	0.99	1.04
Income (ref: lowest income decile)								
Highest income decile	1.13***	0.92***	1.12**	1.01	1.53***	1.11	1.22**	1.01
Assets (ref: lowest asset decile)								
Highest asset decile	0.96**	1.01	1.03	0.90	0.91*	0.90*	1.01	1.12
Homeowner (ref: renter)								
Homeowner	1.05***	0.99	0.97*	1.16***	1.15***	1.06*	1.12***	1.29***
Duration dependence								
γ_{ij}	-0.21***	-1.34**	-0.11***	-1.90***	0.07***	-1.45***	0.52***	-2.55***
Individual-shared frailty								
σ^2_{ij}	1.28***	0.28***	0.34***	1.72***	1.47***	0.00	1.24***	0.50
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	48	23	49	4	34	5	62	1
Sub-log-likelihood (cf. (2.8))[1]	-258,220.58	-267,994.56	-35,513.81	-10,499.87	-72,524.18	-38,216.75	-31,392.24	-6,833.83
Log-likelihood[2]	-594,298.36	-1,089,868.88	-98,854.21	-91,784.15	-187,080.90	-333,568.80	-61,546.94	-198,666.52

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8).[2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

B.5 Mixed Proportional Hazard Model: Child Effect Split by Partner Status

Table B.10: Hazard Ratio Estimates for Transitions to a More Specialized LTC Arrangement

From:	No LTC use	No LTC use	Home-based care		Home-based care	
	Never used LTC before	Ever used LTC before	Physical impairment		Cognitive impairment	
	-	-	Low need	High need	Low need	High need
To:	Home-based care		Institutional care		Institutional care	
Partner (ref: single)						
Partner does not use LTC	0.56***	0.79***	1.01	0.87***	1.69***	1.20***
Partner (ref: partner does not use LTC)						
Partner uses home-based care	3.12***	2.06***	1.18	1.14***	1.25	0.93
Partner uses institutional care	2.58***	1.62***	4.13***	3.05***	2.21*	3.02***
Children (ref: has no children)						
Hazard ratio of having children split by partner category:						
Single	0.91***	0.93***	0.87***	0.92***	0.74***	0.89***
Partner does not use LTC	0.90***	0.94***	0.85*	0.84***	0.81*	0.89**
Partner uses home-based care	0.99	0.96	0.85	1.00	0.92	1.05
Partner uses institutional care	0.98	0.90*	1.01	1.00	0.80	0.99
Income (ref: lowest income decile)						
Highest income decile	0.72***	0.90***	0.98	0.91***	0.97	0.83***
Assets (ref: lowest asset decile)						
Highest asset decile	0.95***	0.91***	0.94	0.92***	1.02	0.83***
Homeowner (ref: renter)						
Homeowner	0.86***	0.87***	0.99	0.93***	0.91**	0.90***
Duration dependence						
γ_{ij}	0.11***	0.03***	-0.30**	-0.46***	-0.15**	0.19***
Individual-shared frailty						
σ^2_{ij}	0.21***	0.43***	2.32***	0.43***	1.30***	1.22***
Individuals	2,888,623	628,283	452,678	592,030	52,687	78,012
Spells	2,888,623	867,987	635,791	829,767	62,357	86,727
Sub-log-likelihood (cf. (2.8))	-1,854,959.50	-927,000.69	-54,627.09	-264,811.19	-15,846.06	-34,266.69
Log-likelihood	-2,905,357.97	-1,449,375.11	-594,386.70	-1,090,016.00	-98,851.68	-91,788.04

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). *5-% **1-% ***0.1-% significance level.

Table B.11: Hazard Ratio Estimates for Transitions to a Less Specialized LTC Arrangement

From:	Home-based care Physical impairment		Home-based care Cognitive impairment		Institutional care Physical impairment		Institutional care Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	No LTC use		No LTC use		Home-based care		Home-based care	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	1.62***	1.39***	2.44**	1.40**	2.10***	1.38***	8.75**	3.50***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	0.48***	0.55***	0.44***	0.93	1.33*	1.43***	1.58	1.70***
Partner uses institutional care	0.52***	0.69***	0.37	1.49	0.16***	0.40***	0.09***	0.37***
Children (ref: has no children)								
Hazard ratio of having children split by partner category:								
Single	1.16***	1.20***	0.93	1.06	1.12***	1.13***	1.14	1.02
Partner does not use LTC	1.09***	1.06***	0.86	1.06	1.15*	1.08**	1.19	0.96
Partner uses home-based care	1.02	1.08	1.44	0.79	0.96	1.02	0.91	0.94
Partner uses institutional care	1.14	1.11	1.60	0.66	0.90	0.93	0.97	0.79
Income (ref: lowest income decile)								
Highest income decile	1.48***	1.41***	1.39**	1.73***	1.72***	1.79***	1.85**	1.83***
Assets (ref: lowest asset decile)								
Highest asset decile	1.02	1.05***	0.82*	1.08	1.10**	1.01	1.01	0.82*
Homeowner (ref: renter)								
Homeowner	1.11***	1.13***	1.14***	1.12**	1.32***	1.24***	1.77***	1.36***
Duration dependence								
γ_{ij}	-1.68***	-1.24***	-0.84***	-0.59***	-4.75***	-3.29***	-4.93**	-6.06***
Individual-shared frailty								
σ^2_{ij}	0.64***	0.89***	2.94***	2.38***	0.39***	0.20***	3.78***	3.98***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	37	25	11	9	13	14	4	4
Sub-log-likelihood (cf. (2.8))[1]	-157,459.69	-221,482.47	-15,212.17	-17,435.66	-22,884.66	-61,847.22	-4,101.59	-13,341.54
Log-likelihood[2]	-594,386.70	-1,090,016.00	-98,851.68	-91,788.04	-187,075.04	-333,553.85	-61,542.01	-198,665.75

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8). [2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

Table B.12: Hazard Ratio Estimates for Transitions Within a LTC Arrangement

From:	Home-based care				Institutional care			
	Physical impairment		Cognitive impairment		Physical impairment		Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	High‘need	Low need	High need	Low need	High need	Low need	High need	Low need
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	0.89***	0.99	1.08	1.57***	2.53***	0.53***	2.86***	0.53*
Partner (ref: partner does not use LTC)								
Partner uses home-based care	1.11**	0.97	1.13	1.49*	1.05	1.57*	1.29	1.96
Partner uses institutional care	1.52***	1.13*	1.62*	0.32	0.28***	2.12***	0.43***	1.86
Children (ref: has no children)								
Hazard ratio of having children split by partner category:								
Single	1.00	0.98**	1.06*	0.88	0.96	0.96	1.02	0.98
Partner does not use LTC	1.02	0.92***	0.93	0.99	1.01	1.02	1.23	1.07
Partner uses home-based care	0.93*	1.06	0.97	0.62**	1.16	1.11	0.90	0.96
Partner uses institutional care	0.93	0.94	0.82	2.02	0.82	0.85	0.64**	0.98
Income (ref: lowest income decile)								
Highest income decile	1.14***	0.92***	1.12**	1.01	1.53***	1.12	1.22**	1.01
Assets (ref: lowest asset decile)								
Highest asset decile	0.96**	1.01	1.04	0.90	0.91*	0.90*	1.00	1.12
Homeowner (ref: renter)								
Homeowner	1.05***	0.99	0.97*	1.16***	1.15***	1.06*	1.12***	1.29***
Duration dependence								
γ_{ij}	-0.21***	-1.34**	-0.11***	-1.90***	0.07***	-1.45***	0.53***	-2.55***
Individual-shared frailty								
σ^2_{ij}	1.28***	0.28***	0.34***	1.72***	1.47***	0.00	1.24***	0.50
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	48	23	49	4	34	5	62	1
Sub-log-likelihood (cf. (2.8)[1]	-258,220.23	-267,991.63	-35,509.91	-10,495.58	-72,532.12	-38,216.39	-31,388.92	-6,834.40
Log-likelihood[2]	-594,386.70	-1,090,016.00	-98,851.68	-91,788.04	-187,075.04	-333,553.85	-61,542.01	-198,665.75

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test); $\sigma^2_{ij} = 0$ (Likelihood-ratio test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8). [2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

B.5.1 Robustness Checks

Table B.13: Hazard Ratio Estimates for Transitions to a More Specialized LTC Arrangement: Cox Model

From:	No LTC use	No LTC use	Home-based care		Home-based care	
	Never used LTC before	Ever used LTC before	Physical impairment		Cognitive impairment	
	-	-	Low need	High need	Low need	High need
To:	Home-based care		Institutional care		Institutional care	
	(1)	(2)	(3)	(4)	(5)	(6)
Partner (ref: single)						
Partner does not use LTC	0.60***	0.82***	1.01	0.80***	1.77***	1.15***
Partner (ref: partner does not use LTC)						
Partner uses home-based care	2.98***	1.95***	1.02	1.17***	1.08	0.94***
Partner uses institutional care	2.34***	1.37***	3.42***	2.64***	1.49***	1.63***
Children (ref: no children)						
Has children	0.93***	0.95***	0.86***	0.91***	0.78***	0.96***
Income (ref: lowest income decile)						
Highest income decile	0.73***	0.92***	0.98	0.90***	0.95	0.86***
Assets (ref: lowest asset decile)						
Highest asset decile	0.98*	0.93***	0.96	0.93***	1.02	0.88***
Homeowner (ref: renter)						
Homeowner	0.89***	0.90***	0.99	0.94***	0.92*	0.92***
Individuals	2,888,623	628,283	452,678	592,030	52,687	78,012
Spells	2,888,623	867,987	635,791	829,767	62,357	86,727
Transition probability (%)	71	70	3	19	11	72

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). However, the model is estimated as a Cox model (baseline hazard $\phi_{ij}(t)$ is not parameterized). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table.

Table B.14: Hazard Ratio Estimates for Transitions to a Less Specialized LTC Arrangement: Cox Model

From:	Home-based care Physical impairment		Home-based care Cognitive impairment		Institutional care Physical impairment		Institutional care Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	No LTC use		No LTC use		Home-based care		Home-based care	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	1.46***	1.21***	1.94***	1.28***	2.02***	1.33***	5.70***	3.14***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	0.52***	0.62***	0.67***	0.81***	1.24***	1.44***	1.29***	1.54***
Partner uses institutional care	0.61***	0.79***	0.62***	1.03	0.16***	0.38***	0.14***	0.33***
Children (ref: no children)								
Has children	1.12***	1.15***	0.94	1.05	1.11***	1.11***	1.10	0.96
Income (ref: lowest income decile)								
Highest income decile	1.41***	1.38***	1.31***	1.56***	1.66***	1.77***	1.68***	1.72***
Assets (ref: lowest asset decile)								
Highest asset decile	1.01	1.04**	0.84**	1.08	1.09**	1.01	0.96	0.84*
Homeowner (ref: renter)								
Homeowner	1.09***	1.11***	1.11***	1.09**	1.30***	1.23***	1.59***	1.31***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	37	25	11	9	13	14	4	4

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). However, the model is estimated as a Cox model (baseline hazard $\phi_{ij}(t)$ is not parameterized). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table.

Table B.15: Hazard Ratio Estimates for Transitions Within a LTC Arrangement: Cox Model

From:	Home-based care				Institutional care			
	Physical impairment		Cognitive impairment		Physical impairment		Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	High'need	Low need	High need	Low need	High need	Low need	High need	Low need
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	0.91***	0.93***	0.97	1.64***	1.87***	0.57***	1.96***	0.57***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	1.07***	1.02	1.08***	1.00	1.19***	1.74***	1.14*	1.90***
Partner uses institutional care	1.25***	1.08***	1.28***	0.65**	0.39***	1.74***	0.51***	1.83***
Children (ref: no children)								
Has children	0.99	0.97***	1.02	0.87**	0.97	0.97	1.01	0.99
Income (ref: lowest income decile)								
Highest income decile	1.05***	0.92***	1.11**	1.01	1.21***	1.13*	1.16**	1.00
Assets (ref: lowest asset decile)								
Highest asset decile	0.96***	1.01	1.03	0.90	0.97	0.90*	1.01	1.12
Homeowner (ref: renter)								
Homeowner	1.03***	0.99	0.96*	1.16***	1.11***	1.07**	1.06**	1.29***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	48	23	49	4	34	5	62	1

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). However, the model is estimated as a Cox model (baseline hazard $\phi_{ij}(t)$ is not parameterized). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table.

Table B.16: Hazard Ratio Estimates for Transitions to a More Specialized LTC Arrangement: No Frailty

From:	No LTC use	No LTC use	Home-based care		Home-based care	
	Never used LTC before	Ever used LTC before	Physical impairment		Cognitive impairment	
	-	-	Low need	High need	Low need	High need
To:	Home-based care		Institutional care		Institutional care	
	(1)	(2)	(3)	(4)	(5)	(6)
Partner (ref: single)						
Partner does not use LTC	0.60***	0.82***	1.03	0.83***	1.79***	1.18***
Partner (ref: partner does not use LTC)						
Partner uses home-based care	3.01***	1.92***	0.99	1.11***	1.07	0.92***
Partner uses institutional care	2.36***	1.38***	3.34***	2.55***	1.48***	1.62***
Children (ref: no children)						
Has children	0.93***	0.95***	0.87***	0.92***	0.78***	0.96***
Income (ref: lowest income decile)						
Highest income decile	0.73***	0.91***	1.00	0.93***	0.96	0.87***
Assets (ref: lowest asset decile)						
Highest asset decile	0.99	0.93***	0.95	0.93***	1.02	0.88***
Homeowner (ref: renter)						
Homeowner	0.88***	0.89***	0.99	0.94***	0.92*	0.92***
Duration dependence						
γ_{ij}	0.09***	-0.06***	-0.41***	-0.57***	-0.33***	-0.77***
Individuals	2,888,623	628,283	452,678	592,030	52,687	78,012
Spells	2,888,623	867,987	635,791	829,767	62,357	86,727
Transition probability (%)	71	70	3	19	11	72
Sub-log-likelihood (cf. (2.8))[1]	-1,858,370.13	-933,150.31	-54,879.61	-265,764.22	-15,925.90	-35,737.84
Log-likelihood[2]	-2,910,097.13	-1,458,275.36	-614,999.87	-1,117,533.80	-99,530.56	-93,735.75

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). However, frailty is restricted to be absent in the model ($\sigma^2_{ij} = 0$). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8). [2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

Table B.17: Hazard Ratio Estimates for Transitions to a Less Specialized LTC Arrangement: No Frailty

From:	Home-based care Physical impairment		Home-based care Cognitive impairment		Institutional care Physical impairment		Institutional care Cognitive impairment	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	No LTC use		No LTC use		Home-based care		Home-based care	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	1.47***	1.21***	1.97***	1.32***	2.03***	1.32***	6.27***	3.23***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	0.52***	0.62***	0.66***	0.79***	1.25***	1.44***	1.33***	1.56***
Partner uses institutional care	0.61***	0.79***	0.62***	1.01	0.16***	0.38***	0.13***	0.31***
Children (ref: no children)								
Has children	1.12***	1.15***	0.94	1.05	1.11***	1.11***	1.10	0.96
Income (ref: lowest income decile)								
Highest income decile	1.42***	1.38***	1.33***	1.57***	1.66***	1.76***	1.68**	1.72***
Assets (ref: lowest asset decile)								
Highest asset decile	1.01	1.04**	0.84**	1.09	1.10**	1.01	0.96	0.84*
Homeowner (ref: renter)								
Homeowner	1.09***	1.11***	1.11***	1.08**	1.30***	1.23***	1.59***	1.31***
Duration dependence								
γ_{ij}	-2.21***	-1.68***	-1.26***	-0.85***	-5.04***	-3.42***	-6.07***	-6.55***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	37	25	11	9	13	14	4	4
Sub-log-likelihood (cf. (2.8))[1]	-163,082.67	-227,575.91	-15,720.47	-17,707.78	-23,064.71	-61,987.04	-4,370.13	-13.875.16
Log-likelihood[2]	-614,999.87	-1,117,533.80	-99,530.56	-93,735.75	-188,224.73	-344,270.82	-62,589.94	-199,783.93

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). However, frailty is restricted to be absent in the model ($\sigma^2_{ij} = 0$). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8). [2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

Table B.18: Hazard Ratio Estimates for Transitions Within a LTC Arrangement: No Frailty

From:	Home-based care				Institutional care			
	Physical need		Cognitive need		Physical need		Cognitive need	
	Low need	High need	Low need	High need	Low need	High need	Low need	High need
To:	High'need	Low need	High need	Low need	High need	Low need	High need	Low need
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partner (ref: single)								
Partner does not use LTC	0.96***	0.93***	0.98	1.68***	2.31***	0.55***	2.08***	0.57***
Partner (ref: partner does not use LTC)								
Partner uses home-based care	1.00*	1.02*	1.08***	0.99	1.25***	1.73***	1.17*	1.89***
Partner uses institutional care	1.17***	1.08***	1.27**	0.64**	0.30***	1.83***	0.48***	1.83***
Children (ref: no children)								
Has children	1.00	0.97***	1.02	0.87**	0.98	0.96	1.01	0.99
Income (ref: lowest income decile)								
Highest income decile	1.11***	0.92***	1.11**	1.01	1.31***	1.12	1.16***	1.01
Assets (ref: lowest asset decile)								
Highest asset decile	0.96***	1.01	1.03	0.91	0.96	0.91*	1.01	1.13
Homeowner (ref: renter)								
Homeowner	1.04***	0.99*	0.96*	1.15***	1.12***	1.06*	1.06**	1.29***
Duration dependence								
γ_{ij}	-0.69***	-1.45***	-0.30***	-2.05***	-0.23***	-1.45***	0.01***	-2.56***
Individuals	452,678	592,030	52,687	78,012	109,297	245,401	34,203	137,897
Spells	635,791	829,767	62,357	86,727	120,625	279,248	36,056	145,327
Transition probability (%)	48	23	49	4	34	5	62	1
Sub-log-likelihood (cf. (2.8))[1]	-271,672.19	-268,988.78	-35,590.79	-10,588.47	-73,004.06	-38,217.66	-31,560.28	-6,835.64
Log-likelihood[2]	-614,999.87	-1,117,533.80	-99,530.56	-93,735.75	-188,224.73	-344,270.82	-62,589.94	-199,783.93

Notes: The estimates are hazard ratios $\exp(\beta_{ij})$ cf. equation (2.5). However, frailty is restricted to be absent in the model ($\sigma^2_{ij} = 0$). A hazard ratio that exceeds unit value implies higher risk, and vice versa. *5-% **1-% ***0.1-% significance level. Null hypotheses: $\beta_{ij} = 0$ (Wald test); $\gamma_{ij} = 0$ (Wald test). Additional controls: gender, year of observation, region of residence, medication use per category, age of entry, and ethnicity. Results for asset and income deciles 2 to 9 are suppressed from the table. [1] The sub-log-likelihood is based on marginal survival probability (2.8). [2] The log-likelihood is the sum of sub-log-likelihoods of the set of competing risks.

B.6 Correlated Frailty and Competing Risks

Suppose that we are interested in two competing event times of somebody who is currently alive without LTC: 'start of LTC' (S) and 'Death' (D). Our analysis stops whichever event occurs first.

Actual event times are latent variables. Assume that α is the sole covariate that determines both event times, and α is unobserved. We follow subjects till time $t^*(\alpha)$, depending on whichever transition comes first. Suppose transition S would actually take place at time t^*_S and transition D at time t^*_D. The observation scheme is:

$$t^*(\alpha) = \begin{cases} t^*_D(\alpha) & \text{if } t^*_D(\alpha) < t^*_S(\alpha) \\ t^*_S(\alpha) & \text{if } t^*_D(\alpha) \geq t^*_S(\alpha), \end{cases}$$

The survival probabilities that draw our interest, are:

$$S_D(t^*(\alpha)|\alpha) = \mathbb{P}(t > t^*(\alpha) \mid t^*_D(\alpha) < t^*_S(\alpha), \alpha)$$
$$S_S(t^*(\alpha)|\alpha) = \mathbb{P}(t > t^*(\alpha) \mid t^*_D(\alpha) \geq t^*_S(\alpha), \alpha),$$

which are the duration distributions conditional upon that event D or S occurs first.

To estimate $\widehat{S_D}(t|\alpha)$ we can naively apply survival function estimation with right censoring. We would treat an observation 'right-censored' at time $t^*(\alpha)$ if $t^*_D(\alpha) \geq t^*_S(\alpha)$, i.e., event S occurs first. The log-likelihood would be as follows:

$$\mathcal{L}(t^*) = log\left\{\widehat{S_D}(t^*(\alpha))\right\} \cdot \mathbf{1}(t^*_D(\alpha) \geq t^*_S(\alpha)) + log\left\{-\frac{\partial \widehat{S_D}(t^*(\alpha))}{\partial t}\right\} \cdot \mathbf{1}(t^*_D(\alpha) < t^*_S(\alpha)). \tag{B.1}$$

Note that treating $\mathbf{1}(t^*_D(\alpha) \geq t^*_S(\alpha))$ as random right censoring ignores its dependence on α. Instead, the event time $t^*(\alpha)$ is correlated with the censoring mechanism. The survival functions of $t^*_D(\alpha)$ and $t^*_S(\alpha)$ must be estimated jointly.

Shared Unobserved α over the Transitions

Suppose we assume a mixed proportional hazard specification for transition S and transition D, no covariates, and exponential hazard for duration dependence. The researcher knows the frailty distribution and will update the log-likelihood function accordingly. The survival distributions (conditional upon α) are (van den Berg, 2001):

$$S_D(t|\alpha) = exp(-\alpha\lambda_D t) = (exp(-\lambda_D t))^\alpha = \left(\widetilde{S_D}(t)\right)^\alpha \qquad S_S(t \mid \alpha) = exp(-\alpha\lambda_S t) = (exp(-\lambda_S t))^\alpha = \left(\widetilde{S_S}(t)\right)^\alpha$$

where λ_S and λ_S are scalar parameters that induce the exponential hazard. $\widetilde{S_D}(t)$ and $\widetilde{S_D}(t)$ the survival probabilities when $\alpha = 1$. The individual random effect α scales these survival probabilities up or down, i.e., the impact of frailty on survival probabilities. The last step holds for any baseline hazard, so that we can generalize the following results.

For log-likelihood estimation we are interested in the joint survival probability. The conditional joint survival probabilities are:

$$S(t_1, t_2|\alpha) = \left(\widetilde{S_D}(t_1)\right)^\alpha \cdot \left(\widetilde{S_S}(t_2)\right)^\alpha$$

The unconditional joint survival probabilities are:

$$S(t_1, t_2) = \int S(t_1, t_2|\alpha)dG(\alpha) = \int \left(\widetilde{S_D}(t_1)\right)^\alpha \cdot \left(\widetilde{S_S}(t_2)\right)^\alpha dG(\alpha) = \int \left(\widetilde{S_D}(t_1) \cdot \widetilde{S_S}(t_2)\right)^\alpha dG(\alpha) \tag{B.2}$$

where G is the distribution function of α. The joint log-likelihood function becomes:

$$\mathcal{L}(t^*) = log\left(\frac{-\partial S(t^*, t^*)}{\partial t_1} \cdot \mathbf{1}(t_D^* \leq t_S^*)\right) + log\left(\frac{-\partial S(t^*, t^*)}{\partial t_2}\right) \cdot \mathbf{1}(t_S^* < t_D^*)$$

Because of shared unobserved heterogeneity α in (B.2), this likelihood is not separable in transition D and S. Hence, we cannot estimate the specification for S and D apart.

Unobserved Term Independent across Transitions

Now, suppose that the unobserved term α determines D and the unobserved term β determines S. α and β are independent. Denote F and H the respective distribution of α and β. The joint survival probability is now separable in D and S:

$$\begin{aligned}
S(t_1,t_2) &= \int\int S(t_1,t_2|\alpha,\beta)dF(\alpha)dH(\beta) \\
&= \int\int \left(\widetilde{S_D}(t_1)\right)^{\alpha} \cdot \left(\widetilde{S_S}(t_2)\right)^{\beta} dF(\alpha)dH(\beta) \\
&= \int \left(\widetilde{S_D}(t_1)\right)^{\alpha} dF(\alpha) \int \left(\widetilde{S_S}(t_2)\right)^{\beta} dH(\beta) \\
&= \mathbb{E}_{\alpha}\left(\widetilde{S_D}\left((t_1)\right)^{\alpha}\right) \cdot \mathbb{E}_{\beta}\left(\widetilde{S_S}\left((t_2)\right)^{\beta}\right).
\end{aligned}$$

Also, the log-likelihood becomes separable in risks S and D:

$$\begin{aligned}
\mathcal{L}(t^*) = log&\left(\frac{-\partial S(t^*,t^*)}{\partial t_1}\right) \cdot \mathbf{1}(t_D^* \leq t_S^*) + log\left(\frac{-\partial S(t^*,t^*)}{\partial t_2}\right) \cdot \mathbf{1}(t_S^* < t_D^*) \\
&= log\left(\frac{-\partial \mathbb{E}_{\alpha}\left(\left(\widetilde{S_D}\left(t^*\right)\right)^{\alpha}\right)\mathbb{E}_{\beta}\left(\left(\widetilde{S_S}\left(t^*\right)\right)^{\beta}\right)}{\partial t_1}\right) \cdot \mathbf{1}(t_D^* \leq t_S^*) \\
&+ log\left(\frac{-\partial \mathbb{E}_{\alpha}\left(\left(\widetilde{S_D}\left(t^*\right)\right)^{\alpha}\right)\mathbb{E}_{\beta}\left(\left(\widetilde{S_S}\left(t^*\right)\right)^{\beta}\right)}{\partial t_2}\right) \cdot \mathbf{1}(t_D^* > t_S^*) \\
&= log\left(\mathbb{E}_{\beta}\left(\left(\widetilde{S_S}\left(t^*\right)\right)^{\beta}\right) \cdot \frac{-\partial \mathbb{E}_{\alpha}\left(\left(\widetilde{S_D}\left(t^*\right)\right)^{\alpha}\right)}{\partial t_1}\right) \cdot \mathbf{1}(t_D^* \leq t_S^*) \\
&+ log\left(\mathbb{E}_{\alpha}\left(\left(\widetilde{S_D}\left(t^*\right)\right)^{\alpha}\right) \cdot \frac{-\mathbb{E}_{\beta}\left(\left(\widetilde{S_S}\left(t^*\right)\right)^{\beta}\right)}{\partial t_2}\right) \cdot \mathbf{1}(t_D^* > t_S^*) \\
&= log\left(\mathbb{E}_{\alpha}\left(\left(\widetilde{S_D}\left(t^*\right)\right)^{\alpha}\right)\right) \cdot \mathbf{1}(t_D^* > t_S^*) + log\left(\frac{-\mathbb{E}_{\alpha}\left(\left(\widetilde{S_D}\left(t^*\right)\right)^{\alpha}\right)}{\partial t_1}\right) \cdot \mathbf{1}(t_D^* \leq t_S^*) \\
&+ log\left(\mathbb{E}_{\beta}\left(\left(\widetilde{S_S}\left(t^*\right)\right)^{\beta}\right)\right) \cdot \mathbf{1}(t_D^* \leq t_S^*) + log\left(\frac{-\partial \mathbb{E}_{\beta}\left(\left(\widetilde{S_S}\left(t^*\right)\right)^{\beta}\right)}{\partial t_2}\right) \cdot \mathbf{1}(t_D^* > t_S^*)
\end{aligned}$$

The first two terms define the log-likelihood contribution due to risk D, similar to equations (2.7) and (B.1). In this particular case, we may treat the occurrence of the other event as random censoring and estimate the specifications for S and D separately.

Parameter Identification of the Mixed Proportional Hazard Model

While we refer to Chapter 5 for the details on the full likelihood estimation, we can here provide heuristic arguments for parameter identification (for an overview, see: van den Berg, 2001). For identification, we resort to variation in hazard and unconditional transition probabilities between individuals and within individuals. First, we observe multiple spells for some individuals, e.g., repeated home-based care use. At time t, ν and $\phi_{ij}(t)$ are the same for both spells, but $\boldsymbol{x}(t)$ can vary across spells. Then hazard rates differ solely due to $\boldsymbol{x}(t)'\boldsymbol{\beta}_{ij}$, and the comparison reveals $\boldsymbol{\beta}_{ij}$. Instead, if $\boldsymbol{x}(t)$ is constant across spells, the repeated event helps identifying $\phi_{ij}(t)$. Namely, the only difference between their unconditional transition probabilities stems from the hazard component $\phi_{ij}(t)$. Second, we have time-varying covariates. Suppose we have two groups, one with $\boldsymbol{x}(t) = \overline{\boldsymbol{x}}$ at any time and another with $\boldsymbol{x}(t) = \overline{\boldsymbol{x}}$ for $t < t^*$ and $\boldsymbol{x}(t) = \widetilde{\boldsymbol{x}} \neq \overline{\boldsymbol{x}}$ for $t \geq t^* > 0$. Then, a difference in their unconditional transition probabilities can be fully attributed to the break in $\boldsymbol{x}(t)$ which, in turn, identifies $\boldsymbol{\beta}_{ij}$. Lastly, $\boldsymbol{x}(t)$ varies between individuals. $\boldsymbol{\beta}_{ij}$ is identified with a similar reasoning as repeated spells but now by comparing distinct individuals under more stringent assumptions (Elbers and Ridder, 1982). Our choice for the Gamma distribution ensures the stricter assumptions are met and guarantees identification of σ^2_{ij}.

Appendix C: Chapter 3

C.1 Constructing a Measure for Lifetime Income

To compute a measure for lifetime income for the households, we follow Knoef et al. (2016). Their approach allows us to include annuity value of household's financial assets. Some households have low income but many assets, e.g., former entrepreneurs, making it indispensable to include the annuity income from financial assets in a lifetime income definition. We measure lifetime income as the average income during retirement plus the annuity value of financial assets.

We use the population tax files on income (2003-2014) and assets (2006-2014). Assets comprise the sum of savings and stock and bond holdings, but exclude home ownership because this is strongly correlated with not being in a nursing home (read: long-term care). Income is measured at the household level, including labor and business income, retirement income (social security benefits, employer-based, and private pension arrangements), social insurance benefits, taxes, and social insurance contributions. Income predominantly consists of retirement income, because we restrict households to have this as their main source of income.

Yet, we do not observe the annuity value of assets, B, which we will therefore impute. We assume that a household bought an annuity when the oldest member was 65. If available, the other member might be younger than 65. The price of the annuity equals the household's current assets A.[1] The annuity yearly pays B if it is a single-person household and $\sqrt{2} \cdot B$ if it is a couple household. $\sqrt{2}$ is an equivalence scale (OECD, 2011) that the OECD applies when comparing income between single and couple households.

The product is actuarially fair: the benefit level B is set such that the expected lifetime

[1] We implicitly assume that households do not save or dis-save after retirement.

benefits and current assets A are equal. Hence the benefit level B is household-specific. Expected lifetime benefits look as follows:

$$\mathbb{E}(Benefit(B)) = \begin{cases} \sum_{n=0}^{99} \frac{1}{(1+r)^n} \left(B \cdot {}_n s_m\right) & \text{if Single man at age 65} \\ \sum_{n=0}^{99} \frac{1}{(1+r)^n} \left(B \cdot {}_n s_w\right) & \text{if Single woman at age 65} \\ \sum_{n=0}^{99} \frac{1}{(1+r)^n} \Big(B \cdot {}_n s_m \cdot (1 - {}_n s_w) + B \cdot {}_n s_w \cdot (1 - {}_n s_m) & \\ \qquad + \sqrt{2} \cdot B \cdot {}_n s_m \cdot {}_n s_w\Big) & \text{if Married couple at age 65,} \end{cases}$$

where n refers to the years since the oldest household member turned 65. ${}_n s_m$ and ${}_n s_w$ are the probabilities that the man or woman in the household survives n years after buying the annuity. $1 - {}_n s_m$ and $1 - {}_n s_w$ are the probabilities that the man or woman died within n years after buying the annuity. The probabilities are gender-, cohort-, and age-specific, and taken from the life tables of Statistics Netherlands.[2] These probabilities are age-specific because couple members might have a different age when the household buys the annuity.

The expected benefits $\mathbb{E}(Benefit(B))$ are the sum of benefits expected in each period. We assume a maximum benefit payout period of 99 years, the length of the life tables. Benefits are deflated using an assumed yearly interest rate $r = 0.02$. Focusing on the case of a single man at age 65, the expected benefit in period n is the product of the household-specific benefit and the probability of being this household type in period n, ${}_n s_w$. Likewise for a single woman. The case for couples is more complex. Households are a single man with probability ${}_n s_m \cdot (1 - {}_n s_w)$, i.e. the man survived until period n while the woman has died. Benefits are scaled up by $\sqrt{2}$ in case of couples, which happens with probability ${}_n s_m \cdot {}_n s_w$, i.e. the man and woman both survive. Survival is assumed to be independent across household members.

The annuity benefit, i.e. annuity value of assets, is found by solving $A = \mathbb{E}(Benefit(B))$ for B. Because assets vary each year in the data, the benefit B is time-varying within a

[2]See: `https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37360ned/table?fromstatweb` [*Retrieved on:* February 18^{th} 2022]

household.

Household's lifetime income is the average sum of household income and the imputed annuity value of assets. However, the size of the household might change during these years, and differs across households. Then, married couples by definition would have high lifetime income. To tackle this problem, we equivalize household's income with equivalence scale $\sqrt{2}$ so to make couples and singles comparable in terms of their income (cf. Attanasio and Emmerson, 2003). Formally, we calculate lifetime income PI_i of household i as follows:

$$PI_i = \frac{\sum_{\tau=1}^{N_i} B_{i\tau} + \frac{y_{i\tau}}{\sqrt{2}} \cdot marstat_{i\tau} + y_{i\tau} \cdot (1 - marstat_{i\tau})}{N_i},$$

where $y_{i\tau}$ is household income in year τ, $B_{i\tau}$ annuity value of assets, N_i the number of panel observations of household i and $marstat_{i\tau}$ an indicator on whether the household is a couple or single person.

C.2 Simulation Procedure

C.2.1 Log-Likelihood Estimation of the Hazard Rates

Suppose we want to estimate the unknown parameters $\boldsymbol{\gamma}_k$, $\boldsymbol{\beta}_k$, and σ_k of the hazard rate $\lambda_k(t, marstat(t); \nu^k, \boldsymbol{\gamma}_k, \boldsymbol{\beta}_k)$, specified in (3.8). We will apply a log-likelihood estimation procedure to estimate the parameters of transition k. We will derive the probability distribution that is input for the individual log-likelihood contribution (we drop index i). To further save on notation, we drop $\boldsymbol{\gamma}_k$ and $\boldsymbol{\beta}_k$; our examples refer to an individual with a given initial marital status, lifetime income group and gender.

Before we derive the probability distribution of interest, we have to discuss the implications of our competing risk setting. Essentially, two transitions are possible at any age, and one will preclude the other from actually occurring. For example, No Long-term Care use $\rightarrow$ Death happens at random age $T_D = t^*$ while No Long-term Care use $\rightarrow$ Long-term Care use would happen at random age $T_L = t^{**} > t^*$. We want to estimate the distribution (transition rate) of both T_L and T_D. Note that the researcher knows $T_L \geq t^*$ but $T_L = t^{**} > t^*$ is hidden information. The competing risks require a log-likelihood function involving the joint distribution of the observed event: $\mathbb{P}(T_D = t^*, T_L \geq t^* \mid \nu^L, \nu^D)$. This distribution simplifies because we assume random effects to be independent across transitions, i.e., $\nu^D \perp \nu^L$: $\mathbb{P}(T_D = t^*, T_L \geq t^* \mid \nu^L, \nu^D) = \mathbb{P}(T_D = t^* \mid \nu^D) \cdot \mathbb{P}(T_L \geq t^* \mid \nu^L)$. Like the distribution function, the likelihood function will split into two sub-likelihoods and we can estimate the transition rates with separate regressions, each for a transition k. The event time T_L would be modelled as randomly right-censored at t^* ($\mathbb{P}(T_L \geq t^* | \ \nu_L)$).

To explain the estimation of a single transition, we look at an example of an individual with two spells of type k. The first spell starts at age $t_{0,1} > 0$, implying a left-truncated observation, for example, because the individual is older than 65 when entering the sample. The other spell starts at age $t_{0,2} > t_{0,1} > 0$. The spells end at ages $t_{0,1} < t_1 < t_{0,2}$ and $t_2 > t_{0,2}$, meaning the first spell ends before the next spell starts. The log-likelihood is based on the joint survival probability of staying in the state until ages t_1 and t_2, given you entered the state at ages $t_{0,1}$ and $t_{0,2}$. As we will show below, the hazard rate (3.8) fully characterizes the distribution T, the random age at transition.

Besides left truncation, our estimation also considers that marital status is a time-varying covariate. In the example, we assume that the individual is married during spell 1, i.e. $marstat(t) = 1$ if $t \leq t_1$. The individual becomes widowed during spell 2 at age t_w: $t_{0,2} < t_w < t_2$, so $marstat(t) = 1$ if $t < t_w < t_2$ and $marstat(t) = 0$ if $t > t_w$.

The first ingredient to construct the log-likelihood is to have the integrated hazard rate m_k, i.e. the transition rate on having made a transition between age 0 and t:

$$\begin{aligned} m_k(t; \nu^k, marstat = x) &= \int_0^t \lambda_k(\tau, marstat = x; \nu^k, \boldsymbol{\gamma}_k, \boldsymbol{\beta}_k) d\tau \\ &= \nu^k \cdot \int_0^t \lambda_k(\tau, marstat = x; \nu^k = 1, \boldsymbol{\gamma}_k, \boldsymbol{\beta}_k) d\tau \\ &= \nu^k \cdot m_k(t; \nu^k = 1, marstat = x), \end{aligned} \tag{C.1}$$

where we can go from step 1 to steps 2 and 3 because the hazard rate is proportional in ν^k. The alternative representations using $\lambda_k(\tau, marstat = x; \nu^k = 1, \gamma_k, \boldsymbol{\beta}_k)$ and $m_k(t; \nu^k = 1, marstat = x)$ have a closed-form solution (see: Bender et al., 2005) and make it easier to derive a closed-form solution for the log-likelihood contribution.

The marital status in (C.1) is assumed to have the fixed value $x = \{0, 1\}$ between age 0 and t, i.e. marital status is time-invariant. The accumulated hazards m_k at the left-truncation points $t = t_{0,1}$ and $t = t_{0,2}$ and end age $t = t_1$ are defined according to (C.1) because marital status only changes after these ages: $t_w > t_{0,2}$. The definition of accumulated hazard at age t_2, however, differs because marital status changes at $t_w < t_2$:

$$
\begin{aligned}
m_k(t_2; \nu^k, \{marstat(s)\}_{s=t_{0,2}}^{t_2}) &= m_k(t_w; \nu^k, marstat = 1) \\
&\quad + m_k(t_2; \nu^k, marstat = 0) - m_k(t_w; \nu^k, marstat = 0)
\end{aligned}
$$

where $\{marstat(s)\}_{s=t_{0,2}}^{t_2}$ denotes the covariate path of marital status between age $t_{0,2}$ and t_2. The accumulated hazard consists of the sum of hazard until t_w when married ($marstat = 1$) plus the hazard accumulated between t_w and t_2 when single ($marstat = 0$).

The joint survival probability of not having made the transition until ages t_1 and t_2 is linked to the integrated hazard rates is:

$$
\begin{aligned}
&\mathbb{P}_k(T_1 > t_1, T_2 > t_2 \mid \{marstat(s)\}_{s=0}^{t_2}, \nu^k) \\
&= \exp\left(-\left\{m_k(t_1; \nu^k, marstat = 1) + m_k(t_2; \nu^k, \{marstat(s)\}_{s=t_{0,2}}^{t_2})\right\}\right),
\end{aligned}
$$

which is the exponential function where the negative sum of accumulated hazards serves as input (see: Bender et al., 2005).

For the left truncation points, we can do the same, i.e. the survival probability of not having made the transition by ages $t_{0,1}$ and $t_{0,2}$:

$$\begin{aligned}&\mathbb{P}_k(T_1 > t_{0,1}, T_2 > t_{0,2} \mid \nu^k, \{marstat(s)\}_{s=0}^{t_{0,2}})\\&= \exp\left(-\left\{m_k\left(t_{0,1}; \nu^k, marstat = 1\right) + m_k\left(t_{0,2}; \nu^k, marstat = 1\right)\right\}\right).\end{aligned}$$

The log-likelihood contribution is based on the joint survival probability of staying in the state until ages t_1 and t_2, given you entered the state at ages $t_{0,1}$ and $t_{0,2}$:

$$\mathbb{P}_k(T_1 > t_1, T_2 > t_2 \mid T_1 > t_{0,1}, T_2 > t_{0,2}, \cdot, \nu^k) = \frac{\mathbb{P}_k(T_1 > t_1, T_2 > t_2 \mid \cdot, \nu^k)}{\mathbb{P}_k(T_1 > t_{0,1}, T_2 > t_{0,2} \mid \cdot, \nu^k)},$$

where for notational convenience we replace the marital histories by a dot $\cdot$.

Lastly, we back out the random effect ν_k, which we do by integrating over its distribution:

$$\begin{aligned}&\mathbb{P}_k(T_1 > t_1, T_2 > t_2 \mid T_1 > t_{0,1}, T_2 > t_{0,2}, \cdot) \qquad \text{(C.2)}\\&= \int_0^\infty \frac{\mathbb{P}_k(T_1 > t_1, T_2 > t_2 \mid \nu^k, \cdot)}{\mathbb{P}_k(T_1 > t_{0,1}, T_2 > t_{0,2} \mid \nu^k, \cdot)} \,\mathrm{d}\Gamma(\nu^k \mid T_1 > t_{0,1}, T_2 > t_{0,2}, \cdot)\\&= \frac{\int_0^\infty \mathbb{P}_k(T_1 > t_1, T_2 > t_2 \mid \nu^k, \cdot)\,\mathrm{d}\Gamma(\nu^k)}{\int_0^\infty \mathbb{P}_k(T_1 > t_{0,1}, T_2 > t_{0,2} \mid \nu^k, \cdot)\,\mathrm{d}\Gamma(\nu^k)}\\&= \frac{\left(\sigma_k^2 \cdot \left\{m_k(t_1; \nu^k = 1, marstat = 1) + m_k(t_2; \nu^k = 1, \{marstat(s)\}_{s=t_{0,2}}^{t_2})\right\} + 1\right)^{-\frac{1}{\sigma_k^2}}}{\{\sigma_k^2 \cdot (m_k(t_{0,1}; \nu^k = 1, marstat = 1) + m_k(t_{0,2}; \nu^k = 1, marstat = 1)) + 1\}^{-\frac{1}{\sigma_k^2}}},\end{aligned}$$

where the final closed-form expression is the probability distribution we use to construct the individual log-likelihood contribution (for the derivation, see Chapter 5). The first step – where we integrate over the conditional distribution of the random effect– reflects dynamic selection. Only a particular share of the initial population survives until these dates, presumably driven by their favorable random effect. Hence, the left-truncated distribution deviates from the initial distribution $\Gamma(\nu^k)$. The second step uses the initial distribution instead, see van den Berg and Drepper (2016) for the justification. The last step arrives at the closed-form solution because m_k analyzed at $\nu^k = 1$ has a closed-form solution itself (see Bender et al., 2005, for the solution of m_k for the Gompertz case).

Note that the current case involves right censoring. We here provided the cumulative probability of staying in a state until a particular age. This refers to the case when we stop observing the individual at ages t_1 and t_2 while the actual transition is not yet made, e.g. due to the end of the observational window or realization of a competing risk (right censoring). Instead, the log-likelihood contribution involves a probability density if the individual actually makes the transition. This is done by taking the derivative of the probability distribution (C.2) with respect to random variable T_1 or T_2 and subsequently multiplying the derivative by -1 (to accommodate that we want a cumulative distribution function, i.e. $<$, instead of a survival function, i.e. $\geq$ probabilities). Chapter 5 provides the log-likelihood contribution for a general case of n spells of an individual.

A final remark involves the value of the log-likelihood function. The survival probability (C.2) involves only transition k but not its competing risk, hence the accompanying log-likelihood is a sub-log-likelihood, particular for transition k. If we add the log-likelihood for the competing risk to this, we obtain the overall likelihood that we effectively maximize. As said, the two sub-log-likelihoods can be optimized separately because the unobservable (random) effect is assumed to be uncorrelated across transitions.

We refer to Honoré (1993) and van den Berg (2001) and the references therein for parameter identification.

C.2.2 Simulation

We use estimated hazard specifications (3.8) and (C.1) to the simulate lifetime duration of long-term care use and the timing of death for 100,000 households. Households initially consist a couple of two members or a single member aged 65 years old. Denote the age of entering the current state by t_0, where $t_0 = 0$ means entry at age 65. We are interested in the subsequent state (not using long-term care, using long-term care, or death) and at what random age $T > t_0$ this transition occurs. We repeat looking for the next state until every individual has died. Finally, we have for each individual a sequence of consecutive states and age at which these states start.

With slight abuse of notation, let the estimates for the integrated hazard rates (C.1) be denoted by $\widehat{m}_k(t;\nu^k = 1, marstat = x) = \widehat{m}_{k,x}(t)$. $\widehat{m}_{k,x}(t)$ refers to an individual with current marital status x who is endowed with a gender, initial marital status, and lifetime income group. Hence, $\widehat{m}_{k,x}(t)$ can differ across individuals. For now we assume x is fixed during life, i.e. we assume initially married individuals to be currently married and assume that they stay married until they die ($x = 1$). Initial singles remain unmarried throughout ($x = 0$). We introduce widowhood later.

Timing of transition k We use $\widehat{m}_{k,x}(t)$ to compute when a transition of type k, e.g. No Long-term Care use → Death, would take place. To this end, we draw a transition time from a conditional survival probability like (C.2): Given that the individual entered the state at age $T > t_0$, the transition k does not occur before age $T > t > t_0$. This gives:[3]

$$\mathbb{P}_k(t \mid t_0, x) = \mathbb{P}(T > t \mid T > t_0, x, k \text{ occurs}) = \frac{\left(\widehat{\sigma_k^2} \cdot \widehat{m}_{k,x}(t) + 1\right)^{-\frac{1}{\widehat{\sigma_k^2}}}}{\left(\widehat{\sigma_k^2} \cdot \widehat{m}_{k,x}(t_0) + 1\right)^{-\frac{1}{\widehat{\sigma_k^2}}}} \sim \mathcal{U}(0,1)$$

Related to our case, Bender et al. (2005) provide the closed-form solution of $\widehat{m}_{k,x}(t)$ when the baseline hazard is of Gompertz form.

The key to the simulation is that survival probability $\mathbb{P}_k(t \mid t_0, x)$ is uniformly distributed itself. Suppose we randomly generate $u \in U(0,1)$ and let $\mathbb{P}_k(t \mid t_0, x) = u$. The value t for which the equation holds, is a randomly generated age t_k at which transition k occurs:

$$t_k = \{\widehat{m}_{k,x}\}^{-1}(\underline{t}), \text{ with: } \underline{t} = \frac{1}{\widehat{\sigma_k^2}} \cdot \left\{u^{-\widehat{\sigma_k^2}} \cdot \left\{\widehat{\sigma_k^2} \cdot \widehat{m}_{k,x}(t_0) + 1\right\} - 1\right\}.$$

[3]Alternatively, we endowed individuals with an individual-specific effect according to $\widehat{\Gamma}_k$ and subsequently simulated their long-term care use and mortality. Our current approach fits age-specific mortality rates and long-term care use rates better.

Hence, we have a closed-form solution to simulate age t_k when transition k would occur.

Our simulation considers that other transitions are possible, i.e. 'Not using Long-term Care $\rightarrow$ using Long-term Care', that might preclude the transition 'Not Using Long-term Care $\rightarrow$ Death' from occurring. We generate a random age t_k for each possible transition. The minimum across these ages defines the next state and the value for t_0 with which we continue the simulation. We end the simulation if the next state is death.

Widowhood So far we assumed that initially married individuals remain married until death. However, one of the two couple members will die first, and the surviving household member becomes single. Becoming single affects the hazard rate $\widehat{m}_{k,x}$ and thereby thus the timing of a transition. While transition paths before widowhood remain unchanged, we modify the simulated transitions for surviving partner after he or she has become widowed. Remarriage after widowhood is not possible.

For this, we distinguish two types of transitions. First, we look at the transition that is the first to occur after widowhood time t_w. If the individual remained married, the transition would take place at simulated age $t_{k,\text{orig}}$. The individual's accumulated hazard is $\widehat{m}_{k,x=1}(t_{k,\text{orig}})$, which is a counterfactual. The true accumulated hazard is the accumulated hazard until widowhood $\widehat{m}_{k,x=1}(t_w)$ complemented with the hazard while being single: $\widehat{m}_{k,x=0}(t_k) - \widehat{m}_{k,x=0}(t_w)$. To incorporate a widowhood effect to t_k, we set the counterfactual and true hazard equal and solve for t_k:

$$
\begin{aligned}
&\widehat{m}_{k,x=1}(t_{k,\text{orig}}) = \widehat{m}_{k,x=1}(t_w) + \widehat{m}_{k,x=0}(t_k) - \widehat{m}_{k,x=0}(t_w) \rightarrow \\
&\qquad t_k = \{\widehat{m}_{k,x=0}\}^{-1}(\underline{t}), \text{ with: } \underline{t} = \widehat{m}_{k,x=1}(t_{k,\text{orig}}) - \widehat{m}_{k,x=1}(t_w) + \widehat{m}_{k,x=0}(t_w).
\end{aligned}
$$

Like earlier, the minimum age across possible transitions determines the next state.

All spells that start after widowhood ($t_0 > t_w$) have a survivor probability as follows:

$$
\frac{(\widehat{\sigma_k^2} \cdot \{\widehat{m}_{k,x=1}(t_w) + \widehat{m}_{k,x=0}(t) - \widehat{m}_{k,x=0}(t_w)\} + 1)^{-\frac{1}{\widehat{\sigma_k^2}}}}{(\widehat{\sigma_k^2} \cdot \{\widehat{m}_{k,x=1}(t_w) + \widehat{m}_{k,x=0}(t_0) - \widehat{m}_{k,x=0}(t_w)\} + 1)^{-\frac{1}{\widehat{\sigma_k^2}}}} \sim \mathcal{U}(0,1),
$$

and the simulated age at transition is:

$$t_k = \{\widehat{m}_{k,x=0}\}^{-1}(\underline{t}), \text{ with:}$$
$$\underline{t} = \frac{1}{\widehat{\sigma_k^2}} \cdot \{u^{-\widehat{\sigma_k^2}} \cdot \{(\widehat{\sigma_k^2} \cdot \{\widehat{m}_{k,x=0}(t_0) - \widehat{m}_{k,x=0}(t_w) + \widehat{m}_{k,x=1}(t_w)\} + 1\} - 1\}$$
$$+ \widehat{m}_{k,x=0}(t_w) - \widehat{m}_{k,x=1}(t_w).$$

Initialization We endow households with initial marital status, long-term care use, and lifetime income according to the empirical distribution of households when the members are aged 65. Sample sizes are provided in Table C.1.

Table C.1: Initial Household Distribution in Simulation ($N = 100,000$)

	No LTC	**Man in LTC**	**Woman in LTC**	**Both in LTC**	**All**[2]	**Share** (%)
All	96,193	1,716	2,047	44	100,000	100.0
Couple						
Bottom Lifetime IQ	3,399	61	71	12	3,543	6.3
2nd Lifetime IQ	7,285	79	102	12	7,478	13.2
3rd Lifetime IQ	12,004	96	121	11	12,232	21.6
4th Lifetime IQ	15,046	92	99	6	15,243	27.0
Top Lifetime IQ	17,868	87	89	3	18,047	31.9
All	55,602	415	482	44	56,543	100.0
Single men						
Bottom Lifetime IQ	2,324	435			2,759	17.4
2nd Lifetime IQ	1,731	327			2,058	13.0
3rd Lifetime IQ	2,505	260			2,765	17.5
4th Lifetime IQ	3,666	177			3,843	24.3
Top Lifetime IQ	4,297	102			4,399	27.8
All	14,523	1,301			15,824	100.0
Single women						
Bottom Lifetime IQ	6,530		751		7,281	26.3
2nd Lifetime IQ	3,938		337		4,275	15.5
3rd Lifetime IQ	4,615		223		4,838	17.5
4th Lifetime IQ	5,542		162		5,704	20.6
Top Lifetime IQ	5,443		92		5,535	20.0
All	26,068		1,565		27,633	100.0

Notes: This table shows the sample distribution at age 65. Long-term care use is measured when the household member is aged 65, also when there is an age difference between couple members. IQ = Income Quintile. [1] The share of an income quintile is not exactly 20% because the lifetime income distribution is determined by all households instead of only those who were aged 65 during the sampling period. [2] The total number of simulated households is 100,000, for which we provide the counts in this table. The actual number of households in the data was higher.

C.3 Demand Curves, WTPs, and Comparative Statics

First, we derive the incentive compatibility constraints, demand curves and willingness-to-pay (WTP) for buying an annuity, LTC insurance, respectively. We have the utility function:

$$V() = U(C_1) + (s(\xi) - l(\xi))U(C_2^h) + l(\xi)U(C_2^l)$$

Let $L = 1$ if consumer buys a LTC insurance; 0 otherwise; $A = 1$ if consumer buys annuity insurance, 0 otherwise. We assume no saving, so $C_1 = W_1 - P_A A - P_L L$; $C_2^h = W_2 + Y \cdot A$; $C_2^l = W_2 + Y \cdot A - X \cdot (1 - L)$. Substitution of these equalities yields the following direct utility function:

$$\begin{aligned} V(A, L; W_1, W_2, X, Y, P_L, P_A, \xi) &= U(W_1 - P_A A - P_L L) \\ &+ (s(\xi) - l(\xi))U(W_2 + Y \cdot A) \\ &+ l(\xi)U(W_2 + Y \cdot A - X \cdot (1 - L)) \end{aligned}$$

A consumer buys LTC insurance if:

$$V(A^*, 1; W_1, W_2, X, Y, P_L, P_A, \xi) - V(A^*, 0; W_1, W_2, X, Y, P_L, P_A, \xi) \geq 0,$$

i.e. the utility when insured exceeds that of being uninsured. Demand for LTC insurance $D_L(P_L|A^*, W_1, W_2, X, Y, P_A)$ is the likelihood that this inequality holds:

$$\begin{aligned} D_L(P_L|\cdot) = \mathbb{P}\Big(&- (U(W_1 - P_A A^*) - U(W_1 - P_A A^* - P_L)) \\ &+ l(\xi)(U(W_2 + YA^*) - U(W_2 + YA^* - X)) > 0\Big) \\ &= \mathbb{P}(IC_L(A^*, W_1, W_2, X, Y, P_L, P_A, \xi) > 0), \end{aligned} \tag{C.3}$$

where IC_L is short-hand notation for the left hand side of the incentive compatibility constraint for LTC insurance. Notice that IC_L (a utility difference) is strictly decreasing in P_L. Therefore, demand $D_L(P_L|\cdot)$ will be strictly decreasing in P_L. We can meaningfully define the WTP as follows:

$$\pi_L(A^*, W_1, W_2, X, Y, P_A, \xi) = \max\{P_L; IC_L(A^*, W_1, W_2, X, Y, P_L, P_A, \xi) \leq 0)\}$$

$\pi_L(\cdot)$ can be solved from the following implicit equation:

$$IC_L(A^*, W_1, W_2, X, Y, \pi_L(\cdot), P_A, \xi) = 0$$

A consumer buys stand alone annuity insurance if $V(1, L^*, W_1, W_2, X, Y, P_L, P_A, \xi) - V(0, L^*, W_1, W_2, X, Y, P_L, P_A, \xi) \geq 0$. The demand curve is the probability that this incentive compatibility constraint holds:

$$\begin{aligned} D_A(P_A|L^*, W_1, W_2, X, Y, P_L) = \mathbb{P}\Big(&- (U(W_1 - P_L L^*) - U(W_1 - P_A - P_L L^*)) \\ &+ s(\xi)(U(W_2 + Y) - U(W_2)) + (1 - L^*)l(\xi)LL(\cdot) > 0\Big) \\ = \mathbb{P}(&IC_A(L^*, W_1, W_2, X, Y, P_L, P_A, \xi) > 0) \end{aligned} \tag{C.4}$$

with $LL(W_2, Y, X) = ((U(W_2 + Y - X) - U(W_2 + Y)) - (U(W_2 - X) - U(W_2)))$. Since $U()$ is strictly concave and $W_2 > 0$, $Y > 0$ and $X > 0$, $LL(W_2, Y, X) > 0$.

Notice that $IC_A(\cdot)$ in (C.4) (a utility difference) is strictly decreasing in P_A. This implies that the demand curve is also decreasing in P_A. We can meaningfully define WTP as follows:

$$\pi_A(L^*, W_1, W_2, X, Y, P_L, \xi) = \max\{P_A; IC_A(L^*, W_1, W_2, X, Y, P_L, P_A, \xi) \leq 0\}$$

In other words, $\pi_A(\cdot)$ can be solved from the following implicit equation:

$$IC_A(L^*, W_1, W_2, X, Y, P_L, \pi_A(\cdot), \xi) = 0$$

Lastly, consider the case of a life care annuity. Suppose that stand alone insurances are not available. We assume no saving, so $C_1 = W_1 - P_{CA} \cdot CA$; $C_2^h = W_2 + Y \cdot CA$; $C_2^l = W_2 + Y \cdot CA + (\rho \cdot Y \cdot CA - X)$. Substitution of these equalities yields the following direct utility function:

$$\begin{aligned} V(CA; W_1, W_2, \rho, Y, X, P_{CA}) &= U(W_1 - P_{CA}CA) + (s(\xi) - l(\xi))U(W_2 + Y \cdot CA) \\ &\quad + l(\xi)U(W_2 + Y \cdot CA + (\rho Y \cdot CA - X)) \end{aligned}$$

A consumer buys a life care annuity if:

$$V(1; W_1, W_2, \rho, Y, X, P_{CA}) - V(0; W_1, W_2, \rho, Y, X, P_{CA}) \geq 0. \tag{C.5}$$

Then, demand for a life care annuity is given by the probability that this incentive compatibility constraint is met:

$$\begin{aligned} D_{CA}(P_{CA}|W_1, W_2, \rho, Y, X) &= \mathbb{P}\Big(- (U(W_1) - U(W_1 - P_{CA})) \\ &\quad + s(\xi)(U(W_2 + Y) - U(W_2)) \\ &\quad + l(\xi)LLL(W_2, \rho, Y, X) \geq 0\Big) \\ &= \mathbb{P}(IC_{CA}(W_1, W_2, \rho, Y, X, P_{CA}, \xi) \geq 0). \end{aligned} \tag{C.6}$$

where

$$LLL(W_2, \rho, Y, X) = ((U(W_2 + Y + (\rho Y - X)) - U(W_2 + Y)) - (U(W_2 - X) - U(W_2)))$$

Since $U()$ is strictly concave and $W_2 > 0$, $Y > 0$ and $X > 0$, $LLL(W_2, \rho, Y, X) > 0$.

Notice that $IC_{CA}(\cdot)$ in (C.6) (a utility difference) is strictly decreasing in P_{CA}. Moreover, $D_{CA}(P_{CA}|\cdot)$ is strictly decreasing in P_{CA}. So, we can meaningfully define the Willingness To Pay (WTP) as follows:

$$\pi_{CA}(W_1, W_2, \rho, Y, X, \xi) = \max\{P_{CA}; IC_{CA}(W_1, W_2, \rho, Y, X, P_{CA}, \xi) = 0\}$$

In other words, $\pi_{CA}(\cdot)$ can be solved from the following implicit equation:

$$IC_{CA}(W_1, W_2, \rho, Y, X, \pi_{CA}(\cdot), \xi) = 0.$$

We now derive the comparative statics of the demand curve considering the premium and correlation between risks l and s. The demand curves (C.3), (C.4), and (C.6) can be written as the following implicit functions:

$$D_L(P_L|\cdot) = \mathbb{P}(l(\xi) \cdot \upsilon_{2,l,L}(A^*, W_2, Y, X) \geq \upsilon_{1,L}(P_L|W_1, P_A, A^*))$$
$$D_A(P_A|\cdot) = \mathbb{P}(l(\xi) \cdot \upsilon_{2,l,A}(L^*, W_2, Y, X) + s(\xi) \cdot \upsilon_{2,s,A}(W_2, Y) \geq \upsilon_{1,A}(P_A|W_1, P_L, L^*))$$
$$D_{CA}(P_{CA}|\cdot) = \mathbb{P}(l(\xi) \cdot \upsilon_{2,l,CA}(W_2, \rho, Y, X) + s(\xi) \cdot \upsilon_{2,s,CA}(W_2, Y) \geq \upsilon_{1,CA}(P_{CA}|W_1)).$$

Note that these demand curves are of the form $\mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \geq \upsilon_1)$ where (s, l) are potentially correlated risks and $\upsilon_{2,l} > 0$, $\upsilon_{2,s} > 0$ and $\upsilon_1 \geq 0$ are scalars determining demand. $\upsilon_{2,l}$ and $\upsilon_{2,s}$ are the utility gains from insurance coverage in period 2 of risks l and s, respectively. υ_1 is the utility loss in period 1 due to paying a premium for the insurance. Obviously, the larger the insurance utility gains $\upsilon_{2,l} > 0$ and $\upsilon_{2,s}$ are, the more likely a consumer will buy insurance. Also, the lower the premium, the smaller the utility loss υ_1 is, and hence the more likely the demand for an insurance product is. Formally:

$$\frac{\partial D_K(P_K|\cdot)}{\partial P_K} = \overbrace{\frac{\mathbb{P}(\cdot)}{\partial \upsilon_{1,K}}}^{>0} \cdot \overbrace{\frac{\partial \upsilon_{1,K}}{\partial P_K}}^{<0} < 0$$

which means that demand is lower if the premium is higher ($K \in (L, A, CA)$).

Next, we ask ourselves: does a demand curve of the form $D(P) = \mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \geq \upsilon_1(P))$ becomes *steeper* or *flatter* if we decrease the correlation θ between risks l and s? Put concretely, we are interested in the comparative statics:

$$\frac{\partial^2 D_L(P_L;\theta)}{\partial P_L \partial \theta}; \qquad \frac{\partial^2 D_A(P_A;\theta)}{\partial P_A \partial \theta}; \qquad \frac{\partial^2 D_{CA}(P_{CA};\theta)}{\partial P_{CA} \partial \theta}.$$

To this end, we have to explicitly derive the demand curve $D(P|\theta) = \mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \geq \upsilon_1|\theta)$ as a function of correlation θ, because that parameter is missing in the current demand function. This requires knowledge of the joint distribution of s and l and its dependence on correlation θ. Also, fixing everything else for our comparative static means that we want to fix the marginal distributions in the population of l and s, and only vary the part of the joint distribution that involves the correlation structure. Define $\Gamma(F_l, F_s)$ to be the set of joint distribution functions with marginals $F_l = \mathbb{P}(l \leq L)$ and $F_s = \mathbb{P}(s \leq S)$. Following Solomon (2022) the correlation structure of interest is:

Definition of a correlation order *Suppose we have two populations $X, Y \in \Gamma(F_l, F_s)$ and have joint CDFs F_X, F_Y, respectively. Solomon (2022) defines the correlation between l and s in population X is less correlated than in population Y or that X precedes Y in correlation order, written as $X \precsim Y$ if and only if:*

$$\mathbb{P}(s \leq S, l \leq L|X) = F_X(S, L) \leq F_Y(S, L) = \mathbb{P}(s \leq S, l \leq L|Y) \textit{ for all } (S, L) \in D_F,$$

so the probability of a pair with low (s, l) is smaller in population X than in Y, implying the correlation is more negative in population X.

Ideally we have the same marginal distribution in F_l and F_s and modify the joint relationship between the two variables only via a correlation parameter. A class of distribution functions that meet these needs including a correlation order, are those of

Farlie-Gumble-Morgenstern form (Denuit and Scaillet, 2004; Solomon, 2022):

$$F(S,L) = F_l(L) \cdot F_s(S) \cdot (1 + \theta \cdot (1 - F_l(L)) \cdot (1 - F_s(S))) \tag{C.7}$$

with $\theta \in [-1,1]$ governing the dependence between the two marginals and $\theta = 0$ implying independent distributions for s and l.

For simplicity, and following Solomon (2022), we assume $l \sim \mathcal{U}(0,1)$ and $s \sim \mathcal{U}(0,1)$ so $F_l(L) = L$ and $F_s(S) = S$. Then:

$$F(S,L) = L \cdot S \cdot (1 + \theta \cdot (1 - L) \cdot (1 - S)) \implies f(S,L) = 1 + \theta \cdot (1 - 2L) \cdot (1 - 2S)$$

To find the demand functions, we require the convolution: $\mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \geq \upsilon_1 \mid \theta) = 1 - \mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \leq \upsilon_1 \mid \theta)$. The closed-form solutions $\mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \leq \upsilon_1 \mid \theta)$, are:

$$\begin{cases}
\frac{\upsilon_1^2}{2\upsilon_{2,l}\upsilon_{2,s}} + \theta \cdot \frac{1}{6} \cdot \frac{\upsilon_1^2}{\upsilon_{2,s}^4} \cdot (\upsilon_1^2 - 2(\upsilon_{2,l} + \upsilon_{2,s}) \cdot \upsilon_1 + 3\upsilon_{2,l}\upsilon_{2,s}), & \text{if } \upsilon_1(P) \leq \min(\upsilon_{2,l}, \upsilon_{2,s}) \\
\frac{\upsilon_1}{\upsilon_{2,s}} - \frac{\upsilon_{2,l}}{2 \cdot \upsilon_{2,s}} + \theta \cdot \frac{1}{6} \cdot \frac{\upsilon_{2,l}}{\upsilon_{2,s}} \cdot \left(1 - 2 \cdot \left(\frac{\upsilon_1 - \upsilon_{2,l}}{\upsilon_{2,s}}\right) - \frac{\upsilon_{2,l}}{\upsilon_{2,s}}\right), & \text{if } \upsilon_1(P) \in [\upsilon_{2,l}, \upsilon_{2,s}] \\
\frac{\upsilon_1}{\upsilon_{2,l}} - \frac{\upsilon_{2,s}}{2 \cdot \upsilon_{2,l}} + \theta \cdot \frac{1}{6} \cdot \frac{\upsilon_{2,s}}{\upsilon_{2,l}} \cdot \left(1 - 2 \cdot \left(\frac{\upsilon_1 - \upsilon_{2,s}}{\upsilon_{2,l}}\right) - \frac{\upsilon_{2,s}}{\upsilon_{2,l}}\right), & \text{if } \upsilon_1(P) \in [\upsilon_{2,s}, \upsilon_{2,l}] \\
1 - \frac{\upsilon_{2,l} + \frac{1}{2}\upsilon_{2,s} - \upsilon_1}{\upsilon_{2,l}} - \frac{1}{2} \cdot \frac{(\upsilon_1 - \upsilon_{2,l})^2}{\upsilon_{2,l}\upsilon_{2,s}} - \theta \cdot \frac{1}{6} \cdot \frac{\upsilon_{2,s}}{\upsilon_{2,l}} \cdot & \\
\left(\frac{\upsilon_1 - \upsilon_{2,s}}{\upsilon_{2,l}} + \frac{\upsilon_1 - \upsilon_{2,l}}{\upsilon_{2,l}} - 3 \cdot \left(\frac{\upsilon_1 - \upsilon_{2,l}}{\upsilon_{2,s}}\right)^2\right) - \theta \cdot \frac{1}{6} \cdot & \\
\frac{\upsilon_{2,s}}{\upsilon_{2,l}} \cdot \left(2 \cdot \left(1 - \frac{\upsilon_{2,s}}{\upsilon_{2,l}}\right) \cdot \left(\frac{\upsilon_1 - \upsilon_{2,l}}{\upsilon_{2,s}}\right)^3 + \frac{\upsilon_{2,s}}{\upsilon_{2,l}} \cdot \left(\frac{\upsilon_1 - \upsilon_{2,l}}{\upsilon_{2,s}}\right)^4\right), & \text{if } \upsilon_1(P) \geq \max(\upsilon_{2,l}, \upsilon_{2,s}),
\end{cases}$$

which depends on the premium level (P) via $\upsilon_1(P)$, with $\upsilon_1'(P) > 0$.

We are interested in the sign of the comparative static:

$$\frac{\partial^2 \mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \geq \upsilon_1(P) \mid \theta)}{\partial P \partial \theta} = \upsilon_1'(P) \cdot - \frac{\partial^2 \mathbb{P}(l \cdot \upsilon_{2,l} + s \cdot \upsilon_{2,s} \leq \upsilon_1 \mid \theta)}{\partial \upsilon_1 \partial \theta}.$$

The relevant part of the comparative static is $-\frac{\partial^2 \mathbb{P}(l\cdot v_{2,l}+s\cdot v_{2,s}\leq v_1 \mid \theta)}{\partial v_1 \partial\theta}$, given by:

$$\begin{cases} -\frac{1}{6}\cdot\frac{v_1}{v_{2,s}^4}\cdot(4v_1^2-6(v_{2,l}+v_{2,s})\cdot v_1+6v_{2,l}v_{2,s}), & \text{if } v_1(P)\leq\min(v_{2,l},v_{2,s}) \\ \frac{1}{6}\cdot\frac{v_{2,l}}{v_{2,s}}\cdot\frac{2}{v_{2,s}}, & \text{if } v_1(P)\in[v_{2,l},v_{2,s}] \\ \frac{1}{6}\cdot\frac{v_{2,s}}{v_{2,l}}\cdot\frac{2}{v_{2,l}}. & \text{if } v_1(P)\in[v_{2,s},v_{2,l}] \\ \frac{1}{6}\cdot\frac{v_{2,s}}{v_{2,l}}\cdot\frac{1}{v_{2,l}v_{2,s}^3}\cdot\left(2v_{2,s}^3-6v_{2,l}v_{2,s}\cdot(v_1-v_{2,l})\right)+\frac{1}{6}\cdot\frac{v_{2,s}}{v_{2,l}}\cdot \\ \quad\frac{1}{v_{2,l}v_{2,s}^3}\cdot\left(6(v_{2,l}-v_{2,s})\cdot(v_1-v_{2,l})^2+4\cdot(v_1-v_{2,l})^3\right), & \text{if } v_1(P)\geq\max(v_{2,l},v_{2,s}) \end{cases}$$

which has sign:

$$\begin{cases} \leq 0, & \text{if } v_1(P)\leq\min\left(\frac{3}{4}\cdot\left(v_{2,l}+v_{2,s}-\sqrt{v_{2,l}^2+v_{2,s}^2-\frac{2}{3}v_{2,l}v_{2,s}}\right),v_{2,l},v_{2,s}\right) \\ \geq 0, & \text{if } v_1(P)\in\left[\frac{3}{4}\cdot\left(v_{2,l}+v_{2,s}-\sqrt{v_{2,l}^2+v_{2,s}^2-\frac{2}{3}v_{2,l}v_{2,s}}\right),\min(v_{2,l},v_{2,s})\right] \\ \geq 0, & \text{if } v_1(P)\in[v_{2,l},v_{2,s}] \\ \geq 0, & \text{if } v_1(P)\in[v_{2,s},v_{2,l}] \\ \geq 0, & \text{if } v_1(P)\in\left[\max(v_{2,l},v_{2,s}),\frac{3}{4}\cdot\left(v_{2,l}+v_{2,s}+\sqrt{v_{2,l}^2+v_{2,s}^2-\frac{2}{3}v_{2,l}v_{2,s}}\right)\right] \\ \leq 0 & \text{if } v_1(P)\in\left[\frac{3}{4}\cdot\left(v_{2,l}+v_{2,s}+\sqrt{v_{2,l}^2+v_{2,s}^2-\frac{2}{3}v_{2,l}v_{2,s}}\right),v_{2,l}+v_{2,s}\right], \end{cases}$$

implying that the impact of θ on the slope of the demand curve changes sign maximally twice. If $\Delta\theta < 0$, then the demand function features a *steeper* decline at high values of P, is *flatter* at intermediate values of P, and has a *steeper* decline at low values of P. The demand curve becomes flatter, because total risk exposure is more homogenous if the correlation is more negative, i.e. θ is lower. However, at high and low premia, there is a steeper decline because extreme risk individuals are still possible, i.e. with a high pair of risks ($s = 1, l = 1$) or a low pair of risks ($s = 0, l = 0$).

C.4 Derivation of Optimal Life Care Annuities

The first-order condition is:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{F}(\rho)}{\partial \rho} = 2 \cdot \mathbb{E}\left(PR(\xi,\rho) \cdot \frac{\partial PR(\xi,\rho)}{\partial \rho} \right), \text{ with:} \\ PR(\xi,\rho) &= \frac{s(\xi) + l(\xi) \cdot \rho}{\mathbb{E}(s(\xi)) + \mathbb{E}(l(\xi)) \cdot \rho} - 1, \end{aligned} \tag{C.8}$$

First, we determine $\frac{\partial PR(\xi,\rho)}{\partial \rho}$. Suppose $\rho \neq -\frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))}$, $\mathbb{E}(s(\xi)) \neq 0$ and $\mathbb{E}(l(\xi)) \neq 0$, then:

$$\begin{aligned} \frac{\partial PR(\xi,\rho)}{\partial \rho} &= \frac{l(\xi) \cdot \{\mathbb{E}(s(\xi)) + \rho \cdot \mathbb{E}(l(\xi))\} - \cdot \mathbb{E}(l(\xi)) \cdot \{s(\xi) + \rho \cdot l(\xi)\}}{\{\mathbb{E}(s(\xi)) + \rho \cdot \mathbb{E}(l(\xi))\}^2} \\ &= \frac{l(\xi) \cdot \mathbb{E}(s(\xi)) - \mathbb{E}(l(\xi)) \cdot s(\xi)}{\{\mathbb{E}(s(\xi)) + \rho \cdot \mathbb{E}(l(\xi))\}^2} \\ &= \omega(\rho)^2 \cdot \left\{ \frac{l(\xi)}{\mathbb{E}(l(\xi))} - \frac{s(\xi)}{\mathbb{E}(s(\xi))} \right\} \text{ and: } \omega(\rho) = \frac{\sqrt{\mathbb{E}(l(\xi)) \cdot \mathbb{E}(s(\xi))}}{\mathbb{E}(s(\xi)) + \rho \cdot \mathbb{E}(l(\xi))} \neq 0. \end{aligned} \tag{C.9}$$

We use this result to solve first-order condition (C.8):

$$\begin{aligned} 0 &= 2 \cdot \mathbb{E}\left(PR(\xi,\rho) \cdot \left. \frac{\partial PR(\xi,\rho)}{\partial \rho} \right|_{\rho=\rho^\star} \right) \\ &= 2 \cdot \mathbb{E}\left(\left\{ \frac{s(\xi) - \mathbb{E}(s(\xi)) + \rho^\star \cdot (l(\xi) - \mathbb{E}(l(\xi)))}{\mathbb{E}(s(\xi)) + \rho^\star \cdot \mathbb{E}(l(\xi))} \right\} \cdot \Omega(\xi) \right) \cdot \omega(\rho^\star)^2 \rightarrow \end{aligned}$$

$$\mathbb{E}\left\{ (s(\xi) - \mathbb{E}(s(\xi))) \cdot \Omega(\xi) \right\} = -\rho^\star \cdot \mathbb{E}\left\{ (l(\xi) - \mathbb{E}(l(\xi))) \cdot \Omega(\xi) \right\} \rightarrow$$

$$\rho^\star = \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\mathbb{E}\left\{ \left(\frac{s(\xi)}{\mathbb{E}(s(\xi))} - 1 \right) \cdot \Omega(\xi) \right\}}{\mathbb{E}\left\{ \left(\frac{l(\xi)}{\mathbb{E}(l(\xi))} - 1 \right) \cdot -\Omega(\xi) \right\}}, \quad \text{with: } \Omega(\xi) = \frac{l(\xi)}{\mathbb{E}(l(\xi))} - \frac{s(\xi)}{\mathbb{E}(s(\xi))}$$

Substituting back $\Omega(\xi)$ gives:

$$\begin{aligned}
\rho^\star &= \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\mathbb{E}\left\{\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}-1\right\}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}-\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}\right\}}{\mathbb{E}\left\{\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}-1\right\}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}-\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}\right\}} \\
&= \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\mathbb{E}\left\{\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}\right\}+\mathbb{E}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}-\mathbb{E}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}-\mathbb{E}\left\{\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}^2\right\}}{\mathbb{E}\left\{\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}\right\}+\mathbb{E}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}-\mathbb{E}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}-\mathbb{E}\left\{\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}^2\right\}} \\
&= \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\operatorname{Cov}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))},\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}-\operatorname{Var}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\operatorname{Cov}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))},\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}-\operatorname{Var}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} \\
&= \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}-\operatorname{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))},\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}{\left\{\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}\right\}^{-1}-\operatorname{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))},\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}. \quad \square
\end{aligned}$$

Note that to get from step two to three we use the identity $\mathbb{E}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}=\mathbb{E}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}=1$.

To examine the behavior of $\rho^\star$ to changes in $\frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))}$, $\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}$ and $\operatorname{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))},\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}$, we can compute the corresponding partial derivatives:

$$\frac{\partial\rho^\star}{\partial\frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))}} = \frac{\mathbb{E}(l(\xi))}{\mathbb{E}(s(\xi))}\cdot\rho^\star = \begin{cases} = 0 \text{ if } & \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} = \operatorname{Corr}\{\cdot\} \\ > 0 \text{ if } & \operatorname{Corr}\{\cdot\} \le 0 \\ > 0 \ \operatorname{Corr}\{\cdot\} > 0 \ \wedge & \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} \in \left(\operatorname{Corr}\{\cdot\}, \frac{1}{\operatorname{Corr}\{\cdot\}}\right) \\ < 0 \text{ elsewhere.} & \end{cases}$$

Note:

$$\frac{\partial\rho^\star}{\partial\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}} = \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))}\cdot\frac{2\cdot\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}-\left(\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}\right)^2\cdot\operatorname{Corr}\{\cdot\}-\operatorname{Corr}\{\cdot\}}{\left(1-\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}\cdot\operatorname{Corr}\{\cdot\}\right)^2}$$

then:

$$\frac{\partial \rho^\star}{\partial \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}} = \begin{cases} = 0 \text{ if } \quad \mathrm{Corr}\{\cdot\} > 0 \quad \wedge \quad \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} = \frac{1}{\mathrm{Corr}\{\cdot\}} \pm \sqrt{\frac{1}{\mathrm{Corr}\{\cdot\}^2} - 1} \\ < 0 \text{ if } \quad \mathrm{Corr}\{\cdot\} > 0 \quad \wedge \\ \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} \notin \left(\frac{1}{\mathrm{Corr}\{\cdot\}} - \sqrt{\frac{1}{\mathrm{Corr}\{\cdot\}^2} - 1},\ \frac{1}{\mathrm{Corr}\{\cdot\}} + \sqrt{\frac{1}{\mathrm{Corr}\{\cdot\}^2} - 1}\right) \\ > 0 \text{ elsewhere.} \end{cases}$$

Lastly:

$$\frac{\partial \rho^\star}{\partial \mathrm{Corr}\{\cdot\}} = \frac{\mathbb{E}(s(\xi))}{\mathbb{E}(l(\xi))} \cdot \frac{\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} - \left(\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}\right)^{-1}}{\left(\left\{\frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}}\right\}^{-1} - \mathrm{Corr}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}, \frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}\right)^2} \tag{C.10}$$

so:

$$\frac{\partial \rho^\star}{\partial \mathrm{Corr}\{\cdot\}} = \begin{cases} = 0 \text{ if } \quad \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} = 1 \\ < 0 \text{ if } \quad \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} < 1 \\ > 0 \text{ if } \quad \frac{\mathrm{SD}\left\{\frac{s(\xi)}{\mathbb{E}(s(\xi))}\right\}}{\mathrm{SD}\left\{\frac{l(\xi)}{\mathbb{E}(l(\xi))}\right\}} > 1. \end{cases}$$

C.5 Descriptive Statistics

Table C.2: Descriptive Statistics

	Mean	Median	S.D.	Min[2]	Max
Individuals (*N = 3,278,797*)					
Uses long-term care (LTC) (%)	39.8				
Passes away (%)	25.3				
Observed duration LTC [1]	2.3	1.3			
Married households (Unbalanced panel; *Panel observations= 5,906,251*)					
Household income[3] (*000s euros*)					
Bottom Lifetime IQ	21.3	21.2	2.2	2.8	41.4
2nd Lifetime IQ	25.8	25.8	2.9	5.7	57.8
3rd Lifetime IQ	31.4	31.4	4.7	6.8	86.1
4th Lifetime IQ	41.1	41.2	7.8	5.9	120.7
Top Lifetime IQ	67.8	61.4	31.8	6.2	1,038.5
All	40.0	33.4	23.2	2.3	1,038.5
Liquid assets (*000s euros*)					
Bottom Lifetime IQ	16.9	8.9	22.3	0.0	336.8
2nd Lifetime IQ	33.6	22.6	39.0	0.0	628.5
3rd Lifetime IQ	53.7	31.3	64.8	0.0	1,056.3
4th Lifetime IQ	85.2	47.6	103.1	0.0	1,753.7
Top Lifetime IQ	396.5	142.3	3,348.3	0.0	156,145.9
All	133.5	37.2	1,630.6	0.0	156,145.9
Single-person households (Unbalanced panel; *Panel observations= 8,073,927)*					
Women (%)	76.6		42.3		
Household income (*000s euros)*					
Bottom Lifetime IQ	14.9	14.6	1.6	0.1	35.4
2nd Lifetime IQ	17.9	18.0	2.1	0.7	51.0
3rd Lifetime IQ	21.8	22.1	3.7	0.7	81.8
4th Lifetime IQ	28.0	28.5	6.3	0.8	113.1
Top Lifetime IQ	43.6	40.7	21.7	0.7	926.7
All	23.9	19.6	13.5	0.1	926.7
Liquid assets (*000s euros)*					
Bottom Lifetime IQ	10.6	5.1	14.5	0.0	342.5
2nd Lifetime IQ	25.2	16.5	31.3	0.0	589.0
3rd Lifetime IQ	43.6	24.1	55.7	0.0	935.1
4th Lifetime IQ	74.1	38.3	91.4	0.0	1,638.7
Top Lifetime IQ	330.4	131.8	1,628.4	0.0	101,383.2
All	83.4	21.4	668.5	0.0	101,383.2

Notes: IQ = Income Quintile; [1] Conditional upon using LTC; [2] Maximum and minimum are the averages of the one hundred highest and lowest values; [3] Income in 2015 prices. Savings and bonds in 2015 prices, stocks inflated with AEX stock-index of 31^{st} of December 2014.

Table C.3: Hazard Rate Estimates

Transition	No LTC → LTC	No LTC → Death	LTC → LTC	LTC → Death
Constant (β_k)	-2.938***	-4.564***	-1.626***	-2.498***
	(0.014)	(0.033)	(0.018)	(0.022)
Single at baseline (β_{1kh})				
Men - 2nd Lifetime IQ	0.008	-0.142***	0.044	-0.031
	(0.021)	(0.052)	(0.027)	(0.035)
Men - 3rd Lifetime IQ	-0.239***	-0.192***	0.511***	0.303***
	(0.021)	(0.049)	(0.027)	(0.035)
Men - 4th Lifetime IQ	-0.459***	-0.261***	0.955***	0.684***
	(0.021)	(0.047)	(0.027)	(0.037)
Men - Top Lifetime IQ	-0.798***	-0.624***	1.249***	0.973***
	(0.021)	(0.048)	(0.027)	(0.040)
Women - Bottom Lifetime IQ	-0.027*	-1.245***	0.874***	-0.365***
	(0.015)	(0.041)	(0.020)	(0.026)
Women - 2nd Lifetime IQ	-0.178***	-1.618***	1.203***	-0.317***
	(0.016)	(0.049)	(0.021)	(0.030)
Women - 3rd Lifetime IQ	-0.360***	-1.615***	1.501***	-0.024
	(0.017)	(0.050)	(0.021)	(0.033)
Women - 4th Lifetime IQ	-0.535***	-1.709***	1.754***	0.227***
	(0.017)	(0.051)	(0.022)	(0.035)
Women - Top Lifetime IQ	-0.723***	-1.988***	1.952***	0.499***
	(0.018)	(0.055)	(0.023)	(0.038)
Married at baseline - currently single (β_{1kh})				
Men - Bottom Lifetime IQ	-0.442***	-0.000	1.066***	0.952***
	(0.019)	(0.049)	(0.026)	(0.034)
Men - 2nd Lifetime IQ	-0.696***	-0.362***	1.245***	1.075***
	(0.018)	(0.045)	(0.024)	(0.032)
Men - 3rd Lifetime IQ	-0.868***	-0.593***	1.351***	1.175***
	(0.018)	(0.045)	(0.024)	(0.032)
Men - 4th Lifetime IQ	-1.074***	-0.662***	1.477***	1.273***
	(0.018)	(0.045)	(0.024)	(0.034)
Men - Top Lifetime IQ	-1.326***	-0.932***	1.602***	1.382***
	(0.019)	(0.045)	(0.025)	(0.037)
Women - Bottom Lifetime IQ	-0.203***	-1.724***	1.459***	-0.648***
	(0.017)	(0.049)	(0.022)	(0.029)
Women - 2nd Lifetime IQ	-0.411***	-1.731***	1.648***	-0.294***
	(0.016)	(0.047)	(0.021)	(0.029)
Women - 3rd Lifetime IQ	-0.643***	-1.672***	1.773***	0.111***
	(0.017)	(0.047)	(0.021)	(0.030)
Women - 4th Lifetime IQ	-0.864***	-1.863***	1.913***	0.440***
	(0.017)	(0.049)	(0.022)	(0.033)
Women - Top Lifetime IQ	-1.079***	-2.013***	2.000***	0.752***
	(0.018)	(0.050)	(0.023)	(0.035)

Table C.3: (continued)

Transition	No LTC → LTC	No LTC → Death	LTC → LTC	LTC → Death
Married at baseline - currently single $(\beta^k_{1kh} + \beta^k_{2kh})$				
Men - Bottom Lifetime IQ	-0.723***	-0.397***	1.546***	1.385***
	(0.016)	(0.038)	(0.022)	(0.029)
Men - 2nd Lifetime IQ	-1.027***	-0.671***	1.727***	1.678***
	(0.016)	(0.037)	(0.021)	(0.028)
Men - 3rd Lifetime IQ	-1.244***	-0.807***	1.821***	1.948***
	(0.016)	(0.036)	(0.021)	(0.028)
Men - 4th Lifetime IQ	-1.448***	-0.958***	1.910***	2.165***
	(0.016)	(0.037)	(0.021)	(0.029)
Men - Top Lifetime IQ	-1.655***	-1.128***	1.949***	2.333***
	(0.016)	(0.037)	(0.022)	(0.030)
Women - Bottom Lifetime IQ	-0.494***	-1.969***	1.738***	-0.362***
	(0.015)	(0.043)	(0.020)	(0.028)
Women - 2nd Lifetime IQ	-0.743***	-1.890***	1.901***	0.091***
	(0.015)	(0.040)	(0.020)	(0.027)
Women - 3rd Lifetime IQ	-0.992***	-1.729***	2.014***	0.688***
	(0.015)	(0.040)	(0.020)	(0.028)
Women - 4th Lifetime IQ	-1.231***	-1.808***	2.115***	1.117***
	(0.016)	(0.040)	(0.020)	(0.029)
Women - Top Lifetime IQ	-1.422***	-1.863***	2.189***	1.617***
	(0.016)	(0.040)	(0.021)	(0.030)
γ_k	0.076***	0.137***	-0.028***	0.075***
	(0.001)	(0.003)	(0.002)	(0.002)
Single at baseline (γ_{kh})				
Men - 2nd Lifetime IQ	-0.003*	-0.012***	0.007***	0.005**
	(0.002)	(0.005)	(0.002)	(0.002)
Men - 3rd Lifetime IQ	0.007***	-0.020***	-0.009***	-0.005**
	(0.002)	(0.005)	(0.002)	(0.002)
Men - 4th Lifetime IQ	0.014***	-0.024***	-0.026***	-0.018***
	(0.002)	(0.004)	(0.002)	(0.002)
Men - Top Lifetime IQ	0.026***	-0.016***	-0.027***	-0.027***
	(0.002)	(0.004)	(0.002)	(0.002)
Women - Bottom Lifetime IQ	0.003**	0.006*	-0.036***	-0.010***
	(0.001)	(0.003)	(0.002)	(0.002)
Women - 2nd Lifetime IQ	0.009***	0.002	-0.043***	-0.012***
	(0.001)	(0.004)	(0.002)	(0.002)
Women - 3rd Lifetime IQ	0.015***	-0.001	-0.051***	-0.022***
	(0.001)	(0.004)	(0.002)	(0.002)
Women - 4th Lifetime IQ	0.021***	0.002	-0.055***	-0.027***
	(0.001)	(0.004)	(0.002)	(0.002)
Women - Top Lifetime IQ	0.025***	0.017***	-0.056***	-0.033***
	(0.001)	(0.004)	(0.002)	(0.002)

Table C.3: (continued)

Transition	No LTC → LTC	No LTC → Death	LTC → LTC	LTC → Death
Married at baseline (γ_{kh})				
Men - 1st Lifetime IQ	0.026***	-0.012***	-0.044***	-0.021***
	(0.001)	(0.003)	(0.002)	(0.002)
Men - 2nd Lifetime IQ	0.038***	-0.011***	-0.048***	-0.028***
	(0.001)	(0.003)	(0.002)	(0.002)
Men - 3rd Lifetime IQ	0.044***	-0.011***	-0.047***	-0.029***
	(0.001)	(0.003)	(0.002)	(0.002)
Men - 4th Lifetime IQ	0.048***	-0.010***	-0.043***	-0.029***
	(0.001)	(0.003)	(0.002)	(0.002)
Men - Top Lifetime IQ	0.052***	-0.002	-0.039***	-0.028***
	(0.001)	(0.003)	(0.002)	(0.002)
Women - 1st Lifetime IQ	0.014***	0.019***	-0.063***	-0.003*
	(0.001)	(0.004)	(0.002)	(0.002)
Women - 2nd Lifetime IQ	0.025***	0.001	-0.066***	-0.016***
	(0.001)	(0.004)	(0.002)	(0.002)
Women - 3rd Lifetime IQ	0.034***	-0.007**	-0.064***	-0.029***
	(0.001)	(0.004)	(0.002)	(0.002)
Women - 4th Lifetime IQ	0.041***	-0.000	-0.062***	-0.038***
	(0.001)	(0.004)	(0.002)	(0.002)
Women - Top Lifetime IQ	0.046***	0.012***	-0.057***	-0.043***
	(0.001)	(0.004)	(0.002)	(0.002)
Frailty: $\ln(\sigma^2)$	-16.369	0.261	-21.026	-2.104
Spells	4,028,551	4,028,551	1,795,027	1,795,027
Uncensored spells	1,425,236	206,997	770,070	622,346
Individuals	3,063,815	3,063,815	1,303,914	1,303,914
Sub-Log-likelihood (C.2)	-4,468,814.2	-1,055,077.0	-1,632,120.1	-1,530,497.4
Log-Likelihood	-5,523,891.2		-3,162,617.5	
Sub-Log-likelihood ($\sigma^2 = 0$)	-4,468,814.2	-1,057,575.3	-1,632,120.1	-1,538,180.9
Log-Likelihood ($\sigma^2 = 0$)	-5,526,389.5		-3,170,301.0	
LR test ($H_0 : \sigma^2 = 0$)	$p > 0.10$	$p < 0.01$	$p > 0.10$	$p < 0.01$

Notes: Significance levels: * 10-%; **5-%; ***1-%. IQ = Income Quintile

C.6 Additional Results

C.6.1 Remaining Life Expectancy and Long-Term Care Use over Marital Status

Table C.4: Remaining Life Expectancy and Long-Term Care Use at Age 65 by PI Quintiles (Men)

(a) Men	All	Bottom	Second	Third	Fourth	Top	ΔTop - Bottom
LE (years)							
All	18.0 (17.9;18.1)	15.3 (15.0;15.5)	16.8 (16.6;17.0)	17.6 (17.5;17.8)	18.4 (18.2;18.6)	19.2 (19.1;19.4)	4.0 (3.7;4.2)
Initial Married	18.6 (18.4;18.7)	16.2 (15.9;16.5)	17.4 (17.2;17.6)	18.1 (17.9;18.3)	18.9 (18.7;19.1)	19.5 (19.3;19.8)	3.4 (3.0;3.7)
Initial Singles	16.1 (15.9;16.2)	14.1 (13.8;14.5)	14.7 (14.3;15.1)	15.5 (15.1;15.9)	16.4 (16.0;16.7)	17.9 (17.6;18.3)	3.8 (3.4;4.3)
LTC (years)*							
All	3.0 (3.0;3.1)	3.8 (3.7;4.0)	3.4 (3.3;3.6)	3.2 (3.1;3.2)	3.0 (2.9;3.0)	2.8 (2.7;2.8)	−1.1 (−1.2;−0.9)
Initial Married	2.9 (2.8;2.9)	3.0 (2.8;3.2)	3.1 (3.0;3.2)	3.0 (2.9;3.1)	2.9 (2.8;3.0)	2.7 (2.6;2.8)	−0.3 (−0.5;−0.1)
Initial Singles	3.7 (3.6;3.8)	4.8 (4.6;5.0)	4.6 (4.4;4.8)	3.8 (3.6;3.9)	3.2 (3.1;3.4)	2.9 (2.8;3.1)	−1.9 (−2.1;−1.6)
Ratio (%)							
All	12 (12;12)	20 (19;20)	15 (15;15)	12 (12;13)	11 (10;11)	9 (9;9)	−11 (−11;−10)
Initial Married	10 (9;10)	12 (11;12)	11 (11;11)	10 (10;10)	9 (9;10)	8 (8;9)	−4 (−4;−3)
Initial Singles	20 (20;21)	30 (29;31)	29 (28;30)	22 (21;22)	16 (16;17)	13 (12;13)	−17 (−18;−16)
Ever use LTC (%)							
All	77 (76;78)	79 (78;80)	79 (78;80)	78 (77;79)	77 (76;78)	75 (75;76)	−3 (−5;−2)

Notes: These are population-averaged measures for the life cycle simulation of 100,000 households. We present the median estimates across 5,000 bootstrapped samples and the 2.5^{th} and 97.5^{th} percentiles between brackets.

Table C.4: Remaining Life Expectancy and Long-Term Care Use at Age 65 by PI Quintiles (continued: Women)

(b) Women	All	Bottom	Second	Third	Fourth	Top	ΔTop - Bottom
LE at age 65 (years)							
All	21.9 (21.8;22.0)	20.1 (19.9;20.3)	21.8 (21.6;22.0)	22.0 (21.8;22.2)	22.2 (22.1;22.4)	22.3 (22.2;22.5)	2.3 (2.0;2.5)
Initial Married	22.5 (22.3;22.6)	22.1 (21.7;22.4)	22.5 (22.3;22.8)	22.4 (22.2;22.6)	22.6 (22.3;22.8)	22.5 (22.3;22.7)	0.4 (0.0;0.8)
Initial Singles	20.7 (20.6;20.8)	19.1 (18.9;19.3)	20.5 (20.2;20.8)	21.0 (20.8;21.3)	21.4 (21.1;21.7)	21.9 (21.6;22.1)	2.8 (2.4;3.1)
LTC (years)*							
All	5.1 (5.1;5.2)	6.0 (5.9;6.2)	5.9 (5.8;6.0)	5.3 (5.2;5.4)	4.8 (4.7;4.9)	4.4 (4.3;4.4)	−1.7 (−1.8;−1.5)
Initial Married	5.1 (5.1;5.2)	6.3 (6.1;6.6)	6.0 (5.9;6.2)	5.4 (5.3;5.5)	5.0 (4.8;5.1)	4.4 (4.3;4.6)	−1.9 (−2.1;−1.6)
Initial Singles	5.1 (5.0;5.1)	5.9 (5.8;6.0)	5.6 (5.5;5.8)	5.1 (5.0;5.3)	4.6 (4.4;4.7)	4.1 (3.9;4.2)	−1.8 (−2.0;−1.7)
Ratio (%)							
All	18 (18;18)	25 (25;26)	21 (21;22)	18 (18;19)	16 (16;16)	14 (13;14)	−12 (−12;−11)
Initial Married	16 (16;16)	21 (21;22)	20 (19;20)	17 (17;18)	15 (15;16)	13 (13;13)	−8 (−9;−8)
Initial Singles	21 (21;22)	27 (27;28)	24 (24;25)	21 (20;21)	18 (17;18)	15 (15;16)	−12 (−13;−12)
Ever use LTC (%)							
All	89 (88;89)	91 (91;92)	91 (91;92)	89 (89;90)	88 (88;89)	86 (85;87)	−5 (−6;−4)

Notes: These are population-averaged measures for the life cycle simulation of 100,000 households. We present the median estimates across 5,000 bootstrapped samples and the 2.5^{th} and 97.5^{th} percentiles between brackets.

C.6.2 Premium Returns

Table C.5: Premium Returns for Different Groups (in %)

Income quintile:	Bottom	Second	Third	Fourth	Top
Household level ($\rho^* = 1.35$)					
Pension Annuity	−8.9	−2.6	−0.6	1.5	3.6
	(−9.6;−8.2)	(−3.2;−1.9)	(−1.1;−0.1)	(1.0;2.0)	(3.2;4.0)
LTC insurance	29.9	17.9	4.1	−6.0	−17.0
	(27.6;31.9)	(16.0;19.7)	(2.7;5.6)	(−7.4;−4.7)	(−18.2;−15.8)
Life care annuity	−1.4	1.4	0.3	0.1	−0.3
	(−2.2;−0.8)	(0.6;2.0)	(−0.4;0.9)	(−0.5;0.7)	(−0.7;0.0)
Single Men ($\rho^* = 2.11$)					
Pension Annuity	−12.0	−8.3	−3.3	1.9	11.8
	(−13.9;−10.1)	(−10.7;−5.9)	(−5.5;−1.3)	(0.1;3.7)	(10.1;13.4)
LTC insurance	29.8	28.9	2.3	−13.8	−21.6
	(24.9;34.8)	(23.0;34.7)	(−2.0;6.7)	(−17.4;−10.0)	(−25.0;−18.3)
Life care annuity	0.0	2.4	−1.7	−2.6	2.2
	(−2.0;1.9)	(−0.3;4.9)	(−4.3;0.8)	(−4.6;−0.4)	(0.8;3.8)
Single Women ($\rho^* = 1.47$)					
Pension Annuity	−7.7	−0.7	1.7	3.4	5.6
	(−8.6;−6.8)	(−2.0;0.5)	(0.5;3.0)	(2.2;4.5)	(4.5;6.9)
LTC insurance	16.0	12.5	1.1	−10.7	−20.8
	(13.8;18.3)	(9.4;15.6)	(−1.7;4.0)	(−13.2;−8.1)	(−23.2;−18.2)
Life care annuity	−1.7	2.6	1.6	−0.1	−1.0
	(−2.6;−0.9)	(1.1;4.1)	(0.1;3.1)	(−1.5;1.2)	(−1.9;0.1)
Married Men ($\rho^* = 11.16$)					
Pension Annuity	−12.9	−6.2	−2.4	1.7	5.3
	(−14.5;−11.4)	(−7.3;−5.2)	(−3.2;−1.5)	(0.9;2.4)	(4.6;6.0)
LTC insurance	4.2	8.5	5.2	0.1	−8.0
	(−1.9;10.1)	(4.1;12.9)	(2.0;8.6)	(−2.7;3.0)	(−10.4;−5.3)
Life care annuity	−3.1	2.1	2.0	0.8	−2.3
	(−6.5;−0.9)	(−0.7;4.7)	(−0.2;4.1)	(−1.1;2.7)	(−3.5;−0.8)
Married women ($\rho^* = 0.08$)					
Pension Annuity	−1.7	0.3	−0.3	0.4	0.0
	(−3.1;−0.4)	(−0.7;1.3)	(−1.1;0.5)	(−0.3;1.1)	(−0.6;0.7)
LTC insurance	26.8	21.1	7.0	−3.4	−15.9
	(22.1;31.6)	(17.8;24.3)	(4.6;9.5)	(−5.5;−1.2)	(−17.8;−14.0)
Life care annuity	−1.3	0.6	−0.2	0.4	−0.2
	(−2.6;0.0)	(−0.2;1.4)	(−0.9;0.6)	(−0.4;1.1)	(−0.6;0.3)

Notes: The table provides the data to Figures 3.2 and 3.4. The premium returns are population-averaged measures for the life cycle simulation of 100,000 individuals. Medians across 5,000 bootstrapped samples and the $2.5^{th}$and 97.5^{th} percentile (in brackets) are shown.

Appendix D: Chapter 4

D.1 Life Cycle Model

D.1.1 Government Budget Constraint

The government collects the taxes and co-payments to finance expenditures on the first pillar pension and LTC provision. Yet, government revenues and spending are not guaranteed to be balanced in the model. To let the government break even, we assume additional fixed transfers of Tr_{SS} and Tr_{LTC}(a tax or subsidy) in each age period. For a household of a given age, the government expenditures on LTC are:

$$\text{LTC}(h_t^m, h_t^f) = \begin{cases} 2 \cdot \text{LTC}_{\text{cost}} & \text{if } \ h_t^m = 2 \ \text{ and } \ h_t^f = 2, \\ \text{LTC}_{\text{cost}} & \text{if } \ h_t^m = 2 \ \text{ or } \ h_t^f = 2, \\ 0 & \text{elsewhere,} \end{cases}$$

where $\text{LTC}_{\text{cost}} = €58,500$ is the cost of an individual stay in a public institution for a year.

Similarly, the government pays first pillar pension:

$$\text{SS}(t, h_t^m, h_t^f) = \begin{cases} 2 \cdot w & \text{if } \ t \geq 65, \ h_t^m \neq 3 \ \text{ and } \ h_t^f \neq 3, \\ 1.4 \cdot w & \text{if } \ t \geq 65, \ h_t^m = 3 \ \text{ or } \ h_t^f = 3, \\ 0 & \text{elsewhere.} \end{cases}$$

These are the expenditures per household and conditional upon age t and health statuses h_t^m and h_t^f. Total, i.e., unconditional, government expenditures GE_{LTC} and GE_{SS} are the expenditures per household weighted by the steady-state distribution on household types $f(\aleph)$, with $\aleph = \aleph^W \cup \aleph^R = (a_t, \theta, \eta_t, \epsilon_t, \text{DB}_t, f_t, h_t^m, h_t^f, t)'$. Then:

$$\text{GE}_{SS} = \sigma_1 \cdot \int_{\aleph} f(\aleph) \cdot \text{SS}(\aleph)\text{d}\aleph, \quad \text{and} \quad \text{GE}_{LTC} = \sigma_2 \cdot \int_{\aleph} f(\aleph) \cdot \text{LTC}(\aleph)\text{d}\aleph,$$

where σ_1 and σ_2 reflect the share of government expenditures financed through dedicated taxes and co-payments. The rest is financed with general taxes and not of interest when balancing the government budget.[1]

To finance these benefits, the government obtains revenue from taxes and co-payments: $\tau_{\text{SS}}(\cdot)$, $\tau_{\text{L}}(\cdot)$, and $m(\cdot)$. Also, there is an additional balancing transfer Tr_x with $x \in (SS, L)$. The transfer is defined as follows:

$$\text{Tr}_x(f) = \begin{cases} 2 \cdot \text{Tr}_x & \text{if } f = \text{ couple} \\ \text{Tr}_x & \text{if } f = \text{ single woman or single man,} \end{cases}$$

and is thus twice as large for couples than for singles.

Government revenues, GR_x, are given by:

$$\text{GR}_{SS}(\text{Tr}_{SS}) = \int_{\aleph} f(\aleph) \cdot (\tau_{\text{SS}}(\aleph) + \text{Tr}_{SS}(\aleph))\,\text{d}\aleph \quad \text{and}$$

$$\text{GR}_{LTC}(\text{Tr}_{LTC}) = \int_{\aleph} f(\aleph) \cdot (\tau_{\text{L}}(\aleph) + m(\aleph) + \text{Tr}_{LTC}(\aleph))\,\text{d}\aleph,$$

which consist of the sum of taxes, co-payments for LTC, and the additional tax (subsidy) that balances the government budget constraint.

The government sets the transfer levels Tr_x according to: $\text{GE}_x = \text{GR}_x(\text{Tr}_x)$, which can be tax or subsidy, depending on whether there is a deficit or a surplus. Appendix D.1.4 explains how we compute these transfers numerically.

[1] We take the values from 2010: $\sigma_1 = 0.664$ and $\sigma_2 = 0.640$, which we computed using aggregate expenditures and revenues reported on: https://www.cbs.nl/nl-nl/nieuws/2019/37/inkomsten-uit-sociale-premies-6-1-miljard-hoger-in-2018, and https:/opendata.cbs.nl/statline/CBS/nl/dataset/84121NED/table?ts=1564565763409, [*both retrieved on*: August 7^{th}, 2023].

D.1.2 Closed-form Solution for Policy Function Iteration

We elaborate here on how the households determine their consumption policy functions. We use the Bellmann maximization principle, which recursively solves the household optimization problem from the last to the first life cycle period. The objective function is the value function in this case. A general form of the value function in any state $\aleph = \aleph^W \cup \aleph^R$ is given by:

$$\begin{aligned} \mathrm{V}(\aleph;\ h_t^m = i, h_t^f = j) = &\max_{c_t, a_{t+1}} \mathrm{u}^f(c_t) \\ &+ \beta \cdot \Big((1 - \pi_{3,3}^{i,j}(t,I)) \cdot \mathbb{E}[\mathrm{V}(\aleph^+)|\aleph] + \pi_{3,3}^{i,j}(t,I) \cdot \mathcal{B}(a_{t+1}) \Big) \\ \text{s.t.} \quad a_{t+1} = &R \cdot a_t + y_t - \tau_{SS} - \tau_L - \tau_G - m_t + \mathrm{Tr}_{SS}(\cdot) + \mathrm{Tr}_{LTC}(\cdot) - c_t \geq 0, \end{aligned} \tag{D.1}$$

where $(i,j) \in \{1,2,3\}$. The Lagrangian optimization problem corresponding to (D.1) reads as:

$$\begin{aligned} \max_{c_t, a_{t+1}, \lambda} \mathcal{L}(\cdot) = \ &\mathrm{u}^f(c_t) + \beta \cdot \Big((1 - \pi_{3,3}^{i,j}(t,I)) \cdot \mathbb{E}[\mathrm{V}(\aleph^+)|\aleph] + \pi_{3,3}^{i,j}(t,I) \cdot \mathcal{B}(a_{t+1}) \Big) \\ &+ \lambda \cdot \{ R \cdot a_t + y_t - \tau_{SS} - \tau_L - \tau_G - m_t + \mathrm{Tr}_{SS}(\cdot) + \mathrm{Tr}_{LTC}(\cdot) - c_t - a_{t+1} \}, \end{aligned} \tag{D.2}$$

which has the following first-order constraints:

$$\frac{\partial \mathcal{L}(\cdot)}{\partial c_t} := u_{c^t}^f - \lambda = 0 \tag{D.3}$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial a_{t+1}} := \beta \cdot \Big((1 - \pi_{3,3}^{i,j}(t,I) \cdot \mathbb{E}[\mathrm{V}_{a_{t+1}}(\aleph^+)|\aleph] + \pi_{3,3}^{i,j}(t,I) \cdot \mathcal{B}_{a_{t+1}}(a_{t+1}) \Big) - \lambda = 0 \tag{D.4}$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial \lambda} := R \cdot a_t + y_t - \tau_{SS} - \tau_L - \tau_G - m_t + \mathrm{Tr}_{SS}(\cdot) + \mathrm{Tr}_{LTC}(\cdot) - c_t - a_{t+1} = 0. \tag{D.5}$$

Note that $\mathrm{V}(\aleph)$ in (D.1) is an optimum, and so is the Lagrangian in (D.2) when analyzed in $c_t(\aleph)$, $a_{t+1}(\aleph)$, and $\lambda(\aleph)$. As a consequence, we can apply the envelope theorem:

$$\mathrm{V}_{a_t}(\aleph) = \frac{\partial \mathrm{V}(\aleph)}{\partial a_t} = \frac{\partial \mathcal{L}(\cdot)}{\partial a_t}\bigg|_{c_t(\aleph), a_{t+1}(\aleph), \lambda(\aleph)} = \lambda(\aleph) \cdot R,$$

which has to hold in the next period as well:

$$\mathrm{V}_{a_{t+1}}(\aleph^{+}) = \frac{\partial \mathrm{V}(\aleph^{+})}{\partial a_{t+1}} = \frac{\partial \mathcal{L}(\cdot)}{\partial a_{t+1}}\bigg|_{c_{t+1}(\aleph^{+}),a_{t+2}(\aleph^{+}),\lambda(\aleph^{+})} = \lambda(\aleph^{+}) \cdot R, \tag{D.6}$$

where $\aleph^{+}$ is the state vector in the next period. Furthermore, (D.3) holds optimally in the future:

$$u^{f^{+}}_{c_{t+1}}\left(c_{t+1}(\aleph^{+})\right) = \lambda(\aleph^{+}). \tag{D.7}$$

Combining (D.6) and (D.7) yields:

$$\mathrm{V}_{a_{t+1}}(\aleph^{+}) = u^{f^{+}}_{c_{t+1}}\left(c_{t+1}(\aleph^{+})\right) \cdot R \tag{D.8}$$

Using (D.8), we build the Euler equation that describes the evolution of consumption and assets over time. We combine (D.8) with (D.3) and (D.4), while (D.5) simultaneously holds (together with the non-negativity constraint of assets). The Euler equation on consumption and bequests (assets) is:

$$\begin{aligned} u^{f}_{c_t}\left(c_t(\aleph)\right) &= \beta\left(\left(1-\pi^{i,j}_{3,3}(t,I)\right)\cdot R\cdot \mathbb{E}[u^{f^{+}}_{c_{t+1}}\left(c_{t+1}(\aleph^{+})\right)|\aleph] \;+\; \pi^{i,j}_{3,3}(t,I)\cdot \mathcal{B}_{a_{t+1}}(a_{t+1}(\aleph))\right) \\ \text{with:}\quad a_{t+1}(\aleph) &= \; R\cdot a_t + y_t - \tau_{SS} - \tau_L - \tau_G - m_t + \mathrm{Tr}_{SS}(\cdot) + \mathrm{Tr}_{LTC}(\cdot) - c_t(\aleph) \;\geq\; 0 \end{aligned} \tag{D.9}$$

This system can be recursively solved if we know the solution for the last period.

D.1.3 Terminal Period Solution

We now solve the dynamic program problem for the terminal (last) period $t = T$. Note that the household will not be around in the next period ($\pi^{i,j}_{3,3}(T,I) = 1$) but can bequeath, where $(i,j) \in \{1,2,3\}$. The terminal period solution of (D.9) in state $\aleph$

reduces to:

$$
\begin{aligned}
u^f_{c_T}\left(c_T(\aleph)\right) &= \beta \cdot \mathcal{B}_{a_{T+1}}(a_{T+1}(\aleph)) \\
a_{T+1}(\aleph) &= R \cdot a_T + y_T - \tau_{SS} - \tau_L - \tau_G - m_T + \text{Tr}_{SS}(\cdot) + \text{Tr}_{LTC}(\cdot) - c_t(\aleph) \\
&= \mu - c_T(\aleph) \ \geq \ 0 \qquad \text{(D.10)}
\end{aligned}
$$

where μ is the total wealth holding at age T that is split over consumption and a bequest.

To solve the system, we have to consider three cases: $\phi = 0$ (no bequest), $\phi \in (0,1)$ (some wealth above threshold c_a is bequeathed), and $\phi = 1$ (all wealth above threshold c_a is bequeathed). The marginal utility of leaving a bequest is:

$$
\mathcal{B}_{a_{T+1}}(a_{T+1}) = \begin{cases} 0 & \text{if } \phi = 0 \\ \frac{\phi}{1-\phi}^{\sigma} \cdot \left(\frac{\phi}{1-\phi} \cdot c_a + a_{T+1}\right)^{-\sigma} & \text{if } \phi \in (0,1) \\ c_a^{-\sigma} & \text{if } \phi = 1. \end{cases}
$$

Also, marginal utility from consumption depends on family structure:

$$
u_{c_T}(c_T) = \begin{cases} c_T^{-\sigma} & \text{if } f_T = \text{ single man or woman} \\ 2 \cdot \left(\frac{1}{\eta}\right)^{1-\sigma} \cdot c_T^{-\sigma} & \text{if } f_T = \text{ couple.} \end{cases}
$$

If $\phi = 0$, the Euler equation in (D.10) becomes:

$$
\begin{aligned}
& u^f_{c_T}\left(c_T(\aleph)\right) = \beta \cdot \mathcal{B}_{a_{J+1}}(a_{J+1}(\aleph)) \rightarrow \\
& u^f_{c_T}\left(c_T(\aleph)\right) > \beta \cdot 0 \rightarrow \\
& \quad c_T(\aleph, \mu) = \mu,
\end{aligned}
$$

where the latter equality stems from the budget constraint in (D.10).

If $\phi = 1$, the Euler equation in (D.10) becomes:

$$\beta \cdot \mathcal{B}_{a_{J+1}}(a_{J+1}(\aleph)) = \begin{cases} c_T^{-\sigma} & \text{if } f_T = \text{ single man or woman} \\ 2 \cdot \left(\frac{1}{\eta}\right)^{1-\sigma} \cdot c_T^{-\sigma} & \text{if } f_T = \text{ couple.} \end{cases}$$

Solving for c_T gives:

$$c_T(\aleph, \mu) = \begin{cases} \min\left(\beta^{-\frac{1}{\sigma}} \cdot c_a,\ \mu\right) & \text{if } f_T = \text{ single man or woman} \\ \min\left(2^{\frac{1}{\sigma}} \cdot \left(\frac{1}{\eta}\right)^{\frac{1}{\sigma}-1} \cdot \beta^{-\frac{1}{\sigma}} \cdot c_a,\ \mu\right) & \text{if } f_T = \text{ couple.} \end{cases}$$

Similarly, we solve the Euler equation for $\phi \in (0,1)$ and get:

$$c_T(\aleph, \mu) = \min\left(\left(\frac{x_1(f_T)}{x_1(f_T) + x_2} \cdot x_1^{-1}(f_T) \cdot c_a + \frac{x_2}{x_1(f_T) + x_2} \cdot \mu\right)\ ,\ \mu\right)$$

with:

$$x_1^{-1}(f_T) = \begin{cases} \beta^{-\frac{1}{\sigma}} & \text{if } f_T = \text{ single man or woman} \\ 2^{\frac{1}{\sigma}} \cdot \left(\frac{1}{\eta}\right)^{\frac{1}{\sigma}-1} \cdot \beta^{-\frac{1}{\sigma}} & \text{if } f_T = \text{ couple,} \end{cases}$$

and $x_2 = \left(\frac{\phi}{1-\phi}\right)^{-1}$.

Note that the bequest size is $a_{T+1}(\aleph, \mu) = \max\left(\mu - c_T(\aleph, \mu)\ ,\ 0\right)$ in all cases.

D.1.4 Numerically Solving the Model

We first discretize the state space and then solve the model along the discrete space.

Discretizing the state space Consider the vector with state variables $\aleph = \aleph^W \cup \aleph^R = (a_t, \theta, \eta_t, \epsilon_t, \text{DB}_t, f_t, h_t^m, h_t^f, t)'$. This vector contains continuous variables a_t, θ, η_t, DB_t, and ϵ_t. Solving the Euler equation for each value is computationally too demanding and we, therefore, discretize these variables while maintaining the core properties of their distribution.

We discretize labor productivity $\theta \sim \mathcal{N}(0, \sigma_\theta^2)$ and the transitory income shock $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ into a five- and three-dimensional grid using Gauss–Hermite quadrature. We discretize the stochastic AR(1)-variable η_t into a time-independent three-state Markov process. We use the decomposition method by Rouwenhorst (1995), which preserves the unconditional mean, the unconditional variance, and the auto-correlation of the actual process. Kopecky and Suen (2010) describes the algorithm in detail. We discretize the second pillar pension benefit on a 12-dimensional exponential grid from 0 to $150,000$ (growth rate $= 0.52$).

Lastly, we discretize assets (a_t) over a grid $\hat{\mathcal{A}}$ from €0 to €1,000,000. The asset grid contains 100 values. To prevent oscillation of the model for asset levels near zero, we take an exponential grid, i.e., we take relatively more low than high values for assets a_t on the grid (growth rate $= 0.05$).

Solving the model We require the probability distribution of assets $a_{t+1}(\aleph)$ and consumption $c_t(\aleph)$ at any age t. Suppose all parameter values are known in the model. We apply policy function iteration to solve the model and then compute the probability distribution.

We start with the closed-form solution of the terminal period T provided in Appendix D.1.3. We hereafter numerically solve the Euler equation system (D.9) from period $T-1$ back to period 1 and calculate the resulting policy functions $c_t(\aleph)$ and $a_{t+1}(\aleph)$.

Next, we compute the distribution of households over the state space $\aleph$. To increase computational speed, we analytically compute the distribution rather than infer this from a simulation (see, e.g., Cagetti, 2003). Furthermore, directly computing the distribution prevents that in an agent-based simulation, it remains unknown for what number of households the model statistics converge.

We compute the state distribution at age t by updating the state distribution at time $t-1$. For this, we assume an initial state distribution at age $t = 25$. The initial household consists of a couple without using LTC. They draw labor productivity level θ from the discrete distribution. We take $a_0 = 0$, $\text{DB}_{25} = 0$, and $\eta_{24} = 0$, so the household initially has no assets, pension accruals, and income shock. This distribution is modified to create a distribution over the state space for age 26. Given the current state $\aleph$ at age 25, we know how many assets any household chooses to possess at age 26 and the conditional probability of ending up in a particular health and income state at age 26. This information (transition matrix) suffices to update the state distribution of $\aleph$ from age 25 to the distribution at age 26. We repeat this procedure until age $t = T = 100$.

These state distributions are also essential to compute the transfer Tr_x that would balance the government budget (see Appendix D.1.1), $x \in (SS, LTC)$. For each state, we know the cost of providing LTC and pension, the paid taxes, and co-payments. We can subsequently compute the expected government revenues and costs. We apply a bisection search to find the level Tr_x that exactly balances the revenues and cost.

D.2 First-stage Estimates

D.2.1 Data and Estimation of the Health Processes

Socioeconomic differences in LTC use and mortality are the primary input in our analysis. To quantify them, we use longitudinal data on LTC use and mortality, a simulation model to compute complete life histories on LTC use and death, and a socioeconomic status measure to stratify the life histories. The data and estimation procedure of the health process closely follows Chapter 3, which we will summarize here.

We use unique registry data from Statistics Netherlands reporting an individual and household key, institutional care use, death, marital status, birth date, and gender for the Dutch population between 2006 and 2014. The data are unique due to their high frequency: the registers daily report whether an individual stays in an institution, i.e., a residential or nursing home, died, and has a partner, i.e., is married, has a partnership contract, or cohabits on a contractual basis. The high frequency of the data allows us to precisely model many short institutional care spells that occur (see Chapter 2 of this thesis). Furthermore, it will enable us to model the effect of marital status on LTC use and mortality precisely from the moment of marital dissolution onward.

We restrict the estimation of the health process to households whose members are both retired, i.e., aged 65 or older and have retirement income as their main income source. The age restriction seems natural as only 1.0% of the 65-year-olds in our sample uses institutional care. To save on the number of heterogeneous groups, and thus state space of the life cycle model, we further restrict to individuals who are or were married at age 65. We observe 2,548,664 individuals and 1,487,109 households.

To construct a socioeconomic measure, we merge this data to household records on income – the sum of couple members' pre-tax income (incl. social transfers and pension income) – and financial assets (savings, stocks, and bonds). The socioeconomic status measure is the average sum of equivalized household income and annuitized financial assets (savings, stocks, and bonds), reflecting lifetime income. This comprehensive measure has the advantage that it considers that after retirement, some households have little income but many assets, e.g., former entrepreneurs (Knoef et al., 2016). We compute lifetime income quintiles $I \in \{1, 2, 3, 4, 5\}$ depending on quintiles of its distribution.[2]

To compute complete life histories on LTC use and death, we use the competing risk model from Chapter 3 that allows for socioeconomic dependencies in risks and explicitly accounts for the spouse as a potential informal care provider. We distinguish three individual states: not using public institutional care ($i = 1$), using public institutional care ($i = 2$), or death ($i = 3$). Home-based care use is not a separate state because its co-payments and, thus, redistributive effects are very limited in the Netherlands (Tenand et al., 2020b). For parsimony, marital status is modeled as a covariate, and not as a separate (sub-)state in the competing risk model. As a first step, we specify and estimate a proportional hazard model for the transition rate λ_{ij} of going from a given state i to state $j \neq i$ at age t (van den Berg, 2001):

$$\lambda_{ij}(t \mid \mathrm{mar}(t), G, I) = \exp\left(\gamma_{ij}(G, I) \cdot t + c_{ij}(G, I) + \beta_{ij}(G, I) \cdot \mathrm{mar}(t)\right) \tag{D.11}$$

[2] An alternative would be to take the level of education, but the register on education is incomplete for older cohorts, implying we have to stick to the current data.

where γ_{ij} is the age effect, c_{ij} is the effect of being single, and $c_{ij} + \beta_{ij}$ is the effect of having a partner (mar(t) = 0: has no partner; mar(t) = 1 has no partner). All coefficients are estimated conditional upon gender G and lifetime income quintile I.[3] We estimate the model following standard log-likelihood inference for duration models (see Chapter 5 in this thesis).

Because we observe the relevant outcomes only between 2006 and 2014, we use the estimates of (D.11) to simulate complete life histories on LTC use, marital status, and mortality. We generate a survival probability and thus a random timing of the transition from i to j:

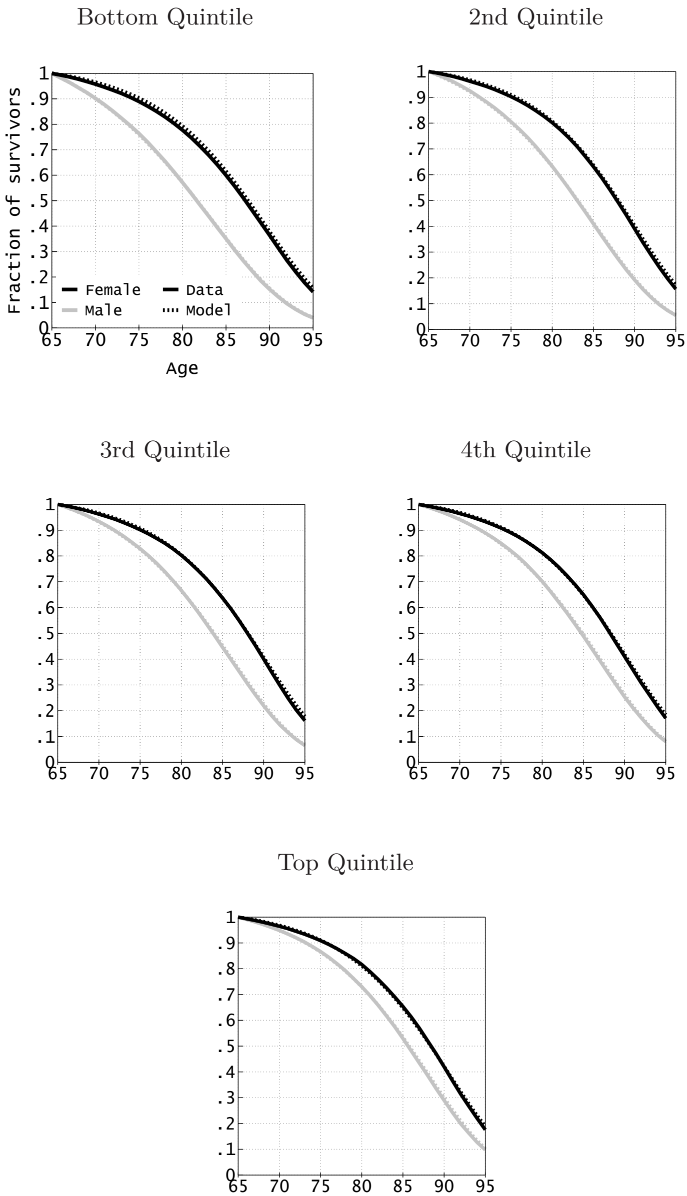
$$S_{ij}(t \mid \text{mar}(t), G, I) = \mathbb{P}(T \geq t,\ j \mid \text{mar}(t), G, I,\ i) \ = \ \exp\left(-\int_0^t \lambda_{ij}(\tau \mid \text{mar}(\tau), G, I) d\tau\right) \tag{D.12}$$

The simulation starts at age 65 with 100,000 households, when both couple members are alive. Each individual can move to two possible destination states. Using (D.12), we draw a transition time for each state. The minimum of the two transition times determines which actual transition occurs. We repeat this procedure for the successive states until both members died. While the simulation is finished for the couple member who dies first, we still have to simulate the life history of LTC use for the surviving partner after widowhood. We use (D.12) but take the dummy value mar(t) = 0 instead of mar(t) = 1. After this last spouse dies, we stop the simulation and have the complete –and dependent– life histories on LTC use and mortality for the two partners.

[3]See Appendix D.2.2 for the fit on LTC use and mortality. We also estimated a model including frailty, but this specification gave a worse fit on LTC use and mortality.

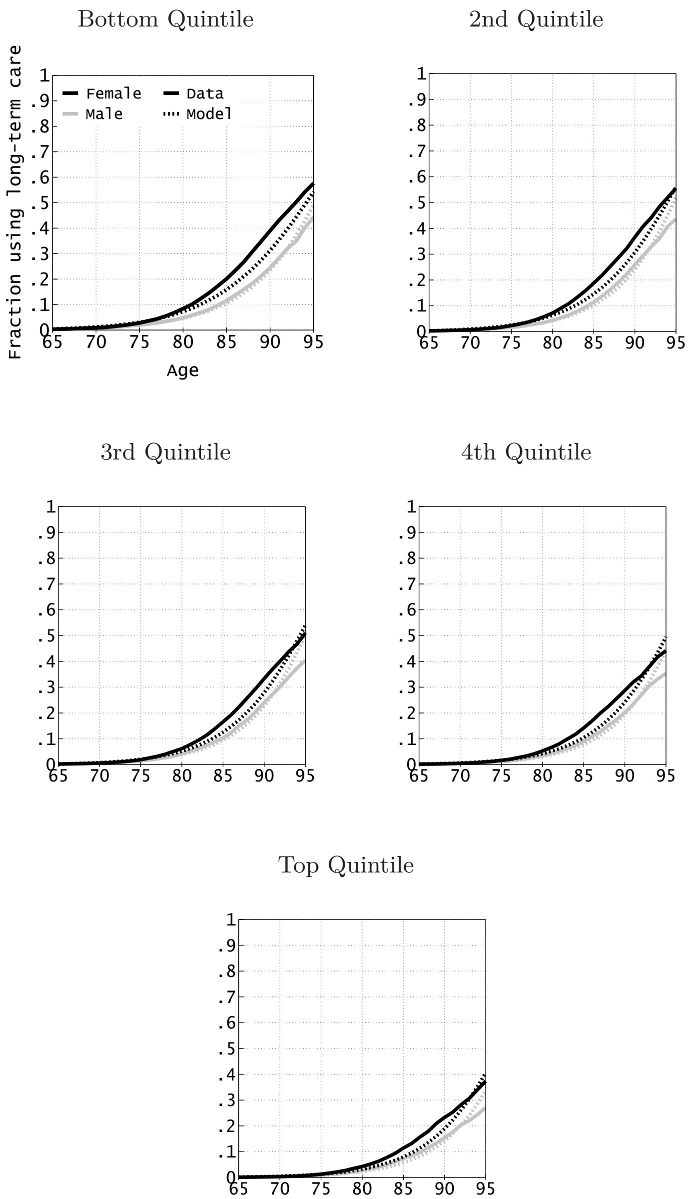
D.2.2 Goodness of Fit of Health Processes

Figure D.1: Goodness of Fit of Survival Curves



Notes: The figure compares the empirical survival curves with their simulated counterpart. The simulated curves are population-averaged measures of a life cycle simulation of 100, 000 households with 1, 000 bootstrapped samples.

Figure D.2: Goodness of Fit of Long-term Care Use

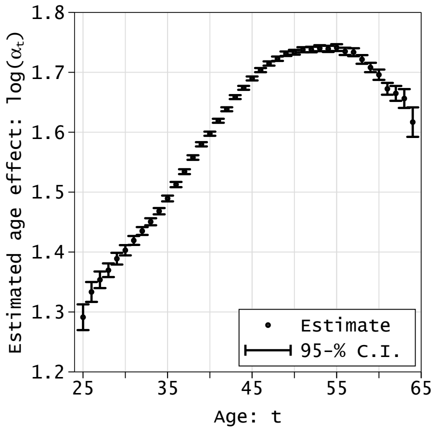


Notes: The figure compares the empirical long-term care curves with their simulated counterpart. The simulated curves are population-averaged measures of a life cycle simulation of 100,000 households with 1,000 bootstrapped samples.

D.2.3 Age Profile on Income

Figure D.3 presents the model estimates for the age profile $\{\bar{c} + \log(\alpha_t)\}_{t=25}^{64}$. $\bar{c}$ is the fixed effect for the 1950 cohort, which we add because we want to tailor the income profile to the 1950 cohort. Figure D.3 displays a familiar hump-shape (cf. Mincer, 1974): income peaks at age 55 and decreases after that. This pattern arises due to the accumulation and decumulation of human capital –working experience– over the life cycle, and households start to work less when retirement nears.

Figure D.3: Estimated Age Profile on Income



Notes: Income is measured in 0000s euros. Parameters are estimated for married households whose oldest member is younger than 65 and born after 1949. Adding $\bar{c}$ implies normalized estimates that refer to the age effect for those born in 1950. Data from the IPO 2001-2014: 77,118 households and 534,006 panel-year observations.

D.2.4 Income Uncertainty

We model household income dynamics as an AR(1) (canonical) process:

$$\begin{aligned} \log(y_t) &= \log(\alpha_t) + \theta + \eta_t + \epsilon_t \\ \eta_t &= \rho \cdot \eta_{t-1} + u_t \\ \theta \sim \mathcal{N}(0, \sigma_\theta^2); \quad \epsilon_t &\sim \mathcal{N}(0, \sigma_\epsilon^2); \quad u_t \sim \mathcal{N}(0, \sigma_u^2); \quad \eta_{24} = 0 \end{aligned}$$

First, we estimate the age effects $\log(\alpha_t)$ by running a fixed effects regression of log household income on age dummies, where each dummy represents a distinct effect $\log(\alpha_t)$. Next, to wash out birth cohort effects, we regress the estimate $\hat{\theta}_i$ on birth year dummies and impute the household's $\hat{\theta}_i$ to the value it would have if born in 1950. We then estimate the uncertain income component $\theta + \eta_t + \epsilon_t$ by minimum distance estimation, minimizing the squared difference between theoretical and empirical moments (cf. Storesletten et al., 2004). Because we have an auto-regressive process with a lag of one year, we match the variance and first-order auto-correlation of the income component.

The assumptions on the persistent income component imply the following process in terms of the past and current shocks:

$$\eta_t \;=\; \rho^{t-24} \cdot \eta_{24} + \sum_{j=25}^{t} \rho^{t-j} \cdot u_j + \epsilon_t, \quad t = 25, ., 64$$

from which the moments

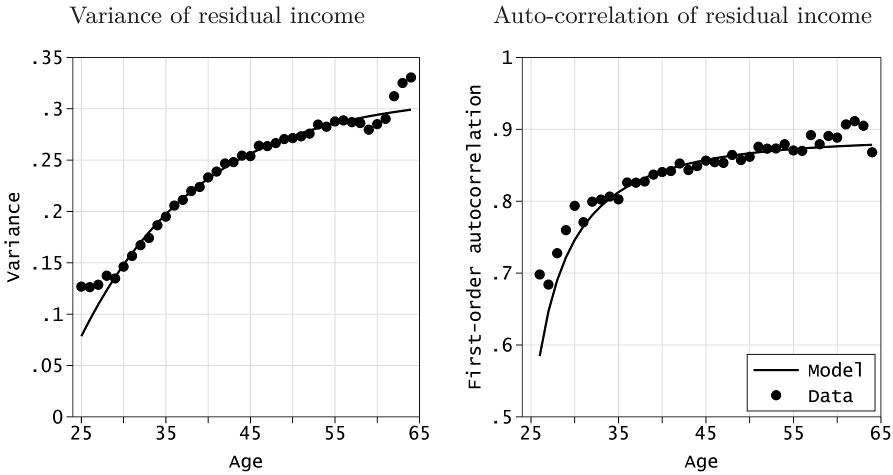
$$\begin{aligned} \mathrm{var}(\theta + \eta_t + \epsilon_t) \;&=\; \sigma_\theta^2 + \rho^{2(t-24)} \cdot \sigma_z^2 + \sum_{j=25}^{t} \rho^{2(t-j)} \cdot \sigma_u^2 + \sigma_\epsilon^2 \\ \mathrm{cov}(\theta + \eta_t + \epsilon_t,\; \theta + \eta_{t-1} + \epsilon_{t-1}) \;&=\; \sigma_\theta^2 + \rho^{2(t-24)-1} \cdot \sigma_z^2 + \sum_{j=25}^{t} \rho^{1+2(t-j)} \cdot \sigma_u^2 \end{aligned}$$

follow, allowing us to identify the moments. Identification follows standard covariance arguments. For further details on identification, we refer to Arellano (2003).

We employ a weighted minimum distance estimator to fit these 79 moments (40 for the variances 39 for the covariances). The objective function is the sum of squared differences between the theoretical and empirical variances and co-variances. Due to the small sample considerations explained in Altonji and Segal (1996), our estimator employs the identity matrix as the weighting matrix. Hence, each moment receives the same weight in the objective function. The estimator, which minimizes the objective function, yields consistent but possibly inefficient estimates. Figure D.3 and Table 4.2

in Section 4.4.1 present the estimates for the structural parameters.

Figure D.4: Fit of the Income Process Before Age 65



Notes: Income measured in 0000s euros. We report the parameters for married households whose oldest member is younger than age 65 and born after 1949. Data from the IPO 2001-2014: 77,118 households and 534,006 panel-year observations.

Figure D.4 shows the goodness-of-fit of the model estimates for the targeted moments. Our model matches the variance and first-order auto-correlation (closely related to first-order autocovariance) of the income shock process well. Notably, the variance of the income shock increases over time, implying more heterogeneity in income when age increases. This is important when constructing heterogeneity in asset profiles with our life cycle model.

D.2.5 Replacement Rates

We compute the replacement rates of survivor pensions using the IPO data restricted to households whose members are all aged 65 and over. Both members must have retirement income as their primary income source. The IPO does not distinguish between occupational pension benefits and income from privately purchased annuities (third pillar), so the replacement rate reflects both occupational and privately-arranged pension benefits. We run a fixed effects regression of log private pension income on year dummies and the family structure: being a couple, a single man, or a single woman. The exponentiated coefficient for singles gives their replacement rate. The estimates for a single man or woman are $\text{rr}_m = 0.93$ (SE: 0.001) and $\text{rr}_f = 0.55$ (SE: 0.005), respectively. The widow's replacement rate means that each euro of a defined pension benefit drops to 55 cents when the female spouse survives. In line with our earlier work van der Vaart et al. (2020), we report $\text{rr}_m > \text{rr}_f$ implied by that men were the prime earner in the households and pension benefits mostly accrued to them.

D.2.6 Tax Function Estimates

For general taxes, we estimate the following specification (cf. Heathcote et al., 2020):

$$\tau_G(y, \cdot) = y - \lambda \cdot y^{1-\tau},$$

which we estimate conditional upon age group (below vs. above age 65) and family structure (married vs. single).

Table D.1 shows the estimates. Our estimates are in the ballpark of Heathcote et al. (2020). Using data from the Congressional Budget Office, they report $\tau \in (0.089, 0.236)$ for the U.S. between 2012-2016. λ is merely a level effect and thus does not have appropriate benchmark values. For dedicated taxes for first pillar pension (τ_{SS}) and

Table D.1: Parameters of the General Income Tax Function τ_G

	Couples		Singles
	Below age 65	Above age 65	Above age 65
λ	1.241	1.157	1.073
	(0.005)	(0.008)	(0.012)
τ	0.185	0.162	0.148
	(0.002)	(0.005)	(0.010)
No. households:	77,118	18,325	14,176
Panel-year observations:	534,006	101,067	64,571

Notes: Income measured in 0000s euros. Estimates for the group younger than 65 restricts to households whose oldest member is younger than 65 and born after 1949. Estimates for the group older than age 65 restricts to households whose youngest member is older than 65 and born before 1950. Standard errors (in parentheses) are clustered at the household level.

LTC provision (τ_L), we estimate the following specification:

$$\tau_x(y,\cdot) = \alpha_{0,x} + \frac{\alpha_{1,x} - \alpha_{0,x}}{1 + e^{-\left(\frac{y-\alpha_{2,x}}{\alpha_{3,x}}\right)}}, \quad x \in \{LTC,\ SS\}, \tag{D.13}$$

which we estimate conditional upon age group (below vs. above age 65) and family structure (married vs. single). $\alpha_{1,x}$ represents the maximum tax amount, which is present in the Dutch system. Table D.2 shows the estimation results.

Table D.2: Parameters of the Dedicated Tax Functions τ_L and τ_{SS}

	Couples		Singles
	Below age 65	Above age 65	Above age 65
Pension income ($x = SS$)			
α_0	-0.255		
	(0.013)		
α_1	0.697		
	(0.002)		
α_2	3.259		
	(0.041)		
α_3	1.566		
	(0.022)		
LTC provision ($x = LTC$)			
α_0	-0.166	-0.060	-0.026
	(0.008)	(0.004)	(0.005)
α_1	0.447	0.378	0.303
	(0.001)	(0.003)	(0.003)
α_2	3.268	3.510	2.578
	(0.004)	(0.001)	(0.002)
α_3	1.599	0.872	0.618
	(0.230)	(0.021)	(0.021)
No. households:	77,118	18,325	14,176
Panel-year observations:	534,006	101,067	64,571

Notes: Income measured in 0000s euros. Estimates for the group younger than 65 are restricted to households whose oldest member is younger than 65 and born after 1949. Estimates for the group older than 65 are restricted to households whose youngest member is older than 65 and born before 1950. Standard errors are clustered at the household level (in parentheses).

D.2.7 Cohort Effects to the Asset Profiles

Akin to estimating the income processes before age 65, we have to deal with cohort effects to observed asset profiles. In the cross-section (a given year), older households are born in an earlier year than younger households and, due to secular income growth, have a lower labor productivity level and pension income. Because of this, asset levels of older cohorts will likely be lower. At the same time, assets of older cohorts may be higher because they include more former entrepreneurs, such as farmers. Computing age profiles of assets unconditionally upon birth cohort would consist of these undesired cohort effects.

To obtain asset profiles without cohort effects, we follow French (2005) and run specifications (4.2a) and (4.2b) with the logarithm of assets a_{it} as outcome:

$$\log(a_{it}) = \log(\alpha_{t,w}) + \theta_{i,w} + \epsilon_{it,w}, \tag{D.14a}$$

where i indexes a household and t is the age of the household, i.e., the age of the oldest household member. This age ranges from 65 to 100. w is a subscript to distinguish these parameters involving assets from those involving income in specifications (4.2a) and (4.2b). To wash out cohort effects, we run the following OLS regression of the predicted fixed effects on birth cohort dummies (cf. French, 2005; De Nardi et al., 2024):

$$\widehat{\theta}_{i,w} = \overline{\theta_w} + \overline{\theta}_{c,w} + \widetilde{\theta}_{i,w},\ c \in \{1905, 1906, ..., 1944, 1945 - 1949\}, \tag{D.14b}$$

where $\overline{\theta_w}$ is the cohort effect of birth years 1945-1949, $\overline{\theta_w} + \overline{\theta}_{c,w}$ is the fixed effect for the other cohorts, and residual $\widetilde{\theta}_i$ is the household-specific effect excluding a cohort effect. To align with the income process before age 65 being tailored to households born in 1950, we take the cohort born between 1945 and 1949, as the reference group. In the ideal econometric scenario, we have $\overline{\theta}_{c,w} = 0$ so no cohort effects. To mimic this, we subtract the estimated cohort effect $\widehat{\widehat{\theta}}_{c,w}$ from the right-hand side of (D.14a):

$$\log(\widehat{a}_{it}) = \log(\widehat{\alpha}_{t,w}) + \widehat{\theta}_{i,w} + \widehat{\epsilon}_{it,w} - \widehat{\widehat{\theta}}_{c,w}. \tag{D.14c}$$

$\log(\widehat{a}_{it})$ is the predicted asset level for the household when they would be born between 1945 and 1949. To allow for distinct age patterns by marital status and lifetime income, we run the regressions for these groups separately.

We exponentiate the assets to get the asset level that is cleaned from cohort effects. While the regression omits zero assets, we re-include them in the 'cleaned' profiles; negative assets and assets above €2,500,000 are dropped.[4][5]

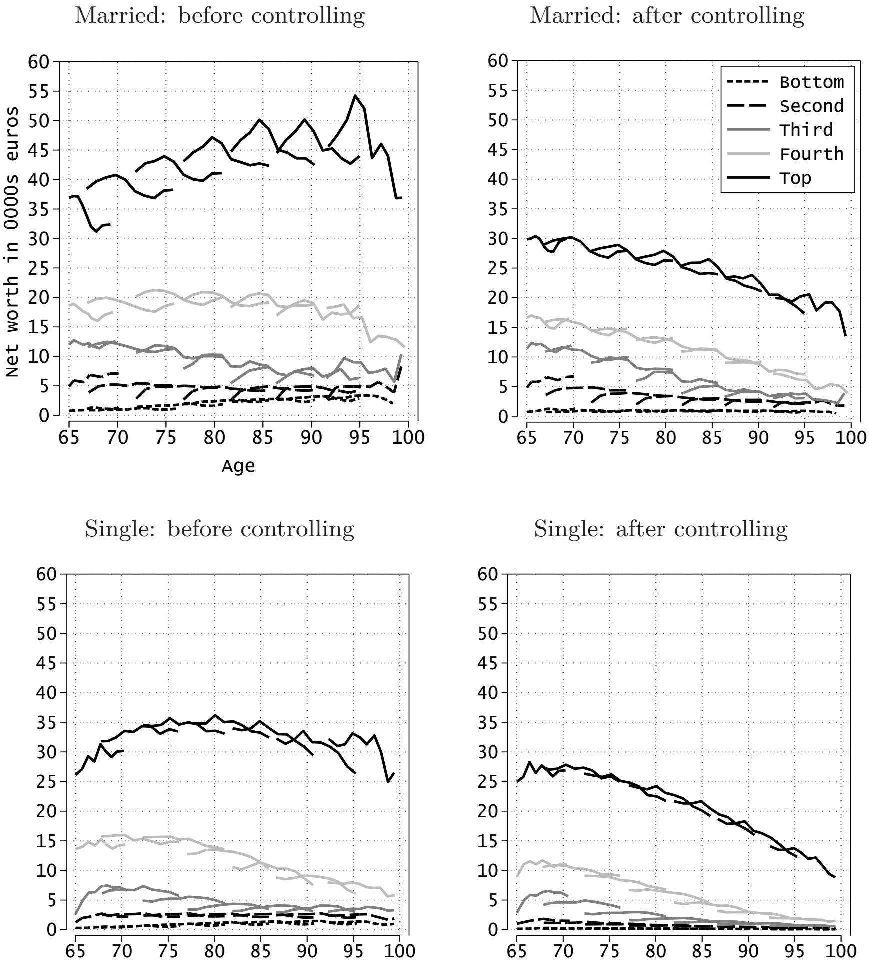
Figure D.5 shows median asset profiles before and after we control for birth cohort effects. Each separate line represents a different birth cohort, depending on the age in 2006. The left panels a. and c., i.e., the raw data, reveal that birth cohort effects are strong, particularly for married households with high lifetime income. Those households have more assets if they are born earlier. Furthermore, within birth cohorts, there seems to be a strong time trend, induced by the period of financial crisis that is part of our observational window.

Using (D.14c), a birth cohort effect is controlled for in panels b. and d.. This reverses the differences between cohorts: the youngest cohorts hold most assets and asset profiles of different cohorts nicely overlap. Also, year trends are less pronounced. As a consequence, we observe households decumulating asset holdings over time. The asset profiles in Figure 4.1 in Section 4.5, which we target, are the data from panels b. and d. unconditional upon birth cohort.

[4]We drop 0.9% of the households and 2.6% of the panel-year observations because of these restrictions.

[5]We also tried Deaton-Paxson dummies, but identifying the effects suffers heavily from multicollinearity. Also, taking levels as outcome could not properly control for many zero assets in the data.

Figure D.5: Asset Profiles Before and After Controlling for Birth Cohort Fixed Effects



Notes: Each line represents the asset profile conditional upon birth cohort and income quintile. We distinguish seven birth cohorts based on the age of the household in 2006: younger than 65; aged 65-69; aged 70-74; aged 75-79; aged 80-84; aged 85-89; and aged 90 and over.

D.3 Second-stage Estimates

D.3.1 Standard Errors of Estimated Preference Parameters

We compute standard errors of $\widehat{\boldsymbol{\delta}}$ by using a matrix $\boldsymbol{D}$ that measures the responsiveness of each moment condition to slightly changing the parameter estimate. Specifically, $\boldsymbol{D}$ is a $k \times 4$ dimensional matrix where the k-th row contains the derivative of the

k-th moment condition: $\frac{\partial\left(M_k^d - M_k^s(\hat{\boldsymbol{\chi}}, \boldsymbol{\delta})\right)}{\partial \boldsymbol{\delta}}$. The variance-covariance $\boldsymbol{V}$ of estimator $\widehat{\boldsymbol{\delta}}$ is documented in De Nardi et al. (2010): $\boldsymbol{V} = (\boldsymbol{D}'\boldsymbol{D})^{-1}(\boldsymbol{D}'\boldsymbol{S}\boldsymbol{D})(\boldsymbol{D}'\boldsymbol{D})^{-1}$, where $\boldsymbol{S}$ is the empirical variance-covariance matrix regarding the data moments. We compute $\boldsymbol{D}$ numerically.

Appendix E: Chapter 5

E.1 Derivation of $\mathcal{L}^{(\overline{D})}$

$\mathcal{L}^{(\overline{D})}$ has a recursive pattern due to its link to the Gamma function. To reach this result, define $y = \sum_{j=1}^{J} M\left(t_j, \boldsymbol{X}_j, \widetilde{\boldsymbol{x}}_j\right)$. Then:

$$\begin{aligned}
\mathcal{L}^{(0)}(y) &= \mathcal{L}(y) = \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}} \\
&= (-1)^0 \cdot \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}-0} \cdot 1
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}^{(1)}(y) &= \frac{\partial \mathcal{L}(y)}{\partial y} = \left(-\frac{1}{\sigma^2} - 0\right) \cdot \sigma^2 \cdot \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}-1} \\
&= -1 \cdot \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}-1} \cdot 1 \cdot \left(0\sigma^2 + 1\right) \\
\mathcal{L}^{(2)}(y) &= \frac{\partial \mathcal{L}(y)}{\partial y^2} = -1 \cdot -1 \cdot \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}-2} \cdot 1 \cdot \left(0\sigma^2 + 1\right) \cdot \left(\sigma^2 + 1\right)
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}^{(3)}(y) &= \frac{\partial \mathcal{L}(y)}{\partial y^3} = -1 \cdot -1 \cdot -1 \cdot \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}-3} \cdot 1 \cdot \left(0\sigma^2 + 1\right) \cdot \left(\sigma^2 + 1\right) \cdot \left(2\sigma^2 + 1\right) \\
&\vdots \\
\mathcal{L}^{(\overline{D})}(y) &= \frac{\partial \mathcal{L}(y)}{\partial y^{\overline{D}}} = (-1)^{\overline{D}} \cdot \left(\sigma^2 \cdot y + 1\right)^{-\frac{1}{\sigma^2}-\overline{D}} \cdot \prod_{q=0}^{\left(\overline{D}-1\right)_+} \left(q\sigma^2 + 1\right)
\end{aligned}$$

E.2 Score Functions and Fisher Information Matrix

Individual Score vectors

We have the unknown parameter vectors $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ and σ^2. Then, by applying the chain rule the score functions are as follows:

$$
\begin{aligned}
\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\beta}} &= \sum_{j=1}^{J} \sum_{s=1}^{S_j} \frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \theta_{j1}\left(t_j^{(s-1)}\right)} \boldsymbol{x}\left(t_j^{(s-1)}\right) = \sum_{j=1}^{J} D_j \boldsymbol{x}\left(t_j^{(S_j-1)}\right) \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \frac{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} \Big\} M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\Big\}}{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} \Big\} M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\Big\} + 1} \boldsymbol{x}\left(t_j^{(s-1)}\right) \\
&+ \frac{1}{\exp\left(\theta_3\right)} \cdot \frac{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right)}{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right) + 1} \boldsymbol{x}\left(t^{(0)}\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\gamma}} &= \sum_{j=1}^{J} \frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \theta_{j2}} \widetilde{\boldsymbol{x}}_j = \sum_{j=1}^{J} D_j \cdot \frac{\frac{\partial \lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)}{\partial \theta_{j2}}}{\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)} \widetilde{\boldsymbol{x}}_j \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \frac{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} \Big\} \frac{\partial M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)}{\partial \theta_{j2}} - \frac{\partial M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)}{\partial \theta_{j2}} \Big\}}{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} \Big\} M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\Big\} + 1} \widetilde{\boldsymbol{x}}_j \\
&+ \frac{1}{\exp\left(\theta_3\right)} \cdot \frac{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \frac{\partial M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right)}{\partial \theta_{j2}}}{\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right) + 1} \widetilde{\boldsymbol{x}}_j
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\theta_3} &= \sum_{q=0}^{\left(\overline{D}-1\right)_+}\frac{q\cdot\exp\left(\theta_3\right)}{q\cdot\exp\left(\theta_3\right)+1}\\
&+\frac{1}{\exp\left(\theta_3\right)}\cdot\ln\left(\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}\Big\}M\left(t_j^{(s)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)-M\left(t_j^{(s-1)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)\Big\}+1\right)\\
&-\left(\frac{1}{\exp\left(\theta_3\right)}+\overline{D}\right)\cdot\frac{\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}\Big\}M\left(t_j^{(s)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)-M\left(t_j^{(s-1)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)\Big\}}{\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}\Big\}M\left(t_j^{(s)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)-M\left(t_j^{(s-1)},\theta_{j1}\left(t_j^{(s-1)}\right),\theta_{j2}\right)\Big\}+1}\\
&-\frac{1}{\exp\left(\theta_3\right)}\cdot\ln\left(\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}M\left(t^{(1)},\theta_{j1}\left(t^{(0)}\right),\theta_{j2}\right)+1\right)+\frac{1}{\exp\left(\theta_3\right)}\cdot\frac{\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}M\left(t^{(1)},\theta_{j1}\left(t^{(0)}\right),\theta_{j2}\right)}{\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}M\left(t^{(1)},\theta_{j1}\left(t^{(0)}\right),\theta_{j2}\right)+1}
\end{aligned}
$$

The score vectors are a product of a scalar and covariate vector. This format allows us to use the efficient `MATA` routine `mlvecsum` to compute these score vectors.

Information Matrices

For the information matrices it is important to note that the score vectors themselves are vectors and consist of sums over the J spells and within spell-variation. The information matrices will contain components that are of the form:

$$\frac{\partial l}{\partial \boldsymbol{\beta_1'} \partial \boldsymbol{\beta_2}} = \sum_{j=1}^{J}\sum_{s=1}^{S_j} c_{sj}\boldsymbol{x}_{1sj}\boldsymbol{x}'_{2sj} + \left(\sum_{j=1}^{J}\sum_{s=1}^{S_j} a_{sj}\right)\left(\sum_{j=1}^{J}\sum_{s=1}^{S_j} b_{sj}\boldsymbol{x}_{1sj}\right)\left(\sum_{j=1}^{J}\sum_{s=1}^{S_j} b_{sj}\boldsymbol{x}'_{2sj}\right),$$

where a_{sj}, b_{sj} and c_{sj} are scalars. We can calculate the first component, the spell-level outer product matrices, with the routine `mlmatsum`. We can calculate the second component, the group-level outer product matrices, with the routine `mlmatbysum`. We obtain:

$$\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'} = \sum_{j=1}^{J}\sum_{s=1}^{S_j}\frac{\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\beta}}}{\partial\theta_{j1}\left(t_j^{(s-1)}\right)}\boldsymbol{x}\left(t_j^{(s-1)}\right)'$$

$$= -\left(\frac{1}{\exp\left(\theta_3\right)}+\overline{D}\right)\cdot\frac{\exp\left(\theta_3\right)}{\Psi_1}\cdot\left(\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}\boldsymbol{x}\left(t_j^{(s-1)}\right)\boldsymbol{x}\left(t_j^{(s-1)}\right)'-\frac{\exp\left(\theta_3\right)}{\Psi_1}\left(\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}\boldsymbol{x}\left(t_j^{(s-1)}\right)\right)\left(\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}\boldsymbol{x}\left(t_j^{(s-1)}\right)'\right)\right)$$

$$+\frac{1}{\Psi_2}\cdot\left(\sum_{j=1}^{J}c_j\boldsymbol{x}\left(t^{(0)}\right)\boldsymbol{x}\left(t^{(0)}\right)'-\frac{\exp\left(\theta_3\right)}{\Psi_2}\sum_{j=1}^{J}c_j\boldsymbol{x}\left(t^{(0)}\right)\sum_{j=1}^{J}c_j\boldsymbol{x}\left(t^{(0)}\right)'\right)$$

$$\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\beta}\partial\theta_3} = \exp\left(\theta_3\right)\cdot\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\beta}\partial\exp\left(\theta_3\right)}$$

$$= \frac{1}{\Psi_1}\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}\boldsymbol{x}\left(t_j^{(s-1)}\right)-\left(\frac{1}{\exp\left(\theta_3\right)}+\overline{D}\right)\cdot\frac{1}{\Psi_1^2}\cdot\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}\boldsymbol{x}\left(t_j^{(s-1)}\right)$$

$$-\frac{1}{\Psi_2}\cdot\sum_{j=1}^{J}c_j\boldsymbol{x}\left(t^{(0)}\right)+\frac{1}{\Psi_2^2}\cdot\sum_{j=1}^{J}c_j\boldsymbol{x}\left(t^{(0)}\right)$$

$$\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\gamma}\partial\theta_3} = \exp\left(\theta_3\right)\cdot\frac{\partial l\left(\boldsymbol{t},\boldsymbol{t}_0,\boldsymbol{X},\widetilde{\boldsymbol{x}},\boldsymbol{\beta},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\gamma}\partial\exp\left(\theta_3\right)}$$

$$= \frac{1}{\Psi_1}\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}'\widetilde{\boldsymbol{x}}_j-\left(\frac{1}{\exp\left(\theta_3\right)}+\overline{D}\right)\cdot\frac{1}{\Psi_1^2}\cdot\exp\left(\theta_3\right)\cdot\sum_{j=1}^{J}\sum_{s=1}^{S_j}c_{sj}'\widetilde{\boldsymbol{x}}_j-\frac{1}{\Psi_2}\cdot\sum_{j=1}^{J}c_j'\boldsymbol{x}\left(t^{(0)}\right)+\frac{1}{\Psi_2^2}\cdot\sum_{j=1}^{J}c_j'\boldsymbol{x}\left(t^{(0)}\right)$$

$$
\begin{aligned}
\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}'} &= \sum_{j=1}^{J} \frac{\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\gamma}}}{\partial \theta_{j2}} \widetilde{\boldsymbol{x}}_j' \\
&= \sum_{j=1}^{J} D_j \cdot \frac{\frac{\partial \lambda_0\left(t^{(S_j)}, \theta_{j2}\right)}{\partial \theta_{j2}^2} \cdot \lambda_0\left(t^{(S_j)}, \theta_{j2}\right) - \Big\} \frac{\partial \lambda_0\left(t^{(S_j)}, \theta_{j2}\right)}{\partial \theta_{j2}} \Big\}^2}{\lambda_0\left(t^{(S_j)}, \theta_{j2}\right)^2} \widetilde{\boldsymbol{x}}_j \widetilde{\boldsymbol{x}}_j' \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \frac{\exp\left(\theta_3\right)}{\Psi_1} \cdot \left(\sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj}'' \widetilde{\boldsymbol{x}}_j \widetilde{\boldsymbol{x}}_j' - \frac{\exp\left(\theta_3\right)}{\Psi_1} \cdot \left(\sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj}' \widetilde{\boldsymbol{x}}_j\right) \left(\sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj}' \widetilde{\boldsymbol{x}}_j'\right)\right) \\
&+ \frac{1}{\Psi_2} \cdot \left(\sum_{j=1}^{J} c_j'' \widetilde{\boldsymbol{x}}_j \widetilde{\boldsymbol{x}}_j' - \frac{\exp\left(\theta_3\right)}{\Psi_2} \sum_{j=1}^{J} c_j' \widetilde{\boldsymbol{x}}_j \sum_{j=1}^{J} c_j' \widetilde{\boldsymbol{x}}_j'\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\gamma}'} &= \left(\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\beta}'}\right)' = \sum_{j=1}^{J} \frac{\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \boldsymbol{\beta}}}{\partial \theta_{j2}} \widetilde{\boldsymbol{x}}_j' \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \frac{\exp\left(\theta_3\right)}{\Psi_1} \cdot \left(\sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj}' \boldsymbol{x}\left(t_j^{(s-1)}\right) \widetilde{\boldsymbol{x}}_j' - \frac{\exp\left(\theta_3\right)}{\Psi_1} \cdot \left(\sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj} \boldsymbol{x}\left(t_j^{(s-1)}\right)\right) \left(\sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj}' \widetilde{\boldsymbol{x}}_j'\right)\right) \\
&+ \frac{1}{\Psi_2} \cdot \left(\sum_{j=1}^{J} c_j' \boldsymbol{x}\left(t_j^{(s-1)}\right) \widetilde{\boldsymbol{x}}_j' - \frac{\exp\left(\theta_3\right)}{\Psi_2} \sum_{j=1}^{J} c_j \boldsymbol{x}\left(t_j^{(s-1)}\right) \sum_{j=1}^{J} c_j' \widetilde{\boldsymbol{x}}_j'\right)
\end{aligned}
$$

$$\frac{\partial l\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)}{\partial \theta_3 \partial \theta_3} = \sum_{q=0}^{\left(\overline{D}-1\right)_+} \frac{q \cdot \exp\left(\theta_3\right)}{\left(q \cdot \exp\left(\theta_3\right)+1\right)^2}$$

$$- \frac{1}{\exp\left(\theta_3\right)} \cdot \ln\left(\Psi_1\right) + \frac{2}{\Psi_1} \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj} - \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \exp\left(\theta_3\right) \cdot \frac{1}{\Psi_1^2} \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} c_{sj}$$

$$+ \frac{1}{\exp\left(\theta_3\right)} \cdot \ln\left(\Psi_2\right) - \frac{2}{\Psi_2} \cdot \sum_{j=1}^{J} c_j + \frac{1}{\Psi_2^2} \cdot \sum_{j=1}^{J} c_j, \text{ with:}$$

$$c_{sj} = M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right),$$

$$c_j = M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right),$$

$$c'_{sj} = \frac{\partial M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)}{\partial \theta_{j2}} - \frac{\partial M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)}{\partial \theta_{j2}},$$

$$c'_j = \frac{\partial M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right)}{\partial \theta_{j2}},$$

$$c''_{sj} = \frac{\partial M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)}{\partial \theta_{j2}^2} - \frac{\partial M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)}{\partial \theta_{j2}^2},$$

$$c''_j = \frac{\partial M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right)}{\partial \theta_{j2}^2},$$

$$\Psi_1 = \exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=1}^{S_j} \Big\} M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) \Big\} + 1,$$

$$\Psi_2 = \exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right) + 1.$$

E.3 Observed Threshold Values Approaching Zero

We will show that $l_{\mathtt{A}}$ and $l_{\mathtt{B}}$ are asymptotically the same if the starting times of the observed population $t_j^{(0)} \to 0$ for all j $\in \{1, .., J\}$. This case of $\boldsymbol{t}_0 \to \boldsymbol{0}$ not only involves the possibility of no left truncation, but also truncation schemes implying that only subjects with exclusively low thresholds are sampled. This for example happens if frailty is shared across many spells which all have to meet a threshold $t_j^{(0)}$; lower $t_j^{(0)}$ increases the likelihood of being sampled as a subject. Proof for $\lim_{\boldsymbol{t}_0 \to 0} l_{\mathtt{A}} = l_{\mathtt{B}}$:

$$
\begin{aligned}
&\lim_{\boldsymbol{t}_0 \to 0} l_{\mathtt{A}} \left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right) = \lim_{\boldsymbol{t}_0 \to 0} \sum_{q=0}^{\left(\overline{D}-1\right)_+} \ln\left(q \cdot \exp\left(\theta_3\right) + 1\right) \\
&+ \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\theta_{j1}\left(t_j^{(S_j-1)}\right)\right)\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)\right) \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \ln\left(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J}\sum_{s=1}^{S_j}\left\{M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\right.\right. \\
&\left.\left. - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\right\} + 1\right) \\
&+ \frac{1}{\exp\left(\theta_3\right)} \cdot \ln\left(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right) + 1\right) \\
&= \sum_{q=0}^{\left(\overline{D}-1\right)_+} \ln\left(q \cdot \exp\left(\theta_3\right) + 1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\theta_{j1}\left(t_j^{(S_j-1)}\right)\right)\right) \\
&+ \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)\right) - \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \\
&\cdot \ln\left(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J}\left\{M\left(\lim_{\boldsymbol{t}_0 \to 0} t_j^{(1)}, \theta_{j1}\left(t_j^{(0)}\right), \theta_{j2}\right) - M\left(t_j^{(0)}, \theta_{j1}\left(t_j^{(0)}\right), \theta_{j2}\right)\right\}\right. \\
&\left. + \exp\left(\theta_3\right) \cdot \sum_{j=1}^{J}\sum_{s=2}^{S_j}\left\{M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right)\right\} + 1\right) \\
&+ \frac{1}{\exp\left(\theta_3\right)} \cdot \ln\left(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(\lim_{\boldsymbol{t}_0 \to 0} t^{(1)}, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right) + 1\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{q=0}^{\left(\overline{D}-1\right)_{+}} \ln\left(q \cdot \exp\left(\theta_3\right) + 1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\theta_{j1}\left(t_j^{(S_j-1)}\right)\right)\right) \\
&+ \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)\right) - \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \\
&\cdot \ln\Bigg(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \left\{ M\left(t_j^{(0)}, \theta_{j1}\left(t_j^{(0)}\right), \theta_{j2}\right) - M\left(t_j^{(0)}, \theta_{j1}\left(t_j^{(0)}\right), \theta_{j2}\right) \right\} \\
&+ \exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=2}^{S_j} \left\{ M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) - M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) \right\} + 1 \Bigg) \\
&+ \frac{1}{\exp\left(\theta_3\right)} \cdot \ln\left(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} M\left(0, \theta_{j1}\left(t^{(0)}\right), \theta_{j2}\right) + 1 \right) \\
&= \sum_{q=0}^{\left(\overline{D}-1\right)_{+}} \ln\left(q \cdot \exp\left(\theta_3\right) + 1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\theta_{j1}\left(t_j^{(S_j-1)}\right)\right)\right) \\
&+ \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)\right) \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \ln\Bigg(\exp\left(\theta_3\right) \cdot 0 + \exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=2}^{S_j} \Big\{ M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) \\
&- M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) \Big\} + 1 \Bigg) \\
&+ \frac{1}{\exp\left(\theta_3\right)} \cdot \ln\left(\exp\left(\theta_3\right) \cdot 0 + 1\right) \\
&= \sum_{q=0}^{\left(\overline{D}-1\right)_{+}} \ln\left(q \cdot \exp\left(\theta_3\right) + 1\right) + \sum_{j=1}^{J} D_j \cdot \ln\left(\phi\left(\theta_{j1}\left(t_j^{(S_j-1)}\right)\right)\right) \\
&- \left(\frac{1}{\exp\left(\theta_3\right)} + \overline{D}\right) \cdot \ln\Bigg(\exp\left(\theta_3\right) \cdot \sum_{j=1}^{J} \sum_{s=2}^{S_j} \Big\{ M\left(t_j^{(s)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) \\
&- M\left(t_j^{(s-1)}, \theta_{j1}\left(t_j^{(s-1)}\right), \theta_{j2}\right) \Big\} + 1 \Bigg) + \sum_{j=1}^{J} D_j \cdot \ln\left(\lambda_0\left(t_j^{(S_j)}, \theta_{j2}\right)\right) \\
&= l_{\mathsf{B}}\left(\boldsymbol{t}, \boldsymbol{t}_0, \boldsymbol{X}, \widetilde{\boldsymbol{x}}, \boldsymbol{\beta}, \boldsymbol{\gamma}\right)
\end{aligned}
$$

with: $t_j^{(0)} = 0;\ \lim_{\boldsymbol{t}_0 \to 0} t_j^{(1)} = \lim_{\boldsymbol{t}_0 \to 0} t_{0j} = 0 = t_j^{(0)};\ t_j^{(S_j)} = t_j,\ \theta_{j1}\left(t\right) = \boldsymbol{x}_j\left(t\right)' \boldsymbol{\beta},$

$\theta_{j2} = \widetilde{\boldsymbol{x}}_j' \boldsymbol{\gamma},$ and $\theta_3 = \ln\left(\sigma^2\right),$

Bibliography

Abbring, J. H. and van den Berg, G. J. (2007). The unobserved heterogeneity distribution in duration analysis. *Biometrika*, 94(1):87–99.

Algemene Rekenkamer (2015). Regionale verschillen in langdurige zorg: mogelijke verklaringen voor zorggebruik van ouderen en chronisch zieken. Policy report.

Altonji, J. G. and Segal, L. M. (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics*, 14(3):353–366.

Ameriks, J., Briggs, J., Caplin, A., Shapiro, M. D., and Tonetti, C. (2018). The long-term-care insurance puzzle: modeling and measurement. NBER Working Papers no. 22726, National Bureau of Economic Research.

Ameriks, J., Briggs, J., Caplin, A., Shapiro, M. D., and Tonetti, C. (2020). Long-term-care utility and late-in-life saving. *Journal of Political Economy*, 128(6):2375–2451.

Ameriks, J., Caplin, A., Laufer, S., and van Nieuwerburgh, S. (2011). The joy of giving or assisted living? Using strategic surveys to separate public care aversion from bequest motives. *The Journal of Finance*, 66(2):519–561.

Arellano, M. (2003). *Panel Data Econometrics.* Oxford University Press, Oxford (United Kingdom).

Attanasio, O. P. and Emmerson, C. (2003). Mortality, health status, and wealth. *Journal of the European Economic Association*, 1(4):821–850.

Auerbach, A. J., Charles, K. K., Coile, C. C., Gale, W., Goldman, D., Lee, R., Lucas, C. M., Orszag, P. R., Sheiner, L. M., Tysinger, B., Weil, D. N., Wolfers, J., and Wong, R. (2017). How the growing gap in life expectancy may affect retirement benefits and reforms. NBER Working Papers no. 23329, National Bureau of Economic Research.

Bagchi, S. (2019). Differential mortality and the progressivity of social security. *Journal of Public Economics*, 177:104044.

Bakx, P., de Meijer, C., Schut, F., and van Doorslaer, E. (2015). Going formal or informal, who cares? The influence of public long-term care insurance. *Health Economics*, 24(6):631–643.

Bakx, P., Van Doorslaer, E., and Wouterse, B. (2023). Long-term care in the Netherlands. NBER Working Papers no. 31823, National Bureau of Economic Research.

Bakx, P., Wouterse, B., van Doorslaer, E., and Wong, A. (2020). Better off at home? Effects of nursing home eligibility on costs, hospitalizations and survival. *Journal of Health Economics*, 73(1):102354.

Balan, A. and Putter, H. (2019). FrailtyEM: an R package for estimating semiparametric shared frailty models. *Journal of Statistical Software*, 90(7):1–29.

Baqui, P., Bica, I., Marra, V., Ercole, A., and van der Schaar, M. (2020). Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *The Lancet Global Health*, 8(8):1018–1026.

Barbi, E., Lagona, F., Marsili, M., Vaupel, J. W., and Wachter, K. W. (2018). The plateau of human mortality: demography of longevity pioneers. *Science*, 360(6396):1459–1461.

Barczyk, D. and Kredler, M. (2018). Evaluating long-term-care policy options, taking the family seriously. *The Review of Economic Studies*, 85(2):766–809.

Bender, R., Augustin, T., and Blettner, M. (2005). Generating survival times to simulate Cox proportional hazards models. *Statistics in Medicine*, 24(11):1713–1723.

Bergeot, J. and Tenand, M. (2023). Does informal care delay nursing home entry? Evidence from Dutch linked survey and administrative data. *Journal of Health Economics*, 92:102831.

Bernheim, B. D. (1987). The economic effects of social security: toward a reconciliation of theory and measurement. *Journal of Public Economics*, 33(3):273–304.

Bhattacharya, J. and Lakdawalla, D. (2006). Does Medicare benefit the poor? *Journal of Public Economics*, 90(1):277–292.

Blundell, R., Graber, M., and Mogstad, M. (2015). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics*, 127:58–73.

Bolin, K., Lindgren, B., and Lundborg, P. (2008). Informal and formal care among single-living elderly in Europe. *Health Economics*, 17(3):393–409.

Bom, J., Bakx, P., Schut, F., and van Doorslaer, E. (2019). The impact of informal caregiving for older adults on the health of various types of caregivers: a systematic review. *The Gerontologist*, 59(5):629–642.

Bonsang, E. (2009). Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, 28(1):143–154.

Bosworth, B., Burtless, G., and Zhang, K. (2016). Later retirement, inequality in old age, and the growing gap in longevity between rich and poor. *Economic Studies at Brookings*.

Boyer, M. M., De Donder, P., Fluet, C., Leroux, M.-L., and Michaud, P.-C. (2020). Long-term care insurance: Information frictions and selection. *American Economic Journal: Economic Policy*, 12(3):134–169.

Braun, R. A., Kopecky, K. A., and Koreshkova, T. (2017). Old, sick, alone, and poor: a welfare analysis of old-age social insurance programmes. *The Review of Economic Studies*, 84(2):580–612.

Braun, R. A., Kopecky, K. A., and Koreshkova, T. (2019). Old, frail, and uninsured: accounting for features of the U.S. long-term care insurance market. *Econometrica*, 87(3):981–1019.

Brown, J. and Finkelstein, A. (2007). Why is the market for long-term care insurance so small? *Journal of Public Economics*, 91(10):1967–1991.

Brown, J. and Finkelstein, A. (2008). The interaction of public and private insurance: medicaid and the long-term care insurance market. *American Economic Review*, 98(3):1083–1102.

Brown, J. and Warshawsky, M. (2013). The life care annuity: a new empirical examination of an insurance innovation that addresses problems in the markets for life annuities and long-term care Insurance. *Journal of Risk and Insurance*, 80(3):677–704.

Browning, M., Chiappori, P.-A., and Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *The Review of Economic Studies*, 80(4 (285)):1267–1303.

Bär, M., Bakx, P., Wouterse, B., and van Doorslaer, E. (2022). Estimating the health value added by nursing homes. *Journal of Economic Behavior & Organization*, 203(1):1–23.

Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics*, 21(3):339–353.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly Journal of Economics*, 112(1):1–55.

Case, A. and Deaton, A. (2017). Mortality and Morbidity in the 21st Century. *Brookings Papers on Economic Activity*, pages 397–443.

Case, A. and Paxson, C. (2005). Sex differences in morbidity and mortality. *Demography*, 42(2):189–214.

CBS (2014). Monitor langdurige zorg. Personen met indicatie zorg met verblijf. Centraal Bureau voor de Statistiek.

Charles, K. K. and Sevak, P. (2005). Can family caregiving substitute for nursing home care? *Journal of Health Economics*, 24(6):1174–1190.

Chen, A., Chen, Y., and Xu, X. (2022a). Care-dependent tontines. *Insurance: Mathematics and Economics*, 106:69–89.

Chen, Y., Persson, P., and Polyakova, M. (2022b). The roots of health inequality and the value of intrafamily expertise. *American Economic Journal: Applied Economics*, 14(3):185–223.

Cherchye, L., Demuynck, T., De Rock, B., Kovaleva, M., Minne, G., Perea, M. D. S., and Vermeulen, F. (2023). Poor and wealthy hand-to-mouth households in Belgium. *Review of Economics of the Household*.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., and Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. *JAMA*, 315(16):1750–1766.

Chodorow-Reich, G. and Coglianese, J. (2021). Projecting unemployment durations: a factor-flows simulation approach with application to the COVID-19 recession. *Journal of Public Economics*, 197(5):104398.

CIZ (2013). Jaardocument 2013. Centrum Indicatiestelling Zorg (Wlz).

CIZ (2014). CIZ indicatiewijzer. Centrum Indicatiestelling Zorg (Wlz).

Coe, N. B., Skira, M. M., and Van Houtven, C. H. (2015). Long-term care insurance: does experience matter? *Journal of Health Economics*, 40(3):122–131.

Colombo, F., Llena-Nozal, A., Mercier, J., and Tjadens, F. (2011). *Help wanted?: Providing and paying for long-term care*. OECD Health Policy Studies. OECD, Paris.

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 2(34):187–220.

Crowther, M. and Lambert, P. (2017). Parametric multistate survival models: flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences. *Statistics in Medicine*, 36(29):4719–4742.

CVZ (2013). Wachtlijstonderzoek AWBZ. Factoren die van invloed zijn op de betrouwbaarheid van wachtlijstinformatie. Policy Report, College voor Zorgverzekeringen.

Davidoff, T. (2009). Housing, health, and annuities. *Journal of Risk and Insurance*, 76(1):31–52.

de Biase, P. and Dougherty, S. (2023). From local to national: Delivering and financing effective long-term care. OECD Working Papers on Fiscal Federalism, OECD.

de Bresser, J., Knoef, M., and van Ooijen, R. (2022). Preferences for in-kind and in-cash home care insurance. *Journal of Health Economics*, 84:102626.

de Bresser, J., van Ooijen, R., and Knoef, M. (2023). Willingness to pay for health-dependent annuities to manage longevity and long-term care risk. Working paper.

De Donder, P., Leroux, M.-L., and Salanié, F. (2022). Advantageous selection with moral hazard (with an application to life care annuities). CESifo Working Paper no. 9764.

de Meijer, C., Koopmanschap, M., d' Uva, T. B., and van Doorslaer, E. (2011). Determinants of long-term care spending: age, time to death or disability? *Journal of Health Economics*, 30(2):425–438.

De Nardi, M. (2004). Wealth inequality and intergenerational links. *The Review of Economic Studies*, 71(3):743–768.

De Nardi, M., French, E., and Jones, J. B. (2010). Why do the elderly save? The role of medical expenses. *Journal of Political Economy*, 118(1):39–75.

De Nardi, M., French, E., and Jones, J. B. (2016). Medicaid insurance in old age. *American Economic Review*, 106(11):3480–3520.

De Nardi, M., French, E., Jones, J. B., and McGee, R. (2021). Why do couples and singles save during retirement? Household heterogeneity and its aggregate implications. NBER Working Papers no. 28828, National Bureau of Economic Research.

De Nardi, M., Pashchenko, S., and Porapakkarm, P. (2024). The lifetime costs of bad health. *Review of Economic Studies (forthcoming).*

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health Affairs*, 21(2):13–30.

Denuit, M. and Scaillet, O. (2004). Nonparametric tests for positive quadrant dependence. *Journal of Financial Econometrics*, 2(3):422–450.

Diepstraten, M., Douven, R., and Wouterse, B. (2020). Can your house keep you out of a nursing home? *Health Economics*, 29(5):540–553.

Donaldson, D. and Pendakur, K. (2004). Equivalent-expenditure functions and expenditure-dependent equivalence scales. *Journal of Public Economics*, 88(1):175–208.

Dostie, B. and Léger, P. T. (2005). The living arrangement dynamics of sick, elderly individuals. *The Journal of Human Resources*, 40(4):989–1014.

Dow, W. H. and Rehkopf, D. H. (2010). Socioeconomic gradients in health in international and historical context. *Annals of the New York Academy of Sciences*, 1186(1):24–36.

Duell, D., Koolman, X., and Portrait, F. (2017). Practice variation in the Dutch long-term care and the role of supply-sensitive care: is access to the Dutch long-term care equitable? *Health Economics*, 26(12):1728–1742.

Duell, D., Non, M., Braam, A. M., Leloup, L., and Portrait, F. (2021). Is the availability of informal care associated with a lower uptake in formal home care? An application to personal care in the Netherlands. *Journal of Long-Term Care*, pages 77–91.

Dynan, K. E., Skinner, J., and Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy*, 112(2):397–444.

Einav, L. and Finkelstein, A. (2023). Empirical analyses of selection and welfare in insurance markets: a self-indulgent survey. *The Geneva Risk and Insurance Review*, 48:167–191.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *The Quarterly Journal of Economics*, 125(3):877–921.

Einav, L., Finkelstein, A., and Mahoney, N. (2022). Producing health: measuring value added of nursing homes. NBER Working Paper no. 30228, National Bureau of Economic Research.

Elbers, C. and Ridder, G. (1982). True and spurious duration dependence: the identifiability of the proportional hazard model. *The Review of Economic Studies*, 49(3):403–409.

Eling, M. and Ghavibazoo, O. (2019). Research on long-term care insurance: status quo and directions for future research. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 44:303–356.

European Commision (2015). The 2015 ageing report: economic and budgetary projections for the 28 EU member states (2013-2060). Publications Office of the European Union.

European Commision (2019). Study on exploring the incidence and costs of informal long-term care in the EU. Publications Office of the European Union.

European Commision (2021). State of Health in the EU - The Netherlands - Country Health Profile 2021. Publications Office of the European Union.

Fehr, H., Kallweit, M., and Kindermann, F. (2013). Should pensions be progressive? *European Economic Review*, 63:94–116.

Finkelstein, A., Luttmer, E. F. P., and Notowidigdo, M. J. (2009). Approaches to estimating the health state dependence of the utility function. *American Economic Review*, 99(2):116–121.

Finkelstein, A. and McGarry, K. (2006). Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*, 96(4):938–958.

Finkelstein, A. and Poterba, J. (2004). Adverse selection in insurance markets: policyholder evidence from the U.K. annuity market. *Journal of Political Economy*, 112(1):183–208.

Fluet, C. and Pannequin, F. (1997). Complete versus incomplete insurance contracts under adverse selection with multiple risks. *The Geneva Papers on Risk and Insurance Theory*, 22(2):81–101.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2):395–427.

Fuino, M. and Wagner, J. (2018). Long-term care models and dependence probability tables by acuity level: new empirical evidence from Switzerland. *Insurance: Mathematics and Economics*, 81(1):51–70.

Garcia-Gomez, P. P. ., Hernandez-Pizarro, H. M., Lopez-Casasnovas, G., and Vidiella-Martin, J. (2019). Unravelling hidden inequities in a universal public long-term care system. Tinbergen Institute Discussion Papers 19-011, Tinbergen Institute.

Goda, G., Shoven, J., and Slavov, S. (2011a). Differential Mortality by Income and Social Security Progressivity. In *Explorations in the Economics of Aging*, volume 5, pages 189 – 204. University of Chicago Press, Chicago.

Goda, G. S., Golberstein, E., and Grabowski, D. C. (2011b). Income and the utilization of long-term care services: evidence from the social security benefit notch. *Journal of Health Economics*, 30(4):719–729.

Gollier, C. and Schlesinger, H. (1995). Second-best insurance contract design in an incomplete market. *The Scandinavian Journal of Economics*, 97(1):123–135.

Gopinath, S., Thomas, J. S., and Krishnamurthi, L. (2014). Investigating the relationship between the content of online word of mouth, advertising, and brand performance. *Marketing Science*, 33(2):241–258.

Gorfine, M. and Zucker, D. (2023). Shared frailty methods for complex survival data: a review of recent advances. *Annual Review of Statistics and Its Application*, 10:51–74.

Gould, G., Pitblado, J., and Poi, B. (2010). *Maximum likelihood estimation with Stata*. Stata Press, Texas, 4th edition.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.

Groneck, M. and Wallenius, J. (2021). It sucks to be single! Marital status and redistribution of social security. *The Economic Journal*, 131(633):327–371.

Gutierrez, R. G. (2002). Parametric frailty and shared frailty survival models. *The Stata Journal*, 2(1):22–44.

Hackmann, M. B. and Pohl, R. V. (2018). Patient vs. provider incentives in long term care. NBER Working Paper no. 25178, National Bureau of Economic Research.

Hall, R. E. and Jones, C. I. (2007). The value of life and the rise in health spending*. *The Quarterly Journal of Economics*, 122(1):39–72.

Hartmans, J., Hengveld, H., Vernhout, G., and IJland, C. (2009). Resultaten verdiepend onderzoek naar wachtlijsten voor verpleging & verzorging en gehandicaptenzorg. Policy report, HHM, Enschede.

Heathcote, J., Violante, G., and Storesletten, K. (2020). Presidential address 2019: how should tax progressivity respond to rising income inequality? 18(6):2715–2754.

Heckman, J. (1981). *The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process*. MIT Press, Cambridge.

Hiedemann, B., Sovinsky, M., and Stern, S. (2018). Will you still want me tomorrow? The dynamics of families' long-term care arrangements. *The Journal of Human Resources*, 53(3):663–716.

Honoré, B. E. (1993). Identification results for duration models with multiple spells. *The Review of Economic Studies*, 60(1):241–246.

Hougaard, P. (2000). Shared frailty models. In *Analysis of Multivariate Survival Data*. Springer, New York.

Hussem, A., Tenand, M., and Wouterse, B. (2020). Publieke middelen voor particuliere ouderenzorg. Netspar Design Paper 143, Netspar.

Ilinca, S., Rodrigues, R., and Schmidt, A. E. (2017). Fairness and eligibility to long-term care: an analysis of the factors driving inequality and inequity in the use of home care for older Europeans. *International Journal of Environmental Research and Public Health*, 14(10):1–15.

Jones, J. B., De Nardi, M., French, E., McGee, R., and Kirschner, J. (2018). The lifetime medical spending of retirees. NBER Working Papers no. 24599, National Bureau of Economic Research.

Jones, J. B. and Li, Y. (2023). Social security reform with heterogeneous mortality. *Review of Economic Dynamics*, 48:320–344.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The statistical analysis of failure time data*. John Wiley & Sons, New York.

Kalwij, A., Alessie, R., and Knoef, M. (2013). The association between individual income and remaining life expectancy at the age of 65 in the Netherlands. *Demography*, 50(1):181–206.

Kalwij, A., Pasini, G., and Wu, M. (2014). Home care for the elderly: the role of relatives, friends and neighbors. *Review of Economics of the Household*, 12(2):379–404.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

Karahan, F. and Ozkan, S. (2013). On the persistence of income shocks over the life cycle: evidence, theory, and implications. *Review of Economic Dynamics*, 16(3):452–476.

Knoef, M., Been, J., Alessie, R., Caminada, K., Goudswaard, K., and Kalwij, A. (2016). Measuring retirement savings adequacy: developing a multi-pillar approach in the Netherlands. *Journal of Pension Economics & Finance*, 15(1):55–89.

Knoef, M., Been, J., Rhuggenaath, J., Goudswaard, C., and Caminada, K. (2017). De toereikendheid van pensioenopbouw na de crisis en pensioenhervormingen. Netspar Design Paper no. 68, Netspar.

Ko, A. (2022). An equilibrium analysis of the long-term care insurance market. *The Review of Economic Studies*, 89(4):1993–2025.

Kok, L., Berden, C., and Sadiraj, K. (2015). Costs and benefits of home care for the elderly versus residential care: a comparison using propensity scores. *The European Journal of Health Economics*, 16(2):119–131.

Kopczuk, W. (2013). Taxation of intergenerational transfers and wealth. In *Handbook of Public Economics*, volume 5, pages 329–390. Elsevier, Amsterdam.

Kopecky, K. A. and Suen, R. M. H. (2010). Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics*, 13(3):701–714.

Lambregts, T. R. and Schut, F. T. (2020). Displaced, disliked and misunderstood: a systematic review of the reasons for low uptake of long-term care insurance and life annuities. *The Journal of the Economics of Ageing*, 17.

Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica*, 47(4):939–956.

Lancaster, T. (1990). *The econometric analysis of transition data*. Cambridge University Press, Cambridge.

Lindahl, M., Palme, M., Massih, S. S., and Sjögren, A. (2015). Long-term intergenerational persistence of human capital: an empirical analysis of four generations. *The Journal of Human Resources*, 50(1):1–33.

Lockwood, L. M. (2018). Incidental bequests and the choice to self-insure late-life risks. *American Economic Review*, 108(9):2513–2550.

Low, H. and Meghir, C. (2017). The use of structural models in econometrics. *Journal of Economic Perspectives*, 31(2):33–58.

Luppa, M., Luck, T., Weyerer, S., König, H.-H., Brähler, E., and Riedel-Heller, S. G. (2010). Prediction of institutionalization in the elderly. a systematic review. *Age and Ageing*, 39(1):31–38.

Maarse, J. and Jeurissen, P. (2016). The policy and politics of the 2015 long-term care reform in the Netherlands. *Health Policy*, 120:241–245.

Mackenbach, J. P., Valverde, J. R., Artnik, B., Bopp, M., Brønnum-Hansen, H., Deboosere, P., Kalediene, R., Kovács, K., Leinsalu, M., Martikainen, P., Menvielle, G., Regidor, E., Rychtaříková, J., Rodriguez-Sanz, M., Vineis, P., White, C., Wojtyniak, B., Hu, Y., and Nusselder, W. J. (2018). Trends in health inequalities in 27 European countries. *Proceedings of the National Academy of Sciences*, 115(25):6440–6445. Publisher: Proceedings of the National Academy of Sciences.

McClellan, M. and Skinner, J. (2006). The incidence of Medicare. *Journal of Public Economics*, 90(1):257–276.

McGarry, K. and Schoeni, R. F. (1997). Transfer behavior within the family: results from the asset and health dynamics study. *The Journals of Gerontology. Series B, Psychological sciences and social sciences*, 52:82–92.

Mehlkopf, R., Nielsen, E., and van Dorssen, Y. (2019). Eenmalige lumpsum bij pensionering kent ook nadelige effecten. *Economisch Statistische Berichten*.

Mincer, J. A. (1974). *Schooling, experience, and earnings*. Columbia University Press.

Missov, T. I., Lenart, A., Nemeth, L., Canudas-Romo, V., and Vaupel, J. W. (2015). The Gompertz force of mortality in terms of the modal age at death. *Demographic Research*, 32(1):1031–1048.

Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., and Brown, J. R. (1999). New evidence on the money's worth of individual annuities. *American Economic Review*, 89(5):1299–1318.

Mommaerts, C. (2024). Long-term care insurance and the family. *Journal of Political Economy (forthcoming)*.

Mot, E. (2010). The Dutch system of long-term care. CPB document 204, Centraal Planbureau.

Munda, M., Rotolo, F., and Legrand, C. (2012). parfm: parametric frailty models in R. *Journal of Statistical Software*, 51(11):1–20.

Murtaugh, C., Spillman, B., and Warshawsky, M. (2001). An annuity approach to financing long-term care and retirement income. *Journal of Risk and Insurance*, 68(2):225–254.

Nakajima, M. and Telyukova, I. (2024). Medical expenses and saving in retirement: the case of U.S. and Sweden. *American Economic Journal: Macroeconomics (forthcoming)*.

Nickell, S. (1979). Estimating the probability of leaving unemployment. *Econometrica*, 47(5):1249.

OECD (2011). *Divided We Stand: Why Inequality Keeps Rising*. OECD Publishing.

OECD (2015). *Health at a glance 2015: OECD indicators*. OECD, Paris.

OECD (2019). *Health at a glance 2019: OECD indicators*. OECD, Paris.

OECD (2021). *Pension at a glance 2021: OECD and G20 Indicators*. OECD, Paris.

OECD (2023). *Pension at a glance 2023: OECD and G20 Indicators*. OECD, Paris.

Oksuzyan, A., Juel, K., Vaupel, J. W., and Christensen, K. (2008). Men: good health and high mortality. Sex differences in health and aging. *Aging Clinical and Experimental Research*, 20(2):91–102.

Pauly, M. V. (1990). The rational nonpurchase of long-term-care insurance. *Journal of Political Economy*, 98(1):153–168.

Paz-Pardo, G. and Galves, J. (2023). Richer earnings dynamics, consumption and portfolio choice over the life-cycle. Working Paper.

Picard, P. (2020). Splitting risks in insurance markets with adverse selection. *Journal of Risk and Insurance*, 87(4):997–1033.

Pijoan-Mas, J. and Ríos-Rull, J. (2014). Heterogeneity in expected longevities. *Demography*, 51(6):2075–2102.

Portrait, F., Lindeboom, M., and Deeg, D. (2000). The use of long-term care services by the Dutch elderly. *Health Economics*, 9(6):513–531.

Poterba, J. M. (2014). Retirement security in an aging population. *American Economic Review*, 104(5):1–30.

Powell, J. L. (1994). Estimation of semiparametric models. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics*, volume 4, pages 2443–2521. Elsevier.

Putter, H., Fiocco, M., and Geskus, R. (2007). Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*, 26(11):2389–2430.

Reichling, F. and Smetters, K. (2015). Optimal annuitization with stochastic mortality and correlated medical costs. *American Economic Review*, 105(11):3273–3320.

Ridder, G. and Tunali, I. (1999). Stratified partial likelihood estimation. *Journal of Econometrics*, 92(2):193–232.

Rodrigues, R., Ilinca, S., and Schmidt, A. E. (2018). Income-rich and wealth-poor? The impact of measures of socio-economic status in the analysis of the distribution of long-term care use among older people. *Health Economics*, 27(3):637–646.

Rondeau, V., Gonzalez, J., Mazroui, Y., Mauguen, A., Diakite, A., Laurent, A., Lopez, A., Krol, A., Sofeu, C., Dumerc, J., Rustand, D., Chauvet, J., and Le Coen, Q. (2022). Package 'frailtypack'.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649.

Rouwendal, J. and Thomese, F. (2013). Homeownership and long-term care. *Housing Studies*, 28(5):746–763.

Rouwenhorst, G. (1995). Asset pricing implications of equilibrium business cycle models. In Cooley, T. F., editor, *Frontiers of Business Cycle Research*, pages 294–330. Princeton University Press, New Jersey, Princeton.

Simon, C. and Blume, L. (1994). *Mathematics for economists*. W.W. Norton & Co., New York, 4th edition.

Smith, J. (2007). The impact of socioeconomic status on health over the life-course. *Journal of Human Resources*, 42(4).

Solomon, A. (2022). Bundled risks in insurance markets. Working paper.

Sovinsky, M. and Stern, S. (2016). Dynamic modelling of long-term care decisions. *Review of Economics of the Household*, 14(2):463–488.

St-Amour, P. (2022). Valuing life over the life cycle. Working paper.

Storesletten, K., Telmer, C. I., and Yaron, A. (2004). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics*, 51(3):609–633.

Tenand, M., Bakx, P., and van Doorslaer, E. (2020a). Equal long-term care for equal needs with universal and comprehensive coverage? An assessment using Dutch administrative data. *Health Economics*, 29(4):435–451.

Tenand, M., Bakx, P., and Wouterse, B. (2023). The impact of co-payments for nursing home care use on use and welfare. CPB Discussion paper, Centraal Planbureau.

Tenand, M., Wouterse, B., Bakx, P., and Bom, J. (2020b). Do co-payments affect nursing home entry? An assessment based on Dutch administrative data. Netspar Design Paper no. 139.

van den Berg, G. (2001). Duration models: specification, identification and multiple durations. In *Handbook of Econometrics*, volume 5, pages 3381–3460. Elsevier, Amsterdam.

van den Berg, G. and Drepper, B. (2016). Inference for shared-frailty survival models with left-truncated data. *Econometric Reviews*, (35):1075–1098.

van der Vaart, J., Alessie, R., and van Ooijen, R. (2020). Economic consequences of widowhood: evidence from a survivor's benefits reform in the Netherlands. Netspar Design Paper no. 160.

van Houtven, C. H. and Norton, E. C. (2004). Informal care and health care use of older adults. *Journal of Health Economics*, 23(6):1159–1180.

van Ooijen, R., Alessie, R., and Kalwij, A. (2015). Saving behavior and portfolio choice after retirement. *De Economist*, 163(3):353–404.

Van Ooijen, R., Alessie, R., and Knoef, M. G. (2015). Health status over the life cycle. Netspar Design Paper no. 62.

Vass, M., Avlund, K., Hendriksen, C., Philipson, L., and Riis, P. (2007). Preventive home visits to older people in Denmark–why, how, by whom, and when? *Zeitschrift für Gerontologie und Geriatrie*, 40(4):209–216.

Vaupel, J. W., Carey, J. R., Christensen, K., Johnson, T. E., Yashin, A. I., Holm, N. V., Iachine, I. A., Kannisto, V., Khazaeli, A. A., Liedo, P., Longo, V. D., Zeng, Y., Manton, K. G., and Curtsinger, J. W. (1998). Biodemographic trajectories of longevity. *Science*, 280(5365):855–860.

Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.

Wang, P., Li, Y., and Reddy, C. K. (2019). Machine learning for survival analysis: a survey. *ACM Computing Surveys*, 51(6):1–36.

Webb, D. C. (2009). Asymmetric information, long-term care insurance, and annuities: the case for bundled contracts. *Journal of Risk and Insurance*, 76(1):53–85.

Wouterse, B., Hussem, A., and Wong, A. (2021). The risk protection and redistribution effects of long-term care co-payments. *Journal of Risk and Insurance*, 89(1):161–186.

Zweifel, P. and Strüwe, W. (1998). Long-term care insurance in a two-generation model. *Journal of Risk and Insurance*, 65(1):13–32.

Samenvatting

Door vergrijzing zijn huishoudens steeds meer zelf verantwoordelijk voor het verzekeren van de gezondheids- en financiële risico's van ouderdom. Waar voorheen zelfs ruimhartige regelingen bestonden om met vervroegd pensioen te gaan, zijn veel pensioenregelingen nu versoberd. Bovendien neemt de generositeit van publieke langdurige zorg af door strengere toelatingseisen voor verpleeghuiszorg en een verhoging van de eigen financiële bijdrage. De hervormingen van langdurige zorg zijn bedoeld om ouderen met een zorgbehoefte langer thuis te laten wonen en om het gebruik van mantelzorg te stimuleren. Hierdoor zijn huishoudens meer genoodzaakt om zelf geld opzij te zetten voor hun oude dag en om deze particulier te verzekeren.

Adequate pensioenen en langdurige zorg vereisen goed inzicht in de verschillen in gezondheidsrisico's tussen huishoudens. Mensen met een lagere sociaal-economische status zijn gemiddeld minder gezond en hebben een kortere levensverwachting. Daarnaast variëren zorgbehoeften en de beschikbaarheid van mantelzorg tussen huishoudens. Er vindt dus herverdeling van (publieke) middelen plaats, omdat huishoudens verschillen in de duur van hun pensioenuitkering en in het gebruik van langdurige zorg. Bovendien keert een particulier pensioen of een verzekering tegen langdurig zorggebruik langer uit aan bepaalde groepen huishoudens, wat kan leiden tot inefficiënties op de verzekeringsmarkt.

In dit proefschrift onderzoeken wij de verschillen in langdurig zorggebruik en sterfte, en de impact ervan op publieke en particuliere verzekeringen.

Hoofdstuk 2: De Determinanten van Langdurige Zorgpaden: Bewijs op basis van Nederlandse Administratieve Gegevens

Dit hoofdstuk bestudeert de invloed van zorgbehoefte, mantelzorg en financiële middelen op de duur van zorgtypen en de overgangen tussen zorgtypen (thuis- of verpleeghuiszorg). Wij gebruiken een overgangsmodel dat wij toepassen op unieke Nederlandse data van het

Centraal Bureau voor de Statistiek over de duur van zorggebruik. De data rapporteert thuis- en verpleeghuisgebruik, en familie- en individuele karakteristieken op (bijna) continue basis. Het model en de data gebruiken wij ook in Hoofdstukken 3 en 4.

Als wij mensen met een fysieke of cognitieve aandoening met elkaar vergelijken, zien wij dat mensen met een fysieke aandoening korter thuis- en verpleeghuiszorg gebruiken en dat zorggebruik vaker tijdelijk is. Daarnaast vinden wij dat mantelzorg de overgang van thuiszorg naar verpleeghuiszorg remt voor mensen met een fysieke aandoening, maar niet voor mensen met een cognitieve aandoening. Hervormingen die inzetten op het gebruik van mantelzorg dienen dus rekening te houden met de aandoening van de zorgvrager. Daarnaast zien wij dat het hebben van meer financiële middelen en een eigen woning (zwaardere) verpleeghuiszorg uitstelt en lichter zorggebruik bespoedigt. Dit wijst op een mogelijke vraag naar particuliere langdurige zorg.

Hoofdstuk 3: Ouderdomsverzekeringen Bundelen Vanwege Sociaal-economische Verschillen in Langdurig Zorggebruik en Sterfte

Vervolgens bestuderen wij averechtse selectie die ontstaat door verschillen in langdurig zorggebruik en sterfte. Averechtse selectie leidt tot een inefficiënt hoge premie en laag aantal verzekerden op de verzekeringsmarkt: alleen mensen met een hogere levensverwachting kopen een pensioenverzekering, en alleen mensen met een hogere verwachte zorgbehoefte verzekeren langdurig zorggebruik. Productbundeling kan averechtse selectieproblemen dempen wanneer de risico's negatief gecorreleerd zijn: langlevenden gebruiken weinig zorg en kortlevenden gebruiken veel zorg, waardoor de totale uitkering in evenwicht is. Wij rapporteren negatieve correlaties voor sociaal-economische groepen.

Wij tonen echter aan dat alléén een negatieve correlatie niet volstaat om averechtse selectie te minimaliseren met productbundeling: het gemiddelde en de spreiding van levensverwachting en langdurig zorggebruik zijn ook van belang. Wij berekenen deze factoren vervolgens op basis van de administratieve data en vinden dat averechtse selectie niet ongedaan gemaakt kan worden door een gebundelde verzekering met uniforme premie. Onze bevindingen tonen het belang van groep-specifieke premies aan.

Hoofdstuk 4: Gezondheidsongelijkheid en de Progressiviteit van Ouderdoms-volksverzekeringen

Hoofdstuk 4 focust op publieke verzekeringen en bestudeert in hoeverre sociaal-economische verschillen in langdurige zorg en sterfte de welvaartsverdeling in Nederland beïnvloeden. Mensen met de hoogste sociaal-economische status verblijven gemiddeld korter in een verpleeghuis en leven langer dan mensen met de laagste sociaal-economische status. Zij betalen dus korter een eigen bijdrage voor zorg, maar ontvangen tegelijkertijd langer pensioen. Uit ons welvaartsmodel blijkt dat zij hierdoor 23.4% meer kunnen consumeren dan degenen met de laagste sociaal-economische status. Een groot deel van deze welvaartswinst wordt verklaard door het nalaten van grotere erfenissen.

Hoofdstuk 5: Schatting van Links Afgekapte Duurmodellen met Gedeelde Niet-Geobserveerde Heterogeniteit ('Left-Truncated Shared Frailty Models')

Hoofdstuk 5 behandelt het gebruikte duurmodel en de schatting daarvan. De uitkomst-variabele is de duur tot het einde van een toestand. De modelspecificatie bevat een niet-geobserveerd effect die hetzelfde is binnen een groep (*frailty*). Wij breiden het model uit met links afgekapte duren (*left truncation*): de steekproef bevat alleen duren die een drempelwaarde overschrijden. De steekproef is hierdoor een dynamische selectie van relatief lange toestandsduren. Wij ontwikkelen een zuivere schattingsmethode die corrigeert voor de selecte steekproef. Daaropvolgende simulaties laten zien dat geschatte duur- en covariaateffecten onzuiver zijn wanneer de dynamische selectie wordt genegeerd. De onzuiverheid neemt echter af als het niet-geobserveerde effect wordt gedeeld binnen een grotere groep. Wij passen onze zuivere schattingsmethode toe in Hoofdstukken 2 tot en met 4.