Long-Term Health Insurance: Theory Meets Evidence*

Juan Pablo Atal† Hanming Fang‡

Martin Karlsson§ Nicolas R. Ziebarth¶

February 25, 2021

Abstract

To insure policyholders against contemporaneous health expenditure shocks and future reclassification risk, long-term health insurance constitutes an alternative to community-rated short-term contracts with an individual mandate. In this paper, we study the German long-term health insurance (GLTHI) from a life-cycle perspective. The GLTHI is one of the few real-world long-term health insurance markets. We first present and discuss insurer regulation, premium setting, and the main market principles of the GLTHI. Then, using unique claims panel data from 620 thousand policyholders over 7 years, we propose a new method to classify and model health transitions. Feeding the empirical inputs into our theoretical model, we assess the welfare effects of the GLTHI over policyholders' lifecycle. We find that GLTHI achieves a high level of welfare against several benchmarks. Finally, we conduct counterfactual policy simulations to illustrate the welfare consequences of integrating GLTHI into a hybrid insurance system similar to the current system in the United States.

Keywords: long-term health insurance; individual private health insurance; reclassification risk, intertemporal incentives, ACG scores, health transitions

JEL Classifications: G22; I11; I18.

---

*We thank Konstantin Beck, Paola Bertoli, Rudi Blankert, Florian Buchner, Keith Ericson, Colleen Carey, David Dillenberger, Stefan Felder, Sebastian Fleitas, Leora Friedberg, Paul Grieco, Ben Handel, Jim Heckmann, Kate Ho, Mathias Kifmann, Amanda Kowalski, Dirk Krueger, Ethan Lieber, Tony Lo Sasso, Claudio Lucarelli, Maria Polyakova, Andreas Ryll, Holger Strulik, Nicholas Tilipman, Harald Uhlig, Jürgen Wasem, Michael Whinston, Joachim Winter, Moto Yogo, Jin Zhou, and Peter Zweifel as well as seminar participants at Boston University, Columbia University, Indiana University, Lund University, Princeton University, University of Chicago, University of Pennsylvania, University of Maryland, University of Wisconsin-Madison, Conference of the American Society of Health Economists (2018), Annual Meeting of the Southern Economic Association (2018), Health Economics Working Group Meeting of the German Economic Association (*Ausschuss für Gesundheitsökonomie*, 2018), Essen Health Conference (2019), Congress of the International Health Economics Association (iHEA 2019), NBER Insurance Working Group Conference (2019), New York Area Health Economics Day (2019), and the GSE Forum in Microeconometrics for very helpful comments and suggestions. We thank representatives of the German Association of Private Health Insurers as well as Frank Bowert for invaluable comments on how well our modeling of German LTHI reflects actuarial practices. We do not have financial interests that would constitute any conflict of interests with this research. Generous funding by the German Federal Ministry of Education and Research (FKZ: 01EH1602A) is gratefully acknowledged.

†University of Pennsylvania, Department of Economics, 133 South 36th Street, Philadelphia, PA 19104, USA. Email: ataljp@econ.upenn.edu

‡University of Pennsylvania, Department of Economics, 133 South 36th Street, Philadelphia, PA 19104, USA; and the NBER. Email: hanming.fang@econ.upenn.edu

§CINCH, University of Duisburg-Essen, Weststadttürme Berliner Platz 6-8, D-45127 Essen, Germany. Email: martin.karlsson@uni-due.de

¶Cornell University, Department of Policy Analysis and Management (PAM), 2218 Martha Van Rensellaer Hall, Ithaca, NY 14850, USA; and the NBER. Email: nrz2@cornell.edu.

1 Introduction

Health insurance contracts sold in the private market tend to be short-term, typically annual, policies. Short-term contracts expose policyholders to potentially large premium fluctuations ("reclassification risk") and can lead to significant welfare losses (Diamond, 1992; Cochrane, 1995). Consequently, for decades, academics and policymakers alike have studied options to regulate such short-term health insurance markets. The standard policy options, such as community-rated premiums and guaranteed issuance regulations, strive to avoid undesirable outcomes like uninsurance, unaffordable premiums for sick individuals and large premium fluctuations following changes in health status (Claxton et al., 2017; Cole et al., 2019). However, these policy options typically also imply a trade-off with unintended consequences such as adverse selection, which must be addressed either through individual mandates or premium subsidies, or both (cf. Akerlof, 1970). The Affordable Care Act (ACA), enacted in the United States in 2010, indeed features community rating, an individual mandate and premium subsidies as its three main pillars (Aizawa and Fang, 2020). At the same time, the inherent trade-offs have lead to passionate debates and lawsuits.

Long-term contracts offer a fundamental alternative to short-term contracts, and provide policyholders with reclassification risk insurance without necessarily triggering adverse selection problems. Under long-term contracts, agents not only receive coverage against contemporaneous medical risk provided by short term contracts, but also coverage against future premium fluctuations through the payment of an additional premium upfront. In theory, a carefully designed long-term contract can minimize the reclassification risk, while ensuring market participation and eliminating adverse selection by leveraging individual's intertemporal incentives (Pauly et al., 1999; Patel and Pauly, 2002; Pauly and Lieberthal, 2008).

In this paper, we study the largest and oldest individual private long-term health insurance market in the world. In Germany, 10 percent of the population (or 8.8 million individuals) hold individual long-term health insurance policies sold by private insurance companies. After risk-rated premium setting at initial enrollment, the policies are guaranteed renewable until death (without an expiration date or enrollment period).[1] All subsequent premium changes have to be community rated; that is, premium changes over the lifecycle are independent of changes in the policyholder's own health status. In fact, given the market regulation, the German long-term health insurance (henceforth GLTHI) foresees the payment of constant real premiums over the lifecycle, regardless of the evolution of an individual's income and health status. As a consequence, the GLTHI contract almost entirely elimi-

[1] Unlike the United States, Germany has no public insurance specifically for people above the age of 65.

nates the reclassification risk—at the expense of relatively high premiums during the early life years ("front-loading").

This paper begins by presenting the main principles and functioning of GLTHI. It is a market that, despite its stable existence for decades, has received very little attention outside Germany. Next, we formulate the theoretical foundations of GLTHI, given the regulatory framework and considering endogenous lapsation of contracts. We show that the evolution of health risk as well as the income profile over the lifecycle are the key empirical inputs to assess the welfare consequences of GLTHI. Then, we leverage detailed claims panel data as well as survey data to construct these inputs.

Specifically, we rely on a unique panel of claims data from one of the largest German private insurers, which includes 620 thousand enrollees over 7 years, spanning all age groups and all of the 16 German federal states.[2] In our next step, in Section 5, making use of the German version of the John Hopkins ACG© software, we propose a novel health risk classification method. This method allows us to categorize and model individuals' expected health risks and to study their health transitions over time. Moreover, because lifecycle income profiles play a crucial role when assessing the welfare effects of GLTHI, we leverage more than three decades of lifecycle income panel data from the representative German Socio-Economic Panel Study (SOEP). For this purpose, we generate household income measures that consider all income streams—including social insurance benefits—and within-household redistribution. Later in the paper, we show that our findings are robust to using more than three decades of lifecycle income data from the U.S. Panel Study of Income Dynamics (PSID).

In Section 6, we use our theoretical and empirical inputs to simulate the economy and to quantify welfare under different contracts. Specifically, we compare the welfare implications of GLTHI to the welfare implications of the (1) first-best contract, which guarantees a constant consumption profile over individuals' lifecycle, (2) a series of risk-rated short-term contracts, and (3) the optimal dynamic contract as characterized in Ghili et al. (2019). We find that the simple GLTHI design generates only small welfare losses compared to the optimal contract. Under various parameterizations and scenarios, replacing the GLTHI contract with the optimal contract would increase welfare by between zero and seven percent. Within a plausible range of parameter values, we find that the welfare gains are smaller than four percent. When delving deeper into an understanding of the underlying mechanisms, we find that, compared to the optimal contract, the GLTHI contract entails less consumption smoothing over the lifecycle, but also less reclassification risk. The welfare loss due to less consumption smoothing is almost entirely offset by better reclassification risk insurance in the GLTHI

[2] For example, the oldest policyholder is 99 years and the most loyal policyholder has been client for 86 years.

contract. These results are robust to the incorporation of private savings, to a wide range of degrees of risk aversion, and to non-time-separable recursive preferences à la Epstein and Zin (1989).

In the final section before we conclude, we discuss the potential implications of an existing real-world private long-term insurance market for U.S. health insurance reform debates. We argue that the U.S. health insurance system, at least prior to the ACA, could be roughly approximated by a hybrid system of private health insurance contracts for the working-age population up to age 64, and payroll tax financed Medicare insurance for those above age 65. In addition, the market for private health insurance contracts is to a first order approximation a 60/40 mixture of employer-sponsored health insurance and short-term contracts. We simulate such a simplified U.S. system to show that transitioning all short-term contracts to long-term contracts would substantially increase welfare. We also find that a hybrid system of private long-term insurance contracts and a single-payer Medicare system achieves *lower* welfare than a genuine system of private long-term contracts over the entire lifecycle (as in GLTHI).

This paper contributes to several literatures. First, it contributes to the literature on dynamic contracts for which vast theoretical work but relatively little empirical evidence exists. Pauly et al. (1995) propose a "guaranteed-renewable" contract with a pre-specified path of premiums that fully eliminates adverse selection and reclassification risk. Similarly, Cochrane (1995) proposes a scheme of severance payments, made after the realization of health shocks, which provides full insurance against reclassification risks. Hendel and Lizzeri (2003) and Ghili et al. (2019) show that the optimal contract only partially insures reclassification risk, because fully eliminating reclassification risks requires large front-loaded payments, preventing consumption smoothing over the lifecycle. Krueger and Uhlig (2006) characterizes the competitive long-term contract that insures the agent against income risk under one-sided commitment. Cole et al. (2019) use a dynamic model of health investments and insurance to study the short and long-term effects of providing social insurance. They find that providing full insurance is suboptimal as the negative dynamic effects on health behavior (and consequently population health) dominate in that setting.

Second, several papers, including Hendel and Lizzeri (2003), Herring and Pauly (2006), Finkelstein et al. (2005), and Atal (2019), investigate empirically the workings of long-term contracts in different contexts. Our paper contributes to this empirical literature by introducing a method of discrete classification of health risks. We base our method on the properties of *homogeneity* and *separation* in the actuarial science literature (see Finger, 2006). Our proposed method is, in our view, a more informative way of discrete classification of health risks than the mostly *ad hoc* method used in the existing literature.

Moreover, our paper relates to previous work on the Germany long-term health insurance market. Hofmann and Browne (2013) describe GLTHI contracts and show that switching behavior in the market is consistent with its incentive structure. Christiansen et al. (2016) empirically study determinants of lapsing and switching behavior. And Baumann et al. (2008) and Eekhoff et al. (2006) discuss the potential effects of higher switching rates on market competition if the capital accumulated through front-loaded payments were to be made portable across insurers. While these two papers discuss a hypothetical reform, Atal et al. (2019) theoretically and empirically study the effects of the actual 2009 portability reform on switching behavior.

2 Institutional Details

Germany has a two-tier health insurance system where a statutory health insurance (SHI) and an individual private health insurance market co-exist. SHI is a public insurance program that covers 90 percent of the population. SHI enrollees and their employers pay income-dependent contribution rates (each pay about 8 percent of the gross wage, up to cap) for a standardized benefit package with very little cost-sharing; as of this writing, SHI enrollees can choose among 109 non-profit sickness funds (Schmitz and Ziebarth, 2017; Bünnings et al., 2019; Bundesministerium für Gesundheit, 2020). However, for historical reasons, select population subgroups can leave the public SHI system *permanently* and fully insure their health risks with long-term health insurance contracts purchased from the private market. Despite the two-tier system, the German system provides almost universal coverage with an uninsurance rate of only around 0.1 percent (German Statistical Office, 2016).

Besides Chile (cf. Atal, 2019), Germany is the only country in the world with an existing private long-term health insurance market. About 8.8 million enrollees, or about 10 percent of the German population, receive health insurance coverage from this market (Association of German Private Healthcare Insurers, 2019b). For historical reasons, GLTHI covers three main population subgroups: (a) the self-employed; (b) high-income earners with annual gross labor incomes above a politically defined federal threshold (2021: €64,350, or about $77,863); and (c) civil servants. These groups can leave the SHI system permanently and insure their health risks privately with a long-term contract (Nuscheler and Knaus, 2005; Hullegie and Klein, 2010; Polyakova, 2016; Panthöfer, 2016). The decision to enter the private market is essentially a lifetime decision. Switching back to SHI is strictly limited, so as to prevent individuals from strategically switching back and forth and gaming the system; the basic principle is "once privately insured, always privately insured" (Schencking, 1999; Innungskrankenkasse Berlin Brandenburg, 2018). We discuss the institutional specifics of this rule and the empirical evidence on the difficulty of switching from GLTHI back to SHI in Appendix A. Hofmann and Browne (2013) and Atal et al. (2019) provide specific details on the individual private market.

The GLTHI market consists of 48 private insurers that sell *comprehensive* as well as *supplemental* insurance coverage (Association of German Private Healthcare Insurers, 2020). The focus of this paper is the comprehensive or "substitutive" (to SHI) insurance. These are always individual non-group policies. In addition to saving the SHI payroll taxes, advantages of getting private GLTHI include a high degree of plan choice as well as actuarially fair premiums in a lifecycle perspective (see below). Compared to the post-ACA era in the U.S., the GLTHI market is less regulated. Applicants can freely choose their level of coverage in terms of benefits and cost-sharing amounts, within some lax limits. This results in thousands of different health plans among the 8.8 million policyholders, most of which are sold across state lines and nationwide. The majority of private insurers operate nationwide and are open to all applicants who opt out of SHI.

Provider Networks. Provider networks and "Managed Care" are unknown in the public and private system in Germany; that is, in either system enrollees are free to choose any providers in the German health care system. Moreover, in both systems, reimbursement rates are centrally determined and do not vary by insurers or health plans. While reimbursement rates for inpatient care are identical in both systems, they are about twice as high for outpatient care in the private market. As a consequence, wait times in the outpatient sector are significantly shorter for the privately insured (Werbeck et al., 2019). Because they do not negotiate rates or build provider networks, private insurers mainly customize health plans and process, scrutinize, and deny claims. Thus, the GLTHI contract primarily constitutes a *pure financial contract* similar to other insurance markets such as life insurance (Fang and Kung, 2020). This specific feature substantially simplifies the welfare analysis of GLTHI.[3]

Guaranteed Renewability and One-Sided Commitment. When individuals apply for a long-term insurance contract, insurers have the right to deny applicants with bad risks coverage or impose pre-

[3] In the spirit of Koijen et al. (2016), one may make the case that a market of private financial long-term contracts reduces the government risk to investors that is driving the "medical innovation premium." Koijen et al. (2016) hypothesize and provide evidence that "government-induced profit risk"—for example, approval regulations—induce investors to demand higher returns on their investment. Compared to public insurance markets, one could argue that private markets and contracts are less prone to such regulatory risk.

existing condition clauses. However, once contracts are purchased, the insurers cannot terminate them. GLTHI contracts are not yearly contracts, but *permanent lifetime contracts* without an end date. In other words, the GLTHI contracts are guaranteed renewable over the lifecycle. However, enrollees can terminate these permanent contracts, e.g. to switch insurers, thus GLTHI is a market with a *one-sided commitment*. Indeed, it is common that enrollees remain insured with their carrier until they die (recall that Medicare does not exist in Germany). For example, in our sample, the policyholders' average age is 46 years and they have been clients for 13 years; the oldest policyholder is 99 years old and one policyholder has been a client of the insurer for 86 years, see Table C1 (Appendix). In addition, whereas the initial premium is risk-rated, all subsequent premium increases are community-rated at the plan level, such that the contract provides lifelong insurance against *reclassification risk*.

Premium Calculation and Old Age Provisions. The initial GLTHI premium is individually underwritten.[4] Premiums consist of several components, and the *Kalkulationsverordnung (KalV)* regulates the exact calculations. The insurers' actuaries carry out the specific calculations which have to be approved by a federal financial regulatory agency (the *Bundesanstalt für Finanzdienstleistungsaufsicht, BaFin*). Specifically, Chapter 1 of the *KalV* specifies that premiums have to be a function of the expected per capita health care claims or *Kopfschäden* (which depend on the plan chosen, age, gender, health risks),[5] the assumed guaranteed interest rate (*Rechnungszins*), the probability to lapse (*Stornowahrscheinlichkeit*), and the life expectancy (*Sterbewahrscheinlichkeit*).

One important and distinct characteristic of the GLTHI market is the legal obligation of insurers to build up *old-age provisions*, typically until age 60 of the policyholder. The old-age provisions accumulated early in the lifecycle serve as the capital to cover higher health expenditures later in the policyholder's lifecycle. Premiums are calculated under the basic principle of a constant lifecycle premium, sufficient to cover expenses over the policyholder's lifecycle (we provide a formal treatment of this principle in Section 3.1). Thus, in young ages, premiums exceed the expected claims; while in old ages, premiums are lower than the expected claims—a phenomenon known as "front-loading"

[4] The only exception is the "Basic Plan" (*Basistarif*). The Basic Plan must be offered by all carriers and is structured after the SHI with the same essential benefits and actuarial values. For the Basic Plan, guaranteed issue exists for people above 55 and those who joined the GLTHI after 2009. The maximum premium is capped at the maximum SHI contribution (2021: €769,16 per month). The legislature mandated the Basic Plan to provide an "affordable" private option for GLTHI enrollees who cannot switch back to SHI, are uninsured, would have to pay excessive premiums, or would be denied coverage. However, the demand for the Basic Plan has been negligible; thus henceforth, we will abstain from it. In 2019, in the entire GLTHI, only 32,400 people, or 0.4 percent, were enrolled in the Basic Plan (Association of German Private Healthcare Insurers, 2020). In our data, only 1,006 enrollees chose the basic plan in 2010.

[5] Gender rating was allowed until December 21, 2012. After this date, for new contracts, all insurers in the European Union (EU) have to provide unisex premiums as the EU Court of Justice banned gender rating as discriminatory (Schmeiser et al., 2014)

in long-term insurance contracts (Hendel and Lizzeri, 2003; Nell and Rosenbrock, 2007, 2009; Fang and Kung, 2020).[6]

Figure 1 illustrate front-loading for four combinations of age at initial enrollment and health risks: high and low health risks, and initial enrollment ages at 30 vs. 50. In this illustration, we assume the health risk types to be constant over the lifecycle.[7] The low risk type (the "healthy") corresponds to a hypothetical individual with no pre-existing conditions; we denote the age profile of her expected health expenditures conditional on survival by the curve $E(m|\text{surv}, \text{low})$. The high risk type (the "sick") corresponds to a hypothetical individual who has 50 percent higher expected health care claims than the low risk type at each age. Her age profile of expected health expenditures conditional on survival is denoted by the curve $E(m|\text{surv}, \text{high})$. Note that $E(m|\text{surv}, \text{low})$ and $E(m|\text{surv}, \text{high})$ would also represent the actuarial fair premiums of short-term spot contracts by age, for low and high risk types, respectively. In Figure 1, $P_{30,\text{low}}$ (respectively, $P_{30,\text{high}}$) are the GLTHI premiums for a low (respectively, high) risk type who first enrolls in a private plan at age 30. Similarly, $P_{50,\text{low}}$ and $P_{50,\text{high}}$ are the premiums if the two types start their initial enrollment much later in life, at age 50.

Figure 1 has the following important features: First, premiums remain stable over individuals' life cycles. Front-loaded premiums dampen the increases of the age-specific premiums for short-term spot contracts via the capital stock built through old-age provisions—the cumulative difference between premiums and expected claims (plus investment returns of the capital stock).[8] Second, premiums are higher for policyholders who joined the GLTHI later in their life, as the expected yearly future expenditures increase with age, and there would be fewer years to build up the old-age provision for those who join the GLTHI late.[9] Third, because of the initial risk rating, high risk types pay higher premiums throughout their lives, relative to the low risk types.[10]

While, theoretically, premiums are constant over individuals' lifecycles, in reality nominal (and

[6] Such front-loading creates a "lock-in" effect, in addition to the lock-in induced by guaranteed renewability (Nell and Rosenbrock, 2008; Atal, 2019). To strengthen consumer power and reduce this lock-in, the German legislature made a standardized portion of these old-age provisions portable across carriers for contracts signed after Jan 1, 2009; see Atal et al. (2019) for an evaluation of this reform. For existing contracts, Atal et al. (2019) do not find a significant impact on external switching rates. However, they find a one-time increase in internal plan switching during the limited six months period from January to June 2009 where portability was granted for existing contracts.

[7] This simplification of stable health risks allows us to illustrate the basic front-loading principle, allowing for a stochastic health status is fundamental to the analysis: First, it allows to show that front-loading can dampen the reclassification risk. Second, an evolving health status means that individuals who start unhealthy may lapse their contract, which introduces (downwards) reclassification risk even if premiums are constant within a given contract. Also, lapsation needs to be taken into account when calculating the premium level. Below we consider evolving health risks extensively.

[8] In 2019, the capital stock built through old-age provisions amounted to € 235 billion ($284 billion) for 8,732,000 policies, or to € 26,918 ($32,570) per policy (Association of German Private Healthcare Insurers, 2019c).

[9] This is not necessarily true when health changes over time. With a stochastic health status, the initial premium may start to decrease at very high ages as, over time, the need to front-load for future health shocks decreases (see Section 6.1.)

[10] Again, this is not necessarily true when health risk may change over time.

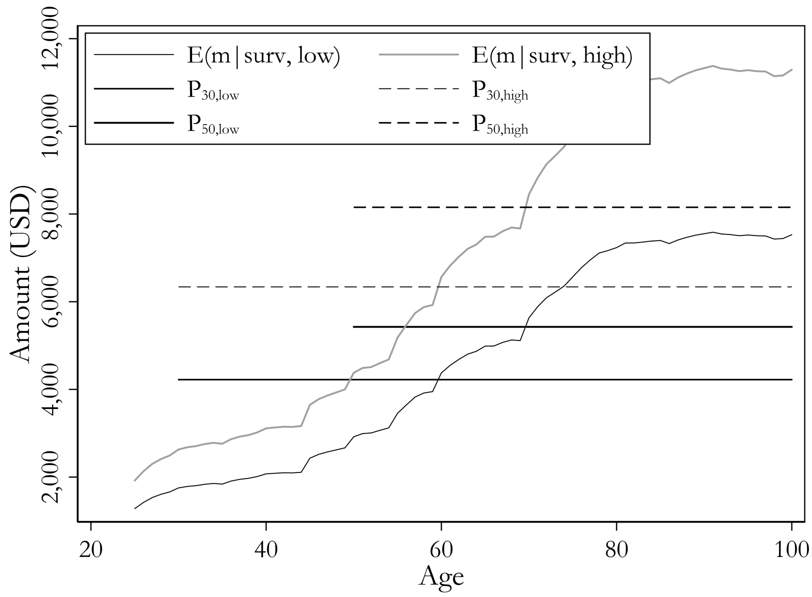


Figure 1: Premiums and Health Expenditures over the Lifecycle in the GLTHI

Source: German Panel Claims Data (see Section 4.1), own calculations, own illustration.

also real) premiums do increase. The main factors that trigger such premium adjustments (*Beitragsanpassungen*) are the following: (i) structural changes in life expectancy; (ii) structural changes in health care consumption; (iii) structural changes in health care prices mostly due to improvement in the quality of medical care, e.g. new expensive drugs or procedures;[11] (iv) structural changes of the economic environment, e.g. through capital markets or new financial regulation. An example of (iv) is the structural and unexpected shift of central banks to a super-low interest rate environment over the past decade; such a structural shift implies a significant decrease in the returns to risk-free capital investment. Because GLTHI insurers (like life insurers) are heavily invested in the bond market, structural premium adjustments are necessary to counter such reductions in investment returns.[12]

Premium adjustments are not only allowed in some cases, but also *required* by the regulatory financial oversight agency *BaFin* to ensure financial stability within the regulatory framework in the *Versicherungsvertragsgesetz (VVG)*, the *Versicherungsaufsichtsgesetz (VAG)*, and the *KalV*.[13] Most insurers have to follow the *Solvency II* reporting requirements. Each year, insurers have to test

[11] The Health Care Reform 2000 (*GKV-Gesundheitsreformgesetz 2000*) introduced a mandatory 10 percent premium surcharge up to age 60 to dampen structural increases in health care spending due to medical progress. This surcharge only applies to GLTHI contracts signed after January 1, 2000 (see article 14 of GKV-Gesundheitsreformgesetz (2000)).

[12] The *KalV* has traditionally capped the assumed return on equity, the so called "guaranteed interest rate" (*Rechnungszins*) for the premium calculation at 3.5 percent. This has been the case for five decades. However, in 2016 for the first time, the average net return on investment has dropped below 3.5 percent, which is why the *German Actuary Association* has issued a new guideline to calculate the new insurer-specific "maximum allowed interest rate" (*Höchstrechnungszins*), see Deutsche Aktuarvereinigung (DAV) (2019).

[13] Effective January 1, 2016 the *KalV* has been replaced by the *Krankenversicherungsaufsichtsverordnung (KVAV)*.

whether their underlying assumptions for their premium and old age provision calculations are still accurate. If they deviate by a certain amount, they have to adjust the premiums, which can result in two-digit premium increases, bad press, and lawsuits (Krankenkassen-Zentrale (KKZ), 2020).[14] However, on average, nominal premium increases have been moderate—in 2018 at 1.8 percent and from 2009 to 2019 at an average nominal rate of 2.8 percent (Association of German Private Healthcare Insurers, 2019a). Most important for our analysis is that, after the initial risk rating, premium adjustments do not depend on enrollees' possibly evolving health status.

3 Lifecycle Premiums and Welfare Measures

In this section, we will first formally derive the lifecycle equilibrium premium of the German Long-Term Health Insurance (GLTHI) contract; we will then discuss a set of welfare measures. In Section 6, we will use these welfare measures to assess the performance of GLTHI relative to several real world and theoretically optimal insurance contract alternatives.

3.1 Lifecycle Premiums in the GLTHI

Let $P_t(\xi_t)$ be the initial premium offered when first signing a GLTHI contract in period t. $P_t(\xi_t)$ depends on the individual's health risk in year t, ξ_t, as GLTHI contracts are individually underwritten at inception (see Section 2). We assume that $\xi_t \in \Xi$ where Ξ is a finite set of health states to be described below. In subsequent periods, each contract is guaranteed-renewable. As such, individuals who sign a contract in period t can renew the contract for the same premium, $P_t(\xi_t)$, in all periods between $t+1$ and T, regardless of the evolution of their health status.

As discussed in Section 2, the contract breaks even in equilibrium, given premium $P_t(\xi_t)$. Consequently, we express $P_t(\xi_t)$ as the solution to a fixed-point problem in which $P_t(\xi_t)$ covers exactly the expected claims of enrollees who *stay in* the contract at premium $P_t(\xi_t)$.

We solve for $P_t(\xi_t)$ recursively, starting from the last period, $t = T$. In the last period T, there is no uncertainty regarding future health shocks and future lapsation. Let m_t denote health care expenditures in year t. Assuming full coverage, it follows that $P_T(\xi_T) = \mathbb{E}(m_T|\xi_T)$.

To calculate the equilibrium premium in $t < T$, we need to consider *endogenous lapsation*. An interesting and practically convenient feature of the GLTHI is that enrollees will lapse their current

[14] All premium adjustments have to be legally checked and approved by 16 independent actuaries who are appointed by the *BaFin*. However, some plaintiffs in lawsuits argue that some of these actuaries would not be sufficiently independent. Other reasons of courts to declare a premium increase as "not justified" were insufficient explanations by the insurers or a deliberate initial underpricing of premiums in the first year to attract enrollees (Krankenkassen-Zentrale (KKZ), 2020).

contract if and only if, given the evolution of their health status, they can obtain a lower premium than their current guaranteed-renewable premium $P_t(\xi_t)$ if they apply for a new policy and switch insurers. Formally, lapsing a contract signed in $t < T$ at the risk-rated premium $P_t(\xi_t)$ occurs at the first $\tau > t$ under health status ξ_τ if $P_\tau(\xi_\tau) < P_t(\xi_t)$, where $P_\tau(\xi_\tau)$ is the premium that the individual can obtain from a new long-term policy at period $\tau > t$ when her health status is ξ_τ.[15]

It is surprising, at least at a first glance, that the policyholder's lapsation decision does *not* depend on the curvature of his/her utility function. To understand this result, it is important to note that the difference in the policyholder's continuation value from holding two guaranteed-premium long-term contracts only depends on the premium difference, because the other determinants of the continuation value, namely health transitions and income dynamics, is independent of what long-term contracts he/she holds; moreover, while the *level* of the difference in values from holding guaranteed-premium contracts with different premiums depends on the curvature of the utility function, the *sign* of the difference does not.[16]

Remark 1 *The lapsation decision under GLTHI is only driven by a comparison between one's current guaranteed premium $P_t(\xi_t)$ and the premium that the policyholder could obtain from a new contract $P_\tau(\xi_\tau)$. Neither risk aversion nor income shocks play any role in the lapsation decision under GLTHI. As GLTHI is a pure financial contract, the lapsation decision is not driven by differentiation in provider networks associated with the policies.*

For a given $t < T$ and $\tau > t$, we denote $\mathbf{P}_{t+1}^{\tau} \equiv \{P_{t+1}(.), ..., P_\tau(.)\}$ as the set of guaranteed premiums from $t+1$ to $t+\tau$. We can then recursively write the break-even GLTHI lifecycle premium for period-t new enrollees with heath state ξ_t, which we denote by $P_t(\xi_t)$, as follows:

$$P_t(\xi_t) = \frac{\mathbb{E}(m_t|\xi_t) + \sum_{\tau>t}^{T} \sum_{z\in\Xi} \delta^{\tau-t}\mathbb{E}(m_\tau|z) \times q_\tau(z|\xi_t, \mathbf{P}_{t+1}^{\tau}, P_t(\xi_t))}{1 + \sum_{\tau>t}^{T} \sum_{z\in\Xi} \delta^{\tau-t} \times q_\tau(z|\xi_t, \mathbf{P}_{t+1}^{\tau}, P_t(\xi_t))}, \quad (1)$$

where the first element of the numerator, $\mathbb{E}(m_t|\xi_t)$, is expected health care costs in period t, given ξ_t; the second element of the numerator is the sum of the expected future health care costs over all remaining life years from t to T. Expected future health care costs are discounted with rate δ, with future spending at period τ weighted by $q_\tau(z|\xi_t, \mathbf{P}_{t+1}^{\tau}, P_t(\xi_t))$, the probability that (i) $\xi_\tau = z$, and (ii)

[15] Note that we abstain from horizontal differentiation across plans, and from switching costs.

[16] This argument also applies when the policyholder's preferences are not time separable, e.g., if they have Epstein-Zin preferences (Epstein and Zin, 1989).

the enrollee does not lapse (or die) between periods t and τ, given the subsequent equilibrium premiums $\mathbf{P}_{t+1}^{\tau}$. These expected lifecycle health care costs are then normalized by the expected number of years not lapsing the contract in the denominator.[17] In other words, in the GLTHI market, the lifecycle premium $P_t(\xi_t)$ equals the average of today's expected health care spending and all expected future health care spending, given the health risk today and in the future, weighted by the likelihood of not lapsing in any of the future time periods until death.

Equation (1) implicitly determines the constant GLTHI equilibrium lifecycle premium for a contract signed in period t when the enrollee's health status is ξ_t. Note that the break-even constraint determines the GLTHI lifecycle premium in any period for different health statuses, considering the likelihood to lapse in future periods. Also note that these lifecycle premiums do not maximize any *ex ante* consumer objective functions; conceptually, they are *not* designed to maximize any welfare criterion.

Remark 2 *The equilibrium premiums of the GLTHI are recursively determined by Equation 1. They do not depend on the policyholder's utility function or lifecycle income profile. Therefore, the GLTHI premiums do not depend on education or other determinants of lifecycle income profiles.*

3.2 Welfare Concepts and the Optimal Dynamic Contract

We use the concept of lifetime utility U to quantify welfare following, e.g., Ghili et al. (2019):

$$U = \mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0} u(c_t)\right)$$

where S_t is an indicator of survival until period t, and c_t is the consumption in period t that is specified by the contract. It may depend on the history of health and income realizations up to t. Expectation is taken over the individual's lifetime health history $(\xi_1, \xi_2, ..., \xi_t)$ and survival.[18]

Certainty Income Equivalent. With a parametric assumption for flow utility $u(.)$, and knowing income y_t, we can summarize welfare with the "*certainty income equivalent*", denoted CE, such that:

$$u(CE) = \frac{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0} u\left(c_t\right)\right)}{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0}\right)} \tag{2}$$

[17] Of course, $q_\tau(z|\xi_t, \mathbf{P}_{t+1}^{\tau}, P_t(\xi_t))$ depends on the evolution of the health status $\xi_{t+1}, ..., \xi_\tau$ and death, conditional on current health status ξ_t. We describe how we model the health risk process in Section 5.

[18] We assume that there are no annuity markets, so mortality risk is still considered.

This simple expression captures the main trade-offs in health insurance design for lifetime welfare. Lifetime utility is higher when consumption is smoothed across health states and across periods.

First-Best. In particular, the *first-best* consumption level is equal to the present discounted value of "net income" $y_t - \mathbb{E}(m_t)$, taking into account mortality risk. This constant optimal consumption level C^* is given by:

$$C^* = \frac{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0}(y_t - \mathbb{E}(m_t))\right)}{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0}\right)} \tag{3}$$

Short-Term Contracts. Under a series of actuarially fair *short-term* contracts, the premium in period t with health status ξ_t will simply be $\mathbb{E}(m_t)$. Thus consumption will be $c_t = y_t - \mathbb{E}(m_t|\xi_t)$, and the certainty equivalent CE becomes:

$$u(CE_{ST}) = \frac{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0} u(y_t - \mathbb{E}(m_t|\xi_t))\right)}{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0}\right)} \tag{4}$$

Optimal Dynamic Contract with One-Sided Commitment. Finally, the optimal dynamic contract with one-sided commitment (by the insurers only), as derived by Ghili et al. (2019) consists of consumption guarantees $\bar{c}_t(\xi_t, \mathbf{y}_t^T)$, that depend not only on health status (like GLTHI) but also on a vector of current and future income $\mathbf{y}_t^T \equiv \{y_t, y_{t+1}, ...y_T\}$. The consumption guarantees can also be written as a series of contracts with guaranteed premium paths:

$$P_\tau(\xi_\tau, y_\tau) = y_\tau - \bar{c}_t(\xi_t, \mathbf{y}_t^T) \tag{5}$$

Compared to the equilibrium GLTHI premium which does not depend on income or health and almost entirely eliminates reclassification risk, the premium of the optimal dynamic contract with one-sided commitment, as in 5, *does* depend on income, and also changes after each health shock. The reason is that the optimal contract penalizes high premiums when the marginal utility of consumption is high. Appendix B provides more details and discussions on the optimal dynamic contract.[19]

[19]Note that, following a similar logic to GLTHI, lapsation occurs if and only the individual is offered a higher consumption guarantee, and thus it does not depend on the utility function. Moreover, this characterization of the optimal long term contract is independent of the preferences, as long as there is time separability (we discuss the case of non-time-separable preferences in Section 6.6, when the contract as characterized by Ghili et al. (2019)) is no longer optimal.

4 Claims and Survey Panel Data from Germany

This section describes the claims panel dataset and the survey panel dataset used in this paper. The main working samples focus on the privately insured in the GLTHI market. We use the claims panel data primarily to estimate individual health transitions and related medical expenditures over the lifecycle. In contrast, we use the survey panel data primarily to estimate individual income dynamics over the lifecycle.

4.1 GLTHI Claims Panel Data

The claims panel data are administrative records and contain the universe of GLTHI contracts and claims between 2005 and 2011 from one of the largest private health insurers in Germany. In total, our data include more than 2.6 million enrollee-year observations from 620 thousand unique policyholders along with detailed information on plan parameters such as premiums, claims, and diagnoses. Atal et al. (2019) provide more details about the dataset. The claims data also contain the *age* and *gender* of all policyholders as well as their occupational group and the age when they first signed a contract with the insurer. We converted all monetary values to 2016 U.S. dollars (USD).

Sample Selection. We focus on primary policyholders. In other words, we disregard children insured by their primary caregivers and those who are younger than 25 years (555,690 enrollee-year observations).[20] Moreover, due to the 2009 portability reform (see footnote 6), we disregard inflows after 2008 (253,325 enrollee-year observations).[21] Our final sample consists of 1,867,465 enrollee-year observations from 362,783 individuals.

Descriptive Statistics. Table C1 (Appendix) presents the descriptive statistics. The mean age of the sample is 45.5 years and the oldest enrollee is 99 years old. Thirty-four percent of the sample are high-income employees, 49 percent are self-employed and 13 percent are civil servants. The majority of policyholders (72 percent) are male, because women are underpresented among the self-employed and high-income earners in Germany. On average, policyholders have been clients of the insurer for 13 years and have been enrolled in their current health plan for 7 years. Ten percent of all policyholders have been with the insurer for more than 28 years and one policyholder has been with the insurer for as long as 86 years, illustrating the existence of a real-world private long-term

[20] Children obtain their own individual risk-rated policies. However, if parents purchase the policy within two months of birth, no risk-rating applies. Under the age of 21, insurers do not have to budget and charge for old-age provisions.

[21] Below we show that the composition of enrollees has remained stable between 2006 and 2011.

health insurance system.[22] Figure A2 shows the distribution of policyholders' age when joining the company. The mass of individuals signs their first GLTHI contract around the age of 30, at a time when most Germans have fully entered the labor market but are still healthy and face reasonable premiums.

Table C1 shows that the average *annual premium* is $4,749 and slightly lower than the average premium for a single plan in the U.S. group market at the time (Kaiser Family Foundation, 2019). Note that the *annual premium* is the total premium—including employer contributions for privately insured high-income earners.[23] The average *deductible* is $675 per year.

In terms of benefits covered, we simplify the rich data and focus on three plan-generosity indicators provided by the insurer. These classify plans into *TOP*, *PLUS*, and *ECO* plans. *ECO* plans lack coverage for services such as single rooms in hospitals and treatments by a leading senior M.D. For *ECO* and *PLUS* plans, a 20 percent coinsurance rate applies if enrollees see a specialist without referral from their primary care physician. About 38 percent of all policyholders have a *TOP* plan, 34 percent a *PLUS* plan, and 29 percent an *ECO* plan. Because these plan characteristics have mechanical effects on claim sizes and correlate with policyholders' age, we control for them in our estimation of health care costs in Section 5.

4.2 Socio-Economic Panel Study

The German Socio-Economic Panel Study (SOEP) is a representative longitudinal survey that started in 1984. It collects annual information at the household and individual level from individuals above the age of 17. Currently, the SOEP surveys more than 20,000 respondents from more than 10,000 households per year (Wagner et al., 2007). We use SOEPlong (SOEP, 2018), and all existing waves as of this writing, from 1984 to 2016, in order to fully exploit the lifecycle dimension of this panel survey.[24] Table C2 (Appendix) provides summary statistics for our SOEP sample. Again, all monetary values are in 2016 USD.

Sample Selection. We leave the representative sample as unrestricted as possible, but exclude observations with missings on core variables such as age, gender, employment status or the insurance

[22] Our insurer doubled the number of clients between the 1980s and 1990s and has thus a relatively young enrollee population, compared to all GLTHI enrollees. Gotthold and Gräber (2015) report that a quarter of all GLTHI enrollees are either retirees or pensioners.

[23] Employers cover roughly one half of the total premium and the self-employed pay the full premium.

[24] Prior to 1990, the SOEP was not in the field in East Germany but started covering East Germans right after the reunification in 1990 (Wagner et al., 2007).

status. Other than that, we only exclude respondents below the age of 25 as many Germans have not entered the labor market before that age.

Income Measures. Our main income measure, *equivalized post-tax post-transfer annual income* accounts for redistribution within households and controls for economies of scale by assigning each individual a needs-adjusted income measure. Specially, *equivalized post-tax post-transfer annual income* sums over all post-tax monetary income flows at the household level, such as income from labor, capital, public and private retirement accounts, or social insurance programs.[25] Then, the total annual post-tax household income is divided by the number of household members, where we use the modified OECD equivalence scale.[26] As Table C2 shows, from 1984 to 2016, the average annual income per household member was $26,433. Note that this measure has positive values for *all* respondents, including those who are not active in the labor market.

For completeness, Table C2 also shows statistics for two additional income measures: *monthly gross wage* and *monthly net wage*. These measures have positive values for all working people with labor earnings (58 percent of observations in Table C2). The SOEP Group generates and provides these individual-level income measures to guarantee consistency over time. As seen in Table C2, the average *monthly gross wage* was $2,940 and the average *monthly net wage* was $1,921 between 1984 and 2016.

Socio-Demographics. Table C2 also provides the summary statistics of all other socio-demographic variables. In the SOEP sample, the average age is 47, and 52 percent are female. About 27 percent are white collar workers, 6 percent are self-employed, and 4 percent are civil servants. 42 percent work full-time and 14 percent part-time.

Below, we differentiate the lifecycle income processes by educational status. We do this because, after age 25, schooling degrees are largely time-invariant and determine lifecycle income substantially. Germany has a three-tier education system: *Ed 13* is one for individuals with the highest schooling degree after 13 years of schooling. *Ed 10* is one for individuals with an intermediate degree after 10 years of schooling. *Ed 8* is one for individuals who earned a degree after 8 or 9 years of schooling.

[25] The SOEP group also generates and provides these single components in a time-consistent manner.

[26] The modified OECD equivalence scale assigns a value of 1 to the household head, 0.5 to other adults, and 0.3 to children up to 14 years of age.

5 Modeling Health Risk and Income over the Lifecycle

5.1 Risk Classification

Risk classification is a key ingredient for calculating the prices of and the welfare from the short- and long-term insurance contracts. The risk classification variable represents the observed risk type of an individual at the beginning of each year. In this section we introduce a procedure that borrows insights from actuarial science, to produce an "efficient" classification. We consider our procedure to be a significant improvement over the approach used in the state-of-the art literature of dynamic contracts.

Following the literature (e.g. Einav et al., 2013; Handel et al., 2015; Ghili et al., 2019), we construct the risk classification variable using the (German version of) the John Hopkins ACG© software, which is routinely used by commercial insurers for underwriting purposes. The ACG© software provides a continuous *risk score* λ_t^*. The commonly-used approach to risk classification would use an ad-hoc criterion to partition the domain of λ_t^* into different risk classes.[27] We depart from the common approach in two key ways: First, we allow the risk class to be a function of current and lagged values of λ_t^*; $\Lambda_t^*(n) \equiv \{\lambda_t^*, \lambda_{t-1}^*, .., \lambda_{t-n-1}^*\}$, where n is determined within our procedure. Our procedure can therefore allow for higher-order dependencies in the health dynamics in a parsimonious way. Second, we propose and implement a method to discretize the vector of scores $\Lambda_t^*(n)$. Our method maximizes an efficiency criterion from the actuarial science literature, that we discuss in detail later (cf. Finger, 2001).

In the first step, we calculate the continuous score λ_t^*, which is the *unscaled total cost predicted risk* variable provided by ACG©. It is based on (a) diagnosis codes (pre-existing conditions and claim diagnoses), (b) costs of treatments, and (c) treatment episode dates. λ_t^* is meant to represent the *expected costs* in year t. In the reference population of publicly insured individuals in Germany, it has a mean of 1.

Figure 2 shows the empirical distributions of λ_t^* for our working sample in 2006 (the first year) and 2011 (the last year). Both distributions are approximately unimodal, and they appear stable over time.[28] Figure 2 also illustrates that the distribution of λ_t^* is heavily skewed and has a long right tail (consistent with stylized facts regarding the distribution of health expenditures, see French and Kelly, 2016). For example, the top percentile of the λ^* distribution has expected health expenditures

[27] For example, Ghili et al. (2019) partition the health statuses measured by λ_t^* into seven mutually exclusive and exhaustive bins, where each bin contain one-seventh of the overall sample.

[28] This suggests that excluding inflows in 2010 and 2011 due to the portability reform, see Section 2, poses no major issue.

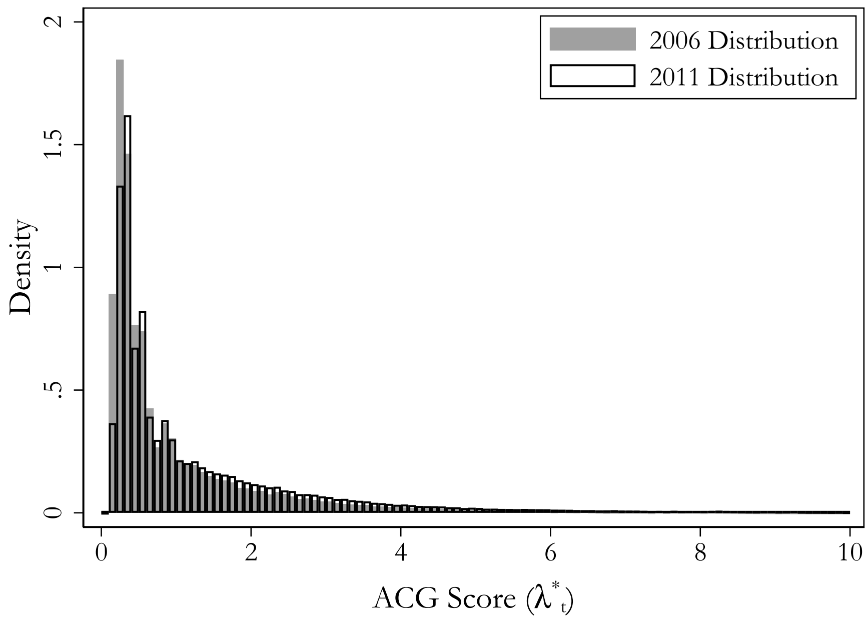


Figure 2: Distribution of λ_t^* in 2006 and 2011

Source: GLTHI claims data, ACG©, own calculations. The distribution of λ_t^* is truncated at 10; but 0.7 percent of the analysis sample have $\lambda_t^* > 10$.

$\mathbb{E}\left(m|\lambda^* \geq \mathbb{P}_{99}\right)$= \$63,422; the second highest percentile has $\mathbb{E}\left(m|\mathbb{P}_{98} \leq \lambda^* < \mathbb{P}_{99}\right) = \$30,027$; and the following three percentiles have $\mathbb{E}\left(m|\mathbb{P}_{95} \leq \lambda^* < \mathbb{P}_{98}\right) = \$19,253$, where $\mathbb{P}_k$ denotes the k-th percentile of the distribution of λ^* plotted in Figure 2.

Next we combine the continuous score λ_t^* and its $n-1$ lags into the vector of scores $\Lambda_t^*(n)$, that we map into K different *risk categories*. These categories will be ultimately combined with the individual's age for the construction of discrete health types. Modeling risk types as a discrete state serves two specific purposes. First, we allow the contract premiums to depend on the risk type. Hence, the granularity in our model should capture the granularity of the information needed by the underwriters, both in the actual environment and in counterfactual scenarios. Second, the model should be parsimonious enough to allow for modeling health dynamics with a reasonable number of parameters.

The considerable skewness in Figure 2 implies that the amount of reclassification risk will strongly depend on the granularity allowed for in the risk classification. We split the task of constructing the risk categories into two sequential problems: (1) For a given number of classes K, and the n most recent values of λ_t^*, define the efficient partitioning of the scores vector $\Lambda_t^*(n)$ into K discrete categories; (2) Find the values of K and n that lead to the best performance of the classification system. We explain the details of each step below.

Efficient Classification. According to the actuarial science literature (cf. Finger, 2001), an efficient risk classification system has two properties: *homogeneity*—meaning that individuals in one risk category are similar in terms of risk, and *separation*—meaning that categories are sufficiently different in terms of expected loss to warrant their specification as being a distinct category.[29]

For any given number of risk categories (K) and number of current and lagged values of λ_t^* (n), we define a *risk classification* as a surjective function $f_K : \Re_+^n \to \{\lambda \in \mathbb{Z} : 1 \leq \lambda \leq K\}$, where $\Re_+^n$ is the state space (i.e. λ_t^* and its $n-1$ lags). Denote this classification function $\lambda_t = f_K(\Lambda_t^*(n))$ where $\Lambda_t^*(n)$ is the vector of the n most recent ACG© scores available for an individual, and $\lambda_t \in \{1, \ldots, K\}$ is the risk category assigned to a person with those ACG© scores. According to Finger (2001), the *efficient* risk classification f_K maximizes the "structure variance" defined as

$$SV(f_K) = \text{Var}(m_t) - \sum_{k=1}^{K} \Pr(\lambda_t = k) \text{Var}(m_t \mid \lambda_t = k), \tag{6}$$

where m_t is individual annual health expenditure. The structure variance $SV(f_K)$ is thus the total variance less the weighted sum of within-class variances of health expenditures. Put differently, the efficient classification maximizes the variance of mean expenditure across groups. Applying the law of total variance to both terms in Equation (6), we can write the structure variance as:[30]

$$SV(f_K) = \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n))) - \sum_{k=1}^{K} \Pr(\lambda_t = k) \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k). \tag{7}$$

Note that the first term in Equation (7) is independent of the classification (as it is independent of the classes λ_t); thus for a given K, finding the efficient classification system is equivalent to finding the classes λ_t that *minimize* the heterogeneity in expected expenditure within risk classes: $\sum_{k=1}^{K} \Pr(\lambda_t = k) \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k)$.

Three things are worth noting about Equation (7). First, only the mean expenditure conditional on ACG© scores $\mathbb{E}(m_t \mid \Lambda_t^*(n))$ matter for the classification system, whereas the dispersion of m_t around this mean is inconsequential. Second, minimizing heterogeneity within classes is incidentally what the k-means clustering method does (Lloyd, 1982; Athey and Imbens, 2019). Thus, we will apply k-means clustering of $\mathbb{E}(m_t \mid \Lambda_t^*(n))$ to determine the efficient classification system. Third, this implies

[29] For instance, given the distribution of λ_t^* in Figure 2, it is easy to see that equally-sized categories are unlikely to be optimal as they would assign similar individuals in terms of λ^* into different categories in the left tail of the distribution, failing the *separation* principle. In addition, it would assign individuals with substantial λ^* differences into identical categories in the right tail of the distribution, failing the *homogeneity* principle.

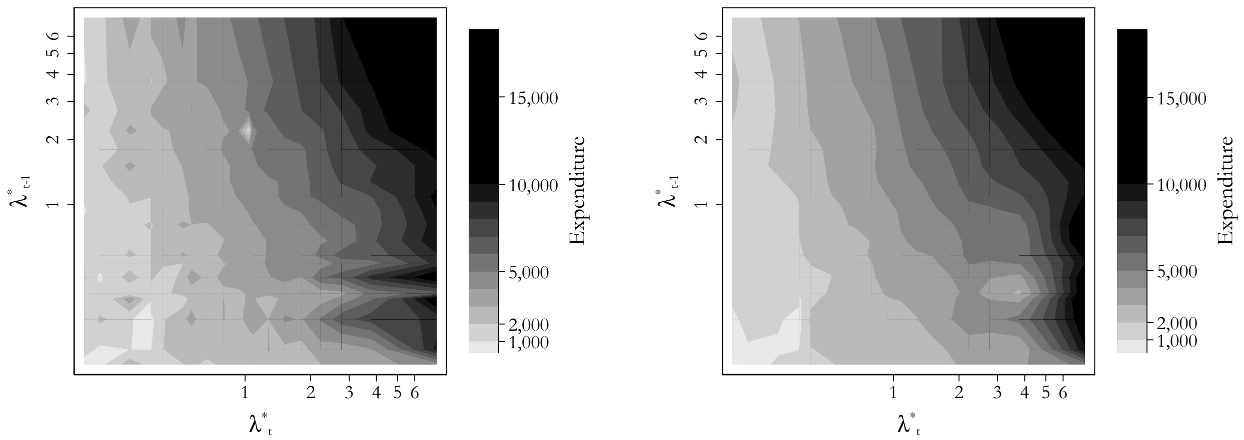
[30] The law of total variance implies $\text{Var}(m_t) = \mathbb{E}(\text{Var}(m_t \mid \Lambda_t^*(n))) + \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n)))$ and $\text{Var}(m_t \mid \lambda_t = k) = \mathbb{E}(\text{Var}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k) + \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k)$.

that the efficient classification also maximizes the coefficient of determination (R^2) in a regression of expenditure on risk class indicators (Kriegel et al., 2017).

Next, we determine the number of risk classes K and the history n (number of lags) of ACG© scores when computing $\mathbb{E}\left(m_t \mid \Lambda_t^*(n)\right)$.

Model selection. The last step of the risk classification system is to perform model selection, i.e., select values for the parameters K and n that determine, respectively, the number of risk classes and how many ACG© scores lags should be included in $\Lambda_t^*(n)$. [31] k-means clustering is an unsupervised learning method; therefore, choosing the correct number of clusters is difficult (Athey and Imbens, 2019). We proceed assuming that the objective $SV(.)$ applies also when determining these parameters. As noted above, this means we can use R^2 as our criterion for model selection.

If $n = 1$ so that $\Lambda_t^*(n) = \lambda_t^*$, the clustering algorithm can be applied to λ_t^* since $\mathbb{E}\left(m_t \mid \lambda_t^*\right) = \mu\lambda_t^*$ (where μ is the global mean expenditure). If, however, previous ACG© scores have explanatory power, $\mathbb{E}\left(m_t \mid \Lambda_t^*(n)\right)$ needs to be estimated. In order to get predictions that are accurate along the entire distribution, including the tails, we use cubic regression splines. Figure 3 provides a comparison of mean expenditure by $\Lambda_t^*(n)$ before and after smoothing for $n = 2$.



(a) Raw Averages

(b) Smoothed Expenditure

Figure 3: Mean Expenditure by Λ_t^*.

Note: The left figure is based on average expenditure within each of 400 cells (ventiles in λ_t^* and λ_{t-1}^*). The right figure uses predicted values from a cubic spline regression. Source: German Claims Panel Data.

Once $\mathbb{E}\left(m_t \mid \Lambda_t^*(n)\right)$ has been estimated for all $n > 1$, we can conduct the k-means clustering in order to maximize the objective function (7). Figure 4 shows how the performance depends on parameters K and n. For all values of n, there is initially a rapid improvement in the predictive

[31] Including lagged ACG© scores is consistent with an underwriting process often covering a relatively long medical history of the applicant (e.g., all diseases of the past 5 years and all surgeries of the past 10 years in case of our insurer).

power when we increase the number of categories K; however, this improvement levels out at quite low levels. Moreover, starting from a classification scheme that uses only the previous year's claims ($n = 1$), there is distinct improvement when we add the previous year ($n = 2$). However, adding a second lag of the ACG© scores brings only marginal improvement in the predictive accuracy. Figure 4 shows that including at least one lag and 7 distinct classes attains the best performance; increasing K or n further yields negligible improvement in performance.[32]

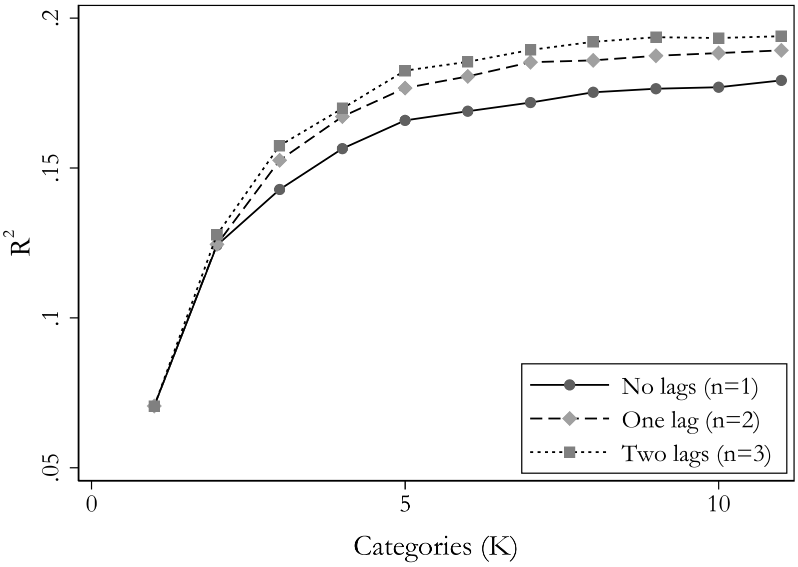


Figure 4: Performance of Alternative Risk Classifications.

Note: Each specification includes 21 age times gender fixed effects, year fixed effects and 79 plan fixed effects. Source: German Claims Panel Data.

Appendix D1 presents a number of robustness checks regarding the efficient classification system. First, we analyze the extent to which results are driven by *outliers* in m_t. It is of course desirable that the classification considers outliers, given their disproportionate contributions to means and variances; however, if the performance of the classification were widely different when they are not considered, it would cast doubt on how well the scheme performs with regard to less extreme risks. Figure D1 (Appendix) plots the performance of different classification systems when using winsorized expenditures. As expected, the topcoding of outliers improves the predictive power of all schemes; however, their relative performance is unaffected by this change.

Second, we compare two different ways of including a longer history of claims. Instead of expanding the information set $\Lambda_t^*(n)$ before discretizing, we consider an alternative based on $\Lambda_t^*(n) = \lambda_t^*$ but where we consider the predictive power of the classification scheme interacted with its lags

[32]We consistently report *unadjusted* R^2. All results are robust to using *adjusted* R^2 instead.

(i.e. a classification based on K^2 classes). Figure D2 (Appendix) provides the results. It shows that our preferred classification with K classes performs only slightly worse than the corresponding interacted classification with K^2 classes.

Third, we acknowledge that increasing n also changes the sample used for estimation. In Figure D3 (Appendix) we compare the performance over different n within the same sample. It shows that our main result is robust to the sample used.

5.2 Estimation of Transition Matrices and Expenditure Risks

Next, we estimate the transition rates between different discrete risk categories λ_t, as well as the mean expenditure by risk categories. We posit that the risk type of individual i at age t, ξ_{it}, depends on the combination of the contemporaneous risk category λ_{it} and age at t (in 5-year bins). That is, $\xi_{it} \equiv (A_{it}, \lambda_{it})$, where A_{it} is an indicator for one of the eleven age groups (five-year bands from age 25 to age 75 and 75+). It is important to note that the ACG© scores are based on an individual's age, so that, in principle, a risk category λ_{it} that uses ACG© scores as input should contain all the information needed to predict mean expenditures. However, ACG© scores are not designed to predict transitions so, in principle, transition matrices may depend on age even after conditioning on λ_{it}. As discussed below, our results confirm these predictions.

Considering that the clustering method generates a set of risk classes of very different sizes, a completely non-parametric estimation for the transition matrices $g(\xi_{it}|\xi_{i,t-1})$ and mean expenditures $\mathbb{E}(m_{it}|\xi_{it})$ is not possible. Instead, we resort to a parametric, yet flexible model. To estimate the transition matrices, we estimate a multinomial logit model for health dynamics specified as:

$$\eta_{it}^{j} = A_{it}\beta_j + L_{it}\gamma_j + h\left(A_{it}, L_{it}; \theta_j\right) + \epsilon_{it}^{j} \tag{8}$$

where η_{it}^{j} represents the log odds for $\lambda_{i,t+1} = j$, for $j \in \{2, \ldots, 8\}$. The category $\lambda_{i,t+1} = 1$ is the reference category and $\lambda_{i,t+1} = 8$ represents death. A_{it} represents i's age groups, and $L_{i,t}$ is a set of indicators for the categories of $\lambda_{i,t}$. In addition, Equation (8) includes $h\left(A_{it}, L_{it}; \theta_j\right)$ which consists of pairwise interactions of A_{it} and L_{it} with the associated parameter vector θ_j.[33]

To model the expected claims based on risk type, we follow a similar approach, but use the predicted values of claims from an OLS regression. In addition to the controls in Equation (8), we also control for a vector of dummies Q_{it} representing health plan generosity $q \in \{ECO, PLUS, TOP\}$. The

[33] We selected the interacted terms sequentially: in each iteration, we include the interaction term with the strongest association with transition rates (based on a χ^2 test), until none of the remaining interaction terms is statistically significant.

base specification is:

$$m_{it} = A_{it}\beta + L_{it}\gamma + Q_{it}\delta + h\left(A_{it}, L_{it}, Q_{it}; \theta\right) + \epsilon_{it} \tag{9}$$

In an iterative process, we add pairwise interaction terms between A_{it}, L_{it}, and Q_{it} (represented by $h\left(A_{it}, L_{it}, Q_{it}; \theta\right)$) to Equation (9) until no remaining term is statistically significant.[34] Hence, we include age groups indicators A_{it} also in the estimation of expected expenditure. As noted above, we should expect that age per-se does not have predictive power in the model for expected expenditures if our risk classification based on ACG© scores is rich and flexible enough.

Descriptive Statistics. Table 1 shows the summary statistics of total claims m by age group. Following Ghili et al. (2019), we decompose the variation of m into two components: the part that is explained by λ, i.e., S.D. of $\mathbb{E}\left(m \mid \lambda\right)$;[35] and the residual variation around the predicted value, i.e., S.D. around $\mathbb{E}\left(m \mid \lambda\right)$.

As expected, mean claims strongly increase in age: they almost double from \$1,996 in age group 25 to 30, to \$3,719 in age group 45 to 50, almost double again to \$7,151 in age group 65 to 70. For enrollees above 75 years, the average amount of claims is \$10,020 (all values are in 2016 U.S. dollars). This age gradient is, however, accounted for by our risk classification. Even though a few age-related parameters in Equation (9) turn out statistically significant, the deviations from mean expenditure within each risk class are economically insignificant. Figure C1 (Appendix) illustrates this point. We interpret it as evidence that our preferred risk classification is rich enough.

Table 2 shows how different age groups are distributed across risk categories λ, and it shows a clear age gradient in health expenditure risk. The probability of being in the lowest risk category, i.e., $\lambda = 1$, declines progressively with age, whereas the share of enrollees in the five highest categories increases in age; the pattern is particularly pronounced for categories $\lambda = 4$ and $\lambda = 5$. Only 1.7 percent of enrollees between 25 and 30 years are in categories $\lambda = 4$ and $\lambda = 5$. This share almost quadruples to 6.2 percent in age group 45 to 50, and then more than quadruples again to 28.6 percent in age group 65 to 70. It is 61 percent for enrollees above 75 years. On the other hand, risk category $\lambda = 7$ clearly represents catastrophic costs and covers at most 0.3 percent of the population in any age group.

[34] The estimation of conditional expenditure given λ_t is based on a subsample of clients with moderately-sized deductibles. The reason is that clients with large deductibles may decide not to submit their claims, which leads to a downward bias in the estimates. This is less of a concern for the risk classification λ_t^*, which is based on a much broader set of information on the clients and on treatment episodes. In Appendix section D2 we provide some descriptives for this subsample, which generally confirm that this assumption is reasonable.

[35] This statistic also corresponds closely to the maximand of the risk classification algorithm, cf. Section 5.1 above.

Table 1: Health Expenditure Claims m by Age Group

Ages	Mean	S.D.	S.D.($\mathbb{E}(m \mid \lambda)$)	S.D.($m - \mathbb{E}(m \mid \lambda)$)
All	4,109	9,451	3,494	8,806
25-	1,996	5,529	1,782	5,234
30-	2,619	6,050	1,938	5,731
35-	2,840	6,312	2,086	5,957
40-	3,119	7,153	2,411	6,734
45-	3,719	8,444	2,946	7,913
50-	4,880	9,866	3,544	9,208
55-	6,517	12,679	4,573	11,825
60-	7,635	18,608	4,299	18,104
65-	7,151	12,753	4,421	11,963
70-	8,355	13,837	5,026	12,892
75-	10,020	13,485	4,490	12,715

Source: German Claims Panel Data. Sample includes all age groups and uses the ACG© scores to construct risk categories λ as explained in Section 5.1.

Table 2: Health Risk Categories λ by Age Group

Age	1 (Healthiest)	2	3	4	5	6	7 (Sickest)
25-30	0.789	0.154	0.039	0.013	0.004	0.001	0.000
30-35	0.740	0.178	0.054	0.020	0.006	0.001	0.000
35-40	0.652	0.225	0.085	0.027	0.009	0.002	0.000
40-45	0.622	0.227	0.103	0.034	0.012	0.003	0.000
45-50	0.539	0.258	0.136	0.046	0.016	0.004	0.001
50-55	0.463	0.263	0.174	0.068	0.024	0.007	0.001
55-60	0.291	0.319	0.232	0.108	0.036	0.011	0.002
60-65	0.184	0.313	0.269	0.155	0.058	0.019	0.003
65-70	0.069	0.291	0.337	0.217	0.069	0.014	0.002
70-75	0.019	0.203	0.347	0.309	0.105	0.015	0.002
75+	0.000	0.092	0.267	0.422	0.188	0.029	0.003

Source: German Claims Panel Data. Sample includes all age groups and uses the ACG© scores to construct risk categories λ as explained in Section 5.1.

Transitions between States. Table 3 displays one-year transition rates between health risk categories for all age groups; the numbers are predicted probabilities based on Equation (8). Two facts emerge from Table 3. First, we find strong persistence in health risk. For instance, an individual with $\lambda_t = 1$ has an 83 percent probability of $\lambda_{t+1} = 1$. The likelihood of staying in the same category between two consecutive years decreases over risk categories but, still, 45 percent of individuals in category 7 remain in category 7 in the next year. Second, despite the high persistence, the likelihood of a severe health shock (and thus the reclassification risk) is non-trivial even when just considering two calendar years. For example, the probability of ending up in risk category 4 in $t + 1$ is 3.6 percent after being category 2 in year t.

Table 3: Health Risk Category Transitions

	λ_{t+1}							
λ_t	1	2	3	4	5	6	7	8 (†)
1	0.831	0.158	0.006	0.003	0.001	0.001	0.000	0.001
2	0.214	0.523	0.215	0.036	0.009	0.001	0.000	0.002
3	0.050	0.179	0.572	0.164	0.029	0.003	0.000	0.003
4	0.024	0.053	0.227	0.541	0.128	0.013	0.001	0.013
5	0.018	0.027	0.035	0.330	0.445	0.104	0.005	0.036
6	0.010	0.018	0.017	0.096	0.294	0.409	0.052	0.104
7	0.002	0.005	0.002	0.027	0.085	0.200	0.452	0.226

Source: German Claims Panel Data. Sample includes all years, all age groups, and uses the ACG© scores to construct risk categories λ as explained in Section 5.1.

The transition rates are highly dependent on age. Tables D1 and D2 (Appendix) show transition matrices for each of the 11 age groups. For example, the probability of remaining in state 1 decreases from 89 percent among 25-year-olds to 18 percent among individuals above 75. Also the probability of recovering, i.e. transitioning from a higher to a lower risk class, is declining in age. Moreover, the mortality rates increase rapidly with age—in particular for states below 7. All these differences are statistically significant. Therefore, allowing for age-dependent transition rates is necessary even though, as noted above, expected expenditure conditional on risk class is constant in age.

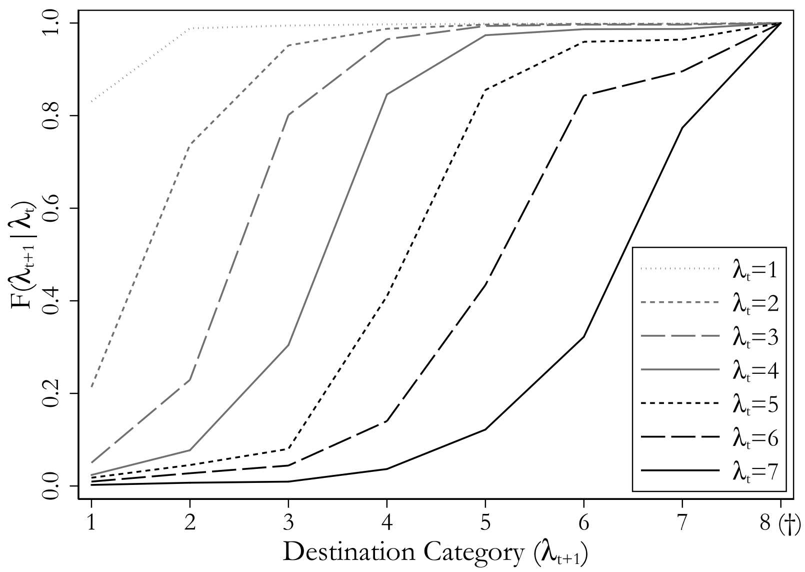


Figure 5: Stochastic Dominance.

Stochastic Dominance. In their characterization of the optimal contract, Ghili et al. (2019) invoke an assumption of stochastic dominance. It requires that transition rates between risk categories—which are represented by the cumulative distribution function $F(\lambda_{t+1} \mid \lambda_t)$—satisfy first-order stochastic dominance in the following sense: if $\lambda'_t > \lambda_t$, then $F(\lambda_{t+1} \mid \lambda'_t) \succ_{FSD} F(\lambda_{t+1} \mid \lambda_t)$. In Figure 5 we show that this property holds for all pairwise combinations of (λ_t, λ'_t) such that $\lambda'_t > \lambda_t$.

5.3 Lifecycle Income Paths

Next, we estimate the lifecycle income paths using 33 years of SOEP panel data. Because individuals may enroll in GLTHI contracts during their entire lifetime, we consider all sources of income beyond wages. Our main income measure is the *equivalized post-tax post-transfer annual income*, which sums over all post-tax income flows at the household level, and then normalizes by the number of household members (see Section 4.2). Using this income measure, we estimate the following individual fixed effects model:

$$log(y_{it}) = \theta_i + f(age_{it}) + \epsilon_{it} \tag{10}$$

where y_{it} stands for our income measure in 2016 U.S. dollars in year t for individual i; and θ_i are individual fixed effects which net out all persistent individual time-invariant income determinants, such as gender, preferences, or work productivity. The flexible function $f(age_{it})$ represents a series of age fixed effects and identifies the main coefficients of interest. They capture the main features of the German lifecycle income profiles from 1984 to 2016.

We estimate this income process separately by educational status for the two following groups: (a) individuals with the highest schooling degree after 13 years of schooling (*Ed 13*), and (b) individuals with an intermediate degree after 10 years of schooling (*Ed 10*).[36] We estimate separate income processes by education groups because lifecycle profiles differ substantially by educational degree (Becker and Chiswick, 1966; Bhuller et al., 2017). As mentioned, the steepness of these lifecycle income profiles will determine the welfare consequences of long-term health insurance to a large extent.

The dashed curves in Figure 6 show the estimated age fixed effects for the two groups. Income rises sharply between age 25 and age 57. Then it decreases substantially until around age 70, from which point it remains relatively flat until death. It is also easy to observe a level difference in income

[36] Germany has three different schooling tracks where the majority of students complete school after 10 years and then start a three-year apprenticeship (cf. Dustmann et al., 2017).

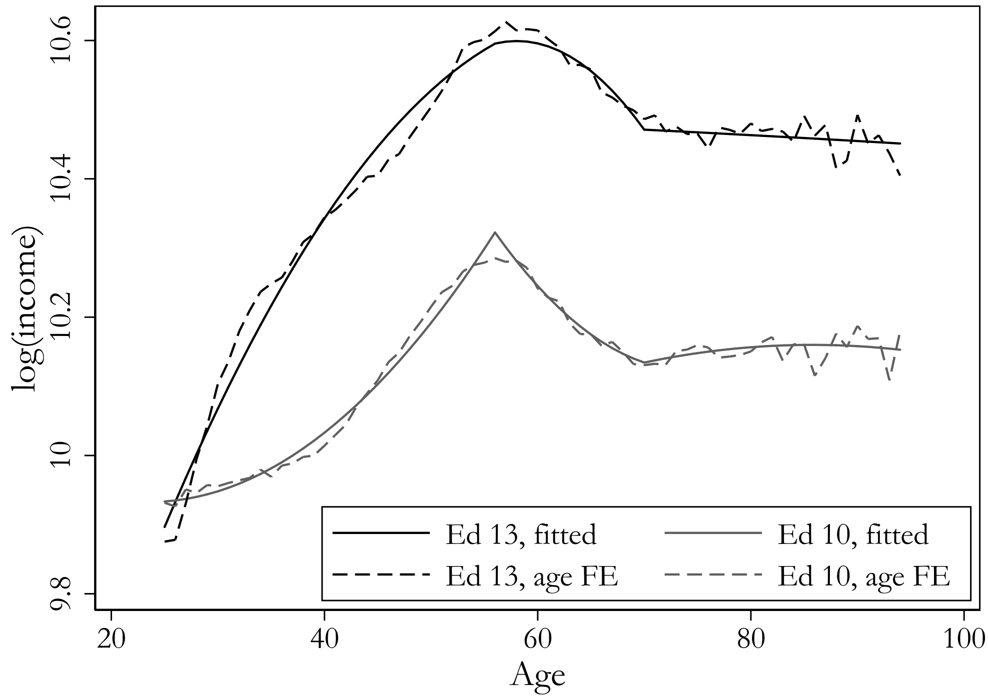


Figure 6: Lifecycle Income Paths Germany, Nonparametric and Fitted.
Source: SOEP (2018), years 1984 to 2016. All values in 2016 USD.

paths between the two educational groups over the entire lifecycle.

Several factors can explain the lifecycle income pattern in Figure 6. First, the labor market entry and subsequent careers significantly increase post-tax income between the main working ages 25 and 55. Second, our income measure includes social insurance benefits, and the German welfare state is known for its generosity. Third, it may be surprising that equivalized household income starts to decrease after age 57 until around age 70. However, especially in the 1980s and 1990s and also today, many Germans retire early (Börsch-Supan and Jürges, 2012); others reduce their working hours, for example, to take care of their grandchildren or provide long-term care for their parents (Schmitz and Westphal, 2017). Finally, the stable permanent income stream from age 70 until death may be explained by the fact that our income measure includes primarily statutory pensions, employer-based pensions and private pensions (Geyer and Steiner, 2014; Kluth and Gasche, 2016; Engels et al., 2017).

We accommodate these lifecycle income patterns by fitting $f(age_{it})$ as piece-wise squared polynomial of age, where we allow the parameters of age and age^2 to differ by education group and across three different age bins: $[25, 56]$, $[56, 70]$ and $70+$. This is illustrated by the two solid curves in Figure 6. Note that the piece-wise squared polynomials fit the empirical lifecycle profiles very well.

6 Main Results

6.1 Equilibrium Lifecycle GLTHI Premiums

After estimating the health risk process, we can calculate the equilibrium GLTHI lifecycle premiums by solving Equation (1) using backwards induction. Note that $P_t(\xi_t)$ in Equation (1) is the guaranteed-renewable premium that an individual with health ξ_t would be offered if she entered a contract in period t in the GLTHI market. Therefore, the equilibrium GLTHI premiums correspond to 490 values: premiums depend on enrollee's current health category $\lambda_t \in \{1, 2, ..., 7\}$, as well as age $t \in \{25, ..., 94\}$. We use a discount factor $\delta = 0.966$ (corresponding to a discount rate of 3.5 percent).

Figure 7 plots the resulting premiums for a handful of the most relevant combinations: $\lambda_t = 1$ and $t \in \{25, ..., 59\}$; $\lambda_t = 2$ and $t \in \{25, ..., 74\}$; $\lambda_t = 3$ and $t \in \{65, ..., 94\}$; $\lambda_t = 4$ and $t \in \{60, ..., 74\}$; $\lambda_t = 5$ and $t \in \{75, ..., 94]\}$. These combinations represent the three most common states for each corresponding age interval.

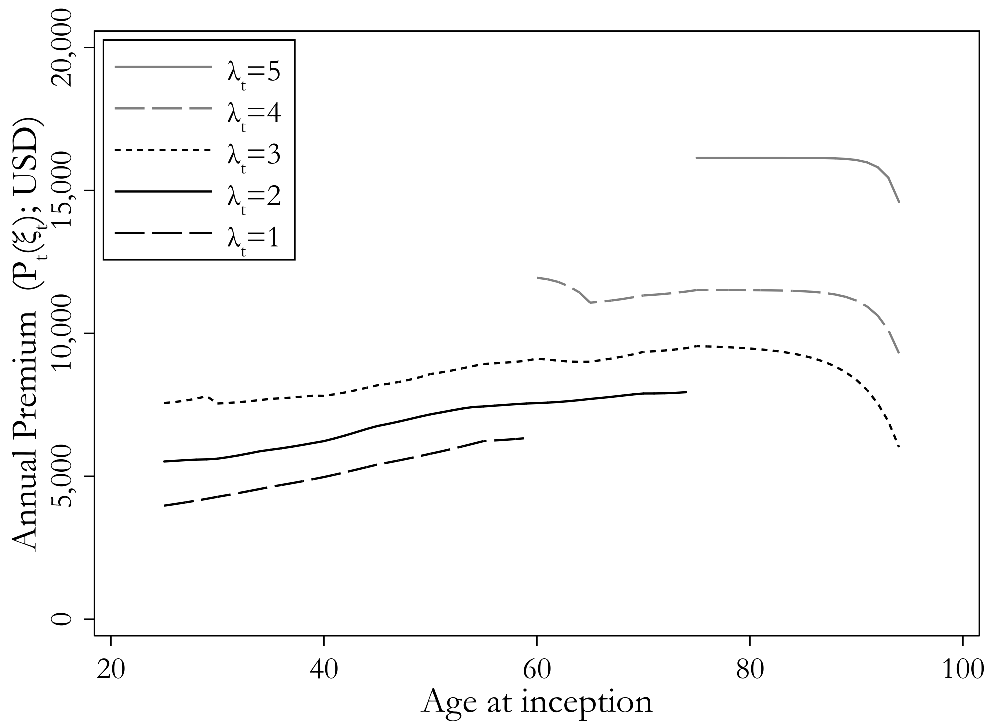


Figure 7: Calibrated Starting Premiums $P_t(\xi_t)$ in the GLTHI

Three forces are at play that determine the lifecycle profile of $P_t(\xi_t)$ in Figure 7. First, $P_t(\xi_t)$ is an increasing function of ξ_t. This is because, for any age, a higher health risk classification is associated with higher current and future health claims (both through their effect on current claims and their effect on health transitions).

Second, starting premiums increase with age for most age ranges. This is because expected health care claims and health transitions depend on age (through the A_t component of ξ_t). As a consequence, the annualized net present value of health care claims of an individual with a given λ_t increases with age for most of the age ranges.

Third, individuals who renew are an adversely selected portion of contract holders, i.e., those who either remain or become sick enough to not get better outside offers in the market. The insurance company breaks even by charging a front-loaded premium that takes into account this dynamic adverse selection. However, for any given health type, the probability of transitioning towards a worst health status in the future decreases with age. Therefore, the need to front-load premiums to fund future negative health shocks *decreases* over the lifecycle. This force explains why $P_t(\xi_t)$ decreases with t when t is sufficiently large.

In Figure E1a and E1b (Appendix), we compare the calibrated and the observed premiums by age at inception. First, we observe positively sloped starting premiums by age over the entire age range, both for the calibrated and the observed premiums. Second, there are clear level differences by health risk such that the starting premiums are a clear function of λ_t—sicker applicants have to pay higher premiums. This rank ordering persists over the entire lifecycle. Third, although the premium levels for sicker individuals are slightly larger in the calibrated than the observed case, the two Figures E1a and E1b show very similar starting premiums by age and health risk.

6.2 Comparison with the Optimal Dynamic Contract

This subsection explicitly compares lifecycle claims, premiums and the amount of front-loading between the GLTHI and the theoretically optimal dynamic contract. In contrast to the GLTHI contract, the optimal contract directly depends on on lifecycle income and the premium paths change after income changes and health shocks (Ghili et al., 2019). It specifies evolving consumption guarantees over the lifecycle where policyholders have time-separable and risk averse preferences (see Appendix B).

Using our empirical health transition and income dynamics, Table 4 illustrates the differences between the GLTHI and the optimal contract by comparing the contract terms at age 25. Panel (a) shows the GLTHI premium and front-loading amounts for a 25 year old by the health status $\lambda_{25} \in \{1, ..., 7\}$. With health status $\lambda_{25} = 1$, she pays a premium of \$3,973, which is \$2,499 in excess of expected claims. Individuals with higher λ's pay higher premiums, but the amount of front-loading decreases. For example, for $\lambda_{25} = 3$ the premium is \$7,563 which includes \$1,545 in front-loading. Note that the amount of front-loading decreases, the worse the current health status is. The reason is that the likelihood of a further health deterioration also decreases, the worse the current health status is. Again, note that the GLTHI premiums do *not* depend on lifecycle income (see Remark 2).

Table 4: Comparing GLTHI Contract to Optimal Contract Terms at Inception

λ_{25}	1	2	3	4	5	6	7
Expected claims	1,473	3,559	6,019	9.302	14,600	24,554	54,930
(a) GLTHI							
Premium	3,973	5,517	7,563	10,363	15,291	24,561	54,930
Front-loading	2,499	1,957	1,545	1,062	691	7	0
(b) Optimal contract *Ed 13*							
Premium	1,895	4,578	6,988	10,103	15,187	24,554	54,930
Front-loading	421	1,019	970	801	586	0	0
(c) Optimal contract *Ed 10*							
Premium	2,571	5,366	7,489	10,307	15,273	24,554	54,930
Front-loading	1,097	1,807	1,471	1,006	673	0	0

Source: German Claims Panel Data, SOEP data. Table shows expected health care claims, starting premiums, and the amount of front-loading by health risk category at age 25, $\lambda_{25} \in \{1, ..., 7\}$. All values in 2016 USD.

Panel (b) of Table 4 compares the premiums and front-loading amount for the optimal dynamic contract for an individual with the highest schooling degree (*Ed 13*) by initial health at age 25. For almost all health states, compared to GLTHI, the initial premiums and front-loading amounts are lower and consumption higher in the optimal dynamic contract. However, the differences in premiums between the GLTHI and the optimal dynamic contract are smaller, the worse the health status at the inception of the contracts. For $\lambda_{25} = 1$ the optimal premium is \$1,895 (vs. \$3,973 for GLTHI) and for $\lambda_{25} = 4$, the optimal premium is \$10,103 (vs. \$10,363 for GLTHI). The optimal contract entails less front-loading than GLTHI because a higher front-loading increases the marginal utility of consumption.

Panel (c) of Table 4 shows the optimal contract for an individual with a schooling degree after 10 years of schooling (*Ed 10*). This individual has a flatter income profile over her lifecycle (see Figure 6), which is why the optimal contract entails a *higher* degree of front-loading for *ED10* eduation group, especially for healthy individuals. In general, the premium and front-loading amounts for *ED 10* with $\lambda_{25} \in \{1, ..., 5\}$ lie between those of the optimal dynamic contracts for *Ed 13* and the GLTHI. Again, the front-loading amount is lower, the sicker the individual is at inception.

Finally, comparing Panels (a)-(c), we see that the GLTHI premiums converge to the optimal premiums for both educational groups for the three sickest health states at inception $\lambda_{25} \in \{5, 6, 7\}$.

6.3 Welfare Results

We now calculate welfare under the different contracts as defined in Section 3.2. We calculate welfare by simulating the economy for a lifecycle of 70 years, from age 25 to age 94 for $N = 500,000$ individuals. Note that, so far, we have not specified the utility function because the GLTHI premium does not hinge on a specific utility function. However, for welfare comparisons, we need to assume some utility function. For the baseline results, we follow the convention and use a constant absolute risk aversion (CARA) utility function of the form:[37]

$$u(c) = -\frac{1}{\gamma}e^{-\gamma c}. \tag{11}$$

In our main results, we use a risk aversion parameter $\gamma = 0.0004$ (cf. Ghili et al., 2019). In Section 6.6, we will explore the robustness of the welfare results with respect to γ, and also under non-time-separable Epstein-Zin preferences.

We provide nine sets of results, corresponding to different assumptions regarding the probability simplex that determines the initial state, $\Delta_0 \in \Delta^7$. Panels (a) to (g) of Table 5 show the results assuming that individuals start in each of the seven possible health states. For instance, Panel (a) assumes that everyone starts in the healthiest state, such that $\Delta_0 = \frac{1}{100}[100, 0, 0, 0, 0, 0, 0]$. Panel (h) assumes that λ_{25} is drawn from the distribution implied by the transition matrix at age 25, given $\lambda_{24} = 1$ (see Table D1, Appendix). By doing so, we accurately replicate the distribution of ξ among the 25-30 age group. In Panel (h), we also assume that individuals cannot start in the worst possible health state, which makes sense given that insurers have the right to deny coverage, and that the public SHI system acts as a fall-back option for young and sick individuals. As discussed in Section 5.3, we stratify the findings by two different education-dependent lifecycle income paths.

Column (1) calculates welfare under the first-best contract as described by Equation (3); Column (2) calculates welfare under a series of short-term contracts, C_{ST} (Equation (4)); Column (3) shows the results under the GLTHI contracts, C_{GLTHI}; and Column (4) calculates the welfare under the optimal dynamic contract, C_{GHHW}. Columns (5) and (6) show the welfare difference between GLTHI and a series of short-term contracts, and between GLTHI and the optimal contract.

Overall, Table 5 shows the following: First, Column (1) shows that welfare in the first-best scenario is always lower for the lower educated (*Ed 10*) and decreases with health at inception. For

[37] The CARA utility function has the convenience of allowing for negative consumption, which occurs when income is lower than the required premium payments, for example, but it also implies that the consumption equivalent may be negative under some contracts, as we will see in Table 5.

Table 5: Benchmarking Welfare under GLTHI

	C^* (1)	C_{ST} (2)	C_{GLTHI} (3)	C_{GHHW} (4)	$\frac{C_{GLTHI}-C_{ST}}{C^*-C_{ST}}$ (5)	$\frac{C_{GHHW}-C_{GLTHI}}{C_{GHHW}}$ (6)
	Panel (a): $\Delta_0 = \frac{1}{100}[100, 0, 0, 0, 0, 0, 0]$					
Ed 10	23,027	-10,058	21,536	22,488	0.955	0.042
Ed 13	34,207	-2,114	26,024	27,726	0.775	0.061
	Panel (b): $\Delta_0 = \frac{1}{100}[0, 100, 0, 0, 0, 0, 0]$					
Ed 10	22,601	-10,807	20,840	21,373	0.947	0.025
Ed 13	33,777	-4,088	24,897	25,570	0.765	0.026
	Panel (c): $\Delta_0 = \frac{1}{100}[0, 0, 100, 0, 0, 0, 0]$					
Ed 10	22,247	-10,713	19,857	20,171	0.927	0.016
Ed 13	33,422	-2,436	23,274	23,622	0.717	0.015
	Panel (d): $\Delta_0 = \frac{1}{100}[0, 0, 0, 100, 0, 0, 0]$					
Ed 10	21,907	-10,811	18,254	18,409	0.888	0.008
Ed 13	33,082	-2,260	20,945	21,101	0.657	0.007
	Panel (e): $\Delta_0 = \frac{1}{100}[0, 0, 0, 0, 100, 0, 0]$					
Ed 10	21,472	-10,941	14,676	14,713	0.790	0.002
Ed 13	32,644	-2,366	16,597	16,645	0.542	0.003
	Panel (f): $\Delta_0 = \frac{1}{100}[0, 0, 0, 0, 0, 100, 0]$					
Ed 10	20,635	-11,172	5,966	5,967	0.539	0.000
Ed 13	31,805	-2,596	7,568	7,574	0.295	0.001
	Panel (g): $\Delta_0 = \frac{1}{100}[0, 0, 0, 0, 0, 0, 100]$					
Ed 10	11,589	-27,085	-27,070	-27,070	0.000	0.000
Ed 13	22,327	-24,631	-24,630	-24,630	0.000	0.000
	Panel (h): $\Delta_0 = \frac{1}{100}[89.11, 10.25, 0.47, 0.11, 0.04, 0.03, 0]$					
Ed 10	22,980	-10,119	21,168	21,945	0.945	0.035
Ed 13	34,159	-2,223	25,088	26,093	0.751	0.039

Source: German Claims Panel Data, SOEP data. Table shows welfare measured by the consumption certainty equivalents in 2016 USD dollars, per capita, per year, separately for two income profiles (see Figure 6). Panels (a) to (g) differentiate by initial health status $\lambda_{25} \in \{1, ..., 7\}$. In Panel (i), we do not allow 25 year olds to be in the worst health risk category. Columns (1) to (4) show welfare according to the (1) first-best (C^*), (2) a series of short-term contracts (C_{ST}), (3) the GLTHI, and (4) the optimal contract (C_{GHHW}). Column (5) shows how much of the welfare gap between (2) and (1) is closed by GLTHI. Column (6) shows the percentage of welfare loss under GLTHI relative to the optimal contract.

example, for individuals with the highest schooling degree who are in the healthiest risk category at age 25, the consumption certainty equivalent is $34,207 per year. This decreases to $22,327 for those 25 year olds who are in the sickest risk category.

Second, Column (2) shows that a series of short-term contracts C_{ST} produces large welfare losses compared to the first-best. For all initial health states at age 25 and for both lifecycle income profiles, the consumption certainty equivalents are always negative.[38]

Third, the GLTHI contract produces substantial welfare gains compared to short-term contracts. Consider Panel (a) for the case when $\lambda_{25} = 1$ at inception, i.e. $\Delta_0 = \frac{1}{100}[100, 0, 0, 0, 0, 0, 0]$. Column (3) shows that, under GLTHI, the consumption certainty equivalent is \$21,536 for *Ed 10* and \$26,024 for *Ed 13*. Column (5) shows that the GLTHI contract closes 96 and 78 percent of the welfare gap between a series of short-term contracts and the first-best for *Ed 10* and *Ed 13* individuals, respectively. Column (4) presents the welfare under the theoretically optimal contract, which is higher than under GLTHI for both education groups. However, the welfare gap between the two is quite small, at only 4.2 and 6.1 percent (Column (6)).

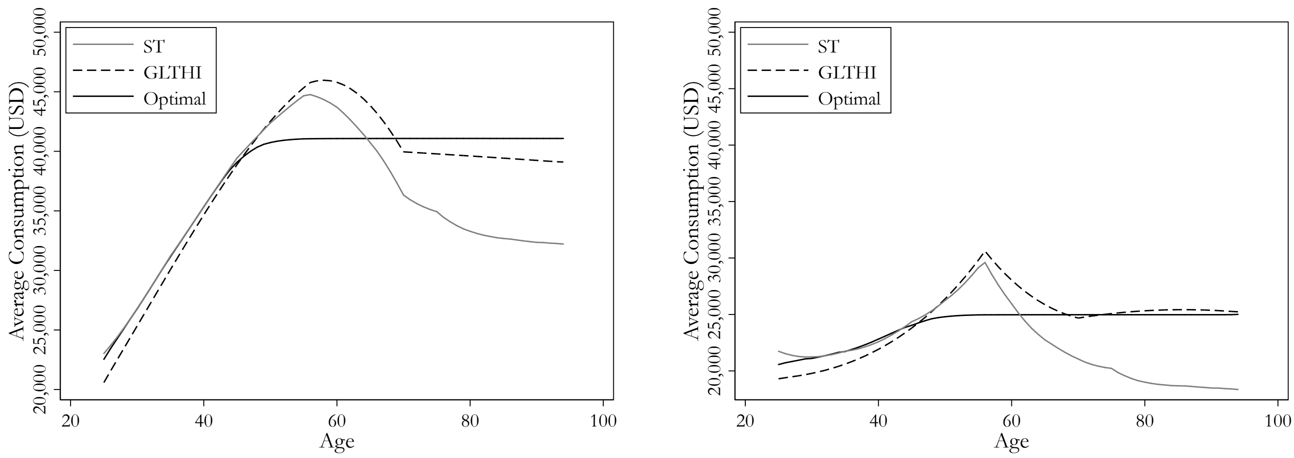
Fourth, when evaluating welfare under different distributions over the initial health states, the findings discussed above turn out to be systematic. The welfare differences between the GLTHI and the optimal contract for initial health states $\lambda_{25} \in \{2, 3, 4\}$ are reported in Panels (b)-(d). Column (6) in these panels show that the welfare differences are only between 0.8 to 3 percent, for both education groups. For very bad initial health states, $\lambda_{25} \in \{5, 6, 7\}$, column (6) in these panels show that the welfare differences are almost identical for both education groups. However, if an enrollee's health status at inception is in the sickest state, $\lambda_{25} = 7$, then both the GLTHI and the optimal GHHW contracts produce negative welfare as measured by the CE, while the welfare is positive under the first best. This highlights the significant negative welfare consequences of *one-sided commitment*, i.e., the inability of enrollees to commit to long-term contracts, together with the inability of consumers to borrow.[39] In Panel (h), where the initial health distributions corresponds to the observed empirical distribution for age-25 enrollees in our sample, we find that the welfare loss under the GLTHI contracts relative to the optimal contract is at most 3.9 percent. In Appendix F, we further explore the robustness of this finding: considering a large number of draws of distributions over starting states, we conclude that the welfare loss is bounded at around six percent for the better-educated group and at around four percent for the less-educated group. This exercise also confirms that the welfare differences between the GLTHI contract and the optimal contract is smaller when the population is less healthy at the beginning, alleviates the concern that our findings are driven by the fact that policyholders in our sample are a relatively healthy subsample of the overall population.

[38] Recall that the CARA utility function as specified by Equation (11) allows for negative consumption.

[39] As is well known, if consumers can borrow, they can "manufacture" commitment power by posting a "bond" with the insurer that equates the discounted sum of expected medical claims (see, e.g., Cochrane (1995)).

6.4 Understanding GLTHI Welfare

Average Lifecycle Consumption and Intertermporal Consumption Smoothing. We now delve deeper into how the short-term contract, the GLTHI contract, and the optimal contract affect individuals' *intertemporal consumption smoothing* and the *consumption volatility* over their lifecycles. Figure 8 plots *average* consumption for these three contracts over the lifecycle, separately for *Ed 10* (Figure 8a), and *Ed 13* (Figure 8b). The figures illustrate the driving forces behind the welfare differences in Table 5.

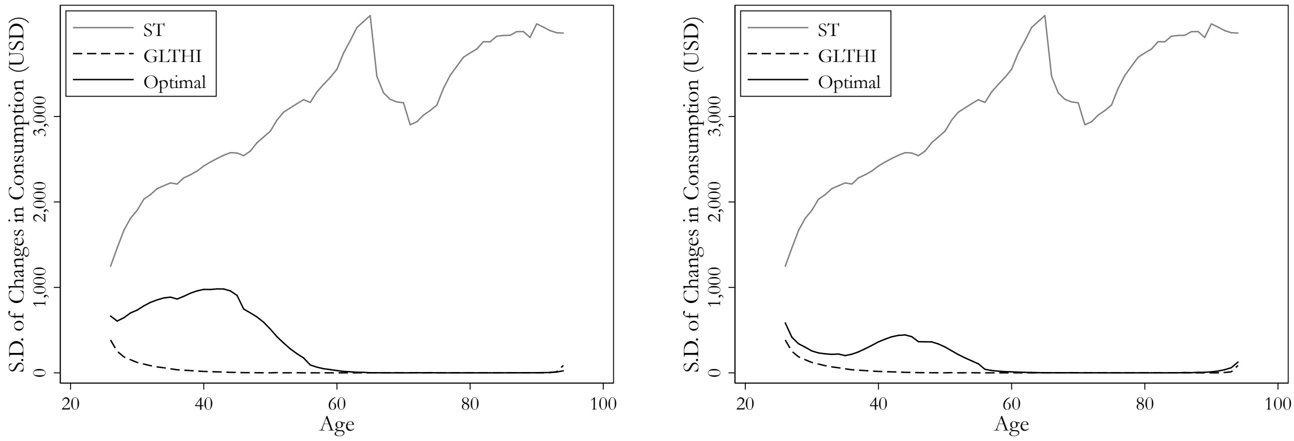


(a) *Ed 13* (b) *Ed 10*

Figure 8: Expected Consumption over the Lifecycle by Education

As shown by the thin solid lines, under a series of short-term contracts, average consumption is simply income (Figure 6) minus expected health care spending (see Equation 4). The average consumption profile is therefore hump-shaped over the lifecycle for both education groups. As shown by the dashed lines, under the GLTHI contract, average consumption has a similar shape, but starts at a lower level and is higher at older ages. This reflects the heavy front-loading of GLTHI up to the early 50s. As shown by the thicker solid lines, under the optimal contract, average consumption accounts for the utility from not only reducing the reclassification risk, but also from the smoothing of consumption over the lifecycle. Hence, the optimal contract implies a much smaller degree of front-loading than the GLTHI contract (Table 4). Thus, compared to GLTHI, the average consumption under the optimal dynamic contract would start at a higher level, particularly for the highly educated who have steeper income profiles and for whom front-loading is costlier. As individuals approach their middle ages, the optimal contract allows to fully smooth consumption, which is illustrated by the straight flat consumption line after around age 40.[40] However, relative to GLTHI, the optimal contract has *more* reclassification risk.

Reclassification Risks. To illustrate the degree of *reclassification risk* over the lifecycle, Figure 9 displays the *standard deviations* of consumption *changes* over the lifecycle for the GLTHI contract, and compares it to both a series of short-term contracts and the optimal contract. (That is, Figure 9 plots, for each age t, the standard deviation of $\Delta C_{i,t} \equiv C_{i,t+1} - C_{i,t}$.)



(a) *Ed 13* (b) *Ed 10*

Figure 9: Simulated Standard Deviation of Consumption Changes

As seen, the GLTHI contract imposes very little reclassification risk as most individuals lock in $P_{25}(\cdot)$ in the first period. The few individuals who switch contracts are those who start with $\lambda_{25} > 1$ and become sufficiently healthier over the lifecycle (such that $P_t(\xi_t) > P_{25}(\xi_{25})$ for some $t > 25$). However, this is a rare event, especially after age 40. On the other hand, the optimal dynamic contract entails consumption bumps early in life. For instance, the consumption guarantee under GHHW increases for individuals who start at $\lambda_{25} = 1$ and remain at $\lambda_{26} = 1$ in the following year. The reason is that a competing insurer can take into account the "good news" regarding future health, contained in the event "$\lambda_{25} = 1$ and $\lambda_{26} = 1$," and offer the individual a higher consumption guarantee, and still break even in expectation. Finally, the standard deviation of consumption changes increases strongly between age 25 and 60 for a series of short-term contracts, then decreases slightly up to age 70 and then increases again until death.

[40]Furthermore, as we will show in Figure 11 below, with a risk aversion parameter of $\gamma = 4 * 10^{-4}$, the welfare differences between GLTHI and GHHW contracts due to differences in the *expected* consumption profiles over the lifecycle are meaningful. Barring differences in reclassification risk across contracts, the lifecycle consumption under the GHHW contracts produces welfare gains of approximately US 2,600 per year.

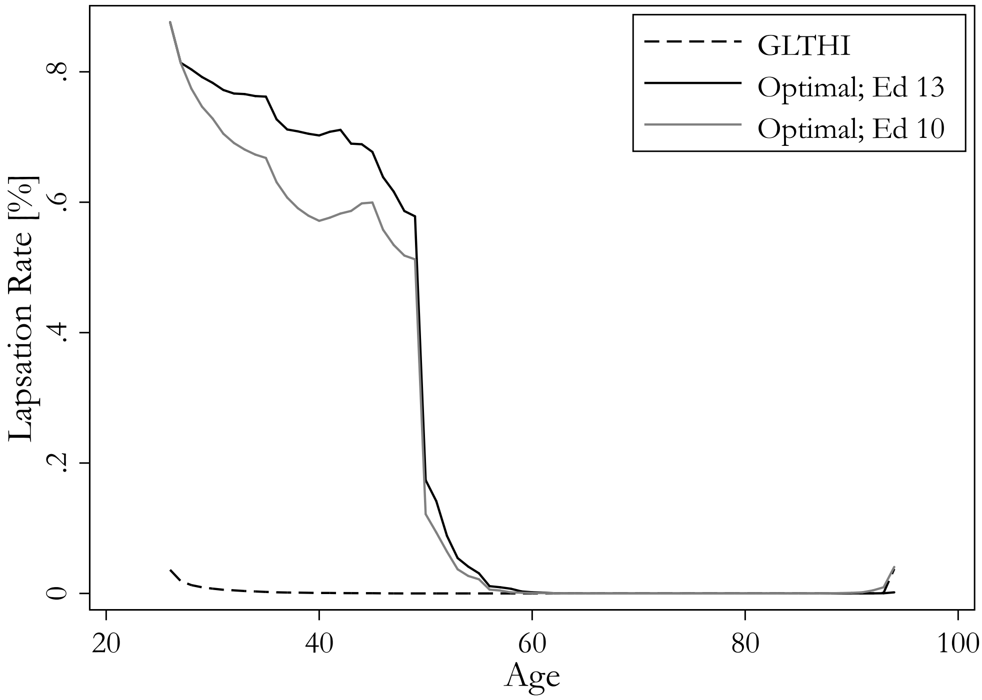


Figure 10: Laspation Rates over the Lifecycle by Education

Figure 10 compares average lapsation rates under each contract.[41] As expected, lapsation from GLTHI is extremely low over the entire lifecycle. In contrast, when expected future health improves, the optimal contract results in higher consumption for the healthiest types (and therefore for sicker types too) early in life. Lapsation in the optimal contract decreases substantially in the late 40s. At this point, most individuals have achieved their consumption plateau. Subsequently, consumption remains constant in order to transfer resources intertemporally and to save for old age.

Summary. Compared to the optimal contract, the GLTHI contract entails too much front-loading and too little consumption volatility and reclassification risks. As income profiles for both education groups tend to rise fast in early ages, compared to the optimal contract, the GLTHI falls short of sufficient intertemporal consumption smoothing, as illustrated by Figure 8. However, the extra front-loading results in a lower standard deviation of consumption changes and much lower lapsation rates than the optimal contract would dictate. Of course, by design, the optimal contract optimally balances these trade-offs and thus—in environments satisfying the conditions required for Ghili et al.

[41]Lapsing under the optimal contract is defined as an increase in the consumption guarantees. As noted by Ghili et al. (2019), optimal contracts impose a "no-lapsation constraint", so that the consumer will always stay in the same contract. However, an increase in the consumption guarantee specified within a contract can be also interpreted as a lapsation from an equivalent set of guaranteed premium paths. Figure 10 uses this interpretation of lapsing.

(2019)'s theoretical characterization—achieve a higher welfare than the GLTHI contract. Our main findings show, however, that the GLTHI contract—despite its simplicity—achieves welfare that is very close to the optimal contracts.

Robustness Varying IES and Risk Aversion. The key welfare trade-offs to understand are intertemporal consumption smoothing vs. consumption volatility. The welfare effect of intertemporal consumption smoothing depends on the *intertemporal elasticity of substitution* (IES), and the welfare effect of consumption volatility depends on *risk aversion*. The time-separable preference assumed so far imposes that IES and risk aversion are parametrically linked. In Section 6.6, we will break the parametric link between IES and risk aversion, and conduct welfare comparisons between different contracts under Epstein and Zin (1989)'s recursive preferences.

6.5 Savings and Welfare

Our main welfare calculations assume that individuals cannot save. This assumption may substantially underestimate the welfare under short-term contracts, and under the GLTHI. As noted above, the GLTHI contracts result in a consumption profile that closely tracks the hump-shaped life-cycle income profile. Moreover, under short-term contracts, individuals experience large premium shocks that could be smoothed with precautionary savings. Hence, this section allows for precautionary savings. We do so by solving a dynamic programming problem of optimal savings with mortality risk as in Yaari (1965). Individuals solve the following maximization problem:

$$
\begin{aligned}
\max_{c_t} \quad & \mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^t u(c_t)\right) \\
\text{s.t.} \quad & a_{t_0} = 0 \\
& a_t \geq 0 \;\; \forall t \\
& a_{t+1} = (1+r)a_t + y_t - c_t - P(\Xi_t)
\end{aligned}
$$

where $P(\Xi_t)$ is the premium in period t as a function of an individual's medical history $\Xi_t \equiv (\xi_1, \xi_2, ..., \xi_t)$, and a_t is the level of assets.

Different contracts result in different mappings between an individual's medical history up to period t and an individual's premium in t. Under a series of short-term contracts, only an individual's current health status matters since $P(\Xi_t) = \mathbb{E}(m_t|\Xi_t) = \mathbb{E}(m_t|\xi_t)$. In contrast, for a GLTHI contract, the entire medical history matters. Due to guaranteed-renewability, $P(\Xi_t)$ is defined recursively: In the first period, $\boldsymbol{\Xi_1} = \xi_1$ and $P(\boldsymbol{\Xi_1}) = P_1(\xi_1)$, where Equation (1) defines $P_t(\xi_t)$. In any period $t > 1$, $P(\boldsymbol{\Xi_t}) = \min\{P(\boldsymbol{\Xi_{t-1}}), P_t(\xi_t)\}$.[42] (Note that, in this optimal consumption problem with savings, there is uncertainty regarding net income $y_t - P(\Xi_t)$ and mortality risk.[43])

For a given lifecycle income profile, the dynamic program provides an optimal consumption policy $C_t^*(\xi_t, a_t)$ where a_t is the level of assets carried into period t. The certainty equivalent (CE) of the dynamic problem is equal to:

$$u(C_{SAV}) = \frac{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0} u(C_t^*(\xi_t, a_t)\right)}{\mathbb{E}\left(\sum_{t=t_0}^{T} S_t \delta^{t-t_0}\right)} \tag{12}$$

Table 6: Welfare by Type of Contract with Savings

	CE_{GHHW}	$CE_{GLTHI,SAV}$	$CE_{ST,SAV}$
Ed 10	21,945	21,177	741
Ed 13	26,093	25,088	4,879

Source: German Claims Panel Data, SOEP data. The distribution of initial health states at age 25 used in this table corresponds to that in Panel (i) of Table 5. All consumption certainty equivalents (welfare) are in 2016 USD per capita, per year.

Table 6 shows the welfare results when allowing for savings, assuming $r = 1/\delta - 1$. Allowing for precautionary savings substantially improves welfare under the series of short-term contracts. Consider the distribution of initial health state $\Delta_0 = \frac{1}{100}[89.11, 10.25, 0.47, 0.11, 0.04, 0.03, 0]$ as in Panel (i) of Table 5. The consumption certainty equivalent increases from $CE_{ST} = -\$10,119$ to $CE_{ST,SAV} = \$741$ for *Ed 10* individuals, and from $CE_{ST} = -\$2,223$ to $CE_{ST,SAV} = \$4,879$ for *Ed 13* individuals. On the other hand, savings do not significantly improve welfare under GLTHI. Intuitively, the GLTHI contract already achieves substantial savings through highly front-loaded premiums. Moreover, as shown in Ghili et al. (2019), with the optimal contract, individuals have no incentives to engage in additional savings. Thus, introducing savings does not affect welfare under the optimal contract.

[42]The state variable in the dynamic program under GLTHI is the guaranteed-renewable premium; its law of motion is given by the probability of qualifying for a lower premium.

[43]Mortality risk implies that individuals may die with positive assets. Therefore, the expected net present value of consumption with optimal savings will be lower than the net present value of resources. Our calculations implicitly assume that individuals do not derive value from bequests.

6.6 Robustness: Risk Aversion, Epstein-Zin Preferences, and Income Profiles

In this section, we investigate the robustness of our main findings in three dimensions. First, we investigate whether our results are robust to the degree of risk aversion, i.e., the parameter γ in the CARA utility function specified by Equation (11). Second, we investigate whether our results are robust to Epstein and Zin (1989)'s recursive preferences where risk aversion and intertemporal elasticity of substitution are separately parameterized. Third, we use U.S. income profiles as an empirical input.

The Degree of Risk Aversion

Under our parametric assumptions on preferences, the GLTHI contracts entail a small welfare loss relative to the optimal dynamic contract. Almost entirely eliminating reclassification risk basically compensates the welfare loss from heavier frontloading in the GLTHI. Following the convention in the literature, our main results assume a level of risk-aversion of $\gamma = 4 \times 10^{-}4$ (cf. Ghili et al., 2019). With this level of risk aversion, an individual would be indifferent between (a) a gamble where she wins \$1,000 with a 50 percent chance and loses \$713 with a 50 percent chance and (b) no gamble, i.e., the status quo. This subsection investigates the robustness of our findings with respect to different levels of γ.

Figure 11 shows the results graphically, where the x-axis spans values of $\gamma \in [5 \times 10^{-}5, 8 \times 10^{-}4]$. For each γ, the y-axis shows the corresponding difference in certainty equivalents as a fraction of the welfare under the optimal contract (see dashed line). As seen, the difference is small when γ is either very low or very high. That is, our main qualitative finding—the simple GLTHI contract can basically achieve similar welfare as the optimal dynamic contract—is robust to the degree of risk aversion, γ.

To investigate the underlying reason for the robustness of the findings with respect to γ, the solid line plots the percentage point differences in welfare when we only focus on *differences in consumption* across the lifecycle. In other words, we eliminate the welfare differences that are due to differences in reclassification risk. As seen, we then find that the welfare gap between GLTHI and the optimal contract increases substantially in γ.[44]

In summary, varying the level of risk aversion affects the performance of GLTHI relative to the optimal contract via two underlying channels. The first is due to differences in lifecycle consumption, where GLTHI clearly falls short, even more so the larger γ; the second is due to differences in

[44] In practice, the line represents the *CE* of consumption after replacing the actual consumption under the optimal contract with the expected consumption at each age, thus eliminating the reclassification risk component of the optimal contract. By contrast, the reclassification risk component of GLTHI is negligible.

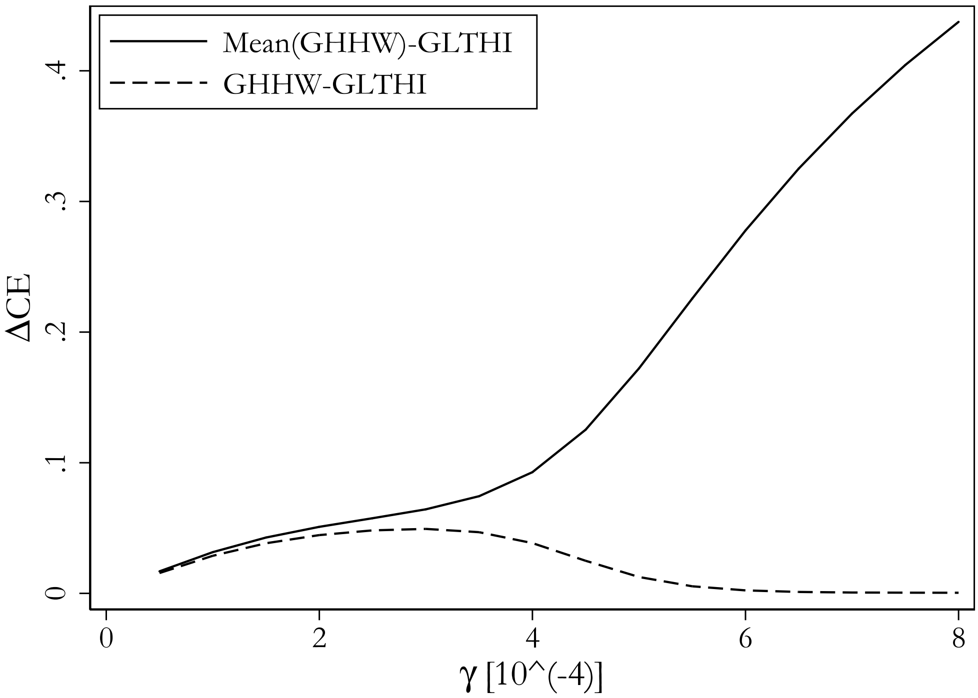


Figure 11: Difference in CE (GHLTI vs. GHHW) by Risk Aversion

Source: German Claims Panel Data, SOEP data. The x-axis shows the level of risk aversion γ. The y-axis shows differences in consumption certainty equivalents (CE) between GLTHI and the optimal contract as a fraction of total possible welfare, in other words, the welfare loss of GLTHI relative to GHHW. The dashed line shows total welfare differences, and the solid line shows only welfare differences due to differences in consumption.

reclassification risk, where GLTHI outperforms the optimal contract, and even more so the larger γ. As we vary γ, these two opposing forces almost completely cancel out.

When risk aversion is close to 0, the GLTHI contract coincides with the optimal dynamic contract. In the extreme case of risk neutrality, the volatility of premiums and the lifecycle shape of expected consumption are irrelevant. For low levels of γ, the lifecycle path of expected consumption is the most relevant factor determining the welfare performance of GLTHI. However, when γ becomes large enough, the elimination of reclassification risk operates in favor of GLTHI. Even though individuals with a large γ strongly prefer smoother consumption, they also dislike the higher associated reclassification risk.

The dashed curve in Figure 11 shows the total welfare gap between the GLTHI and the theoretically optimal contract. The maximal welfare difference between the two across all values of γ is 5 percent when $\gamma = 3 \times 10^-4$.[45]

[45]Under this level of risk aversion, an individual would be indifferent between (a) a gamble where she wins $1,000 with a 50 percent chance or loses $768 with a 50 percent chance, and (b) no gamble.

Epstein-Zin Recursive Preferences

So far, we assumed that a single parameter governs both risk aversion and the intertemporal elasticity of substitution. In this section, we investigate the robustness of our welfare findings when breaking the parametric link between risk aversion, γ, and the intertemporal elasticity of substitution, ψ. In particular, we now assume Epstein-Zin (EZ) preferences (Epstein and Zin, 1989). Preferences are defined recursively as:

$$V_t = F(c_t, R_t(V_{t+1})),$$

with

$$R_t(V_{t+1}) = G^{-1}(\mathbb{E}_t G(V_{t+1}))$$

As in Epstein and Zin (1989), will consider the CES aggregator

$$F(c,z) = \left((1-\delta)c^{1-1/\psi} + \delta z^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}}$$

We embed the same CARA specification used in our main analysis into the EZ preferences by assuming $G(c) = u(c) = \frac{1}{\gamma}e^{-\gamma c}$. In Appendix G, we show that the consumption certainty equivalent can be expressed as:

$$c = \left(\frac{\frac{\left(G^{-1}(\mathbb{E}_0(G(V_{t_0}(\xi_{t_0}))))\right)^{1-1/\psi}}{1-\delta}}{\sum_{j=t_0}^{T} \delta^{t-t_0} S_{t_0}^{j}} \right)^{\frac{1}{1-1/\psi}}. \tag{13}$$

where $\mathbb{E}_0()$ takes expectations with respect to the "birth" state, ξ_{t_0} and S_t^j is the survival probability from t to j.

For each contract, we compute $V_{t_0}(\xi_{t_0})$ numerically *via* backwards induction.

Varying γ and ψ, Figure 12 shows differences in certainty equivalents between GLTHI and the optimal contract. As seen, the welfare differences are small over all the entire range of parameter values. Notice that in Figure 12, with the risk aversion parameter $\gamma = 8E-4$, the GLTHI can even outperform the optimal contract when the intertemporal elasticity of substitution is relatively high. This can occur because the optimal contract in Ghili et al. (2019) is not necessarily the optimal contract under recursive preferences—recall that Ghili et al. (2019)'s theoretical characterization requires that preferences are *time separable*, which Epstein and Zin (1989)'s recursive preferences do not satisfy.

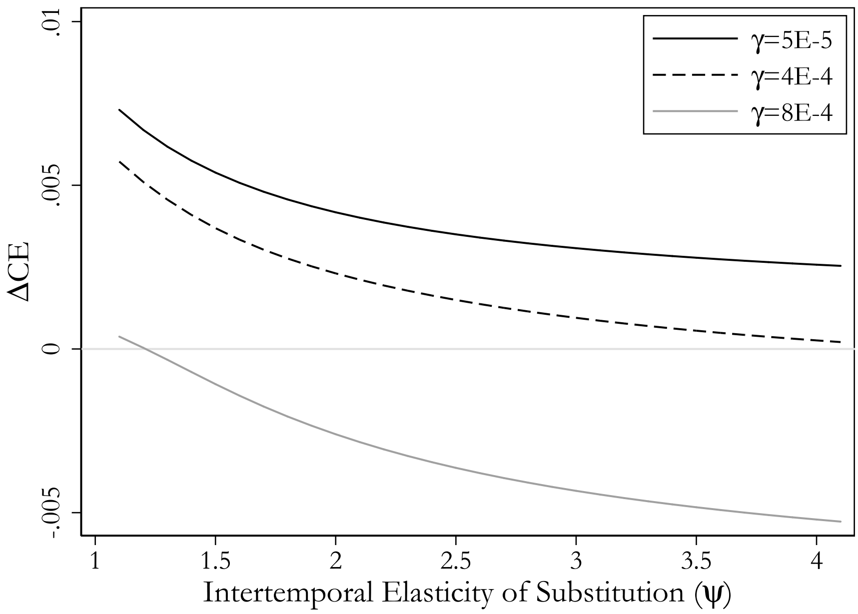


Figure 12: Difference in CE (GHLTI vs. GHHW) by Intertemporal Elasticity of Substitution

Source: German Claims Panel Data, SOEP data. The x-axis shows the level of intertemporal elasticity of substitution ψ. The y-axis shows differences in certainty equivalents (CE) between GLTHI and the optimal contract as a fraction of maximum possible welfare.

Income Profiles

Finally, to test the robustness of our results with respect to the income profile, we apply the lifecycle income pattern for the United States. To this end, we use the Cross National Equivalence Files (CNEF) of the Panel Study of Income Dynamics (PSID). The PSID is the oldest and longest-running panel survey in the world. It has been surveying U.S. families annually since 1968 and, since 1997, biannually (Panel Study of Income Dynamics, 2018). The CNEF harmonizes survey measures across countries and over time (Frick et al., 2007). Using the generated CNEF-PSID variables allows us to follow the exact same income concept (in 2016 USD) and implement the same estimation process than for Germany. That is, we exclude respondents under 25, focus on the years 1984 to 2015, and estimate Equation (10).

Figure C2 (Appendix) shows an increase in the post-tax equivalized income that is very close to the one observed in Germany between ages 25 and 60. However, the decrease in lifecycle income after age 60 is much steeper in the U.S., for both educational groups. Our calculations show that our main findings are also robust to U.S. income profiles: GLTHI contracts would achieve welfare that would only fall 5.8 and 3.5 percent short of the optimal long-term health insurance contract for

Americans with high school and college degrees, respectively.[46]

7 Implications for Reforms to the U.S. Health Insurance System

In this section, we discuss possible implications of our findings for the health care reform debate in the United States. The U.S. system is a mixture of public and private health insurance. Among the working age population below 65, about 60 percent have employer-sponsored health insurance (ESHI) and about 40 percent have either short-term private health insurance or are uninsured (Claxton et al., 2017); Medicare covers people above 65 (and the disabled), financed by payroll taxes. Of course, this system differs from the German health insurance system (see Section 2). ESHI is community-rated at the employee level and essentially long-term—provided that employers and employees do not separate—in which case it resembles the GLTHI.[47] Prior to the ACA, the U.S. individual private health insurance market closely resembled the individually risk-rated short-term contract as described in Section 3.2.[48] Thus, as a first order approximation, pre-ACA, the U.S. system was a mixture of 60 percent GLTHI and 40 percent short-term contracts for workers up to age 65; followed by a Medicare pay-as-you-go system for those 65 and older.

The questions that we ask in this section are: If we were to reform the U.S. health insurance system and replace all private health insurance contracts with individual long-term contracts, followed by Medicare for those 65 and older, by how much could we possibly improve welfare? How would such a hybrid system compare with a system where individuals purchase lifelong long-term insurance until they die?

Let us first consider a public insurance program for people above the age of 65, financed by a proportional tax on income. Although this is a simplified version of the U.S. Medicare program, its structure captures the main effect of Medicare in the context of long-term contracts. The Medicare tax acts as an additional, front-loaded premium during working ages to fund free health insurance for all people above 65, regardless of their health status. Thus, for each education group $Ed \in \{Ed\ 10, Ed\ 13\}$ separately, we assume that the proportional Medicare payroll tax τ_e^* is collected from individuals in this education group. Further, we assume that it covers all health care expenses

[46]The detailed results are available from the authors upon request.

[47]In contrast to long-term contracts, ESHI is subject to the dynamic inefficiency in the incentives to invest in health (see (Fang and Gavazza, 2011)).

[48]However, post-ACA, individual private contracts are community rated—although the ACA still allows insurers to charge older people and smokers more—and thus differ from the short-term contracts described in our paper.

of their education risk pool during the Medicare period (age 65 and above), such that

$$\tau^*_{Ed}\mathbb{E}\left(\sum_{t=25}^{64} S_t\delta^{t-24}y_t|Ed\right) = \mathbb{E}\left(\sum_{t=65}^{94} S_t\delta^{t-24}m_t\right) \tag{14}$$

where S_t is an indicator of survival until period t, y_t is income, m_t medical spending, and δ is the discount rate. In conducting this exercise separately for *Ed 10* and *Ed 13*, we do not allow for cross-subsidization and redistribution between high and low-income earners. By doing so, we can compare the hybrid system to our baseline scenario for the same net present value of resources. Consequently, all welfare consequences are due to intertemporal substitution and reclassification risk, and not due to transfers across individuals of different income levels. To evaluate welfare under the hybrid system, we separately compute a new set of GLTHI premiums, and the consumption guarantees under the optimal contract, *assuming that the terminal period is* $T = 64$.[49]

The consumption certainty equivalent is the constant consumption level that provides the same lifetime utility as those achieved under the hybrid system. Panel (a) of Table 7 shows the welfare results under the hybrid system, separately for *Ed 10* and *Ed 13* lifecycle income profiles. Panel (b) of Table 7 replicates the baseline results without Medicare (and thus the corresponding contracts apply over the entire lifecycle). For illustration purposes, the distribution of initial health states used in the calculations is that of Panel (h) in Table 5, namely, $\Delta_0 = \frac{1}{100}[89.11, 10.25, 0.47, 0.11, 0.04, 0.03, 0]$.

Table 7: Welfare of a Hybrid System of Private Contracts plus "Medicare-Like" Public Insurance

	Ed 10	*Ed 13*
Panel (a): Private Contracts up to 64 + Medicare from 65		
Payroll Tax Rate (%)	4.36	3.12
CE_{GLTHI}	20,349	24,297
CE_{GHHW}	20,740	24,907
CE_{ST}	-11,079	-3,444
Panel (b): Life-Long Private Contracts		
CE_{GLTHI}	21,168	25,088
CE_{GHHW}	21,945	26,093
CE_{ST}	-10,119	-2,223

Source: German Claims Panel Data, SOEP data. The distribution of initial health states used in the calculations is the same as that of Panel (h) in Table 5, namely, $\Delta_0 = \frac{1}{100}[89.11, 10.25, 0.47, 0.11, 0.04, 0.03, 0]$. All consumption certainty equivalents (welfare) are in 2016 USD per capita, per year.

Interestingly, theoretically it is *ambiguous* whether the hybrid system or the private system achieves

[49]For GLTHI, the Medicare payroll tax rates τ^*_{Ed} do not impact the calculation of the equilibrium premiums when $T = 64$ (see Equation (1)). The optimal premiums, however, depend on the income paths (see Equation (15)); we assume that incomes of individuals in education group *Ed 10* and *Ed 13* are taxed at the respective rate τ^*_{Ed} calculated by Equation (14).

higher welfare. The reason is that Medicare is a mandatory public system, and as such, it does not suffer from the one-sided commitment problem that the GLTHI contract needs to address.

Comparing Panels (a) and (b) reveals that welfare under the hybrid system is always lower than under the baseline scenario with lifetime contracts. The reason is as follows: Compared to the optimal contract, the Medicare program reduces consumption at earlier ages (because of the payroll tax), with no substantial changes in the reclassification risk. As seen in Figure 9, the optimal contract involves virtually no reclassification risk after age 65. For similar reasons, the Medicare program does not improve welfare when combined with the GLTHI contract. GLTHI has already *too much* front-loading and *too little* reclassification risk relative to the optimal.

What is more surprising is that the hybrid system also achieves a lower welfare when the private insurance is in the form of short-term contracts (CE_{ST} in Panel (a) vs. in Panel (b)). Because the Medicare provision in the hybrid system substantially decreases consumption volatility at old ages, in principle, introducing a Medicare-like program could increase welfare in an economy with short-term contracts. However, the Medicare tax *decreases* consumption at early ages, when the marginal utility of consumption is high. As Table 7 shows, the latter effect dominates for both income groups. In both cases, introducing Medicare is also welfare decreasing in an economy with short term contracts.

Robustness. The results in Table 7 assume that the Medicare payroll tax during working ages fully covers all medical expenses for the population above 65. In reality, however, Medicare Part B beneficiaries do pay a (subsidized) premium.[50] Premium-free Medicare coverage at old-age increases the tax rate needed to fund the entire program. Therefore, the degree of front-loading increases further. Because our simplified version of Medicare imposes *too much* front-loading, it is instructive to investigate the effect of introducing a Medicare premium with a corresponding decrease in the tax rate. In Appendix H, we illustrate this trade-off between charging a higher Medicare payroll tax for future beneficiaries vs. a higher Medicare premium for current beneficiaries. In conclusion, we find that a higher premium for current beneficiaries increases welfare because it increases consumption at early ages. However, even a very high Medicare premium (such that the Medicare tax is close to zero), combined with either the optimal contract or the GLTHI contract, would *not* achieve the same level of welfare as the optimal contract.

We also test the robustness of the results in Table 7 by allowing for savings in the Medicare envi-

[50] In addition, Medicare Part A imposes substantial cost-sharing, from which we have abstracted throughout in the paper.

ronment. In this economy, individuals are offered the GLTHI premium profile up to age 65, and free Medicare coverage starting at age 65. Such an insurance structure creates incentives to save. As in Section 6.5, we calculate welfare under an optimal level of savings and find a certainty equivalent of $20,672 (*Ed 10*) and $24,656 (*Ed 13*) (detailed results available upon request). This level of welfare is higher than welfare without savings (see Table 7), but still lower than welfare under either a lifetime GLTHI contract or the optimal contract.

8 Conclusion

Pricing regulation in health insurance markets has to trade off reclassification risk, adverse selection, moral hazard as well as consumption smoothing over the lifecycle. Very few countries in the world have organized their health insurance based on private markets—e.g., the U.S., Chile, Switzerland and Germany. The U.S. and Switzerland have traditionally organized their markets as short-term annual contracts with tight community pricing regulation to provide reclassification risk insurance for all citizens. A fundamental alternative is private individual *long-term* health insurance. This paper shows that long-term contracts have the power to leverage individual's intertemporal lifecycle incentives to insure the reclassification risk. We present, discuss and evaluate the basic principles of such real-world market that has been largely overlooked as a fundamental alternative to community-rated short-term health insurance markets: the German individual private long-term health insurance market (GLTHI).

First, we present the basic principles of the market and derive its theoretical lifecycle premiums and welfare effects. We show that GLTHI almost fully eliminates reclassification risk over the lifecycle. However, the low reclassification risk comes at the expense of high premium front-loading resulting in limited intertemporal consumption smoothing. Second, we benchmark lifecycle welfare of the GLTHI contract against several alternative contracts. To that end, we use unique claims panel data of more than half a million GLTHI policyholders along with representative household panel data over more than three decades.

Overall, we find that GLTHI contracts generate substantial welfare gains relative to (risk-rated) short-term contracts. More importantly, we show that German-style long-term health insurance contracts can basically achieve the same welfare as the optimal dynamic contract derived in Ghili et al. (2019). GLTHI contracts almost entirely eliminate the lifecycle reclassification risk, which compensates for the welfare loss from more front-loading relative to the optimal contract. We also show that this finding is robust to alternative degrees of risk aversion and that, for very low and very high degrees of risk aversion, GLTHI welfare converges to the optimal contract. The findings are also robust to different degrees of intertemporal elasticities of substitution and Epstein-Zin preferences, as well as using lifecycle income profiles derived from representative U.S. survey data. The GLTHI contract provides large welfare gains relative to a series of risk-rated short-term contracts as common in the pre-ACA era in the United States. Moreover, we evaluate a combination of long-term contracts and a Medicare-like pay-as-you-go system for people above 65. Such a hybrid system would be superior to the *status quo*, but inferior to a system of long-term contracts over the entire lifecycle.

A practical advantage of the GLTHI contract relative to the theoretically optimal contract is that is does not use policyholder's income in premium setting which, to the extent that incomes are endogenous, avoids potential work disincentives. In addition, the theoretical optimization problem is independent of the exact curvature of citizens' utility function and risk preferences. Market regulation is relatively simple as witnessed by the fact that the GLTHI market has been stable and providing insurance for millions of people for decades. We believe that our findings and these institutional facts strengthen the case of the German long-term contract design as an appealing policy option. We hope that the findings in this paper will inject individual private long-term health insurance as a real-world alternative into the health policy debate, which has largely focused on incremental adjustments to the status quo or the transition to a "single-payer for all" system.

We finish by acknowledging two important and general caveats of long term contracts. First, our results show that neither the German design nor the optimal dynamic contract may be a desirable alternative for some population subgroups. In fact, long term contracts may be highly undesirable for people who are very sick in young ages. From a policy perspective, for those individuals, societies implementing long-term contracts must provide a public alternative—like the co-existing public insurance in the case of Germany.

Second, our theory abstracts from a couple of key features that may have relevant implications for welfare under long-term contracts. First, our model assumes time-consistent individuals. From the perspective of a present-biased consumer, front-loading may render the long-term contracts undesirable, particularly when front-loading is high.[51] In addition, our model abstracts from moral hazard. In the presence of moral hazard, using long-term contracts to minimize reclassification risk could induce inefficiencies in spending. Quantifying the role of moral hazard in long-term contracts

[51] Still, Gottlieb and Zhang (2019) show that with a sufficiently large number of periods, the inefficiencies arising from time inconsistency vanish. With the long-term contract that emerges in the equilibrium with time-inconsistent agents, time-inconsistent agents may achieve the same level of welfare than time-consistent agents.

is an important avenue for future research.

References

Aizawa, N. and H. Fang (2020). Equilibrium labor market search and health insurance reform. *Journal of Political Economy 128*(11), 4258–4336.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

Association of German Private Healthcare Insurers (2019a). *Rechenschaftsbericht der Privaten Krankenversicherung 2018-2019*. Verband der Privaten Krankenversicherung. https://www.pkv.de/service/broschueren/verband/rechenschaftsbericht-2018-2019/, retrieved on February 20, 2020.

Association of German Private Healthcare Insurers (2019b). *Zahlen und Fakten*. https://www.pkv.de/service/zahlen-und-fakten/, retrieved on February 6, 2020.

Association of German Private Healthcare Insurers (2019c). *Zahlenbericht der Privaten Krankenversicherung 2018*. Verband der Privaten Krankenversicherung. https://www.pkv.de/service/broschueren/daten-und-zahlen/zahlenbericht-2018.pdb.pdf, retrieved on February 20, 2020.

Association of German Private Healthcare Insurers (2020). *Zahlenbericht der Privaten Krankenversicherung 2019*. Verband der Privaten Krankenversicherung. https://www.pkv.de/fileadmin/user_upload/PKV/c_Verband/PDF/2020-12_PKV-Zahlenbericht_2019.pdf, retrieved on January 18, 2021.

Atal, J., H. Fang, M. Karlsson, and N. R. Ziebarth (2019). Exit, voice or loyalty? An investigation into mandated portability of front-loaded private health plans. *Journal of Risk and Insurance 86*(3), 697–727.

Atal, J. P. (2019). Lock-in in dynamic health insurance contracts: Evidence from Chile. PIER Working Paper Archive 19-020, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economics 11*(1), 685–725.

Baumann, F., V. Meier, and M. Werding (2008). Transferable ageing provisions in individual health insurance contracts. *German Economic Review 9*(3), 287–311.

Becker, G. S. and B. R. Chiswick (1966). Education and the distribution of earnings. *The American Economic Review 56*(1/2), 358–369.

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017). Life-cycle earnings, education premiums, and internal rates of return. *Journal of Labor Economics 35*(4), 993–1030.

Börsch-Supan, A. and H. Jürges (2012). Disability, pension reform, and early retirement in Germany. In *Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation*, NBER Chapters, pp. 277–300. National Bureau of Economic Research.

Bundesministerium für Gesundheit (2020). Anzahl der Krankenkassen. http://www.gbe-bund.de/, retrieved on February 6, 2020.

Bünnings, C., H. Schmitz, H. Tauchmann, and N. R. Ziebarth (2019). How health plan enrollees value prices relative to supplemental benefits and service quality. *Journal of Risk and Insurance 86*(2), 415–449.

Büser, W. (2012). Private Krankenversicherung: Der Weg zurück in die Kasse. *Deutsches Ärzteblatt 109*(4), A–162–A–162.

Cecu (2018). *Rückkehr aus der privaten in die gesetzliche Krankenversicherung Überblick zum PKV-GKV-Wechsel*. Portal für Finanzen und Versicherungen. https://www.cecu.de/private-krankenversicherung-rueckkehr-gesetzliche.html, last retrieved on October 28, 2018.

Christiansen, M., M. Eling, J.-P. Schmidt, and L. Zirkelbach (2016). Who is changing health insurance coverage? Empirical evidence on policyholder dynamics. *Journal of Risk and Insurance 83*(2), 269–300.

Claxton, G., L. Levitt, and K. Pollitz (2017). *Pre-ACA Market Practices Provide Lessons for ACA Replacement Approaches*. Kaiser Family Foundation. https://www.kff.org/health-costs/issue-brief/pre-aca-market-practices-provide-lessons-for-aca-replacement-approaches/, retrieved on September 14, 2018.

Cochrane, J. H. (1995). Time-consistent health insurance. *Journal of Political Economy 103*(3), 445–473.

Cole, H. L., S. Kim, and D. Krueger (2019). Analysing the effects of insuring health risks: On the trade-off between short-run insurance benefits versus long-run incentive costs. *Review of Economic Studies 86*(3), 1123–1169.

Deutsche Aktuarvereinigung (DAV) (2019). *Der aktuarielle Unternehmenszins in der privaten Krankenversicherung (AUZ): Richtlinie*. https://aktuar.de/unsere-themen/fachgrundsaetze-oeffentlich/2019-10-09_DAV-Richtlinie_AUZ.pdf, retrieved on February 6, 2010.

Diamond, P. (1992). Organizing the health insurance market. *Econometrica 60*(6), 1233–54.

Dustmann, C., P. A. Puhani, and U. Schönberg (2017). The long-term effects of early track choice. *The Economic Journal 127*(603), 1348–1380.

Eekhoff, J., M. Jankowski, and A. Zimmermann (2006). Risk-adjustment in long-term health insurance contracts in Germany. *The Geneva Papers on Risk and Insurance-Issues and Practice 31*(4), 692–704.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on moral hazard in health insurance. *American Economic Review 103*(1), 178–219.

Engels, B., J. Geyer, and P. Haan (2017). Pension incentives and early retirement. *Labour Economics 47*(C), 216–231.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Fang, H. and A. Gavazza (2011, December). Dynamic inefficiencies in an employment-based health insurance system: Theory and evidence. *American Economic Review 101*(7), 3047–77.

Fang, H. and E. Kung (2020). Life insurance and life settlements: The case for health-contingent cash surrender values. *Journal of Risk and Insurance 87*(1), 7–39.

Finger, R. J. (2001). Risk classification. In Casualty Actuarial Society (Ed.), *Foundations of Casualty Actuarial Science* (fourth ed.)., Chapter 6, pp. 287–341.

Finger, R. J. (2006). Risk classification. *Foundations of Casualty Actuarial Science*, 231–276.

Finkelstein, A., K. McGarry, and A. Sufi (2005). Dynamic inefficiencies in insurance markets: Evidence from long-term care insurance. *The American Economic Review 95*(2), 224–228.

French, E. and E. Kelly (2016). Medical spending around the developed world. *Fiscal Studies 37*(3-4), 327–344.

Frick, J. R., S. P. Jenkins, D. R. Lillard, O. Lipps, and M. Wooden (2007). The Cross-National Equivalent File (CNEF) and its Member Country Household Panel Studies. *Journal of Applied Social Science Studies (Schmollers Jahrbuch: Zeitschrift für Wirtschafts- und Sozialwissenschaften) 127*(4), 626–654.

German Statistical Office (2016). Sozialleistungen - Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus). Fachserie 13 reihe 1.1.

Geyer, J. and V. Steiner (2014). Future public pensions and changing employment patterns across birth cohorts. *Journal of Pension Economics and Finance 13*(02), 172–209.

Ghili, S., B. Handel, I. Hendel, and M. D. Whinston (2019). The welfare effects of long-term health insurance contracts. Working Paper 2218, Cowles Foundation Discussion Paper.

GKV-Gesundheitsreformgesetz (2000). Gesetz zur Reform der gesetzlichen Krankenversicherung ab dem Jahr 2000. https://www.bgbl.de/xaver/bgbl/start.xav?start=%2F%2F*[%40attr_id%3D%27bgbl199s2626.pdf], last retrieved on February 10, 2020.

Gotthold, K. and B. Gräber (2015). So kommen Sie zurück in die gesetzliche Kasse. https://www.welt.de/finanzen/verbraucher/article138148142/So-kommen-Sie-zurueck-in-die-gesetzliche-Kasse.html, last retrieved on October 18, 2018.

Gottlieb, D. and X. Zhang (2019). Long-term contracting with time-inconsistent agents. mimeo. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3176354, retrieved February 15, 2020.

Handel, B., I. Hendel, and M. D. Whinston (2015). Equilibria in health exchanges: Adverse selection vs. reclassification risk. *Econometrica 83*(4), 1261–1313.

Hendel, I. and A. Lizzeri (2003). The role of commitment in dynamic contracts: Evidence from life insurance. *The Quarterly Journal of Economics 118*(1), 299–328.

Herring, B. and M. V. Pauly (2006). Incentive-compatible guaranteed renewable health insurance premiums. *Journal of Health Economics 25*(3), 395–417.

Hofmann, A. and M. Browne (2013). One-sided commitment in dynamic insurance contracts: Evidence from private health insurance in Germany. *Journal of Risk and Uncertainty 46*(1), 81–112.

Hullegie, P. and T. J. Klein (2010). The effect of private health insurance on medical care utilization and self-assessed health in Germany. *Health Economics 19*(9), 1048–1062.

Innungskrankenkasse Berlin Brandenburg (2018). *IKK- oder privat versichert: Wichtige Infos zur Entscheidung*. https://www.ikkbb.de/fileadmin/user_upload/doc/Broschueren_fuer_Versicherte/IKKBBoderprivatversichertVers..pdf, last retrieved on October 30, 2018.

Kaiser Family Foundation (2019). *2019 Employer Health Benefits Survey*. https://www.kff.org/health-costs/report/2019-employer-health-benefits-survey/, retrieved on December 14, 2019.

Kluth, S. and M. Gasche (2016). Ersatzraten in der Gesetzlichen Rentenversicherung / Replacement Rates in the German Statutory Pension System. *Jahrbücher für Nationalökonomie und Statistik (Journal of Economics and Statistics) 235*(6), 553–583.

Koijen, R. S. J., T. J. Philipson, and H. Uhlig (2016). Financial Health Economics. *Econometrica 84*, 195–242.

Krankenkassen-Zentrale (KKZ) (2020). *BAP 2020: Neue Private Krankenversicherung – Beitragserhöhung und andere Beitragsanpassungen*. https://www.krankenkassenzentrale.de/wiki/beitragsanpassung#, retrieved on February 6, 2010.

Kriegel, H.-P., E. Schubert, and A. Zimek (2017). The (black) art of runtime evaluation: Are we comparing algorithms or implementations? *Knowledge and Information Systems 52*(2), 341–378.

Krueger, D. and H. Uhlig (2006). Competitive risk sharing contracts with one-sided commitment. *Journal of Monetary Economics 53*(7), 1661–1691.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory 28*(2), 129–137.

Nell, M. and S. Rosenbrock (2007). Die Diskussion über die Portabilität von risikogerechten Transferbeträgen in der Privaten Krankenversicherung. *Zeitschrift für die gesamte Versicherungswissenschaft 96*(1), 39–51.

Nell, M. and S. Rosenbrock (2008). Wettbewerb in kapitalgedeckten Krankenversicherungssystemen: Ein risikogerechter Ansatz zur Übertragung von Alterungsrückstellungen in der Privaten Krankenversicherung. *Perspektiven der Wirtschaftspolitik 9*(2), 173–195.

Nell, M. and S. Rosenbrock (2009). Ein Risikoausgleichsmodell für den Wettbewerb um Bestandskunden in der PKV. *Zeitschrift für die gesamte Versicherungswissenschaft 98*(4), 391.

Nuscheler, R. and T. Knaus (2005). Risk selection in the German public health insurance system. *Health Economics 14*(12), 1253–1271.

Panel Study of Income Dynamics (2018). *Public use dataset*. University of Michigan, Ann Arbor, MI. Produced and distributed by the.

Panthöfer, S. (2016). Risk selection under public health insurance with opt-out. *Health Economics 25*(9), 1163–1181.

Patel, V. and M. V. Pauly (2002). Guaranteed renewability and the problem of risk variation in individual health insurance markets. *Health Affairs Suppl. Web Exclusives*, W280–W289.

Pauly, M., A. Percy, and B. Herring (1999). Individual versus job-based health insurance: Weighing the pros and cons. *Health Affairs 18*(6), 28–44.

Pauly, M. V., H. Kunreuther, and R. Hirth (1995). Guaranteed renewability in insurance. *Journal of Risk and Uncertainty 10*(2), 143–156.

Pauly, M. V. and R. D. Lieberthal (2008). How risky is individual health insurance? *Health Affairs 27*(3), w242–w249.

Polyakova, M. (2016). Risk selection and heterogeneous preferences in health insurance markets with a public option. *Journal of Health Economics 49*, 153 – 168.

Schencking, F. (1999). *Private Krankenversicherung im Unterschied zur Gesetzlichen Krankenversicherung*, pp. 18–32. Wiesbaden: Deutscher Universitätsverlag.

Schmeiser, H., T. Störmer, and J. Wagner (2014). Unisex insurance pricing: Consumers' perception and market implications. In *The Geneva Papers on Risk and Insurance—Issues and Practice*, Volume 39, pp. 102–138.

Schmitz, H. and M. Westphal (2017). Informal care and long-term labor market outcomes. *Journal of Health Economics 56*(C), 1–18.

Schmitz, H. and N. R. Ziebarth (2017). Does framing prices affect the consumer price sensitivity of health plan choice? *The Journal of Human Resources 52*(1), 89–128.

SOEP (2018). *Data for years 1984-2016*. version 33, SOEPlong, 2018 doi:10.5684/soep.v33.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP): Scope, evolution, and enhancements. *Journal of Applied Social Science Studies (Schmollers Jahrbuch: Zeitschrift für Wirtschafts- und Sozialwissenschaften) 127*(1), 139–169.

Werbeck, A., A. Wüebker, and N. R. Ziebarth (2019). Cream skimming by health care providers and inequality in health care access: Evidence from a randomized field experiment. https://www.researchgate.net/publication/335890073_Cream_Skimming_by_Health_Care_Providers_and_Inequality_in_Health_Care_Access_Evidence_from_a_Randomized_Field_Experiment, retrieved February 20, 2020.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies 32*(2), 137–150.

Online Appendix: Not Intended for Publication

Appendix A

Switching from GLTHI to SHI

As mentioned in Section 2, the decision to enter the private market is essentially a "lifetime decision." The basic social insurance principle is: "Once private, always private[ly insured]." Below, we discuss the specific and very limited institutional exemptions for GLTHI enrollees to return to the public SHI system. We also provide empirical evidence on the switching rates.

First, for people above the age of 55, switching back to the public system is essentially impossible, even when their income decreases substantially or when they become unemployed. One of the few options for people above 55 would be to exit the labor force and enroll under the public family insurance of the spouse, if available. Rules for switching back to SHI have been very strict for older employees to avoid strategic switching to the private system when individuals are young and healthy and switching back to the public system when they are old, sick and have little income (and thus low income-dependent contribution rates).

Second, people below the age of 55 can only return to the public system if they become unemployed (and receive UI benefits), or if their gross wage from dependent employment permanently drops below the income threshold. However, assuming an average annual premium of €3,900 (as observed in our data), for an equally high SHI premium (15.5% of the gross wage), annual labor income would need to be as low as €25,000 which does not make sense from a stratical point of view for the overwhelming majority of cases. Moreover, switching to SHI entails loosing the entire old-age provisions which averaged about $33K per policyholder in 2018 (Association of German Private Healthcare Insurers, 2019c). Moreover, switching back to GLTHI in the future would imply risk reclassification.

Third, the self-employed below 55 can only switch to SHI if they give up their business and become an employee with a gross salary below the income threshold (see Social Code Book V, Para. 6 for details of the law, Büser, 2012; Cecu, 2018).

Official statistics show that the absolute number of people who switched from the private to the public system has been relatively stable at around 130,000 since the beginning of the 1990s, which corresponds to around 1.5 percent of the GLTHI market per year.[52] Figure A1 below uses SOEP

[52] As the total number of GLTHI enrollees has steadily increased over the last decades, this implies declining switching rates over time. Several reforms in the last decades are likely to be the cause of these declining switching rates over time: The *Gesundheitsreformgesetz* of December 20, 1988 substantially tightened the possibility of switching for pensioners; the *Gesundheitsstrukturgesetz*, passed on December 21, 1992, also likely affected switching between the systems as it intro-

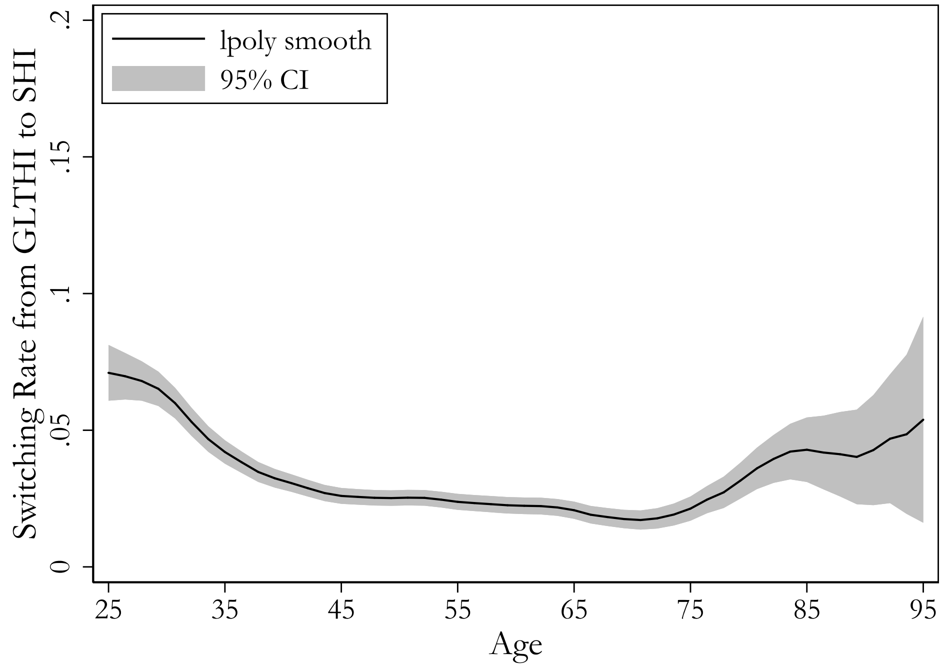


Figure A1: Likelihood to Return to SHI by Age

Source: SOEP (2018), the long version from 1984 to 2016. Epanechnikov kernel, degree 0, bandwidth 2.6.

data to plot switching rates by age. As seen, the likelihood to return to SHI decreases substantially between the age of 25 and 35. We conjecture that this is mostly because people who were privately insured as students enter the labor market and have to enroll in SHI if their gross salaries are below the income threshold. Switching rates remain stable at a low level between age 40 and age 75, and then slightly increase again. Using a fixed effects regression for the probability of switching to SHI among the universe of Germans who were at least once policyholder of a comprehensive private plan, we find very few significant determinants of switching back from the private to the public system. In particular, health care utilization measures (number of hospital nights and doctor visits) are not significant determinants and neither is the equivalized household income. The results of this analysis are available upon request.

Finally, we would like to point out that the historically grown institutional features of the German system induce advantageous selection into the GLTHI. The is almost the case by construction as private insurers have the right to deny coverage (or impose pre-existing condition clauses) to the sick. Hence, the sick basically remain publicly insured with SHI (Nuscheler and Knaus, 2005; Hullegie and Klein, 2010; Polyakova, 2016; Panthöfer, 2016).[53] While the main purpose of our paper is to

duced the free choice of SHI sickness funds, along with other provisions about the regulation of private insurers. Likely due to these and other reforms (e.g. the *GKV-Wettbewerbsstärkungsgesetz* of 2007), the rate as a share of all privately insured has declined in the last decades.

[53]When children of privately parents are also privately insured by their parents, under a family plan or a separately

Figure A2: Age Distribution of Initial Plan Inception

Source: German Claims Panel Data.

present, discuss and evaluate the basic principles of the GLTHI market, it is a real-word possibility for sick people to have a public option in Germany. Confirmed by the welfare analysis, it is clear that GLHI only maximizes welfare when people are relatively healthy at the time of their application. This insight has policy implications, which we discuss in Section 8. If other countries would design a market after the GLTHI and allow insurers to deny coverage (or impose guaranteed issue at all stages but allow risk-rated premiums), then a public option (either direct provision of insurance of premium subsidies) for those who are sick in young ages is necessary to avoid uninsurance and underinsurance. Note that the uninsurance rate in Germany is around 0.1 percent—in 2015, only 69 thousand individuals were without health insurance coverage (German Statistical Office, 2016).

private plan, parents have to pay premiums for each child. These are typically relatively modest as no old-age provisions are built for children. Moreover, if parents sign a private GLTHI contract for their child within two months after birth, risk rating is prohibited. In addition, some insurers offer a relatively unknown "option insurance" (*Optionstarif*) which is mostly sold in combination with supplemental (to SHI basic coverage) private hospital, dental, or travel insurance for which insurers carry out risk ratings. This initial risk rating then purchases the policyholder the option to purchase a GLTHI contract with that specific insurer without another risk rating within 6 to 10 years (and once one becomes eligible to opt out). No official numbers on the practical relevance of this option insurance are available. However, Google Trends yields zero hits, Google Scholar only 17 total hits, and a keyword search in the German *Handelsblatt* (similar to the *Wall Street Journal*) yields only one single hit for the *Optionstarif*, whereas it yielded 152 for the *Basistarif* which covers 0.4% of privately insured (see footnote 4). As a very last point, a more widespread and commonly known option is to put the existing GLTHI contract on hold for a monthly fee (*Anwartschaftsversicherung*), for example, when temporarily moving abroad. When returning to Germany, people with that option can simply re-activate their contract under the old conditions (§ 204 VVG).

Appendix B

Lifecycle Premiums in the Optimal Dynamic Health Insurance Contract (GHHW)

Ghili et al. (2019) study the optimal dynamic health insurance contract that maximizes consumer welfare, subject to break-even, no lapsation, and no borrowing constraints, in an environment where individuals have *time-separable* and *risk averse* preferences subject to stochastic health expenditure shocks. Ghili et al. (2019) show that the optimal dynamic insurance contract provides a *consumption guarantee* $\bar{c}_t(\xi_t, \mathbf{y}_t^T)$ that is a function of enrollees' current health risk and the vector of current and future income $\mathbf{y}_t^T \equiv \{y_t, y_{t+1}, ...y_T\}$. The individual will start consuming $\bar{c}_1(\xi_1, \mathbf{y}_1^T)$ and, over time, the individual's consumption guarantee $\bar{c}$ is bumped up in every period t such that a competing firm can offer a higher guarantee $\bar{c}_t(\xi_t, \mathbf{y}_t^T) > \bar{c}$ and still break-even in expectation.

Analogous to the GLTHI lifecycle premium calculation, $\bar{c}_t(\xi_t, \mathbf{y}_t^T)$ is solved by backwards induction. Specifically, the consumption guarantee in period T is given by $\bar{c}_t(\xi_t, y_T) = y_T - \mathbb{E}(m_T|\xi_T)$. For any $t < T$ and $\tau > t$, denote the set of future equilibrium consumption guarantees $\bar{\mathbf{c}}_{t+1}^{\tau} \equiv \{\bar{c}_{t+1}(.), ..., \bar{c}_{\tau}(.)\}$. Then an algebraic reformulation of the consumption guarantee in Ghili et al. (2019) shows that the equilibrium break-even consumption guarantee under the optimal dynamic contract for an individual purchasing a long-term optimal contract at time t under health status ξ_t is recursively determined by:

$$\bar{c}_t(\xi_t, \mathbf{y}_t^T) = \frac{y_t - \mathbb{E}(m_t|\xi_t) + \sum_{\tau>t}^{T} \sum_{z\in\Xi} \delta^{\tau-t}(y_\tau - \mathbb{E}(m_\tau|z)) \times q_\tau(z|\xi_t, \bar{\mathbf{c}}_{t+1}^{\tau}, \bar{c}_t(\xi_t, \mathbf{y}_t^T))}{1 + \sum_{\tau>t}^{T} \sum_{z\in\Xi} \delta^{\tau-t} \times q_\tau(z|\xi_t, \bar{\mathbf{c}}_{t+1}^{\tau}, \bar{c}_t(\xi_t, \mathbf{y}_t^T))}, \tag{15}$$

where $q_\tau(z|\xi_t, \bar{\mathbf{c}}_{t+1}^{\tau}, \bar{c}_t(\xi_t, \mathbf{y}_t^T))$ is, with some slight abuse of notation, the probability that (i) $\xi_\tau = z$, and (ii) the individual does not lapse (or die) between periods t and τ, given the set of future equilibrium consumption guarantees $\bar{\mathbf{c}}_{t+1}^{\tau}$. As with the GLTHI lifecycle premium, Equation (15) implicitly determines the equilibrium consumption guarantee in period t under health status ξ_t. As noted in Ghili et al. (2019), these consumption guarantees can be re-interpreted as a series of contracts with guaranteed premium *paths* $P_\tau(\xi_\tau, y_\tau) = y_\tau - \bar{c}_t(\xi_t, \mathbf{y}_t^T)$ for $\tau \geq t$; and the consumer would lapse at a time $\tau > t$ under health status ξ_τ whenever $\bar{c}_\tau(\xi_\tau, \mathbf{y}_\tau^T) > \bar{c}_t(\xi_t, \mathbf{y}_t^T)$. That is, a consumer who chose an optimal long-term contract at time t under health status ξ_t will lapse at a future time τ under health status ξ_τ if he/she is able to obtain a new long-term contract from the market that provides higher consumption guarantees.

Remark 3 *The consumption guarantees under GHHW's optimal long-term contracts, recursively characterized by Equation 15, do not depend on the utility function. What is important for the theoretical derivations of the optimal contract is that the consumers' preferences are time separable and exhibit risk aversion.*

Remark 4 *The consumption guarantees under GHHW's optimal long-term contracts, recursively characterized by Equation 15, do depend on income profiles. This implies that the corresponding guaranteed premium paths $P_\tau(\xi_\tau, y_\tau) = y_\tau - \bar{c}_t(\xi_t, \boldsymbol{y}_t^T)$ also depend on the income profiles. Since income profiles differ by education group, the GHHW premiums differ by education group. This differs from the GLTHI premiums (see Remark 2).*

The design of the GLTHI contract differs substantially from the welfare-maximizing GHHW contract, leading to different consumption profiles.[54] On the one hand, GLTHI implies the payment of a constant premium regardless of policyholders' income and the evolution of their health (with the exception of those who become healthy enough to switch to a contract with lower premiums; as shown later, this is a rare occurrence). As a consequence, the GLTHI contract almost completely eliminates the reclassification risk. However, the elimination of reclassification risk comes at the expense of large premium payments at early ages to prevent future premium hikes. These large upfront premiums have negative welfare implications when income is low and the marginal utility of consumption is high at early ages. On the other hand, the optimal dynamic contract involves a path of consumption guarantees (and therefore, a path of premiums) that is income-dependent, and that changes over the lifecycle after health shocks. The reason is that the optimal contract penalizes high premiums when the marginal utility of consumption is high.

[54]In the special case of flat income over the lifecycle, i.e., $y_t = y^0$ for all t, then $y^0 - P_t(\xi_t) = \bar{c}_t(\xi_t)$, cf. Equations (1) and (15). That is, when income is flat over the lifecycle, then the guaranteed premium in GLTHI coincides with the implicit guaranteed premium paths in GHHW.

Appendix C

Descriptive Statistics

Table C1: Summary Statistics: German Claims Panel Data

	Mean	SD	Min	Max	N
Socio-Demographics					
Age (in years)	45.5	11.4	25.0	99.0	1,867,465
Female	0.276	0.447	0.0	1.0	1,867,465
Policyholder since (years)	6.5	5.0	1.0	40.0	1,867,465
Client since (years)	12.8	11.0	1.0	86.0	1,867,465
Employee	0.336	0.473	0.0	1.0	1,867,465
Self-Employed	0.486	0.500	0.0	1.0	1,867,465
Civil Servant	0.132	0.338	0.0	1.0	1,867,465
Health Risk Penalty	0.358	0.480	0.0	1.0	1,867,465
Pre-Existing Condition Exempt	0.016	0.126	0.0	1.0	1,867,465
Health Plan Parameters					
TOP Plan	0.377	0.485	0.0	1.0	1,867,465
PLUS Plan	0.338	0.473	0.0	1.0	1,867,465
ECO Plan	0.285	0.451	0.0	1.0	1,867,465
Annual premium (USD)	4,749	2,157	0	33,037	1,867,318
Annual risk penalty (USD)	157	453	0	21,752	1,867,465
Deductible(USD)	675	659	0	3,224	1,867,465
Total Claims (USD)	3,289	8,577	0	2,345,126	1,867,465

Source: German Claims Panel Data. *Policyholder since* is the number of years since the client has enrolled in the current plan; *Client since* is the number of years since the client joined the company. *Employee* and *Self-Employed* are dummies for the policyholders' current occupation. *Health Risk Penalty* is a dummy that is one if the initial underwriting led to a health-related risk penalty on top of the factors age, gender, and type of plan; *Pre-Existing Conditions Exempt* is a dummy that is one if the initial underwriting led to exclusions of pre-existing conditions. The mutually exclusive dummies *TOP Plan*, *PLUS Plan* and *ECO Plan* capture the generosity of the plan. *Annual premium* is the annual premium, and *Annual Risk Penalty* is the amount of the health risk penalty charged. *Deductible* is the deductible and *Total Claims* the sum all claims in a calendar year. See Section 4.1 for further details.

Table C2: Summary Statistics: German Socio-Economic Panel Study

	Mean	SD	Min	Max	N
Socio-Demographics					
Female	0.5217	0.4995	0	1	530,228
Age	46.9119	17.4922	17	105	530,228
No degree yet	0.058	0.2338	0	1	530,228
Dropout of high school	0.0378	0.1908	0	1	530,228
Degree after 8/9 years of schooling (Ed 8)	0.3619	0.4805	0	1	530,228
Degree after 10 years of schooling (Ed 10)	0.2737	0.4459	0	1	530,228
Degree after 13 years of schooling (Ed 13)	0.1746	0.3796	0	1	530,228
Employment					
Civil servant	0.0393	0.1943	0	1	530,228
Self-employed	0.0624	0.2419	0	1	530,228
White collar	0.2736	0.4458	0	1	530,228
Full-time employed	0.4152	0.4928	0	1	530,228
Part-time employed	0.1402	0.3471	0	1	530,228
Income Measures in 2016 USD					
Monthly gross wage	2,940	2,506	0	215,093	310,460
Monthly net wage	1,921	1,527	0	134511.5	310,460
Individual annual total income	20,361	24,434	0	2,580,000	530,228
Equivalized post-tax post-transfer annual income	26,433	18,731	0	2,155,394	530,228
Insurance and Utilization					
Hospital nights in past calendar year	1.6652	8.3794	0	365	530,228
Doctor visits in past 3 months	2.4941	4.1436	0	99	461,971
Privately insured	1	0	1	1	57,558

Source: SOEP (2018), the long version from 1984 to 2016. Whenever the number of person-year observations is less than 530,228 the question was not asked in all years from 1984 to 2016. For example, *Doctor visits in past 3 months* has only been routinely asked since 1995. *Privately insured* indicates that 57,558/530,228=10.8% of all observations are by people who are insured on the GLTHI market. All income measures have been consistently generated and cleaned by the SOEP team; e.g., *Monthly gross wage* is labeled *labgro* and *Monthly net wage* is labeled *labnet* in SOEP (2018). See Section 4.2 for a detailed discussion of the variables.

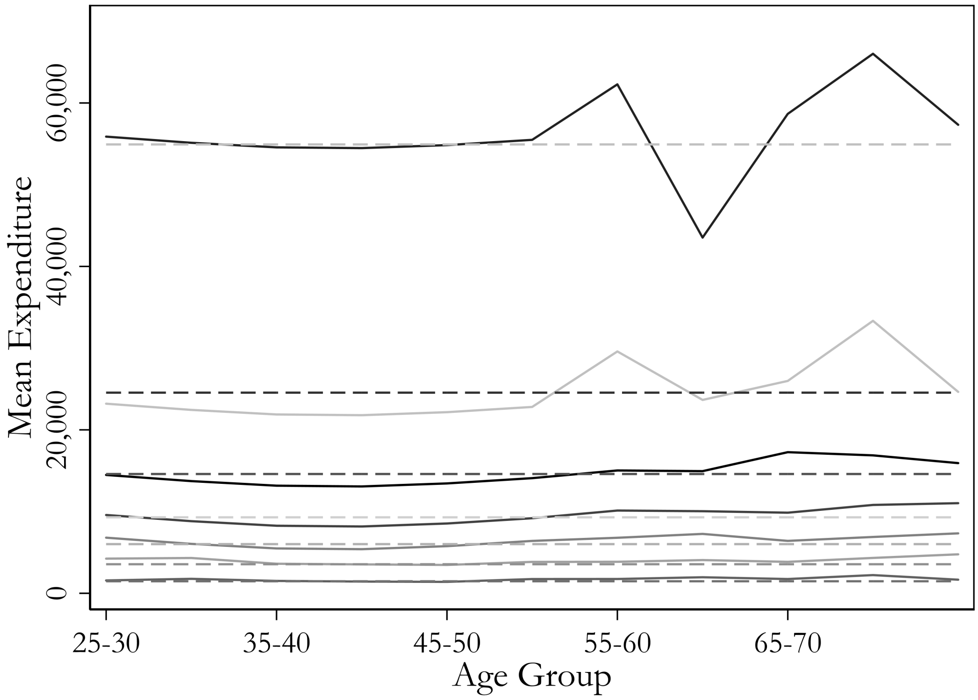


Figure C1: Predicted Health Expenditure

Note: Solid curves represent mean expenditure by age for each risk category λ_t, estimated according to Equation (9) in Section 5.2. The dashed lines represent the corresponding predictions assuming expenditure does not depend on age.

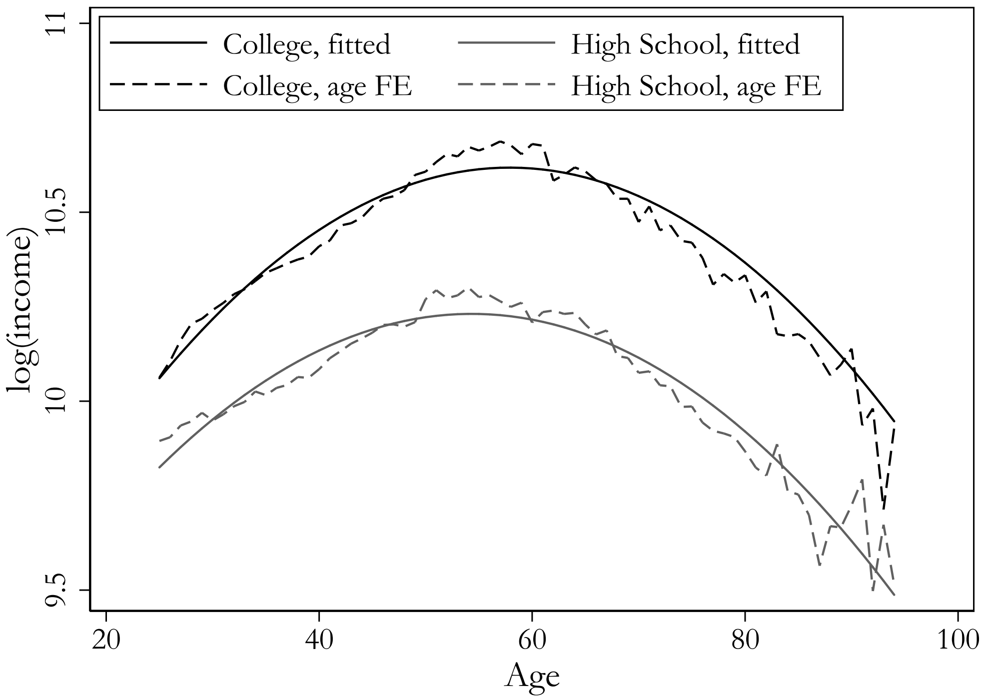


Figure C2: Lifecycle Income Paths for the United States, Nonparametric and Fitted.

Source: Panel Study of Income Dynamics (2018); Frick et al. (2007), years 1984 to 2015. All values in 2016 USD.

Appendix D

D1 Risk Classification: Robustness Checks

We expose the risk classification scheme derived in Section 5.1 to a number of robustness checks.

Winsorizing. First, we analyse the extent to which results are driven by outliers in m_{it}. It is of course desirable that outliers are considered in the classification, given their disproportionate contributions to means and variances; however, if the performance of the classification were widely different when they are not considered, it would cast doubt on how well the scheme performs with regard to less extreme risks. Therefore, we compared the performance of different classification schemes after the top percentile of expenditure had been been winsorized. Results are provided in Figure D1. As expected, the topcoding of outliers improves the predictive power of all schemes; however, their relative performance is unaffected by this change.

Lags of classes. Second, we compare two different ways of including a longer history of claims. Instead of expanding on the information set Λ_t before discretising, we consider an alternative based on $\Lambda_t^* = \lambda_t^*$ but where we consider the predictive power of the classification scheme interacted with its lags (i.e. a classification based on K^2 classes). Results are provided in Figure D2. It compares the two alternatives $q = 0$ and $q = 1$ from above, and in addition an interacted version, where the classification is based on $q = 0$ but this classification scheme is interacted with its lags in the regressions (leading effectively to K^2 classes). Clearly, this alternative has similar, actually even better, predictive power than $q = 1$. However, the variant with $q = 1$ thus achieves similar performance with a much smaller number of classes.

Sample selection. The results in Figure 4 are based on a sample of individuals who are observed over 4 years, since three lags are needed in Λ_{it}^*. In figure D3 we check how robust the finding is to varying the observation window required for sample selection. Sample 1 requires only that m_i and λ_t^* are observed, sample 2, also that λ_{t-1}^* is observed, and sample 3 in addition that λ_{t-2}^* is observed. The results provided in Figure D3 show that the predictive performance is sensitive to the sample used; however, the relative performance between schemes is the same regardless of the sample considered.

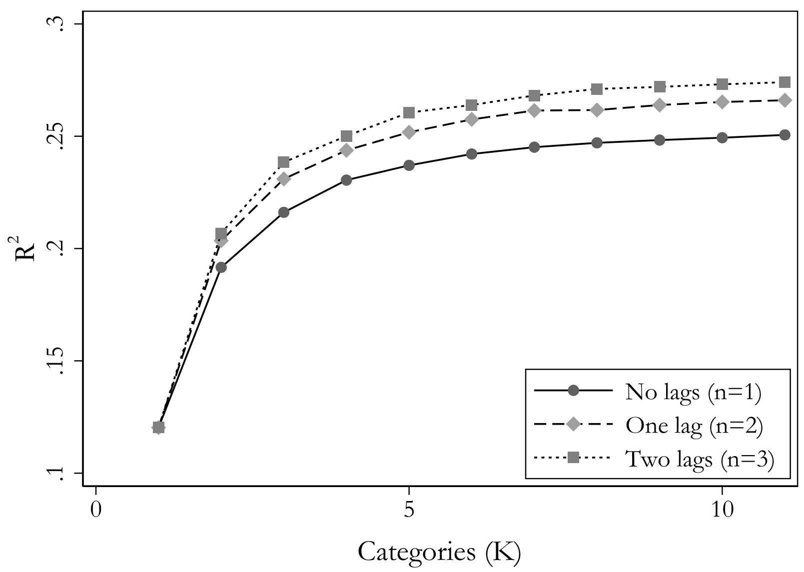


Figure D1: Performance of Alternative Risk Classifications: Winsorized Expenditure.

Note: Each specification includes 21 age times gender fixed effects, 5 year fixed effects and 79 plan fixed effects. Source: German Claims Panel Data.

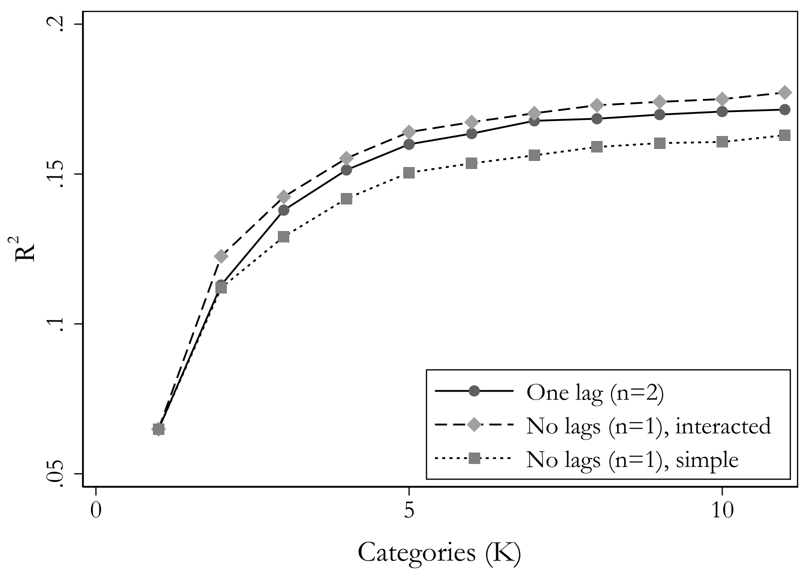


Figure D2: Performance of Alternative Risk Classifications: lags of classification.

Note: Each specification includes 21 age times gender fixed effects, year fixed effects and 79 plan fixed effects. Source: German Claims Panel Data.

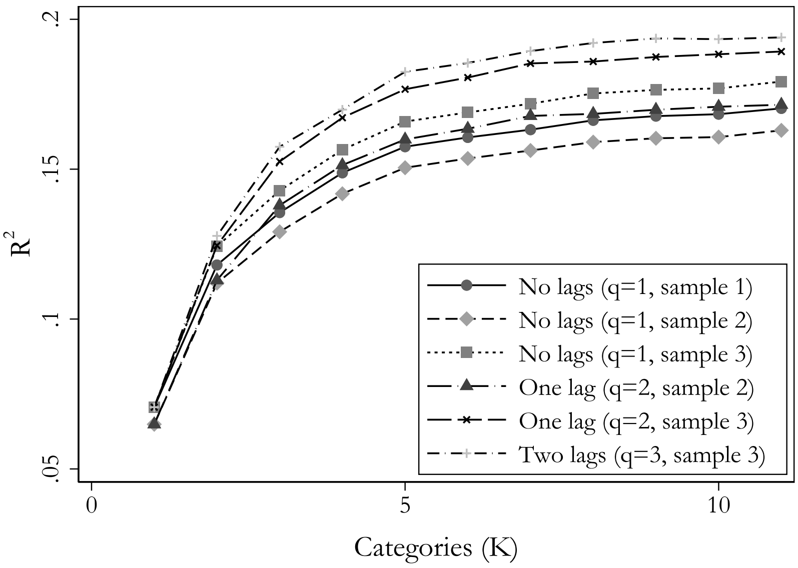


Figure D3: Performance of Alternative Risk Classifications: Different Samples.

Note: Each specification includes 21 age times gender fixed effects, year fixed effects and 79 plan fixed effects. Source: German Claims Panel Data.

Table D1: λ Risk Category Transitions by Age Group—Ages 25–54

		λ_{t+1}							
Age	λ_t	1	2	3	4	5	6	7	8 (†)
	1	0.8907	0.1024	0.0047	0.0011	0.0004	0.0003	0.0001	0.0004
	2	0.3197	0.4257	0.2020	0.0432	0.0077	0.0011	0.0003	0.0003
25-29	3	0.1242	0.2829	0.4104	0.1404	0.0378	0.0043	0.0000	0.0000
	4	0.0892	0.1688	0.2484	0.3917	0.0860	0.0159	0.0000	0.0000
	5	0.0938	0.1250	0.0625	0.3750	0.2917	0.0521	0.0000	0.0000
	6	0.0909	0.0000	0.0455	0.2273	0.3182	0.3182	0.0000	0.0000
	7	0.0000	0.0000	0.0002	0.0045	0.0240	0.1447	0.7619	0.0647
	1	0.8767	0.1145	0.0055	0.0018	0.0009	0.0002	0.0001	0.0003
	2	0.3212	0.4347	0.1909	0.0438	0.0080	0.0006	0.0001	0.0007
30-34	3	0.1241	0.3015	0.4080	0.1409	0.0229	0.0016	0.0000	0.0011
	4	0.1039	0.1640	0.2407	0.3739	0.1032	0.0115	0.0007	0.0021
	5	0.0734	0.0911	0.0506	0.2911	0.3747	0.1089	0.0025	0.0076
	6	0.0422	0.0438	0.0529	0.1678	0.3628	0.2450	0.0525	0.0329
	7	0.0128	0.0115	0.0083	0.0574	0.1545	0.1663	0.4524	0.1368
	1	0.8427	0.1480	0.0055	0.0022	0.0009	0.0002	0.0001	0.0004
	2	0.2798	0.4635	0.2113	0.0360	0.0076	0.0013	0.0000	0.0005
35-39	3	0.1177	0.2379	0.4850	0.1288	0.0275	0.0028	0.0001	0.0002
	4	0.0719	0.0967	0.3055	0.4085	0.0999	0.0158	0.0003	0.0014
	5	0.0743	0.0493	0.0691	0.3402	0.3629	0.0958	0.0039	0.0045
	6	0.0415	0.0331	0.0340	0.1180	0.2958	0.4009	0.0455	0.0312
	7	0.0127	0.0088	0.0054	0.0409	0.1276	0.2757	0.3975	0.1313
	1	0.8514	0.1392	0.0050	0.0024	0.0010	0.0003	0.0001	0.0006
	2	0.2862	0.4666	0.2050	0.0329	0.0075	0.0014	0.0001	0.0003
40-44	3	0.1137	0.2229	0.5134	0.1225	0.0241	0.0022	0.0001	0.0011
	4	0.0790	0.0769	0.2936	0.4213	0.1113	0.0157	0.0003	0.0018
	5	0.0640	0.0392	0.0759	0.3281	0.3763	0.1055	0.0038	0.0072
	6	0.0295	0.0382	0.0342	0.1605	0.2773	0.3613	0.0539	0.0450
	7	0.0081	0.0091	0.0049	0.0502	0.1079	0.2240	0.4247	0.1710
	1	0.8148	0.1736	0.0059	0.0028	0.0012	0.0006	0.0002	0.0009
	2	0.2267	0.5059	0.2229	0.0329	0.0093	0.0013	0.0001	0.0010
45-49	3	0.0653	0.2027	0.5708	0.1309	0.0258	0.0031	0.0001	0.0012
	4	0.0427	0.0712	0.2877	0.4655	0.1153	0.0140	0.0005	0.0029
	5	0.0303	0.0438	0.0475	0.3570	0.3964	0.1101	0.0058	0.0090
	6	0.0153	0.0266	0.0211	0.1118	0.2919	0.4163	0.0607	0.0563
	7	0.0038	0.0057	0.0027	0.0314	0.1021	0.2321	0.4298	0.1923
	1	0.8117	0.1740	0.0056	0.0035	0.0020	0.0008	0.0004	0.0020
	2	0.2283	0.4979	0.2228	0.0377	0.0101	0.0016	0.0002	0.0015
50-54	3	0.0602	0.1799	0.5727	0.1509	0.0317	0.0027	0.0001	0.0018
	4	0.0398	0.0648	0.2660	0.4930	0.1160	0.0155	0.0007	0.0041
	5	0.0274	0.0387	0.0426	0.3666	0.3866	0.1182	0.0075	0.0124
	6	0.0130	0.0222	0.0179	0.1084	0.2688	0.4220	0.0746	0.0732
	7	0.0028	0.0042	0.0020	0.0265	0.0819	0.2049	0.4600	0.2176

Source: German Claims Panel Data. Sample includes all years, 25-30 year old enrollees, and uses the ACG© score as λ.

Table D2: λ Risk Category Transitions by Age Group—Ages 55+

		λ_{t+1}							
Age	λ_t	1	2	3	4	5	6	7	8 (†)
	1	0.7261	0.2537	0.0101	0.0037	0.0020	0.0013	0.0004	0.0027
	2	0.0932	0.6432	0.2123	0.0357	0.0110	0.0018	0.0004	0.0025
55-59	3	0.0002	0.1739	0.6167	0.1690	0.0335	0.0044	0.0001	0.0024
	4	0.0001	0.0637	0.2426	0.5404	0.1287	0.0180	0.0007	0.0058
	5	0.0001	0.0356	0.0363	0.3758	0.4009	0.1282	0.0069	0.0163
	6	0.0000	0.0195	0.0145	0.1061	0.2662	0.4370	0.0650	0.0917
	7	0.0000	0.0037	0.0016	0.0260	0.0813	0.2126	0.4016	0.2732
	1	0.7558	0.2147	0.0145	0.0044	0.0042	0.0019	0.0011	0.0033
	2	0.1023	0.6414	0.1981	0.0387	0.0120	0.0031	0.0004	0.0040
60-64	3	0.0002	0.1612	0.6076	0.1836	0.0394	0.0053	0.0001	0.0028
	4	0.0001	0.0555	0.2243	0.5507	0.1419	0.0204	0.0008	0.0063
	5	0.0001	0.0292	0.0317	0.3610	0.4168	0.1370	0.0075	0.0168
	6	0.0000	0.0153	0.0122	0.0980	0.2660	0.4489	0.0686	0.0910
	7	0.0000	0.0028	0.0013	0.0235	0.0794	0.2136	0.4143	0.2651
	1	0.3707	0.5949	0.0172	0.0076	0.0030	0.0015	0.0009	0.0042
	2	0.0624	0.6492	0.2407	0.0352	0.0065	0.0012	0.0004	0.0045
65-69	3	0.0008	0.1058	0.6561	0.2082	0.0223	0.0013	0.0000	0.0056
	4	0.0002	0.0335	0.2013	0.6242	0.1261	0.0052	0.0005	0.0090
	5	0.0000	0.0128	0.0159	0.3546	0.4985	0.0763	0.0019	0.0400
	6	0.0000	0.0000	0.0107	0.0551	0.4067	0.3517	0.0195	0.1563
	7	0.0006	0.0066	0.0029	0.0264	0.0553	0.1690	0.5289	0.2103
	1	0.3848	0.5793	0.0225	0.0060	0.0011	0.0003	0.0014	0.0048
	2	0.0070	0.6771	0.2554	0.0406	0.0105	0.0012	0.0000	0.0082
70-74	3	0.0001	0.0810	0.6277	0.2599	0.0230	0.0014	0.0001	0.0068
	4	0.0002	0.0115	0.1625	0.6579	0.1404	0.0080	0.0002	0.0195
	5	0.0000	0.0015	0.0184	0.2829	0.5654	0.0736	0.0010	0.0572
	6	0.0000	0.0000	0.0000	0.0327	0.3039	0.4052	0.0065	0.2516
	7	0.0005	0.0056	0.0033	0.0184	0.0172	0.0263	0.7192	0.2094
	1	0.1770	0.5900	0.0442	0.0995	0.0598	0.0063	0.0083	0.0150
	2	0.0006	0.6237	0.2903	0.0471	0.0094	0.0012	0.0000	0.0277
75+	3	0.0000	0.0525	0.5876	0.2988	0.0254	0.0012	0.0000	0.0344
	4	0.0000	0.0029	0.1012	0.6668	0.1623	0.0055	0.0008	0.0605
	5	0.0000	0.0000	0.0060	0.2262	0.5581	0.0837	0.0028	0.1232
	6	0.0000	0.0000	0.0019	0.0206	0.3127	0.4064	0.0225	0.2360
	7	0.0000	0.0000	0.0000	0.0000	0.1111	0.1481	0.4630	0.2778

Source: German Claims Panel Data. Sample includes all years, 25-30 year old enrollees, and uses the ACG© score as λ.

D2 Sample Selection: Robustness Checks

This robustness section focuses on plans with low deductibles. We consider a stricter sample selection rule, where we only include plans with deductibles below $400.[55] These plans have approximately full coverage and thus more reliable information on the universe of health care expenditures. Summary statistics for this subsample are provided in Table D3. A comparison with the numbers in Table C1 makes clear that the two samples are very similar in terms of age, gender and history with the company. On the other hand, the restricted sample has a greater share of employees and civil servants, but a smaller share of self-employed. The plan characteristics are also similar to a great extent—with the obvious exceptions of deductible size and average claims.

Table D3: Summary Statistics: Low-Deductible Plans

	Mean	SD	Min	Max	N
Socio-Demographics					
Age (in years)	44.8	11.8	25.0	99.0	879,468
Female	0.256	0.437	0.0	1.0	879,468
Policyholder since (years)	7.7	5.3	1.0	40.0	879,468
Client since (years)	13.9	11.7	1.0	84.0	879,468
Employee	0.414	0.493	0.0	1.0	879,468
Self-Employed	0.281	0.449	0.0	1.0	879,468
Civil Servant	0.280	0.449	0.0	1.0	879,468
Health Risk Penalty	0.338	0.473	0.0	1.0	879,468
Pre-Existing Condition Exempt	0.015	0.121	0.0	1.0	879,468
Health Plan Parameters					
TOP Plan	0.342	0.475	0.0	1.0	879,468
PLUS Plan	0.397	0.489	0.0	1.0	879,468
ECO Plan	0.261	0.439	0.0	1.0	879,468
Annual premium (USD)	5,208	2,005	0	33,037	879,374
Annual risk penalty (USD)	133	347	0	21,214	879,468
Deductible(USD)	154	164	0	395	879,468
Total Claims (USD)	3,868	9,064	0	2,345,126	879,468

Source: German Claims Panel Data. *Policyholder since* is the number of years since the client has enrolled in the current plan; *Client since* is the number of years since the client joined the company. *Employee* and *Self-Employed* are dummies for the policyholders' current occupation. *Health Risk Penalty* is a dummy that is one if the initial underwriting led to a health-related risk add-on premium on top of the factors age, gender, and plan; *Pre-Existing Conditions Exempt* is a dummy which equals one if the initial underwriting led to a coverage exclusion of services for some conditions. The mutually exclusive dummies *TOP Plan*, *PLUS Plan* and *ECO Plan* capture the generosity of the plan. *Annual premium* is the annual premium, and *Annual Risk Penalty* is the amount of the health risk penalty charged. *Deductible* is the deductible and *Total Claims* the sum all claims in a calendar year. See Section 4.1 for further details.

[55] This is the lowest cutoff for the deductible which gives us a sufficient number of observations to analyze health risk transitions within each age group.

Figure D4 compares the distributions of λ^* in the two samples. As expected, the zero-deductible plans have higher ACG© scores in general.

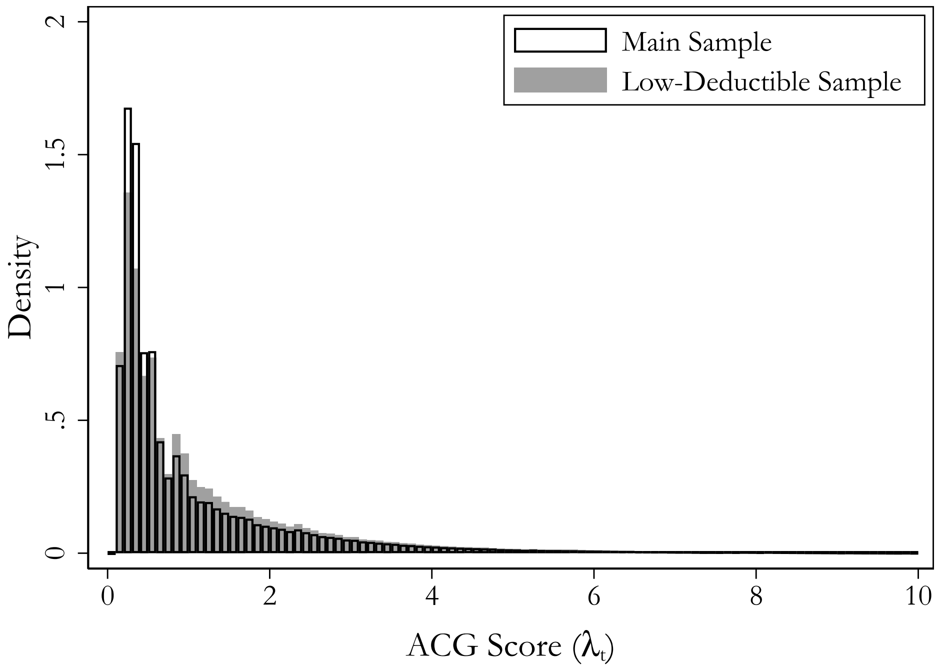


Figure D4: Distribution of λ^* for Main Sample vs. Low-Deductible Plans.

Table D4 shows how clients distribute over different risk categories by age in the low-deductible sample. A comparison with Table 2 confirms that the individuals in the low-deductible sample are in slightly worse health.

Table D4: Health Risk Categories λ by Age Group: Low-Deductible Sample

Age	1 (Healthiest)	2	3	4	5	6	7 (Sickest)
25-30	0.739	0.190	0.049	0.016	0.006	0.001	0.000
30-35	0.672	0.225	0.069	0.025	0.007	0.002	0.000
35-40	0.559	0.282	0.112	0.034	0.011	0.003	0.000
40-45	0.507	0.291	0.141	0.043	0.015	0.003	0.000
45-50	0.406	0.317	0.190	0.060	0.021	0.005	0.001
50-55	0.316	0.311	0.244	0.090	0.030	0.008	0.001
55-60	0.172	0.309	0.320	0.139	0.045	0.013	0.002
60-65	0.093	0.263	0.361	0.190	0.069	0.022	0.003
65-70	0.038	0.200	0.423	0.252	0.072	0.014	0.002
70-75	0.011	0.131	0.403	0.333	0.107	0.015	0.001
75+	0.000	0.055	0.286	0.453	0.179	0.024	0.003

Source: German Claims Panel Data. Sample includes all age groups and uses the ACG© score for the classification.

Table D5 shows the transition probabilities between different health states in the low-deductible sample. The probabilities are very similar to those reported in Table 3.

Table D5: Health Risk Category Transitions: Low-Deductible Sample

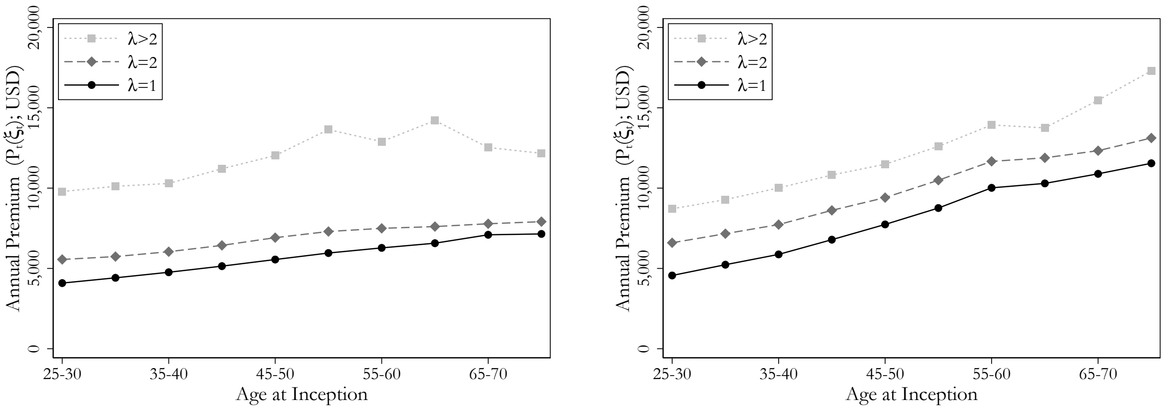
	λ_{t+1}							
λ_t	1	2	3	4	5	6	7	8 (†)
1	0.797	0.192	0.007	0.002	0.001	0.000	0.000	0.001
2	0.186	0.536	0.234	0.033	0.008	0.001	0.000	0.001
3	0.038	0.167	0.602	0.160	0.026	0.003	0.000	0.003
4	0.015	0.041	0.237	0.555	0.126	0.012	0.000	0.014
5	0.014	0.018	0.034	0.339	0.453	0.103	0.004	0.035
6	0.007	0.012	0.016	0.104	0.311	0.401	0.051	0.097
7	0.000	0.000	0.003	0.028	0.113	0.228	0.423	0.204

Source: German Claims Panel Data. Sample includes all years, all age groups, and uses the ACG© score for the classification.

Appendix E

German LTHI Premium Profiles

Figure E1 compares the (a) calibrated and (b) observed premium profiles for individuals entering their plan at different ages. In both figures, the highest category ($\lambda_t > 2$) is a weighted average calculated according to the actual distribution of λ_t in the different age groups.



(a) Calibrated

(b) Observed

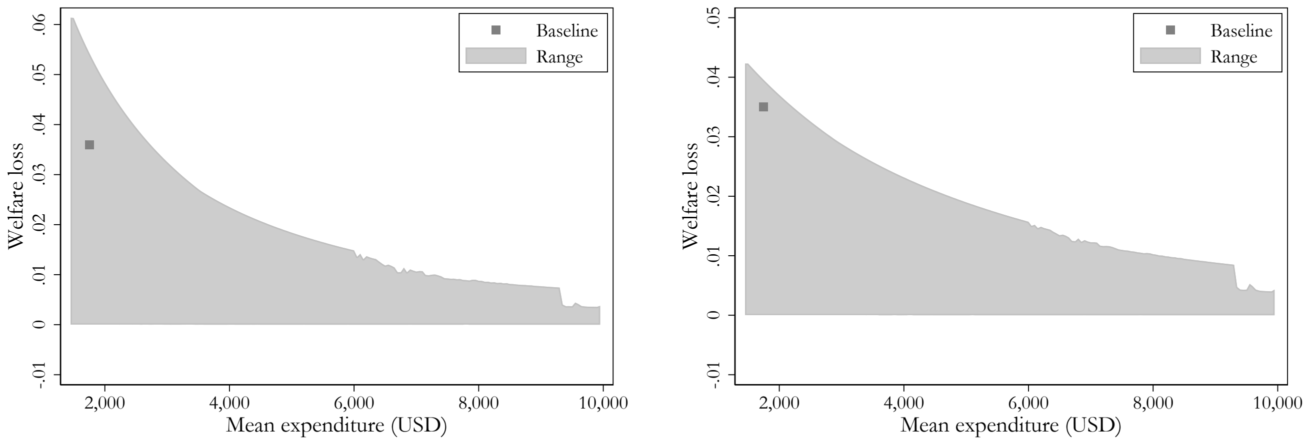
Figure E1: Calibrated vs. Actual Starting Premiums $P_t(\xi_t)$ by Age at Inception

Source: German Claims Panel Data. In Figure E1 (b), the sample includes all years and all health plans, and clients who have been in their contract for 2 to 5 years. We adjusted premiums for the three benefit categories *TOP, PLUS, ECO* and deductible size.

Appendix F

Welfare Results for Different Distribution of Starting States

In order to asses the robustness of results in Table 5 to varying assumptions regarding the distribution of starting states. For this test, we sampled 20 million probability simplices $\Delta_0 \in \Delta^7$ from a Dirichlet distribution with concentration parameters equal to the baseline probabilities coming out of the risk classification procedure. The resulting distribution contains probability simplicies with average health and expected costs quite different from the one that we consider in out baseline scenario in Table 5. For each draw, we calculate certainty equivalents for the various contracts. In Figure F1, we plot the resulting welfare loss (compared with the optimal contract) in relation to the average expenditure associated with each draw. The point "Baseline" corresponds to our baseline estimate in Table 5.



(a) *Ed 13* (b) *Ed 10*

Figure F1: Sensitivity Analysis: Distribution of Starting States.

Note: The figures show maxima and minima of GLHTI welfare losses within increments of $50 of expected expenditure. The underlying distribution is based on 20 million draws from a Dirichlet distribution. For Ed 13, 13 draws were discarded due to GHHW having a CE in a neighborhood of zero; for Ed 10, 15 draws were discarded for the same reason.

According to Figure F1, the welfare loss is bounded above at about 6% welfare loss (4% for the less-educated group). The maximum welfare loss is decreasing in expected expenditure, and the relatively healthy population we consider is in fact relatively close to the maximum.

Appendix G

Certainty Equivalent with CARA-EZ Preferences

We provide the derivation for the formula of the certainty consumption equivalent for Epstein-Zin preferences, provided in Equation (13). Preferences are defined recursively as

$$V_t = F(c_t, R_t(V_{t+1})),$$

with $R_t(V_{t+1}) = G^{-1}(\mathbb{E}_t G(V_{t+1}))$. As mentioned in the main text, we use the CES aggregator for $F(c,z) = \left((1-\delta)c^{1-1/\psi} + \delta z^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}}$, and incorporate the CARA utility function as $G(c) = u(c) = \frac{1}{\gamma}e^{-\gamma c}$.

Throughout we have assumed that utility is zero if the individual is dead. We can re-interpret V_t as the value of being alive in period t. Under that interpretation, one can write preferences recursively as:

$$V_t = \left((1-\delta)\, c_t^{1-1/\psi} + s_t \delta R_t\,(V_{t+1})^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}} \tag{G1}$$

where s_t is the probability of survival between t and $t+1$.

We now derive an expression for the certainty equivalent consumption c for any given value V_t under recursive preferences. Consider the situation in which consumption (while alive) is constant and equal to c. This means that $R_t\,(V_{t+1}) = V_{t+1}$, and therefore we can re-write

$$V_t = \left((1-\delta)\, c^{1-1/\psi} + s_t \delta\,(V_{t+1})^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}} \tag{G2}$$

Replacing the V_{t+1} in Equation (G2) as a function of V_{t+2} yields

$$\begin{aligned} V_t &= \left((1-\delta)\, c^{1-1/\psi} + s_t\delta\left((1-\delta)\, c^{1-1/\psi} + \delta s_{t+1}\,(V_{t+2})^{1-1/\psi}\right)\right)^{\frac{1}{1-1/\psi}} \\ &= \left((1-\delta)\, c^{1-1/\psi} + s_t\delta\,(1-\delta)\, c^{1-1/\psi} + s_t s_{t+1}\delta^2 V_{t+1}^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}} \end{aligned}$$

Iterating forward we can show that

$$\frac{V_t^{1-1/\psi}}{1-\delta} = \sum_{j=t}^{T} c^{1-1/\psi}\delta^{j-t} S_t^j$$

where $S_t^j \equiv \Pi_{k=t}^{j} s_k$ is the survival probability from t to j. Solving for c, we get an expression defining the certainty equivalent:

$$c = \left(\frac{\frac{V_t^{1-1/\psi}}{1-\delta}}{\sum_{j=t}^{T} \delta^{j-t} S_t^j} \right)^{\frac{1}{1-1/\psi}} \tag{G3}$$

Equation (G3) provides the certainty equivalent consumption to a program that provides value V_t.

We are interested in the certainty equivalent taking into account the uncertainty regarding the "birth state" ξ_{t_0}. Denote the value of this lottery V_b. It can be expressed as a function of V_{t_0} (the value at age 25):

$$V_b = G^{-1}(\mathbb{E}_0(G(V_{t_0}(\xi_{t_0})))) \tag{G4}$$

where $\mathbb{E}_0()$ takes expectations with respect to the uncertain "birth" state, ξ_{t_0}.

For each contract, we can compute the value $V_{t_0}(\xi_{t_0})$, for each state ξ_{t_0}, *via* backwards induction. Plugging Equation (G4) into Equation (G3), applied to the initial period t_0 we get the expression in the text.

Appendix H

Trading Off the Medicare Payroll Tax and Medicare Premiums

In this section, we evaluate the welfare consequence of changing the timing of payments into Medicare. Our baseline scenario assumes that Medicare coverage is completely free without any premium. However, the actual Medicare program in the US entails a premium (Part B) and cost-sharing provisions (Part A and B). In the context of our lifecycle model, premiums and cost-sharing provisions backload Medicare expenses by reducing the Medicare tax rate required to fund Medicare.

As a first approach, we maintain the assumption of no cost-sharing, but vary the level of premiums charged during retirement. Specifically, we assume a Medicare premium p has to be paid, starting at age 65. The associated Medicare tax rate $\tau(p)$ is such that the revenue neutrality condition holds

$$\tau(p)\,\mathbb{E}\left(\sum_{25}^{64} S_t \delta^{t-24} y_t\right) = \mathbb{E}\left(\sum_{65}^{94} S_t \delta^{t-24}(m_t - p)\right)$$

It is clear from this equation that a higher premium at old age is compensated by a lower tax rate at younger ages. Figure H1 shows this trade-off, where the x-axis depicts the tax rate that is needed for each premium level depicted on the y-axis.

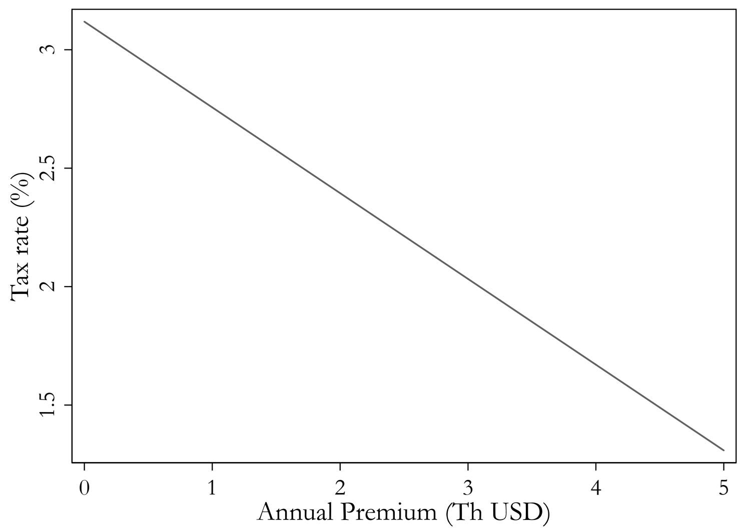


Figure H1: Tax Rate and Medicare Premium

Figure H2 shows welfare for the combined GLTHI + Medicare case, and when charging a Medicare premium in addition to the Medicare tax. The x-axis shows different premium levels, and the y axis shows the welfare consequences.

Three findings emerge from Figure H2: (1) a higher Medicare premium (and thus lower tax rate) is

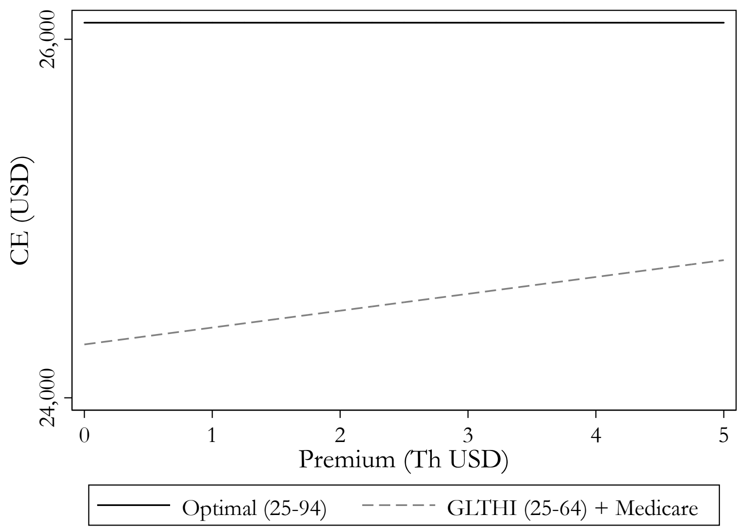


Figure H2: Welfare of GHHW and Medicare with different Premiums

desirable from a welfare perspective, and (2) at any premium level, GHHW does better than GLTHI.

To understand the intuition behind the welfare result in Figure H2, Figure H3 shows the expected lifecycle consumption profiles under (a) GHHW over the entire lifecycle, (b) GLTHI + Medicare with a zero premium and the corresponding tax rate in Figure H1, (c) GLTHI + Medicare with a premium of $5K and the corresponding tax rate in Figure H1.

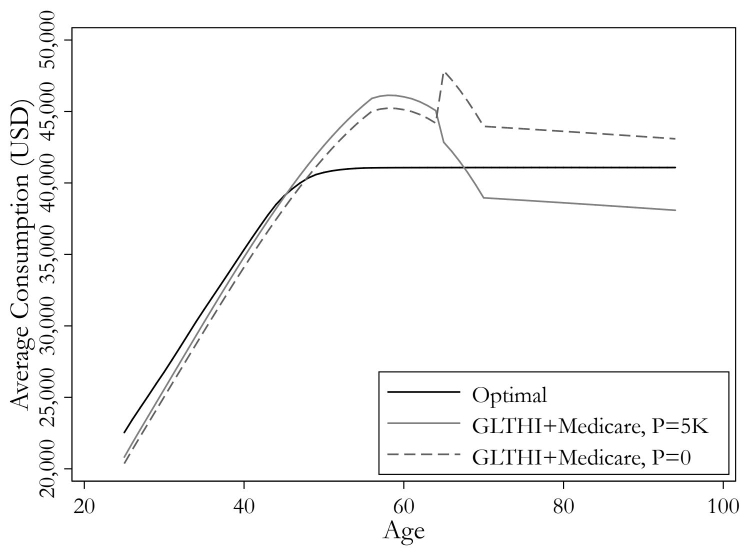


Figure H3: Expected Consumption Profile

Figure H3 illustrates that a higher Medicare premium increases consumption in early ages (because it decreases the tax rate). Under the GLTHI + free Medicare scencario, one observes a sharp increase in consumption at retirement, because individuals stop paying GLTHI premiums and stop paying Medicare taxes. Under the GLTHI + Medicare with a $5K premium scenario, one observes a reduction in consumption at retirement because the Medicare premiums exceeds the GLTHI premium. Figure H3 also illustrates than even a very large Medicare premium (and almost zero Medicare tax) does not outperform GHHW because it fails to achieve the same level of consumption at early ages. Compared with the optimal contract, it still has too much frontloading.