German Long-Term Health Insurance: Theory Meets Evidence*

Juan Pablo Atal† Hanming Fang‡

Martin Karlsson§ Nicolas R. Ziebarth¶

June 18, 2022

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. We study the German long-term health insurance (GLTHI) from a life-cycle perspective. We first present and discuss insurer regulation, premium setting, and the main market principles of the GLTHI. We then empirically estimate the ingredients of a life-cycle model to assess the welfare effects of the GLTHI. We find that the GLTHI achieves a high level of welfare against several benchmarks. Finally, we discuss the implications for the health care reform debate in the U.S.

Keywords: long-term health insurance; individual private health insurance; reclassification risk, intertemporal incentives, ACG scores, health transitions
JEL Classifications: G22; I11; I18.

*We are grateful to the editor, Harald Uhlig, and two anonymous referees for detailed comments that significantly improved the paper. Siqi Li provided outstanding research assistance. We also thank Konstantin Beck, Paola Bertoli, Rudi Blankert, Friedrich Breyer, Florian Buchner, Keith Ericson, Colleen Carey, Mariacristina De Nardi, David Dillenberger, Stefan Felder, Sebastian Fleitas, Eric French, Leora Friedberg, Paul Grieco, Ben Handel, Glenn Harrison, Jim Heckmann, Kate Ho, Mathias Kifmann, Amanda Kowalski, Dirk Krueger, Ethan Lieber, Tony Lo Sasso, Claudio Lucarelli, Maria Polyakova, John Rust, Andreas Ryll, Holger Strulik, Nicholas Tilipman, 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, Georgia State University, Johns Hopkins University, Lund University, Princeton University, University of Chicago, University of Pennsylvania, University of Maryland, University of Wisconsin-Madison, and conference participants at the American Society of Health Economists (2018), Annual Meeting of the Southern Economic Association (2018), Essen Health Conference (2019), Congress of the International Health Economics Association (iHEA 2019), GSE Forum in Microeconometrics (2019), Health Economics Working Group Meeting of the German Economic Association (*Ausschuss für Gesundheitsökonomie*, 2018), NBER Insurance Working Group Conference (2019), New York Area Health Economics Day (2019) and the North American Winter Meeting of the Econometric Society (2022) for very helpful comments and suggestions. We thank representatives of the German Association of Private Health Insurers and Frank Bowert for invaluable comments on how well our modeling of German LTHI reflects actuarial practices, and we thank Friedrich Breyer for connecting us with representatives from the the PKV which provided us the claims data. We have no 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, and the NBER, USA. 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 Economics and Brooks School of Public Policy, 2218 Martha Van Rensellaer Hall, Ithaca, NY 14850, and the NBER, USA; and the ZEW Manheim, Germany. Email: nrz2@cornell.edu.

1 Introduction

In the United States, private health insurance contracts are short-term, typically annual, policies. Without community rating regulation, 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 debated how to regulate short-term health insurance markets. The standard policies, such as community-rated premiums and guaranteed issuance regulations, strive to avoid reclassification risk, uninsurance, and unaffordable premiums for sick individuals (Claxton et al., 2017; Cole et al., 2019). However, these policies also imply a trade-off with unintended consequences such as adverse selection, which are typically addressed either through individual mandates, premium subsidies, or both (cf. Akerlof, 1970). The Affordable Care Act (ACA), enacted in the United States in 2010, 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 continue to lead to passionate debates and lawsuits.

Long-term health insurance contracts offer a fundamental alternative to short-term contracts. They provide policyholders with reclassification risk insurance without necessarily triggering adverse selection problems. Under long-term contracts, agents not only receive protection against contemporaneous health risks, but also against future health risks and premium fluctuations. 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 8.7 million individuals) hold individual long-term health insurance contracts sold by private insurers. After risk-rated premium setting at initial enrollment, the policies are guaranteed renewable until death (without an expiration date or enrollment period).[1] Moreover, all subsequent premium changes have to be community rated; that is, premium changes over the lifecycle are *independent* of changes in the policyholders' health status. In fact, given the market regulation, German Long-Term Private Health Insurance (henceforth GLTHI) foresees the payment of a constant real premium over the lifecycle, regardless of the evolution of an individual's income and health. Consequently, GLTHI contracts almost entirely eliminate the reclassification risk—at the expense of relatively high premiums during the early life years

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

("front-loading"). Further, GLTHI policies are contracts with one-sided commitment: policyholders can lapse their policies and switch to competitors, but insurers cannot cancel the ongoing long-term contracts.

We begin 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 of Germany. Next, we formulate the theoretical foundations of GLTHI and derive equilibrium premiums over the lifecycle, given the regulatory framework and considering endogenous lapsation of contracts. We highlight that the premium characterization of GLTHI is independent of the exact curvature of policyholders' utility functions, which offers a significant practical convenience. We also show that the evolution of income and health over the lifecycle are the key empirical inputs to assess the welfare consequences of GLTHI.

To construct the inputs needed to quantify welfare, we rely on a unique panel of claims data from one of the largest German private health insurers. The data cover 620 thousand policyholders over 7 years, spanning all age groups and all of the 16 German federal states.[2] We use enrollees' current and past diagnoses and claims along with the German version of the John Hopkins ACG software to calculate continuous health risk measures. Next, we propose a novel health risk classification method. This method allows us to categorize and model individuals' expected health risks. Then we study their health transitions at different points in their lifecycles. Moreover, 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.

Next, We use our theoretical and empirical inputs to quantify welfare under different contracts. Specifically, we compare the lifecycle welfare of GLTHI to (1) the 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 health insurance contract as derived by Ghili et al. (2019).

We find that the simple GLTHI performs relatively well when compared to the theoretically optimal contract. Within a plausible range of parameter values, we find that the GLTHI contract brings more than 95 percent of the welfare gain of the optimal contract, against a benchmark represented as a series of short-term contracts. When delving deeper into the underlying mechanisms, compared to the optimal contract, the GLTHI contract entails less consumption smoothing but also less reclassification risk over the lifecycle. The welfare loss due to less consumption smoothing is almost entirely

[2] The oldest policyholder is 106 years old and the most loyal policyholder has been a client for 86 years, illustrating the long-term and lifecycle dimension of these policies.

offset by less reclassification risk in the GLTHI contract. For the special case of flat lifecycle income, the GLTHI contract coincides with the optimal contract. Thus, one could also frame our exercise as quantifying the welfare loss when a social planner offers an optimal contract for flat lifecycle incomes to a population featuring various empirical income profiles.

Our results are robust to allowing for private savings, to a wide range of degrees of risk aversion, to using an initial health distribution that matches the German population as a whole, to the functional form for utility, to allowing for non-time-separable recursive preferences à la Epstein and Zin (1989), and to using U.S. income profiles.

We finish by discussing the potential implications of an existing real-world private long-term health insurance market for U.S. 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 strands of literature. 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. Harris and Holmstrom (1982) and Krueger and Uhlig (2006) study the properties of the competitive long-term contract that insures agents against income risk under one-sided commitment when agents receive endogenous outside offers. 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. Importantly, we build heavily on Ghili et al. (2019), as we use their characterization of the optimal long-term health insurance contract, and perform similar welfare comparisons. A key differentiation from their work is that we investigate a real-life application of long-term contracts using various high-quality claims and survey datasets. Cole et al. (2019)

uses a dynamic model of health investments and insurance to study the short and long-term effects of providing social insurance.

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* methods used in the existing literature.

Moreover, our paper relates to previous work on the German 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 ("old-age provisions") 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. The first tier consists of public health insurance or *Gesetzliche Krankenversicherung (GKV)* and covers 90 percent of the population. GKV enrollees and their employers pay income-dependent contribution rates (each party pays about eight percent of the gross wage, up to a cap) for a standardized benefit package with very little cost-sharing. As of this writing, GKV enrollees can choose between 97 non-profit sickness funds or *Gesetzliche Krankenkassen* (Bünnings et al., 2019; GKV-Spitzenverband, 2022). For historical reasons, select population subgroups can *permanently* opt out of the GKV system and join the second tier: private health insurance or *Private Krankenversicherung (PKV)* which covers the remaining ten percent of the population. Despite the two tiers, the German system provides almost universal coverage with an uninsurance rate of around 0.1 percent (Statistisches Bundesamt, 2020). Further institutional details about the two-tier German health insurance system are in Schmitz and Ziebarth (2017) (English) or Henke (2007) (German).

Besides Chile (cf. Atal, 2019), Germany is the only country in the world with an existing private long-term health insurance market. About 8.7 million policyholders receive GLTHI coverage (Association of German Private Healthcare Insurers, 2022a). During our study period, the GLTHI market consists of 48 private insurers that sell *comprehensive* as well as *supplemental* health policies (Association of German Private Healthcare Insurers, 2020). The focus of this paper is comprehensive insurance.[3]

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. They can decide to leave the public GKV system permanently and join the PKV and a GLTHI contract for the rest of their lives (Hullegie and Klein, 2010; Polyakova, 2016; Panthöfer, 2016).[4] To prevent individuals from strategically switching back and forth, the decision to enter the private market is essentially a lifetime decision. The basic principle is: "Once privately insured, always privately insured" (Schencking, 1999; Innungskrankenkasse Berlin Brandenburg, 2018). Appendix A1 discusses the institutional specifics and the empirical evidence on this principle.

Provider Networks. Provider networks and "Managed Care" are unknown in the public GKV and private PKV system in Germany; that is, in either system enrollees have the free choice of providers. Moreover, in both systems, reimbursement rates are centrally set 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 for GLTHI policyholders. Because they do not negotiate rates or build provider networks, private insurers mainly customize health plans and process, scrutinize,

[3]While comprehensive plans are not, supplemental plans are also available as group policies, primarily through private employers (*Betriebliche Krankenzusatzversicherung*). However, the market is relatively small and only counted 883,400 covered employees for the whole of Germany in 2019 (Association of German Private Healthcare Insurers, 2020).

[4]In 1995, Germany introduced a mandatory long-term care (LTC) insurance system. The structure follows the structure of the two-tier health insurance system with some minor differences: (a) Those who are mandatorily insured in the public GKV system also have to obtain public long-term care insurance (*Gesetzliche Pflegeversicherung, GPV*), which is funded by an additional 3.05% payroll tax for enrollees with children up to the contribution cap, equally split between employees and employers (Bundesministerium für Gesundheit, 2022). (b) Those insured in the private PKV system with GLTHI also have to obtain a private LTC insurance from a private insurer. Like GLTHI, premiums are risk-rated and front-loaded. However, the benefits have to equal the GPV benefits and the maximum premium is capped at the level of the highest possible premium in the public market (after five years as a policyholder) (Bundesministerium für Gesundheit, 2022; Association of German Private Healthcare Insurers, 2022b). However, it is important to note, first, that the LTC benefits and health care benefits are distinct and well-defined, and are separately insured; second, individuals who opt out of the public system can choose to enroll with different private insurers for their LTC and health insurances; third, from the inception, the private GPV system has required that the reserves in an insuree's LTC insurance be fully portable when she switches to another insurer, which was not the case for GLTHI in our study period. While there could definitely be some interesting interactions between LTC insurance and traditional health insurance, for example, one's LTC needs may be affected by the health care she receive and vice versa, modeling these interactions is beyond the scope of this paper and is left for future research.

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 contracts.[5]

2.1 Basic Principles of GLTHI

One-Sided Commitment and Guaranteed Renewability When individuals apply for a GLTHI contract, insurers have the right to deny applicants with bad risks or impose pre-existing condition clauses. However, once contracts are signed, insurers cannot terminate them. GLTHI contracts are not annual contracts, but *permanent lifetime contracts* without an end date. In other words, GLTHI contracts are guaranteed renewable over the lifecycle. However, policyholders can terminate these permanent contracts and switch insurers. Thus, the GLTHI is a market with *one-sided commitment*. It is very common that policyholders keep their GLHTI contract until they die (recall that Medicare does not exist in Germany). For example, in our sample, the policyholders' average age is 46 years, they have been clients of the insurer for 13 years; the oldest policyholder is 99 years old, and one policyholder has been a client for 86 years, see Table A1 (Appendix).

Premium Calculation Whereas the initial GLTHI premium is risk-rated, all subsequent premium increases are community-rated at the plan level. In fact, premiums are calculated under the basic principle of a constant lifecycle premium, sufficient to cover expenses over the policyholder's lifecycle. As this is a central aspect of our analysis, we devote Section 3 to discuss this issue extensively.

3 Lifecycle Premiums in GLTHI

Figure 1 illustrates the principle of constant premiums in the lifecycle governing premiums in the GLTHI market, as in Rosenbrock (2010) or Hofmann and Browne (2013). It uses our claims data as input and showcases four risk-age combinations at initial enrollment: low vs. high health risks and initial enrollment at ages 30 vs. 50. This illustration assumes constant health risk types over the lifecycle.[6] The low risk type (the "healthy") corresponds to a hypothetical individual with no pre-

[5]Compared to public insurance, one could argue that private markets and contracts are less prone to government regulatory risk. Koijen et al. (2016) study the impact of such "government-induced profit risk" on the demand of investors for what they refer to as "medical innovation premium."

[6]Constant health risks are assumed only for the purpose of illustrating the basic front-loading principle. Our empirical analysis below features evolving health risks, which is fundamental: First, it shows that front-loading can dampen the reclassification risk. Second, evolving health implies that individuals who start unhealthy may lapse their contract, which introduces (downwards) reclassification risk even if premiums are constant within a given contract. Also, we must consider lapsation when calculating equilibrium premiums.

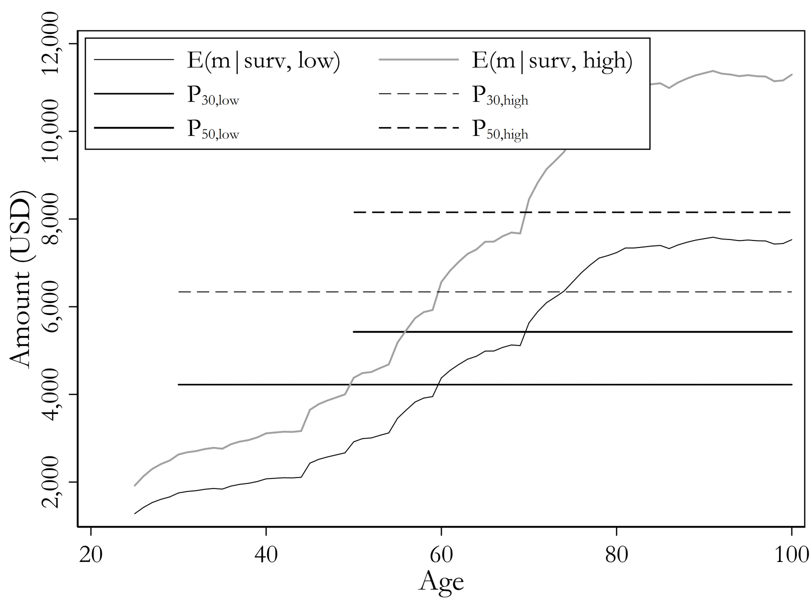


Figure 1: Illustration of GLTHI Lifecycle Premiums and Health Expenditures
Source: GLTHI Claims Panel Data, own calculations and illustration based on Rosenbrock (2010).

existing conditions; we denote her expected lifecycle health expenditures conditional on survival as $E(m|\text{surv, low})$. 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. We denote her expected lifecycle health expenditures conditional on survival as $E(m|\text{surv, high})$. Note that $E(m|\text{surv, low})$ and $E(m|\text{surv, high})$ would also represent the actuarially fair premiums of short-term spot contracts by age, for low and high risk types, respectively. $P_{30,\text{low}}$ ($P_{30,\text{high}}$, respectively) are the GLTHI premiums for a low (high, respectively) risk type who first enrolls in a GLTHI plan at age 30. Similarly, $P_{50,\text{low}}$ and $P_{50,\text{high}}$ are the premiums if the corresponding type joins a GLTHI plan at age 50.

Figure 1 has the following important features: First, in theory, premiums are stable over individuals' lifecycles. Front-loading dampens age-driven premium increases via the legal requirement of building individual capital stocks through so-called old-age provisions (*Altersrückstellungen*). The capital stock is the cumulative difference between premiums and expected claims, plus returns on investment from the capital stock.[7] Second, premiums are higher for policyholders who join a GLTHI plan later in their lives. This is because health care expenditures increase with age, and those who join later in life have fewer years to build up old-age provisions.[8] Third, because of the initial risk

[7]In 2020, the average capital stock built through old-age provisions was €28,195 ($32,425) per enrollee (Association of German Private Healthcare Insurers, 2022a).
[8]This is not necessarily true when health changes over time. With stochastic health, 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.)

rating, high risk types pay higher premiums than low risk types throughout their lives.[9]

3.1 Formal derivation of Lifecycle Premiums in the GLTHI

Let $P_t(\xi_t)$ be the initial premium offered when first signing a GLTHI contract in period t. As GLTHI contracts are individually underwritten at inception, $P_t(\xi_t)$ depends on the individual's health risk in year t, ξ_t. We assume that $\xi_t \in \Xi$ where Ξ is a finite set of health states. 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, though they are free to choose to lapse the existing contract and obtain a new one if a new contract is available at a lower premium.

The contract breaks even in equilibrium, given premium $P_t(\xi_t)$ and the subsequent lapsation decisions of the enrollee. Thus the equilibrium levels of $P_t(\xi_t)$ can be expressed 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 period 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 contract is that enrollees will lapse their current contract if and only if—given the evolution of their health—they can obtain a lower premium than their current guaranteed-renewable premium $P_t(\xi_t)$. Formally, lapsing a contract signed in $t < T$ at the risk-rated premium $P_t(\xi_t)$ occurs at the first $\tau > t$ such that $P_\tau(\xi_\tau) < P_t(\xi_t)$.[10]

It may be surprising that the policyholder's lapsation decision does *not* depend on the curvature of 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. This is because the other determinants of the continuation value—namely health transitions and income dynamics—are independent of the long-term contract. 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

[9] Appendix A2 discusses more institutional details about the premium calculation. It also lists reasons for why, in reality, real premiums may not be flat over the lifecycle.

[10] Note that we abstain from switching costs, and from horizontal differentiation across plans. Horizontal differentiation across plans tends to be minor because, as we explained in Section 2, GLTHI is a pure financial contract.

does not.[11]

Remark 1 *The lapsation decision under GLTHI is only driven by a comparison between one's current guaranteed renewable 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 plays a role in the lapsation decision under GLTHI. As the GLTHI is a pure financial contract, the lapsation decision is not driven by differences in provider networks associated with the policies.*

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

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

where the first element of the numerator, $\mathbb{E}(m_t|\xi_t)$, is expected health care expenditures in period t, given ξ_t; the second element of the numerator is the sum of the expected future health care expenditures over all remaining life years from t to T. Expected future health care expenditures are discounted at rate δ, with future spending at period τ weighted by $q_\tau(z|\xi_t, \mathbf{P}^{\tau}_{t+1}, P_t(\xi_t))$, the probability that (i) $\xi_\tau = z$ and (ii) the enrollee does not lapse (or die) between periods t and τ, given the subsequent equilibrium premiums $\mathbf{P}^{\tau}_{t+1}$. These expected lifecycle health care expenditures are then normalized by the expected number of years not lapsing the contract in the denominator.[12] In other words, the GLTHI lifecycle premium $P_t(\xi_t)$ equals the average of today's expected health care expenditures and all expected future health care expenditures, 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.

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 GLTHI equilibrium premiums 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.*

[11] 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).

[12] Of course, $q_\tau(z|\xi_t, \mathbf{P}^{\tau}_{t+1}, P_t(\xi_t))$ 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.

3.2 Arrow Securities Implementation

Krueger and Uhlig (2006) note that the optimal dynamic long-term contract with one-sided commitment can also be implemented by letting individuals trade one-period state-contingent Arrow securities. As such, long-term contracts can be turned into a sequence of short-term contracts. A main benefit of this alternative implementation is that it may enable competition between insurers rather than exposing the individual to the potential monopoly power of the insurer. A main drawback is the added complexity relative to guaranteed premium profiles under long-term contracts.

The following Lemma characterizes the quantity of securities traded that implement the premium path in the GLTHI.

Lemma 1 *The premium path of the GLTHI contract can be replicated by purchasing actuarially fair short-term insurance contracts supplemented by Arrow securities. The quantity of Arrow securities bought after history $\Xi_t \equiv (\xi_1, \xi_2, ..., \xi_t)$ that pay one dollar in state ξ_{t+1} is equal to*

$$b_t\left(\xi_{t+1} \mid \Xi_t\right) = \begin{cases} 0 \quad \text{if } P_{t+1}\left(\xi_{t+1}\right) < \tilde{P}_t\left(\Xi_t\right) \\ E(m_{t+1}|\xi_{t+1}) - \tilde{P}_t(\Xi_t) + \\ \quad \sum_{\tau>t+1}^{T} \sum_{z\in\mathcal{Z}} \delta^{\tau-(t+1)}[E(m_\tau|z_\tau) - \tilde{P}_t(\Xi_t)]q_\tau\left(z \mid \xi_{t+1}, \mathbf{P}_{t+2}^{\tau}, \tilde{P}_t(\Xi_t)\right) \quad \text{otherwise,} \end{cases} \tag{2}$$

where (a) $\tilde{P}_t(\Xi_t)$ is the GLTHI premium paid in period t after history Ξ_t, and (b) $q_\tau\left(z \mid \xi_{t+1}, \mathbf{P}_{t+2}^{\tau}, \tilde{P}_t(\Xi_t)\right)$ is the probability that (i) $\xi_\tau = z$, and (ii) the policyholder does not lapse (or die) between periods $t+1$ and τ, given the subsequent equilibrium premiums $\mathbf{P}_{t+2}^{\tau}$ and $\tilde{P}_t(\Xi_t)$.

Proof 1 *See Appendix A3.*

Lemma 1 states that the Arrow securities that replicate the GLTHI premiums pay either zero (in the event of lapsing), or the net present value of the difference between the guaranteed-renewable premium of the GLTHI and the premium of a short-term contract (in the event of not lapsing). Appendix A8 makes use of Lemma 1 and our data to calculate the quantity of securities needed to replicate the GLTHI contract (and also the optimal dynamic contract).[13]

3.3 Comparison to the Optimal Dynamic Contract

Under the assumption of time-separable preferences and risk aversion, and one-sided commitment (by insurers), the optimal dynamic health insurance contract as derived by Ghili et al. (2019)

[13] Appendix A4.1 provides an analogous expression for the Arrow securities that replicate the optimal dynamic contract in Ghili et al. (2019).

consists of consumption guarantees, $\bar{c}_t(\xi_t, \mathbf{y}_t^T)$. These depend not only on individuals' health but also on a vector of their current and future incomes $\mathbf{y}_t^T \equiv \{y_t, y_{t+1}, ...y_T\}$. The consumption guarantees can also be written as a series of contracts with guaranteed premium paths:[14]

$$P_\tau(\xi_\tau, y_\tau) = y_\tau - \bar{c}_t(\xi_t, \mathbf{y}_t^T). \tag{3}$$

Compared to the equilibrium GLTHI premium, which does not depend on income and almost entirely eliminates reclassification risk, the premium of the optimal contract *does* depend on income. The reason is that the optimal contract penalizes high premiums when the marginal utility of consumption is high.

Remark 3 *In the special case of flat income profiles over the lifecycle, the GLTHI contract coincides with the optimal contract.*

Appendix A4 provides more details and discussions on the optimal dynamic contract.

4 Claims and Survey Panel Data from Germany

This section describes the GLTHI claims panel dataset and the SOEP survey panel dataset used in this paper. We use the claims data primarily to estimate individual health transitions and expenditures over the lifecycle. We use the SOEP data primarily to estimate individual income dynamics over the lifecycle.[15]

4.1 GLTHI Claims Panel Data

The claims panel data are administrative records. They contain the universe of GLTHI contracts and claims between 2005 and 2011 from one of the largest German private health insurers. 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. The claims data also contain the *age* and *gender* of all policyholders, 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). Atal et al. (2019) provide more details about the dataset.

[14]This characterization of the optimal dynamic contract is independent of the curvature of the individual's utility function provided that it is concave and that her intertemporal preference is time separable. Section 6.6 discusses the case of non-time-separable preferences where the contract as characterized by Ghili et al. (2019) is no longer optimal.

[15]In one extension, we employ a claims dataset from one of the biggest German public insurers to assess the robustness of our results to using an alternative initial health distribution. In yet another extension, we use the PSID survey panel dataset to assess the robust of our results to using lifecycle income dynamics for the United States. See Section 6.6.

Sample Selection. We focus on primary policyholders. In other words, we disregard children and those below 25 years (555,690 enrollee-year observations).[16] Moreover, due to the 2009 portability reform (see footnote 65), we disregard inflows after 2008 (253,325 enrollee-year observations).[17] Our final sample consists of 1,867,465 enrollee-year observations from 362,783 individuals.

Descriptive Statistics. Table A1 (in Appendix A5) 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 underrepresented 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 health insurance system.[18] Figure A2 shows the distribution of policyholders' age at contract inception. The majority of individuals sign 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 affordable premiums.

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.[19] 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. (*Chefarztbehandlung*) that *TOP* and *PLUS* plans offer. For *ECO* and *PLUS* plans, a 20 percent coinsurance rate applies if enrollees see a specialist without referral from their primary care physician, while such coinsurance does not apply for *TOP* plans. 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

[16] The GLTHI market only features individual policies, not family policies; even children have their individual policy. However, if parents purchase the policy for their child within two months of the birth, no risk-rating applies for the child. Under the age of 21, insurers do not have to budget and charge for old-age provisions.

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

[18] 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.

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

have mechanical effects on claim sizes and correlate with policyholders' age, we control for them when modeling health transitions in Section 5.

4.2 Socio-Economic Panel Study

The German Socio-Economic Panel Study (SOEP) is a representative longitudinal survey. Since 1984, it annually surveys about 10,000 households and 20,0000 individuals above the age of 17 (Goebel et al., 2019). 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.[20]

Our main income measure is *equivalized post-tax post-transfer annual income*. It accounts for within-household redistribution 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.[21] Then, the total annual post-tax household income is divided by the number of household members using the modified OECD equivalence scale.[22]

Sample Selection. We leave the representative sample as unrestricted as possible, but exclude observations with missings on core variables such as age, gender, employment, or the insurance status. Other than that, we only exclude respondents below the age of 25.

Descriptive Statistics. Table A2 (Appendix A5) provides summary statistics for our SOEP sample. From 1984 to 2016, the average annual income per household member was $26,433.[23] Note that this measure has positive values for *all* respondents, including those who are not active in the labor market.

For completeness, Table A2 also shows statistics for two additional income measures: *monthly gross wage* ($2,940) and *monthly net wage* ($1,921). These measures have positive values for all working people with labor earnings (58 percent). The SOEP Group generates and provides these individual-level income measures to guarantee consistency over time.

[20]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).

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

[22]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.

[23]Again, all monetary values are in 2016 USD.

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. Forty-two percent work full-time and 14 percent part-time.

5 Modeling Health Risks 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. This section introduces insights from actuarial science to produce an "efficient" risk classification. We consider our procedure to be a significant improvement over the approach used in the state-of-the art literature.

Following the literature (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^* corresponding to the *unscaled total cost predicted risk*. 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 health care 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 GLTHI claims data in 2006 (the first year) and 2011 (the last year). Both distributions are approximately unimodal, and they appear stable over time.[24] It also illustrates that the distribution of λ_t^* is heavily skewed and has a long right tail, consistent with stylized facts of health expenditures distributions (see French and Kelly, 2016).[25]

Our method combines the continuous score λ_t^* and its $n-1$ lags into the vector of scores, $\Lambda_t^*(n) \equiv \{\lambda_t^*, \lambda_{t-1}^*, .., \lambda_{t-n-1}^*\}$, which we then map into K different *discrete risk categories*. 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 used 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 skewness in Figure 2 implies that the amount of reclassification risk will

[24]This also suggests that excluding inflows of new enrollees in 2010 and 2011 due to the portability reform (see Section 2) poses no major issue.

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

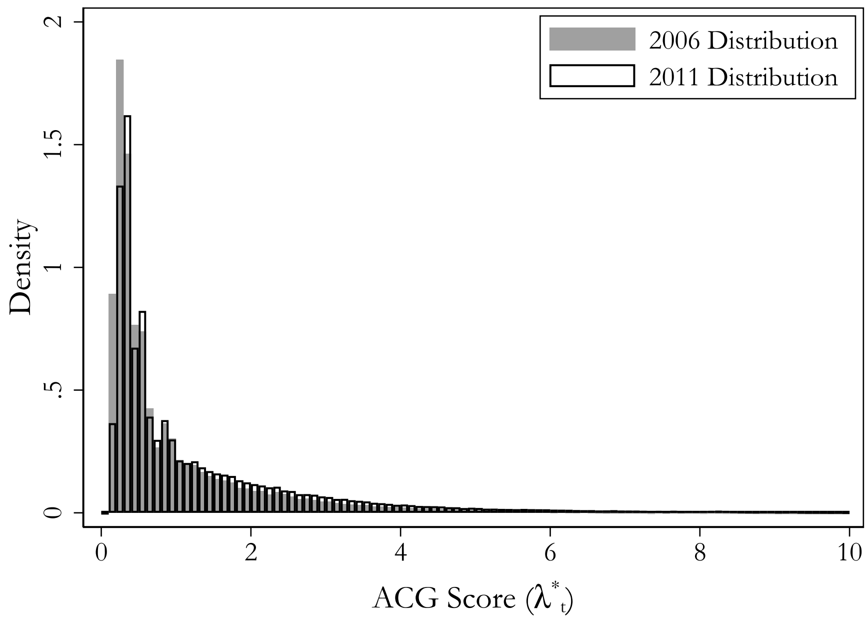


Figure 2: Distribution of λ_t^* in 2006 and 2011

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 $\lambda_t^* > 10$.

strongly depend on the granularity of the risk classification.

The commonly-used approach to risk classification would use an ad-hoc criterion to partition the domain of λ_t^* into different risk classes.[26] 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^*, i.e., $\Lambda_t^*(n)$, where we determine n within our procedure. Our approach 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 $\Lambda_t^*(n)$ into an endogenously determined number of risk categories K and corresponding partitions. The method maximizes an efficiency criterion from the actuarial science literature (cf. Finger, 2001).

More specifically, 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 (Finger, 2001), an efficient risk classification system has two properties: *homogeneity*—meaning that individuals in the same risk

[26] 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.

category have similar risk and *separation*—meaning that the categories have sufficiently different expected claims to justify distinct categories.[27]

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 $\lambda_t \in \{1, \ldots, K\}$ is the risk category assigned to a person with those ACG scores. The *efficient* risk classification f_K maximizes the "structure variance"

$$SV(f_K) = \text{Var}(m_t) - \sum_{k=1}^{K} \Pr(\lambda_t = k) \text{Var}(m_t \mid \lambda_t = k), \tag{4}$$

where m_t is individual annual health expenditures (Finger, 2001). 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 (4), we can write the structure variance as:[28]

$$SV(f_K) = \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n))) - \sum_{k=1}^{K} \Pr(\lambda_t = k) \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k). \tag{5}$$

Note that the first term in Equation (5) 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 additional things are worth noting about Equation (5). First, only mean expenditures 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 that the efficient classification also maximizes the coefficient of determination (R^2) in a regression of expenditures on risk class indicators (Kriegel et al., 2017).

[27] For instance, 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.

[28] 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 $\text{Var}(m_t \mid \lambda_t = k) = \mathbb{E}(\text{Var}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k) + \text{Var}(\mathbb{E}(m_t \mid \Lambda_t^*(n)) \mid \lambda_t = k)$.

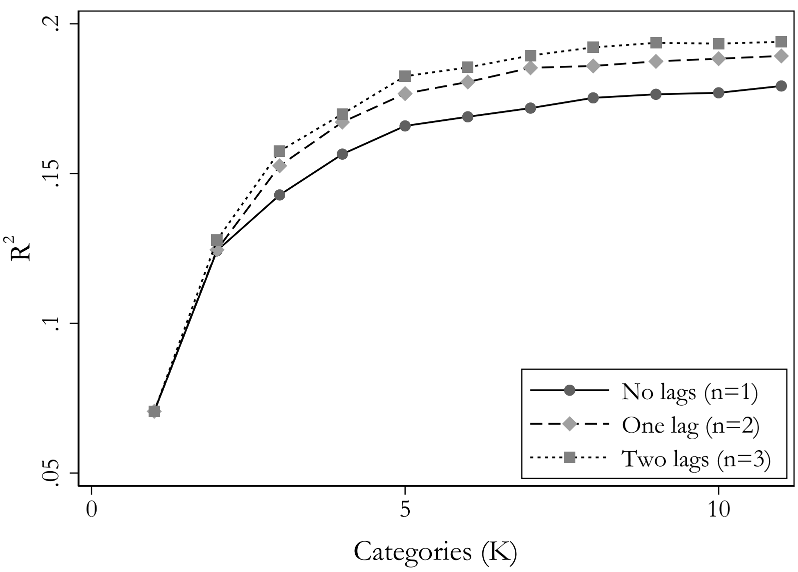


Figure 3: Performance of Alternative Risk Classifications.

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

Model selection. The next step of the risk classification method is to determine the number of risk classes K and the number of lags n of ACG scores when computing $\mathbb{E}\left(m_t \mid \Lambda_t^*(n)\right)$.[29] k-means clustering is an unsupervised learning method; we assume that the objective $SV(.)$ applies also when selecting K and n. As noted, this implies using 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, we need to estimate $\mathbb{E}\left(m_t \mid \Lambda_t^*(n)\right)$. To get accurate predictions along the entire distribution, including the tails, we use cubic regression splines.[30] Once we have estimated $\mathbb{E}\left(m_t \mid \Lambda_t^*(n)\right)$ for all $n > 1$, we conduct k-means clustering to maximize the objective function (5).

Figure 3 shows how the performance depends on parameters K and n. For all values of n, there is initially a rapid improvement in the predictive power when we increase the number of categories K; however, this improvement levels out at relatively low levels of K. Moreover, starting from a classification scheme that uses only the previous year's scores ($n = 1$), there is distinct improvement when we add one lag ($n = 2$). However, adding a second lag of the ACG scores brings only marginal improvement in the predictive accuracy. Figure 3 shows that, beyond including one lag ($n = 2$) and 7 distinct risk classes, increasing K or n further yields negligible improvement in performance. Thus as a compromise of predictive accuracy and model parsimony, we choose $K = 7$ and $n = 2$ (in the spirit

[29] 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).

[30] Figure A3 provides a comparison of mean expenditure by $\Lambda_t^*(n)$ before and after smoothing for $n = 2$.

of Heckman and Burton Singer (1984) and Keane and Wolpin (1997) in their choice of the number of unobserved types in the labor economics literature).[31] [32]

5.2 Estimation of Transition Matrices and Expenditure Risks

Next, we estimate transition rates between different discrete risk categories λ_t and also mean expenditures by risk categories. We posit that the risk type of individual i at age t, $\xi_{it} \equiv (A_{it}, \lambda_{it})$, depends on contemporaneous risk λ_{it} and age, where A_{it} is one of eleven age groups (five-year bands from age 25 to age 75 and 75+).[33]

Considering that the clustering method generates a set of risk classes of very different sizes, we resort to a parametric, yet flexible, model. To estimate the transition matrices, we estimate a multinomial logit model for health dynamics:

$$\eta_{it}^{j} = A_{it}\beta_j + L_{it}\gamma_j + h\left(A_{it}, L_{it}; \theta_j\right) + \epsilon_{it}^{j}, \tag{6}$$

where η_{it}^{j} represents the log odds for $\lambda_{i,t+1} = j$, for $j \in \{2, \ldots, 8\}$ relative to the reference category $\lambda_{i,t+1} = 1$ ($\lambda_{i,t+1} = 8$ 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 (6) 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.[34]

To model the expected claims based on risk type, we use predicted values from an OLS regression. In addition to the controls in Equation (6), we also control for a vector of dummies Q_{it} representing health plan generosity $q \in \{ECO, PLUS, TOP\}$. The base specification is:

$$m_{it} = A_{it}\beta + L_{it}\gamma + Q_{it}\delta + h\left(A_{it}, L_{it}, Q_{it}; \theta\right) + \epsilon_{it}. \tag{7}$$

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 (7) until no remaining term is statistically significant.[35]

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

[32] Section A6 provides a number of additional robustness checks to the classification method, including the role of outliers and an alternative ways of including a longer history of claims. We also assess the extent to which our results are driven by the fact that the sample changes when n changes. We also show that transition rates between risk categories satisfy first-order stochastic dominance as assumed in Ghili et al. (2019).

[33] 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.

[34] We selected the interaction 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.

[35] We use a subsample of policyholders with moderately-sized deductibles to estimate conditional expenditures given λ_t as policyholders with large deductibles may decide not to submit their claims, leading to downward biased estimates. In

Descriptive Statistics for Transition Matrices. Table 1 displays one-year transition rates between health risk categories for all age groups; the numbers are predicted probabilities based on Equation (6). Two facts emerge from Table 1. 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 generally decreases over risk categories but, still, Forty-five 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 1: 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 A3 and A4 (Appendix) show the transition matrices separately 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, 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, expected expenditure conditional on risk class is constant in age.

Descriptive Statistics for Expenditures. Table 2 shows mean expenditures m by age group and the distribution of risk categories within each group.

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

Appendix A6 we provide some descriptive statistics for this subsample, which generally confirm that this assumption is reasonable.

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 (7) turn out statistically significant, the deviations from mean expenditure within each risk class are economically insignificant. Figure A5 (in Appendix A6) illustrates this point. We interpret it as evidence that our preferred risk classification is rich enough and therefore restrict all age effects to zero.

There is also a clear age gradient in health expenditure risk. The probability of being in the lowest risk category, $\lambda = 1$, 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.

Table 2: Health Expenditures and Risk Categories λ by Age Group

			λ_t						
Age	Mean	S.D.($\mathbb{E}(m \mid \lambda)$)	1 (Healthiest)	2	3	4	5	6	7 (Sickest)
25-30	1,996	1,782	0.789	0.154	0.039	0.013	0.004	0.001	0.000
30-35	2,619	1,938	0.740	0.178	0.054	0.020	0.006	0.001	0.000
35-40	2,840	2,086	0.652	0.225	0.085	0.027	0.009	0.002	0.000
40-45	3,119	2,411	0.622	0.227	0.103	0.034	0.012	0.003	0.000
45-50	3,719	2,946	0.539	0.258	0.136	0.046	0.016	0.004	0.001
50-55	4,880	3,544	0.463	0.263	0.174	0.068	0.024	0.007	0.001
55-60	6,517	4,573	0.291	0.319	0.232	0.108	0.036	0.011	0.002
60-65	7,635	4,299	0.184	0.313	0.269	0.155	0.058	0.019	0.003
65-70	7,151	4,421	0.069	0.291	0.337	0.217	0.069	0.014	0.002
70-75	8,355	5,026	0.019	0.203	0.347	0.309	0.105	0.015	0.002
75+	10,020	4,490	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.

5.3 Lifecycle Income Paths

Next, we estimate the lifecycle income paths using 33 years of SOEP panel data (1984-2016). As noted in Section 4.2, our 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 normalized by the number of household members. We estimate the following individual fixed effects model:

$$log(y_{it}) = \theta_i + f(age_{it}) + \epsilon_{it}, \quad (8)$$

where y_{it} is our income measure in 2016 U.S. dollars for individual i in year t; 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.

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 are important determinants of the lifecycle welfare under the different health insurance contracts.

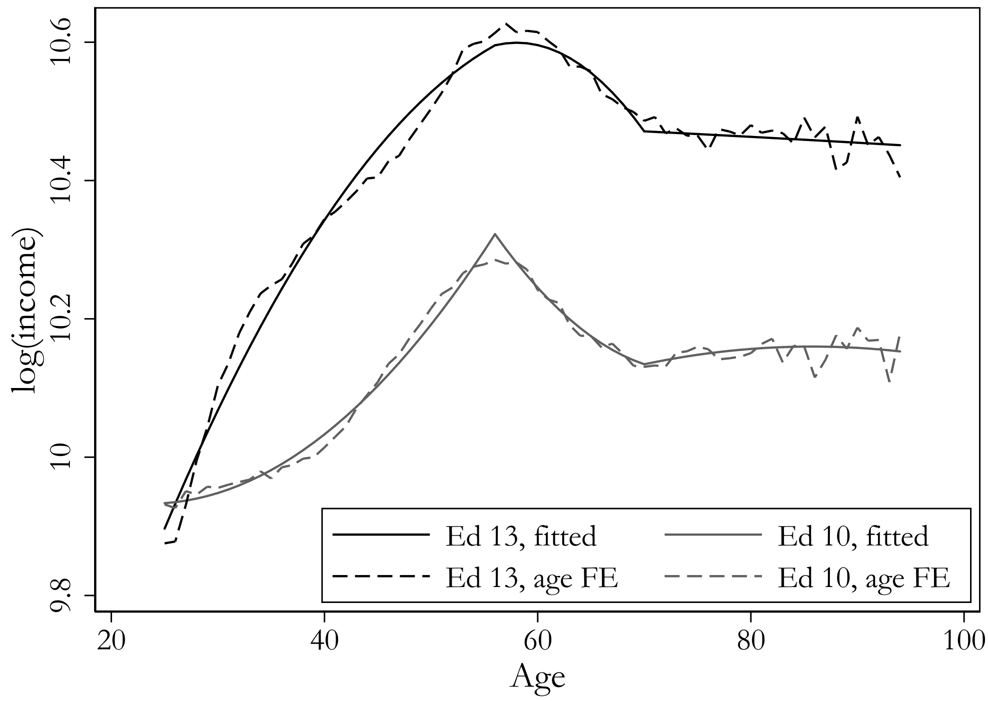


Figure 4: Lifecycle Income Paths Germany, Nonparametric and Fitted.
Source: SOEP (2018), years 1984 to 2016. All values in 2016 USD.

The two dashed curves in Figure 4 show the estimated age fixed effects for people with 10 years

[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).

and 13 years of schooling, respectively. 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. There exists a level difference in income paths between the two educational groups over the entire lifecycle. We accommodate these 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 and across three different age bins: $[25, 56]$, $[56, 70]$ and $70+$. This is illustrated by the two solid lines in Figure 4. Note that the piece-wise squared polynomials fit the empirical lifecycle profiles very well.

Several factors can explain these lifecycle income patterns. 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).

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. We use a discount factor $\delta = 0.966$ (corresponding to a discount rate of 3.5 percent). 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 $t \in \{25, ..., 94\}$.

Figure 5 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.

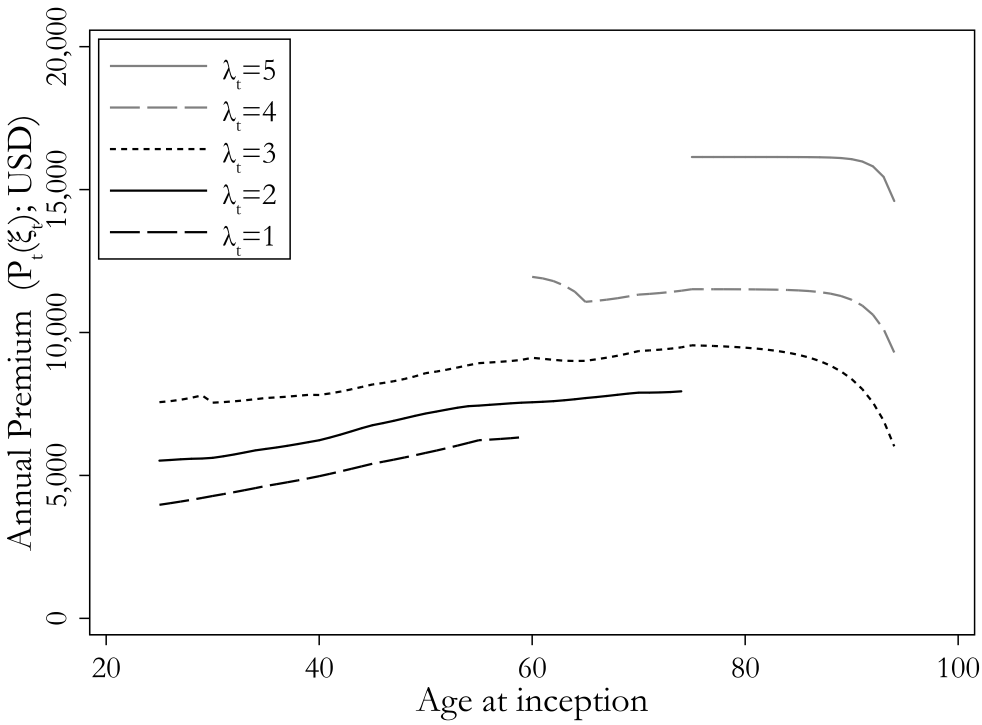


Figure 5: Calibrated Starting Premiums $P_t(\xi_t)$ in the GLTHI

Three forces are at play that determine the lifecycle profile of $P_t(\xi_t)$ in Figure 5. 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 health transitions depend strongly on age (through the A_t component of ξ_t; see Equation (6)). 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 subset 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.

Appendix A7 compares the calibrated and the observed premiums by age at inception for ages between 25 and 75 (where we have enough observations of new policies being issued at different health categories). First, we observe positively sloped starting premiums by age over this 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 calibrated premiums for sicker individuals are slightly larger than observed premiums, the calibrated premiums are in general very similar to the observed ones.[37]

6.2 Comparison with the Optimal Dynamic Contract

This subsection compares lifecycle premiums and the amount of front-loading between the GLTHI and the optimal dynamic contract (Ghili et al., 2019).

Using our empirical health transition and income dynamics, Table 3 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. The reason is that the likelihood of a further health deterioration also decreases, the worse the current health status is.

Panel (b) of Table 3 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.[38] 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. The optimal contract entails less front-loading than GLTHI because a higher front-loading increases the marginal utility of consumption. However, the differences in premiums between the GLTHI and the optimal dynamic contract decrease as the health status at contract inception worsens. 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).

[37] Although we could, in principle, use the observed premiums to evaluate welfare under the GLTHI, there are two reasons to use the calibrated premiums instead: First, our analysis of the market equilibrium under different assumptions requires knowing the premiums for all possible combinations of ξ and t and many of these are completely absent or represented by only a small number of individuals in our data. Second, we compare the welfare properties of GLTHI to some alternatives and want to eliminate possible sources of error in such a comparison; therefore, it is desirable to conduct the welfare comparisons based on premiums generated in an analogous manner.

[38] Recall that the GLTHI premiums do not depend on income profiles, thus do not depend on education levels, while the optimal contract premiums do; see Equations (1) and (3).

Table 3: 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 are in 2016 USD.

Panel (c) of Table 3 shows the optimal contract for an individual with 10 years of schooling (*Ed 10*). This individual has a flatter income profile over her lifecycle (see Figure 4), which is why the optimal contract entails a *higher* degree of front-loading for *Ed 10* education 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.

6.3 Equivalent Arrow Securities

As discussed in Section 3.2, it is possible to replicate the dynamic contracts with one-period Arrow securities. Consider an individual who enters the GLTHI market at age 25 in the healthiest state; $\xi_1 = \Xi_1 = 1$. As shown in Table 3, the GLTHI contract specifies a guaranteed-renewable premium of $P_1(1) = \$3,973$ which represents $\$2,499$ in excess of expected claims. This frontloading can be reinterpreted as the total amount paid for seven Arrow securities $b_1(k|\Xi_1 = 1), k \in 1, ..., 7$, purchased in period 1 and that would pay 1 dollar if the realized health state in period 2 is k. The prices of each security is given by the corresponding transition probability $\pi(k|1)$. Stacking quantities and probabilities in row vectors; i.e., $\mathbf{b}_1(1) \equiv (b_1(1|\Xi_1 = 1), ..., b_1(7|\Xi_1 = 1)), \boldsymbol{\pi}_1(1) \equiv (\pi_1(1|\Xi_1 = 1), ..., \pi_1(7|\Xi_1 = 1))$, we have:

$$\mathbf{b}_1(1) \times \boldsymbol{\pi}_1(1)' \times \delta + E(m_1|\xi_1 = 1) = P_1(1)$$

Applying formula 2 described in Lemma 1 to our data we get:

$$\mathbf{b_1}(1) = 10^3 \times (1.35, 11.88, 20.62, 28.90, 39.57, 60.25, 202.73)$$

In other words, to replicate the premium path under the GLTHI contract, an individual who starts in the healthiest state purchases a short-term contract in the spot market (at a premium $E(m_1|\xi_1 = 1) = 1,473$), plus a quantity $\mathbf{b}_1$ of seven Arrow securities, at a price $\boldsymbol{\pi}(1) \times \delta$.

Consider now an individual in period 2 with history $\Xi_2 \equiv (\xi_1, \xi_2) = (1, 1)$. As $P_1(1) < P_2(1)$, this individual does not lapse between period 1 and period 2, and continues to pay $\tilde{P}_2([1, 1]) = \$3,973$ in period 2. The individual buys securities in the amount of $\mathbf{b}_2(\xi_2 = 1|\Xi_1 = 1))$ such that

$$\mathbf{b}_2(\xi_2 = 1|\Xi_1 = 1)) \times \boldsymbol{\pi}_2(1)' \times \delta + E(m_2|\xi_2 = 1) - b_1(1|\Xi_1 = 1) = \tilde{P}_2([1, 1]),$$

where $\mathbf{b}_2(\xi_2 = 1|\Xi_1 = 1))$ and $\boldsymbol{\pi}_2(1)$ are analogously defined as $\mathbf{b}_1(1)$ and $\boldsymbol{\pi}_1(1)$, respectively. We get $\mathbf{b}_2(\xi_2 = 1|\Xi_1 = 1)) = 10^3 \times (2.74, 13.25, 22.04, 30.41, 41.22, 62.19, 192.92)$.

Note that the amount of securities bought in any given period t for every future realization of the health status depends on age and on the full history of health statuses up to t.

Appendix A8 provides equivalent numerical examples for the optimal contract. As expected, the quantity of securities needed to replicate the optimal contract is smaller than for the GLTHI contract, reflecting the lower degree of front-loading required under the optimal contract.

6.4 Welfare Results

We now calculate welfare under the different contracts. We summarize lifetime utility under each contract with the Certainty Income Equivalent (CE). As a benchmark, we compare the CE of each contract with the first-best consumption level, which is equal to the annualized present discounted value of "net income" $y_t - \mathbb{E}(m_t)$, taking into account mortality risk. Details on these welfare measures are in Appendix A9. We calculate welfare by simulating the economy for a lifecycle of 70 years, from age 25 to age 94 for $N = 500,000$ individuals.

So far, we have not specified the utility function because the premiums in the long term contracts do 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:[39]

$$u(c) = -\frac{1}{\gamma} e^{-\gamma c}. \tag{9}$$

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, to using a constant relative risk aversion (CRRA) utility function, and to using 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 4 show the results assuming that individuals start in each of the seven possible health states (conditional on being alive). 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 A3, 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, C^*; Column (2) calculates welfare under a series of short-term contracts, C_{ST}; Column (3) shows the results under the GLTHI contracts, C_{GLTHI}; and Column (4) calculates welfare under the optimal dynamic contract, C_{GHHW}. Column (5) shows what fraction of the welfare difference between the optimal contract and the first best is recouped by the GLTHI; column (6) shows the percentage welfare difference between the GLTHI and the optimal contract.

Overall, we find in Table 4: 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 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

[39] 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 4.

Table 4: 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): $\Delta_0 = \frac{1}{100}[89.10, 10.25, 0.47, 0.11, 0.04, 0.03, 0]$					
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 4). Panels (a) to (g) differentiate by initial health status $\lambda_{25} \in \{1, ..., 7\}$. In Panel (h), 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.

compared to the first-best. For all initial health states at age 25 and for both lifecycle income profiles, the consumption certainty equivalents are negative.

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. Column (3) shows that, under the GLTHI contract, 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 losses of the GLTHI contract, relative to the optimal contract, are moderate, at 4.2 and 6.1 percent for *ED10* and *ED13*, respectively.

Fourth, comparing Column (6) across panels, we find that the welfare difference between the optimal contract and GLTHI shrinks as health at inception worsens, becoming almost negligible for individuals starting in state 5 or above. We also note that the CE under all contracts is negative for individuals starting in the sickest state, $\lambda_{25} = 7$, even if welfare is positive under the first best. This highlights the significant negative welfare consequences of *one-sided commitment*, i.e., the inability of policyholders to commit to long-term contracts, together with the inability of consumers to borrow.[40]

Finally, Panel (h) provides one of the main results of our analysis. Using initial health distributions that correspond to the observed empirical distributions for those age-25 in our sample, we find that the welfare loss under the GLTHI contract relative to the optimal contract is at most 3.9 percent. Appendix A13 further explores the robustness of this finding: considering a large number of draws of distributions over starting states, we conclude that the welfare loss of the GLTHI relative to the optimal contract is bounded at around six percent for the better-educated *ED13* group and at around four percent for the less-educated *ED10* group.[41]

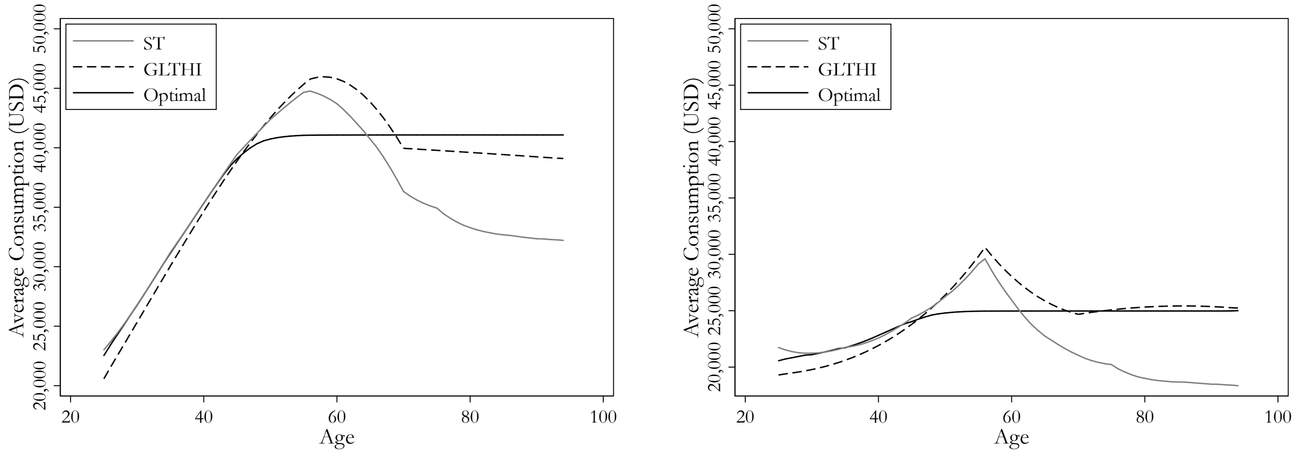
6.5 Understanding GLTHI Welfare

Average Lifecycle Consumption and Intertemporal 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 6 plots *average* consumption for these three contracts over the lifecycle, separately for *Ed 10* (Figure 6a), and *Ed 13* (Figure 6b). The figures illustrate the driving forces behind the welfare differences in Table 4.

As shown by the thin solid lines, under a series of short-term contracts, average consumption is simply income minus expected health expenditures. The average consumption profile is therefore

[40] As is well known, if consumers could borrow, they can "manufacture" commitment power by posting a "bond" with the insurer that equates the discounted sum of expected medical claims and restore first-best outcomes (see, e.g., Cochrane (1995) and Hendel and Lizzeri (2003)).

[41] 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, which 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. Appendix A13 also shows how the welfare losses of GLTHI reported in Table 4 overestimate the welfare gap in a situation where short-term contracts are used as a benchmark.



(a) *Ed 13*

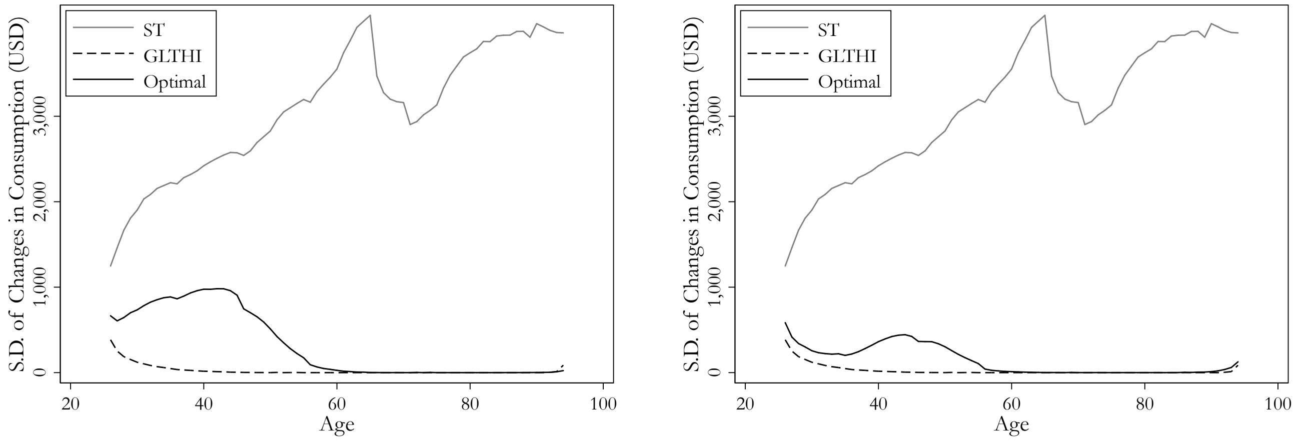
(b) *Ed 10*

Figure 6: Expected Consumption over the Lifecycle by Education

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, 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. Recall that the optimal contract optimally trades off utility from reducing reclassification risk and consumption smoothing over the lifecycle; hence, it implies a much smaller degree of front-loading than the GLTHI contract (see Table 3). Thus, compared to GLTHI, 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.

Reclassification Risks. To illustrate the degree of *reclassification risk* over the lifecycle, Figure 7 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 7 plots, for each age t, the standard deviation of $\Delta C_{i,t} \equiv C_{i,t+1} - C_{i,t}$ across individuals i.)

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 $\lambda_{25} > 1$



(a) *Ed 13*

(b) *Ed 10*

Figure 7: Simulated Standard Deviation of Consumption Changes

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 the optimal contract 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 $\lambda_{26} = 1$," 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.

Relatedly, Figure 8 compares average lapsation rates under each long-term contract.[42] 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.

[42]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 also be interpreted as a lapsation from an equivalent set of guaranteed premium paths. Figure 8 uses this interpretation of lapsing.

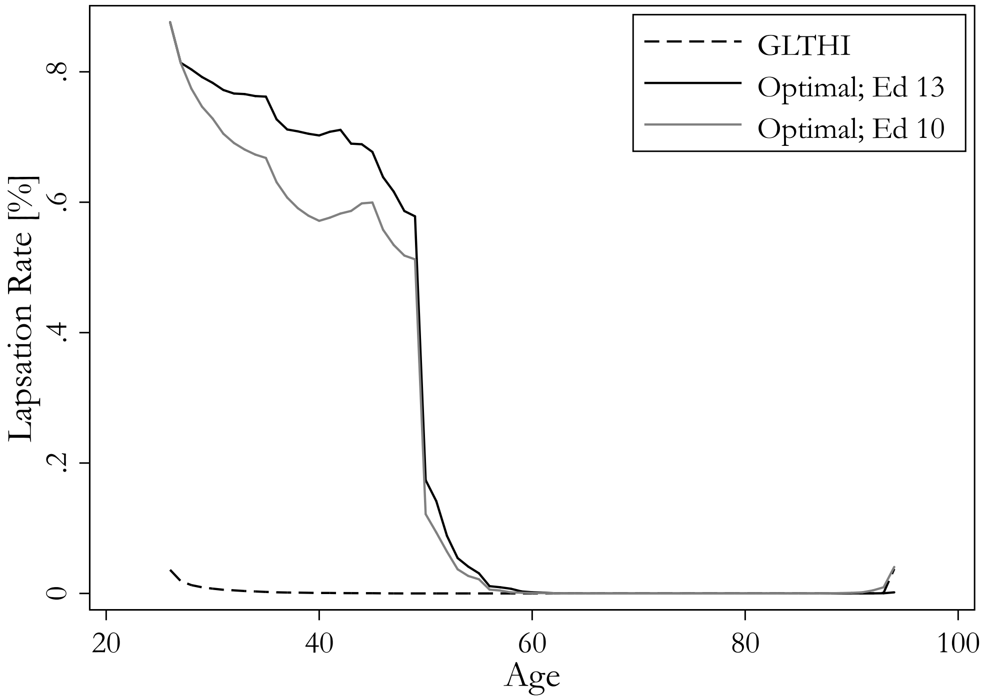


Figure 8: Laspation Rates over the Lifecycle by Education

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 6. However, the extra front-loading results in a lower standard deviation of consumption changes. Of course, by design, the optimal contract optimally balances these trade-offs and thus—in environments satisfying the conditions required for Ghili et al. (2019)'s theoretical characterization such as the one we analyze—achieves a higher welfare than the GLTHI contract. Our main findings show, however, that the GLTHI contract—despite its simplicity—achieves welfare that is close to the optimal contracts.

6.6 Robustness: Savings, Risk Aversion, CRRA preferences, Epstein-Zin preferences, Publicly Insured, and U.S. Income Profiles

We investigate the robustness of our main findings in various dimensions. First, we investigate the role of savings. Second, we quantify how our results change when we vary the degree of risk aversion, i.e., the parameter γ in the CARA utility function specified by Equation (9). Third, we use a CRRA utility function instead of CARA. Fourth, 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. Fifth, we use claims data from one of the biggest German public insurers to replicate our main findings for this population. Finally, we use U.S. income profiles instead of German income profiles.

Savings. 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). We relegate the details to Appendix A10.

Table 5: 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 (h) of Table 4. All consumption certainty equivalents (welfare) are in 2016 USD per capita, per year.

Table 5 shows the welfare results, assuming interest rate from savings $r = 1/\delta - 1$. Allowing for precautionary savings substantially improves welfare under the series of short-term contracts. Consider the same distribution of initial health state as in Panel (h) of Table 4. 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.

Degree of Risk Aversion. Under our parametric assumptions on preferences, the GLTHI contracts entail a welfare loss of less than four percent relative to the optimal dynamic contract. The welfare loss from heavier frontloading in the GLTHI is, to a large extent, compensated by the welfare gains from almost entirely eliminating reclassification risk. Following the literature, our main results assume a level of risk-aversion of $\gamma = 4 \times 10^-4$ (cf. Ghili et al., 2019).[43]

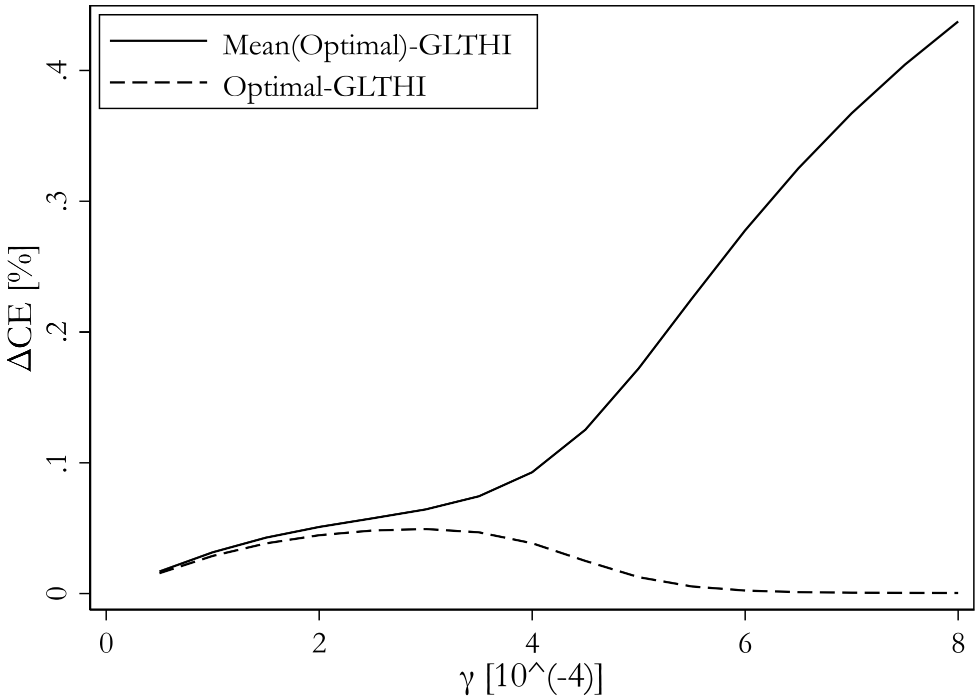


Figure 9: Difference in CE (GHLTI vs. GHHW) by Risk Aversion

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.

Figure 9 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). The dashed curve in Figure 9 shows the total welfare gap between the GLTHI and the optimal contract. The maximal welfare difference between the two across all values of γ is 5 percent when $\gamma = 3 \times 10^-4$.[44] As seen, the welfare gap is small when γ is either very low or very high. That is, our main qualitative finding—the simple GLTHI contract can 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

[43]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.

[44]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.

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. In fact, for $\gamma = 0.0004$ and barring differences in reclassification risk across contracts, the lifecycle consumption under the GHHW contracts produces welfare gains of approximately US $2,600$ per year.[45]

In summary, varying the level of risk aversion affects the performance of the 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 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 zero, 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 the GLTHI contract. 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.

CRRA Preferences. Our main analysis assumes CARA preferences. In this section we evaluate robustness of our main findings when we use CRRA preferences instead, that is $u(c) = \frac{u^{1-\sigma}}{1-\sigma}$. Previous research has suggested that CRRA might represent choices well (see e.g. Chiappori and Paiella 2011). Following Dohmen et al. (2011), we use a coefficient of risk aversion of $\sigma = 4$. As seen, our main conclusions of relatively small welfare differences across both contracts also holds.[46]

[45] 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.

[46] To avoid negative consumption, we impose a consumption floor of \$10,000. This is a binding constraint in approximately 10 out of 10,000 simulated consumption levels.

Table 6: Welfare with CRRA preferences

	C_{GLTHI}	C_{GHHW}	$\frac{C_{GHHW}-C_{GLTHI}}{C_{GHHW}}$
Ed 10	22,148	22,681	0.024
Ed 13	29,949	31,142	0.038

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 (h) of Table 4. All consumption certainty equivalents (welfare) are in 2016 USD per capita, per year.

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.

As in Epstein and Zin (1989), preferences are defined recursively as $V_t = F(c_t, R_t(V_{t+1}))$ with $R_t(V_{t+1}) = G^{-1}(\mathbb{E}_t G(V_{t+1}))$ and we 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 A11, we show that the consumption certainty equivalent can be expressed as:

$$c = \left(\frac{\frac{\left(G^{-1}(\mathbb{E}_0(G(V_{t_0}(\xi_{t_0}))))\right)^{1-1/\psi}}{1-\delta}}{\sum_{j=t_0}^{T} \delta^{t-t_0} S_{t_0}^{j}} \right)^{\frac{1}{1-1/\psi}}, \tag{10}$$

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 10 shows differences in certainty equivalents between the GLTHI and the optimal contract. As seen, the welfare differences are small over all the entire range of parameter values. Notice that in Figure 10, 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.

Publicly Insured. So far, our results leverage a population of GLTHI policyholders. However, several institutional features imply that this population is not representative of the German population

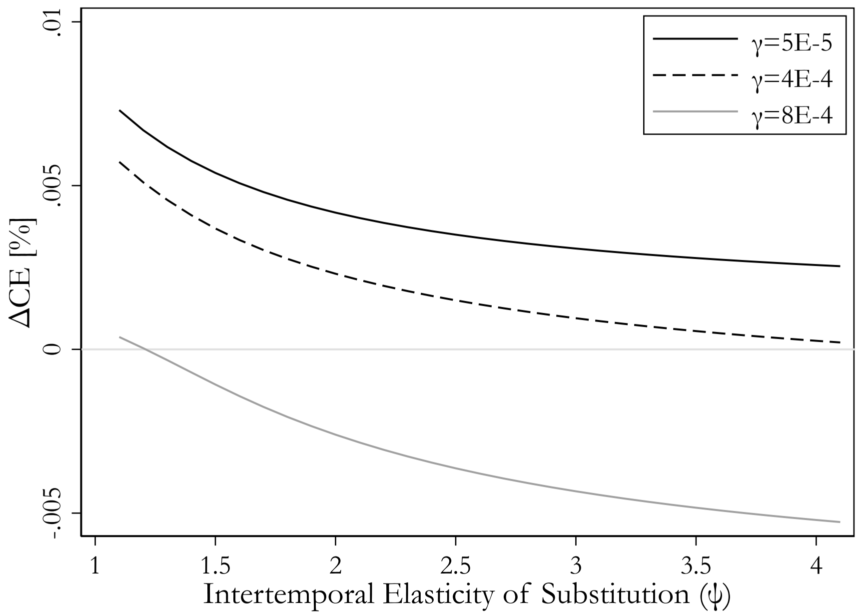


Figure 10: Difference in CE (GHLTI vs. GHHW) by Intertemporal Elasticity of Substitution

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

as a whole (see Section 2). In fact, using the representative SOEP, Table A8 (in Appendix A12) compares socio-demographics, health indicators, and healthcare usage between publicly and privately insured. Overall, we find that privately insured are less likely to smoke, have lower BMIs, and use fewer healthcare services. This section discusses how our results differ for a representative sample of Germans. To do so, we use claims data from one of the biggest public insurers, with more than 5 million enrollees, to compute the distribution of ACG scores among the publicly insured.

Figure A12 compares ACG risk score distributions across the two subpopulations; Table A9 compares average raw scores, and Table A10 shows the distribution by age over risk classes within the sample of publicly insured (all exhibits are in Appendix A12). Overall, the data confirm that the publicly insured are less healthy than the privately insured.[47]

Using these risk scores, Panel (i) of Table A11 shows welfare under the various contracts using the estimated initial probabilities at age 25 for the publicly insured. Approximating reality in Germany, Panel (j) uses a 90/10 mix of publicly and privately insured individuals. Overall, we find that the welfare loss under GLTHI compared to the optimal contract is between 0.9 and 1.4 percent for the population of publicly insured. This is substantially less than for the privately insured in our main

[47] However, when focusing on the 25-35 year olds and reweighting the data to match the age-gender distribution of our GLTHI policyholder sample, the risk scores look much more similar.

analysis. It is also consistent with our previous findings showing that the welfare loss decreases, the worse the initial health states.[48]

Income Profiles. Finally, to test the robustness of our results with respect to the income profile, we use the lifecycle income pattern for the United States estimated from the representative Panel Study of Income Dynamics (PSID).[49] We use the exact same income concept as in our main analysis for German income, 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 (8). Figure A14 in Appendix A14 shows the resulting income profiles. The increase in (post-tax equivalized) income in the U.S. 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. Still, our calculations show that our main findings are robust to U.S. income profiles: GLTHI contracts would achieve welfare that would fall 3.5 and 3.6 percent short of the optimal long-term health insurance contract for Americans with high school and college degrees, respectively.[50]

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

This section discusses possible implications of our findings for the health care reform debate in the United States. The health insurance system in the United States both shares some similarity with and differs substantially from the German system (see Section 2). Like the German system, the U.S. system is a mixture of public and private health insurance: Medicare—a pay-as-you-go public insurance system financed by payroll taxes—covers people above 65 (and the disabled). Yet, 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, are covered by Medicaid, or are uninsured (Claxton et al., 2017). ESHI is community-rated at the employer level and essentially long-term, provided that employers and employees do not separate, thus, it is arguable that it resembles the GLTHI.[51] Before the Affordable Care Act (ACA) was enacted into law in 2010,

[48] Another question is whether risk tolerance differs between GKV and PKV policyholders. Research has found that, at least in West Germany, there is evidence that civil servants are more risk averse than the rest of the population (Fuchs-Schündeln and Schündeln, 2005). This is confirmed by Table A8. Still, Figure A11 plots the risk tolerance distributions for the PKV and GKV populations using the SOEP, and shows that they are very similar. If anything, privately insured are a bit less risk averse, mostly stemming from more mass between 6 and 8 on the 1-10 Likert scale.

[49] The PSID is the oldest and longest-running panel survey in the world. From 1968 to 1996, it surveyed the U.S. families annually, and from 1997, biannually (Panel Study of Income Dynamics, 2018). We use the Cross National Equivalence Files (CNEF), which harmonizes survey measures across years (Frick et al., 2007).

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

[51] However, ESHI emphasizes risk pooling across workers of the same employer instead of intertemporal individual risk pooling as in the GLTHI. As a result, in contrast to long-term contracts, ESHI is subject to the dynamic inefficiency in the

the U.S. individual private health insurance market closely resembled the individually risk-rated short-term contract in Appendix A9.[52] Thus, as a first order approximation, the pre-ACA U.S. system was a mixture of 60 percent GLTHI and 40 percent short-term contracts for workers up to age 65; followed by public Medicare coverage for those 65 and older.

We ask the following questions: If we were to reform the U.S. health insurance system and replace all private health insurance contracts with GLTHI-style 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?

To answer these questions, we implement a public insurance program for people of age 65 and above, financed by a proportional payroll tax on income.[53] Note that the Medicare tax acts as an additional, front-loaded premium during working ages to fund free health insurance for all people 65 and above, 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 of their education risk pool during the Medicare period (age 65 and above), such that

$$\tau_{Ed}^* \mathbb{E}\left(\sum_{t=25}^{64} S_t \delta^{t-25} y_t | Ed\right) = \mathbb{E}\left(\sum_{t=65}^{94} S_t \delta^{t-25} m_t\right) \tag{11}$$

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* $T = 64$.[54]

incentives to invest in health (see Fang and Gavazza (2011), and can lead to job-lock (see Madrian (1994).

[52]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.

[53]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. We abstract away from a myriad of other effects of Medicare, particularly its impact on labor supply decisions.

[54]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 (13)); we assume that

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 4.

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

Education	Payroll Tax (%)	CE_{ST}	CE_{GLTHI}	CE_{HHW}
Panel (a): Private Contracts up to 64 + Medicare from 65 (Hybrid system)				
Ed 10	4.36	-11,124	20,320	20,706
Ed 13	3.12	-3,687	24,220	24,807
Panel (b): Life-long contracts (Baseline results)				
Ed 10	.	-10,119	21,168	21,945
Ed 13	.	-2,223	25,088	26,093

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 4. All consumption certainty equivalents (welfare) are in 2016 USD per capita, per year.

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 7, 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 maybe 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.[55] 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

incomes of individuals in education group *Ed 10* and *Ed 13* are taxed at the respective rate τ^*_{Ed} calculated by Equation (11).

[55]Interestingly, theoretically it is *ambiguous* whether the hybrid system or the private system achieves 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 dynamic contract need to address.

income groups. In both cases, the presence of tax-financed Medicare program for people above 65 also reduces welfare 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.Premium-free Medicare coverage at old-age increases the payroll tax rate needed to fund the entire program therefore increasing the degree of front-loading in working ages. 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 payroll tax rate. In Appendix A15, 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 payroll 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 environment. 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. 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.

Discussion. Any fundamental reform to the U.S. health insurance system is politically difficult. The German-style long-term private health insurance system may offer an attractive reform alternative. First, because competing private insurance companies—rather than the government—offer long-term health insurance contracts, it may be more appealing to a large fraction of American voters than a national health insurance system. Second, the public Medicaid system that currently covers about 20% of the U.S. population can, with expansion, serve as the basis of a public option when the U.S. further develops the private long-term insurance market (Fang and Krueger, 2022). Importantly, this section's findings that a hybrid system combining private long-term contracts for people below 64 with a Medicare-like pay-as-you-go system for those above 65 would result in substantial welfare gains relative to the *status quo* suggests a plausible transition to such a system. Third, not only have

German insurers offered long-term health insurance contracts for decades, but long-term insurance contracts have also been prominently featured in other settings in the United States, for example, in the life and long-term care insurance markets.

These attractive features notwithstanding, there are important differences between the U.S. and the German health insurance markets. For example, currently the U.S. does not have a universal public option for all working age people as Medicaid only covers low-income individuals; thus a tax-financed public option would need to significantly expand Medicaid, or to expand Medicare, or to offer an explicit public option on the ACA Health Insurance Exchanges. Second, as we mention in Section 2, in the United States, different insurers (or even plans) may have different provider networks, while the GLTHI is a pure financial contract. To ensure competition in the long-term private insurance market, insurance plans and provider networks would need to be decoupled to avoid inefficiencies from lock-in (e.g. Atal, 2016).

8 Conclusion

Pricing regulation in health insurance markets has to trade off reclassification risk, adverse selection, moral hazard and 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. Switzerland has traditionally organized its markets as short-term annual contracts with tight community-rating pricing regulation to provide reclassification risk insurance for all citizens; and the US has followed suite after the enactment of the ACA. 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. We show that the GLTHI contract almost fully eliminates reclassification risk over a policyholder's lifecycle. However, the low reclassification risk comes at the expense of high premium front-loading which results in limited intertemporal consumption smoothing. Second, we quantify the lifecycle welfare of the GLTHI contract and contrast it 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. Moreover, we show that under various parameterizations and scenarios, the GLTHI contract brings more than 95 percent of the welfare gain of the optimal contract, against a benchmark represented as a series of short-term contracts. For flat lifecycle income, as for the privately insured civil servants but unlike for high-income earners, the theoretical and the optimal contract coincide. However, for the average German lifecycle income—compared to the theoretically optimal health insurance contract as derived by Ghili et al. (2019)—the GLTHI contract entails less life-cycle consumption smoothing as it requires more front-loading, but it also almost entirely eliminates reclassification risk. In the GLTHI contract, the welfare loss due to less consumption smoothing is almost entirely offset by less reclassification risk. This finding is robust to allowing for private savings, to a wide range of the degrees of risk aversion, to non-time-separable recursive preferences à la Epstein and Zin (1989), to using the population of publicly insured and U.S. lifecycle income profiles. Further, the GLTHI contract provides large welfare gains relative to a series of risk-rated short-term contracts that were common in the pre-ACA era in the United States.

A practical advantage of the GLTHI contract—relative to the theoretically optimal contract—is that is does not rely on policyholders' income changes in premium setting. To the extent that incomes are endogenous, it thus avoids potential work disincentives. Market regulation is relatively simple as witnessed by the fact that the GLTHI market has been stable and has provided insurance for millions of Germans 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.

To further draw lessons for the reform debate in the United States, we evaluate a combination of long-term contracts and a Medicare-like pay-as-you-go system for people above 65. We show that such a hybrid system would result in substantial welfare gains relative to the *status quo*; however, we also find that such a hybrid system with GLTHI-style contracts up to age 64 followed by payroll-taxed financed Medicare for 65 and older is inferior to a genuine system of long-term contracts over the entire lifecycle as in Germany. Thus The German-style long-term private health insurance system may offer an attractive reform alternative for the U.S.

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 discussion abstracts from a couple of key features that may have 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.[56] In addition, our model abstracts from moral hazard. In the presence of moral hazard, using long-term contracts to protect against reclassification risks could also induce inefficiencies in health spending and health investment, similar to what is studied in Cole et al. (2019) for the case of community rating. Quantifying the role of moral hazard in long-term contracts 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 (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.

Association of German Private Healthcare Insurers (2022a). *PKV Zahlenportal*. Verband der Privaten Krankenversicherung. https://www.pkv-zahlenportal.de/werte/2010/2020/12, retrieved February 20, 2022.

Association of German Private Healthcare Insurers (2022b). *So funktioniert die Pflegeversicherung*. Verband der Privaten Krankenversicherung. https://www.pkv.de/wissen/pflegeversicherung/so-funktioniert-die-pflegeversicherung/, retrieved February 20, 2022.

Atal, J. (2016). Lock-in in dynamic health insurance contracts: Evidence from Chile. mimeo. https://sites.google.com/site/juanpabloatal/, last retrieved on September 7, 2019.

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.

[56] 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.

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

Bahnsen, L. and F. Wild (2021). Entwicklung der prämien- und beitragsanpassungen in pkv und gkv 2012-2022. Wip-kurzanalyse, Wissenschaftliches Institut der PKV. http://www.wip-pkv.de/forschungsbereiche/detail/entwicklung-der-praemien-und-beitragseinnahmen-in-pkv-und-gkv-2012-2022.html, retrieved February 10, 2022.

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 (2022). *Die Pflegeversicherung*. https://www.bundesgesundheitsministerium.de/themen/pflege/online-ratgeber-pflege/die-pflegeversicherung.html, retrieved on March 8, 2022.

Bundesministerium für Gesundheit (2022). Gesundheitsberichterstattung des Bundes. http://www.gbe-bund.de/, retrieved on February 6, 2022.

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.

Chiappori, P.-A. and M. Paiella (2011). Relative risk aversion is constant: Evidence from panel data. *Journal of the European Economic Association 9*(6), 1021–1052.

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.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association 9*(3), 522–550.

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.

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 D. Krueger (2022). The affordable care act after a decade: Its impact on the labor market and the macro economy. *Annual Review of Economics 14*, 453–494.

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.

Fuchs-Schündeln, N. and M. Schn̈deln (2005). Precautionary savings and self-selection: Evidence from the german reunification "experiment". *Quarterly Journal of Economics 120*(3), 1085–1120.

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.

GKV-Spitzenverband (2022). Die Gesetzlichen Krankenkassen. https://www.gkv-spitzenverband.de/krankenversicherung/kv_grundprinzipien/alle_gesetzlichen_krankenkassen/alle_gesetzlichen_krankenkassen.jsp, retrieved on February 6, 2022.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The german socio-economic panel (soep). *Jahrbücher für Nationalökonomie und Statistik 239*, 345–360.

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.

Harris, M. and B. Holmstrom (1982). A theory of wage dynamics. *The Review of Economic Studies 49*(3), 315–333.

Heckman, J. and B. Burton Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52 (2)*, 271–320.

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.

Henke, K. D. (2007). Zur dualität von gkv und pkv / the future of private and public health insurance in germany. *Jahrbücher für Nationalökonomie und Statistik 227*(5-6), 502–528.

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.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), 473–522.

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.

Madrian, B. (1994). Employment-based health insurance and job mobility: Is there evidence of job-lock? *The Quarterly Journal of Economics 109*(1), 27–54.

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

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.

Rosenbrock, S. (2010). *Einflussfaktoren für die Entscheidung bei Altersvorsorgeprodukten*.

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.

Schmidt, T. I. (2020). Die beihilfesysteme des bundes und der länder im vergleich. Rechtswissenschaftliches gutachten im auftrag des bundesministeriums fr̈ gesundheit, Bundesministerium für Gesundheit. https://www.bundesgesundheitsministerium.de/service/publikationen/details/vergleich-der-beihilfesysteme-der-bundeslaender-und-des-bundes-aus-rechtlicher-sicht.html, retrieved February 10, 2022.

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.

Statistisches Bundesamt (2020). Sozialleistungen - Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus) 2019. Fachserie 13 reihe 1.1.

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.

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