

German Long-Term Health Insurance: Theory Meets Evidence*

Juan Pablo Atal[†]

Hanming Fang[‡]

Martin Karlsson[§]

Nicolas R. Ziebarth[¶]

June 21, 2024

Abstract

German long-term health insurance (GLTHI) represents the largest market for private long-term health insurance contracts in the world. We show that GLTHI's contract design coincides with the optimal dynamic contract for individuals with constant lifetime income profiles. After estimating the key ingredients of a life-cycle model, we find that, under a variety of assumptions, GLTHI achieves welfare that is at most 4% lower than for the optimal contract. Relative to the gains of replacing short-term contracts with either long-term contract, this welfare gap is smallest when reclassification risk is high.

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, Georgetown 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, and the NBER, 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

[¶]ZEW Mannheim and University of Mannheim, L7, 1, D-68161 Mannheim, Germany Email: nicolas.ziebarth@zew.de

1 Introduction

Short-term health insurance contracts can expose policyholders to potentially large premium fluctuations. Such “reclassification risk” can lead to significant welfare losses (Diamond, 1992; Cochrane, 1995). The standard policies, most notably 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 can generate adverse selection, which is generally addressed through individual mandates, premium subsidies, or both (cf. Akerlof, 1970).¹

Long-term (or *dynamic*) health insurance contracts can provide reclassification risk insurance without triggering adverse selection and without the need for mandates or subsidies. Instead, they leverage individuals’ intertemporal incentives to insure reclassification risk with a schedule of “front-loaded” premiums (Pauly et al., 1995, 1999; Patel and Pauly, 2002; Pauly and Lieberthal, 2008). Although frontloading hinders intertemporal consumption smoothing, long-term contracts that optimally balance those detrimental effects with reclassification risk insurance can generate large welfare gains relative to short-term contracts (Ghili et al., 2024).² However, despite their theoretical appeal, little is known about real-world applications of long-term contracting in health insurance, and how their implementation relates to the theoretical prescriptions.

In this paper, we study the most prominent real-world application of dynamic health insurance contracts: German Long-Term Health Insurance (GLTHI). We provide the following main contributions to the literature. First, we show that, despite their simple design, existing GLTHI contracts share several features with the optimal dynamic contract characterized by Ghili et al. (2024). Second, we estimate the key ingredients of a life-cycle model to evaluate welfare under the GLTHI contracts and compare it to welfare under the optimal contracts. Finally, we highlight the general circumstances under which the performance of optimal contracts is robust to the simplifications of the German design.

Several features make the GLTHI contracts especially appealing to study. First, Germany has the largest and oldest individual private long-term health insurance market in existence.³ Second, contracts in this market are purely financial—with no differentiation in the provider network across

¹For example, the Affordable Care Act (ACA), enacted in the United States in 2010, established health insurance exchanges for non-group plans featuring community rating, an individual mandate, and premium subsidies as its three main pillars (Aizawa and Fang, 2020). The ACA originally introduced an individual “mandate” by imposing a tax penalty for people without insurance (with some exceptions). This was the subject of many legal disputes and controversies. The tax penalty was then set to zero in 2019.

²Naturally, Ghili et al. (2024) study a context where individuals are free to lapse their policies (“one-sided commitment”), and there are borrowing constraints.

³Its roots go back to 1883 and currently it covers almost 9 million individuals. In addition to Germany, the Chilean private health insurance market is the only other private market with comprehensive long-term health insurance.

insurers or plans. The absence of networks makes it well suited for a comparison with the theory of optimal dynamic pricing in health insurance (which also abstracts away from such differentiation across plans). Third, after an initial risk rating, contracts follow the simple principle of guaranteed renewability with constant and income-independent premiums as long as the policyholder decides to renew (Hofmann and Browne, 2013). This simple design is particularly appealing from a policy perspective.

We first present the main principles of GLTHI and show that its simple contract design coincides with the optimal dynamic contracts derived by Ghili et al. (2024) when the policyholder's income profile is flat over the lifecycle. We also use data on premiums to corroborate that the theoretically optimal contract for flat lifecycle incomes matches the observed GLTHI premiums very well. As a consequence, for the typical policyholders with an increasing lifecycle income profile, the simple GLTHI design, relative to that of the optimal dynamic contract, provides more reclassification risk insurance and charges higher premiums early in life, resulting in sub-optimal consumption smoothing over the life-cycle.

We find that welfare under the GLTHI design, as measured by certainty-equivalent yearly consumption, is close to 96 percent of what the optimal contract would achieve. To do so, we first estimate the key ingredients of a model of health expenditures and income over the lifecycle by combining a unique panel of claims data with 33 years of income data from Germany's representative Socio-Economic Panel (SOEP). As a critical component of our analysis, we propose and implement a novel health risk classification method to parsimoniously model individuals' expected health risks and their evolution over the lifecycle, while recovering relevant features of typical health expenditure distributions.

Our welfare results are robust to multiple alternative assumptions. We find welfare differences of a similar magnitude if we consider (i) a wide range of risk aversion, (ii) different functional forms for utility, including Epstein-Zin recursive preferences, (iii) private savings, (iv) an initial health status that matches the German population as a whole (instead of the population of the privately insured), and (v) US income profiles.

We finish by showing that the extent to which the GLTHI design recoups the welfare gains from the optimal dynamic contracts (replacing short-term contracts) critically depends on the degree of the reclassification risk. When policyholders face large reclassification risk, the GLTHI design provides most of the gains from long-term contracting despite being sub-optimal, as its excessive frontloading is largely compensated by its higher insurance against reclassification risk. It is important to note that environments with high reclassification risk are also the cases where long-term contracts are

particularly effective, as the welfare gains from reducing catastrophic reclassification risk are large relative to the welfare losses from premium frontloading. In contrast, in contexts with limited catastrophic losses, due to, e.g., government social safety net programs, the GLTHI contract would recoup a smaller fraction of the gains from optimal long-term contracting.

Our results have policy implications for the US and elsewhere. Our analysis of Germany’s experience shows the effectiveness of a real-world application of long-term contracts in health insurance. It also shows that the simple “flat-income” version of optimal long-term contracts implemented in Germany—a single long-term contract with constant premium guarantees—can recoup much of the overall potential gains of optimal long-term contracting, especially with large reclassification risk. This result strengthens their case as an appealing policy option, particularly in contexts where reclassification risk is not already insured. We see this as a vindication of the virtue of the theory of dynamic contracting in health insurance.

This paper contributes to several strands of literature. First, it contributes to the literature on dynamic contracts, for which there is vast theoretical work, but relatively little empirical evidence. [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 between the insured and the insurer, made after the realization of health shocks, which provides full insurance against reclassification risk. [Harris and Holmstrom \(1982\)](#) and [Krueger and Uhlig \(2006\)](#) study the properties of competitive long-term contracts that insure agents against income risk under one-sided commitment when agents receive endogenous outside offers. [Hendel and Lizzeri \(2003\)](#) and [Ghili et al. \(2024\)](#) show that the optimal dynamic insurance contract only partially insures reclassification risk because reducing it requires frontloaded premium payments that hinder consumption smoothing over the lifecycle. Importantly, we build heavily on [Ghili et al. \(2024\)](#), as we use their characterization of the optimal long-term health insurance contract over the lifecycle and perform similar welfare comparisons. One crucial addition is to show how GLTHI, the most important real-world application of long-term health insurance contracts, corresponds to a version of their theoretically optimal contract—the one that would emerge under flat income profiles—and to quantify the welfare gap from such a sub-optimal implementation of their theory. We also characterize and quantify the one-period state-contingent Arrow Securities that can be used to implement the GLTHI as well as the optimal contract ([Krueger and Uhlig, 2006](#)). Arrow securities have the advantage of fostering insurer competition by overcoming the detrimental lock-in effects of long-term contracts.⁴

⁴We abstract from the effects of long-term health insurance on health transitions in this paper. See [Cole et al. \(2019\)](#) for a quantitative dynamic model of health investments and insurance that studies the short -and long-term effects of

Second, several papers, including [Hendel and Lizzeri \(2003\)](#), [Herring and Pauly \(2006\)](#), [Finkelstein et al. \(2005\)](#), [Fleitas et al. \(2018\)](#), and [Atal \(2019\)](#), empirically investigate the workings of long-term contracts in different contexts. Our paper contributes to this literature by contrasting the empirical applications of health insurance contracts with their optimal design throughout the lifecycle. We also introduce a novel data-driven method for classifying health risks. We base our method on an explicit objective function proposed in the actuarial science literature that maximizes the differentiation between risk classes, invoking the properties of *homogeneity* and *separation* in risk classification (see [Finger, 2006](#)). By better capturing the skewness of health expenditures, our method improves on the *ad hoc* methods for discretizing health risks commonly used in the literature. Critically, by accurately modeling the right tail of health expenditure distributions, we highlight the workings of long-term contracts in insuring individuals against catastrophic risks.

Finally, our paper relates to previous work on the German long-term health insurance market. [Christiansen et al. \(2016\)](#) empirically study the determinants of lapsing and switching behavior. [Baumann et al. \(2008\)](#) and [Eekhoff et al. \(2006\)](#) discuss the potential consequences of higher switching rates on market competition if the capital accumulated through frontloaded payments (“old-age provisions”) were to be made portable across insurers. [Avdic et al. \(2024\)](#) study how cost-sharing in GLTHI affects utilization, and [Atal et al. \(2019\)](#) theoretically and empirically study the effects of the 2009 portability reform in Germany on switching behavior. Closest to our work, [Hofmann and Browne \(2013\)](#) also describe the actuarial principles of premium calculation in GLTHI. They document that contracts involve premium frontloading generating consumer lock-in, that more frontloading is generally associated with lower lapsation rates, and show that low-risk individuals are more likely to lapse. They link these findings with the predictions of the two-period model of optimal dynamic insurance in [Hendel and Lizzeri \(2003\)](#). We explicitly compare the features of the GLTHI contracts with those of the optimal long-term insurance contract over the lifecycle as characterized by [Ghili et al. \(2024\)](#), and show that the two contracts coincide under flat lifecycle income profiles. We also evaluate the welfare consequences of GLTHI contracts.

2 Institutional Background

Germany has a two-tier health insurance system. Since 2009, Germany has an individual mandate that every resident must have health insurance (§193 *Versicherungsvertragsgesetz*).⁵ The first

community-rated health insurance.

⁵Despite the two tiers, the uninsurance rate is low at around 0.1 percent ([Statistisches Bundesamt, 2020](#)). Further institutional details on the two-tier German health insurance system are in [Schmitz and Ziebarth \(2017\)](#) (English) or [Henke \(2007\)](#) (German).

tier consists of public health insurance (*Gesetzliche Krankenversicherung, GKV*) and covers 90 percent of the German population. Every employee whose income is below the politically defined federal threshold (*Versicherungspflichtgrenze*) of € 69,300 in 2024 (about \$74K), is *mandatorily* insured with the public multi-payer system. GKV policyholders can choose between 97 private non-profit sickness funds (*Gesetzliche Krankenkassen*) (Bünnings et al., 2019; GKV-Spitzenverband, 2024). Premiums consist of income-dependent contribution rates (employees and employers each pay about eight percent of the gross wage, up to a cap) for a standardized benefit package with very little cost-sharing. Germany was the first country in the world to introduce such a public health insurance system in 1883 (*Gesetz betreffend der Krankenversicherung der Arbeiter*), whose basic structure has remained intact since its inception.

For historical reasons, select population subgroups can *permanently* opt out of the GKV system and join the second tier: private health insurance (*Private Krankenversicherung, PKV*), which covers the remaining ten percent of the population with about 8.7 million policyholders (Koch, 1971; Association of German Private Healthcare Insurers, 2023). The PKV covers three main population subgroups: the (a) self-employed; (b) high-income earners with gross incomes above the *Versicherungspflichtgrenze*, and (c) civil servants. They can decide to leave the public GKV system permanently to join the PKV and sign a private individual long-term health insurance contract for the rest of their lives (Hullegie and Klein, 2010; Polyakova, 2016; Panthöfer, 2016).^{6, 7} 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 private[ly insured], always private[ly insured]” (Schencking, 1999; Jacobs et al., 2012; Innungskrankenkasse Berlin Brandenburg, 2018). However, in practice, there are (costly) possibilities to switch back to the GKV for people under 55, for example, if one becomes unemployed or (permanently) switches to a private sector job with income below the *Versicherungspflichtgrenze* (see §6, Social Code Book VI).⁸ Note that catastrophic

⁶Civil servants represent a special case as they receive about 50% of their insurance coverage from the states or federal government as their employer—and typically purchase PKV contracts for the remaining part.

⁷In 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, 2024). (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 frontloaded. 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, 2024; Association of German Private Healthcare Insurers, 2024b). 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 made the old-age reserves of LTC insurance fully portable when switching to another insurer, which was not the case for GLTHI in our study period. While interactions between LTC and traditional health insurance could be relevant—for example, one’s LTC needs may be affected by the healthcare she receives and vice versa—modeling these interactions is beyond the scope of this paper.

⁸Note that deliberate short-term income manipulation is not allowed (Krankenkasse, 2024). Also note that it is impossible

health care costs are not an acceptable reason to switch back. Moreover, GLTHI insures reclassification risk and the regulator sets a maximum annual out-of-pocket cost-sharing cap of €5,000 (see §193, Para. 3 *Versicherungsvertragsgesetz*). In practice, cost-sharing is much lower; in our data the mean deductible is €675 and the highest deductible €3,224 (Table A1). Appendix A1 discusses the institutional specifics and the empirical evidence on the general lack of switching from PKV to GKV.

Besides Chile (cf. Atal, 2019), to our knowledge, Germany is the only country in the world with an existing private individual long-term health insurance market. The German Long-Term Health Insurance (GLTHI) market consists of 48 private insurers that sell *comprehensive* as well as *supplemental* coverage (Association of German Private Healthcare Insurers, 2020). The focus of this paper is comprehensive coverage.⁹

Provider Networks. Provider networks and “Managed Care” are largely unknown in both GKV and PKV; that is, in either system, enrollees have the free choice of providers.¹⁰ Moreover, in both systems, reimbursement rates are centrally set and do not vary by insurers or health plans. Although reimbursement rates for inpatient care are identical in both systems, they are approximately twice as high for outpatient care in the PKV. These structurally higher outpatient reimbursement rates result in significantly shorter waiting times for the privately insured (cf. Werbeck et al., 2021).¹¹ Because they do not negotiate rates or build provider networks, private insurers mainly customize health plans and process, scrutinize, and deny claims. Thus, the GLTHI contract mainly constitutes a pure financial contract similar to policies in insurance markets such as life insurance (Fang and Kung, 2020). This specific feature substantially simplifies the welfare analysis of GLTHI contracts and its comparison with the optimal long-term contracts.¹²

One-Sided Commitment and Guaranteed Renewability. When individuals apply for a GLTHI policy, 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

for people above 55 to switch back.

⁹While 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 had 883,400 covered employees for the whole of Germany in 2019 (Association of German Private Healthcare Insurers, 2020).

¹⁰Since 2009, in GKV, sickness funds have to offer select managed care programs such as *Hausarzttarife* where primary care physicians act as gatekeepers. However, these programs are entirely voluntary and only a very small share of GKV insured enrollees in them, see also §73b SGB V and *Gesetz zur Weiterentwicklung der Organisationsstrukturen in der GKV*.

¹¹Other unintended consequences could be overtreatment of the privately insured and explicit discrimination such as separate practice waiting rooms or telephone numbers for GLTHI policyholders. While privately insured individuals customize their cost-sharing, the publicly insured pay precisely defined co-payments that are uniform across sickness funds and set at the federal level, e.g. €10 per night in a hospital, see §61, Social Code Book V.

¹²Compared to public insurance, one could argue that private markets and contracts are less prone to government regulatory risk. Kojien 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.”

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 GLTHI contract until they die: Medicare does not exist in Germany.

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 premium guarantee as long as the individual remains in the contract, as explained in [Hofmann and Browne \(2013\)](#). Since the principles of premium calculation represent a central aspect of our analysis, [Section 3](#) discusses this issue. [Appendix A2](#) provides additional details on premiums, actuarial principles, and regulatory oversight.

3 Lifecycle Premiums in GLTHI

3.1 Formal Derivation of GLTHI Lifecycle Premiums

This section provides a formal treatment of the main principles guiding GLTHI lifecycle premiums: risk-rated initial premiums followed by guaranteed renewability at the same premium as long as the individual stays in the contract. [Appendix A2](#) provides a graphical representation and additional details.¹³

Let $P_t(\zeta_t)$ be the initial premium when first signing a GLTHI contract in period t . As GLTHI contracts are individually underwritten at inception, $P_t(\zeta_t)$ depends on individuals' health status in year t , ζ_t . We assume that $\zeta_t \in \Xi$, where Ξ is a finite set of health states. In subsequent periods, each contract is guaranteed-renewable at the same premium. As such, individuals who sign a contract in period t can renew the contract for the same premium, $P_t(\zeta_t)$, in all periods between $t + 1$ and T , regardless of the evolution of their health status. However, they can lapse the existing contract and sign a new one if a new contract is available at a lower premium.

The contract breaks even in equilibrium, given premium $P_t(\zeta_t)$ and the subsequent lapsation decisions of the enrollee.¹⁴ Thus, the equilibrium levels of $P_t(\zeta_t)$ can be expressed as the solution to a fixed-point problem in which $P_t(\zeta_t)$ covers exactly the expected claims of policyholders who *stay* in the contract at premium $P_t(\zeta_t)$.

¹³ [Hofmann and Browne \(2013\)](#) also explain the guiding principles of premium calculation in GLTHI. We complement [Hofmann and Browne \(2013\)](#) by deriving an explicit formula for how premiums depend on the evolution of health, which we use in the calculations below. This explicit formula enables us to directly compare the GLTHI lifecycle premium calculation with those in the optimal dynamic contract as derived in [Ghili et al. \(2024\)](#).

¹⁴ This is known in the German insurance literature as the *principle of equivalence* (see e.g. [Hofmann and Browne, 2013](#)).

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 status 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 policyholders 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 $P_\tau(\xi_\tau) < P_t(\xi_t)$.¹⁵

Remark 1 *The lapsation decision under GLTHI is only driven by a comparison between the current guaranteed renewable premium $P_t(\xi_t)$, and the premium of a new, alternative, 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, differences in provider networks do not drive the lapsation decision.*

For a given $t < T$ and $\tau > t$, we denote $\mathbf{P}_{t+1}^\tau \equiv \{P_{t+1}(\cdot), \dots, P_\tau(\cdot)\}$ 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 policyholders in period t with health state $\xi_t \in \Xi$, denoted by $P_t(\xi_t)$, as follows:

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

where the first element of the numerator, $\mathbb{E}(m_t|\xi_t)$, is expected health care 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 $q_\tau(z|\xi_t, \mathbf{P}_{t+1}^\tau, P_t(\xi_t))$, the probability that (i) $\xi_\tau = z$ and (ii) the policyholder does not lapse (or die) between periods t and τ , given the subsequent equilibrium premiums \mathbf{P}_{t+1}^τ . These expected lifecycle expenditures are then divided by the expected number of years in the contract.¹⁶ In other words, the GLTHI lifecycle premium offered in period t to an individual in state ξ_t , $P_t(\xi_t)$, equals the annualized expected health care expenditures starting from t and while the individual remains in the contract, conditional on the individual's health status at the inception of the contract.

¹⁵Note 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.

¹⁶Of course, $q_\tau(z|\xi_t, \mathbf{P}_{t+1}^\tau, P_t(\xi_t))$ depends on the evolution of the health status $\xi_{t+1}, \dots, \xi_\tau$ and death, conditional on current health status ξ_t . We describe how we model the evolution of health status and its implications for health expenditures in Section 5.

Note that these lifecycle premiums do not maximize any *ex ante* consumer objective function; 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.*

3.2 Comparison to the Optimal Dynamic Contract

Under the assumption of risk-averse and time-separable preferences for policyholders, one-sided commitment by insurers, and both having identical discount rates, the optimal dynamic health insurance contract as derived by Ghili et al. (2024) 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 $\mathbf{y}_t^T \equiv \{y_t, y_{t+1}, \dots, y_T\}$.

Analogous to the GLTHI lifecycle premium calculation, $\bar{c}_t(\xi_t, \mathbf{y}_t^T)$ can be solved by backwards induction. Specifically, the consumption guarantee in period T is given by $\bar{c}_T(\xi_T, y_T) = y_T - \mathbb{E}(m_T | \xi_T)$. For any $t < T$ and $\tau > t$, denote the set of future equilibrium consumption guarantees $\bar{\mathbf{c}}_{t+1}^\tau \equiv \{\bar{c}_{t+1}(\cdot), \dots, \bar{c}_\tau(\cdot)\}$. An algebraic reformulation of Lemma (2) in Ghili et al. (2024) implies that the equilibrium break-even consumption guarantee under the optimal dynamic contract for an individual purchasing an optimal long-term contract at time t under health status ξ_t is recursively determined by:

$$\bar{c}_t(\xi_t, \mathbf{y}_t^T) = \frac{y_t - \mathbb{E}(m_t | \xi_t) + \sum_{\tau > t} \sum_{z \in \Xi} \delta^{\tau-t} (y_\tau - \mathbb{E}(m_\tau | z)) \times q_\tau(z | \xi_t, \bar{\mathbf{c}}_{t+1}^\tau, \bar{c}_t(\xi_t, \mathbf{y}_t^T))}{1 + \sum_{\tau > t} \sum_{z \in \Xi} \delta^{\tau-t} \times q_\tau(z | \xi_t, \bar{\mathbf{c}}_{t+1}^\tau, \bar{c}_t(\xi_t, \mathbf{y}_t^T))}, \quad (2)$$

where $q_\tau(z | \xi_t, \bar{\mathbf{c}}_{t+1}^\tau, \bar{c}_t(\xi_t, \mathbf{y}_t^T))$ is, with some slight abuse of notation, the probability that (i) $\xi_\tau = z$, and (ii) the individual does not lapse (or die) between periods t and τ , given the set of future equilibrium consumption guarantees $\bar{\mathbf{c}}_{t+1}^\tau$.¹⁷

The optimal and the GLTHI contract have several commonalities. Both contracts break even in expectation. Both contracts entail frontloading to limit (but not eliminate) reclassification risk. Both contracts are risk-rated at inception and provide some form of renewable guarantee. While the

¹⁷ 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 the intertemporal preference is time separable. Section 6.4.1 discusses the case of non-time-separable preferences where the contract as characterized by Ghili et al. (2024) is no longer optimal. It is also important to note that, as shown in Krueger and Uhlig (2006), the optimal contract does depend on preferences if the discount rates differ between insurers and policyholders. They also show that the optimal contract in such case does not take the form of consumption guarantees. Investigating the welfare consequences of alternative contracts when the insured and the insurer differ in their discount factors is an interesting avenue for future research.

GLTHI establishes a constant *premium* that *decreases* when individuals get a better premium in the market, the optimal contract offers a constant *consumption guarantee* that *increases* when individuals get a better consumption guarantee in the market.¹⁸ Moreover, comparing equations (1) with (2), we find that under flat incomes over the lifecycle ($y_t = \bar{y} \forall t$), the premium guarantees under the GLTHI result in the optimal consumption guarantees (which, in the case of flat income over the lifecycle, are independent of the income level). We highlight this important result in the following Remark:

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

3.3 Arrow Securities Implementation

Krueger and Uhlig (2006) note that the optimal dynamic long-term contract with one-sided commitment can be implemented through the trade of 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 created by the lock-in effect embedded in long-term contracts. A main drawback is the added complexity relative to guaranteed premium (or consumption) profiles under long-term contracts.

The following Lemma characterizes the quantity of securities traded to implement the GLTHI premium path:

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 $\Xi_t \equiv (\xi_1, \xi_2, \dots, \xi_t)$ that pay one dollar in state ξ_{t+1} is equal to*

$$b_t(\xi_{t+1} | \Xi_t) = \begin{cases} 0 & \text{if } P_{t+1}(\xi_{t+1}) < \tilde{P}_t(\Xi_t) \\ E(m_{t+1} | \xi_{t+1}) - \tilde{P}_t(\Xi_t) + \\ \sum_{\tau > t+1}^T \sum_{z \in \Xi} \delta^{\tau-(t+1)} [E(m_\tau | z_\tau) - \tilde{P}_t(\Xi_t)] q_\tau(z | \xi_{t+1}, \mathbf{P}_{t+2}^\tau, \tilde{P}_t(\Xi_t)) & \text{otherwise,} \end{cases} \quad (3)$$

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

¹⁸We note that guaranteed premium in the German system decreases when consumers switch to a contract with a lower premium guarantee. Although Ghili et al. (2024) derive the optimal contract imposing “no-lapsation” constraints, they also note that the same outcome could be achieved by contracts that specify a guaranteed “premium path” that provides a constant consumption floor as long as individuals decide to renew. As such, equivalent to GLTHI, increases in the consumption guarantee under the optimal contract would occur when individuals lapse and sign a contract with a new insurer that offers a higher consumption guarantee (via a lower premium).

the subsequent equilibrium premiums P_{t+2}^τ and $\tilde{P}_t(\Xi_t)$.

Proof 1 See Appendix A3.

Lemma 1 states that the Arrow securities to replicate the GLTHI premiums pay either nothing (in the event of lapsing), or the net present value of the difference between the guaranteed-renewable GLTHI premium and the premium of a short-term contract (in the event of not lapsing). Appendix A3.1 provides an analogous expression for the Arrow securities that replicate the optimal dynamic contract in Ghili et al. (2024). Appendix A7 makes use of Lemma 1 and our data to calculate the quantity of securities needed to replicate the GLTHI contract and the optimal dynamic contracts.

4 Claims and Survey Panel Data from Germany

This section describes the GLTHI claims panel data and the SOEP survey panel data. We use the claims data to estimate individual health transitions and expenditures over the lifecycle. We use the SOEP data to estimate individual income dynamics over the lifecycle. The evolution of health expenditures and income paths over the life-cycle are the key inputs to assess welfare under the different health insurance contracts with analyze.¹⁹

4.1 GLTHI Claims Panel Data

The claims panel data 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 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.

Sample Selection. We focus on primary policyholders. In other words, we disregard children and individuals under 25 years of age (555,690 enrollee-year observations).²⁰ Moreover, due to the 2009 portability reform (see footnote A63), we disregard inflows after 2008 (253,325 enrollee-year observations). Our final sample consists of 1,867,465 enrollee-year observations from 362,783 individuals.

¹⁹In one extension in Section 6.4.1, we employ a claims dataset from one of the biggest German *public* insurers to assess the robustness of our results to using alternative initial health distributions. Further, we use the Panel Study of Income Dynamics (PSID) survey to assess the robustness of our results to using lifecycle income paths for the United States.

²⁰The 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.

Descriptive Statistics. Table A1 (Appendix) shows that 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. They have been enrolled in their current health plan for 7 years. Ten percent of all policyholders remain clients for more than 28 years, with one policyholder staying for as long as 86 years, illustrating the existence of a truly *long-term* private health insurance system.²¹ Figure A2 shows the distribution of policyholders' age at contract inception. The majority of individuals sign their first GLTHI contract at around 30 years old, an age 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.²² The average *deductible* is \$675 per year.

In terms of benefits covered, we simplify the rich data and focus on the plan generosity indicator provided by the insurer. It classifies 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 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 (SOEP)

The German Socio-Economic Panel (SOEP) is a representative longitudinal survey. Since 1984, it annually surveys about 10,000 households and 20,000 individuals above the age of 17 (Goebel et al., 2019). We use SOEPlong and all existing waves from 1984 to 2016, to fully exploit the lifecycle dimension of this panel (SOEP, 2018).

²¹Our 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.

²²Employers cover roughly one half of the total premium and the self-employed pay the full premium.

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.²³ Then, the total annual post-tax household income is divided by the number of household members using the modified OECD equivalence scale.²⁴

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) provides summary statistics for our SOEP sample. From 1984 to 2016, the average annual income per household member was \$26,433 (in 2016 USD). Note that this measure has positive values for *all* respondents.

In the SOEP sample, the average age is 47, and 52 percent are female. About 27 percent are white collar workers, six percent are self-employed, and four 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 premiums and welfare of short- and long-term insurance contracts. We rely on insights from actuarial science to produce an “efficient” risk classification. Our method aims to improve upon the state-of-the-art literature by better representing skewed health expenditure distributions. We present below the features of our method and relegate the details to Appendix A5.

Following the economics literature (Einav et al., 2013; Handel et al., 2015; Ghili et al., 2024), we begin by constructing an individual measure for expected health care costs using the German version of the John Hopkins ACG software, which is routinely used by insurers for underwriting. For each individual, the ACG software provides a *continuous risk score* λ_t^* which represents the expected health

²³The SOEP group also generates and provides these single components in a time-consistent manner.

²⁴The 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.

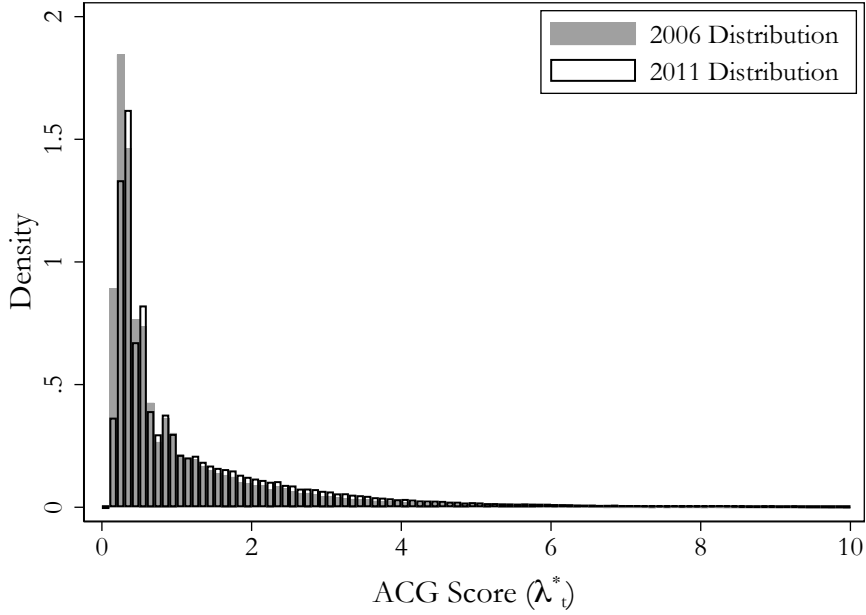


Figure 1: Distribution of λ_t^* in 2006 and 2011

Notes: The figure shows the distribution of ACG scores for years 2006 and 2011. The distribution of λ_t^* is truncated at 10; 0.7 percent of the sample have $\lambda_t^* > 10$. Source: Own calculations using GLTHI claims data and ACG Software.

care costs in year t relative to the mean in the reference population.²⁵ It is based on (a) diagnosis codes (pre-existing conditions and claim diagnoses), (b) costs of treatments, (c) treatment episode dates, and (d) age and gender.

Figure 1 shows the empirical distributions of λ_t^* for our GLTHI claims data in 2006 (the first year in our sample) and 2011 (the last year in our sample). Both distributions are approximately unimodal, and appear stable over time.²⁶ It also illustrates that the distribution of λ_t^* is heavily skewed and has a long right tail. For example, the top percentile of λ^* 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 1. The distribution of expected expenditures thus mirrors the characteristic long right tail of health expenditure distributions (French and Kelly, 2016).

Our method combines the continuous score λ_t^* and its lags, mapping them into K different *discrete risk categories* λ_t . Modeling risk types with discrete categories serves two specific purposes. First, as we allow the contract premiums to depend on the risk type, 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 1 implies that

²⁵In this case, the reference population are the publicly insured individuals in Germany

²⁶This also suggests that excluding inflows of new enrollees in 2010 and 2011 due to the portability reform (see Section 2) poses no major issue.

the amount of reclassification risk will strongly depend on the granularity of the risk classification.

Our main methodological contribution to risk classification modeling is in how to discretize health risk. The commonly-used approach would use an ad-hoc criterion to partition λ_t^* into different discrete categories. We depart from the common approach in two key ways: First, we allow the health status to be a function of current and n lagged values of λ_t^* , where the number of lags n is determined 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 $\{\lambda_{t-v}^*\}_{v=0}^n$ into an endogenously determined number of risk categories K . The method maximizes an efficiency criterion from the actuarial science literature (cf. [Finger, 2001](#)).

We split the task of constructing discrete risk categories into two sequential problems: (1) For a given number of categories K , define the *efficient partition* of the scores vector $\{\lambda_{t-v}^*\}_{v=0}^n$ into K discrete categories. (2) Select the parameters K and n that lead to the best performance of the classification system. We expand on each step below.

Efficient Partitioning. According to the actuarial science literature ([Finger, 2001](#)), an efficient risk classification system has two properties: *homogeneity*—meaning that individuals in the same risk category have similar risk; and *separation*—meaning that the categories have sufficiently different expected claims to justify distinct categories. For instance, in [Figure 1](#), 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. [Appendix A5](#) shows that applying k -means clustering to $\mathbb{E}\left(m_t \mid \{\lambda_{t-v}^*\}_{v=0}^n\right)$ —i.e., to the mean claims by an individual’s ACG scores—yields an efficient classification. In the special case of no lags ($n = 0$), the clustering procedure may target the ACG score directly, as it represents an estimate of $\mathbb{E}(m_t \mid \cdot)$. When lags are included ($n > 0$), $\mathbb{E}(m_t \mid \cdot)$ must be estimated in an intermediate step; [Appendix A5](#) outlines this process.

Parameter Selection. Next, we determine the number of lags n of ACG scores when computing the number of risk categories K .²⁷ k -means clustering is an unsupervised learning method; we assume that the overall objective applies also when selecting K and n . As shown in [Appendix A5](#), this implies using the R^2 of a regression of expenditures on risk category indicators as our criterion for model selection.

²⁷Including 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).

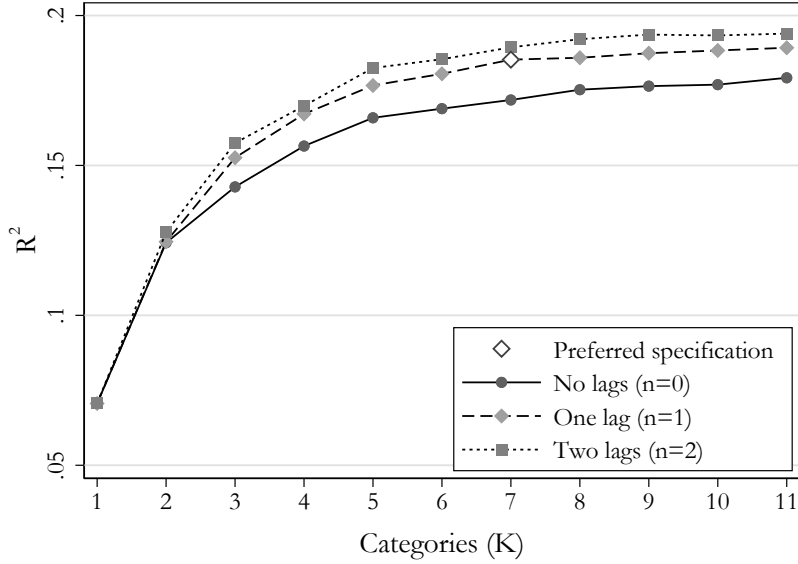


Figure 2: Performance of Alternative Risk Classifications.

Note: The figure displays *unadjusted* R^2 of a regression of expenditures on risk category indicators. All results are robust to using the *adjusted* R^2 . Each specification includes 21 age-gender fixed effects, year fixed effects and 79 plan fixed effects. Source: Own Calculations using German Claims Panel Data.

Figure 2 shows how the performance depends on K and n . For all values of n , initially, the predictive power improves rapidly when we increase the number of categories K ; however, this improvement levels out at around $K = 5$. Moreover, compared to only using the previous year's score ($n = 1$), there is distinct improvement when adding one lag ($n = 2$) whenever $K \geq 3$. However, adding a second lag only marginally improves the predictive accuracy. Figure 2 shows that, beyond including one lag ($n = 2$) and 7 distinct risk categories, 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 of Heckman and Burton Singer, 1984; Keane and Wolpin, 1997, in their choice of the number of unobserved types in the labor economics literature).²⁸

5.2 Estimation of Transition Matrices and Mean Expenditures

We posit that the health status of individual i at age t , $\zeta_{it} \equiv (A_{it}, \lambda_{it})$, depends on contemporaneous risk λ_{it} and age A_{it} , where A_{it} is one of eleven age groups (five-year bands from age 25 to age 75 and 75+).²⁹

As our risk classification generates risk categories of very different sizes, we use a parametric,

²⁸Section A5 provides additional robustness checks, including the role of outliers, including a longer history of claims, and 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. (2024).

²⁹Note 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.

yet flexible, model for transition rates between discrete risk categories λ_t and for mean expenditures by risk category and age. Specifically, to estimate the transition matrices for health dynamics, we estimate a multinomial logit model:

$$\eta_{it}^j = A_{it}\beta_j + L_{it}\gamma_j + h(A_{it}, L_{it}; \theta_j) + \epsilon_{it}^j \quad (4)$$

where η_{it}^j represents the log odds for $\lambda_{it+1} = j$, for $j \in \{2, \dots, 8\}$ relative to the reference category $\lambda_{it+1} = 1$ ($\lambda_{it+1} = 8$ represents death); A_{it} represents i 's age groups, and L_{it} is a set of indicators for the categories of λ_{it} . $h(A_{it}, L_{it}; \theta_j)$ consists of pairwise interactions of A_{it} and L_{it} with the associated parameter vector θ_j .³⁰

To model the expected claims based on risk category, we use predicted values from an OLS regression. In addition to the controls in Equation (4), we also control for Q_{it} representing health plan generosity dummies $q \in \{ECO, PLUS, TOP\}$. The base specification is:

$$m_{it} = A_{it}\beta + L_{it}\gamma + Q_{it}\delta + h(A_{it}, L_{it}, Q_{it}; \theta) + \epsilon_{it} \quad (5)$$

In an iterative process, we add pairwise interaction terms between A_{it} , L_{it} , and Q_{it} (represented by $h(A_{it}, L_{it}, Q_{it}; \theta)$) to Equation (5) until no remaining term is statistically significant.³¹

Descriptive Statistics for Transition Matrices. Table 1 displays the one-year transition rates between health risk categories for all age groups; the numbers are the predicted probabilities based on Equation (4). 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 reclassification is non-trivial even when just considering two subsequent years. For example, the probability of ending up in a different risk category in $t + 1$ (with very different mean expenditures as shown below) is more than 50% for category 2 in year t .

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

³⁰We select 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.

³¹We 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 Appendix A5 we provide some descriptive statistics for this subsample, which generally confirm that this assumption is reasonable.

Table 1: Health Risk Category Transitions

λ_t	$\mathbb{E}(m_t \lambda_t)$	λ_{t+1}							
		1	2	3	4	5	6	7	8 (+)
1	1,473	0.831	0.158	0.006	0.003	0.001	0.001	0.000	0.001
2	3,559	0.214	0.523	0.215	0.036	0.009	0.001	0.000	0.002
3	6,019	0.050	0.179	0.572	0.164	0.029	0.003	0.000	0.003
4	9,302	0.024	0.053	0.227	0.541	0.128	0.013	0.001	0.013
5	14,600	0.018	0.027	0.035	0.330	0.445	0.104	0.005	0.036
6	24,554	0.010	0.018	0.017	0.096	0.294	0.409	0.052	0.104
7	54,930	0.002	0.005	0.002	0.027	0.085	0.200	0.452	0.226

Notes: The table shows expected health expenditures and average transition probabilities across the different health categories. The sample includes all years, all age groups, and uses the ACG scores to construct risk categories λ as explained in Section 5.1. Source: Own calculations based on German Claims Panel Data.

category 1 decreases from 89 percent among 25-year-olds to 18 percent among individuals above 75. Also the probability of recovering, i.e. transitioning from a higher to a lower risk category, declines with 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.

Descriptive Statistics for Expenditures. Table 2 shows the mean expenditures m by age group, and the distribution of risk categories within each group. As expected, the mean expenditures strongly increase in age: they almost double from \$1,996 in the age group 25 to 30, to \$3,719 in the age group 45 to 50, and almost double again to \$7,151 in the age group 65 to 70. For those above 75 years, the average is \$10,020 (all values in 2016 U.S. dollars). This age gradient is, however, accounted for by our risk classification.³²

There is also a clear age gradient in health risk. The probability of $\lambda = 1$ declines progressively with age, whereas the share of enrollees in the five highest categories increases with age. Only 1.7 percent of those in the age group 25 to 30 are in categories $\lambda = 4$ and $\lambda = 5$. This share almost quadruples to 6.2 percent for the age group 45 to 50, and then more than quadruples again to 28.6 percent for the age group 65 to 70. It is 61 percent for enrollees above 75 years. Risk category $\lambda = 7$ clearly represents catastrophic risk and covers at most 0.3 percent in any age group in a given year. Section 6.5 discusses how catastrophic risk affects welfare under the different contracts.

³²As the ACG score depends on age, age-group indicators A_{it} should not have predictive power in the model for expected expenditures if our risk classification based on ACG scores is rich and flexible enough. Even though a few age-related parameters in Equation (5) turn out statistically significant, the deviations from mean expenditure within each risk category are economically insignificant. Figure A6 (in Appendix A5) illustrates this point. We interpret it as evidence that our preferred risk classification is rich enough. Therefore, we restrict all age effects to zero.

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

Age	Mean	S.D.($E(m \lambda)$)	λ_t						
			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

Notes: The table shows average claims and the standard deviation of claims within each risk category and age group for the different health categories. It also shows the fraction of individuals in each category by age group. The sample includes all years, all age groups, and uses the ACG scores to construct risk categories λ as explained in Section 5.1. Source: Own calculations based on German Claims Panel Data.

5.3 Lifecycle Income Paths

Next, we estimate the lifecycle income paths using 33 years of the German Socioeconomic Panel Survey (SOEP 1984-2016). We estimate income over the entire lifecycle as we are interested in assessing the performance of long-term contracts that are guaranteed renewable until death. 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 normalizes by the number of household members. We estimate the following individual fixed effects model:

$$\log(y_{it}) = \theta_i + f(\text{age}_{it}) + \epsilon_{it}, \quad (6)$$

where y_{it} is our income measure in 2016 U.S. dollars for individual i in year t ; and θ_i 's are individual fixed effects which net out persistent individual time-invariant income determinants, such as gender, preferences, or work productivity. The flexible function $f(\text{age}_{it})$ represents a series of age fixed effects and identifies the main coefficients of interest.

Because lifecycle profiles differ substantially by educational degree (Becker and Chiswick, 1966; Bhuller et al., 2017), we estimate the income process by education for (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*).³³

The two dashed curves in Figure 3 show the estimated age fixed effects for people with 10 years

³³Germany has three different schooling tracks where the majority of students complete school after 10 years and then start a three-year apprenticeship (cf. Dustmann et al., 2017).

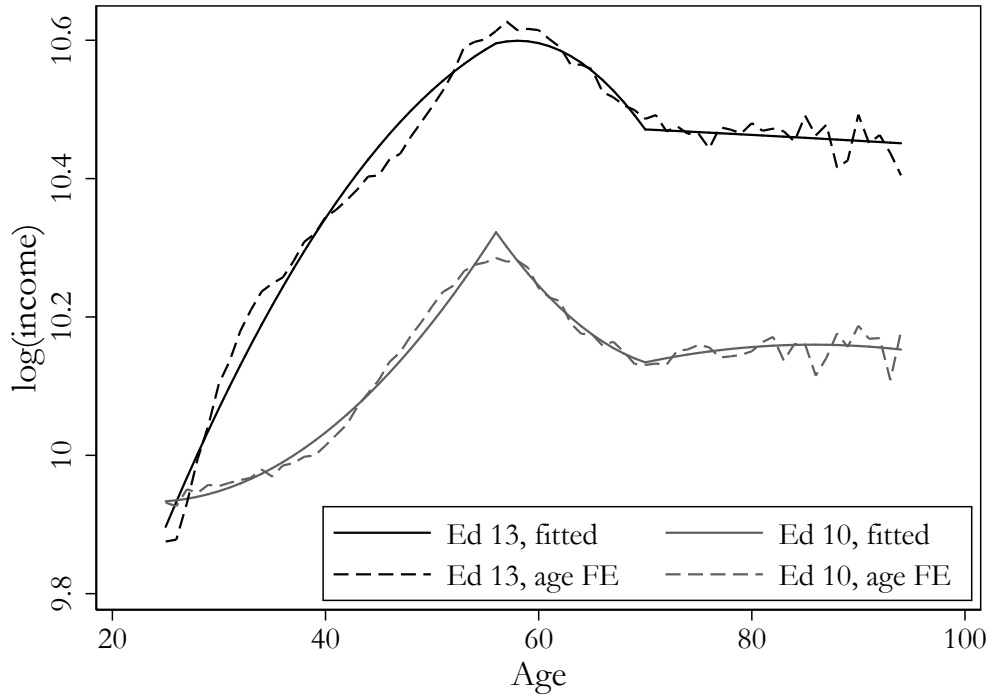


Figure 3: Life-cycle Income Paths Germany, Nonparametric and Fitted.

Notes: The figure shows the predicted income using Equation (6), and the fitted, piece-wise squared polynomial. All values in 2016 USD. Source: Own calculations using SOEP (2018) for the years 1984 to 2016.

and 13 years of schooling, respectively. Income rises sharply between age 25 and age 57 and 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 $f(\text{age}_{it})$ as a 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 3. 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 the 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, equivalized household income starts to decrease after age 57 until around age 70. Especially in the 1980s and 1990s but 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 an annual discount factor $\delta = 0.966$, which corresponds to an annual 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 status 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 an individual's current risk category $\lambda_t \in \{1, 2, \dots, 7\}$, as well as age $t \in \{25, \dots, 94\}$.

Figure 4 plots the resulting premiums calculated using Equation (1) for a handful of the most relevant combinations: $\lambda_t = 1$ and $t \in \{25, \dots, 59\}$; $\lambda_t = 2$ and $t \in \{25, \dots, 74\}$; $\lambda_t = 3$ and $t \in \{25, \dots, 94\}$; $\lambda_t = 4$ and $t \in \{60, \dots, 74\}$; $\lambda_t = 5$ and $t \in \{75, \dots, 94\}$. These combinations represent the most common health categories for each corresponding age interval.

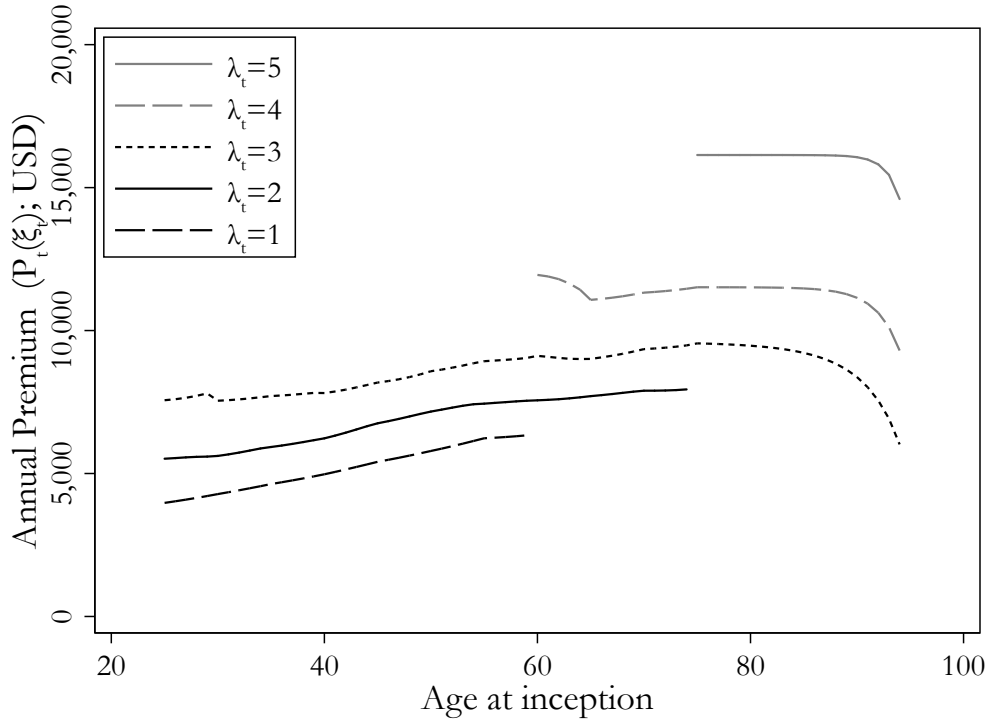


Figure 4: Calibrated Starting Premiums $P_t(\xi_t)$ in the GLTHI.

Notes: Calibrated GLTHI premiums at inception, calculated using Equation (1) and the estimated health status process. Health status ξ is comprised of age and health category λ , as explained in the text. Source: Own calculations using German Claims Panel Data.

Three forces determine the lifecycle profile of $P_t(\xi_t)$ in Figure 4. First, $P_t(\xi_t)$ is an increasing function of ξ_t because for any age, a higher risk category is associated with higher current and future

health care claims.

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 ζ_t , see Equation (4)). 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, starting premiums decrease with age in very old age ranges. The reason is that probability of ever transitioning to a worse category in the future diminishes as one approaches the end of life. Therefore, the need to front-load premiums to fund future negative health shocks *decreases* over the lifecycle. This force explains why $P_t(\zeta_t)$ decreases with t when t is sufficiently large.

Appendix A6 compares the calibrated and the observed premiums by age at inception for ages between 25 and 75.³⁴ Both are very similar in the key aspects that determine welfare over the life-cycle. Calibrated premiums for 25 years-old starting in categories 1 and 2 (99.35 percent of the sample) are very similar to the observed ones (and somewhat smaller than observed premiums for sicker individuals). We also observe positively sloped starting premiums by age over this entire age range, both for the calibrated and the observed premiums. Also, the rank ordering of premiums by health status persists over the entire life-cycle.³⁵

6.2 Comparison between the GLTHI and the Optimal Dynamic Contract

This subsection compares the lifecycle premiums and the amount of frontloading between the flat-income-optimal GLTHI design and the optimal dynamic (GHHW) contract (Ghili et al., 2024). Using our empirical health status transitions and income dynamics, Table 3 illustrates the differences between the flat-income-optimal GLTHI and the optimal contract by comparing the contract terms at age 25. Panel (a) shows the GLTHI premium and the frontloading amounts for a 25-year-old by the health risk category $\lambda_{25} \in \{1, \dots, 7\}$. If $\lambda_{25} = 1$, 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 frontloading decreases. For example, for $\lambda_{25} = 3$ the premium is \$7,563 which includes \$1,545 in frontloading. The reason is that when the current health status worsens, the likelihood of a further

³⁴For those health categories with enough observations of new policies being issued.

³⁵Although 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 ζ and t . Many of these are either completely absent or represented by only a small number of individuals in our data. Second, as we compare the welfare properties of GLTHI to counterfactual, model-based, contracts, we prefer to conduct welfare comparisons based on premiums that were generated in an analogous manner. Note also that the similarities between the observed and theoretical premiums highlighted in the text imply that the welfare under the GLTHI contracts evaluated with the observed premiums should be very close to our results derived from the calibrated premiums. Calibrated starting premiums for the vast majority of individuals are very similar to the observed ones and most people 'lock-in' this premium for their entire life-cycle. Also, note that lapsation rates depend on the rank-order of premiums but not the magnitude of the premium differences.

health deterioration also decreases.

Table 3: Comparing GLTHI Contract to Optimal Contracts Terms at Inception, Age 25

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
Frontloading	2,499	1,957	1,545	1,062	691	7	0
(b) Optimal contract <i>Ed 10</i>							
Premium	2,571	5,366	7,489	10,307	15,273	24,554	54,930
Frontloading	1,097	1,807	1,471	1,006	673	0	0
(c) Optimal contract <i>Ed 13</i>							
Premium	1,895	4,578	6,988	10,103	15,187	24,554	54,930
Frontloading	421	1,019	970	801	586	0	0

Notes: The table shows expected health care claims, starting premiums, and the amount of frontloading by health risk category at age 25, $\lambda_{25} \in \{1, \dots, 7\}$. All values are in 2016 USD. Source: German Claims Panel Data, [SOEP \(2018\)](#).

Panel (b) of Table 3 shows the premiums and the frontloading amounts for the optimal dynamic contract and an individual with 10 years of schooling (*Ed 10*) by initial health at age 25. For all health categories expect the worst, compared to GLTHI, the initial premiums and frontloading amounts are lower in the optimal dynamic contract. As income is low at early ages, the optimal contract entails less frontloading than GLTHI to improve the degree of intertemporal consumption smoothing. 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 \$2,571 (vs. \$3,973 for GLTHI) and for $\lambda_{25} = 4$, the optimal premium is \$10,307 (vs. \$10,363 for GLTHI). Note, however, in contrast to the constant guaranteed premiums under the GLTHI, the premiums under the optimal contracts need to adjust as income evolves to maintain any given level of consumption guarantee.

Panel (c) of Table 3 shows the premiums and the frontloading amounts for the optimal dynamic contract and an individual with 13 years of schooling (*Ed 13*) by initial health at age 25. This individual has a steeper income profile over her lifecycle (see Figure 3), which is why the optimal contract entails a *lower* degree of frontloading than for *Ed 10*, especially for healthy individuals. Again, the sicker the individual is at inception, the lower the frontloading amount.

Equivalent Arrow Securities. As discussed in Section 3.3, it is possible to replicate the dynamic contracts with one-period Arrow securities. Appendix A7 provides numerical examples for the terms of such contracts for the GLTHI contract and the optimal contract. As expected, the quantity of securities needed to replicate the optimal contract is smaller than that to replicate the GLTHI contract. This reflects the lower degree of frontloading required under the optimal contract.

6.3 Consumption Profiles over the Life-cycle

Life-cycle Consumption and Intertemporal Consumption Smoothing. We now explore how the short-term contract, the (flat-income-optimal) GLTHI contract, and the optimal contract affect intertemporal consumption smoothing and reclassification risk over the life-cycle. To do so, we simulate the consumption life-cycle profiles, from age 25 to age 94, for $N = 500,000$ individuals. Figure 5 plots the *average* consumption for these three contracts over the life-cycle, separately for *Ed 10* (Figure 5a), and *Ed 13* (Figure 5b).

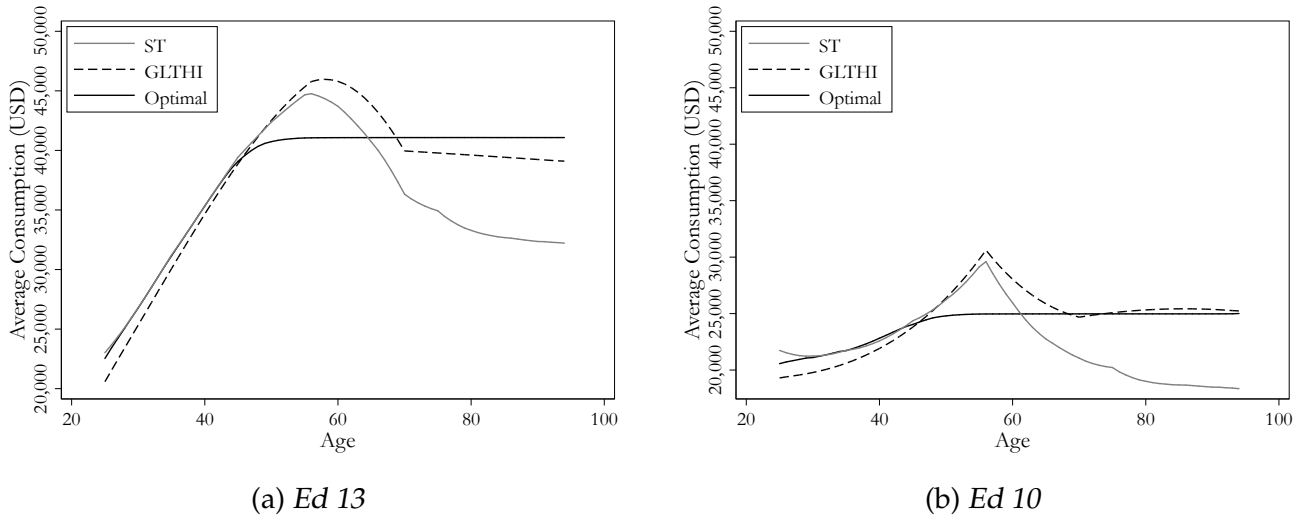


Figure 5: Simulated Average Consumption over the Lifecycle by Education

Notes: The figure displays average simulated consumption under short-term contracts, the GLTHI contract, and the optimal contracts. Panel (a) shows results for the steeper income profile (Ed 13), Panel (b) shows result for the flatter income profile (Ed 10). Source: Own calculations using German Claims Panel Data and SOEP (2018).

Under a series of short-term contracts, average consumption is simply income minus expected health expenditures. As shown by the grey solid lines, the average consumption profile under short term contracts is therefore hump-shaped over the life-cycle 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 frontloading of GLTHI up to the early 50s. As shown by the black 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 frontloading is costlier.

On top of the lower frontloading in the initial periods, studying income profiles over the life-cycle highlights an additional important advantage of optimal long-term contracts. In our setting with hump-shaped income profiles, the optimal contract allows to fully smooth consumption intertemporally after some age, as the straight flat black consumption line after around age 40 illustrates. The optimal contract prescribes higher premiums in the middle ages, not only to finance reclassification

risk insurance, but also to shift resources towards later years when income is low.

Reclassification Risk. To illustrate the degree of *reclassification risk* over the life-cycle, Figure 6 displays the standard deviations of consumption changes over the life-cycle. That is, Figure 6 plots, for each age t , the standard deviation of $\Delta C_{i,t} \equiv C_{i,t+1} - C_{i,t}$ across individuals i .

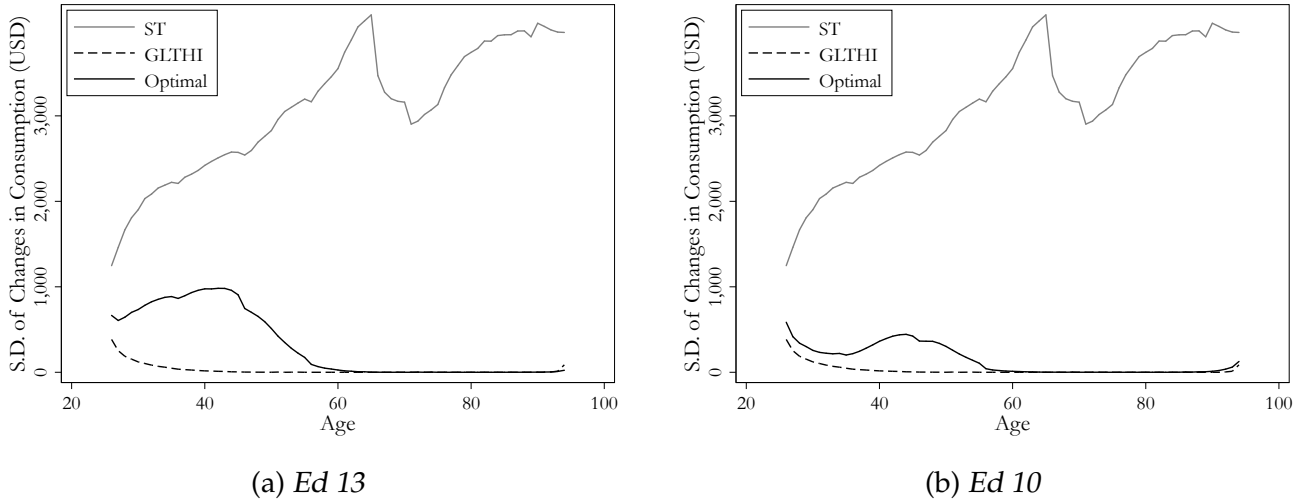


Figure 6: Simulated Standard Deviation of Consumption Changes over the Life-cycle, by Education

Notes: The figure displays the simulated standard deviation of consumption changes under short-term contracts, the GLTHI contract, and the optimal contracts. Panel (a) shows results for the steeper income profile (Ed 13), Panel (b) shows result for the flatter income profile (Ed 10). Source: Own calculations using German Claims Panel Data and SOEP (2018).

As seen, the GLTHI contract imposes very little reclassification risk as most individuals lock in $P_{25}(\cdot)$ in the first period. The few individuals who switch contracts are those who start with $\lambda_{25} > 1$ and become sufficiently healthier over time (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 that considers the “good news” regarding future health, contained in the event “ $\lambda_{25} = 1$ and $\lambda_{26} = 1$,” can 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 A12 (Appendix) shows lapsation rates under each contract.

6.4 Baseline Welfare Results

We now calculate how the differences in consumption profiles over the life-cycle translate into welfare differences across the different contracts. We summarize lifetime utility under each contract with the Consumption Certainty Equivalent (CE). In this section, we focus on the relative welfare difference between the GLTHI contract and the optimal dynamic contracts. We also compare the CE of each long-term contract with welfare under short-term contracts, and 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.

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. For welfare comparisons, we use a constant absolute risk aversion (CARA) utility function of the form:³⁶

$$u(c) = -\frac{1}{\gamma}e^{-\gamma c}. \quad (7)$$

In our main results, we use a risk aversion parameter $\gamma = 0.0004$ (Ghili et al., 2024). In Section 6.4.1, we explore the robustness to (i) the level of risk aversion, (ii) using constant relative risk aversion (CRRA) utility function, and (iii) to using non-time-separable Epstein-Zin preferences, among other robustness tests.

We provide eight sets of results in Table 4, stratifying the findings by the two education groups with different lifecycle income paths. Each corresponds to a different probability simplex that determines the initial category, $\Delta_0 \in \Delta^7$. For instance, Panel (a) assumes that everyone starts in the healthiest category, such that $\Delta_0 = \frac{1}{100}[100, 0, 0, 0, 0, 0, 0]$. Panel (h) shows our baseline results, which assume 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 ζ among the 25 to 30 age group. In Panel (h), we also assume that individuals cannot start in the worst health category. This assumption is motivated by the fact that individuals starting in category 7 are unable to afford actuarially fair premiums. One possibility is a fall-back option for the very sick, e.g. a coexisting public health insurance system.

Column (1) shows welfare under the first-best contract, C^* ; Column (2) shows welfare under the optimal contracts, C_{GHHW} . Column (3) shows welfare under the GLTHI contract, C_{GLTHI} . Column (4) shows welfare under a series of short-term contracts, C_{ST} . Finally, Column (5) shows the percentage

³⁶The CARA utility function has the convenience of allowing for negative consumption, which occurs when income is lower than the required premium payments. However, it also implies that the consumption equivalent may be negative under some contracts (see Table 4).

Table 4: Welfare under different contracts

	(1) C^*	(2) C_{GHHW}	(3) C_{GLTHI}	(4) C_{ST}	(5) $\frac{C_{GHHW}-C_{GLTHI}}{C_{GHHW}}$
Panel (a): $\Delta_0 = \frac{1}{100}[100, 0, 0, 0, 0, 0, 0]$					
Ed 10	23,029	22,489	21,537	-10,156	4.2%
Ed 13	34,210	27,727	26,025	-2,396	6.1%
Panel (b): $\Delta_0 = \frac{1}{100}[0, 100, 0, 0, 0, 0, 0]$					
Ed 10	22,600	21,373	20,840	-10,815	2.5%
Ed 13	33,776	25,569	24,898	-4,074	2.6%
Panel (c): $\Delta_0 = \frac{1}{100}[0, 0, 100, 0, 0, 0, 0]$					
Ed 10	22,244	20,169	19,855	-10,704	1.6%
Ed 13	33,420	23,619	23,273	-2,382	1.5%
Panel (d): $\Delta_0 = \frac{1}{100}[0, 0, 0, 100, 0, 0, 0]$					
Ed 10	21,905	18,409	18,253	-10,844	0.8%
Ed 13	33,079	21,100	20,945	-2,416	0.7%
Panel (e): $\Delta_0 = \frac{1}{100}[0, 0, 0, 0, 100, 0, 0]$					
Ed 10	21,476	14,715	14,679	-10,970	0.2%
Ed 13	32,648	16,646	16,598	-2,473	0.3%
Panel (f): $\Delta_0 = \frac{1}{100}[0, 0, 0, 0, 0, 100, 0]$					
Ed 10	20,639	5,969	5,968	-11,224	0.0%
Ed 13	31,809	7,574	7,568	-2,605	0.1%
Panel (g): $\Delta_0 = \frac{1}{100}[0, 0, 0, 0, 0, 0, 100]$					
Ed 10	11,587	-27,070	-27,070	-27,085	0.0%
Ed 13	22,327	-24,629	-24,629	-24,631	0.0%
Panel (h): $\Delta_0 = \frac{1}{100}[89.10, 10.25, 0.47, 0.11, 0.04, 0.03, 0]$					
Ed 10	22,980	21,945	21,168	-10,119	3.5%
Ed 13	34,159	26,093	25,088	-2,223	3.9%

Notes: The table shows welfare measured by the Consumption Certainty Equivalents in 2016 USD dollars, per capita, per year, separately for two income profiles (see Figure 3). Panels (a) to (g) differentiate by initial health status $\lambda_{25} \in \{1, \dots, 7\}$. 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) the optimal contract (C_{GHHW}), (3) the GLTHI (C_{GLTHI}), and (4) a series of short-term contracts (C_{ST}). Column (5) shows the percentage of welfare loss under GLTHI relative to the optimal contract. German Claims Panel Data, SOEP (2018).

welfare difference between the GLTHI and the optimal contract $\frac{C_{GHHW} - C_{GLTHI}}{C_{GHHW}}$.

Several findings emerge from Table 4. First, comparing Column (1) and Column (4), we find that a series of short-term contracts can produce large welfare losses compared to the first-best. For all initial health categories at age 25 and for both lifecycle income profiles, the CEs are negative. 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.³⁷ It is important to note, however, that the performance of short-term contracts is sensitive to catastrophic risk. Under the baseline scenario, individuals would pay \$54,930 whenever they reach category 7 under short-term contracts. This event, even if unlikely in a given year, has highly detrimental effects on welfare. We come back to this issue in Section 6.4.1.

Long-term contracts can produce substantial welfare gains compared to short-term contracts. Consider Panel (a) for the case where $\lambda_{25} = 1$ at inception. Column (2) shows that, under the optimal contract, the CE is \$22,448 for *Ed 10* and \$27,726 for *Ed 13*. Column (4) shows that, under the GLTHI contract, the CE is \$21,536 for *Ed 10* and \$26,024 for *Ed 13*.

The welfare difference between the optimal and the GLTHI contracts shrinks as health at inception worsens. Column (5) shows that for those starting in the healthiest risk category ($\lambda_{25} = 1$), the CE under the GLTHI contract is 4.2 percent lower for *Ed10* and 6.1 percent lower for *Ed13* compared to the optimal contract. Comparing Columns (5) across panels (a) to (g), we find that this difference becomes almost negligible for individuals starting in health category 5 or worse. Welfare under both long-term contracts (GLTHI and GHHW) is highly driven by the large degree of frontloading when health status at inception deteriorates (see Table 3). In this case, both long-term contracts are ineffective in dealing with reclassification risk.

The welfare difference between the GLTHI and the optimal contract is close to four percent. Panel (h) provides our baseline results using the initial health distribution that correspond to the observed distribution at age 25 in our sample. As most individuals start healthy (99.35% have $\lambda_{25} \in \{1, 2\}$), both long-term contracts produce substantial gains relative to short-term contracts and recoup a large fraction of the welfare loss from short-term contracting. This highlights the large benefits of long-term contracts in providing insurance against reclassification risk.

³⁷As 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 \(2003\)](#)).

6.4.1 Robustness: Preferences, Initial Health Status, Savings, and Income Profiles

We investigate the robustness of our previous result in various dimensions. First, we consider different levels of risk aversion. Second, we use a CRRA utility function instead of CARA. Third, we use Epstein and Zin (1989)'s recursive preferences where risk aversion and intertemporal elasticity of substitution are separately parameterized. Fourth, we use claims data from one of the biggest German public insurers to replicate our main findings for an initial health status representative of the population of Germany, and we further investigate the sensitivity of the results to the initial distribution of health status. Fifth, we allow for savings. Finally, we use U.S. income profiles. We provide details of each robustness analysis below.

Risk Aversion Our baseline results assume risk aversion $\gamma = 4 \times 10^{-4}$. The solid lines in Figure 7 show the percentage point difference in CEs between each long-term contract as we vary the value for γ over the interval $\gamma \in [7.5 \times 10^{-5}, 8 \times 10^{-4}]$. Panel (a) shows the results for the flatter income profile (Ed 10), while Panel (b) for the steeper income profile (Ed 13).

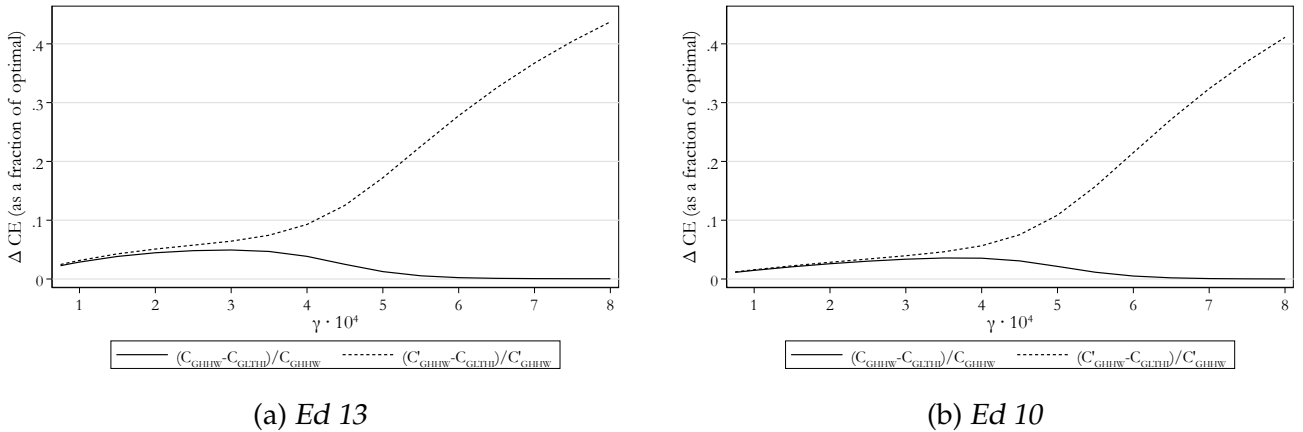


Figure 7: Welfare Difference (GLTHI vs GHHW) by Risk Aversion

Source: The figure shows the differences in Certainty Equivalents (CE) between GLTHI and the GHHW contract as a fraction of welfare under the GHHW contract, for different levels of risk aversion (γ). Panel (a) shows results for the steeper income profile (Ed 13), Panel (b) shows results for the flatter income profile (Ed 10). *Source:* Own calculations based on German Claims Panel Data and SOEP (2018).

We find that the percentage difference in welfare between the GHHW and GLTHI contracts is robust to the degree of risk aversion. The solid curves in the two panels show that the maximal welfare difference between the two contracts across all values of γ is 4.4 percent for Ed 10 and five percent for Ed 13, when $\gamma = 3.7 \times 10^{-4}$ or $\gamma = 3 \times 10^{-4}$, respectively. As expected, the welfare gap is small when γ is small, as reclassification risk and lack of intertemporal consumption smoothing are less relevant with less curvature in the utility function. However, the welfare gap is also small when γ is high.

To understand the mechanisms behind this pattern, we compute welfare under the lifecycle path of *expected* consumption under the GHHW contract—which we denote by C'_{GHHW} —and compare it to welfare under the GLTHI contract. In other words, we eliminate the welfare differences due to differences in reclassification risk, and quantify solely the welfare differences arising from better intertemporal consumption smoothing in the optimal contract. The dashed curves in Figure 7 plot $\frac{C'_{GHHW} - C_{GLTHI}}{C_{GHHW}}$.³⁸

For relatively low levels of γ , C'_{GHHW} is very close to C_{GHHW} , implying that the difference in CE between GLTHI and GHHW are almost exclusively determined by their differences in intertemporal consumption smoothing.³⁹ However, the role of reclassification risk in explaining the differences across contracts is more dominant at high values of γ . For instance, for $\gamma = 0.0004$ and for Ed 13, the average lifecycle consumption path under the optimal contracts produces welfare (C'_{GHHW}) that is 9.3 percent higher than welfare under GLTHI (compared to 3.9 percent higher when considering consumption volatility as in Panel (h) of Table 4). Overall, when γ is large, the welfare benefits of less reclassification risk under GLTHI largely compensate the welfare losses due to worse intertemporal consumption smoothing.

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 intertemporal consumption smoothing, where GLTHI clearly falls short. The larger γ , the bigger this effect. The second is due to differences in reclassification risk, where GLTHI outperforms the optimal contract. Again, the larger γ , the bigger this effect, and the closer is the welfare under GLTHI to the welfare under the optimal contract.

CRRA Preferences. Our main analysis assumes CARA preferences. We investigate the robustness of the findings to using CRRA preferences, that is $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. 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 $\sigma = 4$. We find welfare differences of 2.4% for Ed 10 and 3.8% for Ed 13.⁴⁰

³⁸In principle, one could also eliminate reclassification risk from the consumption path under GLTHI. However, the reclassification risk component of GLTHI is negligible.

³⁹Note that even at these relatively low levels of γ , the lack of intertemporal consumption smoothing decreases welfare substantially, relative to risk neutrality. However, welfare decreases with γ by a similar amount when we eliminate reclassification risk. For instance, for Ed 13, CE_{GHHW} and CE'_{GHHW} decrease by 14 percent and 13 percent compared to risk neutrality when $\gamma = 3 \times 10^{-4}$.

⁴⁰To 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.

Epstein-Zin Recursive Preferences. Our main specification assumes that a single parameter governs both risk aversion and the intertemporal elasticity of substitution. Here we break 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 $F(c, z) = ((1 - \delta)c^{1-1/\psi} + \delta z^{1-1/\psi})^{1/(1-1/\psi)}$. We embed the same CARA specification used in our main analysis into the EZ preferences by assuming $G(c) = u(c) = \frac{1}{\gamma}e^{-\gamma c}$. Appendix [A10.1](#) provides the details of the derivations.

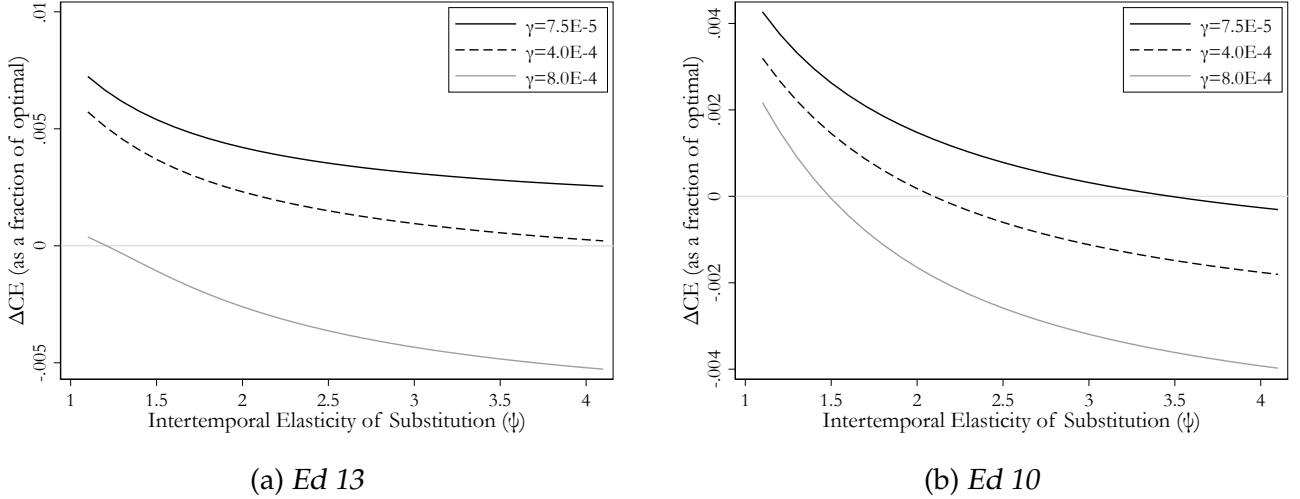


Figure 8: Welfare Differences (GLTHI vs. GHHW) with EZ preferences

Notes: The figure shows the differences in Certainty Equivalents (CE) between GLTHI and the GHHW contract as a fraction of welfare under the GHHW contract with recursive (Epstein-Zin) preferences, for different levels of intertemporal elasticity of substitution ψ , and for different levels of risk aversion (γ). Panel (a) shows results for the steeper income profile (*Ed 13*), Panel (b) shows results for the flatter income profile (*Ed 10*). Source: Own calculations based on German Claims Panel Data and [SOEP \(2018\)](#).

Varying γ and ψ , [Figure 8](#) shows differences in CEs between the GLTHI and the optimal GHHW contract. Each panel corresponds to a different income profile. We show results for both extremes of the risk aversion interval in [Figure 7](#) ($\gamma = 7.5 \times 10^{-5}$ and $\gamma = 8.0 \times 10^{-4}$), as well as our baseline risk aversion parameter ($\gamma = 4 \times 10^{-4}$). As expected, the welfare differences are larger with low levels of risk aversion and low intertemporal elasticity of substitution. The maximum welfare difference is 0.7%, which occurs for the steeper income profile. We also find that the GLTHI can even outperform the optimal contract when the intertemporal elasticity of substitution is relatively high.⁴¹

Initial Health Status Representative of Germany. Our results so far use the estimates of health status for the population of GLTHI policyholders. However, several institutional features imply that this population is not representative of the German population, and likely healthier (see [Section 2](#)).

⁴¹This can occur because the GHHW contract in [Ghili et al. \(2024\)](#) is not necessarily the optimal contract under recursive preferences—recall that [Ghili et al. \(2024\)](#)'s theoretical characterization requires that preferences are *time separable*, which [Epstein and Zin \(1989\)](#)'s recursive preferences do not satisfy.

This section studies the robustness of our results to using a health risk profile that is more representative of the German population as a whole. To do so, we use claims data from one of the biggest public insurers, with more than 5 million enrollees.

Overall, the claims data confirm that the privately insured are healthier. Figure A14 compares ACG risk score distributions across the publicly insured and the publicly insured; 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 the Appendix).⁴²

Using the risk scores and estimated initial probabilities at age 25 for the publicly insured, we find a welfare difference (GLTHI vs. GHHW) of 0.9 and 1.5 percent for Ed 10 and Ed 13, respectively. Appendix A10.2 provides additional details.⁴³

Different Starting States. We further study the robustness of the results to the initial health status by considering a large number of draws of distributions over the initial health categories for the privately insured. Specifically, we sampled 20 million probability simplices $\tilde{\Delta}_r \in \Delta^7$ from a Dirichlet distribution with concentration parameters equal to the baseline probabilities for the privately insured.

We find that the welfare loss of the GLTHI relative to the optimal contract is bounded at around six percent for the highly educated *Ed13* and at around four percent for the intermediate educated *Ed10*. This exercise also confirms that the welfare differences between the GLTHI and the optimal contract is smaller when the population is less healthy at inception, which alleviates the concern that our findings are driven by a relatively healthy subsample of the overall population. Appendix A10.3 provides the details.

Savings. Our main welfare calculations assume that individuals cannot save. Savings could improve welfare under the GLTHI contract, particularly in settings with hump-shaped lifecycle income profiles like the ones we find in the SOEP. Here, we allow for savings, by solving the dynamic programming problem of optimal savings with mortality risk as in Yaari (1965). The details are in Appendix A10.4.

We find that the relative gap between GLTHI and the optimal contract is almost unaffected by savings. Intuitively, the GLTHI contract already forces more savings than desired through highly

⁴²Furthermore, using the representative SOEP, Table A8 (Appendix) shows that privately insured are less likely to smoke, have lower BMIs, and use fewer health care services.

⁴³Another question is whether risk tolerance differs between publicly and privately insured. 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). Table A8 confirms this. Using the SOEP, Figure A13 plots very similar risk tolerance distributions for both populations. 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.

front-loaded premiums. Moreover, under the optimal contract, individuals have no incentives to save additionally as shown in Ghili et al. (2024). Thus, savings do not affect welfare under the optimal contract.

Income Profiles. Finally, to test the robustness with respect to the lifecycle income profile, we now employ the representative Panel Study of Income Dynamics (PSID) for the United States.⁴⁴ To extract lifecycle income profiles, we use the exact same income concept as in our main analysis for German income profiles. Moreover, we use the same estimation process. That is, we exclude respondents under 25, focus on the years 1984 to 2015, and estimate Equation (6) for high-school and college educated. Figure A16 (Appendix) shows the estimated lifecycle income profiles.

In the U.S., the increase in post-tax equivalized income between ages 25 and 60 is very close to the pattern in Germany. However, the decrease in income after age 60 is much steeper in the U.S., for both educational groups. This effect amplifies the beneficial effects of optimal contracts in allowing for consumption smoothing at old ages, as shown in Figure 5. Overall, the GLTHI contract would achieve welfare that would fall 5.8 and 3.5 percent short of the optimal dynamic long-term contract for Americans with high school and college degrees, respectively. To illustrate the role of the hump-shaped income, we calculate the welfare gap between the GLTHI and the GHHW contracts under the PSID income for long-term contracts that would expire at 65 years old.⁴⁵ In such case, the welfare gap between the GLTHI and GHHW contracts is 2.3 percent and 2.0 percent for high school and college income profiles, respectively. Also, allowing for savings in life-long contracts, the welfare gap is 3.7 percent and 3.0 percent for high school and college income profiles, respectively.

6.5 Benchmarking the Loss from Using Flat-Income Optimal Contracts

In our setting, the flat-income optimal GLTHI contracts provide welfare that is around four percent lower than with the optimal contracts. Here we put this welfare difference in perspective, and provide additional lessons regarding when and why the flat-income optimal GHTHI contract can approximate the welfare gains from the optimal contracts.

The benefits of replacing short term contacts with each long-term contract provide a natural benchmark to compare the welfare differences across both long-term contracts. Panel (a) of Table 5 expands our main results by providing additional metrics to assess the welfare gains of long-term over short-term contracts. Columns (1) to (4) replicate the main results for CE under each contract.

⁴⁴The 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).

⁴⁵This exercise is also interesting from a policy perspective if long-term contracts were considered for the US working population while keeping the Medicare program intact.

Column (5) and (6) show what fraction of the welfare difference between the first best and the short-term contract is recouped by the optimal contract and the GLTHI, $\frac{C_{GHHW}-C_{ST}}{C^*-C_{ST}}$ and $\frac{C_{GLTHI}-C_{ST}}{C^*-C_{ST}}$, respectively. Column (7) shows the fraction of the gains from replacing short-term contracts with the optimal (GHHW) contracts that is achieved by replacing short term contract with the GLTHI contract $\frac{C_{GLTHI}-C_{ST}}{C_{GHHW}-C_{ST}}$. Finally, Column (8) shows again the percentage welfare difference between the GLTHI and the optimal contract.

Under our baseline results, the optimal contracts are very effective. They close either 96.9 percent or 77.8 percent of the welfare gap between short-term contracts and the first best for each income profile, respectively (Column 5). This highlights the large benefits of optimal long-term contracts in providing insurance against reclassification risk. Similarly, the GLTHI contract is able to close either 94.5 percent or 75.1 percent of that gap (Column 6). As a consequence, for both income profiles, the GLTHI contract achieves close to 97 percent of the welfare gains from optimal long-term contracting (Column 7). Even though the GLTHI contract is only optimal when income is constant over the life-cycle, its ability to insure *large* reclassification risk, and its particular balance between front-loading and reclassification risk insurance still achieves much of the benefits from optimal long-term contracting.

Table 5: Benchmarking Welfare under GLTHI

	(1) C^*	(2) C_{GHHW}	(3) C_{GLTHI}	(4) C_{ST}	(5) $\frac{C_{GHHW}-C_{ST}}{C^*-C_{ST}}$	(6) $\frac{C_{GLTHI}-C_{ST}}{C^*-C_{ST}}$	(7) $\frac{C_{GLTHI}-C_{ST}}{C_{GHHW}-C_{ST}}$	(8) $\frac{C_{GHHW}-C_{GLTHI}}{C_{GHHW}}$
Panel (a): Baseline Results								
Ed 10	22,980	21,945	21,168	-10,119	96.9%	94.5%	97.6%	3.5%
Ed 13	34,159	26,093	25,088	-2,223	77.8%	75.1%	96.5%	3.9%
Panel (b): Without Catastrophic Losses								
Ed 10	22,979	21,930	21,164	14,504	87.6%	78.6%	89.7%	3.5%
Ed 13	34,159	26,064	25,079	22,123	32.7%	24.6%	75.0%	3.8%

Notes: The table shows welfare measured by the Consumption Certainty Equivalents in 2016 USD dollars, per capita, per year, separately for two income profiles (see Figure 3). Panels (a) uses the baseline contracts as in Table 4. Panel (b) assumes that catastrophic risk is insured by an alternative contract, as explained in the text. Columns (1) to (4) show welfare according to the (1) first-best (C^*), (2) the optimal contract (C_{GHHW}), (3) the GLTHI (C_{GLTHI}), and (4) a series of short-term contracts (C_{ST}). Column (5) and (6) show how much of the welfare gap between (2) and (1) is closed by the optimal contract and the GLTHI, respectively. Column (7) shows the fraction of the potential benefits from long-term contracting achieved by GLTHI. Column (8) shows the percentage of welfare loss under GLTHI relative to the optimal contract. Source: Own calculations based on German Claims Panel Data, SOEP (2018).

The Role of Catastrophic Risk. As previously noted, welfare under short-term contracts is very low in our baseline scenario. Our risk classification method accommodates the heavily skewed dis-

tribution of expenditures with a high-cost but low-probability risk category 7 (Section 5.1). Even though we observe large health expenditures in our data, the assumption that individuals would be subject to such large losses in the counterfactual scenario with short-term contracts may be unrealistic, for instance, due to public safety-net programs, or other forms of protection against catastrophic risks. Relatedly, one may wonder whether *any* mechanism that provides insurance against this catastrophic risk would perform close to the optimal dynamic contract.

We now evaluate the welfare effects under different contracts when catastrophic losses are bounded. Specifically, we recompute welfare under each contract by capping the expenditures of category 7 at those of category 6. To finance that cap, individuals starting at age 25 in $\lambda_{25} = k_0$ pay a yearly fee $\tilde{P}(k_0)$ such that the insurance company breaks-even ex ante.⁴⁶ In particular, premiums under this modified short-term contract are now $P_t(\lambda_t|\lambda_{25} = k_0) = \tilde{P}(k_0) + \mathbb{E}(m_t|\min\{\lambda_t, 6\})$. The long-term contracts are also computed as before, but with the capped health expenditures and the additional yearly fee.⁴⁷ Panel (b) of Table 5 provides the results under this alternative set of contracts.

We find that the gap between the two long-term contracts (GLTHI and GHHW) is robust to catastrophic expenditures (Column 8 of Table 5). It remains below four percent after we eliminate the risk of category-7 expenditures. In fact, welfare under both long-term contracts changes little under this alternative scenario. This is because long-term contracts have the virtue of preventing individuals from paying higher premiums when they transition to high-cost categories. In particular, health expenditures of category 7 only affect welfare by slightly affecting the contract terms (as state 7 only arises with very low probability in the data and the contract terms are set by the risk-neutral insurer). This result highlights that long-term contracts can be very effective in dealing with the risk of future catastrophic losses.⁴⁸

As expected, welfare under short-term contracts increases substantially when we limit the size of the catastrophic risk. As a consequence, bounding catastrophic losses generally results in smaller gains from the optimal GHHW contracts. In particular, Column (5) shows that for the steeper income profile (Ed 13), optimal long-term contracts close 32.7 percent of the gap between the first best and short-term contracts, as opposed to 77.8 percent in our baseline scenario.

Naturally, capping catastrophic losses increases the welfare gap from using the GLTHI instead of the optimal GHHW when measured as a fraction of the gains from optimal long-term relative

⁴⁶ That is, $\tilde{P}(k_0) = \frac{\mathbb{E}(\sum_{t=25}^T S_t \delta^{t-25} \mathbb{1}[\lambda_t=7](\mathbb{E}(m_t|\lambda_t=7) - \mathbb{E}(m_t|\lambda_t=6))|\lambda_{25}=k_0)}{\mathbb{E}(\sum_{t=25}^T S_t \delta^{t-25}|\lambda_{25}=k_0)}$ where $\mathbb{1}[\cdot]$ is the indicator function.

⁴⁷We note that these alternative contracts are not “incentive-compatible,” as individuals who become healthier than expected over time could switch to an alternative contract with a lower fee. The proposed contracts assume commitment in the payment of $P_t(\lambda_t|\lambda_{25} = k_0)$.

⁴⁸It is also interesting to note that welfare under both long-term contracts decreases (slightly) in this alternative scenario, even if we assume that individuals can commit to paying the premium $P_0(\lambda_{25})$ over time, regardless of the evolution of their health status.

to short-term contracting. The reason is simple: The welfare gap between the GLTHI and GHHW long-term contracts stays almost unaltered in *absolute* terms, but the gains from optimal long-term contracts relative to short-term contracting shrinks. In our baseline scenario with large catastrophic losses, GLTHI recoups 97.6 percent (for Ed 10) and 96.5 percent (for Ed 13) of the gains from replacing short term contracts with the optimal long-term contracts (Column (7), Panel (a), Table 5). However, after bounding category-7 expenses, the gains from optimal long-term contracting are more limited, and GLTHI recoups 89.7 percent (for Ed 10) and 75.0 percent (for Ed 13) of those (smaller) gains (Column (7), Table 5).

Both long-term contracts still bring sizable gains beyond those that come from capping catastrophic risk, particularly for flatter income profiles (CE increases from 14,504 to 21,164 under the GLTHI contract and to 21,930 under the GHHW contract). It is also worthwhile noting that, in our empirical setting, both long-term contracts improve welfare substantially relative to the “guaranteed-renewable” contracts of Pauly et al. (1995), which fully eliminate reclassification risk at the expense of even higher frontloading.⁴⁹ These results highlight the virtue of optimal long-term contracting, in that their welfare gains do depend on the intricate trade-offs between intertemporal consumption smoothing and reclassification risk insurance. The particular way in which GLTHI balances this tradeoff, and its shared commonalities with the optimal contract, are important to explain its performance.

The results from eliminating the catastrophic risk are tightly connected to the fact that the German contract provides more reclassification risk insurance at the expense of more frontloading than the optimal dynamic contract. In our setting, much of the welfare loss from short-term contracting would be due to reclassification risk.⁵⁰ Under such situations, both the GLTHI and the optimal GHHW contracts are very effective as they are designed to mitigate reclassification risk using frontloaded premiums. By virtue of practically eliminating reclassification risk, the GLTHI contract achieves much of the benefits of long-term contracting, even though it introduces more frontloading than optimal. In contrast, when the size of the catastrophic risk is bounded, a smaller fraction of the losses from short-term contracting is due to reclassification risk (and a larger fraction is due to the lack of intertemporal consumption smoothing).

The difference across income groups in the relative performance of each contract also relates di-

⁴⁹Results are available from the authors upon request.

⁵⁰To make this point more precisely, we can quantify the relative importance of reclassification risk under our benchmark scenario as well as under the alternative assumptions regarding catastrophic losses by computing the certainty consumption equivalent under a benchmark of no-borrowing and no-savings (C_{NBNS}) and calculating $\frac{C_{NBNS} - C_{ST}}{C^* - C_{ST}}$. This metric quantifies what fraction of the overall welfare losses from short-term contracts ($C^* - C_{ST}$) are due to reclassification risk ($C_{NBNS} - C_{ST}$). The remainder is due to the lack of intertemporal consumption smoothing. We find that in our baseline scenario, for Ed 10 (Ed 13), this metric is equal to 97.3 percent (82.7 percent), whereas after bounding the category-7 catastrophic losses it is 89.8 percent (47.4 percent).

rectly to the source of welfare losses under short-term contracting. For relatively flat income profiles (Ed 10), most of the welfare loss from short-term contracting is due to reclassification risk, even when the size of the catastrophic risk is bounded. As a consequence, the GLTHI contract achieves a large fraction of the potential gains from optimal long-term contracting. For steeper income profiles, however, a larger fraction of the welfare gap comes from poor intertemporal consumption smoothing, especially when catastrophic risk is bounded. In these cases, long-term contracts are in general less effective; and the GLTHI contract achieves a smaller fraction of the potential gains from optimal long-term relative to short-term contracting.

6.6 Summary

In our setting, we find that flat-income optimal (GLTHI) design achieves welfare that is close to 96 percent of the welfare under the optimal contract. This result is robust to several assumptions.

When reclassification risk is high, long-term contracts are very effective and provide large welfare gains relative to short term contracts. In such cases, the welfare gains of replacing short term contracts with the flat-income optimal design represent most of the gains from optimal long-term contracting. However, when catastrophic losses are more limited, the gains from optimal long-term contracting are also more limited, and it is also the case that the flat-income optimal recoups a smaller share of the gains. This pattern is tightly linked to the fact that the flat-income optimal contract provides more reclassification risk insurance at the expense of worse intertemporal consumption smoothing.

7 Conclusion

Regulated private markets for comprehensive health insurance have an important role in several OECD countries—including the US, Chile, Switzerland, and Germany. A central goal of premium rating regulation in such markets is to prevent reclassification risk, which is often achieved through community-rated pricing. For example, Switzerland has a private health insurance market with tight community rating regulations to provide reclassification risk insurance for all citizens. Several large markets and programs in the US also feature community rating, most notably the individual market after the enactment of the Affordable Care Act (ACA). However, community rating regulations, unless combined with individual mandates, tend to exacerbate adverse selection (Geruso and Layton, 2017).

A fundamental, but empirically understudied alternative to regulated short-term contracts are individual *long-term* contracts. They provide policyholders with reclassification risk insurance without necessarily triggering adverse selection. This paper bridges the theoretical literature on long-term

contracts with its most important real-world application—the German individual private long-term health insurance market (GLTHI).

We show that contracts in Germany’s private health insurance share several features with the optimal dynamic contract as recently derived by [Ghili et al. \(2024\)](#). Moreover, we highlight that the two contracts coincide for individuals with a flat lifecycle income profile. More generally, the GLTHI contract almost fully eliminates reclassification risk over a policyholder’s lifecycle. However, the elimination of reclassification risk comes at the expense of higher premium front-loading which, compared to the optimal contract, results in less intertemporal consumption smoothing.

We assess welfare under the GLTHI contracts and compare it to optimal contracts leveraging unique claims panel data from GLTHI policyholders along with household panel data. Key to this analysis, we propose a novel risk classification method based on principles of actuarial science. We find that welfare under the GLTHI design—as measured by certainty-equivalent yearly consumption—is 96 percent of what the income-dependent optimal contract achieves. This result is robust to the degree of risk aversion, to using non-time-separable recursive preferences à la [Epstein and Zin \(1989\)](#), to using the initial health status of the publicly insured in Germany, to using U.S. lifecycle income profiles, and to the extent of catastrophic risk. Overall, the welfare loss due to less consumption smoothing from ignoring income dynamics is partly compensated by less reclassification risk.

We also provide further insights regarding the workings of the flat-income-optimal GLTHI contract relative to the income-dependent optimal dynamic contracts. In our empirical setting, a large share of the welfare loss under the counterfactual scenario of short-term contracts would be due to reclassification risk. The flat-income-optimal and optimal contracts are very effective at improving upon short-term contracts: they mitigate the large reclassification risk with premium frontloading resulting in large welfare gains over short-term contracts. However, in scenarios with more limited catastrophic risk, e.g. if long-term contracts were to coexist with government programs that already insure catastrophic expenses, the additional gains from optimal long-term contracting would be smaller, and the GLTHI design would recoup a smaller share of those incremental gains.

A practical advantage of the GLTHI contract is its simplicity in that, after an initial risk rating, they are guaranteed renewable with constant and income-independent premiums as long as the policyholder does not lapse. This simplicity may explain its proliferation and stable existence in Germany for decades. We see the performance of long-term contracts under this particular design as a virtue of the theory of optimal long-term contracts, which strengthens their case as an appealing policy option. Alternative designs with premium profiles tailored to rising lifecycle income profiles as, for

instance, those observed in life insurance, should also be considered (Hendel and Lizzeri, 2003). In the U.S., so far, the debate has largely focused on either incremental adjustments to the community-rated short-term contracts under the ACA, or the transition to a “single-payer for all” system.

Finally, we acknowledge two important and general caveats of long-term contracts. First, neither the GLTHI nor the optimal dynamic contracts are affordable for those who are sick in young ages. From a policy perspective, societies that implement long-term contracts must offer public insurance for these individuals. Second, our discussion abstracts from a couple of key features that may have implications for welfare under the long-term contracts. First, our model assumes time-consistent individuals. From the perspective of a present-biased consumer, front-loading may render long-term contracts undesirable, particularly when front-loading is high.⁵¹ Furthermore, our model abstracts from moral hazard. In the presence of moral hazard, the use of 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 (2023). *PKV in Zahlen 2022*. Verband der Privaten Krankenversicherung. https://www.pkv.de/fileadmin/user_upload/PKV/3_PDFs/Publikationen/Zahlenbericht_2022.pdf, retrieved April 20, 2024.
- Association of German Private Healthcare Insurers (2024a). *PKV Zahlenportal*. Verband der Privaten Krankenversicherung. <https://www.pkv-zahlenportal.de/>, retrieved April 20, 2024.
- Association of German Private Healthcare Insurers (2024b). *So funktioniert die Pflegeversicherung*. Verband der Privaten Krankenversicherung. <https://www.pkv.de/wissen/pflegeversicherung/so-funktioniert-die-pflegeversicherung/>, retrieved April 24, 2024.
- 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.

⁵¹Still, 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 equilibrium with time-inconsistent agents, time-inconsistent agents may achieve the same level of welfare as time-consistent agents.

- Atal, J. P. (2019). Lock-in in dynamic health insurance contracts: Evidence from Chile. PIER Working Paper Archive 19-020, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.
- Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economics* 11(1), 685–725.
- Avdic, D., S. Decker, M. Karlsson, and M. Salm (2024). No-claim refunds and healthcare use. *Journal of Public Economics* 230, 105061.
- 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 April 10, 2024.
- 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 (2024). Die Pflegeversicherung. <https://www.bundesgesundheitsministerium.de/themen/pflege/online-ratgeber-pflege/die-pflegeversicherung.html>, retrieved April 20, 2024.
- Bundesministerium für Gesundheit (2024). Gesundheitsberichterstattung des Bundes. <http://www.gbe-bund.de/>, retrieved April 20, 2024.
- 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/fachgrundsaeetze-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 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.
- Fleitas, S., G. Gowrisankaran, and A. L. Sasso (2018). Reclassification risk in the small group health insurance market. Technical report, National Bureau of Economic Research.
- 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. Schündeln (2005). Precautionary savings and self-selection: Evidence from the German reunification “experiment”. *Quarterly Journal of Economics* 120(3), 1085–1120.
- Geruso, M. and T. J. Layton (2017, November). Selection in health insurance markets and its policy remedies. *Journal of Economic Perspectives* 31(4), 23–50.
- 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 (2024). Optimal long-term health insurance contracts: Characterization, computation, and welfare effect. *The Review of Economic Studies* 91(2), 1085–1121.
- GKV-Gesundheitsreformgesetz (2000). Gesetz zur Reform der gesetzlichen Krankenversicherung ab dem Jahr 2000. [https://www.bgbl.de/xaver/bgbl/start.xav?start=%2F%2F*\[%40attr_id%3D%27bgbl199s2626.pdf\]](https://www.bgbl.de/xaver/bgbl/start.xav?start=%2F%2F*[%40attr_id%3D%27bgbl199s2626.pdf]), last retrieved on February 10, 2020.
- GKV-Spitzenverband (2024). Die Gesetzlichen Krankenkassen. https://www.gkv-spitzenverband.de/krankenversicherung/kv_grundprinzipien/alle_gesetzlichen_krankenkassen/alle_gesetzlichen_krankenkassen.jsp, retrieved April 20, 2024.
- 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.
- Jacobs, K., F. Breyer, J. Wasem, A. Walendzik, and K.-D. Henke (2012). Dualität aus gesetzlicher und privater krankenversicherung überholt? *wirtschaftsdienst* 92(10), 651–666.
- 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.
- Koch, P. (1971). *Von der Zunftlade zum rationellen Grossbetrieb: kleine Geschichte d. privaten Krankenversicherung in Deutschland*. Verlag Versicherungswirtschaft.
- Koijen, R. S. J., T. J. Philipson, and H. Uhlig (2016). Financial Health Economics. *Econometrica* 84, 195–242.
- Krankenkasse, T. (2024). Wann Beschäftigte in die gesetzliche Krankenversicherung zurückkehren können. <https://www.tk.de/firmenkunden/service/fachthemen/versicherung-fachthema/wann-rueckkehr-in-die-gkv-moeglich-ist-2117276?tkcm=aaus>, retrieved April 23, 2024.
- Krankenkassen-Zentrale (KKZ) (2020). BAP 2020: Neue Private Krankenversicherung – Beitragserhöhung und andere Beitragsanpassungen. <https://www.krankenkassenzentrale.de/wiki/beitragsanpassung#>, retrieved on February 6, 2010.
- Kriegel, H.-P., E. Schubert, and A. Zimek (2017). The (black) art of runtime evaluation: Are we comparing algorithms or implementations? *Knowledge and Information Systems* 52(2), 341–378.
- Krueger, D. and H. Uhlig (2006). Competitive risk sharing contracts with one-sided commitment. *Journal of Monetary Economics* 53(7), 1661–1691.
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28(2), 129–137.
- Nell, M. and S. Rosenbrock (2007). Die Diskussion über die Portabilität von risikogerechten Transferbeträgen in der Privaten Krankenversicherung. *Zeitschrift für die gesamte Versicherungswissenschaft* 96(1), 39–51.
- Nell, M. and S. Rosenbrock (2009). Ein Risikoausgleichsmodell für den Wettbewerb um Bestandskunden in der PKV. *Zeitschrift für die gesamte Versicherungswissenschaft* 98(4), 391.
- Netzwerk der Verbraucherzentralen in Deutschland (2024). So wechseln Sie von der privaten in die gesetzliche Krankenversicherung. <https://www.verbraucherzentrale.de/wissen/gesundheit-pflege/krankenversicherung/so-wechseln-sie-von-der-privaten-in-die-gesetzliche-krankenversicherung-41289>, retrieved April 23, 2024.
- 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. IFRS - gesundheitspolitik beiträge zur 19. wissenschaftstagung des bundes der versicherten und zum workshop 'junge versicherungswissenschaft', versicherungswissenschaftliche studien.
- 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. (2024). Die beihilfesysteme des bundes und der länder im vergleich. Rechtswissenschaftliches gutachten im auftrag des bundesministeriums für gesundheit, Bundesministerium für Gesundheit. <https://www.bundesgesundheitsministerium.de/service/publikationen/details/vergleich-der-beihilfesysteme-der-bundeslaender-und-des-bundes-aus-rechtlicher-sicht.html>, retrieved April 20, 2024.
- 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). Sozialeleistungen - Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus) 2019. Fachserie 13 reihe 1.1.
- Werbeck, A., A. Wöbker, and N. R. Ziebarth (2021). Cream skimming by health care providers and inequality in health care access: Evidence from a randomized field experiment. *Journal of Economic Behavior & Organization* 188, 1325–1350.
- Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies* 32(2), 137–150.