Optimal Long-Term Health Insurance Contracts: Characterization, Computation, and Welfare Effects*

Soheil Ghili†, Ben Handel‡, Igal Hendel§, and Michael D. Whinston¶

August 15, 2022

Abstract

Reclassification risk is a major concern in health insurance where contracts are typically one year in length but health shocks often persist for much longer. While most health systems with private insurers pair short-run contracts with substantial pricing regulations to reduce reclassification risk, long-term contracts with one-sided insurer commitment have significant potential to reduce reclassification risk without the negative side effects of price regulation, such as adverse selection. We theoretically characterize optimal long-term insurance contracts with one-sided commitment, extending the literature in directions necessary for studying health insurance markets. We leverage this characterization to provide a simple algorithm for computing optimal contracts from primitives. We estimate key market fundamentals using data on all under-65 privately insured consumers in Utah. We find that dynamic contracts are very effective at reducing reclassification risk for consumers who arrive to the market in good health, but they are ineffective for consumers who come to the market in bad health, demonstrating that there is a role for the government insurance of pre-market health risks. Individuals with steeply rising income profiles find front-loading costly, and thus relatively prefer ACA-type exchanges. Switching costs enhance, while myopia moderately compromises, the performance of dynamic contracts.

* We thank Hanming Fang, Sebastian Fleitas, Nathan Hendren, Kurt Lavetti, Neale Mahoney, and Stephan Seiler as well as many seminar participants for their comments. All authors are grateful for support from NSF grant SES-1259770. We are grateful to Kurt Lavetti for his guidance with respect to the Utah all-payers claims data. We thank Michael Sullivan for outstanding research assistance.

† Yale School of Management; soheil.ghili@yale.edu

‡ Department of Economics, UC Berkeley; handel@berkeley.edu

§ Department of Economics, Northwestern University; igal@northwestern.edu

¶ Department of Economics and Sloan School of Management, M.I.T.; whinston@mit.edu

1 Introduction

Consumers face substantial health risks over their lifetimes. Much of this risk involves conditions, such as diabetes, heart disease, and cancer, that lead to high expected medical expenses over significant periods of time. These conditions can expose individuals who buy short-term insurance coverage to substantial premium increases – so-called "reclassification risk" – greatly reducing the extent to which their health risks are insured.

Concerns over reclassification risk have received a great deal of public and academic attention in recent years. Markets characterized by managed competition, such as the Affordable Care Act (ACA) exchanges in the U.S. and nationwide exchanges in the Netherlands and Switzerland, emphasize short-term (one-year) insurance contracts and contend with the problem of reclassification risk through community rating and guaranteed issuance, thereby prohibiting discrimination against consumers who have developed pre-existing conditions. Unfortunately, while requiring identical pricing for consumers with different health can eliminate reclassification risk, it can create adverse selection, leading to under-provision of insurance or a need for a byzantine web of regulations to combat that selection [Handel, Hendel and Whinston (2015), Patel and Pauly (2002)].

This paper investigates whether reclassification risk can be managed without strong regulation. Specifically, we investigate the potential welfare gains from long-term contracts. By specifying long-term obligations, such contracts can mitigate reclassification risk without the pricing regulation that leads to adverse selection when contracts are short term. Long-term contracts are common in other less regulated markets, such as life insurance [Hendel and Lizzeri (2003)], and exist in some health insurance markets such as Germany and Chile [Browne and Hoffman (2013), Atal (2016), Atal et al. (2020)]. Further, in the U.S. prior to the ACA, state-level guaranteed renewability regulations introduced some elements of dynamic contracting to health insurance markets, albeit of a restricted form [Patel and Pauly (2002), Marquis and Buntin (2006)]. While there are some practical impediments to long-term contracts in health, which we outline in Section 8, many of those impediments could be overcome with technological or regulatory changes were long-term contracts seen to be sufficiently beneficial.

We make three primary contributions toward understanding the performance of long-term contracts. First, we extend prior work on long-term contracts with one-sided commitment [e.g. Harris and Holmstrom (1982), Hendel and Lizzeri (2003), Pauly, Kunreuther and Hirth (1995), Krueger and Uhlig (2006)] and develop a dynamic model of health insurance contracting that allows for flexible stochastic health processes and long contract durations.[1]

[1] Other prior related work in macro-finance includes Farhi and Werning (2013) and Golosov, Troshkin, and Tsyvinski (2016).

We characterize the form of optimal long-term contracts in a series of theorems, and extend the framework to allow for relevant frictions such as consumer switching costs, consumer myopia, and some forms of consumer self-selection. Second, we use our theorems to establish a simple algorithm for computing optimal long-term contracts as a function of key market fundamentals such as the stochastic health process and the life-cycle path of a consumer's income. This algorithm allows the literature on long-term insurance contracts with one-sided commitment, which has so far been largely theoretical, to expand into empirics. Third, we estimate these market fundamentals using data on all under-65 privately insured individuals from the state of Utah and use our estimates to quantify the potential benefits (or costs) of long-term contracts for this policy-relevant sample. In addition to quantifying the benefits and costs of such contracts, we study positive and normative comparative statics with respect to these fundamentals and also study the welfare impacts of complementary public policies, e.g., public insurance for those entering the insurance system with significant health risks.

We establish several results about the performance of long-term contracts in our empirical analysis. First, optimal long-term health insurance contracts can do a good job insuring individuals who arrive at the market (at age 25 in our analysis) in good health, but do a much worse job at insuring those who are in poor health at that point. The reason is that those in poor health find the front-loading required to insure against reclassification risk very costly given their current needs. This poor performance for those initially in bad health substantially limits the benefits of long-term contracts from an ex ante (at-birth) perspective. We find that, for consumers with net income (equal to income less expected medical expenses) that is flat over their lifetime, dynamic contracts recover 99.4% of the welfare gap between spot (year-to year) contracts and the constrained first-best outcome for a consumer entering the market in the best health state, but only 29.1% of this gap for consumers entering the market in the worst health state.

Second, we find that the value of long-term health insurance contracts is lower for those with steeply rising age-income profiles, who find front-loading costly.[2] This reduction in performance with rising income profiles is particularly acute for those who arrive at the market in poor health. Under an empirically calibrated rising lifetime income, we find that for a consumer who enters the market in the healthiest state dynamic contracts recover 95.1% of the welfare gap between spot contracts and the constrained first-best, reduced slightly from the 99.4% mentioned above. However, for those entering the market in the sickest health state, the 29.1% welfare benefit of dynamic contracts mentioned above goes down to 7.6%.

Third, we find that a complementary policy, government insurance of pre-age-25 health risks, meaningfully improves the performance of long-term contracts. For flat lifetime net

[2]Similarly, long-term contracts perform worse for those who are more borrowing constrained.

income, full government insurance of pre-age-25 health risks increases the ex ante welfare impact of dynamic contracts (using the metric described above and averaging across possible age-25 health states) from 43.4% to 92.6%, averaging across initial health states. The average *one-time* subsidy needed to have all consumers achieve the same lifetime welfare as the healthiest age-25 consumer is $12,050 per consumer. For rising lifetime net income, full government insurance of these risks increases the welfare impact of dynamic contracts from 7.5% to 81.7%, with an average subsidy needed of $7,070 per consumer.

Fourth, we compare the welfare implications of dynamic contracts with full pre-age-25 government health status insurance to those of an ACA-like insurance exchange involving spot contracts with community rating that removes year-to-year reclassification risk. We find that, under flat lifetime income, dynamic contracts perform better (89% vs. 62% welfare gain compared to spot contracts) but that, under our calibrated rising income path, dynamic contracts perform slightly worse (74% vs. 83% welfare gain). Dynamic contracts have a lower % welfare gain as income profiles become steeper in the population, while for the ACA-like exchange the reverse is true.

Fifth, we investigate the role of several empirically relevant consumer choice factors that change the underlying implications of dynamic contracts. We find that switching costs in plan choice can have a positive welfare impact by making the environment more similar to one with two-sided commitment. We find, e.g., that, for consumers with flat net income paths, long-term contracts have higher welfare than an exchange-like environment when switching costs are greater than $1,970 (even without complementary age-25 insurance). A second choice factor we study is consumer myopia, which makes long-term contracts less effective by making front-loading more costly relative to future payouts. We find that, as consumers become more myopic, (i) contracts are less front-loaded (ii) lapsation rates are higher and (iii) myopia affects those with steeper income paths more than those with flatter income paths. But, we also find that, except for when myopia is very strong, it can have relatively small impacts on the welfare implications of long-term contracts.

Taken together, our results paint a detailed picture of the key economic factors that regulators should consider for a potential implementation of long-term contracts. While the optimal long-term contracts that we investigate always outperform basic spot contracts, when comparing to alternative policies supplementary government insurance of early-life health risk is a crucial ingredient for making long-term contracts more effective. Further, when we compare optimal long-term contracts to an ACA-like exchange, optimal long-term contracts perform better when consumers have relatively flat lifetime income paths. While one would want to estimate the micro-foundations of a specific context to draw a sharp recommendation, it is clear that supplementary government insurance of early-life health risks, flatter lifetime income paths, higher consumer switching costs, and lower myopia all help favor long-term contracts relative to the exchanges that saturate current policy.

Related Literature. This paper makes a range of contributions to the theoretical and empirical literatures on dynamic insurance contracts. Our theoretical results generalize work on optimal long-term contracts by Harris and Holmstrom (1983) and Hendel and Lizzeri (2003). Harris and Holmstrom (1983) study optimal long-term labor contracts when risk-averse workers and risk-neutral firms symmetrically learn workers' productivity over time. Aside from the difference in setting from labor markets to health insurance markets, our model allows for a more general stochastic process, which is essential for studying health insurance.[3] While Hendel and Lizzeri (2003) do allow for general health state transitions in their theoretical characterization, it is for a two-period problem, while our framework applies to contracts of arbitrary (finite) durations, which is crucial for studying long-run contracts in health insurance. Moreover, unlike these earlier papers, we provide analyses of (i) consumer self-selection into contracts (ii) consumer inertia and (iii) consumer myopia and also provide methods for computing optimal long-term contracts and assessing their welfare implications.[4]

In addition to the theoretical papers highlighted above, there are a number of related papers that study long-term health insurance empirically. Herring and Pauly (2006) conduct a calibration of guaranteed renewable contracts following the Pauly, Kunreuther and Hirth (1995) model, using data from the Medical Expenditure Panel Survey (MEPS). Browne and Hoffmann (2013) study the long-term contracts present in the German private health insurance (PHI) market and demonstrate that (i) front-loading of premiums generates consumer lock-in, (ii) more front-loading is associated with lower lapsation, and (iii) consumers that lapse are healthier than those who do not. Perhaps most relevant for this paper, Atal et al. (2020) apply our theoretical and computational results to study the welfare implications of the long-term insurance contracts offered in Germany. They find that, in the German context, long-term contracts lead to significant welfare gains above and beyond short-term community-rated contracts. Moreover, these contracts are very simple and are similar in structure to those in our model for flat net income paths. Finally, our analysis also relates to Fleitas et al. (2021), who use our analysis to support their study of the small group

[3]Our model's more general stochastic process, and our computational methods, could also enable study of implicit labor contracts insuring very general forms of productivity shocks.

[4]Our model of long-term health insurance contracting also relates closely to work by Pauly, Kunreuther, and Hirth (1995) and Cochrane (1995). Pauly, Kunreuther, and Hirth (1995) focus on contracts that ensure that an insured can renew future coverage at the same rates that the healthiest possible type would pay (which they term "guaranteed renewable"), while Cochrane (1995) proposes the use of "premium insurance" as a means of insuring against long-term negative shocks to health. In Section 2 and Appendix C, we discuss in depth the theoretical relation of our optimal contracts to those proposed in these two papers and describe the contributions we make relative to these papers. In addition, in those sections we assess the welfare these alternative contracts provide compared to our optimal long-term contracts.

health insurance market in the US, where they document limited pass-through of health risk reclassification onto small group premiums.[5,6] Aside from the fact that our theoretical analysis is used in these two papers, our empirical study is complementary to what these papers examine. Atal et al (2020) study the effects of long-term health insurance contracts in a market where they do exist. Fleitas et al (2021) test for optimal long-term contracts in a market where whether they exist is an empirical question. We examine the counterfactual welfare effects of such contracts in a market where they do not currently exist, highlighting key factors and comparative statics to consider when assessing the potential benefits of these contracts in a given context.

The rest of the paper proceeds as follows. Section 2 describes our model, theoretical results, and computational advances. Section 3 describes the Utah All-Payer Claims Data and our estimates of key primitives. Sections 4-6 describe our main empirical results related to (i) optimal contract characterization (ii) the welfare implications of dynamic contracts (iii) the welfare implications of government insurance for pre-age-25 health risk and (iv) comparisons of dynamic contracts to an ACA-like exchange. Section 7 studies the two empirical extensions of (i) switching costs and (ii) myopia, as well as a robustness analysis related to our risk aversion estimates. Section 8 concludes and considers our results in light of some of the potential practical impediments related to implementing dynamic contracts in practice.

2 Theory

We consider a dynamic insurance problem T periods long, with periods indexed $t = 1, ...T$. In our empirical analysis, periods represent years, with $t = 1$ corresponding to a 25-year old, and $T = 40$ corresponding to a 65-year old, when Medicare coverage would begin in the U.S.

The consumer may incur medical expenses $m_t \in \mathbb{R}$ in each period t, which are uncertain and motivate the desire for insurance. The consumer enters each period t characterized by his *health status* $\lambda_t \in H$, which determines the distribution of that period's medical expenses.

[5] Other related health insurance papers include Atal (2016), which studies the impact of lock-in to an insurance plan on the matching between individuals and health care providers in Chile, and Bundorf, Levin, and Mahoney (2012), who investigate the implications of reclassification risk in a large-employer context in a short-run environment.

[6] There are also a number of papers that empirically study long-term insurance in other insurance markets. Hendel and Lizzeri (2003) examine the structure of life insurance contracts and conclude that these contracts display the features of optimal contracts with one-sided commitment. Finkelstein, McGarry and Sufi (2005) study positive implications of long-term contracting in the context of long-term care markets, and show evidence of adverse retention, namely that healthier consumers lapse from contracts over time, leading to high average costs from those consumers that remain. See Hendel (2016) for a survey of the literature on long-term contracts and reclassification risk.

We take H to be a finite set. In our empirical work, greater λ_t will indicate sicker individuals, so that expected medical expenses $\mathbb{E}[m_t|\lambda_t]$ are strictly increasing in λ_t.

The evolution of the consumer's health status is stochastic, with the probability of health status λ_{t+1} given previous health history $\Lambda_t^1 \equiv (\lambda_1, ..., \lambda_t)$ given by $\overline{f}(\lambda_{t+1}|\Lambda_t^1)$. (Conditional on Λ_t^1, the realization of λ_{t+1} may be correlated with the realization of m_t.) We refer to Λ_t^1 as the consumer's *health state* at the start of period t, and denote the consumer's initial health state by Λ_1^1. The probabilities $\overline{f}(\cdot)$ give rise to the probability $f(\Lambda_{t'}^{t+1}|\Lambda_t^1)$ that any given continuation path $\Lambda_{t'}^{t+1} \equiv (\lambda_{t+1}, ..., \lambda_{t'})$ of health statuses will follow period t starting from health state Λ_t^1. In our empirical work, we will assume that health status transitions are governed by a second-order Markov process, so that $\overline{f}(\cdot)$ can be written as $\overline{f}(\lambda_{t+1}|\lambda_{t-1}, \lambda_t)$, but the results in this section hold more generally.

An individual's health state Λ_t^1 at the start of period t is observed by both the individual and all insurance firms, namely, there is symmetric information and symmetric learning.[7] We assume that the insurance market is perfectly competitive, with risk-neutral firms who discount future cash flows using the discount factor $\delta \in (0, 1)$. A consumer's risk preferences are described by $u(\cdot)$, the consumer's Bernoulli utility function, while the consumer's long-run expected utility is $\mathbb{E}[\sum_t \delta^t u(c_t)]$, where $c_t \in \mathbb{R}$ is the consumer's period t consumption level. Throughout, we assume that $u'(\cdot) > 0$ and that $u''(\cdot) < 0$, which motivates the consumer's desire for insurance. The consumer's income in period t is y_t, and evolves deterministically.[8] Throughout we assume that consumers are unable to borrow to fund premium payments or other expenses.[9] However, as we discuss in Appendix E, one can view the income path $y = \{y_t\}_1^T$ as *post-borrowing* income; so a consumer with a slower growth in y_t over time may be a consumer who is more able to borrow.

In what follows, we will sometimes refer to a consumer's income profile $y \equiv (y_1, ..., y_T)$ and risk preferences $u(\cdot)$ as the consumer's "type" $\theta \equiv (y, u)$.

2.1 Three Benchmarks

We will compare optimal dynamic contracts with one-sided commitment against three natural benchmarks. The first is the efficient, first-best allocation. In this setting, this outcome

[7]Our assumption that all insurers have access to the same information assumes that insurers can properly underwrite new customers. If, instead, an individual's current insurer had better information than other firms, prospective insurers would face an adverse selection problem when attempting to attract lapsing consumers. For the consequences of this type of adverse selection, see, for example, DeGaridel-Thoron (2005).

[8]The model readily generalizes to stochastic income, possibly dependent on the consumer's health status. In this case, the optimal contract would insure both health and income risk. See Appendix I for more details.

[9]A large literature documents credit constraints faced by consumers (Brunnermeier et al. (2012)). In practice, some of these capital market imperfections likely stem from similar factors to those that prevent consumers from committing to make large ex post payments to an insurer. See also the related discussions in the discussions in Diamond (1992), Cochrane (1995), and Pauly, Kunreuther, and Hirth (1995).

involves a constant consumption in all states and periods, equal to the annualized present discounted value of the consumer's "net income" from periods $t = 1$ to T (where the "net income" in period t equals period t income, y_t, less the expectation of period t medical expenses conditional on the consumer's health state Λ_1^1 at the first period of contracting, $\mathbb{E}[m_t|\Lambda_1^1]$). That is, it involves the constant consumption level

$$C^* = \left(\frac{1-\delta}{1-\delta^T}\right) \sum_{t=1}^{T} \delta^{t-1}(y_t - \mathbb{E}[m_t|\Lambda_1^1]). \tag{1}$$

As is well known, if consumers and insurance firms could both commit to a long-term contract given Λ_1^1, the competitive equilibrium would yield this outcome.

At the opposite extreme, long-term contracts may be impossible, leading to single-period "spot" insurance contracts. In a competitive market, in each period t such contracts will fully insure the consumer's within-period medical expense risks at a premium equal to $\mathbb{E}[m_t|\lambda_t]$, the consumer's expected medical expense given his period t health status λ_t. This results in the period t consumption level $y_t - \mathbb{E}[m_t|\lambda_t]$. Because the consumer's period t health status λ_t is ex ante uncertain, this outcome faces the consumer with risk from an ex ante perspective. Given Λ_1^1, the consumer's constant certainty equivalent of this uncertain consumption path is the constant consumption level CE_{SPOT} such that

$$u(CE_{SPOT}) = \left(\frac{1-\delta}{1-\delta^T}\right) \mathbb{E}[\sum_{t=1}^{T} \delta^{t-1} u(y_t - \mathbb{E}[m_t|\lambda_t])|\Lambda_1^1] \tag{2}$$

Finally, in this dynamic setting both insurance and consumption smoothing over time are needed to achieve the first best. Since we will focus on settings in which income y_t is increasing over time and (additional) borrowing is impossible, another natural benchmark is the outcome that would result if the consumer was fully insured within each period (eliminating all ex ante risk) but resources could not be transferred over time, leading to a premium payment in period t of $\mathbb{E}(m_t|\Lambda_1^1)$, the exact expected period-t medical costs at the start of period 1. This certain but time-varying consumption path results in the same welfare as the constant consumption level C^*_{NBNS} ("NBNS" = "No Borrowing/No Savings") such that

$$u(C^*_{NBNS}) = \left(\frac{1-\delta}{1-\delta^T}\right) \sum_{t=1}^{T} \delta^{t-1} u(y_t - \mathbb{E}[m_t|\Lambda_1^1])). \tag{3}$$

Compared to spot contracting, this benchmark eliminates reclassification risk without improving intertemporal allocation.

In our empirical work, we will consider both the expected utility that dynamic contracting and these benchmarks generate at age 25 at the start of contracting (i.e., conditional on Λ_1^1), and also the expected utilities that are implied at birth, factoring in the randomness of the consumer's age-25 health state Λ_1^1.[10]

2.2 Optimal Dynamic Contracts with One-Sided Commitment: Structure

We now turn to the setting in which competitive insurers can offer long-term contracts that they, but not consumers, are committed to. We assume that contracting begins in period 1 (in our empirical setting, at age 25) after Λ_1^1 has been realized. We can view a long-term contract as specifying the consumer's consumption in each period t, c_t, as a function of the consumer's publicly-observed health and medical expense history up through period t including period t's realization of m_t and λ_{t+1}, $[\Lambda_{t+1}^1, (m_1, ..., m_t)]$.[11] The insurer's profit in the period then equals the consumer's income y_t less the sum of period t medical expenses and period t consumption. The lack of commitment by the consumer, however, means that the consumer is free each period to change to another insurer who is offering the consumer better terms.

As in Harris and Holmstrom (1982), without loss of generality we can restrict attention when solving for the optimal contract to contracts in which the consumer never has an incentive to "lapse" in this way: since the new contract the consumer signs following any history must give his new insurer a non-negative expected discounted continuation profit, the consumer's initial insurer could include the same contract continuation in the initial insurance contract and weakly increase its expected discounted profit (lapsation would instead yield the initial insurer a continuation profit of zero). As a result, we can look for an optimal contract by imposing "lapsation constraints" that require that after no history is it possible to offer the consumer an alternative continuation contract that (i) itself prevents future lapsation, (ii) breaks even in expectation, and (iii) gives the consumer a higher continuation utility than in the original contract.

We take a recursive approach to solving this optimal contracting problem. At each date t, we can think of the state as a pair (Λ_t^1, S_t) where Λ_t^1 is the consumer's current health state (which determines future expected medical expenses), and S_t is the absolute value of

[10] When we examine at-birth welfare levels, we compute C^* and C^*_{NBNS} from an at-birth perspective, replacing $\mathbb{E}[m_t|\Lambda_1^1]$ with $\mathbb{E}[m_t]$ in (1) and (3). For spot contracting, we calculate the certainty equivalent annual consumption that generates the same age 25 to 65 welfare as the spot contracting regime, taking account of any risk aversion losses arising because of uncertainty over the consumer's age-25 health state, Λ_1^1. Similarly for the dynamic contracts we discuss in the next subsection, and the managed competition-style exchange which we examine in Section 6.

[11] This formulation assumes, for convenience, that the consumer cannot engage in hidden savings. While we will make this assumption initially, in the end we show that under the optimal contract the consumer has no desire to save. We could also allow consumption to be stochastic conditional on $[\Lambda_{t+1}^1, (m_1, ..., m_t)]$, but this will not be optimal.

the loss that the insurer is allowed to sustain going forward (i.e., S_t is the *subsidy* for future insurance).[12] This is a useful formulation for two reasons. First, after any history Λ_t^1 leading up to period t, continuation of the original contract generates some expected utility to the consumer and some expected loss S_t to the insurer. A necessary condition for an optimal contract, given the consumer's current health state, is that it is not possible to increase the consumers' continuation utility while keeping the insurer's loss equal to S_t. So, the continuation of the contract must itself solve an optimal contracting problem for an insurer who can sustain the loss S_t starting in health state Λ_t^1. Second, the constraint that the contract prevents lapsation can be viewed as saying that the consumer's continuation utility starting in any period t when in health state Λ_t^1 cannot be less than in an optimal contract offered by an insurer who must break even, i.e., who has $S_t = 0$. We denote a generic contract starting in period t for a consumer in health state Λ_t^1 by $c_{\Lambda_t^1}(\cdot)$ and the optimal contract for such a consumer whose type is θ when a subsidy S_t is available by $c_{\Lambda_t^1}^{\theta *}(\cdot|S_t)$. For simplicity, we often write an optimal contract that begins in period t with $S_t = 0$ as $c_{\Lambda_t^1}^{*}(\cdot)$.

More formally, consider the problem that arises if a firm faces a consumer of type θ in health state Λ_t^1 and can sustain, going forward, a (discounted expected) absolute loss of S_t. Let $B^{S_t}(\Lambda_t^1)$ denote the set of period t contracts in health state Λ_t^1 that break even in expectation with a subsidy of S_t if no lapsation occurs, and let $V_{\Lambda_t^1}^{\theta}(c_{\Lambda_t^1}(\cdot))$ be the consumer's discounted expected utility from contract $c_{\Lambda_t^1}(\cdot)$ starting in period t with health state Λ_t^1. In addition, let $c_{\Lambda_t^1|\Lambda_{t'}^{t+1}}(\cdot)$ denote the continuation of contract $c_{\Lambda_t^1}(\cdot)$ starting in period $t' > t$ if the health status realizations between period $t+1$ and t' are $\Lambda_{t'}^{t+1} \equiv (\lambda_{t+1}, ..., \lambda_{t'})$, resulting in a period t health state of $\left\langle \Lambda_t^1, \Lambda_{t'}^{t+1} \right\rangle \equiv (\lambda_1, ..., \lambda_t, \lambda_{t+1}, ..., \lambda_{t'})$. The optimal contract for a consumer of type θ is then described as follows:

Definition 1 *$c_{\Lambda_t^1}^{\theta *}(\cdot|S_t)$ is an optimal contract for a consumer of type θ signed in period t at health state Λ_t^1 with subsidy S_t if it solves the following maximization problem:*

$$\max_{c_{\Lambda_t^1}(\cdot) \in B^{S_t}(\Lambda_t^1)} V_{\Lambda_t^1}^{\theta}(c_{\Lambda_t^1}(\cdot)) \tag{4}$$

$$s.t.\;\; V_{\left\langle \Lambda_t^1, \Lambda_{t'}^{t+1} \right\rangle}^{\theta}(c_{\Lambda_t^1|\Lambda_{t'}^{t+1}}(\cdot)) \geq V_{\left\langle \Lambda_t^1, \Lambda_{t'}^{t+1} \right\rangle}^{\theta}(c_{\left\langle \Lambda_t^1, \Lambda_{t'}^{t+1} \right\rangle}^{\theta *}(\cdot|0)) \text{ for all } \Lambda_{t'}^{t+1} \text{ with } t' > t$$

Note that problem (4) provides a recursive definition of the optimal contract.

Our main characterization result, which we establish in Appendix A, is:

[12] Note that because only the health state Λ_t^1 matters for the distribution of future medical expenses, an optimal contract will not depend on previous medical expense realizations.

Proposition 1 *The optimal contract for a consumer of type θ starting in period t at health state Λ_t^1, denoted by $c_{\Lambda_t^1}^{\theta*}(\cdot)$, is fully characterized by (i) the zero-profit condition and, (ii) for all $t' > t$ and $\Lambda_{t'}^{t+1}$ such that $f(\Lambda_{t'}^{t+1}|\Lambda_t^1) > 0$, the condition that the consumer receives the following certain consumption level:*

$$c_{\Lambda_t^1}^{\theta*}(\langle \Lambda_t^1, \Lambda_{t'}^{t+1} \rangle) = \max\{c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1), \max_{\tau \in \{t+1,\ldots,t'\}} c_{\langle \Lambda_t^1, \Lambda_\tau^{t+1} \rangle}^{\theta*}(\langle \Lambda_t^1, \Lambda_\tau^{t+1} \rangle)\}. \tag{5}$$

Under this contract, the consumer does not wish to secretly save.

In words, the optimal contract $c_{\Lambda_t^1}^{\theta*}(\cdot)$ signed in period t at health state Λ_t^1 offers in each period $t' > t$ after history $\Lambda_{t'}^1 = \langle \Lambda_t^1, \Lambda_{t'}^{t+1} \rangle$ the maximum among the first-period consumption levels offered by all the equilibrium contracts available along the way on continuation health history $\Lambda_{t'}^{t+1}$. Thus, applying this result to the initial contracting in period 1, the optimal contract starting in period 1 offers an initial *consumption floor*, which is then bumped up to "match the market" in later periods $t > 1$ each time the consumer reaches a state in which the market would offer a higher initial consumption floor. The equilibrium contract provides full within-period insurance for the consumer (i.e., consumption in each period is independent of m_t), and partial insurance against reclassification risk, as consumers who have experienced sufficiently bad health states leading up through a given period t [i.e., such that $c_{\Lambda_\tau^1}^{\theta*}(\Lambda_\tau^1) \leq c_{\Lambda_1^1}^{\theta*}(\Lambda_1^1)$ for all $\tau \leq t$] all enjoy the same level of consumption regardless of differences in their period-t health states. Since consumption is always weakly rising over time, the consumer never wishes to save.[13]

To understand the forces leading to Proposition 1, consider the two-period example (with $T = 2$) shown in Figure 1. The consumer starts period 1 in health state $\Lambda_1^1 = \lambda_1$, and can transition to one of two possible period-2 health states, the healthier state $\Lambda_2^{1\prime\prime} = (\lambda_1, \lambda_2'')$ or the less healthy state $\Lambda_2^{1\prime} = (\lambda_1, \lambda_2')$, where $\lambda_2'' < \lambda_2'$. Because the consumer is healthier, the consumer's period-2 outside option has a higher consumption level in state $\Lambda_2^{1\prime\prime} = (\lambda_1, \lambda_2'')$ than in state $\Lambda_2^{1\prime} = (\lambda_1, \lambda_2')$, as indicated by the strict inequality sign on the right side of the figure.

Intuitively, if the lapsation constraint was not binding in the healthier state $\Lambda_2^{1\prime\prime} = (\lambda_1, \lambda_2'')$ but was binding in the less healthy state $\Lambda_2^{1\prime} = (\lambda_1, \lambda_2')$ then consumption would be higher in the healthier state despite the fact that it would be possible to transfer resources to the less healthy state and equalize consumption, raising the consumer's discounted expected utility from the contract. This cannot happen in an optimal contract. Observe as well that consumption in period 1 must equal consumption in the less healthy state if the lapsation constraint does not bind in this state, for otherwise resources could be transferred across

[13]For a formal derivation, which requires specifying what happens if a consumer with hidden savings seeks to buy insurance from a new firm, see Appendix B.

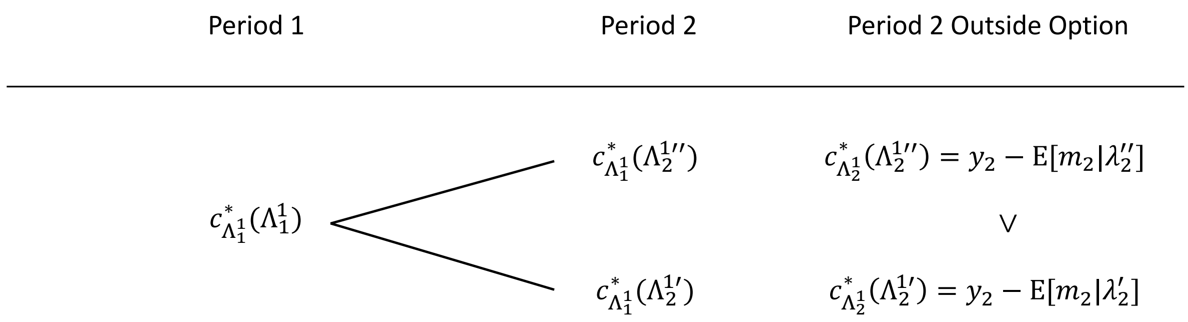

Figure 1: A two-period illustration of the economic forces leading to Proposition 1. The figure shows the contracts that would arise if the no-lapsation constraint was only binding in the less healthy state $\Lambda_2^{1\prime} = (\lambda_1, \lambda_2^\prime)$, where $\lambda_2^\prime > \lambda_2^{\prime\prime}$.

periods to equalize these two consumption levels, again raising the consumer's discounted expected utility from the contract.

More generally, the same logic implies that, in any period of an optimal contract, any state at which the no-lapsation constraint binds must have greater consumption than any state in which it does not bind. Moreover, if we had more than two period-2 states, *every* state in which the constraint does not bind must have the same consumption, which must also equal the consumption level in period 1. With many periods, this same structure exists across all periods: consumption remains constant until the no-lapsation constraint binds, at which point it jumps up, and then stays constant again until the next time the no-lapsation constraint binds.

In this two-period setting, the insurer either sustains a loss or breaks even in expectation in period 2. Thus, for it to also break even in expectation overall, it has to make a positive expected profit in period 1. That is, the consumer initially pays *more* than his expected healthcare expenses. This "front-loaded" amount funds the consumption guarantees the consumer will enjoy in the future. This is key in incentivizing the consumer–who cannot commit to the contract–to stay with the insurer. As we show in the empirical analysis, a steeply rising income over time hurts the performance of dynamic contracts by increasing the tension between front-loading and inter-temporal consumption smoothing.

The guaranteed consumption levels in Proposition 1 are the counterpart for dynamic health insurance contracts of the downwardly-rigid wages in Harris and Holmstrom (1983)'s study of implicit labor contracts, where worker and firm are both learning about the worker's productivity parameter over time from observations of the worker's Normally-distributed output. Relative to their result, aside from the difference in setting, Proposition 1 allows for a much more general stochastic process than the learning process in Harris and Holmstrom's analysis, which is necessary for the study of health insurance.

The optimal contract in Proposition 1 specifies the consumer's consumption levels for each possible health history and prevents lapsation. Importantly, however, the same outcome can alternatively be achieved by means of a much simpler *guaranteed premium path contract* from which the consumer may lapse. Specifically, the consumer is given the option to renew, if he has not yet lapsed, at the guaranteed premium path $p^{\theta*}(\Lambda_1^1) \equiv (p_1^{\theta*}, ..., p_T^{\theta*})$ where $p_t^{\theta*} = y_t - c_{\Lambda_1^1}^{\theta*}(\Lambda_1^1)$ for $t = 1, ...T$, provided that he has always renewed in the past.[14] That is, the guaranteed premium path keeps consumption constant over time, equal to $c_{\Lambda_1^1}^{\theta*}(\Lambda_1^1)$, as long as the consumer sticks with the contract. But if the consumer arrives in a period t with a sufficiently good health state Λ_t^1, he may choose not to renew, instead signing a contract with a new insurer (or renegotiating a contract with the current insurer) that offers guaranteed premium path $p^{\theta*}(\Lambda_t^1) \equiv \{y_\tau - c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1)\}_{\tau \geq t}$ where $c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1) > c_{\Lambda_1^1}^{\theta*}(\Lambda_1^1)$. Such lapses have no effect on the profit of the consumer's initial insurer as that firm was indifferent about whether to match the outside offer.[15]

2.3 Optimal Consumption Guarantees: Characterization and Computation

Proposition 1 describes the structure of an optimal dynamic contract as involving evolving consumption guarantees. The level of these guarantees is then determined by the condition that the insurers offering them must break even. Determining the guarantees that break even, however, is a recursive problem, because at each point in time and health state the profit an insurer earns by offering a guarantee depends on the guarantees that competitive insurers may be willing to offer the consumer in the future. In this section, we describe this recursive condition and note a striking implication of it: the optimal contract does not

[14] This form of contract is the counterpart to the "Annual Renewable Term" life insurance contracts studied in Hendel and Lizzeri (2003).

[15] The recursive formulation also makes clear that this equilibrium outcome can be achieved instead with single-period contracts. A consumer in period 1 with health state Λ_1^1 could purchase a contract that covers all period-1 medical expenses, and that in addition pays the consumer at the start of period 2 the amount that the optimal contract implicitly subsidizes the realized continuation state Λ_2^1. This amount would allow the consumer to buy the long-term continuation contract on the open market. Upon reaching period 2, however, the consumer could instead again buy a one-period policy of this type, and could continue in this manner until period T. [This approach to replicating a long-term contract with a series of short-term contracts is reminiscent of Fudenberg et al. (1990), although our setting is not captured in their model because of the presence of lapsation constraints and the consumer's inability to borrow.]

As noted in Cochrane (1995), such short-term contracts avoid the consumer being locked into an insurer, perhaps resulting in better insurer performance as well as better matching of insurers and consumers when health care networks are bundled with insurance provision. However, such contracts may require that courts can verify the consumer's health state Λ_t^1, while guaranteed premium path contracts do not.

depend on the consumer's risk preferences, $u(\cdot)$, as long as the consumer is risk averse. This condition also serves as the basis for computing optimal contracts in our empirical analysis.

For ease of notation, in this section we denote the initial consumption guarantee offered to a consumer signing a contract in period t with health state Λ_t by $c_t^{\theta *}(\Lambda_t)$. For concreteness, and also anticipating our empirical analysis, we assume in our discussion a second-order Markov process where $\Lambda_t = (\lambda_{t-1,} \lambda_t)$, although our observations here fully generalize. We assume that λ_t can take K possible values in each period t.

Intuitively, for a consumer of type θ, we can first derive, for each possible last-period state $\Lambda_T = (\lambda_{T-1,} \lambda_T)$, the last-period consumption levels $c_T^{\theta *}(\Lambda_T) = y_T - \mathbb{E}[m_T|\lambda_T]$ that would be offered to a consumer in state Λ_T by competitive firms. We then look at each possible state $\Lambda_{T-1} = (\lambda_{T-2,} \lambda_{T-1})$ in period $T-1$. We find $c_{T-1}^{\theta *}(\Lambda_{T-1})$ by doing a binary search over possible values for the consumption guarantee c_{T-1}, looking for the largest c_{T-1} that generates non-negative profits for the insurer, taking account of the fact that the consumer will yield the insurer continuation profits of zero in those states Λ_T in which $c_T^{\theta *}(\Lambda_T) > c_{T-1}$ since the guarantee is either bumped up to match the market or the consumer lapses. We then continue backward in this fashion, with the transitions $f(\cdot)$ being used to generate probabilities that the consumer is in each possible state at each future date (which also generates the probability that the consumer will have lapsed by that date).

More formally, enumerate the K^2 possible combinations of $(\lambda_{t-1,} \lambda_t)$ for each period t by $\{\Delta^s = (\lambda_{t-1,}^s \lambda_t^s)\}_{s=1}^{K2}$. For each period t, we denote by C_t^θ the $(T-t+1) \times K^2$ matrix of first-period consumption guarantees whose (τ, s) element for $\tau \geq t$ and $s \in \{1, ..., K^2\}$ is $c_\tau^{\theta *}(\Lambda^s)$, a consumption guarantee that breaks even for a contract starting in period τ with health state $\Lambda_\tau = \Lambda^s$, given the future guarantees described in Proposition 1 (which are, themselves, contained in C_{t+1}^θ). We start at $t = T$ where, as noted above, $c_T^{\theta *}(\Lambda_T = \Lambda^s) = y_T - \mathbb{E}[m_T|\lambda_t^s]$. This gives us C_T^θ. We then proceed iteratively backwards, deriving C_t^θ given C_{t+1}^θ and the transition probabilities. Specifically, C_t^θ adds an additional row to C_{t+1}^θ; each element (t, s) of this row is the consumption guarantee $c_t^{\theta *}(\Lambda_t = \Lambda^s)$. We derive this guarantee by doing a binary search to find the (unique) value $\bar{c}$ that sets the insurer's expected profit to zero. A key observation that dramatically simplifies computation of the insurer's expected profit given a value of $\bar{c}$ is that, whenever the guarantee is bumped up, the insurer earns an expected continuation profit of zero from that point on. This fact leads to the following lemma characterizing the insurer's expected discounted profit from a consumption guarantee:

Lemma 1 *If consumption guarantee $\bar{c}$ is offered at health state Λ^s, then the expected discounted profit to the insurer will be given by:*

$$\{y_t - \mathbb{E}(m_t|\lambda_t = \lambda_t^s) - \overline{c}\} + \left\{ \sum_{\tau=t+1}^{T} \delta^{\tau-t} \sum_{z=1}^{K^2} [y_\tau - \mathbb{E}(m_\tau|\lambda_\tau = \lambda_t^z) - \overline{c}] \cdot P_\tau(z|\Lambda_t = \Lambda^s, C_{t+1}^\theta, \overline{c}) \right\}, \tag{6}$$

where $P_\tau(z|\Lambda_t = \Lambda^s, C_{t+1}^\theta, \overline{c})$ is the probability that, starting in health state $\Lambda_t = \Lambda^s$ in period t, the health state transitions $(\Lambda_{t+1}, ..., \Lambda_\tau)$ from period t to period τ are such that $\Lambda_\tau = \Lambda^z$ and $c_{t'}^{\theta*}(\Lambda_{t'}) \leq \overline{c}$ for all $t' \in \{t+1, ..., \tau\}$.[16]

Proof of Lemma 1. The first term in curly brackets in (6) is the expected profit in period t given the initial consumption guarantee $\overline{c}$, while the second term in curly brackets is the expected continuation profit from future periods at which the consumption level is still equal to $\overline{c}$. Note that at any health state Λ_τ such that $c_\tau^{\theta*}(\Lambda_\tau) > \overline{c}$ (i.e., the lapsation constraint binds), the insurer's continuation payoff is zero. This is because at any health state where the lapsation constraint binds, the insurer is offering the optimal contract for that state, which we know is zero-profit. Therefore, Equation (6) fully captures the continuation profit to the insurer by taking account of all future health states $\Lambda_{t'}$ at which $c_{t'}^{\theta*}(\Lambda_{t'}) \leq \overline{c}$. ■

Once we compute $c_t^{\theta*}(\Lambda_t = \Lambda^s)$ for all K^2 values of Λ_t, we continue in this iterative manner until we have derived C_1^θ, whose first row gives the initial consumption guarantees offered in period 1 to consumers in each of the K^2 possible period-1 health states.

A striking, and perhaps surprising, feature of condition (6) is that *it does not depend on the degree of the consumer's risk aversion embodied in* $u(\cdot)$ That is, the fact that the consumer is risk averse matters – and is used to prove Proposition 1 – but the optimal contract is the same for any two risk averse consumers who have the same stochastic health process and income path.

Finally, note that, in general, an optimal dynamic contract could be an extremely complicated object, specifying consumption levels along any possible health history (of which there are $\Sigma_{\tau=1}^{T} K^\tau$ ones), potentially making the computation of an optimal contract an intractable problem. However, our theoretical analysis has yielded two extremely useful properties for computation: first, the characterization in Proposition 1 of optimal contracts in terms of consumption guarantees and, second, the fact that the continuation profit of the consumer is zero whenever the consumption guarantee is bumped up to match the market (or, equivalently, if the consumer lapses).

[16]These probabilities are computed using C_{t+1}^θ and the transitions $\overline{f}(\cdot)$.

2.4 Comparison to Pauly, Kunreuther, and Hirth (1995) guaranteed renewable contracts

In an early analysis of the potential for dynamic health insurance contracts, Pauly, Kunreuther, and Hirth (1995) (PKH) proposed what they called "guaranteed renewable contracts" as a solution to prevent reclassification risk.[17] In contrast to the optimal long-term contracts described in Proposition 1, PKH aimed to design a policy that provides full insurance in each period and guarantees that the consumer can renew in the future at the same premium as would be offered to the healthiest consumer type at that age. The idea is that a consumer with such a policy never wishes to lapse and faces no uncertainty in their consumption (i.e, no reclassification risk).

To understand these contracts, consider the simplest case in which $T = 2$ for a consumer who starts off in the healthiest possible state $\Lambda_1^1 = (1)$ (For simplicity, we assume a second-order Markov process, and write the relevant state in period $t > 1$ as $\Lambda_t = (\lambda_{t-1}, \lambda_t)$. We consider the general case in Appendix C.) In period 2 (the last period), the consumer pays a premium equal to $p_2 = \mathbb{E}[m_2|\Lambda_2 = (1,1)]$, the expected medical expenses of the healthiest possible period-2 consumer. In period 1, the consumer pays a premium of

$$\begin{aligned} p_1 &= \mathbb{E}[m_1|\Lambda_1 = (1)] + \delta\{\mathbb{E}[m_2|\Lambda_1 = (1)] - \mathbb{E}[m_2|\Lambda_2 = (1,1)]\} \\ &= \mathbb{E}[m_1|\Lambda_1 = (1)] + \delta \Pr(\Lambda_2 \neq (1,1)|\Lambda_1 = (1))\{\mathbb{E}[m_2|\Lambda_2 \neq (1,1)] - \mathbb{E}[m_2|\Lambda_2 = (1,1)]\} \end{aligned}$$

The first term is the consumer's expected period-1 medical costs (since he starts with $\Lambda_1 = (1,1)$), while the second term is the prepayment of the expected period-2 discount being offered to the consumer (which he enjoys when it turns out that $\Lambda_2 \neq (1,1)$). This prepayment is necessarily (weakly) greater than the prepayment arising in the optimal contract, which promises lower period-2 consumption than the PKH contract in all but the healthiest period-2 state and therefore needs less front loading to enable the insurer to break even.

Unlike the optimal contracts described in Proposition 1, the PKH contracts do not optimally balance the benefits of reducing reclassification risk against the costs of front-loading; for example, the PKH contract is unaffected by a consumer's income profile. The PKH contracts go to the extreme of completely preventing reclassification risk, resulting in a fully deterministic consumption profile but excessively low initial consumption.[18]

[17] Pauly, Kunreuther, and Hirth (1995) refer to their policies as guaranteed renewable contracts, but (as they note) effectively treat them as guaranteed premium path contracts. Actual "guaranteed renewable" contracts often instead merely state that the consumer has a right to renew at a rate at the insurer's discretion, but that must be the same as what the insurer offers to all other consumers in the same policy.

[18] Cochrane (1995) proposes a different insurance scheme to protect consumers from reclassification risk: premium insurance. We discuss this scheme in Appendix C.

2.5 Unobserved Types and Self-Selection

The analysis above assumed that a consumer's lifetime income profile $y = (y_1, ..., y_T)$ and risk aversion, captured in the Bernoulli utility function, were known by both the consumer and all insurers. In reality, this is unlikely to be the case, which could, in principle, pose an important obstacle to these contracts' practical use. In this subsection we show that insurers' failure to possess this information poses no such problem. Specifically, we show that if offered the collection of optimal contracts for all types derived above, presented as guaranteed premium path contracts, consumers will self-select, choosing the optimal contract for their type.[19,20]

Specifically, suppose that there is a set Θ of types in the market where, to recall, a consumer's type $\theta = (y, u)$ includes his income path and risk preferences.[21] As above, a guaranteed premium path contract is a $p = (p_1, ..., p_T)$ that allows the consumer to continue coverage in period t paying premium p_t provided that he has not previously lapsed. As described above, the optimal guaranteed premium path contract for a known type θ starting in period t when the consumer's health state is Λ_t^1 is denoted by the path $p^{\theta*}(\Lambda_t^1) \equiv \{y_\tau - c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1)\}_{\tau \geq t}$, a path that keeps consumption constant [equal to $c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1)$] as income changes from year to year.

Our result is:

Proposition 2 *Suppose that, in each period $t = 1, ..., T$, the menu of optimal guaranteed premium path contracts $\{p^{\theta*}(\Lambda_t^1)\}_{\theta \in \Theta}$ is offered to a consumer in health state Λ_t^1, where $p^{\theta*}(\Lambda_t^1) \equiv \{y_\tau - c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1)\}_{\tau=t}^T$. Then in each period the menu is self-selective and induces no secret savings: that is, if a consumer of type θ agrees to a new contract he chooses that type's optimal contract $p^{\theta*}(\Lambda_t^1)$ and does not secretly save.*

We prove Proposition 2 in Appendix B. The proof proceeds by showing that an insurer's profit is non-negative if a consumer of type θ' chooses the guaranteed premium path contract intended for another type θ. To see the rough idea of why, note that a difference in insurer profit under a guaranteed premium path contract can arise only because of differences in lapsation behavior of the two types. In any state in which type θ would not lapse under its contract, lapsation by type θ' raises insurer continuation profits, which are non-positive

[19]Our discussion above showed that the optimal contract does not depend on the consumer's level of risk aversion embodied in $u(\cdot)$. However, it is still possible that consumers' misrepresentations could depend on $u(\cdot)$, and so we suppose here that both y and u are private information of the consumer.

[20]Note that contracts that instead present the optimal contracts as guaranteed consumption levels (as in Proposition 1), would clearly not induce self-selection as consumers with low lifetime incomes would choose contracts intended for consumers with high lifetime incomes.

[21]Formally, we allow Θ to include all possible income paths $y \in \mathbb{R}_+^T$ to allow for the possibility of secret savings (see footnote 56 in Appendix B).

when facing type θ. On the other hand, in states in which type θ *would* lapse, it is to a guaranteed premium path contract that has lower premiums and breaks even for the new insurer, so the original insurer would have earned non-negative continuation path profits if type θ had stayed (at the higher premiums); an induction argument shows that it also earns non-negative continuation path profits when type θ' stays. With this fact established, Proposition 2 follows: since the contract intended for type θ breaks even for the insurer if type θ' chooses it, it is a feasible contract in the original problem, and so type θ' cannot be better off choosing it than choosing the contract intended for it.[22]

Since insurers cannot offer any type of consumer a greater value than in the optimal contract and still break even, Proposition 2 implies that it is an equilibrium for this menu of contracts to be offered, which results in the same allocation as if consumer types were perfectly observable.[23]

2.6 Consumer Inertia and Myopia

The theory developed above assumed that consumers evaluate long-term insurance contracts according to a canonical rational forward-looking framework. In this section, we extend our results to consider two forms of imperfect assessment by consumers. First, recent evidence suggests that consumers may exhibit substantial inertia in their health insurance choices [see, e.g., Handel (2013)]. We can extend our analysis to consider the effects of consumer inertia, which we model by introducing a switching cost that creates a consumption loss of $\sigma > 0$ if the consumer lapses and switches insurers. The key change this introduces is straightforward: the inequality in the lapsation constraint in the period t problem (4) becomes $V^{\theta}_{\langle \Lambda^1_t, \Lambda^{t+1}_{t'} \rangle}(c_{\Lambda^1_t | \Lambda^{t+1}_{t'}}(\cdot)) \geq V^{\theta}_{\langle \Lambda^1_t, \Lambda^{t+1}_{t'} \rangle}(c^{\theta *}_{\langle \Lambda^1_t, \Lambda^{t+1}_{t'} \rangle}(\cdot| - \sigma))$, since an insurer seeking to induce the consumer to lapse must now incur the cost σ to compensate the consumer for his switching cost.

Second, consumers may exhibit myopia when evaluating insurance contracts that span many years into the future. Myopia has been oft-studied in life-cycle models and can be modeled in myriad ways. We follow the myopia model used in many papers (see, e.g., Aguiar et al. (2020)) and assume that consumers making decisions apply a lower discount factor β than the true (welfare-relevant) discount factor δ modeled earlier in this section.

[22] The proof also establishes that the optimal contracts do not induce secret savings, while also allowing that deviating types might secretly save when choosing contracts not intended for them.

[23] The asymmetric information analysis conducted in this section differs from that in part of the macroeconomics literature on dynamic contracting with two-sided commitment. For example, Atkeson and Lucas Jr (1992) examines efficiency in an environment where there is asymmetric information about "endowment," the equivalent of "health state" in our model. We, on the other hand, assume symmetric information on health states. The equivalent of "income paths" in our model, which is the object of asymmetric information in this section, does not exist in Atkeson and Lucas Jr (1992).

This simple specification means that consumers overweight current utility compared to future utility relative to what their non-myopic selves would want.[24]

To extend our result to incorporate this possibility, we introduce the following definition:

Definition 2 *The "t-period myopic consumption transformation function" is*

$$\psi_t(\cdot) \equiv u'^{-1}[(\frac{\beta}{\delta})^t \times u'(\cdot)]$$

where $u(\cdot)$ is the consumer's Bernoulli utility function.

Note that $\psi_t(\cdot)$ is an increasing function, that $\psi_t(c) \geq c$, and that ψ_t collapses to the identity function if $\beta = \delta$. With this definition, we can extend Proposition 1 as follows to allow for both inertia and myopia:

Proposition 3 *The optimal contract for a consumer of type θ starting in period t at health state Λ_t^1, denoted by $c_{\Lambda_1^1}^{\theta*}(\cdot)$, is fully characterized by the zero-profit condition and, for all $t' > t$ and $\Lambda_{t'}^{t+1}$ such that $f(\Lambda_{t'}^{t+1}|\Lambda_t^1) > 0$, the condition that the consumer receives the following certain consumption level:*

$$\psi_{t'}\left(c_{\Lambda_t^1}^{\theta*}(\langle \Lambda_t^1, \Lambda_{t'}^{t+1} \rangle)\right) = \max\{\psi_t\left(c_{\Lambda_t^1}^{\theta*}(\Lambda_t^1)\right), \max_{\tau \in \{t+1,\ldots,t'\}} \psi_\tau\left(c_{\langle \Lambda_t^1, \Lambda_\tau^{t+1} \rangle}^{\theta*}(\langle \Lambda_t^1, \Lambda_\tau^{t+1} \rangle \,|\, -\sigma)\right)\}. \quad (7)$$

Under this contract, the consumer does not wish to secretly save.

To understand the changes from Proposition 1, consider first inertia. As noted above, inertia makes it less attractive to lapse, as the consumer now incurs cost σ. Firms seeking to induce consumers to switch, must effectively cover this cost, so that a new lapsation-inducing contract is effectively starting with the negative subsidy $-\sigma$. This negative subsidy appears in the last term in expression (7). Nonetheless, Proposition 3 shows that the basic structure of an optimal contract is unchanged when inertia is present and continues to include consumption guarantees. Switching costs enable taking resources away from healthy states in which consumers would have lapsed absent the cost of switching, into bad health states. In Section 7.1 we analyze inertia empirically, and show that it indeed increases consumer welfare from dynamic contracts, achieving the first best if the switching cost is large enough.

[24] An oft-used alternative specification for myopia is present-bias (also called present-focus) where consumers have traditional discounting between all future periods but overweight current utility relative to all future utility (O'Donoghue and Rabin, 2015). Gottlieb and Zhang (2021) study present-bias and dynamic inconsistency in a stylized model of long-term contracting and show that, with one-sided commitment, as the time-horizon of contracting grows the inefficiency generated from present-bias goes to 0. We have computed the Gottlieb-Zhang optimal contracts for our setting and find almost no welfare differences from our baseline fully rational model.

Myopia introduces the transformation functions in expression (7). To understand how myopia changes the optimal contract, consider again the two-period case we examined in Figure 1. Now, when the lapsation constraint does not bind in a period-2 state (λ_1, λ_2'), equalizing marginal utility across periods 1 and 2 no longer involves equal consumption levels: instead consumption should be lower in period 2 than in period 1, satisfying the condition that $(\beta/\delta)u'(c^*_{\Lambda_1^1}(\lambda_1, \lambda_2)) = u'(c^{\theta*}_{\Lambda_1^1}(\lambda_1))$, or equivalently, $\psi_1(c^*_{\Lambda_1^1}(\lambda_1, \lambda_2)) = c^*_{\Lambda_1^1}(\lambda_1)$. However, the same argument that we gave using Figure 1 continues to hold: in the optimal contract, states in which the no-lapsation constraint binds must have greater consumption than states in which it does not bind. Two differences, though, are that myopia (i) leads to optimal guaranteed premium path contracts with premiums that rise faster than the rate of income growth and (ii) leads to increased lapsation rates (with guaranteed-premium-path contracts.) In our empirical analysis in Section 7.2 we will show that myopia indeed reduces the performance of dynamic contracts although they still lead to a non-trivial increase in consumer welfare above the level with spot contracts even under significant myopia (i.e., small β).

3 Empirical Analysis: Setup

Our empirical analysis has two roles. First, it shows how our theoretical model and computation method enable empirically assessing the welfare implications of dynamic contracts with one-sided commitment. Second, and more importantly, our analysis provides us with a set of key empirical takeaways on the potential effects of long-term health insurance contracts and complementary policies.

Specifically, we quantify the welfare implications of a counterfactual system-wide individual market with dynamic contracts where the population has the stochastic health profiles of the entire population of men aged 25-64 in Utah and has risk aversion levels estimated in prior research. This comprehensive health data for an entire population, combined with our model of dynamic contracts, allows us to empirically analyze comparative statics and regulatory decisions that are crucial for understanding the potential benefits of dynamic contracts across a range of counterfactual policy environments.

We have four primary goals in our empirical analysis. The first is to quantify the structure of long-term contracts for consumers as a function of age, income, health status, and risk-aversion. This analysis directly leverages theory results from Section 2, combined with estimated sample micro-foundations described in more depth below. Next, our second is to study the normative implications of these contracts with a focus on how the welfare effects vary depending on a consumer's initial age-25 health state. Understanding the relative and absolute magnitudes of these welfare effects is crucial for assessing the distributional consequences of dynamic contracts as well as the potential benefits of complementary regulations.

Our third goal is to understand the quantitative importance of how the steepness of lifetime income paths impact the potential normative benefits of long-term contracts. Conceptually, steeper income paths are likely to make dynamic contracts less attractive because of the front-loading those contracts require. However, it is important to assess the quantitative importance of this comparative static in order to understand the welfare benefits of these contracts for different types of consumers and for different empirical contexts. As part of this goal, we also want to assess how steepness of income paths over the life-cycle interacts with the initial health state when arriving to the market, to determine the relative importance of these factors.

Our fourth goal is to study how complementary policies can improve the normative benefits of long-term contracts. We focus on a policy that insures consumers against age-25 health risk and investigate (i) the amount of insurance needed to achieve certain welfare benchmarks and (ii) the cost of this insurance to the regulator / taxpayers. Understanding the social costs and benefits for different levels of age-25 health status insurance is an important ingredient for policymakers thinking about implementing dynamic contracts in practice.

In addition to these four primary goals we have several secondary goals in our empirical analysis including (i) comparing dynamic contracts to an ACA-style managed competition year-to-year individual market (ii) understanding the potential importance of switching costs for our positive and normative conclusions (iii) understanding the importance of myopia for our positive and normative conclusions and (iv) understanding the sensitivity of our results to different risk aversion.

3.1 Data and Context

To predict equilibrium contracts and welfare under each regime we need four basic ingredients: (i) expected medical costs conditional on an individual's health status, (ii) the transitions across health states as individuals age, (iii) preferences towards risk, and (iv) income profiles.

We focus on the sample of men that appear in the all-payer claims data from the State of Utah for the years 2013-2015.[25] This dataset includes detailed medical claims for each individual in the state of Utah except for those individuals enrolled in traditional Medicare or those who are uninsured. Our analysis studies all men in the data who (i) are 25-64

[25] These data are utilized as well in Lavetti et al. (2018), which contains a more complete description of the data.

Sample and Total Health Expenditure		
Ages	Population	Mean
All	212,265	4,650
25-29	14,872	2,302
30-34	26,412	2,520
35-39	31,414	2,871
40-44	29,864	3,424
45-49	27,424	4,259
50-54	31,082	5,659
55-59	31,574	7,339
60-64	17,131	8,964

Table 1: Sample statistics for (i) the entire sample of men aged 25-64 used in our equilibrium analysis and (ii) 5-year age buckets within that sample. For each relevant group, "Population" column reports the number of individuals, and "Mean" column reports the average medical cost in 2015.

years old throughout the three-year sample period and (ii) appear in the data each month throughout the sample period (e.g., they have no spells of non-insurance).[26] We make the latter restriction in order to ensure that we can cleanly capture health status transitions, as described in more detail later. We focus on men here for simplicity, since men and women have distinct stochastic health processes. We have also performed our analysis for women, finding broadly similar results for contract structure and welfare.

Table 1 describes our final sample of Utah men, with key descriptive statistics broken down by age. Our sample has 212,265 men who averaged $4,650 in total medical spending in 2015.

3.2 Health States

The most essential part of the data is the available information on the diagnostics (ICD-9 codes) of each individual in the sample. We feed the diagnostic codes as well as other demographics into the ACG software developed at Johns Hopkins Medical School to create individual-level measures of predicted expected medical expenses for the upcoming year

[26]We use age 25 because it is right at / just below the age where children have to be removed from their parents' health plans in the current marketplace. That said, there is no specific other reason that the dynamic contracts market / analysis has to start at this age as opposed to some other nearby age. Since there are only very small differences between the distribution of health states and health-status transitions in the ages around 25, our analysis would be essentially unchanged if starting at another age nearby. Broadly, our key goal is to choose this starting age so that it is young enough that most individuals enter the market in good health, allowing for high pooling / insuring of reclassification risk. See our analysis of pre-age-25 insurance in Section 5 for an extended discussion.

Health Status by Age							
Age	1 (Healthy)	2	3	4	5	6	7 (Sick)
25-29	0.55	0.15	0.11	0.07	0.05	0.04	0.05
30-34	0.44	0.17	0.14	0.08	0.06	0.06	0.05
35-39	0.37	0.17	0.14	0.10	0.08	0.07	0.07
40-44	0.31	0.16	0.15	0.12	0.10	0.08	0.08
45-49	0.17	0.12	0.18	0.16	0.14	0.12	0.11
50-54	0.11	0.10	0.15	0.17	0.17	0.15	0.15
55-59	0	0.11	0.10	0.18	0.19	0.21	0.20
60-64	0	0.08	0.08	0.17	0.19	0.23	0.25
Actuarial costs	837	1,376	1,973	3,052	4,356	6,840	20,507

Table 2: Health status by age in 2015 for our sample, where consumers are divided into 7 bins of their predicted medical spending (determined by their Johns Hopkins ACG predictive score) for the year ahead. The "Actuarial costs" row reports how expected expenses in the upcoming year vary across the consumer health status bins.

relative to the mean of the population.[27] The output is an index that represents the health status of each individual in the population. Since the ACG is used by insurers in their underwriting processes, our empirics are based on similar information about risks that market participants (insurers) have. We denote the ACG index by λ and we refer to λ_{it} as individual i's "health status" at time t.

To ensure meaningful support when we estimate transition matrices, we partition the health statuses into seven mutually exclusive and exhaustive bins that each contain one-seventh of the final sample. Table 2 shows the proportion of individuals in each age group in each of these seven health categories, with bin 1 being the healthiest, and bin 7 being the sickest. In its last row it also shows the expected expenses corresponding to each bin: an individual in the healthiest bin has expected annual medical expenses of $837, while someone in the sickest bin has expected annual medical expenses of $20,507.

3.3 Health State Transitions

The second key input into our empirical analysis is health transitions over time. We model transitions in health status as a second-order Markov process in which the distribution of an individual i's period $t+1$ status $\lambda_{i,t+1}$ is conditional on his health status in the previous two years, $\Lambda_t \equiv (\lambda_{i,t-1}, \lambda_{it})$.[28] Specifically, once we have λ_{it} for every individual and year in the

[27] We use the Johns Hopkins ACG (Adjusted Clinical Groups) Case-Mix System. It is one of the most widely used and respected risk adjustment and predictive modeling packages in the health care sector, specifically designed to use diagnostic claims data to predict future medical expenditures.

[28] We are limited by our data to modeling the dependency of transitions based only on the last two years. Atal et al. (2019) also focus on a second-order Markov process for health state transitions. That said,

sample, we estimate year-to-year transition probabilities $\overline{f}(\lambda_{i,t+1}|\lambda_{i,t-1},\lambda_{it})$ for individuals in five-year age groups (e.g., transitions within cohort 25-30) using the actual transitions of consumers within each age range. (Again, the five-year grouping helps ensure adequate populations in each cell.) The advantage of computing transitions of ACG scores as opposed to medical expense transitions is that the ACG is based on persistent diagnostics. A broken arm probably does not affect significantly future medical expense realizations while asthma does. In other words, the ACG eliminates temporary expenses from the forecast of future expenses.[29] From the estimated probabilities $f(\lambda_{i,t+1}|\lambda_{i,t-1},\lambda_{it})$ we construct the 49-by-49 health state transition matrices [giving the mapping from $(\lambda_{i,t-1},\lambda_{i,t})$ to $(\lambda_{i,t},\lambda_{i,t+1})$] for the five-year age bins from ages 25-65 as the foundation for modeling health state persistence and transitions over time.

The top panel of Table 3 presents an illustrative subset of the estimates of $f(\lambda_{i,t+1}|\lambda_{i,t-1},\lambda_{it})$ for individuals who are between 30 and 35 years old (rounded to the nearest 0.01). Note that for each age group, the full transition matrix has size 49×7. In these tables, however, instead of presenting all of the 49 rows, we present only four, corresponding to $(\lambda_{t-1},\lambda_t)\in\{(1,1),(7,1),(1,7),(7,7)\}$. This presentation is substantially simpler than presenting two 49×7 matrices while, at the same time, illustrates the two main takeaways we would like to emphasize from these transitions: *ceteris paribus*, the probability of transitioning to sicker future health states is higher for (i) sicker current (time t) health states, (ii) sicker past (time $t-1$) health states, and (iii) current illness matters more than past illness. In Appendix G, we show that the same holds for 40 year old individuals.

The persistence embodied in these health state transitions is illustrated in the bottom panel of Table 3. This panel reports the net present value of expected medical expenses for different future periods conditional on the consumer's age-30 health state, focusing on health states in which the consumer's health status was the same at ages 29 and 30. The table shows that while there is significant persistence, much (but not all) of it dissipates after 10 years.[30],

allowing for the last two years (as opposed to only the last year) to influence transitions is indeed crucial. See appendix J for more details on this.

[29]By defining transitions over ACG index states we still may miss potential information on what condition led to the current ACG index that could entail different persistence beyond two years. However, we believe that with the combination of using ACG scores rather than medical expenses, and two-year health states, we capture the health transition process reasonably well. Appendix J includes additional analysis that highlights the value of using a second-order Markov process, with two past years of the ACG index, for health status transitions in our setup, relative to a first-order process and one past year of the ACG index.

[30]The fact that expected costs depend relatively little on the health state 10 years prior is consistent with actuarial mortality tables. There are two kinds of tables: "ultimate" tables are based on attained age only, while "select and ultimate" tables report the death rate not only by attained age, but by years since underwriting (namely, conditional on being in good health at that time). The tables converge as the years since underwriting increase; 10 years after underwriting the rates are quite similar.

λ_{t-1}	λ_t	λ_{t+1} 1	2	3	4	5	6	7
1	1	0.73	0.13	0.07	0.02	0.02	0.01	0.01
7	1	0.63	0.25	0.12	0	0	0	0
1	7	0.30	0.18	0.19	0.06	0.06	0.09	0.12
7	7	0.02	0	0.02	0.02	0.05	0.12	0.77

Health status (age 29 and 30)	Ages 30	31-35	36-40	41-64
1	0.84	1.83	2.39	4.2
4	3.05	3.98	3.41	4.37
7	20.51	13.05	6.31	4.78

Table 3: The top panel gives an example of empirical health status transitions from one year to the next, for 30-35 year old men. The bottom panel reports, for various age ranges, the constant annual medical expenses (in thousands of dollars) such that the present discounted value of these constant annual expenses equals the expected present discounted value of expenses over the age range in question for a Utah man in various age-30 health states.

3.4 Risk Preferences

The third ingredient we need is a consumer's risk aversion, i.e., the degree to which consumption smoothing over different states of the world is valued by consumers. In our main analysis, we use the risk preferences estimated in Handel, Hendel, and Whinston (2015). There we estimate a panel discrete choice model where risk aversion is identified by the choices that households make conditional on their household-specific health expenditure risk for the upcoming year. Consumers have constant absolute risk aversion (CARA) preferences:

$$u(c) = -\frac{1}{\gamma}e^{-\gamma c} \tag{8}$$

where $c = y - p - o$ is consumption (which equals income y less premium payments p and out-of-pocket medical expenses o) and γ is the risk aversion parameter. The mean estimated risk-aversion level is 0.0004, which falls within the range reported in the literature. We also consider the robustness of our conclusions with respect to the degree of risk aversion in Section 7.3.

3.5 Income profiles

The shape of the optimal contract depends on a consumer's income profile. Insurers offer different contracts to consumers with different income profiles to maximize their lifetime

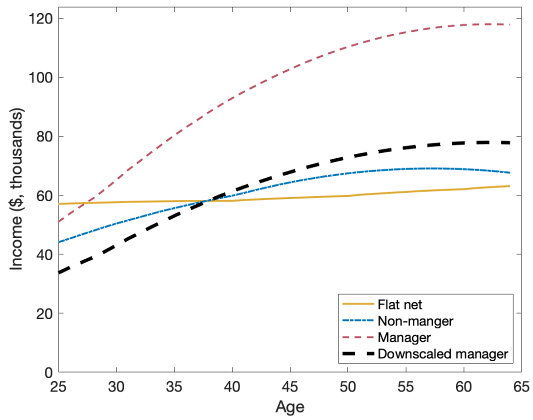

Figure 2: Income Profiles

expected utilities conditional on breaking even and the lapsation constraint. We compute optimal contracts and welfare for several different income profiles that vary in how steeply income rises with age, portrayed in Figure 2. The least steep is a *flat net income profile*, in which the change in income each year equals the change in the population's average medical expenses. With this income profile, there are no intertemporal consumption smoothing motives (for an individual who would pay a premium in each period equal to the population average medical costs), as individuals with flat net income do not want to use the contract as a mechanism to borrow or save, unlike consumers with increasing or decreasing income profiles over time. We also examine several more steeply rising income profiles, based on the income profiles we observed in Handel, Hendel, and Whinston (2015) for managers and non-managers in the firm studied there. Income profiles of managers at the firm were steepest, while those of non-managers were flatter but still steeper than a flat net income profile. We also investigate a downscaled manager income profile that makes the present value of lifetime manager income equal to that of a non-manager (which facilitates comparisons). The flat net income profile has the same net present value as the non-manager and downscaled-manager profiles. Recall that the income path in our model can be interpreted as *net* of any borrowing the consumer can do. Thus, we use these various income paths to illustrate how the expected growth of available resources over time impacts the optimal contract. Of course, in any specific context, the actual level and slope of income paths could be different than those we investigate. However, the income paths we study have values that are similar in magnitude to those in typical U.S. non-Medicaid contexts and should be sufficient to provide meaningful insights into the comparative statics we investigate as a function of income path steepness.

4 Results: Optimal Contracts

Using the data and computational approach described above, in the remainder of the paper we find the consumptions and premiums for empirically-based optimal dynamic contracts with one-sided commitment, and then compare their outcomes to those in various benchmarks. Although our analysis is necessarily stylized given the institutional intricacies of insurance provision in the U.S. (e.g., we don't consider movement between an employer insurance sector and the long-term contract, and don't consider possibilities for medical expense reductions through uncompensated charitable care or bankruptcy), relative to prior work it provides empirical insights into the structure and potential benefits of long-term insurance contracts.

In this section, we first study the structure of these contracts, examining the extent to which they are front-loaded, the degree of reclassification risk they insure, and how these contract characteristics depend on a consumer's income profile. We then turn to welfare analysis.

4.1 Front-loading and Reclassification Risk

We begin by considering the optimal contract for a consumer with flat net income (corresponding to the solid curve in Figure 2, an income profile that creates no borrowing or savings motive when a consumer faces the at-birth ex ante expected medical expenses at each age).

In our context, a contract specifies a premium, or equivalently, a consumption level for each possible history of states at each age from 25 to 65. There are too many histories and concurrent premiums/consumption levels to present: instead we focus on select attributes of the contract. First, we look in detail at the early contract periods, which provide intuition for the form of the equilibrium contract over longer time horizons. The first-period premiums, consumption levels, actuarial costs, and front-loading are presented in the top panel of Table 4 for consumers whose health status at ages 24 and 25 were the same (i.e., with $\lambda_{24} = \lambda_{25}$). In all but the worst state ($\Lambda_{25} = (7, 7)$), premiums are larger than actuarial costs: consumers front-load premiums to transfer utility to future states with negative health shocks. For example, for the healthiest consumer at the beginning of the year ($\Lambda_{25} = (1, 1)$), the premium is \$2,294 despite average costs of only \$837.

The extent of front-loading rises as the consumer's health state worsens to bin 5 out of 7, and then declines to zero for the sickest bin of consumers (7 out of 7). The extent of front-loading depends on both the current state and also on the implications of the current state for future health. While the healthiest type can afford the most front-loading, he might benefit the least. This is why maximum front-loading occurs for consumers in the

First-Year Equilibrium Contract Terms							
	$\lambda_{24} = \lambda_{25}$						
	1	2	3	4	5	6	7
	Flat Net Income						
Premium	2.294	3.472	4.156	6.255	11.444	13.733	20.511
First-Year Costs	0.837	1.375	1.973	3.054	4.358	6.842	20.511
Front-Loading	1.457	2.098	2.182	3.201	7.086	6.891	0
Consumption	54.765	53.586	52.903	50.803	45.615	43.326	36.548
	Downscaled Managers						
Premium	1.165	2.196	2.740	3.487	8.089	10.887	20.511
First-Year Costs	0.837	1.375	1.973	3.054	4.358	6.842	20.511
Front-Loading	0.328	0.821	0.767	0.433	3.732	4.046	0
Consumption	32.520	31.490	30.945	30.199	25.596	22.799	13.175

Table 4: First-year contract terms in the equilibrium long-run contract for men with a flat net income path (top panel) and men with the downscaled manager income path (bottom panel). The table shows first-year premiums, expected costs, the extent of front-loading, and consumption levels (thousands of dollars).

middle of the ex ante health range, rather than for the healthiest consumers. The least healthy, on the other hand, have very high current costs (and, hence, high marginal utility of consumption) compared to their expected future costs (recall Table 3); for them, front-loading isn't worthwhile.

Table 5 presents second-period premiums for the seven possible age-26 health statuses that can follow the seven age-25 health states considered in Table 7. (These age-26 health status realizations λ_{26} give rise to the age-26 health state $\Lambda_{26} = (\lambda_{25}, \lambda_{26})$.) Certain patterns are indicative of the longer-run structure of the contract. First, second-period premiums display extensive pooling which takes place in states for which the lapsation constraint is not binding. For example, if a consumer was in the healthiest possible state at age 25, $\Lambda_{25} = (1, 1)$, all second-year states $\Lambda_{26} = (1, \lambda_{26})$ with $\lambda_{26} > 1$ have premiums of \$2,429. The lapsation constraint does not bind for this consumer when $\lambda_{26} > 1$ because the first period front-loaded amount suffices to make outside offers less attractive than the current premium guarantee. Only when $\lambda_{26} = 1$ does the lapsation constraint bind for this consumer, resulting in a slight premium decrease to \$2,403. Note that these results for premiums map directly to results for second-year consumption levels and guarantees, which are depicted in Appendix G in Table 14.

Second-Year Equilibrium Premiums: Flat Net Income								First-Year
				λ_{26}				Premium
$\lambda_{24} = \lambda_{25}$	1	2	3	4	5	6	7	
1	2.403	2.429	2.429	2.429	2.429	2.429	2.429	2.294
4	2.470	3.063	4.454	6.004	6.390	6.390	6.390	6.255
7	2.385	2.470	2.650	4.975	15.809	14.798	20.511	20.511

Table 5: First- and second-year premiums in the equilibrium long-run contract for men with a flat net income path, as a function of the period 1 health state and period 2 health status (thousands of dollars). Table 14 in the appendix shows the analogous table for contingent consumption.

The lapsation constraint binds for more and more second-year states the sicker the consumer was at the start of the contract. For consumers initially in the sickest health state, $\Lambda_{25} = (7, 7)$, all age-26 health states involve different premiums and, hence, consumption levels: long-run contracts cannot provide any insurance against reclassification risk in year 2 for this consumer as his first-year needs were so great as to preclude any front loading. For this consumer, the long-run contract continuation at age 26 simply matches the best contract he could get on the market given his age-26 health state.

4.2 Effects of Income profiles

The equilibrium contracts offered depend crucially on a consumer's rate of income growth over his lifetime. When income is relatively low early in life, and hence the marginal utility of consumption is relatively high, front-loading is quite costly for utility.

The bottom panel of Table 4 presents first-period (age-25) contract characteristics for "downscaled managers." Recall that, as shown in Figure 2, a downscaled manager income profile proportionally scales down the income of a manager to match the net present value of a non-manager's lifetime income profile. The table makes clear that for downscaled managers, the extent of front-loading is much more limited than in the flat net income case, which translates into less generous consumption guarantees later in life. For example, a downscaled manager who is in the healthiest state ($\Lambda_{25} = (1, 1)$) at age 25 front-loads only \$328, compared to \$1,457 for a consumer with flat net income. Essentially, the rapidly rising income makes paying extra early in life for long-term insurance quite costly, as marginal utility is high early compared to what is expected later in life. Appendix G provides further analysis of this income path by examining second-year premiums and consumption levels.

5 Results: Welfare

We now turn to the welfare analysis of these dynamic contracts. We measure and compare the welfare they achieve to several alternatives. For each market setup and potential income profile considered, we compute a lifetime certainty equivalent. The certainly equivalent represents the constant consumption for the forty years of life from age 25 to 65 that makes the consumer as well off as in a given market setup. Specifically, we compare the certainty equivalent of optimal dynamic contracts with one-sided commitment, denoted by CE_D, to the three benchmarks we have described previously (see Section 2.1 for formal definitions):

(i) The first-best, fully-smoothed consumption C^*, which equates the marginal utility of consumption across periods and states. This is the welfare achievable were long-term contracts with two-sided commitment feasible;

(ii) The certainty equivalent from spot contracts that fully insure event risk in every period and state, but leave reclassification risk across periods fully uninsured, denoted by CE_{SPOT};

(iii) The constant consumption equivalent of the No Borrowing/No Saving constrained first best, in which risk is fully insured in each period but neither borrowing nor saving is possible, denoted C^*_{NBNS}.

5.1 Welfare Effects Conditional on a Consumer's Age-25 Health State

The top panel in Table 6 shows welfare outcomes for Utah men with a flat net income profile who arrive at age 25 in each of the seven health states Λ_{25} in which their age-24 and age-25 health status is the same (i.e., in which $\lambda_{24} = \lambda_{25}$). For each age-25 health state, column (1) reports the annual consumption level C^* in a first-best contract that starts at age 25 given the income profile and expected future medical expenses the consumer faces given his age-25 health state.[31] It ranges from \$54,960 for the healthiest consumer state $\Lambda_{25} = (1, 1)$ to \$49,330 for a consumer in the worst state $\Lambda_{25} = (7, 7)$. Column (2) shows C^*_{NBNS}, the constant consumption equivalent of the constrained first-best outcome that does not allow for intertemporal consumption smoothing.

For consumers with rising net income, C^*_{NBNS} may be a more relevant benchmark of the losses from spot contracting and of how well optimal dynamic contracts with one-sided

[31] In this and the other three tables in this subsection, "ex ante" certainty equivalents are calculated from the perspective of a consumer who arrives at age 25 in a particular health state. Thus, for example, the first-best consumption of a consumer with flat net income will differ across consumers with different health states Λ_{25} because of their differing expected medical costs.

Long Run Welfare Impacts of Dynamic Insurance Contracts							
Init health	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\lambda_{24} = \lambda_{25}$	C^*	C^*_{NBNS}	CE_{SPOT}	CE_{SS}	CE_D	$\frac{C^*_{NBNS}-CE_{SPOT}}{C^*_{NBNS}}$	$\frac{CE_D-CE_{SPOT}}{C^*_{NBNS}-CE_{SPOT}}$
			Flat Net Income				
1	55.14	55.10	45.31	49.89	55.03	0.178	0.994
2	54.96	54.94	45.03	49.39	54.80	0.180	0.986
3	54.84	54.83	44.92	49.43	54.62	0.181	0.979
4	54.36	54.34	44.45	48.22	53.69	0.182	0.934
5	52.86	52.14	42.47	44.70	49.62	0.185	0.739
6	51.51	49.42	41.53	43.49	47.08	0.160	0.703
7	49.33	42.61	39.94	40.26	40.72	0.063	0.291
			Downscaled Managers				
1	55.14	39.07	32.73	34.66	38.76	0.162	0.951
2	54.96	38.47	30.75	33.49	37.87	0.201	0.922
3	54.84	37.94	30.19	33.18	37.26	0.204	0.912
4	54.36	37.07	32.40	34.12	36.13	0.126	0.799
5	52.86	34.23	25.87	27.72	31.31	0.244	0.651
6	51.51	30.82	24.26	26.19	28.47	0.213	0.642
7	49.33	20.56	19.89	19.93	19.94	0.032	0.076

Table 6: Long-run welfare results showing the certainty equivalent annual consumption of different insurance institutions under various initial health states, a discount factor of 0.975, and constant absolute risk aversion equal to 0.0004. The top panel reports results for the case of flat net income while the bottom panel reports results for the downscaled manager income path. Units in columns (1)-(4) are 1000s of dollars.

commitment do at eliminating reclassification risk, since saving and borrowing on their own can greatly improve utility for steep net income profiles. (For a healthy consumer with flat net income, however, this certainty equivalent is very close to C^*.) Column (3) shows welfare under spot contracts for each of these consumers, while Column (6) shows the welfare loss from reclassification risk under spot contracting relative to this benchmark, $\frac{C^*_{NBNS}-CE_{SPOT}}{C^*_{NBNS}}$, a measure that captures solely the loss under spot contracting arising from reclassification risk. This welfare loss is very large: many of these consumers lose roughly 18% of their *lifetime* (age 25-65) certainty equivalent because of reclassification risk.

Column (5) presents the certainty equivalent for dynamic contracts with one-sided commitment, CE_D. As expected CE_D lies between C^*_{NBNS} and CE_{SPOT}. Column (7) shows the fraction of the welfare gap between the No-borrowing/No-saving constrained first-best and spot contracts that these dynamic contracts close, $\frac{CE_D - CE_{SPOT}}{C^*_{NBNS} - CE_{SPOT}}$. Overall, dynamic contracts are very effective at reducing reclassification risk for consumers who arrive at age 25 in excellent health: for a consumer in age-25 health state $\Lambda_{25} = (1, 1)$, dynamic contracts close 99.4% of the gap between C^*_{NBNS} and C_{SPOT}. But they are very ineffective for consumers who arrive at age 25 in poor health. At the extreme, a consumer who arrives at age 25 in the worst health state, $\Lambda_{25} = (1, 1)$, dynamic contracts recover only 29.1% of the welfare loss due to reclassification risk under spot contracting. The reason for this pattern is that consumers who arrive at age 25 in poor health have a high level of current medical expenses, which makes front-loading very costly and therefore greatly limits the effectiveness of dynamic contracts. The bottom panel in Table 6 reports the same welfare statistics but for consumers with the downscaled manager income profile (for results on manager and non-manager income profiles, see the appendix). Welfare losses from reclassification risk (shown, again, in column (6)) are similarly large for consumers with rising income profiles. Comparing column (7) for the cases with rising income profiles to the flat net income profile case shows that rising income profiles reduce the effectiveness of dynamic contracts by making front-loading more costly, since with a rising income profile the marginal utility of a current dollar is larger than the marginal utility of future dollars. The effect is particularly dramatic for consumers in poor health states: for example, for a consumer in state $\Lambda_{25} = (7, 7)$, dynamic contracts reduce the loss from reclassification risk by 29.1% if the consumer has a flat net income profile, but by only 7.6% if he has a downscaled manager income profile.

5.2 Welfare Effects from the Perspective of an Unborn Consumer

We now turn to evaluating the welfare effects of dynamic contracts from the perspective of an unborn consumer who does not know what his age-25 health state will be. We compute these welfare measures using the distribution of age-25 health states that we see empirically among our age-25 Utah men. Table 7 shows the results.

Recall that, to ease comparisons across different income profiles, all profiles we consider have the same net present value of income, except for the manager's profile when it is not downscaled. Column (1) of Table 7 shows the first-best consumption (reflecting the present value of income minus expected medical expenses, now calculated from the perspective of a consumer who does not yet know their age-25 health state), which are \$54,670 for the flat-net, non-manager, and downscaled manager income profiles, and \$85,000 for the manager income profile. Column (2) gives the certainty equivalent of the No-Borrowing/No-Saving full insurance regime, again based on expected medical expenses calculated from the perspective of a consumer who does not yet know their age-25 health state. Column (3) shows welfare outcomes under spot contracts, with no protection for reclassification risk, while column

(4) shows the certainty equivalent of the optimal dynamic contract regime, both from the perspective of an unborn consumer.[32]

The welfare loss from reclassification risk, shown in column (6), ranges between 18.4% and 27.4% of lifetime certainty equivalent (relative to the no-borrowing/no-savings benchmark).[33] For a flat net income profile, column (7) shows that 43.4% of this gap is recovered by dynamic contracts. In contrast, since consumers with steeper income profiles dislike front-loading, dynamic contracts recover a very small portion of this welfare loss for these consumers (7.5% for downscaled managers).

In summary, from the perspective of an unborn Utah male with a flat net income profile, optimal dynamic contracts would be moderately effective at reducing the reclassification risk they face. The moderate effectiveness is due to the poor performance for consumers who arrive at age 25 in poor health, and thus with high immediate consumption needs. Once coupled with more steeply rising income paths, which further accentuate the value of net income when young, dynamic contracts prove rather ineffective from an ex ante welfare perspective.

5.3 Insurance of Pre-age-25 Health Risk

The sensitivity of effectiveness of long term contracts to the initial health state, suggests the government can play a role insuring pre-age-25 health realizations. We consider this in two ways. First, we ask what the expected per capita cost would be for the government to ensure that each consumer's continuation certainty equivalent starting at age 25 is the same as if he had reached age 25 in the healthiest state, $\Lambda_{25} = (1, 1)$. Second, we derive the set of break-even subsidies that most efficiently insure the age-25 health risk these consumers face. This involves finding the subsidy or tax for consumers in each of the 49 age-25 health states such that (i) the government breaks even in expectation, and (ii) the welfare of an unborn consumer is maximized.[34]

Column (10) reports the (one-time) expected per capita cost of subsidies that insure that, under a regime of optimal dynamic contracts that begin at age 25, every consumer has the same certainty equivalent as if he had arrived at age 25 in health state $\Lambda_{25} =$

[32] C^*_{NBNS} in Table 7 equals the expectation of the C^*_{NBNS} values for consumers in the various age-25 health states, evaluated using the empirical distribution of age-25 health states. In contrast, because of risk aversion, CE_D in Table 7 is less than the expectation of the CE_D values for consumers in the various age-25 health states.

[33] In Appendix F we extend this analysis to allow precautionary savings in the spot contracting regime and find similar results. We find that such precautionary savings only closes a small amount of the gap (2.6% to 27.1%) from spot contracts to the no borrowing / no saving benchmark. This is not surprising as the main loss with a rising income path from a lack of inter-temporal smoothing comes from the inability to borrow.

[34] Note that this does not generally yield equal certainty equivalents for the 49 health states because the marginal utility of a dollar subsidy is state-dependent.

Unborn Consumer Welfare Results					
		Certainty Equivalent Welfare			
	(1)	(2)	(3)	(4)	(5)
Income profile	C^*	C^*_{NBNS}	CE_{SPOT}	CE_D	CE_{D+}
Flat net	54.67	54.67	44.35	48.83	53.91
Non-mngr	54.67	47.37	36.96	38.08	45.24
Manager	85.00	55.67	45.44	45.91	54.05
Downs Mngr	54.67	37.68	27.35	28.13	35.79
		Welfare Ratios			
	(6)	(7)	(8)	(9)	(10)
Income Profile	$\frac{C^*_{NBNS}-CE_{SPOT}}{C^*_{NBNS}}$	$\frac{CE_D-CE_{SPOT}}{C^*_{NBNS}-CE_{SPOT}}$	$\frac{CE_D-CE_{SPOT}}{C^*-CE_{SPOT}}$	$\frac{CE_{D+}-CE_{SPOT}}{C^*_{NBNS}-CE_{SPOT}}$	Req. Subsidy
Flat net	0.189	0.434	0.434	0.926	12.05
Non-mngr	0.220	0.108	0.063	0.795	8.81
Manager	0.184	0.046	0.012	0.842	5.47
Downs Mngr	0.274	0.075	0.028	0.817	7.07

Table 7: The top panel shows unborn consumer welfare results showing the certainty equivalent annual consumption of different insurance institutions under various income profiles, a discount factor of 0.975, and constant absolute risk aversion equal to 0.0004. Column (5) shows the certainty equivalent welfare level resulting from a balanced budget scheme that optimally insures the consumer's pre-age-25 health risk prior to the start of dynamic contracting at age 25 while Column (4) shows this same statistic for the unborn consumer, but without age-25 health risk insurance. The bottom panel columns (6)-(9) show key welfare ratios comparing different policy regimes. Column (10) shows the expected one-time subsidy required at age 25 for the consumer to have in all age-25 states the same level of welfare as if he had been in the healthiest possible age-25 health state. Units in columns (1)-(5) and (10) are 1000s of dollars.

$(1, 1)$. This cost ranges from \$5,470 for a manager income profile to \$12,050 for a flat net income profile. Column (5) reports the certainty equivalent achieved instead when optimal dynamic contracts starting at age 25 are coupled with a break-even government insurance scheme that insures against consumers' pre-age-25 health risks, which we label as CE_{D+}, while column (9) shows what fraction of the reclassification risk losses are recovered. When combined with this pre-age-25 insurance, dynamic contracts eliminate roughly 80-90% of reclassification risk. One can think of these policies to insure age-25 risk as similar in spirit to the risk-adjustment, risk-corridor, and reinsurance regulations present in the ACA (and many other current environments) but applied to age-25 consumers choosing dynamic contracts instead of consumers of all ages choosing year-to-year contracts. The results show that long-term dynamic contracts can be rather effective at eliminating reclassification risk in combination with government insurance of consumers' health risks that occur before they reach the insurance market at age 25.

6 Comparison to Managed Competition Exchange

One of the most significant features of the health insurance exchanges created by the ACA was their ban on the pricing of pre-existing conditions. In this section, we examine how dynamic contracts do at eliminating reclassification risk compared to a managed competition-style insurance exchange. As discussed in Handel, Hendel, and Whinston (2015), while the ACA fully eliminated reclassification risk, it created adverse selection, as consumers with differing health could not be differentially priced. This adverse selection led to significant unraveling in the exchange model studied in Handel, Hendel, and Whinston (2015), so that in most cases all consumers ended up obtaining insurance contracts covering only 60% of their expenses.

In this section we compare the welfare achievable with optimal dynamic contracts, both with and without government insurance of pre-age-25 health risk, to the level of welfare that would arise in a managed competition-style insurance exchange.[35] Because consumers end up only partly insured in such an exchange, computing welfare requires as an input the full distribution of health expenses conditional on an individual's health state, rather than just its mean. We have previously estimated this full distribution for the consumers in the Handel, Hendel, and Whinston (2015) sample and we make use of this information here. Specifically, we examine how dynamic contracts would perform for the Handel, Hendel, and

[35]The ACA insured pre-age-25 risk through its ban on pre-existing conditions, so a natural comparison is to dynamic contracts combined with insurance of pre-age-25 risk.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income profile	C^*	C^*_{NBNS}	CE_D	CE_{D+}	CE_{ACA}	$\frac{CE_{D+}-CE_{SPOT}}{C^*_{NBNS}-CE_{SPOT}}$	$\frac{CE_{ACA}-CE_{SPOT}}{C^*_{NBNS}-CE_{SPOT}}$
Flat net	54.62	54.62	52.70	54.12	52.90	0.890	0.622
Non-mngr	54.62	47.25	43.19	45.90	46.44	0.710	0.826
Manager	84.95	55.52	51.25	54.52	54.77	0.777	0.833
Downs Mngr	54.62	37.55	33.37	36.36	36.78	0.740	0.832

Table 8: Long-run welfare results showing the certainty equivalent annual consumption of different insurance institutions, including a managed competition-style insurance exchange (labeled “ACA”), for the large employer sample of Handel, Hendel, and Whinston (2015). Assumes a discount factor of 0.975 for consumers with median estimated constant absolute risk aversion equal to 0.0004. Units in columns (1)-(5) are 1000s of dollars.

Whinston (2015) consumers if they had the transitions that we have estimated for Utah men, and compare it to what the ACA would achieve.[36]

Specifically, we compute welfare under a managed competition-style insurance exchange by imposing (i) one-year contracts, (ii) community rating (no health-state based pricing allowed), (iii) age-based pricing, (iv) a fully enforced mandate, requiring insurance purchase, and (v) insurers that offer plans covering specific actuarial values, with a minimum plan covering 60% of an average individual's spending that the market unravels to. For simplicity, we will refer to this outcome in the rest of this section as the “ACA outcome” and denote its certainty equivalent by CE_{ACA}.

Columns (1)-(4) of Table 8 give the (unborn) certainty equivalents for institutions that were in Table 7, while column (5) gives the corresponding certainty equivalent under the ACA. Comparing columns (3) and (5), dynamic contracts without government insurance of pre-age-25 health risk are worse than the ACA outcome for all income profiles. A primary reason for this difference is that community rating implicitly insures consumers' pre-age-25 health risk.

Column (4), on the other hand, shows the welfare level achievable for this sample when dynamic contracts are coupled with insurance of pre-age-25 health risk. Columns (6) and (7) show how much of the gap between spot contracting and the No-Borrowing/No-Savings

[36] The sample in Handel, Hendel, and Whinston (2015) is too small to estimate second-order Markov transitions, while for the Utah sample we have only mean health costs conditional on health states, not the full distributions. For these reasons we combine the data in the two samples for this analysis. We have examined the effect of dynamic contracts in both the Utah and Handel, Hendel, and Whinston (2015) samples with first-order Markov processes and found similar results. By way of comparison to the male Utah sample, in the Handel, Hendel, and Whinston (2015) sample, the expected costs conditional on each of the seven health statuses (going from healthiest to sickest) are \$1,131, \$2,290, \$3,780, \$3,975, \$5,850, \$10,655, and \$18,554. Comparing to Table 2, Utah expenses were lower in bins 1-6, but higher for consumers with the worst health status.

benchmark is closed by dynamic contracts with insurance of pre-age-25 health risk and the ACA outcome, respectively.[37] Comparing columns (6) and (7) in Table 8 reveals that the managed competition-style insurance exchange environment is preferred to the dynamic contracting environment even with insurance of pre-age-25 health risk for non-managers, managers, and downscaled managers, whose incomes rise over time. For these individuals, the desire to front-load when young and income is relatively low is limited, which reduces the benefits from dynamic contracts. In contrast, individuals with flat net income profiles prefer dynamic contracts with insurance of pre-age-25 health risks to the managed competition-style insurance exchange environment.

7 Extensions

First, we consider how the presence of switching costs would affect the gains from dynamic contracts. Second, we examine how consumer myopia impacts our results. Third, motivated by the concern that risk aversion could differ from typical estimates in the health insurance literature for the large losses created by reclassification risk, we examine the welfare effects of dynamic contracts for lower levels of risk aversion. In the appendix, we also study a number of other issues: (i) how the ability to engage in precautionary savings would affect our conclusions, (ii) the impact of lapsation into uninsurance at an exogenous rate, (iii) limited access to credit markets, and (iv) income uncertainty.

7.1 Inertia

Recent evidence from health insurance markets [Handel (2013), Ho et al. (2016)] points to substantial inertia in insurance choice. As discussed in Section 2.6, switching costs relax the lapsation constraints, which can enhance commitment and the welfare achievable with optimal dynamic contracts.

We leave the analysis of equilibrium consumption levels under switching costs to the appendix and focus, instead, on welfare. Table 9 shows the ex ante (at-birth) welfare achieved by dynamic contracts for different levels of switching costs and our four income profiles in the male Utah sample. As expected, welfare is monotonic in the switching cost. Qualitatively, as switching costs increase from zero to infinity, welfare in the optimal dynamic contract with one-sided commitment approaches the first-best (two-sided commitment) level.

[37] The No-Borrowing/No-Savings outcome is the same as the ACA outcome except that it provides 100% coverage rather than unraveling to 60%; the difference between C^*_{NBNS} and CE_{ACA} therefore reflects the cost of adverse selection under the managed competition-style insurance exchange. Table 8 shows that the cost of adverse selection ranges from roughly \$750 per year for managers, downscaled managers, and non-managers to roughly \$1700 per year for consumer with flat net income profiles.

Switching cost	Flat net	Non-manager	Manager	Downs mngr
0	48.83	38.08	45.91	28.13
10	50.37	41.13	50.46	31.79
50	53.37	47.06	58.95	38.92
100	54.15	51.47	65.58	44.57
500	54.19	54.19	84.51	54.19
C^*	54.67	54.67	85.00	54.67
C^*_{NBNS}	54.67	47.37	55.67	37.68

Table 9: At-birth welfare (in \$1,000s) for Utah men from optimal dynamic contracts (CE_D) under discount factor of 0.975 and risk aversion of 0.0004 for different levels of switching costs and four income profiles.

Notice that it takes extremely large switching costs to achieve welfare close to first best for consumers with steeply rising manager and downscaled-manager income profiles. The reason is that consumption smoothing requires a lot of commitment, especially when the income profile is steep. Thus, an extremely large switching cost is necessary to achieve the first best. Somewhat more moderate switching costs deliver welfare close to the no-borrowing/no-saving benchmark.

Using the same hybrid sample as in Section 6, one can compute, for each income profile, the switching cost that is needed to achieve the same level of welfare as the managed competition-style insurance exchange we consider there. For a flat net income profile, switching costs of \$1,970 suffice. The corresponding numbers for non-manager, manager, and downscaled manager income profiles are, repectively, estimated at \$13,590, \$7,680, and \$10,470. Handel (2013) estimates a mean switching cost of \$2,032 in a static model of choice for the same population as in our Handel, Hendel, and Whinston (2015) sample. Since the dynamic gain from switching likely extends over multiple periods, the comparable value for our model is likely significantly higher than this \$2,032, and so switching costs may well be in the range that make dynamic contracts (without insurance of pre-age-25 health risk and with our baseline risk aversion) preferable to the managed competition-style insurance exchange for some income profiles.[38,39]

[38] Illanes (2017), for example, estimates a lower bound switching cost of \$1200 in a dynamic model of choice for the Chilean pension market; he shows that the estimate from a static model of choice in his sample is \$117.

[39] Note, however, that our ACA model assumes that there are no switching costs; modeling insurance exchange competition with switching costs remains an open issue.

7.2 Myopia

Section 2.6 sets up our framework for studying the implications of consumer myopia for optimal dynamic insurance contracts with one-sided commitment. There are several key takeaways from Proposition 3. First, the qualitative feature of optimal dynamic contracts with one-sided commitment is preserved under myopia: the contracts involve front-loading in exchange for future consumption guarantees. Second, unlike the baseline (i.e., non-myopic) case, under myopia the optimal contract will involve consumption guarantees that diminish over time. Third, given that optimal contracts under myopia turn out to be similar in structure to the no-myopia baseline case, our computation method can be used with slight modifications to numerically analyze equilibria under myopia.

In this section we analyze the performance of optimal contracts under myopia using simulations based on the Utah all-payer claims dataset. To capture the idea that myopia is behavioral, we assume, as is typical in the literature, that the market discount factor δ (instead of consumers' β) is relevant for assessing their welfare. This allows us to capture the fact that consumers act according to the "wrong" discount factor and, consequently, suffer a welfare loss from their myopia.[40]

Figure 3 illustrates the contract features for a young individual in good health and shows how those features change with β. The figure shows several key impacts of myopia. First, the first-year premium and the extent of contract front-loading decrease substantially as myopia increases. For example, a non-myopic individual who is healthy and young has a contract with a premium of \$2,288 in year one, of which \$1,451 is front-loading. An individual with some myopia ($\beta = 0.8$) has a much lower premium (\$1,528) and much lower front-loading (\$691) while an individual with significant myopia ($\beta = 0.2$) has a premium of \$1,004 and minimal front-loading (\$167). Thus, as myopia increases, the extent of front-loading decreases and the dynamic contracts we study insure less against reclassification risk.

Figure 4 shows the extent of annual lapsation from equilibrium optimal dynamic contracts as a function of myopia and lifetime income paths (recall from section 2.2 that there is a *guaranteed premium path* interpretation of our optimal contracts that does involve lapsation in the equilibrium). The left panel of the figure presents annual contract lapsation rates as a function of age and the extent of myopia, for downscaled manager income paths. Lapsation rates always increase, conditional on age, as a function of myopia. But, lapsation rates are fairly high for young individuals even with no myopia, so the extent of lapsation only increases a little for them. However, older individuals who have low lapsation rates with no myopia have much higher lapsation rates as myopia increases since contracts are no longer able to front-load enough to provide the consumption guarantees that keep consumers from

[40] Of course, in practice, the welfare-relevant discount factor for consumers could be different than the insurer's discount factor, which would be straightforward to incorporate here.

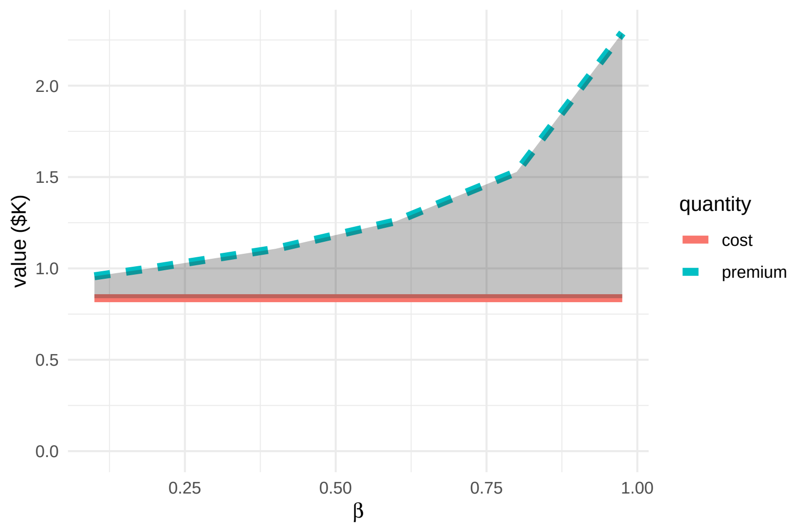

Figure 3: First-year contract terms in the equilibrium long-term contract for men with a flat net income path, showing first-year premiums (blue dashed line), expected costs (red solid line), and the extent of front-loading (shaded area between curves.) This figure is for health status $\lambda_{24} = \lambda_{25} = 1$ and covers range of myopia parameter β from 0.1 to 0.975.

lapsing. The right panel in the figure shows lapsation rates, averaged over age, for different income paths as a function of myopia. The annual lapsation rate is similar and high (≈ 0.9) for all income paths when myopia is high. As myopia decreases, lapsation decreases, for all income paths. But, for flatter income paths this is a big decrease in lapsation while for consumers with steeper income paths, who are less likely to front-load to begin with, the decrease in lapsation as myopia decreases is smaller.

Figure 5 presents our welfare results for cases when the insurer's discount factor δ is equal to 0.975 and consumer's discount factor β varies between 0.1 and 0.975.[41] The figure shows the at-birth welfare impact of dynamic contracts relative to the "no borrowing / no savings" benchmark. We plot the welfare results as a function of (i) myopia and (ii) the steepness of a consumer's income path.

Several results are evident in the figure. First, as expected, the welfare generated by dynamic contracts is decreasing as myopia increases. As consumers become more present-focused, front-loading to ensure continued contract participation is less viable and contracts are less able to insure reclassification risk.

Second, the negative derivative of welfare with respect to myopia is larger in magnitude with flatter income paths. Absent myopia dynamic contracts generate higher value for flatter income paths because front-loading is then more appealing. As myopia increases. consumers

[41] Note that as β moves towards 0, our optimal dynamic contract approaches spot contracts since consumers only consider the current period.

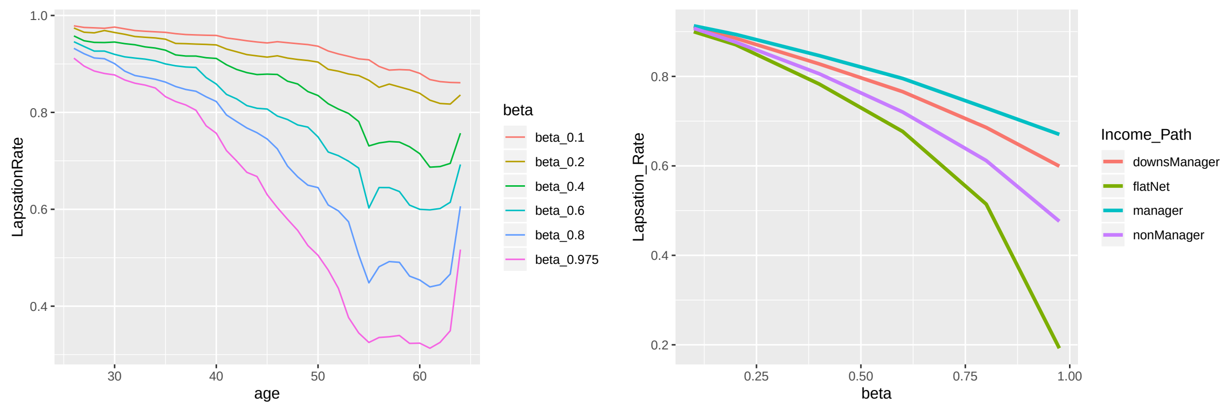

Figure 4: Annual lapsation rates for equilibrium dynamic contracts as a function of myopia (β) and lifetime income paths. The panel on the left shows lapsation rates for different ages for downscaled manager income paths while the panel on the right shows mean lapsation rates over all ages, for each income path.

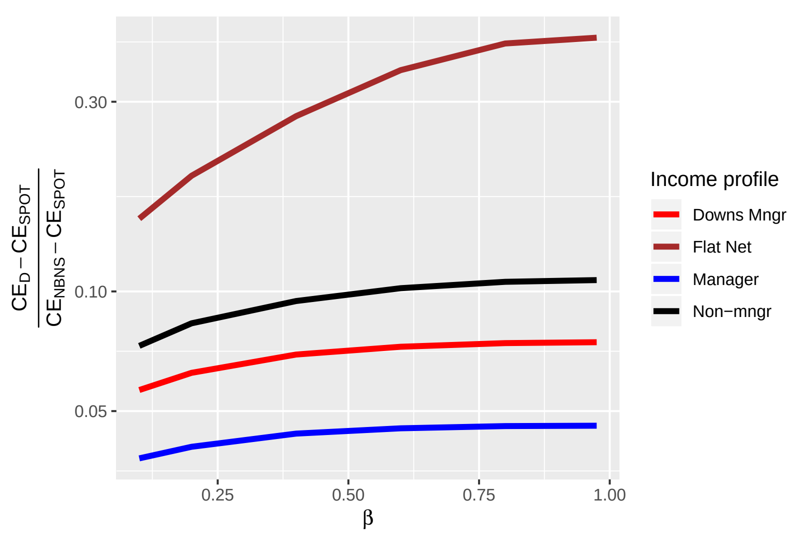

Figure 5: The performance of dynamic contracts, measured by $\frac{CE_D - CE_{SPOT}}{CE_{NBNS} - CE_{SPOT}}$, for different values of customer discount factor β when insurer discount factor δ is 0.975. The vertical axis is on log scale.

with flat net income care increasingly more about the costs of front-loading and increasing less about the consumption guarantees that front-loading allows for. Thus, even though consumers with flat net income paths are better able than consumers with steep income paths to insure against reclassification risk with dynamic contracts for any level of myopia relative to consumers with increasing income, a lot of the welfare gain they could achieve is eroded by myopia.

Third, even when myopia is strong (e.g. $\beta = 0.1$) dynamic contracts still provide non-trivial protection against reclassification risk. This is due primarily to the fact that consumers derive most of their utility from dynamic contracts by insuring against the worst future health risk realizations. Even with some myopia, this desire for risk protection against these catastrophic outcomes is still strong.

7.3 Risk Aversion

So far our analysis has used the risk preferences estimated in Handel, Hendel and Whinston (2015). We now consider the robustness of the analysis with respect to the degree of risk aversion. We are particularly interested in lower risk aversion. A primary reason is that our estimates of risk preferences come from choices among health insurance contracts with out-of-pocket caps (maximum downside exposure) that range between \$3,000 and \$10,000. Since the stakes we study go above this range for some of the policies we investigate, the estimates we use might overstate consumers' risk aversion when considering the larger gamble sizes associated with reclassification risk.[42]

Table 10 presents the welfare comparisons for risk aversion of 0.00008, five times lower than that in our main analysis (Table 7). For a consumer with \$50,000 of consumption, this corresponds to a CRRA risk aversion coefficient of 4, roughly the level suggested in the macro literature on consumption disasters (e.g., Barro (2006)). To put the coefficients in perspective, consider a lottery that assigns the costs associated with each of the seven health statuses, with each having equal probability. For the costs we used in Section 6, our 0.0004 risk aversion coefficient estimate implies a willingness to pay of \$7,222 to avoid the uncertainty associated with this risky prospect. Instead, the lower risk aversion coefficient leads to a willingness to pay of \$1,491.

Lowering risk aversion substantially reduces the loss associated with reclassification risk (captured by the gap between C^*_{NBNS} and CE_{SPOT}). The loss is now between 1.8% and 3.1%, depending on the income profile. For the higher 0.0004 risk aversion, the loss was between 18.4% and 27.4%. Still, a substantial.

[42]See Rabin (2000) for a discussion of issues with CARA and scaling of gambles, and Collier et al (2017) for evidence of differing risk aversion for small versus large stakes decisions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income profile	C^*	CE_{SPOT}	CE_D	C^*_{NBNS}	$\frac{C^*_{NBNS}-CE_{SPOT}}{C^*_{NBNS}}$	$\frac{CE_D-CE_{SPOT}}{C^*_{NBNS}-CE_{SPOT}}$	$\frac{CE_D-CE_{SPOT}}{C^*-CE_{SPOT}}$
Flat net	54.67	52.99	54.38	54.67	0.031	0.827	0.827
Non-mngr	54.67	51.22	52.05	52.72	0.028	0.550	0.239
Manager	85.00	68.45	68.78	69.73	0.018	0.256	0.020
Downs Mngr	54.67	46.44	46.95	47.81	0.029	0.372	0.062

Table 10: Long-run welfare results for Utah men showing the certainty equivalent consumption of different insurance institutions for a discount factor of 0.975 and constant absolute risk aversion equal to 0.00008.

While the loss from reclassification risk is lower, the proportional reductions in reclassification risk from long-term contracting, as captured in column (6) by how much of the gap between the welfare under optimal dynamic contracts and that under the No-Borrowing/No-Saving benchmark, are much larger, ranging between 25.6% for the manager income profile to 82.7% for the flat net income profile. As in our baseline analysis, here with rising income profiles, dynamic contracts without insurance of pre-age-25 health risk leave consumers facing a high share of the reclassification risk. In general, the magnitudes of the welfare effects of dynamic contracts are smaller, but this is only because there is less harm from reclassification risk to begin with; the welfare effects are still meaningfully large. Finally, the comparative statics we investigate with respect to income paths, government insurance of age-25 risk, switching costs, and myopia are directionally similar under both levels of risk aversion.

7.4 Limitations and Other Extensions

In the appendix, we examine a some other extensions/alternative specifications. First, we discuss the role of precautionary savings. Recall that our model of long-term contracts allows consumers to save for possible future negative health shocks, but we show that they do not save in equilibrium because the optimal long-term contract provides that function already. In the appendix we address whether long-term contracts deliver welfare benefits above and beyond effectively helping consumers to save. Our analysis in the appendix shows the answer is yes and quantifies the magnitude by comparing the long-term contracting outcome to the outcome of short-term contracts with dynamically optimal precautionary saving.

As another extension, we examine the impact of consumers' ability to borrow on equilibrium long-term health insurance contracts and on our welfare results, showing how one can interpret the consumer's income path as net of their ability to borrow.

We also discuss the implications for our analysis of introducing stochastic income . We do this both for the case in which uncertainty in income is contractible (similar to that of health) and for the case where there is income uncertainty that cannot be contracted upon. In the former case, most of our theoretical analysis remain unchanged, and long-term contracts now have a greater benefit as they insure income uncertainty as well.[43]

Finally, we theoretically extend our model to consider the impacts of consumers lapsing into uninsurance due to reasons such as irrational decision making or liquidity constraints.

In spite of the extensions we consider both within the paper and in the appendix, our analysis has a number of limitations. First, we do not model non-diversifiable risks faced by insurers and their possible impact on equilibrium contracts. Though there is reason to believe that non-extensive regulatory measures would be able to manage this issue,[44] our study stops short incorporating them into the theoretical model and empirically analyzing their consequences. Additionally, we confine ourselves to theoretical discussions of income uncertainty and lapsation into uninsurance without examining empirically the impact these factors would likely have on the welfare benefits from long-run contracts. Last, we do not consider how the structure of optimal contracts is affected by the presence of various other government insurance programs such as Medicaid and Medicare, or private employer-based insurance, that individuals can endogenously choose to lapse into.

8 Conclusion

In this paper, we have provided a theoretical characterization of optimal dynamic health insurance contracts, shown how to compute these contracts given estimates of primitives (the stochastic health process, consumer income paths, and interest rate), and examined the structure and welfare levels of empirically relevant dynamic contracts using granular all-payers data from Utah. Our empirical results illustrate the key predictions of our framework and provide insights into the potential benefits of long-term health insurance, providing a useful benchmark for longer-run policy discussions of health insurance design.

Among our findings, we show that the welfare that optimal dynamic health insurance contracts could offer for men in Utah depends crucially on (i) whether there is government insurance of pre-age-25 health risk and (ii) the steepness of consumer income profiles. A

[43] An interesting question, is why combined health/income insurance contracts are not commonplace. One possible reason is the difficulty of contracting on income realizations (including, possibly the moral hazard it may induce) which may make any income insurance not based on health or disability difficult. An interesting question, which is beyond the scope of the present paper, is the structure of optimal health insurance contracts when income is correlated with health outcomes but not contractible.

[44] See Atal et al (2020) for a review of how such issues are managed in the German long-term health insurance market.

lack of pre-age-25 health-risk insurance greatly reduces the appeal of dynamic contracts, while, whether or not such insurance is in place, the appeal of these contracts is greater if lifetime income profiles (given available borrowing opportunities) are flatter. With pre-age-25 health risk insurance in place, consumers with flat net income prefer dynamic contracts to the managed competition-style exchange environment we study, but consumers with steeper income profiles prefer the managed competition exchange environment. When we allow for meaningful switching costs (as empirical work has shown are relevant in practice) or lower risk aversion levels dynamic contracts become more effective, while consumer myopia attenuates their benefits.

While our model is stylized in various ways, these results illustrate that there are plausible scenarios where dynamic contracts could improve welfare relative to an ACA-like managed competition-style insurance exchange. However, in practice, unlike in auto insurance or life insurance, explicitly dynamic contracts have been very rare in US health insurance markets.[45] There are some potential practical impediments that are outside the scope of our model that could limit the viability of such contracts, which we now discuss.

One concern is that firms may have difficulty forecasting future medical cost levels, an issue that does not arise to the same degree in markets such as life insurance in which long-term contracts are prevalent. This risk is not fully diversifiable.[46] This issue could be solved (alleviated), by indexing future guaranteed premiums to medical cost inflation indices in a granular manner, something that, e.g., is currently done in the German private health insurance market.

Another potential problem is that consumer lock-in might lead to quality degradation by insurers. This is something that was a major concern in the pre-ACA individual market for insured consumers with pre-existing conditions. While there are a number of ways to regulate product quality (on financial and non-financial dimensions) this is a concern that is potentially difficult to fully resolve. Cutting against this concern is the possibility that insurers' quality incentives would actually be enhanced on some dimensions, as they would have increased incentives to promote long-term health. Lock-in could also reduce a consumer's ability to re-match with firms if firm-specific preferences change (Atal (2015)). This problem would be greatly reduced if health insurance products were purely financial.

[45] However, as we noted in the Introduction, in most states, prior to the ACA insurers faced guaranteed renewability regulations that prevented them from re-pricing a policy to continuing customers on an individual basis [Patel and Pauly (2002)], and such regulations did limit the reclassification risk that consumers faced once enrolled in a policy [Marquis and Buntin (2006), Herring and Pauly (2006), Herring, Song, and Pauly (2008)]. Fleitas, Gowrisannkaran, and Losasso (2018) document a similar fact for the small group insurance market.

[46] The need to forecast could also introduce "winner's curse" type concerns, as firms who attract a lot of business would tend to be those whose forecasts of future medical cost inflation are unreasonably low.

The most serious limitation on the use and benefits of dynamic contracts in the U.S. is the short durations of insurance need in the individual market. Given the current tax-advantage for employer-based insurance, consumers may arrive only when old, or between jobs. For example, in the pre-ACA world, while some consumers purchased individual insurance over long periods of time, many others used it as a short-term solution between employment spells, leading median duration in the individual market in one study to be less than two years [Marquis et al. (2006); see also Herring, Song, and Pauly (2008)].[47] The same is currently true of the ACA individual market exchanges. Short durations greatly reduce the benefits of a long-term contract.[48] In addition, those older consumers newly arriving to the individual market with pre-existing conditions (perhaps because of a job loss) would still face reclassification risk, much as in our discussion of unhealthy 25 year-old consumers in Section 5.1, perhaps necessitating some sort of government insurance (such as high-risk pool subsidies). Removing the employer tax exemption for health insurance is one oft-discussed policy that would help promote the robustness of the individual market, whether in the ACA exchanges or in an individual market for dynamic contracts like we consider here.

Our analysis shows scenarios under which long-term dynamic contracts may be welfare improving relative to a range of alternatives. In practice, several complementary regulations are likely important to help such contracts flourish. One key factor in our analysis that helps dynamic contracts, which was not present pre-ACA, is government insurance of pre-age-25 health risk. Such insurance is crucial to prevent consumers from facing significant pre-age-25 reclassification risk. Outside of our model, it is clear that, as for the ACA exchanges, removing the employer tax exemption will improve robustness of the individual market and meaningfully increase the length of consumer spells in that market. As this suggests, extending our analysis to allow for multiple market layers (e.g. employer markets, Medicare, Medicaid) that exist alongside an individual market with dynamic contracts is an important avenue for future work.

References

[1] Abramitzky, R.(2008) “The Limits of Equality: Evidence from the Israeli Kibbutz,” *Quarterly Journal of Economics* 123: 1111-59.

[2] Aguiar, M., Blis, M. and Boar, C. (2020), “Who are the Hand-to-Mouth?,” working paper.

[47] These short durations may partly explain the apparent absence in the US of explictly dynamic contracts in the pre-ACA world.

[48] Introducing an exogenous probability of break up into our model is equivalent to lowering the discount rate, provided that separation payments upon break up cannot be made; see Proposition 6 in Appendix H.

[3] Akerlof, G. (1970) “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics* 84: 488-500.

[4] Angeletos, G., Laibson, D., Repetto, A., Tobacman, J. and Weiberg, S. (2001) “The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation,” *Journal of Economic Perspectives* 15: 47-68.

[5] Atal, J. P. (2016), “Lock-In in Dynamic Health Insurance Contracts: Evidence from Chile,” working paper.

[6] Atal, J.P, H. Fang, M. Karlsson, and N. Ziebarth (2019), “Long-term Health Insurance: Theory Meets Evidence,” working paper.

[7] Atkeson, A., and R. Lucas Jr (1992), “On efficient distribution with private information,” *The Review of Economic Studies* 59: 427-453.

[8] Barro, R. (2006), “Rare Disasters, Asset Prices, and Welfare Costs,” *Quarterly Journal of Economics* 121: 832-66.

[9] Brown, J., M. Duggan, I. Kuzioemko, and W. Woolsten (2014), “How Does Risk Selection Respond to Risk Adjustment? Evidence from the Medicare Advantage Program,” *American Economic Review* 104: 3335-64.

[10] Browne, M. and A. Hoffmann. (2013) ”One-Sided Commitment in Dynamic Insurance Contracts: Evidence from Private Health Insurance in Germany” *Journal of Risk and Uncertainty* 46: 81-112.

[11] Brunnermeier, M., T. Eisenback and Y. Sannikov ”Macroeconomics with Frictions: A Survey,” working paper.

[12] Bundorf, K., J. Levin, and N. Mahoney (2012), “Pricing and Welfare in Health Plan Choice,” *American Economic Review* 102: 3214-48.

[13] Cardon, J. and I. Hendel (2001), “Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey,” *RAND Journal of Economics* 32: 408-27.

[14] Carlin, C. and R. Town (2009). “Adverse Selection, Welfare and Optimal Pricing of Employer Sponsored Health Plans,” working paper.

[15] Chiappori, P. and B. Salanie (2000), “Testing for Asymmetric Information in Insurance Markets,” *Journal of Political Economy* 108: 56-78.

[16] Cochrane, J. (1995), "Time-Consistent Health Insurance," *Journal of Political Economy* 103: 445-73.

[17] Collier, B, H. Kunreuther, E.O. Michel-Kerjan, and D. Schwartz (2017), "Risk Preferences in Small and Large Stakes: Evidence from Insurance Contract Decisions," working paper.

[18] DeGaridel-Thoron, T. (2005), "Welfare-Improving Asymmetric Information in Dynamic Insurance Markets," *Journal of Political Economy* 113: 121-50.

[19] Farhi, E. and I.Werning (2013), "Insurance and Taxation over the Life Cycle," *Review of Economic Studies* 80: 596-635.

[20] Finkelstein, A, K. McGarry and A. Sufi. (2005) "Dynamic Inefficiencies in Insurance Markets: Evidence from Long-Term Care Insurance," *The American Economic Review* 95: 224-8.

[21] Fleitas, S., G. Gowrisankaran, and A.L. Sasso (2018), "Reclassification Risk in the Small Group Health Insurance Market," working paper.

[22] Fudenberg, D., B. Holmstrom, and P. Milgrom (1990), "Short-term Contracts and Long-term Agency Relationships," *Journal of Economic Theory* 51: 1-31.

[23] Glazer, J., T. McGuire, and J. Shi (2014), "Risk Adjustment of Health Payment Plans to Correct Inefficient Plan Choice From Adverse Selection" NBER working paper no. 19998.

[24] Golosov, Troshkin, and Tsyvinski (2016), "Redistribution and Social Insurance," *American Economic Review* 106: 359-86.

[25] Gottlieb, D. and Zhang, X. (2021), "Long-Term Contracting with Time-Inconsistent Agents," *Econometrica* 89:2 793-824.

[26] Handel, B. (2013), "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review* 103: 2643-82.

[27] Handel, B. and Hendel, I. and Whinston, M. (2015), "Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk," *Econometrica* 83(4): 1261-1313.

[28] Harris, M. and B. Holmstrom (1982), "A Theory of Wage Dynamics," *Review of Economic Studies,* 40: 315-33.

[29] Hendel, I. (2016), "Dynamic Selection and Reclassification Risk: Theory and Empirics," working paper.

[30] Hendel, I. and A. Lizzeri (2003), "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics* 118: 299-327.

[31] Hendren, N. (2013), "Private Information and Insurance Rejections," *Econometrica* 81: 1713-62.

[32] Herring, B. and M.V. Pauly (2006), "Incentive-Compatible Guaranteed Renewable Health Insurance Premiums," *Journal of Health Economics* 25: 395-417.

[33] Herring, B., X. Song, and M. Pauly (2008), "Changes in Coverage in the Individual and Group Health Insurance Markets and the Effect of Health Status," U.S. Department of Health and Human Services Report, Office of the Assistant Secretary for Planning and Evaluation.

[34] Ho, K., J. Hogan, and F. Scott Morton (2016), "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program," NBER working paper no. 21028.

[35] 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: 81-112.

[36] Illanes, G. (2017), "Switching Costs in Pension Plan Choice," working paper.

[37] Koch, T. (2011), "One Pool to Insure Them All? Age, Risk, and the Price(s) of Medical Insurance," working paper.

[38] Lavetti, K., DeLeire, T. and Ziebarth, N. (2018), "How Do Low-Income Enrollees in The Affordable Care Act Exchanges Respond to Cost-Sharing?," working paper.

[39] Marquis, M.S. and M.B. Buntin (2006), "How Much Risk Pooling Is There in the Individual Insurance Market?," *Health Services Research* 41: 1782-1800.

[40] Morris, S. and Postlewaite, A. (2020), "Observational Implications of Non-Exponential Discounting," *Revue Economique* 71: 313-321.

[41] Marquis, M.S., M.B. Buntin, J.J. Escarce, K. Kapur, T.A. Louis, and J.M. Yegian (2006), "Consumer Decision Making in the Individual Health Insurance Market," *Health Affairs* 25: w226-34.

[42] O'Donoghue, Ted, and Matthew Rabin (2015), "Present bias: Lessons learned and to be learned," *American Economic Review* 105: 273-79.

[43] Patel, V. and M.V. Pauly (2002), "Guaranteed Renewability and the Problem of Risk Variation in Individual Insurance Markets," *Health Affairs*.

[44] Pauly, M.V., H. Kunreuther and R. Hirth (1995), "Guaranteed Renewability in Insurance," *Journal of Risk and Uncertainty* 10: 143-56.

[45] Pauly, M.V., A. Percy, and B. Herring (1999), "Individual vs. Job-Based Health Insurance: Weighing the Pros and Cons," *Health Affairs* 18: 28-44.

[46] Rabin, M. (2000), "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica* 68: 1281-92.

[47] Rothschild, M. and J. E. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90: 629-49.