Horizon Effects and Adverse Selection in Health Insurance Markets

Olivier Darmouni and Dan Zeltzer*

December 16, 2017

Abstract

We study how increasing contract length affects adverse selection in health insurance markets. Although health risks are persistent, private health insurance contracts in the United States have short, one-year terms. Short-term, community-rated contracts allow patients to increase their coverage only after risks materialize, which leads to market unraveling. Longer contracts ameliorate adverse selection because both demand and supply exhibit horizon effects. Intuitively, longer horizon risk is less predictable, thus elevating demand for coverage and lowering equilibrium premiums. We estimate risk dynamics using data from 3.5 million U.S. health insurance claims and find that risk predictability falls significantly with horizon. Counterfactuals using these estimates suggest that a reform implementing two-year contracts would increase equilibrium coverage by 12–19 percentage points and yield average annual welfare gains of $600–$900 per person. A third of these effects are driven by insurers' response and the rest by changes in consumer expectations.

Keywords: Insurance markets, adverse selection, contract length

*Olivier Darmouni (omd2109@columbia.edu) is an Assistant Professor of Finance and Economics at Columbia Business School, Columbia University, New York City, New York. Dan Zeltzer (dzeltzer@tauex.tau.ac.il) is a Lecturer (Assistant Professor) of Economics at the School of Economics, Tel Aviv University, Tel Aviv, Israel. This is a substantially updated version of previous drafts. We thank Janet Currie and Sylvain Chassang for their advice. We also thank participants at the Department of Economics and the Center for Health and Wellbeing seminars at Princeton University, Tel Aviv University, the Barcelona GSE Summer Forum, and the iHEA Boston World Congress. This research received financial support from the Program for U.S. Healthcare Policy Research of the Center for Health and Wellbeing at Princeton University. Access to the MarketScan Research Databases was provided by the Truven Health Analytics graduate dissertation support program.

Introduction

While most U.S. health insurance contracts last one year, some illnesses last longer. The combination of predictable individual risk with frequent opportunity to adjust coverage may result in adverse selection. Mitigating selection is central to health insurance regulation. It often involves limiting individual choice, such as the enrollment mandate imposed by the Affordable Care Act (ACA). But with rising premiums on the one hand and enrollment in the marketplaces being concentrated almost exclusively in plans whose out-of-pocket exposure is high on the other hand, selection on both the extensive and intensive margins leaves many individuals with incomplete risk protection.[1]

This study examines how extending the length of health insurance contracts impacts adverse selection. The main contribution is to explore the benefits of increasing contract lengths as an instrument to limit market unraveling. Conceptually, longer contracts reduce selection because over longer horizons individual risk is less predictable. Lower predictability allows for better pooling of risk over time *within* individuals, as opposed to simply *across* individuals. This mechanism impacts both supply and demand, and it generally leads to greater coverage and improved welfare. We study these effects empirically, using administrative claims data on healthcare utilization. We estimate that risk predictability declines by 10%–15% per year due to mean-reversion in healthcare expenditures. Nesting these estimates in a simple model of insurance exchanges suggests that increasing contract length from one to two years would, in equilibrium, increase coverage by 12–19 percentage points and welfare by \$600–900 per person. We also show how these estimated effects change with frictions in consumer choice.

We first show how to analyze an increase in contract length using the framework of Handel et al. (2017). We focus on competitive community-rated insurance markets (like the ACA marketplaces), in which individuals periodically choose between standardized coverage levels. Graphically, increasing contract length (e.g., by extending the time between open enrollment periods) induces a flattening of both cost and demand curves. Intuitively, because individual risk is mean-reverting, individual ability to predict future risk declines, making the risk pools more homogeneous. We show that supply response leads to more coverage, while the demand response depends on how consumer expectations are affected by the policy. In general, the welfare effects of the policy are ambiguous: gains from improved coverage are partly offset by exposing those with low coverage to greater risk.[2]

[1] In 2014, only 14% of enrollees chose one of the two high-coverage plans, i.e., Gold or Platinum. By 2017, high coverage dropped to 5% (See Table A2). On a related note, premiums on these plans have been rising over the years.

[2] Because this paper focuses, for the most part, on the effects of increasing contract length from one year

We then estimate the key parameter—risk predictability—over different horizons, and find that it declines over time in a significant and robust way. We use administrative data with detailed individual healthcare costs and utilization for two separate populations. Data from MarketScan Research Databases cover a working-age population with employer-sponsored insurance, and data from Medicare cover an elderly population with public insurance. We predict individual overall expenditure over different horizons using different sets of predictors. Risk predictability declines significantly within each year: we show that two-years-ahead risk is only 85% as predictable as one year-ahead risk. Surprisingly, while predictability increases when more comprehensive predictors are used, its *decline* over time remains similar. This decline is also similar for the two different populations studied. This estimated decline in predictability reflects mean-reversion in risk. A complementary parameter—the persistence of health shocks—is shown, unsurprisingly, to be higher for the elderly population. Mean-reversion in risks suggests that longer contracts may limit the scope for adverse selection.

Finally, we provide some empirical evidence regarding the potential equilibrium effects of increasing contract length on coverage and welfare. Given the lack of naturally occurring variation in contract length in the United States, we estimate the demand response by calibrating a simple model of regulated insurance exchanges in the line of Handel et al. (2015). Depending on the specification, we predict that such change would increase coverage by 12–19 percentage points over its initial level, yielding an average welfare gain of $600–$900 per person, measured as the ex-ante willingness to pay for longer contracts. However, longer contracts are a mixed blessing: Those who upgrade to high-coverage plans enjoy welfare gains, but those who still remain with lower coverage, even when contracts are extended, suffer from longer exposure to risk without the ability to upgrade coverage. Significantly, we show that these estimates depend on consumer expectations and frictions in plan choice, as emphasized in Handel and Kolstad (2015). About a third of the effect is driven by insurers' response and the rest by changes in consumer expectations.

Practically, our results highlight how the annual periodicity of insurance enrollment in the United States has the implicit cost of increasing the scope for adverse selection. Some scope for selection is inherent in community-rated markets that feature free choice of coverage generosity, which include not only the ACA exchanges but the employer-sponsored health insurance market as well. We show that varying the frequency that coverage can be adjusted is a new instrument for reducing such selection because it allows for better risk-

to two years, we abstract from the issue of commitment. It is important to consider commitment for more long-term contracts, where it is hard to enforce (see Handel et al., 2017; Luz, 2015). However, we assume that enforcing commitment over only a slightly longer period than the current one is not substantially different.

pooling within individuals. It adds to the existing and proposed policy instruments that preserve community rating, which include enrollment mandates, risk adjustment, and premium penalties for lapsed coverage. These instruments all address a trade-off fundamental to the current healthcare policy debate in the United States—between choice and adverse-selection. We show that this trade-off also involves a temporal aspect of consumer choice and that the consequences of such choice are related to temporal properties of the underlying insured risk. Further understanding how consumers value the flexibility to adjust their coverage over time and their actual reaction to such flexibility are important questions for future research.

This study adds to the literature studying adverse selection in health insurance markets and potential policy responses. Previous works have suggested the presence of such selection (Cutler and Zeckhauser, 1998; Culter and Reber, 1998; Einav and Finkelstein, 2010; Heiss et al., 2013; Hendren, 2013) and studied its welfare implications (Einav et al., 2010; Hackmann et al., 2012; Geruso, forthcoming). Bundorf et al. (2012) document preference heterogeneity, which implies that preserving some choice among plans is useful, and show that adverse selection might not be dealt with by setting optimal premiums alone. Their findings highlight why mitigating adverse selection through increased contract length could be helpful in preserving choice while mitigating selection.

This paper contributes to policy discussions that address market breakdown in health insurance markets due to adverse selection. Proposals generally encompass two main ideas: (i) preserving the community rating but introducing other regulations, and (ii) changing pricing regulation and moving away from community rating. This paper is part of the first strand. It focuses on a contract length reform, as opposed to plan standardization (Rice et al., 1997; Ericson and Starc, 2016), risk adjustment mechanisms (Ellis et al., 2000; Glazer and McGuire, 2000; McGuire et al., 2013; Brown et al., 2014; Bundorf et al., 2012; Hackmann et al., 2015; Decarolis and Guglielmo, 2016; Geruso et al., 2016), or participation mandates (Kolstad and Kowalski, 2016).

The second strand is very promising line of research (Handel et al., 2015, 2017). In that case, the most pressing issue is reclassification risk associated with dispensing with community rating. In particular, Handel et al. (2017) study the welfare effect of long-term, risk-rated contracts. In this setting, reclassification risk is a major issue related to pricing regulation reform. However, the reform considered in this paper preserves community rating and in this context, reclassification risk is not a central concern.

In addition, a few works consider the dynamics of risk in health insurance contracts, including Aron-Dine et al. (2015) and Cabral (2016). A few theoretical works exist on adverse selection dynamics in health insurance, in settings radically different from the current regulatory environment, such as Diamond (1992), Cochrane (1995), Luz (2015), and Montanera (2017). This paper, in contrast, studies an environment similar to the current private U.S. insurance marketplaces.

The rest of this paper proceeds as follows. In Section 1 we study the impact of contract lengths on adverse selection in theory. In Section 2 we estimate a key determinant of this impact—the predictability of healthcare cost over different horizons. In Section 3 uses the model with rich nonparametric risk estimates to study the counterfactual implications of longer contract length in equilibrium.

1 Mechanism: Contract Length and Adverse Selection

In this section, we present a simple framework that incorporates varying contract length into the standard model of insurance markets with adverse selection developed by Handel et al. (2015).

We define contract length as the time between two insurance plan choices. Concretely, a contract length reform can be implemented in two ways: either by regulating insurance products by extending their term, or by offering less frequent enrollment periods. For simplicity, we do not distinguish between these two implementations. We focus on contract length as a policy instrument and conduct a comparative statics exercise with respect to the regulatory environment. We consider a reform whereby all contracts are sold for two years instead of a one-year term.[3] We assume there is no coexistence of one-year and two-year contracts, nor are there hybrid contracts.[4]

We study selection among standardized plans of varying generosity that are community-rated, like in the ACA or the employer-sponsored insurance marketplaces.[5] The focus is on analyzing the demand and supply response to contract length reform in order to study the effects of longer contracts on equilibrium coverage and welfare. The objective is to uncover a new potential benefit of extending contract length in community-rated markets. On the other hand, the potential issues associated with longer contracts, e.g. commitment, are well understood (see for instance Handel et al. (2017)). Nevertheless, we make the plausible assumption that one-year and two-year contracts are equally enforceable.[6]

[3]The analysis extends to comparing short- and long-term contracts more generally.

[4]Sustaining an equilibrium with both long-term and short-term contracts seems difficult in practice. For a theoretical example using a dynamic pricing scheme, see Luz (2015).

[5]We choose to focus on the intensive margin of selection within coverage levels, conditional on buying insurance. But a similar logic applies to the extensive margin of selection into buying insurance.

[6]This is not to say that they are perfectly enforceable. Currently, individuals can change their coverage at any time during the year when any one of several qualifying life events occurs (such as marriage or change of employment). However, the incentives to abuse such exemptions are unlikely to be drastically different across

The mechanism we described is fundamentally different from the well-known idea that lifetime contracts chosen behind the veil of ignorance would remove adverse selection. Instead, we study a realistic potential reform that preserves three fundamental features of the current ACA health exchanges: consumer choice over plans, insurer competition, and community rating.

1.1 Health Insurance with Different Contract Lengths

Consider an insurance marketplace in which individuals periodically choose between two health insurance benefit plans—High and Low—with coverage levels $\iota_H > \iota_L$.[7] Denote by p the extra premium for purchasing High. Individuals maximize utility and have an underlying type $\theta_0 \in [\underline{\theta}, \bar{\theta}]$, which determines their propensity to choose High over Low. This underlying type may depend on both individual risk (i.e., healthcare utilization costs) and other characteristics. When utility satisfies the single crossing condition, the aggregate demand for insurance is characterized by the marginal type $\theta^*(p)$ that is indifferent between the plans at price p, and is given by:

$$Q(p) = \int_{\theta^*(p)}^{\bar{\theta}} dF(\theta). \tag{1}$$

We assume supply is regulated: premiums are community rated, rejections are prohibited, and coverage levels are standardized. The marginal cost of providing higher coverage to type θ^* is:

$$\Delta MC(p) = (\iota_h - \iota_l) C(\theta^*(p)), \tag{2}$$

where $C(\theta^*(p))$ is the expected cost of type θ^* during the contract term. To facilitate the comparison between one-year and two-year contracts, assume the premium difference p and costs are annualized. The difference in the average cost between the plans' risk pools is:

$$\Delta AC(p) = \iota_H \int_{\theta^*(p)}^{\bar{\theta}} C(\theta) dF(\theta) - \iota_L \int_{\underline{\theta}}^{\theta^*(p)} C(\theta) dF(\theta). \tag{3}$$

Proposition 1 in Handel et al. (2015) characterizes the Riley equilibrium. As long as adverse selection is severe enough, there is market breakdown and the equilibrium price

one-year and two-year contracts. Handel et al. (2017) provides an in-depth discussion of long-term contracts with imperfect commitment. Given that enforcement frictions are particularly challenging to measure in our context, we leave the question of optimal contract length to future research.

[7] The analysis can be extended to more general contracts. We assume that regulated contract features such as co-insurance rates do not change with contract length, leaving the interaction between these policy instruments for future research.

difference is the lowest break-even p with positive sale of plan L.[8] Because of competition, the premium difference p reflects the difference in average cost across plans in equilibrium, i.e., $\Delta AC(p) = p$.

Adverse selection arises as costlier types are more inclined to purchase higher coverage. In such case, the marginal cost curve is downward sloping; the equilibrium price is generally higher than the efficient price, obtained when $\Delta MC(p) = p$. Namely, adverse selection leads to underprovision of coverage.

Significantly, the key driver of the magnitude of adverse selection is the relevance of individual type for plan choice. Adverse selection is smaller when individuals with different risk types make similar plan choices.

This channel explains the effect of contract length reform on insurance demand and supply. In this setting, *risk predictability* will impact adverse selection. The less predictable risk is, the smaller the scope for selection, because the difference in expected risk $C(\theta_0)$ is smaller across different types θ_0. Graphically, when risk is less predictable, MC is flatter.[9] Moreover, risk predictability depends on contract length, because the time between plan choices determines the horizon over which predictions are to be made. Over longer horizons, individual type θ_0 is a less effective predictor of future risk. As a consequence of this decline in risk predictability, the MC and AC curves are both flatter with longer contracts (Figure 1).

This pivoting of the MC curve reflects an essential property of risk—mean reversion. Intuitively, over longer horizons, the expected risk of different types are slightly more like that of the mean type and therefore slightly closer to each other. We document mean-reversion in health expenditures in detail in the next section. In what follows, we highlight the effects of contract length on supply, demand, coverage, and welfare by focusing on the comparison between one- and two-year contracts. These horizons match our empirical estimates of risk predictability in the next section.

1.2 Equilibrium Effects on Coverage

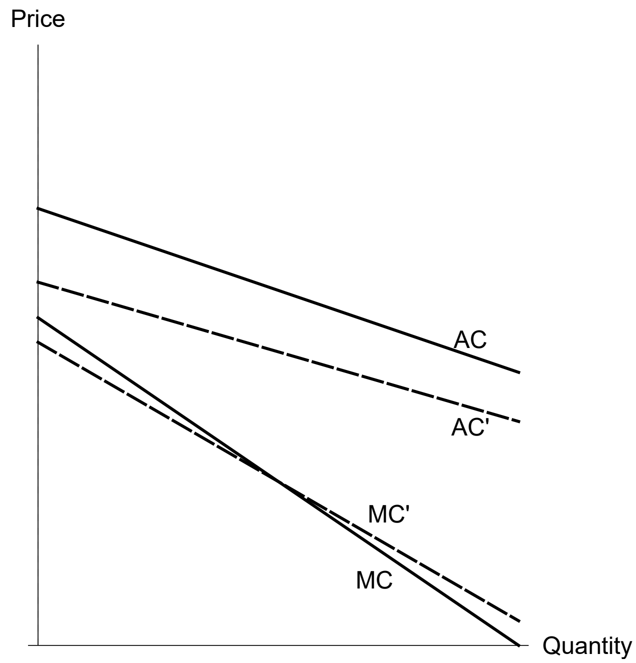
We consider the effect of contract length reform by comparing two regulatory environments: one with one-year contracts and one with two-year contracts. We use our previous analysis to represent this comparative static exercise on one graph and argue that longer contracts change both consumer demand for insurance and its pricing by insurers.

The first main observation is that on the supply side, an increase in contract length increases coverage. Intuitively, coverage increases through a flattening of the AC curve.

[8] Formally, adverse selection is severe enough if $\Delta AC(\theta^{*-1}(\underline{\theta})) > \theta^{*-1}(\underline{\theta})$.

[9] In the extreme case of totally unpredictable risk, $C(\theta_0)$ is constant and therefore MC is fully horizontal. MC also equals AC in this case.

Figure 1: The Effect of Increasing Contract Length on Cost



Notes: Average and marginal cost with one-year contracts (solid) and two-year contracts (dashed). Because of mean reversion, over longer horizons, between-individuals differences are smaller and more within-individual pooling is possible. Therefore, the difference in average cost between risk pools with High and Low coverage, AC, is smaller for all quantities.

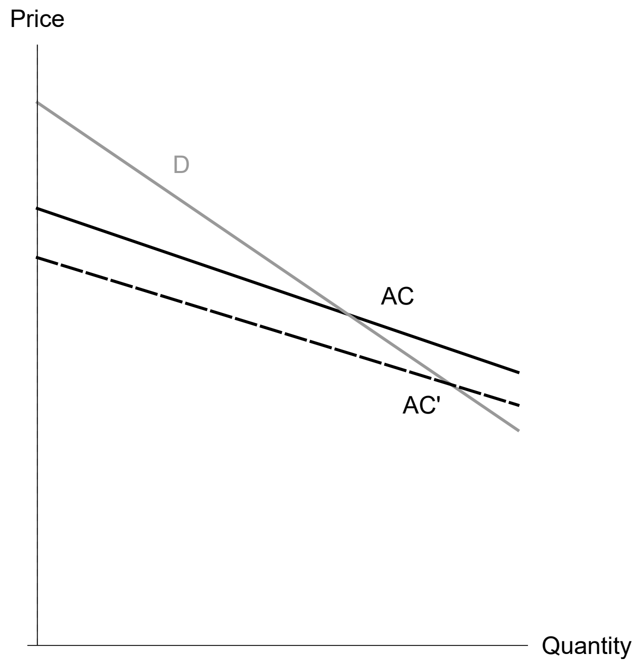
Because AC shifts uniformly towards zero, this result is very general (Figure 2 illustrates this point). As contract length increases, mean reversion in risk leads to a flattening of the marginal cost curve and to a decline in the difference between risk pools for all coverage levels. From the point of an insurer, individuals become more homogeneous. For any given risk pool choosing the High plan at price p, the average risk falls. The difference between high- and low-risk types is more moderate, and consequently the dependence of insurers' profits on the composition of the risk pool is reduced. Under very general conditions, this change in the average cost curve induces an increase in equilibrium coverage.[10]

Increasing contract length reduces average cost by pooling risk over time, *within* an individual. This contrasts with existing contracts that only pool risk *across* individuals. Longer contracts exploit the dynamics of risk: because risk displays some degree of mean-reversion, pooling risk over time reduces the sensitivity of insurers' profits to individual risk types.

Interestingly, the effect of contract length on demand can increase or decrease coverage. Figure 3 illustrates the intuition behind this result. In both cases, an increase in contract length flattens the demand curve. Significantly, this flattening occurs by pivoting around

[10]Specifically, as long as the demand curve crosses the AC curve from above.

Figure 2: Supply Response with Extended Contract Length



a point that has an important economic interpretation. The demand curve pivots counter-clockwise around the valuation of the *representative* type $\bar{\theta}$.

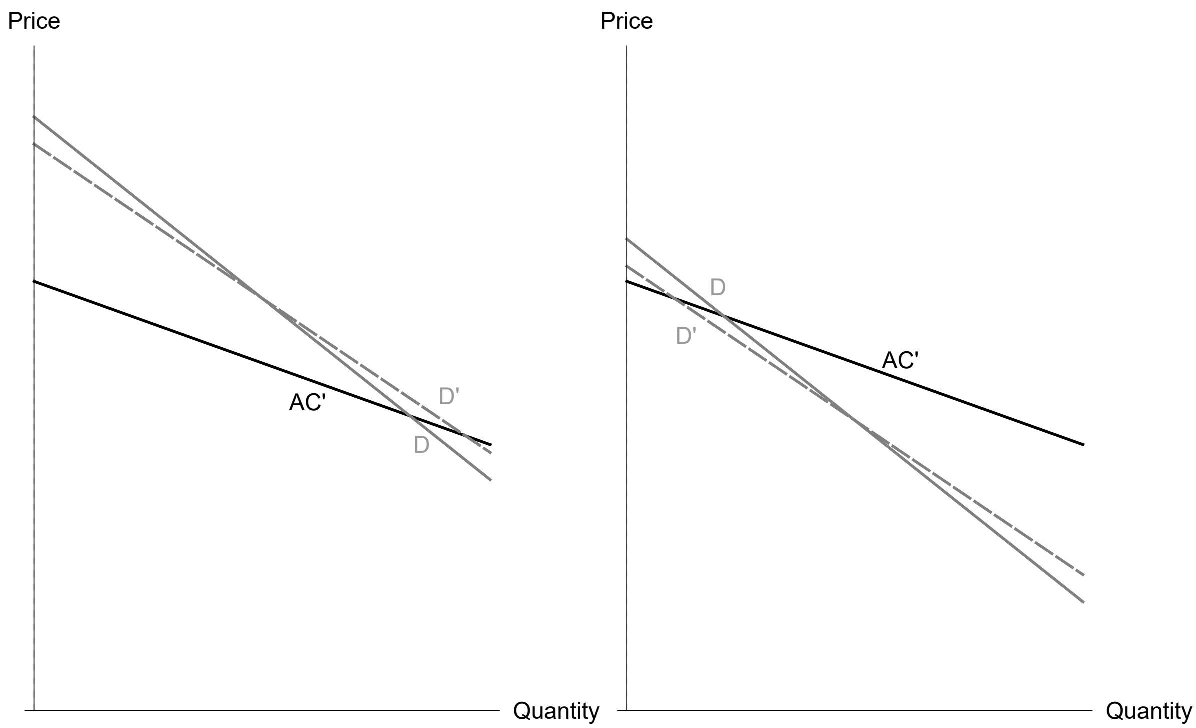
Indeed, mean-reversion in health risk implies that types become effectively more homogeneous with longer contracts. The representative type $\bar{\theta}$ is defined as the type whose value of insurance is left unchanged by extending contract length. In a linear-uniform example, it would correspond to the average type $\mathbb{E}[\theta]$, but in general, it can sit at any point on the risk distribution. The next section shows how to estimate it from the micro-data on health expenditures.

This explains the ambiguous demand channel. Effectively, as types become more homogeneous, individual plan choices tend to mirror the plan choice of the representative type. In the left panel of Figure 3, the representative type originally does not choose the High plan. Therefore, as the contract length increases, fewer individuals choose the High plan. On the other hand, the right panel presents the case in which the representative type initially chooses the High plan. In this case, increasing contract length increases the share of the population choosing the High plan.

1.3 The Effects of Contract Length on Welfare

In general, the welfare effect of the policy is ambiguous. As contract length increases, there are two effects that go in opposite directions. On the one hand, more individuals

Figure 3: Demand Response and Equilibrium with Extended Contract Length



(a) $\theta^* > \bar{\theta}$ (b) $\theta^* < \bar{\theta}$

Notes: Equilibrium with one-year contracts (solid) and two-year contracts (dashed). In case (a) the representative type purchases High coverage; the demand response augments the supply response and further increases equilibrium coverage. In case (b) the representative type does not purchase High coverage; the demand response dampens the supply response. In this example, coverage still increases in equilibrium, but less than in case (a).

choose the High plan, reducing the risk faced by the consumers who upgrade. On the other hand, consumers who stay on the Low plan face greater risk. Intuitively, they will face future interim shocks, which are unrelated to their current type θ_0, but (partly) persist until their next opportunity to upgrade to the High-coverage plan. Increasing the horizon is not favorable to everyone: the healthiest types—those that choose the Low plan even after the reform—can face a welfare loss, as they loose the ability to move to the High plan if they face a bad shock after the first year. The reform can potentially generate ex-ante welfare gains (before types are revealed) but no Pareto improvements ex-post. The size and sign of the welfare effects of the reform are empirical questions that we tackle in detail in Section 3.

1.4 The Role of Consumer Expectations

Note that the demand channel depends on consumers updating their expectations about their risk over the longer contract horizon. However, consumers may not always respond optimally to information (Abaluck and Gruber, 2011). The demand curve would not flatten if consumers ignore the effect of mean-reversion in risk.[11] As emphasized in Handel et al. (2015), there are a number of informational frictions that might prevent customer demand from responding to the policy change. In Section 3, we therefore show how the policy counterfactual changes if consumer expectations are different from those of the econometrician.

Ideally, we need detailed data on expectations to calibrate the demand side response. In fact, without joint choice and survey data we cannot distinguish between inertia, information frictions, or incorrect expectations, as explained in detail in Handel and Kolstad (2015). We therefore model theses frictions in a reduced form in Section 3. Handel and Kolstad further emphasize two empirical challenges. First, frictions bias estimates of risk aversion and the bias can go either way, so ex-ante, it is challenging to know how to correct for it. We will therefore carry out our counterfactual exercise for a wide range of consumer risk aversion. Second, for welfare analysis, one must take a stance on whether these frictions are welfare-relevant.

2 Empirical Estimates of Risk Predictability

We estimate the predictability of risk at different horizons—the key parameters that determine equilibrium coverage change. We estimate these parameters using healthcare utilization and cost data from two different administrative data sources: MarketScan Research Databases and Medicare Cost and Utilization Databases. We find similar results across these

[11] However, note that the positive supply response only depends on insurers' beliefs.

different samples: when moving from a one-year to a two-years horizon, risk predictability diminishes significantly.

2.1 Data

We use administrative data on healthcare utilization and costs. Important for our purpose, these data track individuals over time and, because they are sourced from payors, they are comprehensive. Data of the same granularity are used by insurers for planning and pricing, and thus the internal validity of our estimates is presumably high. The data's external validity is further supported by the use of two distinct sources covering different populations and different periods. The rest of this section describes these data sets.

The first data source is MarketScan Commercial Claims and Encounters Database.[12] These data capture individual clinical utilization, expenditures, and enrollment across inpatient, outpatient, prescription drug, and carve-out services from approximately 45 large employers for 2002–2004. These data represent the medical expenditures of the working-age population, an age profile similar to that of the target population of the ACA marketplaces.

The second data source is Medicare, the federal health insurance program for people who are 65 or older.[13] The data contain claims from Fee-For-Service Medicare beneficiaries. We use the Research Identifiable Files of the Master Beneficiary File and Cost and Utilization databases. These longitudinal databases record, for a 5% sample of Medicare beneficiaries, all healthcare utilization and costs over the period of 2008–2012. For comparability with the population in the ACA marketplaces and the MarketScan sample, both of which cover working-age individuals, we restrict the sample to beneficiaries aged 65–70 when entering the sample.[14]

Table A3 shows summary statistics for each of the samples. The difference in age profiles and utilization patterns work in our favor, as they help us gauge the robustness of our estimates of the decline in risk predictability.

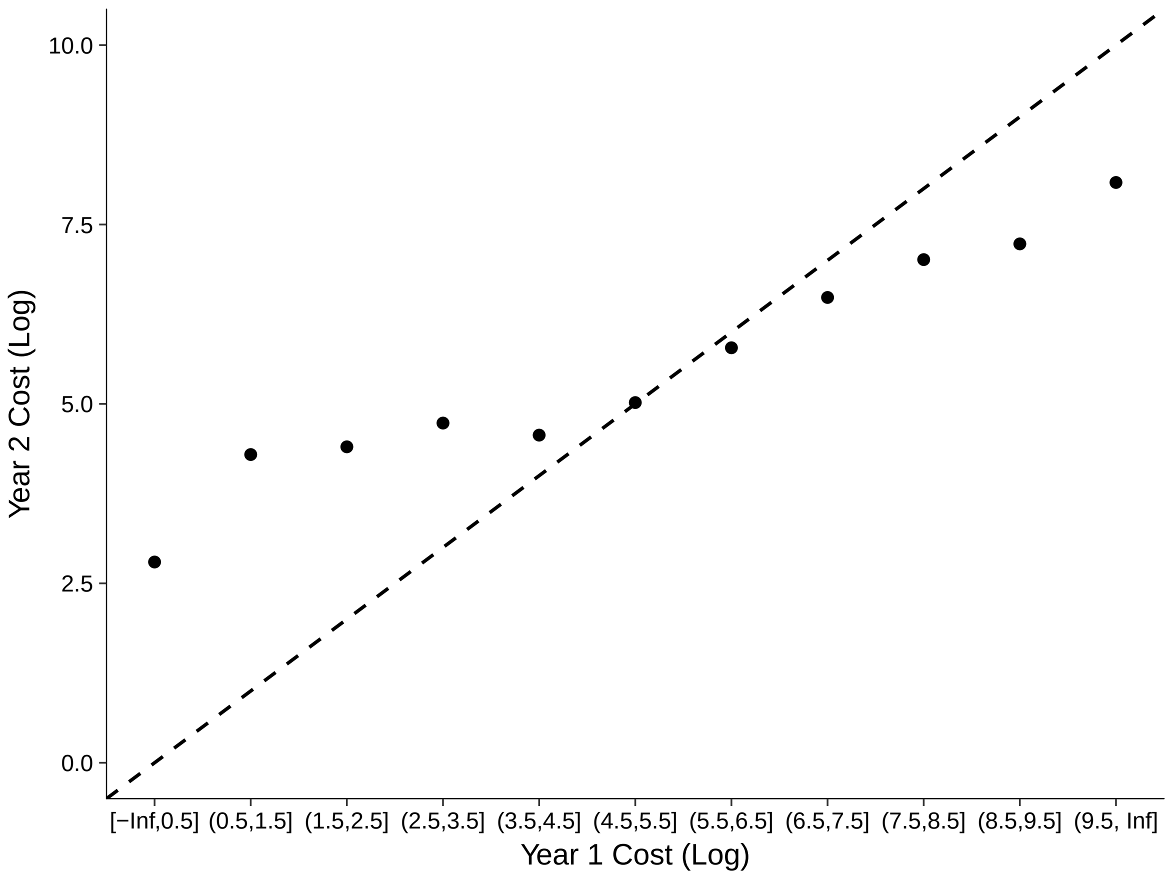
Figure 4 visualizes the key parameter that drives our results: mean reversion in risk (i.e., in healthcare costs). It is a binned scatter plot of annual log costs of our sample of 3.5 million MarketScan working-age enrollees. While risk is persistent—bins preserve their rank order year-over-year—it also exhibits substantial reversion to the mean. Our estimates of (4), presented below, show this reversion holds even when other predictors of risk beyond

[12] MarketScan Research Databases, Truven Health Analytics, Ann Arbor, MI.

[13] We consider only elderly beneficiaries, although Medicare also covers certain younger people with disabilities and people with end-stage renal disease.

[14] For a complete list of utilization and cost measures used in the prediction of risks, see The Master Beneficiary Summary File and Cost and Use segment, `http://www.resdac.org/cms-data/files/mbsf/data-documentation`, accessed June 2016.

Figure 4: Mean Reversion in Individual Healthcare Costs



Notes: Binned scatter plot of individual healthcare costs over two subsequent years. Costs were aggregated at the individual level from administrative claims of 3.5 million working-age employees and their dependents from MarketScan Research Databases. The first-year costs were then binned into 11 equally spaced bins, with each dot representing the mean cost of the subsequent year for the individuals in the bin. The dashed line is the 45-degree line, representing unchanged costs over time. Differences between the means of each pair of bins are all significant (p<0.01).

cost are included. Figure A2 shows the sample distribution of log cost c_1.

2.2 Empirical Model of Risk Predictability over Different Horizons

To compare the predictability of risk over different prediction horizons—one and two years—we study a finite risk process of the following log-linear form:[15]

$$c_1 = \mu + \alpha_1 \theta_0 + \varepsilon_1 \tag{4a}$$

$$c_2 = \mu + \alpha_2 \theta_0 + \beta_1 \varepsilon_1 + \varepsilon_2 \tag{4b}$$

where c is log individual cost (i.e., $C_{t,i} = e^{1+c_{t,i}}$ for $t = 0, 1, 2$), and θ_0, ε_1 and ε_2 are mutually independent, zero mean. We let individual type θ_0 be a function of both the baseline cost C_0 and other characteristics known at $t = 0$, which we denote by X_0. The key parameters determining the effects of changing the contract length in equilibrium are α_1 and α_2. These parameters represent how predictive current risk type θ_0 is of future risks for the next year and two years from now, respectively.

We estimate the model parameters using patient-level data in two steps. As a first step, we estimate an OLS regression of the next year's expenditures on the current information set:

$$c_{1,i} = \mu_1 + \gamma X_{0,i} + \varepsilon_{1,i} \tag{5}$$

where $X_{0,i}$ includes patient characteristics potentially predictive of future costs, such as cost, demographic information, and former utilization that are known initially. To study how predictability varies with time, we "normalize" $\alpha_1 = 1$ by defining the type θ_0 as the demeaned forecast:

$$\hat{\theta}_{0,i} = \hat{\gamma} X_{0,i} - \frac{1}{N} \sum_i \hat{\gamma} X_{0,i} \tag{6}$$

This is the empirical counterpart of the private risk type θ_0 introduced in the general model above.[16] For robustness, we tried different definitions of $X_{0,i}$, ranging from cost-only to detailed utilization, demographics, and the presence of any of multiple chronic conditions. Appendix Table A1 summarizes the different sets of predictors we considered.

As a second step, we estimate α_2, capturing the decline in predictability due to mean-reversion, by regressing expenditures in two years' time on initial risk type $\hat{\theta}_{0,i}$ and the

[15]The model can easily be generalized to an infinite horizon AR model, or even ARIMA. Empirically, we use a log-linear specification that well fits the large skewness in health spending.

[16]For now, we assume risk is one dimensional, although this could be generalized (say, to distinguish between chronic and transitory conditions).

interim cost shock $\hat{\varepsilon}_{1,i}$:

$$c_{2,i} = \mu_2 + \alpha_2 \hat{\theta}_{0,i} + \beta_1 \hat{\varepsilon}_{1,i} + \varepsilon_{2,i} \tag{7}$$

Because we normalized $\hat{\alpha}_1 = 1$, we expect $\hat{\alpha}_2 < 1$, such that risk predictability falls with the horizon. We also expect $0 < \hat{\beta}_1 < 1$, such that interim cost shocks display some degree of persistence. Whenever a longer panel is available, the same method can be extended to estimate risk predictability at horizons longer than two years.

Note that this specification of risk does not explicitly account for moral hazard, i.e., that plan generosity directly affects spending. Moral hazard is less of a problem in Medicare, where the coverage level is fixed. In the other context, this effect leads, if anything, to underestimating mean reversion. Indeed, moral hazard induces persistence (higher α_2) in this setting: after a high realization of first-year cost c_1, one can switch to a more generous plan. Moral hazard thus pushes cost c_2 in the second year upwards, reducing mean-reversion. The fact that the estimated decline in risk predictability turns out to be similar in these two settings that differ in the potential scope for moral hazard, suggests that moral hazard has in fact little influence on our estimates.

2.3 Results

Table 1 presents estimates for the decline in risk predictability between one-year and two-year horizons. Across all specifications and for both samples, the estimated coefficient $\hat{\alpha}_2$ is close to 0.85 and statistically significantly different from 1. That is, at a two-year horizon, risk predictability is only about 85% of that at a one-year horizon. The coefficient capturing the persistence of interim health shock, $\hat{\beta}_1$, is about 0.4 in the MarketScan data and 0.6 in the older Medicare sample, which is to be expected. It is surprising that while predictability increases when more comprehensive predictors are used, its decline over time remains similar across specifications and in different settings.

Risk predictability declines over different horizons for the Medicare sample, where we have five data years, 2008–2012.[17] Table 2 shows estimates of the magnitude of the decline; one-year predictability is normalized to 1. Clearly, both the coefficients and the goodness of fit decrease over time. The constant decrease in predictability over time suggests the difference we focus on—between one-year and two-year horizons—further generalizes to longer periods.

[17]For comparability across horizons, we restrict the sample to include only individuals who were not covered throughout the period observed. (This restriction excludes some attrition due to mortality, but comparing the estimates with Column 1 of Table 1 shows that the impact of the sample restriction on the estimate is negligible.)

Table 1: Predictability: Two-Year Horizon, Different Predictors

	Dependent Variable: *Log Future Medical Spending*		
	Information Set		
	Total Cost	+Demog.	+Utilization
	A. MarketScan Sample		
α_2	0.842	0.868	0.871
	(0.001)	(0.001)	(0.001)
β_1	0.401	0.381	0.374
	(0.001)	(0.001)	(0.001)
R Sqr.	0.294	0.304	0.307
N	3,468,253	3,468,253	3,468,253
	B. Medicare Sample		
α_2	0.857	0.857	0.862
	(0.00140)	(0.00139)	(0.00135)
β_1	0.644	0.643	0.631
	(0.00193)	(0.00193)	(0.00195)
R Sqr.	0.621	0.621	0.623
N	456,482	456,482	456,478

Notes: Prediction of model (1) with log annuitized healthcare expenditure as the risk measure, and different predictor sets, defined in Table A1.

Table 2: Predictability Over Different Horizons

	Dependent Variable: *Log Future Medical Spending*		
	Prediction Horizon		
	2 years	3 years	4 years
α_2	0.858	0.772	0.686
	(0.00185)	(0.00209)	(0.00226)
β_1	0.640	0.543	0.472
	(0.00248)	(0.00266)	(0.00277)
R Sqr.	0.618	0.503	0.407
N	235,927	235,927	235,927

Notes: Prediction of model (1) with log annuitized healthcare expenditure as the risk measure, and different prediction horizons, for a balanced panel of Medicare beneficiaries over 2008–2012.

3 Counterfactuals Under a Contract Length Reform

The previous section estimated how risk predictability declines with the horizon, suggesting that the mechanism highlighted in Section 1 is empirically plausible. Namely, increasing contract length should reduce the scope for adverse selection in community rated markets by extending the horizon over which predictions are made. In this section, we use non-parametric risk estimates together with the model from Section 1 to study the magnitude of the effects of extending contract length on coverage and welfare in equilibrium. We compare the equilibrium level of coverage, premiums, and ex-ante expected utility between the current regime of one-year contracts and an alternative regime with two-year insurance contracts.

Evaluating equilibrium outcomes under different contract horizon regimes requires three components. First, individual risk, and its transition over time, which we directly observe in our claims data for a large sample of individuals. Second, risk preferences, to determine contract choice in response to risk and to evaluate the welfare implications of different equilibrium outcomes. As we do not have sufficiently rich data on plan choice to estimate preferences, we study a range of possible preferences: different functional forms and a range of risk aversion parameters. We also check the importance of consumer expectations, by contrasting coverage choices implied by rational updating of risk types with "myopic ones."[18] Third, an equilibrium concept is required for clearing the market and to determine how premiums and coverage are set. We use the concept of Riley equilibrium, discussed in Section 1. Simulating markets under different scenarios, we compare outcomes and welfare (from an ex-ante perspective) between different contract length regimes.

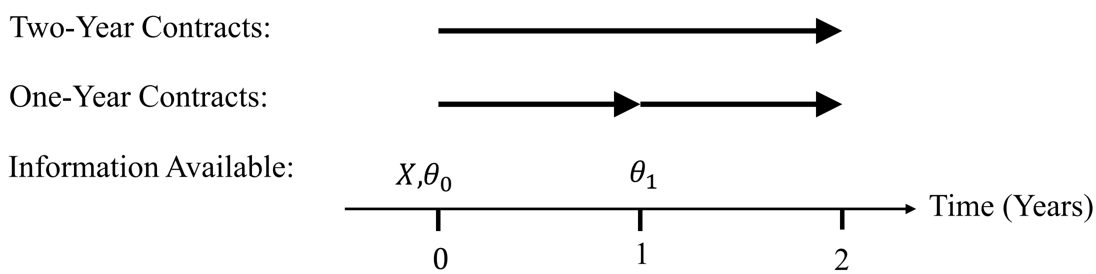
Our estimates suggest that moving to a regime in which contract length is longer by just one year would lead to significant gains in both coverage and welfare, thanks to reduced adverse selection. That is, the average individual would be willing to pay ex-ante to move to a longer contract regime, where they see lower premiums and have on average better coverage. These results hold for a wide range of risk aversion levels and with different specifications of individual utility functions.

While our findings hold for different utility functional forms, risk aversion levels, and assumptions about consumer expectations, they should be interpreted while keeping in mind that our modeling choices are purposely stark, that the risk data we use come from a specific population with employer-sponsored health insurance coverage, and that there is no exogenous variation in contract length to estimate demand response directly. Nevertheless, we believe that these estimates highlight a clear mechanism and provide insight regarding the

[18] Specifically, the "myopic" case assumes that there is no shift in the demand curve following the reform. It is indeed plausible that many consumers would not replicate the econometric analysis of predictability described in Section 2.

merits of extending health insurance contract length. Furthermore, our framework can be flexibly adjusted and reused to study alternative contract structures, preferences, and risk.

Figure 5: Contract Regimes and Information Structure



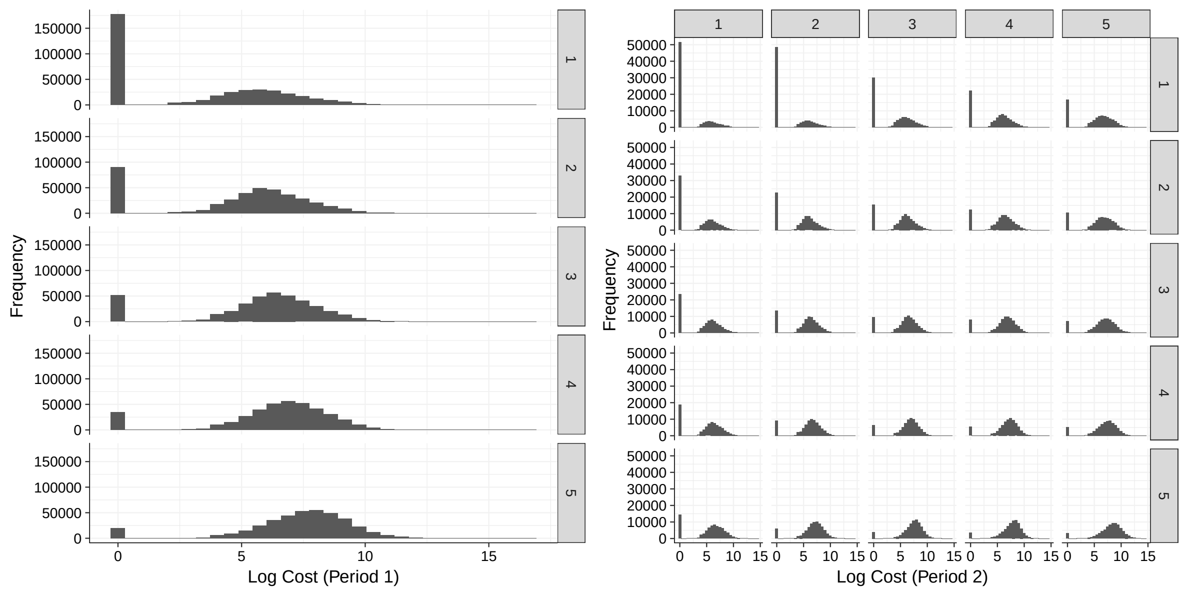
Setup We compare equilibrium outcomes between two regimes: one- and two-year contracts. In both regimes, individuals can periodically adjust their coverage level. In one case, they can do so every year.[19] In the other case, coverage can be adjusted only every other year. The key difference between one-year and two-year contracts is that with one-year contracts, individuals may adjust their second-year coverage levels based on their first-year risk realization, whereas with two-year contracts, they may not. This difference is directly captured in our analysis: in our simulations of coverage choices of one-year contract are annual, and therefore individuals make twice as many choices compared with the case of two-year contracts, where choices are biennial (once every two years). Furthermore, the additional interim choices are allowed to depend on interim information. The timing of choice and information is highlighted in Figure 5. We set the standardized menu of two plans, High and Low, to be actuarially equivalent to the Bronze and Gold plans in the ACA exchanges, by setting the coinsurance rates to be 40% and 10%, respectively ($\iota_H = 0.9$ and $\iota_L = 0.6$).[20]

Risk We estimate the empirical risk distribution corresponding to different risk types non parametrically. We do so in five steps. First, we partition the sample into eight demographic cells (X), based on age group and gender. Second, within each one of eight demographic cells, we estimate the risk model (6) and calculate the (predicted) risk type $\hat{\theta}_0$ for each

[19]This baseline case—annual choice of contracts during an open enrollment period—is currently the most common setup in the United States. It is in place in both the Affordable Care Act health insurance exchanges and most employer-sponsored health insurance markets. Certain qualifying events allow some people to buy contracts for even shorter periods (see Aron-Dine et al., 2015).

[20]This framework can be flexibly adjusted, for example, to acommodate the case when plans also have out-of-pocket spending limits.

Figure 6: Empirical Cost Distribution by Risk Type



Notes: The two panels demonstrate our non parametric estimates of risk distributions using 3.5 million individuals in MarketScan. The left panel shows risk histograms for five quantiles of fitted values of $\hat{\theta}_0$, based on demographic and risk information available before the first period. The right panel shows the additional partition of the type space by $\hat{\theta}_1$, based on first-period information. The grid rows in the right panel correspond to the same risk quantiles in the left panels, whereas columns correspond the the additional partition of the type space based on the first-period information. For illustration purposes, these estimates are based on data from the entire distribution of demographic characteristics. Our analysis performs this partition of the risk type space within each demographic cell separately, for a total of 200 risk types. The mass points at zero capture individuals with no medical claims within the year.

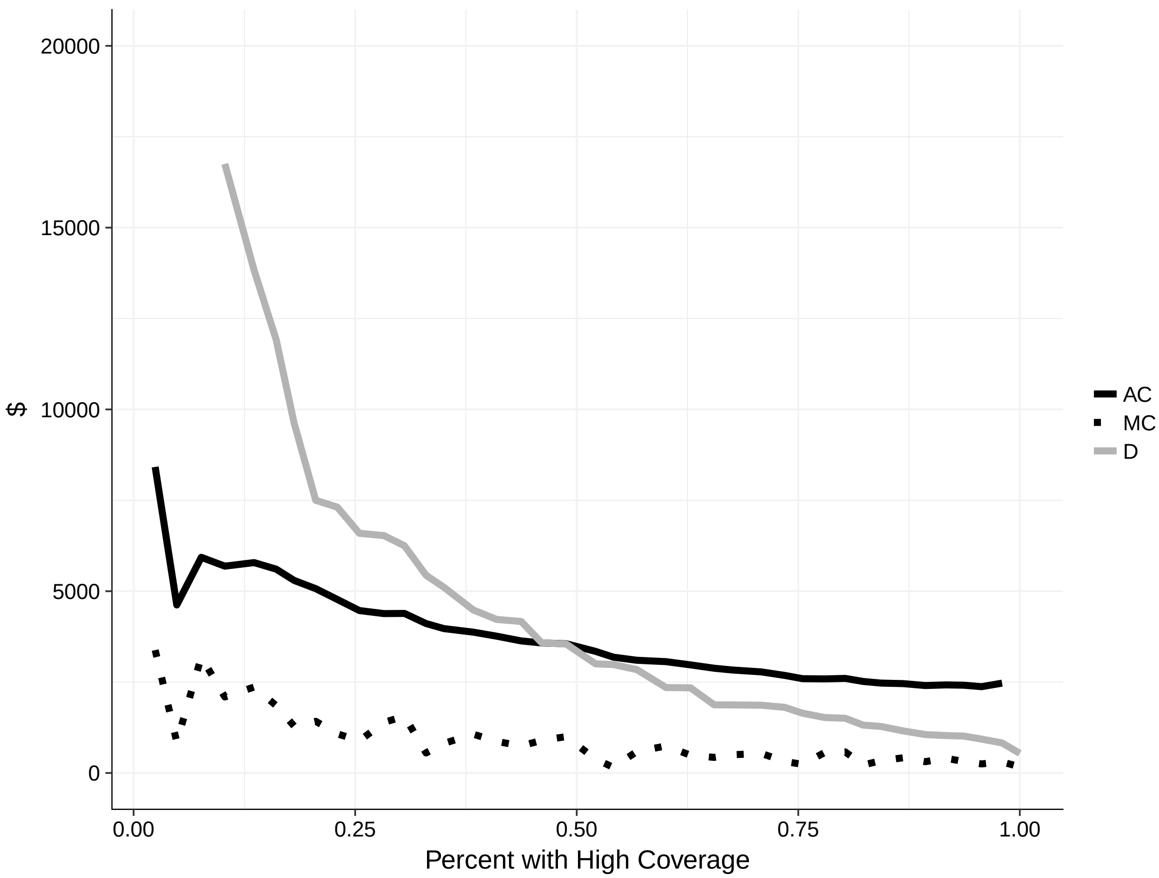
individual. Third, we bin $\hat{\theta}_0$ into five equally sized bins (the risk-type quantiles within each demographic cells X), to obtain a total of 40 risk types, each with its own non parametric distribution of future costs. Fourth, for each type, we bootstrap 10,000 trajectories of risk to obtain the non parametric joint distribution of risk in different periods (weighing the results to preserve the population distribution of types). Fifth, to determine the risk distribution faced by types making interim choice (in the one-period contract regime), we refine the type space based on first-year risk information, to obtain a total of 200 risk types for the second-period contract choices. This increase in the number of types over time reflects the increase in information available to individuals. We use the resulting joint-distribution of risk over the different periods for each type to analyze the different contract regimes. The average risk in our MarketScan population is \$3,177 (in 2003 dollars). However, as is typical of medical expenditure risk, risk varies both across and within demographic cells. The average annual risk by age group and gender is shown in Figure A1.

Supply To quantify the supply-side effects of one- versus two-year contract regimes, we estimate the marginal and average cost curves (namely, ΔMC and ΔAC described in Equations 2 and 3) for different contract lengths using the empirical risk estimates from MarketScan Claims discussed above. Assuming that individuals adversely select coverage, we order individual types by their willingness to pay for coverage (discussed below) and determine the marginal and average costs given individual choices. In this way we calculate premiums for two risk pools: the 90% coverage plan (High coverage) and the 60%-coverage plan, which we assume is the minimal level of coverage allowed under the mandate.

Demand In the absence of exogenous variation in contract length (all contracts in our data are annual), we derive demand by assuming that individuals maximize expected utility given the estimated risk, based on their type. That is, let u be a vNM utility index, and $f(C)$ define a contract—specifying plan premium, and out-of-pocket spending given realized cost C. We assume that an individual of type θ_t chooses coverage $b \in \{L, H\}$ to maximize expected utility $E_t[\sum_{s=1}^{T} u\Big(f_b(C_{t+s}(\theta_t))\Big)]$, where $T = 1, 2$ is the contract length. When, in the one-year contract case, choice has to be repeated, the information set is updated. Conditional on information, choice involves no dynamic aspects, so in the case of repeated choices we only update the information on which choices are based (θ_t). For most of our analysis we use mean-variance utility index $u(c) = \mathrm{E}[c] - \lambda \mathrm{Var}[c]$, and contracts f specified by a pair of premium and coinsurance levels.[21] As the horizons we study are short, we

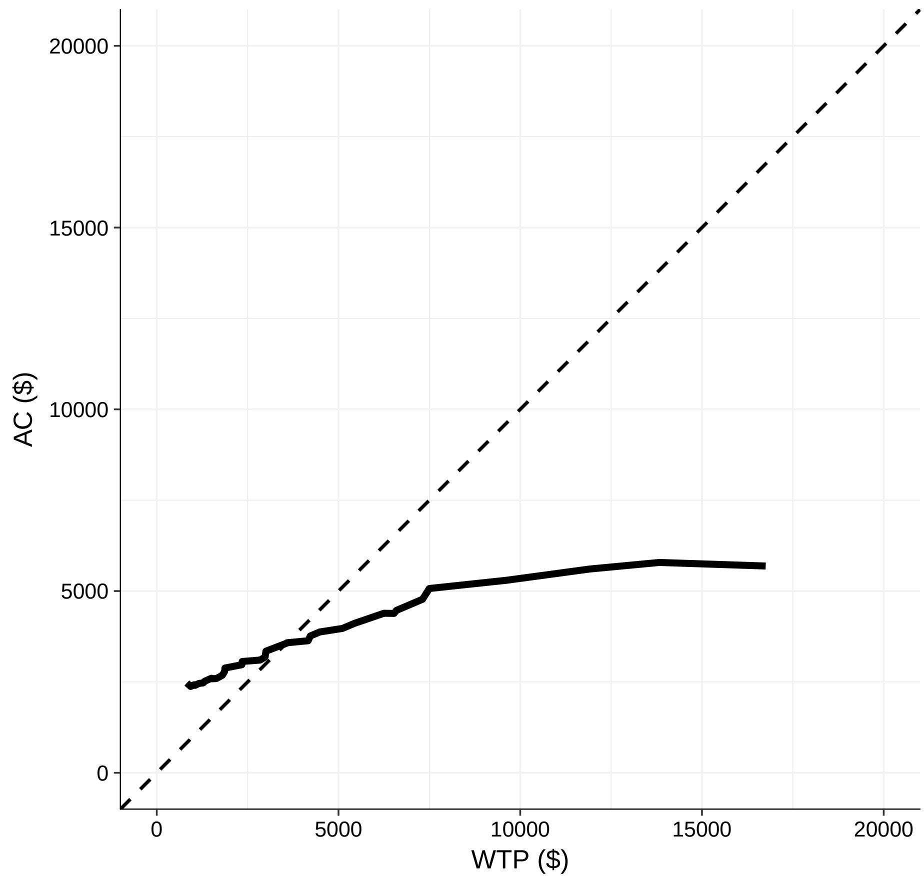
[21] We obtain similar results with CARA utility. The specific functional form specified here can be replaced by others that satisfy the single-crossing property.

Figure 7: Simulated Market Equilibrium: Cost and Demand



Notes: This figure shows estimated ΔMC (green), ΔAC (red), and demand (ΔWTP, blue). These are the empirical counterparts to the schematic figures from Section 1 obtained from one of our counterfactual simulations of the market. Costs are estimated using the empirical distribution of healthcare costs in MarketScan data. Demand is simulated using a mean-variance utility (in the case shown $\lambda = 2 \times 10^{-4}$). We perform such simulations comparing the case of one- and two-year contracts with different utility functions and different levels of risk aversion.

Figure 8: Reilly Equilibrium: Cost and Willingness to Pay for Higher Coverage



Notes: Riley equilibrium in a market for insurance is obtained when AC (denoting ΔAC, the premium difference between the Low- and High-coverage plans) equals WTP (the willingness of the marginal type to pay for higher coverage). AC is based on the non parametric risk estimates from the MarketScan claims database discussed in Section 3. The figure shows an example of this calculation for the first-period market with $\lambda = 2 \times 10^{-4}$. We perform such calculation with one- and two-year contracts, with different utility functions, and with different levels of risk aversion.

assume no discounting.

One potential concern with our baseline model is that demand response to increased contract length may be overstated. In particular, it is plausible that few consumers would replicate the statistical analysis of risk at different horizons described in Section 2. If that is the case, consumers' beliefs would differ from those of insurers or the econometrician. Indeed, as documented before, insurance benefit choices may exhibit inertia or behavioral biases.

To separately evaluate the importance of demand-side change in consumer expectations from the supply-side reaction of insurers, we compare two extreme types of demand: "rational" and "myopic." A rational consumer maximizes the expected utility over the choice horizon based on the current information available and using the best prediction of risk over the contract horizon. Myopic consumers do not update their willingness to pay for longer contracts and treat them as if they were annual contracts from which they could switch at the end of the year. Myopia may capture the inability of some consumers to incorporate the change in incentives induced by longer contract length into their choices. Because equilibria with myopic demand are akin to holding demand fixed at its baseline (one-year) level, they provide an estimate of the net supply-side response to a regime change.

Equilibrium To compare one- and two-year contract regimes, we use the above supply and demand estimates to calculate the Riley equilibrium in markets with different contract lengths. Our supply estimates are pinned down by claims data, but our demand estimates have one degree of freedom—risk aversion. We therefore present results for a range of risk aversion parameters λ that span a range of initial levels of coverage with one-year contracts. For each such initial level with one-year contracts, we compute the counterfactual equilibrium with two-year contracts, holding risk aversion constant. This yields a set of equilibrium pairs, allowing us to compare one- and two-year contracts for each specific level of risk aversion.

We study different levels of risk aversion, which result in different levels of coverage under the baseline, one-period contract regime. Note that with extremely low or extremely high risk aversion, there would be no gains from increasing contract length, as extreme risk aversion would lead either to full market unraveling in both regimens (for sufficiently low risk aversion) or to no unraveling and 100% high-coverage in both regimes (for sufficiently high risk aversion). We therefore focus on the range of risk aversion levels spanning the intermediate (non-trivial) range of partial market unraveling due to adverse selection. The range of coverage levels we study covers the support of the distribution of actual coverage in the ACA exchanges across the United States (see Figure A4).

Figures 7 and 8 demonstrate the calculation of the market equilibrium in the first period of a one-year contract regime with $\lambda = 2\times10^{-4}$ (46.05% of the population buys the 90% plan; the (fair) annual premiums are \$4,458 for this plan and \$881 for the 60%-coverage plan, so $\Delta AC = \$3,577$. The willingness to pay for high coverage of the marginal type is \$3,581; the ultra-marginal type is willing to pay \$3,547, and therefore prefers lower coverage.) Because contracts are community rated, this problem involves no dynamic aspects. Therefore, in the one-year contract regime, we calculate separately the equilibrium in each period (the only difference being the information individuals have when making their choices).

Welfare To compare the welfare gains from extending the contract horizon, we calculate the ex-ante expected utility from the different contract-length regimes, given the estimated risk distribution and the assumed risk preferences. That is, denote $f^*_{s,h}(\theta_t)$ to be the equilibrium plan choice of type θ under regime $h = 1, 2$ (denoting one- and two-year contract regimes). We calculate the expected ex-ante utility $E_\theta[\sum_{s=1}^{2} u\Big(f^*_{s,h}(\theta)(C_s(\theta))\Big)]$. In the one-period contract regime ($h = 1$), this ex-ante utility equals:

$$\sum_{\theta_0}\Big(u\Big(f^*_{1,1}(\theta_0)(C_1(\theta_0))\Big) + \sum_{\theta_1} u\Big(f^*_{2,1}(\theta_1)(C_2(\theta_1))\Big)P[\theta_1|\theta_0]\Big)P[\theta_0]$$

. In the two-period contract regime, it equals:

$$\sum_{\theta_0}\Big(u\Big(f^*_{1,2}(\theta_0)(C_1(\theta_0))\Big) + \sum_{\theta_1} u\Big(f^*_{2,2}(\theta_0)(C_2(\theta_1))\Big)P[\theta_1|\theta_0]\Big)P[\theta_0]$$

. That is, the cost distribution is the same in both regimes, but equilibrium plan choice and premiums differ.[22] In the case of mean-variance utility, which is quasi-linear, the ex-ante willingness to pay for a regime change from one- to two-period contracts is simply the difference between these two values.

When evaluating the case of a myopic demand response to an increased contract length, we still use the same risk process and preferences as in the rational case in our welfare criterion. This approach is consistent with a myopia representing a limitation of rationality of choices, rather than an inherent difference in preference regarding future risks.

Remarks

In principle, it would have been more direct to estimate demand by exploiting quasi-random variation in contract terms and choice. However, the relevant variation for this paper does not exist in the United States: all contracts have the same horizon of one year. Instead,

[22]In particular, the choice of both $f^*_{1,2}$ and $f^*_{2,2}$ is conditioned only on θ_0, making the exposure to risk at period 2 based on choices made with period 0 information.

calibrating a model of insurance choice has the benefit of requiring only data on risk, as estimated in the previous section. The downside is that one must make a parametric assumption about the utility function. See Aron-Dine et al. (2015) for a case in which agents' start of insurance coverage is staggered across the calendar year, introducing a *de facto* small variation in the contract length.

One could use our same methods to study contracts longer than two years and even lifetime contracts (see Cochrane, 1995; Handel et al., 2017). We choose to focus instead on an incremental difference in contract length for two reasons. First, very long contracts present other complicating factors, such as aggregate uncertainty and commitment, that limit the feasibility of implementing them in practice. Second, from an empirical perspective, there is a greater confidence in extrapolating from the current annual contracts to an equilibrium with only slightly longer ones.

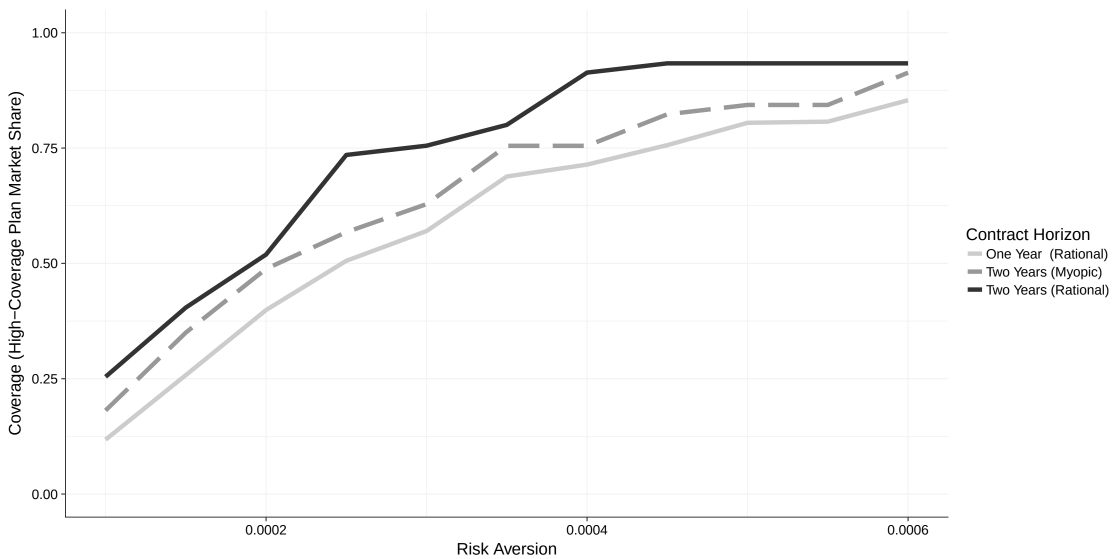
Results

Our counterfactual analysis suggests that a reform implementing two-year contracts (e.g., by lowering the frequency of open enrollment periods) would increase coverage by 12–19 percentage points from its initial level and yield average annual welfare gains of $600–$900 per person (i.e., between 20% and 30% of the insured risk), from an ex-ante perspective. A third of these effects are driven by insurers' response and the rest by changes in consumer expectations.

Counterfactual coverage with two-year contracts is higher for the entire range of risk aversion we considered, which corresponds to coverage level (the market share of the 90% coinsurance plan) of between 10%–85%. Figure 9 shows the average equilibrium coverage over all periods in three cases: (i) one-year contracts with rational demand; (ii) two-year contracts with rational demand; and (iii) two-year contracts with myopic demand, for a range of values of the risk aversion parameter λ. In all cases, a reduction in adverse selection results in better risk pooling. Intuitively, extending the contract length allows for better pooling of risk within—not just across—individuals. We obtain similar results with the CARA utility index (Figure A3). Annual per-person welfare gains, measured in terms of the willingness to pay (WTP) for moving from a one- to two-year contract regime range between $600–$900 (Figure 10 and Table 5).

The differences in coverage and premiums between the one- and two- contract regimes are summarized in Table 3. For different levels of risk aversion, the average market share of the 90% coverage (High) benefit plan under one- and two-year contracts unambiguously increases. The increase is substantial—more than 10 percentage point on average throughout this range. The reduction of adverse selection with increased contract length and the corresponding increase in coverage also results in lower premiums—for both high- *and* low-coverage plans. The last column of Table 3 shows a summary measure of this reduction: the enrollment-weighted decrease in premiums, defined as the decrease in premiums weighted by the enrollment under one-period contracts. Detailed changes in premiums are summarized in Table 4. As expected, the decline in premiums is most substantial when risk aversion is lower, making the initial market unraveling high and the 90% coverage prohibitively expensive.

Figure 9: Coverage Increase with Longer Contract Length

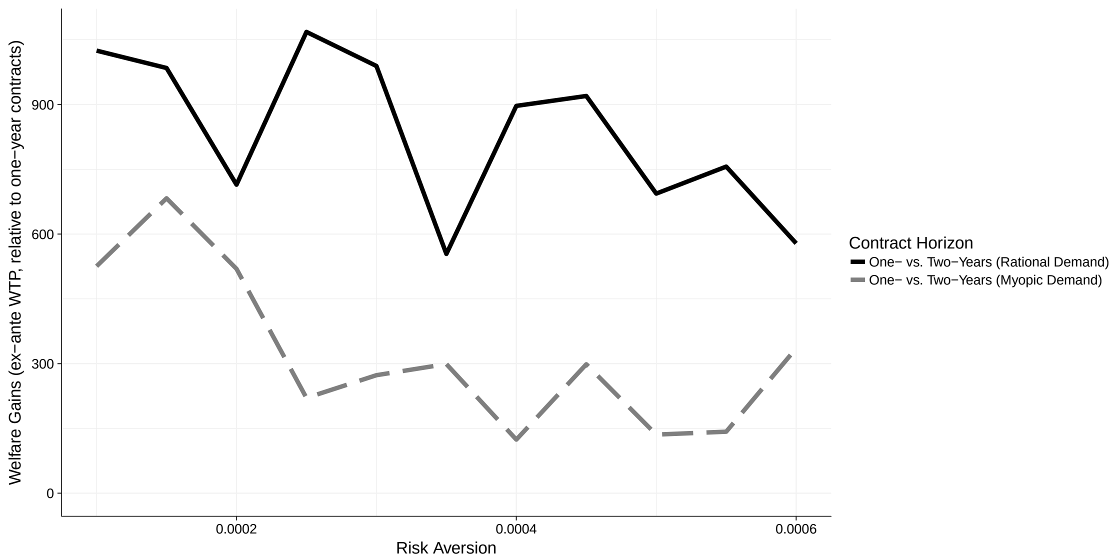


Notes: Coverage (the market share of the 90% coverage benefit plan) with different levels of risk aversion (the x-axis) and different regimes: one-year contracts (gray), two-year contracts (black), and two-year contracts with myopic demand (dashed). In the case of myopic demand, the increase in coverage is solely due to supply-side response to the increase in contract length. This analysis uses Mean Variance utility index. We obtain similar results when using the CARA utility (see Appendix).

In line with our discussion in Section 1, the overall increase in coverage is a combination of demand and supply responses. The dashed lines in Figures 9 and 10 show the coverage and welfare gains with two-year contracts when demand is myopic, i.e. consumers do not update their risk expectations after the reform. These dashed lines thus capture the net supply-side impact of increased contract length. Table 5 summarizes the relative importance of the supply channel. Within the studied range of risk aversion, a significant share of both coverage and welfare gains (between 20%–70%) is due to increased contract length that occur through supply-side effects alone. These findings suggest that even if demand fails to internalize the consequences of an increased contract horizon there would still be significant gains from it as insurers would face a less adversely selected pool of individuals, due to bounded rationality.

Our prediction is therefore that longer contract length will be associated with an overall improved equilibrium coverage and with higher ex-ante welfare. However, this is not a

Figure 10: Willingness to Pay for Longer Contract Length



Notes: The figure shows the average ex-ante willingness to pay for changing the regime from one- to two-year contracts for different levels of risk aversion, with two specifications of demand: rational (Solid) and myopic (Dashed). Myopic consumers choose two-year contracts as if they were one-year contracts, so the welfare gains with myopic demand are solely due to the supply-side response to the increase in contract length.

Table 3: Horizon Effects on Coverage and Premiums

	Coverage by Contract Length		One- versus Two-Year Contracts		
			Coverage Difference		Premium Difference
Risk Aversion	One-Year	Two-Year	Percentage Point Difference	Percent Difference	Enrollment Weighted Difference
1.00 ($\times 10^{-4}$)	%11.8	%25.4	%13.6	%115.6	$-1018
1.50	25.8	40.4	14.6	56.8	-719
2.00	39.9	51.9	12.0	30.2	-430
2.50	50.6	73.5	23.0	45.4	-690
3.00	57.0	75.5	18.5	32.4	-536
3.50	68.8	80.0	11.2	16.3	-289
4.00	71.4	91.4	19.9	27.9	-494
4.50	75.6	93.4	17.7	23.4	-452
5.00	80.5	93.4	12.9	16.0	-314
5.50	80.7	93.4	12.6	15.7	-309
6.00	85.4	93.4	8.0	9.4	-203

Notes: Counterfactual equilibrium coverage (the market share of the 90% coverage plan) with one- and two-year contracts, and the difference in coverage between these scenarios in percentage point and percentage terms, for different levels of risk aversion (λ). Enrollment-weighted difference in premium is the average change in premiums, weighted by the two-year equilibrium market shares (for detailed changes in premiums, see Table 4).

Table 4: Horizon Effects on Premiums

	One-Year Contracts			Two-Year Contracts			Difference	
	Coverage	Premiums		Coverage	Premiums		Premiums	
Risk Aversion ($\times 10^{-4}$)	90-Plan (%)	60-Plan (\$)	90-Plan (\$)	90-Plan (%)	60-Plan (\$)	90-Plan (\$)	60-Plan (\$)	90-Plan (\$)
1.00	11.8	1460	9070	25.4	1297	5542	-163	-3527
1.50	25.8	1196	6306	40.4	1066	4719	-130	-1587
2.00	39.9	990	4989	51.9	929	4217	-61	-772
2.50	50.6	906	4366	73.5	758	3481	-148	-885
3.00	57.0	835	4097	75.5	794	3401	-41	-696
3.50	68.8	781	3647	80.0	744	3295	-37	-352
4.00	71.4	747	3568	91.4	676	3034	-71	-534
4.50	75.6	681	3474	93.4	686	2990	5	-485
5.00	80.5	645	3329	93.4	686	2990	41	-339
5.50	80.7	636	3324	93.4	686	2990	50	-334
6.00	85.4	595	3213	93.4	686	2990	91	-223

Notes: Counterfactual equilibrium coverage (market share of the 90% coverage plan) and premiums, for different levels of risk aversion (λ). Premiums are actuarially fair in competition, and their reduction reflects the increase in coverage and the associated changes to risk pools due to reduced adverse selection. For comparability with the two-year contracts, one-year coverage and premium are averaged over time. All figures are annual.

Table 5: Overall and Supply-Side Horizon Effects on Coverage and Welfare

	Coverage Difference		Welfare Difference		Supply-Side Contribution to Overall Differences	
Risk Aversion ($\times 10^{-4}$)	Supply & Demand (p.p. increase)	Supply Only (p.p. increase)	Supply & Demand (\$ WTP)	Supply Only (\$ WTP)	Coverage %	Welfare %
1.00	13.6	6.3	1025	525	46.4	51.3
1.50	14.6	9.2	984	683	63.0	69.4
2.00	12.0	8.9	714	519	74.2	72.7
2.50	23.0	6.2	1068	221	27.1	20.7
3.00	18.5	5.8	989	273	31.6	27.6
3.50	11.2	6.7	554	299	59.4	54.0
4.00	19.9	4.1	897	124	20.4	13.8
4.50	17.7	6.7	920	298	37.7	32.5
5.00	12.9	3.9	694	136	30.0	19.6
5.50	12.6	3.6	756	142	28.7	18.8
6.00	8.0	6.0	578	336	74.7	58.0

Notes: The table shows coverage differences (the increase in High-plan market share) and welfare differences (the ex-ante willingness to pay (WTP) for moving to two-year contracts) between one- and two-year contracts, for two alternative demand specifications: (i) fully rational consumers ("Supply & Demand"), and (ii) fully myopic consumers ("Supply Only"). The rightmost columns show the ratio of the differences (Supply Only)/(Supply & Demand), thus quantifying the relative contribution of supply-side response to coverage and welfare gains due to increased contract length.

Pareto improvement, as certain types end up committed to contracts they would rather leave. Increasing the horizon is not favorable to everyone: the healthiest types—those that rationally still opt for lower coverage in equilibrium even with a longer horizon—suffer welfare loss, as they are then exposed to greater risk. Their losses counteract some of the gains of the marginal types who obtain High coverage only with a longer horizon, resulting in overall welfare gains. Overall, our results should be interpreted with this trade-off in mind.[23]

4 Conclusion

This paper shows how extending the horizon of health insurance contracts impacts adverse selection in these markets. The main contribution of this paper is to show the implications of the dynamics of health risk *over time*, as opposed to simply its cross-sectional distribution. We show that increasing the contract length is another policy instrument that can be used to reduce selection. Conceptually, we argue that private information is endogenous to contract length because individual risk is harder to predict at longer horizons. This decrease in risk predictability is strongly borne out by the data. Simulating a model of ACA-like exchanges, we find the effect of extending contracts from one to only two years would be to expand coverage substantially, and improve ex-ante average welfare by \$600–\$900 (20%–30% of the covered risk). However, not everyone gains from such change: The gains for some are partly offset by the increased exposure of others to greater risk due to the commitment to longer contracts. While our risk estimates rely on a large sample of non-elderly Americans, it is plausible that different samples would yield a negative change in welfare. However, our analysis provides tools for study and it can be replicated to obtain the magnitude of the welfare gains in other settings.

References

Abaluck, Jason and Jonathan Gruber (2011), "Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program." *The American economic review*, 101, 1180–1210.

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark Cullen (2015), "Moral hazard in health insurance: do dynamic incentives matter?" *Review of Economics and Statistics*, 97, 725–741.

Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston (2014), "How does risk selection respond to risk adjustment? New evidence from the Medicare advantage program." *American Economic Review*, 104, 3335–64.

[23]In addition, there may be other implicit costs related to longer contracts.

Bundorf, M Kate, Jonathan Levin, Neale Mahoney, et al. (2012), "Pricing and welfare in health plan choice." *American Economic Review*, 102, 3214–48.

Cabral, Marika (2016), "Claim timing and ex post adverse selection." *The Review of Economic Studies*.

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

Culter, David M and Sarah J Reber (1998), "Paying for health insurance: The trade-off between competition and adverse selection." *Quarterly Journal of Economics*, 433–466.

Cutler, David M and Richard J Zeckhauser (1998), "Adverse selection in health insurance." In *Forum for Health Economics & Policy*, volume 1.

Decarolis, Francesco and Andrea Guglielmo (2016), "Insurers' response to selection risk: Evidence from Medicare enrollment reforms." *NBER Working Paper 22876*.

Diamond, Peter (1992), "Organizing the health insurance market." *Econometrica: Journal of the Econometric Society*, 1233–1254.

Einav, L and A Finkelstein (2010), "Selection in insurance markets: Theory and empirics in pictures." *The Journal of Economic Perspectives*, 25, 115–138.

Einav, Liran, Amy Finkelstein, Mark R Cullen, et al. (2010), "Estimating welfare in insurance markets using variation in prices." *The Quarterly Journal of Economics*, 125, 877–921.

Ellis, Randall P et al. (2000), "Risk adjustment in competitive health plan markets." *Handbook of Health Economics*, 1, 755–845.

Ericson, Keith M Marzilli and Amanda Starc (2016), "How product standardization affects choice: Evidence from the Massachusetts health insurance exchange." *Journal of Health Economics*, 50, 71–85.

Geruso, Michael (forthcoming), "Demand heterogeneity in insurance markets: Implications for equity and efficiency." *Quantitative Economics*.

Geruso, Michael, Timothy J Layton, and Daniel Prinz (2016), "Screening in contract design: Evidence from the ACA health insurance exchanges." *Working Paper*.

Glazer, Jacob and Thomas G McGuire (2000), "Optimal risk adjustment in markets with adverse selection: an application to managed care." *The American Economic Review*, 90, 1055–1071.

Hackmann, Martin B, Jonathan T Kolstad, and Amanda E Kowalski (2012), "Health reform, health insurance, and selection: Estimating selection into health insurance using the Massachusetts health reform." *The American Economic Review*, 102, 498–501.

Hackmann, Martin B, Jonathan T Kolstad, and Amanda E Kowalski (2015), "Adverse selection and an individual mandate: When theory meets practice." *The American Economic Review*, 105, 1030.

Handel, Ben, Igal Hendel, and MD Whinston (2017), "The welfare impact of long-term contracts." Working Paper.

Handel, Ben, Igal Hendel, and Michael D Whinston (2015), "Equilibria in health exchanges: Adverse selection versus reclassification risk." *Econometrica*, 83, 1261–1313.

Handel, Benjamin R and Jonathan T Kolstad (2015), "Health insurance for "humans": information frictions, plan choice, and consumer welfare." *The American Economic Review*, 105, 2449–2500.

Heiss, Florian, Adam Leive, Daniel McFadden, and Joachim Winter (2013), "Plan selection in Medicare part D: Evidence from administrative data." *Journal of Health Economics*, 32, 1325–1344.

Hendren, Nathaniel (2013), "Private information and insurance rejections." *Econometrica*, 81, 1713–1762.

Kolstad, Jonathan T and Amanda E Kowalski (2016), "Mandate-based health reform and the labor market: Evidence from the Massachusetts reform." *Journal of Health Economics*, 47, 81–106.

Luz, Vitor Farinha (2015), "Dynamic competitive insurance." Discussion paper. University of British Columbia.

McGuire, Thomas G, Jacob Glazer, Joseph P Newhouse, Sharon-Lise Normand, Julie Shi, Anna D Sinaiko, and Samuel H Zuvekas (2013), "Integrating risk adjustment and enrollee premiums in health plan payment." *Journal of Health Economics*, 32, 1263–1277.

Montanera, Daniel (2017), "Selection-proof health insurance exchanges." *Mimeo*.

Rice, Thomas, Marcia L Graham, and Peter D Fox (1997), "The impact of policy standardization on the medigap market." *Inquiry*, 106–116.

A Appendix

Table A1: Risk Predictors Used in X_i^t in Different Specifications

Predictors	*Cost*	*Demographics*	*Utilization*
total annual cost	v	v	v
cost by category (outpatient vs. inpatient)		v	v
age, sex		v	v
event counts, duration, and all other measures (see Descriptive Statistics Table for details)			v

Table A2: Choice of Coverage Level in the ACA Exchanges

		Percent Enrolled			
Metalic Level	Avg. Coverage	2014	2015	2016	2017
Bronze or Silver	60%–70%	86%	90%	92%	95%
Gold or Platinum	80%–90%	14%	10%	8%	5%
Enrollment (Millions)		6.3	10.2	11.1	12.2

Notes: National average enrollment levels in the marketplace exchanges, by actuarial value. Actuarial value is the average share of spending paid for by the plan. Source: CMS, KFF. Excluding 1% catastrophic and unknown levels.

Figure A1: Average Annual Cost by Age Group and Gender

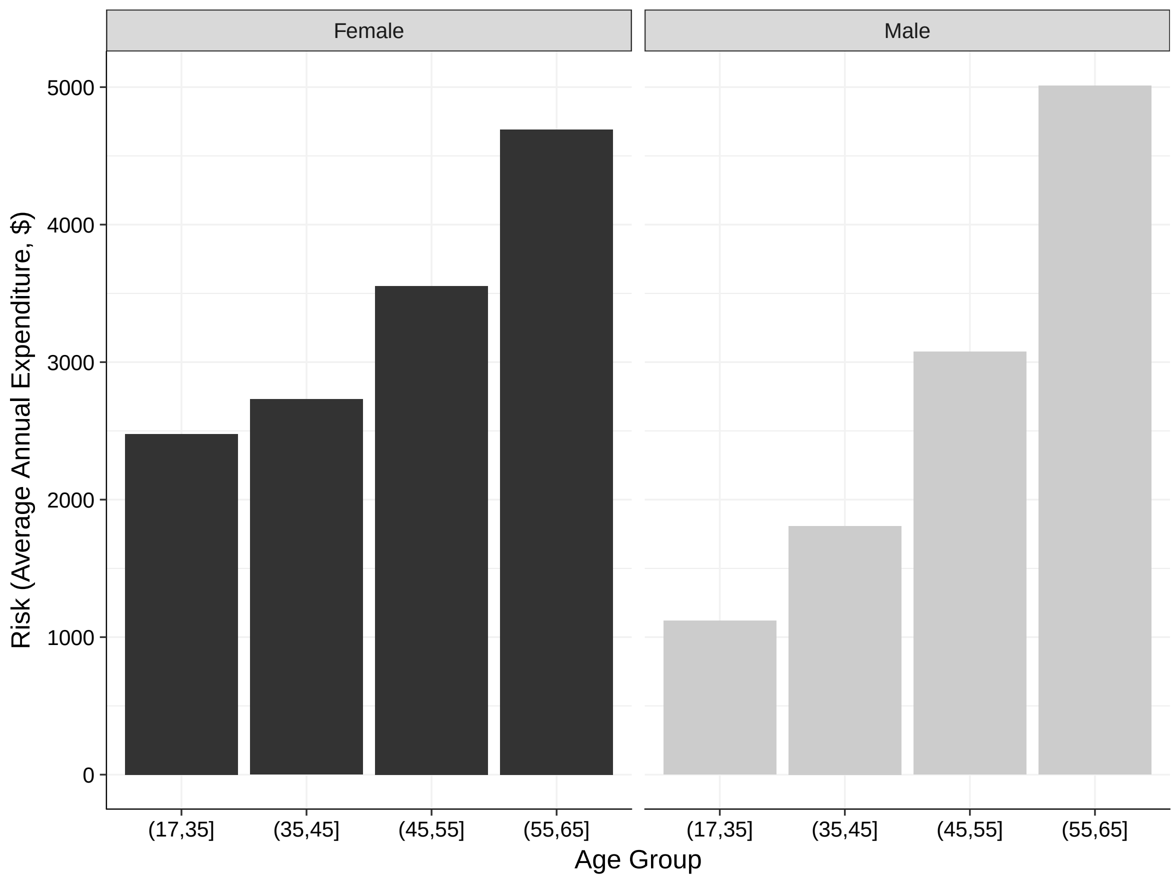
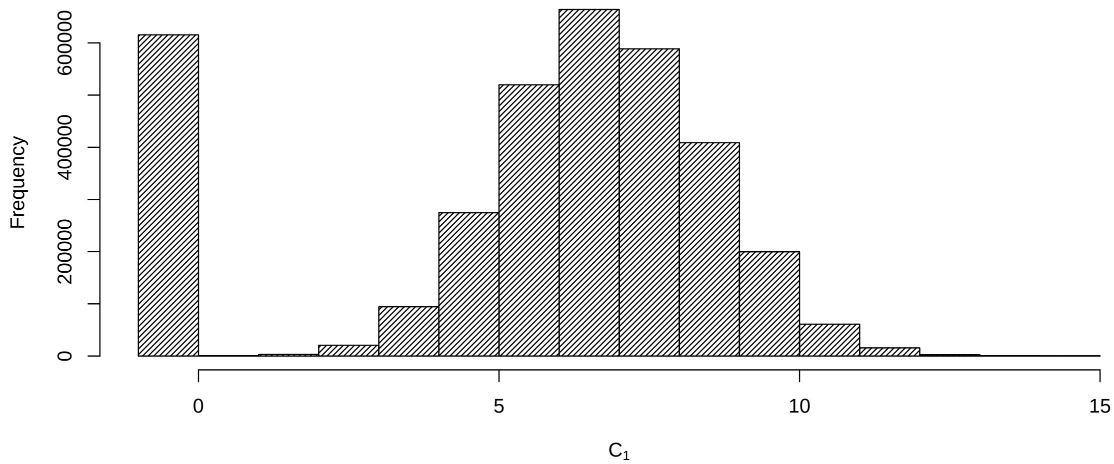


Table A3: Descriptive Statistics: Cost and Use

	mean	sd
A. MarketScan		
Total Medical Costs	2,350	10,375
Total Inpatient Costs	757	7,474
Gender (Male)	0.465	0.498
Age	42.4	12.9
Inpatient Events	0.066	0.332
Outpatient Events	16.9	30.9
Observations	10,338,484	
B. Medicare		
Parts A and B Total Costs	4,667	15,021
Part A and B Annuatized Total Costs	6,003	23,968
Part D Total Prescription Costs	808	2,501
Gender (Male)	0.472	0.499
Age	66.29	1.710
Ambulatory Surgery Center Events	0.105	0.605
Part B Drug Events	1.395	5.516
Evaluation and Management Events	2.147	10.43
Anesthesia Events	0.181	0.714
Dialysis Events	0.0536	1.043
Other Procedures Events	2.387	9.841
Imaging Events	1.996	4.961
Tests Events	5.778	13.40
Durable Medical Equipment Events	1.050	4.485
Other Part B Carrier Events	0.690	6.705
Part B Physician Events	2.953	5.403
Part D Events	12.23	25.02
Acute Inpatient Covered Days	0.539	3.519
Other Inpatient Covered Days	0.115	2.065
Skilled Nursing Facility Covered Days	0.362	4.876
Hospice Covered Days	0.272	6.801
Hospital Outpatient Visits	2.658	11.20
Hospital Outpatient Emergency Room Visits	0.134	0.633
Inpatient Emergency Room Visits	0.0663	0.380
Home Health Visits	0.854	14.07
Parts AB Total Costs	4667.7	15021.1
Part AB Annuatized Total Costs	6003.3	23968.0
Part D Total Prescription Costs	808.1	2501.9
Log Annuatized Cost	5.294	3.792
Demeaned Log Annuatized Cost	-0.587	3.779
Part D Fill Count	16.02	30.16
Acute Inpatient Stays	0.118	0.527
Other Inpatient Stays	0.00798	0.115
Skilled Nursing Facility Stays	0.0150	0.179
Hospice Stays	0.00499	0.0805
Hospital Readmissions	0.0195	0.222
Observations	1,289,776	

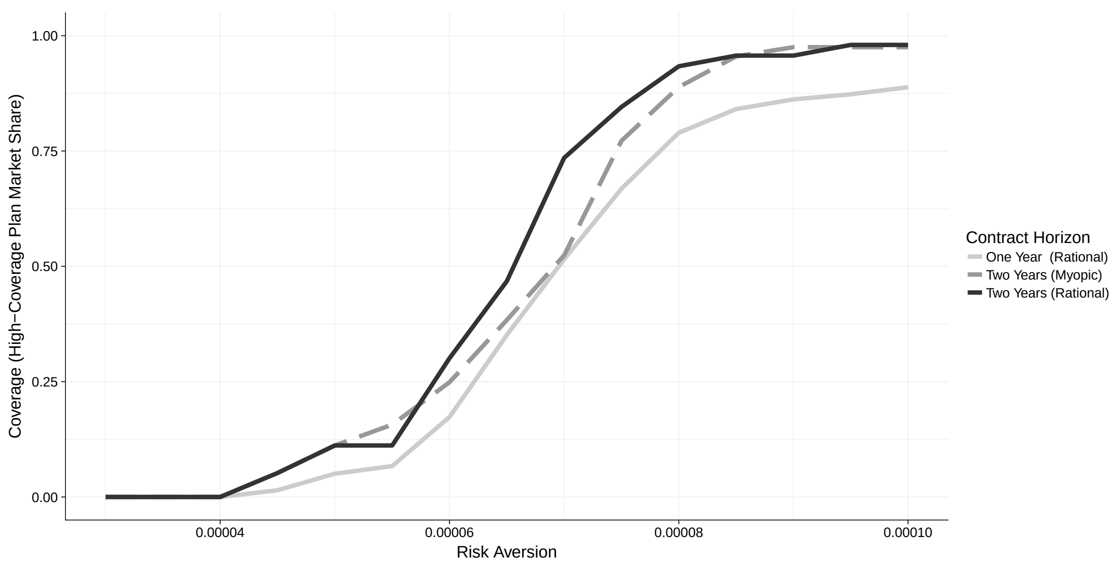
Notes: Risk measures and predictors for both samples. See Section 2.1 for details.

Figure A2: Empirical Cost Distribution



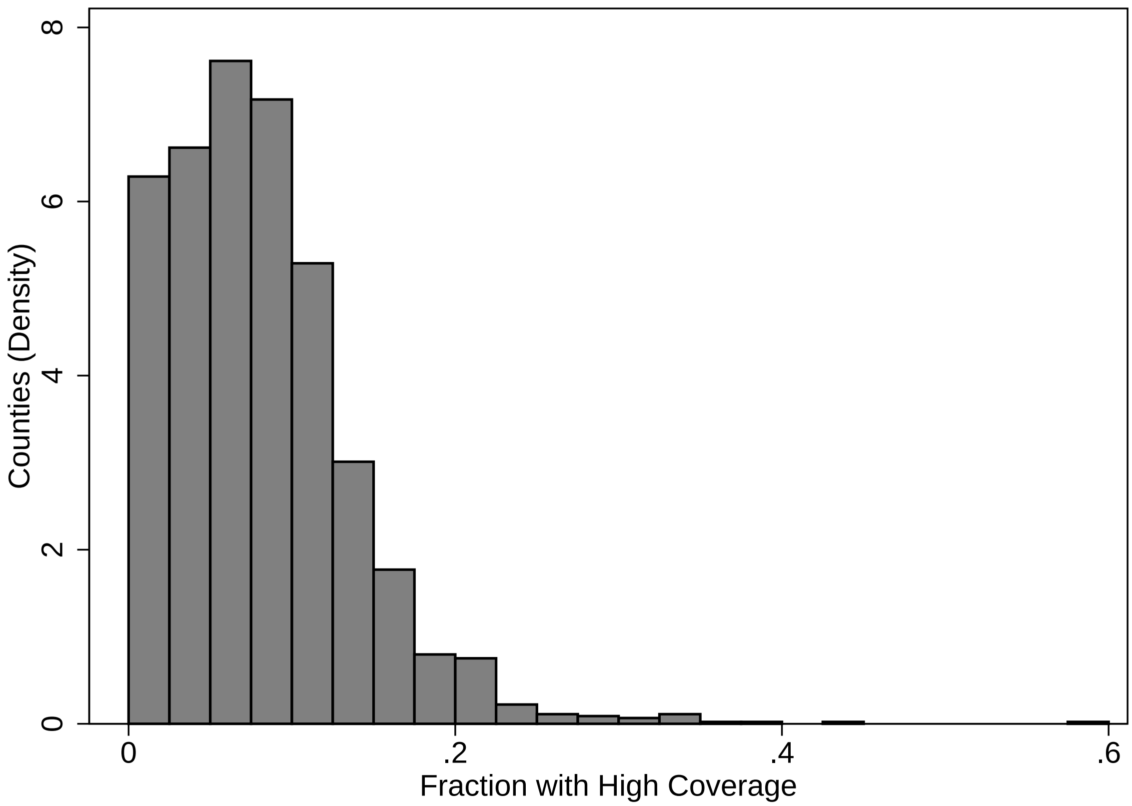
Notes: Histogram of the empirical distribution of log annual total healthcare costs (denoted c_1 in (4a)). The mass at zero had no expenditure during the year of coverage. Data source: MarketScan Research Databases.

Figure A3: Coverage Increase with Longer Contract Length (CARA Utility)



Notes: Coverage (the market share of the 90% coverage benefit plan) with different levels of risk aversion (the x-axis) and different regimes: one-year contracts (gray), two-year contracts (black), and two year contracts with myopic demand (dashed). This figure shows the analysis with CARA utility. In the case of myopic demand, the increase in coverage is solely due to supply-side response to the increase in contract length.

Figure A4: Distribution of High Coverage



Notes: This histogram shows the distribution of county-level coverage for 1,807 counties in the 37 states with federal exchanges (counties are the smallest unit used for pricing in the ACA marketplaces). High coverage is defined as coverage above the Silver level (i.e., actuarial value of 80% or more). Source: Data on 2016 enrollment figures for qualifying health plan selections by metal level as of February 22, 2015, from data.cms.gov. Retrieved January 2017.