Claim Timing and Ex Post Adverse Selection*

Marika Cabral†

February 26, 2016

Abstract

Many health care treatments are not urgent and may be delayed if patients so choose. Because insurance coverage is typically determined by the treatment date, individuals may have incentives to strategically delay treatments to minimize out-of-pocket costs. The strategic delay of treatment—a particular form of moral hazard— can be an important source of subsequent adverse selection, in which ex ante identical individuals select insurance coverage based on their differing accumulation of previously delayed treatments. This paper investigates these forces empirically in the context of the missing market for dental insurance. Using rich claim-level data, my analysis reveals that approximately 40% of individuals strategically delay dental treatments when incentivized to do so, and this flexibility in delaying treatment can explain why the market for dental insurance has largely unraveled. More generally, the counterfactual analysis suggests features such as open enrollment periods and contracting on pre-existing conditions may be helpful tools in overcoming adverse selection in insurance contexts where the timing of uncertainty is not contractible.

*For their useful comments and feedback, I thank Jérôme Adda, three anonymous referees, Michael Boskin, Tim Bresnahan, Liran Einav, Gopi Shah Goda, Matthew Harding, Caroline Hoxby, Amanda Kowalski, Jonathan Levin, Neale Mahoney, Isabelle Sin, William Gui Woolston, and Ali Yurukoglu. Helpful feedback was provided by seminar participants at American Society of Health Economists Conference, Brown, Columbia, RAND, STATA Texas Empirical Micro Conference, Stanford, UBC, UC Berkeley Haas School of Business, UC Irvine, University of Pittsburgh, University of Rochester Simon School of Business, USC, UT Austin, Wharton, and Yale School of Management. I am also very grateful to Mark Cullen for his helpful advice and for his support in obtaining data for this project. In addition, I would like to thank Brenda Barlek and Martin Slade for their assistance in acquiring data and explaining the environment. The Alcoa data were provided as part of an ongoing service and research agreement between Alcoa, Inc. and academic institutions to perform jointly agreed-upon ongoing and ad hoc research projects on workers health, injury, disability, and health care, and research using this data is supported by a grant from the National Institute on Aging (Disease, Disability and Death in an Aging Workforce, NIH/NIA, 1 R01 AG026291).

†UT Austin and NBER, marika.cabral@austin.utexas.edu.

In most non-emergency medicine, there is a time lapse between the recognition that a particular treatment is necessary and the actual treatment itself. This time lapse can often be controlled by the patient, so that treatment is delayed if the patient wishes it to be. Because insurance coverage is typically determined by insurance status on the treatment date, such control over the timing of treatment can generate substantial problems for insurance. If people can delay a treatment (once they know it will occur) just long enough to buy more insurance in anticipation of it, severe adverse selection may result.[1] In the extreme case where all treatments can be easily delayed, individuals can opt out of insurance, accumulate untreated problems, buy generous coverage and get treatment for the entire stock of problems they have accumulated, and then opt out of insurance again; in such cases, these forces can lead insurance markets to completely unravel. While traditional adverse selection is driven by heterogeneity in ex ante risk types, this "ex post adverse selection" may arise even in an ex ante homogeneous population, as individuals become differentiated at different points in time by the amount of treatment they have postponed (the amount of "timing moral hazard" they have engaged in).[2] Ex post adverse selection has different policy implications than traditional moral hazard and traditional adverse selection. While the insurance coverage period is of little consequence in the context of traditional adverse selection (or moral hazard), the best policy response to timing moral hazard and ex post adverse selection may include waiting periods, pricing pre-existing conditions, open enrollment periods, and eligibility restrictions based on prior insurance status.

Health care treatments span a spectrum of urgency. While treating a heart attack is extremely costly to delay, most orthopedic surgery and many dental treatments can be delayed easily for quite some time. Although most health care treatments are covered by medical insurance, historically certain categories of health care have been excluded from coverage by medical insurance policies, most notably dental care. Interestingly, the market for standard medical insurance and the market for dental insurance look quite different. Unlike the market for medical insurance, the market for stand-alone dental insurance in the United States has almost completely unraveled. Only about 40% of individuals have any dental coverage,[3] and those with coverage actually have little insurance against dental risk. The typical dental "insurance" policy provides very incomplete coverage owing to a low annual maximum benefit, which is on average $1,100. Above this maximum benefit, "insured" individuals must pay the full cost of services.[4] Although dental care can involve considerable uncertainty and financial cost, available policies tend to offer no coverage for large, urgent dental expenditures. Perhaps because the available policies provide so little insurance, almost no one takes them up except through an employer. That is, if it were not for the tax subsidy that allows people to pay for dental premiums with pre-tax dollars when they enroll in an employer-sponsored plan

[1] It may be easiest to explain this concept through an example. Suppose an individual suffers from an illness that necessitates treatment but this treatment is not particularly urgent. Since health insurers typically use the date of treatment in contracting (rather than the date the illness began), the individual may have incentives to delay the treatment until he has signed up for better insurance coverage. It is often the case that health insurers cannot contract on the date that uncertainty is resolved (i.e., they cannot perfectly underwrite pre-existing conditions). This contracting limitation could arise because of regulations (e.g., the ban on underwriting pre-existing conditions in the Affordable Care Act), or it could arise from technological limitations on identifying pre-existing conditions (e.g, the fact that much of diagnosis in health care settings relies on self-reported pain or self-reported limitations in daily activities).

[2] The strategic delay of treatment can be thought of as a form of moral hazard. In much of the literature, "moral hazard" is often used to characterize any hidden action in a contractual relationship that creates unobservable contingencies, about which information is needed in order to design the first-best efficient contract. In the context of insurance, this broad definition of moral hazard refers to the adverse effects, from the insurer's point of view, that result from incomplete contracting on individual's actions that are potentially influenced by insurance.

[3] Calculation based on self-reported insurance status in the Medical Expenditure Panel Survey, 2007.

[4] Typically we think of insurance as a contract that caps the losses of an individual in the case of a bad financial shock; available dental "insurance" contracts, however, do just the opposite: they cap the per-individual losses of the insurer.

(as opposed to post-tax dollars when they pay for care out of pocket), dental "insurance" might not exist at all.

In this paper, I use the missing market for dental insurance as a setting to investigate market unraveling generated by timing moral hazard and ex post adverse selection. This is a natural first context to look for evidence of strategic delay and ex post adverse selection for several reasons. First, it is a relatively simple context in which many treatments are delayable for weeks, months or longer.[5] At the same time, there are extremely urgent and large dental expenditures for which insurance would be very valuable. Second, dental care is insured in isolation separate from other health care treatments. Because ex ante it is likely that dental procedures are particularly delayable, we might expect to find large perverse effects arising from the delayability of treatment. Third, there is very little insurance for dental risk even though people seem to demonstrate a high degree of risk aversion in this domain. This puzzle is explored in this paper. Ultimately, the results in this paper suggest that observed market unraveling can be explained by treatment delay and the resulting adverse selection. Fourth, despite the significant market unraveling present in this context, there exists data to study this risk empirically owing to the presence of subsidized employer-provided fringe benefits. From the perspective of a researcher, it is fortunate that this tax subsidy exists as companies that offer these incomplete subsidized fringe benefits collect enrollment and dental spending data in order to review claims.[6] There has been very little prior empirical research explaining the non-existence of insurance markets, perhaps because of the difficulty in obtaining data related to missing markets. One of the contributions of this paper is that I can analyze how asymmetric information may play a role in explaining a largely missing insurance market.

While the lack of insurance in the $110 billion dental care market may be of interest in its own right, the findings in this paper may shed light on how this source of asymmetric information can affect insurance markets more generally.[7] The source of asymmetric information studied in this paper is not limited to dental or other health insurance contexts. This dynamic connection between moral hazard and adverse selection may play a role in any insurance setting where (1) individuals can control the timing of the insured financial cost after the realization of an "event" that necessitates this cost and (2) the date of the financial cost, rather than the date of the "event," is used to determine coverage. In other words, these incentives are relevant in contexts where the timing of claims is manipulable and the timing of risky events is not contractible. These two conditions are often met in the cases of medical and dental insurance.[8] In these contexts, it is often challenging or impossible for insurers to contract on the date that an individual begins to feel ill or begins to experience pain.[9] Insurers also demonstrate awareness that strategically delaying treatment is often possible. For example, health insurers selling products on the individual market have a long history of using instruments such as waiting periods and pricing pre-existing conditions to combat delayed health treatments.[10] The counterfactual policy analysis in this paper within the dental context

[5] A number of articles in dental journals describe the flexibility in timing dental care (see, e.g., Guay (2006), Jeffcoat (2004)).

[6] Since employees are often asked to pay a premium for these policies, it is also possible to estimate a willingness to pay for these incomplete benefits and extrapolate from this information to quantify the value of the missing market.

[7] Total US dental spending in 2012 was $111 billion according to Wall, Nasseh and Vujicic (2014).

[8] In the case of medical and dental insurance, individuals can often delay costly treatment after the realization of an event (a health problem), and coverage is typically determined by insurance enrollment on the treatment date. In contrast, strategic timing plays little role in auto insurance where the date of a typical event, an auto accident, is generally contractible so insurers need not rely on the date of the associated auto repairs to determine coverage.

[9] Sometimes this date is not observable, and in other cases, regulations limit the ability of insurers to price pre-existing conditions.

[10] In light of the recent restrictions the Affordable Care Act has placed on the use of waiting periods and pricing pre-existing conditions, strategically delayed treatments and subsequent adverse selection may be an increasingly important policy issue in broader health insurance contexts. While the analysis in this paper is limited to the dental context, I discuss the potential importance of timing moral hazard and ex post adverse selection in broader insurance settings further in Section 7.

suggests that less frequent open enrollment periods and contracting on pre-existing conditions may be helpful tools in contexts where the timing of uncertainty is not contractible. At the same time, the results indicate that such features are likely to be inconsequential in contexts like automobile insurance, in which the timing of uncertain events (e.g., auto accidents) is typically very contractible.

The empirical analysis in this paper employs a rich and complete claim-level data from one large firm, Alcoa Inc. The data are particularly useful for at least four reasons. First, the firm offers two vertically differentiated dental plans, where the primary difference between the plans is the size of the maximum benefit, \$1,000 or \$2,000. This feature allows me to look for evidence of selection across these options. Second, because the firm's less generous dental insurance plan is free for employees, no one opts out of dental benefits. Therefore, the data contain all dental claims for all employees for all years they are with the company. Third, the data include claims for treatments that were not reimbursed because the costs were above the annual maximum benefit. Thus, the data are not censored. Fourth, during the data period (2004-2007), the firm changed the coverage period twice (going from one-year coverage to two-year coverage and then back to one-year coverage). This feature is very useful in the analysis, as will become clear below.

Analyzing the strategic delay of treatment and the subsequent incentive to buy more insurance coverage is inherently challenging. The decision about when to time a treatment, conditional on the recognition that treatment is necessary, is a complex, dynamic decision. A person needs to take into account his inventory of untreated events (if any), the insurance choices that will be available to him in the future, and the expected premium of these future insurance options. In addition, we (the econometricians) do not observe when an event that requires treatment is realized or the costs of delaying this treatment. Given this challenging environment, I adopt two complementary analytic approaches. My first approach is a series of reduced-form tests for the patterns in the data that theory predicts will exist when strategic timing and ex post adverse selection are important. The advantage of this approach is that I find clear and transparent evidence of both strategic claim delay and dynamic asymmetric information. This evidence also makes us aware of the variation that ultimately identifies the more structural analysis, which is my second approach.

For example, in my first analytic approach, I look for a spike in dental treatment (and spending) in the beginning of the calendar year among only those people who likely had an incentive to delay treatment because they were likely to have exceeded the previous year's maximum benefit. This analysis reveals overall striking patterns consistent with treatment delay. Further evidence reveals that there is considerable heterogeneity across procedures in these patterns, which suggests that some procedures may be systematically easier to delay than others. In addition, I inspect the data for evidence of dynamic asymmetric information. To do this, I look for a positive correlation between claims and choice of more generous coverage, a classic symptom of asymmetric information (Chiappori and Salanie, 2000). My analysis shows there is strong evidence of asymmetric information in this setting and that much of this asymmetric information operates within-household, over time, consistent with ex post adverse selection.

The tests employed in my first approach provide clear evidence of strategic treatment delay and dynamic asymmetric information. However, this evidence alone does not allow me to analyze policies that may address this asymmetric information, such as contracting on pre-existing conditions and open enrollment periods. This motivates my second analytic approach: an empirical model that explicitly links strategic claim delay to the adverse selection it can cause. The model I develop allows me to quantify the relative importance of ex post adverse selection (as opposed to traditional adverse risk selection) and to explore counterfactual policies. In the model, individuals realize dental events and then decide how much of the associated treatment to delay until the following year, knowing they will receive a stochastic draw of events in the following year that depends on their risk type. The model makes explicit the cost of delay (e.g., ongoing pain or cognitive costs). The key model primitives I estimate are observed and unobserved heterogeneity in delay costs and dental risk. Intuitively, the distribution of risk is identified by the distribution of total claims in the data (without regard for year-to-year timing), and the distribution of delay costs is identified by how individuals close to their maximum benefit allocate claims between adjacent years. Note that this latter source of identification is one of the key correlations examined in my first, more descriptive approach.

The model estimates suggest that approximately 40% of individuals strategically delay claims from one year to the next when they have incentives to do so. The estimates also reveal interesting demographic heterogeneity: women delay treatments more often than men, and older individuals delay treatments more often than younger individuals. In addition, the model estimates suggest there is considerable heterogeneity across types of claims in the propensity to delay. Using the model estimates, I extrapolate beyond the firm's benefits to investigate the broader implications of this strategic behavior. Specifically, the counterfactual analysis investigates the missing market for comprehensive dental insurance revealing two key lessons. First, ex post adverse selection alone causes more unraveling and reduces welfare more than traditional adverse selection alone in this setting. Relative to an equilibrium with only traditional adverse selection, an equilibrium with only ex post adverse selection is associated with 61% lower insurance enrollment and 15% lower per-capita welfare. The counterfactual analysis illustrates that even if insurers could perfectly price ex ante risk types (eliminating traditional adverse selection), the ability individuals have to delay treatment could alone explain substantial unraveling in this setting. Second, reducing the frequency of open enrollment periods, thereby effectively lengthening the commitment to contracts, can substantially improve welfare. If insurers can risk-adjust premiums to reflect ex ante risk types, simply moving from an annual open enrollment period to an open enrollment period once every two years would more than double insurance enrollment and increase welfare by 31%. Further reducing the frequency of open enrollment periods to once every five years would increase insurance enrollment by nearly eight-fold and increase welfare by 68% relative to annual enrollment. Overall, I show that the strategic delay of claims and ex post adverse selection is one explanation for why so few people have any dental coverage in the US and why available policies offer so little insurance. More generally, the results illustrate that features such as infrequent open enrollment periods and contracting on pre-existing conditions may overcome market unraveling in insurance contexts where the timing of risk is not always contractible.

The remainder of the paper proceeds as follows. Sections 1 and 2 describe, respectively, the related literature and the data. Section 3 discusses the theoretical incentives for delaying treatment and selecting coverage ex post in the context of the studied firm's dental benefits. Section 4 presents evidence that people behave according to these incentives. Section 5 describes the empirical model—its setup, identification, estimation, and results. In Section 6, I analyze counterfactual policies, and I discuss robustness of the findings. Lastly, I conclude in Section 7 by summarizing my findings and describing the potential relevance for broader insurance markets.

1 Related Literature

This paper contributes to several distinct literatures. Building on the seminal work of Akerlof (1970) and Rothschild and Stiglitz (1976), a growing body of recent empirical literature has focused on identifying and quantifying the impact of asymmetric information in insurance markets. Much of this literature builds on the work of Chiappori and Salanie (2000), who outline a robust set of tests for asymmetric information. The results of many subsequent papers that use some version of this positive correlation test have been mixed: some papers find little evidence of asymmetric information in particular markets (e.g., Chiappori and Salanie (2000), Cardon and Showalter (2001)), while some studies find evidence of asymmetric information in other markets (e.g., Finkelstein and Poterba (2004), Cohen (2005)). A number of recent papers go beyond testing for the presence of asymmetric information and seek to quantify the effects of asymmetric information in insurance markets.[11] See Einav, Finkelstein and Levin (2010) for a comprehensive review of this literature. The present paper contributes to this literature in several ways. While the prior literature focuses almost exclusively on static asymmetric information, this paper studies a source of asymmetric information that is inherently dynamic. This paper also identifies and explores the consequences of a connection between two classes of asymmetric information that have traditionally been viewed as distinct: moral hazard and adverse selection.[12,13] This paper finds that this previously unexplored dynamic connection between moral hazard and adverse selection may motivate many insurance market features seen in reality, such as waiting periods and open enrollment periods. In addition, this paper extends the literature by investigating whether asymmetric information may explain a missing insurance market.

Although there are many risks for which individuals cannot purchase insurance, there has been very little empirical work investigating missing insurance markets.[14] One practical reason for this gap in the literature is that it is difficult to study a market that does not exist because finding data on realized risks is difficult and estimating willingness-to-pay is even more challenging. This paper begins to fill this gap by studying the missing market for dental insurance. I overcome the traditional data limitations by leveraging the fact that employers offer tax-subsidized dental fringe benefits, allowing me to estimate costs and extrapolate from this data to estimate willingness-to-pay for a non-existent market for comprehensive insurance. While many papers looking at the effect of asymmetric information have concluded that adverse selection has small welfare consequences, prior studies have restricted attention to relatively well-functioning insurance markets (that is, insurance markets that exist for which there are data). In this paper, I find relatively large consequences from asymmetric information that can plausibly explain why one insurance market fails to exist. This contributes to the literature on asymmetric information because it illustrates that asymmetric information may have important consequences in harder to study missing insurance markets and may even explain why some risks are uninsured.

[11] Some recent papers in this literature include Cardon and Showalter (2001), Cohen and Einav (2007), Einav, Finkelstein and Schrimpf (2010), Bundorf, Levin and Mahoney (2012), Einav, Finkelstein and Cullen (2010), Carlin and Town (2010) , Sydnor (2010), Handel (2013), Bajari, Hong and Khwaja (2006), Hendren (2013), and Lustig (2010)). For a more comprehensive review of this literature, see Einav, Finkelstein and Levin (2010).

[12] The delayed treatment and the associated adverse selection studied in this paper suggests a new theoretical channel for why a type of moral hazard can generate adverse selection. Prior empirical work estimating adverse selection lumps together all sources of adverse selection (including adverse selection based on delayed treatments), and one contribution of this work is drawing an empirical distinction between sources of adverse selection that have different policy implications. Prior theoretical models (and structural empirical models) do not allow for the type of asymmetric information I study here in that they assume adverse selection is a pure hidden information issue distinct from any hidden action in the form of treatment timing. Relative to models of asymmetric information in the prior literature, the contribution of this paper is to document this new theoretical channel linking a type of moral hazard with the adverse selection it generates and to evaluate the welfare consequences of policy interventions that target this source of asymmetric information (e.g., decreasing the frequency of open-enrollment periods, pricing pre-existing conditions, etc).

[13] A related topic is explored in a recent paper by Einav et al. (2013). The authors investigate adverse selection based on individuals' static health care utilization elasticity (or static "moral hazard type"). In contrast, this paper investigates a dynamic connection between adverse selection and moral hazard that can arise in a population even with identical preferences including moral hazard types. Instead, the adverse selection in this setting arises dynamically as people become differentiated at different points in time by the amount of treatment they have delayed from prior periods. One example of how these concepts differ is that the asymmetric information studied in this paper would predict that individuals would switch plans over time even if the available menu of contracts remained fixed. This would not be predicted in the presence of the static asymmetric information studied by Einav et al. (2013).

[14] In a related recent paper, Hendren (2013) investigates insurance rejections in existing non-group insurance markets and the extent to which they can be explained by asymmetric information.

This paper contributes to the literature on dynamic inefficiencies in insurance. The prior literature on dynamic inefficiencies has primarily focused on reclassification risk and the difficulty of insuring long-term risk because of evolving risk types (e.g., Cochrane (1995), Hendel and Lizzeri (2003), Finkelstein, McGarry and Sufi (2005), Handel, Hendel and Whinston (2015)).[15] Such dynamic inefficiencies arise in the absence of enforceable lifetime contracts because individuals are typically not fully insured against becoming a bad risk and being reclassified into a higher risk group associated with higher insurance premiums. Reclassification risk is an inefficiency resulting not from asymmetric information, but instead from dynamically evolving risk known to both the insurer and the individual. While the inefficiency I study in this paper also exists in the absence of enforceable lifetime contracts, the source of the inefficiency here stems from the non-contractibility of the underlying uncertainty (and thus the reliance on the timing of claims to determine insurance coverage). Thus, in contrast to the literature on reclassification risk, the dynamic inefficiency studied in this paper arises because of the inherent asymmetric information between insurers and consumers that occurs when insurers cannot observe the resolution of uncertainty and individuals can both delay claims and re-evaluate insurance decisions. The results in this paper extend the literature on dynamic inefficiencies by demonstrating that dynamic asymmetric information can generate large dynamic inefficiencies even in settings in which underlying fundamental risk is not evolving over time, providing another motivation for designing markets and contracts to address limitations on long-term contracting.

This paper also contributes to the literature on the elasticity of health care utilization. Traditionally, researchers estimating the elasticity of health care utilization have adopted a static view of health care demand.[16] The typical assumption is that the patient decides to consume health care this period or never based on a single static price of care. In reality, patients typically have a richer set of options available: to treat a health problem now, to postpone the treatment to some later date, or to forgo treatment altogether (postpone treatment indefinitely). Although some early theoretical papers have pointed out the difficulty in estimating elasticities when transitory demand is present,[17] there has been little empirical work that explicitly takes this dynamic prospective into account. There are at least two recent notable exceptions: Card, Dobkin and Maestas (2008) show that the near-elderly delay care in anticipation of better coverage from Medicare and Aron-Dine, Einav and Finkelstein (2013) show that individuals are not completely myopic when it comes to the price they respond to in health care settings. This paper builds on prior work by estimating the degree to which treatments can be delayed with claim-level data and by using these estimates to explore the consequences of this form of elasticity for adverse selection in the broader insurance market.

This paper contributes to the literature on the sophistication of individual decision-making in insurance and related contexts. While there is a sizable literature on enrollment decisions in related contexts, most prior studies have focused on plan enrollment induced by forces external to individuals: changes in prices

[15]Some other sources of dynamic inefficiencies have also been studied. For instance, Fang and Gavazza (2011) study the under-investment in health that can occur in an employer-provided health insurance system due to labor market frictions and employee turnover.

[16]The gold standard in this literature are the studies based on the Rand Health Insurance Experiment, a large-scale experiment from the mid-1970s in which households were randomly assigned to health insurance plans with differing levels of benefits. Using data from this experiment, Manning et al. (1987) estimate the price elasticity for overall health care spending to be around -0.2. Using more recent non-experimental data, Kowalski (2010) identifies the price elasticity of medical care utilization using a novel instrumental variable approach. Using the injury of a family member as an exogenous shift in the marginal price for health care, the author estimates the price elasticity of utilization to be around -2.3, which is an order of magnitude higher than that found in the Rand experiment. In addition to the many theoretical and econometric reasons for the difference in these estimates discussed by Kowalski (2010), the strategic timing of care may be an alternative explanation for some of this discrepancy.

[17]Earlier papers by Arrow (1975) and Metcalf (1973) point out the difficulty of inferring long-run policy responses from temporary experiments (like the Rand HIE) due to transitory demand (potentially induced by the sort of strategic timing I study in this paper). Indeed, the Rand insurance experiment revealed some data that are consistent with the strategic postponement of care. Manning et al. (1986) show that the responsiveness of spending to insurance coverage is larger in the first year of the experiment than in subsequent years, and this elevated first-year responsiveness is more pronounced in dental spending than in other health care spending.

(e.g., Buchmueller and Feldstein (1997), Cutler and Reber (1998), Handel (2013)), changes in defaults (e.g., Madrian and Shea (2001), Goda and Manchester (2010), and changes in the set of options (e.g., Strombom, Buchmueller and Feldstein (2002), Abaluck and Gruber (2011)).[18] This paper contributes to this literature by investigating delaying claims and plan switching based on forces internal to individuals (in particular, their realized, untreated health problems). Much of the prior work reveals that people are somewhat myopic when it comes to changes in external forces related to coverage choices. In contrast, this paper reveals that people are quite strategic with respect to internal reasons to behave strategically. This finding may help us understand choice behavior in insurance contexts and decision environments more generally.

Finally, the present paper is also contributes to the broader literature on dynamic incentives by connecting this literature with the literature on asymmetric information in insurance markets. Economists have long been concerned with how incentives can distort intertemporal behavior on many margins: unemployment durations (e.g.,Katz and Meyer (1990), Chetty (2008)), retirement timing (e.g., Rust and Phelan (1997)), disability-induced labor force exits (e.g., Gruber (2000), Golosov and Tsyvinski (2006)), and re-timing of income to avoid taxes (e.g., Stiglitz (1985), Burman and Randolph (1994), Goolsbee (2000)). The analysis in this paper illustrates that strategic interemporal behavior and the associated adverse selection can create large distortions in private insurance markets, even leading insurance markets to completely unravel in some settings.

2 Data

I use rich, claim-level data from Alcoa Inc., a self-insured, multinational manufacturing company that employs approximately 48,000 individuals in 40 states across the US. The company offers dental benefits to employees and their dependents, and covers approximately 110,000 individuals through dental benefits annually. The data contain claim-level information regarding dental and medical spending as well as insurance coverage choices from 2004 to 2007. In addition, the data contain basic demographic information for employees, including wage, sex, age, and job tenure. The data include claims for all employees and insured dependents.

Each claim contains information on the total claim cost, out-of-pocket expenses, insurance payment, date of service, and procedure codes. While the data contain financial information for each separately billed procedure, I aggregate this information to the individual-visit level, and "claims" in the remainder of the paper will refer to the individual's total billed procedures within a visit.[19] While most of the analysis uses this claim-level data, the disaggregated procedure-level data is used to explore heterogeneity as discussed further below. A crucial feature of the claims data is that it contains information on *all* dental spending for insured individuals. All claims submitted to the insurance administrator are reported in the data, including unpaid claims after the individual has exhausted his annual benefits.

[18] More closely related to this paper, a few descriptive studies have estimated correlations between health insurance switching and changes in health status that are potentially consistent with strategically delayed treatments (e.g., Cutler, Lincoln and Zeckhauser (2010), Tchernis et al. (2006)).

[19] There are typically multiple procedure codes billed for a dental visit. Some dental visits may necessarily involve multiple procedure codes because of the complexity of treatment, while other visits may contain multiple procedures by the patient's choice. While ideally a claim would be defined as a related collection of procedure codes that cannot be unbundled by patient choice, the data do not contain the information necessary to make this distinction. In the remainder of the paper, the word "claim" is used to describe the total procedures claimed for an individual on a particular date, at a particular location of care (e.g., outpatient, inpatient), and by a particular type of provider (e.g., dentist, dental specialist).

2.1 Definition of Baseline Samples

Dental insurance options vary within the company by employee benefit groups. Employees are divided into benefit groups that reflect the firm's subsidiary business model, as well as employee occupation, job location, and union membership. The analysis in this paper focuses on the most common dental benefit menu that was offered to approximately 70% of employees over the observed years. While the company introduced this menu in 2004, the company rolled out this dental benefit menu over a number of years because of staggered union contract expirations.

Employee demographic information is described in Table 1 for each of the samples analyzed in this paper. The first column describes all employees who were ever observed in the data over the available years, 2004-2007. The second column describes the employees who have the relevant benefit menu offered to them at some point during the observed years; these employees along with their associated dependents will be referred to as the "baseline sample." In Section 4, the baseline sample is used to look for evidence of claim timing and dynamic asymmetric information. The third column summarizes the employees in the sample used to estimate the empirical model. This sample is restricted to employees and dependents associated with employees who were offered the relevant dental benefits in the two years used in the estimation of the model,[20] and this sample will be referred to as the "restricted sample." [21]

From inspecting Table 1, one can see that the median employee tenure is about ten years. The majority of employees are male, and about 40% of employees live in rural areas. The median wage is around $46,000, and the median employee age is 45 years. Approximately 75% of employees choose to enroll dependents in dental insurance. The individuals in the restricted sample look a bit different from those in the overall company population; fewer of those in the restricted sample are unionized,[22] their earnings are slightly higher on average, and the median age is a bit higher. One can compare the employee demographic information against a representative sample of employed individuals with dental insurance in the US. The last column in Table 1 displays some descriptive statistics for the sample of individuals in the Medical Expenditure Panel Survey (2007) who are continuously employed and enrolled in dental insurance throughout 2007.[23] The median age and mean wage look broadly similar in the company and the overall employee population. Compared to the overall employee population, a much larger fraction of the company employees are male and unionized. Dental spending of individuals in the baseline sample and the overall US population are compared in detail in Appendix A.

2.2 Description of Plan Details

Table 2 describes the dental insurance benefits of Plan L and Plan H, the two plans available on the relevant benefit menu. This table reports the percentage of dental expenditures that the company will pay below the annual maximum benefit by dental claim category: basic care, major care, oral surgery, and preventive care. Plan H has a $2,000 annual individual maximum benefit, while Plan L has a $1,000 annual maximum benefit per covered individual. Once individuals reach this annual maximum benefit, the insurer does not

[20] Conditional on being an employee in a given year, 82.5% of employees remain with the company in the following year; other employees part with the company due to termination, suspension, and retirement.

[21] Because some specifications of the empirical model include covariates based on prior utilization, a further sample restriction is made to construct the restricted sample. In addition to individuals being in the data during the two-year period used to estimate the model, the individuals must have been in the data the year prior as well to be included in the restricted sample.

[22] Those in the restricted sample are offered the relevant dental benefits in both 2005 and 2006. Many of the unionized employees in the company do not meet this criteria because the company only switched them to the relevant benefit menu after 2005 due to union contract expiration dates.

[23] This is not a perfect comparison group as the individuals in this sample may have obtained dental insurance through sources other than their own employer (for example, their spouses' employer).

reimburse for subsequent dental treatment. Because this maximum benefit is at the individual level (as opposed to the household level), an individual's dental spending doesn't affect the price for dental care faced by other household members.

The company varied the length of commitment to this annual insurance coverage over the time period studied. In particular, the company "locked" employee dental insurance decisions for two years, 2005-2006, and later reverted to annual dental coverage decisions. Coverage decisions are made during November of the year preceding the calendar year for which the coverage first applies. By November, it is likely that individuals will know almost all their dental problems for the current calendar year and will generally know whether they have delayed claims to the next year. It is important to note that even during the years with locked insurance coverage, the insurance terms still included an annual maximum benefit that applied in each calendar year. This means that, even during a locked period, individuals have the incentive to delay claims from one year to the next if they require dental treatment that puts them at the individual annual maximum benefit threshold during the first year of the locked period.

Similar to the majority of firms that offer dental insurance to employees, the company subsidizes this insurance in the sense that premiums are lower than the average cost of insured individuals. Plan L is available to all employees and dependents at no cost, while Plan H is available for a premium. It is convenient that Plan L coverage is free to employees as this means there is universal coverage for dental care among employees, so all employee dental usage is recorded by the company and available in the data. The premium for Plan H depends on the chosen coverage tier: employee-only coverage, employee plus family coverage, employee plus spouse coverage, or employee plus children coverage.[24] In addition, the Plan H annual premium varies across benefit groups and over time. In 2004, the average Plan H premium for family coverage was just under $150 while the average premium for Plan H single coverage was around $50. In 2005, average Plan H premiums increased to $200 and $65 for family and employee-only coverage respectively. Plan H premiums remained roughly constant for the remainder of the sample period. Appendix A displays the average premium for Plan H by coverage tier over the years.

Employees that select Plan H coverage pay the associated premium with pre-tax income. Additionally, some out-of-pocket dental expenditures may also be paid with pre-tax income saved in tax-advantaged accounts offered by the company.[25] Because the data does not include information on which out-of-pocket expenses were paid with funds from tax-advantaged accounts, premiums and out-of-pocket expenses are treated symmetrically throughout the paper except where otherwise noted. As will become clear later, the estimation of the empirical model is not sensitive to the tax treatment of premiums.[26]

The reimbursement of dental spending under the two plans varies somewhat with the category of care. Examples of these categories are given in the employee dental plan information brochure: basic care (e.g., fillings, root canal therapy), major care (e.g., bridgework, dentures), preventive care (e.g., exams, cleanings,

[24]Because the Plan H premium varies with the coverage tier, it is possible that some dependents go uninsured in some years. The selective enrollment of dependents is another potential avenue for adverse selection (in addition to plan choice) and is dealt with in detail in Section 4 and Appendix A.

[25]Employees may use Flexible Spending Account funds or Health Savings Account funds to pay for out-of-pocket dental or medical expenditures. Roughly 19% of employees opt to contribute to Flexible Spending Accounts. Analysis presented in Appendix A demonstrates that the main claim-timing patterns are equally present among employees that do not contribute to Flexible Spending Accounts as among those that do contribute, indicating that the variation used for identification is driven by the dental insurance plan nonlinearities not Flexible Spending Account incentives.

[26]The estimation is insensitive to the tax treatment of premiums because (1) premiums remain stable over the time period used to estimate the model, and (2) the model is estimated during a period in which coverage decisions were locked. The only avenue through which the tax treatment can affect the model is through the calibration of the risk aversion parameter. See Section 5 for details on the calibration of the risk aversion parameter. Robustness analysis presented in Appendix C reveals that the main estimates and lessons from the counterfactual analysis are robust to a wide range of risk aversion values.

emergency pain treatments), and oral surgery (e.g., removal of impacted teeth). Table 2 displays the breakdown of claims and spending for the baseline sample by inferred category of care.[27] Approximately 57% of claims are for preventive care, while 42% claims are for basic care. However, 63% of dental spending is basic care spending, while 32% is preventive care spending. The remaining spending and claims are for major care and oral surgery. The estimation of the empirical model uses individual claim-level data on the reported total and insurer dental spending to determine the precise budget set of each individual relative to the applicable maximum benefit to accurately capture the incentives for delaying treatment. However, for simplicity, the more descriptive analysis abstracts from these categories of care and measures out-of-pocket spending using a plan's average coinsurance rate, γ_j, defined as the average percentage of spending the company reimburses across the categories of care below the maximum benefit.[28] The average coinsurance rate for Plan L is 88.3% below the maximum benefit of $1,000, and the average coinsurance for Plan H is 89.9% below the maximum benefit of $2,000.

Figure 1 plots individual out-of-pocket expenditures (excluding premiums) as a function of total individual expenditures using the baseline sample plan average coinsurance rates. For Plan j, one can write this function as follows:

$$\mathrm{OOP}_j(x) = x - \min(\gamma_j x, b_j). \tag{1}$$

Because this calculation excludes premiums and the two plans are vertically differentiated, in the figure the out-of-pocket expenditures on Plan L are greater than the out-of-pocket expenditures on Plan H for any given level of total expenditures. Inspecting the figure, one can see that the main difference between the plans is the difference in annual maximum benefits. One can calculate the amount of individual dental spending it would take to exhaust the annual maximum benefit of each plan (the spending levels that correspond to the location of the kink points in Figure 1). An individual facing the average coinsurance rates above would exhaust the annual maximum benefit by spending around $1,133 ($= \frac{1{,}000}{0.883}$) on Plan L or $2,225 ($= \frac{2{,}000}{0.899}$) on Plan H. To get a sense of how much dental spending one would need to be better off ex post having Plan H coverage, one can compare the Plan H premium to the out-of-pocket spending differences displayed in Figure 1. Given the employee-only premium of $65, a single employee would be better off *ex post* under Plan H if he had more than approximately $1,184 of dental spending.

2.3 Description of Dental Claims and Spending

Figure 2 panel a displays the distributions of annual individual dental expenditures and claim cost for the baseline sample. While the majority of claims cost less than $200, the sparse right tail of claims reaches a few thousand dollars. The mean cost of a claim is $148 with a standard deviation of $161. The mean annual individual dental spending in the baseline sample is $272 (median $123), while approximately 37% of individuals have no dental expenditures in a given year. The right tail of dental expenditures is thin with the maximum observed annual spending around $20,000. Figure 2 panel b displays the right tail of annual individual dental spending by plan. Inspecting the figure, one can see that the percent of individuals with dental spending falls sharply near the point at which individuals would exhaust the maximum benefit of each plan, around $1,133 for Plan L and $2,225 for Plan H. There also seems to be some bunching of individuals near the level of spending it would take to exhaust the annual maximum benefit of Plan L.

[27] Claim categories are inferred by combining procedure codes and the claim reimbursement information. The average out-of-pocket spending to total spending ratio is calculated for each procedure code, and these codes are then classified into care categories. This process left less than 5% of procedures with unclassifiable codes. Claims with these codes are omitted from the statistics on the percentage of procedures and spending by category.

[28] This is basically a weighted average of the coinsurance rates for the different categories of care. Specifics of this calculation are described in Table 2.

Though Figure 2 is created using the baseline sample, Appendix A includes figures that demonstrate that the annual expenditure and claim cost distributions look very similar in the restricted sample.

Table 3 displays descriptive statistics by plan enrollment for the samples used in the following sections. The baseline sample includes 104,636 individuals across the years, while the restricted sample contains 29,559 individuals. Approximately 79% of household-years are enrolled in Plan H in the baseline sample and 74% of household-years are enrolled in Plan H in the restricted sample. A large fraction of individuals on both plans have zero expenditures in a given year. Thirty-eight percent of individuals on Plan H in the baseline sample have no dental expenditures in a given year despite paying a premium for Plan H coverage. The large fraction of individuals and families with no dental spending selecting Plan H indicates that there is some uncertainty in dental spending and many individuals value the available coverage of this dental uncertainty despite the low annual maximum benefit.[29] Individuals enrolled in Plan H have higher expenditures than those enrolled in Plan L: $85 more expenditures on average in the baseline sample, and $66 more expenditures in the restricted sample. In the baseline sample, 3.3% of individuals enrolled in Plan L reach the $1,000 annual maximum benefit, while 3.9% of Plan L enrollees in the restricted sample reach this maximum benefit. Approximately 1% of individuals enrolled in Plan H reach its $2,000 maximum benefit in either sample. The timing incentives explored in the following section may explain why relatively few individuals have expenditures that reach or exceed the annual maximum benefit. Significantly more individuals get within $200 of exhausting the individual annual maximum benefit. Among those on Plan L, 5.5% in the baseline sample and 6.3% in the restricted sample get close to the maximum benefit in a given year, while among those on Plan H, 1.5% in the baseline sample and 1.7% in the restricted sample get close to the maximum benefit.

3 Claim Delay and Selection Incentives

I discuss two general incentives for delaying claims in the present context and then discuss how delaying claims may lead to the subsequent incentive to select more generous insurance coverage ex post. Before continuing, it is worth reiterating an important distinction for the following discussion and the remainder of the paper: an "event" is a problem that requires treatment (for example, a dental cavity), and a "claim" is the treatment of an event (for example, the associated filling).[30] After realizing an event, an individual may decide to strategically delay the associated treatment (claim).

Two insurance incentives may motivate individuals to strategically delay claims. First, per-period non-linearities in insurance coverage, in this context the annual individual maximum benefit, can incentivize individuals to delay claims. When treatment costs exceed the annual maximum benefit threshold in a given year, an individual has an incentive to delay costs beyond this maximum until his benefits reset next January. In this way, the individual gains coverage for treatment that he would have paid completely out-of-pocket otherwise. Second, the opportunity to select more generous insurance coverage after the realization of events may motivate individuals to delay claims. In the environment studied presently, the opportunity to select more insurance coverage, to switch from Plan L to Plan H, may motivate individuals to delay more claims than they would otherwise delay if they were restricted to remain on Plan L. The intuition behind this is simple. An individual enrolled in Plan L who requires a lot of treatment beyond the maximum benefit, would have an incentive to delay some amount of treatment even if he was restricted to remain on Plan L, as in the period when the company locks coverage. However, if he had the opportunity to sign up for

[29]This demonstrated risk aversion can be seen more clearly through inspecting data for those who select employee-only coverage. Among those employees that select employee-only Plan H coverage, 43% have no claims in a given year.

[30]Since all spending is observed in the data, I make no distinction between treatment and claims in this discussion.

Plan H in the coming year, he may want to delay more treatment to take advantage of the higher Plan H maximum benefit. After delaying claims due to these incentives, individuals may have a subsequent incentive to select more generous insurance coverage. This subsequent ex post adverse selection incentive stems from the fact that an individual who has postponed treatment anticipates future treatment costs which generally increases his valuation of Plan H benefits. Appendix A contains a simple example to illustrate these incentives more explicitly.

In reality, many frictions may prevent individuals from delaying care or selecting more generous insurance ex post. For example, individuals may find it costly to delay claims because of pain or inconvenience associated with delaying treatment. In addition, cognitive limitations could inhibit both strategic claim delay and subsequent plan switching. Results described in Section 4 show that these frictions are not too large because there is evidence of both a substantial amount of strategic treatment delay in the data and evidence of asymmetric information associated with plan switching. Using an empirical model, Section 5 estimates that the extent of strategic timing is substantial, and these estimates are then used in Section 6 to investigate various counterfactual policies that may address this source of asymmetric information.

4 Evidence of Incentivized Behavior

This section examines the data for evidence of behavior encouraged by these timing and selection incentives. First, I show that coverage choices and claim realizations are positively correlated, indicative of some asymmetric information in this context. Consistent with ex post adverse selection, the results suggest that much of the asymmetric information in this context operates dynamically within-household, over time. Second, direct evidence of strategic claim delay is presented. Taking advantage of the annual maximum benefit contract feature, I use several tests to show that incentivized individuals postpone treatment until just after the commencement of a new calendar year (at which point their benefits reset).

4.1 Asymmetric Information

A central theoretical prediction in many models of asymmetric information is a positive correlation between coverage and claim realization (Chiappori & Salanie 2000). This positive correlation can arise for several reasons. Traditional moral hazard incentives may lead people with more comprehensive insurance to have more discretionary dental spending. People who are ex ante more risky may select more coverage, as in traditional adverse risk selection. Ex post adverse selection may lead to a positive correlation in claims and coverage over time as individuals sign up for more insurance in anticipation of delayed claims. Following previous empirical studies (e.g., Chiappori & Salanie 2000, Finkelstein & Poterba 2004), I look for evidence of positive correlation in coverage choice and claim realization to test the joint hypothesis that there is some type of asymmetric information in this setting.[31] To do this, I estimate the following equation:

$$Claims_{h,t} = \gamma + \alpha Choice_{h,t} + X_{h,t}\beta + \epsilon_{h,t}. \tag{2}$$

The basic idea behind this test is to look for a positive correlation between claim realization, $Claims_{h,t}$, and choice of the more generous insurance option, $Choice_{h,t}$, conditional on the household information priced by the insurer, $X_{h,t}$. The test for positive correlation is then a test of the sign and significance of α. To distinguish between the within- and across-household correlation in claims and coverage, the equation

[31] While the focus of the positive correlation analysis is to test for any source of asymmetric information, the model and counterfactual analysis in the following sections highlight the importance of distinguishing between these sources of asymmetric information to understand the welfare consequences of various market interventions.

above is estimated both with and without household fixed effects taking advantage of the panel nature of the data.

Let $Choice_{h,t}$ indicate that household h is enrolled in Plan H in year t. To ensure that the analysis focuses on vertically ranked insurance options, I limit the sample to those households that insure the same dependents under dental and medical insurance in each year.[32] Let $Claims_{h,t}$ be the amount of money the company would have paid in claims for household h in year t, had the household been enrolled in Plan H in year t (regardless of the actual household enrollment).[33] The covariates $X_{h,t}$ are those exogenous household characteristics the company uses to price $Choice_{h,t}$. In this employer-provided insurance setting, very little information is priced. Because there is some variation in the premium menu employees receive based on occupation and location, I control for the premium menu in the regression above.[34]

The results are reported in Table 4. The point estimate for the specification without household fixed effects indicates a very significant positive correlation between claim realization and coverage choice. Insurer expenditures per household under Plan H are on average \$215 higher in the first specification for households enrolled in Plan H than for households enrolled in less coverage. Because this estimate confounds sources of asymmetric information that would cause within-household and across-household correlation in claims and choices, Equation 2 is re-estimated with household fixed effects. The positive and significant point estimate for α in this specification indicates that households have more claims when the households select more coverage. In addition to the sign, the magnitude of α is notable: the within-household coefficient is more than 65% of the size of the overall coefficient, meaning a substantial amount of the positive correlation in this environment is within-household correlation, over time.[35]

Overall, the positive correlation estimates reveal that households switching from Plan L to Plan H increase their claiming behavior after switching (or correspondingly, households decrease claiming behavior after switching from Plan H to Plan L). To more directly look for evidence of ex post adverse selection, I investigate the correlation between switching to more generous coverage and household monthly claims. If households delay treatment in anticipation of switching to more generous coverage, we would expect to see that household claims would increase directly after switching to more generous coverage but that this elevation in claims would be short-term (as opposed to a long-run shift in the household's dental risk). Let $Claims_{h,m}$ be the amount of money the company would have paid in claims for household h in month m, had the whole household been enrolled in Plan H (regardless of the actual household enrollment). Figure 3

[32] Fixing the number of dependents covered by an employee, the company dental insurance plans, Plan L and Plan H, are vertically ranked. However, employees may select the number of family members to insure and have an incentive to select this number carefully due to premium variation based on the number of insured family members. This ability to select the coverage tier means that in reality there are more than two coverage options for most employees, and some of these options are not ranked vertically. Appendix A demonstrates the results are qualitatively very similar under alternative definitions of $Choice_{h,t}$ and alternative samples which take into account the dependent enrollment avenue for selection.

[33] $Claims_{h,t}$ is calculated by applying the Plan H cost-sharing rules to the spending of each household. In this context, this measure is more desirable than coarser measures of claims because it can capture the finer differences between the plans that may cause asymmetric information to be important. Examples of coarser measures of claims include a claim indicator or claim count. Using a measure similar to the one used here, Chiappori et al. (2006) suggest that the following condition is a test for positive correlation that is robust to many permutations in the plan differences, utility framework, and competitive environment: $\int R_H(d)\mathrm{dF}_H(d|X_i) \geq \int R_H(d)\mathrm{dF}_L(d|X_i)$, where $R_H(d)$ is the insurance payout for someone with d dental expenditures enrolled in Plan H and F_j, $j \in \{H, L\}$, is the distribution of dental expenditures for those who choose Plan j. When this condition holds, the authors say there is evidence of "relevant" asymmetric information. The test outlined in Equation 2 is simply this condition in this setting.

[34] Dental coverage tier is also included as an exogenously priced household characteristic in $X_{h,t}$ since the sample in this specification is restricted to households that treat their family composition as fixed for the purpose of insurance enrollment. Analysis in Appendix A demonstrates the results are similar when employing alternative sample restrictions and sets of controls.

[35] One might be worried that the difference between the within- and across-household estimation is due to compositional differences between the households that are in the sample for longer or short periods of time (perhaps because of differing job turnover rates). This is not the case. In Appendix A, I demonstrate that the results are qualitatively similar when the sample is restricted to those households that remain with the company and enrolled on a relevant plan throughout the data period.

plots the mean of $Claims_{h,m}$ for those households that switch to more generous coverage restricting attention to the year before and after the switch. The figure demonstrates that among households who switch to more generous coverage, there is a large increase in average claims directly after switching, and this increase in claims is relatively short-lived as average claims return to prior levels by the end of the calendar year. Claims increase by 62% between December and January coincident with the switch in coverage, and this increase is statistically significant (with an F-stat of 25.32). This difference is not simply due to seasonality in claiming behavior. Claims during January directly after the switch are 28% higher than in the January prior to switching coverage, and this difference is statistically significant with an F-stat of 7.65. By the end of the calendar year after switching, the claims return to levels similar to those seen prior to the switch; claims in the last month of the calendar year before and after the switch are statistically indistinguishable (the p-value associated with this difference is 0.23).[36] Although there may be alternative explanations for these patterns, this evidence suggests that ex post adverse selection may play a role in this environment.

4.2 Strategic Claim Delay

To look for evidence of strategic claim delay, the annual individual maximum benefit feature of the dental plans is exploited. When individuals are close to exhausting the annual maximum of their plan, they may have an incentive to delay treatment until the benefits reset next January. Note that the analysis in this section focuses on individual-level data because the annual maximum benefit applies to individual-level spending (as opposed to household-level spending). I inspect monthly dental spending in adjacent years for this predicted pattern of elevated beginning-of-year expenditures for individuals who are incentivized to delay claims. The sample here is limited to those individuals who were covered under the company dental benefits during adjacent years, who had positive expenditures across these years, and who were enrolled in Plan L during the first of these years. Because individuals with more expenditures across the two years are more likely to have been incentivized to delay claims between the years, spending patterns are separately examined for those with a lot of expenditures across the two years (expenditures exceeding $1,400), and those with less expenditures across the two years (expenditures less than $1,400). Figure 4 plots the average fraction of total individual expenditures by month, and this fraction is normalized to one so that a flat line at one would indicate that spending is, on average, equally distributed across the months. Among those more incentivized individuals, there is a large increase in the fraction of spending beginning in January of year 2, and this elevation in spending persists for the following six months. These incentivized individuals did 32% more spending in the first 3 months of year 2 than would be predicted if spending were equally distributed across time, and this difference is statistically significant (with an associated p-value < 0.001). In addition, there is an associated relative dip in spending among these people at the end of year 1. This elevation of claims in the beginning of year 2 is not simply due to seasonality. Spending in the first quarter of year 2 is 29% higher than spending in the first quarter of year 1 (with an associated p-value < 0.001).[37] The patterns in this figure suggest that incentivized individuals delay some treatments from the end of year 1 to the beginning of year 2 to get more coverage for these treatments. There is comparatively

[36] Appendix A displays the corresponding regression coefficients for Figure 3.

[37] In this setting, you can think about the baseline seasonality patterns as being established by two potential control groups: (i) the individuals themselves in years in which they are less likely to be marginal to the maximum benefit, and (ii) other individuals who have relatively low spending for whom there is virtually no incentive to strategically time care. Taking advantage of the two control groups described above, we can see that this elevation of claims in the beginning of year 2 is not simply due to seasonality: (i) We see that spending in the first quarter of year 2 is 29% higher than spending in the first quarter of year 1 among these same individuals (with an associated p-value < 0.001). (ii) We see that the difference in spending between the first quarter of year 1 and the first quarter of year 2 is much larger for the incentivized group than for the less incentivized group: 29% vs. 1%. Leveraging both control groups, the implied difference-in-difference estimate tells us that the differential increase in the incentivized group is 28% during the first quarter of year 2 (with an associated p-value < 0.001). See Appendix A for the complete regression results corresponding to Figure 4.

little monthly variation in dental expenditures among the less incentivized individuals. In Appendix A, I display the analogous regression results, and I demonstrate that these patterns are qualitatively unchanged when the cutoff between low- and high-spending individuals is moved by a few hundred dollars or when controlling for year 2 dental coverage.[38,39] Further analysis in Appendix A leverages variation in whether insurance decisions were locked for two years or unlocked, and this analysis illustrates that the annual maximum benefit (applicable in both the locked and unlocked periods) and the opportunity to switch insurance coverage (applicable in only the unlocked periods) are both important incentives for delaying treatment.

One might wonder to what extent are some types of procedures are delayed systematically more often than others. Next, I present evidence regarding the role of heterogeneity across procedures in contributing to the patterns seen in Figure 4. To do this, I use disaggregated procedure-level data and consider heterogeneity associated with two different classifications of procedures. The first classification relies on the degree to which the procedure is over-represented in the first quarter of year 2 among individuals with the incentive to delay claims (those with expenditures exceeding $1,400 over the two year horizon). Specifically, I classify procedures as follows. I begin with the data on procedures for those incentivized individuals used to construct Figure 4. For each procedure in the data, I calculate (i) the monthly frequency of the procedure relative to all procedures and (ii) the associated T-statistic on a test that this monthly frequency is the same across the first quarter in year 2 and the remainder of the two-year period. The resulting T-statistic measures the degree to which a procedure is over-represented in the first quarter of year 2 relative to the remaining quarters. I use the quartiles of this T-statistic distribution as a measure of the "time-sensitivity" of a procedure. There are a few ways to see that this classification reveals meaningful heterogeneity across procedures. First, this classification yields expected patterns among the types of procedures characterized as more or less time-sensitive. Inspecting Appendix Table A1, one can see that the most time-sensitive procedures according to this characterization include root canal therapy and removal of an impacted tooth, while the least time-sensitive procedures include crowns and partial dentures. Second, when the analysis in the prior figure is replicated using the subsets of more and less time-sensitive procedures, as displayed in Figure 5, we see that those procedures that are classified as relatively time-sensitive do not have a spike in January, while those classified as relatively less time-sensitive are associated with a large spike in January.[40] Overall, these patterns suggest that some procedures are delayed systematically more often than others.[41]

In addition to the above classification of procedures, I also investigate heterogeneity with respect to a second classification that relies on the billing code hierarchy used for dental care (the current dental

[38] As one would expect, those individuals who switch from Plan L to Plan H between the adjacent years display higher displacement between the end of year 1 and the beginning of year 2 than the non-switchers display. Still, those who do not switch also display displacement though somewhat smaller in magnitude. Of course, even non-switchers have incentives to delay in this context because of the annual maximum benefit feature of coverage.

[39] Another prediction of ex post adverse selection is that the more incentivized individuals are more likely to switch to Plan H than the less incentivized individuals. Indeed, the high spending individuals (who are relatively more incentivized in the sense of Figure 4) do switch to the more generous plan at a higher frequency. Approximately 26% of individuals in the relatively more incentivized group switch to Plan H in the second year, while only 15% in the relatively less incentivized group switch to Plan H in the second year.

[40] Figure 5 displays patterns consistent with individuals responding to insurance incentives by delaying claims for procedures that are not time-sensitive from the end of one calendar year to the beginning of the next year. Among relatively incentivized individuals (with overall spending exceeding $1,400), spending on procedures characterized as less time-sensitive is 50% higher in the first quarter of year 2 and 31% lower in the final quarter of year 1 than would be predicted if spending were instead equally distributed across time; this pattern is statistically significant, with an associated p-value < 0.001. As with the overall spending patterns, these patterns are robust to using spending of less incentivized individuals (those with overall spending less than $1,400) to net out underlying seasonality in dental spending. Appendix A displays the analogous regression results that correspond to the claim timing figures.

[41] If we instead saw that the spending patterns are very similar across procedures in different quartiles of "time-sensitivity", we might conclude that there is little variation across procedures that could explain the patterns of delay in Figure 4.

terminology, or CDT classification). The coding system groups procedure codes into a few aggregate categories: Preventive, Diagnostic, Restorative, and Other (including Endontics, Periodontics, Prosthodontics, and Oral Surgery). Focusing on procedures within each of these categories, Figure 6 displays the corresponding timing of spending among those more and less incentivized to delay claims. Even with this relatively coarse categorization, we see there is some heterogeneity in the delayed treatment across the categories. This figure suggests that preventive and diagnostic procedures are delayed from one year to the next less often than restorative or other procedures. It is not surprising that there is little evidence of delay in diagnostic and preventive procedures for several reasons. First, many individuals follow a regular yearly or bi-yearly schedule with respect to preventive cleanings and diagnostic x-rays, and there is little therapeutic benefit to piling up extra cleanings or x-rays even if a regularly scheduled visit is missed. Second, the company's insurance plans are structured such at there is basically no financial benefit to piling preventive and diagnostic procedures, as covered preventive cleanings are limited to two per calendar year and covered diagnostic x-rays are limited to two partial mouth x-rays per calendar year. More generally, it is very common for employer-provided dental plans to place annual limits on the number of covered preventive cleanings and x-rays.

Taken together, Figure 5 and Figure 6 collectively indicate that systematic differences across procedures may explain some of the delay seen in this context. The empirical model in the following section allows for analogous heterogeneity across claims in the costs of delay.

5 Empirical Model: Setup, Identification, Estimation, and Results

The prior section reveals direct evidence from the data consistent with strategic treatment delay and dynamic asymmetric information. However, this evidence alone cannot inform us of the relative importance of ex post adverse selection verses traditional sources of adverse selection in explaining market unraveling within the dental context. Thus, I develop and estimate a model that precisely specifies the determinants of the decision to delay claims making explicit the costs of delaying treatment, which may include pain and cognitive costs. Using this framework, I estimate heterogeneity in dental risk and delay costs using correlations similar to those displayed in Figure 4 of the prior section. Section 6 then uses the model estimates of the delay elasticity and the heterogeneity in risk as inputs to investigate the relative importance of ex post adverse selection and to explore counterfactual policies.

5.1 Model Setup

5.1.1 Overview

Though ex post adverse selection is of central importance in the counterfactual analysis, the focus of the model and estimation is on the economic primitives that affect the decision to delay claims between two adjacent years. This modeling approach takes advantage of two characteristics specific to this environment. First, the two-year period during which the company locked insurance coverage allows me to estimate the model without extra assumptions or data requirements that would be needed to model endogenous plan switching. Using data from the locked period, 2005-2006, the model is used to estimate observed and unobserved heterogeneity in delay costs and dental risk abstracting from plan switching.[42] Second, the

[42]Because plan switching decisions rely on individuals' expectations about plan offerings many years into the future, modeling such behavior with a short panel would require fairly heroic assumptions about expectations into the distant future. For this reason, I focus the delay decision between adjacent years during which insurance coverage is locked. In contrast to the plan switching decision, this delay decision requires very little foresight and thus it is possible to convincing capture all the relevant information in a stylized model using the short panel that is available.

incentive to switch plans in this environment is aligned with the incentive to delay claims from one year to the next. Though the model is estimated using the locked period during which there is no insurance plan switching, this second feature means that the estimated delay frictions are able to capture some of the frictions that may inhibit optimal insurance plan switching more generally. Under some additional assumptions, which are made clear later in the paper, the estimated distributions of dental risk and delay costs are sufficient for policy analysis related to claim timing and ex post adverse selection. Third, because coverage decisions are fixed during the estimation period, insurance terms are solely at the individual level. This means there are no interactions among family members' incentives to delay treatment (barring income effects), which allows the model to focus on individual-level decisions for delaying treatment.[43] While the main estimation focuses on individual-level delay decisions, household-level coverage decisions are used in the calibration of the risk aversion parameter which is an input in the model estimation, as discussed in more detail below.

I first outline the model for a single individual, and then explain how heterogeneity is incorporated. In the model, the individual decides how much treatment to postpone after realizing events (in this case, dental problems). To be clear on the terminology used here, the "timing of events" or the "event date" will refer to the date the individual becomes aware of an event and the necessary associated treatment.[44] The timing of the model is outlined below.

1. The individual realizes first-year dental events.
2. The individual decides how much treatment in dollars to delay until the second year, m. He makes this decision according to his cost of delaying claims, $c(\alpha_i, m)$. He treats the remaining events for which he has not delayed treatment.
3. In year 2, the individual realizes new events, and he treats these new events along with any treatment delayed from the previous year.

In this baseline model, claims (treatments) cannot be delayed beyond year 2, and claims are not carried over to year 1 from prior years. In other words, I abstract from initial and terminal condition issues.[45] While this simplifying assumption is employed in the baseline specification, Appendix C describes and estimates an alternative specification that relaxes this assumption by allowing individuals to consider the possibility of delaying claims in the future when deciding on the amount of claims to delay in the present. This alternative specification yields results very similar to the baseline estimates.

[43] The main estimation of the empirical model takes the insurance enrollment decisions as fixed, focusing on the incentives to delay care. All of the incentives for delaying care in this setting are at the individual-level arising from the individual annual maximum benefit. For this reason, the model focuses on the individual's decisions for delaying treatment. Note that focusing on an individual's optimization problem (instead of a household's optimization problem) is very general in this context for a few key reasons. There are no interactions among family members' incentives to delay treatment barring income effects. However, the assumption of CARA utility means income does not impact risk preferences. Thus, under very few additional assumptions, this means that a model where the household maximizes a household-level CARA utility function by choosing the amount each individual delays would give the same solution as the individual model. A complete derivation of this is in Appendix B.

[44] It is the date of event recognition that is important for strategic delay, not the date the problem was initially acquired (if different from the date of recognition).

[45] Also, implicit in this setup is the assumption that the only form of moral hazard is delaying treatment. That is, there are no optional treatments (or treatments that may be delayed indefinitely) in the model. This assumption is made to focus on claim timing rather than purely static moral hazard. In reality, some dental procedures are very optional. (The most obvious example of optional dental procedures are those done for purely cosmetic reasons that have become increasingly popular in recent years. Because cosmetic dentistry is not covered by the dental insurance offered by the company, these procedures are excluded from the analysis.) While this assumption may exclude some forms of moral hazard, much of what is traditionally thought of as moral hazard in this context has a timing element as well. Intuitively, if an individual does an "extra" procedure today, this crowds out the probability that he will do the same procedure for the same problem tomorrow. Returning to the two-period model, the assumption of no optional treatment implies that treating events is mandatory. An individual will treat his events when it is advantageous to do so, but ultimately he will treat them within the two years.

There are two components of the empirical model: the cost model and the decision model. The cost model specifies the frequency of events and costliness of treatment, while the decision model specifies the incentives to postpone treatment from year 1 to year 2.

5.1.2 Cost Model

The cost model describes both the frequency of dental events and the financial cost associated with treating these events. When individuals strategically delay treatment, the timing of events diverges from the timing of claims. Because the identification of the model relies on quantifying the extent of this divergence, it is important to describe the underlying frequency of events. One major challenge in this setting is that event timing is not directly observed as the data contains only claim information available to the insurer. Though the timing of claims puts some bounds on the timing of events, I need to place some additional restrictions on the timing of events in order to estimate the distribution of delay costs.

The assumption I make is that dental events arrive independently over time where the rate of arrival depends on the individual's "risk type," λ_i.[46] This conditional independence assumption implies that the number of events received by the individual in year t, n_t, is governed by a Poisson distribution[47]: $n_t \sim$ Poiss(λ_i).

The individual's decision to postpone claims will depend on the number of events he expects to receive the following year and the expense of treating these events. For each event l, the treatment cost, c_l, is assumed to be an independent draw from the "cost intensity" distribution, $c_l \sim \mathrm{G}_{a_i}$, representing the empirical distribution of claim costs in the data. In the estimation, this distribution is allowed to vary with age, a_i, in a categorical manner to capture the fact that the types of treatments done by the middle-aged differ from those done by children.[48] Using this notation, the total cost of treating the events that arrive for the individual in year t can be written as follows:

$$d_t = \Sigma_{l \leq n_t} c_l. \tag{3}$$

5.1.3 Decision Model

The decision model describes how the individual chooses the amount of treatment to delay between the adjacent years. Suppose the individual is enrolled in Plan j over the adjacent years, and the individual has Constant Absolute Risk Aversion per-period utility, $u(c) = -exp(-rc)$, with risk aversion r.[49]

The individual maximizes the sum of his first-year utility and his expected second-year utility by choosing m, the dollar value of delayed claims, from M, the set of possible delay amounts. The individual does not discount his second period utility, and the individual is permitted to save and borrow without interest,

[46] It is important again to highlight the distinction between events and claims. Although events (e.g., cavities) are assumed to be independent over time, claims (e.g., fillings) can be serially correlated for a number of reasons including individuals' decisions to delay claims based on insurance incentives.

[47] Though at first blush this conditional independence assumption may seem restrictive, the model is robust to some correlation in events. For example, the model accommodates some correlation in events through allowing for rich heterogeneity in risk types.

[48] For this purpose, age is discretized into the following categories: Less than 19 years of age, 19-30 years of age, 31-40 years of age, 41-50 years of age, and over 50 years of age.

[49] The CARA per-year utility form is convenient as any component of consumption that remains constant across the two years will be ignored in the individual's optimization problem. I assume the individual does not expect changes in income over the two years so income drops out of the individual's decision to postpone dental claims. Although premiums change slightly over the locked period ($5 on average), I assume that individuals treat premiums as constant across the two years as well.

denoted as s below.[50] The individual's optimization problem is below:

$$\max_{m\in M,s} \quad u(-OOP_j(d_1-m)-c(\alpha_i,m)-s)+E_{\tilde{d_2}}[u(-OOP_j(\tilde{d_2}+m)+s)|\lambda_i,G_{a_i}]. \tag{4}$$

In the first year, the individual pays out-of-pocket expenses associated with his incurred dental claims, $\text{OOP}_j(d_1-m)$. The individual also pays a "delay cost" to postpone m dollars worth of claims, $c(\alpha_i,m)$, and the individual is permitted to save, s. In the second year, the individual pays dental out-of-pocket expenses that are a function of the delayed claims and the second year events, $\text{OOP}_j(d_2+m)$. In addition, he receives savings from the previous year, s. When the individual makes his decision to postpone treatment, he does not yet know what dental events will arrive in the second year or how costly these events will be to treat. Thus, the second term in Equation 4 is the expected second-year utility where the expectation is taken over d_2, the cost of treating events that arrive in year 2. This expectation is conditional on the information the individual knows about the distribution of d_2: his risk type, λ_i, and the cost intensity distribution, G_{a_i}. Note that the decision model accounts for the nonlinearities of the individual's budget along two important dimensions: (i) non-linearities with respect to within-year dental spending (due to the annual maximum benefit as captured by the out-of-pocket cost function) and (ii) non-linearities with respect to time (due to the reset of the benefits at the start of each calendar year).[51]

The individual may find it costly to delay treating dental events for many reasons, including physical pain caused by delaying a treatment, discomfort with postponing a treatment recommended by a dental professional, and inconvenience associated with maximizing insurance coverage of dental needs. To rationalize these frictions, I assume the individual pays the following delay cost:

$$c(\alpha_i,m)=\alpha_i\text{I}(m>0). \tag{5}$$

This cost is equal to α_i for any positive amount of delayed treatment, and the interpretation of this form of the delay cost is discussed further below.

If individuals could delay any continuous amount of claims, this model would predict a large mass of annual dental spending corresponding to the level that exactly exhausts the maximum benefit. In the data, however, many individuals have dental spending just short of exhausting the maximum benefit, or dental spending just above the maximum benefit. This pattern suggests that individuals face some rigidities in the dollar value of treatment they may delay, so it is conceptually important for the model to include some restrictions on the possible delay amounts, M, and I make two such restrictions.[52] It should be emphasized that although it is conceptually important to make some restriction, the particular restrictions I make may not correspond to the actual restrictions individuals face; I make these particular restrictions only for computational simplicity. First, I assume the individual can delay expenses only claim by claim. For example,

[50] It is assumed that individuals have an interior solution for savings. Thus, I can calculate the optimal savings as a function of the remaining parameters and substitute this into the objective function, meaning that savings data are not needed for estimating the model. See Appendix B for a further description of how optimal savings affects the estimation of the model.

[51] The model accounts for the non-linearity with respect to within-year dental spending using the out-of-pocket cost function (see Figure 1). The non-linearity in the budget with respect to time is the focus of the analysis, and the novel piece of the model relative to the prior literature. Non-linearity with respect to time is accounted for within the model because individuals choose the amount of care to delay by maximizing the sum of the current period utility and their expected future utility which explicitly accounts for the fact that their benefits will reset in January.

[52] Absent any restrictions on M, a mass of people exactly at the benefit exhaustion point would be predicted because of the form of the cost function above. That is, because the cost function is not increasing in the amount of treatment delayed, individuals will either delay everything beyond the maximum or nothing. Practically, it is possible to assume a different cost functional form instead of adding additional rigidities in the possible delay amounts, M, to rationalize the data. However, restricting M in addition to the cost function captures the separate forces that lead to the degree of bunching we see in the data: the degree to which people can delay treatment for intrinsic reasons related to personal or claim characteristics (captured in the cost function) and the limitations on bunching due to the fact that claims are not divisible (captured by the restrictions on M).

this prevents the individual from delaying half of a cavity filling from year 1 to year 2. Second, I assume the individual cannot manipulate the claim order. This assumption means that the sequence of claims in the data reflects the sequence of events the individual received.

5.1.4 Heterogeneity

Observed and unobserved heterogeneity is introduced in two parameters: the risk type, λ_i, and the delay cost α_i. The cost intensity distribution, G_{a_i}, varies with age in a categorical manner. This means the ex ante heterogeneity in total dental spending (conditional on age) stems from heterogeneity in the frequency of events (which is governed by the risk type λ), and not in the cost per event. In this context, the risk type can be thought of as a measure of one's overall dental health (conditional on age), where those with better health (lower risk) receive fewer events on average.

The parameters λ_i and α_i are known to the individual but unobserved by the researcher. The delay cost, α_i, is assumed conditionally independent of the risk type, λ_i, conditional on the observables included in the modeling of each of these parameters as described below. Risk type, λ_i, is assumed to come from a lognormal distribution with parameters $\mu_\lambda(X_i)$ and σ_λ:

$$\lambda_i \sim \text{lognormal}(\mu_\lambda(X_i), \sigma_\lambda). \tag{6}$$

The distribution of risk types is allowed to vary with individual characteristics, X_i, including age, sex, and prior utilization measures based on the year just prior to the estimation period,

$$\mu_\lambda(X_i) = X_i\beta_\lambda. \tag{7}$$

In addition to age and sex, the baseline specification also includes prior utilization measures to increase precision. In particular, these measures of prior utilization include an indicator for having any visit in the prior year and categorical variables on spending in the prior year, indicating that spending was in one of the following categories: \$0-500, \$500-1000 or \$1000+.[53]

For the delay cost, I assume that individuals can either freely delay claims ($\alpha = 0$) or cannot delay claims ($\alpha = \infty$) between the two years. In other words, individuals who can delay claims (for whom $\alpha = 0$) do so strategically as in the optimization problem outlined above assuming it is costless to delay claims. Thus, the delay cost distribution can be written as follows:

$$\alpha_i = \begin{cases} 0 & \text{with probability } p_\alpha(X_i, Z_{c_i}) \\ \infty & \text{with probability } 1 - p_\alpha(X_i, Z_{c_i}). \end{cases} \tag{8}$$

The focus of the estimation is then to characterize the probability that an individual can optimally delay claims, $p_\alpha(X_i, Z_{c_i})$. If we integrate this probability across the distribution of characteristics in the population (X_i, Z_{c_i}), then we obtain the fraction of individuals who optimally delay claims.

There are multiple ways to interpret the estimated heterogeneity in the delay cost. The variation in delay costs could be driven by individual characteristics such as sophistication in navigating insurance incentives or pain tolerance. Alternatively, the variation in delay costs could be driven by characteristics of marginal claims such as urgency. In reality, the cost of delaying claims probably has some component that is claim-specific but common across individuals and some component that is individual-specific but common across claims. The delay cost distribution, summarized by the probability of costless delay $p_\alpha(X_i, Z_{c_i})$, is allowed to vary with individual characteristics, X_i, (such as age and sex) as well as characteristics of the

[53]The estimation sample is restricted to those households that were in the data for three years: the two-year estimation period and the year just prior to the estimation period. Thus, the prior claim experience information is available for all individuals in the estimation sample.

marginal claim, Z_{c_i},[54]

$$p_\alpha(X_i, Z_{c_i}) = X_i\beta_\alpha + Z_{c_i}\delta_\alpha. \quad (9)$$

Allowing for heterogeneity across these two dimensions allows me to investigate the degree to which delay costs are associated with certain observable individual characteristics and claim characteristics. I estimate various specifications allowing for heterogeneity in claim characteristics analogous to the dimensions of heterogeneity discussed in Section 4.[55]

5.2 Identification

The risk type distribution parameters, $(\beta_\lambda, \sigma_\lambda)$, are identified by the variation across individuals in the total number of claims observed during the two years (without regard for year-to-year timing). The delay cost distribution parameters, $(\beta_\alpha, \delta_\alpha)$, are identified by the division of claims between the two years among individuals with claims close to the maximum benefit in the first year (who likely had an incentive to delay claims). Notice this source of identification is very similar to the evidence depicted in Figure 4 (discussed in Section 4).

Intuitively, only individuals with the incentive to delay claims give us any information about the delay costs. In this environment, individuals have an incentive to delay claims if and only if they receive year 1 events that put them beyond the maximum benefit. Because events are not observed directly, I must use the model, in combination with the claims data, to infer which individuals may have received events that would have put them at the plan maximum benefit in the first year. These "potential delayers" are those individuals who have first-year claims that either exceed the maximum benefit or would have exceeded the maximum benefit if their first claim from year 2 had instead been claimed in year 1.[56] In the sample used to estimate the model, 2.1% are potential delayers.[57] For those who are potential delayers, the model determines the likelihood of observing the data given the possible delay costs. Intuitively, an increase in the fraction of potential delayers with many claims beyond the year 1 maximum benefit would lead to a lower estimate of p_α (the fraction who strategically delay claims if incentivized).

In the counterfactual analysis, the estimated parameters governing the delay cost distribution are applied to the entire sample. Though the delay parameters are identified by these incentivized individuals, the estimated parameters are relevant in the entire sample under two maintained assumptions: (1) conditional independence between risk types and delay costs and (2) plan enrollment is unrelated to delay costs. I estimate specifications that include many observable characteristics to relax the first assumption by al-

[54]For each individual close to the maximum benefit, the marginal claim is the claim that would have put the individual just over the annual maximum benefit if claimed during the first year of the estimation period.

[55]Section 4 explores heterogeneity using disaggregated procedure-level data. In contrast, the empirical model is estimated with claim-level data, where each claim aggregates all procedures done by the individual on a particular date (as discussed in Section 2). To investigate claim heterogeneity in the empirical model, claims are classified using measures analogous to the procedure-level measures investigated in Section 4. For example, the claim-level analog for the first measure is defined as follows. For each claim, I calculate the mean procedure T-statistic of all procedures done within that claim and take the quartiles of the distribution of this mean as a measure of claim-level "time-sensitivity". For the second claim-level measure of heterogeneity, I classify claims based on the procedures within that claim using a priority ordering: Restorative, Other, Preventive, Diagnostic. Though there are ex ante reasons why one would want to de-prioritize diagnostic procedures (as these are often done in combination with other procedures), alternative specifications reveal that the priority ordering of the remaining procedure types for this classification does not effect the basic results.

[56]The identification of these "potential delayers" relies on the definition of possible delay amounts (M). Because the definition of M implies the sequence of claims is fixed, this allows one to conclude that an individual did not consider delaying claims if the sum of his first-year dental spending plus the first claim in year 2 does not reach the maximum benefit. To illustrate this point, I suppose the opposite and show this leads us to a contradiction. In particular, suppose that such an individual did delay claims. Then, in the first year this individual must have received the event associated with the first claim in year 2, since the sequence of claims is fixed. However, since this claim would not have pushed him over the maximum benefit, it would have been suboptimal to delay this claim to year 2 and give up benefits, leading to a contradiction as this is inconsistent with the data. Thus, according to the model, only "potential delayers," as defined in the text, may have considered delaying claims.

[57]Among individuals over 18 years of age, 2.8% are potential delayers.

lowing for rich observable correlation between risk types and delay costs.[58] Additional analysis presented in Appendix C shows the model estimates and the counterfactual analysis are also robust to relaxing the second assumption.

5.3 Estimation

The main estimation focuses on two sets of parameters: the distribution of risk types, parameterized by $(\beta_\lambda, \sigma_\lambda)$, and the distribution delay costs, parameterized by $(\beta_\alpha, \delta_\alpha)$. These parameters are estimated taking the empirical cost intensity distribution, G_{a_i}, as given.

Let $\Theta = (\beta_\lambda, \sigma_\lambda, \beta_\alpha, \delta_\alpha)$. Define $claims_i$ as the sequence of claims observed for individual i over the two years, and define D_i as the division of this sequence of claims into those done in year 1 and those done in year 2. The individual's contribution to the likelihood of Θ can be written as follows:

$$l_i(\Theta|claims_i, D_i) = \int P(D_i|\lambda, \alpha, claims_i) dF(\lambda, \alpha|\Theta). \tag{10}$$

Because λ_i and α_i are known to the individual but not observed by the researcher, the likelihood must integrate over the distribution of these latent parameters, $F(\lambda, \alpha|\Theta)$.

Recall that λ_i and α_i are assumed conditionally independent in the model. Thus, the individual's contribution to the likelihood can be re-written as,

$$\begin{aligned} l_i(\Theta|claims_i, D_i) = \quad & p_\alpha(X_i, Z_{c_i}) \int P(D_i|\lambda, \alpha = 0, claims_i) dF(\lambda|\beta_\lambda, \sigma_\lambda, X_i) \\ + \quad & (1 - p_\alpha(X_i, Z_{c_i})) \int P(D_i|\lambda, \alpha = \infty, claims_i) dF(\lambda|\beta_\lambda, \sigma_\lambda, X_i). \end{aligned} \tag{11}$$

To describe the conditional probabilities in this likelihood, $P(D_i|\lambda, \alpha, claims_i)$ for $\alpha = 0$ or ∞, it is necessary to define a few additional objects. Let the sequence $claims_i$ be written as: $c_1, ..., c_{n_1}, c_{n_1+1}, ..., c_{N_1+N_2}$, where n_t is the number of dental events received in year t, and N_t is the number of claims in year t. The number of first-year events, n_1, is unobserved though we can bound this random variable by the number of first-year claims, $n_1 \geq N_1$. The decision model defines the probability of the observed division of claims D_i conditional on receiving n_1 first-year events, and the cost model defines the probability of receiving n_1 first-year events.

The probability of observing an individual's claims given that his delay cost is infinite is simply the probability that his division of events between the two years mimics his division of claims observed in the data,

$$P(D_i|\lambda, \alpha = \infty, claims_i) = P(n_1 = N_1|\lambda) P(n_2 = N_2|\lambda). \tag{12}$$

On the other hand, the probability of observing an individual's claims given that he can freely postpone claims ($\alpha_i = 0$) is simply the probability that the individual received any combination of events that would have led to the observed division of claims between the two years given optimal behavior and costless delay,

$$\mathrm{P}(D_i|\lambda, \alpha = 0, claims_i) = \sum_{k=N_1}^{N_1+N_2} P(D_i|\lambda, \alpha = 0, n_1 = k, claims_i) P(n_1 = k|\lambda) P(n_2 = N_1 + N_2 - k|\lambda). \tag{13}$$

The individual's maximization problem enters this expression through,

$$P(D_i|\lambda, \alpha = 0, n_1 = k, claims_i) = \mathrm{I}(D_i \text{ is optimal}|\lambda, \alpha = 0, n_1 = k, claims_i). \tag{14}$$

The term described in Equation 14 indicates when the division of claims in the data (D_i) can be rational-

[58] Any residual positive correlation between the delay cost and risk type would lead the model to underestimate the propensity to delay claims.

ized by the decision model given the parameters (λ, α), the sequence of claims ($claims_i$), and a particular division of events, indicating that the individual knows the first $n_1 = k$ events arrived in the first year while future events are uncertain. In Equation 13, this term is multiplied by the probability of observing the division of events according to the cost model, $P(n_1 = k|\lambda)P(n_2 = N_1 + N_2 - k|\lambda)$. This product of probabilities is calculated for all the potential combinations of events, and the sum of these terms is the right-hand side of Equation 13 above.

Aggregating across individuals, the complete likelihood for Θ can be written as follows:

$$L(\Theta|\text{claims,D}) = \prod_i l_i(\Theta|claims_i, D_i). \tag{15}$$

The method of Maximum Simulated Likelihood is used to estimate Θ. In practice, 25 draws per simulation per individual are used to approximate the integral over the distribution of λ while the independent distribution for α enters analytically as described in Equation 11.[59] To back out the optimal delay decision for an individual, it is necessary to evaluate the individual's second-year expected utility conditional on his risk type and the postponed treatment. In order to speed computation, I calculate this expectation using linear interpolation over a grid of λ and m values.

An input into the estimation described above is the coefficient of absolute risk aversion, which enters through the optimal delay decision. Using household coverage decisions among the firm's insurance plans, the baseline risk aversion parameter, 2.2×10^{-3}, is calibrated such that the predicted share of households in each plan matches the observed plan shares in the data, where this calibration is done under the assumption that households have a common risk aversion coefficient and households choose plans by maximizing the expected utility of the household without accounting for future coverage prospects.[60] This baseline value of risk aversion is similar in magnitude to prior risk aversion estimates over similar ranges of financial risk: deductible choice in auto insurance (Cohen and Einav (2007)) and deductible choice in homeowner's insurance (Sydnor (2010)). At the same time, prior studies investigating contexts that involve a much wider range of financial risk (such as studies focused on broader health insurance markets) have often found smaller risk aversion estimates. See Appendix C for a more detailed comparison of the baseline risk aversion parameter and prior estimates from other contexts. As is standard in the literature evaluating the welfare consequences of asymmetric information, I use individual decisions in this context to infer preferences used

[59] Because one would expect the risk type distribution to be bounded, λ is bounded above at 17 in the estimation, the maximum number of observed claims. This is a conservative bound for λ. If $\lambda = 17$, this means that an individual is so risky that he receives 17 dental events on average annually.

[60] The details of this calibration are as follows. Preliminary estimates for the risk type parameters $(\beta_\lambda, \sigma_\lambda)$ are obtained by maximizing the above likelihood assuming no delay. (As expected, the resulting risk type parameters from this partial estimation are almost identical to those obtained in full estimation because the risk aversion parameter only enters through the delay decision, and thus the identification of the risk type parameters $(\beta_\lambda, \sigma_\lambda)$ is largely independent of the risk aversion value.) Risk types, λ_i, are then simulated for each individual using these estimated parameters. The expected utility for each household h in each plan j, $\mathrm{E}u(-p_j - \Sigma_{i \in h} OOP_j(d_i)|\lambda_i \text{for } i \in h)$, is calculated assuming CARA utility, and household decisions are predicted assuming households choose the plan that maximizes this static utility. The baseline risk aversion parameter is then chosen such that the observed plan shares are equal to the predicted plan shares. There are two things worth nothing about this calibration. First, the calibration assumes that dental events received among individuals within the same household are conditionally independent, conditional on λ_i for these individuals. Note that this actually allows for rich correlation in claims among household members to the extent that household members have correlated λ_is. In the estimation, correlation in risk types among household members is captured to the extent that household members have correlated prior utilization (as variables on prior utilization are included in the specification of risk types in the empirical model). In practice, the estimation captures much of the correlation in claims among family members. (One easy way to see this is to compare the overall and within-household variation in the actual and the predicted number of dental claims. The overall standard deviation across the population in the actual and the predicted number of dental claims is 1.35 and 1.23, respectively. The mean within-household standard deviation in the actual and the predicted number of dental claims is 0.86 and 0.89, respectively.) Second, this calibration assumes household make static insurance decisions. Because saving/borrowing and claim delay may make Plan H less attractive in reality, this calibrated risk aversion parameter may underestimate the true degree of risk aversion in the population. Additional analysis presented in Appendix C illustrates that the model estimates and counterfactual results are robust to a wide range of alternative risk aversion values.

to conduct welfare analysis. However, if decisions are influenced by behavioral biases unrelated to classical risk aversion, the calibrated risk aversion may not reflect true preferences which could potentially impact the results. For instance, if the calibrated risk aversion overstates the true risk aversion in this population, the counterfactual analysis would understate the unraveling that strategic behavior induces, which would bias against the conclusions I draw from the baseline counterfactual analysis. To investigate the sensitivity of the results to the level of risk aversion, I present additional analysis in Appendix C demonstrating that the model estimates and main lessons from the counterfactuals are unchanged when considering a range of alternative risk aversion values.

5.4 Results

The parameter estimates from the model are displayed in Table 5, along with bootstrapped standard errors.[61] The parameters $(\beta_\lambda, \sigma_\lambda)$ describe the heterogeneity of risk types in the sample. Analysis presented in Appendix A illustrates that the implied annual spending fits the data quite well. Table 5 describes the parameter estimates from several specifications that consider heterogeneity across various individual and claim attributes. There are some notable patterns of observable variation in dental risk. Across the specifications, parameter estimates indicate that men have between 6% and 9% fewer dental events than women. The estimates reveal that dental risk is increasing in age, with the parameter estimates across the specifications indicating that those over age 50 have between 26% and 37% more events than those in the omitted category of individuals less than 19 years of age. Specifications (2) through (4) account for additional heterogeneity based on utilization in the year prior to the sample period. An indicator for having a prior dental visit and categorical variables on the level of spending in the year prior to the sample period are strongly associated with the frequency of dental events.

Table 5 columns (1) and (2) present estimates from the most parsimonious specifications in which the delay probability is constant across individuals (regardless of individual or procedure characteristics). Columns (3) and (4) present estimates from a specifications that allow for additional heterogeneity. The final row of Table 5 presents the fraction of individuals who optimally delay claims according to the model estimates, where this is obtained by integrating the estimated delay probability $p_\alpha(X_i, Z_{c_i})$ across the distribution of characteristics (X_i, Z_{c_i}) in the population close to the annual maximum benefit. Across the specifications, the estimates suggest that roughly 40% of individuals strategically delay claims when insurance incentives encourage them to do so.[62] Note that while the precise implied fraction of individuals delaying claims ranges from 36% and 45% across the specifications, these estimates are not statistically distinguishable from one another. As discussed further below, analysis in Appendix C demonstrates that the main results of the counterfactual analysis are robust to using delay propensity estimates spanning the range of the specifications displayed in Table 5.

In addition to estimating the overall delay rigidities, the model estimates displayed in columns (3) and (4) shed light on dimensions of heterogeneity across individuals and claims. Men are less likely to delay claims than women; the percent of men who delay claims when incentivized is 25 percentage points lower. Across the age distribution, those aged 50 and above have the highest propensity to delay claims. Interestingly, these demographic patterns of heterogeneity in delay line up with heterogeneity of reported

[61] The bootstraped standard errors are calculated using based on 100 bootstrap samples.

[62] In other words, the model estimates imply that approximately 40% of individuals close to the maximum benefit in the first of the adjacent years delayed claims as in the optimization problem described in equation 4 with costless delay. Within the context of the model, this means that roughly 40% of individuals strategically chose to delay any number of claims (between zero and the total number of dental problems received in the first year), while the remaining 60% of individuals did not optimize on this margin and simply treated all the problems that arose in the first year. So, the estimate of the delay probability can be interpreted as describing the fraction of individuals who respond to insurance incentives by re-timing claims when it is optimal to do so.

treatment delay in both dental and broader medical settings from the Medical Expenditure Panel Survey (MEPS). See Appendix A for a detailed description of this supplemental evidence on reported delay in the MEPS.

Specification (3) considers heterogeneity in the propensity to delay claims by the claim "time-sensitivity" measure defined earlier. The coefficient estimates suggest that claims that are classified as the least time-sensitive are associated with a delay propensity that is roughly 24 percentage points higher than those claims that are in the interquartile range of time-sensitivity; in addition, those claims are classified as the most time-sensitive are associated with a 9 percentage points lower delay propensity than those claims in the interquartile range. Specification (4) considers heterogeneity in the delay rigidities by claim category: Preventive, Diagnostic (excluded category), Restorative, Other (includes Endontics, Periodontics, Prosthodontics, and Oral Surgery). The coefficient estimates are consistent with the descriptive evidence in Figure 6. Relative to preventive and diagnostic claims, restorative care (such as cavity fillings) and other claims are more likely to be delayed to the following year. I present additional analysis in Appendix C which demonstrates that the qualitative patterns in the model estimates are robust to several additional alternative specifications.

The delay probability estimates inform us about the heterogeneity in the delay propensity at a point in time; the source of this heterogeneity (specifically whether it is tied to persistent individual characteristics or changing procedure characteristics) may influence the effectiveness of policies aimed at addressing this asymmetric information. The estimates in Table 5 columns (3) and (4) indicate that both individual characteristics and procedure characteristics are likely important determinants of delay rigidities. As described in Section 6, the counterfactual analysis investigates the impact of counterfactual policies allowing for delay rigidities to vary with both demographic characteristics and idiosyncratic factors (such as the urgency of dental problems that arise in a particular period).

6 Counterfactual Analysis

The parameter estimates are used to investigate the impact of strategic timing and the associated ex post adverse selection on insurance enrollment, insurer costs, and welfare. In the interest of exploring broader questions relating to the overall insurance market, the counterfactual analysis focuses on the impact of adverse selection outside the scope of the insurance options available within the firm. In particular, the counterfactual analysis investigates the market for comprehensive dental insurance, a product that in practice does not exist, to explore potential explanations for this market's unraveling. I compare insurance enrollment, insurer costs, and welfare in four different contracting scenarios: insurers contract on no individual information, insurers contract on ex ante risk types, insurers contract on pre-existing events, and insurers contract on both pre-existing events and risk types. This analysis allows me to isolate the impact of ex post adverse selection (resulting from the non-contractibility of pre-existing events), and compare the impact of this selection to the impact of traditional adverse risk selection (resulting from the non-contractibility of ex ante risk types). In addition to investigating the impact of underwriting various information, I investigate the impact of enrollment frequency restrictions on the viability of comprehensive dental insurance by comparing annual enrollment to less frequent enrollment. Although it probably goes without saying, it is important to keep in mind that the model estimates come from a particular insurance setting, so one should exercise the appropriate amount of caution in interpreting the counterfactual analysis based on these estimates.

Before continuing, it is worth highlighting how the welfare impact of strategic delay is analyzed below.

In principle, the strategic delay of treatment can lead to welfare losses through two channels: (i) ex post adverse selection and the associated market unraveling that this behavior induces, and (ii) inefficiencies directly associated with the delay of treatment in it of itself (e.g., pain associated with delaying treatment, the long run health consequences of delaying treatment, etc.). The aim of the welfare analysis in this paper is to focus on the first channel, ex post adverse selection generated by strategically delaying claims, putting aside any welfare cost created by the delay action in it of itself. Note that the specification of the delay cost as binary, either costless delay or infinitely costly delay, is in line with this aim. Estimating the welfare implications of the second channel is a very challenging issue (and one that is not taken on in this paper). For example, estimating a more flexible model of delay costs would not be sufficient for quantifying the welfare impact of the delay behavior in it of itself because observationally equivalent rigidities in this setting (e.g., lack of individual sophistication and pain) can have very different welfare consequences. Even beyond estimating the costs individuals internalize at the time of delay, there may also be long run health costs of delay that people do not internalize but that affect welfare in important ways (e.g., "internalities"). While the short panel of data and variation used in this paper does not allow me to estimate these long-run costs of delay, the effect of delaying care on long run health is certainly an interesting topic for future work. To the extent that the welfare costs of the second channel are important, the welfare analysis presented below understates the potential gains from alternative contracting environments that discourage treatment delay, such as allowing insurers to contract on pre-existing conditions and restricting the frequency of open enrollment periods.

6.1 Setup

Below, I explain how the model is extended to look at the broader counterfactuals of interest. The time horizon considered in each counterfactual scenario is ten years, where individuals decide whether to purchase full insurance or go without insurance in each period.[63] In each period, individuals maximize the sum of their current year utility and the future expected utility by deciding whether to purchase insurance, whether to delay claims, and optimally choosing savings/borrowing. For each period $t < T$, the individual solves the following recursive problem:

$$V_t(m_{t-1}, s_{t-1}|\mathcal{J}) = \tag{16}$$
$$\max_{Ins_t \in I_t}(1 - Ins_t)\mathrm{E}[\max_{s_t, m_t \in M_t} u(-d_t + m_t - m_{t-1} - s_t + s_{t-1} - \alpha_{i,t}\mathrm{I}(m_t > 0)) + V_{t+1}(m_t, s_t|\mathcal{J})]$$
$$+ (Ins_t)\mathrm{E}[\max_{s_t} u(-p - s_t + s_{t-1}) + V_{t+1}(0, s_t|\mathcal{J})],$$

where m_t is the amount of treatment delayed from period t to the following period, s_t is the savings in period t, and Ins_t indicates the individual has insurance in period t. In the final period, T, the individual's decision is simply:

$$V_T(m_{T-1}, s_{T-1}|\mathcal{J}) = \max_{Ins_T}(1 - Ins_T)E_{d_T}[u(-d_T - m_{T-1} + s_{T-1}] + (Ins_T)u(-p + s_{T-1}). \tag{17}$$

To make this problem computationally tractable, I assume individuals who choose to delay claims in period t, delay all their untreated events, $m_t \in M_t = \{0, d_t + m_{t-1}\}$. In addition to the state variables (m_{t-1}, s_{t-1}), the individual's maximization problem depends on the conditions of the counterfactual, represented by $\mathcal{J} = \{p, M_t, I_t, \ \forall t\}$, which summarizes the annual premium for insurance, the information insurers can use in contracting, and the frequency of insurance enrollment. For example, in counterfactuals that allow insurers to contract on pre-existing events, claim delay is effectively shut down (in which case

[63]The focus in these counterfactuals is on individual insurance coverage decisions (as opposed to household insurance decisions) because typically individuals can make separate insurance enrollment decisions in non-group insurance markets, without restrictions related to the coverage of other household members.

$M_t = \{0\} \; \forall t$). On the other hand, a counterfactual with open enrollment periods every five years would be associated with restricted insurance options represented by $I_{t \in \{1,6\}} = \{0, 1\}, \; I_{t \notin \{1,6\}} = \{Ins_{t-1}\}$. For each counterfactual situation examined, Table 6 details the corresponding restrictions on choice and insurer contracting within the maximization problem described above.

A key input in the individual's decision to delay treatment in period t is the cost of delay, $\alpha_{i,t}$. The counterfactuals are estimated under the assumption that individuals receive an independent draw from the conditional delay cost distribution in each year along with their draw of events (after their insurance coverage decision in that year):

$$\alpha_{i,t} = \begin{cases} 0 & \text{with probability } p_\alpha(X_i) \\ \infty & \text{with probability } 1 - p_\alpha(X_i), \end{cases} \tag{18}$$

where $p_\alpha(X_i)$ is the probability of costless delay conditional on the individual's demographic characteristics. The baseline counterfactual analysis uses the parameter estimates from Table 5 column (3), where the probability of costless delay used here ($p_\alpha(X_i)$) is found by integrating over claim characteristics using the empirical distribution of claims conditional on the individual's age category ($p_\alpha(X_i) = \int_z p_\alpha(X_i, z) F(z|a_i)$)[64] With this specification, an individual's draw from the delay cost distribution in a given period depends on both demographic characteristics (as the delay probability is a function of X_i) and idiosyncratic factors such as the urgency of procedures (as evidenced by the fact that the delay cost is probabilistic even after controlling for demographics).

This set up for the counterfactual analysis implies individuals could choose to delay claims for multiple years so long as the delay cost is zero in each relevant year. In other words, this means that individuals can potentially delay treating events for multiple years, but there is a $(1 - p_\alpha(X_i))$% chance in any given year that the individuals' accumulation of events will become urgent and will need to be treated immediately.[65] The baseline counterfactuals assume that individuals face no other frictions related to switching plans beyond the delay frictions captured by $\alpha_{i,t}$.[66] Additional analysis presented in Appendix C illustrates that the results are robust to allowing for additional switching frictions above and beyond the delay frictions estimated in the empirical model.

Note that this framework for the counterfactual analysis allows for the investigation of two sources of asymmetric information: static adverse selection arising from variation in ex ante risk types and dynamic ex post adverse selection arising from claim delay. The counterfactuals abstract from other potential sources of dynamic asymmetric information, such as that arising from evolving long-run risk.[67]

Constraining premiums to be constant across the ten years of available insurance, I solve for the equilibrium in each counterfactual situation numerically assuming insurers break even. Details of this compu-

[64] The age categories used here are: less than 19 years of age, 19-30 years of age, 31-40 years of age, 41-50 years of age, and over 50 years of age.

[65] This effectively means the ex ante probability of delaying claims for multiple consecutive years decreases in an exponential way with the number of years (the probability of being able to delay for n consecutive years is $p_\alpha(X_i)^n$).

[66] Under some additional assumptions, the estimated delay cost parameters are sufficient to capture rigidities related to both strategic treatment delay and ex post adverse selection. Some typical explanations of suboptimal plan switching are that individuals lack information about plan details or lack the cognitive sophistication necessary to navigate complex incentives. These explanations and many others used to explain switching costs are also potentially important explanations for the suboptimal delay of claims. If the frictions that prevent one from optimally switching plans also prevent one from optimally delaying treatment, the estimated fraction of individuals who optimally delay claims is a lower bound on the fraction of individuals who, given the choice, would have switched plans if it were optimal to do so. Thus, under some assumptions, the estimated delay cost distribution can be viewed as a sufficient statistic that summarizes the frictions relevant to evaluating the impact of policies addressing ex post adverse selection.

[67] The counterfactuals consider a ten year period, over which it may be reasonable to assume that dental risk types are stable. While in principle it would be possible to incorporate long-run transitions in dental health risk types into the counterfactuals, such an analysis would be quite speculative in practice since the data covers a short time period and thus cannot inform us of the long-run evolution of risk types.

tation are given in Appendix B. In addition to considering counterfactuals within an unsubsidized market, I consider counterfactuals under various premium subsidies to allow us to contextualize the relative effectiveness of hypothetical interventions like more extensive underwriting of risk or reducing the frequency of open enrollment periods relative to premium subsidies, a common intervention aimed at overcoming adverse selection in health insurance contexts.

In the baseline counterfactuals, I assume the insurance market is perfectly competitive, and I consider the case of no administrative loading meaning the first best outcome with risk averse agents is full insurance. Note that the counterfactuals abstract from any potential benefits from choice as modeled individuals have homogeneous and stable preferences, and the insurance market is assumed to be competitive.[68] Thus, the focus of the counterfactual analysis is to shed light on the costs of choice relative to the first best of full insurance.

In each counterfactual scenario, I calculate the equilibrium insurance enrollment, annual mean insured costs, and annual per-capita welfare. Because the counterfactuals abstract from insurer market power, I measure welfare as consumer surplus net of any applicable subsidy costs. Consumer surplus is calculated using a certainty equivalent measure. Let the certainty equivalent in a counterfactual scenario $\mathcal{J}$, $CE_i^{\mathcal{J}}$, be defined as

$$\sum_{t=1}^{T} u(CE_i^{\mathcal{J}}) = V_{i,1}(0,0|\mathcal{J}). \tag{19}$$

In words, this expression says the individual is ex ante indifferent between either accepting the certainty equivalent in each period or accepting the expected value of his utility over the entire time horizon solving the dynamic optimization problem above under counterfactual conditions $\mathcal{J}$. Consumer surplus is then measured as the difference between the certainty equivalent in the counterfactual scenario relative to the certainty equivalent in a setting when no insurance is available, $CE_i^{no\ ins}$. Welfare is defined as the annual per-capita consumers surplus netting out the cost of providing a subsidy (if applicable),

$$\text{Welfare}^{\mathcal{J}} = \frac{1}{n}\sum_{i=1}^{n}(CE_i^{\mathcal{J}} - CE_i^{no\ ins}) - (1+\gamma)\text{AnnualPerCapitaSubsidyCost}^{\mathcal{J}}. \tag{20}$$

In counterfactuals that consider a premium subsidy, this measure nets out the per-capita cost of the subsidy adjusted to account for the shadow cost of raising public funds (γ), where $\gamma = 0.25$ is used in the baseline calculations. In other words, the counterfactuals consider a hypothetical subsidy that comes from the government (as would be the case in a tax subsidy), where there is some deadweight loss associated with raising money to cover this subsidy, captured by γ. Note that the inclusion of this shadow cost of public funds means that an increase in the subsidy may either help or hurt welfare, depending on whether the welfare gains from increased insurance enrollment outweigh the welfare costs of raising money to fund the subsidy.

It is worth noting a few key differences in the welfare calculation within this dynamic insurance model relative to typical static insurance models. First, the individual's dynamic optimization problem means that an individual's *ex ante* welfare depends on the option to buy more insurance coverage in the future in the case that he has delayed treatments even if he doesn't *ex post* end up buying insurance. Thus, in contrast to a static model, the *ex ante* welfare benefit of insurance is not limited to those who *ex post* bought insurance. Second, assumptions on liquidity can play a key role in welfare calculations in insurance markets. While

[68]Following most of the empirical literature on insurance, the counterfactuals consider a competitive insurance market as a benchmark.

studies that focus on static insurance markets typically assume individuals are hand-to-mouth, the baseline counterfactuals here are estimated under the assumption that individuals can save and borrow quite freely within the time horizon of the counterfactuals. This assumption captures two important features of this setting: (i) individuals usually have quite a bit of liquidity to smooth small to moderate shocks, and (ii) individuals may be relatively more liquidity constrained in some periods (where the model captures this variation in effective liquidity through variation in the length of the remaining time horizon over which individuals can smooth future shocks).

6.2 Information Contractibility

The first set of counterfactuals explores the market for comprehensive dental insurance in different contracting settings. Specifically, four different contracting scenarios are considered: insurers contract on no information (where both traditional adverse selection and ex post adverse selection are present), insurers contract on risk type but not pre-existing events (only ex post adverse selection is present), insurers contract on pre-existing events but not risk type (only traditional adverse selection is present), and insurers contract on both risk type and pre-existing events (symmetric information with no selection).

For each counterfactual scenario, Table 7 displays the equilibrium percent of insured individual-years, the annual mean insured cost, and annual per-capita welfare both in terms of dollars and in terms of percent relative to the first best of symmetric information. The columns of the table describe the contracting environment while the rows describe the outcome variable of interest. The symmetric equilibrium first best (reflected in column 4 in the ”No Subsidy” rows) is associated with annual per-capita welfare of $79. In other words, the “maximum welfare at stake” in this setting is $79 per-capita annually. While the maximum annual per-capita welfare in this setting is modest in absolute terms, it is substantial in relative terms as it represents 33% of annual per-capita dental expenditures.[69] To contextualize the welfare estimates obtained in the various counterfactuals, Table 7 presents welfare estimates both in terms of annual dollars per-capita and as a percent relative to the first best welfare obtained in the case of symmetric information. As will become clear through the discussion of the results below, the welfare costs of asymmetric information in this setting, though modest in absolute terms, represent substantial losses relative to the maximum welfare at stake in this setting (the welfare under the first best of symmetric information).

When insurers cannot contract on all available consumer information, the results in Table 7 suggest that the market suffers from severe adverse selection. First consider an unsubsidized market. When insurers can contract on no information, we see that the market largely unravels with only 2.0% of individual-years insured, for a mean cost of $1,809 and an associated annual per-capita welfare of $28, which represents a 64% reduction in welfare compared to the maximum potential welfare under the first best. It is worth noting that in this scenario where insurers contract on no information, we see that some non-negligible fraction of the first best ex ante welfare (36%) is achieved even though only 2% of individual-years are insured ex post. In other words, the market unraveling is more dramatic in terms of insurance enrollment compared to welfare. There are two key reasons for this. First, there is a lot of heterogeneity in the willingness-to-pay for insurance both across individuals and within-individual, across years. (Note that the willingness-to-pay for insurance varies within-individual at different points in time based on variation in delayed treatments and variation in effective liquidity because of variation in the length of the remaining time horizon over which to smooth future shocks.) Second, as noted above, in this dynamic model of insurance, individuals

[69] This may be more appropriately thought of as a lower bound on the first-best welfare in this context since, as described above, the welfare analysis assumes individuals have access to a lot of liquidity. If instead, individuals are assumed to be hand-to-mouth, the value of insurance in this setting would be considerably larger than under the baseline assumptions.

place *ex ante* value on the opportunity to buy insurance coverage in the future in the case that they have delayed treatment even if they do not end up buying insurance *ex post*.

The counterfactuals also allow us to isolate the impact of ex post selection (column 2) and contrast this with the impact of traditional adverse selection in isolation (column 3). The results indicate that unraveling is more dramatic under ex post adverse selection than under traditional adverse selection on ex ante risk types. Once again consider an unsubsidized market. Starting from a baseline of insurers contracting on no information (column 1), eliminating traditional adverse selection by contracting on risk types slightly more than doubles insurance enrollment and increases annual per-capita welfare by $14, or 50% of the baseline value. In contrast, eliminating ex post adverse selection by contracting on pre-existing events leads to a six-fold increase in insurance enrollment and a 76% increase in welfare relative to the baseline value. Relative to the equilibrium with only traditional adverse selection (column 3), the equilibrium under ex post adverse selection alone (column 2) is associated with 61% lower insurance enrollment and 15% lower annual per-capita welfare.

The remainder of the results in Table 7 contextualize the effect of hypothetically underwriting more information (e.g. ex ante risk types or pre-existing events) relative to the effect of premium subsidies, a common intervention aimed at improving welfare in insurance markets. Under symmetric information, subsidies only harm welfare by making costly transfers to people who are inframarginal. When there is asymmetric information, premium subsidies can increase welfare. Consider the setting when insurers can contract on no information. In this setting, a 25% premium subsidy only yields 10.3% of individual-years insured and 63% of the per-capita welfare that would be obtained under the symmetric information first best. Starting from a setting where insurers can contract on no information, hypothetically allowing insurers to contract on pre-existing events yields insurance enrollment increases and welfare improvements that exceed those associated with the introduction of a 25% premium subsidy.

6.3 Enrollment Period Frequency Restrictions

Limiting the frequency of open enrollment periods may limit the adverse selection induced by strategic timing. I investigate four scenarios with different open enrollment period frequencies: annual enrollment, enrollment once every two years, enrollment once every five years and lifetime enrollment. In the case of annual enrollment, for example, an individual decides on a contract just before the start of each year. In the case of enrollment every two years, an individual chooses a contract for the first two years just before the start of year 1, and so on. To isolate the impact of enrollment frequency on ex post adverse selection, I calculate an equilibrium in each scenario assuming the insurer contracts on risk types but not on pre-existing events. Note that in this setting the first best is achieved under one single lifetime open enrollment period.[70]

For each counterfactual scenario, Table 8 displays the equilibrium percent of insured individual-years, the annual mean insured cost, and annual per-capita welfare both in terms of dollars and in terms of percent relative to the first best of lifetime contracts. Under annual open enrollment periods, the unsubsidized market suffers from severe adverse selection with only 4.8% of individual-years insured and annual per-capita welfare of $43, or 54% of the per-capita welfare under the first best. Comparing the first column

[70] As Cochrane (1995) points out, there are reasons why lifetime contracts might not achieve the first best. For example, lifetime contracts may not be enforceable, or lifetime contracts may not be the first best if individual preferences for insurance evolve over time. In the counterfactual analysis, none of these reasons apply as contracts are assumed to be enforceable and risk-types are assumed to be fixed over time. In this way, the counterfactual results can be thought of as capturing the costs of choice flexibility relative to lifetime contracts, while the benefits (deriving from better matching to evolving preferences or easier enforcement) are not considered but could in principle be used to weigh against these estimated costs.

to the remaining columns, one can see that reducing the frequency of enrollment can encourage comprehensive coverage and improve welfare. Relative to annual contracts, two-year commitment to contracts causes insurance enrollment to more than double and increases annual per-capita welfare by $13. While this welfare gain is modest in absolute terms, this gain is substantial in relative terms, representing 17% of the maximum welfare at stake in this setting and amounting to a 31% increase over the baseline welfare under annual enrollment. Further reducing the frequency of open enrollment periods to once every five years would increase insurance enrollment by nearly eight-fold and increase welfare by 68% relative to annual enrollment. Again, this welfare gain, while modest in absolute terms, is large in relative terms representing 36% of the maximum welfare at stake in this context.

Interestingly, this analysis reveals that much of the welfare benefit of lifetime contracts can be achieved by instituting an open enrollment period once every five years. While the analysis abstracts from some potential benefits short-term contracting in reality (that could arise if individual risk preferences evolve over time or commitment issues made lifetime contracts unenforceable), the counterfactual results suggest that an open enrollment period every five years could provide individuals quite a bit of flexibility in choice and still avoid much of the unraveling that would arise because of ex post adverse selection.

The counterfactuals also allow one to compare the effectiveness of lengthening commitment to contracts in overcoming market unraveling relative to premium subsidies. Reducing the frequency of open enrollment periods to once every two years increases insurance enrollment and welfare by more than the introduction of a 25% premium subsidy. Even less frequent open enrollment periods further increase insurance enrollment and welfare.

6.4 Robustness

To ensure that the model estimates and main lessons from the counterfactual analysis are not too sensitive to the assumptions used to get the baseline estimates, the model is re-estimated under alternative assumptions on the relevant time horizon, risk aversion value, plan selection, and plan switching frictions. The details of these alternative specifications are discussed in Appendix C and the results are reported in Appendix Table C1. The implied mean probability of strategic delay ranges from 0.29 to 0.54 across the alternative specifications. Using these alternative estimates and estimates from alternative specifications in Table 5, the counterfactual analysis is repeated, and the results are qualitatively unchanged. See Appendix C for a more detailed discussion of the robustness analysis.

7 Conclusion

The strategic timing of claims can cause inefficiencies in insurance markets. Using claim-level data, I find clear patterns that suggest individuals strategically delay dental treatment when insurance incentives encourage them to do so. I then develop and estimate a model that explicitly links this strategic delay of claims to the adverse selection it creates. The counterfactual analysis reveals that strategic delay of treatment and the associated ex post adverse selection is severe in this setting and causes more unraveling than traditional adverse selection based on heterogeneity in ex ante risk types. Overall, my results indicate that severe adverse selection may arise in settings where insured costs are elastic with respect to timing and the timing of the underlying risk is not contractible. In addition, the results suggest that decreasing the frequency of open enrollment periods (thereby effectively lengthening the commitment to contracts) can improve welfare in contexts where the timing of uncertainty is not contractible.

While it is important to emphasize that the specific estimates in this paper should not be interpreted as applying to broader health insurance settings, there are several pieces of evidence which suggest that this type of asymmetric information may play an important role in broader health insurance contexts. First, findings from other studies suggest that many health care treatments are not particularly urgent and individuals strategically delay health care treatments when they anticipate better insurance coverage in the future (e.g., Card, Dobkin and Maestas (2008), Ellimoottil et al. (2014), Manning et al. (1986), Einav, Finkelstein and Schrimpf (2015)).[71] Second, the Affordable Care Act (ACA) potentially exposes broader health insurance markets to a large degree strategic timing and ex post adverse selection because of three features: the elimination of excessive waiting periods, the elimination of medical underwriting, and the creation of insurance exchanges in which people have access to a wide range of vertically and horizontally differentiated coverage options.[72] My results suggest that instituting open-enrollment periods within the newly created insurance exchanges may be an key step to combat adverse selection. The optimal frequency of such enrollment periods in health insurance (annual, biannual, etc.) is an important topic for future research.

Another naturally related topic for future work to explore is the bundling of treatments for the purpose of insurance coverage. While the main purpose of health insurance is to allow individuals to pool risk, much of this risk pooling can break down when individuals can re-evaluate insurance decisions frequently. My analysis in this paper illustrates that this breakdown can be particularly dramatic when many insured treatments are not urgent. Within health care, treatments span the spectrum of urgency. While treatment for a heart attack is extremely urgent, knee replacements can often be delayed for years. Still, all health care spending, urgent or not, is typically covered by the same insurance product, with the notable exceptions being the historical exclusion of dental and vision services. One could imagine many other ways to bundle (or unbundle) health care services for the purpose of insurance. Ex ante, it is not obvious what an optimal grouping of services would look like from an efficiency perspective. Should urgent and less urgent types of treatments be separately insured? If so, how should the design of these insurance products differ? Alternatively, if the risk of urgent care is large enough, can we obtain more efficient insurance for all treatments by bundling all care together for the purpose of insurance? The Affordable Care Act mandates that health insurance policies cover certain procedures, some of which are not typically urgent (e.g., preventive cancer screenings, dental care for children, etc). The welfare impact of mandating that relatively non-urgent procedures be insured jointly with urgent procedures is a important and policy-relevant topic for future work to explore.

[71] For example, Card, Dobkin and Maestas (2008) show that near-elderly individuals delay both inpatient and outpatient care until just after they turn 65, at which point they are eligible for Medicare. Using a difference-in-difference design to study Massachusetts health care reform, Ellimoottil et al. (2014) show the increase in insurance coverage caused by the reform lead to a short-run rise in discretionary surgeries (such as knee replacements, hip replacements, back surgery), with this effect concentrated among individuals most likely to have been uninsured prior to the reform. Manning et al. (1986) show that the responsiveness of spending to insurance coverage was larger in the first year of the Rand Health Insurance Experiment than in subsequent years, and this elevated first-year responsiveness was present in both dental and broader health care spending. Einav, Finkelstein and Schrimpf (2015) show evidence consistent with the strategic delay of prescription drug utilization in response to insurance terms that reset at the start of the calendar year.

[72] Massachusetts health care reform provides an interesting case study highlighting the potential impact of strategic treatment delay and subsequent ex post adverse selection within broader health insurance. Like the ACA, the Massachusetts reform eliminated excessive waiting periods and eliminated most medical underwriting. In addition, when the Massachusetts law was first put into effect, there was no specified open enrollment period, which meant that individuals could effectively buy and drop coverage whenever they would like. Some have documented that there was a lot of apparent strategic selection during the first few years after the reform (perhaps driven by the strategic timing of treatments) in which individuals sign up for insurance coverage for a month or two, run up abnormally large health care bills, and then drop coverage (see, for example, Lazar (2010), Welch and Giesa (2010)). In response, lawmakers in Massachusetts passed a law a few years later restricting enrollment to two annual periods in 2011 and just one annual period starting in 2012 (Massachusetts Legislation Chapter 288, August 10 2010). Perhaps because of this lesson learned in Massachusetts, many of the subsequent health insurance exchanges under the ACA have implemented open enrollment periods restricting how often individuals can re-evaluate their insurance decisions.

References

Abaluck, Jason, and Jonathan Gruber. 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review*, 101(4): 1180–1210.

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, 84(3): 488–500.

Aron-Dine, Aviva, Liran Einav, and Amy Finkelstein. 2013. "The RAND Health Insurance Experiment, Three Decades Later." *Journal of Economic Perspectives*, 27(1): 197–222.

Arrow, Kenneth. 1975. "Two Notes on Inferring Long Run Behavior from Social Experiments." *RAND Corporation Publication*, P-5546.

Bajari, Patrick, Han Hong, and Ahmed Khwaja. 2006. "Moral Hazard, Adverse Selection and Health Expenditures: A Semiparametric Analysis." National Bureau of Economic Research Working Paper 12445.

Buchmueller, Thomas C., and Paul J. Feldstein. 1997. "The effect of price on switching among health plans." *Journal of Health Economics*, 16(2): 231 – 247.

Bundorf, M. Kate, Jonathan D. Levin, and Neale Mahoney. 2012. "Pricing and Welfare in Health Plan Choice." *American Economic Review*, 102(7): 3214–3248.

Burman, Leonard E, and William C Randolph. 1994. "Measuring Permanent Responses to Capital-Gains Tax Changes in Panel Data." *American Economic Review*, 84(4): 794–809.

Card, David, Carlos Dobkin, and Nicole Maestas. 2008. "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare." *American Economic Review*, 98(5): 2242–58.

Cardon, James H, and Mark H Showalter. 2001. "An examination of flexible spending accounts." *Journal of Health Economics*, 20(6): 935 – 954.

Carlin, Caroline, and Robert J. Town. 2010. "Adverse Selection: The Dog that Didn't Bite." *miemo*.

Chetty, Raj. 2008. "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy*, 116(2): 173–234.

Chiappori, Pierre-Andre, and Bernard Salanie. 2000. "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy*, 108(1): pp. 56–78.

Chiappori, Pierre-André, Bruno Jullien, Bernard Salanié, and Francois Salanie. 2006. "Asymmetric information in insurance: General testable implications." *The RAND Journal of Economics*, 37(4): 783–798.

Cochrane, John H. 1995. "Time-Consistent Health Insurance." *Journal of Political Economy*, 103(3): 445–73.

Cohen, Alma. 2005. "Asymmetric Information and Learning: Evidence from the Automobile Insurance Market." *The Review of Economics and Statistics*, 87(2): pp. 197–207.

Cohen, Alma, and Liran Einav. 2007. "Estimating Risk Preferences from Deductible Choice." *American Economic Review*, 97(3): 745–788.

Cutler, David, Bryan Lincoln, and Richard Zeckhauser. 2010. "Selection stories: Understanding movement across health plans." *Journal of Health Economics*, 29(6): 821 – 838.

Cutler, David M., and Sarah J. Reber. 1998. "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection." *The Quarterly Journal of Economics*, 113(2): pp. 433–466.

Einav, Liran, Amy Finkelstein, and Jonathan Levin. 2010. "Beyond Testing: Empirical Models of Insurance Markets." *Annual Review of Economics*, 2: 311–336.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen. 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *Quarterly Journal of Economics*, 125(3): 877–921.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2010. "Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence From the U.K. Annuity Market." *Econometrica*, 78(3): 1031–1092.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2015. "The Response of Drug Expenditure to Non-Linear Contract Design: Evidence from Medicare Part D." *forthcoming in the Quarterly Journal of Economics*.

Einav, Liran, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen. 2013. "Selection on Moral Hazard in Health Insurance." *American Economic Review*, 103(1): 178–219.

Ellimoottil, Chandy, Sarah Miller, John Z Ayanian, and David C Miller. 2014. "Effect of Insurance Expansion on Utilization of Inpatient Surgery." *JAMA surgery*, 149(8): 829–836.

Fang, Hanming, and Alessandro Gavazza. 2011. "Dynamic Inefficiencies in an Employment-Based Health Insurance System: Theory and Evidence." *American Economic Review*, 101(7): 3047–77.

Finkelstein, Amy, and James Poterba. 2004. "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market." *Journal of Political Economy*, 112(1): 183–208.

Finkelstein, Amy, Kathleen McGarry, and Amir Sufi. 2005. "Dynamic Inefficiencies in Insurance Markets: Evidence from Long-Term Care Insurance." *American Economic Review*, 95(2): 224–228.

Goda, Gopi Shah, and Colleen Flaherty Manchester. 2010. "Incorporating Employee Heterogeneity into Default Rules for Retirement Plan Selection." National Bureau of Economic Research Working Paper 16099.

Golosov, Mikhail, and Aleh Tsyvinski. 2006. "Designing Optimal Disability Insurance: A Case for Asset Testing." *Journal of Political Economy*, 114(2): 257–279.

Goolsbee, Austan. 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy*, 108(2): 352–378.

Gruber, Jonathan. 2000. "Disability Insurance Benefits and Labor Supply." *Journal of Political Economy*, 108(6): 1162–1183.

Guay, Albert H. 2006. "The differences between dental and medical care: implications for dental benefit plan design." *Journal of the American Dental Association*, 137 (6): 801–806.

Handel, Benjamin, Igal Hendel, and Michael Whinston. 2015. "Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk." *Econometrica*, 83(4): 1261–1313.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review*, 103(7): 2643–82.

Hendel, Igal, and Alessandro Lizzeri. 2003. "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance." *The Quarterly Journal of Economics*, 118(1): pp. 299–327.

Hendren, Nathaniel. 2013. "Private Information and Insurance Rejections." *Econometrica*, 81(5): 1713–1762.

Jeffcoat, Marjorie K. 2004. "A double-edged sword: insurance A to Z." *Journal of the American Dental Association*, 135(7): 842–844.

Katz, Lawrence F., and Bruce D. Meyer. 1990. "The impact of the potential duration of unemployment benefits on the duration of unemployment." *Journal of Public Economics*, 41(1): 45 – 72.

Kowalski, Amanda E. 2010. "Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care." *NBER Working Paper*.

Lazar, Kay. 2010. "Short-term Insurance Buyers Drive up Cost in Massachusetts: Lawmakers Look to Close Loopholes." *Boston Globe*. Printed June 30, 2010.

Lustig, Joshua. 2010. "Measuring Welfare Losses from Adverse Selection and Imperfect Competition in Privatized Medi- care." *mimeo*.

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *The Quarterly Journal of Economics*, 116(4): 1149–1187.

Manning, Willard G., Howard L. Bailit, Bernadette Benjamin, and Joseph P. Newhouse. 1986. "The Demand for Dental Care: Evidence from a Randomized Trial in Health Insurance." *Rand publication series*.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *The American Economic Review*, 77(3): pp. 251–277.

Metcalf, Charles E. 1973. "Making Inferences from Controlled Income Maintenance Experiments." *The American Economic Review*, 63(3): pp. 478–483.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *The Quarterly Journal of Economics*, 90(4): 629–649.

Rust, John, and Christopher Phelan. 1997. "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets." *Econometrica*, 65(4): pp. 781–831.

Stiglitz, Joseph E. 1985. "The General Theory of Tax Avoidance." *National Tax Journal*, 38: 325–338.

Strombom, Bruce A, Thomas C Buchmueller, and Paul J Feldstein. 2002. "Switching costs, price sensitivity and health plan choice." *Journal of Health Economics*, 21(1): 89 – 116. Health Plan Choice.

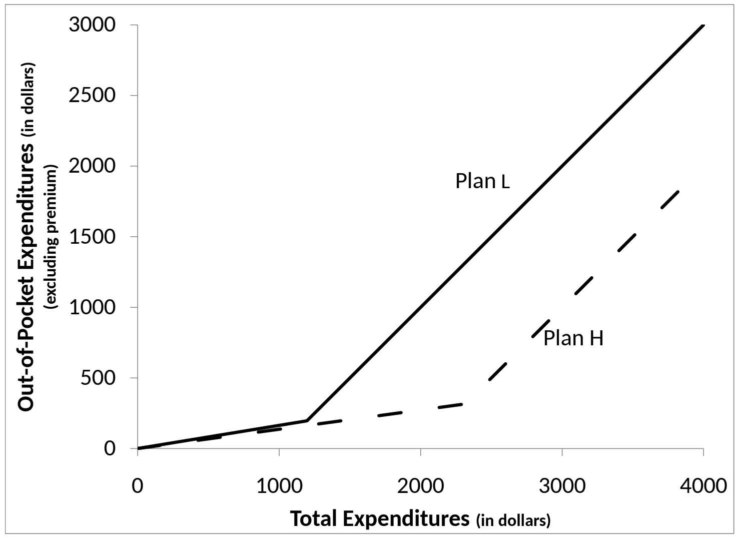
Sydnor, Justin. 2010. "(Over)insuring Modest Risks." *American Economic Journal: Applied Economics*, 2(4): 177–99.

Tchernis, Rusty, Sharon-Lise T. Normand, Juliana Pakes, Peter Gaccione, and Joseph P. Newhouse. 2006. "Selection and Plan Switching Behavior." *Inquiry*, 43(1): 10–22.

Wall, Thomas, Kamyar Nasseh, and Marko Vujicic. 2014. "U.S. Dental Spending Remains Flat Through 2012." *Health Policy Institute, American Dental Association Research Brief*.

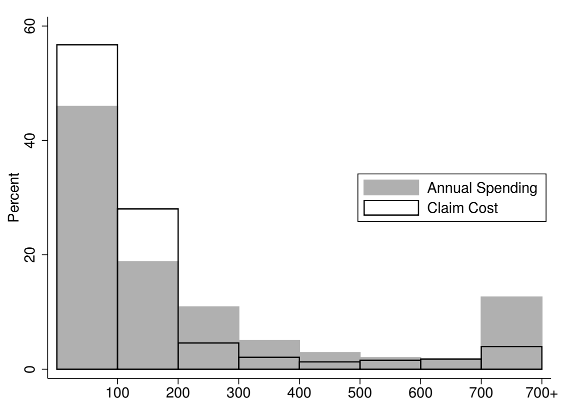
Welch, Dianna K., and Kurt Giesa. 2010. "Analysis of Individual Health Coverage in Massachusetts Before and After the July 1, 2007 Merger of the Small Group and Nongroup Health Insurance Market." *Report for Massachusetts Division of Insurance*.

Figure 1: Individual Out-of-Pocket Spending as a Function of Total Spending by Plan

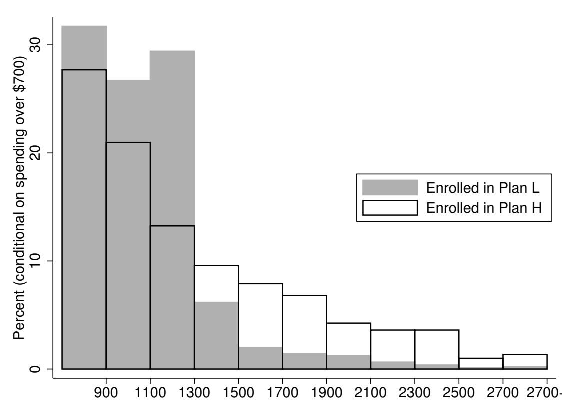


Notes: The above is a plot of annual out-of-pocket spending (excluding premiums) per individual as a function of total annual spending by plan, using the unconditional average coinsurance for the baseline sample below the annual individual maximum benefit. The details of the average coinsurance calculation are described in Table 2. Based on these out-of-pocket cost functions with the average coinsurance rate across all ages, it would take approximately $1,184 of dental spending for a single coverage employee facing $65 Plan H premium to be indifferent ex post between the two plans. The kinks in this figure are at $1,133 (=$1,000/0.883) for Plan L, and $2,225 (=$2,000/0.899) for Plan H (these are the levels of total spending that correspond to exhausting the $1,000 and $2,000 maximum benefits, respectively, given the coinsurance paid by the insurer under the maximum benefit, 88.3% and 89.9% respectively). Above these values of total dental spending, the individual pays the full cost of care.

Figure 2: Annual Individual Total Expenditures and Claim Cost for the Baseline Sample



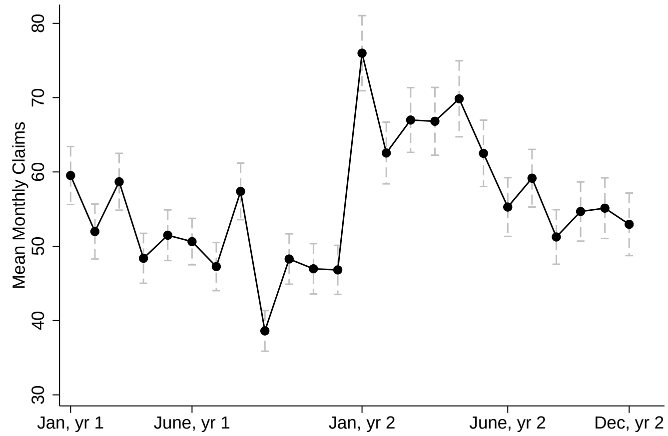
(a) Annual expenditures and claim-level costs



(b) Annual expenditures exceeding $700

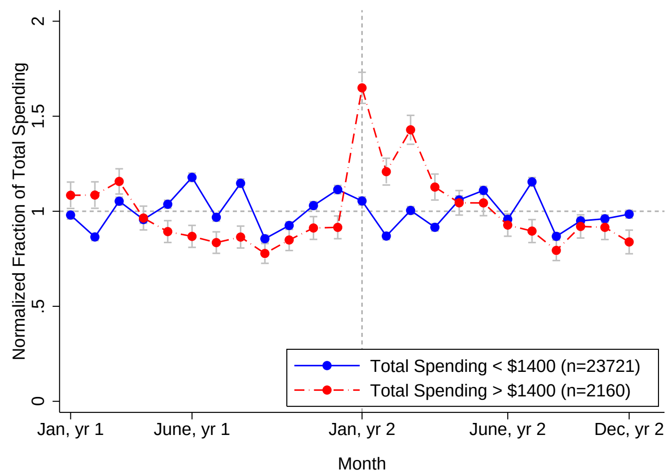
Notes: Panel (a) above displays the distribution of claim cost (costs within one day) and annual total dental spending for the baseline sample. Observations are pooled across years to create this histogram. Thirty-seven percent of individual-years are exactly at zero dollars of annual dental spending. Panel (b) shows the distribution of annual dental spending for the right tail of the baseline sample by plan. The percent here is calculated based on the percent of individual-years out of all individual-years on the given plan conditional on having spending exceeding $700 in that year.

Figure 3: Monthly Claims Among Those Switching to More Generous Coverage



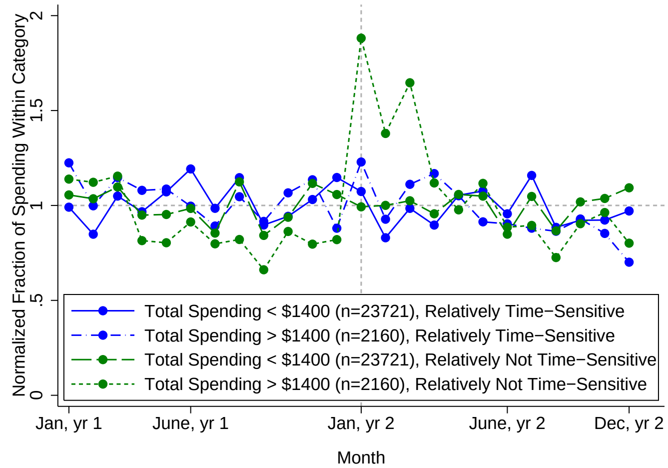
Notes: The figure displays mean monthly dental claims applying the cost-sharing rules of Plan H. The sample is limited to households that switch to more generous coverage within the sample used in the regressions in Table 4, and the sample is further restricted to the data in the year just before and the year just after this switch. In total, the number of household-month observations used to construct this figure is 42,528 observations. The figure also displays $+/-$ the standard error. Appendix A reports the analogous regression results.

Figure 4: Strategic Claim Delay



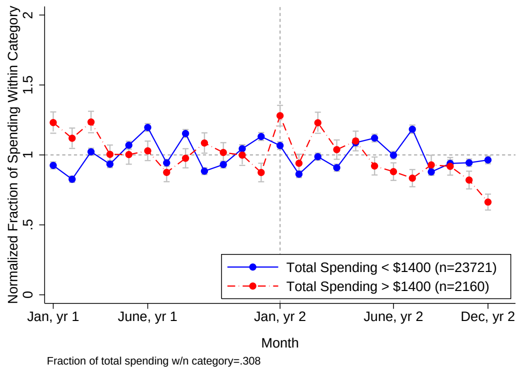
Notes: The figure displays the average normalized monthly individual dental expenditures as a fraction of the total expenditures across two adjacent years, the average value of (monthly spending $\times$ 24/total spending). The figure also displays $+/-$ the standard error. The sample used to create this figure is restricted to those who were insured with the company for two adjacent years (2004-2005, 2005-2006, or 2006-2007) and were enrolled in Plan L (which has a maximum benefit of $1,000) in the first of these adjacent years. Individuals with no spending across the two years are dropped. This series is displayed separately for those with overall expenditures less than $1,400 (those who probably did not have the incentive to delay claims) and those with overall expenditures exceeding $1,400 (those who were more likely to have the incentive to delay claims). For those with high overall expenditures, one can see a spike beginning in January of the second year. Appendix A contains alternative figures with different cutoff values to identify incentivized individuals and alternative figures that control for year 2 plan choice. The qualitative patterns remain the same in these alternative figures.

Figure 5: Strategic Claim Delay: heterogeneity by procedure time-sensitivity measure

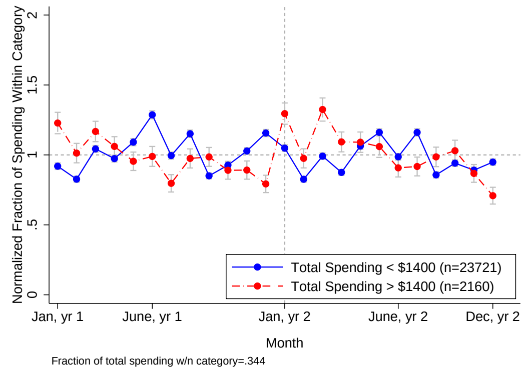


Notes: For each category of procedures, this figure displays the average normalized monthly individual dental expenditures on procedures in that category as a fraction of the total expenditures in that category across two adjacent years, the average value of (monthly category spending $\times$ 24/total category spending). The sample used to create this figure is restricted to those who were insured with the company for two adjacent years (2004-2005, 2005-2006, or 2006-2007) and were enrolled in Plan L (which has a maximum benefit of $1,000) in the first of these adjacent years. Individuals with no spending across the two years are dropped. This series is displayed separately for those with overall expenditures (across all categories) less than $1,400 (those who probably did not have the incentive to delay claims) and those with overall expenditures (across all categories) exceeding $1,400 (those who were more likely to have the incentive to delay claims). For those with high overall expenditures, one can see a spike among relatively not time-sensitive procedures beginning in January of the second year. The categorization of procedures for the purpose of this figure is described in detail in Section 4.

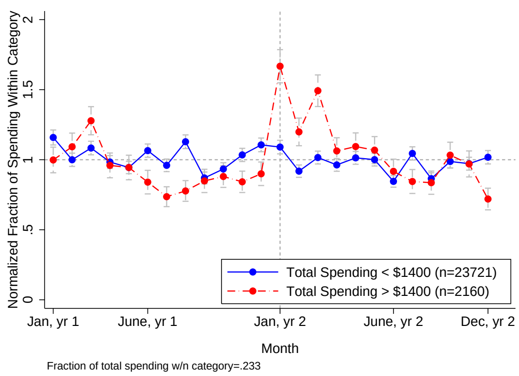
Figure 6: Strategic Claim Delay: heterogeneity by procedure type classification



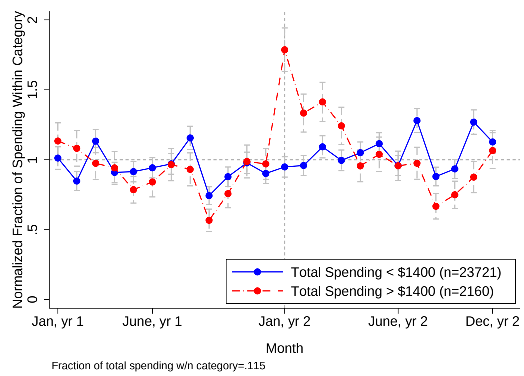
(a) Diagnostic



(b) Preventive



(c) Restorative



(d) Other

Notes: For each category of procedures, this figure displays the average normalized monthly individual dental expenditures on procedures in that category as a fraction of the total expenditures in that category across two adjacent years, the average value of (monthly category spending $\times$ 24/total category spending). The figure also displays $+/-$ the standard error. The sample used to create this figure is restricted to those who were insured with the company for two adjacent years (2004-2005, 2005-2006, or 2006-2007) and were enrolled in Plan L (which has a maximum benefit of $1,000) in the first of these adjacent years. Individuals with no spending across the two years are dropped. This series is displayed separately for those with overall expenditures (across all categories) less than $1,400 (those who probably did not have the incentive to delay claims) and those with overall expenditures (across all categories) exceeding $1,400 (those who were more likely to have the incentive to delay claims). The categorization of procedures for the purpose of this figure is described in detail in Section 4.

Table 1: Description of Employees and Sample Restrictions

	All Employees	Employees in Baseline Sample	Employees in Restricted Sample	US employees with dental Insurance
Employee-years	181,552	98,054	25,054	
(unique employees)	(64,466)	(38,028)	(12,530)	
Male	76%	72%	70%	53%
Age (median)	45	45	47	44
Rural	39%	38%	38%	
Job Tenure (median)	10	10	14	
Union	35%	6%	0.4%	7%
Wage				
Mean	$51,043	$56,602	$58,506	$48,943
Median	$45,652	$47,634	$50,360	$39,000
Dental Coverage Tier				
emp single	23%	24%	24%	
emp + spouse	17%	23%	29%	
emp + children	7%	10%	9%	
emp + family	53%	43%	38%	

Notes: In the left panel above, all the statistics are for employees not the associated dependents. The employee-year level of observation is used when calculating the mean and median statistics. The "All Employees" column describes all employee-years for all employees that were with the company at any point between 2004 and 2007. The "Employees in Baseline Sample" column describes all employees who ever had the relevant benefit menu and were employed for the entire relevant calendar year. These employees along with their associated dependents make up the "Baseline Sample" used in Section 4 to identify evidence of strategic claim timing. The "Employees in Restricted Sample" column describes employees that are in the baseline sample and were employed from 2004-2006. These employees along with their associated dependents make up the "Restricted Sample" used in the estimation of the empirical model. The "% Rural" is the percent of the sample that lives in a municipality characterized as rural by the 2000 US Census. For comparison, the "Employed US Population with Dental Insurance" column lists some descriptive statistics for the sample of people in the 2007 Medical Expenditure Panel Survey (MEPS) who were continuously employed and reported having dental insurance throughout 2007. All the values for these employees are for the year 2007. Because the MEPS does not indicate the source of dental coverage, this MEPS sample includes both employees that obtained coverage through their own employer and employees that obtained coverage from other sources (for example, through a spouse's employer). Overall, approximately 40% of the people in the MEPS 2007 report having dental insurance.

Table 2: Description of Dental Insurance Benefits by Plan

	Plan Coverage		Categories of Care[a]	
	Plan L	Plan H	% of Total Spending[b]	% of Total Claims[b]
Coinsurance Below Annual Maximum Benefit[c]				
Preventive Care[d]	100%	100%	32%	57%
Basic Care[e]	85%	85%	63%	42%
Major Care[e]	50%	50%	5%	1%
Oral Surgery[e]	50%	100%	<1%	<1%
Average Coinsurance[f]	88.3%	89.9%		
Annual Maximum Benefit	$1,000/person	$2,000/person		

[a] Claim categories are inferred by combining the procedure codes and claim reimbursement information. The average out-of-pocket spending to total spending ratio is calculated for each procedure code, and these codes are then classified into the care categories above. This process left less than 5% of claims with unclassifiable codes, and these claims are omitted from the statistics on the percentage of claims and spending by category. In addition to the plan differences noted above, Plan H provides orthodontia coverage for children under 18 years of age up to a separate lifetime maximum benefit of $1,500. In the empirical analysis, I use the annual maximum benefit feature to identify strategic claim timing. Since orthodontia is not subject to this annual maximum benefit, I exclude orthodontia claims from the analysis. Only a small fraction of households have any orthodontia claims and these households are dropped from the analysis.

[b] The "% of total spending" is the percent of total dental spending for each care category for individuals in the baseline sample, and the "% of total claims" is the percent of total claimed procedures for each care category for claims submitted by those in the baseline sample. The usage of the word "claims" in the heading of this table differs from the usage throughout the paper. Throughout the paper, I use "claims" to describe the total claimed procedures within an individual visit. In contrast, here "claims" is at a more disaggregated level, as an individual may have claimed procedures that span multiple care categories above within one visit.

[c] The displayed percentages are the percent of expenditures paid by the insurer for care in each of the above categories below the annual maximum benefit. Beyond the annual maximum benefit, all dental spending is the responsibility of the patient.

[d] The company places some limits on the annual amount of covered preventive cleanings and diagnostic X-rays. For example, covered patients may have up to two preventive cleanings and two partial mouth X-rays reimbursed within one calendar year.

[e] These types of care are subject to an annual individual deductible—$50 for Plan L and $25 for Plan H. This deductible is smaller than the vast majority of expenses in any of these categories, so one can think of it as simply a factor which increases the coinsurance rate for these services. This deductible is taken into account when calculating the average coinsurance rate.

[f] The "Average Coinsurance" rate displayed in the table is the average percent of expenditures paid by the company below the annual individual maximum benefit, where the average is taken over all care done below the annual maximum benefit accounting for the small annual deductible.

Table 3: Descriptive Statistics

	Baseline Sample		Restricted Sample	
	Plan L	Plan H	Plan L	Plan H
% Household-year observations	21%	79%	26%	74%
Mean % individuals with zero claims	37%	38%	30%	36%
Individual dental expenditures				
Mean	$206	$291	$230	$296
Median	$114	$125	$138	$130
Std Dev	$314	$467	$328	$480
% Individuals reached maximium benefit	3.3%	1.0%	3.9%	1.1%
% Individuals within $200 of maximium benefit	5.5%	1.5%	6.3%	1.7%
# Unique individuals	104,636		29,559	
# Unique households	38,028		12,530	

Notes: The "Baseline Sample" column describes all employees and dependents who ever had the relevant benefit menu and were associated with employees who remained with the company for the entire relevant calendar year. This sample is used in Section 4 to identify evidence of claim timing. The "Restricted Sample" column describes employees and dependents that are in the baseline sample and are associated with employees who were employed from 2004-2006 and selected dental coverage from the relevant menu. This sample is used in the estimation of the empirical model. The "% Household-years" is the percent of household-year observations on each plan. The "Mean % individuals with zero claims" is the percent of individual-years with zero claims on the relevant plan. The individual dental expenditure statistics are calculated across all individual-year observations. The "% Individuals reached maximum benefit" is the percent of individual-years that exceed the level of total spending that would exhaust the maximum benefit of the plan in which they were enrolled. The "% Individuals within $200 of maximum benefit" is the percent of individual-years that would exhaust the annual individual maximum benefit of the relevant plan with $200 dollars more in total dental spending. The maximum benefit for Plan L is $1,000 and for Plan H is $2,000.

Table 4: Testing for Asymmetric Information

	(1) $\text{Claims}_{h,t}$	(2) $\text{Claims}_{h,t}$
$\text{Choice}_{h,t}$	214.9	144.2
	(5.60)	(12.08)
Fixed Effects		
Premium Menu	x	x
Coverage Tier	x	x
Household		x
Dep Var		
Mean	650	650
Std Dev	779	779
N	86,062	86,062

Notes: The table above presents OLS regressions results where the level of observation is the household-year. The dependent variable, "$\text{Claims}_{h,t}$," is the amount the company would reimburse for household h's dental expenses had the household been enrolled in Plan H during year t (regardless of the actual enrollment of the household). This amount is calculated by applying the Plan H cost-sharing rules to the dental spending of the household using the appropriate average coinsurance rates. The variable "$\text{Choice}_{h,t}$" indicates whether household h enrolled in Plan H in year t. To ensure the insurance options are vertically differentiated, the sample is restricted to those households in the baseline sample that enroll the same dependents in medical and dental insurance in each year. Alternative specifications displayed in Appendix A demonstrate that the results are very similar when using different sample restrictions and alternative definitions of "$\text{Choice}_{h,t}$". All specifications include fixed effects for dental insurance premium menus, which vary slightly across employee benefit groups. Robust standard errors are clustered at the household level are reported.

Table 5: Estimates from Empirical Model

	(1)		(2)		(3)		(4)	
Risk Type, λ								
$\mu_\lambda(x)$								
Constant	0.1566	(0.0114)	-0.4046	(0.0164)	-0.4016	(0.0184)	-0.4107	(0.0165)
Male	-0.0995	(0.0118)	-0.0637	(0.0099)	-0.0604	(0.0115)	-0.0638	(0.0094)
Age (18,30]	-0.0167	(0.0031)	-0.0023	(0.0005)	-0.0023	(0.0008)	-0.0031	(0.0007)
Age (30,40]	0.0517	(0.0107)	0.011	(0.0042)	0.0143	(0.0062)	0.0074	(0.0034)
Age (40,50]	0.2248	(0.0143)	0.1685	(0.0132)	0.1422	(0.0150)	0.1727	(0.0148)
Age (50+)	0.3136	(0.0148)	0.2435	(0.0127)	0.2443	(0.0146)	0.2621	(0.0135)
priorTotal 500-1000			0.0806	(0.0199)	0.0989	(0.0217)	0.0810	(0.0174)
priorTotal 1000+			0.1139	(0.0189)	0.1487	(0.0239)	0.0756	(0.0160)
priorVisit			0.8783	(0.0178)	0.8754	(0.0171)	0.8852	(0.0163)
σ_λ	0.6793	(0.0068)	0.5467	(0.0079)	0.5518	(0.0084)	0.5502	(0.0087)
Delay, p_α (x, z)								
Constant	0.3663	(0.0549)	0.3623	(0.0626)	0.3987	(0.0232)	0.3350	(0.0466)
Male					-0.2814	(0.0267)	-0.2618	(0.0494)
Age (18,30]					-0.0199	(0.0065)	-0.0225	(0.0051)
Age (30,40]					0.0088	(0.0027)	0.0148	(0.0039)
Age (40,50]					-0.0263	(0.0060)	-0.0369	(0.0081)
Age (50+)					0.1883	(0.0535)	0.1435	(0.0290)
Least "time-sensitive"					0.2416	(0.1189)		
Most "time-sensitive"					-0.0908	(0.0108)		
D0: Preventive							0.0014	(0.0003)
D2: Restorative							0.1812	(0.0336)
Other: Endontics, Periodontics, Prosthodontics, Oral Surgery							0.1784	(0.0562)
Implied Mean p_α	0.3663	(0.0549)	0.3623	(0.0626)	0.4485	(0.0618)	0.4017	(0.0355)

Notes: The parameter estimates from the empirical model are reported above along with bootstrapped standard errors using 100 bootstrap iterations. Estimation of $(\beta_\lambda, \sigma_\lambda, \beta_\alpha, \delta_\alpha)$ is done using the method of maximum simulated likelihood, taking the empirical cost intensity distribution and risk aversion as given. The coefficient of absolute risk aversion used is 2.2×10^{-3}, the value calibrated to match the plan shares in the data as described in Section 5. The cost intensity distribution is the empirical distribution of claim costs for the restricted sample conditioned on age in a categorical manner: less than 19 years of age, 19-30 years of age, 31-40 years of age, 41-50 years of age, and over 50 years of age. Estimation details are in Section 5. The final row reports the "Implied Mean p_α" among individuals close to the maximum benefit whose division of claims helps to identify the delay parameters (described as "the potential delayers" in Section 5).

Table 6: Description of Counterfactuals

	Counterfactuals and associated restrictions on pricing and choice		
	(1)	(2)	(3)
	p	M_t	I_t
Contractible Information (Table 7)			
None	p same across individuals	$M_t = \{0, d_t + m_{t-1}\}$	$I_t = \{0, 1\}$
Risk Type	p varies by risk type (λ)	$M_t = \{0, d_t + m_{t-1}\}$	$I_t = \{0, 1\}$
Pre-Existing Events	p same across individuals	$M_t = \{0\}$	$I_t = \{0, 1\}$
Both Risk Type and Pre-Existing Events	p varies by risk type (λ)	$M_t = \{0\}$	$I_t = \{0, 1\}$
Choice Frequency (Table 8)			
Annual	p varies by risk type (λ)	$M_t = \{0, d_t + m_{t-1}\}$	$I_t = \{0, 1\}$
Every 2 years	p varies by risk type (λ)	$M_t = \{0, d_t + m_{t-1}\}$	$I_t = \{0, 1\}$ if $t \in \{1, 3, 5, 7, 9\}$; $I_t = I_{t-1}$ otherwise.
Every 5 years	p varies by risk type (λ)	$M_t = \{0, d_t + m_{t-1}\}$	$I_t = \{0, 1\}$ if $t \in \{1, 6\}$; $I_t = I_{t-1}$ otherwise.
Lifetime	p varies by risk type (λ)	$M_t = \{0, d_t + m_{t-1}\}$	$I_t = \{0, 1\}$ if $t = 1$; $I_t = I_{t-1}$ otherwise.

Notes: For each counterfactual scenario considered, this table presents the formal conditions corresponding to the the decision problem described in Section 6, Equation 16. The first four rows describe the conditions in the contractible information counterfactuals (presented in Table 7). When pre-existing events are contractible, firms can exclude pre-existing events (delayed treatment) from coverage which removes any incentive to delay claims. Thus, in the counterfactuals associated with contractible pre-existing events, in equilibrium no one will delay claims ($M_t = \{0\}$). The final four rows of the above table describe the conditions in the choice frequency counterfactuals (presented in Table 8).

Table 7: Contractible Information Counterfactuals

	Contractible Information							
	None		Risk Type		Pre-Existing Events		Both Risk Type and Pre-Existing Events	
	(1)		(2)		(3)		(4)	
Insurance Enrollment (%)								
No Subsidy	2.0%		4.8%		12.3%		100%	
10% Subsidy	4.5%		7.3%		17.9%		100%	
25% Subsidy	10.3%		11.4%		33.1%		100%	
Annual Mean Insured Cost ($)								
No Subsidy	1,809		1,197		576		241	
10% Subsidy	1,288		987		526		241	
25% Subsidy	889		805		435		241	
Annual Per-Capita Welfare ($, % of first best)								
	$	%	$	%	$	%	**$**	**%**
No Subsidy	28.4	36.0%	42.6	53.8%	50.1	63.2%	79.1	100%
10% Subsidy	38.8	49.0%	48.3	61.1%	53.5	67.7%	73.1	92.4%
25% Subsidy	49.5	62.6%	51.0	64.4%	56.3	71.1%	64.1	81.0%

Notes: In the three panels above, each cell represents an equilibrium of a different counterfactual scenario in a market for comprehensive dental insurance. “Insurance Enrollment” is the percent of individual-years insured. The “Annual Mean Insured Cost” is the average cost across the individual-years insured. The ”Annual Per-Capita Welfare” is the welfare measure discussed in the text (the mean certainty equivalent associated with that equilibrium relative to a world with no insurance available). The “No Information” column represents a scenario when the insurer can price no information about the individuals. The “Risk Type” column represents the scenario when the insurer can price individuals’ risk types, λ in the model. The “Pre-Existing Events” column represents the case when the insurer can contract on pre-existing events (delayed treatments). In this case, there are no delayed treatments in equilibrium. The “Both Risk Type & Pre-Existing Events” column represents the case in which the insurer and individual have symmetric information, and this information is contractible. The values above are for the calculated equilibrium (details in Appendix B).

Table 8: Choice Frequency Counterfactuals

	Enrollment Period Frequency							
	Annual (1)		Every 2 years (2)		Every 5 years (3)		Lifetime (4)	
Insurance Enrollment (%)								
No Subsidy	4.8%		13.3%		37.7%		100%	
10% Subsidy	7.3%		17.6%		62.2%		100%	
25% Subsidy	11.4%		55.7%		95.2%		100%	
Annual Mean Insured Cost ($)								
No Subsidy	1,197		591		328		241	
10% Subsidy	987		524		294		241	
25% Subsidy	805		333		246		241	
Annual Per-Capita Welfare ($, % of first best)								
	$	%	$	%	$	%	**$**	**%**
No Subsidy	42.6	53.8%	55.8	70.5%	71.4	90.2%	79.1	100%
10% Subsidy	48.3	61.1%	58.5	74.0%	71.9	90.8%	73.1	92.4%
25% Subsidy	51.0	64.4%	64.1	80.9%	63.8	80.6%	64.1	81.0%

Notes: In the three panels above, each cell represents the equilibrium of a different counterfactual scenario in a market for comprehensive dental insurance. In each simulation, it is assumed that insurers can price risk type (λ), but cannot contract on pre-existing events (delayed treatments). "Insurance Enrollment" is the percent of individual-years insured. The "Annual Mean Insured Cost" is the average cost across the individual-years insured. The "Annual Per-Capita Welfare" is the welfare measure discussed in the text (the mean certainty equivalent associated with that equilibrium relative to a world with no insurance available). An equilibrium is calculated (details in Appendix B) for each choice frequency restriction: annual insurance selection, every 2 years, every 5 years, and lifetime.