Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D

Maria Polyakova*

January 2015

Abstract

I take advantage of regulatory and pricing dynamics in Medicare Part D to empirically explore interactions among adverse selection, switching costs, and regulation. I first document novel evidence of adverse selection and switching costs within Part D using detailed administrative data. I then estimate a contract choice and pricing model that quantifies the importance of switching costs for risk-sorting. Conceptually, how switching costs affect selection depends on evolution of contract space relative to initial conditions. In Part D, switching costs help sustain an adversely-selected equilibrium and mute the ability of ACA policies to improve risk allocation in this important marketplace.

JEL classification numbers: D12, I13, I18, G22

Keywords: Medicare, Adverse Selection, Health Insurance Regulation, Switching Costs, ACA

1 Introduction

Outsourcing a public health insurance benefit to private insurers creates a familiar trade-off between potential efficiency gains from competition and potential efficiency losses from adverse

*Department of Health Research and Policy, Stanford, maria.polyakova@stanford.edu. This paper is the first chapter of my MIT dissertation. I am indebted to Amy Finkelstein, Nancy Rose, Stephen Ryan, and Jonathan Gruber for their guidance and support throughout this project. I thank Jason Abaluck, Isaiah Andrews, Andreas Breiter, Francesco Decarolis, Glenn Ellison, Sara Fisher Ellison, Kate Easterbrook, Michael Greenstone, Nathan Hendren, Gaston Illanes, Ray Kluender, Brad Larsen, Anna Mikusheva, Manisha Padi, Maxim Pinkovskiy, Jim Poterba, Adam Sacarny, Brad Shapiro, Annalisa Scognamiglio, Stefanie Stantcheva, Michael Whinston, Joachim Winter and seminar participants at MIT, Georgetown, Yale, UIC, University of Wisconsin-Madison, Stanford, Toulouse School of Economics, Bocconi, Sciences Po, and WUSTL, for very helpful comments, suggestions and discussions. I also thank Jean Roth at NBER and Sarah Brunsberg at ResDAC for their help in navigating the Medicare data, as well as the Max Planck MEA for their hospitality.

selection. We know much less, however, about the role of other market imperfections in competitive insurance settings. In this paper, I utilize the institutional environment of Medicare Part D prescription drug insurance program, to empirically analyze how switching frictions affect the re-distribution of risks across insurance contracts in response to market evolution, as well as how they alter the consequences of regulatory interventions that directly change the contract space.

Medicare Part D is a large insurance program introduced in 2006 that currently enrolls approximately 32 million beneficiaries, with annual federal spending of around $63 billion. This heavily subsidized benefit is administered entirely by private insurers that are extensively regulated. The insurers offer a variety of plans, making individual contract choice a prominent feature of the program.

Motivated by the evolution of the regulatory and pricing environment of Medicare Part D, I use the detailed administrative records on individual-level enrollment decisions, ex ante health risk, and realized ex post spending in the program, to explore how switching costs interact with adverse selection and with the regulatory interventions that exogenously change the relative generosity of plans. The paper argues that, conceptually, whether switching costs ameliorate or exacerbate adverse selection depends crucially on how both the relative prices *and* the relative coverage generosity of contracts evolve in comparison to the initial conditions. For instance, in an environment where, absent switching frictions, the degree of adverse selection would have declined over time, switching costs will have a counter-intuitive effect of exacerbating adverse selection. While the classical models of adverse selection leave little space for the possibility of adverse selection declining over time, in practice regulatory interventions or insurer decisions in complex marketplaces may well lead to such developments. As a specific example consider a situation, where in the presence of a less and more generous contract the regulator mandates the less generous contract to increase its coverage.

The observation that the interaction between switching frictions and adverse selection depends on the evolution of the contract space has direct policy implications. First, my findings suggest that implementing a nudging policy in the Part D environment would not lead to the unraveling of the market, as conventional wisdom may suggest. Second, health insurance markets are frequently characterized by regulatory interventions that change contract features - some of these policies directly attempt to reverse adverse risk-sorting dynamics. Taking the example of contract space changes envisioned in Medicare Part D under the Affordable Care Act, the paper demonstrates that switching costs may mute the effectiveness of policy instruments designed to ameliorate adverse selection.

The paper is structured as follows. The key pieces of Part D's institutional environment that are necessary for the subsequent empirical analysis are summarized in Section 2. Section 2 also describes the sample of the nationwide administrative data from the Centers of Medicare and Medicaid. In Section 3, I turn to the descriptive analysis of the data.

The data suggest that despite numerous regulatory provisions designed to mute adverse selection in this market, both cross-sectional and dynamic selection exist in Medicare Part D.[1] Surprisingly, given the extent of policy attention to this concern, adverse selection in Medicare Part D has received little attention in the empirical literature.[2] To fill this gap, I document several pieces of evidence for the presence of adverse selection. First, I present the results of the cross-sectional positive correlation tests as described in Chiappori and Salanie (2000). These tests suggest a large degree of asymmetric information: the most generous plans attracted individuals whose annual drug spending was more than a standard deviation above the spending in the least generous plans. Although it is typically hard to disentangle selection from moral hazard in the positive correlation tests, Medicare Part D setting provides me with a way of addressing this problem. I use a rich set of moral hazard-free ex ante diagnostic information about the beneficiaries to recompute the positive correlation tests, isolating the lower bound of the cross-sectional adverse selection. These tests reveal that the most generous contracts in

[1]It is important to note that I evaluate and define adverse selection as the inherent sorting of risks across contracts. I do not explicitly consider whether this inherent risk-sorting is ex-post successfully mitigated in insurers profits though a multitude of CMS risk-adjustment and reinsurance policies that the federal government has put in place to combat selection in Part D. We can think about the exercise as testing for whether there is in fact a reason for the government to be implementing these policies. Implicitly, however, the evidence of selection spirals demonstrates that the existing mechanisms to combat adverse selection have been imperfect and ineffective in insuring access to generous coverage on the market.

[2]To the best of my knowledge, no systematic analysis of selection in the spirit of the established literature testing for asymmetric information has been conducted in this important setting.

Medicare Part D attracted individuals with substantially higher (about $730 a year) expected spending. Consistent with these cross-sectional results is also the evidence of dynamic adverse selection, which we see in the rapid unraveling of the most generous contracts.

I then turn to the question of whether Medicare Part D exhibits switching costs, which may be an important factor in understanding the over time dynamics of the risk-allocation patterns highlighted above. While the literature has focused on the "inconsistency" of individual choices relative to a perfectly rational benchmark (e.g. Abaluck and Gruber 2011, 2013; Heiss et al. 2013; Ketcham et al. 2012), I do not assess the efficiency of choices here; rather, I document that, whichever choices individuals make, changing their initial choices appears costly.

Identifying switching costs or "true" state dependence separately from unobserved heterogeneity is a notoriously difficult problem (e.g., Heckman, 1991). Medicare Part D's institutional setting and administrative records, however, provide nearly ideal variation and data for documenting the presence of structural state dependence. First, there is substantial variation in the choice set that is available in the program over time. Second, every year, new beneficiaries enter the program, as they become eligible for Medicare upon turning 65. Using this variation, I show that choices of the "continuing" cohorts persistently reflect the market conditions of the year in which these individuals made their first choices, while the choices of the otherwise similar newly entering cohorts are different and reflect the current market conditions. This approach to documenting switching costs is in the spirit of Handel (2013)'s approach to documenting inertia in employer-provided health insurance. Consistent with the cross-cohort choice comparison, I also find that existing cohorts are less price-sensitive than the newly entering cohorts. These findings provide micro-level identification for the presence of switching costs in Part D and thus support the hypothesis made in the earlier literature based primarily on aggregate data (Ericson, 2013) that inertia is an important feature of this market.[3]

With the reduced-form evidence of adverse selection and switching costs in hand, I proceed to a utility-based choice model of the individual preferences for insurance plans in Section 4.

[3]In contemporaneous work, Ho, Hogan, and Scott Morton (2013) document that switching is rare in Part D and analyze channels that may be driving switching, as well as the implications of these market features for supply-side behavior.

I estimate a flexible choice model that admits heterogeneous preferences and private information about health risk, and allows me to quantify switching costs, adverse selection, and the individuals' willingness to pay for different contract features conditional on their risk. Importantly, the presence of off-equilibrium contracts in the first two years of the data, allows me to estimate preferences for a variety of coverage levels that reduce the necessity to extrapolate the model far out-of-sample in the counterfactual analyses. The estimation results, discussed in Section 5.1, suggest that switching costs are large and critical for explaining enrollment paths over time. The estimates further suggest that information about risk plays an important role in determining individual choices, which is consistent with the descriptive evidence that the self-selection of beneficiaries in the Part D environment leads to adverse selection.

Counterfactual analyses in Sections 5.2 and 5.3 utilize the estimated choice model to explore the complex interactions among adverse selection, switching costs, and minimum standard policy regulations in the Medicare Part D setting. The textbook model of adverse selection in insurance hinges on the idea that costlier individuals will choose the most comprehensive plan; this plan will experience high costs and will have to increase its premium. In turn, the least costly individuals will drop out of the plan and the spiral will continue unraveling. It then seems intuitive that if we introduce a switching friction into this model, it may stop the unraveling process right at the first loop: if the friction is sufficiently high, no individuals will leave the most comprehensive plan despite the increase in its price. Indeed, Handel (2013) provides an empirical example of switching costs restricting the selection spiral in an employer-provided health insurance setting. In that setting, the relative price of a more generous contract increases over time, but individuals do not re-adjust their choices due to switching costs. This evidence is sometimes interpreted as confirming the intuition that switching costs *must* be helpful in ameliorating adverse selection.[4] In general, however, the direction of the effect of switching costs on selection is ambiguous. If the relative price of the more comprehensive contract fell instead of increasing over time, the result would have likely been the opposite.

[4]Handel (2013), however, points out that the direction of the results is specific to the considered environment and enrollee population.

In practice, the latter possibility of relative premiums, or relative coverage, falling, is especially relevant in more complex insurance settings, such as public health insurance exchanges. In such settings, a large number of competing plans are differentiated both vertically and horizontally, their prices and characteristics are affected by a host of regulatory interventions, and, due to complex subsidy provisions and strategic pricing, the relative premiums of plans may not perfectly reflect the actuarial differentials among them. The combination of these factors implies that in such markets the relative price of a more comprehensive contract, for example, may go down rather than up, or its characteristics may adjust relative to the other contracts in a way that favors less acute sorting. For example, in Medicare Part D, we observe that the government incrementally decreases the relative generosity of the most generous plans overtime by mandating higher coverage levels for less generous plans. In such cases, the switching cost would have the unintuitive effect of exacerbating the adverse selection concern. This conceptual mechanism is illustrated in Figure 1. In Section 5.2, I discuss two empirical equivalents to Figure 1. These empirical examples use a significantly simplified, yet realistic, version of the Part D contract space that illustrate the opposite effects that switching frictions may have on selection. In the same section, I proceed to describe the full counterfactual simulation that considers all contracts offered in Medicare Part D. In this simulation I find that Part D market evolved in a way that made switching costs support, rather than mute, adverse selection. Specifically, I find that without the switching friction, the difference in the average risk between the least and the most generous contracts would have been 21% lower. While this conclusion incorporates a standard assumption that contract characteristics, except for price, follow the pattern observed in the data (in the sense that I do not let choices of coverage levels be endogenous to inertia or other market forces in the model), this assumption is potentially less stringent relative to a standard product-market setting, because some product characteristics here are driven by incremental exogenous changes to the Standard Defined Benefit that likely do not depend on inertia.

Indeed, a common channel driving the changes in the relative prices and generosity of contracts in public health insurance settings is regulatory intervention. The fact that switching costs alter the response of risk-sorting to contract changes implies that switching costs will also significantly alter how policy instruments used by the government to regulate this market will impact the allocation of risk. In Section 5.3, I explore this hypothesis on the example of Part D's minimum standard regulation. This regulation - the so-called Standard Defined Benefit (SDB) - specifies the minimal amount of coverage that the participating insurers have to offer in their plans. The regulation is dynamic, in that the government adjusts the minimum requirements every year by raising the standard deductible levels and increasing the coverage limits. Considering this particular regulation is highly policy-relevant, as the provisions of the Affordable Care Act (ACA) envision significant expansion of the minimum benefit to eventually eliminate the infamous "donut hole". My simulations suggest that switching costs will significantly mute the ability of the minimum standard regulation to change the distribution of risks across contracts. For example, I estimate that without switching costs, the expansion of the minimum standard by "filling" the donut hole would substantially reduce adverse selection; with switching costs, however, this expansion has little effect on sorting. In the spirit of the theory of second-best, these results demonstrate the importance of accounting for the interaction among different market failures in health insurance markets and how the correction of one market failure, such as demand-side frictions, may change the effect that regulatory instruments targeted at correcting a different market failure, such as adverse selection, have on the market.

The analysis in this paper is related to several literatures. First, the paper is related to the growing body of literature that analyzes the Medicare Part D program. Most of this literature has focused on assessing the rationality of individual decisions (Heiss et al. 2010, 2013; Abaluck and Gruber 2011, 2013; Kesternich et al. 2013; Ketcham et al. 2012; Kling et al. 2012). Heiss et al. (2009) note the stark growth in Part D premiums for generous coverage and acknowledge that adverse selection may be an important concern in this highly regulated setting. They allow for selection in their simulations of a life-cycle choice model in an environment stylized to reflect the key features of Medicare Part D. Several papers in this literature have suggested that switching costs may be present in the program. In addition to the analysis in Ericson (2013) that documents evidence of insurer pricing strategies consistent with the presence of inertia, Miller and Yeo (2012) assume that the Medicare Part D market exhibits consumer inertia and include a switching cost parameter in their choice model estimated on the market-level data. Further, Abaluck and Gruber (2013) in their analysis of choice inconsistencies allow for switching costs separately from other choice imperfections. Ketcham et al. (2012), on the other hand, suggest that inertia is not of key concern in Part D. Some of the other work analyzing Medicare Part D has looked at the impact of Part D on prescription drug consumption (Yin et al. 2008; Duggan and Scott Morton 2010); the role of low income subsidy regulation (Decarolis, 2013); the welfare of reducing choice (Lucarelli et al., 2012); and the moral hazard response to non-linearities of the contracts (Einav et al., 2013). Duggan, Healy, and Scott Morton (2008) provide an extensive overview of the program's design.

Second, beyond the Medicare Part D setting, the paper is related to the growing empirical literature that analyzes asymmetric information, regulation, contract design, and welfare in both employer-provided and public health insurance settings. Einav, Finkelstein, and Levin (2010) provide a systematic overview of the literature in this vein. The current paper builds upon the insights in the work on the interaction between adverse selection and minimum standard regulation in Finkelstein (2004), as well as on the interaction between adverse selection and inertia in an employer-provided insurance setting in Handel (2013).

Third, the paper also relates to the growing literature on inertia and defaults in a variety of public finance settings: among recent examples are Chetty et al. (2013), Beshears et al. (2013), Hastings et al. (2013) and Luco (2013), who document such patterns in retirement savings accounts in Denmark, the US, Mexico, and Chile respectively, and Nosal (2012), who estimates switching costs in Medicare Advantage. Finally, methodologically and conceptually, the paper is related to a broad literature in industrial organization that assesses the impact of switching costs and incumbent advantages on market outcomes in a variety of settings (Farrell and Klemperer 2007 provides a survey), as well as the vast literature on private market regulation (Joskow and Rose 1989; Armstrong and Sappington 2007).

2 Institutional Setting and Data

Basics of Medicare Part D

Medicare is a public health insurance program for the elderly and disabled in the U.S. Until 2006, standard Medicare insurance, so-called Parts A and B, covered hospital and physician services, but not prescription drugs. In 2006, Medicare Part D prescription drug coverage was launched as part of the Medicare Modernization Act of 2003, becoming the largest expansion of Medicare since its introduction in 1965. While Medicare bears the greater share of Part D costs (CBO projects 2013 spending on Part D to be $63 billion or 2% of the 2013 budget outlays), the actual administration of the Rx benefit and part of the actuarial risk have been outsourced to private insurers. In 2012, Part D covered around 32 million beneficiaries (Hoadley et al., 2012) - 62% of these were enrolled in stand-alone prescription drug plans (PDPs), which are the focus of the current paper.

Medicare Part D coverage is voluntary. Eligible individuals have to actively enroll in one of more than 30 stand-alone Rx plans offered in their state of residence during annual open enrollment period or when they first become eligible, e.g. turn 65. Once enrolled, beneficiaries pay premiums on the order of $400 – $500 a year, and in return insurers pay for prescription drug purchases subject to a deductible, co-payments or co-insurance, and coverage limits. Beneficiaries stay in their chosen plan for a year and may change their choice during the open enrollment period next year. If beneficiaries make no changes to their plan choice in subsequent years, CMS will continue enrolling them in their first chosen plan unless it is terminated by the insurer. The fact that individuals self-select into plans and have a "default" plan if they do not take any action after their first enrollment will be important for my analysis.

Data – baseline sample

I utilize the detailed administrative data provided by CMS that comprises a 20% random sample of Medicare beneficiaries nationwide for years 2006-2009. The data provides basic demographic and detailed health information about the beneficiaries, the characteristics of all Part D plans available in each region of the country, the enrollment choices of the beneficiaries, and subsequent prescription drug spending for those who enrolled in Part D. I make a number of restrictions to the original sample of 40.3 million beneficiary-year observations to isolate the part of the market where 65 year old and older enrollees self-sort into a cleanly observable set of Part D contracts. These restrictions bring the sample down to 5.3 million individual-year observations.[5] This baseline sample has individuals that chose to enroll in a stand-alone prescription drug plan and did not receive any additional subsidies from the government that would have distorted the monthly premiums or cost-sharing characteristics of the choices. Since in the econometric choice model I need to observe individual's choices over consecutive years, I also construct a panel sub-sample of the baseline sample. This sub-sample contains individuals whose choices and utilization can be observed from the first year they enter the program to 2009. The panel sub-sample has 3.7 million beneficiary-year observations on approximately 1 million unique individuals.

Table 1 provides the summary statistics of the observed demographic and risk related variables for the full sample, the baseline sample and the panel sub-sample. The individuals in the baseline sample are on average 76 years old, 64% female, predominantly white (95%), with annual average drug spending of about $1,900. The panel sub-sample looks very similar, albeit a year younger on average and with slightly lower annual spending. In comparison to the full sample that includes beneficiaries eligible for both Medicare and Medicaid, non-enrollees, and those who qualified for Medicare before turning 65, the baseline sample has individuals that are older, more often white and more often female. Overall, the baseline sample has somewhat healthier individuals compared to the full Medicare population, as measured by the average risk score of 0.9, which is 10 % below the population average that by construction is 1.

Regulatory environment and the nature of the observed contract space

The nature of Medicare Part D's contract space is driven by a minimum standard regulation. Medicare has designed a so-called Standard Defined Benefit (SDB) for the Part D program and insurers are required to provide coverage that gives at least the same actuarial value as the SDB. The SDB has an unusual design that features a relatively low deductible, flat co-insurance rate

[5] More details on data construction are available in Appendix; Table A.1

of 25% up to the initial coverage limit (ICL) and subsequent "donut hole", or coverage gap, that has a 100% co-insurance until the individual reaches the catastrophic coverage arm of the contract. Figure 2 illustrates what these cost-sharing features imply for individual out-of-pocket spending.

A crucial feature of the institutional setting, which generates cross-sectional variation in contract characteristics, is that insurers are allowed to adjust and/or top up the SDB contract design. As a result, contracts offered by Part D insurers are highly multidimensional and vary on a variety of characteristics that differentiate them from the SDB minimum. Some of this differentiation is purely financial - contracts can change cost-sharing thresholds, co-pay and co-insurance levels, and may offer coverage in the "donut" hole. Other differentiating features are more related to the quality of the insurance provider.

Despite the multi-dimensionality of contracts and official counts of more than 1,500 contracts in the Part D program, there are three stylized facts about this market that haven't been emphasized in the literature, but allow me to simplify the description of the contract space. First, each insurer in practice offers the same menu of 2-3 contracts in all Part D regions in which it operates. Second, insurers tend to adjust only premiums, but not other contract features across different regions. Third, insurers tend to keep the "types" of contracts in their menu fixed over time, adjusting the key characteristics only to the SDB policy changes. Using these three stylized facts, I classify all contracts into four types. Contracts that offer the standard-defined-benefit level of deductible and initial coverage limit are classified as *Type 1* contracts. *Type 2* contracts offer a reduced deductible (usually reduced all the way to zero), but still the standard level of the initial coverage limit. *Type 3* contracts offer a reduced deductible and partial coverage (usually coverage of generics) in the gap beyond the ICL, while *Type 4* contracts offer a reduced deductible and full coverage in the gap.

In addition to the cross-sectional complexity of the contract space, there is substantial over time variation in the characteristics and premiums of the plans. The time-series changes are the outcome of both the market-driven adjustments of contract characteristics by insurers, as well as of annual changes in SDB regulation. We observe in Figure 2 the annual adjustments to the levels of standard deductible and coverage limit. In addition to the regulatory adjustments of contract features, there were substantial additional changes in premiums that differed greatly across plans. These differential adjustments to premiums of different contract types resulted in relative premiums changing substantially over time. For example, the relative premiums for *Type 3* contracts (relative to *Type 1)* increased by 25%, while the relative premiums for *Type 2* fell over time by 73%. These market dynamics motivate the question, pursued in this paper, about the role of switching costs in determining the allocation of risks across contracts in an environment where the relative attractiveness of choices evolves due to market forces and regulatory interventions.

3 Descriptive evidence

3.1 Adverse selection in Medicare Part D

The theory of adverse selection suggests that more comprehensive insurance contracts are likely to attract an adversely selected risk pool of individuals with higher health risks. In practice, the literature does not always find evidence of adverse selection in insurance markets (Finkelstein and McGarry, 2006), which suggests that the presence of selection is ultimately an empirical question. In this section, I document novel empirical evidence for the presence of adverse selection in Medicare Part D. To begin with, I present the cross-sectional correlation test as described in Chiappori and Salanie (2000) using the ex-post realized drug expenditures. This exercise detects the presence of asymmetric information and follows the well-established testing literature reviewed in Einav et al. (2010). I then provide two pieces of additional evidence that help disentangle selection from moral hazard, which is a common concern in health insurance. First, I repeat the Chiappori and Salanie (2000) test using ex-ante information about individuals' health summarized in a risk score index. Second, I document the presence of two selection death spirals in the early years of the program. It is important to note that in this section I evaluate and define adverse selection as the inherent sorting of risks across contracts. I do not explicitly consider whether this inherent risk-sorting is ex-post successfully mitigated in insurers profits though a multitude of risk-adjustment and reinsurance policies that the federal government has put in place to combat selection in Part D.[6] We can think about the exercise as testing for whether there is in fact a reason for the government to be implementing these policies. Implicitly, however, the evidence of selection spirals demonstrates that the existing mechanisms to combat adverse selection have been imperfect and ineffective in insuring access to generous coverage on the market.

Figure 3 illustrates the first positive correlation test graphically. It plots the average realized drug spending in each region by contract type in 2006. We can see the stark differences in the expenditures in the more and less generous contracts. The differences are especially striking for 2006, as it contains data for several *Type 4* contracts that have realized average annual spending on the order of $\$4,000$ in all regions, as compared to $\$1,500$ in the *Type 1* contracts. The differences in expenditures shrink slightly in later years, as the high risks of the unraveled *Type 4* contracts become integrated into the rest of the market, but there still remains a substantial difference between *Type 1* and *Type 3* contracts (graphical representation of other years is available in Appendix).

Table 2 presents the formal specification of the test for the presence of asymmetric information. The test is done using the baseline sample pooled for years 2007-2009 that includes observations for which risk scores and claims are accurately measured for the whole year. Since insurers are allowed to price the same contracts differently in different regions, the test controls for region fixed effects. The regression specification takes the following form:

$$Y_{irt} = \alpha_r + \delta_t + \sum_{k=2}^{k=4} \beta_k \mathbf{1}\{ContractType_{it} = k\} + \epsilon_{irt} \tag{1}$$

where i indexes individuals, r indexes regions and t indexes years. I use realized total drug

[6]There are three key policies that the government uses to ex-post adjust insurer profitability from bad risk pools. First, government subsidies to insurers are adjusted to account for enrollees' risk scores. Second, at very high spending level,s the government directly picks up 80% of the bill, leaving the insurer liability limited to 15% with the remaining 5% paid by the patient. Thirds, so-called risk corridors ensure that insurance companies do not suffer large overall losses at the end of one fiscal year of operations.

spending as the first outcome variable in Column (1). Since the goal of the exercise is to test whether higher risk individuals sort into more generous contracts, the spending variable does not account for the cost-sharing provisions of the plans. The results suggest that more generous contracts have higher spenders. For instance, contracts with full gap coverage attract individuals with realized drug spending that is more than a standard deviation higher than in the plans with minimum standard coverage.

Since the correlation tests that use the realized spending as the outcome variable capture both adverse selection and moral hazard, I repeat the testing exercise using risk scores as the outcome variable in Column (2). Risk scores are constructed using ex-ante diagnostic information from Medicare Part A/B (hospital and physician) claims and therefore do not reflect any effects of plan structure on spending that may show up in the tests that use drug spending as the outcome variable. CMS uses risk scores to adjust subsidy payments to insurers.[7] Using the risk score measure gives qualitatively similar results, although the magnitude of differences is smaller. To give a more meaningful interpretation to the results with risk scores, in Column (3) I project drug spending onto risk scores and use the projection as the outcome variable in the correlation test. This exercise expresses the risk scores in dollars of expected spending. Since risk scores do not have the moral hazard aspect, this projection gives me the *lower bound* on how much of the estimated differences between the realized spending in different types of plans can be attributed to adverse selection on ex ante observed (and thus potentially priceable) risk. The results remain qualitatively similar. Although this measure accounts for only about a third of the differences that were observed in the first regression, the differences are still large in the absolute sense, corresponding to more than 100% of the average annual premiums. Figure 3 illustrates the moral hazard-free results graphically. Instead of comparing just the average expected risk, this figure plots the whole empirical CDF of the ex-ante risks in different types of contracts in year 2006. We can clearly see that, indeed, the whole distribution of risks in the

[7] This implies, if risk adjustment system were perfect, the implications of adverse selection documented here on premiums would have been mitigated by CMS payments across insureres. As noted above, in that case we could think about the exercise as testing for whether there is in fact a reason for the government to be implementing these policies.

more generous contracts is shifted towards having more mass of higher risks.[8]

In addition to this evidence of cross-sectional adverse selection, Medicare Part D illustrates rarely observed evidence of dynamic adverse selection. Despite the extensive efforts of the regulator to incentivize private insurers to offer Part D contracts with full gap coverage, such contracts were discontinued after 2007. The selection spiral happened twice, since different insurers attempted offering full gap coverage in 2006 and 2007, both of them discontinuing full gap coverage after one year of operation. *Type 4* plans with full gap coverage experienced annual claims that were about twice as high as the claims in plans without gap coverage. Subsequently, their premiums nearly doubled and enrollment dropped. Up until today no plans in Medicare Part D program offer full coverage in the gap.

The Medicare Part D environment thus illustrates a rare setting, where an off-equilibrium plan was in practice offered on the market and then rapidly unraveled. This unique feature of the data allows me to recover preferences (in demand estimation of Section 4.1) for levels of coverage that would otherwise have had to be extrapolated from far out of sample.

3.2 Switching costs in Medicare Part D

One objective of this paper is to analyze how risk-sorting among contract types changes in response to regulatory and market-driven adjustments in the contract menu. We would expect changes in the relative prices or generosity of the available insurance contracts over time to induce a re-allocation of risks among them. In the presence of high switching costs, however, such re-allocation may be completely muted. In this section I begin exploring this interaction mechanism by first documenting evidence for the presence of switching costs in Medicare Part D.

In general, documenting evidence of switching costs is challenging, since we need to distinguish between the "structural" state dependence and unobserved persistent individual heterogeneity (Heckman, 1981a,b; Heckman and Singer, 1986; Heckman, 1991; Honoré, 2002; Honoré and Tamer, 2006; Dube et al., 2010). Several features of the Medicare Part D environment,

[8]Risk CDF for years 2007-2009 are available in Appendix Figure A.3

however, render themselves well to such analysis. First, due to the regulatory changes, entry and exit of plans, as well as substantial market-driven over time adjustment in prices and characteristics of contracts (some of it strategic, see Decarolis (2013)), there is pronounced non-stationarity in the observed contract space of the program. Second, the Part D environment allows observing all individuals making their first choice in 2006, as well as younger individuals making their first choices in years 2007-2009 from the adjusted contract menus. In this section, I utilize these features of the Part D environment to provide descriptive evidence of choice behavior consistent with the presence of switching costs as separate from persistent individual heterogeneity.[9] This approach to documenting switching costs is in the spirit of Handel (2013)'s approach to documenting inertia in employer-provided health insurance.

I find four descriptive patterns in the data consistent with the presence of significant switching costs in Part D. First, in each year of the program about 90% of individuals enroll in their "default" plans; for individuals whose default plans significantly change their financial characteristics (and thus their type) this probability is still around 80%. Since premiums and contract characteristics change substantially from year to year even if plans do not change their "type", the high persistence in choices suggests that switching costs may be present; this evidence alone, however, could just point to very persistent preferences.

As the second piece of evidence, I compare the choices of the newly entering and existing enrollees in different years. The results are recorded in Table 3. I focus on the individuals that can be tracked from their first entry continuously to 2009 and whose default plans' types were not changed by insurers throughout the observed years. This isolates individuals whose choices are not conflated with substantial supply-induced re-classification of plan types. Two patterns are pronounced in the data and consistent with the idea that switching costs play an important role. First, enrollment shares over time for a given cohort tend to be closely related to the choices and market conditions of the first year in which the cohort entered the program. Second, the choices of different cohorts in the same year are very different. For

[9]Section 4.1 then presents a more formal discussion of the issues related to the identification of the state dependence parameter in dynamic discrete-choice panel data models.

example, comparing the choices in 2008 of the cohort that entered in 2006 and the cohort that entered in 2008, we see that the 2008 cohort is almost twice less likely to enroll in the least generous *Type 1* plan than the 2006 cohort in 2008: 10% vs 19 % enrollment share. Another persistent difference in choices is visible for the 2007 cohort, which in 2007 was much more likely to select the *Type 3* plan with partial gap coverage than cohorts entering in other years.

Comparing the choices of different cohorts by their choice of different insurer brands rather than contract types paints a similar picture. Figure 4 records the enrollment shares of the two biggest insurers in the sample for each year 2006-2009. The enrollment shares are shown separately for the 65 year olds, who are entering the program anew, and older enrollees with incumbent plans. We see a striking difference in the 2009 choices. In this year, one of the insurers ("Insurer B") - which substantially increased its premiums in 2009 - lost almost all of its market share with the new enrollees. Only about 5% of the 65 year olds chose to enroll in the plans offered by this insurer. Among the continuing cohorts, its enrollment share also fell, but not nearly as dramatically. It remained higher than 20%, implying that in 2009 the existing cohorts were four times more likely to be enrolled with Insurer B than the new enrollees.

As the final piece of evidence, I use a simple conditional logit regression to test whether there are statistically significant differences in the price sensitivity of the cohorts of new and continuing enrollees.[10] Under the null hypothesis of no switching costs, we would expect the coefficients on plan premiums for new (65 y.o.) and existing enrollees of similar age (66 - 70 y.o.) in the same year to be very close to each other. The estimates allow me to reject this null. I find that price sensitivity is significantly higher in magnitude for 65 year olds than for all cohorts of 66-70 year olds in years 2007-2009. This does not hold in 2006 when beneficiaries of all ages are entering the program anew. Furthermore, the estimates of the price coefficient are virtually identical for each age group among 66-70 year olds, suggesting that the difference between the estimated price sensitivity for the new and continuing cohorts is not driven by age differences per se, but instead are related to the lack of switching costs for the 65 year old beneficiaries.

[10] The exact specification is discussed in the notes of Table 4 that reports the results of this regression.

4 Empirical model

4.1 Specification and identification of the contract choice model

The descriptive evidence in Section 3 has documented that adverse selection and switching costs are present in Medicare Part D. This evidence alone, however, doesn't allow quantifying the economic significance of these market imperfections, or the extent to which they may impact regulatory policies. To address these questions, I formulate an econometric model of how individuals choose which contract to enroll in. The model specifies the choice decision as a function of the information about the individual health risk, the switching cost, and heterogeneity in individuals' preferences for different features of the contracts. The model takes a contract-valuation approach rather than a realized utility approach as has been utilized by Abaluck and Gruber (2011, 2013) in which Part D contract characteristics were projected into expected out-of-pocket spending under certain assumptions about individual's expectations. The contract-valuation approach does not include the expected out-of-pocket spending as an explanatory variable, as it attempts to make fewer assumptions on how individuals interpret the financial features of the contracts and on the individuals' information set about risk at the time of choice.[11] This comes at the cost of not recovering deeper utility primitives, such as risk aversion. Revealed valuation of contract characteristics, however, is sufficient for assessing how choices and risk allocation would change with costless switching and in response to policy-driven changes in the contract space that are reasonably within the realm of observed contract characteristics. These are exactly the type of counterfactuals that I consider in Sections 5.2 and 5.3.

Specification Each year t an individual i who lives in region r and is enrolled in the Medicare Part D stand-alone prescription drug program chooses among J_r plans offered by B insurers. Each insurer b typically offers a menu of up to three plans of different types in each

[11]Both the contract-valuation and the realized utility approaches to modeling of health insurance choices are common in the literature. Einav et al. (2010) discuss the trade-offs between these approaches, illustrating the contract-valuation approach on the work of Bundorf et al. (2012) on pricing in the employer-provided health insurance.

region. The plans of the same type offered by the same insurer are likely to have the same characteristics, but different premiums in different regions. Some plans are offered only in a subset of regions. These two features of the program imply that J_r varies by region. The plans in each J_r can be projected into the same set of observed characteristics. While in principle the plans could be characterized by a very high-dimensional vector of characteristics available from the administrative records, in practice I have to take into account which characteristics of the plans are feasibly observed by beneficiaries when they are making their choice. I let individual i's utility from choosing plan j (where "plan" is region-specific, so r-indexing is suppressed) in year t be given by:

$$u_{ijt} = -\alpha p_{jt} + \beta_{it}\phi_{jt} + \gamma_{it}\mathbf{1}\{\text{Default}\}_{ijt} + \epsilon_{ijt} \tag{2}$$

$$\epsilon_{ijt} \sim \text{iid Type 1 EV}$$

Utility thus depends on the annual premiums charged by the plan in a given region in a given year p_{jt}, the characteristics of the plan ϕ_{jt}, and whether j was a default plan for individual i in year t, where the default plan is usually the plan chosen in $t-1$.[12] Individuals are assumed to choose a plan that gives them the highest utility. An important assumption that is implicit in this formulation of the utility function is that individuals are myopic in their choice of plans. In other words, I am assuming away a possibility that individuals, for example, expect a plan that is cheap today to become very expensive tomorrow and thus choose a suboptimal plan today to avoid paying the switching cost tomorrow. Further, this formulation assumes that individuals choose the option with the highest "perceived" utility, which may not necessarily correspond to the highest "objective" valuation of plans as financial contracts (indeed, Abaluck and Gruber (2011, 2013) suggest that beneficiaries are choosing their plans inconsistently with the objective efficiency frontier). For the analysis of risk-allocation and choices in counterfactual scenarios, however, this "subjective" utility is exactly the object of interest.

[12] To construct the "default" variable I use the administrative records of which plans were renewed over time, which plans were discontinued and which plans were consolidated. In cases where plans were renewed or consolidated, CMS would default individuals in the same (if renewed) or designated new (if consolidated) plan if individuals took no action to change their choices. In rare cases when plans were terminated, individuals are recorded to have no default plan in that year.

I let the characteristics component of the utility function ϕ_{jt} include the following:

$$\begin{aligned}\beta_{it}\phi_{jt} &= \beta_{1it}\text{Deductible}_{jt} + \beta_{2it}\text{ICL}_{jt} + \beta_{3it}\mathbf{1}\{\text{Partial coverage in gap}\}_{jt} + \\ &+ \beta_{4it}\mathbf{1}\{\text{Tiered Cost Sharing}\}_{jt} + \beta_{5it}\mathbf{1}\{\text{LIS eligible plan}\}_{jt} + \beta_{6i}\mathbf{1}\{\text{Brand}\}_{j}\end{aligned} \tag{3}$$

This specification assumes that conditional on insurer, differences in plans can be accounted for by the deductible level, the initial coverage limit and a set of indicators on whether or not the plan offers partial coverage in the gap, whether the plan uses fixed dollar co-payments or co-insurance percentage, and whether the plan is eligible to enroll individuals with low-income subsidy. The included characteristics capture a substantial amount of variation among plans, since many features that are not explicitly included, such as the quality of services, pharmacy network quality, and the generosity of drug formularies are insurer-level rather than plan-level characteristics. The features of the plans included in this specification correspond to the information that individuals had readily available from front-end consumer advertisement materials by a typical insurer on its 1-3 different contracts.

Beneficiaries in health insurance markets differ in two key ways - in their preferences for different contract features, as well as in how costly they are for the insurers. Individual preferences, in turn, may reflect both individual health risk as well as horizontal tastes and risk aversion that may or may not be correlated with the expected drug spending. To capture these features of insurance demand in the model, I first allow for rich observed heterogeneity of preferences in the specification of marginal utility from contract features. Importantly, I allow individual preferences to depend on the individual's health risk. I use the full set of demographics observed in the data - age, gender, and race; as well as proxies for expected spending - risk scores and an additional flag for having end-stage renal disease diagnosis, which identifies especially high risks. Vector D_{it} records this demographic information and risk measures: $D_{it} = \{\text{age}_{it}, \text{gender}_i, \text{race}_i, \text{risk score}_{it}, \text{esrd indicator}_{it}\}$. The reduced-form evidence for adverse selection in Section 3.1 suggests that individuals may have more private information about their expected spending than what is accounted for by risk scores. Thus, I additionally allow for unobserved heterogeneity in preferences for the key financial features of contracts. This is achieved by specifying random coefficients on three key features of the contracts: deductible, initial coverage limit, and partial gap coverage. The distribution of the random coefficients is assumed to be normal. I interpret this unobserved heterogeneity as at least partially likely stemming from the private information about health risk not captured in risk scores as well as from the heterogeneity in risk aversion.[13]

The assumption behind this specification of heterogeneity is that individuals are aware of their previous medical diagnoses and what these diagnoses typically imply for drug expenditures. In other words, the mean of random coefficients is allowed to depend on risk scores, but not on the realized spending. This advantage of such specificationversus the one that would include the observed ex-post spending is twofold: first, it takes an ex-ante choice perspective and does not impose that individuals have exact knowledge about their ex-post realized spending, but rather have a general understanding that certain diagnostic groups on average cause higher drug spending; second, it removes the concern about moral hazard in the choice model.[14] An additional dimension of heterogeneity is conceivable with respect to the switching cost. If switching costs are, for instance, interpreted as search costs, we may think that, for example, older and sicker individuals have higher switching costs. To capture this possibility, I allow the switching costs to differ among individuals of different demographic groups and risk types. All in all, the coefficients on the contract characteristics and the lagged dependent variable are specified as follows:

$$\beta_{it} = \pi^{\beta} D_{it} + \psi_i^{\beta}, \text{ where } \psi_i^{\beta} \sim N(\psi^{\beta}, \sigma^2) \tag{4}$$

$$\gamma_{it} = \pi^{\gamma} D_{it} + \psi^{\gamma} \tag{5}$$

[13]Given the results in Abaluck and Gruber (2011, 2013), it is also plausible that some of the estimated heterogeneity captures the lack of information or incomplete understanding of contract features by enrollees. To the extent that counterfactual policies considered do not alter the mapping of contract characteristics into individuals' assessment of contracts and choices, but rather ask how individuals would reassess their choices if encouraged to choose again, we can remain agnostic about the exact sources of heterogeneity from the policy-making perspective.

[14]To verfiy that this choice of specification does not drive the results, I have estimated the model with ex-post spending as part of the mean of random coefficients, with and without risk scores as additional mean shifters. The outcomes of the model and of the counterfactuals do not depend on this specification choice.

Assuming that an individual chooses the plan that maximizes his or her utility, the model allows expressing the probability of the beneficiary choosing different plans in his or her choice set as a function of parameters. I then use the maximum likelihood estimation approach to find the values of the parameters that best rationalize observed choices.[15] Since in the data I can track the same individuals making several consecutive choices, the estimation utilizes this panel structure, explicitly modeling the probability of a sequence of choices. While assuming the extreme value distribution for the taste shocks produces a closed-form probability expression conditional on the realization of the random coefficients, the unconditional probability involves integrating out the normally-distributed random coefficients. The latter implies that there is no analytic closed-form solution for the probability integral that is part of the log-likelihood function. Thus, the model is estimated using a simulated maximum likelihood (MSL) procedure as described in Train (2003, 2009) and Hole (2007).

Identification. The identification of the parameters relies on several unique features of the data. First, to recover the switching costs parameter γ, we have to consider two issues: distinguishing between the "spurious" versus "structural" state-dependence and the initial conditions problem.[16] The inclusion of the unobserved individual heterogeneity through random coefficients into the model addresses the first issue in a way that is standard in the literature. The assumption is that the normal distribution in random coefficients correctly captures the heterogeneity, and thus the lagged dependent variable parameter estimates the "structural" part of state-dependence. Moreover, in 2007, 2008, and 2009, there were cohorts of 65 year olds that first became eligible for Medicare and entered the Part D program anew without switching costs. This implies that in years 2007-2009 of the data I observe individuals choosing with

[15] As these derivations are standard for a mixed logit model, I omit the details in the paper.

[16] The concern in the first issue is that the lagged dependent variable in the utility function, which is capturing the switching cost, will be correlated with (or rather directly a function of) an individual-specific preference parameter. To illustrate, in a generic binary non-linear dynamic panel model, this would imply that in $y_{it} = 1\{\beta x_{it} + y_{i,t-1}\gamma + \alpha_i + \epsilon_{it} > 0\}$, $y_{i,t-1}$ is a function of α_i. Thus, if α_i is unaccounted for and left in the unobserved part of the utility function, the identification assumptions about ϵ_i will be violated. The literature on the non-linear dynamic panel data discusses the two broad approaches to this problem - assuming a parametric distribution for the unobserved individual heterogeneity ("random effects"), or trying to difference out the individual effects without functional form assumptions ("fixed effects"), where the latter approach encounters a lot of challenges given the non-linear nature of the model. Honoré (2002) and Honoré and Tamer (2006) provide an excellent discussion.

and without switching costs from the same menu of contracts. The latter feature greatly aids in separating the persistent individual heterogeneity from the switching friction. The initial conditions problem does not arise in the current setting, as I observe the first choices for all individuals in the estimation sample, because the year of the program's launch is recorded in the data.

In addition, the descriptive evidence in Sections 2 and 3 suggests that there is substantial variation in the prices and the characteristics of plans in each year of the program. Such variation is important, since if the environment were very stable, we couldn't expect to observe any changes in choices either with or without costly switching.[17] The cross-sectional variation in the non-price features offered by different plans (such as zero deductible) is generated by the insurers' strategy of offering menus of several vertically-differentiated contracts. This strategy is stable over time, suggesting no contemporaneous responses to aggregate demand shocks. A lot of time-series variation in non-price contract features is generated by the changes in the minimum standard policy that annually adjusts the standard deductible and initial coverage limits. The variation in premiums stems from two sources. First, insurers set different relative prices for the two or three contract types in their contract menus.[18] Second, insurers set different prices for the same contracts in different geographic regions. All these features of the data combined, allow for the identification of preferences separately from the switching costs that are hard to distinguish in more typical observational choice data settings.

Naturally, the rich cross-sectional and time-series variation in the premiums of insurance contracts comes from observational data and not pricing experiments, suggesting that endogeneity concerns are warranted. The stochastic component of the utility function ϵ_{ij} may include unobserved characteristics of contracts that are correlated both with premiums and individual choices, leading to an omitted variables bias. For example, an insurer could be advertising a

[17] Dube et al. (2010) discuss the importance of observing variation in the choice set for the identification of structural state-dependence; they utilize promotions as generating such variation.

[18] Some of the variation further reflects the complex price-setting mechanism. The premiums faced by individuals reflect only a small fraction of the actual prices charged by the insurance providers to Medicare. These premiums are constructed through a bidding mechanism that ties the individual premiums to the average of prices charged by all insurers to Medicare; thus, insurers do not know exactly which premiums individuals will face for the plans in advance of setting prices.

particular contract in its menu more than other contracts and setting the price of this contract higher/lower because of that. Note that conditioning on the insurer-specific fixed effects does not resolve this issue if insurers are advertising a particular contract in their menu. One example of such situation in the Medicare Part D setting is the endorsement of selected contracts by a well-known third party. Several contracts in the portfolio of one insurer were endorsed by the American Association of Retired Persons; as the names of the contracts and insurers are not observed in the administrative data due to commercial privacy restrictions, we cannot include a variable measuring this endorsement directly. At the same time, it would be natural to assume that the AARP endorsement both leads individuals to select this contract with a higher probability and allows the insurer to raise prices either to exploit the less elastic demand, or to cover costs that may arise from the marketing relationship with the AARP. Having such unobserved characteristics in the stochastic portion of the random utility specification violates the assumption of no correlation between the observed and unobserved components of utility. The standard approach in the Medicare Part D literature (Abaluck and Gruber (2011); Heiss et al. (2013)) has been to assume that the rich observed characteristics capture all the relevant information about choices. In this paper I utilize an instrumental variables strategy based on the observations of contracts' expenditures in the administrative data to improve upon this approach.

To correct for potential endogeneity, we need an instrumental variable that affects the contract premiums, but is not correlated with the contract characteristics not observed in the utility function. Since the costs of insurance contracts depend almost entirely on the prescription drug claims submitted by their enrollees, I use mean lagged realized claims as an instrument for current premiums. As Medicare regulation suggests, and the pricing model in Section 4.2 confirms empirically, contract prices are strongly conditionally and unconditionally correlated with lagged mean realized claims in the contract. To utilize the lagged mean realized claims in each contract as an instrumental variable, we need to assume that the variation in this variable is independent of the unobserved contract characteristics conditional on the observed contract characteristics. For example, we need to assume that the AARP endorsement does not affect the level of realized risks in the contract. This assumption appears plausible, as long as we believe that the unobserved characteristics do not screen risks. The latter is indeed very likely, as the reduced-form analysis of Section 3.1 suggests that risk-screening happens primarily on the gap coverage margin. To operationalize the instrumental variables estimation, I utilize the control function approach.[19]

4.2 Empirical model of contract pricing

To analyze the effects of switching costs and the minimum standard regulation on risk-sorting in Medicare Part D in Sections 5.2 and 5.3, I need to account for how insurers may adjust the contracts they offer in response to the hypothesized changes in the environment. One standard approach at the supply-side analysis would be to assume that premiums are determined as an outcome of a pricing game, such as Bertrand, and are set as a mark-up over the marginal cost. Making an assumption about the type of game that insurers play and deriving mark-ups using standard first-order-condition inversion is problematic in the Part D setting due to the substantial regulatory intervention into insurers' price-setting. The premiums that individuals face on this market are not set directly, but are the outcomes of a "bidding" mechanism run by Medicare. This mechanism determines the payments that insurers get from the enrollees in premiums and from the government in subsidies. Building a pricing model that accounts for this regulatory mechanism in Medicare Part D is beyond the scope of this paper and is pursued

[19] Petrin and Train (2009). Formally, the premium for contract j is a function of observed contract characteristics ϕ_j, a variable that affects premiums, but doesn't otherwise affect the choice decisions z_j , and the remaining unobserved term κ_j (Section 4.2 discusses the details of the pricing function in the Part D setting): $p_j = f(\phi_j, z_j, \kappa_j)$. The endogeneity concern arises if κ_j is correlated with ϵ_{ij} in the utility function. Assuming linearity and additive separability of the unobserved component we have: $p_j = \lambda_\phi \phi_j + \lambda_z z_j + \kappa_j$, which is the first stage familiar from a linear IV model. As the choice model is not linear in price, however, the 2SLS technique cannot be applied. One alternative is the control function approach. The idea of this approach is to empirically estimate κ_j and condition on it (or its function) explicitly in the utility function. In practice, κ_j is calculated as the residuals of the first-stage regression of premiums on the observed contract characteristics included in the utility function and the lagged claims instrument. In the second step, a linear control function $CF = \omega\hat{\kappa_j}$ is included into the utility function:$u_{ijt} = -\alpha p_{jt} + \beta_{it}\phi_{jt} + \gamma_{it}\mathbf{1}\{\text{Default}\}_{ijt} + \omega\hat{\kappa_j} + \epsilon'_{ijt}$. I assume that the stochastic part of the utility function ϵ'_{ijt} has an iid extreme value component that is independent of κ_j with the remaining component distributed jointly normal with κ_j. This assumption returns a mixed logit model with mixing over the selected characteristics of the contract as well as the error component (Villas-Boas and Winer, 1999).

in follow-up work.[20]

Instead, here I utilize a hedonic-style approach to empirically relate the premiums faced by the beneficiaries to the risk portfolio and the characteristics of plans. To alleviate a concern about potentially radically different pricing techniques in the first year of the program due to "loss-leader" strategies that, anecdotally, may have been pursued by some insurers in Part D (and if so, quickly reversed in 2007 as the pricing time series would suggest), I only utilize data on 2007-2009 price that should reflect a more stable pricing equilibrium that is more applicable to counterfactuals without consumer inertia. While this approach does not impose an explicit model of supply-side pricing incentives, it does rely on the information about how the regulator outlines the pricing process for the plans. When insurance plans submit their annual bids to the Part D program, Medicare requires them to "justify" the economic validity of the bids. Participating insurers have to provide information about the spending experienced by the current enrollees in a given plan in the previous year and how the plan projects these spending will change in light of any planned changes in plan characteristics (usually those driven by changes in the minimum standard regulation).[21] I therefore include the moments of the lagged spending distribution and the key financial characteristics of the plans as the primary components of this pricing regression. Medicare allows plans to include administrative costs and desired profit margins for the plans, which I assume are insurer-specific and so can be picked up by insurer fixed effects. The full specification takes the following form:

$$\begin{aligned} E[Y_{jbt}|.] &= \alpha_b + \delta_r + {M'}_{jbt-1}\beta + \\ &+ \gamma_1 Ded_{jbt} + \gamma_2 ICL_{jbt} + \gamma_3 1\{PartialGap\}_{jbt} \end{aligned} \tag{6}$$

where j indexes plans (where "plan" is region-specific), b indexes insurers (brands), r indexes

[20] Decarolis, Polyakova, Ryan "The Welfare Effects of Supply-Side Regulations in Medicare Part D", 2014 working paper

[21] For example, from CMS 2007 regulation ("Call Letter"): "In order to prepare plan bids, PDP Sponsors will use HPMS to define their plan structures and associated plan service areas and then download the PBP and Bid Pricing Tool (BPT) software. For each plan being offered, PDP Sponsors will use the PBP software to describe the detailed structure of their benefit packages and the BPT software to define their bid pricing information. Each formulary submitted by April 17, 2006, must accurately crosswalk to a plan (or set of plans) defined during the bid process. The combination of the PBP and BPT for a plan comprises a bid."

34 Part D regions, t indexes years. Y_{jbt} is the annual premium charged to the beneficiaries by plan j of insurer b in year t. Vector M contains several moments of the distribution of drug expenditures experienced by plan j in year $t-1$, including the mean, the standard deviation, the inter-quartile range and the tail percentiles. The moments of the lagged risk distribution and the key characteristics of the plans together with the region and insurer fixed effects account for 80% of the variation in the data on premiums over years 2007-2009. As expected, plans with higher lagged mean annual spending, or plans that offer a lower deductible, higher ICL or partial coverage in the gap are more expensive.[22]

5 Results

5.1 Parameter estimates and assessment of fit

Parameter estimates Table 5 records the detailed parameter estimates of the preferred model specification with the control function instrumental variables approach.[23] The willingness-to-pay estimates in the following discussion are calculated by dividing the parameter estimate for a given plan characteristic or demographic interaction by the coefficient on price. The first estimate of interest is the magnitude and the heterogeneity of the switching cost. The switching cost is estimated to be large, but not to vary among demographic and risk groups in an economically significant way. For example, a 75 year old white female with no end stage renal disease and an average health risk with expected total drug spending of about $\$2,190$ is estimated to face a switching cost of $\$1,164$; while an 80 year old white male with no ESRD and twice the expected risk is estimated to face a switching cost of $\$1,253$.[24] While there is some

[22] More detailed results are reported in Appendix Table A.5.

[23] Table 6 compares parameter estimates from several specifications of the choice model. We observe that instrumental variables strategy slightly increases the magnitude of the price coefficient, as we would expect if the unobserved characteristics in the utility function are positively correlated with the premium and the choice probability.

[24] The estimated order of magnitude is roughly similar to other findings in the health insurance literature. Handel (2013) estimates the switching costs to be about $\$2,000$ in the context of employer-provided health insurance; Nosal (2012) estimates the switching cost in Medicare Advantage health plans for seniors to be about $\$4,000$. Using a different choice model, Abaluck and Gruber (2013) estimate the switching costs in Medicare Part D to be on the order of $\$600-\$700$, which is lower than what my estimates suggest in the willingness-to-pay

risk-related heterogeneity in the magnitude of the switching friction, it is not large enough to suggest that whether lower or higher risks tend to stay in their plans could drive the selection patterns. I estimate that switching cost are increasing in the observed risk at the rate of \$74 for an additional unit of risk score; in other words, an individual with twice as high expenditure risk has a 6% higher switching cost. The switching cost is also increasing with age at the rate of about \$3 per year.

The second set of estimates sheds light on the role of information about risk for the choice of contracts. Consistent with the reduced-form tests for adverse selection in Section 3, I find that beneficiaries with higher health risk value the generosity of coverage more than individuals with lower risks. For instance, individuals with a risk score that is twice the Medicare average, are willing to pay \$22 more, on average, for each additional \$100 of the initial coverage limit, about \$120 more to enroll in plans that have fixed co-pays rather than co-insurance, and about \$230 more, on average, to be in a plan that offers partial gap coverage, than otherwise observationally identical beneficiaries with average risk. Moreover, consistent with the hypothesis that beneficiaries have private information about their expected spending beyond the diagnostic information accounted for by the risk scores, I estimate a large degree of heterogeneity in the valuation of the key financial features of the contracts. For instance, the estimate of the standard deviation in the valuation of the partial gap coverage is \$250 dollars. This implies that even for individuals within the same demographic and risk group, the differences in the valuation of the coverage in the gap may be very large. Put differently, even if my diagnostic information does not suggest that I am a particularly high risk individual, I may have private information that my expected spending is going to fall into the donut hole, or I may be very risk averse, warranting me to pay around 50% more in premiums to have a contract with partial coverage in the gap.

Model fit and descriptive patterns generated in the model To assess whether the model can capture the key patterns in the data, for each beneficiary-year observation I simu-

terms, although they similarly find that beneficiaries are roughly 500% more likely to choose a plan that is the "default" plan for them in a given year. Using aggregate data and a dynamic demand model, Miller and Yeo (2012) estimate the switching cost in Part D to be \$1,700, which is higher than my estimates.

late contract choices using the estimated parameters and the observed information about each beneficiary-year. The model performs well in capturing individual preferences for specific plans as well as the average expected risk in each plan.[25] As the counterfactual simulations in the next two sections rely on the ability of the model to capture the role of switching costs and the allocation of risks across contracts of different types, I perform two more checks of the model that test these two aspects. In particular, I generate the key descriptive patterns from Section 3 in the model. First, I check whether the model accurately predicts that individuals are likely to choose their default contracts, or, in other words, whether the estimated switching cost has a meaningful magnitude. The model predicts that, for example, in 2007, 86% of enrollees would choose their "default" plans, which is close to the observed share of 90%. Second, I check whether the model is able to predict the differences in the distribution of risks across contracts of different generosity. This tests whether the model can generate the same adverse selection patterns as the ones observed in the data. Figure 5 plots the simulated empirical CDFs of risk by the type of contract. The graph clearly shows the adverse selection patterns, with more generous contracts enrolling higher expected risks. The distribution predicted by the model's simulated enrollment choices look quantitatively similar to the distributions of risk in the data. In the counterfactual simulations of the next two sections, I will analyze how this distribution changes in response to the reductions in the switching costs and regulatory interventions that shift the features of the minimum standard.

5.2 Quantifying the effect of switching costs on risk allocation

The estimated model of contract choice and contract pricing in Medicare Part D allows me to explore how switching costs are altering the allocation of risks among contracts over time. I begin with the analysis of two distinct examples of the Part D contract space development that resemble the conceptual mechanisms in Figure 1.[26] These examples illustrate two conceptual

[25]Figure 5 illustrates the result of this exercise. Table A.3 further summarizes the fit of the choice model. I report the simulated and the observed values for three moments in the data - enrollment shares, average ex-post spending, and average risk scores along the 4-type plan typology and for top 2 insurer brands.

[26]More detail and graphical representation of the empirical examples is available in Appendix 7.3.

points. First, whether switching costs ameliorate or exacerbate adverse selection depends crucially on the evolution of the relative generosity and the relative prices of contracts in relation to the contract characteristics at the first choice incidence. Second, accounting for changes in the contract space beyond price adjustments may have important implications for the conclusions we draw about the interaction between adverse selection and switching costs. In particular, in insurance settings akin to Medicare Part D, insurers can control when they let contracts enter the market and can change their contracts on a variety of dimensions in addition to premiums. These margins of adjustment are crucial for the distribution of risks and the interaction with switching costs. Moreover, they imply that in an environment where adjustments to the contract space are undertaken simultaneously by a large number of insurers along several different dimensions, the net effect of the switching costs on the allocation of risks is an inherently empirical question. Indeed, the simulation results for the evolution of the full contract space in Medicare Part D presented below suggest that, costly switching could have helped support an adversely selected equilibrium in this environment. While this conclusion in principle hinges on the assumption that contract characteristics, except for price, followed the pattern observed in the data (in the sense that I do not let choices of coverage levels be endogenous to inertia or other market forces), relative to a standard product-market setting, some product characteristics decisions here were driven by incremental exogenous changes to the Standard Defined Benefit that likely do not depend on inertia. The counterfactual exercises in this section are similar in spirit to the analysis in Handel (2013) in the employer-provided insurance setting with no regulatory intervention.

Simulation of risk allocation with costly and costless switching on the full contract space Table 7 illustrates the results of a counterfactual nudging policy simulation that completely shuts down inertia. The counterfactual is implemented as a policy shock in year 2009 of the program. While in practice the complete elimination of inertia may be impossible, where we can imagine an intervention that dropped all existing coverage required individuals to choose their plans anew. A policy shock style simulation rather than a re-simulation of the market from 2006 onwards without inertia allows this counterfactual to be more realistic in assessing a potential policy reform and it also makes it more compatible with the simulation of the ACA policy in the next section. I allow for two scenarios in the current simulation. One scenario takes the observed contract premiums as given; the other scenario allows insurers to endogenously adjust premiums to the new sorting patterns according to the contract pricing model described in Section 4.2. Both scenarios take the other features of the contracts, as well as the observed individual demographics and expected spending, as given. Table 7 focuses on four moments of the data - enrollment shares, average expected spending, average ex-post spending, and average premiums. All results are aggregated to the three types of plans. I present the results for year 2009, which serves as a summary statistic of the market development with switching costs since 2006.

I perform three simulations of the model. The first one, marked with A, simulates the model with estimated switching costs, taking the observed prices and pre-2009 choices as given.[27] This step creates a baseline that takes into account the simulation error and which I use instead of the actual observed 2009 outcomes as a comparison benchmark in analyzing the scenarios without switching costs.[28] Baseline enrollment shares, risk sorting, and prices paid, closely resemble the descriptive evidence presented earlier. *Type 2* plans with reduced deductible and no gap coverage have the highest enrollment share - 72% and about average (in-sample) risk profile of $\$1,926$ expected annual spending. *Type 3* plans with partial gap coverage are adversely selected with the average expected spending of $\$2,368$. The difference in the average risk between the *Type 1* and *Type 3* plans is substantial - expected spending in type 3 plans is 22% or \$526 higher. *Type 3* plans, with just 9% enrollment share, have annual premiums that are on average \$400 higher than in plans without partial gap coverage.

The next simulation, marked with B, sets the switching cost parameter γ in the utility function to zero for all individuals, which eliminates the inertia channel. This simulation shows

[27]As part of specification checks, I tested the alternative approach of simulating the model from 2006 onward rather than taking the observed lagged choices as given in 2009. While the simulation error accumulates starker over several simulation periods in this case, this doesn't change the analysis in a substantive way. The baseline approach pursued in the main text renders itself better to the interpretation of the switching cost reduction as a sudden policy shock in one year. This point is irrelevant for the simulations without switching costs, since lagged choices do not enter the utility function.

[28]As discussed in Section 5.1, the simulated baseline is close to the observed moments in the data.

what enrollment shares and risk sorting would look like absent choice frictions if insurers kept their prices at the observed 2009 levels.[29] Two key changes are notable relative to the baseline with switching costs. First, there is some change in enrollment shares, as with costless switching individuals can actively respond to adjustments in the price and other characteristics of contracts relative to 2006. Enrollment in *Type 1* plans increases by 5 percentage points, as enrollees respond to lower premiums available for these plans. Respectively, the enrollment in *Type 2* and *Type 3* plans decreases. Beneficiaries also select cheaper contracts within these types of plans, paying on average \$27 and \$75 less in annual premiums. Second, the simulation predicts smoother distribution of the average expected spending among plans. Both *Type 1* and *Type 3* plans experience changes in average risks that move them closer to the average expected spending in the sample. Thus, even before allowing the insurers to respond to the change in sorting patterns by adjusting their prices (which we would expect to amplify these effects), we see that removing switching costs mutes adverse selection in this setting in the sense of balancing the distribution of risks across plans.

In the last simulation, I allow insurers to respond to the changes in the sorting patterns induced by the removal of switching costs. Insurers are assumed to adjust their premiums according to the stylized pricing model outlined in Section 4.2. To calculate new prices I use the simulation of demand for individual contracts without switching costs to calculate the new allocation of risks, which in conjunction with the pricing equation gives me new prices. The demand simulation uses the existing draws of random coefficients for the individuals, so that those individuals who, for instance, were simulated to have a high idiosyncratic preference for coverage in the gap, continue having this preference - this in turn affects risk-sorting and prices. Allowing prices to adjust to the sorting without switching costs amplifies the results in B. Individuals now pay on average an additional \$25 less (\$100 less relative to the baseline

[29] While this scenario may appear very unrealistic for a competitive market and its primary purpose is to illustrate the separate effects of risk-sorting response and supply-side adjustments, the regulatory environment in Medicare Part D is such that no changes in the prices of contracts from the individual's point of view is in fact conceivable. Since CMS regulates how bids by insurers get translated into beneficiary premiums, adjusting this mechanism so as to freeze relative prices faced by individual beneficiaries in addition to policies that reduce switching costs is, in principle, possible.

simulation) for the most generous *Type 3* plans with partial gap coverage that now enroll slighly lower average expected risk, which triggers an increase in their enrollment share relative to the simulation in B. Risk sorting among plans becomes even less acute and all of them move even closer to the average. The relative average risk between the *Type 1* minimum standard plans and *Type 3* plans with partial gap coverage falls to $414, which is a 21% decrease in the risk difference relative to the baseline simulation. Figure 6 plots the counterfactual risk CDFs by the type of plan for this simulation. The graph demonstrates that the decrease in risk difference holds throughout the whole distribution of risks and not only for the mean.

5.3 Effect of minimum standard regulation on the allocation of risks with costly and costless switching

The analysis in the preceding section is consistent with the idea that switching costs significantly alter the response of risk-sorting to changes in insurance contracts. It also demonstrates that the effect of switching costs on adverse selection depends critically on the exact evolution of the contract space relative to the initial conditions. For public health insurance environments, this insight has important policy implications. Multiple regulatory interventions that directly affect the contract space are ubiquitous in such environments. The fact that switching costs alter the response of risk-sorting to contract changes implies that switching costs will also significantly alter how regulatory interventions impact the allocation of risk.

I explore this hypothesis on the Part D's minimum standard regulation. The policy experiment that I consider corresponds to the changes in the Part D regulation envisioned under the Affordable Care Act (ACA). One provision of the ACA is that the standard plan in Part D will no longer have the gap in coverage. This "filling" of the donut hole would then fill the contract space gap that resulted from the unraveling of plans with full donut hole coverage on the private market. The Centers for Medicare and Medicaid Services has been concerned with the availability of Part D coverage that goes beyond the minimum coverage requirements since the inception of the program. CMS recognized that such coverage may be valuable for beneficiaries with costly health conditions, but also that the existing risk-adjustment and reinsurance policies may not be sufficient to overcome adverse selection and the resulting profitability concerns of insures in offering such coverage. CMS has experimented with the reinsurance schemes that were supposed to encourage private provision of coverage in the gap.[30] These schemes were potentially sufficient to keep *Type* 3 plans with some coverage in the gap, but, as we have seen, plans with full coverage were not available on the market starting with 2008. Beyond pure choice availability concerns, the research in public health additionally suggested that the "donut" hole could be harmful for beneficiaries with chronic conditions (Stuart et al., 2005; Zhang et al., 2009; Gu et al., 2010). All this contributed to the gap in coverage becoming a prominent issue in public debate and in the media. In light of the failure of the private market to provide the generous coverage, the goal of closing the coverage in the gap via regulatory decree became an explicit part of the political agenda during elections and later a part of the health care reform law. While this policy will certainly decrease out-of-pocket spending on drugs for those beneficiaries that would have ended up in the donut hole, in my counterfactual simulation I show that the efficiency and cost of this change in the minimum coverage requirements could be substantially improved if it were combined with strong nudging policies.

To implement this counterfactual simulation I force all plans to have the level of 2009 catastrophic coverage threshold ($\$6,145$) as their initial coverage limit (originally $\$2,700$). This implies that *Type 2* and *Type 3* plans become identical in terms of their key financial characteristics, as both of them have reduced deductible and are imposed to have the same coverage limit. Even though the financial characteristics of these contracts are now very close, there is still some residual horizontal differentiation across the individual plans within each group, such as which insurer offers the plans and whether the insurer utilizes a co-insurance or co-pay cost-sharing structure. I start the counterfactual policy analysis by quantifying how changes in the minimum standard policy would impact risk sorting across contracts in the status quo with costly switching. I then consider how switching costs are altering this impact.

[30] For related discussions, see for example "CMS Instructions for the Part D Payment Demonstration", May 2005.

The counterfactual simulations suggest that switching costs play a key role in determining the effect of this policy change on risk allocation. Ex ante, we would expect that, since there is no gap coverage dimension of differentiation among plans anymore, this policy should eliminate the acute selection that took place on the gap coverage margin. Indeed, the results reported in Table 8 suggest that exactly this effect takes place, but only in the scenarios, where I shut down the inertia channel. In these scenarios, all types of plans converge to having practically identical pool of risks of around $\$1,950$ in expected spending. In the scenario with costly switching, however, the simulated policy does not change the distribution of risks. Figure 7 plots the counterfactual risk CDFs for this simulation, demonstrating that the result holds throughout the whole distribution of risks and not only for the differences in mean risk. In essence, the presence of switching frictions sustains "artificial" adverse selection, despite the regulation-induced convergence in the generosity of contracts.[31]

6 Conclusion

In this paper I have documented evidence of adverse selection and switching costs in a highly regulated Medicare Part D environment using a parsimonious classification of the contract space and detailed administrative data. I have also shown that switching costs have important interactions with adverse selection in this complex health insurance program. In particular, I have shown that in this environment the initial conditions led switching costs to support an adversely selected equilibrium over time, in the sense that different types of plans would have had more similar average risks if switching were costless. In considering an important channel that drives the changes in Part D contracts over time - the minimum standard policy, I have shown that the market imperfection of costly switching plays an important role in determining the risk-sorting outcomes of this policy intervention, which is targeted at correcting a different market failure - adverse selection. Specifically, I find that in the presence of switching costs,

[31] At the same time, the presence of switching frictions makes it less likely that firms will just discontinue adversely selected contracts in response to the regulatory change. Loosing beneficiaries, even the high-cost ones, may be detrimental to profits, as market shares are hard to gain in this market with sticky consumers.

tightening the minimum standard requirement by "filling" the donut hole is unlikely to have the intuitive effect of balancing risks across different contracts. It is worth emphasizing that while the magnitude of interaction between selection and inertia appears moderate under the pre-ACA regulatory conditions, the counterfactual analyses demonstrate that the issue of locked-in risk distribution due to stickiness of choices is going to have an important effect on the outcomes of ACA reforms in Medicare Part D, and would have benefited from concurrent policies encouraging re-selection of plans.

More broadly, this paper argues that in considering the polices that may improve consumer choice by eliminating or decreasing switching frictions in the increasingly common public health insurance settings with regulated competition, we have to take into account the nuanced interconnections of the different market imperfections with the regulatory instruments targeted at correcting them. The caution, of course, comes from the caveat that in this work I represented the reaction of insurers to policy changes in a very stylized way. Expanding the model to allow insurers to endogenously react to adverse selection and regulation in their choices of contract characteristics in a competitive setting provides a fruitful area for future research.

References

Abaluck, J., Gruber, J., 2011. Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. The American Economic Review 101 (June), 1180–1210.

Abaluck, J., Gruber, J., 2013. Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance. NBER Working Paper no. 19163.

Armstrong, M., Sappington, D. E. M., 2007. Recent Developments in the Theory of Regulation. In: Armstrong, M., Porter, R. (Eds.), Handbook of Industrial Organization, Vol. 3. No. 06. Elsevier B.V., pp. 1557–1700.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., 2013. Who uses the roth 401(k), and how do they use it? NBER Working Paper No. 19193.

Bundorf, K., Levin, J., Mahoney, N., 2012. Pricing and Welfare in Health Plan Choice. American Economic Review 102(7), 1–38.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T., Olsen, T., 2013. Active vs. passive decisions and crowdout in retirement savings accounts: Evidence from denmark. NBER Working Paper No. 18565.

Chiappori, P.-A., Salanie, B., 2000. Testing for Asymmetric Information in Insurance Markets. The Journal of Political Economy 108 (1), 56–78.

Decarolis, F., 2013. Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design? Mimeo, Boston University.

Dube, J.-P., Hitsch, G. J., Rossi, P. E., 2010. State dependence and alternative explanations for consumer inertia. RAND Journal of Economics 41, 417–445.

Duggan, M., Healy, P., Scott Morton, F., 2008. Providing Prescription Drug Coverage to the Elderly: America's Experiment with Medicare Part D. Journal of Economic Perspectives 22(4), 69–92.

Duggan, M., Scott Morton, F., 2010. The Effect of Medicare Part D on Pharmaceutical Prices and Utilization. The American Economic Review 100 (1), 590–607.

Einav, L., Finkelstein, A., Levin, J., 2010. Beyond Testing: Empirical Models of Insurance Markets. Annual Review of Economics 2, 311–336.

Einav, L., Finkelstein, A., Schrimpf, P., 2013. The Response of Drug Expenditure to Non-Linear Contract Design: Evidence from Medicare Part D. NBER Working Paper no. 19393.

Ericson, K., 2013. Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange. American Economic Journal: Economic Policy (forthcoming).

Farrell, J., Klemperer, P., 2007. Coordination and Lock-In: Competition with Switching Costs and Network Effects. In: Armstrong, M., Porter, R. (Eds.), Handbook of Industrial Organization, Vol. 3. No. 06. Elsevier B.V., pp. 1967–2072.

Finkelstein, A., 2004. Minimum standards, insurance regulation and adverse selection: evidence from the Medigap market. Journal of Public Economics 88 (12), 2515–2547.

Finkelstein, A., McGarry, K., 2006. Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. American Economic Review 96(4), 938–958.

Gu, Q., Zeng, F., Patel, B., Tripoli, L., 2010. Part D coverage gap and adherence to diabetes medications. The American Journal of Managed Care 16(12), 911–918.

Handel, B. R., 2013. Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. The American Economic Review (forthcoming).

Hastings, J., Hortacsu, A., Syverson, C., 2013. Advertising and competition in privatized Social Security: The case of Mexico. NBER Working Paper No. 18881.

Heckman, J., 1981a. Structural Analysis of Discrete Data with Econometric Applications. Cambridge, Mass.: MIT Press, Ch. Statistical Models for Discrete Panel Data, pp. 114–78.

Heckman, J., 1981b. Structural Analysis of Discrete Data with Econometric Applications. Cambridge, Mass.: MIT Press, Ch. The Incidental Parameters Problem and the Problem of Initial Condition in Estimating a Discrete Time-Discrete Data Stochastic Process, pp. 179–95.

Heckman, J., 1991. Identifying the hand of past: Distinguishing state dependence from heterogeneity. American Economic Review 81(2), 75–79.

Heckman, J., Singer, B., 1986. Handbook of Econometrics. Handbooks in Economics series, book 2 Amsterdam; Oxford and Tokyo: North-Holland;, Ch. Econometric Analysis of Longitudinal Data, pp. 1689–763.

Heiss, F., Leive, A., McFadden, D., Winter, J., 2013. Plan Selection in Medicare Part D: Evidence from Administrative Data. Journal of Health Economics (forthcoming).

Heiss, F., McFadden, D., Winter, J., 2009. Regulation of Private Health Insurance Markets: Lessons from Enrollment, Plan Type Choice, and Adverse Selection in Medicare Part D. NBER Working Paper No. 15392.

Heiss, F., Mcfadden, D., Winter, J., 2010. Mind the Gap ! Consumer Perceptions and Choices of Medicare Part D Prescription. In: Wise, D. (Ed.), Research Findings in the Economics of Aging. The University of Chicago Press, pp. 413–481.

Ho, K., Hogan, J., Scott Morton, F., 2013. Insurer (In)Efficiency in the Medicare Part D Program. Mimeo, Columbia University and Yale University.

Hoadley, J., Cubanski, J., Hargrave, E., Summer, L., Huang, J., 2012. Medicare Policy data spotlight Medicare Part D: A First Look at Part D Plan Offerings in 2013. KFF Report (October 2012).

Hole, A. R., 2007. Fitting mixed logit models by using maximum simulated likelihood. The Stata Journal 7-3, 388–401.

Honoré, B., 2002. Nonlinear models with panel data. Portuguese Economic Journal 1(2), 163–179.

Honoré, B., Tamer, E., 2006. Bounds on parameters in panel dynamic discrete choice models. Econometrica 74(3), 611–629.

Joskow, P., Rose, N., 1989. The Effects of Economic Regulation. In: Schmalensee, R., Willig, R. (Eds.), Handbook of Industrial Organization, Vol. II. Elsevier Science Publishers B.V., pp. 1449–1506.

Kesternich, I., Heiss, F., McFadden, D., Winter, J., 2013. Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. Journal of Health Economics (2013).

Ketcham, J. D., Lucarelli, C., Miravete, E. J., Roebuck, M. C., 2012. Sinking, Swimming, or Learning to Swim in Medicare Part D. The American Economic Review 102 (6), 2639–2673.

Kling, J. R., Mullainathan, S., Shafir, E., Vermeulen, L. C., Wrobel, M. V., 2012. Comparison Friction: Experimental Evidence from Medicare Drug Plans. The Quarterly Journal of Economics 127 (1), 199–235.

Lucarelli, C., Prince, J., Simon, K., 2012. The Welfare Impact of Reducing Choice in Medicare Part D: A Comparison of Two Regulation Strategies. International Economic Review 53 (4), 1155–1177.

Luco, F., 2013. Switching Costs and Competition in Retirement Investment. Mimeo, Northwestern University.

Miller, D. P., Yeo, J., 2012. Estimating Dynamic Discrete Choice Models of Product Differentiation: An Application to Medicare Part D with Switching Costs. Mimeo, Clemson University and Singapore Management University.

Nosal, K., 2012. Estimating Switching Costs for Medicare Advantage Plans. Mimeo, University of Mannheim.

Petrin, A., Train, K., 2009. A control function approach to endogeneity in consumer choice models. Journal of Marketing Research XLVI.

Stuart, B., Simoni-Wastila, L., Chauncey, D., 2005. Assessing The Impact Of Coverage Gaps In The Medicare Part D Drug Benefit. Health Affairs.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press.

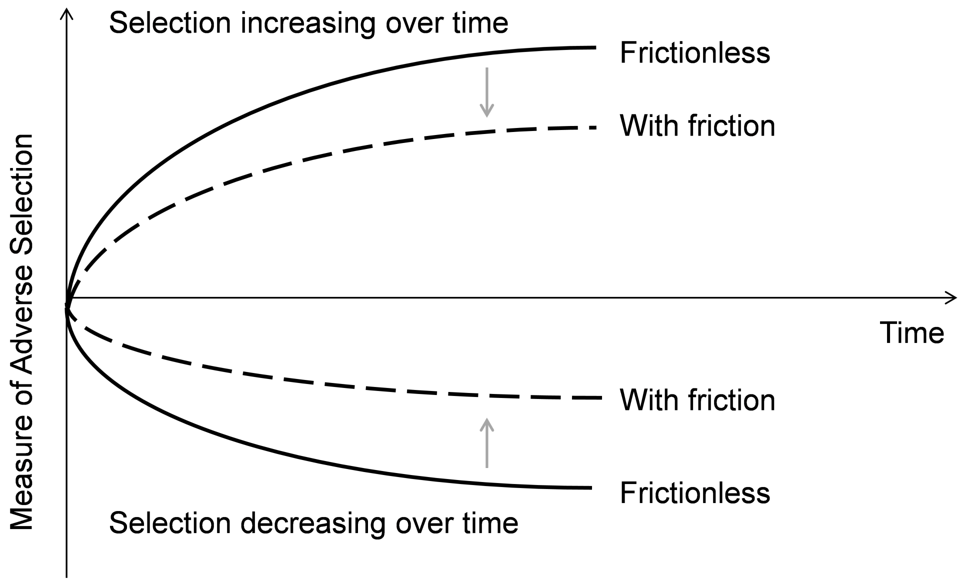
Train, K., 2009. Discrete Choice Methods with Simulation, 2nd Edition. Cambridge University Press.

Villas-Boas, J. M., Winer, R., 1999. Endogeneity in brand choice models. Management Science 45(10), 1324–1338.

Yin, W., Basu, A., Zhang, J. X., Rabbani, A., Meltzer, D. O., Alexander, G. C., 2008. The Effect of the Medicare Part D Prescription Benefit on Drug Utilization and Expenditures. Annals of Internal Medicine 148, 169–177.

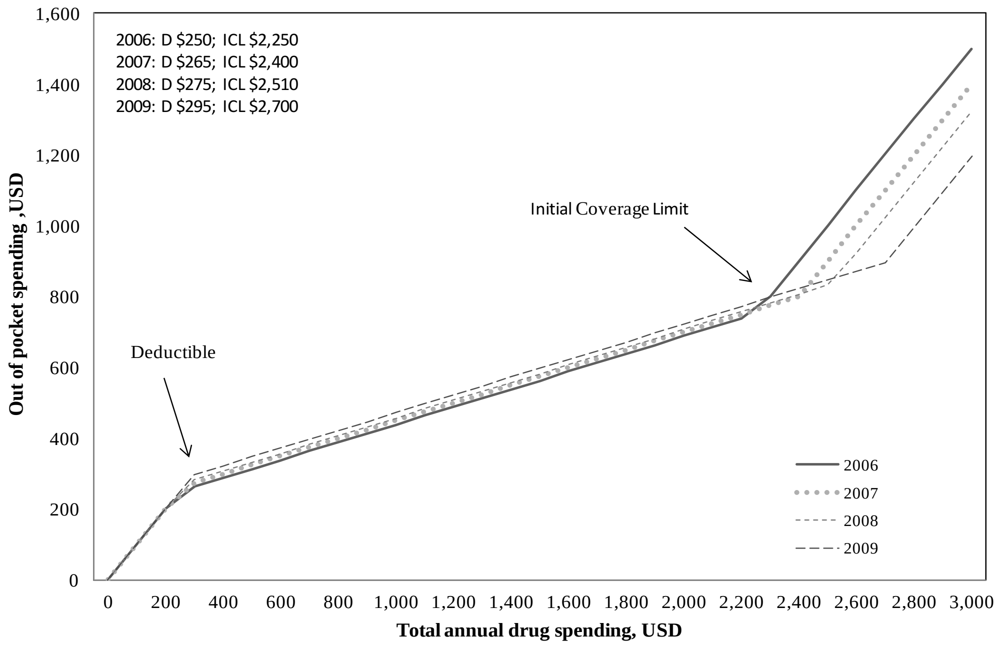
Zhang, Y., Donohue, J., Newhouse, J., Lave, J., 2009. The Effects Of The Coverage Gap On Drug Spending: A Closer Look At Medicare Part D. Health Affairs 28(2), 317–325.

Figure 1: **Conceptual framework: switching costs and adverse selection dynamics**



This stylized graph illustrates that the effect of switching costs on a market with selection depends on the inherent market dynamics relative to the initial conditions. Switching costs slow down the adjustment of the economy to equilibria with different levels of adverse selection. The textbook version of the adverse selection model suggests that adverse selection in a frictionless environment necessarily grows worse over-time, in which case switching frictions slow down the unraveling spiral. In practice, however, in the presence of multi-dimentional insurance contracts, market power, preference heterogeneity and exogenous regulatory interventions into the contract space, adjustments to equilibria with less adverse selection are equally likely. In the latter scenario, switching frictions would slow down this market dynamic, effectively exacerbating adverse selection.

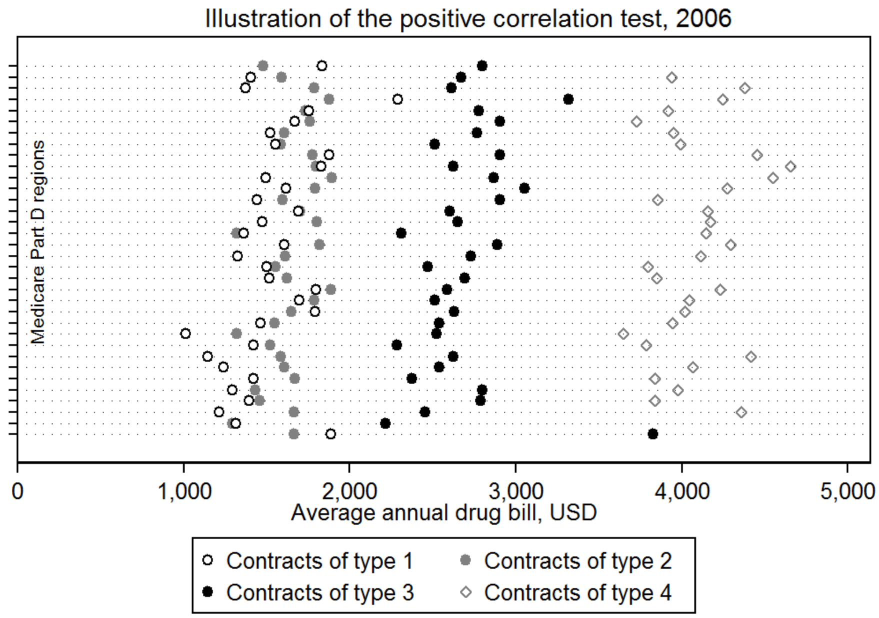
Figure 2: **Minimum standard policy in Medicare Part D: shape of the Standard Defined Benefit in 2006-2009**



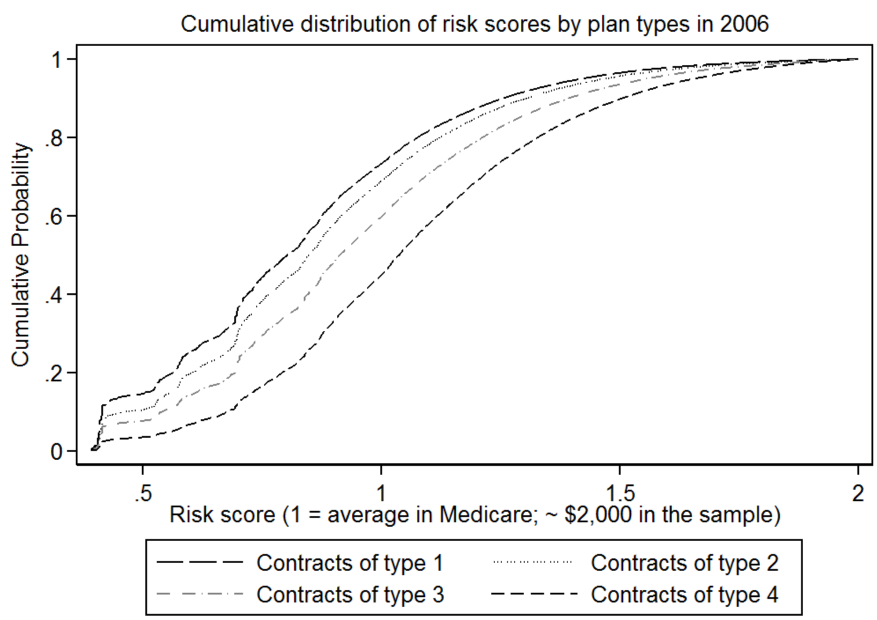
Insurers in the Medicare Part D program are required to provide coverage that gives at least the same actuarial value as the Standard Defined Benefit (SDB). The SDB design features a deductible, a co-insurance rate of 25% up to the initial coverage limit (ICL) and the subsequent "donut hole" that has a 100% co-insurance until the individual reaches the catastrophic coverage arm of the contract. The graph illustrates these features of the SDB by mapping the total annual drug spending into the out-of-pocket expenditure. Consider an individual, who in 2006 was in an SDB contract, and purchased prescription drugs for a total of $3,000. Out of this amount, the individual would pay the deductible of $250, then 25% of the next $2,000 up to the ICL of $2,500, and then 100% of the next $750 in the gap, for a total out of pocket spending of $1,500. As the figure illustrates, the generosity of the SDB changed over time. For example, an individual spending $3,000 on drugs in 2009 would face the out-of-pocket expenditure of less than $1,200.

Figure 3: **Adverse selection in Medicare Part D**

Panel A: ***Positive correlation tests for the presence of asymmetric information***

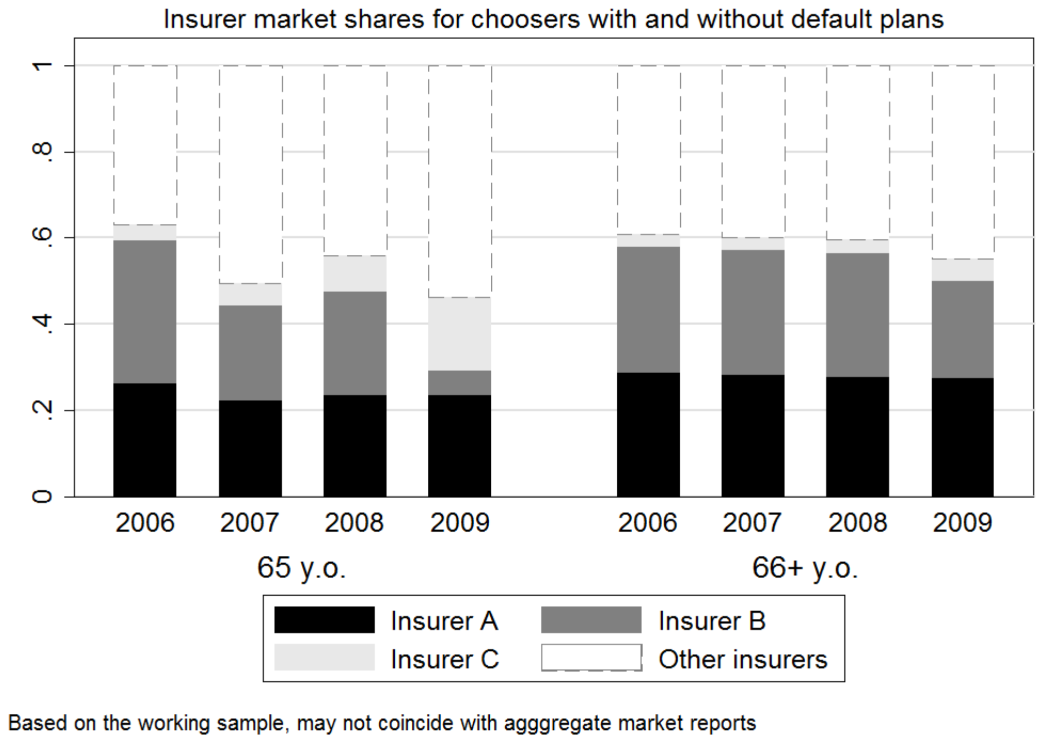


Panel B: ***Distribution of risks by type of plan using risk score measure***



*Panel A: A*verage annual drug expenditures in different types of Medicare Part D plans are calculated separately within each geographic region. Baseline sample. The spending of individuals with enrollment shorter than 12 months (primarily 65 year olds) was extrapolated to the full year. *Panel B:* Baseline sample. Risk scores are based on lagged diagnostic information and not on drug expenditures. The stochastic ranking of the distribution functions visualizes the riskier pool of enrollees recruited by plans with more generous coverage.

Figure 4: **Evidence of switching costs: over time development of insurers' enrollment shares for new and continuing beneficiaries**



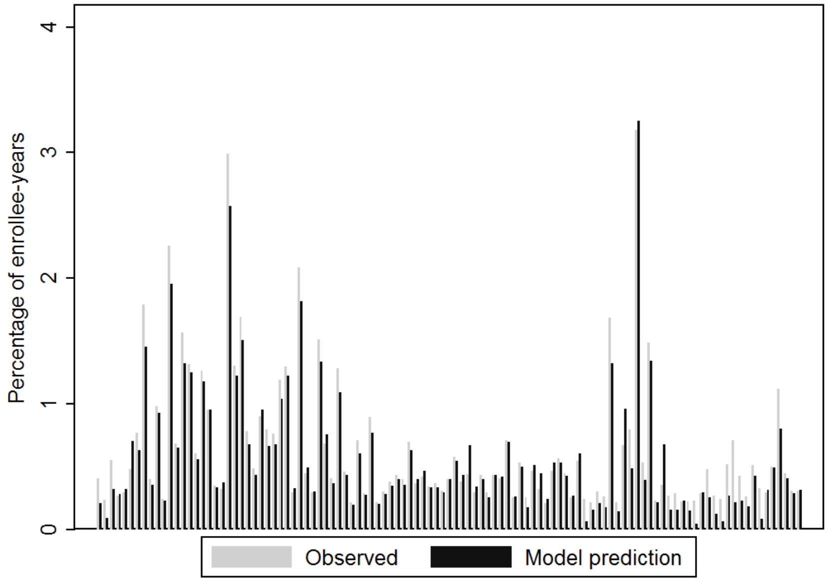
The graph uses data of the baseline sample. We observe that the choices of the 65 year old individuals newly entering the program, who by definition do not have incumbent plans, are much more volatile and responsive to the market conditions over time, than choices of the individuals in the existing cohorts, who usually have the default option of their incumbent plan available. Insurer "identities" here are constructed using contract encryption in the administrative data. Because of the data encryption, separate insurers may have been identified with error. The corresponding commercial identities of the insurance companies are not known to the author.

Aggregate share of continuing enrollees choosing that same plan as in $t-1$ (includes all insurers and all plans):

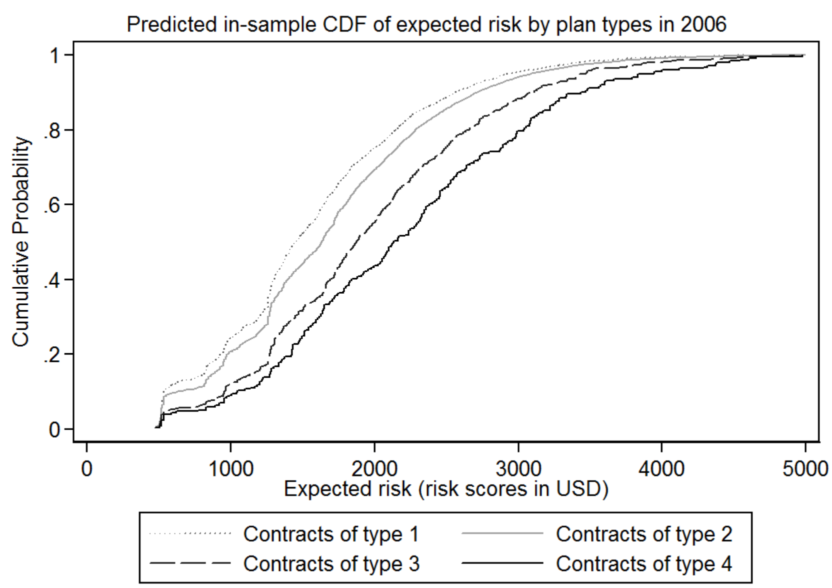
	2007	2008	2009
Probability of choosing default plan for 66+ y.o. enrollees	89.9 %	88.7 %	89.1 %
N	1,089,978	1,162,545	1,194,036

Figure 5: **Model fit and descriptive evidence simulated in the model**

Panel A: ***Simulated enrollment shares***



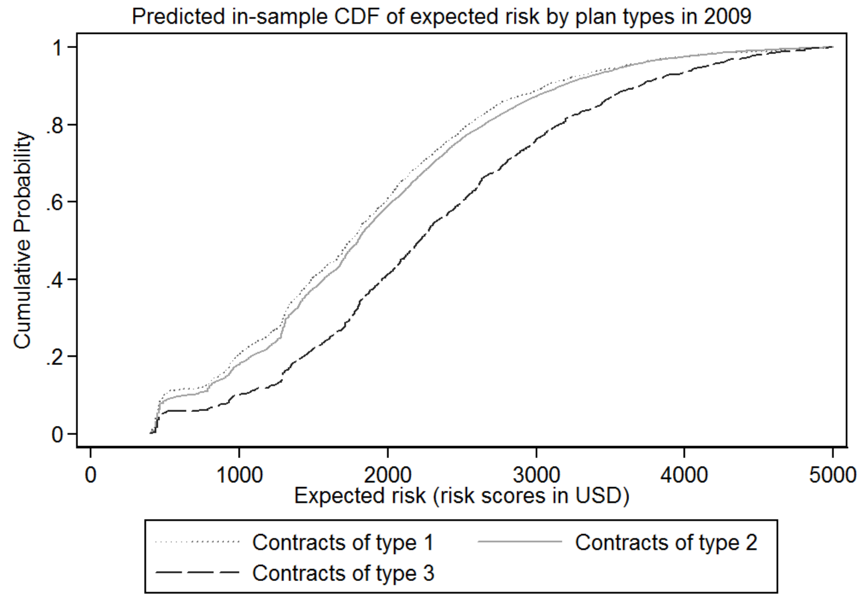
Panel B: ***Simulated distribution of risks***



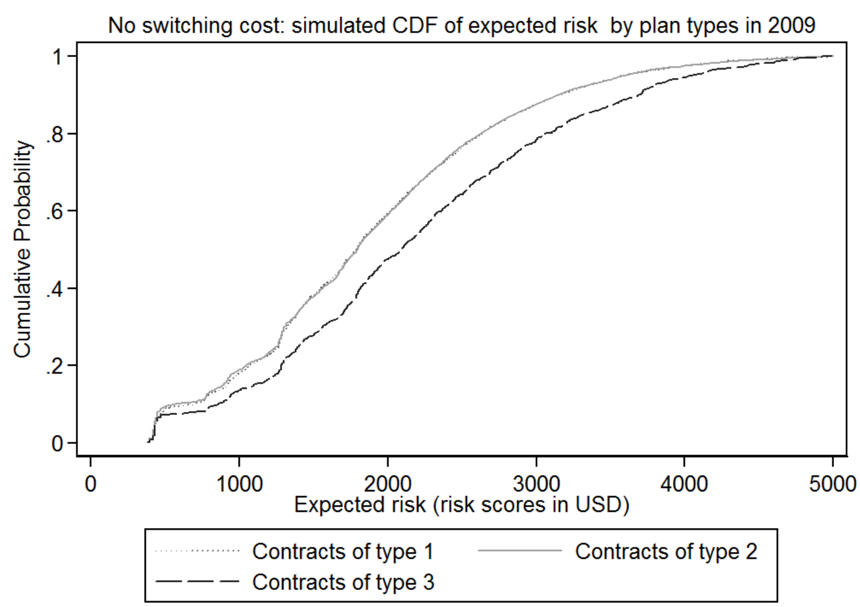
To construct the simulated enrollment and risk distribution, the coefficient estimates of the choice model together with simulated random components were used to find the contract with the highest utility in each individual's choice set. The risk score associated with each individual was then used to plot the distribution of risks. *Panel A* graph displays only top 90 out of 2,357 contracts - each pair of bars in the graph represents a different Medicare Part D plan ("plan" is region-specific).

Figure 6: **Counterfactual risk allocation without switching costs: baseline versus counterfactual distribution of risks by contract type in 2009**

- *Panel A: Baseline simulated distribution of risks*



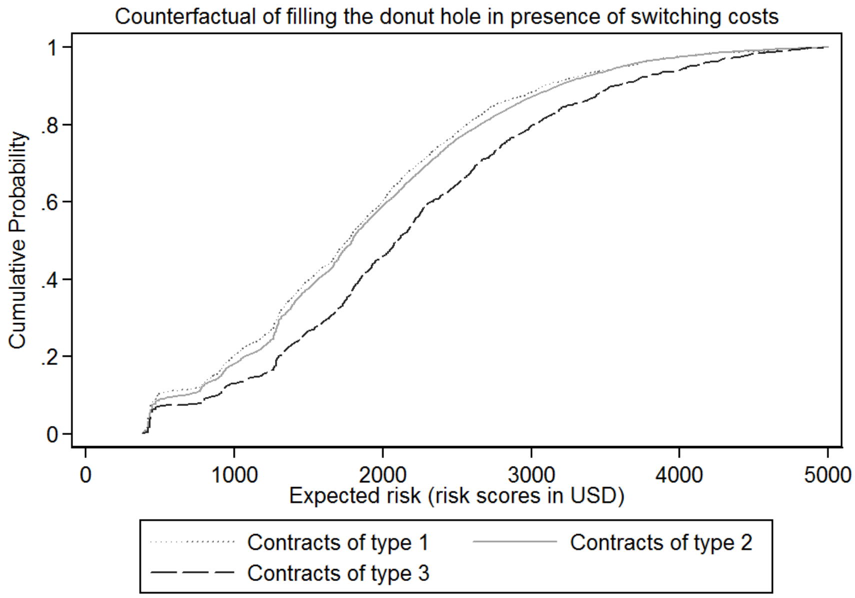
- *Panel B: Distribution of risks without switching costs (with endogenous re-pricing)*



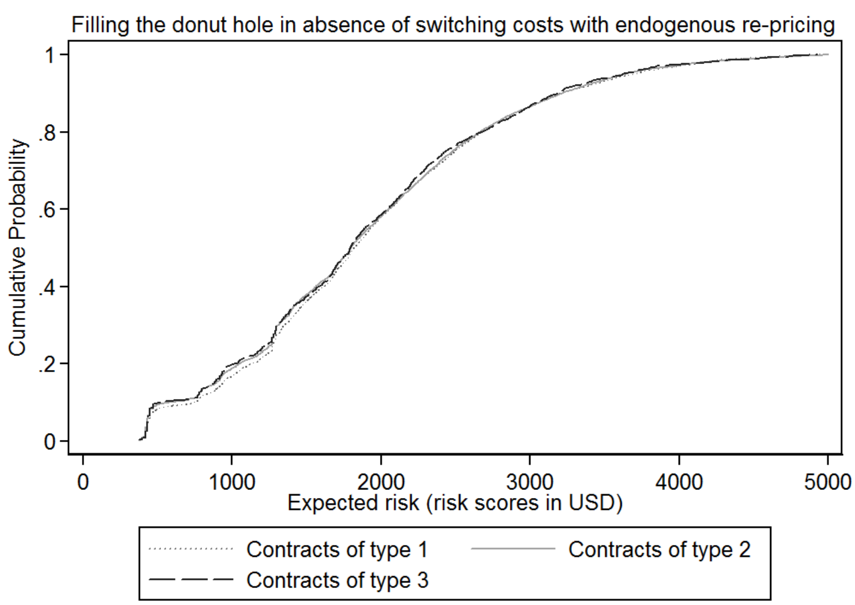
In-sample simulation of the choice model. *Panel A* keeps the estimated switching cost parameter γ. *Panel B* sets $\gamma = 0$. *Panel B* also incorporates price-adjustment by insurers in response to changes in risk-sorting, using the pricing model in Section 4.2.

Figure 7: **Minimum standard counterfactual: the role of switching costs in determining the effect of the ACA policy on the distribution of risks among contracts**

- *Panel A: Distribution of risks* ***with*** *switching costs; "filling the donut hole"*



- *Panel B: Distribution of risks* ***without*** *switching costs; "filling the donut hole"*



In-sample simulation of the choice model with and without switching costs under the counterfactual policy of raising the minimum standard to completely fill the "donut" hole as envisioned under the Affordable Care Act.

Table 1: **Summary statistics: full sample; baseline sample; panel sub-sample**

	Full sample	Baseline sample	Panel sub-sample
2006			
N	9,999,538	1,221,252	871,818
Age (σ)	72 (12)	76 (8)	75 (7)
Female	0.59	0.65	0.65
White	0.84	0.95	0.96
ESRD	0.01	0.003	0.001
Risk score*	n/a	0.89 (0.34)	0.86 (0.31)
Annual drug spending (σ)	n/a	1,518 (1,899)	1,449 (1,733)
2007			
N	10,176,611	1,307,966	911,403
Age (σ)	72 (12)	76 (8)	75 (7)
Female	0.59	0.63	0.65
White	0.84	0.95	0.96
ESRD	0.01	0.003	0.002
Risk score (σ)	n/a	0.90 (0.35)	0.88 (0.32)
Annual drug spending (σ)	n/a	1,883 (2,407)	1,832 (2,227)
2008			
N	10,369,814	1,356,861	954,494
Age (σ)	72 (12)	76 (8)	76 (7)
Female	0.58	0.63	0.65
White	0.83	0.95	0.96
ESRD	0.01	0.003	0.002
Risk score (σ)	n/a	0.91 (0.36)	0.90 (0.34)
Annual drug spending (σ)	n/a	1,907 (2,648)	1,869 (2,479)
2009			
N	9,781,213	1,365,239	998,014
Age (σ)	71 (12)	76 (8)	76 (8)
Female	0.55	0.63	0.65
White	0.83	0.95	0.96
ESRD	0.01	0.003	0.003
Risk score (σ)	n/a	0.92 (0.36)	0.92 (0.35)
Annual drug spending (σ)	n/a	1,950 (2,973)	1,947 (2,934)

* Risk scores are indices summarizing individual medical history from Medicare Parts A and B. They are used by Medicare for risk-adjustment and are scaled to be 1 for the average Medicare risk. The calculation of risk scores using Medicare A/B diagnostic records and Part D RxHCC software was generously provided by Amy Finkelstein and Ray Kluender.

Table 2: **Evidence of adverse selection: positive correlation tests using realized ex-post drug expenditures and diagnosis-based risk scores**

$$Y_{irt} = \alpha_r + \delta_t + \sum_{k=2}^{k=4} \beta_k \mathbf{1}\{ContractType_{it} = k\} + \epsilon_{irt}$$

	(1) Annual drug spending	(2) Risk score	(3) Risk score projected to USD
Contracts of type 1	reference category		
Contracts of type 2	-2.047	0.00927	24.98
	(71.83)	(0.00730)	(20.89)
Contracts of type 3	1213.4***	0.146***	415.0***
	(105.5)	(0.0107)	(30.46)
Contracts of type 4	3081.3***	0.260***	728.5***
	(70.71)	(0.00521)	(14.57)
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
N	3,892,280	3,892,280	3,892,280
Mean Y	1948.7	0.920	1948.7
St. dev. Y	2712.2	0.357	1018.7

Standard errors in parentheses clustered at the region level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The regressions test for the correlation between the generosity of the chosen insurance contract and the measures of individual risk conditional on geographic regions, which are used for the pricing of contracts. The results are based on the pooled data of the baseline sample that spans years 2007-2009. Details of contract classification typology are discussed in the main text. Outcome in column (1) is measured as annual recorded ex-post drug spending without accounting for cost-sharing. Outcome in column (2) are risk scores based on the diagnostic information from Medicare A/B (1=average risk in Medicare). In column (3), the outcome variable is the predicted value from the regression: $E[AnnualDrugSpending_i|.] = \alpha + \beta RiskScore_i$ run separately for each cross-section. The specification in column (3) thus considers the part of drug spending variation that is attributable to the observed medical diagnosis and thus forecastable risk and doesn't contain moral hazard effects. The results in column (3) can therefore be interpreted as the lower bound of the adverse selection portion of the overall level of asymmetric information.

Table 3: **Evidence of switching costs: choice patterns in 2006-2009 tracked for cohorts entering in different years**

Cohorts of 65 year olds whose incumbent plans were not re-classified into a different type by the insurer										
	65 y.o. in 2006				65 y.o. in 2007			65 y.o. in 2008		65 y.o. in 2009
A. *Enrollment shares*	2006	2007	2008	2009	2007	2008	2009	2008	2009	2009
Contracts of type 1	22 %	22 %	19 %	17 %	17 %	15 %	14 %	10 %	11 %	12 %
Contracts of type 2	73 %	73 %	77 %	79 %	72 %	75 %	77 %	82 %	82 %	82 %
Contracts of type 3	4 %	5 %	5 %	4 %	11 %	11 %	10 %	7 %	7 %	5 %
N	37,500	37,500	37,500	37,500	35,759	35,759	35,759	40,960	40,960	43, 520
B. *Incremental premium*	**in year 2006**				**in year 2007**			**in year 2008**		**in year 2009**
Contracts of type 2	$138				$125			$54		$37
Contracts of type 3	$375				$360			$410		$469

Panel A shows enrollment shares in each year across three types of plans for cohorts of 65 year olds entering the program in different years. The sample includes only individuals, whose incumbent plans were not re-classified in their type by the supply-side throughout the observed enrollment time. The choices of a given cohort are recorded for all subsequent years of available data. The calculation is based on the panel sub-sample data, as described in the data construction appendix. The table shows raw enrollment shares as observed in each year subject to the classification of contracts into the 4-type topology. The choices of cohorts show persistence over time and differ from the choices of newly entering cohorts within the same year. The difference is especially apparent between the first two and the last two years of the data. Panel B adds information about the development of the relative annual premiums over-time. The premiums are reported as increments relative to the *Type 1* contract with SDB deductible and no coverage in the gap. To reflect the market conditions, the premiums are constructed as averages weighted by enrollment of all 65 year olds in the respective years.

Table 4: **Evidence of switching costs: price sensitivity estimates for individuals with and without incumbent plans**

	Age of beneficiaries					
Price coefficient [p-value]	65	66	67	68	69	70
	Baseline	Interaction	Interaction	Interaction	Interaction	Interaction
2006	**-0.003**	0.0001	0.0002	0.0006	-0.0001	0.0006
	[0.000]	[0.809]	[0.683]	[0.386]	[0.876]	[0.321]
2007	**-0.003**	**0.0018**	**0.0012**	**0.0011**	**0.0013**	**0.0010**
	[0.000]	[0.002]	[0.035]	[0.031]	[0.002]	[0.040]
2008	**-0.003**	**0.0022**	**0.0023**	**0.0021**	**0.0019**	**0.0020**
	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.001]
2009	**-0.010**	**0.0072**	**0.0085**	**0.0090**	**0.0085**	**0.0084**
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

The price coefficients are estimated using the following random utility specification:

$$\begin{aligned} u_{ij} &= -\alpha_{65}p_{ij} + \alpha_{66}p_{ij}\mathbf{1}\{Age=66\} + \alpha_{67}p_{ij}\mathbf{1}\{Age=67\} + \\ &+ \alpha_{68}p_{ij}\mathbf{1}\{Age=68\} + \alpha_{69}p_{ij}\mathbf{1}\{Age=69\} + \alpha_{70}p_{ij}\mathbf{1}\{Age=70\} + brand_j + \epsilon_{ij} \end{aligned}$$

$\epsilon_{ij} \sim iid$ Type 1 EV. The specification includes fixed effects for eight largest insurers. The estimates use separate cross-sectional parts of the data sample that is used later to estimate the full choice model. The sample is restricted to only include individuals that are 65-70 years old. The estimates show that in the later years of the program the price sensitivity of new and existing enrollees diverges in the direction that is consistent with the hypothesis of substantial switching costs - enrollees with incumbent plans appear significantly less price sensitive (and similarly so across different 66+ ages) than newly entering beneficiaries. Standard errors (not reported) are clustered at the regional level; p-values in square brackets - differences from the baseline of 65 year olds significant at <5% are marked in bold font. Reported are coefficients on premiums in the utility function and not marginal effects.

Table 5: **Detailed parameter estimates of the preferred choice model specification (#4 in Table 6)**

			Interaction with demographics									
Plan characteristic	Estimate	s.e.	Risk score	s.e.	White	s.e.	Female	s.e	Age	s.e	ESRD	s.e.
in $100 or binary												
Annual premium	**-0.498**	0.011	-	-	-	-	-	-	-	-	-	-
Deductible - μ	**-1.238**	0.127	0.038	0.037	-0.023	0.057	0.013	0.025	**0.008**	0.002	-0.039	0.258
Deductible - σ	**0.470**	0.024	-	-	-	-	-	-	-	-	-	-
Initial Coverage - μ	**-0.168**	0.027	**0.109**	0.007	**0.021**	0.012	0.003	0.005	0.0003	0.0003	**0.099**	0.036
Initial Coverage - σ	**0.087**	0.005	-	-	-	-	-	-	-	-	-	-
Eligible for LIS, 1/0	**-1.093**	0.241	**-0.125**	0.072	-0.089	0.106	0.017	0.047	**0.026**	0.003	-0.683	0.563
Fixed co-pays, 1/0	**-1.538**	0.394	**0.589**	0.119	**-0.343**	0.188	0.003	0.076	**0.019**	0.005	**-1.521**	0.789
Partial gap, 1/0 - μ	**-2.153**	0.334	**1.126**	0.089	0.105	0.150	-0.059	0.065	**0.016**	0.004	0.055	0.540
Partial gap, 1/0 - σ	**1.238**	0.054	-	-	-	-	-	-	-	-	-	-
Default plan, 1/0	**5.091**	0.262	**0.369**	0.066	**-0.815**	0.125	-0.067	0.046	**0.015**	0.003	-0.394	0.381
Insurer fixed effects	Yes (10)											
Log-likelihood	-59,224											
N total observations	2,435,171											
N unique individuals	12,769											
N choice situations	47,770											
Max N alternatives	66											
N of regions	34											

The MSL estimates are based on the control function IV specification of the model as described in the main text. To reduce the computational burden, the model was estimated on a sub-sample of individuals. This allows me to include panel observations for these individuals for all fours years of the data as well as to directly specify each contract in the individuals' choice set in each year without any aggregation. To construct the estimation sample, I restricted the panel sub-sample to include individuals marked with CMS 5% sample flag (which amounts to taking a 25% draw of the panel sub-sample, since the original data represented a 20% population sample of Medicare). I then took a 5% random draw of that data in a way that preserves the original panel structure. The resulting dataset is not different from the original panel sub-sample in a statistically meaningful way. The table reports estimates of utility parameters and not marginal effects. Coefficients significant at 10% level are marked in bold font.

Table 6: **Contract choice model specifications**

	Non-IV (1)	IV (2)	Non-IV (3)	IV (4)
Annual premium, \$100	-0.3911	-0.4464	-0.4148	-0.4984
	(0.0074)	(0.0111)	(0.0078)	(0.0110)
Deductible, \$100	-1.3633	-1.3745	-1.2329	-1.2377
	(0.1307)	(0.1322)	(0.1266)	(0.1266)
sigma	0.5787	0.5904	0.4802	0.4704
	(0.0250)	(0.0253)	(0.0232)	(0.0238)
x Risk	0.0190	0.0139	0.0393	0.0384
	(0.0373)	(0.0377)	(0.0367)	(0.0367)
Initial Coverage Limit, \$100	-0.1584	-0.1589	-0.1650	-0.1679
	(0.0234)	(0.0239)	(0.0262)	(0.0268)
sigma	0.0679	0.0731	0.0815	0.0866
	(0.0045)	(0.0046)	(0.0053)	(0.0053)
x Risk	0.0934	0.0938	0.1052	0.1087
	(0.0059)	(0.0061)	(0.0066)	(0.0068)
Partial coverage gap, 1/0	-1.8659	-1.7814	-2.0850	-2.1528
	(0.2990)	(0.2948)	(0.3358)	(0.3336)
sigma	0.8055	0.6929	1.2640	1.2380
	(0.0611)	(0.0728)	(0.0522)	(0.0542)
x Risk	1.0315	1.0339	1.0954	1.1261
	(0.0814)	(0.0810)	(0.0897)	(0.0895)
Default plan, 1/0	5.4487	5.6093	5.0675	5.0914
	(0.2533)	(0.2596)	(0.2584)	(0.2618)
x Risk	0.2291	0.2227	0.3589	0.3687
	(0.0643)	(0.0658)	(0.0655)	(0.0664)
Observations	2,435,171	2,435,171	2,435,171	2,435,171
Likelihood at convergence	-62,470	-62,379	-59,291	-59,224
Number of insurer fixed effects	3	3	10	10
Switching cost for 75y.o. female, av. risk	\$1,506	\$1,330	\$1,392	\$1,164

The MSL estimates are based on the model specification in the main text. To reduce the computational burden, the model was estimated on a random sub-sample of individuals that preserved the panel structure of the data,. The table reports estimates of utility parameters and not marginal effects. Reported are only the key estimates; the model also includes other contract parameters, demographic interactions and fixed effects as discussed in the main text. The IV specification uses the Control Function approach with lagged plan-level realized claims as the instrument for contract premiums (see Table A.5). Specifications in (3) and (4) add additional insurer fixed effects.

Table 7: **Counterfactual simulations measuring the interaction between adverse selection and switching costs under the observed regulatory regime**

Year 2009 outcomes	Type 1	Type 2	Type 3
Enrollment share			
A. Baseline model prediction, observed prices	18%	72%	9%
B. No switching cost; observed prices	23%	70%	7%
C. No switching cost; endogenous re-pricing	26%	64%	9%
Average risk (risk scores in USD)			
A. Baseline model prediction, observed prices	$1,842	$1,926	$2,368
B. No switching cost; observed prices	$1,892	$1,930	$2,355
C. No switching cost; endogenous re-pricing	$1,907	$1,917	$2,321
Average ex-post drug spending			
A. Baseline model prediction, observed prices	$1,741	$1,881	$2,924
B. No switching cost; observed prices	$1,881	$1,939	$2,325
C. No switching cost; endogenous re-pricing	$1,915	$1,914	$2,330
Enrollment-weighted average premium			
A. Baseline model prediction, observed price	$407	$439	$842
B. No switching cost; observed prices	$350	$412	$767
C. No switching cost; endogenous re-pricing	$313	$456	$742

The table presents the results of a simulation that analyzes the interaction between switching costs and selection patterns conditional on the observed non-price contract features. The reported results include the model's prediction for the baseline with switching costs as well as two counterfactuals without switching cost for one year in the program - 2009. Counterfactual simulation of the baseline marked with A takes premiums and contract defaults in 2009 as given. The idea of this baseline is to replace the observed choices in 2009 with simulated predictions, so as to account for the simulation error in the interpretation of the counterfactuals. Counterfactual simulation in B takes premiums in 2009 as they were observed on the market and shuts down the switching cost channel in the utility function. Counterfactual simulation marked with C allows premiums to adjust to the new sorting of individuals when switching costs are not present. These counterfactual premiums are calculated using the pricing model discussed in the text. Since the model for premiums assumes that insurers adjust prices in accordance with lagged expenditures in plans, switching costs only change pricing-relevant sorting in 2008. Price simulation thus takes the premiums in 2006 and 2007 as given and re-calculates premiums for 2008 and 2009. The choices without switching costs in 2008 are then simulated using the new prices, which in turn affect the simulation of prices and subsequent choices in 2009.

Table 8: **Simulation of the Affordable Care Act - "filling the donut hole" policy - with costly and costless switching**

Year 2009 outcomes	Type 1	Type 2	Type 3
Enrollment shares by contract type			
A. With switching cost, observed prices			
Baseline predicted 2009 enrollment	18%	72%	9%
Minimum standard without donut hole in 2009	19%	72%	9%
B. No switching cost, observed prices			
Baseline predicted 2009 enrollment	23%	70%	7%
Minimum standard without donut hole in 2009	24%	70%	5%
C. No switching cost, endogenous re-pricing			
Baseline predicted 2009 enrollment	26%	64%	9%
Minimum standard without donut hole in 2009	40%	47%	13%
Average risk (risk scores in USD)			
A. With switching cost, observed prices			
Baseline predicted 2009 sorting	\$1,842	\$1,926	\$2,368
Minimum standard without donut hole in 2009	\$1,870	\$1,939	\$2,246
B. No switching cost, observed prices			
Baseline predicted 2009 sorting	\$1,892	\$1,930	\$2,355
Minimum standard without donut hole in 2009	\$1,940	\$1,962	\$1,884
C. No switching cost, endogenous re-pricing			
Baseline predicted 2009 sorting	\$1,907	\$1,917	\$2,321
Minimum standard without donut hole in 2009	\$1,965	\$1,947	\$1,935

Enrollment and risk-sorting under the Affordable Care Act policy of filling the donut hole. To simulate the policy, I impose an increase in the Initial Coverage Limit to \$6,154 for all plans - this amount corresponds to the Catastrophic Coverage Limit in 2009. This structure emulates the contract structure of a plan with no coverage in the gap as envisioned under the ACA as well as the contract structure that was used in plans with full coverage in the gap in 2006-2007. Expected risk is the risk score in USD. Endogenous re-pricing model as described in Section 4.2. In the counterfactual simulation, *Type 2* and *Type 3* contracts are effectively identical, as all contracts now have the same initial coverage limit (i.e. all have full coverage in the gap). Thus, the distinction among the contract types is based on the baseline simulation.

FOR ONLINE PUBLICATION APPENDIX

7 Appendix

7.1 Conceptual framework: interaction between adverse selection and switching cost in the presence of regulatory intervention

A stylized model of insurance contract choice below highlights the key economic channels that are analyzed empirically in the paper. Consider a mass of beneficiaries, each described by a pair of characteristics - the individual's risk type r, as well as risk preferences and other demographic or idiosyncratic factors that may affect the individual's preference for insurance together denoted with ϕ. For simplicity, assume that the individual faces a choice between two insurance contracts that differ only in their deductible. The more generous contract H has a zero deductible and a premium p_H, while the less generous contract L has a deductible $d > 0$ and a premium $p_L < p_H$.

Assuming the separability of prices in the indirect utility function and letting $v(d, \phi, r)$ denote the valuation of a contract with deductible d by individual (ϕ, r), we arrive at a standard choice problem in a differentiated goods environment. Individual (ϕ, r) chooses contract L if:

$$v(0, \phi, r) - v(d, \phi, r) < p_H - p_L$$

$$\triangle v(d, \phi, r) < p$$

where p denotes the relative price. Suppose that for any given level of the deductible, the valuation of an insurance contract is increasing in risk r, i.e. $\frac{\partial v(d,\phi,r)}{\partial r} > 0$ and preferences such as risk aversion, i.e. $\frac{\partial v(d,\phi,r)}{\partial \phi} > 0$, while the valuation is decreasing in the deductible for a given (ϕ, r), i.e. $\frac{\partial v(d,\phi,r)}{\partial d} < 0$. Suppose further that the valuation and prices are such that the "market is covered" in the sense that all individuals find it optimal to buy one of the insurance contracts rather than to remain uninsured.[32] Then, there exists an individual of type $(\hat{\phi},\hat{r})$ who is indifferent between the two contracts, i.e. $\triangle v(d, \hat{\phi},\hat{r}) = p$. The average risk that contract L expects to get after individuals choose between the two contracts is $E[r|\triangle v(d, \phi, r) < \triangle v(d, \hat{\phi},\hat{r})]$.

Now suppose we introduce an exogenous shock to the model that changes the features of the contract space. Consider, for instance, a one-dimensional minimum standard policy that only sets the maximum allowed deductible $\bar{d}$. Assume further that the less generous contract sets its deductible d to always equal the maximum deductible set by the government: $d = \bar{d}$. The more generous contract, at the same time, always keeps zero deductible. This simplification implies that I am not modeling how insurers originally decide whether to offer the minimum standard or zero deductible, taking these decisions as given and stable from the policy perspective.

Now suppose the government changes its policy and increases the maximum allowed deductible from d to $d' > d > 0$ and nothing else changes. In particular, suppose for a moment that relative prices remain the same p. Individuals that were choosing contract L before, will

[32]While this assumption is certainly restrictive and eliminates an important extensive margin on which the minimum standard may affect the market (Finkelstein, 2004), the empirical model in this paper focuses on the effects of the minimum standard on the intensive margin, across different levels of contract generosity, and thus I focus on this aspect of the question in this stylized model as well.

switch to contract H under the new policy if now:

$$\triangle v(d', \phi, r) > p$$

The risk pool of switchers from the less to the more generous contract under the new policy but without price adjustment is: $E[r|\triangle v(d, \phi, r) < p \text{ and } \triangle v(d', \phi, r) > p]$. Whether this re-sorting results in higher or lower risk in contract L depends on whether the effect of risk on valuation grows faster at a higher deductible than the effect of non-risk preferences on valuation under a higher deductible. In other words, it depends on the relationship between $\frac{\partial^2 v(.)}{\partial r \partial d}$ and $\frac{\partial^2 v(.)}{\partial \phi \partial d}$.

Now suppose that individuals face a switching cost γ. This cost may be heterogeneous across individuals and correlate both with individual preferences ϕ and risk type r. Let γ be a function of individual characteristics $\gamma(\phi, r)$. With the switching friction individuals that were choosing contract L before the policy change, will switch to contract H under the new policy if:

$$\triangle v(d', \phi, r) > p + \gamma(\phi, r)$$

The switching cost has the effect of diminishing and tilting the set of beneficiaries that are indifferent between switching to H and staying in L. The first order effect is that the presence of the switching friction slows down the re-sorting process, as now fewer consumers react to the change in the contract space. The second-order tilting effect is that whether relatively higher or lower risks tend to stay in contract L rather than change to H in the presence of switching cost will depend on the partial and cross-partial derivatives of the switching cost with respect to risk r and preferences ϕ.

Allowing insurers to adjust prices to the new regulation and sorting patterns that are distorted by the switching costs produces theoretically ambiguous results that depend on the relationship between contract valuation and risk. For example, with a higher regulated deductible, the relative price will increase because a higher deductible mechanically reduces the liability of contract L. This, in turn tightens the switching constraint $\triangle v(d', \phi, r) > p' + \gamma(\phi, r) > p + \gamma(\phi, r)$, which can further decrease or increase the risk depending on the individual value function. Overall, the direction of change in sorting patterns induced by the change in the contract space are ambiguous if we allow for switching costs and allow insurers to adjust prices in response to changes in selection patterns. The effect that the regulation has on the allocation of risks across contracts will depend on the partial and cross-partial derivatives of the valuation and switching costs with respect to risk and preferences. The choice model in Section 4 estimates these inter-dependencies in Medicare Part D empirically and uses the estimates to simulate the role of switching costs in shaping the risk-sorting across contracts in response to market-driven and regulatory changes in contracts.

7.2 Construction of the empirical sample from Medicare administrative data

I restrict the sample to individuals of age 65 and older residing within 34 Medicare Part D regions or 50 states (Medicare combines some states into the same PDP market), who did not die in the reference year and were originally entitled to Medicare because of old age rather than disability. In other words, I do not include individuals, who may become eligible for Medicare before they turn 65 as part of their SSDI benefit. I further drop observations on individuals that were dual eligible for Medicare and Medicaid in the reference year, since these individuals are assigned to plans by CMS rather than choosing plans on their own. This brings the sample down to 25.6 million beneficiary-year observations. I then eliminate individuals that did not enroll in Part D or were enrolled in Medicare Advantage (or another managed care) option that combines prescription drug coverage with healthcare insurance.

Most differences between the panel sub-sample and the baseline comes from the way CMS draws its 20% random sample of the Medicare population. These samples are only partially based on panel draws and thus not all individuals are observed in every year. For details on the CMS sampling procedures see the Chronic Condition Data Warehouse User Manual v.1.7. Some individuals in the panel sub-sample will be lost if they change from a PDP to a Medicare Advantage prescription drug plan simultaneously with switching from the "traditional" Medicare to the HMO system. Moreover, it is possible that some individuals leave the Part D program altogether; this option is likely to be very rare, however, since these beneficiaries would then face premium penalties if they decide to re-enter the program at a later date. Lastly, some observations will be lost in the panel sub-sample due to individuals dying in years 2007-2009.

7.3 Stylized examples of switching costs altering risk sorting within a simplified contract menu

I provide two stylized examples using a significantly simplified version of the Part D contract space to illustrate the opposite effects that switching frictions may have on selection. Each example considers a menu of three contracts of different types and simulates the evolution of choices and risk selection among these three contracts, assuming that they were the only choices available to all individuals on the market for four years. The two different menus I consider were actually observed on the market. For each, I do the simulation exercise twice - assuming either costly or costless switching. For the simulations, I use the estimated parameters of the model, including the switching cost parameter, the observed demographics, and the risk scores for all individuals in the estimation sample.

Consider the contract menu in the first example. In this menu, the insurer offers *Type 1* and *Type 2* plans in the first year of the market, adding a *Type 3* plan in the second year. Panel A of Figure A.6 plots the development of the relative premium between the *Type 3* and *Type 1* plan for the three years when both are available. The relative premium is quite low in the first year *Type 3* contract is introduced; nevertheless, in the scenario with costly switching very few individuals take up this more comprehensive plan. This is in stark contrast to the scenario with costless switching, in which the enrollment in the *Type 3* plan jumps to 36%. Importantly, we see that with the jump in the enrollment share, costless switching also results in adversely selected enrollment in the *Type 3* plan relative to *Type 1*. With costless switching, selection gets more acute in 2009, as the relative premium of the contract increases and enrollment drops significantly. In the scenario with switching costs, neither the enrollment, nor the relative risk change much over time. The enrollment in *Type 3* contract stays at below 5% level and the relative risk stays substantially below the relative risk in the scenario with costless switching. This setting thus illustrates the natural intuition that switching costs should be muting adverse selection.

Now consider a different contract menu example. In this example, the insurer offers a generous *Type 4* plan priced at a substantial premium over the *Type 1* plan in the first year, which has a very low premium. *Type 4* plan collects the highest risks on the market in year one. In year two, the insurer demotes *Type 4* plan to be *Type 3* and raises its premium, while also raising the premium of *Type 1* plan. Figure A.6 illustrates that the relative premium between the most and the least generous plans of this insurer first rises and then falls over time. The same figure illustrates the simulated enrollment and risk-sorting patterns for this contract menu. We see that in this setting, where a very generous plan collected the highest risk in the first year, switching costs support the large differential in risk between the most and least comprehensive plan over time. The simulation of choices suggests that the difference in risks would have been lower in a counterfactual with costless switching. In other words, in this scenario of contract menu evolution, switching costs exacerbate, rather than mute, selection.

7.4 Additional minimum standard counterfactuals

Effect of minimum standard regulation on the allocation of risks with costly and costless switching

Local deviation in the minimum standard level. In this counterfactual (results reported in Table A.6) , I change the 2009 minimum standard to the 2006 level. This implies a reduction in the deductible offered by all *Type 1* plans in 2009 from the observed \$295 to \$250, and a reduction in the ICL from the observed $\$2,700$ to $\$2,250$ for all three types of plans. This exercise simulates the short-run effects of a counterfactual policy shock in 2009. Ex ante, we would expect that absent choice frictions and price adjustment, *Type 1* enrollment share should increases and its risk pool worsen, since *Type 1* contracts with a lower deductible at the same prices become more attractive relative to *Type 2* contracts. With endogenous re-pricing of contracts, the counterfactual policy has a theoretically ambiguous effect. To test which direction the empirical simulation takes, I consider three scenarios of the choice environment in this policy experiment. One with switching costs and observed prices, another without switching costs, but still observed prices kept fixed, and lastly the no switching cost scenario where I allow insurers to adjust prices to the new sorting and regulatory conditions. Note that I do not allow for any other changes in the contract space, and take the entry and exit of plans, as well as the changes of other characteristics between 2008 and 2009 as given. The effect of this policy deviation is small, but follows the predicted direction. In the baseline scenario that keeps switching costs and observed prices, *Type 1* enrollment share moves from 18% to 19%. The effect on enrollment is only somewhat starker in the other two scenarios where I shut down the inertia channel. This local deviation in the minimum standard policy has a very small, but expected direction for risk-distribution.

Loosening the minimum standard. Here I analyze the case where the standard defined benefit would have been a high-deductible plan, which is a policy that had been considered by the government as a possible option during the regulatory design of Medicare Part D. I simulate the market outcomes for a policy shock that forces all *Type 1* plans to have $\$1,000$ deductible in 2009, but keeps other features of *Type 1* and other contract types the same. The results are reported in Table A.6. I again consider three scenarios: with switching costs at observed prices and without switching costs with and without price adjustment. Not surprisingly, the response to this policy is much starker than to the previous local deviation in the minimum standard. In the high deductible scenario, *Type 1* plans loose a substantial share of their enrollment. It drops from 19% to 10% in the scenario with switching costs and from 23% and 26% to 12% and 20% in the scenarios with no switching costs with and without price adjustment respectively. The effect on the allocation of risk is small, even in scenarios that completely shut down the inertia channel and re-price the contract. This is not too surprising, given that in the reduced-form evidence of adverse selection we found relatively little selection on the deductible margin. Although the observation that *Type 1* plans substantially loose their enrollment share under observed prices is natural, since increasing the deductible makes the plans much less generous at the same price, the full equilibrium result is less intuitive. We could have expected that with a price adjustment that makes the high deductible plan significantly cheaper than the next available alternative (the simulated annual premium drops to \$45), there would exist substantial demand for such a "catastrophic coverage" plan. Evidently, for the beneficiaries whose observed choices were used to estimate the model and who selected into Part D insurance to begin with, the valuation of the first-dollar coverage is very high. In this pool of beneficiaries, introducing a high deductible plan amplifies the tendency to a pooling equilibrium on the contracts with low deductible - a tendency that has been observed in this program over time as the regulator has been increasing the SDB deductible every year. Although not modeled in the current setting, it is possible that the existence of a relatively cheap high deductible plan would attract more individuals on the extensive margin of the program, potentially improving the overall welfare in Part D.

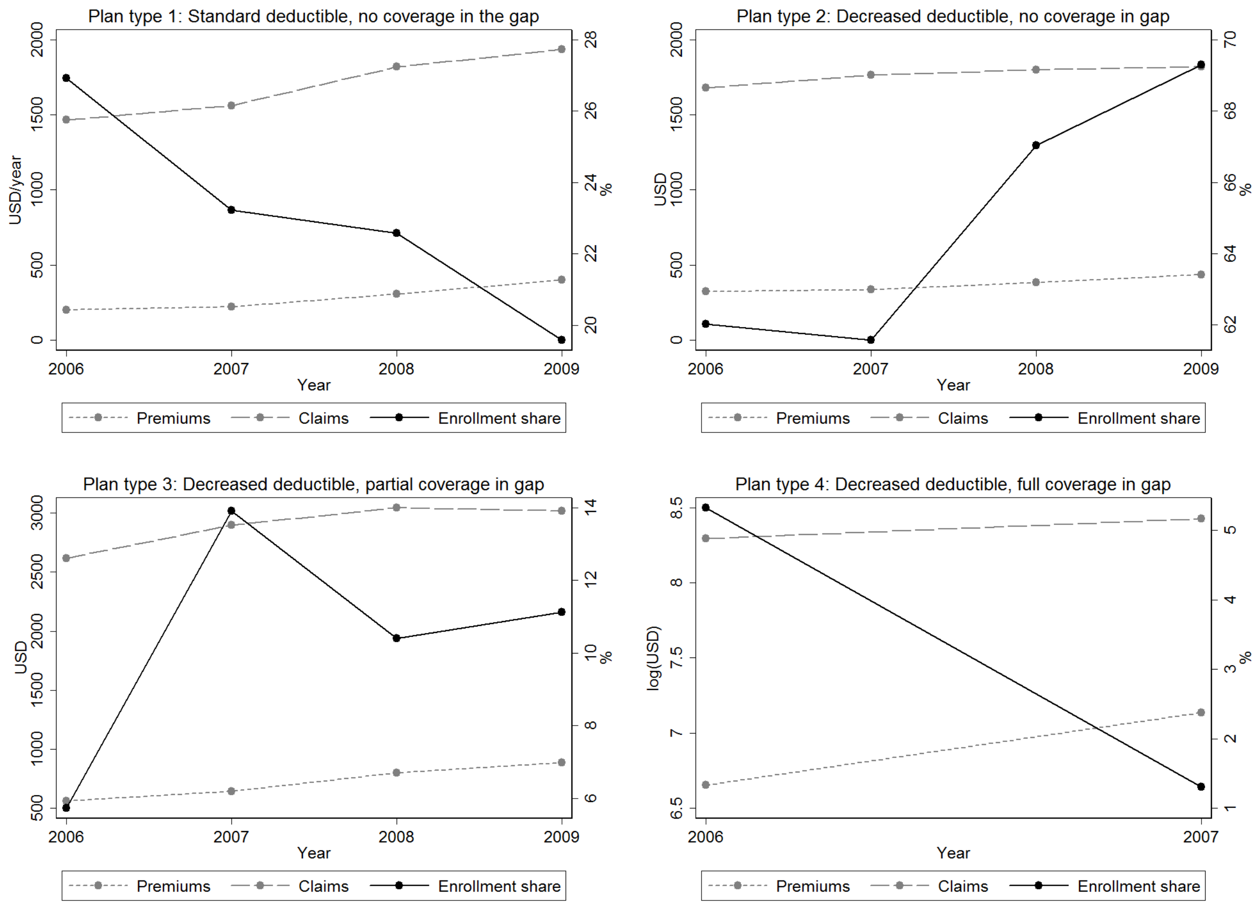
Table A.1: **Construction of the baseline sample**

	2006		2007		2008		2009	
Full sample, N	9,999,538	*100%*	10,176,611	*100%*	10,369,814	*100%*	9,781,213	*100%*
Keep age 65+ within 50 states	8,385,276	*84%*	8,511,573	*84%*	8,658,693	*83%*	8,066,696	*82%*
Drop if died in the reference year	7,982,664	*80%*	8,111,023	*80%*	8,249,112	*80%*	7,714,002	*79%*
Drop if dual eligible any month of year	6,839,959	*68%*	6,952,339	*68%*	7,087,638	*68%*	6,637,418	*68%*
Keep if Medicare b/c of old age	6,412,259	*64%*	6,505,996	*64%*	6,619,029	*64%*	6,178,410	*63%*
Keep PDP enrollees [a]	1,797,409	*18%*	1,739,617	*17%*	1,800,364	*17%*	1,611,820	*16%*
Drop recipients of premium subsidies	1,551,253	*16%*	1,597,567	*16%*	1,668,923	*16%*	1,505,854	*15%*
Drop RDS and missing risk scores	1,221,252	*12%*	1,307,966	*13%*	1,356,861	*13%*	1,365,239	*14%*
Baseline sample	**1,221,252**	*12%*	**1,307,966**	*13%*	**1,356,861**	*13%*	**1,365,239**	*14%*
Panel sub-sample	871,818	*9%*	911,403	*9%*	954,494	*9%*	998,014	*10%*

The table shows the restrictions to the original sample of 20 % Medicare beneficiaries that were imposed to get to the baseline sample. The key restriction was to drop observations on individuals who didn't enroll in any Part D plan or enrolled in Part D through their managed care plan rather than through a stand-alone prescription drug plan (PDP). For years 2007-2009, I kept only individuals who were enrolled in a PDP for the whole year with the exception of the 65 year olds - this excludes those individuals who were allowed to join the plan outside of the open enrollment period because they e.g. changed their state of residence. In 2006, given the different special open enrollment period, many individuals were not enrolled for all 12 months and so I keep all individuals who initiated enrollment at some point during 2006 and didn't leave in subsequent months of 2006.

[a] Mainly drops those who did not enroll in Part D at all and those who enrolled in Medicare Advantage or other Part D coverage options.

Figure A.1: **Evolution of premiums, claims and enrollment shares by type of plan**



The data in the panels is based on the baseline sample for years 2006-2009. The calculation of average premiums and claims is enrollment-weighted. Claims include the total annual drug spending. To ensure comparability of the observations, the spending of individuals with enrollment shorter than 12 months (primarily 65 year olds) was extrapolated to the full year. The panel for *Type 4* plan illustrates the significant drop in enrollment and contemporaneous increase in the premiums and claims in the plans with full coverage in the gap. The claims and premiums for this plan type are recorded on log-scale for visual clarity. Contract space over time is split using the 4-type classification of plans. Contracts of *Type 2* with a reduced deductible and no gap coverage had the highest enrollment, greater than 60%, in all years. The enrollment share in these contracts increased over time by almost 8 percentage points, mirroring the decrease of 8 percentage points in the enrollment share of the *Type 1* contracts with SDB deductible. The enrollment share in *Type 3* contracts is more volatile over time. Average premiums and average spending have remained relatively stable in *Type 2* contracts, while both grew in *Type 1*, *Type 3*, and especially *Type 4* contracts.

Figure A.2: **Asymmetric information in Medicare Part D: years 2006-2009**

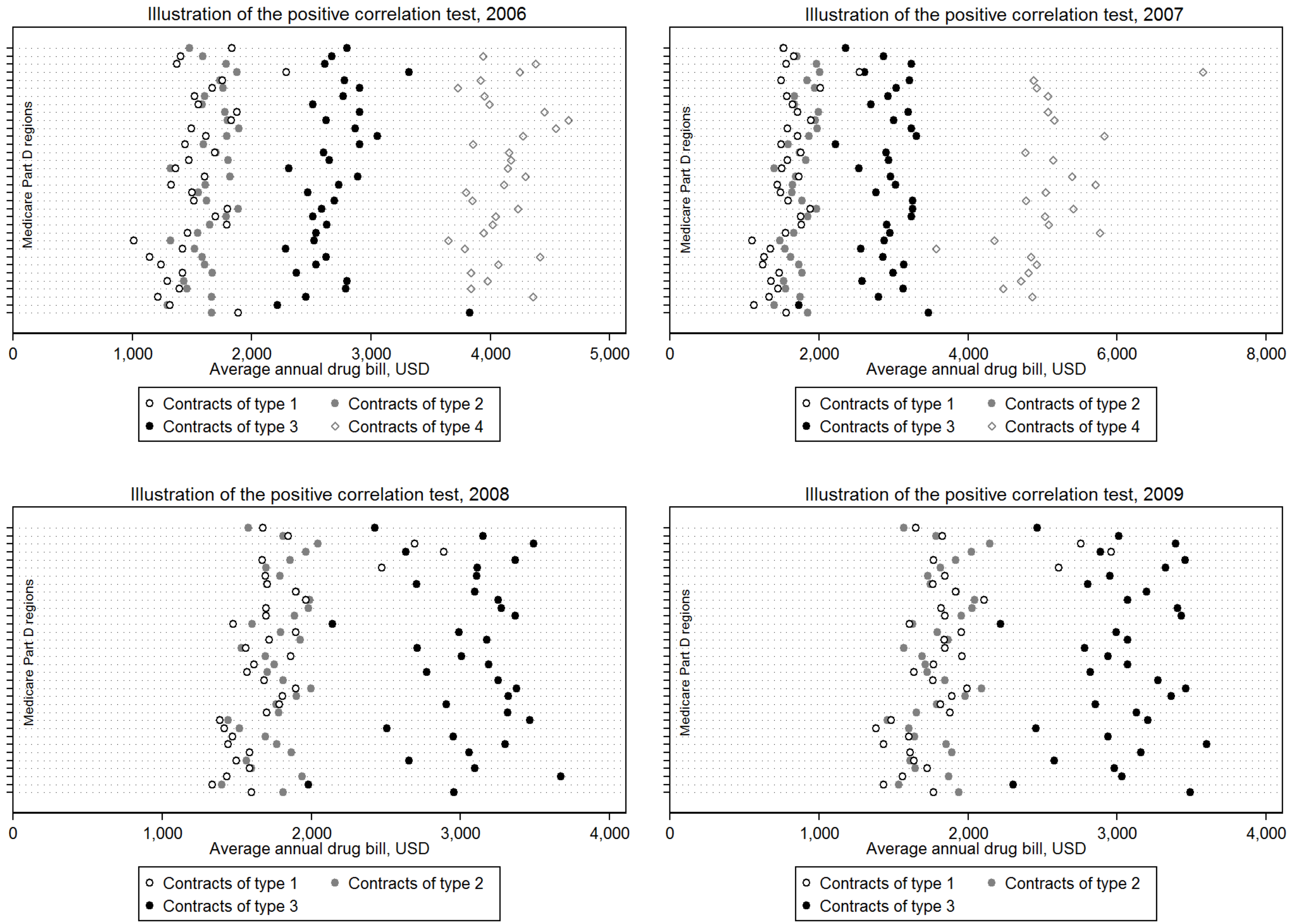


Table A.2: **Evidence of switching costs: share of enrollees that choose their "default" plan in a given year**

	2007	2008	2009
A. All plans			
Probability of choosing default plan for 66+ y.o. enrollees	89.9 %	88.7 %	89.1 %
N	1,089,978	1,162,545	1,194,036
B. Default plan in t was re-classified in its type vs $t-1$			
Probability of choosing default plan for 66+ y.o. enrollees	73.4 %	78.9 %	87.9 %
% of sample	14 %	7 %	4 %
C. Default plan in t had the same type as in $t-1$			
Probability of choosing default plan for 66+ y.o. enrollees	92.6 %	89.5 %	89.3 %
% of sample	86 %	93 %	96 %

The table shows the share of individuals in each year that choose to remain in their "default" plan among those for whom default plans could be defined. The data is a sub-sample with two-year panel observations from the baseline sample of enrollees. The observations do not include 65 year olds, as these by definition do not have default plans. A default plan flag is constructed by using plan cross-walks provided by CMS and requires observing the policy in which the individual was enrolled in two consecutive years. The default plan enrollment flag is set to equal one in two cases. First, and most common, if the individual enrolled in the plan with exactly the same CMS id in year $t-1$ and t. Second, if in year t individual enrolled in a plan that is not identical to the plan id in $t-1$, but was recorded as a plan consolidating the original $t-1$ plan, in which case the CMS policy is to default the individual into the consolidating plan if the individual takes no action of choosing a different option. Panels B and C define plan "type" according to the 4-type topology used in the main text. These panels divide the sample of beneficiaries into those who did and did not experience a significant supply-induced change in their incumbent plan. Panel B excludes observations for individuals whose plans were terminated and no default option was available. These individuals comprise less than 1% of the sample.

Figure A.3: **Adverse selection in Medicare Part D: years 2006-2009**

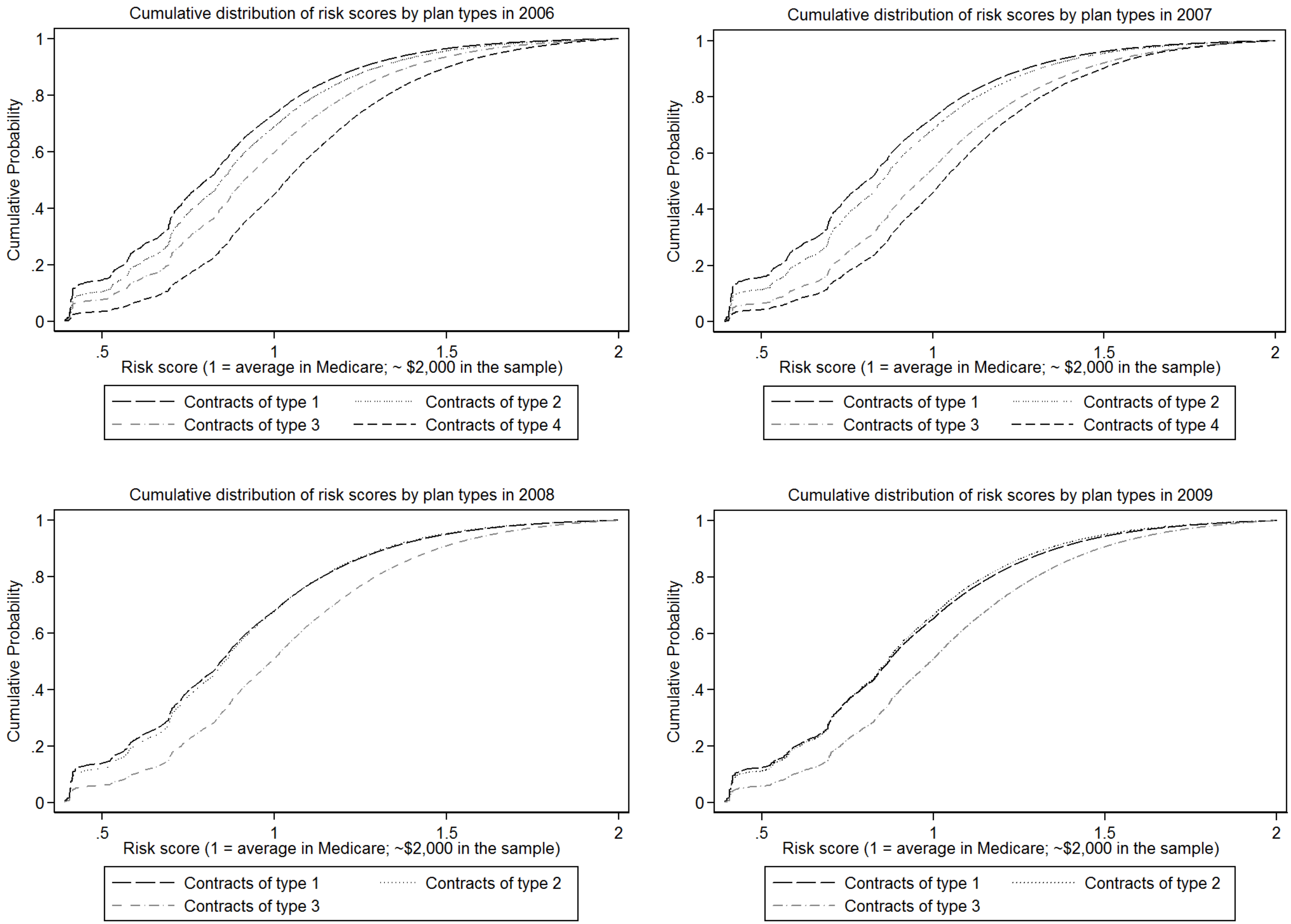
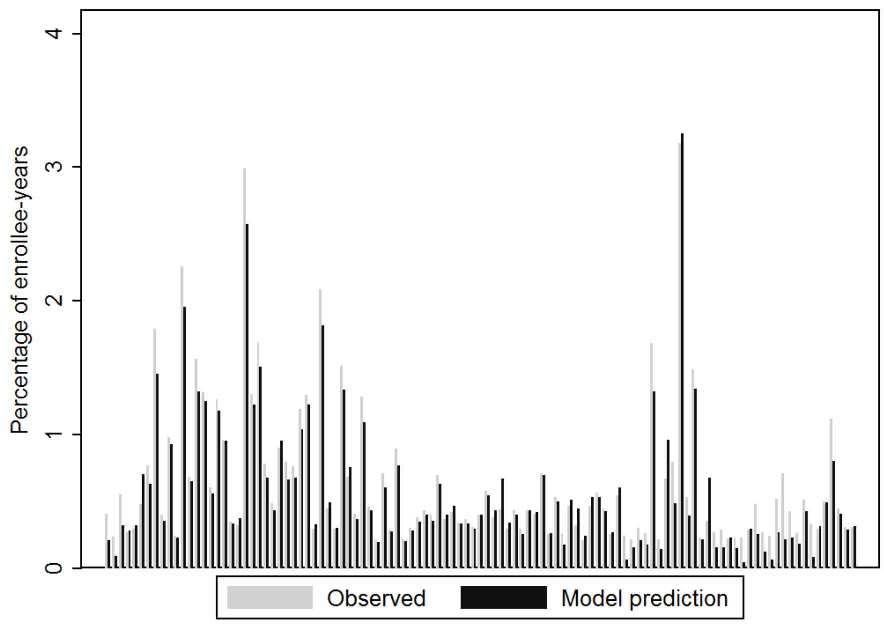


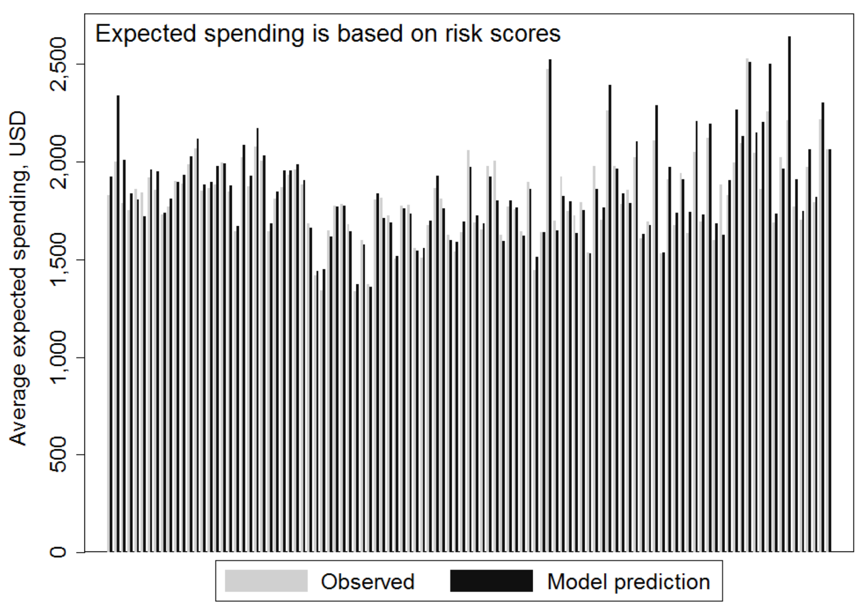
Table A.3: **Choice model fit: summary statistics by contract type and insurer for enrollment and risk distribution moments**

	Enrollment		Average drug spending		Average risk score	
	Observed	Model	Observed	Model	Observed	Model
Contracts of type 1	21.71%	20.71%	$1,580	$1,614	0.85	0.84
Contracts of type 2	65.93%	69.78%	$1,743	$1,819	0.88	0.89
Contracts of type 3	10.78%	8.88%	$2,846	$2,716	1.01	1.03
Contracts of type 4	1.58%	0.63%	$3,971	$2,658	1.04	1.08
Insurer A	29.65%	30.77%	$1,915	$1,939	0.92	0.92
Insurer B	27.10%	25.91%	$1,621	$1,614	0.86	0.85

The table compares three within-sample predicted and observed moments in the data: 1) Enrollment shares in different types of plans and in different insurer brands; 2) Average drug spending in different types of plans and in different insurer brands and 3) Average risk scores in different types of plans and in different insurer brands. The data is pooled over time and regions. To simplify the contract space, the comparison is made at the 4-type plan aggregation and at brand-level aggregation for the top 2 insurers. A more disaggregated fit of the model is illustrated in Figure A.4. For the risk scores and drug bills, "predicted" measures refer to the sorting of the observed risks and expenditures as suggested by the simulation of the choice model.

Figure A.4: **In-sample fit of the choice model**





To construct simulated enrollment, the estimated coefficients of the choice model together with simulated random components were used to find the contract with the highest utility in each individual's choice set. The observed risk scores of the individuals predicted to enroll in different plans were used to compute the average predicted risk. Each pair of bars in the graph represents a different Medicare Part D plan ("plan" is region-specific). The graphs display only top 90 out of 2,357 contracts.

Table A.4: **Basic descriptive evidence generated in the model: share of enrollees choosing the "default" option**

Includes all plans	2007	2008	2009
1. Share observed in the baseline sample			
Probability of choosing default plan for 66+ y.o. enrollees	89.9 %	88.7 %	89.1 %
N	1,089,978	1,162,545	1,194,036
2. Share observed in the estimation sample			
Probability of choosing default plan for 66+ y.o. enrollees	89.9%	89.5%	89.6%
N	11,170	11,640	12,197
3. Share predicted in the estimation sample			
Probability of choosing default plan for 66+ y.o. enrollees	86.3%	84.8%	86.3%
N	11,170	11,640	12,197

This tables reports the simulation of the baseline descriptive evidence on the switching rates as documented in Table A.2 using the estimated contract choice model.

Figure A.5: **Key descriptive evidence generated in the model: empirical distribution of expected spending by type of plan over time; in-sample data**

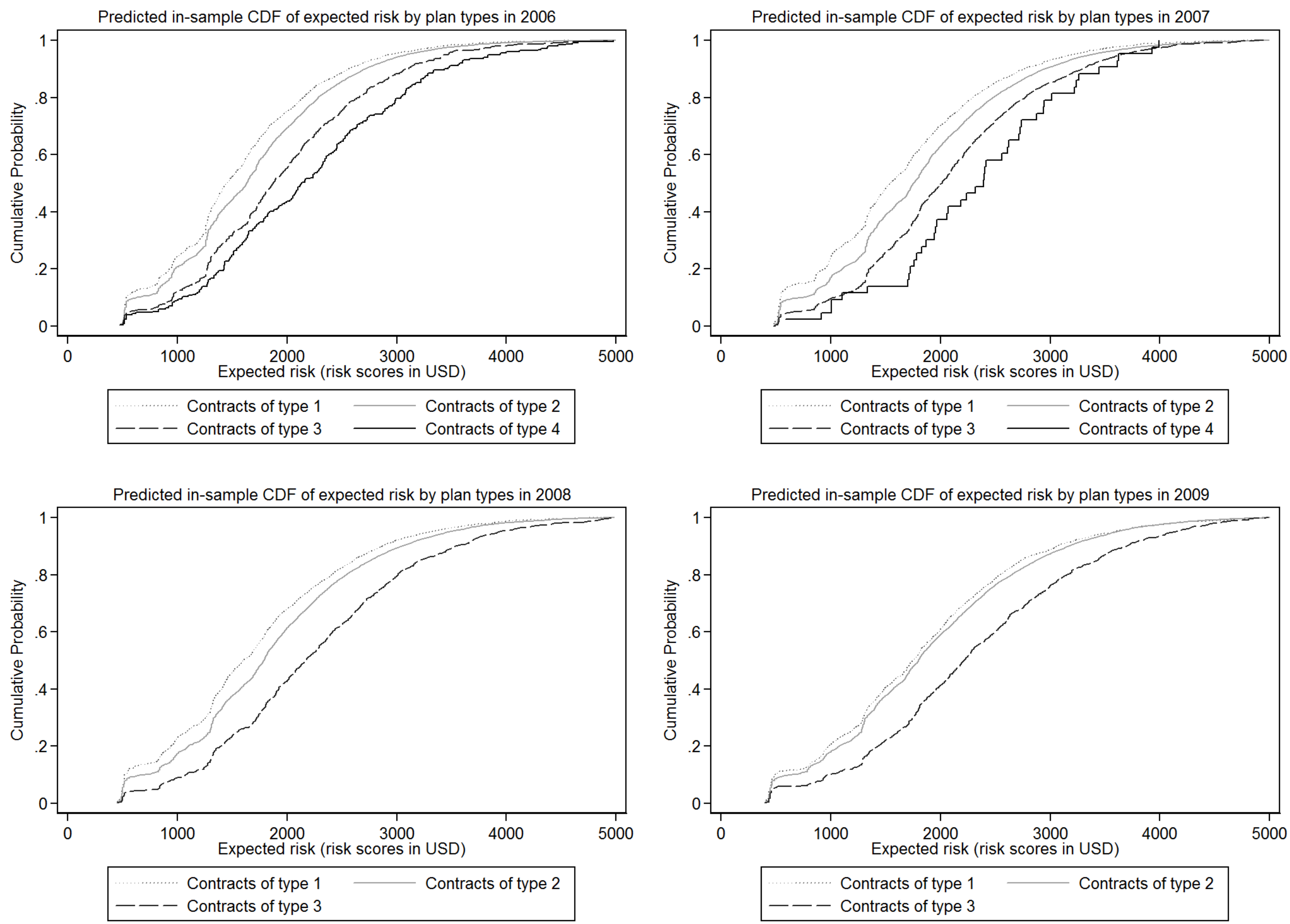


Table A.5: **Pricing model used for the simulation of premiums in the counterfactual scenarios**

$$E[Y_{jbt}|.] = \alpha_b + \delta_r + M'_{jbt-1}\beta + \gamma_1 Ded_{jbt} + \gamma_2 ICL_{jbt} + \gamma_3 1\{PartialGap\}_{jbt}$$

where j indexes region-specific plans, b indexes insurers (brands), r indexes 34 Part D regions, t indexes years

	(1) Annual premium, USD
Lagged mean spending	0.132***
	(0.00992)
Deductible amount, USD	-0.489***
	(0.0262)
ICL amount, USD	0.312***
	(0.0198)
Partial coverage in the gap, 1/0	293.9***
	(11.89)
Insurer FE	Yes
Region FE	Yes
N	2566
Mean Y	540.2
St. dev. Y	253.3
R-squared	0.802

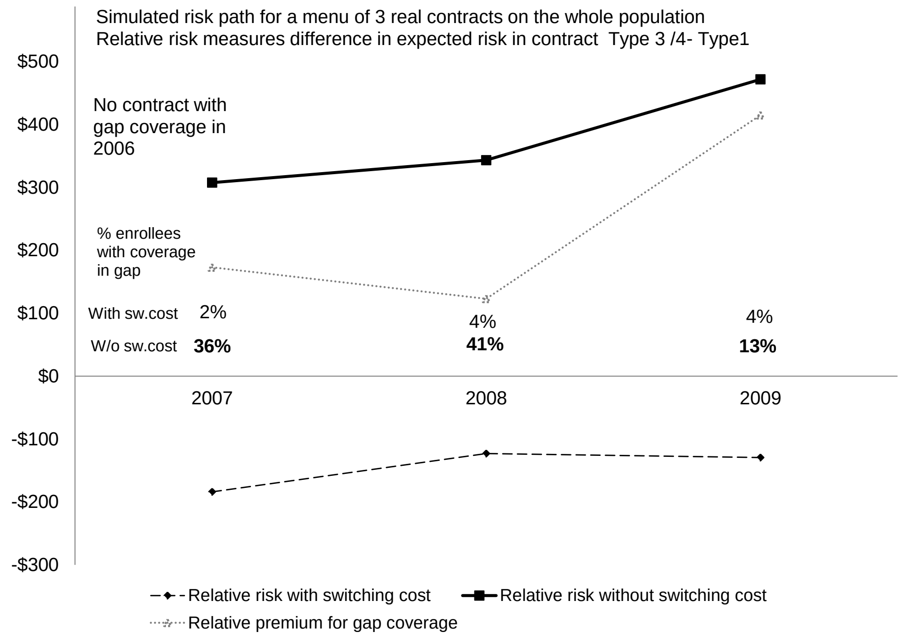
Clustered standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The pricing regression is estimated on a dataset that records, for all prescription drug plans, the annual premium, the mean, the standard deviation and other moments of the lagged drug spending distribution in the plan (by plan enrollees in the baseline sample). The data also records the key financial characteristics of the plans - the deductible, the ICL and the gap coverage indicator of each plan in the program for years 2007-2009. For the cases where plans changed their ID over time due to mergers, I use Medicare plan cross-walk to match plans. The regression output doesn't report the coefficients on the set of fixed effects, as well as on the standard deviation, the kurtosis, the inter-quartile range, the 95th and 5th percentiles of the lagged distribution of realized expenditures, but these variables are included in the regression.

Figure A.6: **Stylized interaction between switching costs and selection**

- *Panel A: Example of switching costs muting adverse selection*



- *Panel B: Example of switching costs supporting adverse selection*

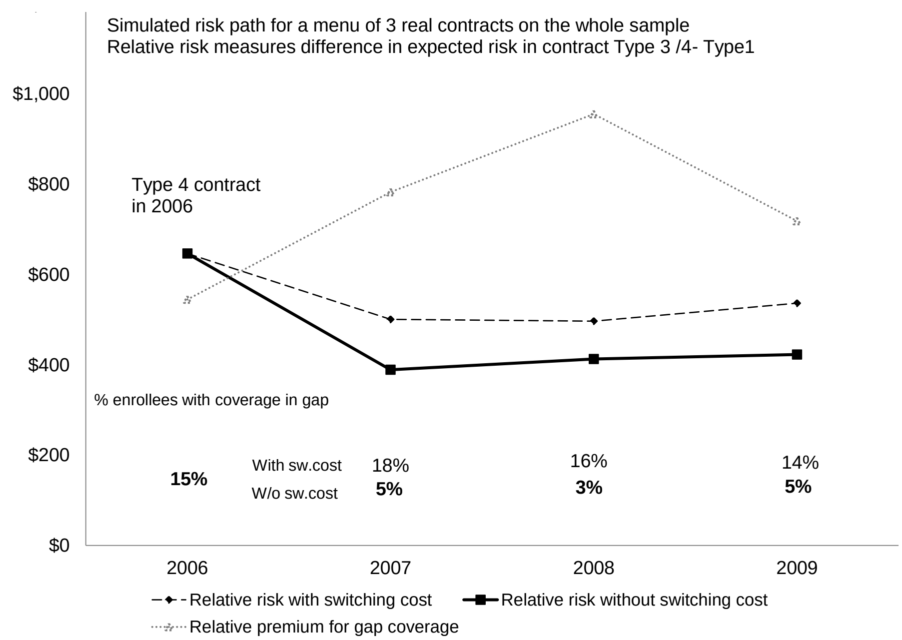


Table A.6: **Counterfactual minimum standard policies with costly and costless switching**

Year 2009 outcomes	Type 1	Type 2	Type 3
Enrollment shares			
A. With switching cost, observed prices			
Baseline predicted 2009 enrollment	18%	72%	9%
Local deviation	19%	71%	9%
Loosening	10%	80%	10%
B. No switching cost, observed prices			
Baseline predicted 2009 enrollment	23%	70%	7%
Local deviation	25%	67%	7%
Loosening	12%	80%	8%
C. No switching cost, endogenous re-pricing			
Baseline predicted 2009 enrollment	26%	64%	9%
Local deviation	27%	64%	9%
Loosening	20%	72%	9%
Average risk (risk scores in USD)			
A. With switching cost, observed prices			
Baseline predicted 2009 sorting	$1,842	$1,926	$2,368
Local deviation	$1,838	$1,929	$2,366
Loosening	$1,854	$1,914	$2,369
B. No switching cost, observed prices			
Baseline predicted 2009 sorting	$1,892	$1,930	$2,355
Local deviation	$1,896	$1,929	$2,366
Loosening	$1,901	$1,925	$2,326
C. No switching cost, endogenous re-pricing			
Baseline predicted 2009 sorting	$1,907	$1,917	$2,321
Local deviation	$1,901	$1,923	$2,311
Loosening	$1,948	$1,910	$2,320

Enrollment and risk-sorting under two counterfactual minimum standard policies. Counterfactual 1 changes *Type 1* deductible from $295 to $250 and ICL in all plans to $2,250 from $2,700. Counterfactual 2 increases the deductible in *Type 1* plans to $1,000. Expected risk is the risk score in USD. Endogenous re-pricing model as described in Section 4.2.