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

Maria Polyakova*

Conceptual framework: interaction between adverse selection and switching cost in the presence of minimum standard regulation

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

*Department of Health Research and Policy, Stanford University, 150 Governor's Lane, Stanford, CA 94305, and National Bureau of Economic Research (e-mail: maria.polyakova@stanford.edu).

insurance contracts rather than to remain uninsured.[1] 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 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

[1]While this assumption is certainly restrictive and eliminates an important extensive margin on which the minimum standard may affect the market, 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.

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 III 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.

Construction of the empirical sample from Medicare's 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.

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.

Empirical model: specification checks and fit of the choice model; point estimates of the pricing regression

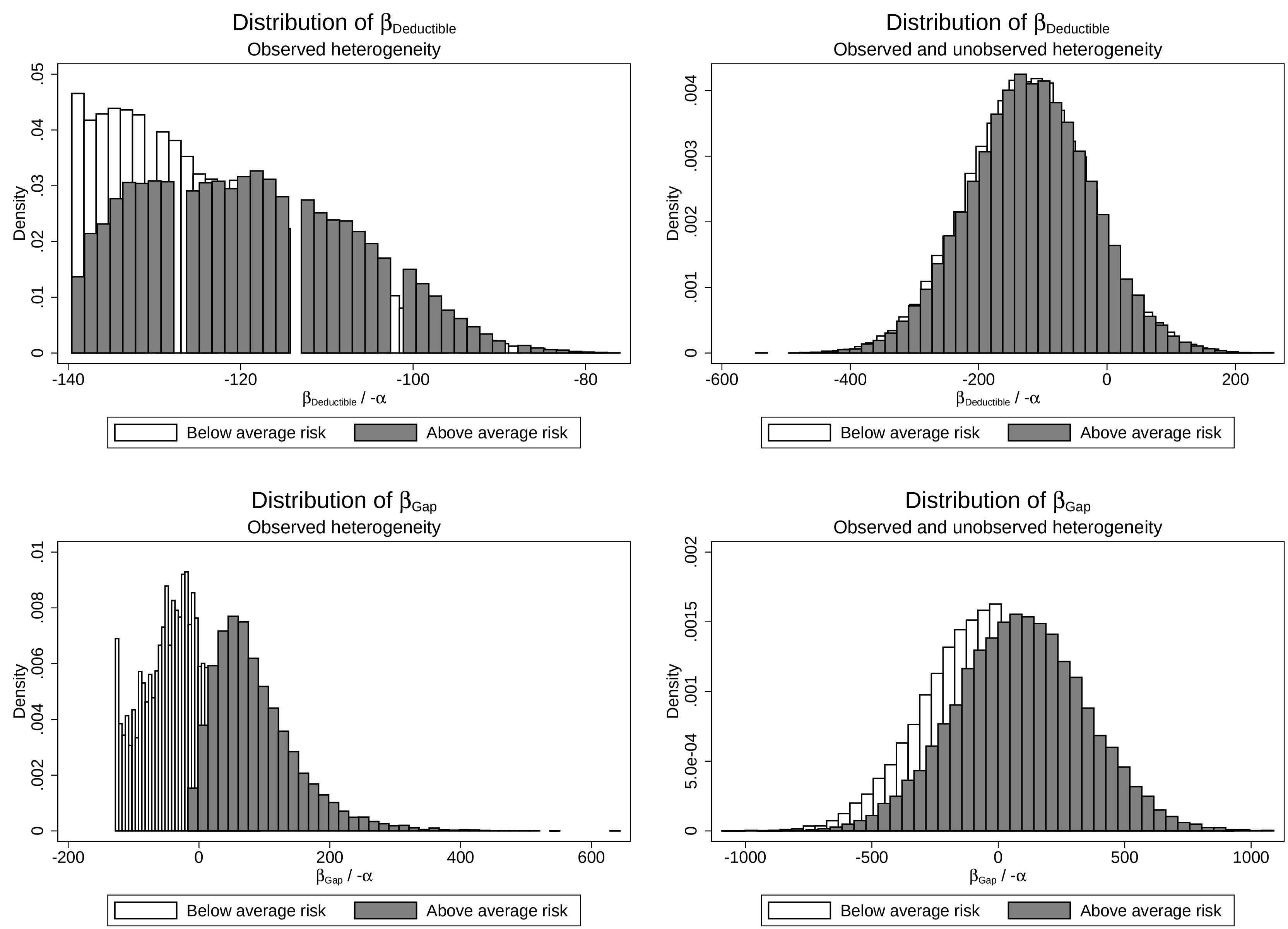
The following set of tables and figures report additional results related to the empirical model in Section III. Table A.2 reports several alternative specifications of the choice model. The main differences across the specification are whether there is observed and unobserved heterogeneity, and if unobserved heterogeneity is present, how it is specified. I also report several specifications that do not use the control function IV to assess how the inclusion of the instrumental variable changes the results. Figure A.1 adds to the reported point estimates by plotting the estimated distributions of observed and unobserved heterogeneity for the baseline model specification. Tables A.3 and A.4 and Figures A.2 and A.3 report several metrics and simulations of model for for the baseline choice model. Table A.5 reports the results of the pricing regression.

Table A.2: **Contract choice model specifications**

	Type of heterogeneity					
	No	No	Observed	Unobserved	Unobserved	LN Unobsered
	(1)	(2)	(3)	(4)	(5)	(6)
Annual premium, \$100	-0.3727	-0.4042	-0.4299	-0.4148	-0.5272	-0.5052
	(0.0070)	(0.0096)	(0.0099)	(0.0078)	(0.0126)	(0.0123)
Deductible, \$100, μ	-0.4329	-0.4405	-0.8430	-1.2329	-1.1415	0.305 [-1.36]
	(0.0104)	(0.0105)	(0.1110)	(0.1266)	(0.1331)	(0.0906)
σ	-	-	-	0.4802	0.4112	0.304 [0.44]
				(0.0232)	(0.0249)	(0.0303)
x Risk	-	-	0.0715	0.0393	0.0529	0.0703
			(0.0329)	(0.0367)	(0.0389)	(0.0362)
ICL, \$100	0.0305	0.0325	-0.0839	-0.1650	-0.1631	-1.648 [-0.19]
	(0.0017)	(0.0017)	(0.0179)	(0.0262)	(0.0276)	(0.1268)
σ	-	-	-	0.0815	0.0849	0.402 [0.087]
				(0.0053)	(0.0055)	(0.0408)
x Risk	-	-	0.0812	0.1052	0.1125	0.1061
			(0.0045)	(0.0066)	(0.0071)	(0.0062)
Partial coverage in gap, 1/0	0.4102	0.4717	-1.4298	-2.0850	-2.0823	0.55 [-1.734]
	(0.0297)	(0.0323)	(0.2754)	(0.3358)	(0.3634)	(0.1856)
σ	-	-	-	1.2640	1.3023	0.979 [3.55]
				(0.0522)	(0.0589)	0.1826
x Risk	-	-	0.8983	1.0954	1.1030	1.0729
			(0.0762)	(0.0897)	(0.0976)	(0.0913)
Default plan, 1/0	5.7324	5.7330	5.1096	5.0675	4.8150	5.0449
	(0.0213)	(0.0213)	(0.2413)	(0.2584)	(0.2917)	(0.2762)
x Risk	-	-	0.2365	0.3589	0.4033	0.3788
			(0.0612)	(0.0655)	(0.0741)	(0.0700)
Heterogeneity in preferences for specific insurers	No	No	Yes (observed)	Yes (obs on top 2)	Yes (obs+ unobs)	Yes (obs + unobs on top 2)
Control Function IV	No	Yes	Yes	No	Yes	Yes
Number of rand. coefficients	0	0	0	3	13	5
Observations	2,435,171	2,435,171	2,435,171	2,435,171	2,435,171	2,435,171
SL at convergence	-60,179	-60,167	-59,536	-59,291	-58,655	-59,106
γ, 75 y.o. female, av.risk	**5.73**	**5.73**	**5.70**	**5.78**	**5.44**	**5.63**
$\frac{\gamma}{\alpha}$, 75 y.o. female, av.risk	**\$1,538**	**\$1,418**	**\$1,326**	**\$1,392**	**\$1,032**	**\$1,114**

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. In Column (6) the square brakets report the median or standard deviation of the random coefficients based on the point-estimates for the mean and s.d. of the natural logarithm of the coefficients.

Figure A.1: **Distributions of observed and unobserved heterogeneity in valuation of deductible and gap coverage**



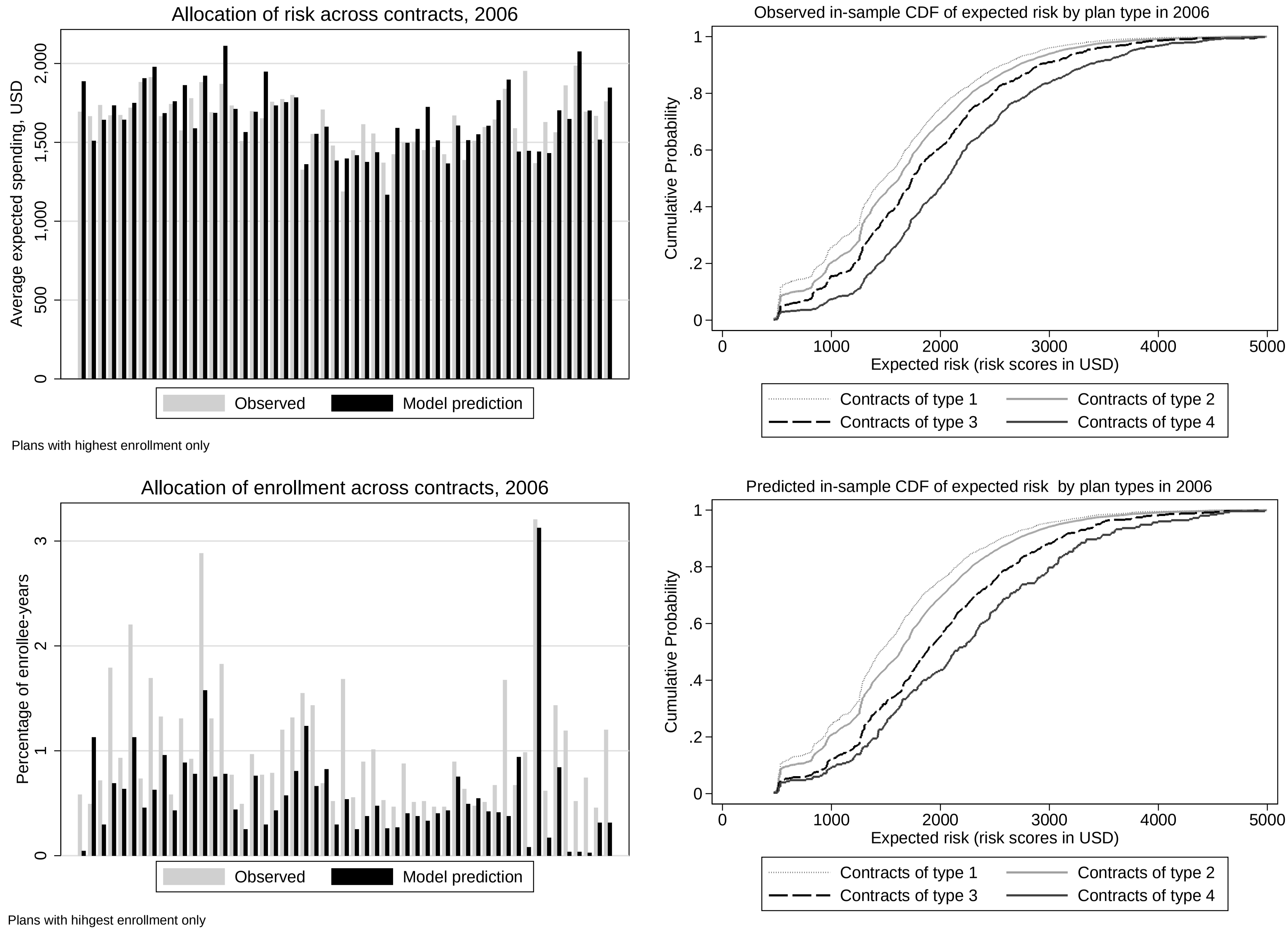
Distributions are based on the estiamted mean and standard deviation of random coefficients from the baseline model specification in Table 4. The distributions are simulated in the estimation sample, and hence reflect the observed distribution of demographics.

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

	Enrollment		
	Observed	Model simulation with observed defaults	Model simulation without observed defaults
Contracts of type 1	21.71%	20.71%	24.16%
Contracts of type 2	65.93%	69.78%	68.34%
Contracts of type 3	10.78%	8.88%	6.94%
Contracts of type 4	1.58%	0.63%	0.56%
Insurer A	29.65%	30.77%	28.78%
Insurer B	27.10%	25.91%	22.32%
	Average risk score		
	Observed	Model simulation with observed defaults	Model simulation without observed defaults
Contracts of type 1	0.85	0.84	0.86
Contracts of type 2	0.88	0.89	0.89
Contracts of type 3	1.01	1.03	1.00
Contracts of type 4	1.04	1.08	1.08
Insurer A	0.92	0.92	0.91
Insurer B	0.86	0.85	0.86

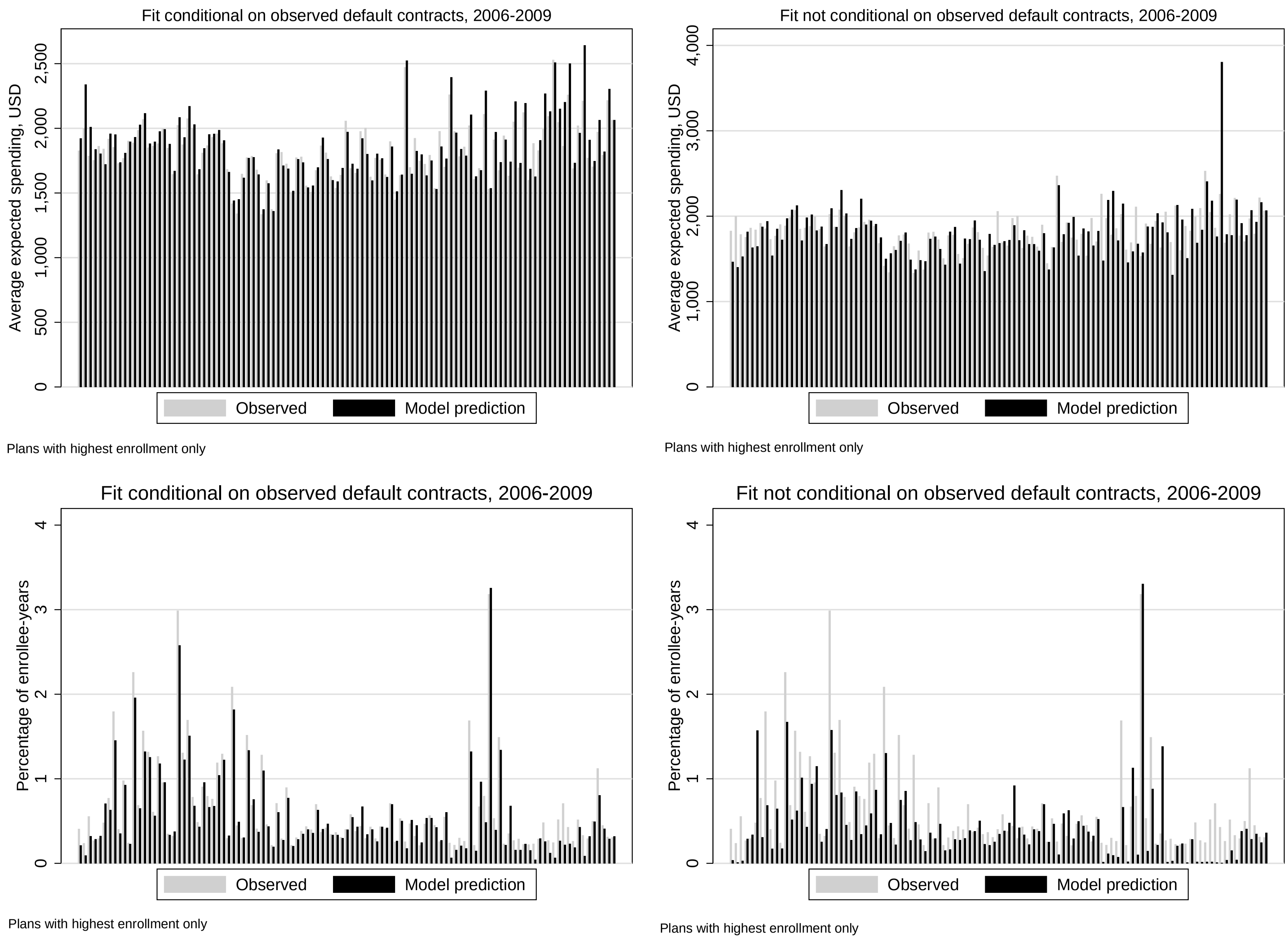
The table compares three within-sample predicted and observed moments in the data: 1) Enrollment shares in different types of plans and in different insurer brands; 2) Average drug spending in different types of plans and in different insurer brands and 3) Average risk scores in different types of plans 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.2. 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.2: **In-sample fit of the choice model: fit in the first year of the program (2006)**



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.

Figure A.3: **In-sample fit of the choice model: panel fit with and without observed defaults**



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**

	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 (conditional on observed defaults)			
Probability of choosing default plan for 66+ y.o. enrollees	86.3%	84.8%	86.3%
N	11,170	11,640	12,197
4. Share predicted in the estimation sample (not conditional on observed defaults)			
Probability of choosing default plan for 66+ y.o. enrollees	89.4%	88.5%	88.9%
N	11,170	11,640	12,197

This tables reports the simulation of the baseline descriptive evidence on the switching rates in the contract choice model.

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.

Model-free evidence from Section II for all years in the data

Figure A.4: **Positive correlation tests: years 2006-2009**

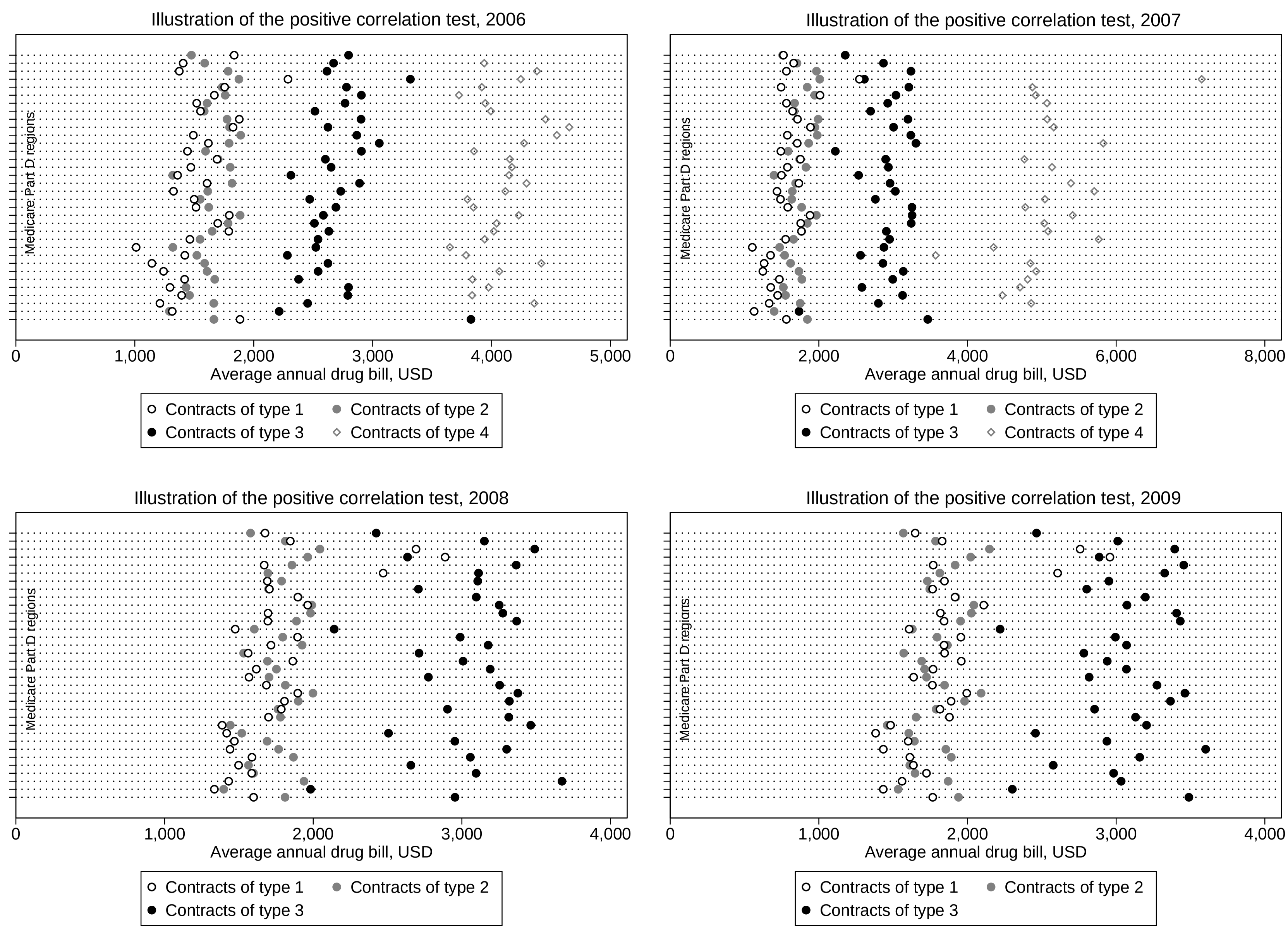


Figure A.5: **Distribution of risks by type of plan: years 2006-2009**

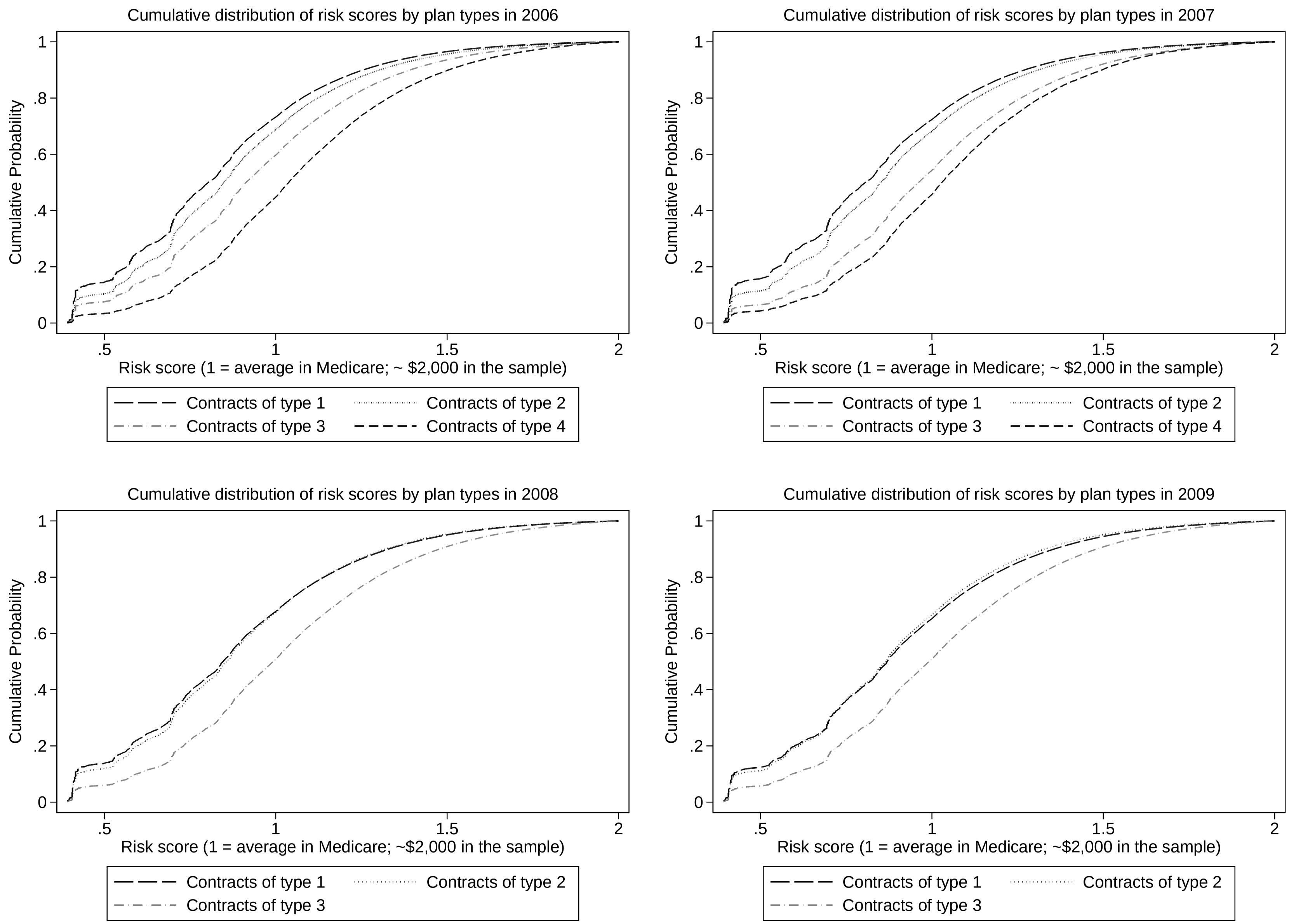
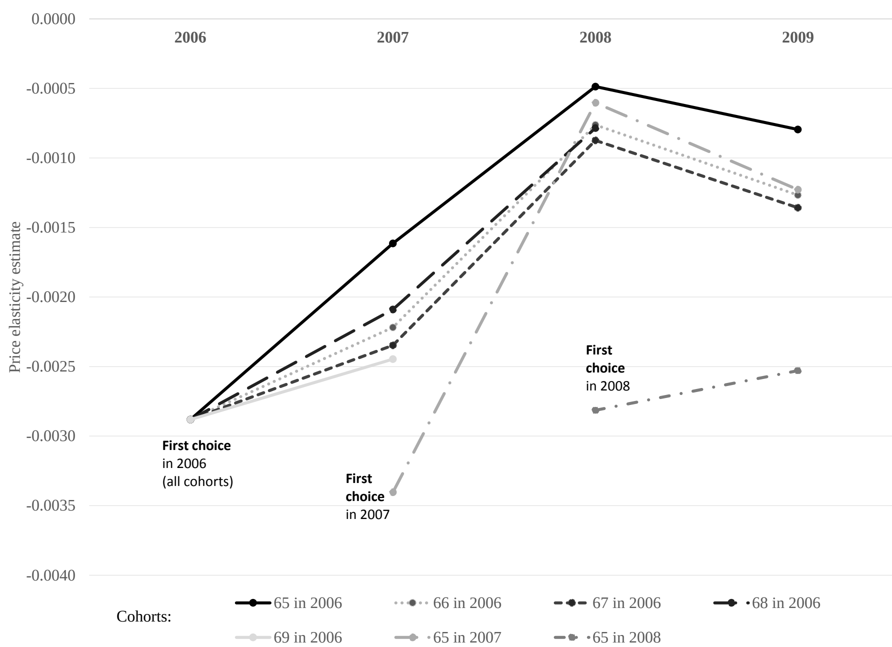


Figure A.6: **Evidence of switching costs: price sensitivity estimates by cohorts over time**



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. 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. 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 graph plots the (sum of) coefficients on premiums in the utility function and not marginal effects.