Final Report for Empirical Exercise 3

Ka Yan CHENG

2022-11-04

Contents

1 Introduction

In this report, we are going to replicate some of the results of

Marzilli Ericson, K. M. (2014). Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange. American Economic Journal: Economic Policy, 6(1), 38-64.

Dataset used in the paper is generously offered by the authors in here.

Here is the Github Repo containing all necessary files to reproduce results in this report.

2 Replication of Table 1

Recreate the table of descriptive statistics (Table 1) from @ericson2014.

Ans:

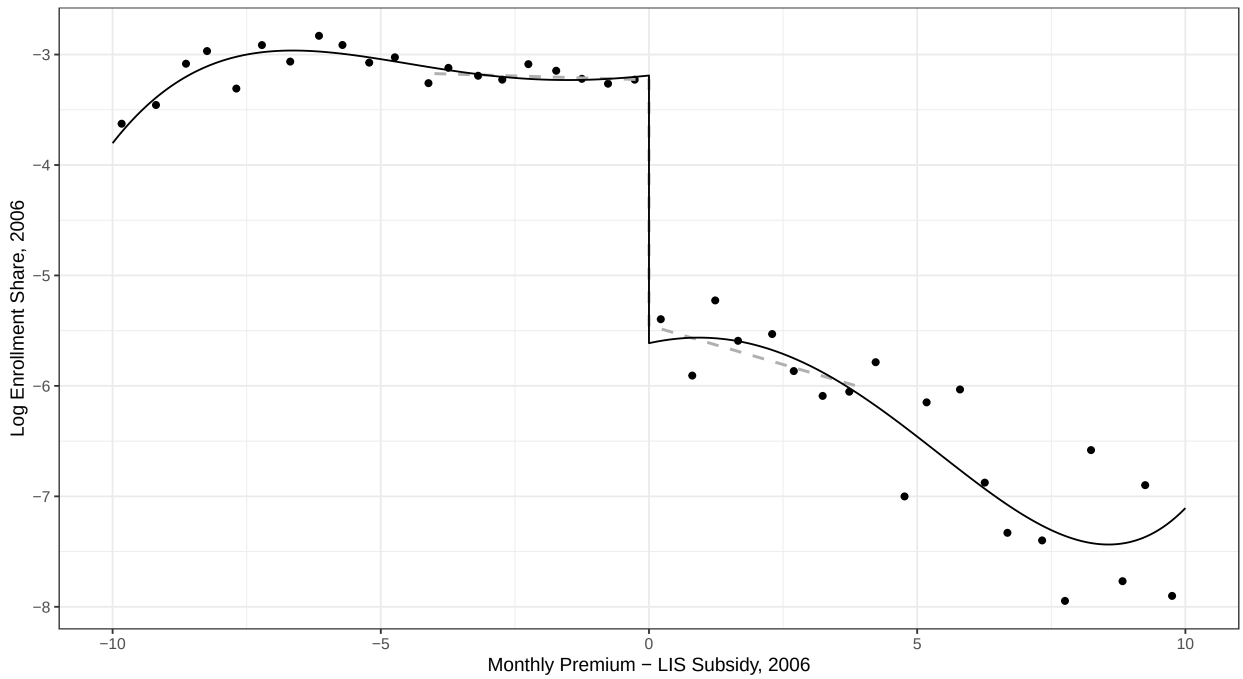
Table 1: Descriptive Statistics of Medicare Part D Plans

	2006	2007	2008	2009	2010
Mean monthly premium	$ 37	$ 40	$ 36	$ 30	$ 33
	(13)	(17)	(20)	(5)	(9)
Mean deductible	$ 92	$ 114	$ 146	$ 253	$ 118
	(116)	(128)	(125)	(102)	(139)
Fraction enhanced benefit	0.43	0.43	0.58	0.03	0.69
Fraction of plans offered by firms already offering a plan ...					
... in the US	0	0.76	0.98	1	0.97
... in the same state	0	0.53	0.91	0.68	0.86
N Unique Firms	51	38	16	5	6
N Plans	1429	658	202	68	107

3 Replication of Figure 3

Recreate Figure 3 from @ericson2014.

Ans:



4 Replication of Figure 3

@calonico2015 discuss the appropriate partition size for binned scatterplots such as that in Figure 3 of Ericson (2014). More formally, denote by $\mathcal{P}_{-,n} = \{P_{-,j} : j = 1, 2, ...J_{-,n}\}$ and $\mathcal{P}_{+,n} = \{P_{+,j} : j = 1, 2, ...J_{+,n}\}$ the partitions of the support of the running variable x_i on the left and right (respectively) of the cutoff, $\bar{x}$. $P_{-,j}$ and $P_{+,n}$ denote the actual supports for each j partition of size $J_{-,n}$ and $J_{+,n}$, such that $[x_l, \bar{x}) = \bigcup_{j=1}^{J_{-,n}} P_{-,j}$ and $(\bar{x}, x_u] = \bigcup_{j=1}^{J_{+,n}} P_{+,j}$. Individual bins are denoted by $p_{-,j}$ and $p_{+,j}$. With this notation in hand, we can write the partitions $J_{-,n}$ and $J_{+,n}$ with equally-spaced bins as

$$p_{-,j} = x_l + j \times \frac{\bar{x} - x_l}{J_{-,n}},$$

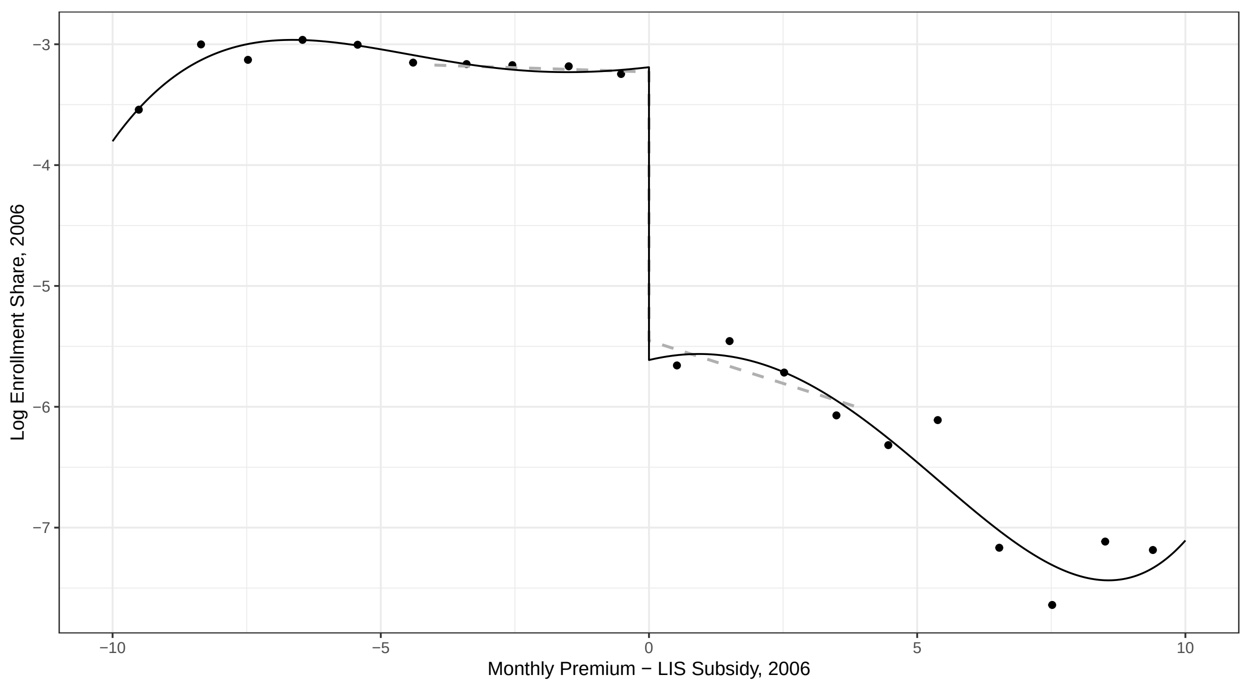
and

$$p_{+,j} = \bar{x} + j \times \frac{x_u - \bar{x}}{J_{+,n}}.$$

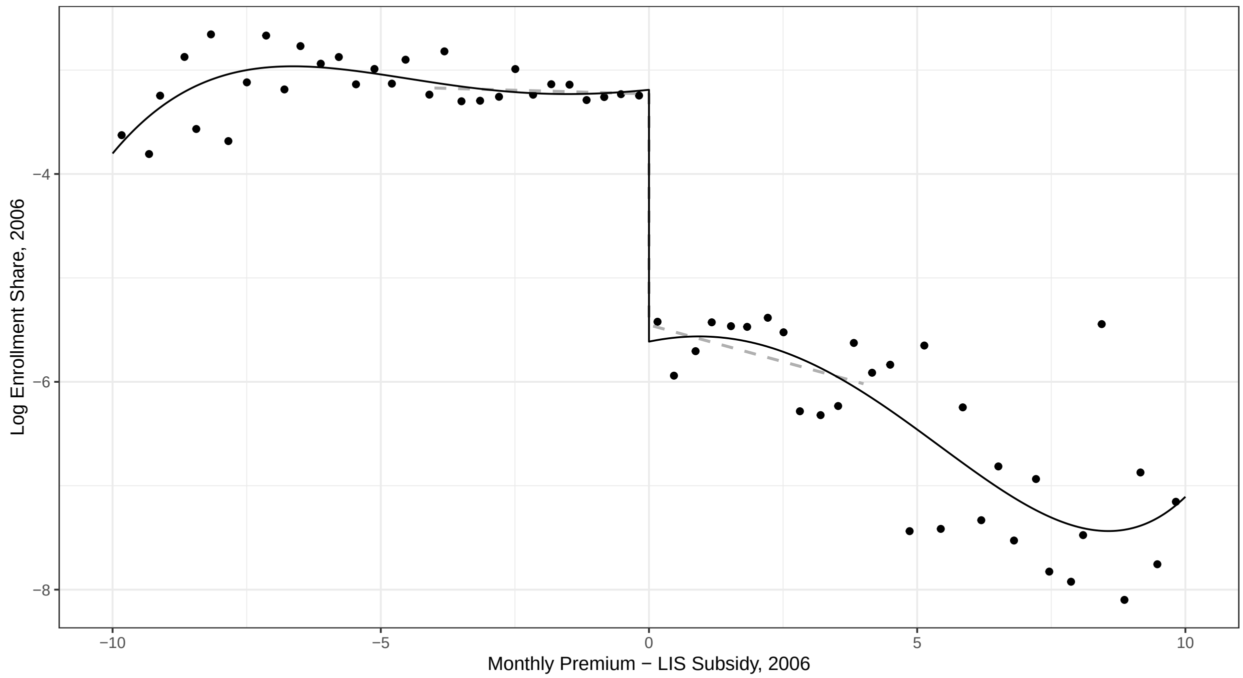
Recreate Figure 3 from Ericson (2014) using $J_{-,n} = J_{+,n} = 10$ and $J_{-,n} = J_{+,n} = 30$. Discuss your results and compare them to your figure in Part 2.

Ans:

For $J_{-,n} = J_{+,n} = 10$:



For $J_{-,n} = J_{+,n} = 30$:



−4
−6
−8
Log Enrollment Share, 2006
−10
−5
0
5
10
Monthly Premium − LIS Subsidy, 2006

5 Optimal bin number

With the notation above, @calonico2015 derive the optimal number of partitions for an evenly-spaced (ES) RD plot. They show that

$$J_{ES,-,n} = \left\lceil \frac{V_-}{\mathcal{V}_{ES,-}} \frac{n}{\log(n)^2} \right\rceil$$

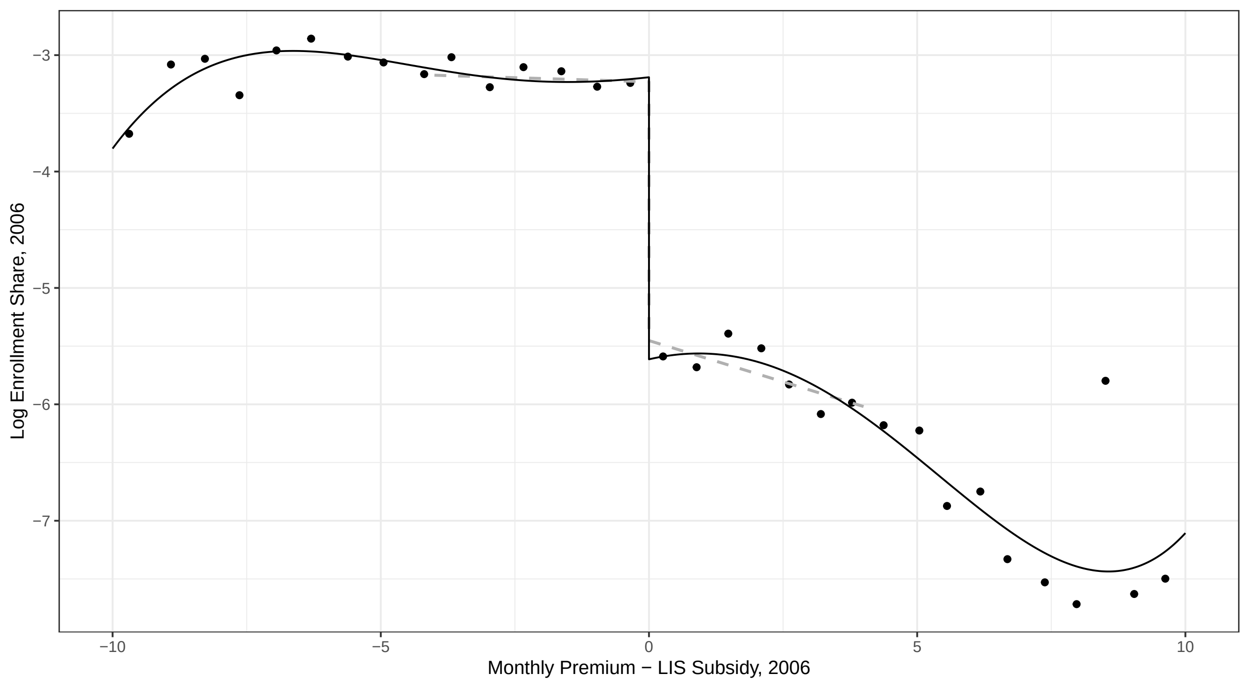
and

$$J_{ES,+,n} = \left\lceil \frac{V_+}{\mathcal{V}_{ES,+}} \frac{n}{\log(n)^2} \right\rceil,$$

where V_- and V_+ denote the sample variance of the subsamples to the left and right of the cutoff and $\mathcal{V}_{ES,\cdot}$ is an integrated variance term derived in the paper. Use the `rdrobust` package in `R` (or `Stata` or `Python`) to find the optimal number of bins with an evenly-spaced binning strategy. Report this bin count and recreate your binned scatterplots from parts 2 and 3 based on the optimal bin number.

Ans:

The optimal bin number is: 15 and 17 on the left hand side and right hand side of the benchmark repectively.



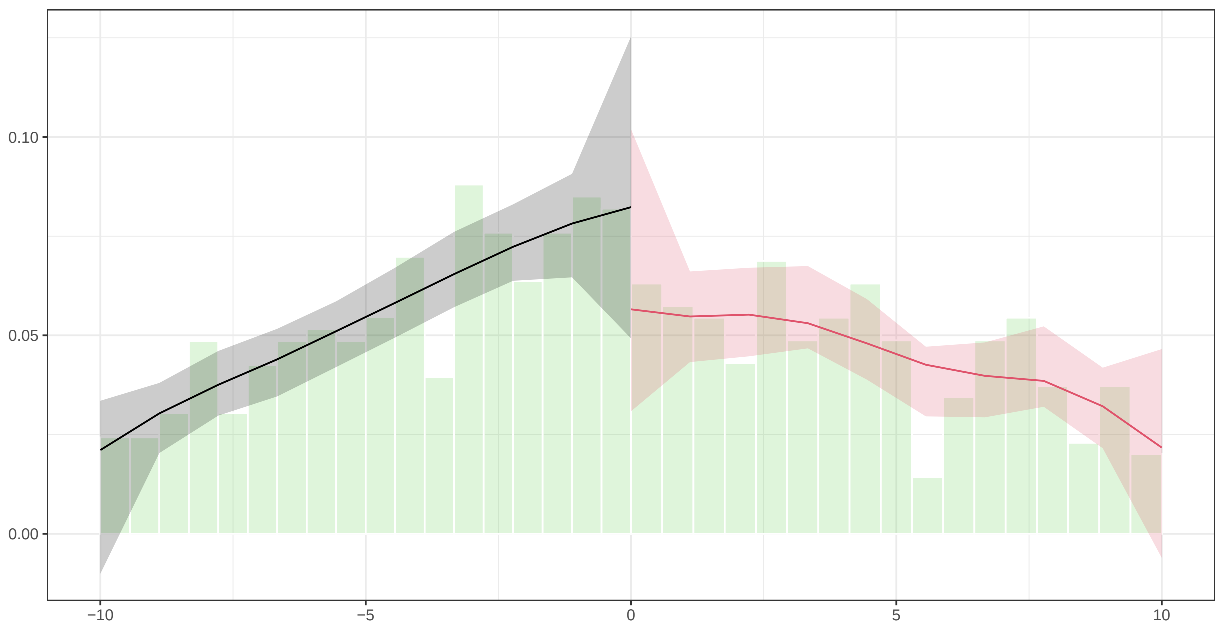
6 Manipulation test

One key underlying assumption for RD design is that agents cannot precisely manipulate the running variable. While "precisely" is not very scientific, we can at least test for whether there appears to be a discrete jump in the running variable around the threshold. Evidence of such a jump may suggest that manipulation is present. Provide the results from the manipulation tests described in @cattaneo2018. This test can be implemented with the `rddensity` package in `R`, `Stata`, or `Python`.

Ans:

The p-value of the test is 0.45, therefore evidence suggests that manipulation exists.

Graphical illustration:



From the above graph, we can see that near the cut-off, confidence intervals to the left and to the right largely overlapped with each other. It provides evidence towards the existence of manipulation visually.

7 Replication of Table 3 - Panel A and B

Recreate Panels A and B of Table 3 in @ericson2014 using the same bandwidth of $4.00 but without any covariates.

Ans:

Table 2: Effect of LIS Benchmark Status in 2006 on Plan Enrollment

	2006	2007	2008	2009	2010
Panel 1: Local linear, bandwidth $4					
Below Benchmark, 2006	2.224***	1.332***	0.902***	0.803*	0.677
	(0.283)	(0.267)	(0.248)	(0.362)	(0.481)
Premium - Subsidy, 2006					
... Below Benchmark	-0.014	-0.077	-0.073	-0.170	-0.215*
	(0.032)	(0.088)	(0.116)	(0.105)	(0.088)
... Above Benchmark	-0.142+	-0.033	0.049	0.074	0.049
	(0.078)	(0.110)	(0.163)	(0.170)	(0.202)
N	306	299	298	246	212
R2	0.576	0.325	0.131	0.141	0.124
Panel 2: Polynomial, bandwidth $4					
Below Benchmark, 2006	2.349***	1.206**	0.697+	0.238	0.152
	(0.279)	(0.387)	(0.394)	(0.516)	(0.633)
Premium - Subsidy, 2006	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic
N	306	299	298	246	212
R2	0.577	0.327	0.137	0.163	0.140

8 Minimal coverage error (CE)-optimal bandwidths

@calonico2020 show that pre-existing optimal bandwidth calculations (such as those used in @ericson2014) are invalid for appropriate inference. They propose an alternative method to derive minimal coverage error (CE)-optimal bandwidths. Re-estimate your RD results using the CE-optimal bandwidth (`rdrobust` will do this for you) and compare the bandwidth and RD estimates to that in Table 3 of @ericson2014.

Ans:

For local linear models:

	2006	2007	2008	2009	2010
Est. coefficient	2.507	-1.010	-1.365	1.382	1.141
Standard error	(0.471)	(0.616)	(0.931)	(0.478)	(0.553)
Optimal bandwidth	0.906	2.368	1.943	3.160	4.304

For quadratic models:

	2006	2007	2008	2009	2010
Est. coefficient	2.894	-0.982	-1.811	0.884	0.926
Standard error	(0.593)	(0.784)	(1.082)	(1.087)	(0.939)
Optimal bandwidth	1.044	3.375	3.128	2.254	3.791

9 Instrumental variable

Now let's extend the analysis in Section V of @ericson2014 using IV. Use the presence of Part D low-income subsidy as an IV for market share to examine the effect of market share in 2006 on future premium changes.

	IV
lnS	-0.178***
	(0.033)
Num.Obs.	4123
R2	-0.572

10 Discussions

Discuss your findings and compare results from different binwidths and bandwidths. Compare your results in part 8 to the invest-then-harvest estimates from Table 4 in @ericson2014.

11 Conclusion

Reflect on this assignment. What did you find most challenging? What did you find most surprising?