Assignment 1, Health Econ 2

Paul George

2022-09-20

1. Provide and discuss a table of simple summary statistics showing the mean, standard deviation, min, and max of hospital total revenues and uncompensated care over time.

We do not have reliable data from 1997-2002 for uncompensated care, that data seems to not be reported. As such, I will focus my analysis on the years 2003-2021, given the overarching point of this assignment is regarding uncompensated care. As we can see below, total revenue increases steadily over time, whereas uncompensated care increases until 2010, and then sharply decreases in 2011 and remains somewhat stable from 2011 until 2011.

Uncompensated Care by Year

in millions USD

year	N	Total Uncompensated Care	sd	min	max
2003	2782	11.245	18.997	0.000	145.869
2004	2692	12.454	19.186	0.000	144.864
2005	2498	13.924	20.119	0.000	146.614
2006	2359	16.202	21.695	0.000	147.561
2007	2348	17.824	22.863	0.000	141.751
2008	2400	19.857	25.192	0.000	146.677
2009	2454	21.428	26.504	0.000	146.813
2010	2680	21.814	26.500	0.000	147.751
2011	4069	11.708	20.105	0.000	144.268
2012	4198	11.729	20.129	0.000	146.766
2013	4171	12.472	21.359	0.000	147.795
2014	4138	10.709	19.547	0.000	145.885
2015	4085	8.653	17.085	0.000	141.638
2016	4038	9.418	18.758	0.000	147.310
2017	4024	10.939	20.472	0.000	144.808
2018	3977	11.977	22.284	0.000	147.518
2019	3855	12.666	22.751	0.000	148.044
2020	3871	12.786	23.031	0.000	147.984
2021	3162	12.531	23.182	0.000	147.864

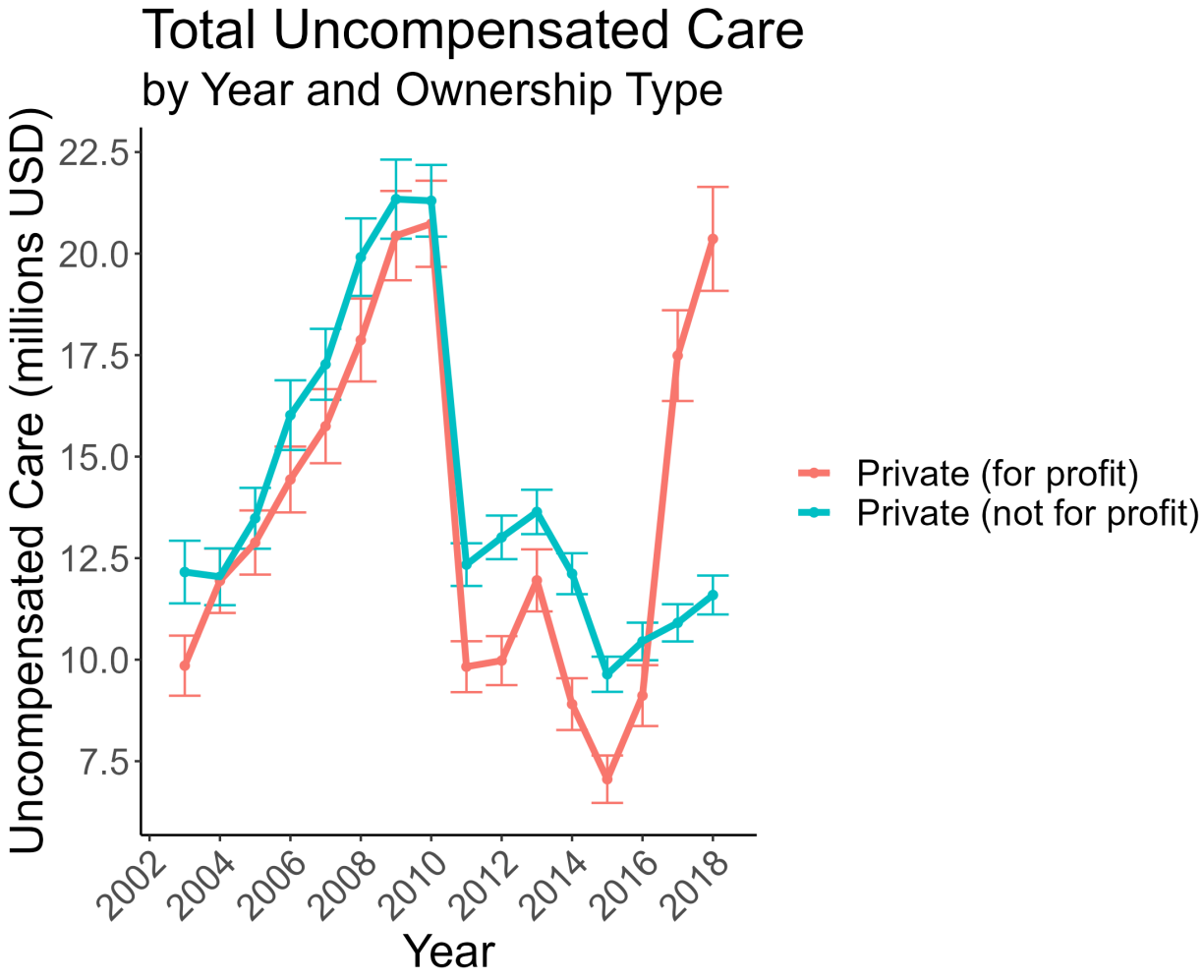
Total Revenue by Year

in millions USD

year	N	Total Revenue	sd	min	max
2003	5745	185.034	293.458	0.048	2,440.415
2004	5828	202.418	318.329	0.154	2,525.756
2005	5852	218.625	342.245	0.000	2,498.876
2006	5790	233.576	355.711	0.205	2,539.261
2007	5755	250.168	376.424	0.064	2,509.578
2008	5760	267.956	402.808	0.000	2,531.531
2009	5727	287.612	427.161	0.119	2,539.917
2010	5712	300.156	442.593	0.307	2,510.376
2011	5683	317.105	463.657	0.003	2,516.088
2012	5706	326.927	475.575	0.214	2,531.756
2013	5645	336.051	489.770	0.095	2,529.477
2014	5618	348.548	502.707	0.007	2,533.635
2015	5585	366.206	522.515	0.009	2,524.713
2016	5556	378.138	535.077	0.085	2,539.722
2017	5462	375.565	524.972	0.125	2,526.277
2018	5386	384.958	535.837	0.283	2,541.196
2019	5318	389.973	540.656	0.000	2,541.045
2020	5284	387.441	539.848	0.019	2,538.235
2021	4462	403.920	559.765	0.000	2,537.481

.

2. Create a figure showing the mean hospital uncompensated care from 2003 to 2019. Show this trend separately by hospital ownership type (private not for profit and private for profit).



.

3. Using a simple DD identification strategy, estimate the effect of Medicaid expansion on hospital uncompensated care using a traditional two-way fixed effects (TWFE) estimation.

Table: Two Way Fixed Effects Models

in millions USD

Model	Estimate	Std..Error	p value	N
Model 1	−7.13	0.22	0.00	52819
Model 2	−8.13	0.27	0.00	41594
Model 3	−5.32	0.56	0.00	21096
Model 4	−7.31	0.85	0.00	19296

.

In the above table, Model 1 is the full sample, Model 2 limits to 2014, model 3 to 2015, and model 4 to 2016. The estimates are roughly the same, showing that there is a significant decrease in uncompensated care for hospitals in states that expanded. However, the standard errors increase from Model 1 to Model 4, which could be a function of the decreasing number of observations from Model 1 to 4.

4. Estimate an "event study" version of the specification in part 3.

Event Study Estimation		
Year Before Expansion in State	Model 1	Model 2
-10	0.8540 (1.690)	1.381 (1.910)
-9	0.8666 (1.659)	1.403 (1.927)
-8	0.7547 (1.431)	1.234 (1.620)
-7	0.0176 (1.268)	-0.2200 (1.453)
-6	-0.1067 (1.258)	-0.6623 (1.606)
-5	-0.6622 (1.100)	-1.688 (1.611)
-4	-0.6138 (0.9442)	-1.393 (1.465)
-3	-0.3026 (0.6149)	0.9861 (0.7571)
-2	-0.2336 (0.4245)	0.5583 (0.5768)
0	-2.546*** (0.5659)	-2.814** (0.8085)
1	-6.216*** (0.9992)	-5.066*** (1.325)
2	-8.748*** (1.211)	-7.096*** (1.242)
3	-10.95*** (1.643)	-11.26*** (2.064)
4	-12.95*** (2.130)	-13.83*** (2.693)
Fixed Effects	----------	----------
Provider Number	Yes	Yes
Year	Yes	Yes
	_____	_____
SE Clustered	by: State	by: State
Observations	52,819	41,594
R2	0.73139	0.72547
Within R2	0.03271	0.03455

Note that we can see similar trends prior to the cutoff, and decreasing values of our β coefficient after the cutoff.

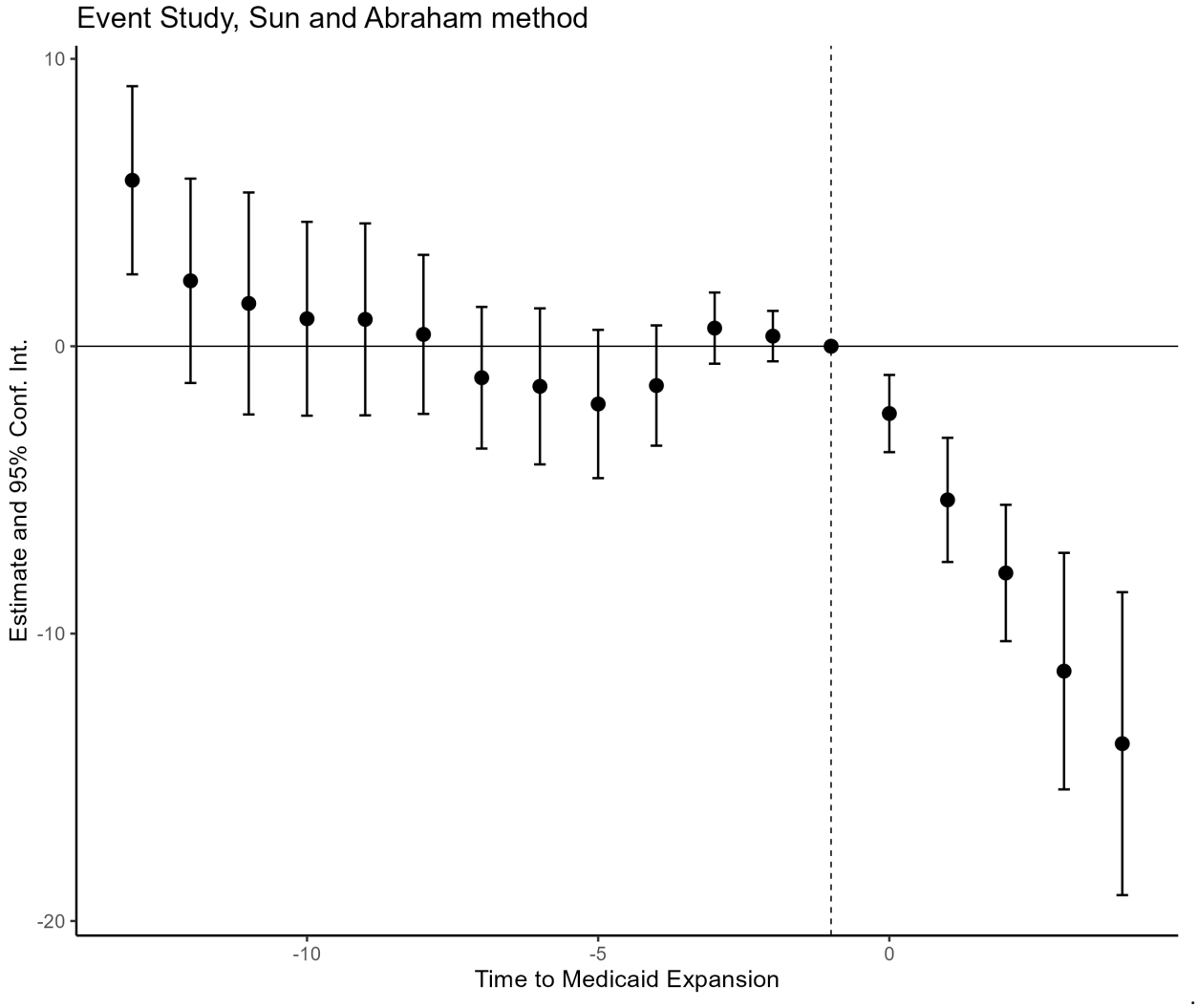
5. Sun and Abraham (SA) show that the coefficients in Equation (2) can be written as a non-convex average of all other group-time specific average treatment effects; provide an alternative event study using the CS estimator.

Table: Event Study

Sun and Abraham method, Expansion Years 2014, 2015, 2016

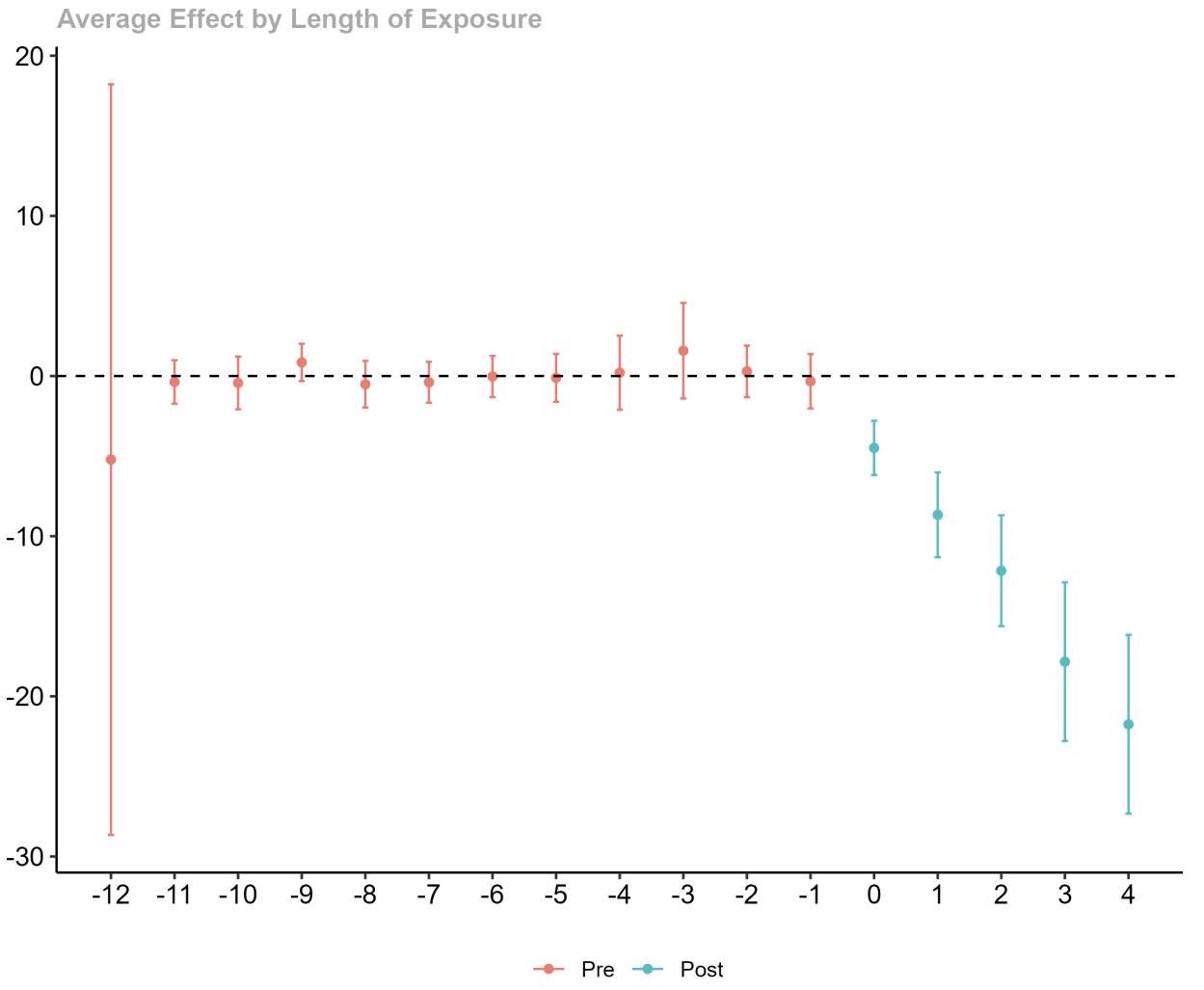
Year	Estimate	Standard Error	p value
-13	5.78	1.67	0.00
-12	2.28	1.81	0.22
-11	1.49	1.97	0.45
-10	0.96	1.72	0.58
-9	0.94	1.70	0.59
-8	0.41	1.41	0.77
-7	−1.10	1.26	0.39
-6	−1.40	1.39	0.32
-5	−2.01	1.32	0.13
-4	−1.37	1.07	0.21
-3	0.63	0.63	0.32
-2	0.35	0.45	0.44
0	−2.34	0.69	0.00
1	−5.35	1.10	0.00
2	−7.89	1.21	0.00
3	−11.31	2.10	0.00
4	−13.83	2.69	0.00

6. Present an event study graph based on the results in part 5.



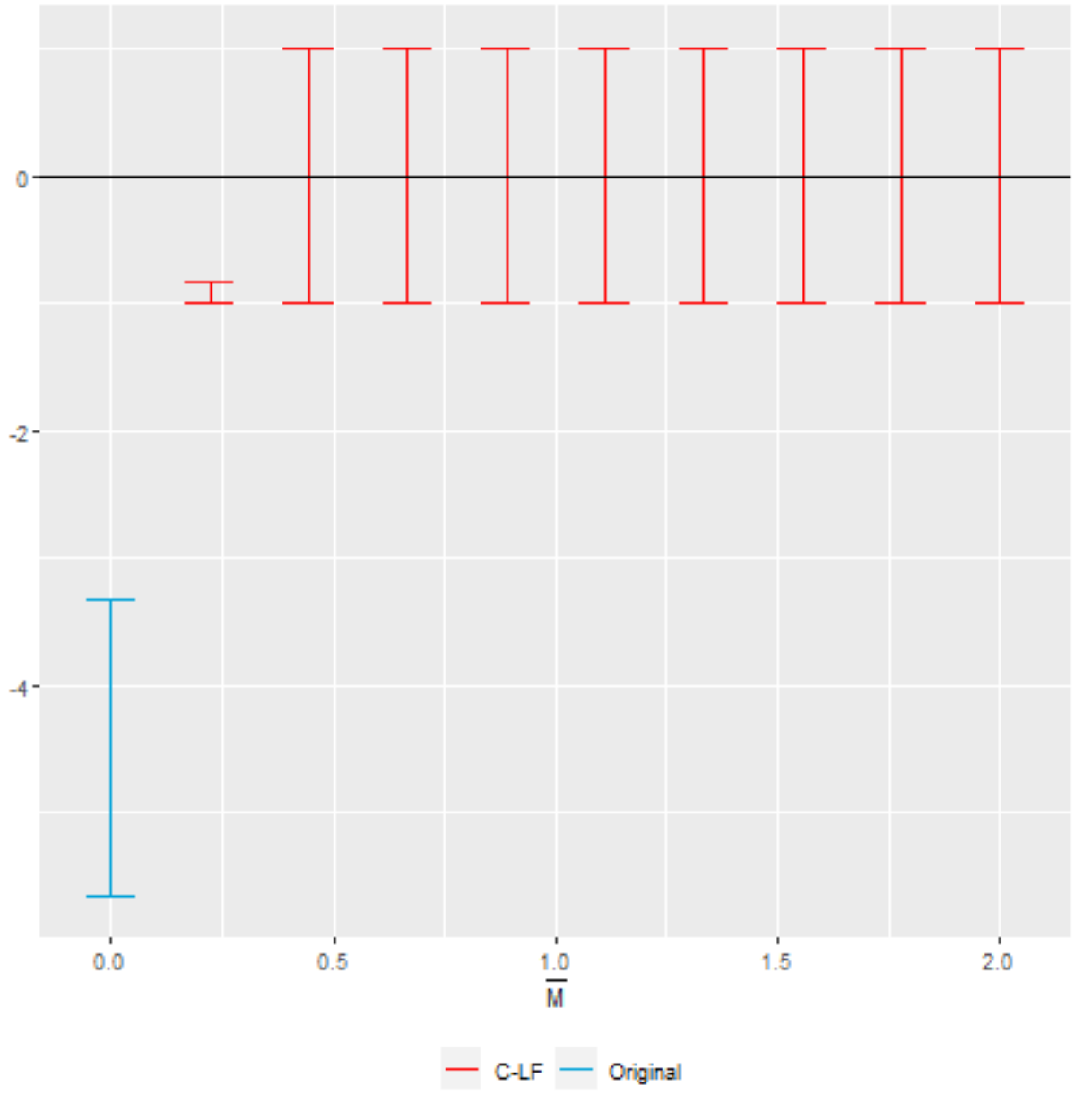
7. Callaway and Sant'Anna (CS) offer a non-parametric solution that effectively calculates a set of group-time specific differences.

We will aggregate the group-time effects into a single event study plot, but averaging the average treatment effects at different lengths of exposure to the treatment.



8. Using the HonestDiD package in R or Stata, present a sensitivity plot of your CS ATT estimates.

Note that in the above graph, there does seem to be a clear inflection at year -1. Nonetheless, we will do a sensitivity analysis using Rambachan and Roth (2021) for violation of parallel trends.



.

9. Discuss your findings and compare estimates from different estimators (e.g., are your results sensitive to different specifications or estimators? Are your results sensitive to violation of parallel trends assumptions?).

As we can see above, the different estimators actually give very similar results. Even when choosing different years to evaluate, we continue to see that Medicaid expansion is clearly associated with decreases in uncompensated care, though, as expected, standard errors increase with fewer observations. Surprisingly, the RR specification seems to show a violation of parallel trends, though given the results prior, I'm concerned that is a mistake in my implementation of RR as opposed to a true result (though in several attempts to de-bug question 8, I was unsuccessful).

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

I found the initial part, the data wrangling, to be most frustrating. I was not particularly surprised by this, as in other projects with large, public datasets I have found this aspect to be very difficult and time-consuming. I think that's an under-appreciated aspect in working with these datasets, is the time and care needed to correctly unpack the data. I was pleasantly surprised at how easy it was to implement Callaway and Sant'Anna and Abraham and Sun; the R packages are very convenient and easy to use (as opposed to RR, at least in my experience). If I were an econometrician, that would be a key takeaway from this assignment; that if I want my estimator to be widely used in practice, it is important to have easy-to-use, well-documented support for the actual implementation. Finally, I was quite surprised how quickly uncompensated care drops off at time = 0. I would have expected a several year delay, as oftentimes it takes several years for a policy that is passed to come into actual implementation. But the event studies above show a quick and clear reduction in uncompensated care. Finally, I found it quite difficult to convert everything into R-markdown, but perhaps this has laid the groundwork for future work, because the final product is quite nice.