Do school principals respond to public scrutiny? New survey evidence from Australia

Michael Coelli and Gigi Foster
(and Andrew Leigh?)

ACE 2017
Sydney

What are we doing?

- Using an education-related policy experiment to explore the effects of school accountability
- Scope: Australian schools of all types and sectors
- 2010: My School goes live. School-level national test scores made publicly accessible.
- Our novel contribution: pre and post-My School surveys to all school principals
- Main question:
  - How have schools changed their policies or practices in response to My School?

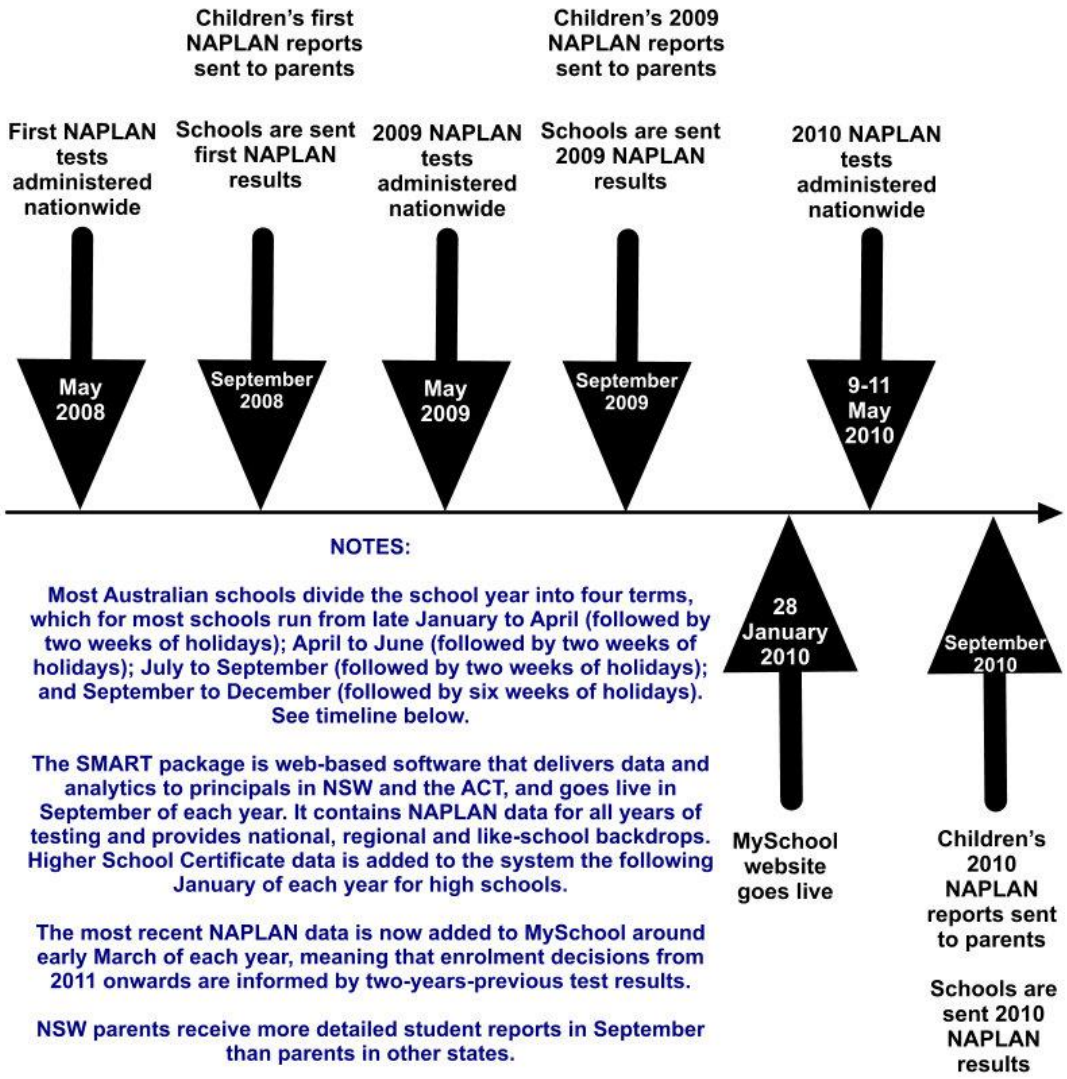
The conceptual underpinning

- Reducing information barriers in the education market should lead to more pressure on schools through…
  - Students and parents voting through enrolment decisions
  - State departments of education applying pressure to lower-performing schools (to some extent already happening pre-My School)
- Schools that look worse on My School should in theory have a stronger incentive to invest in score improvements
- Caveat: Apparent score improvements can be generated through low-hanging-fruit mechanisms unrelated to actual learning improvements

Why this project is not a breeze

- The measurement of both degree of predicted influence, and outcome, is fiddly
  - Multi-dimensionality of national tests
  - Numerical translation of a visual display of relative scores
- Survey-sourced data, with all its potential biases

MySchool Timeline



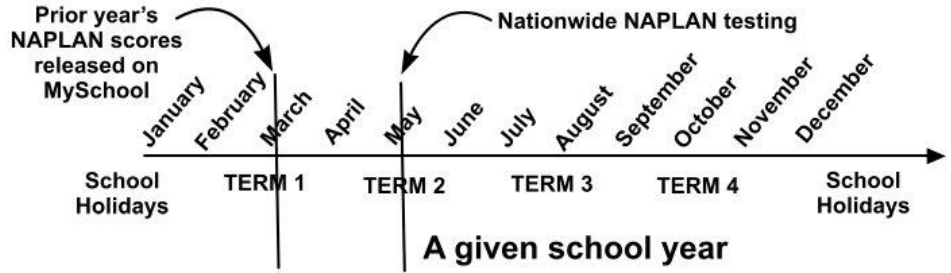
NOTES:

Most Australian schools divide the school year into four terms, which for most schools run from late January to April (followed by two weeks of holidays); April to June (followed by two weeks of holidays); July to September (followed by two weeks of holidays); and September to December (followed by six weeks of holidays). See timeline below.

The SMART package is web-based software that delivers data and analytics to principals in NSW and the ACT, and goes live in September of each year. It contains NAPLAN data for all years of testing and provides national, regional and like-school backdrops. Higher School Certificate data is added to the system the following January of each year for high schools.

The most recent NAPLAN data is now added to MySchool around early March of each year, meaning that enrolment decisions from 2011 onwards are informed by two-years-previous test results.

NSW parents receive more detailed student reports in September than parents in other states.



Administrative data from ACARA

- School sector; enrolment; rural/urban; grades offered
- NAPLAN score average for each of the five domains, for each grade cohort in each of years 2008 through 2012
- Information drawn directly from ACARA that was used to build the “traffic-lights” display on MySchool that is arguably the easiest way to compare performance across schools

School profile
School finances
NAPLAN
Results in graphs
Results in numbers
Results in bands
Student gain
Similar schools
VET in schools
Local schools

Randwick Public School, Randwick, NSW

Bookmark this page

Results in numbers

The National Assessment Program – Literacy and Numeracy (NAPLAN) assesses all students in Australian schools in Years 3, 5, 7 and 9. For more information visit the NAPLAN website.

The chart below displays average NAPLAN scores for each domain. The selected school's scores are displayed in blue. Also displayed are average scores for statistically similar schools (SIM) and all Australian schools (ALL). The coloured bars indicate whether the selected school's scores are above, close to, or below the other scores.

2008 | 2009 | 2010 | 2011 | 2012 | **2013**

Colour Scheme: Red & Green ▼ Submit

Alternate view: Results in graphs

	Reading		Persuasive Writing		Spelling		Grammar and Punctuation		Numeracy	
Year 3	474 461 - 487		463 452 - 474		457 445 - 469		495 481 - 509		462 451 - 474	
	SIM 477 468 - 486	ALL 419	SIM 451 442 - 459	ALL 416	SIM 450 441 - 458	ALL 411	SIM 482 472 - 492	ALL 428	SIM 444 436 - 452	ALL 397
Year 5	543 529 - 557		512 498 - 525		535 522 - 548		554 539 - 570		530 517 - 543	
	SIM 548 539 - 556	ALL 502	SIM 514 505 - 522	ALL 478	SIM 529 521 - 537	ALL 494	SIM 549 540 - 558	ALL 501	SIM 535 527 - 543	ALL 496

How to interpret this chart

SIM schools serving students from statistically similar backgrounds

ALL Australian schools' average

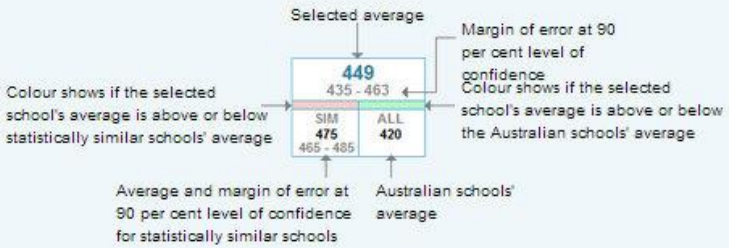
Student population below reporting threshold

Year level not tested

Selected school's average is

- **substantially above**
- **above**
- **close to**
- **below**
- **substantially below**

- average of schools serving students from statistically similar socio-educational backgrounds (SIM box)
- average of all Australian schools (ALL box)



Tell me more

Survey data

- In each of 2009 and 2012: a core module plus one of four additional modules sent to school principals
  - In 2009: sent to the principals of all schools in Australia (response rate = 20.5%)
  - In 2012: sent to principals of all responding schools from the 2009 survey (response rate = 58%)
- Questions replicating and/or customizing those appearing in Rouse et al (2013)

Rouse, Cecilia Elena, Jane Hannaway, Dan Goldhaber and David Figlio (2013). Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure. *American Economic Journal: Economic Policy* 5 (2): 251-281.

What we do

- Follow Rouse et al. (2013) method, in relation to combining survey responses into **domains** of policies/practices (normalize, then average within domain), and in terms of econometric approach
- Australia-specific quality indicator of each school: its average (across all domains and grades) normalized mean NAPLAN score in 2008 and 2009

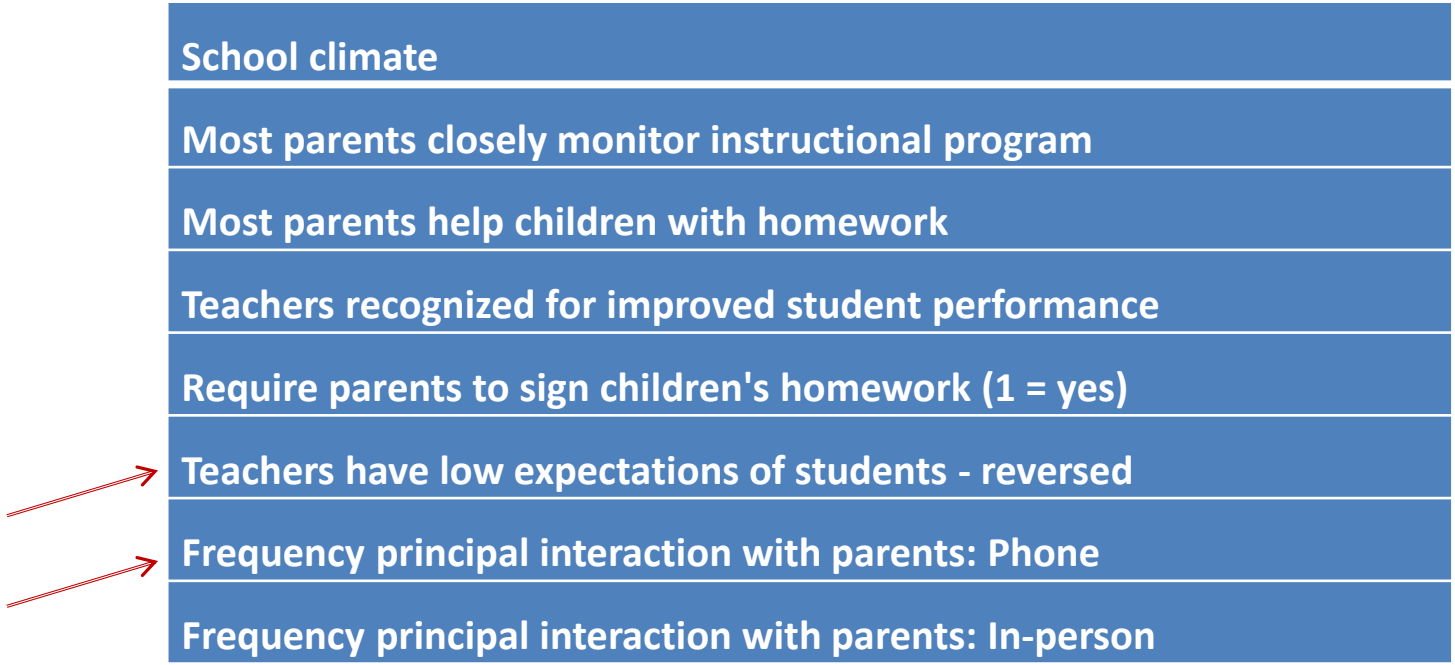
$$P_{jt} = \alpha_j + \beta_j \cdot SQ_{t-1} + X_{t-1} \cdot \Gamma_j + \varepsilon_{jt}$$

$$\beta_d = \frac{1}{J} \sum_{j=1}^{J} \frac{\beta_j}{\sigma_j}$$

What's in X? State, sector, location, size, Indigenous %, LBOTE %, policy in 2009

- SUR within each domain to generate covariances across the policy-level betas with which to correct the standard errors of the domain-level betas

Low-Performing Students	What special measures does your school take to try to improve the performance of low-performing students?	Five answers suggested (of which multiple could be selected): -Recommend to parents that the student repeat the grade -Additional tutoring during regular school hours -Before-school or after-school tutoring paid for by the school -Saturday classes -Develop an individual learning plan for the student - Other policy
Lengthening Instructional Time	For students attending the middle grade taught at your school, what time does the school day typically start/finish?	
Reduced Class Size for Subject	In order to give extra attention to particular subjects, such as maths, some schools use additional school staff to teach that subject so that the effective class size (number of students per teacher) for that subject is smaller than it is for other subjects. What is the typical number of students per teacher when <student type> receive regular instruction in smaller sections: in <subject>?	Student type is gifted students, students with academic difficulties, or ESL students. Subject type is maths, reading, or writing.
Narrowing of Curriculum	How much time do your students typically spend on the following subjects in an average week?	Subjects (with time counted separately) are maths, writing, science, reading, art and music, social studies, and PE/sport. Minimum time indicator used in domain construction.
Low-Performing Teachers	What special measures does your school take to try to improve low-performing teachers?	Five answers suggested (of which multiple could be selected): -Supervise teachers more closely -Assign an aide to teachers -Assign teachers to mentors or leading teachers -Provide additional professional development for individual teachers -Coaching from yourself -Other policy
Teacher Resources	In an average week, about how many hours per week do your teachers get for class preparation?	
Teacher Incentives	(A) Does your school use any of the following incentives to reward teacher performance? (B) Which of the following types of compensation does your school use to reward teacher performance?	(A) Four answers suggested (of which multiple could be selected): -Special leadership position/assignment (mentor teacher, curriculum development) -Choice of class -Release time from teaching -Attendance at conferences and workshops (B) Three mutually-exclusive answers suggested: -Permanent increase to base salary -One off performance bonus -Both types of rewards
School Climate	No questions that match those in Rouse et al. (2013)	
Control	Indicate how much actual influence each of the following actors has on decisions concerning the following activities in your school: <area>	Actors include (with influence counted separately): -Teachers at this school -State/Territory education department Areas include: -Establishing curriculum -Evaluating teachers -Hiring new full-time teachers -Deciding how this school's budget will be spent


School climate
Most parents closely monitor instructional program
Most parents help children with homework
Teachers recognized for improved student performance
Require parents to sign children's homework (1 = yes)
Teachers have low expectations of students - reversed
Frequency principal interaction with parents: Phone
Frequency principal interaction with parents: In-person

	2009 survey	2009 survey	p-value for difference
Variable	Respondents	Non-respondents	
Number of students	375.1	380.8	0.539
	(353.1)	(354.3)	
ICSEA score	1,004.5	999.8	0.057
	(91.5)	(103.3)	
Indigenous (%)	6.90	8.51	0.000
	(14.48)	(18.01)	
LBOTE in 2010 (%)	16.42	17.30	0.151
	(22.95)	(23.83)	
Government (%)	70.5	71.4	0.458
Catholic (%)	19.2	17.9	0.203
Independent (%)	10.3	10.7	0.614
Primary (%)	69.2	70.4	0.322
Secondary (%)	16.6	15.1	0.128
Combined (%)	14.2	14.5	0.783
Metropolitan (%)	53.8	55.0	0.342
Provincial (%)	38.9	38.1	0.489
Remote (%)	4.6	3.7	0.079
Very remote (%)	2.8	3.3	0.249
Average standardised scores	-0.065	-0.073	0.552
	(0.460)	(0.537)	
Observations	1,890	7,351	

	2009 and 2012 survey	2012 survey	P-value for difference
Variable	Respondents	Non-respondents (of 2009 respondents)	
Number of students	404.0	360.1	0.010
	(369.3)	(347.9)	
ICSEA score	1,008.5	997.2	0.010
	(89.1)	(94.7)	
Indigenous (%)	6.96	8.41	0.049
	(14.27)	(15.21)	
LBOTE (%)	17.35	14.94	0.025
	(23.75)	(21.85)	
Government (%)	66.6	72.6	0.006
Catholic (%)	22.1	15.7	0.001
Independent (%)	10.8	9.7	0.447
Primary (%)	67.3	72.9	0.010
Secondary (%)	18.4	13.9	0.010
Combined (%)	15.0	13.5	0.381
Metropolitan (%)	56.4	49.6	0.004
Provincial (%)	36.7	42.3	0.014
Remote (%)	4.4	4.8	0.626
Very remote (%)	2.5	3.2	0.380
Average standardised scores	-0.067	-0.110	0.071
	(0.468)	(0.484)	
Observations	1,076	784	

Domain: 2009 policy settings by "quality", relative to schools with no dark red indicators	0< dark red <20%	20%+ dark red
Policies to improve low-performing students	-0.00571	-0.0301
	(0.0322)	(0.0275)
Lengthening instructional time	-0.00831	-0.298***
	(0.102)	(0.0856)
Reduced class size for subject	0.0307	0.0455
	(0.196)	(0.160)
Narrowing of curriculum	0.0921	-0.0740
	(0.0886)	(0.0770)
Policies to improve low-performing teachers	0.0302	0.0983**
	(0.0486)	(0.0415)
Teacher resources	0.0410	-0.302**
	(0.173)	(0.143)
Teacher control	-0.0719	-0.192*
	(0.132)	(0.116)
State control	0.205*	0.259**
	(0.115)	(0.101)
Principal control	-0.0149	-0.216*
	(0.130)	(0.114)
School climate	-0.105**	-0.192***
	(0.0510)	(0.0435)

Domain: 2009 settings by sector	Independent	Catholic
(relative to government schools)		
Policies to improve low-performing students	0.165***	0.0226
	(0.0341)	(0.0263)
Lengthening instructional time	0.710***	0.551***
	(0.0830)	(0.0653)
Reduced class size for subject	0.0177	0.0131
	(0.155)	(0.128)
Narrowing of curriculum	-0.107	-0.0843
	(0.0683)	(0.0596)
Policies to improve low-performing teachers	0.0417	0.0557*
	(0.0402)	(0.0310)
Teacher resources	0.651***	-0.113
	(0.217)	(0.173)
Teacher incentives		-0.079**
		(0.031)
Teacher control	-0.159	-0.119
	(0.115)	(0.0907)
State control	-1.165***	-0.865***
	(0.108)	(0.0853)
Principal control	0.258**	0.322***
	(0.116)	(0.0914)
School climate	0.0904*	0.147***
	(0.0490)	(0.0377)

	Government		Independent	Catholic
Domain: Changes 2009-2012 by sector	Primary	all		
Policies to improve low-performing students	-0.0162	-0.0009	-0.0103	-0.0701*
	(0.0196)	(0.0180)	(0.0588)	(0.0387)
Lengthening instructional time	-0.0003	-0.0022	0.1252	0.0882*
	(0.0222)	(0.0311)	(0.0880)	(0.0466)
Reduced class size for subject	0.1005	0.0718	0.3929*	-0.1962
	(0.1244)	(0.1000)	(0.2203)	(0.1909)
Narrowing of curriculum	0.0290	-0.0371	-0.1873*	0.1724*
	(0.0512)	(0.0455)	(0.1082)	(0.1046)
Policies to improve low-performing teachers	0.0937***	0.0693***	-0.0051	-0.0319
	(0.0281)	(0.0229)	(0.0484)	(0.0363)
Teacher resources	0.1643	0.1002	-0.0358	0.0396
	(0.1353)	(0.1228)	(0.2572)	(0.2304)
Teacher incentives			-0.0673	0.0153
			(0.0652)	(0.0476)
Teacher control	0.0024	-0.0069	-0.2499*	-0.0617
	(0.0786)	(0.0614)	(0.1412)	(0.1070)
State control	-0.1255	-0.188***	-0.2795	0.0748
	(0.0682)	(0.0616)	(0.1734)	(0.1031)
Principal control	0.0535	0.0704	-0.0084	0.1313
	(0.0787)	(0.0642)	(0.1727)	(0.1215)
School climate	-0.0057	-0.013	-0.0939	-0.0285
	(0.0278)	(0.0243)	(0.0624)	(0.0478)

Estimates of coefficients on school quality from the SUR regressions: Government primary schools ONLY	Quality wrt similar schools		Quality wrt all schools	
Policies to improve low-performing students	-0.011	-0.065	-0.076	-0.069
	(0.0912)	(0.1032)	(0.0503)	(0.0895)
Lengthening instructional time	-0.068	-0.128	0.040	-0.120
	(0.0998)	(0.1036)	(0.0624)	(0.0933)
Reduced class size for subject	0.159	0.174	0.173	0.202
	(0.3397)	(0.3770)	(0.2416)	(0.3896)
Narrowing of curriculum	-0.230	-0.389**	0.013	-0.379**
	(0.2030)	(0.1937)	(0.1192)	(0.1547)
Policies to improve low-performing teachers	0.054	-0.007	-0.080	0.009
	(0.0948)	(0.0971)	(0.0528)	(0.0848)
Teacher resources	0.118	0.285	0.071	0.266
	(0.5413)	(0.6197)	(0.2898)	(0.6206)
Teacher control	0.135	-0.392	-0.024	-0.103
	(0.2682)	(0.3929)	(0.1881)	(0.2348)
State control	-0.345	-0.101	-0.163	0.184
	(0.2646)	(0.3218)	(0.2280)	(0.2101)
Principal control	-0.242	-0.270	-0.078	-0.271
	(0.2805)	(0.3771)	(0.1452)	(0.2083)
School climate	0.098	0.256*	0.287***	0.169
	(0.1450)	(0.1507)	(0.0732)	(0.1504)
covariates	N	Y	N	Y

Estimates of coefficients on school quality from the SUR regressions: All schools	Quality wrt similar schools	Quality wrt all schools
Policies to improve low-performing students	-0.043	-0.051
	(0.0662)	(0.0591)
Lengthening instructional time	0.105	0.078
	(0.1046)	(0.0964)
Reduced class size for subject	-0.139	-0.125
	(0.2619)	(0.2654)
Narrowing of curriculum	-0.108	-0.168
	(0.1509)	(0.1354)
Policies to improve low-performing teachers	-0.003	-0.012
	(0.0634)	(0.0572)
Teacher resources	0.194	0.160
	(0.3720)	(0.2946)
Teacher incentives	0.2313	0.1951
	(0.1768)	(0.1810)
Teacher control	-0.082	-0.041
	(0.1390)	(0.1227)
State control	0.054	0.089
	(0.1131)	(0.1001)
Principal control	-0.080	-0.107
	(0.1345)	(0.1180)
School climate	0.203**	0.125
	(0.0990)	(0.0900)
Covariates	Y	Y

Discussion

- Our evidence suggests that poorly-performing Australian govt primary schools narrow the curriculum, but some poorly-performing schools' climates get "worse" from 2009 to 2012.
- Further regressions with sector-specific interactions indicate the climate effect is confined to government and independent schools.
- Few changes overall from 2009 to 2012, but large differences in baseline levels of P&P across Australian schools
- *For comparison*: Rouse et al. (2013) find that relative to other schools, the worst-graded schools, in their words, "are more likely to focus on low-performing students, lengthen the amount of time devoted to instruction, adopt different ways to organize the day and learning environment of the students and teachers, increase resources available to teachers, and decrease principal control"
  - Particular policies at play in Florida: scheduling of classes and teacher preparation time, and increased financial control by the school district as opposed to the principal or other players.

What's next

- Look at some additional questions we included in our surveys that have no analog in Rouse et al. (2013)
- A bit of fiddling with subsample-specific analysis
- Finish the paper