Milestone 6

Molly Chiang 4/3/2020

Grapic

Below is a beautiful graphic which uses the data from Marshall 2015 and is based on Figure 3 in the paper. I used inspiration/guidance from (Gary King 2000) to create it.

Proportion of Cohorts Voting Conservative Before and After 1947 Modeled after Figure 3 from Marshall 2015

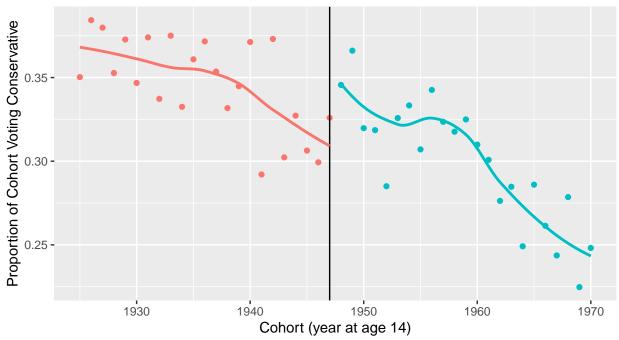


Figure 1: This figure illustrates the jump in proportion of a cohort, or class, of British individuals (classified by which they were 14), after the education reform in 1947 which increased the year at which you could legally school from 14 to 15.

Overview of Replication Paper

All analysis for this paper is available in my github repo.¹.

This paper by John Marshall uses data from the results of the 1947 high school leaving age reform in Great Britain, to analyze how additional years in high school affect political preferences (Marshall 2015). This paper looked specifically in how additional years in high school effected voting for the Conservative Party (Marshall 2015). In 1947, Great Britain changed the high school leaving age from 14 to 15, this induced almost half the student population to stay in school for at least 1 or 2 more years (Marshall 2015). Data from the 10 British elections between 1947 and 2010 was then used to compare voters young enough to have been effected by the reform to those who were too old to have been effected, using regression discontinuity (Marshall 2015). Regression discontinuity is usually used for determining if a program/treatment is effective, and essentially is a pretest-posttest program-comparison group design strategy (Trochim 2020). Regression

¹Link to Github Repo

discontinuity is unique in that individuals are assigned to one of two groups, just based on if they are on either side of a pre-determined cut-off (Trochim 2020). The results of data analysis revealed staying in high school for longer substantially increased likelihood to vote for the Conservative Party (staying one extra year increased probaility of voting Conservative by almost 12 percentage points) (Marshall 2015). This supports the previously studied fact that high school is extremely pertinent to political opinions later in life (Marshall 2015). In addition, this significant finding indicates the education reform of 1947 may have had an even greater affect on politics and election results nationwide over many years than ever expected (Marshall 2015).

Extension

I believe there are two routes I could take with a proposed extension of the work in "Education and Voting Conservative: Evidence from a Major Schooling Reform in Great Britain" by John Marshall.

First, I could look at other factors that may have been influenced by the 1947 leaving age reform. This would involve doing a similar regression discontinuity analysis of other variables around the year 1947, but in R instead of stata. For example, did this reform lead to an uptick in college educated people? Or in another vein, how did the distribution of voting in all parties, liberal, labour, and conservative change after the reform, do we see a corresponding downtick in the other parties to go along with the observed uptick in conservative voting? In addition, I think that I could, using some of the adivce from RAOS, find more exciting/visually pleasing ways to display the results of these regression analyses.

On the other hand, I could use this data to look at the effect of other education reforms that have happened over the course of British history. There was a big education reform in England in 1988 ("History of Education in England" 2020). This reform including introducing a national curriculum and corresponding national curriculum assessments, as well as modifying religious education ("History of Education in England" 2020). Although this reform did not result in increased years of school (so it would not be able to be compared to Marshall's main findings in that way), I think it could be interesting if this reform in content and uniformity of the British public education had any effect on voting preferences of the first classes to appreciate this reform (just as Marshall analyzed those who first appreciated the 1947 reform). Again, there could be issues in that this reform required more institutional changes than an easy age cut-off, and thus likely took longer to implement, but the results could still be interesting. In addition, the results could shed light on Marshall's finidings and if they are specifically correlated with more years in school, or are perhaps a response to government intervention via school reforms.

In addition to the 1988 education reform, the school leaving age was increased to 16 years of age in officially in 1996 ("History of Education in England" 2020). This may pose an issue because Marshall's data only goes to 2010 (so we would have much more data on pre-reform individuals than post-reform individuals), but would be very, very interesting addition to the findings of Marshall (2015), as it could really support the idea that additional years of school being associated with more conservative voting, if there was a similar uptick around this year. Additionally, if it did not show an uptick, it could show that perhaps Marshall discovered a unique effect of the 1947 British education reform, but not of more years of education in general.

Replication Process

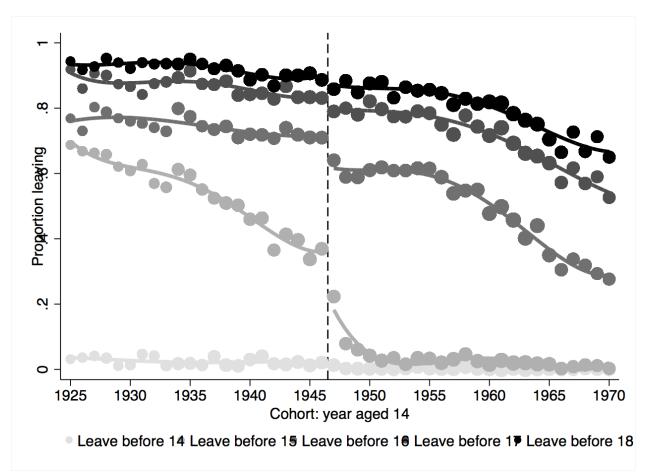
I was able to replicate nearly all of the figures in the paper in stata. The code I ran and the output figures are included in the appendix. In terms of the tables, I was able to run all of the code in stata almost directly from Marshall's replication code (some aspects I had to change because some of the functions he used have been updated since he wrote the paper). However, I had a lot of trouble going from the raw results of the models Marshall ran to the polished tables I saw in his paper, as only code for the models were included in his .do file. Thus, I have included screenshots of the raw code and the relevant aspects of the output that would be included in the table in the paper. I hope to learn more basics of stata in order to compile and nicely display these raw outputs from Marshall's code.

For brevity, I did only included the screenshots from tables 1 and 2 and not table 3 (as it would be a lot of seperate screenshots to include), but I have included the table 3 code, which was run in the same way. Table 4 was not able to be replicated as there was no code for table 4 included in Marshall's replication code.

Most coefficients exactly or almost exactly matched Marshall's published results. A few were a few decimal places off (potentially due to different iterations of the regression), and a even fewer were quite different, which still needs to be investigated.

This paper relies heavily on rdrobust models, and the reference material for learning about this function and its package is from (Sebastian Calonico 2020).

Appendix



```
*** Figure 1: Trends in school leaving age
twoway (lpoly leave_18 yearat14 if yearat14<-1947 & yearat14>=1925, lcolor(gs14) clwidth(thick) degree(4)) ///
(lpoly leave_18 yearat14 if yearat14>=1947 & yearat14<=1970, lcolor(gs14) clwidth(thick) degree(4)) ///
(scatter meanleave_18 yearat14 if yearat14>=1925 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs14)) ///
(lpoly leave_19 yearat14 if yearat14<-1947 & yearat14>=1925, lcolor(gs11) clwidth(thick) degree(4)) ///
(lpoly leave_19 yearat14 if yearat14>=1947 & yearat14<=1970, lcolor(gs11) clwidth(thick) degree(4)) ///
(scatter meanleave_19 yearat14 if yearat14>=1925 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs11)) ///
(lpoly leave_110 yearat14 if yearat14>=1947 & yearat14>=1925, lcolor(gs7) clwidth(thick) degree(4)) ///
(scatter meanleave_110 yearat14 if yearat14>=1947 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs7)) ///
(lpoly leave_111 yearat14 if yearat14>=1925 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs7)) ///
(lpoly leave_111 yearat14 if yearat14>=1947 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs7)) ///
(lpoly leave_111 yearat14 if yearat14>=1947 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs7)) ///
(scatter meanleave_111 yearat14 if yearat14>=1945 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs5)) ///
(lpoly leave_112 yearat14 if yearat14>=1945 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs5)) ///
(lpoly leave_112 yearat14 if yearat14>=1947 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gs5)) ///
(scatter meanleave_112 yearat14 if yearat14>=1947 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(black)), ///
(scatter meanleave_112 yearat14 if yearat14>=1945 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(black)), ///
(scatter meanleave_112 yearat14 if yearat14>=1945 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(black)), ///
(scatter meanleave_12 yearat14 if yearat14>=1945 & yearat14<=1940 [weight=weight_14], msize(s
```

Figure 1: Marshall Figure 1 code

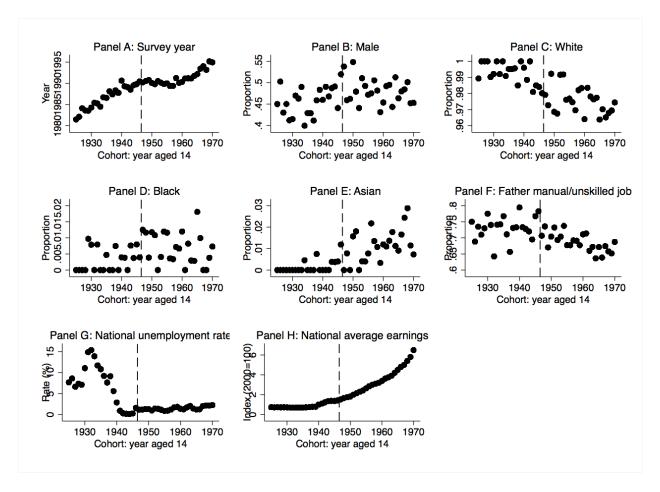


Figure 2: Marshall Figure 2

```
capture by capture and the search of the sea
```

Figure 3: Marshall Figure 2 code

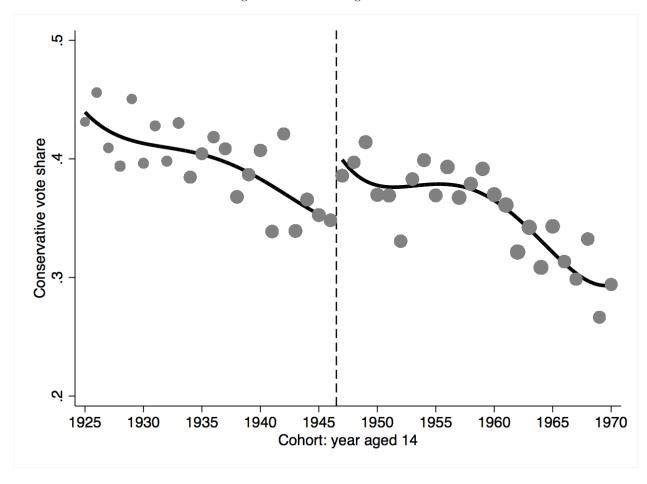


Figure 4: Marshall Figure 3

```
*** Figure 3: Reduced form
twoway (lpoly con yearat14 if yearat14>=1925 & yearat14<1947, lcolor(black) clwidth(thick) degree(4)) ///
(lpoly con yearat14 if yearat14>=1947 & yearat14<=1970, lcolor(black) clwidth(thick) degree(4)) ///
(scatter meancon14 yearat14 if yearat14>=1925 & yearat14<=1970 [weight=weight_14], msize(small) mcolor(gray)), ///
graphregion(fcolor(white)) clolor(white)) ylab(,nogrid) ytitle(Conservative vote share) xtitle(Cohort: year aged 14) xline(1946.5, lcolor(black) lpattern(dash)) ///
yscale(range(.2 .5)) ylabel(.2[0.1]0.5) xlab(1925[5]1970) legend(off)
```

Figure 5: Marshall Figure 3 code

Table 1. Estimates of Schooling's Effect on Voting Conservative

			<u> </u>					
	Years of Schooling LLR (1)	Attend University LLR (2)	Vote Conservative LLR (3)	Vote Conservative LLR IV (4)	Vote Conservative OLS (5)	Vote Conservative OLS (6)	Vote Labour LLR IV (7)	Vote Liberal LLR IV (8)
Post-1947 reform	.381*** (.076)	.009 (.013)	.044** (.020)					
Years of schooling	, ,	, ,	, ,	.116** (.056)	.021*** (.002)		071 (.052)	021 (.043)
8th year of schooling				, ,	, ,	020 (.036)	, ,	, ,
9th year of schooling						Baseline		
10th year of schooling						.126*** (.013)		
11th year of schooling						.213***		
12th year of schooling						(.014) .289*** (.017)		
13th year of schooling						.306***		
14th year of schooling						.281***		
Outcome range	0-40	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
Outcome mean Outcome standard devi-	10.6	.11	.38	.38	.36	.36	.36	.19
ation	1.86	.32	.49	.49	.48	.48	.48	.39
First stage F-statistic	25.4	.5		25.4			25.4	25.4
Observations	11,068	11,068	11,068	11,068	16,757	16,757	11,068	11,068

Figure 6: Original Marshall Table 1

Column 1	Coef.	Std. Err.	Variable	0bs	Mean	Std. Dev.	Min	Max
Coln	. 4112	.09145	leave	11,068	10.58312	1.858667	0	4
mn 2	Coef.	Std. Err.	Variable	Obs	Mean	Std. Dev.	Min	Ma
Column	.03414	.01798	uni	11,068	.1145645	.31851	0	,
ın 3	Coef.	Std. Err.	Variable	0bs	Mean	Std. Dev.	Min	Max
Column 4 Column 3	.04548	.04548 .01829		11,068	10.58312	1.858667	0	40
	Coef.	Std. Err.	Variable	0bs	Mean	Std. Dev.	Min	Max
Colur	.11581	. 05539	leave	11,068	10.58312	1.858667	0	40
Column 5 	Coef.	Robust Std. Err.	Variable	0bs	Mean	Std. Dev.	Min	Max
<u></u> 5–	.021135	.0018537	con	16,757	.3587158	. 4796379	0	1
Column 6	10 .1 11 .2	200463 .0364474 257948 .0134744 126346 .0139522	Variable 	0bs	Mean	Std. Dev.	Min	Max
	13 . 3	.3061406 .0176145		16,757	.3587158	. 4796379	0	1
		Std. Err.	Variable	0bs	Mean	Std. Dev.	Min	Max
nn 7	Coef.	Stu. Ell.						
Column 7	Coef. 06892	.05082	lib	11,068	.1857608	.3889311	0	1
Column 8 Column 7			lib Variable	11,068 Obs		. 3889311 d. Dev.		1 ax

```
*** Table 1: Main estimates
rdrobust leave yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
sum leave if yearat14>=1933 & yearat14<=1961
rdrobust uni yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
sum uni if yearat14>=1933 & yearat14<=1961
rdrobust con yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust con yearat14, c(1947) fuzzy(leave) p(1) q(2) kernel(tri) bwselect(mserd)
sum leave if yearat14>=1933 & yearat14<=1961
areg con leave male white black asian sagesq-sagequart syearat14 syearat14sq syearat14cub syearat14quart, ro a(survey)
areg con ib9.leave male white black asian sagesq-sagequart syearat14 syearat14sq syearat14cub syearat14quart, ro a(survey)
summ con if e(sample)
rdrobust lab yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(mserd) fuzzy(leave)
sum lib if yearat14>=1933 & yearat14<=1961
rdrobust lib yearat14>=1933 & yearat14<=1961
```

Figure 7: Marshall Table 1 Code

Table 2. Schooling, Conservative Voting, and Income-Based Mechanisms

	Nonmanual Worker (below 60)	Vote Conservative (below 60)	Vote Conservative (60 or Above)	Conservative Partisan	Decided before Campaign
	(1)	(2)	(3)	(4)	(5)
A. Reduced form estimates:					
Post-1947 reform	.075**	.049**	.006	.039*	.041**
	(.030)	(.024)	(.029)	(.022)	(.019)
B. IV estimates:					
Years of schooling	.144***	.111**	.030	.092*	.103*
_	(.055)	(.056)	(.153)	(.052)	(.053)
Outcome range	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
Outcome mean	.46	.38	.35	.37	.74
Outcome standard deviation	.50	.48	.48	.48	.44
Bandwidth	13.9	22.6	15.0	13.9	12.5
First stage F-statistic	28.1	28.3	2.7	26.4	23.4
Observations	6,086	10,152	4,589	9,711	9,510

Figure 8: Original Marshall Table 2

n 1	Coef.	Std. Err.	Coef.	Std. Err.	0bs	Mean	Std. Dev.	Min	Max
Column 1	.10474	.0502	. 15754	.06094	6,086	. 4574433	.4982266	0	1
Column 2	Coef.	Std. Err.	Coef.	Std. Err.	0bs	Mean	Std. Dev.	Min	Max
ပိ	.02718	.05559	.06984	.06284	10,152	.375591	. 484299	0	1
Column 3	Coef.	Std. Err.	Coef.	Std. Err.	0bs	Mean	Std. Dev.	Min	Max
Colur	.02756	. 04706	.81987	4.6518	4,589	.3497494	. 4769427	0	1
Column 4	Coef.	Std. Err.	Coef.	Std. Err.	0bs	Mean	Std. Dev.	Min	Max
Col	.03894	. 0224	.08817	. 05303	9,711	.369169	. 4826046	0	1
Column 5	Coef.	Std. Err.	Coef.	Std. Err.	Obs	Mean	Std. Dev.	Min	Max
Col	.04599	.0212	.1049	.05416	9,510	.7413249	. 4379297	0	1

Figure 9: Replicated Table 2

```
*** Table 2: Raising social class, Heterogeneity by age (above 60), Become a Conservative partisan, and Decide before the electoral campaign
rdrobust nonmanual yearat14 if age<60, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust nonmanual yearat14 if age<60, fuzzy(leave) c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
sum nonmanual if age<60 & yearat14>=1934 & yearat14<=1960

rdrobust con yearat14 if age<60, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust con yearat14 if age<60, c(1947) p(1) q(2) kernel(tri) bwselect(mserd) fuzzy(leave)

rdrobust con yearat14 if age>=60, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust con yearat14 if age>=60, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust con yearat14 if age>=60, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust conpart yearat14>=1932 & yearat14<=1962

rdrobust conpart yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust conpart yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(mserd)
rdrobust perm yearat14>=1935 & yearat14<=1959
```

Figure 10: Marshall Table 2 Code

```
*** Table 3 Panel A: Economic policy preferences
rdrobust taxspendself yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
rdrobust welfaretoofar yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) rdrobust redist yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
rdrobust gender_not_too_much yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
sum gender_not_too_much if yearat14>=1931 & yearat14<=1963</pre>
rdrobust econ_values yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
sum econ_values if yearat14>=1930 & yearat14<=1964
rdrobust taxspendself yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
rdrobust welfaretoofar yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
rdrobust redist yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
rdrobust gender_not_too_much yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
rdrobust econ_values yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
*** Table 3 Panel B: Non-economic policy preferences
rdrobust crime_rights_scale yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) rdrobust leave_europe yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
rdrobust end_priv_edu yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
rdrobust abortion_too_far yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
sum abortion_too_far if yearat14>=1933 & yearat14<=1961 rdrobust raceequ_too_far yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK)
sum raceegu too far if yearat14>=1927 & yearat14<=1967
rdrobust crime_rights_scale yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave) rdrobust leave_europe yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
rdrobust end_priv_edu yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
rdrobust abortion_too_far yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave) rdrobust raceequ_too_far yearat14, c(1947) p(1) q(2) kernel(tri) bwselect(IK) fuzzy(leave)
```

Figure 11: Marshall Table 3 Code

References

Reference material to create this bibliography comes from (J Allaire 2016)

Gary King, Jason Whittenberg, Michael Tomz. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." Midwest Political Science Associatio. https://www-jstor-org.ezp-prod1. hul.harvard.edu/stable/pdf/2669316.pdf.

"History of Education in England." 2020. https://en.wikipedia.org/wiki/History_of_education_in_England.

J Allaire, Jonathan McPherson, Yihui Xie. 2016. "Bibliographies and Citations." https://rmarkdown.rstudio.com/authoring bibliographies and citations.html.

Marshall, John. 2015. "Education and Voting Conservative: Evidence from a Major Schooling Reform in Great Britain." Southern Political Science Association. https://www-journals-uchicago-edu.ezp-prod1.hul. harvard.edu/doi/pdfplus/10.1086/683848.

Sebastian Calonico, Max H. Farrell, Matias D. Cattaneo. 2020. "Package 'Rdrobust'." https://cran.r-project.org/web/packages/rdrobust/rdrobust.pdf.

Trochim, Prof William M.K. 2020. "The Regression-Discontinuity Design." https://socialresearchmethods.net/kb/regression-discontinuity-design/.