

Gender, Education, and Voting Conservative

Molly Chiang

5/8/2020

Abstract

Marshall (2015) shows the causal effect of additional years of schooling on voting conservative in his analysis of voting records before and after the British 1947 school-leaving age reform. I successfully replicated Marshall's code, except an update in the `rdrobust` package led to slightly different coefficients. In an extension of Marshall's work I investigated treatment effect heterogeneity between genders. Running `rdrobust` and creating regression discontinuity figures on male and female subsets of the data revealed Marshall's effect of more years of education increasing the likelihood of voting conservative was stronger in women than men. This finding could complicate Marshall's argument that more education leads to higher income then to more conservative political opinions and perhaps reveals something about the differing effect of education on men and women.

1 Introduction

Marshall (2015) 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 affected 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 affected by the reform to those who were too old to have been affected, 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 discontinuity is unique in that individuals are assigned to one of two groups, just based on if they are on either side of a predetermined 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 probability of voting Conservative by almost 12 percentage points), controlling for continuity of all other covariates (Marshall 2015). Marshall does further analysis to support a hypothesis where increased education leads to increased income which leads to voting more conservatively (Marshall 2015). This supports the previously studied fact that high school is extremely pertinent to political opinions later in life, and that more education generally leads to higher income and thus voting more conservative (Devereux and Hart 2010), (Meltzer and Richard 1981). 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).

The first aspect of this project was a replication of Marshall's results. His code and data are publically available on the Harvard Dataverse. In order to replication Marshall's results, I ran his original code—with a few modifications—in `stata`. Marshall's figures were able to be replicated by running `stata` code, but replicating the tables was met with some difficulty. First, the `rdrobust` package has been updated since 2015 and thus some of the arguments Marshall used have since been deprecated (Sebastian Calonico 2017). I was able to replace old arguments with their 'updated' versions, according to an update from the package authors in 2017 (Sebastian Calonico 2017). Changing this argument allowed the code to be run, but changed the values of the calculated coefficients by a bit. In addition, the code to go from raw output to polished table was not included in the replication code on the Dataverse, so I just analyzed the raw output I was able to replicate. I also replicated Marshall's figure 3 in R (R Core Team 2019). All code for the replication is available in my github repo¹.

¹Link to Github Repo

After replication, my extension of Marshall’s findings look at the effect of his findings when broken down by gender. In order to perform this treatment effect heterogeneity analysis, I subsetting Marshall’s data into a solely male and solely female. With this breakdown of the data, I created regression discontinuity figures and ran `rdrobust()` just as Marshall had with the entire data set.

The results of the figures and table with `rdrobust()` outputs revealed the effect of more years of education increasing the probability of voting conservative was much stronger in women than men. Overall, students affected by the reform are 4.5 percentage points more likely to vote conservative. However, males of this group are 3.9 percentage points and females 5.3 percentage points more likely to vote conservative. These results do not fit perfectly with Marshall’s argument that more education leads to higher income and then to more conservative political opinions, as his analysis is focused in the 1940’s, when women (even educated) had little space for economic advancement. This extension thus evidently creates many more questions for investigation on the intersection of gender, education, and political views.

Over the course of this paper, I will contextualize Marshall’s 2015 paper in a literature review, and dive more deeply into explaining my replication process and extension process.

2 Literature Review

This paper is written as a response to mixed literature on education and voting preferences. On one hand, it has been suggested that education leads to more socially liberal attitudes. In 1959, Lipset proposed more education is associated with more liberal attitudes and support of democracy (Lipset 1959). Hyman and Wright support that analysis saying education leads to students thinking with a fundamentally liberal fashion (Hyman and Wright 1979).

On the other hand, Devereux and Hart showed in 2010—using similar data and the same regression discontinuity analysis design—that additional years of schooling as a result of the 1947 reform increased wage by 5-15% (Devereux and Hart 2010). In addition Meltzer and Richard (1981) have shown that higher wage earners prefer low income tax and government spending, policies aligned more closely with the British Conservative party (Meltzer and Richard 1981).

However, Marshall’s paper was the first to directly address the causal effects of voting conservative and more years of education. This paper also used these past findings about the effect of education to help explain Marshall’s findings.

3 Replication Process

I was able to replicate all of the figures in stata (see appendix for this process). In addition, I replicated Marshall’s figure 3 (his main finding) in R. This paper relies heavily on `rdrobust` models, and the reference material for learning about this function and its package is from (Sebastian Calonico 2020).

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. Namely, when using `rdrobust`, the `bwselect` option `IK` and the `h()` argument are no longer functional, and have been replaced with an updated version `bwselect(mserd)`, which I used in all the code. This modification was necessary because of the changes that have been made to the `rdrobust` package since 2015, and did change my values a bit from Marshall’s.

Although I was able to replicate the values of Marshall’s tables by running his stata file (with modifications), I was unable to go 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. The summary statistics for all of the tables (number of observations, outcome mean, etc) were replicated perfectly. However, switching the bandwidth selection did result in changes to the results of the `rdrobust` functions. Some values differed more than others from

Marshall’s values after changing to `bwselect(mserd)`, but it seems when fuzzy regression discontinuity was run, the values differed less.

Finally, I focused on replicating Marshall’s main figures and tables, and did not replicate all of his thorough robustness and continuity checks. For example, Marshall ran the McCrary test for continuity, ran his `rdrobust` commands with different order polynomials, and did placebo cutoff checks, all of which I accepted to be successful and support his data.

4 Extension

Marshall (2015) demonstrated, using the 1947 school-leaving age reform in Great Britain and regression discontinuity design, that as years of schooling increased, so did one’s likelihood to vote conservative. His explanation for this finding was that increased schooling leads to higher income and then more conservative voting based on economic policies. In my extension of his work, I examined treatment effect heterogeneity, specifically how Marshall’s results differed when subsetting his data for males and females.

Creating regression discontinuity figures on just the male and female subsets of the data revealed the effect of more years of education increasing the probability of voting conservative was much stronger in women than men (Figure 1). These figures were created with fourth-degree polynomial fits as was used in Marshall (2015). In addition, when running `rdrobust`, the coefficient for the female subset was considerably higher than the coefficient of the male subset (with the both gender coefficients being intermediate).

In addition to creating regression discontinuity figures with fourth-degree polynomials, I also created the figures with linear fits (Figure 2). This choice was guided by Gelman and Zelizer (2015), where it was argued 3rd and 4th degree polynomials, although commonly used for regression discontinuity, can lead to overfitting and the validating of causal inferences that are arbitrary and substantively implausible (Gelman and Zelizer 2015). However, in this case, even when switching to linear fits, the discontinuity around 1947 is clear, both overall (see appendix), for females and even males where the polynomial fit indicated little discontinuity.

Gender Breakdown of Schooling’s Effect on Voting Conservative

Results of regression discontinuity analysis of a conservative vote around the 1947 school-leaving age reform in Great Britain

	All-Gender	Male	Female
Estimated Effect of 1947 Reform	0.045	0.039	0.053
Standard Error	0.019	0.033	0.031
p-value	0.018	0.228	0.089
Number of Observations	29396	13745	15661

The RDD I ran for the extension was based on what Marshall (2015) ran. I set my dependent variable to `con` (0 or 1 for if the individual voted conservative) and the independent variable to `yearat14` (yearat14 of 1947 is the first class of individuals who were forced to stay in school an extra year). I set the cutoff to the year of the reform, 1947, and like Marshall, set the order of the point-estimator local-polynomial to 1, the order of the bias-correction local-polynomial to 2, and the kernel function to triangular. Finally, differing from Marshall, I set `bwselect = "mserd,"` as that is the updated version of the now-deprecated `bwselect(IK)`. The table with the results of these regression contains the standard-form coefficients and their associated standard error and p-values.

The coefficient when running RDD is an indicator of causality of the forcing variable (in our case, year when the individual was 14 - if you were 14 in 1947 you would have been affected by this program) on the observed variable (voting conservative) around the cutoff point (1947). Thus demonstrating years in school causes more conservative voting patterns in females more than it does in men. Thus, overall, students affected by the reform are 4.5 percentage points more likely to vote conservative.

Differing Conservative Vote Share in Men and Women Before and After 1947

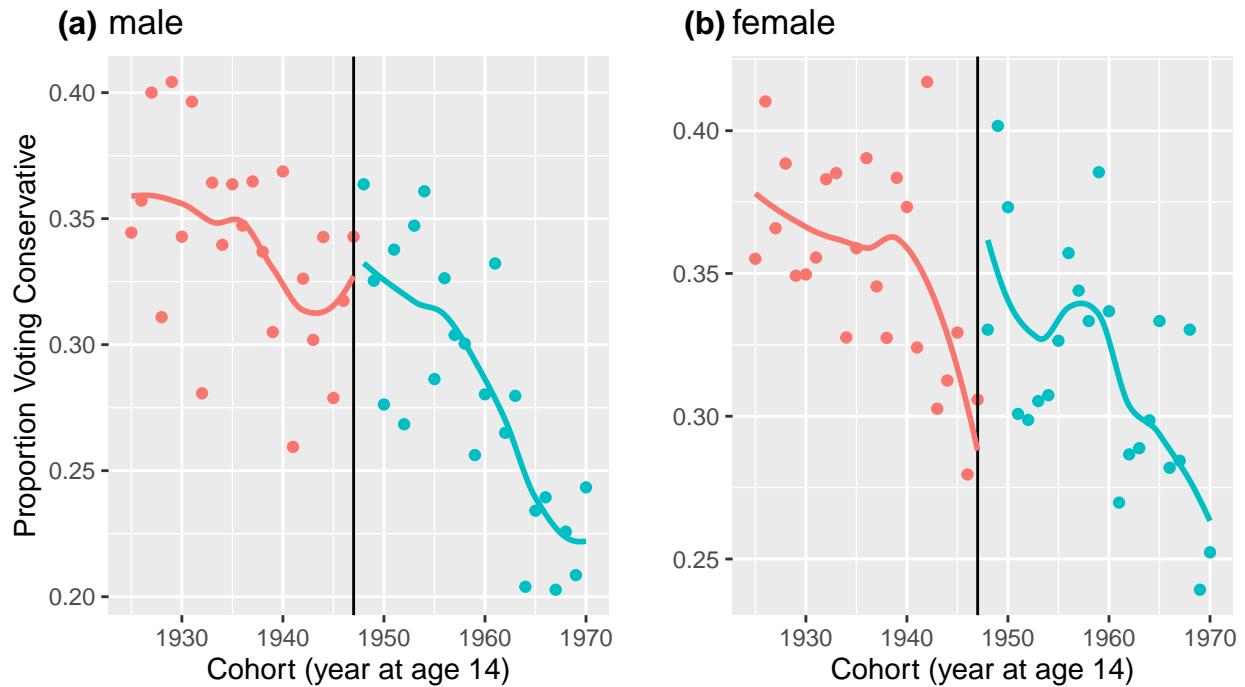


Figure 1: The change in proportion of individuals in each class voting conservative around the implementation of the British school-leaving age reform in 1947, subset by gender. Curves represent fourth order polynomial fits on either side of the 1947 discontinuity in (a), as was done in Marshall (2015).

Linear Fit of Conservative Vote Share Around 1947

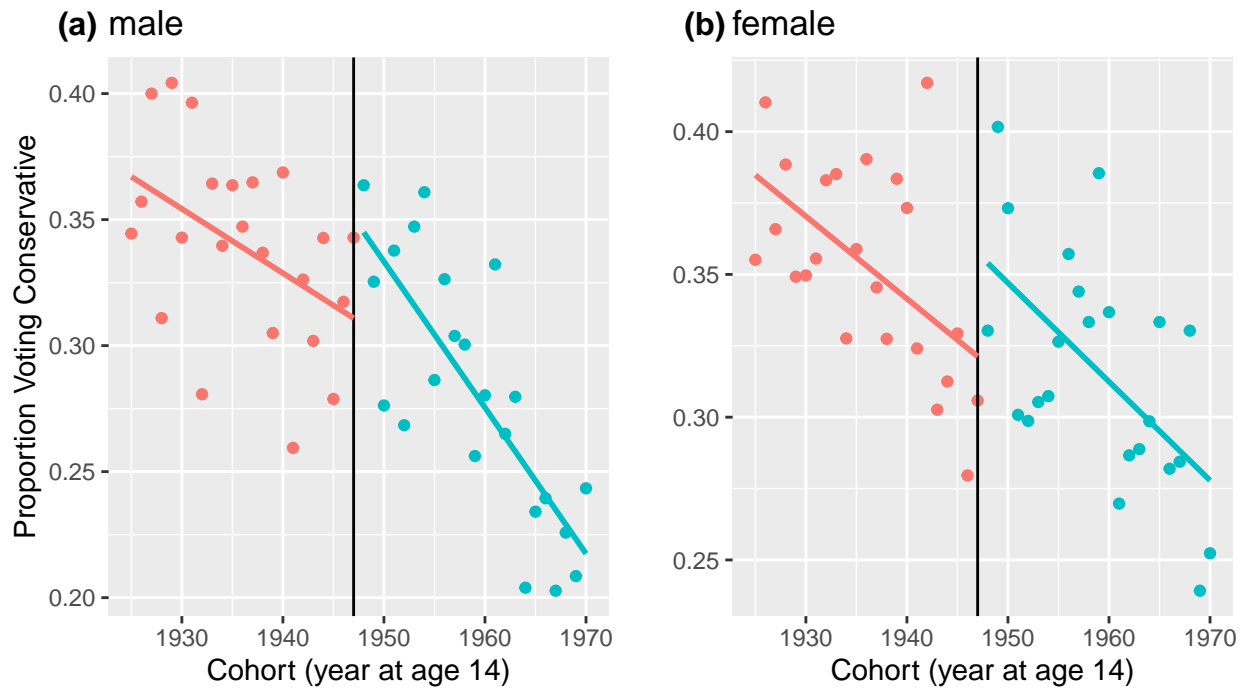


Figure 2: The change in proportion of individuals in each class voting conservative around the implementation of the British school-leaving age reform in 1947, subset by gender. Curves represent linear fits on either side of the 1947 discontinuity as recommended by Gelman (2015).

However, males of this group are 3.9 percentage points and females 5.3 percentage points more likely to vote conservative. Although neither of the male or female coefficients had a p-value less than .05 (traditionally indicating statistical significance), accepting uncertainty in our model and analyzing the magnitude of these p-values, we see the female coefficient p-value is an order of magnitude smaller, indicating the effect on females is more significant than on males.

The results of this extension complicate Marshall’s argument that more education leads to higher income and then to more conservative political opinions. Women did not have a lot of space for income advancement in the 1940s and 50s in comparison to men, and thus Marshall’s explanation would likely assume women would thus have a smaller increase in voting conservative in response to more education, but this extension shows the opposite.

In addition, this extension perhaps reveals something about the differing effect of education on men and women and creates questions and areas of more study on gender, education, and political views.

5 Conclusion

Marshall (2015) first proved that education attainment significantly increased after the 1947 school-leaving age reform in Great Britain. Then, using this reform and regression discontinuity design, demonstrated as years of schooling increased, so did one’s likelihood to vote conservative, ensuring other covariates maintained continuity around this cutoff. He postulated these findings were explained by a hypothesis that more schooling leads to higher income and then more conservative voting based on economic policies. This hypothesis is supported by previous research that schooling leads to increased income (Devereux and Hart 2010) and that higher wage earners align more closely with the British Conservative party (Meltzer and Richard 1981).

I was able to replicate all of Marshall’s figures directly in stata, and I replicated Marshall’s figure 3 (his main finding) in R as well. For his tables, an update in the rdrobust package led to the deprecation of bwselect(IK) and h() arguments, which I replaced with the updated version bwselect(mserd), but the update changed the values of the results a bit. Other than this change, I was able to run and replicate all of the code and results for Marshall’s tables in stata (rdrobust calls and summary statistics). However, I was not able to go from results of regression calls to his polished tables as that code was not included in his .do file.

In my extension Marshall (2015), I examined treatment effect heterogeneity of gender via regression discontinuity figures and rdrobust() calls. The results revealed the effect of more years of education increasing the probability of voting conservative was stronger in women (5.3 percentage points) than men (3.9 percentage points). In addition, creating regression discontinuity figures with linear fits, as recommended by Gelman and Zelizer (2015), and on either side of the cutoff still maintained a clear jump in proportion of a cohort voting conservatively in both men and women.

These results are perhaps in conflict with Marshall’s argument that more education leads to higher income and then to more conservative political opinions, as his analysis is focused in the 1940’s, when women (even educated) had little space for economic advancement in comparison to men. Marshall’s explanation would likely assume women would have a smaller increase in voting conservative in response to more education since they had less economic opportunity, but this extension shows the opposite.

In addition, this extension perhaps reveals something about the differing effect of education on men and women. Are women’s political views more responsive to education? Women are generally more liberal than men, how does this result fit into this narrative? This extension certainly creates questions and opens new areas of study on the relationship between gender, education, and political views.

6 Appendix

I was able to replicate all of Marshall’s figures and table code directly in stata. That stata code is available at this project’s repo. Here I have a screenshot of Marshall’s key figure (Figure 3, from page 389) that

I replicated in stata and a version in R that I created. In addition, I performed the linearity test on my replicated figure, as suggested by Gelman and Zelizer (2015).

Marshall (2015) Figure 3 Replication and Linear Test

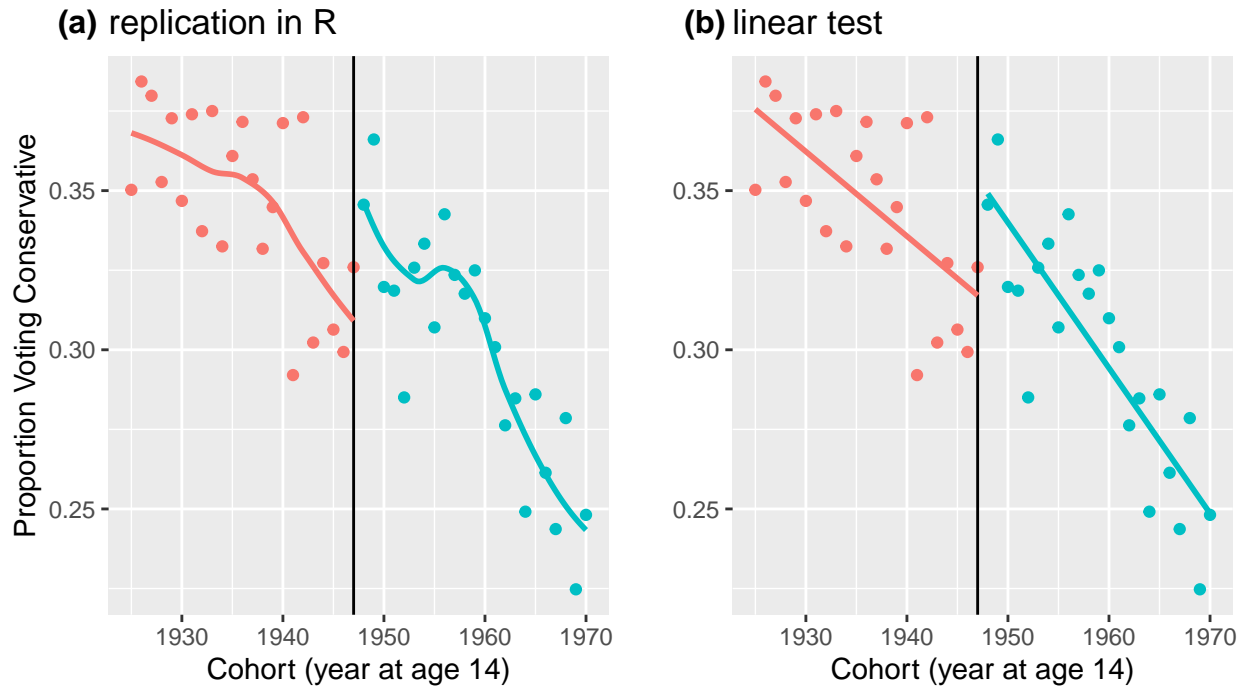


Figure 3: This figure is a replication of Figure 3 (Marshall 2015) in R instead of Stata, it illustrates the jump in proportion of a cohort, or class, of British individuals (classified by the year at which they were 14), after the education reform in 1947 which increased the year at which you could legally leave high school from 14 to 15.

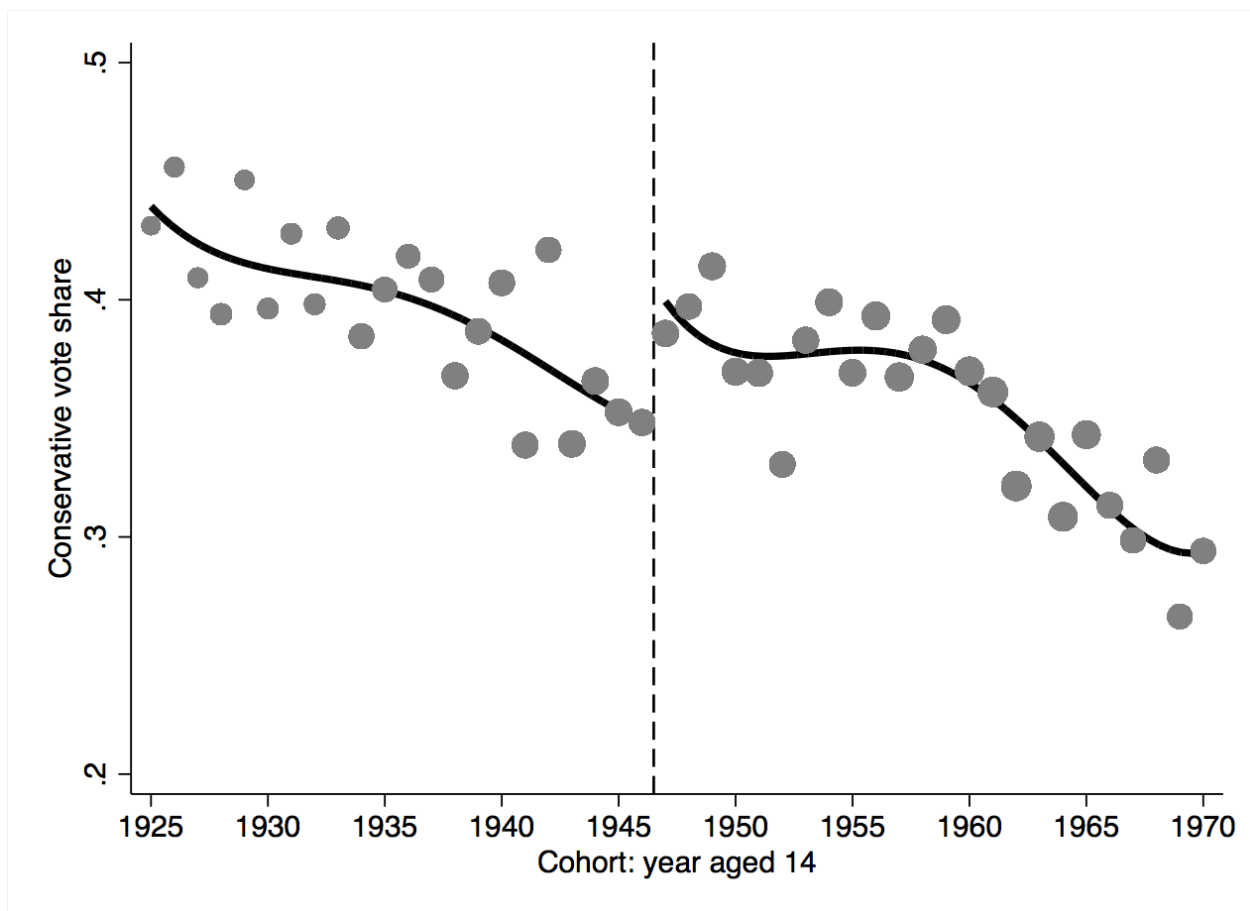


Figure 1: Marshall Years of Schooling and Voting Conservative (Figure 3) Replicated in Stata

7 References

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

Devereux, Paul, and Robert Hart. 2010. “Forced to Be Rich? Returns to Compulsory Schooling in Britain.” *Economic Journal*.

Gelman, Andrew, and Adam Zelizer. 2015. “Evidence on the Deleterious Impact of Sustained Use of Polynomial Regression on Causal Inference.” *Research; Politics*.

Hyman, Herbert, and Charles Wright. 1979. “Education’s Lasting Influence on Values.” University of Chicago Press.

J Allaire, Jonathan McPherson, Yihui Xie. 2016. “Bibliographies and Citations.” https://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html.

Lipset, Seymour Martin. 1959. “Some Social Requisites of Democracy: Economic Development and Political Legitimacy.” *American Political Science Review*.

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>.

Meltzer, Allan, and Scott Richard. 1981. “A Rational Theory of the Size of Government.” *Journal of Political Economy*.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Sebastian Calonico, Max H. Farrell, Matias D. Cattaneo. 2017. “Rdrobust: Software for Regression-Discontinuity Designs.” *Midwest Political Science Association*.

———. 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/>.