Impact of Compulsory Schooling on Socioeconomic Outcomes

An Analysis Using Regression Discontinuity Design

Jasmine Jia
May 10, 2024

Introduction




Purpose:

- Investigate the transformative impact of compulsory education on various facets of life including educational attainment, earnings, marital status, and parental status by the age of 45.

Brief Overview:

- **Policy Implementation:**
  - **Date Enacted**: September 1, 1951.
  - **Requirement**: School attendance is mandatory until age 16 for those born on or after this date.
  - **Goal**: Establish a baseline education for all individuals.

Introduction (continued)

- **Policy significance:**
  - Justifying and optimizing public expenditures on education.
  - Necessitating adjustments in social policies including housing, welfare, and healthcare systems.
  - Narrowing the gap of inequality by ensuring every child receives a fundamental level of education.
- **Endogeneity concerns:**
  - Self-selection bias
  - The risk of reverse causality
  - Omitted variable bias

Research Design

Fuzzy vs. Sharp RDD:

- Fuzzy RDD accounts for imperfect compliance with the policy.
- Unlike sharp RDD, the assignment to treatment at the cutoff is not strictly deterministic, allowing for variability.

Instrumental Variable (IV) Approach & Two-stage Least Squares (2SLS Estimation):

- **Definition and Role**:
  - **Z (Instrument)**: The cutoff date (September 1, 1951) acts as an instrument.
  - **X (Treatment Variable)**: Indicates whether a student left school at or after age 16.
  - **Y (Outcome Variables)**: Focus on log earnings, marital status, parental status at age 45, and years of schooling attained.
- **IV Conditions**:
  - **Inclusion Condition**: Strong predictor of the treatment, ensuring that the instrument affects the treated group.
  - **Exclusion Condition**: Ensures the instrument affects the outcome only through the treatment.

Research Design (continued)

Control vs. Treatment Groups:

- **Segmentation Based on Cutoff**:
  - Control Group: Born before September 1, 1951.
  - Treatment Group: Born on or after September 1, 1951.
  - Assumes comparability for those near the cutoff, differing only in policy impact.

Construction of the Running Variable:

- **Definition**: Measures days between each individual's birthdate and the policy implementation date.
- **Binning for Analysis**:
  - It is segmented into 90-day intervals around the cutoff.
  - Enhances analysis precision and handles data granularity.

First Stage

```
t test of coefficients:

                                Estimate   Std. Error t value            Pr(>|t|)
(Intercept)                  0.869285271  0.001840167  472.39 <0.0000000000000002 ***
runvar                       0.000000524  0.000002331    0.22                0.82
subject_to_policy            0.130714729  0.001840167   71.03 <0.0000000000000002 ***
I(subject_to_policy * runvar) -0.000000524  0.000002331   -0.22                0.82
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

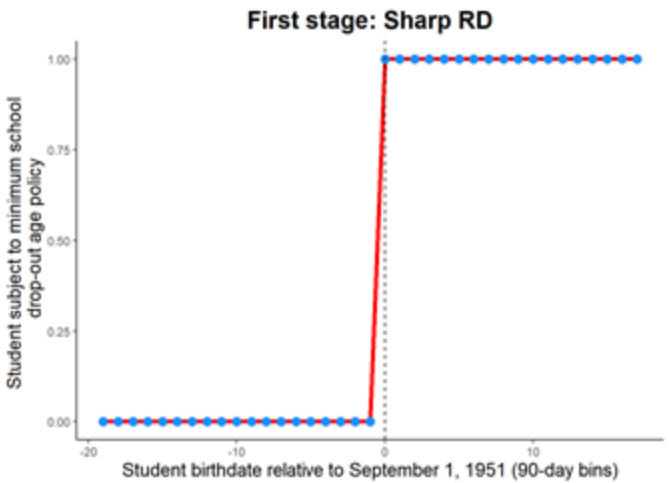
First Stage Analysis:

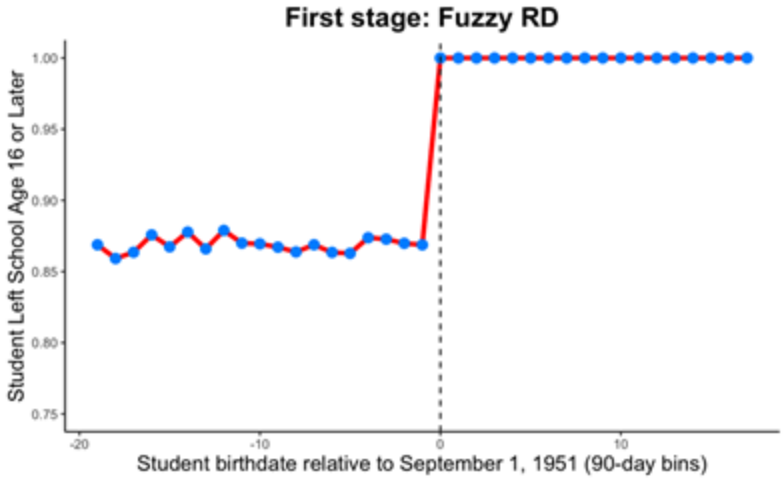
- **OLS Specification:**

$$(1)\ LeftSchAge16OrLater_i = \alpha_0 + \alpha_1 runvar_i + \beta_1 SubjectToPolicy_i + \alpha_2 (runvar_i * SubjectToPolicy_i) + \epsilon_i$$

- **Regression Results**:
  - **Findings**: Students that were subject to the policy were 13.1% more likely to leave school at age 16 or later ($p < 0.001$).
  - **Statistical Evidence**: T-statistic on the policy indicator is 71.03, indicating a robust first-stage relationship.

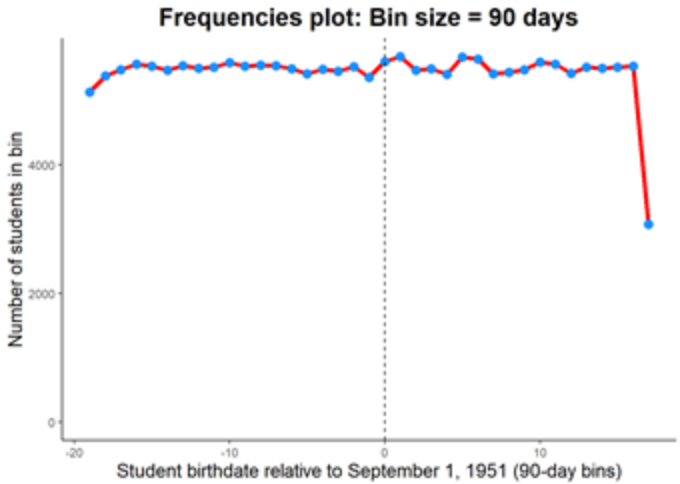
First Stage Visual Assessments





First Stage: Assumptions

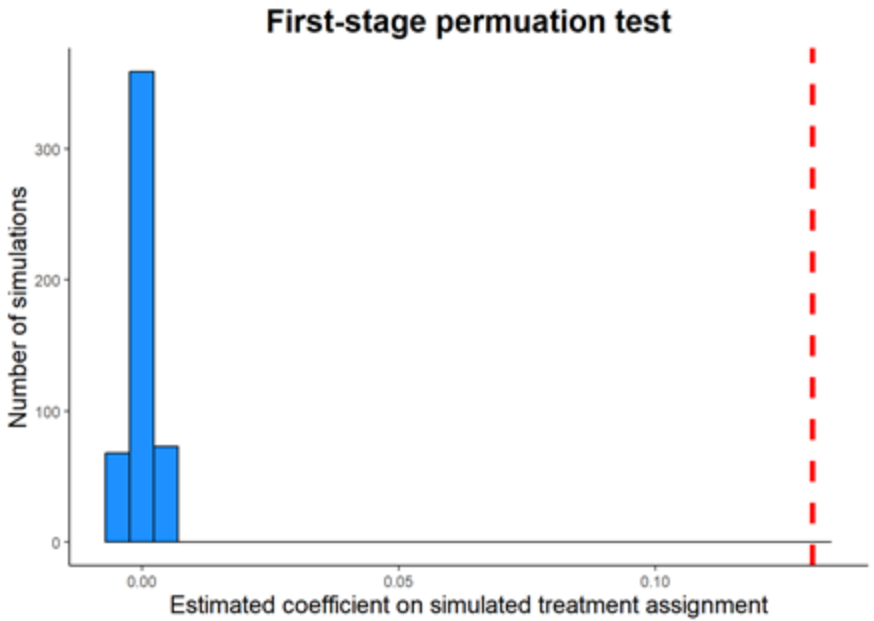
(1) No bunching



(2) No other discontinuous changes at the cutoff

No covariates in our analysis

First Stage: Permutation Test



Years of Schooling

```
Coefficients:
                                Estimate   Std. Error t value            Pr(>|t|)
(Intercept)                 13.068417584  0.014442631  904.85 <0.0000000000000002 ***
runvar                       0.000000573  0.000014674    0.04                0.97
subject_to_policy            0.190122089  0.020720728    9.18 <0.0000000000000002 ***
I(subject_to_policy * runvar) 0.000029191  0.000021939    1.33                0.18
```

First Stage Analysis:

- $$(2)\ Years\ Of\ Schooling_i = \alpha_0 + \alpha_1 runvar_i + \beta_1 SubjectToPolicy_i + \alpha_2(runvar_i * SubjectToPolicy_i) + \epsilon_i$$

- **Regression Results**:
  - **Findings**: Students that were subject to the policy on average completed 0.19 more years of schooling than student not subject to the policy ($p < 0.001$).
  - **Statistical Evidence**: T-statistic on the policy indicator is 9.18, indicating a robust first-stage relationship.

Earnings

```
Coefficients:
                                  Estimate Std. Error t value             Pr(>|t|)
(Intercept)                     10.9394943  0.0371873  294.17 < 0.0000000000000002 ***
left_sch_age16_or_later          0.1335516  0.0397821    3.36              0.00079 ***
I(left_sch_age16_or_later * runvar)  0.0009451  0.0000417   22.68 < 0.0000000000000002 ***
runvar                          -0.0006127  0.0000387  -15.82 < 0.0000000000000002 ***
```

Regression Equation and Results:

- **I**

$$(3)\ Earnings45_i = \beta_0 + \beta_1 runvar_i + \beta_2 \widehat{LeftSchoolAfter16}_i + \beta_3 (\widehat{LeftSchoolAfter16}_i * runvar_i) + \epsilon_i$$

  - Purpose: To isolate the direct effect of schooling beyond age 16 on earnings at age 45.
- **Key Findings**:
  - **Primary Coefficient (β2)**: The coefficient for staying in school after age 16 is 0.134, indicating a 13.4% increase in earnings at age 45 compared to peers who left school earlier.
  - **Runvar Coefficient (β1)**: The rate of change in earnings for those who did not stay in school until age 16 is -0.062 percentage point.
  - **Runvar + Interaction Term (β1 + β3)**: The rate of change in earnings for those who left school after age 16 is 0.033 percentage points.

Earnings

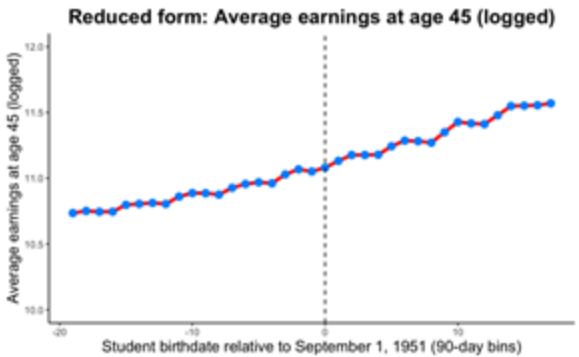
Threshold Indicator

- **Cutoff Date**: The Vertical dashed line at '0' on the x-axis represents the September 1, 1951 cutoff.

- **Division**: Splits the cohort into those required (post-cutoff) and not needed (pre-cutoff) to stay in school until at least age 16.

Trend Analysis

- **Before the Threshold**:
  - There is a slight, consistent increase in logged earnings as one approaches the threshold from the left.
  - Indicates that pre-policy factors have already begun impacting earnings.
- **After the Threshold**:
  - Noticeable, pronounced increase in earnings right after crossing the threshold.
  - Strong evidence of the positive impact of extended schooling on earnings.



Statistical Significance

- **Key Highlight**: Statistically significant increase in earnings for those subjected to compulsory schooling, affirming the policy's effectiveness.

Marital Status

```
Coefficients:
                                  Estimate Std. Error t value             Pr(>|t|)
(Intercept)                      0.2598050  0.0311468    8.34 < 0.0000000000000002 ***
left_sch_age16_or_later          0.1543565  0.0333201    4.63            0.0000036 ***
I(left_sch_age16_or_later * runvar) 0.0000281  0.0000349    0.81                 0.42
runvar                           -0.0000251  0.0000324   -0.77                 0.44
```

Regression Equation and Results:

- **E**

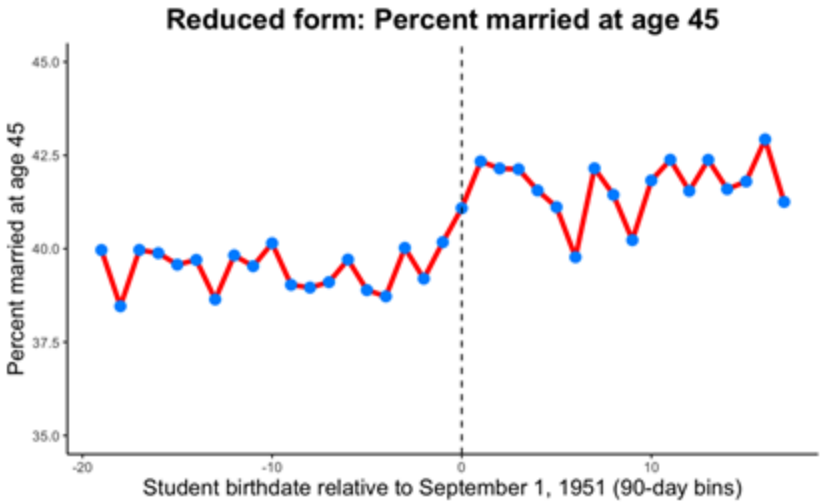
$$(4)\ Married45_i = \beta_0 + \beta_1 runvar_i + \beta_2 \widehat{LeftSchoolAfter16}_i + \beta_3 (\widehat{LeftSchoolAfter16}_i * runvar_i) + \epsilon_i$$

  - Purpose: Examine how staying in school beyond age 16 influences marital status at age 45.
- **Key Findings**:
  - **Primary Coefficient (β2)**: Statistically significant increase of 15.4% in the likelihood of being married at age 45 for those who stayed in school past age 16.
  - **Runvar Coefficient (β1)**: The rate of change in probability of marriage for those who did not stay in school until age 16 is -0.00251%, however this term is not statistically significant and therefore negligible.
  - **Runvar + Interaction Term (β1 + β3)**: The rate of change in probability of marriage for those who left school after age 16 is 0.0003% and is also not statistically significant.

Marital Status

Reduced Form Plot Analysis:

- **Visual Description**: The plot shows a noticeable jump in the percentage of individuals married at age 45 as they cross the policy threshold.
- **Before the Threshold**: Relatively stable or minor fluctuations in marital rates.
- **After the Threshold**: Clear increase, indicating a positive impact of extended schooling on



Children

```
Coefficients:
                                  Estimate Std. Error t value             Pr(>|t|)
(Intercept)                      0.4848990  0.0297866   16.28 < 0.0000000000000002 ***
left_sch_age16_or_later          0.1937781  0.0318650    6.08          0.0000000012 ***
I(left_sch_age16_or_later * runvar) -0.0000432  0.0000334   -1.29                 0.20
runvar                           0.0000411  0.0000310    1.32                 0.19
```

Regression Equation and Results:

- **E**

$$(5)\ HasKids45_i = \beta_0 + \beta_1 runvar_i + \beta_2 \widehat{LeftSchoolAfter16}_i + \beta_3(\widehat{LeftSchoolAfter16}_i * runvar_i) + \epsilon_i$$

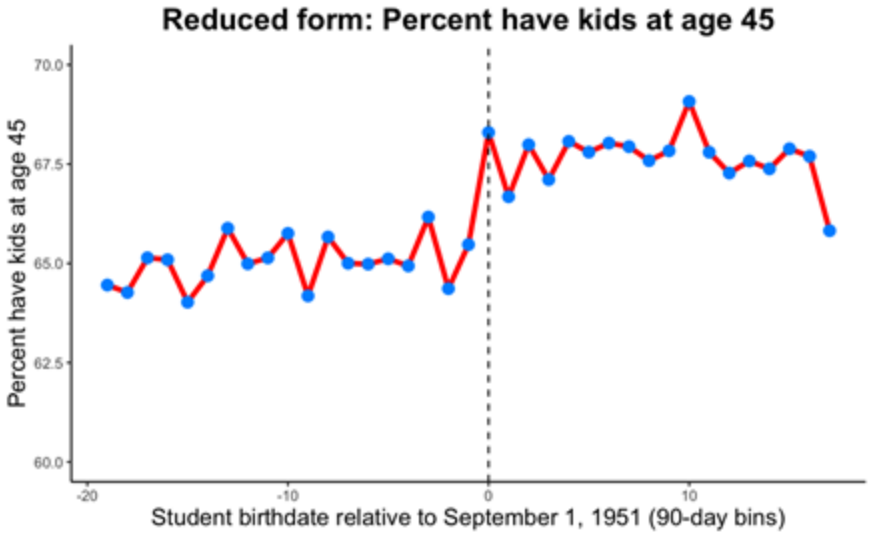
  - Purpose: Examine how remaining in school beyond age 16 influences the likelihood of having children by age 45.
- **Key Findings**:
  - **Primary Coefficient (β2)**: Statistically significant increase of 19.4% in the likelihood of having children at age 45 for those who stayed in school past age 16.
  - **Runvar Coefficient (β1)**: The rate of change in probability of having children for those who did not stay in school until age 16 is 0.0041%, however this term is not statistically significant and therefore negligible.
  - **Runvar + Interaction Term (β1 + β3)**: The rate of change in probability of having children for those who left school after age 16 is ~0% and is also not statistically significant.

Children

Reduced Form Plot Analysis:

- **Visual Description**: The plot shows fluctuations in the percentage of individuals having children at age 45 across the policy threshold.

- **Observations**:
  - **Before the Threshold**: Varied rates with some fluctuations, generally showing a stable or slight trend in parental status.
  - **After the Threshold**: Although there are ups and downs, the overall trend does not show a clear increase; rather, it highlights the complexities and potentially other factors influencing the decision to have children besides education.

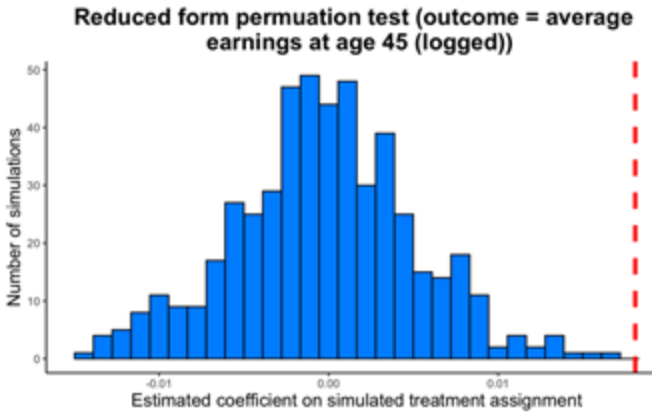


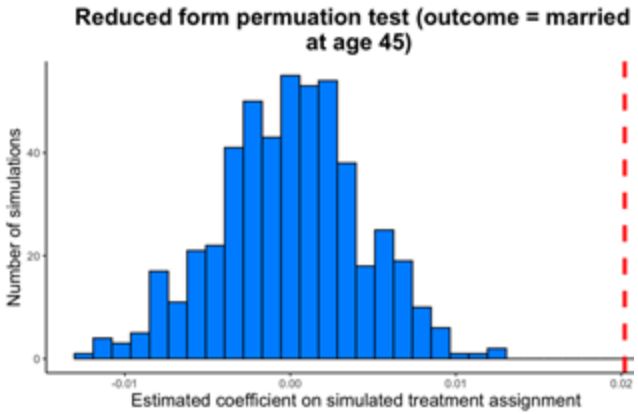
Permutation Test

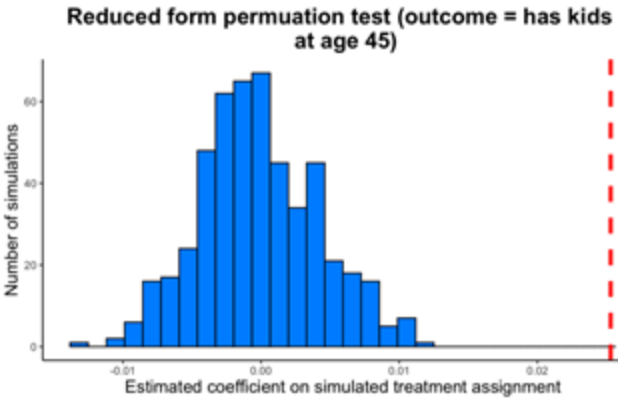
Permutation Testing:

- **The objective of Testing**:
  - To validate the robustness of the estimated causal effects of schooling on various outcomes (earnings, marital status, children) by randomized treatment assignments.
- **Permutation Test Results**:
  - **Plot Explanation**: Histogram showing the distribution of the estimated coefficients on simulated treatment assignment.
  - **Key Observation**: The absolute coefficient (indicated by the dashed red line) is significantly to the right of the distribution, suggesting that the observed effect of schooling on education is not due to random variation but likely represents an actual causal effect.
  - **Statistical Significance**: The distance of the actual coefficient from the center of the distribution emphasizes its statistical significance beyond what would be expected by chance.
  - **Implications**: This significant deviation from zero in the real data supports the hypothesis that staying in school past the age of 16 has a substantial and statistically significant positive effect on earnings, marital status, and having children.

Permutation Test







Conclusion

The effects of extended schooling beyond age 16 on education, earnings, marital status, and presence of children are both positive and statistically significant..

Policy Relevance

- **Insight**: The plot illustrates that higher educational attainment due to the policy correlates with increased earnings.
- **Policy Impact**: Reinforces the economic value of extending mandatory schooling.

Broader Impact

- Our results advocate for policies that encourage prolonged educational engagement.
- The societal benefits of investing in longer schooling durations are big.