

# Education and Voting

Josh (Yueru) Li, Yican Liu, Shiqi Zhou

April 27, 2022

## Abstract

Policy makers and economists have long held the view that education will lead to democracy prosperity by increasing civic participation. However, recent empirical researches have yielded inconclusive results regarding the magnitude of and the channels through which education affect voter turnout. In this paper, we quantify the causal effect of education on voter turnout in two cross sections, year 2004 and 2008. We then decompose the causal effect to investigate popular theories regarding the mechanisms through which education affects on voting behavior. Two popular theories are the “absolute education” model and the “education as a proxy” model. We conclude that the causal effect of education is a composite of both theories.

*Keywords:* political economy; voter turnout; education; civic participation

# 1 Introduction

Political participation is the foundation of democratic politics and the main mechanism through which citizens influence government decisions. Therefore, scholars have devoted a lot of effort to understanding the determinants of political participation. The conventional view holds that education could lead to democratic prosperity (Dewey, 1923; Lipset, 1959; Friedman, 1962; Glaeser et al., 2007) and shape people's political altitude (Cantoni et al., 2017). Lipset (1959) argues that a high level of education could be considered a necessary condition for democracy. And Friedman (1962) proposes that only when citizens have a minimum degree of literacy and widely accepted values can a society have democracy, and education can contribute to both.

Micro empirical research provides supporting evidence for this conventional view. Empirical evidence suggests that a higher level of education has a positive effect on civic participation, especially voter turnout, which will lead to democratic prosperity (Blais, 2000; Campbell et al., 1980; Nie et al., 1996; Verba and Nie, 1987; Verba et al., 1995; Wolfinger and Rosenstone, 1980). Verba and Nie (1987) link socioeconomic status to political participation. And Wolfinger and Rosenstone (1980) break socioeconomic status down into income and education attainment and find that education has stronger effect than income.

However, this conventional view is not supported by macro data, as shown in Figure 1. According to data from the Census Bureau's Current Population Survey (CPS), high school and college graduation rates in the U.S. have been increasing steadily since the 1960s. The percentage of population who have at least high-school degree rose from 41.1 percent in 1960 to 90 percent in 2020, while the that with college degree rose from 7.7 percent in 1960 to 37.5 percent in 2020. But voter turnout peaked at 80 percent in the 1800s and then has fallen to 60 percent in the 1900s. Moreover, voter turnout has maintained at 50-60 percent in the whole twentieth century and has not risen significantly with rising levels of aggregate education attainment in the United States. This paradox is defined by Brody (1978) as "puzzle of participation." The inconsistency between micro and macro evidence encourages more researchers to study the relationship between education and voter turnout. Some scholars revised the conventional view and put forward new models on the mechanisms through which education affect voter turnout. Other scholars believe that previous micro evidence only established the correlation between education and voter turnout, and the identification of the causal relationship needs to be carried out based on improved research methods.

In this paper, we quantify the causal effect of education on voter turn out and decompose the effect of education into various possible channels through to better understand various models proposed by scholars on how education might affect voter turnout. We start our analysis with a thorough and comprehensive revisit on Tenn (2007)'s identification strategy with the National Longitudinal Study of Youth 97 (NLSY97) data set. This data set provides detailed personal infor-

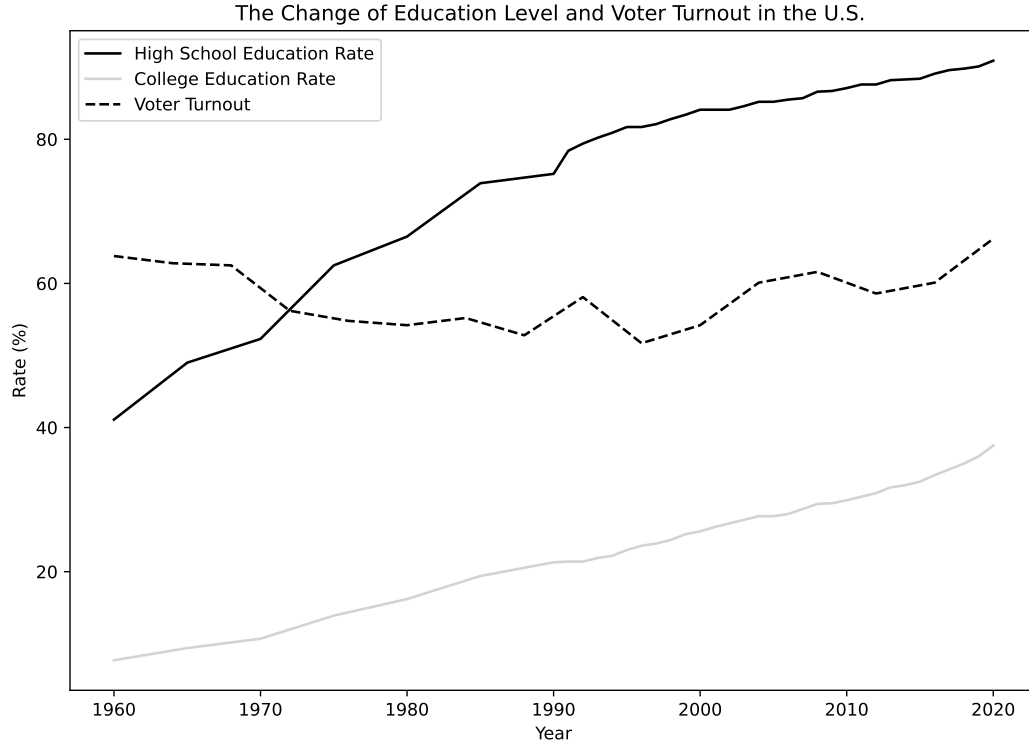


Figure 1: Education Level Growth and Voter Turnout

mation such as education, political activities' participation, income, and so on. As in Tenn (2007), we compare the individuals who have similar age and education levels in the current year and/or next year. We shall discuss more in detail about Tenn (2007)'s method in the identification strategy section. Controlling for other individual variables such as household income, gender, and so on, we find a significant positive causal effect of education year on voter turn-out.

Exploiting the long panel structure of NLSY 97, we propose a complementary identification strategy. We group individuals by their final education attainment to account for selection issue. Our complementary identification strategy also generates significant positive causal effect of education on voting probabilities.

We test three channels through which education may affect voter turn out. First possibility we test is that higher degree, not schooling years, affects voter turnout. We test this hypothesis by revisiting our identification strategies with indicators for high school, college, or graduate school degree instead of linear education year. Our results show that while discrete degree attainment do have significant effect on voter turn out, education year remains statistically significant. The second channel we explore is parental education. We include parental education (father's education

and mother’s education) into our main specifications and find that father’s high education has a significant positive impact on people’s voting turnout, while mother’s education does not. The effect of schooling year, however, remains statistically significant. Finally, we investigate the effect of school environment exposure on people’s voting behavior. We show that losing student status is correlated with a lower probability of voting and the educational effect on voting diminishes over time after graduation. The effect of education year remains statistically significant.

We proceed the rest of our paper as follows. In Section 2, we summarize the related literature about the causal effects of education on voter turnout, both theoretical overview and empirical strategies. In Section 3, we offer two identification strategies that we use in this paper. Section 4 describes the data and the main descriptive statistics. We presents our main results in Section 5. Section 6 discusses three channels for explaining the education effect. In Section 7, we provide a conclusive overview of our results.

## 2 Literature Review

### 2.1 Theoretical Overview

In the previous studies, there are three main explanatory mechanisms to explain how education affects voter turnout. The first mechanism is the conventional view, which can also be called the absolute education model (Dewey, 1923; Lipset, 1959; Friedman, 1962; Glaeser et al., 2007). The absolute education model holds that education and political participation have direct causality and explains the mechanism of education’s effect on politics through human capital. Education reduces the perceived cost of political participation for potential voters by improving civic skills and political knowledge (Nie et al., 1996; Verba et al., 1995; Wolfinger and Rosenstone, 1980). Also, it has been suggested that education stimulates potential voters’ political beliefs. That is, education increases potential voters’ belief in their ability to participate in political activities and their belief that political participation does influence government behaviors (Galston, 2001; Jackson, 1995). Therefore, with the conventional view, the more educated people will be more likely to vote, which is consistent with past cross-sectional microdata evidence (Blais, 2000; Campbell et al., 1980; Nie et al., 1996; Verba and Nie, 1987; Verba et al., 1995; Wolfinger and Rosenstone, 1980). Therefore, with the conventional view, the more educated people will be more likely to vote, which is consistent with past cross-sectional microdata evidence. However, the absolute education model has some drawbacks. First, Kam and Palmer (2008) argues that the micro evidence in previous studies only verifies the correlation between education and voter turnout but does not determine the causal relationship. Second, there is also debate about whether the effects of education on voter turnout are linear (Persson, 2015). Some argue that education matters while others believed that it is the higher education degree (such as college education or its equivalent) that

influences voting behavior. Third, the macro evidence is inconsistent with the theoretical inference and the micro evidence under the absolute education model. Although increasing aggregate education attainment in the United States has improved civic literacy, it cannot explain the voter turnout variation (Langton and Jennings, 1968; Niemi and Junn, 2005).

The second main explanatory mechanism is the “education-as-a-proxy” model, also called as “pre-adult socialization model”. This model is the extreme alternative to the absolute education model, which revises the causal inference in the absolute education model that education directly influences political behavior. According to the pre-adult socialization model, education does not directly affect political participation, but some factors before adulthood affect the educational choice and political participation in adulthood at the same time. Hess and Torney (1967) find that children with higher intelligence are more likely to participate in political activities, while intelligence is largely heritable (Bouchard Jr and McGue, 2003). Similarly, personality traits are thought to be heritable or formed early in life (Caspi et al., 2005). Also, it has been shown that personality traits influence educational attainment (Poropat, 2009) and political participation (Schoen and Schumann, 2007). In addition, other studies believe that family environment and parents’ political attitudes also affect education and political participation at the same time (Persson, 2014; Gidengil et al., 2019). In the pre-adult socialization model, the empirical results of studying the relationship between education and political participation by taking education as an independent variable cannot infer the causal relationship between them. It is no longer education that influences political participation, but rather unobservable pre-adult factors. Therefore, with pre-adults factors taken into consideration, how to determine whether education can directly influence political participation has become a new problem.

The third mechanism is called the relative education model. This model is quite different from the way causalities were established in the first two mechanisms. The relative education model holds that education affects people’s rank in society and the individual’s relative education level or social status determines the individual’s voting behavior (Nie et al., 1996). Therefore, the higher aggregate education level in a society does not indicate higher voter turnout, which answers the “puzzle of participation” proposed by Brody (1978). Also, some scholars argue that the emergence of the “puzzle of participation” is due to the possible negative effects of other social factors, such as the complexity of politics, and the decline of education quality, which offsets the positive factors brought by education (Carpini, 1997).

To summarize, previous studies have had problems in determining whether there is a causal relationship between education and politics, mainly because it is difficult to empirically test these three possible mechanisms. In this paper, we will mainly consider the absolute education mechanism and the pre-adult socialization mechanism, and the possibility that both mechanisms exist simultaneously, to determine whether education is a direct or a proxy factor for political partici-

pation.

## 2.2 Empirical Methodologies

Methodologically, it is difficult to assess the causal effects of education. Because education attainments are rarely exogenous, single cross-sectional data on education records and voting behaviors rarely provides data detailed enough for causal inference. Exploiting rich panel data sets, previous studies have tried to overcome the endogeneity of education and unobserved individual differences (such as family-level factors or pre-adult factors): instrumental variables (Dee, 2004; Milligan et al., 2004) and the quasi-natural experiment (Berinsky and Lenz, 2011; Kam and Palmer, 2008; Pelkonen, 2012; Tenn, 2007).

Dee (2004) discusses two instrumental variables, the compulsory education laws (Acemoglu and Angrist, 2000) and the distance to a higher education institution (Card, 1993). Dee (2004) finds that education level has a significant positive impact on voter turnout. Milligan et al. (2004) also use compulsory education laws as an instrumental variable to study voter turnout in the United States and The United Kingdom. They find that completing high school in the United States had a positive effect on voting, but not in the UK. However, after controlling for enrollment, the impact of education in the United States is greatly reduced. The conclusions of these two studies are consistent with the inference of the absolute education mechanism. However, there may be missing variables in both instrumental variable approaches. In addition, Akbulut (2007) finds that compulsory education law is no longer an effective instrumental variable since late twenty century because education has been valued more and more people have been chosen to return to school to pursue higher degrees.

Quasi-natural experiments are designed to isolate the effect of education on voting behavior through the artificial construction of “randomly assigned” control and treatment groups by policy changes or reasonable assumptions. In 1959, an education reform that increases the compulsory education years was implemented randomly in different parts of Norway at different times. Pelkonen (2012) uses the data before and after this reform which forms a quasi-natural experiment. And Pelkonen (2012) finds that education has no direct impact on political participation. Berinsky and Lenz (2011) use the quasi-natural data from the policy of conscription by lottery during the conscription period in Vietnam. Young men were randomly assigned to the army by lottery during the conscription period in Vietnam and attending college could postpone their military service. Therefore, the policy of conscription by lottery formed a quasi-natural experiment and ensured the exogeneity of education. And from the results of Berinsky and Lenz (2011), education level has no direct effect on political participation. However, this research only includes men which may cause biased results.

Kam and Palmer (2008) first proposed the use of observable variables for grouping matching,

Table 1: Sample Grouping at Year  $t$ 

Group	Age <sub><math>t</math></sub>	Edu <sub><math>t</math></sub>	Age <sub><math>t+1</math></sub>	Edu <sub><math>t+1</math></sub>
Group1	22	15	23	Not Important
Group2	22	16	23	Not Important
Group3	21	15	22	15
Group4	21	15	22	16

to artificially form the control group and treatment group. From their results, there was also no significant difference in turnout between college participants and non-participants. However, one drawback in Kam and Palmer (2008)’s study is that unobserved variables are still ignored. Tenn (2007) takes another route and compares individuals who are 1 year apart from each other in both ages and education attainment and limit the attention to the younger people who will attain another year of education the next year. The identification assumption is that the constructed groups of people differ only in 1 year of education. Two problems arise with this approach. The identification strategy of Tenn (2007) limits the target of study to enrolled students and new graduates, and does not allow one to examine the long run effect of education on political participation. Tenn (2007) also arrives the similar results with Pelkonen (2012), Berinsky and Lenz (2011), and Kam and Palmer (2008), which support the pre-adult socialization mechanism.

### 3 Identification Strategies

#### 3.1 Tenn (2007)’s Strategy

Tenn (2007)’s ingenious idea allows one to almost perfectly control for potential selection issues in educational attainment. Its mechanism is simple: we compare individuals who will be of certain age and have certain education levels in the next year to similar individuals in the current year. We assume that after controlling for observable characteristics, the two groups would have similarly distributed unobserved characteristics, which resolves the potential selection issue. The identification assumption hinges on people of adjacent ages not having significant generational gaps in their cognitive traits. More specifically, Table 1 provides a simple illustration of grouping.

Under grouping illustrated by Table 1, there is no variation in education year in each group. Under the above mentioned identification strategy, we merge group 1 and group 3, and merge group 2 and group 4. It is the variation of education in the second group that allows us to identify the causal effect of education.

We estimate the following multinomial logit model on our data:

$$\text{Pr}_{it} = \Lambda(\lambda \text{Edu}_{i,t} + \alpha \text{Group FE} + \beta' \text{Controls}_{i,t} + e_{it})$$

It is worth noting that education here is used as an umbrella term and should be interpreted generously. As we’ve discussed in the introduction, it might be the extra year of immersion in a politically active environment that schools tend to be (Tenn (2007)). It might be that students are induced to vote by an extra year of schooling that teaches them the importance of civic participation. It should be noted that we cannot separately identify these two possibilities and method does not allow us to detect a possibility that the effect of education only materialize much later in peoples’ lives.

### 3.2 Our Complementary strategy

Tenn (2007)’s analysis is limited by the short panel he has at the time of analysis. The CPS data only provides observations on the same individual for up to 2 consecutive years, hence, while he could implement the above identification strategy, he also had to limit the targets of study to just students and new graduates. Exploiting the long panel nature of NLSY97, we can control for the final education attainment and extend the study to both students and professionals.

We need a second identification assumption to exploit the long panel structure provided by NLSY97. Our new identification strategy relies on the assumption that after controlling for a rich set of personal characteristics and family background, individuals who choose to attain the same level of education attainment would have unobservables that are independently distributed from the education year at each point in their lives. This identification assumption would also allows us to also control for age in our analysis.

For this original identification strategy of ours, we estimate the following multinomial logit model

$$\text{Pr}_{it} = \Lambda(\lambda \text{Edu}_{i,t} + \alpha \text{Final Education at 2019} + \beta' \text{Controls}_{i,t} + e_{it})$$

We perform selection on observables to achieve conditional independence between education and unobservables.

## 4 The Data

The data we use comes from the National Longitudinal Study of Youth 97 (NLSY97) data set. Summary statistics for two snapshots of the data set, one for year 2004 and another for year 2008 are presented in Table 2. Notice that we have the same number of observations at the two time points, so we have a balanced panel and don’t need to worry about attrition problem.



Table 2: Sumamry Statistics

Variables	Mean	S.D.	P10	Median	P90	Obs.
Year of 2004						
Vote	0.52	0.49	0	1	1	8,984
EduYear	13.22	2.33	10	13	16	8,984
Income (/1000)	46.91	71.35	2	25	112	6,827
Female	0.50	0.50	0	1	1	8,984
TestScore (/1000) (1997)	386.34	320.33	0	341	864	5,700
Age	21.93	1.39	20	22	24	8,984
Move	0.15	0.35	0	0	1	8,984
Married	0.13	0.34	0	0	1	8,984
GovAid	0.08	0.27	0	0	0	8,984
EduDad	13.00	3.45	9	12	17	4,355
EduMom	12.70	3.58	9	12	16	6,165
Year of 2008						
Vote	0.34	0.47	0	0	1	8,984
EduYear	13.92	2.93	10	13	18	8,984
Income (/1000)	47.90	57.59	3	33	110	6,875
Female	0.50	0.50	0	1	1	8,984
TestScore (/1000) (1997)	386.34	320.33	0	341	864	5,700
Age	25.93	1.39	24	26	28	8,984
Move	0.13	0.33	0	0	1	8,984
Married	0.25	0.43	0	0	1	8,984
GovAid	0.12	0.32	0	0	1	8,984
EduDad	13.00	3.45	9	12	17	4,355
EduMom	12.70	3.58	9	12	16	6,165

The “move” variable is coded to be a binary variable where a 1 refers to a cross state or cross country “move” in the current year or the year before. “move”s across counties within the same state are treated as not moving. This is justified by the different voter registration processes within different states. We speculate that a new comer to a state would take some time to adjust, process the paper work and are thus likely to go through all the paper work to vote immediately.

For the “income” variable, we use household income as most people don’t have income yet in 2004. We replicated our results using individual data for 2008 and yield similar results. So we stick to individual data for both 2004 and 2008.

The variable “GovAid” is a binary variable that takes the value of 1 if the individual has received some form of government aid during the four years before the election year.

The variable “test score” codes the result of the AFQT score that the NLSY97 survey collected. We use it to control for personal capabilities, which is commonly accepted in labor economics Altonji and Pierret (2001).

We see that in our sample, the overall voter turn out was much higher in 2004, at 52%, than in 2008, at 35%. Note that the national voter turnout in 2004 was 60% while the national voter turnout in 2008 was 62% McDonald (2016a). Since our sample consists of a particular age group at two election years. We further look at turnout rates by age group. In 2004, individuals in our sample were aged 19 – 24, while in 2008, they were aged 23 – 28. The voter turnout rates for the corresponding age groups were around 44% in 2004 and around 52% in 2008, McDonald (2016b). So we speculate the our estimates would be more representative for the 2004 election year.

## 5 Main Results

The first two columns show the regression result for our basic specification, which only measures correlation between education and voting probabilities. They are multinomial logit regressions with rich controls. We observe in our data set that higher education level is associated with a higher probability to vote. The estimated coefficient on education is 0.259 using data in year 2004 election and is 0.164 using data in 2008 election. They are both statistically significant. In marginal effect terms, during the 2004 election, the mean marginal effect of education is 5.2% while during the 2008 election the mean marginal effect of education is 3.9%. We emphasize that these are correlation values and should not be interpreted causally.

It is worth noting that other variables that are correlated with probability to vote significantly are “gender”, “income”, “AFQT test score”, “marriage” and “move”. The coefficient on “move” is only significant for election in 2004 and the coefficient on “marriage” is only significant for the election in 2008. This is a general trend for other specifications we’ve estimated for causal inference as well. It is also worth noting that while the coefficient on “move” is not statistically significant in the 2008 estimate, the coefficient on “marriage” is. That is, people who are married are more likely to vote. We leave a discussion on this issue in interpretation of our causal inference specifications.

We add in “Gov Aid” as a control because we speculate that people receiving government aids would be more likely to vote because they felt the importance of having a functional government that provides policies that people need. At the same time, people who go through the documentation to get government may also be more willing to go through the paper work to vote. Without carefully examination of the causal channel of receiving government aid, we include it as a control because the unemployment rate and eligibility to government aid tend to be correlated with education level.

Table 3: Education and Voting Behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EduYear	0.259*** (0.016)	0.164*** (0.012)	0.340*** (0.067)	0.161*** (0.023)	0.185*** (0.019)	0.137*** (0.016)	0.477** (0.169)	0.218* (0.099)
Income	0.013** (0.004)	0.019*** (0.005)	0.012** (0.004)	0.017*** (0.005)	0.012** (0.004)	0.017*** (0.005)	0.003 (0.006)	0.025 (0.013)
Age	-0.002 (0.021)	0.001 (0.021)	-0.161 (0.108)	-0.086 (0.112)	0.012 (0.022)	0.003 (0.021)	-0.381 (0.197)	-0.131 (0.218)
Female	0.247*** (0.058)	0.231*** (0.058)	0.247*** (0.059)	0.222*** (0.058)	0.217*** (0.059)	0.238*** (0.058)	0.404*** (0.110)	0.313* (0.155)
Move	-0.301*** (0.081)	0.142 (0.083)	-0.307*** (0.082)	0.144 (0.084)	-0.347*** (0.082)	0.123 (0.083)	-0.225 (0.150)	0.165 (0.203)
Married	0.071 (0.087)	0.193** (0.065)	0.091 (0.088)	0.191** (0.066)	0.048 (0.088)	0.184** (0.066)	-0.36 (0.224)	0.101 (0.186)
TestScore	0.012*** (0.001)	0.004** (0.001)	0.012*** (0.001)	0.003* (0.001)	0.010*** (0.001)	0.003* (0.001)	0.013*** (0.002)	0.003 (0.003)
GovAid	0.251* (0.106)	0.036 (0.09)	0.271* (0.108)	0.003 (0.091)	0.238* (0.107)	0.038 (0.091)	0.247 (0.307)	-0.249 (0.262)
Cons	-4.019*** (0.483)	-3.507*** (0.557)	-2.001 (2.325)	-1.387 (3.002)	-3.316*** (0.564)	-2.967*** (0.633)	0.637 (0.514)	-1.455 (6.204)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Race FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenns Control			Yes	Yes			Yes	Yes
Final Education					Yes	Yes		
Year	2004	2008	2004	2008	2004	2008	2004	2008
Sample	Full	Full	Full	Full	Full	Full	Enrolled	Enrolled
N	5676	5700	5660	5678	5674	5698	1762	837

Note: Standard errors in parentheses, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

## 5.1 Revisit Tenn (2007) with NLSY97

We first group every student based on the above outlined identification strategy. Notice that some individuals could be grouped into multiple groups. For example, an individual of age 23 can be used to compare to individuals of age 22 and age 24. More information about the data manipulation can be found in Table 3.

We replicate Tenn (2007)'s identification for both elections years in Column 3 and 4 of Table 3. We get statistical significant estimates for coefficient on education year, household income, gender,

“move” and “AFQT test score”. The estimated coefficient on education is 0.34 for 2004 election and 0.16 for 2008 election. In mean marginal effect terms, an extra year of education, on average, increases the probability of voting by 7.0% and 3.5% for 2004 and 2008 elections respectively.

In comparison, the correlation estimates over estimate the effect of education on voting in 2008 and underestimates the effect in 2004. Over the four years span, the number of individuals in our sample have decreases by more than 50% <sup>3</sup>. It is important to note that Tenn (2007)’s identification strategy only allows one to test the effect of education on students and new graduates. To understand why the direction of bias has changed comparing the two estimates, we switch to our complementary identification strategy and exploit the long and balanced panel we have from 2004 to 2008.

## 5.2 Controlling for final educational attainment

Controlling for final education allows us to extend the analysis to not just students, but also professionals who have graduated for more than 1 year. Column 5 and 6 of Table 3 show the basic specification of the logit model we estimate. The coefficients on education are statistically significant at 0.185 for 2004 and 0.137 for 2008. In mean marginal effect terms, an extra year of education, on average, increases one’s probability to vote by 3.9% and 3.0% for year 2004 and 2008 respectively. We also get statistical significant estimates for coefficient on education year, household income, gender, “move” and “AFQT test score”.

Compare to the above estimates, it is clear that the correlation estimate over estimates the mean marginal effect of one extra year of education for both education year. Correlation regression over estimates the marginal effect by 33% in 2004 and 30% in 2008. It appears that the endogenous selection of education accounts of around 30% of the observed correlation between education and voting. The consistent overestimation of about 30% between the two election year gives us confidence in our identification strategy.

## 5.3 Comparing Two identification strategies

All regression results with Tenn (2007)’s model and our complementary model show that education years have direct effects on voter turnout, which support the absolute education mechanism. Moreover, as we’ve discussed in the introduction and in Section 5.1, the estimated mean marginal effect from Tenn (2007)’s model is a combination of two effects: short run effect of education that affects students’ cognitive abilities and awareness and the effect of recent immersion in a politically active environment. Therefore, it is likely to overestimate the effect of education. We can see this comparing the estimates from Tenn (2007)’s model and our complementary model. While in both election year, Tenn (2007)’s model predicts a higher mean marginal effect, the over-prediction is more severe in 2004 when a higher proportion of the sample was still in school. The number in

3 is likely an underestimate of the evolution of student status in the sample. This is because a proportion of students in 2008 would be enrolling in graduate school or professional programs, both of which are different in nature compared to high school and college and should be considered as professional experiences.

Another worth noting fact is the interchanging statistical significance of “move” for 2004 and 2008 regressions. This is a trend across all specifications except for when we look just at the student sub-sample. Our initial speculated reason for the significance of a recent cross-state move is the required paperwork associated with registering in a new state. This paperwork process may seem more daunting for younger individuals when they have less experience with documentation. By the time of 2008, most of them have acquired much experience dealing with complicated documentation processes with various institutions including with the local government, so their voting decisions are less likely to be affected by the required paper work.

For the marriage variable, we see from Table 2, the proportion of individuals who are married doubled from 2004 to 2008. There are two potential channels through which this affect our estimate for marriage. Firstly, this change in sample composition allows us to more accurately estimate the effect of marriage. Secondly, marriage may have different influence on people at different ages in their lives.

## **6 Some Channels for Explaining the Education Effect**

### **6.1 Education Year: Natural Cutoff and Linearity Assumption**

One alternative channel of education effect is that only higher education (college or equivalent) affects voter turnout, not schooling years (Persson, 2015). So, we use dummy variables to indicate the education levels instead of schooling years in regression. We regress voter turnout in 2004 and 2008 on years of education, education dummy variables and both of them with Tenn (2007)’s model and our complementary model, respectively. In effect, this is equivalent to dropping the linearity assumption in the effect of education. Table 4, Column 2, represents the regression results only including education dummy variables with Tenn (2007)’s model in 2004. The results indicate that both high school and college education experience significantly increase the likelihood of potential voters voting. The estimated coefficients of high school and college education experience are 0.53 and 0.48, respectively. One possible explanation for this result is that, compared with graduation education, high schools and colleges offers more political education (Hansen, 2016). And the active campus atmosphere of high school and college may increase voter turnout by exposing people to more information about politics. Moreover, we reach the same conclusion by using data from 2008 or our complementary model 4.

We also including both education years and education dummy variables in regression. In gen-

Table 4: Collage Attendance and Voting Behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
EduYear	0.340*** (0.067)		0.270*** (0.068)	0.161*** (0.023)		0.126*** (0.026)	0.185*** (0.019)		0.119*** (0.023)	0.137*** (0.016)		0.105*** (0.018)
HighSchool		0.534*** (0.119)	0.524*** (0.118)		0.297* (0.128)	0.230 (0.128)		0.586*** (0.146)	0.465** (0.148)		0.547** (0.189)	0.456* (0.190)
College		0.480*** (0.113)	0.416*** (0.114)		0.382*** (0.113)	0.230 (0.118)		0.626*** (0.100)	0.378*** (0.110)		0.602*** (0.132)	0.330* (0.139)
Graduate		0.025 (0.264)	-0.032 (0.264)		0.260* (0.121)	0.059 (0.127)		-0.038 (0.247)	-0.254 (0.251)		0.087 (0.143)	-0.117 (0.147)
Income	0.012** (0.004)	0.011* (0.004)	0.011* (0.004)	0.017*** (0.005)	0.016** (0.005)	0.016** (0.005)	0.012** (0.004)	0.011* (0.004)	0.011* (0.004)	0.017*** (0.005)	0.017*** (0.005)	0.016** (0.005)
Age	-0.161 (0.108)	-0.043 (0.103)	-0.163 (0.108)	-0.086 (0.112)	-0.041 (0.109)	-0.088 (0.112)	0.012 (0.022)	0.038 (0.022)	0.019 (0.022)	0.003 (0.021)	0.010 (0.021)	0.005 (0.021)
Female	0.247*** (0.059)	0.240*** (0.059)	0.230*** (0.059)	0.222*** (0.058)	0.210*** (0.058)	0.212*** (0.059)	0.217*** (0.059)	0.213*** (0.059)	0.205*** (0.059)	0.238*** (0.058)	0.226*** (0.058)	0.222*** (0.058)
Move	-0.307*** (0.082)	-0.319*** (0.082)	-0.321*** (0.082)	0.144 (0.084)	0.126 (0.084)	0.136 (0.084)	-0.347*** (0.082)	-0.367*** (0.082)	-0.348*** (0.082)	0.123 (0.083)	0.100 (0.083)	0.123 (0.084)
Married	0.091 (0.088)	0.061 (0.089)	0.072 (0.089)	0.191** (0.066)	0.164* (0.066)	0.181** (0.066)	0.048 (0.088)	0.000 (0.088)	0.041 (0.088)	0.184** (0.066)	0.149* (0.065)	0.173** (0.066)
TestScore	0.012*** (0.001)	0.011*** (0.001)	0.010*** (0.001)	0.003* (0.001)	0.003* (0.001)	0.002 (0.001)	0.010*** (0.001)	0.010*** (0.001)	0.009*** (0.001)	0.003* (0.001)	0.004** (0.001)	0.002 (0.001)
GovAid	0.271* (0.108)	0.255* (0.109)	0.281** (0.109)	0.003 (0.091)	-0.009 (0.091)	-0.001 (0.092)	0.238* (0.107)	0.223* (0.107)	0.255* (0.108)	0.038 (0.091)	0.016 (0.091)	0.036 (0.091)
Cons	-2.001 (2.325)	-0.426 (2.308)	-1.64 (2.330)	-1.387 (3.002)	-0.677 (2.939)	-1.04 (3.012)	-3.316*** (0.564)	-2.256*** (0.547)	-3.215*** (0.580)	-2.967*** (0.633)	-2.009** (0.612)	-3.173*** (0.646)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Race FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tennis Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Final Education							Yes	Yes	Yes	Yes	Yes	Yes
Year	2004	2004	2004	2008	2008	2008	2004	2004	2004	2008	2008	2008
N	5660	5660	5660	5678	5678	5678	5674	5674	5674	5698	5698	5698

Note: Highschool indicates whether the observation ever went to high school, with 1 indicating the observation did. College and Graduate have similar meaning. Standard errors in parentheses, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

eral, both education years and education dummy variables have significant positive effects on voter turnout. Comparing Column 1 with 3 in Table 4, we find that the mean marginal effect of education year decreases from 7% to 5.6% after including education dummy variables. That is, the different education experience explains part of the education effects on voter turnout. But, when using data in 2008 for regression, we find that the coefficients of education dummy variables is no longer statistically significant (in Column 6 and 12, Table 4). The results match our expectations. We believe that the positive effect of high school and college dummy variables on voter turnout is mainly due to curriculum education and school environment, and is a short-term impact. So, the effect of high school and college education experiment will decrease as the observations leave school. In our data, there are only 1762 and 837 observations enrolled in school in 2004 and 2008,

respectively. Therefore, the short-term positive effect of high school and college dummy variables have largely reduced in 2008, which results in the insignificance of coefficients of education dummy variables.

To summarize, education has effects on voter turnout through two channels, one is the increase of education years, and the other is the different types of education. And the positive effects of education years are long-term, while the effects of different types of education are short-term.

## 6.2 Parental Education

Another channel through which our estimates can be biased is parental education. Researchers have found that higher parental education contribute to higher child education attainment (White and Mistry, 2016; Chevalier, 2004). Assuming education does contribute positively to political participation, individuals with better educated parents are more likely to be immersed in a politically active environment since a very young age. The rich family background information in NLSY97 allows us to control for parents' education. The results are shown in Table 5.

Parental education coefficients are statistically significant in all but one specifications. This result indicates that education also has effects on voter turnout as a proxy factor of parents' education, which support the pre-adult socialization mechanism. Consistently, when we control for both parents' education, the coefficient on father's is statistically significant but the coefficient on mother's education is not. This is suggestive evidence that fathers play a bigger role in shaping civic participation of their children. But within our framework, we interpret them as controls and only interpret the coefficient on education causally.

For both Tenn (2007)'s model and our complementary model, in both 2004 and 2008, the effect of the education level is still statistically significant regardless of whether we control for both residential parents' education level or the highest education level among the parents, with the exception for Tenn (2007)'s model in year 2004. Given that the estimate is significant in all 8 specifications but 1, and the value is actually close to being statistically significant at 5% with a  $t$  score of 1.74. We believe that we cannot reject the hypothesis that individuals' self education has statistically significant impact on voting behavior.

We calculate the mean marginal effect of education in the new specifications in Table 5. Under Tenn (2007)'s model, for election year 2004, the mean marginal effect of education are 3.5% and 6.4% respectively for controlling for both parents' education and the highest education among parents. For election year 2008, the mean marginal effect are 3.5% and 3.4% respectively. With the exception of the estimate for controlling for both parents' education for 2004 election, the mean marginal effects remain stable with minor decreases compared to our main specification.

The mean marginal effect of education in our complementary models are calculated to be 3.7% for both types of parent education controls for election year 2004. For election year 2008, the mean

Table 5: The Impact of Parental Education on Voting Behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EduYear	0.174 (0.100)	0.324*** (0.071)	0.162*** (0.030)	0.160*** (0.024)	0.185*** (0.026)	0.183*** (0.020)	0.110*** (0.020)	0.132*** (0.016)
EduDad	0.037** (0.014)		0.045** (0.014)		0.030* (0.014)		0.040** (0.014)	
EduMom	0.008 (0.013)		0.009 (0.012)		0.009 (0.013)		0.010 (0.013)	
ParentEduHigh		0.018* (0.009)		0.025** (0.009)		0.017 (0.009)		0.025** (0.009)
Income	0.011* (0.005)	0.012** (0.004)	0.014* (0.006)	0.015** (0.005)	0.010* (0.005)	0.012** (0.004)	0.014* (0.006)	0.016** (0.005)
Age	-0.027 (0.143)	-0.151 (0.112)	-0.036 (0.147)	-0.052 (0.116)	0.01 (0.029)	0.029 (0.022)	0.002 (0.027)	-0.001 (0.021)
Female	0.322*** (0.077)	0.257*** (0.061)	0.197** (0.075)	0.230*** (0.060)	0.277*** (0.077)	0.225*** (0.061)	0.196** (0.074)	0.252*** (0.060)
Move	-0.311** (0.102)	-0.322*** (0.084)	0.105 (0.106)	0.136 (0.087)	-0.352*** (0.101)	-0.369*** (0.084)	0.082 (0.105)	0.114 (0.086)
Married	0.295** (0.114)	0.115 (0.092)	0.241** (0.082)	0.216** (0.068)	0.235* (0.113)	0.063 (0.091)	0.235** (0.080)	0.209** (0.067)
TestScore	0.010*** (0.002)	0.011*** (0.001)	0 (0.002)	0.002 (0.001)	0.008*** (0.002)	0.009*** (0.001)	-0.001 (0.002)	0.002 (0.001)
GovAid	0.224 (0.146)	0.297** (0.111)	0.088 (0.121)	0.071 (0.094)	0.227 (0.143)	0.269* (0.110)	0.144 (0.120)	0.103 (0.093)
Cons	-2.974 (2.954)	-2.202 (2.422)	-3.372 (4.014)	-2.658 (3.142)	-3.635*** (0.748)	-3.721*** (0.596)	-2.919*** (0.817)	-3.048*** (0.663)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Race FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenns Control	Yes	Yes	Yes	Yes				
Final Education					Yes	Yes	Yes	Yes
Year	2004	2004	2008	2008	2004	2004	2008	2008
N	3453	5323	3453	5337	3460	5336	3481	5351

Note: Standard errors in parentheses, \* p &lt; 0.05, \*\* p &lt; 0.01, \*\*\* p &lt; 0.001.



Table 6: Year Since Graduated from School

Variable Names	(1)	(2)	(3)	(4)
EduYear	0.281*** (0.072)	0.144*** (0.029)	0.126*** (0.029)	0.115*** (0.020)
YearSinceGrad	-0.072** (0.026)	-0.015 (0.016)	-0.072** (0.027)	-0.024* (0.010)
Income	0.013** (0.004)	0.017*** (0.005)	0.012** (0.004)	0.017*** (0.005)
Age	-0.103 (0.110)	-0.069 (0.113)	0.067* (0.030)	0.021 (0.023)
Female	0.250*** (0.059)	0.221*** (0.058)	0.220*** (0.059)	0.236*** (0.058)
Move	-0.309*** (0.082)	0.143 (0.084)	-0.348*** (0.082)	0.121 (0.083)
Married	0.106 (0.089)	0.193** (0.066)	0.061 (0.088)	0.188** (0.066)
TestScore	0.012*** (0.001)	0.003* (0.001)	0.010*** (0.001)	0.003** (0.001)
GovAid	0.280** (0.108)	0.003 (0.091)	0.252* (0.107)	0.039 (0.091)
Cons	-2.282 (2.330)	-1.475 (3.002)	-3.587*** (0.574)	-3.019*** (0.634)
Region FE	Yes	Yes	Yes	Yes
Race FE	Yes	Yes	Yes	Yes
Tenns Control	Yes	Yes		
Final Education			Yes	Yes
Year	2004	2008	2004	2008
N	5660	5678	5674	5698

Note: Standard errors in parentheses, \*  $p < 0.05$ , \*\*  $p < 0.01$ ,  
\*\*\*  $p < 0.001$ .

marginal effects are calculated to be 2.4 and 2.9 for controlling for education of both parents and controlling for the highest education of parents respectively. We see that controlling for parents' education decreases the estimates from the main specification. But the mean marginal effect of education remains statistically significant.

### 6.3 Controlling for School Environment Exposure Effect

In Section 5.3, we discussed how our complementary identification, exploiting the long panel nature of NLSY97, provides identification of the effect of education many years after graduation. The concern we have about Tenn (2007)'s model is that it mostly captures the effect of education on people who are still staying in or haven't left schools long enough. Given the politically active environments that schools tend to be, we worry that our estimated result will be heavily influenced by the school exposure effect. We justify our much lower estimated mean marginal effect of education in our complementary model following the above analysis. In this section, we formalize this idea and control for the years lapsed since graduation from the final institution they attended. The results are presented in Table 6.

The specifications in Column 1 and 2 controls for years since graduation in Tenn (2007)'s model. Notice that in Tenn (2007)'s framework, we can only study people who have graduated from all school in the past year. So the point estimate on year since graduation in the first two columns are effectively controlling for student status. We observe the losing student status is correlated with a lower probability of voting, but the estimated effect of education remains statistically significant. The mean marginal effect of education are calculated to be 5.8% and 3.6% for election in 2004 and 2008 respectively. Compared to the basic specification of Tenn (2007)'s model, the point estimates decrease with decrease in 2004 estimate to be most significant. This aligns with our analysis that the 2004 cross-section contains more students than the 2008 cross section.

The specifications in Column 3 and 4 controls for years since graduation in our complementary model. We observe the same trend and the new mean marginal effect of education are 2.6% and 2.5% in election year 2004 and 2008 respectively. Down from 3.9% and 3.0%, respectively, from our original specification. This result suggests that the mean marginal effect of education has been relatively stable from 2004 to 2008. This discovery, together with the rising education level trend documented in Figure 1, lends itself in support of the hypothesis that it is the higher relative education level that is causing people to be more likely to vote.

## 7 Conclusion

On top of quantifying the causal effect of education on voting, we examine whether education is a direct or a proxy factor in the causal relationship between education and voter turnout. Our empirical methodology, Tenn (2007)'s model and our complementary model, are well suited to addressing this question since both models isolate the causal effect of education by controlling unobserved individual differences. In addition, we try to decompose the effect of education by exploring some possible channels to further explain the effect of education: (i) marginal effects of education year with unobserved characteristics controlled; (ii) education affecting voter turnout

as dummy variables instead of continuous education years; (iii) the impact of parental education on voter turnout; and (iv) considering the school environment exposure effect by control years since graduation.

Our results show that education is both a direct and proxy factor in the causal relationship between education and voter turnout. With group controlled in Tenn's model and our complementary model, education year has significant direct effects on voting behavior, which support the absolute education mechanism. As parental education is taken into consideration, we find that the mean marginal effect of education decreases while parental education explains part of the positive effect of education on voter turnout. This result matches the inference of the pre-adult socialization mechanism. Therefore, the effect of education on voter turnout is through both the absolute education mechanism and the pre-adult socialization mechanism.

Moreover, as a direct factor, education has effects on voter turnout through two channels, one is the increase of education years, and the other is the different types of education. But as people leave school, the effects of different types of education reduce over time, while the effect of education years is still statistically significant. These results suggest that the positive effects of education years are long-term, while the effects of different types of education are short-term.

We recognize two limitations in our analysis. First, we do not include individuals who are not able to vote to analyze the education effect on them because the lacking of data on voting. Second, we do not consider the relative education mechanism in our analysis. Only when these issues addressed in the future, will we have better understanding on the relationship between education and voter turnout. As a consequence, we haven't been able to provide a justification for the "puzzle of participation". This is, however, not a major concern because it is unlikely that education is the only determinants of voter turnout and a more comprehensive research is needed to address the "puzzle of participation".

## References

- Acemoglu, D. and Angrist, J. (2000). How large are human-capital externalities? evidence from compulsory schooling laws. *NBER macroeconomics annual*, 15:9–59.
- Akbulut, M. (2007). Does compulsory school attendance affect schooling and earnings? revisited.
- Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.
- Berinsky, A. J. and Lenz, G. S. (2011). Education and political participation: Exploring the causal link. *Political Behavior*, 33(3):357–373.
- Blais, A. (2000). *To vote or not to vote?: The merits and limits of rational choice theory*. University of Pittsburgh Pre.
- Bouchard Jr, T. J. and McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of neurobiology*, 54(1):4–45.
- Brody, R. A. (1978). The puzzle of political participation in america. *The new American political system*, pages 287–324.
- Campbell, A., Converse, P. E., Miller, W. E., and Stokes, D. E. (1980). *The american voter*. University of Chicago Press.
- Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., and Zhang, Y. J. (2017). Curriculum and ideology. *Journal of Political Economy*, 125(2):338–392.
- Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling.
- Carpini, M. X. D. (1997). Education and democratic citizenship in america. by norman h. nie, jane junn, and kenneth stehlik-barry. chicago: University of chicago press, 1996. 268p. 16.95 paper. *American Political Science Review*, 91(4):971–972.
- Caspi, A., Roberts, B. W., and Shiner, R. L. (2005). Personality development: Stability and change. *Annu. Rev. Psychol.*, 56:453–484.
- Chevalier, A. (2004). Chevalier, arnaud, parental education and child’s education: A natural experiment (may 2004).
- Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, 88(9):1697–1720.
- Dewey, J. (1923). *Democracy and education: An introduction to the philosophy of education*. macmillan.

- Friedman, M. (1962). Capitalism and freedom. Technical report.
- Galston, W. A. (2001). Political knowledge, political engagement, and civic education. *Annual review of political science*, 4(1):217–234.
- Gidengil, E., Tarkiainen, L., Wass, H., and Martikainen, P. (2019). Turnout and education: Is education proxying for pre-adult experiences within the family? *Political Science Research and Methods*, 7(2):349–365.
- Glaeser, E. L., Ponzetto, G. A., and Shleifer, A. (2007). Why does democracy need education? *Journal of economic growth*, 12(2):77–99.
- Hansen, J. H. (2016). Residential mobility and turnout: The relevance of social costs, timing and education. *Political Behavior*, 38(4):769–791.
- Hess, R. D. and Torney, J. V. (1967). The development of political attitudes in children.
- Jackson, R. A. (1995). Clarifying the relationship between education and turnout. *American Politics Quarterly*, 23(3):279–299.
- Kam, C. D. and Palmer, C. L. (2008). Reconsidering the effects of education on political participation. *The Journal of Politics*, 70(3):612–631.
- Langton, K. P. and Jennings, M. K. (1968). Political socialization and the high school civics curriculum in the united states. *American political science review*, 62(3):852–867.
- Lipset, S. M. (1959). Democracy and working-class authoritarianism. *American Sociological Review*, pages 482–501.
- McDonald, M. P. (2016a). Presidential voter turnout rates, 19789-present. <http://www.electproject.org/national-1789-present>. Accessed: 2022-04-23.
- McDonald, M. P. (2016b). Voter turnout demographics. <http://www.electproject.org/home/voter-turnout/demographics>. Accessed: 2022-04-23.
- Milligan, K., Moretti, E., and Oreopoulos, P. (2004). Does education improve citizenship? evidence from the united states and the united kingdom. *Journal of Public Economics*, 88(9):1667–1695.
- Nie, N. H., Junn, J., Stehlik-Barry, K., et al. (1996). *Education and democratic citizenship in America*. University of Chicago Press.
- Niemi, R. G. and Junn, J. (2005). *Civic education: What makes students learn*. Yale University Press.

- Pelkonen, P. (2012). Length of compulsory education and voter turnout—evidence from a staged reform. *Public Choice*, 150(1):51–75.
- Persson, M. (2014). Testing the relationship between education and political participation using the 1970 british cohort study. *Political Behavior*, 36(4):877–897.
- Persson, M. (2015). Education and political participation. *British Journal of Political Science*, 45(3):689–703.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, 135(2):322.
- Schoen, H. and Schumann, S. (2007). Personality traits, partisan attitudes, and voting behavior. evidence from germany. *Political psychology*, 28(4):471–498.
- Tenn, S. (2007). The effect of education on voter turnout. *Political Analysis*, 15(4):446–464.
- Verba, S. and Nie, N. H. (1987). *Participation in America: Political democracy and social equality*. University of Chicago Press.
- Verba, S., Schlozman, K. L., and Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.
- White, E. S. and Mistry, R. S. (2016). Parent civic beliefs, civic participation, socialization practices, and child civic engagement. *Applied Developmental Science*, 20(1):44–60.
- Wolfinger, R. E. and Rosenstone, S. J. (1980). *Who votes?* Yale University Press.