Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1152

Education and the Gender Gap in U.S. Presidential Elections

Cameron J. Arnzen
Brown University

Sarah R. Cohodes
University of Michigan and NBER

Women in the United States have outpaced men in both voter participation and educational attainment in recent decades. Since education is closely tied to political participation, we consider these trends in tandem and assess how much of the gender gap in voting is attributable to educational attainment as opposed to cognitive skills, noncognitive skills, college preparation behaviors, and demographics. Using comprehensive educational data from Massachusetts students matched with voter records, we estimate a Blinder–Oaxaca–Kitagawa decomposition to understand how these factors contribute to gender voting differentials. In our data, young women outvote young men by 5.8 percentage points in the first presidential election in which they can vote after having potentially completed college. We find that just under 50 percent of this gap in voting can be explained by differences in educational attainment and civic returns to attainment by gender. If we also consider gender differences in college preparation, we explain two-thirds of the gender gap in voting. These results suggest that a significant portion of the gender gap in voting can be attributed to the rise in women's education.

VERSION: September 2025

Suggested citation: Arnzen, Cameron J., and Sarah R. Cohodes. (2025). Education and the Gender Gap in U.S. Presidential Elections. (EdWorkingPaper: 25-1152). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/7md9-7z05

Education and the Gender Gap in U.S. Presidential Elections

Cameron J. Arnzen
Brown University

Sarah R. Cohodes
University of Michigan and NBER

September 2, 2025

Word Count: 9,954 Words (Main Text)

Abstract

Women in the United States have outpaced men in both voter participation and educational attainment in recent decades. Since education is closely tied to political participation, we consider these trends in tandem and assess how much of the gender gap in voting is attributable to educational attainment as opposed to cognitive skills, noncognitive skills, college preparation behaviors, and demographics. Using comprehensive educational data from Massachusetts students matched with voter records, we estimate a Blinder–Oaxaca–Kitagawa decomposition to understand how these factors contribute to gender voting differentials. In our data, young women outvote young men by 5.8 percentage points in the first presidential election in which they can vote after having potentially completed college. We find that just under 50 percent of this gap in voting can be explained by differences in educational attainment and civic returns to attainment by gender. If we also consider gender differences in college preparation, we explain two-thirds of the gender gap in voting. These results suggest that a significant portion of the gender gap in voting can be attributed to the rise in women's education.

Key Words: voting, political participation, gender gaps, educational attainment

---

Author Information: Cameron J. Arnzen (Corresponding Author), Brown University, Providence, RI, USA Email: cameron_arnzen@brown.edu. ORCID: 0000-0001-9076-2655
Sarah R. Cohodes, Gerald R. Ford School of Public Policy, University of Michigan, Ann Arbor, MI, USA. Email: scohodes@umich.edu. ORCID: 0000-0002-7921-7020

Statements and Declarations

Acknowledgments: The authors thank Matthew Deninger, Elana McDermott and the staff of the Massachusetts Department of Elementary and Secondary Education for facilitating the data access that made this project possible. The authors are also grateful for feedback from Jeff Henig, Robert Y. Shapiro, Ansley Erickson, Michael G. Miller, Charlotte Cavaillé, D. Sunshine Hillygus, and participants at the Association for Public Policy Analysis & Management Fall Conference 2022. Special thanks also go to James Feigenbaum, Jacob Brown, and Ryan Enos.
Competing Interests: The authors have no conflicts of interest to declare.
Ethical Research Statement: This study was approved by the Institutional Review Boards at Teachers College Columbia University and University of Michigan.
Data Availability: This paper uses confidential, student-level administrative data from the Massachusetts Department of Elementary and Secondary Education, and the authors are unable to share this publicly. The authors can provide all computer code and instructions on how researchers can apply to access the data.

1 Introduction

Women in the US were disenfranchised for over a century, and even after the passage of the 19th Amendment, their political participation lagged men's for decades. For the past forty years, however, women have outvoted men in every national election (Burns et al., 2018; Cascio and Shenhav, 2020; Stauffer and Fraga, 2022). Despite similar rates of registration, women outvoted men by 3.2 percentage points in the 2024 presidential election, with 91.3 million women (66.9 percent of those eligible) casting a ballot but only 82.6 million men (63.7 percent of those eligible) doing so (Center for American Women and Politics, 2022).

Women's exercise of the franchise is consequential. Initially, opponents of women's suffrage argued that women would simply vote as their husbands did, resulting in no substantive electoral changes. However, in recent decades, women have increasingly identified as liberal while men's ideological identification has remained relatively stable (Saad, 2024). Similarly, women have increasingly gravitated toward the Democratic Party and men toward the Republican Party (Box-Steffensmeier et al., 2004; Gillion et al., 2020). This divergence in ideological and partisan identities is also reflected in gender gaps on a number of issue preferences (Buser et al., 2020; Cascio and Shenhav, 2020; Fernández et al., 2019; Herek, 2002; Ranehill and Weber, 2022). Given these divergent attitudes, women's higher electoral participation rates have the potential to shape American politics.

How did women's voting rates not simply catch up to but surpass men's? The resource model of political participation suggests that individually held stocks of resources such as time, money, and civic skills can explain participation in politics (Brady et al., 1995). This model has also been used to explain the gender gap in voter turnout (Cascio and Shenhav, 2020; Burns et al., 2002; Burns, 2007; Burns et al., 2018; Schlozman et al., 1994; Verba et al., 1997), with the differential voting rates attributed to men's and women's different endowments of participation-enhancing resources. With respect to the gender gaps in voter turnout, this model implies that, prior to the 1980s, women lacked many of the resources that men had to vote (time, power, and access to the polls) and that, over the course of recent decades, women have gained resources that have supported their participation and helped them surpass men in their voting rates. However, women still lag behind men in many of the resources that supposedly matter for civic engagement, such as wages, political interest, and access to power. Thus, the resource model would suggest that changes in some other resources, and/or how they matter for civic engagement, can explain the changes in voter dynamics by gender.

One resource that might help explain some of the rapid rise in women's turnout is educational attainment. In the same period over which higher proportions of women have turned out to vote, women have also outpaced men in educational attainment (Goldin et al., 2006; Goldin and Katz, 2010). As of 2021, 46 percent of women ages 25 to 34 had earned a bachelor's degree, while only 36 percent of men in the same age range had done so (Parker, 2021). These parallel national trends,

in conjunction with the well-documented association between education and political participation (Brady et al., 1995; Nie et al., 1996; Sondheimer and Green, 2010; Wolfinger and Rosenstone, 1980), have led some scholars to conclude that women's voter participation gains should be attributed to their increasing educational attainment (Burns et al., 2018; Cascio and Shenhav, 2020). However, without data that contain both information about individuals' educational backgrounds and their registration and turnout histories, it is difficult to verify exactly how much of the voting gap is attributable to educational attainment as opposed to other individual-level and societal-level factors. We attempt to do so here.

The link between education and political participation is well established (Wolfinger and Rosenstone, 1980; Verba et al., 1995; Nie et al., 1996; Sondheimer and Green, 2010). People with more schooling tend to vote more, so it is likely the case that women's educational gains contribute to their greater presence at the polling booth. However, the literature on political participation widely acknowledges that education plays a central role as both a resource and a means of acquiring other resources (Verba et al., 1978; Brady et al., 1995). Nevertheless, a variety of causal analyses demonstrate that plausibly exogenous increases in educational attainment result in increases in voting (Dee, 2004; Milligan et al., 2004; Sondheimer and Green, 2010; Oreopoulos and Salvanes, 2011). Relatively few studies that do estimate causal impacts go beyond years of schooling to evaluate whether skills gained via education increase voter turnout or whether education provides a gateway to other resources like income and employment or social networks that in turn affect voting behavior.[1]

In this paper, we examine how much of women's higher rates of voting in recent elections is attributable to their higher educational attainment as opposed to other factors. We do so by decomposing the contributions of educational attainment, cognitive and noncognitive abilities, college preparation behavior, and demographics to voter turnout. Using longitudinal data consisting of comprehensive student records of Massachusetts public high-school students matched to voting records, we follow 10th graders in the late 2000's through their high-school and college trajectories and the first presidential election in which they could vote after potential college attendance. In Massachusetts at this time, young women and men register to vote at similar rates, however, women outvote men by 5.8 percentage points, a gender gap slightly larger than the participation gaps in more recent presidential elections.

We build on recent literature on gender gaps in education and political participation. Burns et al. (2018) and Cascio and Shenhav (2020) most directly connect American women's increases in voting participation to their levels of education using surveys and data from the Current Population Survey (CPS). These studies illuminate the parallel national trends in women's education and

[1]Some recent, notable exceptions unpack the effects of education beyond years of schooling. Holbein (2017) finds that a school-based social and emotional learning intervention increases voting. In a study of Democracy Prep, a charter school network focused on civic engagement, Gill et al. (2020) find that an enhanced, civic-focused educational experience boosts students' subsequent electoral turnout. Likewise, Cohodes and Feigenbaum (2021), find significant impacts of Boston charter school attendance on voter turnout.

participation but are unable to connect these trends at the individual level over time. Surveys and the CPS data have the benefit of national coverage, but are limited by self-reporting and a lack of data on individuals' ability. Our Massachusetts data is more limited in context but allows us to directly connect individuals' voting behavior with their prior educational experiences, enabling us to account for rich measures of individual experiences and ability. Our methodological approach is most similar to that of Dassonneville and Kostelka (2021), who use rich, individual-level data from the European Election Studies project to explore gender gaps in voter turnout in subnational European elections.

Our primary focus is documenting education's role in the gender gap in voter turnout, but we also document the relationships between a wide array of education-related factors and voting for young people. In doing so, we also contribute to scholarly debates about whether educational attainment is a *cause* or a *proxy* for other factors that increase voting behaviors (Kam and Palmer, 2008). Put differently, women's higher voting rates may be due to skills gained via education or to resources connected to higher education like earnings and social networks. While we cannot distinguish between every possibility here, a key feature of our linked education-voting dataset is our ability to account for individuals' cognitive and noncognitive skill endowments separately from educational attainment.

One critique of education as a *cause* of voting, is that selection into college is not random (Berinsky and Lenz, 2011; Kam and Palmer, 2008). We cannot fully account for selection in our context, but we still advance this scholarly debate by separately accounting for college preparation via indicators for taking the SAT, completing AP coursework, and completing an optional college-ready curriculum. Though not perfect, these measures, alongside our measures of ability, help us better understand the plausible contribution of educational attainment as opposed to confounding factors associated with selection into higher levels of education (e.g., preparation or ability).

Our Massachusetts sample brings evidence to bear on these debates. By matching students' educational records to their voter records, we find that most student skills, college preparation, and education levels have substantively similar civic returns for women and men. We also find that differences in education levels indeed explain a larger proportion of the gender gap in voter turnout than gender differences in any other characteristics. Additionally, we show that women's advantages in educational attainment can be partly explained by higher rates of college preparation (e.g., taking the SAT) and high-school graduation, however such experiences matter differently for men and women in terms of college attendance. Taken together, this findings suggest that gender differences appear more likely to manifest in the process of accumulating education rather than in the process of education translating into civic behaviors.

This paper proceeds as follows. First, in Section 2, we describe the history of the gender gap in voter turnout and various theories for its emergence. In Section 3, we detail the data we use to investigate the relationship between the gaps in voting and educational attainment, and we explain

the decomposition methods we use. We present the estimates from the linear decomposition model for voter turnout in Section 4, followed by the results from the decomposition of the gender gap in education in Section 5. We conclude in Section 6 by considering the political implications of current educational attainment trends.

2 Gender Gaps in Voter Turnout

Though the 19th Amendment enfranchised white women in 1920, women lagged behind men on all forms of political engagement for many subsequent decades, despite gradual increases in participation (Burns, 2007; Burns et al., 2018). Additionally, women of color were largely prevented from voting until the passage of the Voting Rights Act in 1965 and, in many cases, still faced voter suppression in subsequent years. In national elections, the shares of women and men who vote reached parity only by the 1980 presidential and 1986 midterm elections (Conway, 1991, 2000). In every presidential and midterm election since (Burns, 2007; Burns et al., 2018; Center for American Women and Politics, 2022), women have outvoted men, as shown in the first panel of Figure 1.

Empirical explanations of political participation generally employ the resource model of political participation (Brady et al., 1995): Resources such as income, employment, civic interest, and education—endowments of which have historically favored men—enable political participation. In this model, individually held resources both motivate and equip individuals to participate in politics. For example, if income matters for donating to political campaigns, individuals with higher incomes would be likelier to donate. A variety of additional resources, such as civic skills and interest (Carreras, 2018), verbal reasoning abilities (Hillygus, 2005), and noncognitive skills (Holbein, 2017), have been theorized to complement the resources highlighted by the basic model. Some explanations have focused on how these resources are acquired and then utilized, including through social networks (Schlozman et al., 1999), social norms (Glynn et al., 2009), and other processes of political socialization (Bos et al., 2021). Despite the various explanations for political participation, many questions remain as to *which* resources matter, *how* they matter, and *where* they are acquired.

These puzzles manifest in the relationship between gender and political participation: While women now outvote men, they still trail on other forms of participation such as donating to campaigns and running for office (Burns et al., 2017; Coffé and Bolzendahl, 2010). These gender gaps in political participation also vary by race and ethnicity (Brown, 2014). Men still receive higher incomes, are likelier to hold full-time jobs, and express greater interest in politics—all of which would weigh in favor of men's participation in the resource framework (Brady et al., 1995).

The one resource on which women have gained an advantage is education. In the last decade, the overall share of women with a college degree has overcome that of men, as shown in the top panel of Figure 2. This parity represents the culmination of decades of women outpacing men at lower levels of educational attainment. Since the 1950s, young women have graduated high-school

and attended some form of college at higher rates than men (Goldin et al., 2006). The second panel of Figure 2 shows young women have far outpaced young men in educational attainment. These gains for young women have undermined the decades-long head start men had in educational attainment. Young women began to earn bachelor's degrees at higher rates than young men in the 1980s (Goldin et al., 2006), and since then, their educational attainment has continued on an upward trend. Approximately 44 percent of women aged 25–29 hold a bachelor's degree today, while 35 percent of men do. Thus, along with the changes in women's social and economic standing in recent decades, women's educational gains are the perhaps the likeliest factor within the resource framework to explain the reversal of the gender gap in voter turnout.

Other explanations for the gender gaps in political participation point to civic and socioeconomic resources, personality traits, and the broader political environment in which these factors matter. Regarding civic resources, one of the older explanations of gender deficits in participation is that women participated less out of "apathy" toward politics (Verba et al., 1978). Gender gaps in other forms of political engagement such as contacting elected officials, working for parties, and attending protests have also been attributed to women's lower interest in politics broadly (Carreras, 2018). Attempting to further explain gender gaps in civic skills, work on political socialization argues that politics is a field dominated by men, uninviting to women. Bos et al. (2021) propose the concept of gendered political socialization, whereby children perceive politics to be a male-dominant space, which shapes their own interest in participation. This work also acknowledges gender roles within and outside the home that have shaped women's participation. These explanations tie deficits in participation to children's socialization into gender norms and expectations from a young age (Verba et al., 1997; Burns, 2007; Coffé and Bolzendahl, 2010).

It may also be the case that women express their political interest and participation in different dimensions. Hooghe and Stolle (2004) find that 14-year-olds express similar levels of anticipated participation in politics but prefer different paths of participation: Girls express higher interest in volunteering and canvassing, for example, while boys have higher levels of interest in more radical or direct acts such as running for office. Fridkin and Kenney (2007) find that political preferences in terms of partisanship and ideology can be detected in eighth grade. These nuanced explanations of gendered interest in politics acknowledge the role of participatory resources as well as cultural and gender norms. Overall, these civic skill–related accounts may explain why women less frequently run for office or participate via other political acts such as contacting an elected official or being an active member of a political party (Coffé and Bolzendahl, 2010), but they do not explain the increases in women's voting behavior.

Another strand of explanation focuses on socioeconomic resources. This work emphasizes the importance of socioeconomic factors such as time and money. Schlozman et al. (1999) argue that more men hold full-time jobs while employment is associated with political participation, knowledge, and efficacy. Some of these studies even allege that if women had the same levels of these resources as

men, no gaps would exist (Schlozman et al., 1994). These types of studies often rely on associations derived from cross-sectional survey data of self-reported factors. Nonetheless, men are likelier than women to be employed full-time, earn higher incomes, and sacrifice less professional status for family priorities (Bailey and DiPrete, 2016), suggesting that these resources are unlikely to explain the reversal of the gender gap in voter turnout. Given that men still have higher socioeconomic resources and professional opportunities, it is also unlikely that these factors can explain women's increased voting rates either.

Other contributors to the gender gap in turnout have been offered in recent years. Wang (2014) suggests that gendered psychological differences make a difference in voting and argues that conscientiousness and emotional stability (present at higher rates in women) predict turnout for women but not for men. Carreras (2018) argues that women have a higher sense of "civic duty" than men. These studies both rely on survey data to argue that the gender gap in voter turnout is associated with some skill, ability, or resource that women have more of than men.

While the broader relationship between gender and political participation has many facets (Stauffer and Fraga, 2022; Brown, 2014), accounts of gender differences in civic interest, civic skills, socioeconomic factors, and personality traits fail to fully explain how women typically outvote men, leaving educational gains as a likely explanation for a significant portion of the gap. Relying on a logic quite similar to that of a Blinder–Oaxaca–Kitagawa (BOK) decomposition, Burns, Schlozman, Jardina, Shames, and Verba (2018) apply the logic of the resource model of participation to explore whether changes in *levels* of resources or their *effects* explain the changes in various types of participation. Reviewing trends in turnout alongside trends in resources, the authors theorize that changing education levels are uniquely associated with the disappearance and subsequent reversal of the gender gap in voter turnout. They also explore the explanatory potential of other gendered dynamics, such as the presence of women in politics. The authors conclude that differences in *levels* of resources by gender explain most of the gaps in participation, noting, "the single most important transformation for diminishing the gender difference in political participation has been the reversal of rates of educational attainment among younger cohorts of women and men" (Burns et al., 2018).

We build on this work, and the work of Cascio and Shenhav (2020), by formally testing the importance of gendered educational attainment for voter turnout using individual-level data, documenting how much of the gender gap in turnout can be attributed to educational attainment. We separately account for measures of cognitive and noncognitive skills to consider gendered selection into higher education.

3 Data and Methods

3.1 Data

To better understand how much of the gender gap in voter turnout might be attributable to educational attainment as opposed to cognitive and noncognitive abilities, and demographics, we match detailed information about Massachusetts students with their voting records. The primary data for this study come from two sources: the Massachusetts Department of Elementary and Secondary Education (DESE) and state voter records. DESE provided student information from its Student Information Management System (SIMS), including student names, demographic information, participation in special education, enrollment in a subsided lunch program (free and reduced price lunch), and high-school graduation. DESE also provided information on students' test scores on the Massachusetts Comprehensive Assessment System (MCAS) exams in mathematics and English/language arts,[2] SAT exam and Advanced Placement (AP) course records from the College Board, completion of the MassCore curriculum,[3] and college enrollment and degree information from the National Student Clearinghouse.[4] The sample is limited to Massachusetts students, however our rich data offer a more detailed account of individuals educational experiences and abilities far beyond the measures typically available in nationally-representative surveys such as the Current Population Survey.

To compile the voter turnout data, we combine three Massachusetts voter files from 2012, 2015, and 2018. The voter files come from commercial vendors and include voters' name, date of birth, address, registration date, and participation in general elections. To ensure we capture students who potentially move out of state for college or another reason, we include voter files from the neighboring states of Connecticut, Maine, New Hampshire, New York, Rhode Island, and Vermont in 2018; these files include voter records for presidential elections also collected from commercial vendors. Following similar studies using voter files matched with administrative records (e.g., Akee et al. (2020a); Holbein (2017); Sondheimer and Green (2010), we match students to their voter records by name and date of birth using a combination of exact and fuzzy matching techniques detailed in Appendix Section A.1.

As demonstrated in the match rates in Appendix Table A.1, approximately 83 percent of the students in our sample match to at least one voter record. This means that 17 percent of the students never registered to vote, registered in states not in our data, or are not matched because of inconsistencies across our data sources. However, even in the 2020 election, with its record-high

[2]MCAS scores are standardized to have mean zero and standard deviation one by grade, subject, and year. A small proportion of students do not have MCAS results. For these students, we impute the mean MCAS scores and generate an indicator variable for missing test scores. We include this variable in all analyses but do not report the associated coefficients.

[3]MassCore is a state-recommended program of study that aligns high-school coursework with college and workforce expectations.

[4]Small subsets of students were missing information from the College Board and MassCore completion. For these binary indicators, we code missing data as zeroes.

turnout, only 72.7 percent of the eligible population was registered to vote (Fabina, 2021). Given this, and the high match rate with voter registrations in our data, we are likely to pick up any voting behavior in our sample. One concern regarding the voting data is that we may be less likely to find matching voter records for women because of marriage and subsequent name changes. However, the relative youth of our sample makes this unlikely.[5] Additionally, we employ fuzzy matching techniques to account for minor spelling differences or common nicknames. Individuals who do not match to a voter file are counted as having never voted, as is standard practice in the use of voting records (e.g., Akee et al. (2020b); Ansolabehere and Hersh (2012); Holbein (2017); Holbein et al. (2022); Sondheimer and Green (2010)).

We restrict our sample to the 220,530 students Massachusetts public-school students who were 10th graders in the years 2006, 2007, and 2008. We focus on voter turnout in the 2016 Presidential Election for these students. This strategy lets us follow the students for at least 6 years after projected high-school graduation (8 years after 10th grade), allowing time for the students to have completed a bachelor's degree before the year for which we observe a voting outcome.[6] We focus on "voted in the first presidential election after potential college attendance" to account for whether a student cast a ballot in the first election in which she was eligible to do so after we allow for time for her to have completed her education.

3.2 Descriptive Statistics

Summary statistics for young men and women and the p-values from difference-of-means tests are displayed in Table 1. Each value represents the average for that variable; aside from the MCAS scores and the noncognitive skill index, all the variables are indicator variables and can be interpreted as the proportion of the sample that meets that condition. Men and women are similar in many of their demographic characteristics (Panel A). However, men are likelier than women to be enrolled in special education in 10th grade. We see some divergence in high-school academic experiences (Panel B). Men score considerably lower on the English/language arts (ELA) MCAS exam, in line with previous work that has demonstrated links between verbal reasoning skills and voter turnout (Hillygus, 2005; Nie and Hillygus, 2008). Men also take longer to graduate high-school.

We also observe a difference in noncognitive skills. Our measure of noncognitive skills is modeled after that in Jackson (2018): We construct a noncognitive skill index using the first

[5] If a woman changes her name upon marriage, it may be difficult to match her to a voter file. This to not a major concern in our sample, as the average age at first marriage for women in Massachusetts is 30.1 years, while our primary outcome of interest, whether students vote in the first presidential election occurring after they have potentially completed college, is measured when the respondents are approximately 23 to 26 years old. Any potential undermatching due to women changing their last name to a greater extent than men would also result in our estimation of a smaller gap than actually exists.

[6] We focus on the election most proximate to potential high-school completion to avoid matching issues associated with women getting married and changing their names.

principal component of three variables: 10th-grade suspensions, 10th-grade attendance, and on-time-progression to 11th grade.[7] Researchers often use these behavioral outcomes as proxies for noncognitive skills (Gershenson, 2016; Holbein and Ladd, 2017; Jackson, 2018; Jackson et al., 2020). As Holbein and Hillygus (2020) note, a variety of psychosocial factors, separate from academic achievement, matter for later-life political participation. Notably, girls score approximately a standard deviation higher on this measure than boys, in line with work on the gendered dynamics of noncognitive skills (Jacob, 2002; DiPrete and Jennings, 2012).

We also observe gender difference in college preparation (Panel C). Girls were about 8 percentage points more likely to take an Advanced Placement course and 10.7 percentage points more likely to take the SAT. Girls are also slightly more likely to graduate having completed MassCore curriculum requirements. These disparities reflect wide gender gaps in both interest in pursuing higher education as well as preparation for success in college.

Striking gender differences—in education levels (Panel D) and political participation (Panel E)—emerge later in life. In line with the national trends in educational attainment, the men in our sample tend to be less educated than the women (Figure 2): 13.3 percent of men have no high-school diploma, while this figure only 9 percent for women. Similarly, 18.9 percent of men only have a high-school diploma, while this is the true of only 12 percent of women. The proportions of men and women who completed some college and the proportion who earned an associate's degree are approximately the same. However, a much larger proportion of women than of men hold a bachelor's degree: 42.2 percent of women in the sample hold at least a bachelor's degree, but only 30.5 percent of men do, as measured 6 years after expected high-school graduation.

For voting outcomes (Panel E), there is little gender difference in voter registration: 83 percent of both genders registered to vote in at least one state at some point in our data. However, when it comes to actually casting a ballot, women are likelier to follow through. In our primary outcome of interest, whether an individual voted in the first presidential election after we allow sufficient time for college completion, women outvote men by 5.8 percentage points.

Our first step in understanding the factors that contribute to the gender gap in voter turnout is estimating models of voting as a function of demographics, educational experiences, and educational attainment to determine how these characteristics contribute to voting behavior. Table 2 illustrates, without differentiating by gender, the role that these factors play in predicting voter turnout in the first presidential election after individuals have had a chance to attend and complete college.

For each of the models, the gender gap in voter turnout is captured by the coefficient for *Female* in the top row. We begin by using only demographic characteristics and participation in school-based programs to explain voting (Column 1) and progressively add academic experiences as summarized by test scores and the noncognitive skill index (Column 2) and educational attainment (Column 3). Column 4 adds high-school fixed effects, which account for the schools that individuals

[7] Jackson (2018) also includes GPA, which we omit due to data differences.

attended and absorb any possible heterogeneous effects for different types of schools.

In this regression taxonomy, a reduction in the coefficient on the indicator for *Female* when we add a variable indicates that the factor has predictive power for gender differences in voter turnout. Model 1 shows a gender gap of approximately 5.4 percentage points when we include only demographics, which indicates that little of the 5.8-percentage-point difference in voting by gender observed in our data can be estimated by gendered differences in student characteristics. Adding student skills (cognitive and noncognitive abilities) reduces the coefficient on *Female* to approximately 4.4 percentage points, indicating that these variables have some explanatory power. Model 3, which includes variables for college preparation factors, as measures of aptitude and interest. The addition of these variables reduces the estimated gender gap by 1 percentage point, comparable to student skills. The addition of education levels in Model 4 reduces the explanatory power of gender in our models to 2.2 percentage points, confirming the importance of educational attainment. Adding high-school fixed effects in Model 5 yields remarkably similar estimates to those in Model 4, indicating that these patterns play out within schools, not across them. Thus, we move forward with a variation of Model 4 to decompose the contribution of different factors to the gender voting gap. We now turn to the methods we use connect these gaps in voting to the gaps we observe in education levels, college preparation, and student skills.

3.3 BOK Decomposition

We use a BOK-style decomposition to separate the factors that contribute to the gender differences in voter turnout. Such models are often used in economics and political science to "decompose" relationships between variables by group. Most famously, this decomposition method has been applied to labor-market outcomes to understand group differences in earnings and employment (Blinder, 1973; Oaxaca, 1973; Kitagawa, 1955). This approach has been used to similarly explore gender gaps in political participation. For example, (Dassonneville and Kostelka, 2021) decomposes the gender gap in turnout for European Parliament elections, finding that gender differences in political interest explain a large portion of the gap. Other decompositions on gender gaps in political participation have focused on political skills and turnout in other European contexts (Dow, 2009; Kostelka et al., 2019). In education, BOK decompositions have been used to explore gender gaps in college attendance (Jacob, 2002), primary-school achievement (Golsteyn and Schils, 2014), and GPA distributions (Fortin et al., 2015). In political science, in addition to explaining gender voting gaps as described above, these models have been used to explain partisan differences in evaluations of inflation (Bachmann et al., 2021) and the amount of political polarization due to changes in voter positioning versus politician positioning (Kertzer, 2020).

BOK decompositions approximate counterfactuals across specified group differences such as race, gender, or party identification. Estimated coefficients for one group from linear regression models are applied to the characteristics of another group to "decompose" gaps in outcomes,

indicating how much of the gap would disappear if both groups had the same characteristics and the same response to those characteristics. The observed gaps are thus broken down into the portions explicable by the measurable differences in characteristics and by the differential response to those characteristics and the portion that remains unexplained.

The data employed here are well suited for a BOK decomposition of the gender gap in voter turnout. First, our longitudinal data include a variety of education-specific factors at an individual level over time. Second, we match these educational records with voting records for two cohorts of students from the entire state of Massachusetts, generating data for 220,530 individuals matched with voting records, which yields a large dataset. However, it is important to note that BOK decompositions are not causal estimates, as the factors that explain a difference in groups may be causal factors or reflect an association with a related variable. Given that we are interested in the relative importance of educational attainment, student skills, college preparation, and demographics for the gender gap in voter turnout, a BOK decomposition is well-suited for our inquiry.

We present a threefold BOK decomposition of the drivers of the gender gap in voter turnout into differences in the educational characteristics of men and women, differences in how men and women respond to educational experiences, or an unaccounted-for factor. BOK decompositions quantify the gap that results from estimating a model separately for two groups. Thus, we first explore whether significant gender differences exist in the determinants of voting for women and men by estimating the linear probability model below separately for men and women. We model voting as a function of some combination of observed and unobserved factors:

$$V_{ig} = X'_{ig}\beta_g + \varepsilon_{ig}, \tag{1}$$

where i represents the individual and g the gender group. voter turnout is reflected in V_{ig}, the primary outcome of interest: whether an individual votes in the first presidential election after time is allowed for college graduation.[8] A vector of potential determinants of voting, X_{ig}, includes educational attainment levels, school experiences as summarized by test scores and noncognitive skills, and student demographics, while ε_{ig} represents the error term: unobserved factors that contribute to voting.

The general logic of decompositions is that we can break down the gap by estimating the models separately by group. We thus make use of the coefficients from the estimates for each gender from Equation 1, $\hat{\beta}_f$ and $\hat{\beta}_m$, with f and m denoting the group outcomes, characteristics, and coefficients for women and men separately. Thus, the difference in voter participation by gender is:

$$\bar{V}_f \; - \; \bar{V}_m \;\; = \;\; \bar{X}'_f \; \hat{\beta}_f \; - \; \bar{X}'_m \; \hat{\beta}_m \; . \tag{2}$$

We can further decompose the gender gap in voter turnout into three components: the variation

[8]For our BOK linear decomposition, we follow the general empirical framework from Jacob (2002) and the threefold structure from Hlavac (2014).

explicable by 1) differences in observable characteristics and 2) the characteristic's importance for the outcome of voting and 3) the remaining proportion left unexplained:

$$\bar{V}_f - \bar{V}_m = \underbrace{(\bar{X}_f - \bar{X}_m)'\hat{\beta}_m}_{\text{Endowments}} + \underbrace{\bar{X}'_m\,(\hat{\beta}_f - \hat{\beta}_m)}_{\text{Coefficients}} + \underbrace{(\bar{X}_f - \bar{X}_m)'(\hat{\beta}_f - \hat{\beta}_m)}_{\text{Interaction}} \tag{3}$$

The endowments term reflects the gender differences in voter turnout due to the differences in characteristics between men and women, such as education levels, student skills, college preparation, and demographics. The coefficients term accounts for different "returns" to these characteristics by gender group. While in a twofold decomposition, the returns to characteristics are assumed to be the same across groups, in a threefold decomposition, these returns can vary, as some of the explanatory factors of interest in our data appear to. In our case, noncognitive skills, for example, may matter more for women's voting than for men's. Similarly, we can explore the relative importance of college completion, accounting for gender difference in college preparation and interest. Last, the interaction term accounts for the cross-group differences in coefficients and endowments occurring concurrently.

4 Results: Education's Role in the Gender Gap in Voter Turnout

We first consider whether there are differences by gender in the linear probability model, as in Equation 1. Table 3 reports the estimated coefficients, which can be interpreted as percentage point increases in an individual's likelihood of voting, from our selected model separately by gender. The coefficient estimates for each gender are substantively similar for most variables other than race. Black and Hispanic women are likelier to vote than their white peers when we account for the other determinants in the models, while Asian women are less likely to vote than white women. However, men of all minority groups have a lower probability of voting than their white counterparts. These estimates affirm the importance of focusing intersectional identities while studying voter turnout gaps (Brown, 2014; Stauffer and Fraga, 2022), though they largely mirror broader participation trends in in American elections (Center, 2025).

Commonalities between the two models show that both women and men who receive free or reduced price lunch in 10th grade are less likely to vote than their peers not in this program, though this matters slightly more for men. Similarly, English language learners are less likely to vote than their peers, though this matters slightly more for women. Students with 10th-grade MCAS ELA scores one standard deviation higher than the mean have an approximately 1.8-percentage-point higher probability of voting among both women (2.0) and men (1.6). Both genders experience substantively similar returns to math scores as well. Though the difference between the returns for noncognitive skills is statistically significant, the gap amounts to only a 0.5 percentage point difference in the likelihood of voting across genders.

Despite the gender differences in college preparation for men and women illustrated in Table 1, the civic returns to these experiences appear similar across gender. Though women have a slightly higher return to taking an AP course, taking the SAT, and completing the MassCore curriculum (women have about a 1 percentage point advantage for each), these differences are not statistically significant.

Most informatively for our purposes, the coefficients for educational attainment levels are broadly similar across gender. Compared to individuals with a high-school diploma (the excluded category), people of either gender without a high-school degree are less likely to vote, though this is not a statistically significant relationship for women. Conversely, attainment of each additional level of education is associated with a significant, positive impact on the likelihood of voting—generally of similar magnitude—for both men and women. Compared to those with a high-school diploma, men with no diploma a 5.8-percentage-point decrease in voting likelihood. Attending some college without obtaining a degree predicts a 8.8-percentage-point increase in voting likelihood for women and a 7.1-percentage-point increase for men. Earning an associate's degree is associated with a 16.8-percentage-point increase in voting likelihood for women and a parallel 17.2-percentage-point increase for men over the likelihood of their peers with only a high-school diploma. Earning a bachelor's degree or more is associated with an increase of approximately 19.9 percentage points in the probability of voting for women and of 15.3 percentage points for men. Completing a bachelor's degree or higher has the largest differential return by gender (a 4.6-percentage-point higher return for women), though the coefficient for both women and men is the largest of predictors included. Taken together, the variables for educational attainment explain significant portions of voting likelihood for both women and men, even after accounting for demographics, student skills, and college preparation.

Differences exist between the coefficients for women and men, suggesting that some factors matter more for one gender than the other in terms of how they predict later-life voting. In the case of demographics, these coefficients likely account for how the broader political environment invites participation differently for individuals based on race and gender (See Slaughter et al. (2024) for example in the case of Black Women's unmatched engagement). In the case of skills and experiences, however, differences can illuminate barriers. For example, while women score higher on the noncognitive skills index (Table 1), these skills appear to matter more for men than women (though only slightly). In the case of educational attainment, gendered differences appear for some college and for bachelor's degrees. These differences suggest that a threefold decomposition, which allows endowments to vary in their importance by gender (returns) is ideal.

The broad similarities by gender in the predictive power of educational attainment and sign of the effects suggest that levels of education matter more for voting than do gendered differences in the returns to education. We next turn to the BOK decomposition to formalize this intuition. Table 4 breaks down the components of the gender gap in voter turnout estimated in Equation 3.

Approximately 43.8 percent of women voted in the first election after enough time had elapsed for them to have attended college, while this figure is 38 percent for men, resulting in a 5.8-percentage-point gap in voter turnout by gender. This estimate is slightly larger than the national gap in Figure 1, which may reflect differences in age.

Table 4 reports the results of the decomposition of the gender gap in voter turnout grouped by category, while Figure 3 details the specific contributions of individual characteristics and their returns separately by endowment (characteristic) and coefficient (returns). In Figure 3, bars to the right of 0 increase the gender gap in voting (favor women), while bars to the left decrease the gender gap (favor men). We find that 57 percent of the 5.8-percentage-point gender gap (approximately 3.3 percentage points) can be explained by differences in endowments by gender while differential returns account for an additional 32 percent of the gap and the interaction of the two the final 11 percent.

Little of the gender gap is due to differences in demographics (in terms of endowments). Women and men have similar demographics except for participation in special education—women are just over 8 percentage points less likely to receive special education services (Table 1). However, once other variables are accounted for, special education students vote at higher rates (Tables 3 and 4). Thus, women as a group have demographics that make them slightly less likely to vote.

Differences in endowments of cognitive and noncognitive skills together account for 8 percent of the total gap. Their dynamics follow a logic similar to that of the role of special education, albeit with a contrasting impact: Women outscore men on ELA and noncognitive skills (Table 1), and thus, women's educational experiences in high-school make them more likely to vote, though these factors account for a relatively small share of the overall gap.

Gender differences in college preparation endowments account for roughly 15 percent of the gender gap in voter turnout. As illustrated in Table 1, women are more likely to have taken an AP course and to have taken the SAT. This difference likely represents both differences in interest and preparation for college, as well as a variety of economic, social, and political factors that shape educational choices and similarly contribute to voting.

Educational attainment is the most important endowment factor in explaining the gender gap. Differences in educational attainment endowments between women and men are responsible for approximately 36 percent of the total gender gap in voter turnout. This confirms the hypothesis in Burns et al. (2018) and Cascio and Shenhav (2020) that the key component of the rise in women's exercise of the franchise is the growth in women's educational attainment. Figure 3 further illustrates that much of the estimate for educational attainment is due to women's higher rates of earning a bachelor's degree. Indeed, the endowment estimate for a bachelor's degree or more is the largest of the endowment estimates.[9]

[9]The Presidential Election of 2016 was an election with high interest and an unusual campaign. To assess whether our results are driven by this particular election, we conduct a robustness check with a larger data set which includes an additional panel of students who attended 10th grade in 2002, 2003, and 2004 matched with voting in the 2012

Additionally, given the differences in coefficients for women and men in Table 3, we use a threefold BOK decomposition, which also estimates the amount of the gender gap in voter turnout that can be explained by different returns to the characteristics of each group. In other words, this model allows factors to matter differently such a educational attainment to matter differently for men and women. This is the "Coefficients" section of 4 and the lower half of Figure 3. As with endowments, bars to the left of 0 theoretically shrink the gender gap (favor men), while bars to the right increase the gender gap in voter turnout (favoring women). Our model estimates that gendered differences in the relationship between some explanatory factors and voting account for approximately 32 percent of the voter turnout gap. Specifically, the coefficients for student demographics matter most. The positive share of the gap for student demographics is due to the fact that Black and Hispanic women far outvote Black and Hispanic men as seen in Figure 3. These larger gender gaps align with modern trends in American elections (Center, 2025). Gendered returns to cognitive skills, noncognitive skills, and college preparation matter little for the gender gap in voter turnout, though returns to educational attainment do matter slightly more for women than for men.

Also reported here is the coefficient for the intercept term, which accounts for the *unexplained* factors contributing to the gender voter turnout gap (both endowments and coefficients). The negative coefficient implies that factors unexplained in our data make it *less* likely for women to vote. This is consistent with the idea that men still benefit from greater resources on several dimensions—but that the differential educational attainment endowment and returns to race counteract and offset those forces.

By carrying forward women's endowment and return estimates for educational attainment and applying them to men in Panel A of Figure 4, we illustrate how much of the gender gap would disappear if men were to achieve the same level of and returns to education as women.[10] The shaded region illustrates that if men had education levels similar to women's, and that education translated to participation at in the same manner, the voter turnout gap between the genders would be halved.

Additionally, because our measures of college preparation are likely proxies for interest in college (e.g., students are more likely to take the SAT if they plan to attend college), we estimate how much of the gender gap would remain if women and men had similar college preparation *and* educational attainment levels and returns. As Panel B of Figure 4 shows, holding all else constant, the gender gap in voter turnout would be reduced by by nearly two-thirds.

Presidential Election to allow for college completion in Appendix Table A.3, Appendix Table A.4 and Appendix Table A.5. While this provides a larger sample across two elections, we lack the additional cohort's measures for SAT-taking, AP Course enrollment, or MassCore curriculum completion. For this sample, the portion of the gap explained by educational attainment absorbs the gender difference in college preparation, explaining 51 percent of the gender gap in voting. All other decomposition results remain similar.

[10]This is an approximation based on our data, so it does not account for the fact that education may have mattered differently in the 1980s, for example.

Figure 4 shows that, even if men had the same education levels and preparation as women, the gap would not entirely disappear but would significantly shrink. Other factors such as the racial dynamics we discuss here still contribute to women's greater presence at the polls. However, educational attainment is responsible for a significant portion of the gap. Further, educational attainment can be shaped through policies that increase college-going, such as emphasizing affordability for broad audiences (Hyman, 2020), demystifying college costs for low income students (Dynarski et al., 2021), or direct admissions policies (Delaney et al., 2023). Given that women are increasingly likelier to hold a bachelor's degree than men, we can reasonably expect the gender gap in voter turnout to continue to grow absent some intervention for men or major change in the determinants of voting. The importance of educational attainment for the gendered differences in voting leads us to next consider the determinants of college-going.

5 Gendered Differences in College-Going

Given that different levels of educational attainment by gender explain approximately a third of the gender gap in voter turnout, we investigate the roots of the gender gap in educational attainment. The bottom panel of Figure 2 underscores the importance of this work: Young women are approximately 10 percentage points likelier than young men to have a bachelor's degree—a gap that is growing and will continue to reshape the educational differences for the general population. In this Massachusetts sample, 79 percent of girls pursued at least some college, while 68 percent of boys did so, in a pattern that mirrors that in Figure 2.

Massachusetts offers a unique case for which to examine educational attainment: 46.6 percent of individuals 25 or older hold at least a bachelor's degree, compared to 35 percent of the United States (Census, 2021). Massachusetts also consistently scores as the highest state in most educational rankings. However, even in a state with a robust education system and a well-educated population such as Massachusetts, young women are still far outpacing men.

We attempt to illuminate the roots of the gender voter turnout gap by conducting an exercise similar to the one above for voter turnout, this time decomposing the gap in college attendance using the pre-college student-level measures available to us in our data: cognitive skills, noncognitive skills, college preparation, and demographics. Table 5 details the summary results of this decomposition, while Figure 5 illustrates the detailed results. Similar to our findings for voting, different endowments across gender explain approximately 60 percent of the gender gap in college attendance. Of the portion that can be explained, gender differences in endowments of college preparation and rates of high-school completion are responsible for nearly all of the explained variation in the gap. Figure 5 shows that women's higher ELA scores, rates of taking the SAT, and high-school graduation rates advantage them relative to men in terms of college attendance. Demographics and noncognitive skills play minimal roles. These results imply that college-going differences by gender are not immutable and greater preparation for boys in high-school, such as

policies that support boys' high-school completion and/or increase their interest in college, could increase their college-going rates to some degree.

The coefficients portion of Table 5 reflects differential returns by gender. Demographics, cognitive skills, and noncognitive skills all appear to operate similarly for women and men in terms of college attendance. However, college preparation is far more important for college attendance for men than women, actually shrinking the gender gap in college attendance by roughly 50 percent. Put simply, men appear to only prepare for college if they plan to attend college. Figure 5 shows that, for example, taking the SAT predicts college attendance for men by 26 percentage points compared to only 16 percentage points for women. While women are more likely to take the SAT in the first place (Table 1), it is a stronger signal of college attendance for men. Conversely, graduating high-school has greater returns for women: women are both more likely to graduate high-school and have greater returns to high-school graduation in term of converting high-school graduation to college attendance. However, roughly 9 percent of the explanation for college attendance remains outside the scope of the academic, demographic, and socioeconomic factors present in our data.

A variety of potential explanations for these divergent findings could include gender norms around schooling, gendered expectations about career choice, gendered opportunity costs of attending college, and gendered returns to degree holding. Our analyses are limited to administrative data on student- and school-level measures, in contrast to the analyses by Jacob (2002), who uses survey data to also find that the opportunity costs of college attendance, gendered expectations for employment, and noncognitive abilities are important factors in the gender gap in college attendance, though the magnitude of gap he decomposed several decades ago was less than half that of the gap we find in our data.

We have shown that educational attainment accounts for a significant portion of the gender gap in voter turnout, and we have shed light on some of the forces behind the gap in educational attainment. Girls' better performance in high-school plays a role, but so do other factors. Since men have similar civic returns to educational attainment, inducing them to attend college at rates similar to women's would likely improve their voter turnout. However, we cannot say exactly what must change for young men to be induced to attend college at higher rates. Our analyses point to gendered interest in higher education (commitment to taking the SAT and completing high-school) as a likely factor.

6 Conclusion

Women's educational gains in recent decades have been hypothesized to be the cause of the gender gap in voter turnout (Burns et al., 2018; Cascio and Shenhav, 2020). These explanations are in line with the well-documented finding in the United States that education predicts voting. With data directly connecting educational attainment, academic performance, noncognitive measures, college preparation, and voting, we are able to confirm that a third of the gender gap in voter turnout is

indeed explained by women's higher levels of education alone. However, education also has greater returns for women's later-life voting than for men's. Taking both into account explains about half of the gender gap in voting—far more than any other factor in our data. This finding broadly suggests that education has played a pivotal role in women's greater presence at the ballot box in recent decades.

Our findings also show that differences by race—the higher turnout of Black and Hispanic women in contrast to fewer differences by race for men—contribute to the gender gap in voter turnout. This is consistent with the finding of Brown (2014) that Black women have different historical relationships with political parties, shaping their participation differently (Slaughter et al., 2024).

Our findings broadly suggest that educational attainment, while the most significant factor contributing to gender gap in voter turnout, is far from the only factor. Some of gender gap in voting is due to unexplained factors that make it *less* likely for women to vote. These factors may include gendered social, political, and economic environments that differently shape voting norms. Future work should further investigate why some factors matter different by gender. While differences in returns to demographics likely signal broader gendered political environments, differences in returns to noncognitive skills and educational attainment raise several questions about how these skills translate into action.

Nevertheless, we provide credible estimates of how much educational attainment—as opposed to a variety of other factors—contributes to the contemporary gender gap in voter turnout, providing direct evidence to support the work of Burns et al. (2018) and Cascio and Shenhav (2020). However, though we offer compelling evidence that educational attainment is the largest factor contributing to the gap in voter turnout, the gender gap in voting is preceded by the gender gap in college attendance. Approximately three-quarters of what induces girls to attend college at greater rates than men remains unexplained by our comprehensive administrative data.

While women have reversed the gender gaps in both voter turnout and educational attainment, women still trail men in other forms of political participation such as donating to campaigns and running for office (Burns et al., 2017). The persistence of gaps in other forms of participation shed light on the unique relationship between educational attainment and the act of voting. Further research should explore the extent to which educational attainment shapes other forms of political engagement, as well as additional contexts.

These findings contribute to our broader understanding of the relationship between education and voting. Gender matters, but in ways that likely shape both educational attainment and, in turn, voting behavior. Higher proportions of women show up at polling booths and in college classes, with this divergence in behavior showing no sign of disappearing absent some broader intervention for men. Despite over a century of disenfranchisement in conjunction with persistent economic, social, and political discrimination, women have made steady educational and civic gains in recent decades. In light of the trends in Figure 2, our findings lead us to believe that, unless the gender

gap in educational attainment begins to shrink, we can reasonably expect women to continue to outvote men for the foreseeable future.

References

Akee, R., W. Copeland, J. B. Holbein, and E. Simeonova (2020a). Human capital and voting behavior across generations: Evidence from an income intervention. *American Political Science Review 114*(2), 609–616.

Akee, R., W. Copeland, J. B. Holbein, and E. Simeonova (2020b). Human capital and voting behavior across generations: Evidence from an income intervention. *American Political Science Review 114*(2), 609–616.

Ansolabehere, S. and E. Hersh (2012). Validation: What big data reveal about survey misreporting and the real electorate. *Political Analysis 20*(4), 437–459.

Bachmann, O., K. Gründler, N. Potrafke, and R. Seiberlich (2021). Partisan bias in inflation expectations. *Public Choice 186*(3), 513–536.

Bailey, M. J. and T. A. DiPrete (2016). Five decades of remarkable but slowing change in us women's economic and social status and political participation. *RSF: The Russell Sage Foundation Journal of the Social Sciences 2*(4), 1–32.

Berinsky, A. J. and G. S. Lenz (2011). Education and political participation: Exploring the causal link. *Political Behavior 33*(3), 357–373.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436–455.

Bos, A. L., J. S. Greenlee, M. R. Holman, Z. M. Oxley, and J. C. Lay (2021). This one's for the boys: How gendered political socialization limits girls' political ambition and interest. *American Political Science Review*, 1–18.

Box-Steffensmeier, J. M., S. De Boef, and T.-M. Lin (2004). The Dynamics of the Partisan Gender Gap. *American Political Science Review 98*(3), 515–528.

Brady, H. E., S. Verba, and K. L. Schlozman (1995). Beyond ses: A resource model of political participation. *American Political Science Review*, 271–294.

Brown, N. E. (2014). Political participation of women of color: An intersectional analysis. *Journal of Women, Politics & Policy 35*(4), 315–348.

Burns, N. (2007). Gender in the aggregate, gender in the individual, gender and political action. *Politics & Gender 3*(1), 104–124.

Burns, N., A. Jardina, and N. Yadon (2017). Women as a force in electoral politics. In *The Oxford Handbook of US Women's Social Movement Activism*.

Burns, N., K. L. Schlozman, A. Jardina, S. Shames, and S. Verba (2018). What's happened to the gender gap in political participation?: How might we explain it? In *100 Years of the Nineteenth Amendment: An Appraisal of Women's Political Activism*, pp. 69–104. Oxford University Press.

Burns, N., K. L. Schlozman, and S. Verba (2002). *The Private Roots of Public Action: Gender, Equality, and Political Participation*. Harvard University Press.

Buser, T., G. Grimalda, L. Putterman, and J. van der Weele (2020). Overconfidence and gender gaps in redistributive preferences: Cross-country experimental evidence. *Journal of Economic Behavior & Organization 178*, 267–286.

Carreras, M. (2018). Why no gender gap in electoral participation? a civic duty explanation. *Electoral Studies 52*, 36–45.

Cascio, E. U. and N. Shenhav (2020). A century of the american woman voter: Sex gaps in political participation, preferences, and partisanship since women's enfranchisement. *Journal of Economic Perspectives 34*(2), 24–48.

Census, U. (2021). American community survey. Technical report, United States Census.

Center, P. R. (2025). Behindtrump's 2024 victory, a more racially andethnically diverse voter coalition. Technical report, Pew Research Center.

Center for American Women and Politics (2022). Gender differences in voter turnout.

Coffé, H. and C. Bolzendahl (2010). Same game, different rules? gender differences in political participation. *Sex roles 62*, 318–333.

Cohodes, S. and J. J. Feigenbaum (2021). Why does education increase voting? evidence from boston's charter schools. Technical report, National Bureau of Economic Research.

Conway, M. M. (1991). The study of political participation: Past, present. *Political Science: Political behavior 3*, 31.

Conway, M. M. (2000). Gender and political participation. *Gender and American Politics*, 74–87.

Dassonneville, R. and F. Kostelka (2021). The cultural sources of the gender gap in voter turnout. *British Journal of Political Science 51*(3), 1040–1061.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics 88*(9-10), 1697–1720.

Delaney, J. A., T. M. Trinh, and T. K. Odle (2023). Direct admissions: Policies and principles. *College and University 98*(3), 75–80.

DiPrete, T. A. and J. L. Jennings (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research 41*(1), 1–15.

Dow, J. K. (2009). Gender differences in political knowledge: Distinguishing characteristics-based and returns-based differences. *Political Behavior 31*(1), 117–136.

Dynarski, S., C. J. Libassi, K. Michelmore, and S. Owen (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review 111*(6), 1721–1756.

Fabina, J. (2021). Despite pandemic challenges, 2020 election had largest increase in voting between presidential elections on record. *US Census Bureau 29*.

Fernández, R., S. Parsa, and M. Viarengo (2019). Coming Out in America: AIDS, Politics, and Cultural Change. Technical report, National Bureau of Economic Research.

Fortin, N. M., P. Oreopoulos, and S. Phipps (2015). Leaving boys behind gender disparities in high academic achievement. *Journal of Human Resources 50*(3), 549–579.

Fridkin, K. L. and P. J. Kenney (2007). Examining the gender gap in children's attitudes toward politics. *Sex Roles 56*(3), 133–140.

Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy 11*(2), 125–149.

Gill, B., E. R. Whitesell, S. P. Corcoran, C. Tilley, M. Finucane, and L. Potamites (2020). Can charter schools boost civic participation? The impact of Democracy Prep Public Schools on voting behavior. *American Political Science Review 114*(4), 1386–1392.

Gillion, D. Q., J. M. Ladd, and M. Meredith (2020). Party Polarization, Ideological Sorting and the Emergence of the US Partisan Gender Gap. *British Journal of Political Science 50*(4), 1217–1243.

Glynn, C. J., M. E. Huge, and C. A. Lunney (2009). The influence of perceived social norms on college students' intention to vote. *Political Communication 26*(1), 48–64.

Goldin, C. and L. F. Katz (2010). *The race between education and technology*. Harvard University Press.

Goldin, C., L. F. Katz, and I. Kuziemko (2006). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic perspectives 20*(4), 133–156.

Golsteyn, B. H. and T. Schils (2014). Gender gaps in primary school achievement: a decomposition into endowments and returns to iq and non-cognitive factors. *Economics of Education Review 41*, 176–187.

Herek, G. M. (2002). Gender Gaps in Public Opinion About Lesbians and Gay Men. *Public Opinion Quarterly 66*(1), 40–66.

Hillygus, D. S. (2005). The missing link: Exploring the relationship between higher education and political engagement. *Political behavior 27*(1), 25–47.

Hlavac, M. (2014). oaxaca: Blinder-oaxaca decomposition in r. *Available at SSRN 2528391*.

Holbein, J. B. (2017). Childhood skill development and adult political participation. *American Political Science Review 111*(3), 572–583.

Holbein, J. B., C. P. Bradshaw, B. K. Munis, J. Rabinowitz, and N. S. Ialongo (2022). Promoting voter turnout: An unanticipated impact of early-childhood preventive interventions. *Prevention Science*, 1–12.

Holbein, J. B. and D. S. Hillygus (2020). *Making young voters: Converting civic attitudes into civic action*. Cambridge University Press.

Holbein, J. B. and H. F. Ladd (2017). Accountability pressure: Regression discontinuity estimates of how No Child Left Behind influenced student behavior. *Economics of Education Review 58*, 55–67.

Hooghe, M. and D. Stolle (2004). Good girls go to the polling booth, bad boys go everywhere: Gender differences in anticipated political participation among american fourteen-year-olds. *Women & Politics 26*(3-4), 1–23.

Hyman, J. (2020). Can light-touch college-going interventions make a difference? evidence from a statewide experiment in michigan. *Journal of Policy Analysis and Management 39*(1), 159–190.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy 126*(5), 2072–2107.

Jackson, C. K., S. C. Porter, J. Q. Easton, A. Blanchard, and S. Kiguel (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American Economic Review: Insights 2*(4), 491–508.

Jacob, B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education review 21*(6), 589–598.

Kam, C. D. and C. L. Palmer (2008). Reconsidering the effects of education on political participation. *The Journal of Politics 70*(3), 612–631.

Kertzer, J. D. (2020). Re-assessing elite-public gaps in political behavior. *American Journal of Political Science*.

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the american statistical association 50*(272), 1168–1194.

Kostelka, F., A. Blais, and E. Gidengil (2019). Has the gender gap in voter turnout really disappeared? *West European Politics 42*(3), 437–463.

Milligan, K., E. Moretti, and P. Oreopoulos (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics 88*(9-10), 1667–1695.

Nie, N. and D. S. Hillygus (2008). Education and democratic citizenship. In *Making Good Citizens: Education and Civil Society.*, pp. 30–57. Yale University Press.

Nie, N. H., J. Junn, and K. Stehlik-Barry (1996). *Education and democratic citizenship in America*. University of Chicago Press.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.

Oreopoulos, P. and K. G. Salvanes (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives 25*(1), 159–84.

Parker, K. (2021). What's behind the growing gap between men and women in college completion?

Ranehill, E. and R. Weber (2022). Gender Preference Gaps and Voting for Redistribution. *Experimental Economics 25*, 845–875.

Saad, L. (2024). U.S. Women Have Become More Liberal; Men Mostly Stable. Technical report, Gallup.

Schlozman, K. L., N. Burns, and S. Verba (1994). Gender and the pathways to participation: The role of resources. *the Journal of Politics 56*(4), 963–990.

Schlozman, K. L., N. Burns, and S. Verba (1999). " what happened at work today?": A multistage model of gender, employment, and political participation. *The Journal of Politics 61*(1), 29–53.

Slaughter, C., C. Crowder, and C. Greer (2024). Black women: Keepers of democracy, the democratic process, and the democratic party. *Politics & Gender 20*(1), 162–181.

Sondheimer, R. M. and D. P. Green (2010). Using experiments to estimate the effects of education on voter turnout. *American Journal of Political Science 54*(1), 174–189.

Stauffer, K. E. and B. L. Fraga (2022). Contextualizing the gender gap in voter turnout. *Politics, Groups, and Identities 10*(2), 334–341.

Verba, S., N. Burns, and K. L. Schlozman (1997). Knowing and caring about politics: Gender and political engagement. *The journal of politics 59*(4), 1051–1072.

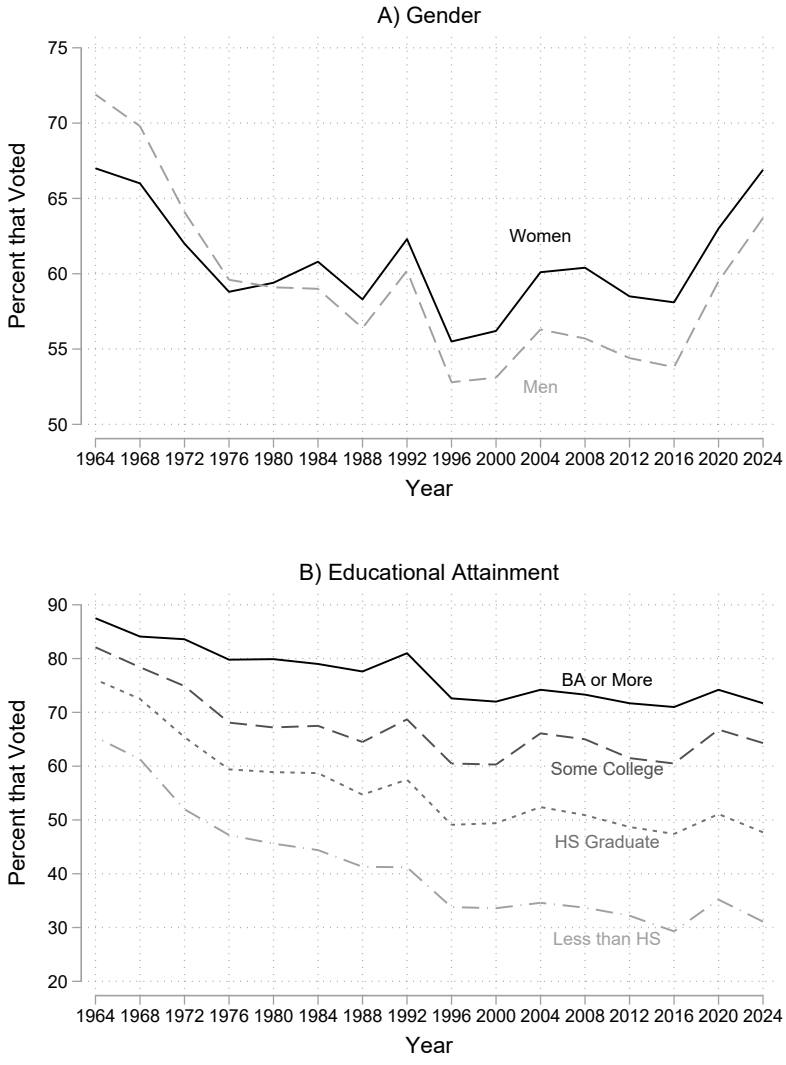
Verba, S., N. H. Nie, and J.-o. Kim (1978). *Participation and political equality: A seven-nation comparison*. Cambridge university press.

Verba, S., K. L. Schlozman, and H. E. Brady (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Wang, C.-H. (2014). Gender differences in the effects of personality traits on voter turnout. *Electoral Studies 34*, 167–176.

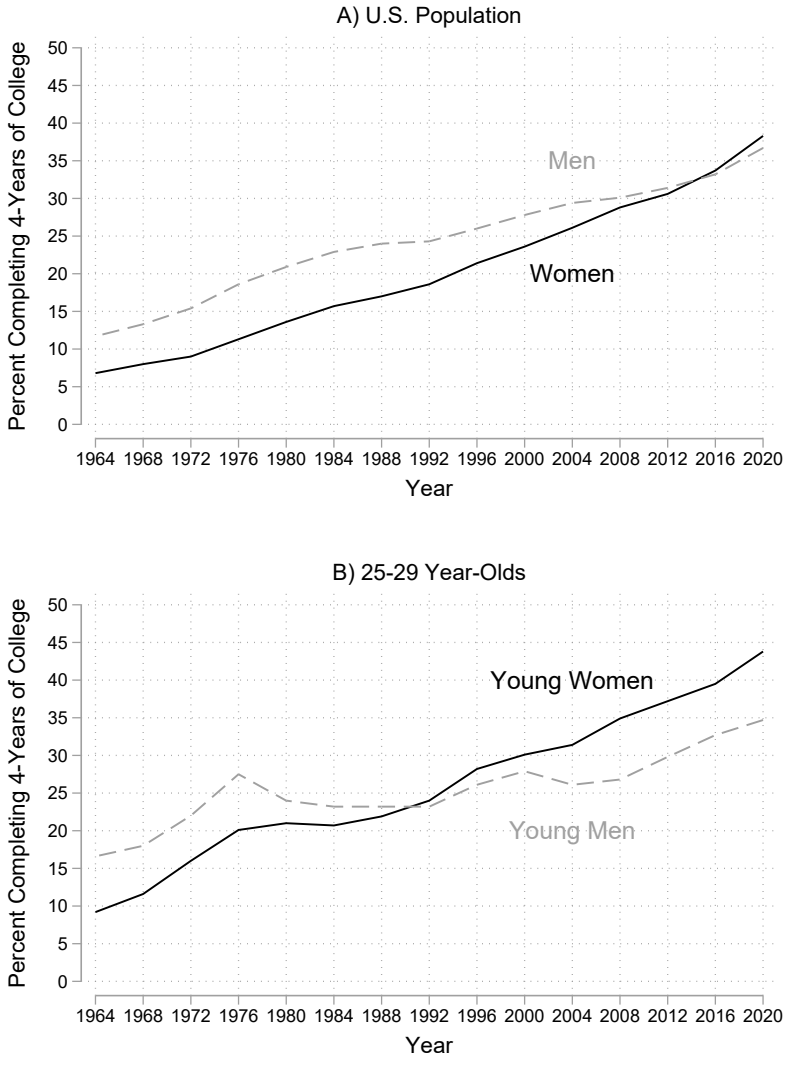
Wolfinger, R. E. and S. J. Rosenstone (1980). *Who votes?* Yale University Press.

Figure 1: Voter Turnout Gaps by Gender and Education, 1964 to 2024



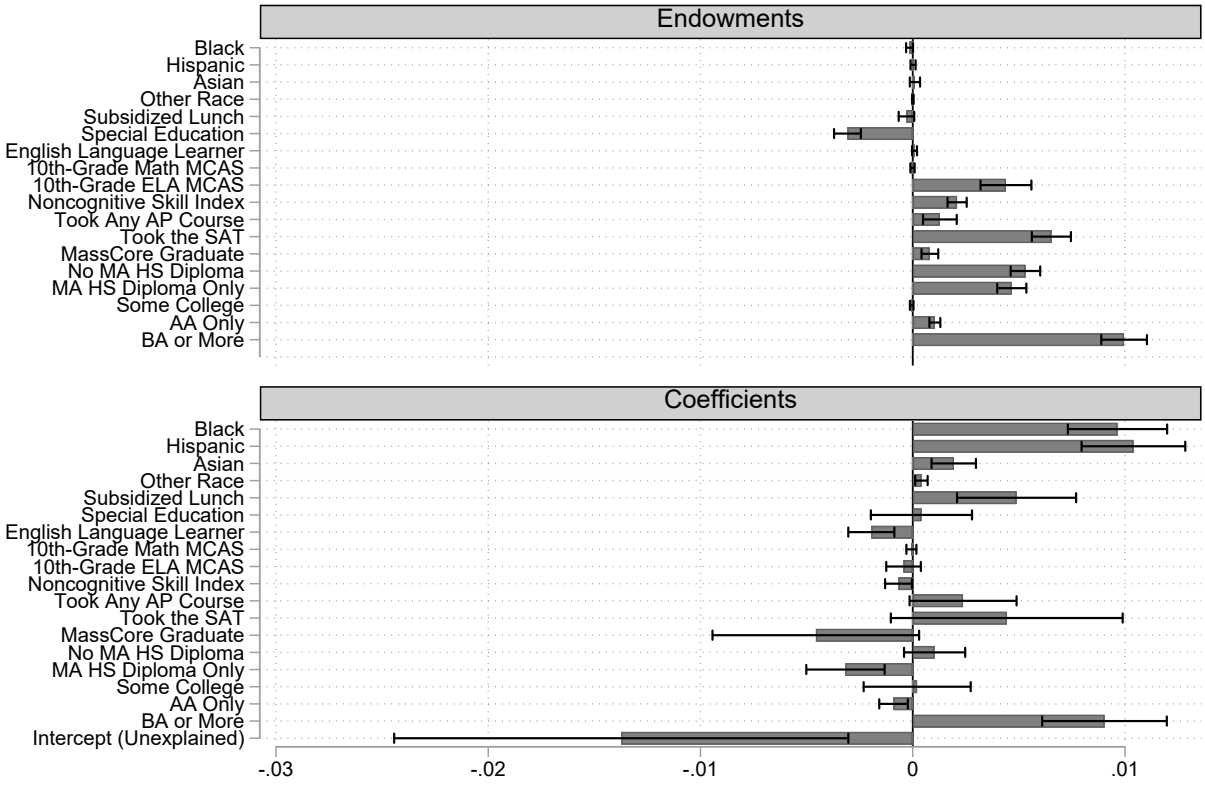
Notes: This figure shows voting rates between 1964 and 2024 by gender (Panel A) and education level (Panel B). Source: U.S. Current Population Survey

Figure 2: Population with College Degree by Gender, 1964 to 2020



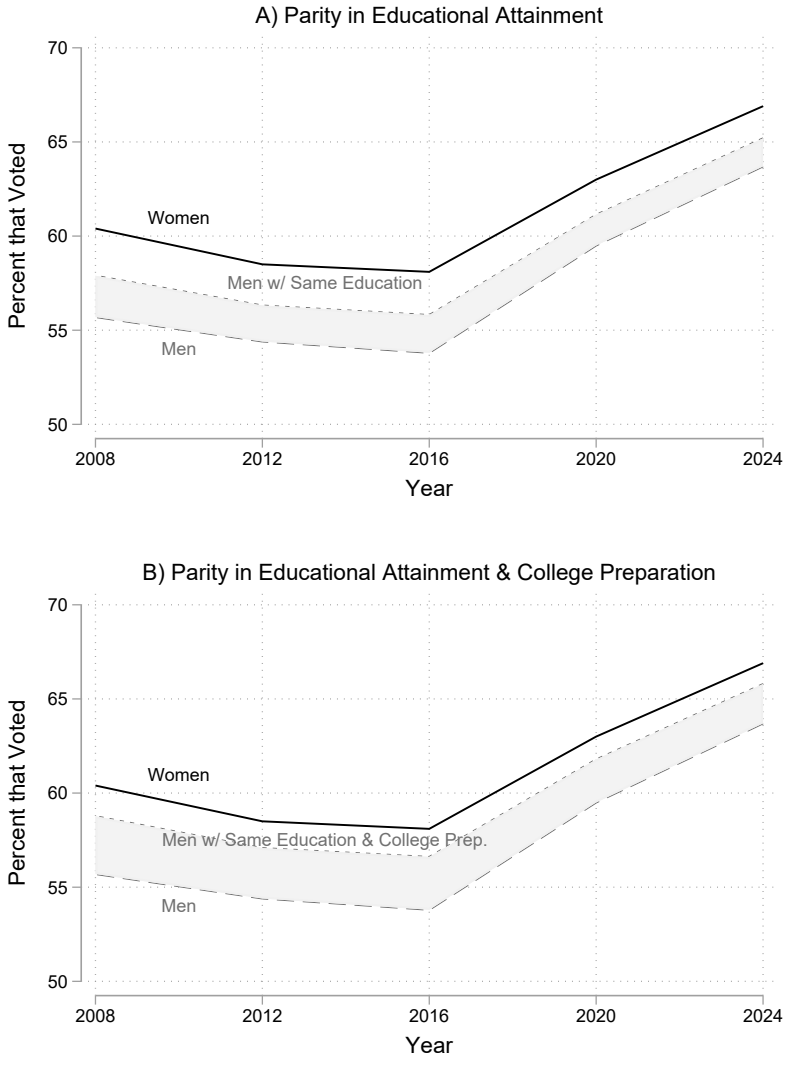
Notes: This figure shows college completion rates between 1964 and 2020 by gender for the U.S. population (Panel A) and for 25-29 year-olds (Panel B). Source: U.S. Current Population Survey

Figure 3: Detailed Threefold BOK Decomposition of Gender Gap in Voter Turnout



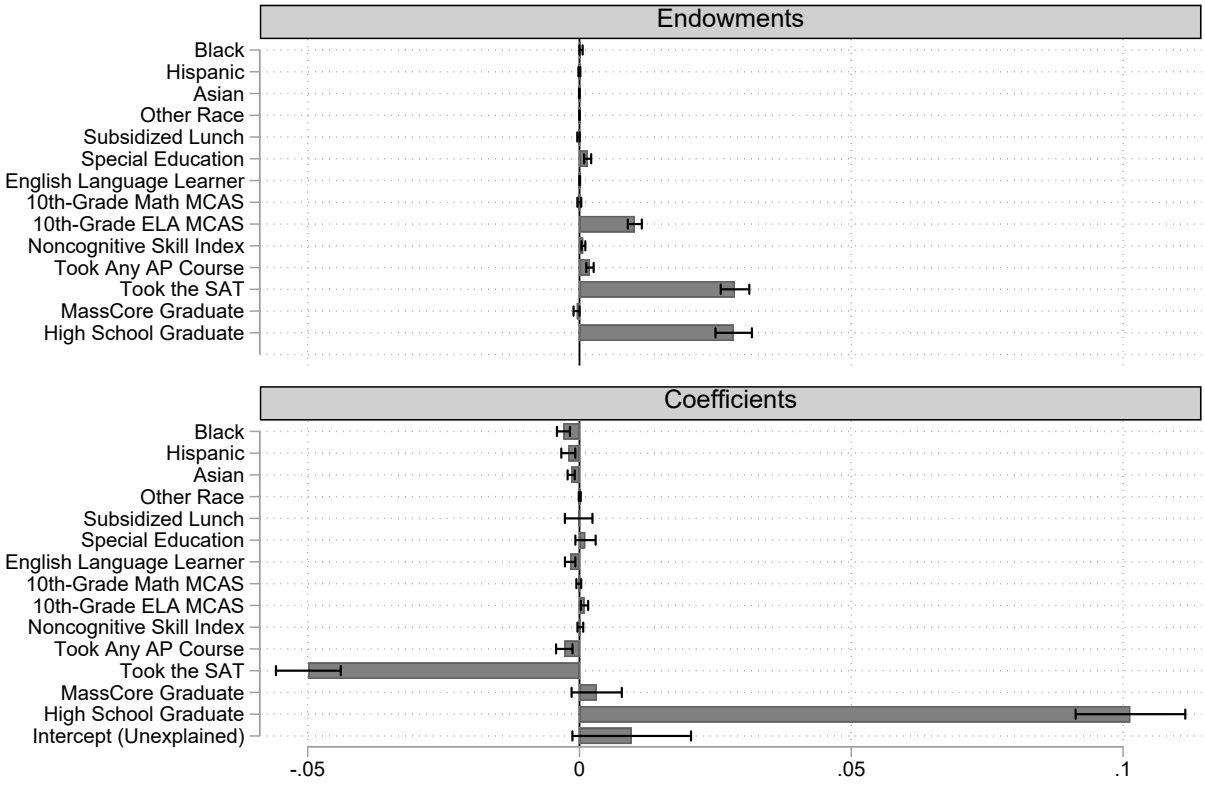
Notes: This figure details the portion of the gender gap in voting that can be explained by endowments of individual factors by gender and differential returns to those factors (coefficients).

Figure 4: Gender Gap in Voter Turnout with Women's Educational Behavior Applied to Men



Notes: This figure projects the magnitude of the gender voting gap using estimates from the decomposition if men and women had the same educational attainment and returns to education (Panel A). Panel B adds similar college preparation to the projection.

Figure 5: Detailed Threefold BOK Decomposition of Gender Gap in College Attendance



Notes: This figure details the portion of the gender gap in college attendance that can be explained by endowments of individual factors by gender and differential returns to those factors (coefficients).

Table 1: Summary Statistics by Gender

	All (1)	Women (2)	Men (3)	p-value (4)
(A) Student Characteristics				
White	0.752	0.751	0.754	0.029
Black	0.089	0.091	0.088	0.001
Hispanic	0.108	0.108	0.108	0.917
Asian	0.047	0.047	0.047	0.986
Other Race	0.004	0.004	0.004	0.380
Subsidized Lunch	0.252	0.253	0.251	0.181
Special Education	0.158	0.116	0.200	0.000
English Language Learner	0.053	0.052	0.054	0.045
(B) High-School Experiences				
10th-Grade Math MCAS	0.025	0.022	0.028	0.038
10th-Grade ELA MCAS	0.026	0.147	-0.091	0.000
Noncognitive Skill Index	0.001	0.051	-0.047	0.000
Graduated MA HS in 4 Years	0.768	0.803	0.735	0.000
Graduated MA HS in 5 Years	0.795	0.823	0.768	0.000
(C) College Preparation				
Took Any AP Course	0.252	0.291	0.213	0.000
Took the SAT	0.586	0.640	0.533	0.000
MassCore Graduate	0.594	0.613	0.574	0.000
(D) Educational Attainment				
No MA HS Diploma	0.112	0.090	0.133	0.000
MA HS Diploma Only	0.155	0.120	0.189	0.000
Some College	0.323	0.315	0.330	0.000
AA Only	0.047	0.053	0.043	0.000
BA or More	0.363	0.422	0.305	0.000
(E) Voting Outcomes				
Registered to Vote by Age 19	0.350	0.351	0.350	0.681
Ever Registered to Vote	0.834	0.839	0.829	0.000
Voted in First Possible Pres. Election After Pot. College	0.408	0.438	0.380	0.000
N	225,875	111,154	114,721	

Notes: This table reports the summary statistics of the sample of Massachusetts student records matched with voter files. The fourth column shows the p-value from a difference-of-means tests between the two groups. Education levels in Panel D are exclusive and non-overlapping, so individuals are counted only in their highest education category. MCAS is the statewide standardized given to students in grades 3-8 and 10. The noncognitive skill index is an index of measures used to approximate noncognitive skills: attendance in 10th grade, on-time progression to 11th grade, and a school suspension indicator.

Table 2: Voting as a Function of Education

	Vote				
	(1)	(2)	(3)	(4)	(4)
Female	0.054***	0.044***	0.034***	0.022***	0.021***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Black	-0.009	0.034***	0.020***	0.017***	-0.020***
	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)
Hispanic	-0.034***	0.019***	0.013**	0.014***	-0.010**
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Asian	-0.051***	-0.083***	-0.096***	-0.098***	-0.118***
	(0.011)	(0.009)	(0.008)	(0.008)	(0.007)
Other Race	-0.064***	-0.020	-0.021	-0.017	-0.022
	(0.015)	(0.015)	(0.014)	(0.014)	(0.015)
Subsidized Lunch	-0.129***	-0.072***	-0.059***	-0.047***	-0.053***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)
Special Education	-0.054***	0.027***	0.033***	0.038***	0.038***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
English Language Learner	-0.123***	-0.057***	-0.064***	-0.061***	-0.072***
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
10th-Grade Math MCAS		0.042***	0.018***	0.007***	0.006***
		(0.002)	(0.002)	(0.002)	(0.002)
10th-Grade ELA MCAS		0.040***	0.027***	0.017***	0.019***
		(0.002)	(0.002)	(0.002)	(0.002)
Noncognitive Skill Index		0.038***	0.025***	0.019***	0.021***
		(0.001)	(0.001)	(0.001)	(0.001)
Took Any AP Course			0.038***	0.022***	0.020***
			(0.004)	(0.004)	(0.004)
Took the SAT			0.108***	0.063***	0.057***
			(0.003)	(0.003)	(0.003)
MassCore Graduate			0.026***	0.012***	0.017***
			(0.003)	(0.003)	(0.003)
No MA HS Diploma				-0.052***	-0.045***
				(0.004)	(0.004)
Some College				0.077***	0.077***
				(0.003)	(0.003)
AA Only				0.166***	0.170***
				(0.006)	(0.006)
BA or More				0.173***	0.175***
				(0.005)	(0.005)
School Fixed Effects					Yes
R-Squared	0.032	0.066	0.077	0.090	0.100
N	225,875	225,875	225,875	225,875	225,875

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays linear probability models of whether an individual voted in the first presidential election after enough time had lapsed for her to have attended college. The top-line coefficients represent the gender gap in voting given the different model specifications. To address potential issues with multicollinearity between the variables, we implement LASSO (least absolute shrinkage and selection operator) estimates for our final model. For all the tuning parameters used, including cross-validation, adaptive LASSO, and plug-in methods, the ordinary least squares (OLS) model produces the lowest mean squared error, suggesting that all variables tested should be included in the model.

Table 3: Determinants of Voting for Women vs. Men

	Vote			
	(1)	(2)	(3)	(4)
	Women	Men	Difference	p-value
Demographics				
Black	0.071***	-0.037***	0.108	0.000
	(0.007)	(0.006)		
Hispanic	0.059***	-0.030***	0.090	0.000
	(0.007)	(0.005)		
Asian	-0.077***	-0.117***	0.040	0.000
	(0.010)	(0.008)		
Other Race	0.027	-0.059***	0.086	0.005
	(0.023)	(0.019)		
Subsidized Lunch	-0.038***	-0.055***	0.017	0.000
	(0.004)	(0.004)		
Special Education	0.039***	0.037***	0.002	0.736
	(0.005)	(0.004)		
English Language Learner	-0.077***	-0.045***	-0.032	0.000
	(0.010)	(0.007)		
Student Skills				
10th-Grade Math MCAS	0.004	0.008***	-0.004	0.264
	(0.003)	(0.002)		
10th-Grade ELA MCAS	0.020***	0.016***	0.004	0.284
	(0.003)	(0.002)		
Noncognitive Skill Index	0.016***	0.021***	-0.005	0.024
	(0.002)	(0.002)		
College Preparation				
Took Any AP Course	0.028***	0.016***	0.011	0.064
	(0.005)	(0.005)		
Took the SAT	0.068***	0.059***	0.009	0.112
	(0.004)	(0.004)		
MassCore Graduate	0.008*	0.017***	-0.009	0.065
	(0.004)	(0.004)		
Educational Attainment				
No MA HS Diploma	-0.033***	-0.058***	0.025	0.001
	(0.007)	(0.005)		
Some College	0.088***	0.071***	0.017	0.005
	(0.005)	(0.004)		
AA Only	0.168***	0.172***	-0.004	0.697
	(0.009)	(0.008)		
BA or More	0.199***	0.153***	0.046	0.000
	(0.006)	(0.006)		
R^2	0.082	0.096		
N	111,154	114,721		

*p<0.1; **p<0.05; ***p<0.01
Standard errors in parentheses

Notes: This table displays linear probability models of whether an individual voted in the first presidential election after enough time had lapsed for her to have attended college. The first column represents the determinants of voting for women, while the second column represents the determinants of voting for men. Column 3 denotes the difference between the coefficients in Column 1 and Column 2, while Column 4 reports the p-value for whether the difference is statistically significant.

Table 4: Threefold BOK Decomposition of Gender Voting Gap

Decomposition Information		Share of the Gap
Probability of Voting: Women	0.438	
Probability of Voting: Men	0.380	
Gender Gap in Voting	**0.058**	**100%**
Endowments	0.033	56.68%
Student Demographics	-0.003	-5.69%
Cognitive Skills	0.005	6.96%
Noncognitive Skills	0.002	3.60%
College Preparation	0.009	14.82%
Educational Attainment	0.021	36.00%
Coefficients	0.019	32.31%
Student Demographics	0.026	44.27%
Cognitive Skills	-0.001	-1.70%
Noncognitive Skills	-0.001	-1.15%
College Preparation	0.002	3.84%
Educational Attainment	0.006	10.67%
Intercept (Unexplained)	-0.014	-23.62%
Interaction	0.006	11.01%
N (Women)	111,154	
N (Men)	114,721	

Notes: This table shows individuals' probabilities of voting in the first presidential election after the lapse of enough time for them to have attended college. The gap between these probabilities is then decomposed into the portion explained by differences in characteristics and the portion that remains unexplained, which accounts for the differential returns to these characteristics. The third component of the decomposition is the interaction between the two endowments and coefficients. Student demographics include dummy variables for Black, Hispanic, Asian, other race, free and reduced-price lunch status, special education status, and English language learner status. Cognitive skills include scores on the MCAS math and ELA in 10th grade. Noncognitive skills are represented by a noncognitive skill index composed of attendance, suspension, and on-track progression into 11th grade. College Preparation includes a set of dummy variables for whether the student took the Advanced Placement exam, the SAT, or the MassCore college prep curricula. Educational attainment includes a set of exclusive dummy variables for no high-school diploma, high-school diploma only, some college, associate's degree, and bachelor's degree or more. Standard errors are clustered at the school level.

Table 5: Threefold BOK Decomposition of Gender College Gap

Decomposition Information		Share of the Gap
Probability of Attending College: Women	0.788	
Probability of Attending College: Men	0.677	
Gender Gap in College Attendance	**0.111**	**100%**
Endowments	0.067	60.08%
Student Demographics	0.002	1.43%
Cognitive Skills	0.010	9.17%
Noncognitive Skills	0.001	0.65%
College Preparation	0.026	23.25%
High School Graduate	0.028	25.58%
Coefficients	0.053	47.59%
Student Demographics	-0.007	-6.53%
Cognitive Skills	0.002	1.61%
Noncognitive Skills	0.000	0.14%
College Preparation	-0.053	-47.69%
High School Graduate	0.101	91.40%
Intercept (Unexplained)	0.010	8.66%
Interaction	-0.008	-7.66%
N (Women)	111,154	
N (Men)	114,721	

Notes: This table shows probabilities for college attendance for men and women. The gap between these probabilities is then decomposed into the portion explained by differences in characteristics and the portion that remains unexplained, which accounts for the differential returns to these characteristics. The third component of the decomposition is the interaction between the two endowments and coefficients. Student demographics include dummy variables for Black, Hispanic, Asian, other race, free and reduced-price lunch status, special education status, and English language learner status. Cognitive skills include scores on the MCAS math and ELA in 10th grade. Noncognitive skills are represented by a noncognitive skill index composed of attendance, suspension, and on-track progression into 11th grade. College Preparation includes a set of dummy variables for whether the student took the Advanced Placement exam, the SAT, or the MassCore college prep curricula. Standard errors are clustered at the school level.

APPENDIX

Education and the Gender Gap in Voter Turnout

A.1 Description of Matching of Student Data to Voter Files

A.1.1 Data Source 1: Student Records

As described in the text, the Massachusetts DESE provided information on students' names, demographic information, participation in special education, English learner status, free and reduced-price lunch receipt, high-school graduation, test scores, SAT and Advanced Placement (AP) participation, and college enrollment and graduation. The important information for the purpose of our matching to voter rolls was the students' names and dates of birth. For the matching, we limited the data to these identifying variables and a student ID. All student records associated with a student ID from the Massachusetts Student Information Management System (SIMS) database were used, including duplicate records with different information for the same student ID (such as a record from one school that might have a middle initial and one from another school that might not). This provides the most comprehensive opportunity for matching with voter records, which can then be collapsed on student ID later.

A.1.2 Data Source 2: Voter Records

The pool of voter records is comprised of a variety of separate voter files merged into a single file to encompass multiple presidential elections for multiple states. Specifically, we use voter files collected in 2018 for Massachusetts and nearby states: Connecticut, Maine, New Hampshire, New York, Rhode Island, and Vermont. These comprehensive voter files contain voter records for the 2008, 2012, and 2016 presidential elections. While the records do not cover all possible states to which Massachusetts students may eventually move and where they may register to vote, we cast a wide net. The New England states have a regional agreement whereby students can attend public institutions of higher education in any of these partnering states without paying out-of-state tuition (`https://nebhe.org/tuitionbreak/`). Verified by our student records, fewer than 10 percent of the students in our sample attended college outside these states (as seen in Online Appendix Table A.1). All state voter records include the voter's name, and the majority of records (including all Massachusetts records) include the date of birth.

As with state voter records in general, some fields in the commercial vendor voter files are verified and validated, while others are less consistent. While we have high confidence in voter first name and last name, the availability of fields such as middle name, date of birth, and gender vary based on state voter record requirements and commercial vendor data quality. For primary matching purposes, the measure for date of birth, a standard variable used in matching administrative records, varies across and within some state voter files. Fortunately, the voter files from 2018 include measures of confidence in date of birth accuracy for each state: valid complete date, valid year and month or date, valid year, or missing birth date. Importantly, the birth dates for the state most central to our argument, Massachusetts, have high validity: 4.04 million of its 4.05 million voters have verified birth dates. Similarly, nearly all of the voter records in the Connecticut, New York, and Rhode Island voter files have verified birth dates. The New Hampshire voter file is missing nearly 20 percent of voters' dates of birth, while the birth dates for Maine include information only on year of birth, and the Vermont file has varying levels of birth date information. Thus, we use matching protocols to account for this variation as detailed below.

A.1.3 Matching Protocols

To ensure the matching was as accurate as possible, we implemented a number of matching procedures as follows:

1. Exact matches on first name, last name, and date of birth between the student records and Massachusetts voter records were declared matches and set aside.

2. For fuzzy matches (with minor discrepancies between two fields) on first name, last name, and date of birth between the student records and Massachusetts voter records, we employed two measures of "distance" between string variables in the matching process to determine likely matches: the Jaro–Winkler distance (JWD) and the cosine string distance (CSD). The following criteria (in order) were then used to make fuzzy matches, which were then also set aside:

   (a) Require exact matches on first name and last name and required two of the birth day, birth month, and birth year to match; require the birth year to be off by no more than two years; require the middle initial to match; if a middle name is reported in both sources, require the middle name to be within 0.1 on the JWD.

   (b) Require exact matches on the first name and date of birth; require last names to be within 0.2 on the JWD or 0.2 on the CSD with $q = 1$; require last names to be within 0.5 on the CSD with $q = 3$.

   (c) Require exact matches on the last name and date of birth; require first names to be within 0.2 on the JWD or 0.2 on the CSD with $q = 1$; require first names to be less than 1 on the CSD with $q = 4$ or agree on soundex code or within 0.2 on the JWD.

   (d) Require exact matches on the date of birth; require the first name to be within 0.2 on the JWD; require last name to be within 0.2 on the JWD; require last names to be less than 1 on the CSD with $q = 4$ or the sum of the JWD for the first and last name to be less than 0.15; require gender to match.

   (e) Require exact matches on the last name and date of birth; require the first name to match the middle name from the SIMS record to voter file or from the voter file to SIMS record; require the first letter of the first name to match the first letter of the middle name (in both directions). This captures students with reversed first and middle names between their SIMS record and voter file.

   (f) Require exact matches on the first and last name; require the year of birth to match; require the day of birth to match the month of birth (in both directions). This captures students whose day and month of birth are transposed.

3. Exact matches on first name, last name, and date of birth between the all student records (even those that matched to a Massachusetts voter file) and voter records for *states other than Massachusetts* were declared matches and set aside. These included even student records that previously matched with Massachusetts voter files to account for individuals who might have later moved (so that we retained initial votes in Massachusetts).

4. Fuzzy matches were made on first name, last name, and date of birth, allowing for minor discrepancies between the student records and voter records from *states other than Massachusetts*. The following fuzzy matching criteria (in order) were used to determine matches, which were then also set aside before we progressed to the next matching technique:

(a) Records with only a valid year and month or day of birth: Require exact matches on the first name, last name, and gender; require the middle initial to match; require the birth year and birth month to match; and if a middle name is reported in both data sources, require it to be within 0.1 on the JWD.

(b) Records with only a valid birth year: Require an exact match on the first name, last name, and gender; require the middle initial to match; require the birth year to match; and if a middle name is reported in both data sources, require the middle name to be within 0.1 on the JWD.

(c) Records with a missing birthday in the voter files, yet the first and last names in SIMS are unique: Require an exact match on the first name, last name, and gender; require the middle initial to match; and if a middle name is reported in both data sources, require the middle name to be within 0.1 on the JWD.

All student records not matched with a voter record through this process above were coded as not having registered to vote and thus not having voted. Because our outcome of interest is whether an individual voted, our coding of nonmatches as having not voted—even if they did vote in a different state or simply did not match for other reasons—potentially attenuates our estimates.

After we established numerous voter outcome variables for an individual's having ever voted in each election and in each state measured, the student records were collapsed onto the unique SIMS student ID, preserving all of the voting records along with the most complete SIMS student record for each student ID. The result of this matching process, additionally subset to just students who were in 10th grade in the years listed above, is a dataset of 445,740 individuals with a variety of student records (independent variables) and voting outcomes (dependent variables). This matching process allows us to reasonably link comprehensive administrative student data with extant administrative voter records.

Appendix Table A.1 displays the rates of match between student and voter records. Of the 445,740 student records in our sample, broken out by the states when the students attended college, approximately 80 percent match to a voter record in one of the states covered by the voter files. This table shows a few important checks on the matching process. First, regardless of the state where a student went to college, vast majorities of Massachusetts students still later appear in Massachusetts voter files. Second, students who attended college in a different state (a primary reason for moving across state lines for young people) do appear in voter files for those states at expected rates. For example, 79 percent of Massachusetts students who attended college in Connecticut still show up in the Massachusetts voter files. However, 12.6 percent of the Massachusetts students who attended college in Connecticut still show up in the Connecticut voter files. Except in the Massachusetts voter files, college state and state of registration have the second highest match rates. The third check on this matching process is the low match rates in most states for Massachusetts students who did not go to college in the respective states. For example, the rest of the Massachusetts students who attended college in Connecticut appear in the other state voter files quite rarely, which likely reflects those who moved to other states after college. This shows that the matching processes are likely accurate across state lines.

Table A.1: Presence in State Voter Files

College State	N	Massachusetts	Connecticut	Maine	New Hampshire	New York	Rhode Island	Vermont
Massachusetts School Students								
Massachusetts	132,894	0.896	0.007	0.006	0.010	0.015	0.011	0.003
Connecticut	2,090	0.872	0.097	0.005	0.007	0.054	0.010	0.004
Maine	1,140	0.804	0.008	0.257	0.017	0.032	0.008	0.005
New Hampshire	8,322	0.872	0.007	0.010	0.092	0.008	0.008	0.006
New York	7,993	0.871	0.013	0.008	0.008	0.182	0.010	0.004
Rhode Island	8,489	0.891	0.009	0.007	0.007	0.019	0.100	0.002
Vermont	3,761	0.891	0.010	0.017	0.015	0.041	0.007	0.060
Other States	19,929	0.744	0.007	0.007	0.006	0.039	0.010	0.003
No College	60,559	0.683	0.004	0.004	0.006	0.009	0.011	0.001
All	225,875	0.822	0.007	0.007	0.010	0.020	0.013	0.003

Notes: This table shows the rates at which students in our primary panel of Massachusetts 10th graders appear in the voter files for Massachusetts and six nearby states by state of college attended.

Table A.2: Determinants of College Attendance for Women vs. Men

	Attend Any College			
	(1)	(2)	(3)	(4)
	Women	Men	Difference	p-value
Demographics				
Black	0.036***	0.067***	-0.031	0.000
	(0.005)	(0.007)		
Hispanic	0.027***	0.044***	-0.017	0.002
	(0.006)	(0.006)		
Asian	0.003	0.034***	-0.031	0.000
	(0.007)	(0.007)		
Other Race	0.009	0.002	0.007	0.708
	(0.013)	(0.016)		
Subsidized Lunch	-0.036***	-0.038***	0.001	0.803
	(0.003)	(0.004)		
Special Education	-0.020***	-0.026***	0.006	0.211
	(0.004)	(0.004)		
English Language Learner	-0.052***	-0.023***	-0.029	0.000
	(0.009)	(0.007)		
Student Skills				
10th-Grade Math MCAS	0.024***	0.031***	-0.007	0.006
	(0.002)	(0.002)		
10th-Grade ELA MCAS	0.031***	0.038***	-0.008	0.001
	(0.002)	(0.002)		
Noncognitive Skill Index	-0.013***	-0.016***	0.004	0.029
	(0.002)	(0.002)		
College Preparation				
Took Any AP Course	0.008***	0.023***	-0.015	0.000
	(0.003)	(0.004)		
Took the SAT	0.152***	0.251***	-0.099	0.000
	(0.005)	(0.008)		
MassCore Graduate	-0.015	-0.023*	0.007	0.119
	(0.006)	(0.009)		
Education				
High School Graduate	0.737***	0.607***	0.130	0.000
	(0.006)	(0.009)		
R-Squared	0.431	0.434		
N	111,154	114,721		

*p<0.1; **p<0.05; ***p<0.01
School-clustered standard errors in parentheses

Notes: This table displays linear probability models of whether an individual attended any college. The first column represents the predictors of college attendance for women, while the second column represents the predictors for men. The third column denotes the difference between the coefficients in Column 1 and Column 2, while Column 4 reports the p-value for whether the difference is statistically significant.

Table A.3: Summary Statistics by Gender - Full Sample

	All (1)	Women (2)	Men (3)	p-value (4)
(A) Student Characteristics				
White	0.752	0.751	0.754	0.029
Black	0.089	0.091	0.088	0.001
Hispanic	0.108	0.108	0.108	0.917
Asian	0.047	0.047	0.047	0.986
Other Race	0.004	0.004	0.004	0.380
Subsidized Lunch	0.252	0.253	0.251	0.181
Special Education	0.158	0.116	0.200	0.000
English Language Learner	0.053	0.052	0.054	0.045
(B) High-School Experiences				
10th-Grade Math MCAS	0.025	0.022	0.028	0.038
10th-Grade ELA MCAS	0.026	0.147	-0.091	0.000
Noncognitive Skill Index	0.001	0.051	-0.047	0.000
Graduated MA HS in 4 Years	0.768	0.803	0.735	0.000
Graduated MA HS in 5 Years	0.795	0.823	0.768	0.000
(C) Educational Attainment				
No MA HS Diploma	0.131	0.106	0.155	0.000
MA HS Diploma Only	0.161	0.128	0.192	0.000
Some College	0.323	0.319	0.327	0.000
AA Only	0.045	0.050	0.040	0.000
BA or More	0.341	0.397	0.286	0.000
(D) Voting Outcomes				
Registered to Vote by Age 19	0.292	0.292	0.292	0.789
Ever Registered to Vote	0.800	0.791	0.808	0.000
Voted in First Possible Pres. Election After Pot. College	0.382	0.407	0.357	0.000
N	444,459	218,586	225,873	

Notes: This table reports the summary statistics of the sample of Massachusetts student records matched with voter files. The fourth column shows the p-value from a difference-of-means tests between the results for the two groups. (B) Educational experiences include a variety of school-related factors. MCAS stands for Massachusetts Comprehensive Assessment System, the statewide exams given to students in grades 3 through 8 and in grade 10. Noncognitive Skill Index is an index of measures used to approximate noncognitive skills: attendance in 10th grade, on-time progression to 11th grade, and a school suspension indicator. (C) Education levels are exclusive and non-overlapping, so individuals are counted only in their highest education category.

Table A.4: Determinants of Voting for Women vs. Men - Full Sample

	Vote			
	(1)	(2)	(3)	(4)
	Women	Men	Difference	p-value
Demographics				
Black	0.119***	-0.000	0.119	0.000
	(0.008)	(0.006)		
Hispanic	0.069***	-0.020***	0.090	0.000
	(0.006)	(0.004)		
Asian	-0.076***	-0.120***	0.044	0.000
	(0.008)	(0.006)		
Other Race	0.026	-0.037**	0.063	0.019
	(0.021)	(0.015)		
Subsidized Lunch	-0.032***	-0.054***	0.022	0.000
	(0.003)	(0.003)		
Special Education	0.028***	0.037***	-0.009	0.033
	(0.004)	(0.003)		
English Language Learner	-0.081***	-0.043***	-0.038	0.000
	(0.008)	(0.006)		
Student Skills				
10th-Grade Math MCAS	0.010***	0.011***	-0.002	0.470
	(0.002)	(0.002)		
10th-Grade ELA MCAS	0.022***	0.020***	0.003	0.261
	(0.002)	(0.002)		
Noncognitive Skill Index	0.026***	0.024***	0.002	0.231
	(0.002)	(0.001)		
Educational Attainment				
No MA HS Diploma	-0.040***	-0.064***	0.024	0.000
	(0.004)	(0.003)		
Some College	0.098***	0.086***	0.013	0.001
	(0.003)	(0.003)		
AA Only	0.176***	0.191***	-0.015	0.060
	(0.006)	(0.006)		
BA or More	0.222***	0.189***	0.033	0.000
	(0.005)	(0.005)		
R^2	0.073	0.087		
N	218,586	225,873		

*p<0.1; **p<0.05; ***p<0.01
School-clustered standard errors in parentheses

Notes: This table displays linear probability models of whether an individual voted in the first presidential election after enough time had lapsed for her to have attended college. The first column represents the determinants of voting for women, while the second column represents the determinants of voting for men. Column 3 denotes the difference between the coefficients in Column 1 and Column 2, while Column 4 reports the p-value for whether the difference is statistically significant.

Table A.5: Threefold BOK Decomposition of Gender Voting Gap - Full Sample

Decomposition Information		Share of the Gap
Probability of Voting: Women	0.407	
Probability of Voting: Men	0.357	
Gender Gap in Voting	**0.050**	**100%**
Endowments	0.030	59.61%
Student Demographics	-0.003	- 6.33%
Cognitive Skills	0.005	10.04%
Noncognitive Skills	0.002	4.76%
Educational Attainment	0.025	51.13%
Coefficients	0.016	32.10%
Student Demographics	0.024	48.17%
Cognitive Skills	0.001	1.08%
Noncognitive Skills	-0.000	-0.16%
Educational Attainment	0.006	11.72%
Intercept (Unexplained)	-0.014	-28.70%
Interaction	0.004	8.29%
N (Women)	218,586	
N (Men)	225,873	

Notes: This table shows individuals' probabilities of voting in the first presidential election after the lapse of enough time for them to have attended college. The gap between these probabilities is then decomposed into the portion explained by differences in characteristics and the portion that remains unexplained, which accounts for the differential returns to these characteristics. The third component of the decomposition is the interaction between the two endowments and coefficients. Student demographics include dummy variables for Black, Hispanic, Asian, other race, free and reduced-price lunch status, special education status, and English language learner status. Cognitive skills include scores on the MCAS math and ELA in 10th grade. Noncognitive skills are represented by a noncognitive skill index composed of attendance, suspension, and on-track progression into 11th grade. Educational attainment includes a set of exclusive dummy variables for no high-school diploma, high-school diploma only, some college, associate's degree, and bachelor's degree or more. Standard errors are clustered at the school level.