Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1152

Education and the Gender Voting Gap

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 differences in educational attainment, differential returns to education, or other, non-education related elements. 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, we observe young women outvoting young men by 5 percentage points in the first presidential election in which they can vote after having potentially completed college. We find that just over 50 percent of this gap in voting can be explained by differences in educational attainment by gender. These results broadly suggest that a significant portion of the gender gap in voting can be attributed to the rise in women's education. If men's educational levels reached those of women, we would expect the gender voting gap to shrink significantly.

VERSION: March 2025

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

Education and the Gender Voting Gap*

Cameron J. Arnzen†
Brown University

Sarah R. Cohodes‡
University of Michigan and NBER

March 5, 2025

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 differences in educational attainment, differential returns to education, or other, non-education-related elements. 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, we observe young women outvoting young men by 5 percentage points in the first presidential election in which they can vote after having potentially completed college. We find that just over 50 percent of this gap in voting can be explained by differences in educational attainment by gender. These results broadly suggest that a significant portion of the gender gap in voting can be attributed to the rise in women's education. If men's educational levels reached those of women, we would expect the gender voting gap to shrink significantly.

*Acknowledgements: 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, and Michael G. Miller and from participants at the Association for Public Policy Analysis & Management Fall Conference 2022. Special thanks also go to James Feigenbaum, Jacob Brown, and Ryan Enos. This study was approved by the Institutional Review Boards at Teachers College Columbia University and University of Michigan.

†Annenberg Institute at Brown University, 164 Angell St., 2nd Floor, Providence, RI 02906. Email: cameron$_arnzen@brown.edu$

‡Gerald R. Ford School of Public Policy, University of Michigan, 735 S. State St., Ann Arbor, MI 48109. Email: scohodes@umich.edu.

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). Women outvoted men by 3.5 percentage points in the 2020 presidential election, with 82.2 million women (68.5 percent of those eligible) casting a ballot but only 72.5 million men (65 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 do, resulting in no substantive electoral changes. However, in recent decades, trends in public opinion show that women increasingly identify as liberal while men's ideological identification has remained relatively stable (Saad, 2024). Similarly, women have increasingly gravitated toward the Democratic 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 (Fernández et al., 2019; Herek, 2002; Cascio and Shenhav, 2020; Buser et al., 2020; 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., 1995a). This model has also been used to explain the gender gap in voter turnout (Schlozman et al., 1994; Verba et al., 1997; Burns et al., 2002; Burns, 2007; Burns et al., 2018; Cascio and Shenhav, 2020), 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 political participation, this work 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 their endowments of many of the resources that supposedly matter for civic engagement, such as wages 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. How these resources translate to participation remains a subject of debate, particularly as women's social, economic, and political standing has been greatly enhanced in recent decades.

In addition to the broad societal progress toward gender equity in recent decades, 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 begun to turn out to vote, women have also far outpaced men in educational attainment (Goldin et al., 2006; Goldin and Katz, 2010; Cascio and Shenhav, 2020). 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 trends, in conjunction with the well-documented association between education and political participation (Wolfinger and Rosenstone, 1980; Verba and Almond, 1963; Brady et al., 1995a; Nie et al., 1996; Sondheimer and Green, 2010), have led some scholars to conclude that women's gains in their exercise of the franchise should be attributed to their increasing educational attainment (Burns et al., 2018; Cascio and Shenhav, 2020). However, without data that directly link educational records to voting outcomes, it is difficult to verify exactly how much of the voting gap is attributable to educational attainment as opposed to other individual and societal factors. We attempt to do so here.

In this paper, we examine how much of women's higher rates of voting in recent elections can be explained by their higher educational attainment. We also account for cognitive and noncognitive skills in case the differences in education and voting are codetermined by differences in such skills. Using longitudinal data consisting of student records of Massachusetts public high school students matched to voting records, we follow two cohorts of 10th graders through their high school and college trajectories and the first presidential election in which they could vote after potential college attendance.[1] In the elections in these data, women outvote men by 5 percentage points, a gender gap consistent with the participation gaps in more recent presidential elections.

We build on recent literature on gender gaps in attitudes and behaviors and the literature on education and political participation. Burns et al. (2018) and Cascio and Shenhav (2020) most directly connect women's increases in voting participation to their levels of education. Burns et al. (2018) explore broad trends in women's political participation in the past 100 years, relying on a variety of compiled surveys, mostly from Pew Research. Cascio and Shenhav (2020) explore the same question, relying on data from the Current Population Survey and Gallup polls. These studies illuminate the parallel national trends in women's education and participation but are unable to connect these trends at the individual level.

Our unique, comprehensive data allow us to formally test the link between education and voting using student-level, longitudinal data. Initially, we construct a linear probability model of voting in the first presidential election after a cohort would have completed college. Focusing on factors that help explain the 5-percentage-point gender gap in voter turnout, we find that the gap significantly shrinks when we control for individuals' demographic characteristics and cognitive and noncognitive skills. However, the gap shrinks most when we account for educational attainment, a finding that motivates our further analyses. Next, we estimate the linear probability models separately for men

[1] The first cohort includes students who were 10th graders in 2002, 2003, and 2004, while the second includes students in 10th grade in 2006, 2007, and 2008. We follow the first cohort to the 2012 election, a time span that allows at least 6 years for them to graduate if they attended college; we follow the second cohort to the 2016 election, again allowing at least 6 years for them to graduate college.

and women, finding broad similarities in civic returns to the aforementioned factors: Most of the characteristics impact later-life voting similarly for men and women, with educational attainment playing the most significant role. Finally, using a Kitagawa–Oaxaca–Blinder-style decomposition technique, we decompose the gender gap in voting to more precisely estimate how much of the gap can be attributed specifically to the gender differences in educational attainment and other educational factors and how much is due to other, broader factors not captured by our individual-level data. We find that women's higher educational attainment explains over half of the gender gap in voting—a significantly larger proportion of the gap than the share explicable by any other factor. After we account for demographics and cognitive and noncognitive skills, only 40 percent of the gap remains unexplained, though we can also precisely estimate that approximately 30 percent of the gender gap in voting can be attributed to gendered returns to individual characteristics (for example, Black women outvote Black men for reasons our data cannot capture, though we can rule out that this residual gap for these groups arises from differences in educational attainment and cognitive and noncognitive skills). In short, women's higher levels of education explain a significant portion of the gender gap in voting when we account for other factors.

To explore why women pursue higher levels of education than men—a difference theorized to be the foundation of the voting gap—we also decompose the gender gap in college attendance using the measures of demographics and cognitive and noncognitive skills available in our data. In contrast to Jacob (2002), who finds that similar educational experiences explain 96 percent of college attendance, our educational data explain only approximately 31 percent of the gender gap in college attendance.[2] Taken together, the results from our analyses suggest that, though educational attainment explains a large portion of the gender *voting* gap, a significant portion of the gender gap in *educational attainment* remains unexplained by our data. As with the gender voting gap, broader social, economic, and political factors also contribute to the gender gaps in college attendance and educational attainment.

This paper proceeds as follows. First, in Section 2, we describe the history of the gender gap in voting 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 voting 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]Jacob (2002) decomposes a 4-percentage-point gap in college attendance, while the gap in our data is of 10 percentage points.

2 Gender and Education Gaps in Voter Turnout

2.1 Explanations for the Gender Gap in Voting

Though the 19th Amendment enfranchised white women in 1920, women lagged men on all forms of political engagement for many subsequent decades, despite gradual increases in participation (Burns, 2007; Burns et al., 2018). Additionally, even under the 19th Amendment, women of color were largely prevented from voting until the passage of the Voting Rights Act in 1965 and, in many cases, still faced barriers to voting because of voter suppression. In national elections, the shares of women who vote 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 Figure 1.

Empirical explanations of political participation generally employ the resource model of political participation (Brady et al., 1995b): Resources such as income, employment, civic interest, and education—endowments of which have historically favored men—have been argued to foster 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 (Coffé and Bolzendahl, 2010; Burns et al., 2017). 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., 1995b).

The one resource on which women have gained an advantage is education. Despite women's exclusion from many institutions of higher education for much of the 20th century, in recent decades, women's levels of educational attainment have surpassed men's. Recently, the overall share of women with a college degree has surpassed 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). 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 voting.

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 voting. 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 explanations for a portion of 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 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.

These factors may explain some of the current gender gap in turnout. In example, women's rise in employment in recent decades likely plays a role; however, if employment rates and workplace socialization fully explain gender differences in voting behaviors (Schlozman et al., 1999), we would still expect men to outvote women today. However, most studies do not explain why we might expect changes in women's employment, personality traits, or sense of civic duty to occur over time, underscoring that the broader political environment that differently engages these traits is likely to have changed. For example, (Stauffer and Fraga, 2022) find that women outvote men consistently in most electoral arenas but that this gap is affected by electoral factors such as race competitiveness, candidate gender, and racial composition of the electorate. Additionally, noting that much of the literature treats these resources and skills as equally important to all groups, Brown (2014) applies an intersectional lens to analyze political participation, arguing that a variety of group identities shape how women are mobilized to vote.

While the broader relationship between gender and political participation has many facets, accounts turning on 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[3] 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 voting. 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).

While women's social, economic, and political standing has been enhanced in the last century, given the consistent and positive association between educational attainment and participation, it

[3]This group consists of some of the architects of the general resource model theory and colleagues who helped develop and apply the theory over time.

is reasonable to expect that the education gap is likely responsible for a significant portion of the gender voting gap. Indeed, while the top panel of Figure 2 demonstrates the broad reversal of the education gap by gender for the general population, it notably hides underlying changes driving this reversal. Specifically, as demonstrated in the second panel of Figure 2, 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. The increasing gender gap in college attendance for young people is likely to continue to inflate the gender gap in voter turnout. We shed light on how much this gap might reshape the electorate in this paper.

Cascio and Shenhav (2020) also describe changes in women's political participation, issue preferences, and partisanship in the last century. They demonstrate that the gap in voter turnout inverted from a nearly 10-point participation deficit for women in 1940 to an advantage of approximately 4 percentage points in 2016. Investigating this change in behavior across cohorts, they find that higher turnout among younger cohorts of women over the years accounts for the new gender gap. These are the same young women who are increasingly college educated over time. Both Burns et al. (2018) and Cascio and Shenhav (2020) cite women's education gains as the source of their increased exercise of the franchise.

We build on this work by formally testing the explanatory power of gendered educational attainment for voting using individual-level data, documenting how much of the gender gap in turnout can be attributed to educational attainment. We also explore measures of educational experiences beyond degrees held or years of formal schooling to explore whether measures of cognitive or noncognitive skills can explain the gender gap in voter turnout.

2.2 Voting Gaps by Education Level

People with more education vote at higher rates than those with less education. This 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), so it is likely the case that women's educational gains are related to their disproportionate presence at the polling booth. The bottom panel of Figure 1 shows the education gap in voter turnout for presidential elections in the past sixty years. Over time, the voting gaps by educational attainment have widened, with the most educated voting at nearly double the rate of the least educated. Differential rates of political participation by education may skew the democratic process by weighting electoral outcomes toward elites (Gilens and Page, 2014). Given these stark gaps, it is no surprise that 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., 1995b).

Our investigation of the relationship between gender, education, and voting is further complicated by the fact that we still have much to learn about *how* education shapes voting. Nearly all studies that employ the logic of the resource model of political participation underscore the importance of education as both a resource itself *and* as a process through which other resources are developed (Burns, 2007; Burns et al., 2018). Some 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). However, relatively few studies that do estimate causal impacts go beyond years of schooling to evaluate *how* education increases the likelihood of voting. As education also affects the acquisition of other resources that increase political participation such as income/employment, civic skills, and social networks, it is difficult to tease out what aspect of education increases voter participation (Tenn, 2007; Kam and Palmer, 2008; Berinsky and Lenz, 2011).

Some recent, notable exceptions further 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), using lottery admissions from Boston charter schools, find significant impacts of charter school attendance on voting, which they argue are driven through education-induced noncognitive skills. Focusing more on higher education, Bell et al. (2024) find that the type of college an individual attends also matters for the instillation of voting habits; in particular, attendance of a high-voting college increases an individual's likelihood of voting even well after college. Bell et al. (2024) find negligible differences in voting rates between STEM and non-STEM college students, in contrast to the idea that curriculum choices as a result of interest in politics drive voter behavior.

3 Data and Methods

3.1 Data

To understand how much of the gender gap in voting is attributable to educational attainment as opposed to other factors, 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, free and reduced-price lunch receipt, 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,[4] SAT exam and Advanced Placement (AP) course records from the College Board, and college enrollment and degree information from the National Student Clearinghouse. These data allow us to explore a wide array of education-related factors as determinants of later-life political participation.

To compile the voter turnout data, we combine three Massachusetts voter files from 2012, 2015, and 2018. The resulting data span all national elections from 2008 to 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 between 2008 and 2016 also collected from commercial vendors. 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.

Approximately 80 percent of the students in our sample match to at least one voter record.[5] This means that approximately 20 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 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.[6] 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.

We restrict our sample to Massachusetts public-school students in two panels of students. The first panel includes students who were 10th graders in the years 2002, 2003, and 2004 and the second students who were 10th graders in the years 2006, 2007, and 2008. We observe voter turnout in the 2012 presidential election for the first panel and the 2016 presidential election for the second panel. This strategy lets us follow each cohort for 6 years after projected high-school graduation (8

[4] 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.

[5] Some students match to multiple voter records in different states because of movement between multiple elections, as we discuss in the descriptive statistics.

[6] 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.

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. We define several outcome variables of interest. We use "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. We also construct measures of "ever registered to vote" and "registered to vote by the age of 19" for additional analysis. Pooling the panels yields a sample of 444,447 students.

3.2 Descriptive Statistics

Summary statistics for men and women and the p-values from difference-of-means tests are displayed in Table 2. 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 voting (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 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).

The most striking gender differences—in education levels (Panel C) and political participation (Panel D)—emerge later in life. In line with the gendered trends in educational attainment, the men in our sample tend to be less educated than the women: 15.5 percent of men have no high-school diploma, while this figure only 10.6 percent for women. Similarly, 19.2 percent of men only have a high-school diploma, while this is the true of only 12.8 percent of women. The proportions of men and women who completed some college (but did not earn a degree) 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: 39.7 percent of women in the sample hold at least a bachelor's degree, but only 28.6 percent of men do, as measured 6 years after expected high-school graduation. Consistent with the national trends we presented in Figure 2, the young women in our sample are far more educated than the young men.

[7] Jackson (2018) also includes GPA, but we omit this component because of data differences.

For voting outcomes (Panel D), there is no gender difference in voter registration: 80 percent of both genders registered to vote in at least one state at some point in our data. Registration by age 19, signaling the intention to participate just out of high school, is also the same for men and women, at 29 percent. 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 percentage points.

Our first step in understanding the gender gap in voting is estimating models of voting as a function of demographics, educational experiences, and educational attainment to determine whether these characteristics contribute to voting behavior. Table 3 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 voting 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.

In this regression taxonomy, a reduction in the coefficient on the indicator for *Female* when we add a variable indicates that the factor has explanatory power for gender differences in voting. Model 1 shows a gender gap of approximately 4.5 percentage points when we include only demographics, which indicates that little of the 5-percentage-point difference in voting by gender can be explained by gendered differences in student characteristics. Adding test scores and noncognitive skills reduces the coefficient on *Female* to approximately 3.7 percentage points, indicating that these variables have some explanatory power. Model 3, which includes education levels, reduces the explanatory power of gender in our models to the greatest extent, to 1.8 percentage points, confirming the importance of educational attainment. Adding high-school fixed effects in Model 4 yields remarkably similar estimates to those in Model 3, indicating that these patterns play out within schools, not across them.[8] Thus, we move forward with a variation of Model 3 in the decomposition.[9] We now turn to the methods we use connect these gaps in voting to the gaps we observe in education levels.

[8] Appendix Table A.3 through Appendix Table A.8 include binary variables for school type (public, charter, magnet, alternative, etc.) replicating all of our analyses. However, as our primary focus is on educational attainment, we omit the coefficients on these school-type variables from our primary analyses, as gender and racial differences in enrollment by school type vary but explain little more of the gap in voting than do other demographic variables.

[9] 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.

3.3 BOK Decomposition

We use a BOK-style decomposition to separate the factors that contribute to the gender differences in voting. 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 estimate group differences (Blinder, 1973; Oaxaca, 1973; Kitagawa, 1955). 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, these models have been used to explain a wide range of topics including gender differences in political knowledge (Dow, 2009), partisan differences in evaluations of inflation (Bachmann et al., 2021), and even the amount of political polarization due to changes in voter positioning versus politician positioning (Kertzer, 2020). The methods have been used elsewhere to explore gender gaps in political participation, focusing on gendered interest in politics and turnout in European countries (Dow, 2009; Kostelka et al., 2019; Dassonneville and Kostelka, 2021), though previous work on gender and voting decomposes survey data and cross-sectional data, while we consider administrative data here.

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.

One criticism of such decompositions is that the estimates derived from them are only as good as the data underlying the results being decomposed. The data employed here are well suited for a BOK decomposition of the gender gap in voting. 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 444,447 individuals matched with voting records, which yields a large, representative 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.

We present a threefold BOK decomposition of the drivers of the gender gap in voting 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. Voting 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.[10] 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 voting into three components: the variation 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 voting due to the differences in characteristics between men and women, such as education levels, academic factors, 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. In our case, noncognitive skills, for example, may matter more for women's voting than for men's. 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 Voting

We consider whether there are differences by gender in the linear probability model, as in Equation 1. Table 4 reports the estimated coefficients from our selected model separately by gender.[11]

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

[11] Additionally, we conduct a separate robustness check using the measures of AP course enrollment and SAT scores in Appendix Table A.2 for the subset of students for whom we have these measures. The general effects

The coefficient estimates for each gender are similar for most variables other than race. Differences emerge by gender along student demographic characteristics. 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 calls such as those by Brown (2014) and Stauffer and Fraga (2022) for researchers to devote particular attention to intersectional identities while studying voting gaps.

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 without subsidized lunch. Similarly, English language learners are less likely to vote than their peers. Students with 10th-grade MCAS ELA scores one standard deviation higher than the mean have an approximately 2.1-percentage-point higher probability of voting among both women (2.2) and men (2.0). Both genders experience similar returns to math scores and noncognitive skills, as well.

Most informatively for our purposes, the coefficients for educational attainment levels are 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. 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, women who did not earn a diploma show a 4-percentage-point decrease and men with no diploma a 6.4-percentage-point decrease in voting likelihood. Attending some college without obtaining a degree predicts a 9.8-percentage-point increase in voting likelihood for women and a 8.6-percentage-point increase for men. Earning an associate's degree is associated with a 17.6-percentage-point increase in voting likelihood for women and a parallel 19.1-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 22.2 percentage points in the probability of voting for women and of 18.9 percentage points for men.

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 thus turn to the BOK decomposition to formalize this intuition. Table A.6 breaks down the components of the gender voting gap estimated in Equation 3. Approximately 41 percent of women voted in the first election after enough time had elapsed for them to have attended college, while this figure is 36 percent for men, resulting in a 5-percentage-point gap in voting by gender. This estimate is slightly larger than the national gap in Figure 1, which may reflect differences in age.

---

of educational attainment remain similar to those in our main analyses. We also replicate these analyses using school-type indicator variables as a robustness check on our linear probability models in Appendix Table A.5. The coefficients for educational attainment levels are remarkably similar across models, though the school-type indicators appear to draw some of the explanatory power away from other demographic variables.

Table A.6 reports the results of the decomposition of the gender gap in voting estimated from linear probability models. We find that 60 percent of the 5-percentage-point gender gap (approximately 3 percentage points) can be explained by differences in endowments by gender while the differential returns account for 32 percent of the gap and the interaction of the two the final 8.25 percent. We detail the contribution of demographics and specific educational factors below our discussion of the endowment and coefficient portions of the decomposition.

Differences in educational attainment endowments between women and men are responsible for approximately 51 percent of the total gender gap in voting. 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. Student demographics and skills explain a portion of the gap, as well, though they work in opposite directions. Women and men have similar demographics except for participation in special education—women are 8 percentage points less likely to receive special education services (Table 1). However, as shown in Tables 3 and 4, in our multivariable model when other variables are accounted for, special education students vote at higher rates. Thus, women as a group have demographics that make them less likely to vote.[12] Cognitive and noncognitive skills work in the opposite direction, accounting for approximately 15 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 likely to vote, though these factors account for a much smaller share of the gap than educational attainment. Thus, educational attainment is the most important endowment factor in explaining the gender gap.

Additionally, the threefold BOK decomposition estimates the amount of the gender gap in voting that can be explained by different returns to the characteristics of each group. This is the "Coefficients" section of A.6. Our model estimates that gendered differences in the relationship between some explanatory factors and voting account for approximately 32 percent of the voting gap. Specifically, the coefficients for student demographics matter most. The positive share of the gap for student demographics means that women of color far outvote men of color, which accounts for approximately 48 percent of the gender gap. Students of different racial backgrounds are similarly represented among men and women, but being in the Black, Hispanic, or other race categories increases voting for women but not for men. Gendered returns to skills matter little for the gender gap in voting, 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*

[12]To further explore the contributions of individual characteristics and their returns, Appendix Figure A.1 illustrates the contributions of each individual variable separately by endowment (characteristic) and coefficient (returns). We see that a variety of characteristics and their coefficients matter but that the characteristic that matters most, by far, is education. This mirrors the general consensus that a variety of factors matter for voting but that education is foremost among them (Smets and van Ham, 2013; Willeck and Mendelberg, 2022).

factors contributing to the gender voting 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 educational attainment endowment and returns to race counteract and offset those forces.

Table A.6 illustrates that educational attainment can explain a significant portion of the gender gap in voting. We also show that a significant portion of the gender gap in voting is explicable by the intersection of race and gender: Women of color are much likelier to cast a ballot than men of color, with demographic factors accounting for a similar portion of the gender gap as educational attainment.

These estimates allow us to empirically demonstrate how much of the gender gap educational attainment is responsible for, as education levels have significantly increased for women in recent decades. Carrying forward the endowment estimate for education levels (the first part that represents only the difference in education levels across gender), in Figure 3, we illustrate how much of the gender gap would disappear if men were to achieve the same level of education as women.[13] The shaded region illustrates that if men had education levels similar to women's, the voting gap between the genders would be halved.

Figure 3 shows that, even if men had the same education levels and returns to education as women, the gap would not entirely disappear but would significantly shrink. Other factors such as the racial dynamics we discuss here likely contribute to women's greater presence at the polls. However, educational attainment is responsible for a significant portion of the gap. 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 appear to explain approximately half 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 appears to be growing and will continue to reshape the educational differences for the general population. In this sample, 76.6 percent of girls pursued at least some college, while 65.3 percent of boys did so, in a pattern that mirrors that in Figure 2. We thus conduct an exercise similar to the one above for voting, this time decomposing the gap in college attendance using a variety of the student-level measures available to us in our data: demographics, MCAS scores, and

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

the noncognitive skill index. This exercise may help illuminate the roots of the gender voting gap through the gender educational attainment gap.

We estimate linear probability models for the binary outcome of attending any college in Table A.7. We find that Black and Hispanic women and men are both generally likelier to attend college than their white peers, when we control for test scores and noncognitive skills. Students enrolled in free and reduced-price lunch programs are less likely to attend college, as are students enrolled in special education. Unsurprisingly, students who score higher on standardized exams in math and ELA are likelier to attend college. Additionally, students who score higher on the noncognitive skill index (reflecting attendance, on-time progression, and suspensions), are likelier to attend college, even when test scores are included in the regression.

To the extent that college enrollment reflects the achievement variables and demographic and socioeconomic factors illustrated in Table A.7 (which may vary by gender), these factors should predict a significant portion of the college attendance gap. Table 7 details the summary results of this decomposition, while Appendix Table A.2 illustrates the detailed results. In contrast to our findings for voting, different endowments across gender explain only approximately 32 percent of the gender gap in college attendance. Of the portion that can be explained, student MCAS scores are responsible for nearly all of the explained variation in the gap. Demographics and noncognitive skills play a role, but to a lesser extent. These results imply that college-going differences by gender are not immutable and greater preparation for boys in high school could increase their college-going rates to some degree.

However, much of the explanation for college attendance remains outside the scope of the academic, demographic, and socioeconomic factors present in our data. The coefficients portion of Table A.7 reflects only limitedly differential returns by gender to demographics and educational experiences. However, a great deal of variation remains in the intercept term, which reflects factors that are unexplained by our set of variables but induce girls into college at higher rates than boys. A variety of potential explanations for this 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 finds that the opportunity costs of college attendance, gendered expectations for employment, and noncognitive abilities can explain nearly all (98 percent) of 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.[14]

[14]We also conduct analyses of the gender gap in college attendance for Massachusetts students, incorporating economic and opportunity cost measures similar to those in Jacob (2002). Jacob's measures represented economic factors at the state level, while our measures from the American Community Survey (college premiums for girls and boys, weekly earnings for high-school graduates, and local unemployment rates) are at the Massachusetts county level. However, in contrast to Jacob's findings, only the unemployment rate is statistically significant in these models. The inclusion of these measures makes virtually no difference for the decomposition results. This is likely because of the lack of variation in these measures *within* Massachusetts, in contrast to the cross-state measures utilized by Jacob.

Jacob (2002) emphasizes the role that noncognitive skills play in college attendance—a factor that has also been explicitly linked to voting. Jacob finds that, when girls' higher noncognitive skills are accounted for alongside cognitive skills, family background, and employment expectations, the gender gap in college attendance disappears. This finding is potentially connected to voting, as Holbein (2017) and Cohodes and Feigenbaum (2021) find evidence that increased noncognitive abilities cause increases in later-life voting behavior. We find that the noncognitive skills we measure do matter for both voting and college-going, albeit to a much lesser extent than in Jacob (2002). This could be because of differences in our measures of such skills, as Jacob (2002) employs more detailed survey data than we do, but could also be due to societal changes over time.

We have shown that educational attainment accounts for a large portion of the gender gap in voting, but we have less evidence on 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.

6 Conclusion

Women's educational gains in recent decades have been hypothesized to be the cause of the gender gap in voting (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 education and voting, we are able to confirm that half of the gender gap in voting is indeed explained by women's higher levels of education. 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 voting. This is consistent with the finding of Brown (2014) that Black women have different historical relationships with political parties, shaping their participation differently.

The lack of differences by gender in voter registration suggests that education matters for dynamics beyond the barriers in the process of registering to vote. Given that young women are far outpacing young men in educational attainment, our findings suggest that the voter turnout gap will continue to grow absent some intervention to improve young men's educational attainment.

However, our findings broadly suggest that educational attainment, while likely the most significant factor contributing to gender gap in voting, is far from the only factor. Gendered racial voting trends also play a considerable role in voting, and 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.

While the unexplained portions of these gaps warrant further exploration, we provide credible estimates of how much educational attainment 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 explains a majority of the voting gap, 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.

Nevertheless, 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

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.

Bell, D., J. Feng, J. B. Holbein, and J. Smith (2024). Do stem students vote? 12. *American Educational Research Journal 61*(1), 48–73.

Bell, D., J. B. Holbein, S. J. Imlay, and J. Smith (2024). Which colleges increase voting rates? Technical report, IZA Discussion Papers.

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 (1995a). Beyond ses: A resource model of political participation. *American Political Science Review*, 271–294.

Brady, H. E., S. Verba, and K. L. Schlozman (1995b). Beyond ses: A resource model of political participation. *American political science review 89*(2), 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.

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.

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.

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.

Gilens, M. and B. I. Page (2014). Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on politics 12*(3), 564–581.

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. 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.

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, K. Stehlik-Barry, et al. (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.

Smets, K. and C. van Ham (2013). The embarrassment of riches? a meta-analysis of individual-level research on voter turnout. *Electoral Studies 32*(2), 344–359.

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.

Tenn, S. (2007). The effect of education on voter turnout. *Political Analysis*, 446–464.

Verba, S. and G. Almond (1963). *The civic culture.*

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.

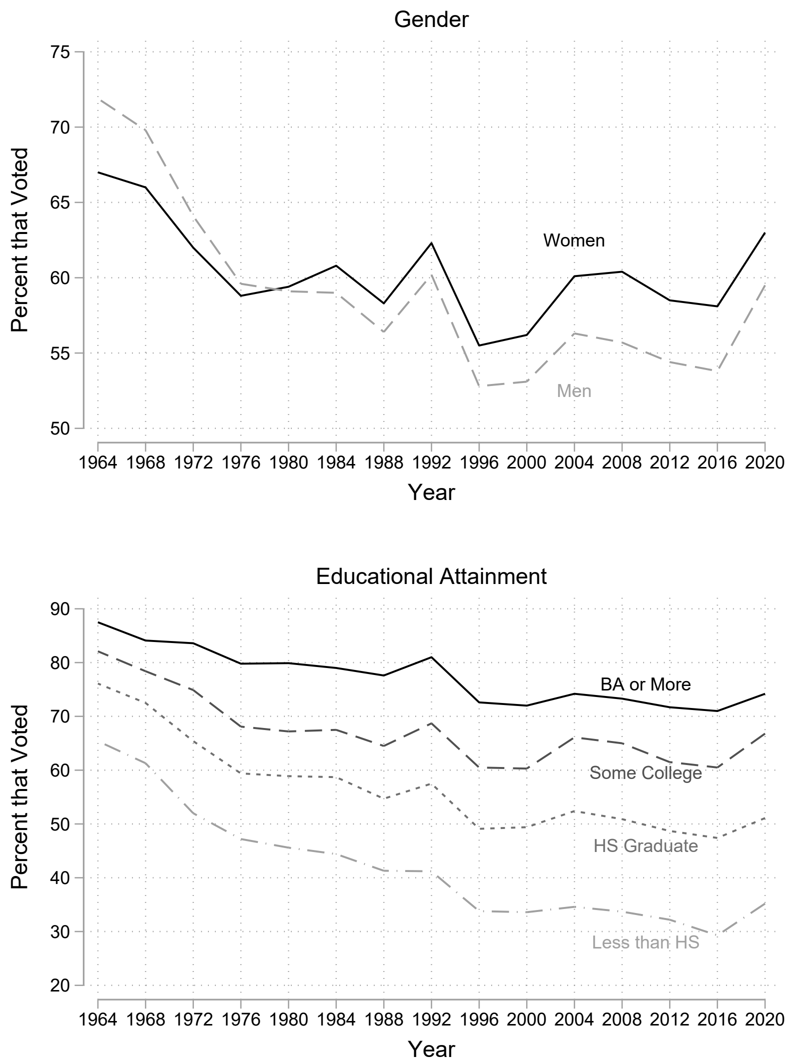
Verba, S., K. L. Schlozman, and H. E. Brady (1995). *Voice and equality: Civic voluntarism in American politics.* Harvard University Press.

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

Willeck, C. and T. Mendelberg (2022). Education and Political Participation. *Annual Review of Political Science 25*, 89–110.

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

Figure 1: Voting Gaps by Gender and Education, 1964 to 2020

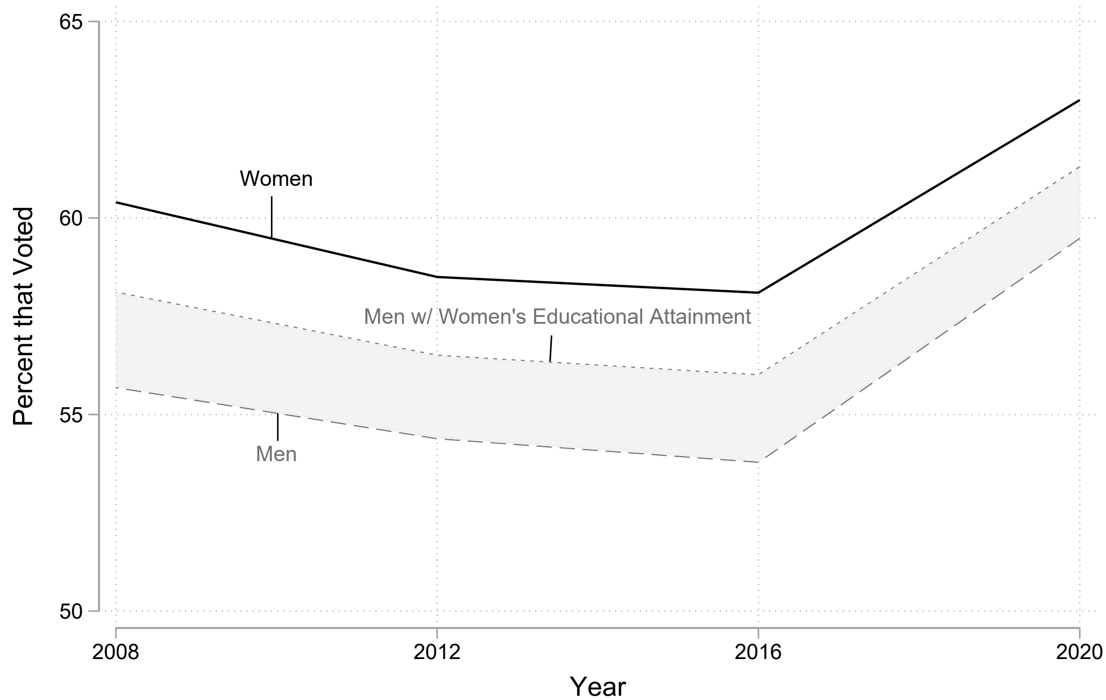


Notes: This figure shows the change in voting rates between 1964 and 2020. The top panel disaggregates voter turnout by gender, showing the reversal of the gender gap in voting since the 1980s. The bottom panel disaggregates voter turnout by education level, showing that the participation gaps by educational attainment have widened in recent decades. Source: U.S. Census Bureau Current Population Survey

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

Notes: This figure shows the change in college completion rates between 1964 and 2020 by gender, showing the elimination of the gender gap in college completion in the 2020s for the general population. The second panel shows the change in college completion rates between 1964 and 2020 by gender for individuals ages 25 to 29, demonstrating that young women have attained higher levels of education since 1990. Source: U.S. Census Bureau Current Population Survey

Figure 3: Voting Gap if Men and Women Had the Same Educational Attainment and Returns to Education



Notes: This figure projects how much of the gender voting gap would disappear if men and women had the same educational attainment and returns to education given the estimates from our decomposition, estimated for the time period covered by our data. The shaded portion reflects that approximately half of the gap is due to gender differences in educational attainment.

Table 1: MA Student Presence in State Voter Files

College State	N	MA	CT	ME	NH	NY	RI	VT
Massachusetts	251,526	0.870	0.007	0.006	0.011	0.015	0.011	0.003
Connecticut	3,922	0.818	0.090	0.007	0.006	0.052	0.015	0.003
Maine	2,092	0.741	0.008	0.282	0.016	0.032	0.007	0.004
New Hampshire	15,396	0.828	0.006	0.012	0.092	0.010	0.008	0.005
New York	15,674	0.822	0.014	0.009	0.009	0.170	0.009	0.004
Rhode Island	17,629	0.850	0.009	0.006	0.008	0.019	0.102	0.002
Vermont	6,744	0.850	0.009	0.019	0.016	0.040	0.009	0.052
Other States	39,087	0.677	0.007	0.008	0.006	0.038	0.011	0.002
No College	130,594	0.650	0.004	0.005	0.007	0.009	0.012	0.002
All	445,229	0.784	0.007	0.007	0.011	0.020	0.014	0.003

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

Table 2: Summary Statistics by Gender

	All (1)	Women (2)	Men (3)	p-Value (4)
(A) Student Characteristics				
White	0.752	0.751	0.754	0.030
Black	0.089	0.091	0.088	0.001
Hispanic	0.108	0.108	0.108	0.915
Asian	0.047	0.047	0.047	0.987
Other Race	0.004	0.004	0.004	0.380
FRPL	0.252	0.253	0.251	0.183
Special Education	0.158	0.116	0.200	0.000
English Language Learner	0.053	0.052	0.054	0.044
(B) High-School Experiences				
10th-Grade Math MCAS	0.025	0.022	0.028	0.040
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	0.769	0.803	0.736	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	0.784
Ever Registered to Vote	0.800	0.791	0.808	0.000
Voted in First Pres. Election After Pot. College	0.382	0.407	0.357	0.000
N	444,447	218,578	225,869	

Notes: This table details the summary statistics of our sample of Massachusetts student records matched with voter files for the whole sample and for women and men separately. The fourth column shows the statistical significance (p-value) from difference-of-means tests between the results for the two groups. (C) Education levels are exclusive and nonoverlapping, so individuals are counted only in their highest education category. 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.

Table 3: Regression Taxonomy of Voting as a Function of Education Variables

	Vote			
	(1)	(2)	(3)	(4)
Female	0.045***	0.037***	0.018***	0.017***
	(0.003)	(0.003)	(0.003)	(0.003)
Black	0.020***	0.067***	0.059***	0.013***
	(0.006)	(0.006)	(0.006)	(0.004)
Hispanic	-0.035***	0.023***	0.024***	-0.001
	(0.005)	(0.005)	(0.005)	(0.004)
Asian	-0.072***	-0.092***	-0.099***	-0.122***
	(0.009)	(0.006)	(0.006)	(0.006)
Other Race	-0.052***	-0.013	-0.007	-0.012
	(0.012)	(0.012)	(0.012)	(0.013)
FRPL	-0.107***	-0.062***	-0.043***	-0.048***
	(0.004)	(0.003)	(0.003)	(0.002)
Special Education	-0.053***	0.022***	0.033***	0.034***
	(0.004)	(0.003)	(0.003)	(0.002)
English Language Learner	-0.125***	-0.063***	-0.062***	-0.075***
	(0.008)	(0.007)	(0.006)	(0.007)
10th-Grade Math MCAS		0.038***	0.011***	0.009***
		(0.002)	(0.002)	(0.001)
10th-Grade ELA MCAS		0.038***	0.020***	0.021***
		(0.002)	(0.002)	(0.001)
Noncognitive Skill Index		0.040***	0.025***	0.027***
		(0.001)	(0.001)	(0.001)
No MA HS Diploma			-0.057***	-0.054***
			(0.003)	(0.003)
Some College			0.090***	0.088***
			(0.002)	(0.002)
AA Only			0.179***	0.183***
			(0.005)	(0.004)
BA or More			0.203***	0.199***
			(0.004)	(0.004)
School Fixed Effects	No	No	No	Yes
R-Squared	0.024	0.058	0.080	0.091
N	444,447	444,447	444,447	444,447

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

School-clustered standard errors in parentheses

Notes: This table displays results from various 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. The noncognitive skill measure is an index of measures used to approximate noncognitive skills (first principal component): attendance in 10th grade, on-time progression to 11th grade, and a school suspension indicator.

Table 4: Determinants of Voting for Women vs. Men

	Vote	
	(1)	(2)
	Women	Men
Demographics		
Black	0.119***	0.000
	(0.008)	(0.006)
Hispanic	0.069***	-0.020***
	(0.006)	(0.004)
Asian	-0.076***	-0.120***
	(0.008)	(0.006)
Other Race	0.026	-0.037**
	(0.021)	(0.015)
FRPL	-0.032***	-0.054***
	(0.003)	(0.003)
Special Education	0.028***	0.037***
	(0.004)	(0.003)
English Language Learner	-0.081***	-0.043***
	(0.008)	(0.006)
High-School Experiences		
10th-Grade Math MCAS	0.010***	0.012***
	(0.002)	(0.002)
10th-Grade ELA MCAS	0.022***	0.020***
	(0.002)	(0.002)
Noncogntive Skill Index	0.026***	0.024***
	(0.001)	(0.001)
Educational Attainment		
No MA HS Diploma	-0.040***	-0.064***
	(0.004)	(0.003)
Some College	0.098***	0.086***
	(0.003)	(0.003)
AA Only	0.176***	0.191***
	(0.006)	(0.006)
BA or More	0.222***	0.189***
	(0.005)	(0.005)
R^2	0.073	0.087
N	218,578	225,869

*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.

Table 5: Threefold BOK Decomposition of Gender Voting Gap

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	59.65%
Student Demographics	-0.003	- 6.34%
Cognitive Skills	0.005	10.13%
Noncognitive Skills	0.002	4.76%
Educational Attainment	0.025	51.11%
Coefficients	0.016	32.09%
Student Demographics	0.024	48.15%
Cognitive Skills	0.001	1.10%
Noncognitive Skills	-0.000	-0.16%
Educational Attainment	0.006	11.73%
Intercept (Unexplained)	-0.014	-28.74%
Interaction	0.004	8.25%
N (Women)	218,578	
N (Men)	225,869	

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.

Table 6: Determinants of College Attendance for Women vs. Men

	College Attendance (Any)	
	(1)	(2)
	Women	Men
Demographics		
Black	0.072***	0.099***
	(0.006)	(0.008)
Hispanic	0.018**	0.032***
	(0.008)	(0.009)
Asian	0.005	0.060***
	(0.007)	(0.009)
Other Race	-0.000	-0.008
	(0.014)	(0.015)
FRPL	-0.059***	-0.069***
	(0.004)	(0.004)
Special Education	-0.084***	-0.069***
	(0.005)	(0.005)
English Language Learner	-0.083***	-0.028***
	(0.009)	(0.008)
High-School Experiences		
10th-Grade Math MCAS	0.062***	0.095***
	(0.002)	(0.003)
10th-Grade ELA MCAS	0.072***	0.087***
	(0.002)	(0.002)
Noncognitive Skill Index	0.061***	0.053***
	(0.002)	(0.002)
N	218,578	225,869
R^2	0.223	0.254

*p<0.1; **p<0.05; ***p<0.01

School-clustered standard errors in parentheses

Notes: This table shows linear probability models for the outcome variable college attendance (which includes attendance of any type of institution of higher education). The first column represents the determinants of voting for women, while the second column represents the determinants of voting for men. FRPL indicates free or reduced-price lunch status.

Table 7: BOK Decomposition of Gender College Gap

Decomposition Information		Share of the Gap
Probability of Attending College: Women	0.764	
Probability of Attending College: Men	0.651	
Gender Gap in College Attendance	**0.113**	**100%**
Endowments	0.036	31.58%
Student Demographics	0.006	5.08%
Cognitive Skills	0.025	22.03%
Noncognitive Skills	0.005	4.58%
Coefficients	0.080	70.24%
Student Demographics	-0.010	-8.61%
Cognitive Skills	0.004	3.43%
Noncognitive Skills	-0.000	-0.32%
Intercept (Unexplained)	0.086	75.74%
Interaction	-0.002	- 1.82%
N (Women)	218,578	
N (Men)	225,869	

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. Standard errors are clustered at the school level.

APPENDIX

Education and the Gender Gap in Voting

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

The primary data take advantage of the link between the Massachusetts Department of Elementary and Secondary Education (DESE) data and voter data created for use in Cohodes and Feigenbaum (2021). This appendix describes the data in more detail and the process of matching the DESE data to voting records, which is drawn from Cohodes and Feigenbaum (2021), with permission from the authors.

The primary difference between the matched student–voter data used in Cohodes and Feigenbaum (2021) and the subset in this study is the scope of students and years included. The Cohodes and Feigenbaum (2021) study focuses on Boston charter school lottery participants, while this study focuses on all Massachusetts students who were in 10th grade in 2002, 2003, 2004, 2006, 2007, and 2010 matched with voting records covering the 2008, 2012, and 2016 presidential elections.

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.1.1.1 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 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.1.1.2 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 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: Summary Statistics by Panel

	All (1)	Panel 1 (2)	Panel 2 (3)	p-Value (4)
(A) Student Characteristics				
Female	0.492	0.491	0.492	0.685
White	0.752	0.764	0.741	0.000
Black	0.089	0.088	0.091	0.000
Hispanic	0.108	0.099	0.116	0.000
Asian	0.047	0.046	0.048	0.004
Other Race	0.004	0.003	0.005	0.000
FRPL	0.252	0.216	0.287	0.000
Special Education	0.158	0.157	0.160	0.009
English Language Learner	0.053	0.046	0.060	0.000
(B) High-School Experiences				
10th-Grade Math MCAS	0.025	0.035	0.016	0.000
10th-Grade ELA MCAS	0.026	0.036	0.016	0.000
Noncognitive Skill Index	0.001	-0.193	0.188	0.000
Graduated MA HS in 4 Years	0.769	0.757	0.781	0.000
Graduated MA HS in 5 Years	0.795	0.778	0.812	0.000
(C) Education Levels				
No MA HS Diploma	0.131	0.151	0.112	0.000
MA HS Diploma Only	0.161	0.166	0.155	0.000
Some College	0.323	0.323	0.323	0.862
AA Only	0.045	0.042	0.047	0.000
BA or More	0.341	0.318	0.363	0.000
(D) Voting Outcomes				
Registered to Vote by Age 19	0.292	0.232	0.350	0.000
Ever Registered to Vote	0.800	0.764	0.834	0.000
Voted in First Pres. Election After Pot. College	0.382	0.354	0.408	0.000
N	444,447	218,579	225,868	

Notes: This table details the summary statistics of our sample of Massachusetts student records matched with voter files for each of the panels separately. The fourth column shows the statistical significance (p-value) from a difference-of-means tests between the two groups. (C) Education levels are exclusive and nonoverlapping, so individuals are counted only in their highest education category. FRPL indicates free or reduced-price lunch status. 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.

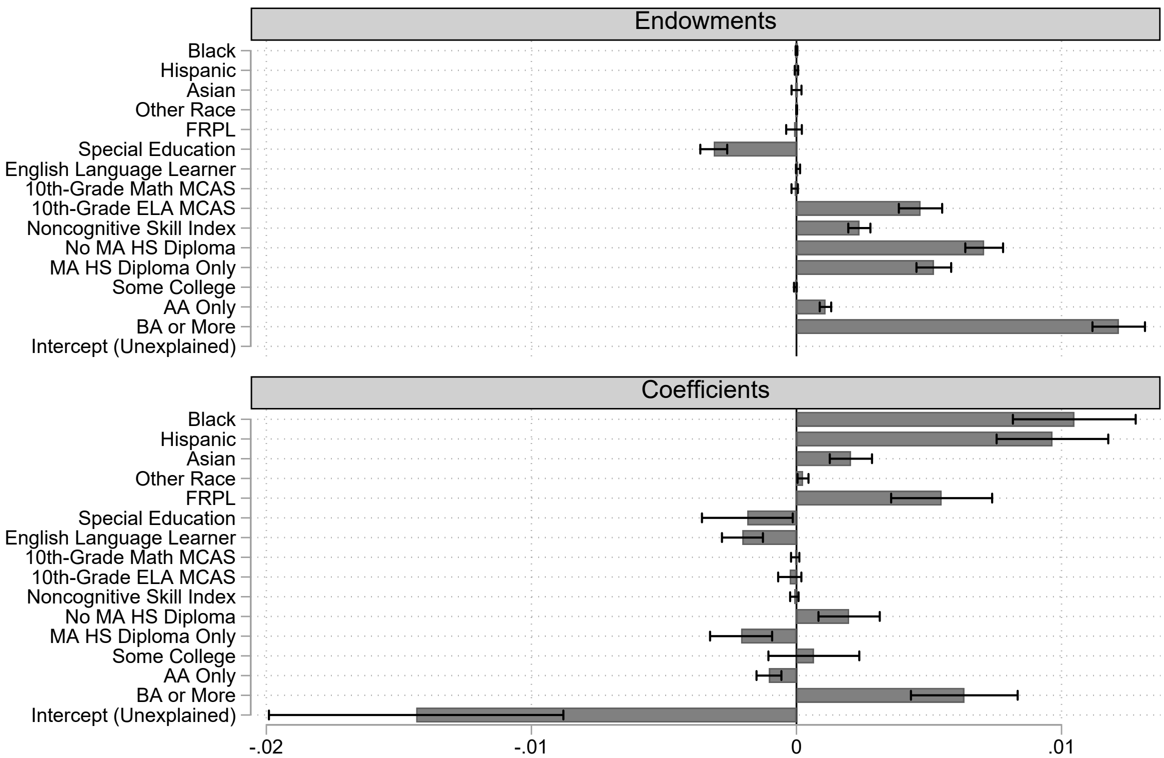
Table A.2: Determinants of Voting for Women vs. Men (with SAT & AP Measures)

	Vote	
	(1)	(2)
	Women	Men
Black	0.069***	-0.039***
	(0.007)	(0.007)
Hispanic	0.058***	-0.029***
	(0.007)	(0.005)
Asian	-0.078***	-0.118***
	(0.011)	(0.008)
Other Race	0.025	-0.059***
	(0.023)	(0.020)
FRPL	-0.038***	-0.058***
	(0.004)	(0.004)
Special Education	0.041***	0.039***
	(0.005)	(0.004)
English Language Learner	-0.079***	-0.046***
	(0.010)	(0.007)
10th-Grade Math MCAS	0.004	0.008***
	(0.003)	(0.002)
10th-Grade ELA MCAS	0.020***	0.016***
	(0.003)	(0.002)
Noncognitive Skill Index	0.019***	0.024***
	(0.002)	(0.002)
Took Any AP Course	0.028***	0.017***
	(0.005)	(0.005)
Took the SAT	0.070***	0.062***
	(0.004)	(0.004)
No MA HS Diploma	-0.034***	-0.065***
	(0.006)	(0.005)
Some College	0.088***	0.070***
	(0.005)	(0.004)
AA Only	0.167***	0.172***
	(0.009)	(0.008)
BA or More	0.199***	0.153***
	(0.006)	(0.006)
R^2	0.095	0.105
N	109,120	111,639

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

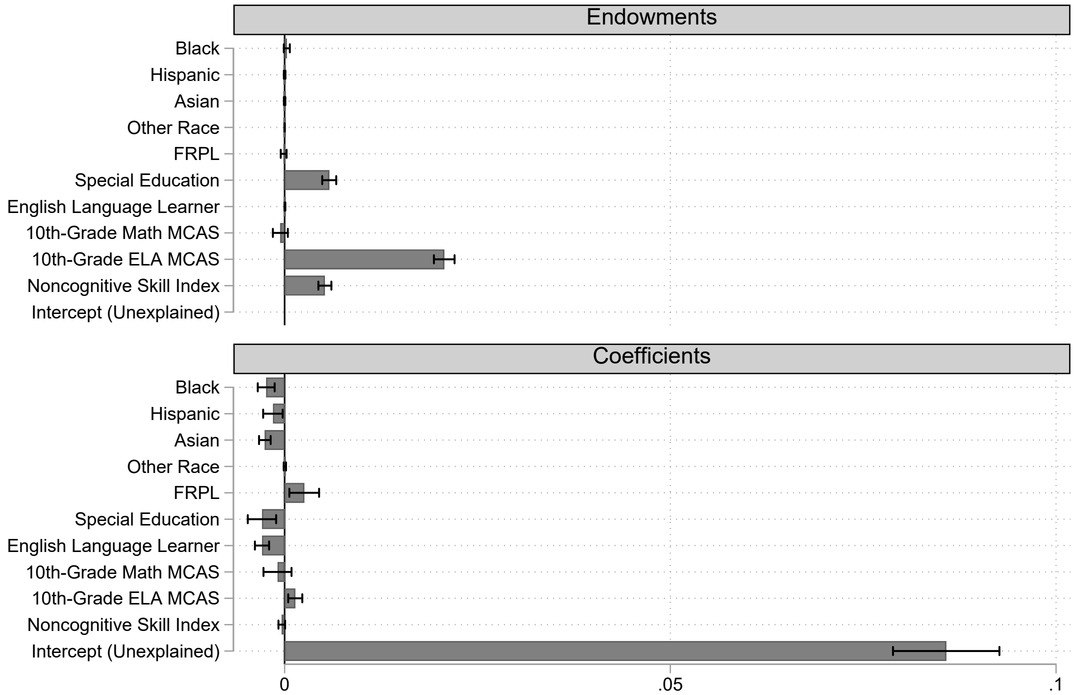
Notes: This table shows linear probability models for the outcome variable of voting in the first presidential election after potential college completion. The first column represents the determinants of voting for women, while the second column represents the determinants of voting for men. These models include dummy variables for whether the student took the SAT and whether she took any Advanced Placement (AP) courses, with the size of the samples accordingly significantly reduced from those in Table 2, but without significant changes in the coefficients or R-squared values. 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.

Figure A.1: Detailed Threefold BOK Decomposition of Gender Gap in Voting



Notes: This figure shows a detailed breakdown the portion of the gender gap in voting that can be explained by individual explanatory factors by gender. The top panel illustrates the proportion of the gap that can be explained by endowments, while the bottom panel illustrates the proportion of the gap that remains unexplained (coefficients).

Figure A.2: Threefold BOK Decomposition of Gender Gap in College Attendance



Notes: This figure shows a detailed breakdown of the portion of the gender gap in college attendance that can be explained by individual explanatory factors by gender. The top panel illustrates the proportion of the gap that can be explained by endowments, while the bottom panel illustrates the proportion of the gap that remains unexplained (coefficients).

Table A.3: Summary Statistics by Gender w/ MA High School Type

	All (1)	Women (2)	Men (3)	P-Value (4)
(A) Student Characteristics				
White	0.752	0.751	0.754	0.030
Black	0.089	0.091	0.088	0.001
Hispanic	0.108	0.108	0.108	0.915
Asian	0.047	0.047	0.047	0.987
Other Race	0.004	0.004	0.004	0.380
FRPL	0.252	0.253	0.251	0.183
Special Education	0.158	0.116	0.200	0.000
English Language Learner	0.053	0.052	0.054	0.044
(B) High School Experiences				
10th Grade Math MCAS	0.025	0.022	0.028	0.040
10th Grade ELA MCAS	0.026	0.147	-0.091	0.000
Noncogntive Skill Index	0.001	0.051	-0.047	0.000
Graduated MA HS in 4 Years	0.769	0.803	0.736	0.000
Graduated MA HS in 5 Years	0.795	0.823	0.768	0.000
(C) High School Type				
Traditional Public School	0.813	0.828	0.800	0.000
Charter School	0.012	0.014	0.011	0.000
Magnet School	0.049	0.051	0.048	0.000
Alternative School	0.005	0.005	0.005	0.823
Vocational School	0.118	0.100	0.135	0.000
Innovation/Pilot School	0.002	0.003	0.002	0.000
(D) 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
(E) Voting Outcomes				
Registered to Vote by Age 19	0.292	0.292	0.292	0.784
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	444,447	218,578	225,869	

Notes: This table details the summary statistics of our sample of Massachusetts student records matched with voter files for the whole sample, women, and men separately. The fourth column shows the statistical significance (p value) from a difference of means tests between the two groups. (C) High School Type is based on 10th grade enrollment. (D) Education Levels are exclusive and non-overlapping, so individuals are only counted in their highest education category. The MCAS stands for Massachusetts Comprehensive Assessment System, which are the statewide exams given to students in grades 3 through 8 as well as in grade 10. Noncognitive measure 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 A.4: Regression Taxonomy of Voting as a Function of Education Variables w/ MA High School Type

	Vote				
	(1)	(2)	(3)	(4)	(5)
Female	0.045***	0.037***	0.018***	0.018***	0.017***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Black	0.020***	0.067***	0.059***	0.041***	0.013***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.004)
Hispanic	-0.035***	0.023***	0.024***	0.016***	-0.001
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Asian	-0.072***	-0.092***	-0.099***	-0.105***	-0.122***
	(0.009)	(0.006)	(0.006)	(0.005)	(0.006)
Other Race	-0.052***	-0.013	-0.007	-0.011	-0.012
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
FRPL	-0.107***	-0.062***	-0.043***	-0.047***	-0.048***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
Special Education	-0.053***	0.022***	0.033***	0.034***	0.034***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
English Language Learner	-0.125***	-0.063***	-0.062***	-0.062***	-0.075***
	(0.008)	(0.007)	(0.006)	(0.007)	(0.007)
10th-Grade Math MCAS		0.038***	0.011***	0.010***	0.009***
		(0.002)	(0.002)	(0.001)	(0.001)
10th-Grade ELA MCAS		0.038***	0.020***	0.021***	0.021***
		(0.002)	(0.002)	(0.001)	(0.001)
Noncognitive Skill Index		0.040***	0.025***	0.026***	0.027***
		(0.001)	(0.001)	(0.001)	(0.001)
No MA HS Diploma			-0.057***	-0.058***	-0.054***
			(0.003)	(0.003)	(0.003)
Some College			0.090***	0.089***	0.088***
			(0.002)	(0.003)	(0.002)
AA Only			0.179***	0.179***	0.183***
			(0.005)	(0.005)	(0.004)
BA or More			0.203***	0.202***	0.199***
			(0.004)	(0.004)	(0.004)
Charter School				0.061***	
				(0.014)	
Magnet School				0.065***	
				(0.012)	
Alternative School				0.093***	
				(0.019)	
Vocational School				-0.009	
				(0.007)	
Innovation/Pilot School				0.093***	
				(0.027)	
School Fixed Effects	No	No	No	No	Yes
R-Squared	0.024	0.058	0.080	0.082	0.091
N	444,447	444,447	444,447	444,447	444,447

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
School-clustered standard errors in parentheses

Notes:This table displays results from various 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. The noncognitive skill measure is an index of measures used to approximate noncognitive skills (first principal component): attendance in 10th grade, on-time progression to 11th grade, and a school suspension indicator. School-type variables account for 10th-grade enrollment.

Table A.5: Determinants of Voting for Women vs. Men w/ MA High-School Type

	Vote	
	(1)	(2)
	Women	Men
Demographics		
Black	0.096***	-0.012**
	(0.008)	(0.006)
Hispanic	0.061***	-0.026***
	(0.006)	(0.004)
Asian	-0.084***	-0.124***
	(0.007)	(0.006)
Other Race	0.019	-0.039***
	(0.021)	(0.015)
FRPL	-0.036***	-0.057***
	(0.003)	(0.003)
Special Education	0.029***	0.037***
	(0.004)	(0.003)
English Language Learner	-0.083***	-0.043***
	(0.009)	(0.006)
High-School Experiences		
10th-Grade Math MCAS	0.007***	0.011***
	(0.002)	(0.002)
10th-Grade ELA MCAS	0.022***	0.020***
	(0.002)	(0.002)
Noncognitive Skill Index	0.027***	0.024***
	(0.002)	(0.001)
MA High-School Type		
Charter School	0.053***	0.065***
	(0.018)	(0.015)
Magnet School	0.079***	0.048***
	(0.011)	(0.014)
Alternative School	0.112***	0.070***
	(0.012)	(0.024)
Vocational School	-0.033***	0.006
	(0.011)	(0.006)
Innovation/Pilot School	0.120***	0.049***
	(0.029)	(0.018)
Educational Attainment		
No MA HS Diploma	-0.042***	-0.064***
	(0.004)	(0.003)
Some College	0.096***	0.086***
	(0.003)	(0.003)
AA Only	0.175***	0.192***
	(0.006)	(0.006)
BA or More	0.218***	0.191***
	(0.005)	(0.005)
R^2	0.073	0.087
N	218,578	225,869

*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.

Table A.6: Threefold BOK Decomposition of Gender Voting Gap w/ MA High-School Type

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	60.22%
Student Demographics	-0.003	- 6.45%
Cognitive Skills	0.005	10.12%
Noncognitive Skills	0.002	4.82%
MA High-School Type	0.000	0.36%
Educational Attainment	0.026	51.40%
Coefficients	0.015	29.64%
Student Demographics	0.022	44.93%
Cognitive Skills	0.001	1.09%
Noncognitive Skills	-0.000	-0.26%
MA High-School Type	-0.003	-6.90%
Educational Attainment	0.005	10.93%
Intercept (Unexplained)	-0.010	-20.15%
Interaction	0.005	10.11%
N (Women)	218,578	
N (Men)	225,869	

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.

Table A.7: Determinants of College Attendance for Women vs. Men w/ MA High-School Type

	College Attendance (Any)	
	(1)	(2)
	Women	Men
Demographics		
Black	0.070***	0.091***
	(0.005)	(0.006)
Hispanic	0.025***	0.038***
	(0.008)	(0.009)
Asian	0.002	0.050***
	(0.006)	(0.007)
Other Race	-0.000	-0.009
	(0.014)	(0.015)
FRPL	-0.051***	-0.057***
	(0.004)	(0.004)
Special Education	-0.083***	-0.067***
	(0.005)	(0.004)
English Language Learner	-0.088***	-0.035***
	(0.009)	(0.008)
High-School Experiences		
10th-Grade Math MCAS	0.059***	0.088***
	(0.002)	(0.002)
10th-Grade ELA MCAS	0.068***	0.080***
	(0.002)	(0.002)
Noncognitive Skill Index	0.063***	0.057***
	(0.002)	(0.001)
MA High-School Type		
Charter School	0.027***	0.031**
	(0.008)	(0.013)
Magnet School	-0.026**	-0.046***
	(0.012)	(0.014)
Alternative School	-0.010	-0.064***
	(0.016)	(0.019)
Vocational School	-0.098***	-0.175***
	(0.010)	(0.010)
Innovation/Pilot School	-0.025**	-0.007
	(0.012)	(0.057)
N	218,578	225,869
r-Squared)	0.227	0.269

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

Notes:This table shows linear probability models for the outcome variable college attendance (which includes attendance of any type of institution of higher education). The first column represents the determinants of voting for women, while the second column represents the determinants of voting for men. FRPL indicates free or reduced-price lunch status.

Table A.8: BOK Decomposition of Gender College Gap w/ MA High-School Type

Decomposition Information		Share of the Gap
Probability of Attending College: Women	0.764	
Probability of Attending College: Men	0.651	
Gender Gap in College Attendance	**0.113**	**100%**
Endowments	0.040	35.58%
Student Demographics	0.006	5.12%
Cognitive Skills	0.023	20.21%
Noncognitive Skills	0.006	4.96%
MA High-School Type	0.006	5.29%
Coefficients	0.077	67.97%
Student Demographics	-0.010	-8.89%
Cognitive Skills	0.004	3.43%
Noncognitive Skills	-0.000	-0.24%
MA High-School Type	0.012	10.24%
Intercept (Unexplained)	0.072	63.43%
Interaction	-0.004	- 3.55%
N (Women)	218,578	
N (Men)	225,869	

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. Standard errors are clustered at the school level.