4-12-2022

Effects of education on political engagement in rural Burkina Faso

Elodie Djemaï

Michael J. Kevane

Effects of education on political engagement in rural Burkina Faso*

Elodie Djemaï
Université Paris-Dauphine
Université PSL, IRD, CNRS, LEDa
Place du Maréchal de Lattre de Tassigny
75775 Paris cedex 16 France
elodie.djemai@dauphine.psl.eu

Michael Kevane
Santa Clara University
Dept. of Economics
Santa Clara, CA USA 95053
mkevane@scu.edu

This version: April 12, 2022

Abstract. Many African countries are both consolidating their democratic institutions and continuing to expand mass primary schooling. In this context, citizens may be interested in the broad general effect of education on political engagement. Recent social science work estimating this effect has not arrived at consensus, with researchers suggesting the relationship may be context dependent, and could vary from positive to negative. We apply an instrumental variable (IV) approach, using Afrobarometer surveys in Burkina Faso over the 2008-2019 period, merged with data on the timing of school establishment at the village level. Individual schooling attainment is instrumented by whether a school was established in the village of residence when the person was seven years old. The data is finer than recent papers that estimate the relationship using national-level quasi-experiments where education access changed across birth cohorts or where an indirect proxy measure of education access varied across regions and birth cohorts. We find that the relationship appears to differ by gender: men exhibit a substantial negative effect of education on engagement, while women exhibit no sizable relationship. The null effect for women may be due to low power, as there is less variation in education outcomes for women in rural areas. The results suggest that gender may be an important mediator of the direction and magnitude of the complex relationship between education and political engagement in polities with low overall levels of schooling.

Journal of Economic Literature classification: I25, D72, O55.

Keywords: Education, Voting, Natural Experiment, Africa, Burkina Faso.

* We are grateful to Clément d'Haultfoeuille for insightful discussions at an early stage of the paper, to participants at the Santa Clara University Economics Department seminar and participants at the Oxford Center for the Study

of African Economies annual conference poster session for helpful comments, and to Horacio Larreguy, and Pierre André for comments on an earlier draft. All errors of course remain responsibility of the authors. Data Availability Statement: The data used in this article can be obtained upon request to Afrobarometer (afrobarometer.org) and from the Institut géographique du Burkina Faso (IGB) which can provide a CD-Rom that geo-locates all villages in the country. Additional replication materials will be provided in the Online Appendix if the manuscript is accepted for publication. Declarations of interest: none.

1 Introduction

The advent and spread of mass schooling in the 19th and 20th centuries coincided with the movement towards open, democratic societies where most citizens became eligible to vote in free, fair, and secret ballot elections. The two processes were linked, as educators incorporated civics into primary and secondary school curricula, and as school boards were important local elections, often with broader suffrage than elections for higher political office. Social scientists have thus been interested in the relationship between the acculturation that children and young adults undergo in formal schooling and their voting behavior and other expressions of political engagement later in life, especially when significant fractions of the population are not literate. Intuition suggests that the effect may be dependent on both the electoral context and the specific acculturation that happens in and is associated with formal schooling. Citizens who are more educated may be acculturated to view political engagement as an obligation and virtue of citizenship, but they may also be more likely to engage in critical thinking, and take a "rational" approach to evaluating the marginal benefits of voting relative to marginal costs. Moreover, political regimes and political parties may make political engagement more or less rewarding for educated citizens relative to uneducated citizens. To take an extreme case, one might imagine that in a political context where power is exercised by a peasant-based social movement, book lovers learn to keep a low profile.

The acculturation and changed incentives to be politically engaged induced by formal schooling may vary by gender, in different societies and in different contexts. Communities with social norms that proscribe women from working outside the home or household farm may both discourage their education and their political engagement. Indeed, many early democracies, such as the United Kingdom, the United States, and France, limited the vote to men for more than a century. While after independence most former colonies in Africa adopted universal suffrage, in practice in most countries social norms discouraged women's participation (Amoateng, Kalule-Sabiti, and Heaton (2014); Tripp (2001)).

Estimating the relationship between schooling and political engagement, and how that link might vary by gender, is difficult. Both schooling attainment and engagement may be influenced by unobserved variables. For example, community-level norms might influence both individual engagement and also individual schooling decisions. Localities with strong norms favoring engagement might also be strongly in favor of formal education. Communities more likely to be politically engaged may also be more likely to provide, through collective action processes, opportunities for education.

There have, however, been several recent empirical analyses estimating the relationship between schooling and political engagement. Most relevant for this paper, André and Maarek (2020) used individual-level data and dates of school founding to estimate the effects of education on political engagement in Mali. Croke et al. (2016) and Larreguy and Marshall (2017) used Afrobarometer survey data to investigate the causal effect of education on political participation in Zimbabwe and Nigeria, respectively. They argued that relatively unanticipated large expansions of universal primary schooling (in 1980 and 1976, respectively) could be interpreted as natural experiments. Other papers have used similar quasi-experimental or natural experimental contexts to estimate causal effects of education expansion on political engagement (Blaydes (2006); Wantchekon, Klašnja, and Novta (2015); Dang (2019); Parinduri (2019)). The question has also been addressed at an aggregate level using cross-country variation in experiences to estimate the effects of education on electoral participation and a variety of attitudes regarding democracy and political engagement (Acemoglu et al. (2005); Paglayan (2021); Diwan and Vartanova (2020)). A number of papers have estimated the micro-level relationship by leveraging experimental settings where levels of education attainment varied with program treatments in the context of randomized control trials (Friedman et al. (2016); Kuenzi (2006)).

In this paper, we estimate the effect of education on political engagement in Burkina Faso, separately for men and for women. We use five waves of Afrobarometer surveys collected in Burkina Faso over the 2008-2019 period Afrobarometer (2021). This individual-level data is merged with administrative data on school establishment from 1900 to 2004 and with geographic data on the locations of all villages in Burkina Faso.

Burkina Faso may be a good setting to estimate the effects of education on engagement and thus add to the growing evidence base. Access to primary schools, among the lowest in the world, accelerated in the early 1980s. The political context over the 1991-2020 period was at the "semi" end of semi-authoritarianism, with gradual and significant liberalization of political space after a 1998 social movement protested the assassination of independent journalist Norbert Zongo, and especially after a street uprising in 2014 that ousted longtime president Blaise Compaoré. Women have had both lower levels of schooling and lower levels of political participation (as measured by voting participation and other indicators).

Since individual schooling attainment may be correlated with unobserved individual or community propensities to be politically engaged, we use the timing of school establishment in the locality of residence of the individual, relative to the individual's age, as an instrumental variable. That is, the estimation strategy is to consider access to schooling as an instrumental variable for educational attainment, in estimating the relationship between schooling attainment and political engagement, including voting behavior. As there are several respondents for each location in the sample, each individual born at a different time, the method controls for likely persisting locality-level relationships between local preferences of valuing schooling and local preferences about political participation.

We conclude, from the range of estimates, that the magnitude of the relationship between education attainment and political engagement is quite negative for men, and apparently negligible for women. As with most statistical analysis of observational data, there is a veritable "garden of forking paths" regarding choices made by researchers (Gelman and Loken (2014)). We present a range of results, and eschew selecting a "preferred" specification based on post-hoc results. The sizable negative relationship for men may simply be an unlikely and possibly idiosyncratic artifact of the particular sample, and the lack of statistically significant relationship for women may be indicative that the statistical test lacks power. We find that the Afrobarometer sample for a single country such as Burkina Faso (about 4,300 individuals observed in rural areas, with about half in each gender, by design) is likely large enough, for estimation techniques such as instrumental variables in a repeated cross-section setting, to detect reasonable sized effects separately for men. For women, however, there is less variation in the education outcome, and so the statistical test of the relationship is somewhat under-powered for relevant effect sizes.

The remainder of the paper is organised as follows. Section 2 discusses the relationship between schooling and political engagement and Section 3 presents the context of Burkina Faso regarding schooling access and achievement over time and democratic elections. Section 4 describes the Afrobarometer survey and the school location data. The estimation strategy is presented in Section 5 and the regression results appear in Section 6. Extensive robustness checks are presented in several appendices. Power considerations are investigated in Section 7. Section 8 supplements the analysis of the effects of education on political engagement with an analysis of the effects of education on political attitudes and preferences. Section 9 concludes.

2 Schooling and political engagement

Formal education in most societies is designed to deliver both critical reasoning skills and a bundle of attitudes, ideas, or memes that acculturate young persons. This bundle when unpacked might consist of patriotism, civic virtues, idealism, nationalism, knowledge about history and society, elitism, class consciousness, tolerance, respect for differences in values, etc. To these might be added change in executive-level cognitive capabilities, psychological development of feelings of self-efficacy, empowerment, and confidence. Some education programs are intended to have a meta-habit of considering a variety of contradictory thoughts (ideas, or meanings) and approaching decisions by thinking and evaluating them in the light of contextual knowledge; that is, specific attitudes and preferences to be acculturated are not intended but rather are emergent properties. Moreover, schooling is also about obedience and discipline: showing up on time, completing assignments on time, taking tests, being quiet when the teacher is talking, not questioning the teacher, etc. Because of the complexity and bundled nature of schooling, political psychology does not have a set of presumptively valid hypotheses concerning the general effect of formal schooling on political engagement (Basu (2002); Bruch and Soss (2018); Lieberman and

Zhou (2020); Soboleva (2020); Wolak (2020)).

A complication in theorizing the effect of schooling on political engagement is the temporal remove between the occasion of schooling and the occasions of political engagement. Schooling occurs during a person's youth, typically from age 6 to 18, and, for some, into university. Political engagement for most people starts after age 18, when persons become eligible to vote, and continues through life, and indeed evolves through life. As Larreguy and Marshall (2017):p. 400 noted: "it is important to differentiate the skills and values learned at school (...) from downstream effects such as increased income, community interactions, and empowered social status." The suggested differentiation is likely more difficult than envisaged. Education attainment by age 18 influences the path of life experiences and lifelong learning of a person, which in turn influences the person's (usually) unobserved mindset as an older person (which is likely still influenced by their mindset developed during their years of pre-adult schooling). The reasoning skills, and the bundle of attitudes, capabilities, and ideas acquired during schooling presumably initiate a trajectory of continuity and change of mental habits, predispositions, and knowledge components over the adult life-cycle. Schooling clearly changes, for example, the economic and social opportunities for individuals. An effect of these changed opportunities is likely higher lifetime income. Higher income levels may change the mental calculus of opportunity costs of engagement and preferences for political parties who may have platforms that address the interests of different income categories. Moreover, age and life experiences produce wisdom (e.g., the common saying in West Africa that when an elder dies it is as if a library had burned to the ground). The trajectory of acquisition of wisdom is presumably influenced by schooling attainment. The capability and habit of reading, for example, might be expected to change the wisdom of a person. Literate persons, for example, have relatively easier access as adults to newspapers and books, which may be very efficient transmitters of memes, thus shaping mindsets.

The underlying premise of some research has been similar to that of the "historical persistence" literature: events that happened long ago persisted and left legacies in other time periods (Cirone and Pepinsky (2021)). The "effect" of schooling attainment on political engagement is then a shortcut phrasing for a process that may include many mechanisms. Education may have a direct effect in changing attitudes towards political engagement, an indirect effect in setting a trajectory for changing mindset during adulthood, and an indirect effect in changing socioeconomic circumstances of adult life. Numerous other indirect mechanisms may be activated by formal education: different networks of classmates, different urban migration patterns, and improved access to professional occupations, each with their own acculturation processes. The literature at present is unable to separate the various mechanisms that link schooling to political engagement outcomes.

Most studies on political engagement include gender as a possible determinant of engagement and find a gender gap in favor of men being more likely to participate than women, even though the gap tends to disappear in high-income countries. Using the 2005 Afrobarometer surveys, Coffe and Bolzendahl (2011) found a significant gender gap in collective action engagement (attending a community meeting and joining others to raise an issue), but no gender gap in terms of registration to vote, in most of the 18 countries considered. Isaksson, Kotsadam, and Nerman (2014) found evidence along the same lines; pooling the Afrobarometer survey data from 20 African countries, women tended to be less likely to vote and to join with others raise an issue, the gender gap being four times larger in the former than the latter. In the case of Burkina Faso, Ozdemir, Ozkes, and Sanver (2021) found no significant relationship between gender and either voter registration or turnout in Burkina Faso.

The question is not only whether women and men have different behaviors on average, but also how to explain the gender gap and whether the effects of determinants of engagement are different across gender. In high-income countries, the gender gap is mostly explained by resource endowments (e.g., education), employment, and differences based on religious affiliations. Using the Afrobarometer surveys collected in 20 African countries, Isaksson, Kotsadam, and Nerman (2014) suggested that clientelism, restricted civil liberties, economic development, and gender norms were important drivers of the participatory gender gap. They found that the level of education was not significantly related to the probability of voting in the previous election while it was significantly correlated with the response to a question about whether the person had "raised an issue" with others in the community. However the size of the correlation did not differ by gender. This result is in line with additional analysis from Larreguy and Marshall (2017), who mentioned that their instrumented coefficient of education on civic and political engagement did not vary by gender. André and Maarek (2020) also split their sample by gender, and found that the effect of education was no longer significant for each of their 10 measures of political participation.

3 Schooling and democratic elections in Burkina Faso

Burkina Faso (formerly Upper Volta) became an independent country in 1960, after more than 60 years of French colonial rule. Complex political machinations led to the election in the national territorial assembly of Maurice Yaméogo as president prior to independence. Yaméogo then proceeded to dismantle open electoral competition, replacing it with a one-party state (as in many other countries of the era). General Sangoulé Lamizana toppled Yaméogo in 1966 through a bloodless coup, following urban unrest and strikes. There were several legislative and presidential elections, and several constitutional referenda, during the unstable period of quasi military-civilian rule that lasted 25 years from 1966 until 1991. In 1991, President Blaise Compaoré, who had taken power with others in a military coup in 1983, initiated a gradual but apparently steady transition to civilian rule. Compaoré retired from the military, and most leadership posts were given to civilians. His semi-authoritarian regime was, however, backed by a strong presidential guard unit of the military, and endured until a popular uprising in 2014. A brief transition government with joint civilian-military leadership handed power back to an elected civilian government in January

2016. A low-level insurgency that began in the north of the country has led to more than 1.5 million displaced persons by 2022, and disrupted the 2020 presidential and legislative elections in several regions in the north and east. President Roch Marc Chistian Kaboré, re-elected in 2020, was ousted in a coup d'état in January 2022. The coup was met with widespread apathy among the citizenry.

For presidential elections, turnout during the Compaoré regime was fairly low: 25% in 1991, 56% in 1998, 58% in 2005, and 55% in 2010. Turnout was similar for national legislative elections held in 1992, 1997, 2002 and 2007 (Elischer (2013); Santiso and Loada (2003); Loada and Santiso (2002)). The 2006 countrywide municipal elections where rural councils were first elected had turnout of 49%. The joint legislative and municipal elections of 2012 had a turnout rate of 76%. Turnout was 60% in the first post-Compaoré presidential and legislative elections of 2015, with 3,309,988 voting out of 5,517,015 registered voters. Turnout was considerably lower in the presidential and legislative elections of 2020, with 2,972,590 votes recorded for 5,918,844 registered voters, for turnout of about 50%.

There has been limited research on the determinants of voter registration and turnout in Burkina Faso. Ozdemir, Ozkes, and Sanver (2021) used original survey data to estimate the determinants of registration and voting in the 2015 elections in Burkina Faso. They suggested that socio-demographic variables such as income, education, and marital status were significant determinants of voter registration, but only ethnicity was a significant determinant of turnout on election day.

Burkina Faso had relatively low rates of primary and secondary schooling into the 1990s, typical for former French colonies of the Sahel region. The French colonial authorities had little inclination to spread schooling widely. Some factions in the colonial administration opposed the Catholic missionaries who sought to expand schooling. The missionaries were constrained, in any case, by the limited number of White Fathers taking up missionary work in the colony. When Protestant schools started spreading in the 1930s, the Catholic missionaries often allied with sympathetic administrators to slow the spread of Protestant schools. According to Maxime Compaoré, as reported by Nabaloum (2012), the first formal non-Quranic schools in the colony were established by French military authorities in 1898 in Bobo-Dioulasso and Boromo, followed in 1899 by a school in Ouagadougou. Subsequently, schools were established in Koupela, Léo and Koury in 1900, in Dori in 1901, in Gaoua in 1902, and in Tenkodogo in 1903. By 1920 there were 17 primary schools : 4 *écoles régionales* in Ouagadougou, Bobo-Dioulasso, Dédougou, Dori, 2 *écoles de village à deux classes* in Koudougou and Ouahigouya, and 11 *écoles de village à 1 classe* (in Tenkodogo, Fada N'Gourma, Diapaga, Say, Kaya, Téra, Léo, Diébougou, Gaoua, Boromo, and Banfora). In 1947, the colony had 89 primary schools. Of these 27 were private, and 10 were girls schools. By 1960, according to Maxime Compaoré, the territory had 224 public primary schools and 130 private primary schools.

After independence, the first presidents, Maurice Yaméogo, himself a former schoolteacher, and

Sangoulé Lamizana, who had only finished primary education level (typical for ex-colonial military officers), appeared to have been relatively uninterested in investing in schooling. Around the late 1970s, the groundwork for more rapid expansion of primary schools was laid. The military coups and period of interregnum in the early 1980s were followed by the 1983 revolutionary regime of Captain Thomas Sankara, who might have been expected to accelerate even further the expansion of schooling. However, Sankara was opposed by one of the two important teachers unions, and he fired more than 1,000 teachers in 1984, slowing the rate of primary school expansion from what it might have been.

We use the terms education and schooling interchangeably. In some countries and regions, religious schooling with curricula regulated by the government has been as important as public, secular schooling. Burkina Faso has seen both Christian-oriented private schools and Muslim-oriented private schools (*écoles Franco-Arabes, madrasas*), and so potential differential effects might have been, and possibly remain, significant. However, the large majority of primary schools have been and remain secular public schools in the entire country.

4 Data description

4.1 Afrobarometer survey

We use five rounds of the nationally representative Afrobarometer survey collected in Burkina Faso (2008, 2012, 2015, 2018, 2019). The sampling design was in two steps. First, enumeration areas were randomly selected, and then in each sampled enumeration area, eight respondents were randomly selected to participate in the survey. 1200 respondents were surveyed in each round, for a total of 6,000 respondents. For each of the enumeration areas included in the survey rounds, we observe the name of the locality and the GPS coordinates. This enables us to match the localities to the school list, and so determine whether there is a primary school in the sampled villages and if so, the date of establishment.

The Afrobarometer surveys are repeated cross sections, and the sample is independently drawn in each wave, so that the probability of observing an enumeration area in more than one round is rather small. It turns out that we observe a total of 576 villages: 459 villages are observed only once, 76 villages are observed in two rounds, 21 in three rounds, 20 in four rounds and 0 in five rounds. If instead we use the GPS coordinates to define the enumeration area, we observe a total of 578 unique GPS points: 500 are observed only in one round, 66 are observed in two rounds, 12 in three rounds, and 0 in four or five rounds. Figure 1 shows the locations of each village included in at least one of the five survey rounds.

The analysis sample for this paper consists of rural residents only, and those born on or after 1950, leaving 4,520 respondents. We exclude residents of large towns and cities (the two principal large cities of Ouagadougou or Bobo-Dioulasso account for about 70% of those excluded) because respondents in almost all urban areas were born after school founding dates, which in most towns

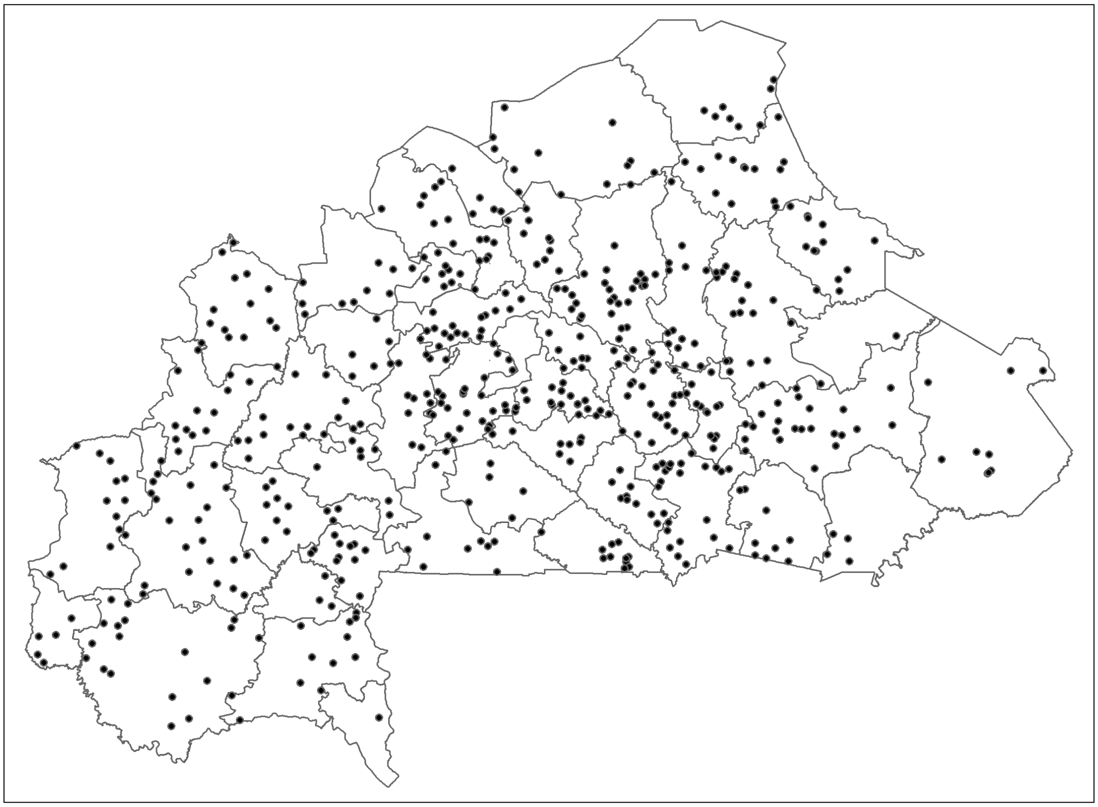

Figure 1: Location of enumeration areas included in Afrobarometer surveys, Burkina Faso

Source: Afrobarometer, five rounds (2008, 2012, 2015, 2018, 2019). Province boundary lines are shown (45 provinces). As discussed in text, most villages are observed in only one round of the survey.

and cities were before 1940. We include the few towns that saw their first schools established after 1940. The sample is further restricted to only include respondents born after 1950 in the analysis sample.

4.2 School and village geolocations

The *Institut géographique du Burkina Faso* (IGB) provides a CD-Rom with the official list of all village units (administrative and informal), the latitude and longitude of each, and shapefiles for *communes*, provinces, and regions. There are almost 11,000 unique administrative entities listed for the 13 regions, 45 provinces, and 351 *communes* (including the two urban communes of Ouagadougou and Bobo-Dioulasso). A *commune* is the French language designation for the lowest level of national administration, and is equivalent to the old designation of *département*. A typical rural commune consists of about 30 villages, with population of about 30,000 people, usually with a central village as the seat of local administration (known as the *chef-lieu*).

Data on the year of school opening comes from a document, *Répertoire des écoles publiques*

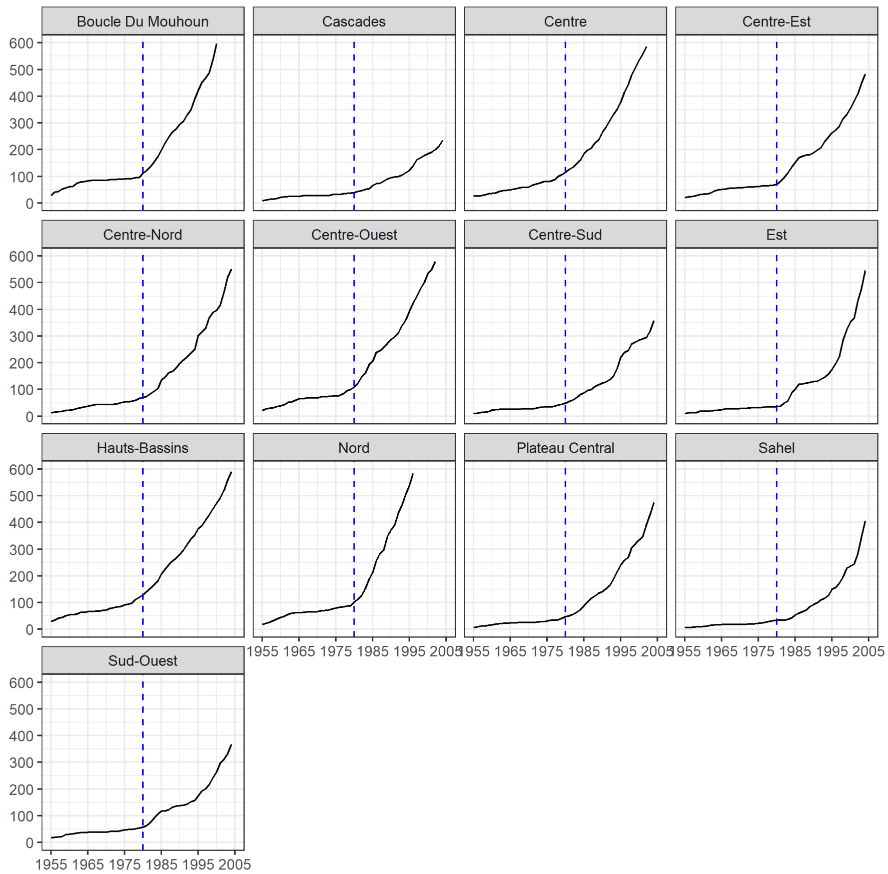

Figure 2: Cumulative number of schools established in Burkina Faso, by region, 1955-2004

Note: The 194 schools established before 1955 are counted as 1955 in the graph. Source: *Répertoire des écoles publiques et privées du Burkina Faso en 2004-2005* produced by the *Direction des études et de la planification of the Ministère de l'enseignement de base et de l'alphabétisation* (MEBA) in March 2006.

et privées du Burkina Faso en 2004-2005 (called *Répertoire* hereafter) produced by the *Direction des études et de la planification* of the *Ministère de l'enseignement de base et de l'alphabétisation* (MEBA) in March 2006. The document contains a listing of 6,913 public and private primary schools in the country, whether the school was urban or rural, the number of students and classrooms in each school, and the opening date of the school. There are no missing values in the report for founding dates. The number of schools listed accords exactly with another annual report produced by the MEBA on the number of schools in the country in 2005. It seems unlikely that any established primary school, whether public or private, would have not been included in the list.

Over the 6,913 schools, 85.5% were public schools, 7.0% arabic or franco-arabic schools, 4.2% private secular schools, and 3.0% private Catholic or Protestant schools. There were six levels in primary education in Burkina Faso, but most rural schools only had three classrooms, with each classroom teaching to two levels. In our data, if we restrict the sample of schools to the 5,447 rural schools, in 2005 about 15% had six classrooms (a few had more than six), about 8% had between four and five classrooms, about 50% had three classrooms, 13% had two classrooms, and 14% had a single classroom. Over the period 1898-1955, 194 schools were established in the country, while from 1955 to 2004 and additional 6,719 schools were established. There was a break in the time series around 1980 when school establishment accelerated. (Statistical tests of breaks in the time series at the national level and by region confirm the clear visual pattern.) Figure 2 shows that the regional pattern of school establishment over the period 1955 to 2004 was quite varied. Many regions did see acceleration around 1980, but some did not. Regions in the southwest of the country (Cascades, Centre-Sud, and Sud-Ouest) and in the east and northeast of the country (Centre-Est, Est, Sahel) saw fewer schools established in general, and took longer to accelerate. Several core regions saw dramatic acceleration around 1980 (Boucle du Mouhoun, Centre-Nord, Centre-Ouest, Haut-Bassins, Nord).

4.3 Merging the Afrobarometer, IGB, and school founding data

Appendix A describes the extensive data cleaning that was necessary for the school list to be matched to both Afrobarometer respondent villages and to geo-coordinates.

We first merge the Afrobarometer data with the official IGB list of village names. We verify that the latitude and longitude for those that match are reasonably close. There are a few choices that must be made. Some of the Afrobarometer sites are listed as two places (literally "place1 + place2"). We chose the first place if that matched and the second place if the first did not match (records do not distinguish which place the person resided in). Spellings are not standardized in Burkina Faso, and the many different unwritten languages mean that there are many different spellings of place names. We used a place list concordance to match place names to standardized spellings. We then merge the 6,913 school village names from the *Répertoire* to the 6,000 merged Afrobarometer and IGB data. Again, there are similar issues.

Separately, we merge the list of Afrobarometer place names with the IGB list of place names. That is then merged with the *Répertoire*. Then for every Afrobarometer place name and every year after 1950, we construct a variable equal to the number of primary schools within 30 kilometers of the village, weighting each school by the simple Haversine distance between the village and the school. The weight for each school is the inverse of the distance squared. (We add 1 km to each distance, to handle the 0 km cases.) This weighted distance measure is then merged back to the Afrobarometer data for each respondent.

An assumption for this analysis is that a school, once established, does not close. It is possible that there were schools that had opened and then closed, and thus were not on the list as they were not used during the school year 2004-2005. Public school closings appear to have been rare, and the occasional public school closing does make national news. (Since 2016, an insurgency in many parts of the country, but especially in northern regions, has closed several hundred schools and displaced more than 1.5 million persons.) It is possible that some private schools were not on the list because they were opened and then closed before 2004, since occasionally the government inspects schools and declares some to be non-conforming with existing rules regarding school facilities and instructional capabilities.

A better measure of school accessibility would take into account geographic obstacles (a forest or river) that make it unlikely that a child would attend a primary school. The size of the schools would also matter. As noted earlier, many of the first schools were simple one-room affairs, and later schools were often simple straw shelters that shaded the students but offered little protection from wind and dust. These one-classroom schools may or may not have welcomed students from other villages. In some cases, even the one-classroom schools were under-subscribed as families preferred to have their children continue to work. In other places, one-room schools were crowded and pupils from other villages were not accepted.

In summary, we calculate two measures of access. First, we use the timing of the establishment of a primary school in the person's village itself. The "treatment" variable takes on value one if there was a primary school in the village of residence when the person was seven years old (and thus eligible for schooling). A second more complex treatment measure is the weighted sum of schools around the village when the person was seven years old, where the weights are the inverse of the square of the distance.

4.4 Descriptive statistics from Afrobarometer

Education The main explanatory variable of interest is education attainment. In the Afrobarometer, respondents are asked to report their level of education on a 10-point scale: 0 no formal education, 1 some informal education, including Koranic schooling, 2 some primary education, 3 primary completion, 4 some secondary, 5 completed secondary, 6 some post-secondary, 7 some university, 8 graduation from university level, and 9 some post-graduate work. The left panel of Figure 3 shows the distribution of the original scale.

We change this initial scale in two ways. First, we top-censored the scale because only 1.3% of the respondents report a level strictly higher than secondary school completion. Secondary school completion alone in rural areas also remained very low (1.8%) through the time period considered in the paper (here completing secondary stands for the seven years after the primary school). Second, even though the proportion of respondents who report informal education is sizable (11.9%), we do not retain this distinction in the further analysis. Informal education in this context refers usually to informal sites of Islamic instruction (often called *maktab*, *kuttāb* or *daar*), where students learn to memorize the Qur'an and write in Arabic. Students in *maktab* also learn the *hadith* and of course life lessons from their imam. Our scale for education level is then a six-point scale: 0 no formal education or some informal education, 1 some primary education, 2 primary completion, 3 some secondary, 4 completed secondary, 5 some post-secondary, some or complete

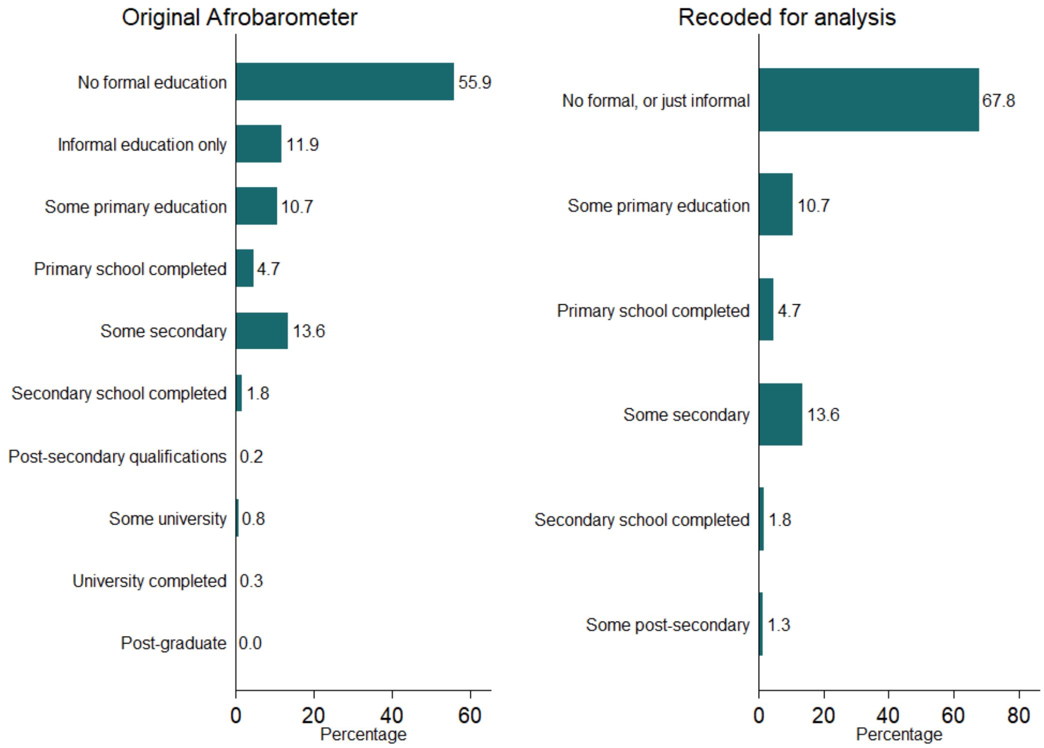

Figure 3: Distribution of scale of education level in Afrobarometer sample, original and recoded for analysis

Source: Afrobarometer data, various rounds. Sample: Rural residents born after 1950.

university, and for some post-graduate. The right panel of Figure 3 shows the distribution of these indicators of attainment.

We also construct three dummy variables: a dummy variable for having some education; a dummy variable for whether the person completed primary school; and a dummy variable for whether the person had some secondary or more.

Table 1 reports descriptive statistics for rural residents born after 1950, pooling the survey rounds. The average values are also calculated by gender. The Afrobarometer sampling strategy sought gender balance in every enumeration areas, but there are slightly more females than males as we removed the respondents born before 1950. In terms of schooling attainment in rural areas, about 32% of the respondents had some schooling, 21% had completed primary, and 17% of respondents had some secondary. The mean education level was quite low, 0.75. Women had lower schooling than men, in the sample.

As shown in Figure 4, the low overall average education attainment levels mask large differences across birth cohorts. In the left panel, the proportion of respondents having done some schooling, by 3-year birth cohorts, is plotted for the 1950-2000 cohorts. Men consistently were about 10-20

Table 1: Descriptive statistics of sample, according to gender

	(1) All	(2) Male	(3) Female	(4) p-value
Year of birth	1978.31	1975.63	1980.86	0.000
	(12.250)	(12.537)	(11.400)	
School before age 7	0.43	0.40	0.47	0.000
	(0.496)	(0.490)	(0.499)	
Number of schools, weighted by distance	1.03	0.94	1.11	0.002
	(1.802)	(1.756)	(1.842)	
Closest school distance (in km.)	5.20	5.92	4.50	0.000
	(6.418)	(6.842)	(5.897)	
Some schooling	0.32	0.36	0.28	0.000
	(0.467)	(0.481)	(0.450)	
Completed primary	0.21	0.24	0.19	0.000
	(0.411)	(0.428)	(0.392)	
Some secondary	0.17	0.18	0.15	0.003
	(0.373)	(0.388)	(0.359)	
Education level	0.75	0.85	0.65	0.000
	(1.259)	(1.333)	(1.176)	
Index of engagement	2.03	2.29	1.78	0.000
	(1.044)	(0.920)	(1.093)	
Voted in recent election?	0.74	0.80	0.68	0.000
	(0.439)	(0.397)	(0.468)	
Member of community group?	0.44	0.47	0.40	0.000
	(0.496)	(0.499)	(0.491)	
Participate in community meeting?	0.67	0.77	0.58	0.000
	(0.469)	(0.421)	(0.494)	
Join with others on issue?	0.60	0.71	0.50	0.000
	(0.489)	(0.454)	(0.500)	
Observations	4520	2189	2331	4520

Notes: Afrobarometer Burkina Faso data, various rounds, rural residents, birth year after 1950. Standard deviations in parentheses. Participation in community meeting only observed in rounds 4-7. Final column gives the p-value for t-test of difference in means between the two groups.

percentage points more likely to have had some schooling. The right panel gives the average education attainment level, by gender and birth cohort. Again, the discrepancies by gender are visible. For the sake of comparison, education attainment levels and completion rates in urban areas, who are not included in the analysis sample, were considerably higher than in rural areas. In urban areas, for the Afrobarometer sample, overall about 60% of respondents had completed primary, and 25% had completed secondary. The proportions of respondents who had completed secondary education went from close to zero for those born during the 1950s to 30% for those born in the 1990s in urban areas, but only about 5% for those in rural areas completed secondary, even in the 1990s.

Access to school The variable *School before age 7* in Table 1 is a dummy variable taking on value one if there was a primary school in the village of residence established before the person was seven years old. Here the schools considered are either public or private primary schools.

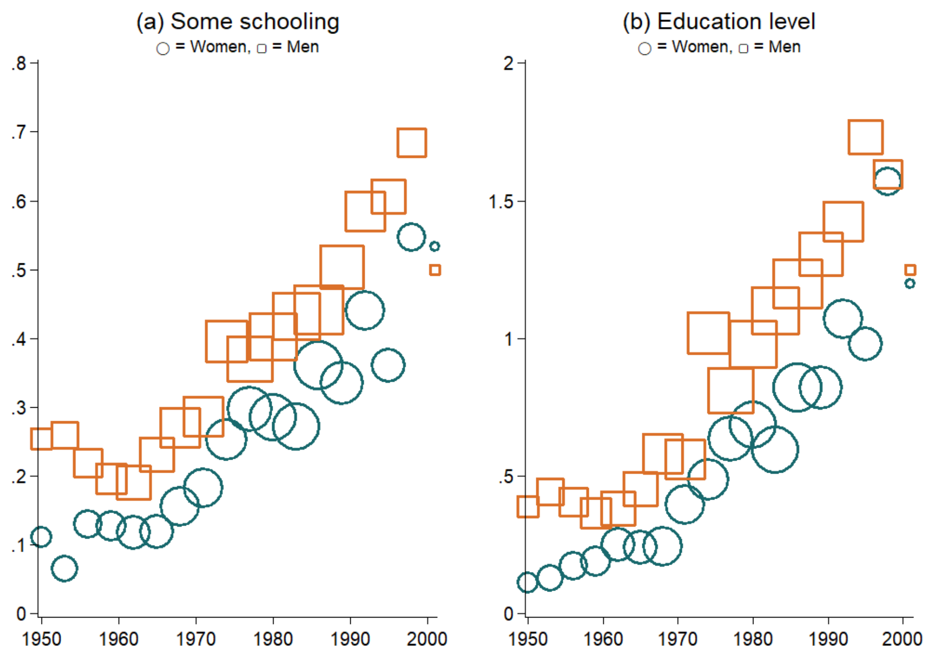

Figure 4: Rate for having done some schooling, and average education level, by gender and binned birth year cohort, Burkina Faso

Source: Afrobarometer data, various rounds, rural residents, birth year after 1950. Respondents grouped into 3-year bins by birth year. Markers are sized by number of respondents; there are fewer respondents in the early and later birth years.

On average, about 43 percent of rural respondents born after 1950 had schools in their villages when they were age seven. The variable *Number of schools, weighted by distance* is the number of primary schools in a radius of 30 km around the village of residence at the time the person was age seven, weighted by dividing by the distance squared (we add 1 km to the distance so as to not divide by 0). This is 1.03 for the entire sample, suggesting that on average over the time period the typical person had reasonable access to school not far from their village (the number varies considerably over time and by village). The average age of males was about five years higher than females. Since women are on average younger in the sample, it is not surprising that the values of these schooling access variables are higher. There are more schools proximate to the village of residence at age seven of women than on men. Likewise, the table reports the distance, in kilometers, to the nearest school. This averages 5 km overall, and is somewhat higher for men and lower for women, again due to the difference in average age of the two kinds of respondents.

Political engagement The Afrobarometer survey contains a variety of measures of political engagement. We focus on an index that aggregates three measures of engagement. The measures are: (1) did the person vote in the previous election; (2) had the person participated in a community meeting (scaled from 1 to 4); and (3) whether the person met with other persons to raise an issue. The two questions about participating were: *Here is a list of actions that people sometimes take as citizens. For each of these, please tell me whether you, personally, have done any of these things during the past year. If not, would you do this if you had the chance: Got together with others to raise an issue? Attended a community meeting?*. The participation questions are recoded to be dummy variables.

The question of whether the respondent voted in an election referred to the last election that took place before each survey round. In round 4 (2008), individuals were asked about the legislative elections of May 2007; in round 5 (2012) about the presidential elections of November 2010; in round 6 (2015) about the legislative and municipal elections of 2012; and in rounds 7 and 8 (2018 and 2019) about the presidential and legislative elections of 2015. The question had a large number of possible responses, ranging from the person being too young to vote, to not having registered to vote, to having decided not to vote, to not finding the polling station, to actually voting. We have recoded the answers to be a simple dummy variable for whether the person voted. If the person indicated in response that they were too young to vote, and the person's age was 21 or younger at the time of the survey, we have recoded that variable as missing, since the question is not pertinent. If they said they were too young to vote but they were older than 21 at the time of the survey, we have recoded them as not voting. About 130 (or 3.7% of respondents in the analysis sample) were indeed too young to have voted.

The index of engagement is the simple sum of the three dummy variables.

In addition, we shall also examine a question about the extent of membership in a community group (scaled from 1 to 4), that was asked in rounds 4-7 but not in round 8. The question about membership in a community group was as follows: *Let's turn to your role in the community. Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an official leader, an active member, an inactive member, or not a member: Member of voluntary association or community group.* This membership question is also recoded to be a dummy variable.

About 74% of respondents (who were not too young to vote) in rural areas said they had voted (see Table 1). There is little variation over time. 76% of the respondents said they had voted in round 4, 72% in round 5, 73% in round 6 and 74.5% in rounds 7 and 8. Other forms of political engagement were less common. Overall, about 67% said they had participated in a community meeting; 60% said they had "joined with others to discuss an issue"; and 44% of respondents said they were members of a community group (this question is only available in rounds 4 to 7, not in round 8). Men were much more politically engaged than women, by these measures, for the sample respondents.

Graphical relationship between education and political engagement behavior Figure 5 shows graphically that education level is negatively correlated with the index of engagement, for both women and men, if one does not control for any confounds. The plot displays the mean index of

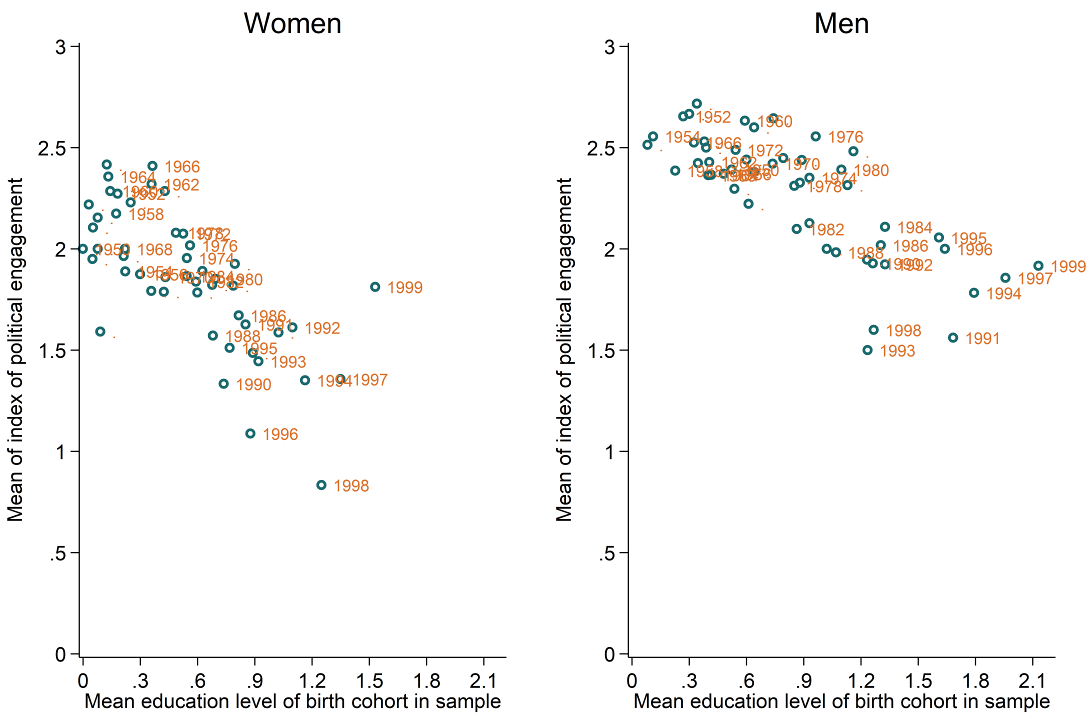

Figure 5: Mean of index of political engagement by gender, according to mean education level, by birth cohort, Burkina Faso

Source: Afrobarometer data, various rounds, rural residents only. Respondents grouped by birth year. There are fewer respondents in the early and later birth years. Select markers are labelled by birth cohort.

engagement according to the mean education level in each birth year cohort. As can be seen, the different birth cohorts have very different mean levels of education, and also different mean levels of the index of engagement. Women have lower levels of education and lower levels of political engagement.

There are several possible confounds in Figure 5 that are important to recall. First, mean education levels for the cohort is closely correlated with the birth year of the cohort. Older generations are on average less educated, as seen earlier in Figure 4. Different birth cohorts have more than just differences in education affecting their propensity to be civically engaged. Older people may have attained greater civic-mindedness by virtue of their life experiences, on average, and so may be more motivated to be politically engaged. Second, the round of the Afrobarometer survey is not controlled for, and even though the overall turnout levels are rather similar over time, the composition of the voters might differ across different elections. Third, education is endogenous and may be correlated with propensity to be politically engaged, and so the estimate of the relationship (here simply a visual relationship) may be biased. It may be that education is associated with more engagement, but education is also positively associated with a third variable associated with less engagement. Or education may be negatively associated with a third variable associated with more engagement. For example, larger villages might have more education and education might induce more people to be engaged, but larger villages have something else about them that induces lower engagement. Education attainment may be correlated with time preference or a variety of cognitive capabilities or dispositions, and these may be correlated with a propensity to be politically engaged. These various confounds may vary by gender.

An instrumental variables approach with geographic and birth cohort dummy variables might be more credible in estimating the causal effect of education on political engagement. The remainder of the paper explores the issue of determining whether the substantial variation in political engagement is plausibly attributed to the effects of attending school (and perhaps having been differently acculturated and having different economic and social opportunities).

5 Estimation strategy

We wish to estimate an average causal effect of educational attainment on the propensity to be politically engaged in Burkina Faso. How much more or less engaged are female and male citizens who have attained higher levels of schooling?

Let Y be the index of political engagement (or alternatively the various components of the index, such as having voted in a recent election, having joined with others to raise an issue, or having joined a community group). Let E be a measure of education attainment. E may be measured in several ways: (1) having been enrolled in primary school; (2) having completed primary school; (3) number of years of education completed, etc.

If education were randomly distributed across individuals, we could estimate the average effect of education on political engagement using the following model:

$$Y_{itv} = \alpha + \beta E_{itv} + \gamma X_{itv} + \varepsilon_{itv} \tag{1}$$

We observe Y, E, and X, for each individual i born in year t living in village v. X includes a set of covariates, such as dummy variables indicating survey rounds. Each survey makes reference to a different election and this may matter in the individual decision to be politically engaged. We create a binary variable taking value one if the respondent is surveyed after the ouster in 2014 of President Blaise Compaoré and zero otherwise. We estimate the regressions separately for men and for women.

Schooling attainment is evidently not randomly distributed across the population, and likely to be associated with unobserved personal, family, village, and generational characteristics that might also determine political engagement. We therefore adopt an instrumental variable (IV) estimation strategy to correct for likely omitted variable bias. Following the lead of other researchers, we use school establishment as an IV for educational attainment (see Card (1995); Card (1999); Currie and Moretti (2003); Duflo (2001) for the pioneering approaches, and André and Maarek (2020) for similar estimation using data from neighbouring Mali).

The main instrumental variable, denoted by Z, is a binary variable that takes value one if the respondent lives in a village where there was a school when she was eligible for school (i.e. seven years old), and zero if the respondent lives in a village where there was no school when she was of school-age. We also use as IV the number of primary schools around the village, each school weighted by the inverse of the square of distance from the village. The Afrobarometer data unfortunately does not include place of birth or childhood, so current residence is used as a proxy for location during schooling years. Data on migration status is not collected in the Afrobarometer. The self-selection concern is reasonable: perhaps individuals or parents who desire high education attainment move to village shortly after a school is established, thus the educated in a village are different for other reasons compared with those in the village just before the school was established. The bias introduced may not be large, however. Burkina Faso historically had very low migration rates compared with other countries. Furthermore, we restrict our sample to the rural respondents, so the estimates may not suffer from the bias that would arise due to rural-urban migration.

The normal instrumental variable model estimated, expressed as a two-stage model, is:

$$\begin{cases} E_{itv} = \phi + \theta Z_{tv} + \pi X_{itv} + \nu_{itv} \\ \\ Y_{itv} = \alpha + \beta \widehat{E}_{itv} + \gamma X_{itv} + \varepsilon_{itv} \end{cases} \tag{2}$$

For the identification approach to be valid, the instrumental variable has to be correlated with the level of education and not correlated with the error term of the second-stage equation ε. The first correlation is discussed when presenting the results from the first-stage estimations in Section 6.1. The second requirement is the exclusion restriction: the instrument should have no effect on the outcome other than through the first-stage channel. The exclusion restriction is conditional on all covariates, and does not assume the unconditional orthogonality of the instrument and the outcome. It is therefore important to control for community-level variables to account for possible endogenous placement of the schools and variation across cohorts as education levels increased over time. All of the regressions will control for location dummies, alternatively defined to be at the commune or province levels, and year of birth dummies, either as annual dummies or five-year birth cohort dummies. Ideally the location dummy variables would be at the village level and the birth cohort would be an the annual level, but as noted earlier there are, for most villages, only 8 observations per village, and so few respondents “straddle” the treatment of being of school age before and after the establishment of the school. We return to this issue below. The actual estimation we perform then is given by equation (3):

$$\begin{cases} E_{itvl} = \phi + \theta Z_{tvl} + \pi X_{itvl} + \delta_t + \theta_l + \nu_{itvl} \\ \\ Y_{itvl} = \alpha + \beta \widehat{E}_{itvl} + \gamma X_{itv} + \delta_t + \theta_l + \varepsilon_{itvl} \end{cases} \tag{3}$$

where we observe Y, E, Z, and X, for each individual i born in year t living in village v located in locality l (commune or province). We therefore exploit two sources of variation: birth cohort and location. We observe within-birth cohort variation as people from the same cohort had different access to primary schools as they lived in different locations. We observe within-location variation as people from a given location are from different birth cohorts, and some were too old to benefit from school construction, while others were of school age or were born after the school was built and thus had access to the school.

Marshall (2016) suggested that the use of a continuous measure of education such as number of years of education or highest education level would be preferred over the use of a binary variable that would capture a particular level of education (e.g. having completed primary school, or secondary school). The policy reform or the construction of new school induces children to go to school and their enrolment in primary school then conditions their likelihood of pursuing each additional year of education. As a result, one could argue that the construction of primary school has an effect not only on the propensity to be enrolled in primary school but also in higher levels of education. Thus the exclusion restriction is not satisfied if we restrict the measure of education to having attended or completed primary school, because the construction of the school affects political engagement by other channels than just completing primary (i.e. by continuing on to secondary).

However a drawback of restricting attention to continuous measures, given the data we have, is that using the education level assumes that any one-unit increase in education induces a rise or decrease in the probability of political participation by β percentage point. For instance everything else being equal, the difference between those who have no education and those who attended without completing primary school is equal to the difference between those who attended without completing secondary school and those who had completed secondary school. This assumption is rather strong, and we thus consider both the continuous education attainment variable and several binary education outcomes.

Our first-stage equation being close to a difference-in-difference equation, this estimation strategy had been labelled 'instrumented difference in differences' (DDIV) by Hudson, Hull, and Liebersohn (2017). DDIV estimates reflect the impact of education attainment on the final outcomes for the individuals who would not have attained much schooling had access remained difficult in their locality. It is an estimate of the average response of those who are treated. In other words it consists in a IV strategy where the first-stage equation relies on a difference-in-difference style estimator of the effect of the treatment on the endogenous variable.

Duflo (2001) is a key example of the DDIV approach. She estimated the income returns to education. She argued that children in localities in Indonesia who were eligible to attend primary school after the INPRES school expansion program were more likely to reach a higher level of educational attainment than the children in older cohorts. Since most primary school programs group children by age in classes, and social stigma for older children from being in classes with much younger children is high, older children were less likely to benefit from newly constructed primary schools located in their district. While school construction was likely not random, Duflo argued that the timing of the school construction program relative to the age of any particular child was likely to be as good as random. Duflo used a treatment variable aggregated at the district level, making the implicit assumption that the construction program was uniformly distributed within each district.

Several researchers interested in the question of the effects of education on political engagement have implemented estimation strategies similar to Duflo's DDIV strategy. Larreguy and Marshall (2017) examined the Universal Primary Education policy reform (and accompanying secondary school construction program) to extend free schooling to all children in Nigeria in 1976. Other researchers have used a simple difference strategy at the national level; i.e. when a country changes education policy in a way that affects all students across the country, the cohort affected has different outcomes from the cohort that was too old to benefit from the policy change. In particular, Croke et al. (2016) used a regression discontinuity equation for the first stage regression. They examined the effects of reforms undertaken in Zimbabwe in 1980 to extend schooling access to black residents (after the end of white rule). André and Maarek (2020) estimated the effects of schooling with a very similar setup as ours, using data from Mali.

6 Results

6.1 Effects of school establishment on schooling outcomes

Table 2 presents descriptive statistics (means and standard deviations) of relevant variables where respondents are grouped according to whether they were of age to be eligible for school (seven years old) before or after a school was established in their village of residence. This variable is the main "treatment" variable in the analysis. Treated respondents are those for whom there was a school, when they were seven years old, in their village of current residence. The table makes clear that treated respondents, both women and men, had substantially larger education outcomes; primary completion rates, for example, were 29% compared with 11% for women, and 40% compared to 14% for men. Engagement outcomes also differed, again both for women and for men. Those of school age after a school was established scored lower on the engagement index:, 1.68 compared with 1.87 for women, and 2.15 compared with 2.39 for men. All of the components of the index were similarly lower, for both men and women (not reported). Those treated were less likely to vote, less likely to participate in a community meeting, and less likely to join with others to discuss a problem. Also, those educated were less likely to be a member of a community group.

Table 2: Descriptive statistics of sample, according to whether school eligible after or before school established, according to gender

	(1) All	(2) School exists when age 7	(3) No school when age 7	(4) p-value
	Panel A: Women			
Some schooling	0.28	0.41	0.18	0.000
	(0.450)	(0.492)	(0.384)	
Completed primary	0.19	0.29	0.11	0.000
	(0.392)	(0.455)	(0.310)	
Some secondary	0.15	0.24	0.08	0.000
	(0.359)	(0.427)	(0.271)	
Education attainment	0.65	0.98	0.38	0.000
	(1.176)	(1.355)	(0.925)	
Index of engagement	1.78	1.68	1.87	0.000
	(1.093)	(1.096)	(1.083)	
Observations	2331	1065	1215	2280
	Panel B: Men			
Some schooling	0.36	0.56	0.24	0.000
	(0.481)	(0.497)	(0.427)	
Completed primary	0.24	0.40	0.14	0.000
	(0.428)	(0.490)	(0.348)	
Some secondary	0.18	0.32	0.10	0.000
	(0.388)	(0.468)	(0.294)	
Education attainment	0.85	1.41	0.50	0.000
	(1.333)	(1.548)	(1.024)	
Index of engagement	2.29	2.15	2.39	0.000
	(0.920)	(0.991)	(0.859)	
Observations	2189	866	1309	2175

Notes: Afrobarometer Burkina Faso data, various rounds, rural residents, birth year after 1950. Standard deviations in parentheses. Final column gives the p-value for t-test of difference in means between the two groups.

The differences in average outcomes do not, however, account for cohort effects. Respondents born after schools were established are of course on average younger, by about 12 years. Neither do the simple differences in means account for geographic differences and survey round effects.

We turn, then, to multivariable regression analysis. We examine first the magnitude of the relationship at the individual level between education attainment and being "treated." We have two different measures for being treated. One is there being a school established in the village of current residence when the respondent was of an age eligible for schooling, which we set at seven years old. The other is the number of schools around the village when the respondent was age seven, each weighted by the inverse of the square of the distance from the school to the village.

Tables 3 and 4 present in a compact display the regression coefficients of interest for a variety of specifications and for four outcome variables, for men only and then for women only. Standard errors in all specifications are clustered at the village level. The four outcome variables are a

Table 3: Effects of school in village or near village, when individual of school age, on education outcomes, various specifications, men only

	(1) Some education	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Dummy variable of school in village			
No cohort or locality controls	0.322***	0.259***	0.229***	0.924***
	(0.024)	(0.023)	(0.022)	(0.076)
Birth year, province	0.249***	0.190***	0.151***	0.695***
	(0.027)	(0.024)	(0.023)	(0.078)
5-year cohort, commune	0.218***	0.169***	0.136***	0.618***
	(0.028)	(0.024)	(0.023)	(0.079)
Birth year, village	0.136***	0.096**	0.057	0.369***
	(0.043)	(0.040)	(0.038)	(0.125)
	Panel B: Number of schools weighted by proximity			
No cohort or locality controls	0.083***	0.076***	0.071***	0.262***
	(0.009)	(0.006)	(0.007)	(0.026)
Birth year, province	0.070***	0.061***	0.052***	0.212***
	(0.008)	(0.007)	(0.006)	(0.025)
5-year cohort, commune	0.071***	0.061***	0.055***	0.214***
	(0.010)	(0.010)	(0.009)	(0.034)
Birth year, village	0.032***	0.021**	0.027***	0.075**
	(0.012)	(0.010)	(0.010)	(0.030)

Notes: Each coefficient and standard error in the table is for a separate regression, according to the specification (row) and outcome measure (column) and measure of schooling access (panel). For Panel A the coefficient is that of the dummy variable indicating if a school was present, for Panel B the coefficient is on the variable representing the count of schools, each weighted by the inverse distance squared, where distance is the measure in kilometers of the distance from the village of residence to the village of the school. Standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `reghdfe`. * p <0.10,** p <0.05,*** p <0.01. Data is from Afrobarometer survey, Burkina Faso, rural residents born after 1950, various rounds, men only, and administrative data on year of school establishment.

binary variable for whether the person had some formal education, a binary for having completed primary school, and a binary variable for having attended at least some years of secondary school, and lastly the scale of schooling attainment, ranging from 0 to 5 as displayed in Figure 3. All specifications include a dummy variable for whether the respondent is surveyed during one of the post-Compaoré rounds.

We refer to the specifications as: (1) No cohort or locality dummy variables; (2) Birth year cohort and province dummy variables; (3) Five-year birth cohort and commune dummies; (4) Birth year cohort and village dummy variables. The specifications differ in terms of the geographic and birth year dummy variables that control for unobserved correlates. Dummy variables for geographic areas may be at the commune level (for the 209 communes in the data, the median number of respondents is 16, while the minimum is 5 respondents) or at the province level (45 clusters with median size of 99 respondents). We can also use the village level as the geographic

Table 4: Effects of school in village or near village, when individual of school age, on education outcomes, various specifications, women only

	(1) Some education	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Dummy variable of school in village			
No cohort or locality controls	0.238***	0.187***	0.162***	0.618***
	(0.022)	(0.020)	(0.019)	(0.062)
Birth year, province	0.175***	0.131***	0.103***	0.429***
	(0.023)	(0.020)	(0.019)	(0.062)
5-year cohort, commune	0.126***	0.081***	0.055***	0.273***
	(0.027)	(0.023)	(0.020)	(0.068)
Birth year, village	0.003	-0.025	-0.037	-0.060
	(0.038)	(0.033)	(0.031)	(0.097)
	Panel B: Number of schools weighted by proximity			
No cohort or locality controls	0.064***	0.054***	0.048***	0.177***
	(0.009)	(0.009)	(0.008)	(0.028)
Birth year, province	0.052***	0.045***	0.040***	0.146***
	(0.009)	(0.007)	(0.007)	(0.025)
5-year cohort, commune	0.048***	0.043***	0.037***	0.140***
	(0.010)	(0.008)	(0.008)	(0.027)
Birth year, village	0.016*	0.014	0.024*	0.060*
	(0.009)	(0.011)	(0.013)	(0.033)

Notes: Each coefficient and standard error in the table is for a separate regression, according to the specification (row) and outcome measure (column) and measure of schooling access (panel). For Panel A the coefficient is that of the dummy variable indicating if a school was present, for Panel B the coefficient is on the variable representing the count of schools, each weighted by the inverse distance squared, where distance is the measure in kilometers of the distance from the village of residence to the village of the school. Standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `reghdfe`. * p <0.10,** p <0.05,*** p <0.01. Data is from Afrobarometer survey, Burkina Faso, rural residents born after 1950, various rounds, women only, and administrative data on year of school establishment.

locality even though there are a small number of respondents in each cluster (526 rural village clusters, with median number of respondents of 8, maximum of 32, and minimum of 5). Separating by gender means there are typically only 4 respondents in a gender-village cluster, and adding year of birth cohort dummy variables further reduces the possibility of variation, to the point that there are typically zero or one observation per sub-group.

The various approaches find sizable and statistically significant effects of school establishment on schooling outcomes. For the first specification, with neither geographic nor birth cohort dummy variables, the effects are sizable. Including geographic and birth cohort dummy variables reduces the magnitudes of the estimated coefficients, suggesting that places and birth cohorts that had relatively more respondents attain higher levels of schooling after a school was established were also more likely to be places where respondents of school age before a school was established attained high schooling (perhaps by going to school in neighboring villages, or because these villages attracted educated persons to settle in the villages). Yet even when controlling for year of birth and village, the effects are statistically significant and sizable, for men. For example, in Table 3 it can be seen that having a school in the village when a young boy is seven years old is associated with an increase in the probability of having some education of 0.14 (that is, by 14 percentage points), increasing the probability of completing primary school by 0.096 (almost 10 percentage points), and increasing the measure of attainment by 0.37. This is with the most "demanding" specification, with village dummy variables and birth year cohort dummy variables. The effect on completing secondary is not significant in that specification, but is substantial in the other specifications.

In Panel B the estimated magnitudes of the effects are similar. A one unit increase in the inverse-distance weighted number of schools in proximity is associated with a 3-8 percentage point increase in the probability of attaining some education (depending on the specification), and a similar sized increase in the probability of completing primary school.

The effects are smaller for women, as can be seen in Table 4. For the 5-year cohort and commune locality specification, having a school in the village when a young girl is seven years old is associated with an increase in the probability of having some education of 0.126 (that is, by 12 percentage points), increasing the probability of completing primary school by 0.081 (8 percentage points), and increasing the measure of attainment by 0.27. The effect on completing secondary is 0.055 (about half the size of the effect for men). In Panel B for women the estimated magnitudes of the effects are also smaller than for men. A one unit increase in the inverse-distance weighted number of schools in proximity is associated with about a 5 percentage point increase in the probability of attaining some education, and a similar increase in the probability of completing primary school, in the five-year birth cohort and commune dummies specification. The birth year and village dummies specification does not yield statistically significant coefficients for the women sample.

The effects of school access being smaller for women compared to those for men is unsurprising. Even if access to school reduces the cost of schooling, there is still a school enrolment cost for the families, that might favor their boys in a context of resource constraints and in the absence of compulsory school enrolment. In Nigeria, Larreguy and Marshall (2017) find that the first-stage effect of the universal primary education reform on education level is 40% lower for women. In previous analyses on Zimbabwe (Djemai, Samson, and Renard (2022)), the differential effect of the 1980 reform by gender has been found to be the opposite: greater for women than for men mostly because primary school became compulsory and the level of education was greater for men than for women before the reform.

There are many other variable coding and specification variations possible for these first stage regressions. Some specifications could exclude the partially treated, defined as those who were 8, 9 or 10 years old when the school was established in the village. This group could have benefited from the newly established school if they started school later in age or if they switched from being enrolled in a school that was further away from their village to the newly established school. For the latter, the cost of being enrolled is reduced and continued enrollment in primary school is likely to be exogenously increased as a result. Alternatively, one could assign partial schooling access to those older than 7 when schools were established in the village or in the vicinity and include that variable separately. First-stage regression and second-stage regressions results are highly stable whether the partially treated are included or excluded from the sample (not reported here).

Being of school age after a school was established could be interacted with a dummy variable for whether the person was of school age after 1980 (when government promotion of schooling accelerated). When the locality cluster is the commune or province, time-varying or fixed covariates at lower geographic levels could be included to control for some potential confounds (e.g., distance to commune village seat, usually the commercially important village of the area with more public infrastructure; village population over time). Characteristics of schools (public, private, how many classrooms) might also be controlled for. In general, the "first stage" results yield similar estimates of the effects of schooling access on education outcomes in other specifications and variable coding choices (results not reported).

This first-stage is crucial for the credibility of second-stage estimates of the effect of schooling levels of individuals on their political engagement. We report in Appendix B, Tables B1 and B2, the F-statistics for excluded instruments for the various specifications. We see that the F-statistics are generally quite large, and many above the various rules of thumb proposed in the past, except for the case of the birth year and village dummies specification, where the F-statistic is quite small. This last specification clearly generates a "weak instruments" problem, and so we do not use it in estimating the second stage.

In Appendix B we also report event study graphs that provide one visual check for the parallel trends assumption that is used to gauge the validity of difference-in-differences estimation. The event study graphs plot the coefficients with specifications that include ten leads and ten lags, with the furthest lead and lag accumulating the school establishment dummy prior to or after the lag or lead. As can be seen, coefficients on lagged indicators of school establishment are not significant. Those respondents in villages who were school age eligible well before a school was established were not seeing rising education levels, relative to those of similar age in villages where a school would be established even later, or not at all. Coefficients on schooling outcomes for those who were around seven years old or younger at the time of schooling establishment are positive and with confidence intervals above zero. The effect seems to diminish, though, as those who are school-age-eligible more than five years prior to when a school was established do not see as large or as statistically significant effects.

For further discussion of the first stage, we include in Appendix C the results of using the Probit specification for the three first-stage regressions that have binary outcome variables. The marginal effects are basically all statistically significant and the same magnitudes as in the linear probability model. We also discuss the two-way fixed effects (TWFE) specification and various recent modifications to that approach, in the context of the structure of the Afrobarometer sample. Similarly, we explain why a regression discontinuity approach does not work well with the Afrobarometer data in our case.

We turn, then, to estimates of the effects of education attainment on political engagement.

6.2 Effects of schooling outcomes on political engagement

The expansion of primary school infrastructure appears to have significantly increased the education attainment and primary completion rates of residents of rural areas. In this section, we explore the effects of education on the index of political engagement. We present the results using the specifications with various combinations of birth cohort and location dummies. We do not consider in the rest of the paper the specifications that control for birth year and village dummies as the instrumental variables (school in village or number of schools near village) exhibit low explanatory power in the first stage.

Before turning to the IV estimates, we note that in Appendix D, Table D1 we present the results from the naive estimation of the correlations between the index of political engagement and the education measures (the coefficients in OLS equation with only the control for survey round). The table also includes the coefficient from the simple OLS for each component of the index, and for the indicator of being a member of a community group (rounds 4-7 only). In all cases, the relationship with education is negative, and is mostly significant for women. Table D2 presents the reduced-form estimates with the binary treatment variable for having a school in the village when school aged in Panel A, and the weighted number of schools in Panel B. Column 1 reports the estimated coefficients for the whole sample, and Columns 2 and 3 for men and women respectively. Access to primary schools is negatively related with the index of political engagement, and significantly so, in the specification with no cohort or locality controls. When controlling for cohort and locality dummies, the negative effect persists mostly for men.

Table 5 presents the IV estimates with the first stage the various specifications of Table 3 with outcome variable the index of political engagement, for men only. Panel A shows results for the instrumental variable of whether there was a school established in the village of residence prior to the respondent being school age eligible. Panel B shows results for the instrumental variable indicating the number of schools in proximity to the village of residence, when the respondent was school age eligible. Each panel reports the coefficients of interest for Equation (3), estimated with a different education measure (as indicated in the columns). The specifications vary in the controls for birth cohort and locality. A dummy variable for the later post-2014 survey rounds is also included in all specifications.

In general, the results are consistent across the different measures of education and specifications. For men, the effect of education is negative and often statistically significant even with the birth and locality controls. That is, the results suggest that education, instrumented by access to

Table 5: Effects of education on index of political engagement, with various measures of education outcomes instrumented with indicators of access to schooling, for men only

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: IV dummy variable of school in village			
No cohort or locality controls	-0.760***	-0.952***	-1.089***	-0.263***
	(0.159)	(0.203)	(0.243)	(0.056)
Birth year, province	-0.281	-0.373	-0.462	-0.102
	(0.199)	(0.258)	(0.334)	(0.071)
5-year cohort, commune	-0.433*	-0.561*	-0.697	-0.152*
	(0.257)	(0.333)	(0.430)	(0.090)
	Panel B: IV number of schools weighted by proximity			
No cohort or locality controls	-0.926***	-1.042***	-1.132***	-0.296***
	(0.200)	(0.258)	(0.286)	(0.074)
Birth year, province	-0.461**	-0.539*	-0.650*	-0.154*
	(0.231)	(0.277)	(0.343)	(0.079)
5-year cohort, commune	-0.473*	-0.573	-0.649	-0.159
	(0.287)	(0.361)	(0.426)	(0.100)

Notes: Dependent variable is index of political engagement, the sum of three dummy variables (whether voted in the prior election, whether participated in a meeting, and whether joined with others to raise an issue). Data on the outcome is from the Afrobarometer Burkina Faso data, rural residents born after 1950, various rounds, men only. Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include a dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

school, was associated with lower levels of political engagement. The magnitudes are fairly large: in the specifications without birth and locality controls, the effects of having some education, completing primary, or completing secondary are all associated with reductions in the index of engagement of about one unit; while once birth cohort and geographic locality are controlled for the effects of the education dummy variables falls to about .5 units. The engagement index is on a maximum three point scale, so an effect of .5 units is substantial. For specifications that use the weighted number of schools in the proximity as instrumental variable, the education effects remain negative, substantial, and often statistically significant even when using the birth cohort and geographic control variables.

For women, as reported in Table 6, the effects are statistically significant when not using any geographic controls and birth cohort controls. In that specification, the effects are negative and basically the same magnitude as for men. For the specifications with birth cohort and locality controls, results are statistically significant in Panel B with birth year and province, and roughly similar magnitudes as for men.

Table 6: Effects of education on index of political engagement, with various measures of education outcomes instrumented with indicators of access to schooling, women only

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: IV dummy variable of school in village			
No cohort or locality controls	-0.741***	-0.969***	-1.124***	-0.291***
	(0.212)	(0.281)	(0.323)	(0.083)
Birth year, province	-0.075	-0.101	-0.125	-0.031
	(0.291)	(0.389)	(0.479)	(0.118)
5-year cohort, commune	0.116	0.180	0.251	0.053
	(0.413)	(0.643)	(0.899)	(0.190)
	Panel B: IV number of schools weighted by proximity			
No cohort or locality controls	-1.000***	-1.213***	-1.365***	-0.364***
	(0.208)	(0.282)	(0.342)	(0.084)
Birth year, province	-0.545*	-0.645*	-0.735*	-0.196*
	(0.315)	(0.385)	(0.436)	(0.116)
5-year cohort, commune	-0.066	-0.076	-0.094	-0.023
	(0.472)	(0.545)	(0.665)	(0.166)

Notes: Dependent variable is index of political engagement, the sum of three dummy variables (whether voted in the prior election, whether participated in a meeting, and whether joined with others to raise an issue). Data on the outcome is from the Afrobarometer Burkina Faso data, rural residents born after 1950, various rounds, women only. Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include a dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

6.3 Effects of school outcomes on the components of index of engagement

Now we estimate the causal effect of schooling on each of the underlying four indicators of political engagement. We present results of whether the education outcomes partly induced by access to primary schools had effects on the four components in Table 7 for the men sample and in Table 8 for the women sample. Both tables have as outcome variables whether the person voted in the previous election (Panel A), met with other persons to raise an issue (Panel B), was a member of a community group (Panel C, for rounds 4-7 only) and participated in a community meeting (Panel D). The original survey questions on membership, participation, and joining with others had four responses, here they have been recoded to be binary variables. Each panel reports the coefficients of interest for Equation (3), estimated with a different education measure (as indicated in the columns). The specifications vary in the controls for birth cohort and locality. Dummy variable for the survey rounds are also included in all specifications as in the previous tables. Both tables report the estimated IV coefficients when the instrumental variable is whether there was a school established in the village of residence prior to the respondent being school age eligible. Results are reasonably similar with the use of the weighted number of schools as the instrumental variable and are reported in Appendix Tables E1 and E2.

Table 7: Effects of education on the four measures of political engagement, with various measures of education, and with presence of school in the village when school eligible as instrumental variable, men sample

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Voted in the previous election			
No cohort or locality controls	-0.218***	-0.276***	-0.315***	-0.076***
	(0.063)	(0.079)	(0.094)	(0.022)
Birth year, province	0.124	0.159	0.206	0.043
	(0.082)	(0.109)	(0.138)	(0.029)
5-year cohort, commune	0.078	0.099	0.126	0.027
	(0.109)	(0.142)	(0.178)	(0.038)
	Panel B: Joined with others to raise an issue			
No cohort or locality controls	-0.234***	-0.289***	-0.328***	-0.081***
	(0.070)	(0.089)	(0.102)	(0.025)
Birth year, province	-0.159	-0.206	-0.259	-0.056
	(0.100)	(0.131)	(0.168)	(0.036)
5-year cohort, commune	-0.117	-0.149	-0.186	-0.041
	(0.120)	(0.155)	(0.195)	(0.042)
	Panel C: Member of a community group			
No cohort or locality controls	-0.314***	-0.392***	-0.450***	-0.113***
	(0.081)	(0.105)	(0.121)	(0.030)
Birth year, province	-0.277**	-0.380**	-0.497**	-0.105**
	(0.121)	(0.170)	(0.229)	(0.047)
5-year cohort, commune	-0.295*	-0.404*	-0.546*	-0.112*
	(0.155)	(0.223)	(0.312)	(0.060)
	Panel D: Participated in a meeting			
No cohort or locality controls	-0.317***	-0.396***	-0.446***	-0.111***
	(0.065)	(0.082)	(0.095)	(0.023)
Birth year, province	-0.246***	-0.327***	-0.406**	-0.089***
	(0.091)	(0.121)	(0.159)	(0.033)
5-year cohort, commune	-0.362***	-0.468***	-0.580***	-0.128***
	(0.123)	(0.160)	(0.212)	(0.044)

Notes: Afrobarometer Burkina Faso data, rural residents born after 1950, five rounds (except in Panel C, rounds 4-7). Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * $p < 0.10$,** $p < 0.05$,*** $p < 0.01$.

In general, the results are consistent across the four outcome variables and different measures of education. For all four outcomes, the specification without birth cohort and locality dummy variables yields negative and statistically significant coefficients of the education variable. That is, the estimations suggest that education, instrumented by access to school, is associated with lower levels of political engagement. These results hold for both men and women. For instance, having been enrolled in school decreases the probability of voting by 22 percentage points for men (column 1, Panel A of Table 7) and, very similarly, by 16 percentage points for women (column 1, Panel A of Table 8).

Once cohort and locality controls are introduced, however, the estimated coefficients are never statistically significant for women. For men, some of the individual components have statistically significant negative coefficients. For two of the outcomes (membership in a community group and participation in community meetings), the coefficients remain negative and significant when using province-level geographic controls and year of birth cohort controls, and when using commune dummies and 5-year birth cohorts. But for the other two specifications and outcomes (voting behavior and joining with others to raise an issue), the instrumented effect of education is not significant. These results are roughly similar across all four measures of education.

Table 8: Effects of education on the four measures of political engagement, with various measures of education, and with presence of school in the village when school eligible as instrumental variable, women sample

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Voted in the previous election			
No cohort or locality controls	-0.161*	-0.210*	-0.243*	-0.063*
	(0.087)	(0.114)	(0.132)	(0.034)
Birth year, province	0.134	0.178	0.219	0.054
	(0.117)	(0.156)	(0.192)	(0.047)
5-year cohort, commune	0.083	0.128	0.177	0.038
	(0.169)	(0.262)	(0.364)	(0.077)
	Panel B: Joined with others to raise an issue			
No cohort or locality controls	-0.269***	-0.344***	-0.399***	-0.104***
	(0.090)	(0.116)	(0.135)	(0.035)
Birth year, province	-0.048	-0.065	-0.083	-0.020
	(0.137)	(0.186)	(0.237)	(0.057)
5-year cohort, commune	0.055	0.088	0.130	0.026
	(0.207)	(0.333)	(0.494)	(0.098)
	Panel C: Member of a community group			
No cohort or locality controls	-0.289**	-0.371**	-0.416**	-0.112**
	(0.114)	(0.146)	(0.163)	(0.044)
Birth year, province	-0.076	-0.101	-0.123	-0.031
	(0.173)	(0.230)	(0.276)	(0.070)
5-year cohort, commune	-0.119	-0.160	-0.219	-0.050
	(0.273)	(0.366)	(0.498)	(0.113)
	Panel D: Participated in a meeting			
No cohort or locality controls	-0.354***	-0.453***	-0.525***	-0.137***
	(0.095)	(0.122)	(0.141)	(0.037)
Birth year, province	-0.152	-0.206	-0.262	-0.062
	(0.137)	(0.186)	(0.234)	(0.056)
5-year cohort, commune	0.007	0.011	0.016	0.003
	(0.208)	(0.330)	(0.485)	(0.097)

Notes: Afrobarometer Burkina Faso data, rural residents born after 1950, five rounds (except in Panel C, rounds 4-7). Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

7 Geographic and birth cohort dummies, and power considerations

Overall, we find that the causal effect of education on political engagement is likely negative for men, and zero for women. For men, the effect is negative for the index of engagement and for the four separate measures of political engagement. For women the results appear to be zero effects on all of the measures of engagement, once controls are included. The results suggest that for Burkina Faso there is little reason to think there is a sizable positive relationship between education levels of a person and their likelihood of being politically engaged (as measured in the Afrobarometer surveys). The "null" results for women could be due to misspecification and to lack of statistical power to detect a sizable effect, so we turn to those issues.

As noted earlier, estimation results with a specification including birth year cohorts and geographic dummy variables at the village level are not reported here, because the great majority of villages only have 8 observations, with dummy variable for birth cohort, and split by gender, most observation cells are empty. The estimated coefficients in such specifications are close to zero and never statistically significant. This raises the general issue of statistical power. If there were 80 observations in each village, presumably the only specification that mattered would be the one with village level dummy variables.

Other empirical analyses using the Afrobarometer data recognize this small sample problem and the trade-off implied. Briggs (2019) estimated that the presence of a foreign aid project within 50 km of a locality reduced support for incumbent presidents, using Afrobarometer data for Nigeria, Senegal, and Uganda (the three countries for which aid data was available). He reported two specifications, one using country dummy variables and the other using region dummy variables for the highest administrative level of the country. His discussion of the choice of the level of dummy variables is worth quoting in full: "Comparing within region-rounds increases the likelihood that I am comparing across groups of people that are similar except for their aid status, but many region-rounds lack observations in at least one of the three aid categories ... so relatively few observations contribute identifying variation when using region-round fixed effects. This is much less of an issue when using country-round fixed effects, but in this case I am less likely to be comparing groups of people that are similar aside from their aid status because I allow comparisons across regions within the same country and survey round. Neither approach is obviously better and so I report results from both approaches." T. Knutsen and Kotsadam (2020), Cha (2020), and Watkins (2021) estimate quite similar relationships between various measures of aid to a locality and various political outcomes.

C. H. Knutsen et al. (2017), to take another example, estimated that industrial mines locating in proximity to villages and towns likely increased local corruption. They used a variety of corruption related questions in the Afrobarometer surveys for 33 countries. They coded a dummy variable for whether an industrial mine was located within 50 km of the respondent's residence. For rural residents, then, the empirical strategy was quite similar to ours. Their locality dummy variables were set at the country level in one specification, and at the 50km radius around a mine in another specification. They noted that the latter "conservative" specification, "... has the advantage of only comparing the exact same areas before and after mine opening. The disadvantage is that we only have Afrobarometer observations before and after mine openings for a few locations, making this a demanding test." Indeed, while with country level dummy variables all four estimated effects of an active mine on corruption were statistically significant, with the mine locality dummy variable only one of the coefficients remained statistically significant.

Similarly, Konte and Vincent (2021) used Afrobarometer data to estimate that industrial mining worsened perceptions of the quality of public services and worsened respondents' optimism about future living conditions. Their empirical specification used only country level dummy variables.

Choi, Laughlin, and Schultz (2021) used Afrobarometer data to estimate that individuals who gained access to mobile internet coverage were less likely on average to identify with a national identity instead of an ethnic identity. As dummy variables they used country and administration level (the highest level for each country), and also 0.5 $\times$ 0.5 degree grid cells of approximately 55 $\times$ 55 square km, in the various specifications.

In our linear models, the causal effect of education on the index of political engagement and the causal effect of education on the underlying four indicators of political engagement is not significant for women when we control for birth cohorts and locality of residence, and for men in certain specifications. The estimated null effects may be due to a true zero effect or to a lack of power to detect a reasonable sized effect. The standard definition of power is one minus the probability of Type II error, where Type II error is when the "true" β in Equation (3) is different from zero but we fail to reject the null hypothesis (Campbell and Gustafson (2018); Walker et al. (2017); Parinduri (2019)). We may evaluate the power of the estimation approach using the R package `ivmodel`, developed by Kang et al. (2021). The package implements the power formula developed by Freeman, Cowling, and Schooling (2013). The formula appears to be appropriate to our setting; Kang et al. (2021) use as their example data the analysis of Card (1995), who pioneered the estimation approach used here. Card (1995) estimated the income returns to schooling with an instrumental variable measuring distance to a college; the sample size he used, about 3,000 observations, was similar to that used here.

Figures 6 and 7 show that for the three education indicators considered (some schooling, completed primary, and education level), and with whether there was a school in the village at the time the person was eligible for school as the instrumental variable, the estimation appears to be adequately powered for relevant effect sizes for the sub-sample of rural men surveyed in the Afrobarometer, but not for the sub-sample of women. As we have seen in Table 5, the estimated coefficient for men is about -0.5 for the two binary indicators and about -0.3 for the index of schooling level. The different effect sizes (here not standardized, but rather the magnitude of the coefficient) are on the x-axes, and the power is indicated on the y-axis. The power is calculated for the men-only sample, with about 2,100 observations.

In Figure 6 when education is measured by having some schooling, we exceed 80% power if the size of the effect on the political engagement index is in the -0.7 to -0.4 range, in specifications with no geographic or birth year controls. We have about 80% power for -0.7 to -0.5 range

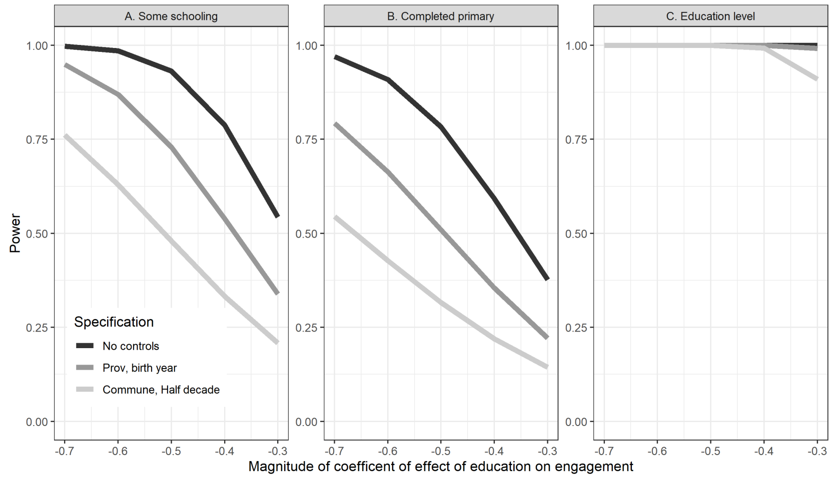

Figure 6: Power calculations for different education variables (endogenous) using schooling access as instrument and outcome variable the index of political engagement, men sample

Notes: Power calculated under various effect sizes (magnitudes of coefficients) using R package `ivmodel` developed by Kang et al. (2021), using Afrobarometer data, various rounds, rural residents born after 1950, men only (sample size about 2,100).

range with province dummies and birth year cohorts, and with commune dummies and five-year birth cohorts we are below 80% power for effect sizes below -0.7. When the education measure is whether the person completed primary, power declines, since there is less variation in the sample in this indicator. But power remains above 80% for relevant coefficient magnitudes in specifications with no geographic or birth year controls. Specifications with commune dummies and five-year birth cohorts are under-powered for coefficient effects below -0.7. When the interval variable for education attainment is used as the education indicator, the effect is for a one unit increase in the indicator, and the estimated coefficient range as seen in Table 5 was between -0.5 and -0.1. Power appears to be adequate for coefficient magnitudes of -0.3 or greater, since there is considerably more variation in this indicator of education. Overall, these specifications appear to have reasonable power to detect economically and politically relevant effects.

The picture is quite different when we calculate power using the women sample. Despite the sample being slightly larger (2,168 observations), as can be seen in Figure 7, power is lower for the binary education variables. Power is only adequate for the very large effect sizes for some of the specifications (i.e., the specification with no controls). When the interval variable for education attainment is used as the education indicator, power appears to be adequate for coefficient magnitudes of -0.3 or greater, in the specification with province and birth year cohort dummy variables. But for the others, there simply is not enough variation in education outcomes, within birth and geographic units, to estimate the coefficient of interest with much precision.

Appendix F contains similar power analysis for the regressions with the separate components

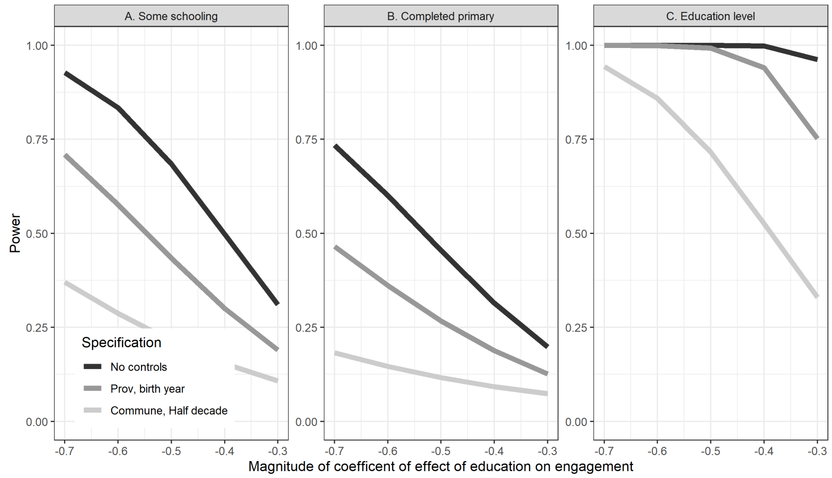

Figure 7: Power calculations for different education variables (endogenous) using schooling access as instrument and outcome variable the index of political engagement, women sample

Notes: Power calculated under various effect sizes (magnitudes of coefficients) using R package `ivmodel` developed by Kang et al. (2021), using Afrobarometer data, various rounds, rural residents born after 1950, women only (sample size 2,168).

of the engagement index, for men (the results for women are quite similar). The analysis suggests that specifications with the education level variable have considerably more power than those with the dummy variables for schooling attainment, that specifications with no geographic or birth cohort controls are adequately powered, while specifications with geographic and birth cohort controls are only powered for larger effects.

8 Attitudes towards democracy

We now digress and consider attitudes and preferences regarding democracy as the preferred system of government, and views about the functioning of democracy in Burkina Faso. These attitudes now are the outcome variables, instead of the responses about political engagement. The analysis may offer insight into why more educated rural men are less likely to be politically engaged. In large-sample analyses of World Values Survey and Afrobarometer, more education is typically found to be associated with more positive attitudes towards democracy as the "better" system of government, though the correlations are not credibly identified as causal (Fuchs-Schündeln and Schündeln (2015); Aquino (2015)). We use the same specification as in the previous section, with the presence of a school in the village of residence when the person was school eligible as an instrumental variable for the respondent's level of education.

Respondents to the Afrobarometer in Burkina Faso were asked whether they were satisfied with the democracy in the country at the time of the survey (from 1: not at all, to 4: very satisfied). Over the five rounds of survey, 11.3% respondents were not satisfied at all, 30.6% not satisfied, 39% rather satisfied, and 19.1% very satisfied. Panel A in Tables 9 and 10 uses as the dependent variable a binary variable equal to one if the respondent declared being satisfied or very satisfied, and Panel B uses the categorical variable. In both cases, satisfaction is significantly lower for educated men compared to less educated men, for all three specifications. For women, the education does not significantly influence the level of satisfaction toward the level of democracy.

In addition, respondents were asked about their views about the level of democracy in Burkina Faso at the time of the survey over a four-point scale: 1 not a democracy, 2 a democracy with minor concerns, 3 a democracy with major concerns, 4 a full democracy. 8% of the respondents considered that Burkina Faso was not a democracy, at the other extreme, 28.2% viewed the country as a full democracy. Panel C uses the 4-point scale perception measure as dependent variable. Once again, the effect of education for men is negative and significant across all specifications and all measures of education. For women, the effect is negative as well. It is significant in the first two specifications and for every measure of education.

Lastly, people were also asked about their preferences for democratic or non democratic regime. Over the five rounds, 9.8% reported that they preferred non-democratic regimes, 73.7% democratic regimes, and 16.5% declared that the type of regime did not matter for people like them. Panel D estimates the IV coefficient of education on the probability of declaring a preference for a democratic regime. There is no significant effect, not surprisingly given the limited variation in the survey responses.

Overall, the main results are that more educated men appeared to be less satisfied with the level of democracy in Burkina Faso than less educated men. Moreover, educated men were less likely to see Burkina Faso as a full democracy, but there was no significant effect of education on their preference for a democratic regime. For women, on the other hand, education was not a significant and consistent predictor of attitudes towards democracy.

Table 9: Effects of education on the attitudes, with various measures of education, and with presence of school in the village when school eligible as instrumental variable, men sample

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Satisfied with the level of democracy			
No cohort or locality controls	-0.216***	-0.265***	-0.301***	-0.075***
	(0.066)	(0.081)	(0.093)	(0.023)
Birth year, province	-0.309***	-0.405***	-0.507***	-0.110***
	(0.102)	(0.134)	(0.173)	(0.036)
5-year cohort, commune	-0.275**	-0.364**	-0.457**	-0.098**
	(0.127)	(0.171)	(0.218)	(0.045)
	Panel B: Satisfaction [1;4]			
No cohort or locality controls	-0.489***	-0.602***	-0.682***	-0.169***
	(0.124)	(0.152)	(0.175)	(0.043)
Birth year, province	-0.656***	-0.860***	-1.077***	-0.234***
	(0.195)	(0.259)	(0.333)	(0.069)
5-year cohort, commune	-0.605**	-0.802**	-1.005**	-0.216**
	(0.239)	(0.327)	(0.416)	(0.086)
	Panel C: Views about democracy in Burkina Faso			
No cohort or locality controls	-0.463***	-0.560***	-0.630***	-0.158***
	(0.123)	(0.148)	(0.168)	(0.041)
Birth year, province	-0.690***	-0.879***	-1.080***	-0.239***
	(0.190)	(0.249)	(0.314)	(0.066)
5-year cohort, commune	-0.789***	-1.002***	-1.225***	-0.271***
	(0.255)	(0.338)	(0.427)	(0.088)
	Panel D: Preferences for democracy			
No cohort or locality controls	-0.088	-0.110	-0.124	-0.031
	(0.062)	(0.076)	(0.086)	(0.021)
Birth year, province	-0.015	-0.019	-0.025	-0.005
	(0.087)	(0.115)	(0.144)	(0.031)
5-year cohort, commune	0.033	0.044	0.053	0.012
	(0.111)	(0.145)	(0.181)	(0.039)

Notes: Afrobarometer Burkina Faso data, rural residents born after 1950, various rounds. Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

Table 10: Effects of education on the attitudes, with various measures of education, and with presence of school in the village when school eligible as instrumental variable, women sample

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Satisfied with the level of democracy			
No cohort or locality controls	-0.051	-0.064	-0.075	-0.020
	(0.093)	(0.118)	(0.137)	(0.036)
Birth year, province	-0.021	-0.027	-0.035	-0.008
	(0.136)	(0.181)	(0.234)	(0.055)
5-year cohort, commune	0.152	0.233	0.357	0.070
	(0.235)	(0.366)	(0.581)	(0.110)
	Panel B: Satisfaction [1;4]			
No cohort or locality controls	-0.194	-0.244	-0.283	-0.074
	(0.175)	(0.221)	(0.256)	(0.067)
Birth year, province	-0.286	-0.380	-0.492	-0.117
	(0.265)	(0.354)	(0.460)	(0.108)
5-year cohort, commune	-0.057	-0.087	-0.133	-0.026
	(0.423)	(0.646)	(0.989)	(0.195)
	Panel C: Views about democracy in Burkina Faso			
No cohort or locality controls	-0.341**	-0.426**	-0.492**	-0.129**
	(0.158)	(0.197)	(0.227)	(0.059)
Birth year, province	-0.551**	-0.709**	-0.901**	-0.217**
	(0.244)	(0.320)	(0.408)	(0.097)
5-year cohort, commune	-0.547	-0.778	-1.093	-0.235
	(0.392)	(0.558)	(0.793)	(0.167)
	Panel D: Preferences for democracy			
No cohort or locality controls	-0.026	-0.032	-0.037	-0.010
	(0.089)	(0.111)	(0.128)	(0.034)
Birth year, province	-0.025	-0.033	-0.041	-0.010
	(0.131)	(0.174)	(0.216)	(0.052)
5-year cohort, commune	0.221	0.333	0.481	0.099
	(0.212)	(0.327)	(0.484)	(0.096)

Notes: Afrobarometer Burkina Faso data, rural residents born after 1950, various rounds. Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

9 Conclusion

Patterns of political engagement by citizens likely influence the quality of economic policymaking, especially in the provision of public goods and the choices made regarding policies that have broad rather than concentrated benefits (Khemani et al. (2016)). The relationship between education and political engagement has been a perennial concern in political economy, as democracies around the world ebb and flow. Recent work has suggested the relationship may be quite variable, both in terms of different social groups exhibiting different tendencies and in terms of tendencies changing significantly as political contexts change. Some of this research has used individual survey data and quasi-experimental estimation approaches, as we do (André and Maarek (2020); Croke et al. (2016); Larreguy and Marshall (2017); Dang (2019); Blaydes (2006); Kuenzi (2006); Parinduri (2019); Mattes and Mughogho (2009)). A few papers have leveraged randomized control trials to examine the relationship (Friedman et al. (2016)). Some papers approach the relationship at a more aggregated level (Acemoglu et al. (2005); Paglayan (2021)).

This paper has matched information about the timing of establishment of primary schools at the village level to nationally-representative household survey data from Burkina Faso to analyse the relationship between education and political engagement. The timing of school establishment relative to the age of each respondent provides an instrumental variable for the individual's education attainment. Our results suggest that formal schooling in the contexts of low overall education levels, as is typical of many African countries, probably has more of a negative effect on political engagement than one might have thought. This negative finding appears to be clear for men, but the low level of variation in schooling outcomes for women means our data and estimation approach are under-powered for women. The finding of a negative relationship may be specific to the data used here, and estimates of the relationship with data from other countries and other electoral contexts may vary. Given the continued expansion of formal education, and the pronounced trend away from democratic governance in many West African countries with recent military-led coups d'etat, the finding of a negative relationship between education and political engagement should be cause for concern for citizens of Burkina Faso and other countries of the region.

References

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2005). "From education to democracy?" In: *American Economic Review* 95(2), pp. 44–49.

Afrobarometer (2021). *Burkina Faso, Rounds 4, 5, 6, and 7*. Available at www.afrobarometer.org.

Amoateng, Acheampong Yaw, Ishmael Kalule-Sabiti, and Tim B. Heaton (2014). "Gender and changing patterns of political participation in sub-Saharan Africa: evidence from five waves of the Afrobarometer surveys". In: *Gender and Behaviour* 12(3), pp. 5897–5910.

André, Pierre and Paul Maarek (2020). *Education, social capital and political participation: Evidence from school construction in Malian villages*. THEMA (Théorie Economique, Modélisation et Applications), Université de Cergy-Pontoise.

Aquino, Jakson Alves de (2015). *The effect of exposure to political institutions and economic events on demand for democracy in Africa*. Afrobarometer, Working paper No. 160.

Basu, Alaka Malwade (2002). "Why does education lead to lower fertility? A critical review of some of the possibilities". In: *World Development* 30(10), pp. 1779–1790.

Blaydes, Lisa (2006). *Who votes in authoritarian elections and why? Determinants of voter turnout in contemporary Egypt*. Mimeo, presented at Annual Meeting of the American Political Science Association. Philadelphia, PA, August.

Briggs, Ryan C. (2019). "Receiving foreign aid can reduce support for incumbent presidents". In: *Political Research Quarterly* 72(3), pp. 610–622.

Bruch, Sarah K and Joe Soss (2018). "Schooling as a formative political experience: Authority relations and the education of citizens". In: *Perspectives on Politics* 16(1), pp. 36–57.

Callaway, Brantly and Pedro HC Sant'Anna (2020). "Difference-in-differences with multiple time periods". In: *Journal of Econometrics*. Publisher: Elsevier.

Campbell, Harlan and Paul Gustafson (2018). "The validity and efficiency of hypothesis testing in observational studies with time-varying exposures". In: *Observational Studies* 4(1), pp. 260–291.

Card, David (1995). "Using geographic variation in college proximity to estimate the return to schooling". In: *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. Ed. by E. K. Grant L. N. Christofdes and R. Swidinsky. University of Toronto Press, pp. 201–222.

Card, David (1999). "The causal effect of education on earnings". In: *Handbook of Labor Economics*. Ed. by Orley C. Ashenfelter and David Card. Vol. 3. Elsevier, pp. 1801–1863.

Cha, Sujin (2020). *Does foreign aid raise awareness of corruption? Evidence from Chinese aid in 30 African states*. Tech. rep. AidData Working Paper.

Champagne, Eric and Ben Mamadou Ouedraogo (2011). "Decentralization in Burkina Faso: A policy reform process in slow motion". In: *Decentralization in Developing Countries: Global Perspectives on the Obstacles to Fiscal Devolution*. Ed. by Jorge Martinez-Vazquez and Francois Vaillancourt. Edward Elgar, Cheltenham, UK, pp. 99–116.

Choi, Donghyun Danny, Benjamin Laughlin, and Anna E Schultz (2021). *Mobile communication technology and national identity in Sub-Saharan Africa*. OSF Preprints. June 17.

Cirone, Alexandra and Thomas B. Pepinsky (2021). "Historical persistence". In: *Annual Review of Political Science* 25.

Coffe, Hilde and Catherine Bolzendahl (2011). "Gender gaps in political participation across Sub-Saharan African nations". In: *Social Indicators Research* 102(2), pp. 245–264.

Colclough, Christopher, Pauline Rose, and Mercy Tembon (2000). "Gender inequalities in primary schooling: The roles of poverty and adverse cultural practice". In: *International Journal of Educational Development* 20(1), pp. 5–27.

Correia, Sergio (2015). *Singletons, cluster-robust standard errors and fixed effects: A bad mix.* Mimeo, Duke University.

Croke, Kevin, Guy Grossman, Horacio A. Larreguy, and John Marshall (2016). "Deliberate disengagement: How education can decrease political participation in electoral authoritarian regimes". In: *American Political Science Review* 110(3), pp. 579–600.

Currie, Janet and Enrico Moretti (Nov. 2003). "Mother's education and the intergenerational transmission of human capital: Evidence from college openings". In: *The Quarterly Journal of Economics* 118(4), pp. 1495–1532.

Dang, Thang (2019). "Quasi-experimental evidence on the political impacts of education in Vietnam". In: *Education Economics* 27(2), pp. 207–221.

De Chaisemartin, Clément and Xavier D'Haultfoeuille (2020). "Two-way fixed effects estimators with heterogeneous treatment effects". In: *American Economic Review* 110(9), pp. 2964–96.

De Chaisemartin, Clément and Xavier D'Haultfoeuille (2021). *Difference-in-Differences estimators of intertemporal treatment effects*. Mimeo.

Diwan, Ishac and Irina Vartanova (2020). "Does education indoctrinate?" In: *International Journal of Educational Development* 78, p. 102249.

Djemai, Elodie, Anne-Laure Samson, and Yohan Renard (2022). *Mothers and fathers : Education, co-residence and child health*. Document de travail du LEM / Discussion paper LEM.

Duflo, Esther (2001). "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment". In: *American Economic Review* 91(4), pp. 795–813.

Elischer, Sebastian (2013). *Political parties in Africa: Ethnicity and party formation*. Cambridge University Press.

Freeman, Guy, Benjamin J. Cowling, and C. Mary Schooling (2013). "Power and sample size calculations for Mendelian randomization studies using one genetic instrument". In: *International Journal of Epidemiology* 42(4), pp. 1157–1163.

Friedman, Willa, Michael Kremer, Edward Miguel, and Rebecca Thornton (2016). "Education as liberation?" In: *Economica* 83(329), pp. 1–30.

Fuchs-Schündeln, Nicola and Matthias Schündeln (2015). "On the endogeneity of political preferences: Evidence from individual experience with democracy". In: *Science* 347(6226), pp. 1145–1148.

Gelman, Andrew and Eric Loken (2014). "The statistical crisis in science: data-dependent analysis–'A garden of forking paths'–explains why many statistically significant comparisons don't hold up". In: *American Scientist* 102(6), pp. 460–466.

Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics*, pp. 254–277.

Hudson, Sally, Peter Hull, and Jack Liebersohn (2017). *Interpreting instrumented difference-in-differences*. Mimeo.

Isaksson, Ann-Sofie, Andreas Kotsadam, and Måns Nerman (2014). "The gender gap in African political participation: Testing theories of individual and contextual determinants". In: *The Journal of Development Studies* 50(2), pp. 302–318.

Kang, Hyunseung, Yang Jiang, Qingyuan Zhao, and Dylan S Small (2021). "Ivmodel: an R package for inference and sensitivity analysis of instrumental variables models with one endogenous variable". In: *Observational Studies* 7(2), pp. 1–24.

Keane, Michael and Timothy Neal (2021). *A practical guide to weak instruments*. Discussion Papers 2021-05c, School of Economics, The University of New South Wales.

Khemani, Stuti, Ernesto Dal Bó, Claudio Ferraz, Frederico Shimizu Finan, Corinne Louise Stephenson Johnson, Adesinaola Michael Odugbemi, Dikshya Thapa, and Scott David Abrahams (2016). *Making politics work for development: Harnessing transparency and citizen engagement*. The World Bank.

Knutsen, Carl Henrik, Andreas Kotsadam, Eivind Hammersmark Olsen, and Tore Wig (2017). "Mining and local corruption in Africa". In: *American Journal of Political Science* 61(2), pp. 320–334.

Knutsen, Tora and Andreas Kotsadam (2020). "The political economy of aid allocation: Aid and incumbency at the local level in Sub Saharan Africa". In: *World Development* 127, p. 104729.

Konte, Maty and Rose Camille Vincent (2021). "Mining and quality of public services: The role of local governance and decentralization". In: *World Development* 140, p. 105350.

Kuenzi, Michelle T. (2006). "Nonformal education, political participation, and democracy: Findings from Senegal". In: *Political Behavior* 28(1), pp. 1–31.

Larreguy, Horacio and John Marshall (2017). "The effect of education on civic and political engagement in nonconsolidated democracies: Evidence from Nigeria". In: *Review of Economics and Statistics* 99(3), pp. 387–401.

Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack R. Porter (2021). *Valid t-ratio inference for IV*. Working paper, National Bureau of Economic Research.

Lieberman, Evan and Yang-Yang Zhou (2020). "Self-efficacy and citizen engagement in development: Experimental evidence from Tanzania". In: *Journal of Experimental Political Science*, pp. 1–18.

Loada, Augustin and Carlos Santiso (2002). *Elections historiques au Burkina Faso: Vers une maturité démocratique?* Mimeo, Centre pour la gouvernance démocratique Burkina Faso.

Mahieu, Sylvie and Serdar Yilmaz (2010). "Local government discretion and accountability in Burkina Faso". In: *Public Administration and Development* 30(5), pp. 329–344.

Marshall, John (2016). "Coarsening bias: How coarse treatment measurement upwardly biases instrumental variable estimates". In: *Political Analysis* 24(2), pp. 157–171.

Mattes, Robert and Dangalira Mughogho (2009). *The limited impacts of formal education on democratic citizenship in Africa.* Afrobarometer, Working paper No. 109.

Mittag, Nikolas (2012). *TWFE: Stata module to perform regressions with two-way fixed effects or match effects for large datasets.* Boston College Department of Economics.

Ozdemir, Ugur, Ali Ozkes, and Remzi Sanver (2021). *Ability or motivation? Voter registration and turnout in Burkina Faso.* Working Paper.

Paglayan, Agustina S (2021). "The non-democratic roots of mass education: evidence from 200 years". In: *American Political Science Review* 115(1), pp. 179–198.

Parinduri, Rasyad A. (2019). "Does education increase political participation? Evidence from Indonesia". In: *Education Economics* 27(6), pp. 645–657.

Santiso, Carlos and Augustin Loada (2003). "Explaining the unexpected: electoral reform and democratic governance in Burkina Faso". In: *The Journal of Modern African Studies* 41(3), pp. 395–419.

Soboleva, Irina (2020). "Efficacy, openness, ingenuousness: Micro-foundations of democratic engagement". PhD thesis. Columbia University.

Staiger, D. and J.H. Stock (1997). "Instrumental variables regression with weak instruments". In: *Econometrica* 65, pp. 557–586.

Sun, Liyang and Sarah Abraham (2020). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics*, pp. 175–199.

Tripp, Aili Mari (2001). "Women and democracy: The new political activism in Africa". In: *Journal of Democracy* 12(3), pp. 141–155.

Walker, Venexia M., Neil M. Davies, Frank Windmeijer, Stephen Burgess, and Richard M. Martin (2017). "Power calculator for instrumental variable analysis in pharmacoepidemiology". In: *International Journal of Epidemiology* 46(5), pp. 1627–1632.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta (2015). "Education and human capital externalities: evidence from colonial Benin". In: *The Quarterly Journal of Economics* 130(2), pp. 703–757.

Watkins, Mitchell (2021). "Foreign aid projects and trust in political institutions". In: *Governance.*

Wolak, Jennifer (2020). "Self-confidence and gender gaps in political interest, attention, and efficacy". In: *The Journal of Politics* 82(4), pp. 1490–1501.

Appendix A

This appendix describes the data cleaning process applied to the list *Répertoire des écoles publiques et privées du Burkina Faso en 2004-2005* produced by the *Direction des études et de la planification* of the *Ministère de l'enseignement de base et de l'alphabétisation* (MEBA) produced in March 2006. The goal of the data cleaning was to be able to merge the school list to the Afrobarometer villages and to geo-locate all schools by their village.

Two features of the list require researcher data coding choices. The first is that there is no variable for 'village' which is the basic unit of geographic location for Burkina Faso. The second is that the school lists are organized by their *Circonscription d'Education de Base* (CEB) which was the basic geographic unit for administration of schools of the MEBA. Each CEB was supervised by a *Direction Provinciale* (DPEBA) and then by a *Direction Régionale* (DREBA). The schools are not grouped by the political administrative unit of the *commune*. A *commune*, in Burkina Faso, is the primary or lowest layer for geographically defined administrative entities. Under the decentralization law of 2006, every locality in the country is in the jurisdiction of a commune with an elected council (Champagne and Ouedraogo (2011); Mahieu and Yilmaz (2010)). Urban communes often include some rural villages within their boundaries. Most of the CEB overlap with communes, but in many instances the CEB might be located in an urban commune and the inspectors will have responsibility for some urban schools and some rural schools in the rural commune neighboring the urban commune.

We deal with these issues in two ways. First, we wrote code that strips away all common words, numbers, punctuation and extraneous letters from school titles. We call that the "imputed village name." Very frequently, the words that remained were indeed the name of the village (i.e., "Ecole de Tanghin, garcons sud" became "Tanghin", almost certainly the name of the village where the school was located). But sometimes they were clearly not (i.e., "1200 logements" is the name of a residential district in Ouagadougou, and not of a village). Since the list classified some schools as urban, we determined the set of 30 communes that included 10 or more urban schools. We then merged that list with the schools list: if the imputed village name matched the village name listed as part of the commune, then it was retained. If the imputed village name did not however match any village name in the urban commune, then the urban commune became the name of the village (i.e., "1200 logements"became Ouagadougou).

Misclassifications here are not likely to be significantly troublesome, since as we see below we coded an individual's access to schooling as a distance radius (1km, 5km, 10km, 30km), and many communes would thus include the other villages in the commune in this radius. What we worry about, though, is that a village with a school with a peculiar name may have been mistakenly listed as urban (the school list indicates urban or rural, but there is no explanation for the criterion), and so was assigned to the town of the commune, rather than the village. Our understanding of school names and village names suggest this problem was likely to be rare.

To deal with the second problem, that some CEB encompassed more than one commune, we carried out a similar exercise. We created a list of all villages with unique names, and their provinces and communes. We then merged that list onto the school list, using the village name and province as the match. If the CEB of the school list did not match the commune from the IGB list, than the CEB was replaced with the IGB commune.

With these two modifications to the list, we obtained a list of 6,917 schools with founding dates, by village (or town), commune, province, and region. We then matched these to the IGB geographic spatial database, by village (or town), commune, province, and region. There were 5,524 that matched exactly. We then conducted a fuzzy match, and visually inspected these matches, to obtain about 1,000 more matches. We thus had a list of 6,500 schools that were correctly geo-located. The remaining 400 schools for which the school name and CEB did not permit an automated match were manually inspected and matched to a village and commune. There were a variety of patterns in this manual matching, and thus a few rules of thumb for coding led to some consistency in the choices of matches.

First, some villages in Burkina Faso have numbers or directions indicating the village, for example in some irrigation projects or resettlement zones. So villages might be Rakaye 1, Rakaye 2, Rakaye 3, or Solle Sud and Solle Nord, etc. In the other list the village might simply be Rakaye or Solle. In these cases, we match to the number 1 village or to the north village.

Second, some villages also are described by the predominant ethnic group, in areas where ethnic groups are more segregation. So there might be Waria Mossi, Waria Peulh, and Waria Yarce. The school village might be Waria. In this case, we usually assign the school to the Waria Mossi village. This assignment may of course have some error, but since we use two measures of school establishment (in the village and within a distance radius) these villages, which are usually fairly proximate, will be counted in their respective distance radii.

Third, there is considerable non-standardization of spelling for all of the local languages of Burkina Faso, and village names are almost all local names. People from the same village may spell their village in different ways. So a village might have many different spellings: Ouarankuy, Warankui, Ourankui, Warounkuy, Walankuy, Oualranko. Further complexity is added when consonants in the French spelling do not correspond to consonant sounds in local languages: r, l, and lr may all stand for the same sound in a local language. Moreover, many village names are variants of locally-important words or phrases (the equivalent of “Big Tree,” “Tall Trees,” and “Tree Grove”) and so may be pronounced similarly in local languages, so their written names may overlap. A poor spelling of a village name might end up being more close to the correct spelling of another village name. This poses a challenge for an iterative algorithm to correct names and match names because changing a name on a second list that has a name that reasonably matches then changes the names available to match to a third and fourth list. So two villages on two different lists, Walankuy and Ouarankuy, are matched and standardized to Ouarankuy. But a third list has a village Wargankuy that might have been a good match for Walankuy but instead matches to Walugaye.

Appendix B: F-statistics for "first stage" regressions

This appendix presents results from the various "first-stage" specifications discussed in Section 6.1. The outcome variables are: a binary variable for whether the person had any schooling; a binary if completed primary school; a binary variable for some secondary; and the scale value of schooling attainment. We first present two summary tables with the F-statistics for the first-stage regressions presented in section 6.1.

Tables 3 and 4 present results from regressions where the principal explanatory variable is either a dummy variable for whether the respondent was school age eligible after school establishment in the village of residence (Panel A), or the weighted average of the number of primary schools when the person was of school age (Panel B), for men and for women, respectively. Each specification has a dummy for whether the Afrobarometer survey round was post-2014. There are also two sets of dummy variables included, cohort of birth (year, or 5-year group) and locality of residence (at either the province, commune, or village level), as indicated. As noted in the text, the results in Tables 3 and 4 suggest that the effect of school establishment on schooling attainment was sizable. Persons who were born after a school was established saw substantial increases on average in all four measures of schooling outcomes, relative to those born earlier.

The relevant F-statistics for the variables that constitutes the two possible instrumental variables (primary school access in the village, or inverse-distance weighted average of number of schools) are displayed in Table B1. The table for the male sample suggests that access to primary school within the village or proximity to neighboring schools produces a reasonably convincing first-stage F-statistic, where "convincing" ranges from the value greater than 10 that remains a common rule of thumb (Staiger and Stock (1997)) to the more recent reminders that F-statistics on the order of 100 are needed for proper hypothesis testing (Lee et al. (2021); Keane and Neal (2021)). In the case of the village-level dummy variables, the estimated coefficients of interest approach zero and are often no longer statistically significant. The village geographic unit typically only includes 8 persons, and with controls for gender, survey round and birth cohort there is often no variation in whether respondents were of school age before or after school establishment. The F-statistics for this most demanding specification is below 10 for both instrumental variables.

For the women sample, Table B2 suggests again that access to primary school within the village or proximity to neighboring schools produces a reasonably convincing first-stage F-statistic. Again, the specification with the village-level adn birth-year level dummy variables has lower F-statistics and here close to zero, so these specifications are not used for the IV estimation.

Specifications with both variables included as explanatory variables (primary school access in the village and inverse-distance weighted number of primary schools) produced about the same F-statistics for men and women samples. There was no gain in precision of estimates.

Specifications were estimated using the Stata command `reghdfe`. The command drops singletons in clusters of two-way fixed effects. For this reason, the sample sizes may vary slightly

(Correia (2015)).

Table B1: F-statistics of various specifications of first-stage regressions of effects of school in village or near village, when individual of school age, on education outcomes, men only

	(1) Some education	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Dummy variable of school in village			
No cohort or locality controls	177.6	122.2	107.3	146.6
Birth year, province	86.8	62.7	44.3	78.4
5-year cohort, commune	62.5	48.6	33.7	60.9
Birth year, village	9.8	5.6	2.2	8.7
	Panel B: Number of schools weighted by proximity			
No cohort or locality controls	86.0	138.6	112.8	101.4
Birth year, province	72.8	81.6	69.4	69.4
5-year cohort, commune	48.3	37.4	38.4	40.1
Birth year, village	7.2	4.1	7.5	6.2

Notes: Each F-statistic in the table is for a separate regression, according to the specification (row) and outcome measure (column) and measure of access to schools (panel). All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition) and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `reghdfe`. Data is from Afrobarometer survey, Burkina Faso, rural residents born after 1950, various rounds, men only, and administrative data on year of school establishment.

Table B2: F-statistics of various specifications of first-stage regressions of effects of school in village or near village, when individual of school age, on education outcomes, women only

	(1) Some education	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Dummy variable of school in village			
No cohort or locality controls	119.6	88.4	75.5	98.1
Birth year, province	59.3	40.8	29.8	48.4
5-year cohort, commune	22.2	12.1	7.6	15.9
Birth year, village	0.0	0.6	1.4	0.4
	Panel B: Number of schools weighted by proximity			
No cohort or locality controls	46.9	37.7	34.8	39.0
Birth year, province	34.6	37.6	32.3	35.4
5-year cohort, commune	22.5	30.8	22.4	27.2
Birth year, village	3.0	1.5	3.6	3.3

Notes: Each F-statistic in the table is for a separate regression, according to the specification (row) and outcome measure (column) and measure of access to schools (panel). All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition) and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `reghdfe`. Data is from Afrobarometer survey, Burkina Faso, rural residents born after 1950, various rounds, women only, and administrative data on year of school establishment.

Figure B1 displays event study graphs for the three different specification (no controls, half-decade and commune controls, and birth-year and province controls). The outcome variable is the education attainment level. The event study graph plots the coefficients from specifications that include ten leads and ten lags, with the furthest lead and lag accumulating the school establishment dummy prior to or after the lag or lead. Coefficients on lagged indicators of school establishment are not significant, except for one, and are centered on zero, for the most part. Coefficients on schooling outcomes for those who were around seven years old or younger at the time of schooling establishment are positive and with confidence intervals above zero. The effect seems to diminish, though, and the lead coefficient for 5 years (the effect on education attainment for those who were 2 years old when the school was established) is zero when we add the geographic and birth cohort dummies, though it is higher for subsequent leads. The results are quite similar when using the dummy variable for whether the person had some schooling.

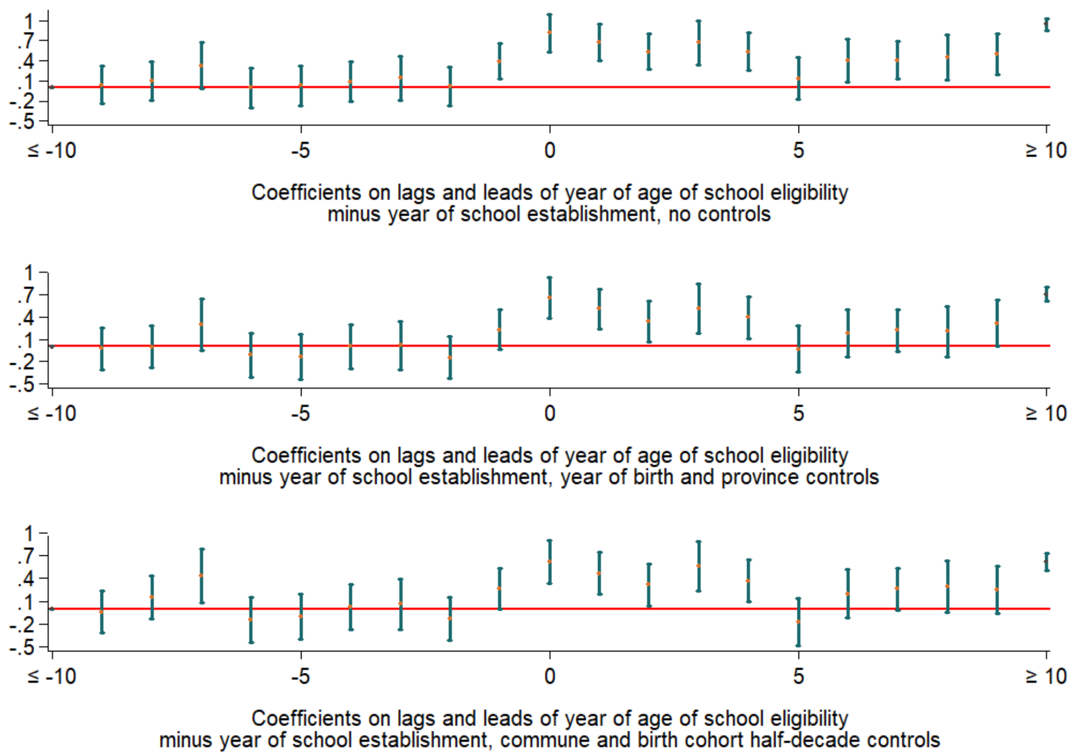

Figure B1: Event study style graph of coefficients of lags and leads of treatment variable effect on education attainment level

Source: Afrobarometer data, various rounds, rural residents only, birth year after 1950. The sample used in the regression estimation pooled men and women together, and included a dummy variable control for whether the person was male.

Appendix C: Alternative first-stage regression models

In this appendix we discuss three alternative specification strategies for the first stage regression. These are the Probit specification, the two-way fixed effects specification, and the regression discontinuity specification.

Probit Tables C1 and C2 present the results from probit regressions for the binary education outcome variables (whether the person had any schooling; whether the person completed primary school; and whether had some secondary education). The probit specifications were estimated with Stata command `probit`. The coefficients presented are the marginal effects calculated by the Stata option `margins, dydx`. As in the linear probability model of Tables 3 and 4, the coefficients on the two indicators of access to schooling (school in the village, and inverse-distance weighted number of schools) are sizable and statistically significant. The probit specifications yield a somewhat better fit in particular for the specification with village geographic fixed effects along with birth year controls, compared with the linear probability model. The results reinforce the validity of our instrumental variables.

Table C1: Probit estimate of effects of school in village or near village, when individual of school age, on education outcomes, various specifications, men only

	(1) Some education	(2) Completed primary	(3) Some secondary
	Panel A: Dummy variable of school in village		
No cohort or locality controls	0.298***	0.238***	0.210***
	(0.019)	(0.019)	(0.018)
Birth year, province	0.224***	0.180***	0.151***
	(0.022)	(0.020)	(0.020)
5-year cohort, commune	0.209***	0.180***	0.157***
	(0.024)	(0.023)	(0.023)
Birth year, village	0.162***	0.142**	0.102
	(0.058)	(0.062)	(0.069)
	Panel B: Number of schools weighted by proximity		
No cohort or locality controls	0.095***	0.082***	0.075***
	(0.005)	(0.005)	(0.005)
Birth year, province	0.096***	0.083***	0.066***
	(0.008)	(0.007)	(0.007)
5-year cohort, commune	0.090***	0.084***	0.072***
	(0.009)	(0.008)	(0.009)
Birth year, village	0.082***	0.086***	0.068**
	(0.026)	(0.028)	(0.030)

Notes: Each coefficient and standard error in the table is for a separate regression, according to the specification (row) and outcome measure (column) and measure of schooling access (panel). For Panel A the coefficient is that of the dummy variable indicating if a school was present, for Panel B the coefficient is the variable representing the count of schools, each weighted by the inverse distance squared, where distance is the measure in kilometers of the distance from the village of residence to the village of the school. Standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `probit`. * p <0.10,** p <0.05,*** p <0.01. Data is from Afrobarometer survey, Burkina Faso, rural residents born after 1950, various rounds, men only, and administrative data on year of school establishment.

Table C2: Probit estimate of effects of school in village or near village, when individual of school age, on education outcomes, various specifications, women only

	(1) Some education	(2) Completed primary	(3) Some secondary
	Panel A: Dummy variable of school in village		
No cohort or locality controls	0.229***	0.182***	0.158***
	(0.020)	(0.019)	(0.018)
Birth year, province	0.167***	0.138***	0.118***
	(0.020)	(0.020)	(0.019)
5-year cohort, commune	0.128***	0.101***	0.089***
	(0.024)	(0.023)	(0.023)
Birth year, village	0.046	0.070	0.076
	(0.059)	(0.079)	(0.098)
	Panel B: Number of schools weighted by proximity		
No cohort or locality controls	0.092***	0.076***	0.066***
	(0.006)	(0.006)	(0.006)
Birth year, province	0.089***	0.077***	0.067***
	(0.007)	(0.007)	(0.007)
5-year cohort, commune	0.077***	0.067***	0.057***
	(0.010)	(0.009)	(0.010)
Birth year, village	0.019	0.016	0.016
	(0.027)	(0.037)	(0.050)

Notes: Each coefficient and standard error in the table is for a separate regression, according to the specification (row) and outcome measure (column) and measure of schooling access (panel). For Panel A the coefficient is that of the dummy variable indicating if a school was present, for Panel B the coefficient is the variable representing the count of schools, each weighted by the inverse distance squared, where distance is the measure in kilometers of the distance from the village of residence to the village of the school. Standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `probit`. * p <0.10,** p <0.05,*** p <0.01. Data is from Afrobarometer survey, Burkina Faso, rural residents born after 1950, various rounds, women only, and administrative data on year of school establishment.

Two-way Fixed Effects Another possible specification is a two-way fixed effect (TWFE) strategy that is a generalization of the difference-in-differences (DID) strategy (Mittag (2012); Sun and Abraham (2020); Goodman-Bacon (2021); De Chaisemartin and D'Haultfoeuille (2020)). School construction programs take place slowly over time, or reforms affect different localities at different times, and so the change in access or intensity of access is best described as a 'staggered treatment.' The standard DID setting with two groups and two time periods may not apply to settings where multiple groups are observed over a large time period and where all treated groups do not receive the treatment at the same time (staggered adoption design).

For all pairs of consecutive time periods $t-1$ and t, the DID compares the outcome evolution among groups whose treatment changes from $t-1$ to t (switchers) and groups whose treatment does not change. We might expect that the effect of building a new school in the village is different depending on the year of construction. The effect might be very high at the beginning of the period when there are very few alternative schools in the surrounding villages, or on the opposite very small at the beginning of the period because the perceived benefits of school enrolment were lower in the 1960s than those in the 2000s and the government was not encouraging school attendance at that time.

The TWFE strategy is to use all localities without a new school (or policy change) at a particular time as controls for the treatment localities where new schools are established or where schooling policies are changed. In the 'naive' TWFE no regard is given to the quite different nature of the localities in the control group. These include: 'never switch' localities that will never get schools or policy reforms; 'later switch' localities that will have schools established or reforms implemented later relative to the treatment locality; and 'previously switched' localities that had schools established or reforms adopted prior to the treatment locality, and where an assumption is made that recent treatment will have a different effect on outcomes than treatment in earlier years. Thus, the average effect of school availability or schooling reform is estimated by calculating a weighted average of the differences between those individuals in a locality eligible for schooling in a particular year, and similarly aged children in other localities in that same year (who may be in one of the three different kinds of localities).

Several papers have proposed a variety of TWFE estimators that do take into account the compositional issues discussed above, and have applied them in the context of using the TWFE estimate as a first stage in a two-stage least squares approach. These papers relax the assumption that the treatment effect is constant over time. Callaway and Sant'Anna (2020); Sun and Abraham (2020); De Chaisemartin and D'Haultfoeuille (2021) have proposed estimators that apply when the treatment is binary and staggered and there are dynamic heterogeneous effects. The main difference between Sun and Abraham (2020) and De Chaisemartin and D'Haultfoeuille (2021) estimators is the control group used in the analysis. Sun and Abraham (2020) use as controls the never-treated groups (or the groups treated in the very last period if all groups are treated by the end of the period) while De Chaisemartin and D'Haultfoeuille (2021) consider all the not-yet-treated observations in a larger control group.

The De Chaisemartin and D'Haultfoeuille (2021) estimator, denoted by DID_M is the average of DIDs across all pairs of periods. Under a parallel trends assumption, DID_M is an unbiased and consistent estimator of the average treatment effect among switchers, at the time period when they get the treatment. In staggered adoption designs where additional groups receive the treatment and already treated groups remain treated, the estimator of dynamic effects is similar to the DID_M estimator, except that it makes use of long differences of the outcome (e.g. from t-1 to t+1) rather than first differences. In other words it can compute estimators of switchers' dynamic treatment effects, one time period or more after the switchers have started receiving the treatment. This is relevant in our case to estimate and discuss the size of the instantaneous effect and the longer-term effects.

The De Chaisemartin and D'Haultfoeuille (2021) estimator does not perfectly apply to our setting for several reasons. First, all groups are observed at unevenly spaced intervals. Within a village, we do not observe respondents from all the birth years of the panel (1950-2000). Second, in some villages, we do not observe someone eligible the year right before the construction, the given year, and the year after the construction. These observations are required to compute estimators based on first differences. For instance, in the village Bouere in region Hauts-Bassins, the school was established in 1985 and there is no respondent who turned seven years old in 1984 and in 1985. The sample includes six respondents born in 1968, 1978, 1981, 1987, 1988, and in 1996. As we do not observe the outcome of the relevant cohorts, this location can not be used to compute the DID_M estimator. The majority of villages would then be dropped, in implementing this estimator. Third, even though within a village, we might observe respondents eligible around the year of school establishment, at most it would be two respondents for each year and the law of large numbers does not hold. This would require us to make the very strong assumption that these two respondents are representative of their respective birth cohort. Fourth, the size of our sample within groups (village, cohort) and the number of switchers we have, do not allow us to estimate dynamic effects for more than two years after the school establishment (same for the placebo estimation of the effects before the school establishment to check for parallel trends). Last, in Burkina Faso, as in other places (see Colclough, Rose, and Tembon (2000) for Guinea), it is often the case that children start school later than the official starting age. It is then difficult to argue that children born one or two years before the first cohort who benefit from the school construction were excluded from that school. (The difference between gross and net enrolment rates in places like Burkina Faso reflects this phenomenon of repeated school years and delayed school entry. The gross enrolment rate is on average 7-15 percentage points larger than the net enrolment rates, depending on the year of observation of enrolment rates).

Regression Discontinuity Another alternative specification is a regression discontinuity design (RDD) strategy where the identification relies on the discontinuity around the cutoff point. In our setting, the running variable is then equal to the birth year + 7 - the year of school opening and is equal to zero for the cohort who become eligible the year of the school opening. The cutoff point is when the running variable is equal to zero and we should find that a discontinuity at the cutoff, meaning that there is a sharp difference in school attainment between the children who become eligible for schooling (turn seven years old) on the year of school opening and the children who become eligible one year before schooling opening.

It turns out that we do not see a discontinuity if we graph the schooling outcomes over the running variable values, and this holds for all education measures: having some schooling, having completed primary school, having some secondary or the education level. This might be due to the catching-up of the children who were 8 or 9 when the school was established in their village and may have benefited from this increased access to school. We might then withdraw the partially treated from the analysis but were already working with a small sample as there are 71 observations at the cutoff year, and about 37 observations on average (median of 31) for each value of the running variable, in the pooled sample, men and women included.

Appendix D: OLS and reduced-form equations

Table D1: Correlation between measures of political engagement and schooling outcomes, by gender

	(1) Some education	(2) Completed primary	(3) Some secondary	(4) Education attainment
1. Index, men sample	-0.031	-0.074	-0.099*	-0.023
	(0.041)	(0.047)	(0.056)	(0.015)
1. Index, women sample	-0.202***	-0.214***	-0.288***	-0.082***
	(0.052)	(0.062)	(0.068)	(0.020)
2. Voting, men sample	-0.052***	-0.056**	-0.064**	-0.021***
	(0.019)	(0.022)	(0.025)	(0.007)
2. Voting, women sample	-0.117***	-0.122***	-0.127***	-0.047***
	(0.024)	(0.028)	(0.030)	(0.009)
3. Joining issue, men sample	0.006	-0.012	-0.021	-0.001
	(0.020)	(0.022)	(0.026)	(0.007)
3. Joining issue, women sample	-0.043**	-0.042*	-0.070**	-0.015*
	(0.022)	(0.025)	(0.028)	(0.008)
4. Part. meeting, men sample	-0.009	-0.034*	-0.046*	-0.010
	(0.018)	(0.020)	(0.023)	(0.007)
4. Part. meeting, women sample	-0.066***	-0.089***	-0.125***	-0.032***
	(0.022)	(0.027)	(0.029)	(0.009)
5. Part. in community group, men sample	0.046*	0.028	-0.020	0.007
	(0.024)	(0.027)	(0.029)	(0.009)
5. Part. in community group, women sample	-0.020	-0.071**	-0.105***	-0.018*
	(0.026)	(0.030)	(0.033)	(0.010)

Notes: Dependent variable is alternatively index of political engagement, or each of the four dummy variables (whether voted in the prior election, whether participated in a meeting, whether joined with others to raise an issue, and whether a member of a community group (only available for rounds 4-7)). Data on the outcome is from the Afrobarometer Burkina Faso data, rural residents born after 1950, various rounds. Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include a dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and are estimated with Stata command `reghdfe`. * $p < 0.10$,** $p < 0.05$,*** $p < 0.01$.

Table D2: Reduced form equation on index of political engagement, with the instrumental variables

	(1)	(2)	(3)
Sample	Whole	Men	Women
	Panel A: IV dummy variable of school in village		
No cohort or locality controls	-0.239***	-0.231***	-0.175***
	(0.035)	(0.044)	(0.049)
Birth year, province	-0.038	-0.069	-0.014
	(0.038)	(0.047)	(0.055)
5-year cohort, commune	-0.017	-0.091*	0.017
	(0.041)	(0.052)	(0.059)
	Panel B: IV number of schools weighted by proximity		
No cohort or locality controls	-0.077***	-0.078***	-0.063***
	(0.010)	(0.017)	(0.013)
Birth year, province	-0.028***	-0.033**	-0.029*
	(0.009)	(0.015)	(0.016)
5-year cohort, commune	-0.012	-0.034*	-0.003
	(0.013)	(0.019)	(0.023)

Notes: Dependent variable is index of political engagement, the sum of three dummy variables (whether voted in the prior election, whether participated in a meeting, and whether joined with others to raise an issue). Data on the outcome is from the Afrobarometer Burkina Faso data, rural residents born after 1950, various rounds. Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include a dummy variable for post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `reghdfe`. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix E: Instrumental variable regressions using proximity to school as instrumental variable

Table E1: Effects of education on the four measures of political engagement, with various measures of education, and with weighted number of schools when school eligible as instrumental variable, men sample

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Voted in the previous election			
No cohort or locality controls	-0.342***	-0.403***	-0.449***	-0.115***
	(0.059)	(0.071)	(0.082)	(0.020)
Birth year, province	0.092	0.108	0.149	0.032
	(0.064)	(0.078)	(0.106)	(0.023)
5-year cohort, commune	0.051	0.059	0.080	0.017
	(0.092)	(0.110)	(0.145)	(0.032)
	Panel B: Joined with others to raise an issue			
No cohort or locality controls	-0.230***	-0.265***	-0.294***	-0.077***
	(0.072)	(0.086)	(0.096)	(0.025)
Birth year, province	-0.048	-0.057	-0.076	-0.017
	(0.091)	(0.108)	(0.145)	(0.032)
5-year cohort, commune	0.031	0.037	0.048	0.011
	(0.106)	(0.124)	(0.161)	(0.036)
	Panel C: Member of a community group			
No cohort or locality controls	-0.295***	-0.343***	-0.380***	-0.101***
	(0.080)	(0.095)	(0.105)	(0.028)
Birth year, province	-0.169	-0.212	-0.292	-0.063
	(0.104)	(0.132)	(0.186)	(0.039)
5-year cohort, commune	-0.264*	-0.320*	-0.438*	-0.096*
	(0.143)	(0.179)	(0.257)	(0.053)
	Panel D: Participated in a meeting			
No cohort or locality controls	-0.300***	-0.349***	-0.385***	-0.101***
	(0.063)	(0.076)	(0.085)	(0.022)
Birth year, province	-0.145*	-0.177*	-0.232*	-0.052*
	(0.077)	(0.092)	(0.125)	(0.027)
5-year cohort, commune	-0.219**	-0.257**	-0.334*	-0.076**
	(0.109)	(0.126)	(0.171)	(0.038)

Notes: Afrobarometer Burkina Faso data, rural residents born after 1950, five rounds (except in Panel C, rounds 4-7). Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

Table E2: Effects of education on the four measures of political engagement, with various measures of education, and with weighted number of schools when school eligible as instrumental variable, women sample

Education measure	(1) Some schooling	(2) Completed primary	(3) Some secondary	(4) Education attainment
	Panel A: Voted in the previous election			
No cohort or locality controls	-0.250***	-0.320***	-0.380***	-0.097***
	(0.078)	(0.101)	(0.121)	(0.030)
Birth year, province	0.103	0.136	0.168	0.042
	(0.100)	(0.131)	(0.163)	(0.040)
5-year cohort, commune	0.080	0.111	0.164	0.035
	(0.141)	(0.199)	(0.295)	(0.062)
	Panel B: Joined with others to raise an issue			
No cohort or locality controls	-0.367***	-0.459***	-0.543***	-0.141***
	(0.077)	(0.097)	(0.116)	(0.030)
Birth year, province	-0.091	-0.119	-0.149	-0.037
	(0.110)	(0.144)	(0.180)	(0.044)
5-year cohort, commune	-0.013	-0.019	-0.028	-0.006
	(0.170)	(0.237)	(0.352)	(0.074)
	Panel C: Member of a community group			
No cohort or locality controls	-0.312***	-0.396***	-0.449***	-0.120***
	(0.097)	(0.123)	(0.138)	(0.038)
Birth year, province	-0.083	-0.110	-0.133	-0.033
	(0.136)	(0.180)	(0.216)	(0.055)
5-year cohort, commune	-0.010	-0.013	-0.018	-0.004
	(0.205)	(0.263)	(0.382)	(0.086)
	Panel D: Participated in a meeting			
No cohort or locality controls	-0.381***	-0.477***	-0.565***	-0.147***
	(0.080)	(0.100)	(0.118)	(0.031)
Birth year, province	-0.095	-0.124	-0.154	-0.038
	(0.111)	(0.145)	(0.179)	(0.044)
5-year cohort, commune	0.104	0.145	0.214	0.045
	(0.165)	(0.232)	(0.349)	(0.072)

Notes: Afrobarometer Burkina Faso data, rural residents born after 1950, five rounds (except in Panel C, rounds 4-7). Coefficients of the education measure are reported and their standard errors, clustered at the village level, are in parentheses. All specifications include dummy variable for two post-2014 survey rounds (after popular uprising leading to democratic transition), and birth cohort and geographic dummy variables as indicated, and are estimated with Stata command `ivreghdfe`. * p <0.10,** p <0.05,*** p <0.01.

Appendix F: Power

We may evaluate the power of the estimation approach for the four separate components on the index of political engagement using the R package `ivmodel`, developed by Kang et al. (2021), that implements the power formula developed by Freeman, Cowling, and Schooling (2013). Figure F1 shows the calculated power for specifications with the binary outcome variable indicating whether the person participated in a meeting. Power is calculated for the three education indicators considered (some schooling, completed primary, and education level), with whether there was a school in the village at the time the person was eligible for school as the instrumental variable, and different sets of control variables, as indicated. The coefficients (effects) considered are in the range of those estimated in Tables 7 and 8. These are about -0.3 for the index of schooling level.

The different effect sizes (here not standardized, but rather the magnitude of the coefficient) are on the x-axes, and the power is indicated on the y-axis. The power is calculated for the men-only sample, with about 2,100 observations. The power calculations for the other components (whether voted, and whether joined with others to discuss an issue) are also presented, as well as for whether member of a community group (available for rounds 4-7 only, and so with a smaller sample size).

As can be seen, the estimations appear to be well-powered on the whole, with power for coefficients of -0.3 being either 80% or higher for many specifications with some schooling or the education level as explanatory variable, but less often for the variable of whether completed primary schooling, as there is less variation in that indicator relative to the other two schooling indicators.

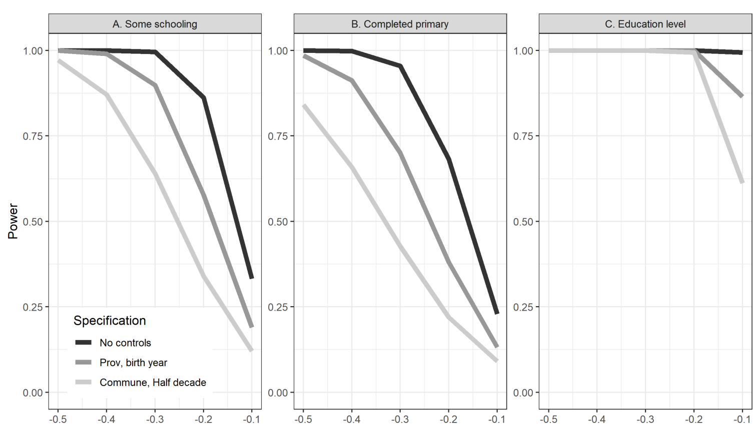

Figure F1: Power calculations for different education variables (endogenous) using schooling access as instrument and outcome variable *whether participated in a meeting*

Notes: Power calculated under various effect sizes using R package `ivmodel` developed by Kang et al. (2021), using Afrobarometer data, various rounds, rural residents born after 1950, men only, sample size about 2,100.

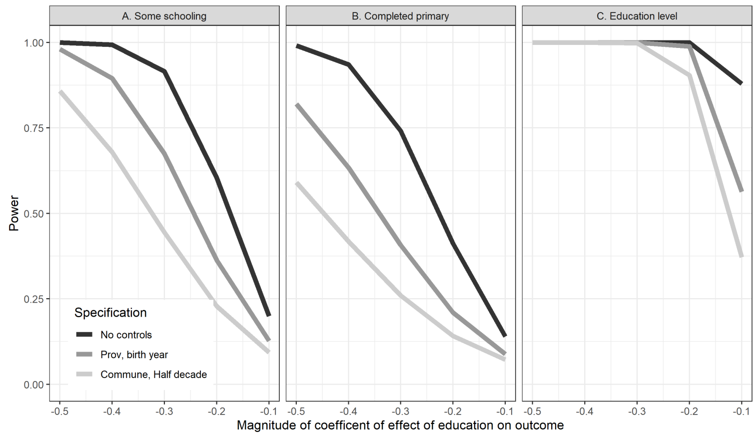

Figure F2: Power calculations for different education variables (endogenous) using schooling access as instrument and outcome variable *is a member of a community group*

Notes: Power calculated under various effect sizes using R package `ivmodel` developed by Kang et al. (2021), using Afrobarometer data, various rounds, rural residents born after 1950, men only, sample size about 1,650.

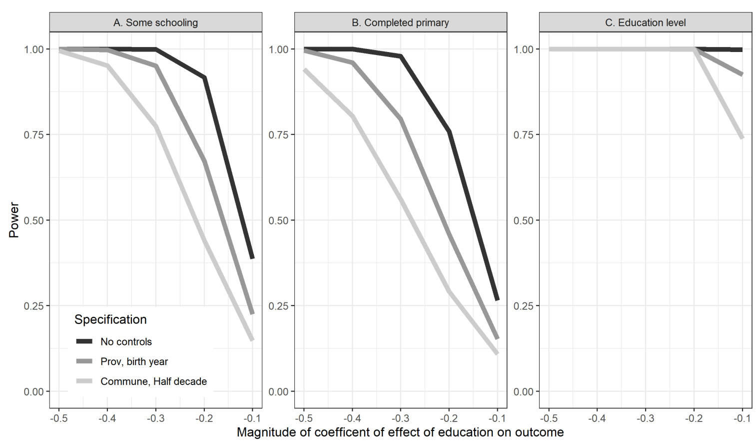

Figure F3: Power calculations for different education variables (endogenous) using schooling access as instrument and outcome variable *whether voted*

Notes: Power calculated under various effect sizes using R package `ivmodel` developed by Kang et al. (2021), using Afrobarometer data, various rounds, rural residents born after 1950, men only, sample size about 2,100.

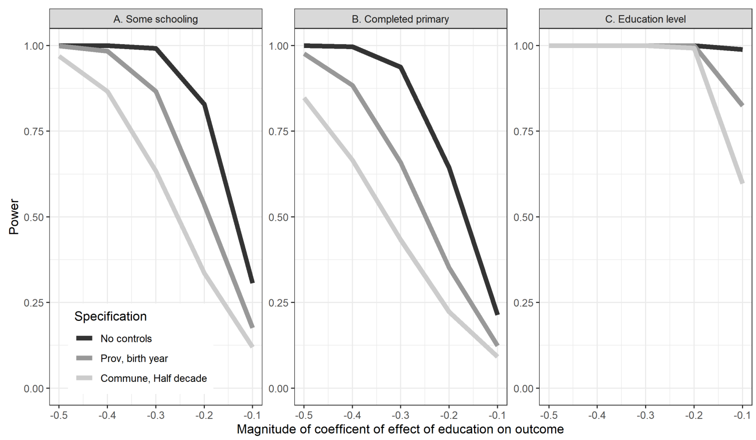

Figure F4: Power calculations for different education variables (endogenous) using schooling access as instrument and outcome variable *whether joined with others to discuss an issue*

Notes: Power calculated under various effect sizes using R package `ivmodel` developed by Kang et al. (2021), using Afrobarometer data, various rounds, rural residents born after 1950, men only, sample size about 2,100.