N m'a faamu[*]

Boosting Learning Through Bilingual Education: Evidence from Mali

Mathilde Col[†]

Abstract

Despite increased school enrollment across Sub-Saharan Africa, learning outcomes remain critically low. In this paper, I explore the replacement of colonial languages as the primary medium of instruction in Africa as an educational intervention designed to enhance student learning. This study evaluates the long-term impacts of Mali's 1999 national bilingual education reform, which introduced 11 local languages into the primary school curriculum as languages of instruction. Exploiting spatial variation in policy implementation and using a difference-in-differences approach with commune-level data, the study finds that exposure to bilingual education significantly improves local language literacy by 30% and French literacy by 10%, and increases school attendance by 5 percentage points. These gains are especially pronounced among women and individuals with educated mothers. Contrary to concerns over national cohesion, the policy does not exacerbate ethnic tensions. The findings contribute to the literature on large-scale bilingual education and the political economy of language use in Africa, suggesting that integrating local languages into education systems can enhance human capital without undermining nation-building.

JEL Codes: 055: I25: I28

Keywords: Bilingual education, Learning, Mali, Difference-in-Difference

[*] "I don't understand" in Bamanakan

[†] University of Bordeaux. https://mathildecol.github.io/

1 Introduction

Learning outcomes in sub-Saharan Africa (SSA) remain alarmingly low despite significant increases in school enrollment over recent decades (UNESCO, 2013). The 2018 PASEC assessment revealed that nearly 60% of students in grade 4 and grade 6 lacked fundamental mathematics and reading skills (CONFEMEN, 2020).[1] This disconnect between educational access and actual learning highlights the multiple barriers impeding human capital development in the region. While recent research has primarily addressed supply-side constraints, including inadequate educational resources and teacher capacity (Bold et al., 2017; Glewwe and Muralidharan, 2016), less attention has been devoted to the critical role of the language of instruction in the quality of education.

The widespread use of non-native languages for instruction represents a significant yet understudied barrier to effective learning across SSA. As of 2022, eighteen African countries continue to use colonial languages as their official instructional medium beginning in the earliest grades. The remaining countries have largely adopted bilingual education systems, where teachers initially deliver content in local languages while gradually transitioning to foreign languages, the colonial language, throughout the primary cycle. This progressive approach begins with instruction primarily in the vehicular languages before increasing the use of the foreign language until it becomes the exclusive medium of instruction by the end of primary school. A large body of literature in linguistics highlights the benefits of learning in the mother tongue (Benson (2002) and Cummins (2000); see Rolstad, Mahoney, and Glass (2005) and Sakaryalı, Bal, and Yıldırım (2024) for meta-analysis reviews). Indeed, when the linguistic distance between a child's mother tongue and the language used as the medium of instruction is significant, as is often the case in Africa, effective learning is hindered because a foundational understanding in a familiar language is essential for cognitive development (UNESCO, 2016). Only a few papers in empirical economics have investigated this issue and causally estimated the effect of the language of education on learning and economic returns to years of education. In a randomized controlled study in Cameroon, Laitin, Ramachandran, and Walter (2019) show that grades 1 and 3 students taught in the local language had more than one standard deviation better test scores in mathematics and English. The same conclusions were drawn from two field experiments in South Africa (Mohohlwane et al., 2023). However, no study focuses on a bilingual education program at scale. The only policy extensively studied is the introduction of Oromo in the primary education system in Ethiopia, instead of Amharic (Ramachandran, 2017; Seid, 2016). They find that mother tongue instruction increases school attendance, but they do not exploit a radical change from a colonial language to a local

[1] 14 countries were part of the last PASEC (*Programme d'Analyse des Systèmes Educatifs de la CONFEMEN*) evaluation, making it one of the largest international tests in SSA.

one.

This paper provides the first empirical evidence on the long-term effects of a national education policy that introduces local languages instead of a colonial language on human capital accumulation. I exploit local variation in the implementation of Mali's 1999 linguistic education reform, leveraging granular commune-level data to estimate the causal effect of bilingual instruction on human capital accumulation. Mali is one of the last SSA countries where primary school enrollment is still insufficient. Learning levels are also extremely low, with only 30% of the adult population literate in 2020. Until 1999, French was the only language of instruction in primary schools, despite a high linguistic diversity with 63 living languages, including 57 stable and institutionalized (Eberheard, Simons, and Fennig, 2025). In 1999, the Ministry of Education enforced a nationwide reform of the primary education system that introduced 11 national languages as medium of instruction.

Using school-level information on the implementation of the bilingual education (BE) curriculum from the 2011 national census, I estimate the BE supply at the commune level as the share of schools offering bilingual education (BE) over the total number of schools. Exploiting the local variation in this ratio at the commune level, I use a difference-in-differences strategy, comparing birth cohorts in high-intensity BE communes to those in low-intensity BE communes before and after the implementation of the reform. I rely on the 2018 LSMS survey to have an accurate measure of human capital accumulation.

I find that individuals with greater exposure to the linguistic reform demonstrate a 10% higher likelihood of French literacy and a 30% higher local language literacy in adulthood. School attendance and success in the primary school leaving exam also increase by 6 and 4 percentage points, respectively. I further show that changes in school supply do not drive these effects, and thus can be attributed to the BE aspects of the reform. I provide suggestive evidence that the positive effects are concentrated on women, as girls are less exposed to the official language before primary education (Benson, 2005; Hovens, 2002). Moreover, mothers' education plays a key role in explaining the effectiveness of bilingual education, highlighting an important channel related to parental investment. These findings remain robust across alternative difference-in-differences specifications, including those that relax the assumption of treatment homogeneity or account for variations in treatment timing. One of the main obstacles to introducing local languages in education is a lack of political will. Using Afrobarometer data, I show that bilingual education in Mali did not harm the nation-building process by exacerbating ethnic feelings.

This paper contributes to four strands of literature. First, it provides new empirical evidence on the learning benefits of bilingual education. While previous research demonstrates substantial returns to primary education in sub-Saharan Africa, particularly for women (Psacharopoulos and

Patrinos, 2004), little is known about the effectiveness of alternative teaching approaches, such as bilingual education. This study takes an important first step by examining adult learning and schooling outcomes, suggesting avenues for future research on long-term wage effects. Additionally, it moves beyond simply measuring years of education by incorporating a quality dimension through its analysis of literacy in multiple languages.

Second, I contribute to the recent and growing literature evaluating education policies at scale. Implementation is critical to understanding the differential impacts of education programs (Angrist and Meager, 2023). However, scaling up successful interventions remains challenging. Evidence on nationwide bilingual education policies shows mixed results due to implementation difficulties (Piper, Zuilkowski, and Ong'ele, 2016). This study adds to the literature by documenting evidence from a successfully scaled-up policy intervention.

Third, I provide additional evidence on an understudied beneficial educational intervention. A recent meta-analysis finds that mother tongue instruction shows promising results in controlled experimental settings (Evans and Mendez Acosta, 2021). Bilingual education should be considered alongside other educational resources (Glewwe and Muralidharan, 2016; Mbiti et al., 2019), as I show that effectiveness depends on maintaining appropriate student-teacher ratios. Further research is needed to identify other complementarities that may explain the success or failure of scaled-up bilingual education programs.

Finally, this paper contributes to the nation-building literature in Africa by examining a policy that promoted the use of local languages. Nation-building built on language unification in Europe in the 19th century (Blanc and Kubo, 2023; Clots-Figueras and Masella, 2013). So far, evidence in sub-Saharan Africa is limited to the study by Pengl, Roessler, and Rueda (2022), which shows that ethnic groups exposed to African language publications exhibit more salient ethnic behaviors. I expand on this work by looking at an education policy promoting African languages.

This paper is organized as follows. In Section 2, I document the context surrounding the 1999 reform. Section 3 details the data used for the empirical analysis. Section 4 presents the conceptual framework. I explain the main empirical strategy in Section 5 and show the results in Section 6, including results on political outcomes. I address the main concerns and perform additional robustness checks in Section 7.

2 Bilingual education and the 1999 reform in Mali

2.1 Historical background

There is a long-lasting history of bilingual education in the country that started right after the independence.[2] In 1960, experiments started using local languages to expand literacy for adults who did not attend school under colonial rule (UNESCO, 1963). Following positive results from these experiments, the use of local languages expanded to primary education at the onset of the 1980s (Ba, 2009; Diarra, 2020).

The first pilots of bilingual education combining French and another local language in primary schools started in the late 1970's (see Section A.1 for more details). They expanded gradually from Segou to other regions, and from using only Bamanankan to introducing other local languages in education. Although the impact on test scores was limited to null (Maurer, 2007; Skattum, 2010; Traoré, 2001), these experiments provided crucial insights for policymakers during the 1999 expansion of the bilingual education program.

2.2 The 1999 bilingual education reform

Expansion In 1999, the Ministry of Education passed a law to scale up the bilingual education experiment to all public primary schools as an essential component of the PRODEC (*Programme Décennal de Développement de l'Education*), a countrywide educational reform aiming at expanding primary schooling coverage (Loua, 2017). The main objective of this reform was to build schools massively. Consequently, school supply increased rapidly, from 2,600 schools in 1998 to almost 10,000 in 2008. The bilingual education reform was implemented at the school level at the start of the 2001 school year. In similar contexts, demand has been identified as one of the primary obstacles to the expansion of bilingual education (Piper, Zuilkowski, and Ong'ele, 2016; Ramachandran and Rauh, 2022). However, because of the information campaign that took place in 1999, community demand was high at the onset of the reform.[3] In 2002, there were 2,110 bilingual classes, and 666 schools with at least one bilingual track throughout the country. The bilingual curriculum was available in 11 additional languages (Bamanankan, Songhay, Tamasheq, Soninke, Dogon, Fulfude, Bomu, Syenara, Mamara, Bozo, and Khassonke) and counted 121,734 enrolled students. Following a rapid expansion in the early 2000s, the number of bilingual schools reached a peak of 2,530 in

[2] During the colonization period, like any other country under the French administration, the official language of education and administration was French, and the use of the local languages in the schooling environment was severely punished (De Gaston, 2011).

[3] To boost the demand for bilingual education, information sessions were organized at the district level, and short information messages were displayed through 34 local radios (MEN, 2003)

2005, accounting for approximately one-third of the total number of schools. Since then, a lack of funding and political will has led to stagnation in the number of schools following the bilingual education curriculum. In recent years, this number has even declined, and bilingual education is implemented in only 25% of the primary schools.

Curriculum content Bilingual education is a pedagogical approach that relies on knowledge of the familiar local language to ease the transition to the foreign language (Mohohlwane et al., 2023). In practice, students are expected to spend the first years of primary education learning the basics of writing and reading, as well as other subjects such as Mathematics, in the local vehicular language.[4] At the same time, teachers introduce French progressively to finish at the end of the primary cycle with French-only instruction. Table A.1 details the use of French and the local language for each grade, as specified in the official curriculum. During the first two years, the curriculum is primarily in the local language chosen. Over the next two years, French is gradually introduced, so that by the end of grade 4, half of the curriculum is taught in French and the other half in the local language. During the last two years preceding the secondary cycle, French becomes the dominant language. Officially, French remains the sole language of instruction in the second cycle of primary education, as well as in secondary and tertiary education.

Implementation Qualitative evidence points to a demand-led bottom-up implementation process. To open a bilingual education track in a school, local leaders must request approval from the local committee responsible for expanding bilingual education.[5] In collaboration with the community, this committee would select the primary language to be used in the bilingual class by teachers and students as the language of instruction. Schools would receive textbooks in the selected language, and teachers would undergo a short additional training course to adapt to this new curriculum (20 days in theory) (MEN, 2003). Hence, the provision of bilingual education does not follow a random assignment, and communities that requested the opening of a bilingual education track differ from the others on many observable and unobservable characteristics.

Official reports document the mixed quality of the policy expansion. In practice, the vast majority of teachers undergo additional training (Diarra, 2013).[6] However, little is known about the training quality, as the bilingual education expansion followed a very decentralized process

[4] Figure A.1 shows an example of a textbook for grade 1 students, fully written in Bamanankan.

[5] The opening of a bilingual education track in a school can arise from two different situations: either the school is newly built as part of the PRODEC program, or the school switches from the monolingual to the bilingual education curriculum.

[6] In 2000, 9 training centers were created. In 2001, 26 centers were operating with 3.775 teachers trained. In 2002, this number increased to 34 centers with 3,608 additional teachers trained (MEN, 2003).

(Ba, 2009). Reports also point to the long delay in textbook provision, which is attributed to the lengthy printing process (MEN, 2003).

3 Data

3.1 Roll-out of the bilingual education program

To document the progressive expansion of the bilingual education policy throughout the country after its initial implementation in 2001, I exploit a census conducted in 2011 by linguistics experts on behalf of the Ministry of Education (Diarra, 2013). This census covers all schools that were declared officially as "bilingual" at the 2011 school start and reports at the school level whether the bilingual curriculum is still used, or was abandoned for the French-only curriculum.[7] Out of the 3,784 bilingual schools that opened since 2001, 83% still had a bilingual education track in 2011.

Due to the civil war at the time of data collection, the census took place only in the Bamako, Kayes, Koulikoro, Segou, and Sikasso regions, covering only one-third of Mali's territory. Figure A.4 depicts this coverage visually. However, as 78% of the population lives in these regions, I argue that this census covers a large fraction of the school supply, reducing the external validity threat for this study (INSTAT-Mali, 2017).

Figure A.2 shows the evolution of the number of bilingual schools from 1994 to 2011.[8] Before the 2000s, the bilingual education supply was close to zero. Following the official introduction of bilingual education into the curriculum in 2000, the proportion of bilingual schools increased progressively to reach 25% of all primary schools.

Qualitative evidence suggests that community schools were more likely than other schools to adopt the bilingual education curriculum (Diarra, 2020). These specific public primary schools are opened as a result of a community initiative; their number rapidly grew during the 1990s as a response to the low school supply. Using the name of the school in the 2011 census, I find that only 5% of the bilingual schools are community schools.[9]

[7] If the school still has a bilingual class, the census reports whether it is used for all grades or only the first ones. Moreover, information about the local language used at the school level is not available in the census.

[8] Table A.3 details which geographical level, information available, and source used for bilingual education supply for every year presented in Figure A.2.

[9] I consider a school as a community school if it is labeled as such in its name: with the words "EC" or "communautaire". Hence, this number is a lower bound of the true fraction of community schools among the bilingual education supply.

3.2 Population census

School supply Using census data from 1998 and 2009, I construct a panel dataset on school supply at the commune level.[10][11] Figure A.2 provides an overview of its evolution over time, showing that the number of schools improved rapidly during the 2000s.[12] I use the total school supply information to create an indicator of exposure to bilingual education at the commune level: the ratio between the number of bilingual schools as officially listed in the bilingual education census and the total number of schools.

Commune characteristics The main level of the analysis of this study is at the commune level.[13] Using the 2009 census data, I derive some characteristics of communes that I expect to be essential when analyzing the efficiency of bilingual education. Specifically, I use the number of speakers for every language at the commune level to derive an ethnolinguistic diversity index. I also exploit the main language spoken in the community as a proxy for the language chosen to be the new language of instruction for the first grades.

3.3 Human capital

Measurement of literacy I am interested in the impact of bilingual education on human capital accumulation. I use literacy as a proxy for education quality. Literacy is a widespread but poorly measured indicator in many surveys, as it is often overestimated. For instance, in the Demographic and Health Surveys (DHS) conducted in developing countries, respondents were automatically considered literate after a certain number of years of education, usually after completing the primary cycle (Sandefur, 2017). However, in many sub-Saharan countries, the Learning-Adjusted Years of Schooling (LAYS) indicates that language proficiency requires more than just finishing primary education when accounting for the quality of schooling (Filmer et al., 2018). I use the 2018 Living Standards Measurement Study (LSMS), conducted by the World Bank in 2018, to

[10] I thank Flore Gubert for providing the panel dataset on the public infrastructures that were used in Chauvet et al. (2015).

[11] The administrative system in Mali is decomposed as follows: a region is composed of a set of districts, a district includes different communes, and a commune aggregates various villages.

[12] I address this potential confounder and further discuss it in 7.2.

[13] Information scarcity in the data does not allow me to perform the same analysis at the village level. Moreover, even if these data were available, the identification strategy would rely on a strong assumption: that people attend the school in their village and do not attend a neighboring school. Substantial evidence suggests that people are willing to attend other schools to match their language preferences (Ramachandran and Rauh, 2022).

address this issue. Indeed, the LSMS survey provides a measure of literacy based on writing and reading skills in French, the local vehicular language, and another language, for all respondents.

Other outcomes Besides literacy, I also consider other educational outcomes: school attendance, and whether the respondent has the primary school diploma.[14] I do not consider the number of school years as the linguistics literature documents two opposite mechanisms that could affect the time spent at school (Benson, 2002): bilingual education affects positively the retention rate of students, but it is also expected to decrease the repetition rate. Moreover, as the youngest birth cohort in my sample turned 15 in 2018, some may not have completed their education at the time of the survey. Due to data limitations, the completion of primary education is also not considered in this analysis[15]; instead, obtaining the end-of-primary-school diploma is used as a proxy for completion of the first cycle.

4 Theory of change

4.1 Primary hypotheses (PH)

PH1: Bilingual education increases learning I expect to observe positive long-term effects of bilingual education on the accumulation of human capital, partly through better learning. The literature has shown in highly controlled and limited settings that learning in a familiar language increases literacy in both the local language and the foreign one (Benson, 2002; Cummins, 2000; Mohohlwane et al., 2023). Learning begins from the first moment of primary education, rather than being delayed until students develop proficiency in the language of instruction. Furthermore, past evidence points that bilingual education connects better writing and speech (Hovens, 2002).[16]

[14] In Mali, primary education starts at the age of 7 officially, for 6 years for the first primary cycle and 3 years for the second primary cycle. At the end of this first cycle, until 2010, students passed an exam called the *Certificat d'études primaires* (CEP) to pass to the second primary cycle.

[15] In the raw LSMS data, for this survey round in this country only, the variable about education attainment is incomplete.

[16] Another advantage is linked to a particular linguistic feature of Sub-Saharan Africa: the spread of "mixed" languages, also known as *creoles* or *pidgins*, that take words or structures from different local and foreign languages (Calvet, 2010). Learning a second language through a familiar language reduces the risk of code-switching (i.e., using words from another language in the middle of a sentence), allowing the skill to transfer quickly from one language to the other (Cummins, 2000).

PH2: Bilingual education increases schooling Bilingual education is expected to increase school attendance and, to some extent, the number of schooling years. Indeed, bilingual schools can attract students compared to monolingual ones when the language used at school is close to the home language (Ball et al., 2024). It also reduces the repetition rate and drop-out, leading to higher promotion rates (Patrinos and Velez, 2009). However, it could also increase reluctance to attend school among minority language groups, as their mother tongue is unlikely to be used in the bilingual education class.

4.2 Secondary hypotheses (SH)

SH1: Women benefit more from bilingual education Qualitative evidence shows that girls benefit more from bilingual education than boys (Benson, 2002; Hovens, 2002). The primary mechanism behind these positive results is that girls are less exposed to the foreign language than boys in the home environment (Q'Gara and Kendall, 1996). Hence, by bridging the gap between the school and home environments, bilingual education yields higher human capital accumulation for girls (Benson, 2005).

SH2: Areas more linguistically diverse are less prone to positive returns to bilingual education Which language to choose is a key implementation challenge in the context of a nation-wide bilingual education policy (Piper, Zuilkowski, and Ong'ele, 2016). In the case of the 1999 Malian reform, the vehicular language that replaces French in schools is chosen among the 11 national languages, all of which are documented. Figure A.3 gives a broad picture of the distance existing between these languages: 9 languages out of 11 are from the same language family. However, even if the linguistic distance between the official language spoken in the school and the mother tongue may be reduced (Laitin and Ramachandran, 2022), community members speaking a minority language that is not chosen as the new language of instruction would still face the same understanding challenge. Hence, I expect that in a very linguistically diverse area, using only one local language at school might exclude more students than in a linguistically homogeneous area, reducing the aggregated benefits of bilingual education.

SH3: Returns to bilingual education are higher in rural areas Related to the previous hypothesis, I expect to see higher benefits of bilingual education in rural areas, as these are less linguistically diverse than cities. Urban areas concentrate different ethnolinguistic groups speaking different languages, where French can act as a *lingua franca* (Calvet, 2010).

5 Empirical strategy

5.1 Naive approach

A simple approach to the research question consists of looking at the increase in literacy (in both the local language and French) and schooling that occurs when one more bilingual school is opened in the commune. To do so, I estimate the following regression:

$$Y_{i,y,c,d} = \beta log(\text{Nb. bilingual schools}_c) + \alpha log(\text{Nb. schools}_c) + \theta_d + u_{i,y,c,d} \tag{1}$$

$Y_{i,y,c,d}$ is the outcome for individual i, born in year y, living in the commune c located in the district d. θ_d captures district fixed effects, and standard errors are clustered at the LSMS cluster level. The number of bilingual schools in the commune is reported in the 2011 census conducted by the Ministry of Education, and the number of schools is derived from the 2009 national household census. The coefficient β captures the effect of one additional bilingual school, while keeping the number of schools fixed.

Results are presented in Table B.1. One more bilingual school in the commune is associated with a 14 percentage point (pp) increase in local language proficiency, and a 5pp increase when the language is French. For schooling, the effect is approximately 10pp. Reassuringly, the number of schools seems to matter for all outcomes, except for literacy in the local language, as the majority of these schools don't use it for teaching.

However, this analysis is purely exploratory, and results are subject to omitted variable bias and selection, as the location of (bilingual) schools is not random. I detail the identification strategy designed to overcome this issue in the next section.

5.2 Identification strategy

Difference-in-difference I leverage the introduction of the reform at the 2000 school start and the heterogeneity in the coverage of bilingual education throughout the country to assess its impact on long-term educational outcomes. Using the year of birth of adult individuals assessed in the LSMS, I can infer when they started school. In this aspect, if they first attended school before the 2000 school year, it is unlikely that they were exposed to bilingual education (Figure A.2). Given that the average age at school entry in the LSMS sample is 6.6, I consider that individuals born in 1993 and before likely attended monolingual schools (i.e., schools with French-only education). To avoid capturing the effects of other reforms, I restrict the sample to 10 years around the reform implementation, specifically individuals born from 1983 to 2003, aged 15 to 35 in 2018. Then, for each LSMS cluster, I map the supply of bilingual education using the 2011 census data on bilingual schools and 2009 census on school supply, and obtain a panel data set at the enumeration area level with the year of birth as the key time dimension.

Treatment definition I define exposure to bilingual education (BE) at the commune c level as follows:

$$\text{Share of BE}_c = \frac{\text{Nb. bilingual schools}_c}{\text{Nb. schools}_c}$$

I include in the numerator all schools officially declared as bilingual education schools in the 2011 census, and in the denominator the school supply as given by the 2009 census.[17] By definition, the exposure measure is bounded between 0 and 1.[18]

Figure A.5 shows the share of bilingual schools among the total school supply at the commune level, along with the localization of the LSMS cluster. Bilingual education provision is not random and follows a clear geographical pattern, with some districts and regions concentrating most of the bilingual education supply. Looking at the linguistic areas depicted in Figure B.10, I observe that the supply of bilingual education is high in locations where Bamanankan is the main language. As a result, I cannot use the national median exposure to bilingual education, i.e., the fraction of bilingual schools in the total school supply, as a threshold to categorize communes into low vs. high intensity, mimicking the design used in Duflo (2001). Instead, I use the median ratio of bilingual schools over the total number of schools within each district as a threshold and consider a commune with a ratio below this median to be a low-exposure area. Similarly, I consider a commune with a fraction of bilingual education higher than the district median as highly exposed to bilingual education. Figure 1 displays the geographical coverage of the binary treatment.

Sample selection One concern about the sample considered is that migration already took place at the time of the survey. In 2011, migration, both internal and external, was estimated to affect almost 20% of the population (Sougane, 2014). It may impact my empirical specification in two ways. First, I do not observe in the LSMS sample migrants who left the locality to live abroad. However, if anything, this should yield a downward bias in the results: if bilingual education enhances learning skills, particularly in French, then individuals who benefited significantly from the bilingual curriculum leveraged these new skills to migrate before the LSMS survey took place. Hence, the estimated effect would be a lower bound of the true impact of bilingual education.

[17] I further consider the fraction of schools that dropped out of the bilingual curriculum in the section 6.4.

[18] Because the school supply is measured through the population census that took place 2 years prior to the 2011 bilingual education census, some ratios exceeded 1 for a few communes. For these communes, I bounded the ratio to 1.

Second, I may wrongly attribute a high exposure to bilingual education to an individual who migrated after their childhood. The level at which the analysis is done reduced this threat: 98.8% of individuals in the sample are reported being born in the same commune.

5.3 Empirical specification

Two Way Fixed-Effects (TWFE) By leveraging the reform implementation timing and the differential exposure to bilingual education, I estimate the following regression:

$$Y_{i,y,c,d} = \alpha_1 \mathbb{1}[\text{y} \geq 1994]_{i,c,d} + \alpha_2 \mathbb{1}[\text{c} \in \text{BS}]_{i,c} + \beta \mathbb{1}[\text{y} \geq 1994]_{i,c,d} \times \mathbb{1}[\text{c} \in \text{BS}]_{i,c} + \theta_d + u_{i,y,c,d} \quad (2)$$

$Y_{i,y,c,d}$ is the outcome for individual i, born in year y, living in the commune c located in the district d. BS represents the set of communes considered as highly exposed to bilingual education relative to the median district ratio of bilingual schools. α_1 captures the effects of being born after 1994, i.e., starting school after the implementation of the bilingual education reform. α_2 isolates the differences in the outcome that might exist between communes with a high share of bilingual schools compared to communes with a low share. The coefficient of interest is β, which captures the effect of being highly exposed to bilingual education (BE) compared to low exposure after the reform. I use district-fixed effects θ_d to control for district unobservable characteristics and differences in bilingual education supply (Figure A.5). I use the LSMS cluster provided in the survey to cluster standard errors, and I use household weights also present in the survey.[19]

Testing key assumptions Following the difference-in-difference literature, I test for the parallel trends (PT) and the no anticipation (NA) assumptions using an event study. To do so, I estimate the following regression:

$$Y_{i,y,c,d} = \sum_{t=1983}^{2003} \alpha_{1,t} \mathbb{1}[t = y_{i,c,d}] + \alpha_2 \mathbb{1}[\text{c} \in \text{BS}]_{i,c} + \sum_{t=1983}^{2003} \beta_t \mathbb{1}[t = y_{i,c,d}] \times \mathbb{1}[\text{c} \in \text{BS}]_{i,c} + \theta_d + u_{i,y,c,d} \quad (3)$$

The same definitions as in Equation (2) apply. t represents the year of birth and covers the period considered in this analysis, i.e. from 1983 to 2003. I allow the coefficient β_t to vary in the pre and post-treatment period according to the year of birth y. I also implement multiple robustness tests in Section 7 to reinforce the credibility aspect of the PT and NA assumptions.

[19] Given that a commune is roughly equal to one LSMS cluster (except for Bamako), and that Bamako represents 60 clusters out of 325, I prefer using the LSMS cluster level rather than the commune level for clustering.

An additional key assumption made in the specification is the homogeneity of the treatment effect. This assumption is likely not to hold, as the treatment at the commune level is relative to a median value at the district level. In other words, I assign commune A in district 1, where 70% of schools are bilingual, to the low-exposure bilingual education group. In contrast, commune B in district 2, where only 20% of schools officially use local languages for instruction, can be assigned to the high-exposure bilingual education group. This apparent contradiction arises because the median ratio of bilingual education supply is 75% in District 1 but only 15% in District 2. I employ different strategies to show that this is unlikely to bias my estimate. First, I control for this by adding a fixed effect at the district level in Equation 2. Second, I release the assumption of treatment effect homogeneity in Section 7.1 by employing different approaches.

5.4 Descriptive statistics

Table 1 describes the main outcomes for the sample, at the individual and the commune level (Columns 1 to 3). The sample includes 8,636 individuals, with half women and half urban. Less than 50% of the sample attended school, and the same fraction is literate in French (either writing or reading). Finally, only one individual out of 10 can write or read in the local language. Columns 4 to 6 provide additional insights into the significant differences between the sample of communes less exposed to bilingual education and those highly exposed to it before the reform implementation. The main difference arises from local language proficiency: individuals in locations with more bilingual schools after 2001 are 3pp more literate in the local language. This "balancing" further highlights the non-random assignment of bilingual education throughout the country and the need for a quasi-experimental design.

6 Results

6.1 Human capital accumulation (PH)

Learning Figure 2 plots estimated coefficients from Equation 3 explaining BE effects on learning outcomes in the local language. It seems that introducing officially local languages in education has a significant, positive, and persistent impact on writing and reading in the vehicular language. Table 2 shows the TWFE results of estimating Equation 2: adults who had access to more bilingual schools when they were children are 4pp more likely to be literate (both writing and reading) in their local language (Columns 1 and 2). Given that only 13% of the sample is literate in the local language, being more exposed to bilingual education increases learning by approximately 30%. These initial sets of results can be viewed as a first stage: even though it is likely that teachers had already used local languages before their introduction into the curricula in 1999, their official enactment increased their use in the schooling environment, thereby boosting literacy in the local language.

Figure B.1 shows that these skills in the local languages partially transferred to French for the exposed students. This linguistic mechanism has been documented by Cummins (2000). Columns (3) and (4) of Table 2 provide additional evidence on it: on average, the literacy skills increased by 5pp (i.e., a 10% increase of the literacy rate), both in writing and reading, with the introduction of the bilingual education reform. These results are less significant and smaller in magnitude than those for the local language. A potential explanation is that some skills may not have been fully transferred from one language to the other because the children attended school for only 4 years.

In addition to literacy in the two languages, LSMS data provide the most appropriate variable for a placebo test: literacy skills in a language that is neither French nor the main language of the community. Indeed, introducing bilingual education should not affect this outcome at all, as this language is likely not to be used at all in bilingual schools in this area. Figure B.3 shows the event-study estimate and provides additional evidence supporting the identification strategy, as I do not observe any change in writing and reading skills after introducing bilingual education.

Schooling Introducing familiar languages to students makes school more attractive: as shown in Column (1) of Table 3, school attendance increased by 12% for cohorts that got access to more bilingual schools (+ 6pp). They are also more likely to hold a primary education diploma, suggesting that the duration of schooling is also affected. This result aligns with the linguistics literature, which shows that reducing the linguistic distance between the home and school environments incentivizes more parents to send their children to school (Benson, 2004).

The pre-trend coefficients for all the outcomes are small and jointly non-different from zero (see Figures 2, B.1, and B.2). All these pieces of evidence prove that the assumption of parallel trends between communes that were more and less exposed to bilingual education holds.

6.2 Gender (SH)

Previous findings document that girls benefit more from bilingual education than boys, because they are less exposed to the colonial language in their early childhood (Benson, 2005). Figure 3 shows heterogeneity results by sex and provides evidence supporting this qualitative result.[20] Women who had more access to bilingual schools as children are 8pp more likely to be literate

[20] Parallel trends at the subgroups levels are shown for literacy outcomes in Figure B.4. No anticipation effect is observed for both men and women. Event studies for other education outcomes give similar conclusions.

in French and attend school.[21] They are also 6pp more likely to obtain the end-of-primary-school diploma. Relative to the sample mean, it represents respectively an increase of 15% in the probability of being literate in French and attending school, and 20% in obtaining the CEP. In comparison, boys do not get any returns from introducing education in a familiar language. Figure B.5 displays graphically how girls catch up to boys thanks to bilingual education. While both girls and boys display similar levels of local language literacy, boys are still better off learning French, even after these local languages are introduced in primary schools.

Another difference related to gender can be observed at the parental level: whether the mother is educated or not matters a lot for the efficiency of the new bilingual curriculum (Figure B.6). The mechanism at play relates to the fact that mothers have been shown to invest more in their daughters' human capital compared to fathers, according to a recent experiment done in Uganda by Dizon-Ross and Jayachandran (2023). Consistent with their findings, I find that father education does not play a role in explaining differentials impacts, suggesting that educated fathers are not key actors in the bilingual education curriculum implementation (Figure B.6). With a bilingual curriculum, mothers are more invested, as they are now more familiar with the language spoken at school and in the textbooks. This is a novel result, as the interaction between maternal investment and bilingual education has not been previously documented in the literature.

6.3 Language and urbanization (SH)

Linguistic diversity Building on the theory of change outlined in Section 4.2, I hypothesize that the impact of bilingual education scales with the proportion of students whose mother tongue aligns with the chosen language of instruction. To test this, I stratify the sample by quartiles of linguistic diversity at the commune level, measured using a Herfindahl-Hirschman index derived from the 2009 census data on language use.[22] Contrary to expectations, empirical results in Figure B.7 indicate that high linguistic diversity does not impede the effectiveness of bilingual education.

The selection of the instructional language may still play a critical role, independent of overall linguistic diversity. As illustrated in Figure B.10, Bamanankan—spoken as a first language by 4 million Malians and as a second language by an additional 10 million—serves as a *lingua franca* across linguistic communities. However, regression results in Table B.2 (Panel 1) reveal that the positive effects of bilingual education are not driven by Bamanankan-dominant areas. Thus, neither linguistic diversity nor the dominance of a particular language emerges as a decisive factor in the success of bilingual education programs. Urbanization.

[21] The effects for women are significantly different at the 5% level of the impacts for men.

[22] Linguistic diversity is calculated as $1-\sum s^2$, where s represents the share of speakers for each language reported in the census.

Urbanization I further explore whether rural households derive greater benefits from local-language instruction. Panel (2) in Table B.2 presents mixed findings: local language proficiency improves only in urban areas, with no discernible effect on French literacy. To address potential endogeneity in urban classification — since urban status in the 2018 LSMS survey may reflect changes induced by the reform — I refine the analysis by using historical road access as a proxy for urbanization. Drawing on road network data from Müller-Crepon (2023) and Müller-Crepon, Hunziker, and Cederman (2021), I measure each cluster's distance to the nearest road as of 2000, applying a 10-kilometer threshold (see Figure B.11).[23] Heterogeneity analysis, reported in Table B.2 (Panel 3), shows that the positive impacts of local-language instruction on both learning and schooling outcomes are not concentrated in areas further from roads. Moreover, local language proficiency seems to improve in locations with better road access. These results challenge the assumptions on linguistic diversity in Section 6.2 and highlight potential unequal returns to bilingual education, particularly for rural households.

6.4 Mechanisms

Continuity in the bilingual education curriculum Leveraging data on bilingual status from the 2011 school census, I identify schools that maintained local language instruction to estimate the differential impacts of the reform along this dimension. In communes where schools rapidly abandoned the bilingual curriculum after initial implementation, the expected effects would be minimal or absent. This outcome is predictable because if schools quickly reverted from bilingual to French-only instruction, no birth cohort would have experienced full exposure to bilingual education throughout all primary grades as originally intended.

Table B.2 (Panel 4) presents a comparative analysis of the effects of bilingual education between communes where at least one school abandoned the bilingual program and those where schools maintained the implementation of the language reform.[24] Of the 167 communes analyzed, 100 experienced at least partial reversion from bilingual to monolingual (French-only) instruction in their schools. Positive learning and schooling outcomes are concentrated in areas with a continuous implementation of the linguistic education reform. This finding aligns with the literature showing that longer exposure to the familiar tongue is key for bilingual education to be effective (Ball et al., 2024).

[23] Urban status and proximity to roads in 2000 are positively and significantly correlated (coefficient: 0.47, SE: 0.05). Importantly, the treatment variable for bilingual education is uncorrelated with road access at baseline.

[24] The median switching rate at the potential cutoff point is zero, making the intensive and extensive margin analyses equivalent in this context.

Number of students per school Using 2011 administrative data on student enrollment and teacher staffing per district, I examine how classroom conditions influence reform outcomes. Table B.8 shows that the benefits of bilingual education are concentrated in communes with lower student-teacher ratios. This result suggests that above a certain class size, the choice of the language might not matter as learning is already impeded by too many students with heterogeneous levels in the same class (Gazeaud and Ricard, 2024). As a consequence, the effectiveness of bilingual instruction may depend significantly on classroom resources and teacher capacity.

6.5 Political costs

Language unification has been shown to be a key component in nation-building (Blanc and Kubo, 2023; Clots-Figueras and Masella, 2013). Consequently, one major obstacle to introducing local languages in education is the lack of political will among policymakers. The fear is that promoting local languages in schools would increase ethnic salience and threaten national identity.

Using the last two survey rounds from the Afrobarometer survey (2018 and 2020), I provide the first estimate of this political cost on nation-building. To do so, I retrieve for every enumeration area provided in the dataset its commune, and assign a treatment status according to the share of bilingual education in this area. A few questions in the Afrobarometer relate to ethnic identity, nation-building, and language. In particular, I look at two dimensions: salience of ethnicity and cultural practices. For the first outcome, I use two questions: (1) salience of ethnicity over national identity, and (2) preference for a neighbor from another ethnic group.[25] For the second outcome, I look at whether the respondent feels comfortable (1) wearing traditional clothes, and (2) speaking her mother tongue.

I estimate the same equation 2 and find evidence supporting that bilingual education does not harm nation-building. Regression results can be found in 4. Figure B.9 displays the same results, relative to the outcome mean. Due to age and location restrictions, the sample size is significantly reduced, increasing the magnitude of any detectable effect.[26] However, the global picture is clear: introducing local languages in education in Mali did not give rise to a strong political ethnic sentiment. It even increased by 10pp the probability that the respondent expressed a preference for national over ethnic identity. Moreover, the nation-building process did not come

[25] For interpretation purposes, I dichotomize these two outcomes. Details about the dichotomization are given in the notes of Figure B.9 or Table 4.

[26] The estimations are run on a sample of 850 individuals for Columns (1) and (2), and 466 for Columns (3) and (4) of Table 4. Indeed, as for the LSMS sample, I restricted the Afrobarometer sample to 10 years around the reform. Additionally, due to the availability of data on bilingual schools, I have limited this analysis to Southern Mali.

at a cost of invisibilizing the different ethnic groups, as seen in the results on cultural practices.

The closest study to this topic finds that bilingual education (Catalan and Spanish) in Catalonia increased the Catalan feeling (Clots-Figueras and Masella, 2013). Hence, these preliminary findings shed light on other mechanisms that could be at play in SSA, limiting the diffusion of ethnic sentiment.

7 Robustness checks

7.1 Relaxing the homogeneous treatment effect assumption

In recent years, the difference-in-differences estimation strategy has garnered significant attention, leading to substantial methodological advancements through the development of various estimators (see Callaway and Sant'Anna (2021) and de Chaisemartin and D'Haultfœuille (2022) for a non-exhaustive survey of the recent literature). In particular, these papers have shown that an important underlying assumption regarding the validity of the TWFE estimator is treatment effect homogeneity. Given the nature of the treatment under consideration, this assumption is unlikely to hold. One of the main recent updates in the difference-in-difference literature relates to continuous treatment definition, instead of relying on a binary measure (Callaway, Goodman-Bacon, and Sant'Anna, 2024). Because of the spatial correlation of the share of bilingual education, using this estimator is not recommended.[27] Indeed, comparing different levels of bilingual education provision would capture just differences between geographical areas, as shown in Figure C.5.

Another way to release this assumption is to look at the effect heterogeneity according to the median share of bilingual schools at the district level, which serves as a reference point for creating the binary treatment detailed in Section 5. I divide my sample into quartiles and show the results in Figure C.1 for all outcomes. The first result is that for all quintiles of median BE provision, the treatment effect is either null or positive. Second, in communes where at least 75% of schools are bilingual (the highest quartile), the estimated impact on all outcomes is zero. This is likely because the widespread provision of bilingual education in these areas does not allow for a clear threshold for comparing low- and high-exposure communes, as students are faced with an education market overwhelmingly dominated by bilingual education.

Finally, I consider the continuum of bilingual education supply by looking at how far the commune is from the district median, i.e., the cutoff point. Indeed, comparing communes that

[27] The same conclusion applies when considering using bins of treatment. This econometric strategy relies on the assumption that within bins (i.e., categories of bilingual education supply in which every commune receives a close "dose" of treatment), the treatment effect is homogeneous.

are well below the median and those that are well above the median should provide more striking results than comparing communes that are close to the cutoff point. I test this hypothesis in Table C.2 by splitting my sample into two.[28] As expected, I find stronger results for the subsample of communes far from the cutoff point. For communes whose bilingual education supply is close to the district median, I find no significant effects and smaller coefficients. This suggests that provision of bilingual education in these areas is less significantly different between the treatment and control group (i.e., respectively, communes above and below the district median of bilingual education supply).

7.2 School construction

The main objective of the PRODEC reform was to increase school supply by massively building schools. Figure C.6 maps the growth rate of the school supply at the commune level: the building effort was widespread across the Southern regions of Mali, and did not target one specific area.

Expansion of school supply can be a confounder of the impact of bilingual education, as the 1999 reform included both components. I test this hypothesis by first looking at the correlation between bilingual education supply and the growth rate of school supply in Figure C.2, exploiting the number of schools in 1998 and 2009 as reported in the census. I do not see any evidence that communes where more schools were built during this period were also communes with more bilingual schools. Second, I estimate Equation 2 with the number of schools as the outcome (Table C.4). I find that in communes with more bilingual schools, fewer schools were built from 1998 to 2009. Hence, I interpret the previous as a lower bound of the real impact of bilingual education, given that these areas were relatively less equipped with school infrastructures over time. This evidence suggests that the two components of Mali's 1999 education reform—bilingual curriculum and school construction—were not implemented in the same areas, but rather seen as complementary across different regions of the country.

Finally, I examine the differential effects of bilingual education intensity across communes with varying exposure to school construction programs. I find that introducing a bilingual curriculum was effective where fewer schools were built during the first decade of the 2000s (Table C.3, Panel 1). This evidence suggests that the benefits of bilingual education do not stem from increased school infrastructure, but rather from the program itself.

[28] I use the median distance to divide the sample of communes into two subsamples, which is 9pp above and below the district median. It corresponds roughly to the 25th and 75th percentiles of the distribution of bilingual education supply at the commune level.

7.3 Alternative treatment definitions

The treatment definition described in Section 5 is relative, given that it depends on the district median cutoff. This empirical strategy is chosen to address the unequal distribution of bilingual education across districts. I propose an alternative treatment definition based on a simpler, absolute cutoff: whether the commune has more bilingual than monolingual schools, i.e., whether the share of bilingual schools exceeds 50%. As most regions in the West and East parts of Southern Mali are well below this threshold, I focus only on districts in the center of the region. Figure C.3 presents the map of the selected areas, for which the dominant local language is Bamanankan. I estimate the same equation as in Equation 2, except I use a 50% cutoff to classify a commune as highly exposed to bilingual education. Results are presented in Table C.6. The main difference with results presented in Table 2 and Table 3 emerges when looking at the impact on local language literacy. In this setting, the effect of bilingual education is null and precisely estimated. This finding suggests that proficiency in Bamanankan might not be affected, as some individuals already speak and write this *lingua franca*.

District-level data on bilingual education expansion enables a staggered difference-in-difference analysis using the Borusyak, Jaravel, and Spiess (2024) imputation estimator (see Table A.3) .[29] I define treatment year as when a district exceeds its median share of bilingual education over the years, creating time variation from 1994 to 2005. Figure C.4 shows writing skills outcomes at the district level (n=28), which yields less precise estimates than the commune-level analysis (n=167). Despite this limitation, results confirm two key findings: parallel trends hold in the pre-treatment period, and bilingual education positively impacts writing skills in both French and local languages (significant at 10%), with stronger effects in local languages.[30] As expected, effect sizes are smaller than in the TWFE commune-level analysis due to reduced precision in identifying affected populations.

Then, I test an alternative treatment definition: instead of using the median ratio of bilingual schools at the district level, I take the median at the regional level. Similarly, I use region-fixed effects instead of district-level fixed effects. I see in Table C.1 that the coefficients do not vary much from the results in Table 2 and Table 3. The standard errors are predictably larger due to reduced precision in the estimation strategy. The only discrepancy from the previous rein the local language, suggesting that individuals targeted by the reform might already have been proficient in

[29] Following the procedure suggested in Roth et al. (2023), I prefer the imputation estimator over the Callaway and Sant'Anna (2021) as event studies displayed in Figures B.1, 2 and B.2 show that we should not worry about violation of the parallel trends assumption, and I do not expect important serial correlation between birth cohorts.

[30] Results hold for the writing skills in the local language when I take a 95% confidence interval.

it before the reform was implemented.

7.4 Other robustness tests

Additional controls I test whether results are robust to the inclusion of additional controls in Table C.5. In the first two panels, I add respectively the ratio of the number of children theoretically at school if school attendance were complete to the number of schools, and the number of schools in the commune.[31] In the first panel, I add fixed effects related to the linguistic areas, as reported in the census. I find that the results are robust to the inclusion of these additional variables.

Migration I further investigate whether bilingual education had differential effects based on individuals' migration history. Table C.3 (Panel 2) shows that only non-migrants experienced significant benefits from bilingual education. Those who had previously migrated showed no measurable improvements in literacy outcomes following the reform. This finding suggests that educational continuity and consistent exposure to the bilingual program may be critical factors in its effectiveness, with disruptions caused by migration potentially limiting the reform's impact.

Heterogeneity on school attendance I expect the positive effects to be concentrated among individuals who attended school. The results of this test are presented in Panel (3) of Table C.3. The interaction coefficients capture the additional treatment effects for enrolled students, indicating that the positive impacts are primarily driven by this subgroup. However, since school attendance is not orthogonal to the provision of bilingual education, these results should be interpreted with caution.

Restrict the sample to villages far from the communal borders Nearly half of all communes are located within 2km of neighboring commune boundaries.[32] Residents in these border areas could potentially enroll their children in schools across commune lines. When analyzing only communes situated at least 2km from borders, as shown in Table C.7, the coefficients are not significant anymore because of a loss of statistical power, but extremely close to what I find in Tables 2 and 3, strengthening the original findings.

[31] The number of children theoretically at school if school attendance were complete is computed as the number of individuals aged between 7 and 12 in the 2009 census.

[32] I take the Euclidean distance between the LSMS cluster and the border of the closest commune as my main measure of distance. I do not consider the random displacement of GPS points in my analysis as a limitation, as the random offset procedure ensures that the enumeration area stays within the lowest administrative unit (Michler et al., 2022).

Potential confounders I also examine potential confounders that could affect the treatment variables. Specifically, I assess whether bilingual education influences (i) migration patterns or (ii) school entry age. If bilingual education fosters migration, my sample would be highly selective. Likewise, if it encourages earlier school enrollment, comparing individuals by birth year would become problematic. Table C.8 provide suggestive evidence that is unlikely to be the case. Indeed, there is no correlation between treatment assignment and either (i) birthplace in the commune or (ii) age at school entry among those who attended school.

8 Limitations and future avenues

One important limitation of this paper concerns the definition of the treatment, which is relative to each district. In other words, what constitutes "treated" and "untreated" may vary substantially depending on the district context. I have begun to address this concern in the subsection 7.3, particularly in Table C.6, where I restrict the analysis to districts in the central part of Mali with significant bilingual education provision. This restriction makes it possible to have an absolute definition of treatment, ensuring that the treated and control units are defined consistently across districts. Nevertheless, I ultimately rely on the relative treatment definition because it allows me to preserve a larger and more diverse sample, thereby extending the scope of inference to other ethnic groups and to districts where the language of instruction differs from Bambara. This broader perspective is particularly valuable for understanding the policy's potential reach across different linguistic and sociocultural settings, even if it comes at some cost to internal validity. A natural next step would be to systematically test the robustness of the results to alternative treatment definitions. Specifically, the analysis could be reproduced using only the subset of districts already considered in Table C.6, where the treatment can be defined in absolute rather than relative terms. In addition, implementing a continuous measure of treatment intensity in these districts would help refine the definition of exposure, moving beyond a binary treatment distinction (Chaisemartin, D'Haultfœuille, and Vazquez-Bare, 2024).

A second limitation relates to the potential for selection bias in the establishment of bilingual schools, which I use *ex-post* to define exposure to bilingual education. Unfortunately, data on this selection process are limited. Future research should therefore focus on identifying the determinants of selection into treatment. Two potential avenues are particularly promising. First, commune-level electoral outcomes in 1997–1998, as documented by Chauvet and Mercier (2014), could serve as a proxy for political proximity to district authorities, helping to capture whether politically aligned communes were more likely to obtain bilingual schools. Second, data on local fiscal capacity from Sanoh (2015) could provide a measure of administrative and financial capability, reflecting a commune's ability to implement new education programs. Both political connections and local state capacity are plausible factors influencing the allocation of bilingual schools. Incorporating these variables into the analysis, as instruments or controls, would therefore represent an important step toward addressing potential selection in treatment assignment.

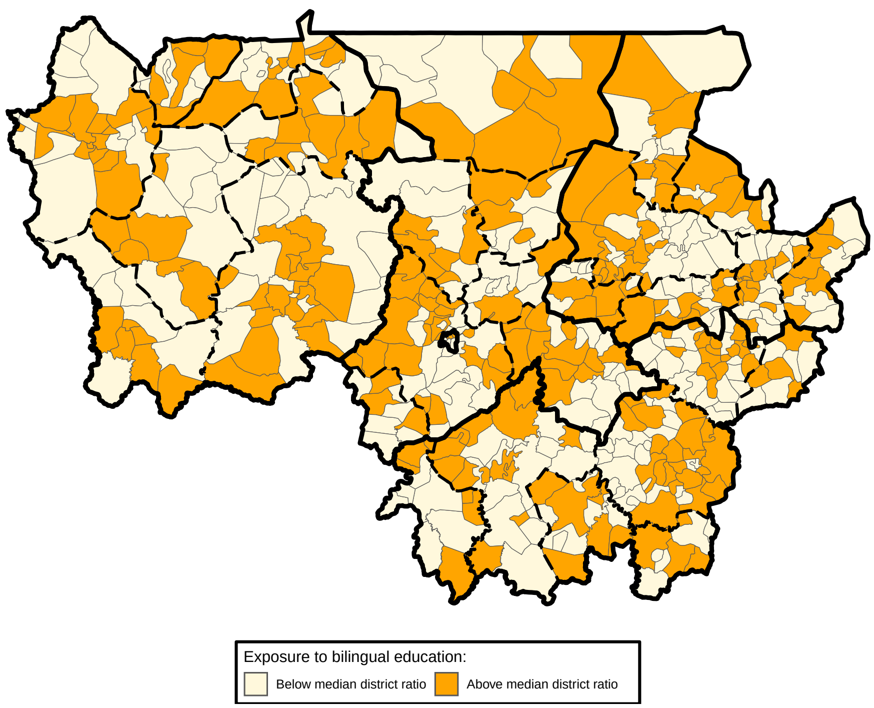
9 Conclusion

More and more African countries use local languages instead of colonial languages as the main medium of primary education instruction. However, the literature on such at-scale policies is still scarce. languages

This paper estimates the returns to bilingual education in Mali using implementation data on the current bilingual education program. The impacts are high, both in terms of learning and schooling. Girls are the ones who benefit the most from these additional new languages of instruction. All of these learning gains do not come at a cost in terms of nation-building. Section D provides insights into the monetary cost of such an intervention. While very preliminary, these computations suggest that bilingual education is very cost-effective.

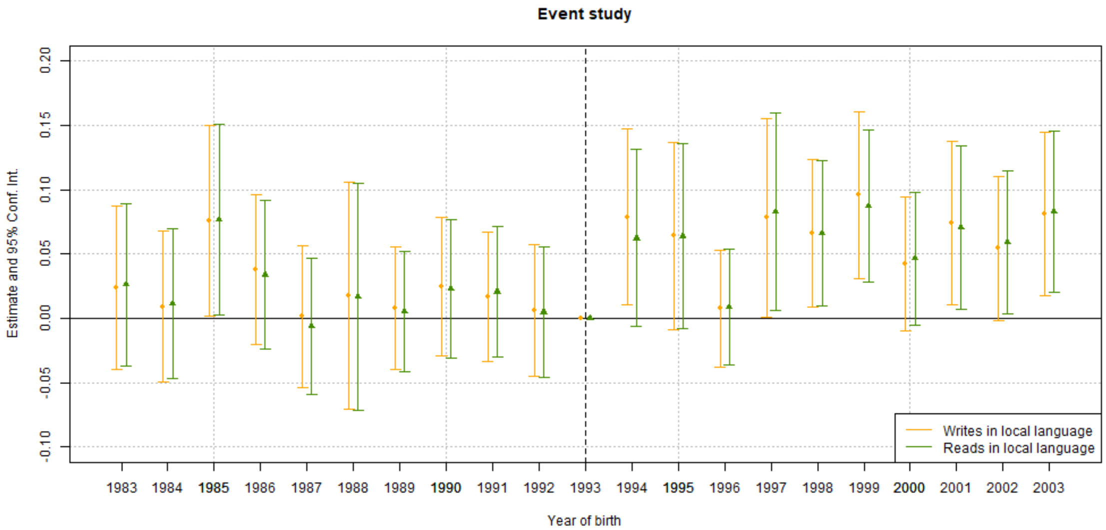
These findings have strong policy implications. In particular, they suggest that specific requirements are necessary for bilingual education to be effective, such as small class sizes. Further research is needed to understand the complementarity between basic school inputs and bilingual curricula, which would provide policymakers with more precise estimates of bilingual education's benefits under various implementation contexts.

Main Figures



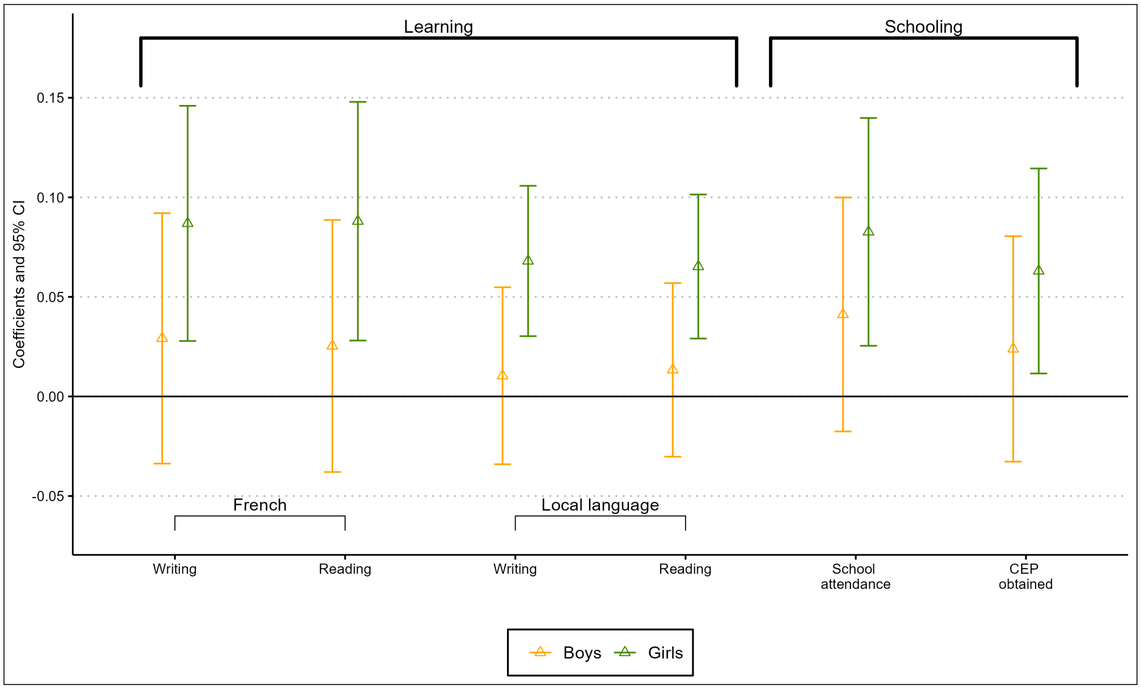
Notes: Solid black lines represent the regional borders, dashed black lines represent the district borders, and gray lines represent the communal borders. The two colors indicate the treatment status at the commune level, as defined in Subsection 5.2. This figure maps only the following regions that are considered in the empirical analysis: Bamako, Kayes, Koulikoro, Segou, and Sikasso.

Figure 1: Map of the treatment status by commune



Notes: Point estimates and 95% confidence intervals are derived from an event study regression over the 1983-2003 birth cohorts from which I extracted only the estimates of the interaction terms. I use district fixed effects to capture variation in the median of bilingual education share across districts. To obtain these estimates, I perform two distinct regressions: one with whether the individual knows how to write in the local language as the dependent variable (in orange), and one with whether the individual knows how to read in the local language as the dependent variable (in green). I cluster the standard errors at the LSMS cluster level, which corresponds roughly to the commune level. Individuals born in 1993 are the last birth cohort to be taught entirely in French (i.e., not treated) represented by a black vertical dashed line.

Figure 2: Event study regression for literacy in the local language



Notes: Point estimates and 95% confidence intervals are derived from a TWFE linear regression on several outcomes: (1) and (2) are writing and reading literacy in French, (3) and (4) in the local language, and (5) (6) and (7) are school achievements indicators. I use district fixed effects to capture the fact that the median of bilingual education share varies across districts. I cluster at the LSMS cluster level, which corresponds roughly to the village level. To obtain heterogeneous effects, I interact the DiD coefficient in Equation 2 with sex.

Figure 3: TWFE coefficient estimates of bilingual education effects, by sex

Main Tables

Characteristics	Mean	SD	N_1	Diff. in means	p-value	N_2
A. Individual characteristics						
Female (0/1)	0.55	0.5	8636	-0.04	0.03	4200
Age	23.7	6.12	8636	0.13	0.27	4200
Muslim (0/1)	0.95	0.22	8636	-0.02	0.2	4200
Urban (0/1)	0.49	0.48	8636	0.06	0.3	4200
Attended school (0/1)	0.51	0.5	8635	0.04	0.25	4199
Number of schooling years	4.01	6.25	8636	0.56	0.27	4200
Literate in French (0/1)	0.52	0.5	8633	0.06	0.17	4198
Literate in the local language (0/1)	0.13	0.33	8623	0.03	0.03	4195
B. Commune characteristics						
Number of students in school age	616.65	893.63	167	-32	0.74	167
Number of primary schools	24.74	35.01	167	-8.54	0.04	167
Share of bilingual schools	0.45	0.32	167	0.28	0	167
Linguistic HHI	0.46	0.15	165	0.01	0.76	165

Notes: (0/1) indicates a dummy variable. "SD" stands for standard deviation. All descriptive statistics are computed using weights provided in the LSMS survey. N_1 stands for the number of non-missing observations when computing these descriptive statistics. Column (3) presents the difference in means between the high-exposed and low-exposed to bilingual education sample (individuals and communes). Column (4) displays the p-value associated with the mean difference. The coefficients and p-values are estimated using weights provided in the LSMS survey, on the subsamples of individuals born before 1994, and clustering standard errors at the survey location level for individuals. The size of the sample used for estimating the mean difference between high and low-exposed samples is described in Column 5, N_2. Fixed effects at the district level are used in all these estimations.

Table 1: Description of the sample

Language tested	Local language		French	
Dependent Variables:	Writing	Reading	Writing	Reading
Model:	(1)	(2)	(3)	(4)
Variables				
Born after 1994	0.008	0.007	0.138***	0.137***
	(0.012)	(0.011)	(0.020)	(0.020)
High exposed commune	0.023	0.022	0.034	0.032
	(0.015)	(0.015)	(0.037)	(0.037)
Born after 1994 × High exposed commune	0.039**	0.039**	0.055*	0.054*
	(0.018)	(0.017)	(0.028)	(0.028)
Fixed-effects				
District FE	Yes	Yes	Yes	Yes
Fit statistics				
Mean of Y	0.123	0.118	0.489	0.484
Observations	8,636	8,623	8,636	8,633
R^2	0.04228	0.04123	0.15238	0.14913

Clustered (LSMS cluster) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. The level of analysis is at the individual level, i.e. one observation is one person. The differences in the number of observations is related to missingness of the outcome of interest in the raw LSMS data. I use the weights given in the LSMS data, and cluster the standard errors at the enumeration area. Clustering at the treatment level, i.e. the commune, produces the same results.

Table 2: TWFE results on learning outcomes

Dependent Variables:	Attended school	Obtained primary school diploma
Model:	(1)	(2)
Variables		
Born after 1994	0.156***	0.073***
	(0.019)	(0.016)
High exposed commune	0.023	0.028
	(0.032)	(0.032)
Born after 1994 × High exposed commune	0.059**	0.041*
	(0.026)	(0.025)
Fixed-effects		
District FE	Yes	Yes
Fit statistics		
Mean of Y	0.495	0.305
Observations	8,635	8,635
R^2	0.12179	0.12921

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. The level of analysis is at the individual level, i.e. one observation is one person. The differences in the number of observations is related to missingness of the outcome of interest in the raw LSMS data. I use the weights given in the LSMS data, and cluster the standard errors at the enumeration area. Clustering at the treatment level, i.e. the commune, produces the same results.

Table 3: TWFE results on schooling outcomes

Dependent Variables:	Doesn't dislike having a neighbor from another ethnic group	Feels more Malian than from her ethnic group	Feels confortable... Wearing her traditional or cultural dress in public	Feels confortable... Speaking her mother tongue in public
Model:	(1)	(2)	(3)	(4)
Variables				
Born after 1994	-0.088**	-0.026	-0.036	-0.044
	(0.036)	(0.039)	(0.026)	(0.032)
High exposed commune	0.012	-0.022	-0.051*	-0.037
	(0.021)	(0.041)	(0.028)	(0.031)
Born after 1994 × High exposed commune	0.076	0.099*	0.077**	0.077
	(0.046)	(0.059)	(0.038)	(0.049)
Fit statistics				
Mean of Y	0.917	0.816	0.955	0.943
Observations	854	851	466	467
R^2	0.04905	0.04081	0.10228	0.07983

Clustered (Afrobarometer EA) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The first outcome is derived from the question Q86B, and the second from the question Q82B, as coded in the codebook for the round 8 of Afrobarometer; I assigned a value of 0 for answers between 1 and 2, and 1 for answers between 3 and 5. The third outcome is derived from the question Q82D and the fourth from the question Q82C. The model estimated is a linear regression. The level of analysis is at the individual level, i.e. one observation is one person. The differences in the number of observations is related to missingness of the outcome of interest in the raw Afrobarometer data. I use district and survey round fixed-effects. I use the weights given in the Afrobarometer data, and cluster the standard errors at the enumeration area for every round of survey data collection. Clustering at the treatment level, i.e. the commune, produces the same results.

Table 4: TWFE results on ethnic salience outcomes

References

Angrist, Noam and Rachael Meager (2023). “Implementation Matters: Generalizing Treatment Effects in Education”. en. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.4487496.

Angrist, Noam et al. (2023). “Improving Learning in Low- and Lower-Middle-Income Countries”. en. In: *Journal of Benefit-Cost Analysis* 14.S1, pp. 55–80. DOI: 10.1017/bca.2023.26.

Ba, Oumar Issiaka (2009). *Une histoire de l’enseignement au Mali - Entre réforme et réticences*. L’Harmattan. La Sahélienne.

Ball, Mary-Claire et al. (July 2024). “Effective bilingual education in Francophone West Africa: constraints and possibilities”. en. In: *International Journal of Bilingual Education and Bilingualism* 27.6, pp. 821–835. DOI: 10.1080/13670050.2023.2290482.

Bank, World (2000). *Document d’évaluation du projet d’appui a l’amélioration des apprentissages dans les écoles fondamentales*. Tech. rep. 1960 1 -MLI.

Benson, Carol (Apr. 2004). “Do We Expect Too Much of Bilingual Teachers? Bilingual Teaching in Developing Countries”. en. In: *International Journal of Bilingual Education and Bilingualism* 7.2-3, pp. 204–221. DOI: 10.1080/13670050408667809.

— (2005). *Girls, educational equity and mother tongue-based teaching*. Tech. rep. UNESCO Office Bangkok, Regional Bureau for Education in Asia, and the Pacific.

Benson, Carolyn J. (Dec. 2002). “Real and Potential Benefits of Bilingual Programmes in Developing Countries”. In: *International Journal of Bilingual Education and Bilingualism* 5.6. Publisher: Routledge _eprint: https://doi.org/10.1080/13670050208667764, pp. 303–317. DOI: 10.1080/13670050208667764.

Blanc, Guillaume and Masahiro Kubo (2023). “French”. en. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.4636225.

Bold, Tessa et al. (Nov. 2017). “Enrollment without Learning: Teacher Effort, Knowledge, and Skill in Primary Schools in Africa”. en. In: *Journal of Economic Perspectives* 31.4, pp. 185–204. DOI: 10.1257/jep.31.4.185.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (Jan. 2024). “Revisiting Event Study Designs: Robust and Efficient Estimation”. en. In: *Papers*. Number: 2108.12419 Publisher: arXiv.org.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. Sant’Anna (Feb. 2024). *Difference-in-differences with a Continuous Treatment*. en. Tech. rep. w32117. Cambridge, MA: National Bureau of Economic Research, w32117. DOI: 10.3386/w32117.

Callaway, Brantly and Pedro H. C. Sant’Anna (Dec. 2021). “Difference-in-Differences with multiple time periods”. In: *Journal of Econometrics*. Themed Issue: Treatment Effect 1 225.2, pp. 200–230. DOI: 10.1016/j.jeconom.2020.12.001.

Calvet, Louis-Jean (1993). *Évaluation des écoles expérimentales en bambara dans la région de Ségou*. Tech. rep. République du Mali. Ministère de l’éducation de base. Institut pédagogique national.

— (2010). *Histoire du français en Afrique. Une langue en copropriété ?. Organisation internationale de la francophonie*. fr. Editions Ecriture.

Chaisemartin, Clément de, Xavier D’Haultfœuille, and Gonzalo Vazquez-Bare (May 2024). “Difference-in-Difference Estimators with Continuous Treatments and No Stayers”. en. In: *AEA Papers and Proceedings* 114, pp. 610–613. DOI: 10.1257/pandp.20241049.

Chauvet, Lisa and Marion Mercier (Aug. 2014). “Do return migrants transfer political norms to their origin country? Evidence from Mali”. In: *Journal of Comparative Economics* 42.3, pp. 630–651. DOI: 10.1016/j.jce.2014.01.001.

Chauvet, Lisa et al. (Apr. 2015). “Migrants’ Home Town Associations and Local Development in Mali: Migrants’ HTAs and local development in Mali”. en. In: *The Scandinavian Journal of Economics* 117.2, pp. 686–722. DOI: 10.1111/sjoe.12100.

Clots-Figueras, Irma and Paolo Masella (Aug. 2013). “Education, Language and Identity”. In: *The Economic Journal* 123.570, F332–F357. DOI: 10.1111/ecoj.12051.

CONFEMEN (2020). *PASEC 2019. Qualité des systèmes éducatifs en Afrique Subsaharienne Francophone. Performance et Environnement de l’Enseignement-Apprentissage au Primaire*. Tech. rep. Programme d’Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC).

Cummins, Jim (2000). *Language, power, and pedagogy: bilingual children in the crossfire*. en. Bilingual education and bilingualism 23. Clevedon [England] ; Buffalo [N.Y.]: Multilingual Matters.

De Gaston, William (Oct. 2011). "Le signal communicationnel français face à la langue Évé : Pour une phénoménologie de la communication". fr. In: *Communiquer. Revue de communication sociale et publique* 6. Number: 6, pp. 57–70. DOI: 10.4000/communiquer.452.

de Chaisemartin, Clément and Xavier D'Haultfœuille (June 2022). "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey". en. In: *The Econometrics Journal*, utac017. DOI: 10.1093/ectj/utac017.

Diarra, Abdramane (2020). "Curriculum Bilingue dans l'enseignement fondamental au Mali : Etat des lieux de sa mise en œuvre". PhD thesis. France: Université Grenoble-Alpes.

Diarra, Abou (2013). *Etats des lieux du curriculum de l'enseignement fondamental au Mali*. Rapport final Curriculum OMAES.

Dizon-Ross, Rebecca and Seema Jayachandran (Dec. 2023). "Detecting Mother-Father Differences in Spending on Children: A New Approach Using Willingness-to-Pay Elicitation". en. In: *American Economic Review: Insights* 5.4, pp. 445–459. DOI: 10.1257/aeri.20220159.

Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment". en. In: *The American Economic Review* 91.4, pp. 795–813.

Eberheard, David M., Gary F. Simons, and Charles D. Fennig (2025). *Ethnologue Global Dataset (eds. 2025)*. Section: 2016-03-22 16:44:43.547.

Evans, David K and Amina Mendez Acosta (Jan. 2021). "Education in Africa: What Are We Learning?" In: *Journal of African Economies* 30.1, pp. 13–54. DOI: 10.1093/jae/ejaa009.

Filmer, Deon et al. (2018). "Learning-Adjusted Years of Schooling (LAYS). Defining a New Macro Measure of Education". In: *World Bank Policy Research Working Paper*.

Gazeaud, Jules and Claire Ricard (Jan. 2024). “Learning effects of conditional cash transfers: The role of class size and composition”. In: *Journal of Development Economics* 166, p. 103194. DOI: 10.1016/j.jdeveco.2023.103194.

Glewwe, P. and K. Muralidharan (2016). “Improving Education Outcomes in Developing Countries”. en. In: *Handbook of the Economics of Education*. Vol. 5. Elsevier, pp. 653–743. DOI: 10.1016/B978-0-444-63459-7.00010-5.

Hovens, Mart (Oct. 2002). “Bilingual Education in West Africa: Does It Work?” en. In: *International Journal of Bilingual Education and Bilingualism* 5.5, pp. 249–266. DOI: 10.1080/13670050208667760.

Hutchison, J, Abdramane Diarra, and J Poth (1990). *Évaluation de l’expérimentation en langues nationales dans l’enseignement fondamental en République du Mali*. Tech. rep. USAID.

INSTAT-Mali (2017). *Enquête Modulaire et Permanente auprès des Ménages*.

Laitin, David D. and Rajesh Ramachandran (May 2022). “Linguistic diversity, official language choice and human capital”. In: *Journal of Development Economics* 156, p. 102811. DOI: 10.1016/j.jdeveco.2021.102811.

Laitin, David D., Rajesh Ramachandran, and Stephen L. Walter (Oct. 2019). “The Legacy of Colonial Language Policies and Their Impact on Student Learning: Evidence from an Experimental Program in Cameroon”. en. In: *Economic Development and Cultural Change* 68.1, pp. 239–272. DOI: 10.1086/700617.

Loua, Seydou (Sept. 2017). “Les grandes réformes de l’école malienne de 1962 à 2016”. fr. In: *Revue internationale d’éducation de Sèvres* 75, pp. 34–40. DOI: 10.4000/ries.5917.

Maurer, Bruno (2007). *De la pédagogie convergente à une didactique intégrée.* L’Harmattan.

Mbiti, Isaac et al. (Aug. 2019). “Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania”. en. In: *The Quarterly Journal of Economics* 134.3, pp. 1627–1673. DOI: 10.1093/qje/qjz010.

MEN (2003). *Projet d’Appui à l’Amélioration des Apprentissages dans les Ecoles Fondamentales (PAAA)*. Tech. rep. Crédit n°3318 MLI. Ministère de l’Education Nationale - Secrétariat Général.

Michler, Jeffrey D. et al. (Sept. 2022). "Privacy protection, measurement error, and the integration of remote sensing and socioeconomic survey data". In: *Journal of Development Economics* 158, p. 102927. DOI: 10.1016/j.jdeveco.2022.102927.

Mohohlwane, Nompumelelo et al. (2023). "Reading Skills Transfer Best from Home Language to a Second Language". en. In: *CGD Working Paper. Washington, DC: Center for Global Development* 647.

Müller-Crepon, Carl (July 2023). "State reach and development in Africa since the 1960s: new data and analysis". en. In: *Political Science Research and Methods* 11.3, pp. 663–672. DOI: 10.1017/psrm.2021.60.

Müller-Crepon, Carl, Philipp Hunziker, and Lars-Erik Cederman (Feb. 2021). "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa". EN. In: *Journal of Conflict Resolution* 65.2-3. Publisher: SAGE Publications Inc, pp. 563–590. DOI: 10.1177/0022002720963674.

Patrinos, Harry Anthony and Eduardo Velez (Nov. 2009). "Costs and benefits of bilingual education in Guatemala: A partial analysis". en. In: *International Journal of Educational Development* 29.6, pp. 594–598. DOI: 10.1016/j.ijedudev.2009.02.001.

Pengl, Yannick I., Philip Roessler, and Valeria Rueda (Feb. 2022). "Cash Crops, Print Technologies, and the Politicization of Ethnicity in Africa". en. In: *American Political Science Review* 116.1, pp. 181–199. DOI: 10.1017/S0003055421000782.

Piper, Benjamin, Stephanie S. Zuilkowski, and Salome Ong'ele (Nov. 2016). "Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya". en. In: *Comparative Education Review* 60.4, pp. 776–807. DOI: 10.1086/688493.

Psacharopoulos, George and Harry Anthony Patrinos (Aug. 2004). "Returns to investment in education: a further update". en. In: *Education Economics* 12.2, pp. 111–134. DOI: 10.1080/0964529042000239140.

Q'Gara, Chloe and Nancy Kendall (1996). *Beyond Enrollment: A Handbook For Improving Girls' Experiences In Primary Classrooms*. en. Tech. rep. ABEL.

Ramachandran, Rajesh (Oct. 2017). "Language use in education and human capital formation: Evidence from the Ethiopian educational reform". en. In: *World Development* 98, pp. 195–213. DOI: 10.1016/j.worlddev.2017.04.029.

Ramachandran, Rajesh and Christopher Rauh (Dec. 2022). “The Imperium of the Colonial Tongue? Evidence on Language Policy Preferences in Zambia”. en. In: *Journal of African Economies* 32.1, pp. 52–80. DOI: 10.1093/jae/ejac002.

Rolstad, Kellie, Kate Mahoney, and Gene V. Glass (Sept. 2005). “The Big Picture: A Meta-Analysis of Program Effectiveness Research on English Language Learners”. en. In: *Educational Policy* 19.4, pp. 572–594. DOI: 10.1177/0895904805278067.

Roth, Jonathan et al. (Aug. 2023). “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature”. en. In: *Journal of Econometrics* 235.2, pp. 2218–2244. DOI: 10.1016/j.jeconom.2023.03.008.

Sakaryalı, Özge, Mazhar Bal, and Yakup Yıldırım (Oct. 2024). “Investigating the Relationship Between Mother Tongue Education and Creativity: A Systematic Review at the K-12 Level”. en. In: *Sage Open* 14.4, p. 21582440241307806. DOI: 10.1177/21582440241307806.

Sandefur, Justin (2017). *How USAID’s Demographic and Health Surveys Overestimate Literacy around the World and How to Fix That.*

Sanoh, Aly (Feb. 2015). “Rainfall Shocks, Local Revenues, and Intergovernmental Transfer in Mali”. In: *World Development* 66, pp. 359–370. DOI: 10.1016/j.worlddev.2014.08.022.

Seid, Yared (Dec. 2016). “Does learning in mother tongue matter? Evidence from a natural experiment in Ethiopia”. In: *Economics of Education Review* 55, pp. 21–38. DOI: 10.1016/j.econedurev.2016.08.006.

Skattum, Ingse (Jan. 2010). “L’introduction des langues nationales dans le système éducatif au Mali : objectifs et conséquences”. fr. In: *Journal of Language Contact* 3.1, pp. 247–270. DOI: 10.1163/19552629-90000013.

Sougane, Arouna (2014). “Migrations et transferts : un état des lieux”. fr. In: *Le Mali contemporain*. Ed. by Joseph Brunet-Jailly, Jacques Charmes, and Doulaye Konaté. IRD Éditions, pp. 613–642. DOI: 10.4000/books.irdeditions.21263.

Traoré, Samba (2001). “La Pédagogie Convergente : son expérimentation au Mali et son impact sur le système éducatif”. fr. In: *UNESCO, Bureau International d’Education.* Monographies-Innodata.

Trefault, Thierry (1999). *L'Ecole malienne à l'heure du bilinguisme: deux écoles rurales de la région de Ségou.* Agence intergouvernementale de la francophonie. Langues et développement.

UNESCO (1963). "Impulsions nouvelles à l'éducation". In.

UNESCO (2013). *THE GLOBAL LEARNING CRISIS. Why every child deserves a quality education.* Tech. rep. UNESCO.

UNESCO (2016). *If you don't understand, how can you learn?* Policy Paper. UNESCO.

Appendix

A Context

A.1 History of bilingual education prior to the 1999 policy

First-generation experimental schools In 1979, four schools opened a class using bilingual education in Bamanankan only, the main vehicular language in Mali (Calvet, 1993; Traoré, 2001). Teachers used the local language only in the first four years of primary education. Positive evaluations of the bilingual program at the beginning of the experiment led to its rapid expansion. A few years later, almost 100 schools operated in the four main languages: Bamanankan, Songhay, Tamasheq, and Fulfulde (Hutchison, Diarra, and Poth, 1990; Skattum, 2010). Lack of monitoring, teacher training, and sufficient budget led to the end of the first experimentation at the beginning of the 1990s (Skattum, 2010).

The *Pédagogie Convergente* This new experiment, developed and piloted by a Belgian linguistics center, the CIAVER (Centre International Audiovisuel d'Études et de Recherches), started in 1987 with two classes in Segou in Bamanankan. Following a positive evaluation of the pilot, it rapidly expanded in 1994 to other languages and had grown to approximately 100 schools by 1997 (Traoré, 2001). Unlike the first experimental schools, the 6 years of primary education were entirely taught in the local language, with a gradual introduction to French once students had fully mastered their mother tongue.

Qualitative and quantitative evaluations conducted during the experiment provided mixed evidence about its results (Skattum, 2010; Traoré, 2001). Maurer (2007) also reported implementation issues that hindered the program's scale-up.

A.2 Figures

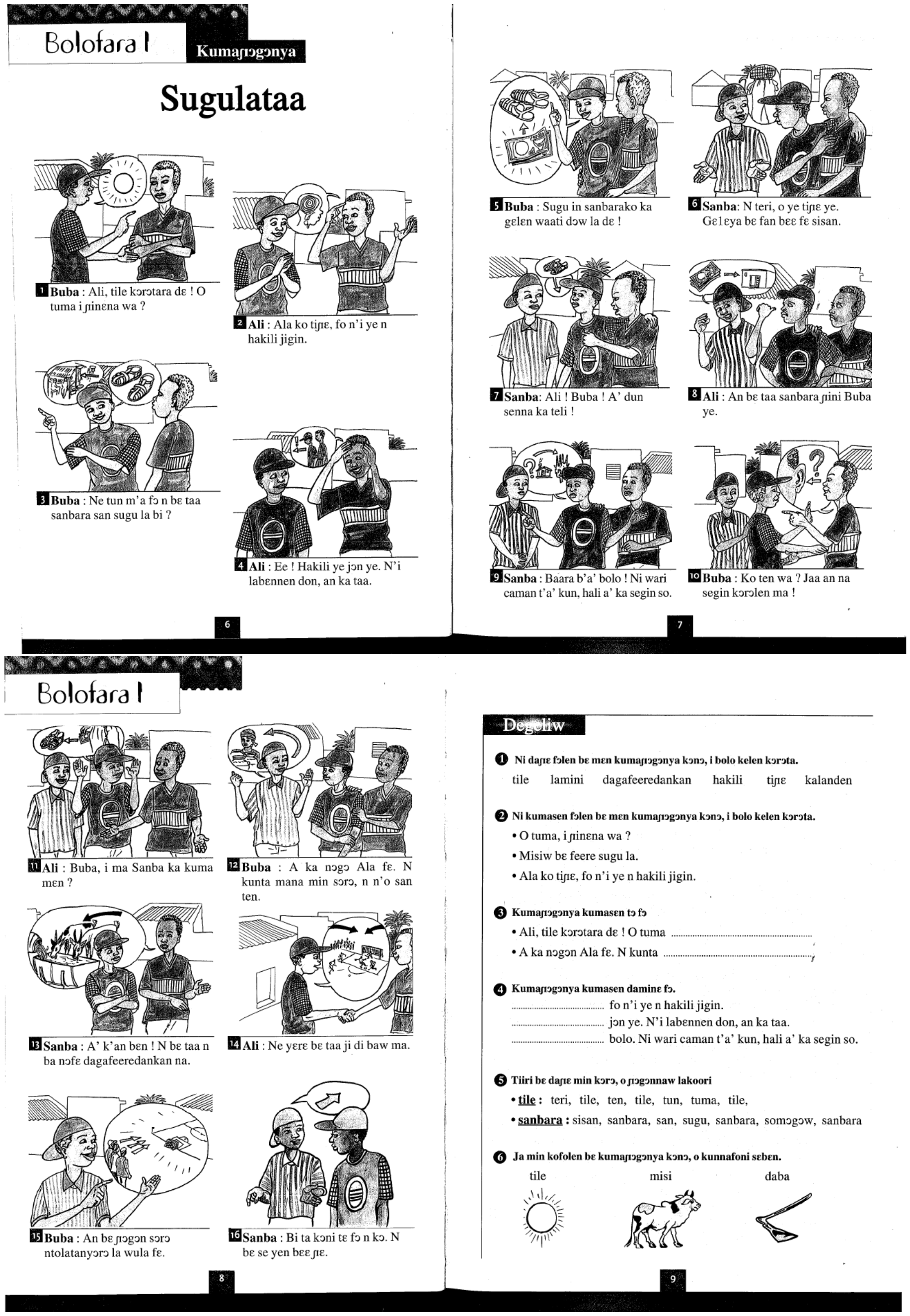
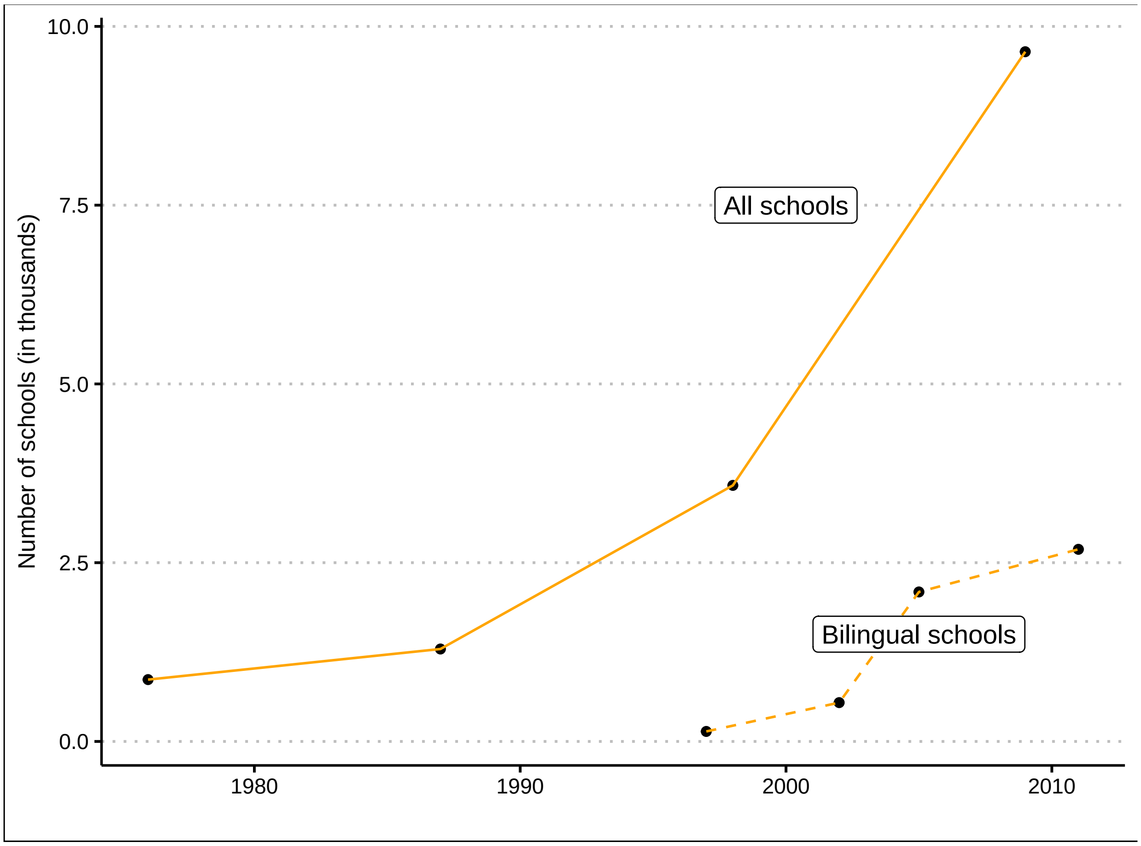


Figure A.1: Examples of a textbook in Bamanankan for grade 1 students



Notes: The orange line represents the school supply, regardless of the languages of instruction used in the school. The school supply was inferred from the national census data. The orange dotted line represents the number of bilingual schools. The number of bilingual schools was retrieved from archival sources from the Bilingual Education monitoring section. The four points in time are the only years for which I have accurate data (see Table A.3).

Figure A.2: Time evolution of the number of schools

Notes: This figure presents the different language trees for the 12 languages that can be used as a language of instruction, as stated in the 1999 bilingual education reform in Mali. Ethnologue is the source used to create the trees (Eberheard, Simons, and Fennig, 2025). Yellow cells indicate languages introduced in the 1999 PRODEC reform, while green cells relate to language families. Family connections are indicated with black lines. The four language families on the left-hand side of the figure are the higher-level language families as listed in Ethnologue.

Figure A.3: Language trees for the languages of instruction in Mali

A.3 Tables

Primary education	**Local language**	**French**
Grade 1	100%	0%
Grade 2	75%	25%
Grade 3	50%	50%
Grade 4	50%	50%
Grade 5	25%	75%
Grade 6	25%	75%

Table A.1: Language repartition in the bilingual curriculum

Notes: This table comes from the work of Diarra (2020) and indicates the official time spent teaching in French and the local language at each grade of the primary cycle.

Region	**Language(s)**
Kayes	Soninke, Bamanankan, Fulfulde, Khassonke
Koulikoro	Bamanankan, Syenara, Mamara
Sikasso	Bamanankan, Syenara, Mamara
Segou	Bamanankan, Bomu, Mamara
Mopti	Fulfulde, Dogon, Bamanankan, Songhay, Bozo
Tombouctou	Songhay, Tamasheq, Fulfulde
Gao	Songhay, Tamasheq
Kidal	Tamasheq
Bamako	Bamanankan

Table A.2: Language choice per region

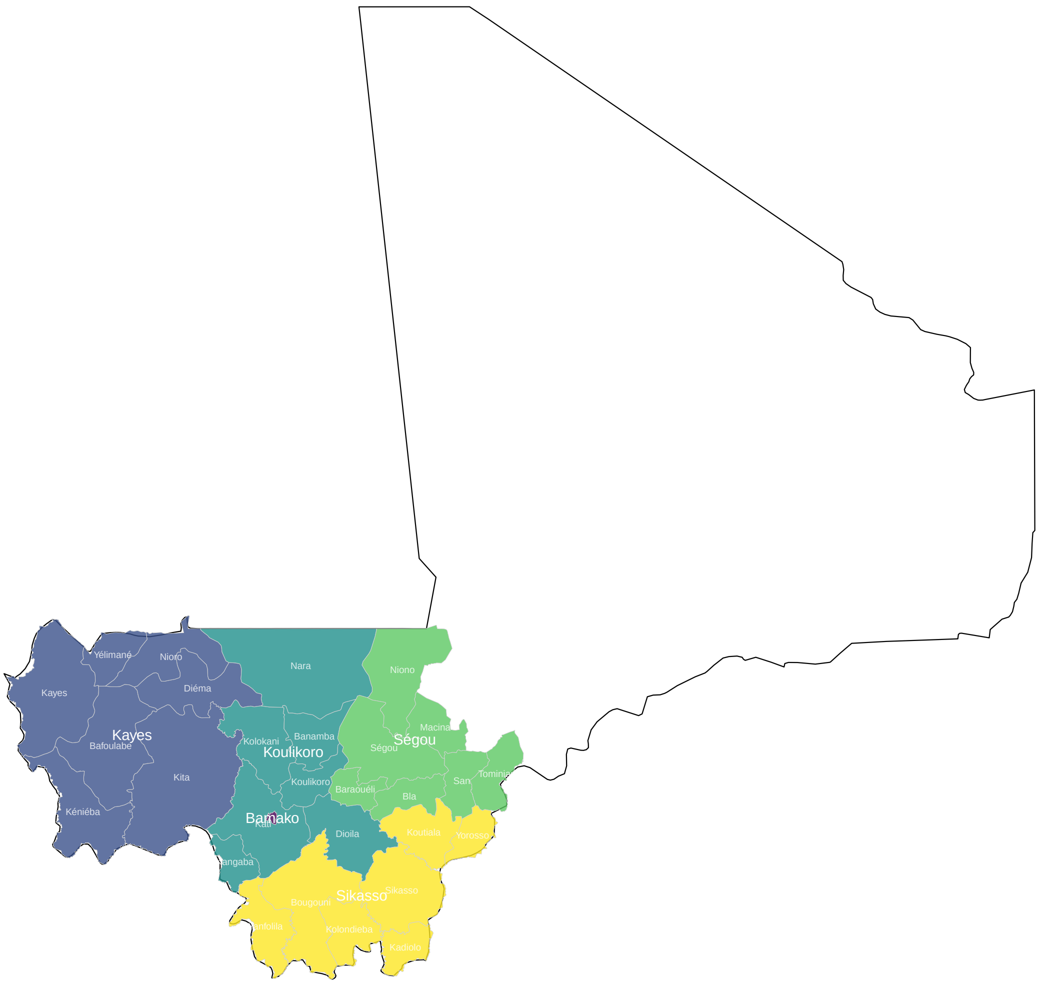
Notes: This table comes from the work of Diarra (2020) and indicates the local language(s) chosen to be languages of instruction for at least one school in each region.

Year	**Geographical level**	**Information**	**Source**
1994-1997	District	Number of BS per language	Trefault (1999) Diarra (2020)
2002	District	Number of BS	MEN (2003)
2005	Region District	Number of BS per language Number of BS	Official reports
2011	Commune	List of BS with info. on their status	Diarra (2013)

Table A.3: Year, geographical level, and information available for bilingual education

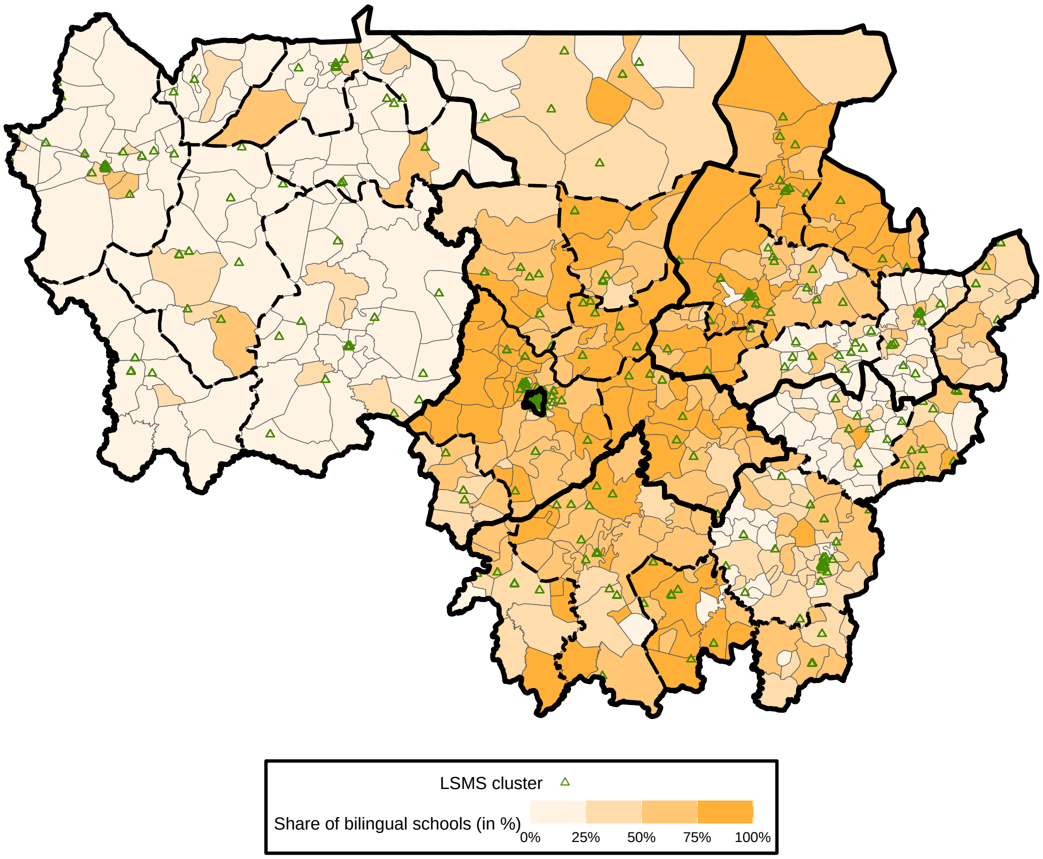
Notes: "BS" stands for bilingual schools. "1994-1997" stands for "from 1994 to 1997". This table provides information on data retrieved regarding bilingual education supply in Mali. To document the expansion of bilingual education, I use original data from several reports published by the Bilingual Education Department within the Ministry of Education between 2002 and 2011 (Diarra, 2013; MEN, 2003). I also use data on the first experiments from a PhD thesis in linguistics (Diarra, 2020) and a field qualitative study conducted at the end of the 1990s (Trefault, 1999). From the various sources detailed previously, I can document four periods: the first experiment from 1994 to 1997, followed by 2002, 2005, and 2011.

A.4 Maps



Notes: This map presents the regions and districts covered by the 2011 census on bilingual schools, and the rest of the Malian territory. Regions are indicated in colors, while districts are delimited in gray. The black lines represent the country's borders.

Figure A.4: Map of Mali with regions covered by the 2011 bilingual education census

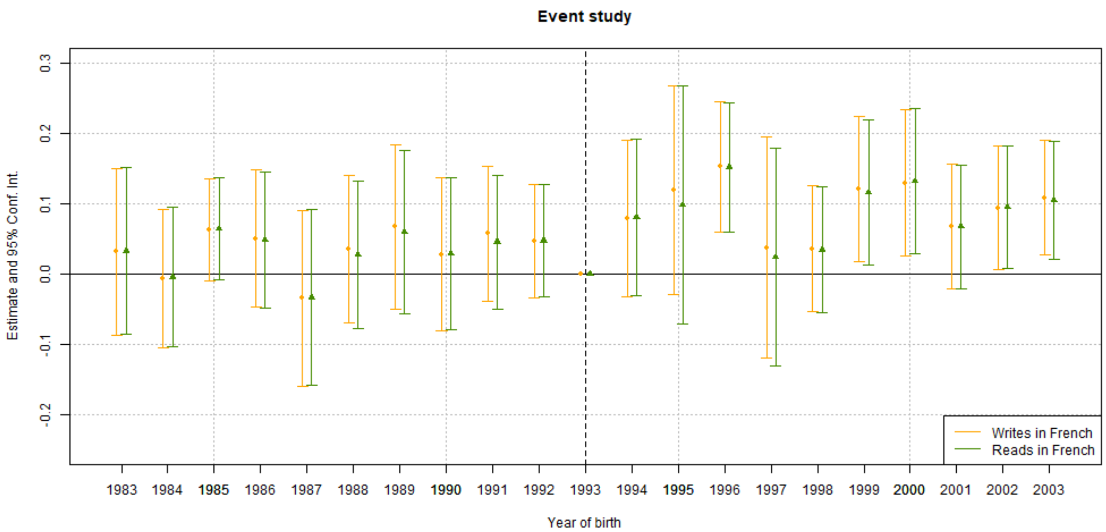


Notes: Solid black lines represent the regional borders, dashed black lines represent the district borders, and gray lines represent the communal borders. The color represents the intensity in the fraction of bilingual schools in each commune: the whiter the commune, the less the provision of bilingual education in the commune. The green triangles represent the location of LSMS clusters used in the analysis. This figure maps only the following regions considered in the empirical analysis: Bamako, Kayes, Koulikoro, Segou, and Sikasso.

Figure A.5: Bilingual education supply in 2011 and LSMS clusters by commune

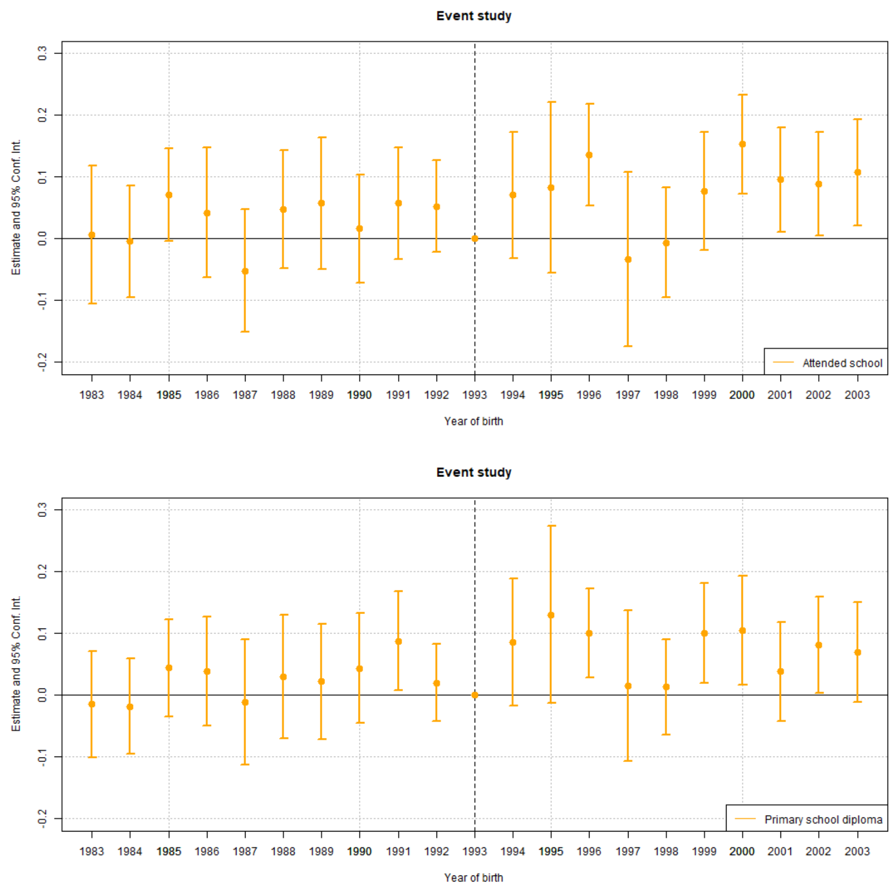
B Results

B.1 Figures



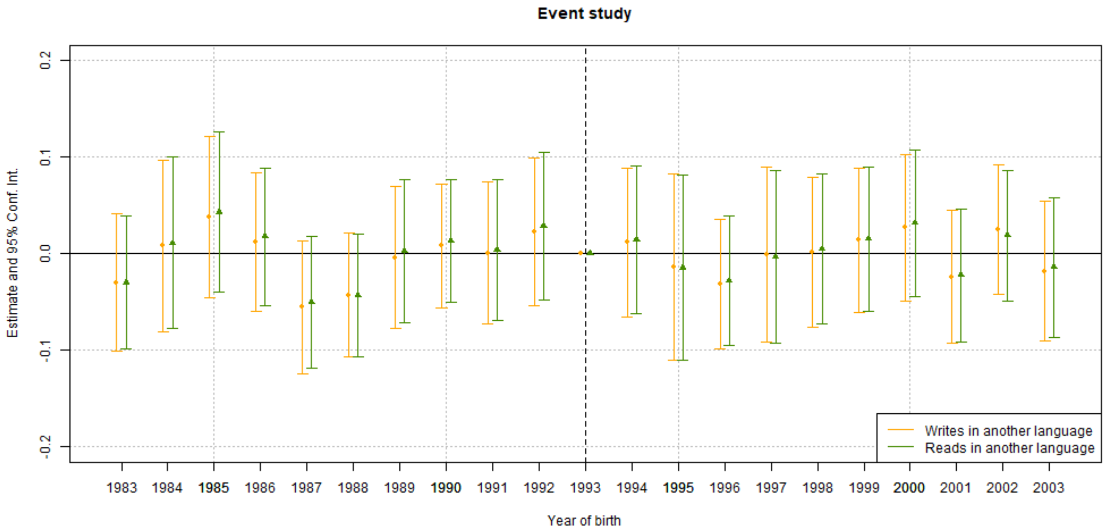
Notes: Point estimates and 95% confidence intervals are derived from an event study regression over the 1983-2003 birth cohorts, specified in Equation 3. In addition to time fixed effects, I use district fixed effects to capture that the median of bilingual education share varies across districts. To obtain these estimates, I perform two distinct regressions on two separate outcomes: (i) whether the individual knows how to write in French as the dependent variable and (ii) whether the individual knows how to read in French as the dependent variable. I cluster at the LSMS cluster level, which corresponds roughly to the village level. Individuals born in 1993 are the last birth cohort to be taught entirely in French, i.e., not treated, represented by a black dashed vertical line.

Figure B.1: Event study results for literacy in French



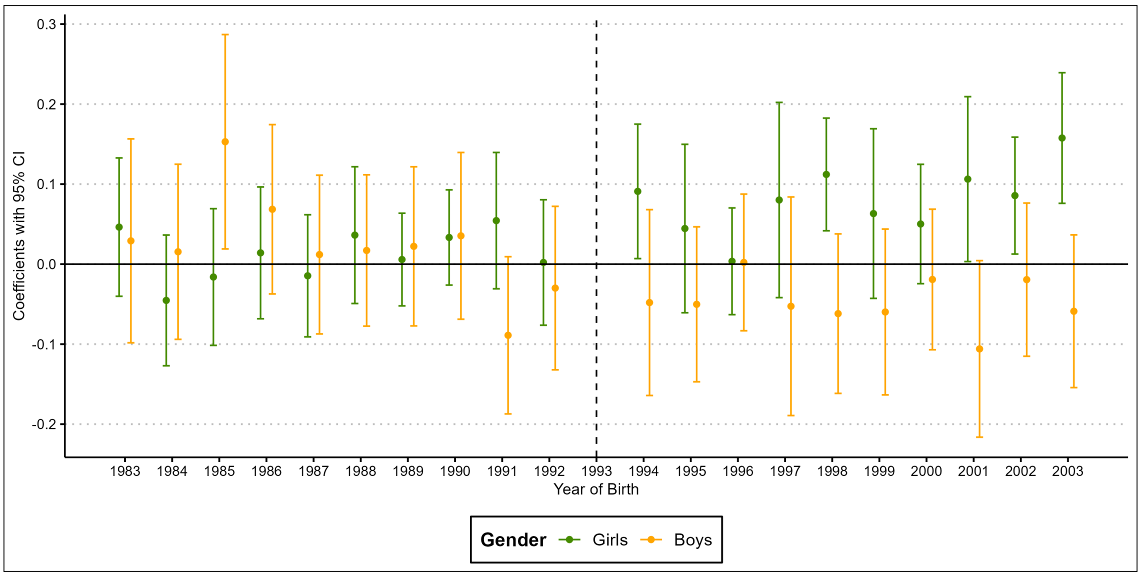
Notes: Point estimates and 95% confidence intervals are derived from an event study regression over the 1983-2003 birth cohorts, specified in Equation 3. In addition to time fixed effects, I use district fixed effects to capture that the median of bilingual education share varies across districts. To obtain these estimates, I perform two distinct regressions on two separate outcomes: (i) whether the individual attended school, and (ii) whether the individual obtained the end-of-primary-school exam, called CEP. I cluster at the LSMS cluster level, which corresponds roughly to the village level. Individuals born in 1993 are the last birth cohort to be taught entirely in French, i.e., not treated, represented by a black dashed vertical line.

Figure B.2: Event study results for schooling outcomes



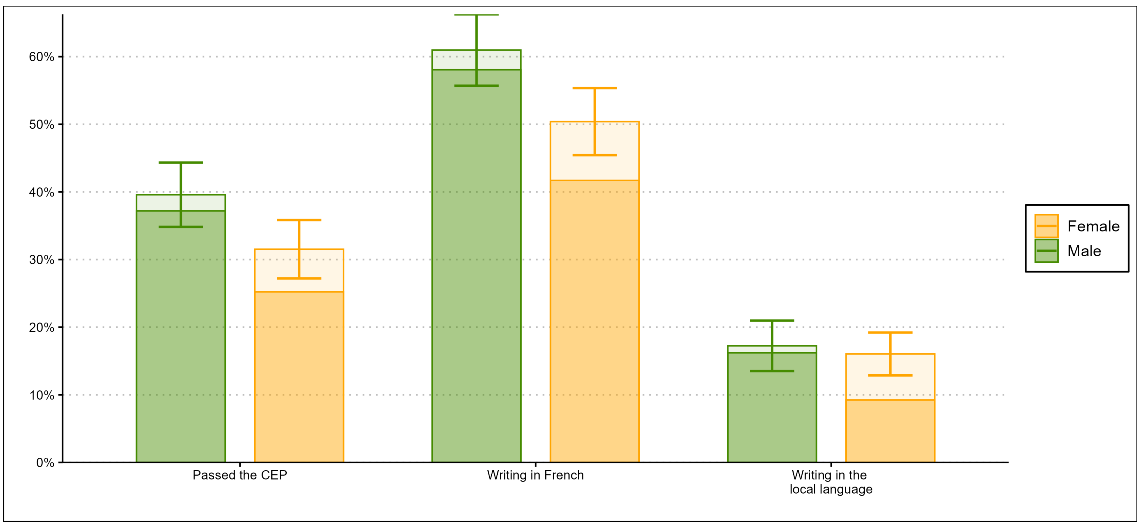
Notes: Point estimates and 95% confidence intervals are derived from an event study regression over the 1983-2003 birth cohorts, specified in Equation 3. In addition to time fixed effects, I use district fixed effects to capture variation in the median bilingual education share across districts. To obtain these estimates, I perform two distinct regressions on two separate outcomes: (i) whether the individual knows how to write in another language as the dependent variable (in orange), and (ii) whether the individual knows how to read in another language as the dependent variable (in green). I cluster the standard errors at the LSMS cluster level, which corresponds roughly to the village level. Individuals born in 1993 are the last birth cohort to be taught entirely in French (i.e., not treated), represented by a black vertical dashed line.

Figure B.3: Event study results for literacy in another language



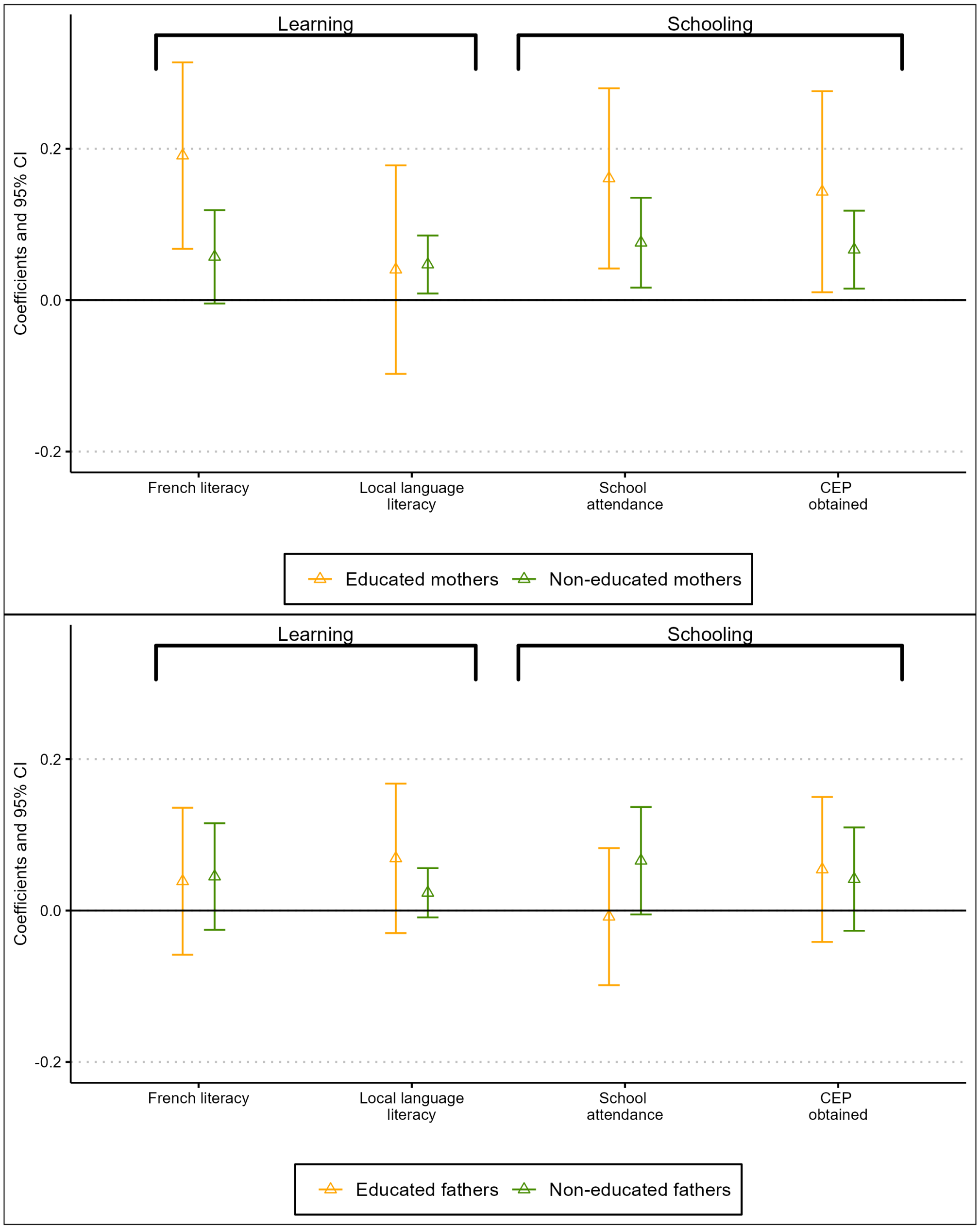
Notes: Point estimates and 95% confidence intervals are derived from an event study regression over the 1985-2003 birth cohorts, specified in Equation 3. In addition to time fixed effects, I use district fixed effects to capture variation in the median bilingual education share across districts. The outcome considered here is whether the individual knows how to read and write in the local language. I interact the α_t and β_t coefficients with a dummy indicating the respondent's gender. I cluster the standard errors at the LSMS cluster level, which corresponds roughly to the village level. Individuals born in 1993 are the last birth cohort to be taught entirely in French (i.e., not treated), represented by a black vertical dashed line.

Figure B.4: Event study results for literacy in the local language, by gender



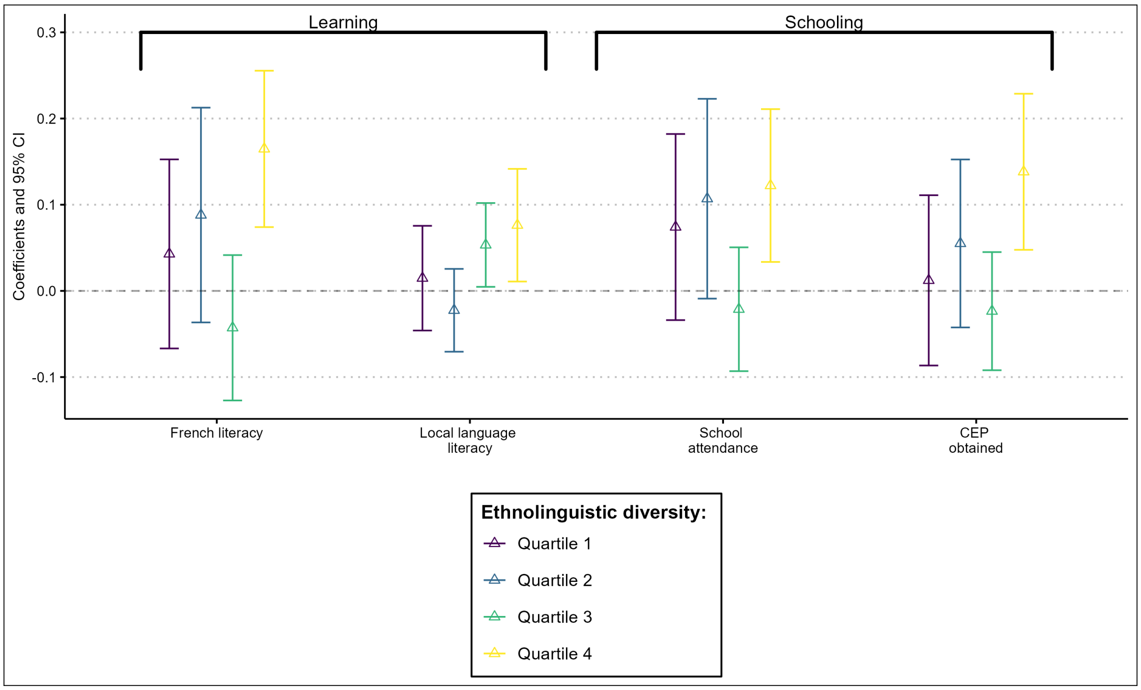
Notes: The solid bars represent the outcome mean at baseline, for the three outcomes described on the x-axis. The shadowed areas on top of the solid bars illustrate the effect of bilingual education on each of these outcomes, with the corresponding 95% confidence intervals. Each of these results is depicted at the gender level. Point estimates and 95% confidence intervals are derived from a TWFE linear regression, with the interaction dummy interacted with a gender dummy.

Figure B.5: TWFE coefficient estimates of bilingual education effects relative to the outcome means, by gender



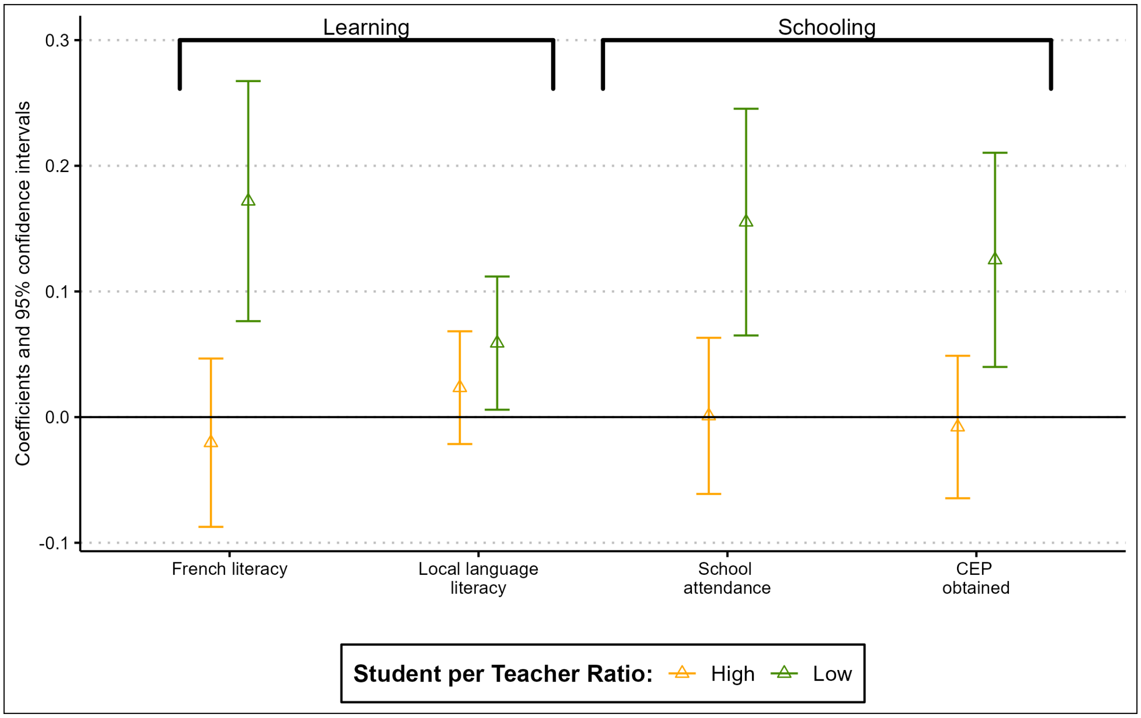
Notes: Point estimates and 95% confidence intervals are derived from a TWFE linear regression on several outcomes: Point estimates and 95% confidence intervals are derived from a TWFE linear regression on several outcomes: (1) and (2) are literacy measures (i.e., ability to write and read) in French and in the local language, and (3) (4) and (5) are school achievements indicators. I use district fixed effects to capture the fact that the median of bilingual education share varies across districts. I cluster at the LSMS cluster level, which corresponds roughly to the commune level. To obtain heterogeneous effects, I interact the DiD coefficient in Equation 2 with a dummy variable equal to one if the mother or father attended school, and I also add this dummy variable to the set of fixed effects.

Figure B.6: TWFE coefficient estimates of bilingual education effects, by parental education



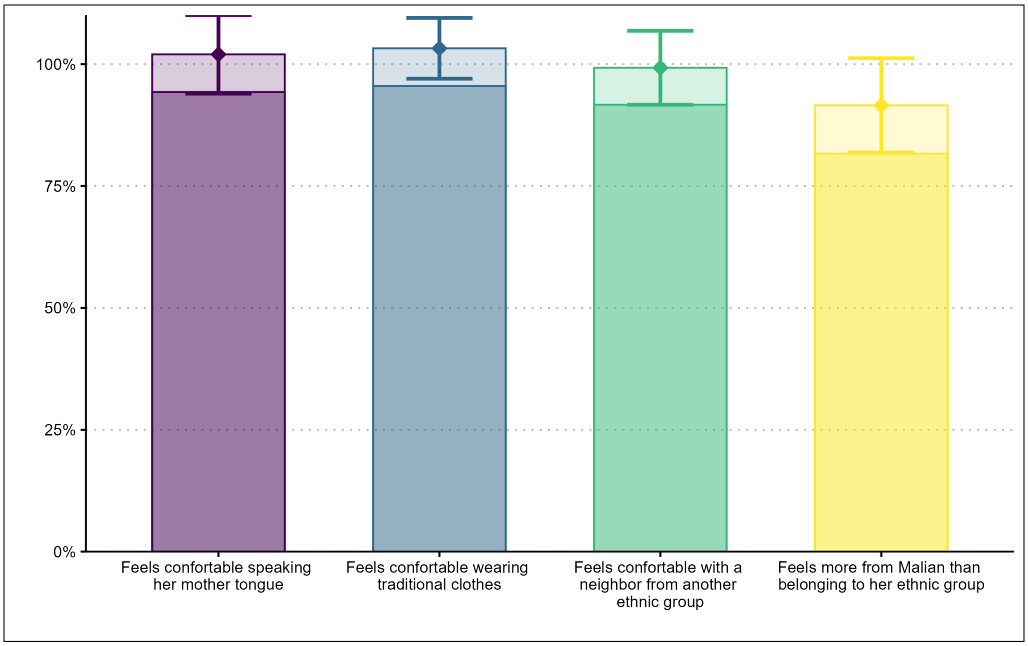
Notes: Point estimates and 95% confidence intervals are derived from a TWFE linear regression on several outcomes: (1) and (2) are literacy measures (i.e., ability to write and read) in French and in the local language, and (3) (4) and (5) are school achievements indicators. I use district fixed effects to capture the fact that the median of bilingual education share varies across districts. I cluster at the LSMS cluster level, which corresponds roughly to the commune level. To obtain heterogeneous effects, I interact the DiD coefficient in Equation 2 with quartiles of the ethnolinguistic diversity index, and add these quartiles to the set of fixed-effects as well.

Figure B.7: TWFE coefficient estimates of bilingual education effects, by linguistic diversity



Notes: Point estimates and 95% confidence intervals are derived from a TWFE linear regression on several outcomes. In addition to time fixed effects, I use district fixed effects to capture that the median of bilingual education share varies across districts. I cluster at the LSMS cluster level, which corresponds roughly to the village level. To obtain heterogeneous effects, I interact the DiD coefficient with a dummy variable indicating whether the student per teacher ratio at the district level is higher than the median.

Figure B.8: TWFE coefficient estimates of bilingual education effects, by the student-per-teacher ratio



Notes: The solid bars represent the outcome mean at baseline, for the four outcomes described on the x-axis. The shadowed areas on top of the solid bars illustrate the effect of bilingual education on each of these outcomes, with the corresponding 95% confidence intervals. Point estimates and 95% confidence intervals are derived from a TWFE linear regression, described in Table 4. The first outcome is derived from question Q86B, and the second from question Q82B, as coded in the codebook for Round 8 of Afrobarometer. I assigned a value of 0 for answers between 1 and 2, and 1 for answers between 3 and 5. The third outcome is derived from question Q82D, and the fourth from question Q82C.

Figure B.9: TWFE coefficient estimates of bilingual education effects on ethnic outcomes

B.2 Tables

	Local language		French		Attended	Primary school
Dependent Variables:	Writing	Reading	Writing	Reading	school	diploma
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Nb. bilingual schools (log)	0.048***	0.049***	0.138***	0.135***	0.106***	0.105***
	(0.015)	(0.014)	(0.024)	(0.024)	(0.022)	(0.021)
Nb. schools (log)	-0.005	-0.005	0.055**	0.057**	0.055**	0.048*
	(0.015)	(0.015)	(0.028)	(0.028)	(0.024)	(0.026)
Fixed-effects						
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Mean of Y	0.123	0.118	0.489	0.484	0.495	0.305
Observations	8,636	8,623	8,636	8,633	8,635	8,635
R^2	0.04330	0.04283	0.16919	0.16596	0.11773	0.15112

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. The logarithm of the number of schools is taken, to which one is added to avoid infinite values when this number is zero. The level of analysis is at the individual level, i.e., one observation is one person. The differences in the number of observations are related to missingness of the outcome of interest in the raw LSMS data. I use district fixed effects for all estimations. I use the weights given in the LSMS data, and cluster the standard errors at the enumeration area. Clustering at the treatment level, i.e., the commune, produces the same results.

Table B.1: Naive estimation results

Dependent Variables:	French literacy	Local language literacy	Attended school	Primary school diploma
Model:	(1)	(2)	(3)	(4)
Panel 1 – Heterogeneity variable: Bambara as the main language of the location				
HV	0.163*** (0.060)	0.032 (0.023)	0.121** (0.052)	0.127** (0.053)
DiD	0.106* (0.064)	0.037 (0.039)	0.109* (0.058)	0.058 (0.052)
DiD x HV	-0.070 (0.069)	0.001 (0.040)	-0.066 (0.061)	-0.024 (0.056)
Panel 2 – Heterogeneity variable: Urban location				
HV	0.326*** (0.040)	0.055*** (0.020)	0.264*** (0.035)	0.275*** (0.037)
DiD	0.054 (0.033)	0.017 (0.020)	0.058* (0.032)	0.025 (0.029)
DiD x HV	-0.028 (0.042)	0.058** (0.029)	-0.022 (0.042)	0.024 (0.037)
Panel 3 – Heterogeneity variable: Location less than 10km from a road (in 2000)				
HV	0.152*** (0.026)	0.027* (0.015)	0.147*** (0.026)	0.115*** (0.021)
DiD	0.038 (0.042)	0.018 (0.022)	0.042 (0.037)	0.016 (0.034)
DiD x HV	0.043 (0.060)	0.046* (0.026)	0.044 (0.050)	0.060 (0.052)
Panel 4 – Heterogeneity variable: All BS in the commune are still bilingual in 2011				
HV	-0.116** (0.058)	-0.009 (0.018)	-0.102** (0.045)	-0.073 (0.052)
DiD	-0.024 (0.035)	0.035 (0.022)	-0.019 (0.034)	-0.007 (0.029)
DiD x HV	0.135*** (0.045)	0.007 (0.028)	0.137*** (0.043)	0.083** (0.038)
Fixed-effects				
District FE	Yes	Yes	Yes	Yes

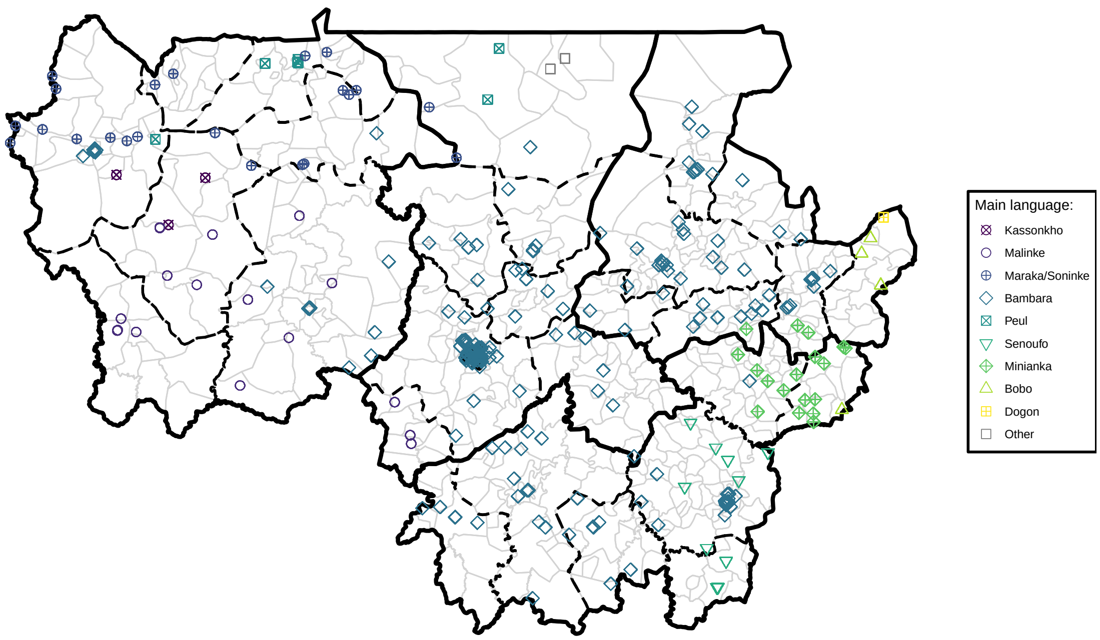
Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: "BS" stands for bilingual schools. All dependent variables are binary outcomes. The model estimated is a linear regression. Literacy in one language is equal to one when the individual knows how to write and read in this language. The level of analysis is at the individual level, i.e., one observation is one person. I use the weights provided in the LSMS data and cluster the standard errors at the enumeration area level. Clustering at the treatment level, i.e., the commune, produces the same results. HV stands for heterogeneity variable. DiD stands for the difference-in-difference interaction coefficient. I report only this coefficient, along with the heterogeneity variable coefficient, detailed in the first column, and the interaction coefficient between the DiD dummy and the heterogeneity dummy.

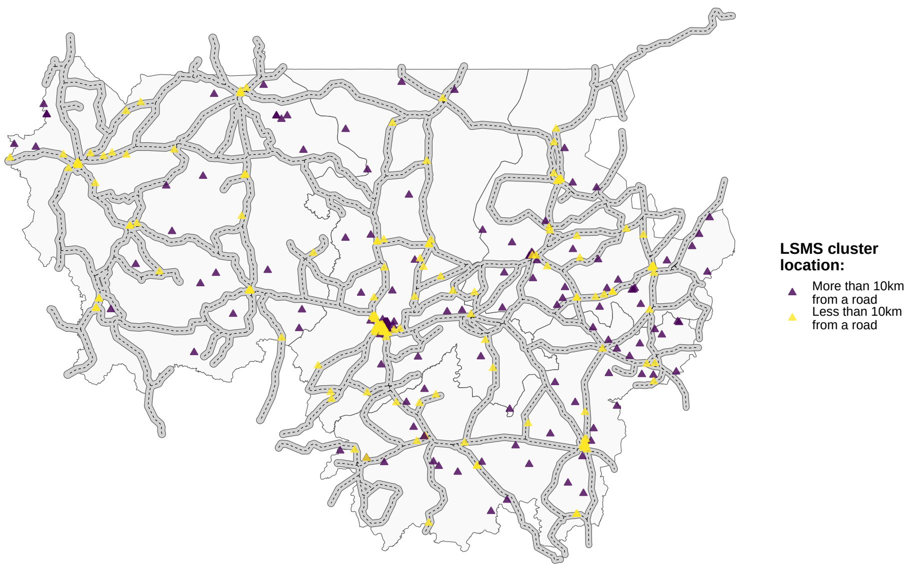
Table B.2: Heterogeneity results on language, urban, road access, and continuity in the bilingual education provision

B.3 Maps



Notes: Solid black lines represent the regional borders, dashed black lines represent the district borders, and gray lines represent the communal borders. The dots are the LSMS cluster locations. Colors and shapes indicate the dominant language of the commune in which these clusters are located, as given in the 2009 census.

Figure B.10: Map of LSMS clusters with the main language spoken

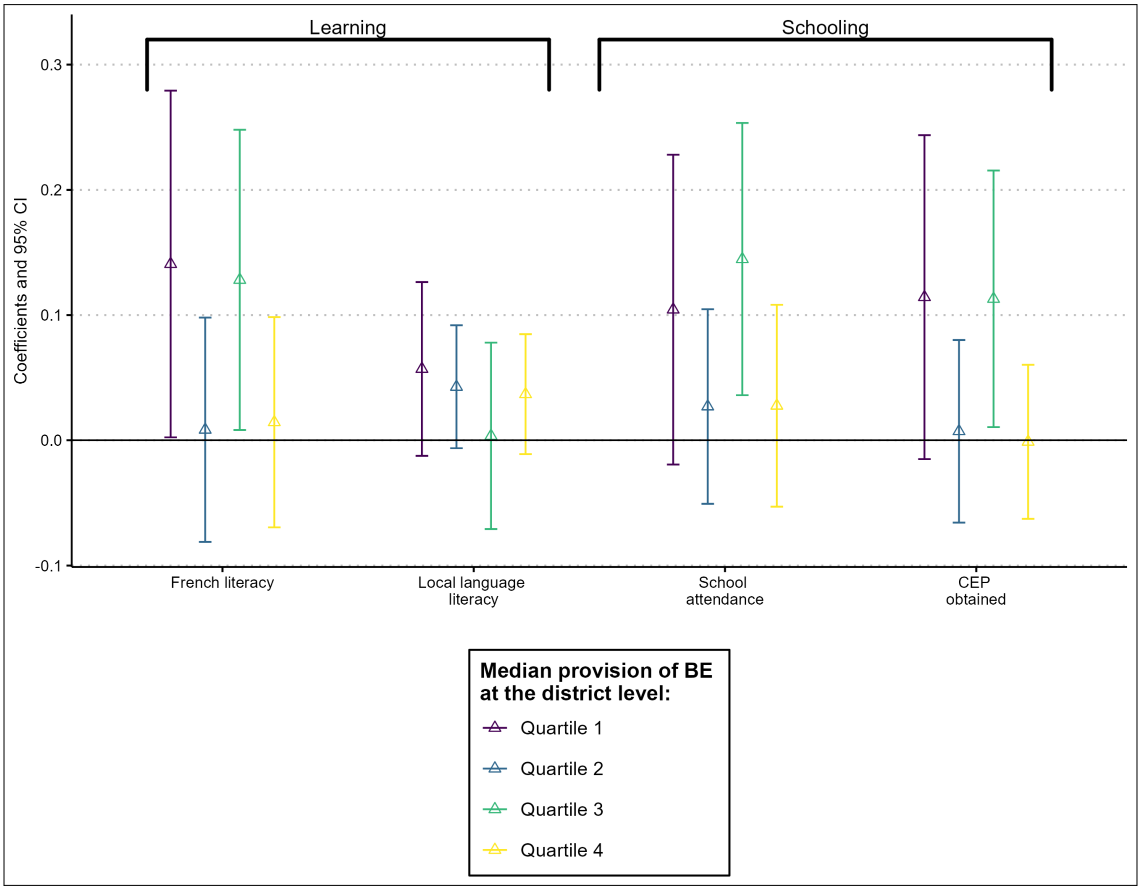


Notes: Solid black lines represent the regional borders. Roads represented on the map are taken from the Michelin map from 2000 (Müller-Crepon, 2023; Müller-Crepon, Hunziker, and Cederman, 2021). The dots represent the LSMS cluster locations, and their colors vary according to whether they are located within 10 kilometers (as the crow flies) of the closest road.

Figure B.11: Map of road access to LSMS clusters

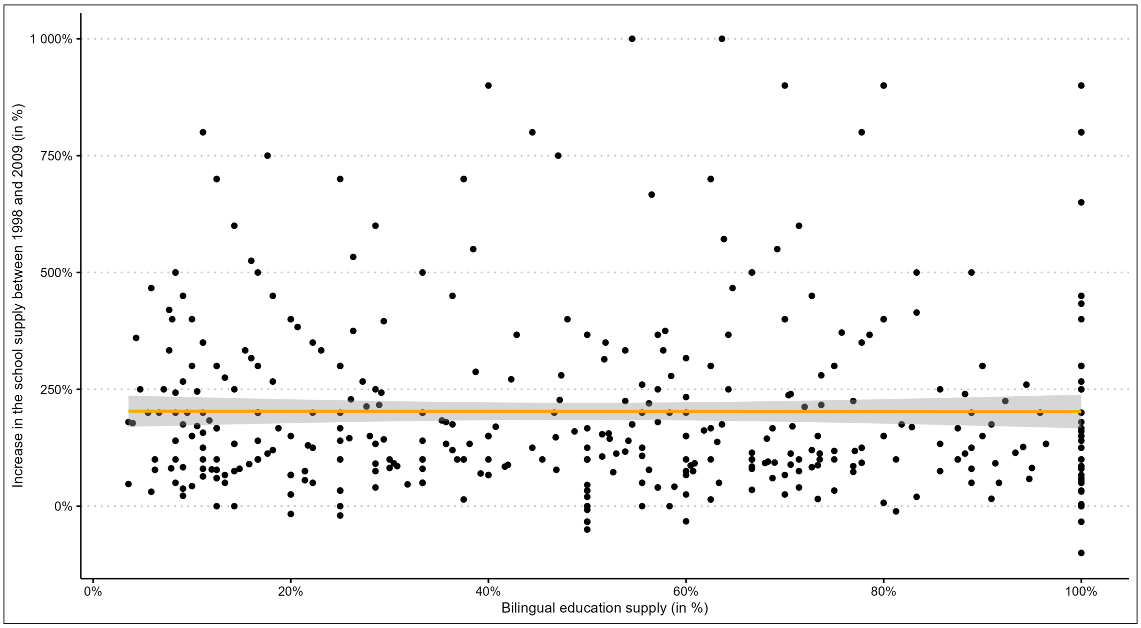
C Robustness checks

C.1 Figures



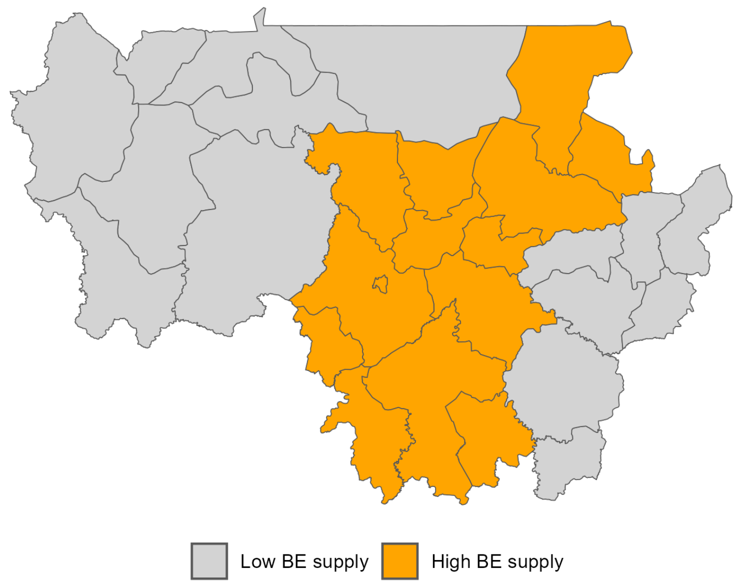
Notes: Point estimates and 95% confidence intervals are derived from a TWFE linear regression on several outcomes: (1) and (2) are literacy measures (i.e., ability to write and read) in French and in the local language, and (3) (4) and (5) are school achievements indicators. I use district fixed effects to capture the fact that the median of bilingual education share varies across districts. I cluster at the LSMS cluster level, which corresponds roughly to the commune level. To obtain heterogeneous effects, I interact the DiD coefficient in Equation 2 with quartiles of bilingual education share, and add these quartiles to the set of fixed-effects as well.

Figure C.1: TWFE coefficient estimates of bilingual education effects, by the median share of bilingual education provision



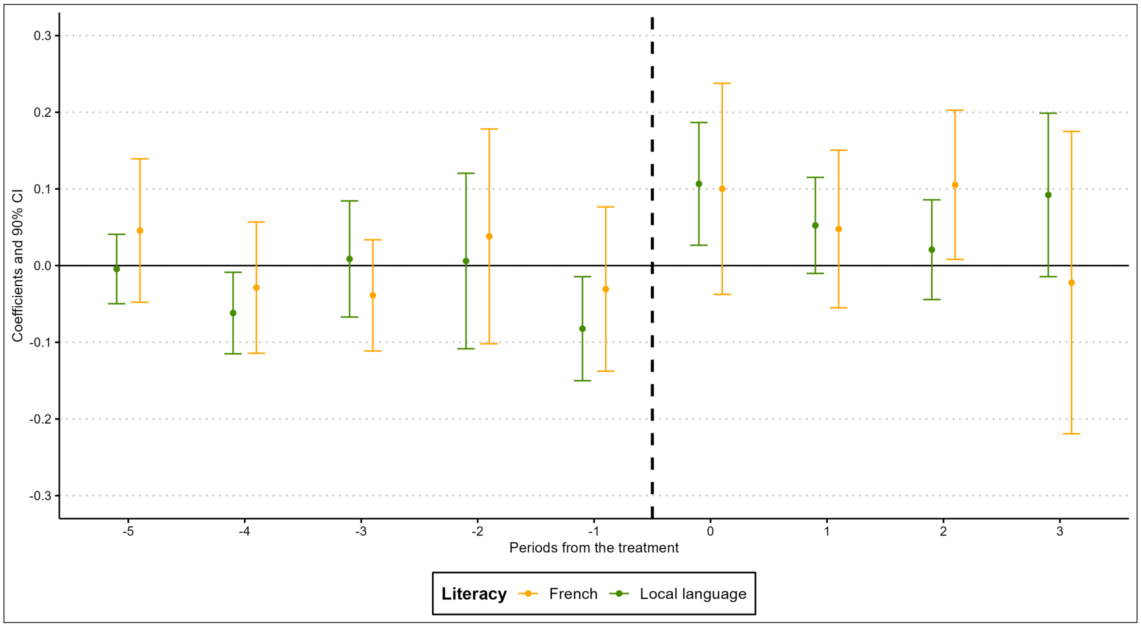
Notes: One black dot corresponds to one commune in the sample. The y-axis depicts the increase in the number of schools from 1998 to 2009 as given in the census, in percentage. The x-axis represents the share of bilingual schools among the total number of schools within the commune. The orange line is obtained from regressing the share of bilingual schools on the increase in schools. The shadowed area represents the 95% confidence interval of the estimated regression coefficient.

Figure C.2: Correlation between school building and bilingual education supply



Notes: BE stands for Bilingual Education. The districts highlighted in orange are districts with a relatively high share of bilingual schools in the total school supply. These orange districts are used for performing the analysis described in Section 7.3.

Figure C.3: Map of districts considered in the analysis described in Section 7.3



Notes: Point estimates and 90% confidence intervals are derived from a staggered analysis using Borusyak, Jaravel, and Spiess (2024). The dotted black vertical line indicates the last pre-treatment period. The two sets of points and coefficient estimates are obtained through two separate regressions. The treatment is estimated at the district level. I use region-fixed effects, and I cluster at the LSMS cluster level, which corresponds roughly to the village level.

Figure C.4: Staggered analysis on French and local language literacy

C.2 Tables

	Local language		French		Attended	Primary school
Dependent Variables:	Writing	Reading	Writing	Reading	school	diploma
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Born after 1994	0.017	0.015	0.142***	0.140***	0.160***	0.076***
	(0.013)	(0.012)	(0.020)	(0.020)	(0.019)	(0.017)
High exposed commune	-0.0004	-0.0007	-0.003	-0.004	-0.012	-0.006
	(0.015)	(0.014)	(0.037)	(0.036)	(0.032)	(0.031)
Born after 1994 × High exposed commune	0.019	0.022	0.050*	0.051*	0.058**	0.039
	(0.019)	(0.018)	(0.030)	(0.030)	(0.028)	(0.026)
Fixed-effects						
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Mean of Y	0.123	0.118	0.489	0.484	0.495	0.305
Observations	8,636	8,623	8,636	8,633	8,635	8,635
R^2	0.00695	0.00671	0.11743	0.11475	0.08517	0.09628

Clustered (LSMS cluster) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. The level of analysis is at the individual level, i.e. one observation is one person. The differences in the number of observations is related to missingness of the outcome of interest in the raw LSMS data. I use region fixed-effects, as the treatment uses the regional median as a reference point. I use the weights given in the LSMS data, and cluster the standard errors at the enumeration area. Clustering at the treatment level, i.e. the commune, produces the same results.

Table C.1: TWFE results with region fixed-effects

Dependent Variables:	French literacy	Local language literacy	Attended school	Primary school diploma	French literacy	Local language literacy	Attended school	Primary school diploma
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Born after 1994	0.206***	0.019	0.194***	0.089***	0.110***	6.54×10^{-5}	0.142***	0.066***
	(0.028)	(0.020)	(0.039)	(0.021)	(0.024)	(0.014)	(0.021)	(0.021)
High exposed commune	0.079**	0.009	0.076	0.076**	0.015	0.039*	-0.007	0.006
	(0.038)	(0.028)	(0.050)	(0.035)	(0.063)	(0.022)	(0.049)	(0.053)
Born after 1994	0.011	0.022	0.020	0.003	0.060	0.046**	0.076**	0.065*
× High exposed commune	(0.040)	(0.029)	(0.047)	(0.032)	(0.038)	(0.022)	(0.034)	(0.034)
Sub-sample	Close	Close	Close	Close	Far	Far	Far	Far
Fixed-effects								
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	3,064	3,064	3,064	3,064	5,572	5,572	5,571	5,571

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. Literacy in one language is equal to one when the individual knows how to write and read in this language. The level of analysis is at the individual level, i.e. one observation is one person. The differences in the number of observations is related to missingness of the outcome of interest in the raw LSMS data. I use the weights given in the LSMS data, and cluster the standard errors at the enumeration area. Clustering at the treatment level, i.e. the commune, produces the same results. I define the sub-sample of commune far (close) to the the median when it has a bilingual education supply at least (below) 9 percentage points higher or lower than the median. 9pp corresponds to the median distance to the median. I use district fixed effects in all regressions.

Table C.2: TWFE results, by distance to the district median

Dependent Variables:	French literacy	Local language literacy	Attended school	Primary school diploma
Model:	(1)	(2)	(3)	(4)
Panel 1 – Heterogeneity variable: High school construction rate in the location				
HV	-0.081** (0.032)	-0.015 (0.019)	-0.099*** (0.033)	-0.062** (0.025)
DiD	0.109*** (0.035)	0.063*** (0.022)	0.091*** (0.034)	0.082** (0.032)
DiD x HV	-0.183*** (0.048)	-0.086*** (0.027)	-0.148*** (0.049)	-0.126*** (0.041)
Panel 2 – Heterogeneity variable: Respondent born in the commune				
HV	0.013 (0.018)	-0.034*** (0.012)	-0.013 (0.020)	-0.004 (0.018)
DiD	-0.012 (0.040)	-0.001 (0.026)	-0.030 (0.038)	0.002 (0.035)
DiD x HV	0.084** (0.041)	0.052* (0.028)	0.113*** (0.041)	0.050 (0.038)
Panel 3 – Heterogeneity variable: Attended school				
HV	0.774*** (0.016)	0.164*** (0.015)		
DiD	-0.011 (0.019)	0.0007 (0.015)		
DiD x HV	0.034 (0.024)	0.050** (0.025)		
Fixed-effects				
District FE	Yes	Yes	Yes	Yes

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. The level of analysis is at the individual level, i.e., one observation is one person. I use the weights provided in the LSMS data and cluster the standard errors at the enumeration area level. Clustering at the treatment level, i.e., the commune, produces the same results. "HV" stands for heterogeneity variable. "DiD" stands for the difference-in-difference interaction coefficient. I report only this coefficient, along with the heterogeneity variable coefficient, detailed in the first column, and the interaction coefficient between the DiD dummy and the heterogeneity dummy.

Table C.3: Heterogeneity results on school construction and migration

Dependent Variable:	Number of schools in the commune
Model:	(1)
Variables	
High exposed commune	-0.418
	(0.755)
2009 census	13.9***
	(3.42)
High exposed commune × 2009 census	-5.59*
	(2.77)
Fixed-effects	
District FE	Yes
Mean of Y	11.56
Fit statistics	
Observations	918
R^2	0.31844

Clustered (District FE) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The model estimated is a linear regression. The level of analysis is at the year-commune level, i.e., one observation is one commune at a given year. All data on the number of schools is taken from the 1998 and 2009 censuses. "2009 census" is a dummy equal to one when the number of schools is taken from the 2009 census, and zero when it comes from the 1998 census. It is equivalent to the dummy "Born before 1994" in Equation 2, as the number of schools in 1998 was before the 1999 PRODEC reform.

Table C.4: TWFE results on the number of schools

	Local language literacy	French literacy	Attended school	Primary school diploma
Model:	(1)	(2)	(3)	(4)
Panel 1 – Controlling for the number of school-age children per school				
Born after 1994	0.136***	0.006	0.156***	0.072***
	(0.020)	(0.011)	(0.019)	(0.016)
High exposed commune	0.019	0.014	0.012	0.016
	(0.041)	(0.015)	(0.034)	(0.034)
DiD	0.054*	0.039**	0.059**	0.041*
	(0.028)	(0.018)	(0.026)	(0.025)
Average number of	0.001	0.0009**	0.001	0.001
school-age children per school	(0.0010)	(0.0004)	(0.0009)	(0.0008)
Panel 2 – Controlling for the number of school-age children per school				
Born after 1994	0.144***	0.008	0.163***	0.079***
	(0.019)	(0.011)	(0.018)	(0.016)
High exposed commune	0.092***	0.036**	0.080***	0.077***
	(0.028)	(0.015)	(0.028)	(0.022)
DiD	0.047*	0.037**	0.052**	0.035
	(0.028)	(0.017)	(0.026)	(0.024)
Number of schools in the commune	0.002***	0.0004**	0.002***	0.001***
	(0.0003)	(0.0002)	(0.0002)	(0.0003)
Panel 3 – Adding linguistic areas fixed-effects				
Born after 1994	0.006	0.136***	0.155***	0.074***
	(0.011)	(0.020)	(0.019)	(0.016)
High exposed commune	0.021	0.018	0.019	0.015
	(0.016)	(0.040)	(0.034)	(0.034)
DiD	0.039**	0.054*	0.060**	0.040
	(0.017)	(0.028)	(0.027)	(0.024)
Linguistic area FE	Y	Y	Y	Y

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: “DiD” stands for the coefficient associated with the interaction between high exposed commune and being born after 1994. All dependent variables are binary outcomes. The model estimated is a linear regression. The level of analysis is at the individual level, i.e. one observation is one person. I use the weights given in the LSMS data, use fixed-effects at the district level, and cluster the standard errors at the enumeration area. Literacy in one language is equal to one when the individual knows how to write and read in this language. The average number of school-age children per school is obtained by taking the ratio of the total number of children between 7 and 12 in 2009 (census) divided by the number of schools. I define linguistic areas by using the main language spoken in the commune in the census.

Table C.5: TWFE results with additional controls

Dependent Variables: Model:	French literacy (1)	Local language literacy (2)	Attended school literacy (3)	Primary school diploma (4)
Variables				
Born after 1994	0.093***	0.001	0.129***	0.059**
	(0.026)	(0.018)	(0.024)	(0.024)
High exposed commune	-0.045	0.022	-0.048	-0.024
	(0.050)	(0.026)	(0.046)	(0.038)
Born after 1994	0.071**	0.006	0.059*	0.028
× High exposed commune	(0.034)	(0.022)	(0.032)	(0.030)
Fixed-effects				
District FE	Yes	Yes	Yes	Yes
Fit statistics				
Observations	5,104	5,104	5,104	5,104

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. Literacy in one language is equal to one when the individual knows how to write and read in this language. The level of analysis is at the individual level, i.e., one observation is one person. The differences in the number of observations are related to missingness of the outcome of interest in the raw LSMS data. I use the weights provided in the LSMS data and cluster the standard errors at the enumeration area level. Clustering at the treatment level, i.e., the commune, produces the same results. I define a commune as highly exposed to bilingual education when the number of bilingual schools in this location exceeds half of the total school supply. The sample is limited to districts highlighted in Figure C.3.

Table C.6: TWFE results using a 50% bilingual education supply cutoff

Dependent Variables:	Local language literacy	French literacy	Attended school	Primary school diploma
Model:	(1)	(2)	(3)	(4)
Variables				
Born after 1994	0.174***	0.011	0.199***	0.089***
	(0.029)	(0.015)	(0.029)	(0.021)
High exposed commune	0.160***	0.046**	0.148***	0.128***
	(0.037)	(0.020)	(0.037)	(0.027)
Born after 1994 × High exposed commune	0.063	0.033	0.046	0.035
	(0.040)	(0.025)	(0.040)	(0.036)
Fixed-effects				
District FE	Yes	Yes	Yes	Yes
Fit statistics				
Mean of Y	0.434	0.112	0.460	0.260
Observations	4,183	4,183	4,183	4,183
R^2	0.16701	0.07334	0.15884	0.13867

Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: All dependent variables are binary outcomes. The model estimated is a linear regression. The level of analysis is at the individual level, i.e. one observation is one person. The differences in the number of observations is related to missingness of the outcome of interest in the raw LSMS data. I use district fixed-effects. I use the weights given in the LSMS data, and cluster the standard errors at the enumeration area. Clustering at the treatment level, i.e. the commune, produces the same results. The sample is restricted to clusters located at least 2km from the borders of the communes.

Table C.7: TWFE results restricting the sample to communes far from the communal borders

Dependent Variables:	Born here (0/1)	Age at the school entry
Model:	(1)	(2)
Variables		
High exposed commune	0.004	-0.003
	(0.003)	(0.041)
Fixed-effects		
District FE	Yes	Yes
Fit statistics		
Mean of Y	0.989	6.60
Observations	7,727	4,121
R^2	0.00713	0.08627

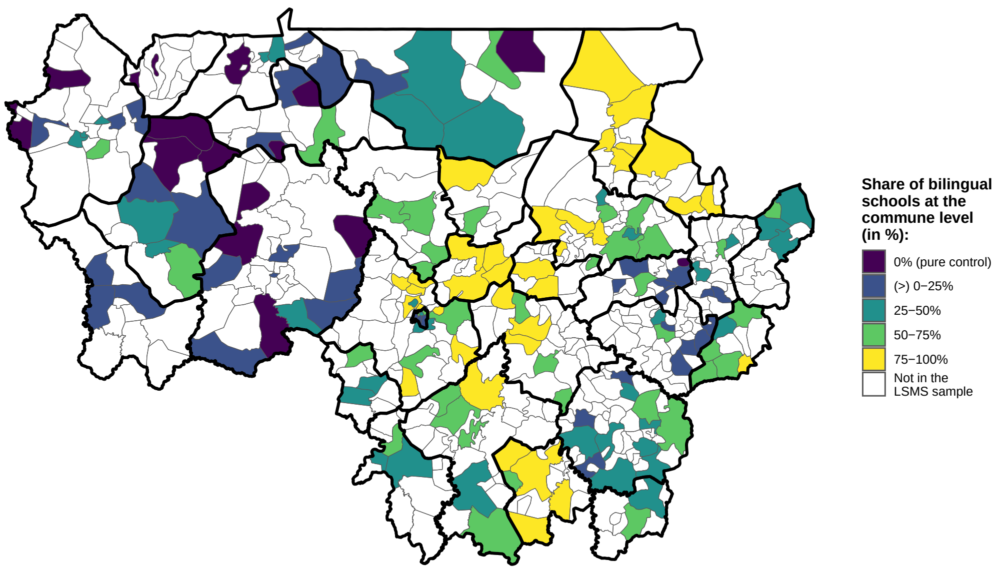
Clustered (LSMS cluster) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: I use a linear regression specification to obtain the coefficient estimates. I use district fixed effects, and I cluster at the LSMS cluster level, which corresponds roughly to the village level, and use household weights given in the LSMS survey. The mean of the outcomes is a weighted mean, using the same weights. The dependent variables are in Column (1) whether the individual was born in the surveyed village and in Column (2) the age when the individual started school for individuals who went to school (explaining the decrease in the number of observations).

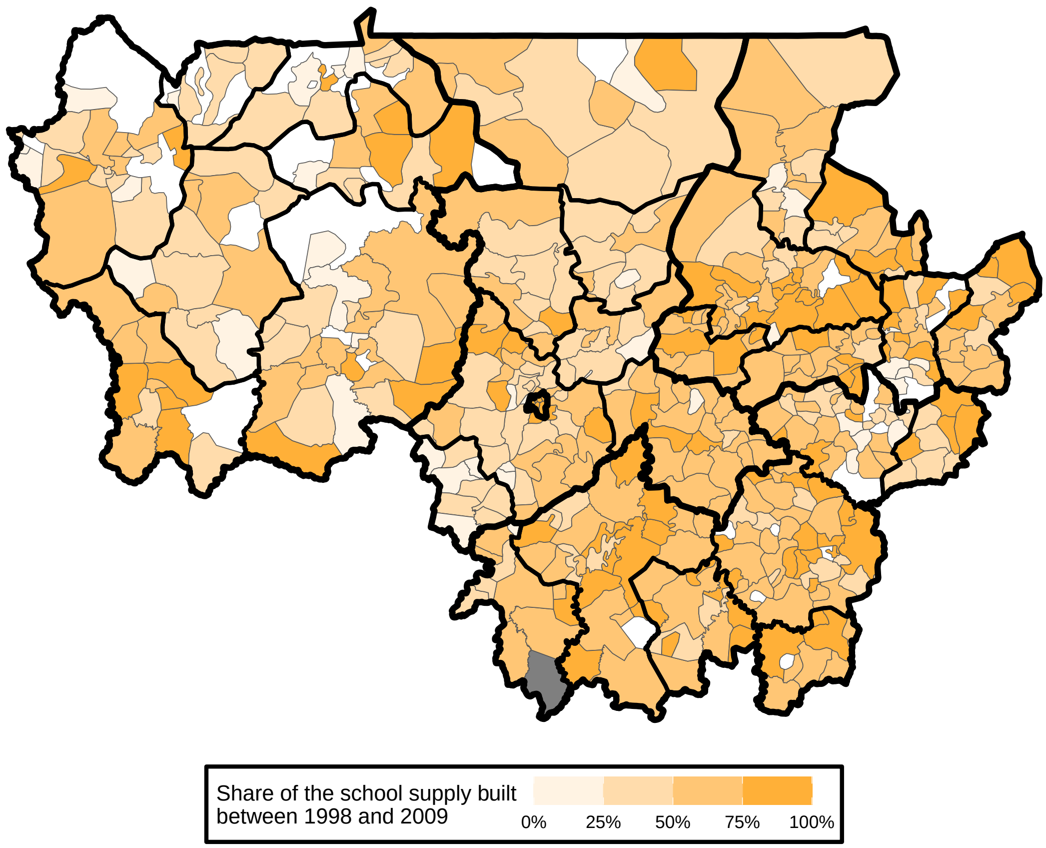
Table C.8: Correlation between treatment and potential confounders

C.3 Maps



Notes: Solid black lines represent the district borders, and gray lines the communal ones. Colors indicate the fraction of the bilingual education supply for communes in the LSMS sample.

Figure C.5: Map of binned bilingual education supply



Notes: Solid black lines represent the regional borders, dashed black lines represent the district borders, and gray lines represent the communal borders. The color represents the growth rate of the education supply at the commune level: the whiter the commune, the fewer schools were built between 1998 and 2009, in comparison to the existing school supply in 1998.

Figure C.6: Map of the education growth rate between 1998 and 2009 at the commune level

D Cost effectiveness analysis

Finally, I benchmark the estimated cost per student of bilingual education at scale in the case of Mali. To do so, I use World Bank reports to assess the implementation costs and get information on the target population (Bank, 2000). The total costs cover production and distribution of textbooks and teaching materials, the training of teachers, school principals, inspectors, and educational advisors, and the evaluation of students' and teachers' academic performance ($3.55 million). This cost is estimated for 400 bilingual schools. From the 1998 census, I infer that one-fourth of the primary school supply is affected by the reform at its onset in the early 2000s, representing roughly 350.000 students. To do so, I use the number of children between 7 and 12 from 2000 to 2002, as they are explicitly the targeted population in the report, i.e., individuals born between 1988 and 1995 in the 2009 census. The estimated cost per targeted student is then $10.6, well below the average cost of $17.97 of the most effective interventions in SSA, in the recent review done by Angrist et al. (2023).