Long-Lasting Effects of Exposure to Bible Translations: Evidence from Sub-Saharan Africa*

Vinicius Okada da Silva[1], M. Noelia Romero[2], Abigail Stocker[3], and Rebecca Thornton[4]

[1,3]Department of Economics, University of Illinois at Urbana-Champaign

[2]Department of Economics, Universidad de San Andres

[4]Department of Economics, Hankamer School of Business, Baylor University

December 5, 2023

Abstract

This paper evaluates the impact of early life exposure to Bible translations on education. To estimate causal effects and avoid issues related to selection into translation and mission location, we use the timing of translation relative to an individual's year of birth to compare educational outcomes across cohorts of individuals within language groups with and without exposure to a Bible translation in their mother-tongue language. We analyze data from a representative sample of approximately 75,000 adults in 13 sub-Saharan African countries using the Demographic and Health Surveys. Our difference-in-differences strategy accounts for the differential timing of Bible translations and the trends in educational outcomes over time within each language group. Individuals born ten years after the first Bible translation are 12 percentage points more likely to be literate later in life and acquire 1.3 additional years of education than those born before the translation. Effects do not vary by proximity to missions (either Catholic or Protestant), distance to a printing press, urban area, or religious faith. We provide the first causal evidence of the impact of Bible translations on education.

Keywords: literacy, Bible translations, languages, missions, ethnicity, cohort.

*We thank Marieke Kleemans, Mary Arends-Kuenning, and the seminar participants at the Applied Micro Research Lunch Seminar, Graduate Seminar at the Department of Economics (University of Illinois at Urbana Champaign), the Inequality, Demography, Health, and Religion (IDHR) workshop at Baylor University, the University of Maryland, and the Collaborative for Econometrics and Integrated Development Studies (CEIDS) seminar for their comments. We also thank attendees at sessions at the CIES 2023 Conference and seminar at the Department of Economics (Universidad de San Andres). Contact: Okada Da Silva: vo10@illinois.edu; Romero: marianr2@illinois.edu; Stocker: stocker3@illinois.edu; Thornton: Rebecca_Thornton@baylor.edu.

1. Introduction

A large literature has generally found a positive impact of Christian missions on educational outcomes, mainly due to the role missions have had in providing formal schools, access to the printing press, and hospitals, as well as new beliefs stressing the importance of literacy (Boppart et al., 2014; Calvi, Hoehn-Velasco, et al., 2022; D. Kim, 2020).[1] Some have argued that access to the Bible has played a role in the positive impact of Christianity on educational outcomes, but previous papers have been unable to disentangle the effect of exposure to the Bible from the larger confounding set of missionary and Christian influences (Becker and Woessmann, 2009; Moilanen and Sommerseth, 2021; Mosher, 2016; Feld, 2022).[2] The Bible could affect educational outcomes either through the availability of print material or through the ideas contained within it. In this paper, we examine the impact of one aspect of the Bible – translation into local languages – on long-term educational outcomes (adult literacy, primary school completion, and educational attainment) in Sub-Saharan Africa.

To evaluate the effect of exposure to Bible translations, we construct a dataset of Bible translations over time in Sub-Saharan Africa from the largest atlas of languages, the *Ethnologue* (David M. Eberhard et al., 2023), and combine it with information from historical ethnic group characteristics from the Ethnoatlas datasets (Bahrami-Rad et al., 2021; Kirby et al., 2016). These combined historical datasets allow us to describe the ethno-language group characteristics of languages with early translations before contact with colonial powers. We measure adult literacy through reading assessments, and our data on literacy and educational attainment comes from round 7 of the Demographic and Health Surveys (DHS) for men and women living in 13 Sub-Saharan African countries.

We estimate the causal effect of exposure to a Bible translation in one's mother-tongue on later-life education outcomes. Simply comparing individuals whose language was translated (or translated earlier) with those whose language was not translated would result in bias due to non-random selection into translation. Historical accounts point to missionaries choosing to translate the languages of the largest ethno-language groups first, for example. In Uganda, missionaries in the late 19th century decided to translate the Bible first into Luganda, the language of the largest ethnic group in Uganda, the Baganda people (Phares Mukasa. Mutibwa, 2016; Tuma and Phares Mukasa Mutibwa, 1978).[3] Using our data, we document empirical evidence

[1] Christianity has been found to be positively associated with educational outcomes (Odén, 1975; Guttormsson, 1990), such as reading (Moilanen and Sommerseth, 2021), and overall cognitive skills (Boppart et al., 2014; Fernihough and Henderson, 2015).

[2] Out of the numerous papers studying the impact of religion or Christianity on educational outcomes, only two papers explicitly examine the impact of the number of Bibles or Bible schools and find no overall effect, in Africa and China (Bai and Kung, 2015; Nunn, 2010). Those papers are somewhat limited by their identification strategy despite trying to control for selection into mission settlement. We utilize a cohort-exposure strategy.

[3] One of the first Protestant missionaries in Ethiopia in the 1840s, Johann Ludwig Krapf, translated the Gospels into the language of the Oromo people, who comprise the largest ethnic group in Ethiopia. Krapf did this because he believed that

of the non-random selection into Bible translation and show that ethno-language groups with earlier Bible translations rely more on cultivation and less on gathering, hunting, and fishing than groups with later or no Bible translations.

Given the selection into language translation, we turn to a novel identification strategy to measure the impact of exposure to a Bible translation by year of birth. Using an event-study difference-in-differences estimation strategy that accounts for the differential timing of Bible translation, we compare individuals across languages and cohorts. We compare the adult outcomes of cohorts with and without exposure to a Bible translation in their mother-tongue language, using the timing of the translation relative to the individual's year of birth. Our specification relies on variation across languages, years of Bible translation, and individuals' birth years. Our estimates account for the differential timing of Bible translations across languages and the increasing trends in literacy and educational attainment over time within each ethno-language group (Le Nestour et al., 2021; Callaway and Sant'Anna, 2021). Our results do not rely on the endogenous location of missions, since we combine cohort and time variation instead of using the geographic variation common in the existing literature that studies effects of Christianity based on mission location (Jedwab et al., 2022). To our knowledge, our data on Bible translations and our empirical approach are novel, making this paper the first to measure the effect of a Bible translation on literacy and education.

We find positive effects of exposure to a Bible translation on education. For individuals who are born approximately ten years after the first Bible translation in their mother tongue, we find an increase of 12 percentage points in the probability of being literate as an adult relative to individuals born before translation. We also find that exposure to a Bible translation results in an additional 1.3 years of schooling and an increase of 17 percentage points in the probability of completing primary school for those born ten years after translation relative to those born before translation. Overall, these findings are consistent across multiple specifications, including Sun and Abraham (2021) estimates, two-way fixed effects estimates, imputation estimates (Borusyak et al., 2021), and Callaway and Sant'Anna (2021) estimates, as well as across various sample restrictions.

Finally, we estimate the effects on the probability of holding Christian or Muslim beliefs, and we do not find any effects on the probability of adhering to Christianity or Islam. However, we do find some suggestive evidence of substitution toward Protestant beliefs and away from Catholic beliefs ten years after translation, though our estimated effects are not statistically different than zero and show some evidence of pre-trends.

The availability of the Biblical text in indigenous languages could affect literacy and education in several ways. First, most literacy acquisition occurs during primary school, approximately between the ages of

converting the Oromo to Protestantism would be key to the conversion of the rest of northeast Africa (Vilhanova, 2006). The same missionary started a Bible translation into Swahili, which was completed in 1891 would give another large ethnic group in East Africa access to the Scripture in their mother tongue (Vilhanova, 2006).

5-15 years old.[4] Increases in literacy could be caused by the increase in education provided by the first Christian schools in Africa (Wantchekon et al., 2015). Local-language translations of the Bible may have also resulted in additional investments by Christian missions such as additional schools, teachers, or text materials. Second, mother-tongue translations also contribute to the development of local orthographies, which in turn may have resulted in additional local-language print material availability. Literacy could increase through direct access to written texts and the desire to be able to read these texts. Finally, local language Bibles may change local beliefs, resulting in increased investment in education. We are unable to determine if the increases in literacy and education are due to increased quality or quantity of schools or simply increased investment. However, we do find that the results are consistent across proximity to missions and whether the region is predominantly Muslim, and effects are similar for men and women.

Our paper contributes to the literature in economic history regarding the long-term effect of Christian missions. Missionary activities are associated with a package of inputs such as new ideas and beliefs, schools, hospitals, Bible translations, and printing presses. We build on previous findings regarding the role of missionaries in human capital accumulation in Africa (Nunn, 2014; Cagé and Rueda, 2016; Becker, Rubin, et al., 2021; Wantchekon et al., 2015; Okoye and Pongou, 2023; Cappelli and Baten, 2017; Alesina et al., 2020; Cagé and Rueda, 2020; Huillery, 2009; Doyle et al., 2020), India (Calvi and Mantovanelli, 2018), China (Bai and Kung, 2015; Ma, 2021), and South America (Valencia Caicedo, 2018; Alston et al., 2022; Waldinger, 2017; Gómez-i-Aznar, 2022).[5] This paper also relates to the literature that analyzes the effects of access to the Bible and the printing press in medieval Europe (Becker and Woessmann, 2009; Rubin, 2014). Our paper contributes by measuring the impact of access to a text in one's mother tongue on literacy as a first step in human capital accumulation and is the first to study the direct impacts of Bible translations on human capital accumulation.

We are also the first to document quantitatively that the earlier translations of the Bible are positively correlated with ethnic group characteristics related to higher dependency on hunting and agriculture and lower dependency on gathering and nomadic behavior (Bai and Kung, 2015; Moilanen and Sommerseth, 2021; Nunn, 2010). This evidence of selection into translation of some ethno-language groups expands the understanding of group characteristics associated with current language survival and Christian mission expansion (Buzasi, 2015; Jedwab et al., 2022).

The remainder of the paper is organized as follows. In Section 2, we describe our contemporaneous and historical data and summarize the analytical sample. The empirical approach is described in Section 4,

[4] In the 1990s, 62.7 percent of women with 5 years of schooling are literate, while 76.4 percent of men with at least 5 years of education are literate in Sub-Saharan Africa (Le Nestour et al., 2021)

[5] Existing work find that exposure to missions increases Christian beliefs and attitudes in Africa (Nunn, 2010; Jedwab et al., 2018; Becker, Rubin, et al., 2021; R. D. Woodberry, 2012), as well as leading to inter-generational effects on trust in institutions (Wantchekon et al., 2015; Meier Zu Selhausen and Weisdorf, 2016; Alesina et al., 2020; Ricart-Huguet, 2022).

which begins with the definition of treatment exposure across cohorts within a language receiving a Bible translation, follows with our difference-in-difference specification, and explains the assumptions and threats to identification. In Section 5, we show our main findings regarding the impact of access to a Bible translation on educational outcomes and robustness. Section 6 discusses additional estimations and a heterogeneity analysis. Finally, in Section 7 we discuss the implications and conclude.

2. Data

This section introduces our contemporary and historical data sources and sets the stage with the background and context of African Bible translations. In our contemporary data, we measure our primary and secondary outcomes of interest: literacy (measured using a literacy test), education (measured through years of education and primary school completion), and religious affiliation. In our historical data, we obtain information on years and languages of Bible translations, precolonial characteristics of various ethnic groups,and the locations of Catholic and Protestant missions. We wrap up this section with a summary of the linked datasets and analytical samples.

2.1. Contemporary Data: Demographic and Health Surveys (DHS)

To study the effect of exposure to a Bible translation in one's mother tongue on adult literacy and educational outcomes, we use round 7 of the Demographic and Health Surveys (DHS) for 13 countries in sub-Saharan Africa.[6] We use data for individuals born between 1948 and 2005, containing a representative sample of married women between 15 and 49 years old in each country. Additionally, we use the DHS men's survey for data on the male members of the households sampled in the women's survey.[7]

To construct Bible translation exposure, we use the age and ethnicity of the respondent reported in the DHS.[8] Respondents classify themselves into one of 138 ethnic groups. We assign individuals to languages based on these ethnic groups, a process which is explained in more detail below. Because ethnicity is used to determine treatment status, we restrict our analysis to the 13 countries that contain data on ethnicity: Benin, Chad, Ethiopia, Gambia, Ghana, Guinea, Kenya, Malawi, Mali, Nigeria, Senegal, Sierra Leone, and Uganda.[9]

Our first outcome of interest is literacy. In all of the round 7 surveys, interviewers assessed reading ability

[6] The DHS round 7 was completed between 2014 and 2020 in our sample of 13 Sub-Saharan countries.
[7] The survey design means that the sample of men is not nationally representative of men between 15 and 59 years old.
[8] This variation is explained in further detail in Section 5.
[9] In all, the DHS sample includes 306,591 individuals in 13 countries where we have non-missing literacy, GPS data, and ethnicity. Figure B.1 shows the geographic distribution of the available DHS clusters in these 13 countries.

using a card with sentences in various languages.[10] The respondents chose which language they wanted to take the literacy test in, and they were asked to read the sentence in their chosen language. Literacy is reported in terms of whether an individual can read a whole sentence, can read part of a sentence, or is not able to read at all. Our main literacy outcome is full literacy, which we define as an individual's ability to read the entire sentence; we also present results on partial literacy, which we define as an individual's ability to read at least part of a sentence, in Appendix B.

We also estimate the effect of a Bible translation on educational attainment, measured in years of education completed and an indicator for whether primary school was completed, which are reported by each DHS respondent. In Appendix B, we also present the effects on whether an individual completed any years of education and whether an individual completed secondary school or higher.

In addition to variables related to literacy, we use information from the DHS on the reported religious affiliation of each individual. We combine all Christian beliefs into a single category, including Protestant, Catholic, and Orthodox beliefs, to create an indicator for being a Christian.[11] We use this variable to examine the impact of the Bible translations on Christian affiliation.

Finally, we exploit geographic variation using the Demographic and Health Surveys information on the locations of households by sampled cluster. The DHS has two types of GPS locations: urban (with a 2-kilometer radius) and rural (with a 5-kilometer radius). Each cluster contains, on average, 36 observations, ranging from a minimum of one to a maximum of 123 (see Figure B.2). We use this information to match respondents to their closest historical Catholic or Protestant mission.

2.2. Historical Data

In this subsection, we describe our historical data, beginning with data on Bible translations and the historical characteristics of ethnic groups with and without translations during the pre-colonial time. Then, we briefly describe the data on mission locations.

2.2.1. Bible Translations

Our data on Bible translations come from the 16th edition of the *Ethnologue: Languages of the World* by David M. Eberhard et al. (2023). This atlas has information on the year of the first Bible translation for each specific language, the portion of the Bible translated (New Testament or full Bible), whether the language is written or oral, and the status of the language, which indicates how widely the language is used.[12]

[10] Previous rounds of the DHS assumed literacy for individuals who had already completed secondary school or higher.

[11] The categories that we combine to represent Christian beliefs are Catholic, Methodist, Protestant, Orthodox, Christian, Pentecostal, Anglican, Presbyterian, Roman Catholic, Seventh Day Adventist, Baptist, Evangelical, and Jehovah's Witness.

[12] Note that the *Ethnologue* provides information on full Bible translations and New Testament translations. However, the information on New Testament translations is only provided if the language does not have a full Bible translation. In other

Figure 1 shows the timing of translation into different languages in Sub-Saharan Africa. Up until the 1980s, the number of Bible translations completed was limited. After the 1980s, we see a significant increase in the number of African languages translated. By the end of the 1990s, about a hundred languages in Sub-Saharan Africa had at least some portion of the Bible translated. Our analysis uses this variation in timing of modern Bible translations across languages to identify effects on educational outcomes, focusing on languages that received translations between 1980 and 1988 as our treated language groups.

Bible translations were not assigned randomly to languages, so we check for selection into translation. We use a database of precolonial characteristics of ethnic groups provided by Kirby et al. (2016), which combines cultural data from the *Ethnographic Atlas* by Murdock (1959), the Binford Hunter-Gatherer Dataset, the Standard Cross-Cultural Sample, and the Western North American Indians datasets (Lowes, 2020). The data is available for 1291 societies across the globe, including in our countries of interest in Africa. This data includes cultural traits for each society, and we focus on reliance on hunting, gathering, fishing, and agriculture; marital structure; agricultural intensity; major crop types; and settlement patterns.

To understand the selection into an earlier Bible translation, we divide our language sample into languages translated before 1970 (45 languages), languages translated after 1970 (86 languages), and languages that were never translated (7 languages).[13] Table B.1 summarizes the precolonial characteristics of these language groups by translation timing. For example, of the groups that received a translation prior to 1970, 77.8 percent of the groups were in the lowest category for gathering dependence. The table shows that ethno-language groups with earlier translations had lower dependence on gathering, hunting, and fishing; higher agricultural dependence; higher prevalence of polygamous marriage; and higher incidence of village or town settlements. These results provide evidence that there are systematic differences between languages that were translated earlier and later (or never translated), indicating selection of languages into translation. This means that a simple comparison of outcomes between ethno-language groups with earlier and later translations is likely biased and could potentially overstate any causal effect of Bible translation. To deal with this selection into early Bible translation, we exploit the ethnic group cohort-level treatment to estimate the impact of exposure to a Bible during schooling on adult literacy and education.[14]

Our main specification includes XXX later-translated languages, with all of the translation years falling between XXX and XXX. This section includes a brief overview of the process of translation for each of these languages, which are listed in Table XXX.

words, we have information on the first full Bible translation in each language if it is available. If there is no full Bible translation, the *Ethnologue* provides information on the first New Testament translation, if there is one.

[13]We use the year 1970 because is almost the middle point of the range of year of birth in our DHS individual data. Additionally, as shown in Figure 1, after 1980s there is a clear change in the trend of translations.

[14]In Appendix A, we present a simple exercise of exploiting only the language level and mission geographic variation to estimate the effect of the earlier Bible translations on literacy. After accounting for pre-colonial characteristics of the language groups, we find a positive association between Bible translations and literacy of 3 percentage points.

Many of the translations considered in our analysis were translated by Protestant groups. For example, the Lango and Acholi Bibles were published by the Bible Society of Uganda, the Somali and Sara Bibles were published by the Bible Society of Chad.

2.2.2. Missions

To understand one potential channel of the effect of Bible translations on education, we match combine the geographical locations of clusters in the DHS with the historical locations of Christian missions in Africa. We obtain data on mission locations from Nunn (2010) and Cagé and Rueda (2020). The information on the location of the missions in these two papers comes from three historical maps. Nunn (2010) digitized the map from Roome (1924), which also includes information on Bible translations.[15] Cagé and Rueda (2020) uses mission location on Protestant mission from *Geography and Atlas of Christian Missions* (Beach, 1903). The Catholic mission information used in Cagé and Rueda (2020) comes from official Vatican sources with a map by Streit (1913) in the 1913 *Atlas Hierarchicus*.

The locations of the 665 Catholic and Protestant mission stations are shown in Figure 2. The data from Nunn (2010) provides the sites of 394 missions, and the data from Cagé and Rueda (2020) provides the locations of an additional 271 missions. In total, we have 224 Catholic missions and 441 Protestant missions. The data from Cagé and Rueda (2020) also includes information on whether or not each mission has a printing press onsite. Although this data is only available for Protestant missions, in our heterogeneity analysis we estimate the effects of access to a Bible translation for individuals whose closest mission does not have a printing press.

2.3. Linking Datasets and Analytical Sample Summary

To examine the impact of exposure to a mother-tongue Bible translation, we link our contemporary data on educational outcomes and ethnicity data with mission locations, Bible translation year by language, and ethnic group characteristics before the arrival of colonial powers. We do this in a two-step process.

In step one, we use the package "Linking Ethnic Data from Africa" (LEDA) by Müller-Crepon et al. (2021) to match Bible translations with ethnic groups through languages. Each ethnic group in LEDA is associated with several languages. We match each language to data on the year of Bible translation from the *"Ethnologue: Languages of the World"* (David M Eberhard et al., 2021). This gives us several potential Bible translation matches for each ethnic group in the LEDA package. However, our analysis requires a one-to-one match between ethnic groups and Bible translations. To be conservative and obtain this match,

[15] Roome, William R. M. 1924. "Ethnographic Survey of Africa: Showing the Tribes and Languages; also the Stations of Missionary Societies [map]."

we choose the earliest Bible translation for each ethnic group if there are multiple potential matches.

In step two, we match each individual in the DHS to their respective ethnic group in LEDA, since the DHS does not directly report language. This results in a one-to-one match between ethnic groups in the DHS and Bible translations without losing observations from any ethnic group. For simplicity and as a result of this match, we use ethnic group and language group interchangeably in the rest of the paper.[16]

Next, we match mission locations to the DHS data using the GPS data for both the mission locations and the clusters to assign the closest mission to each cluster.[17] Finally, we match each ethnic group in the DHS to the ethnic groups in the Ethnoatlas to obtain information on the characteristics of each group. These two steps result in 138 ethno-language groups with matched Bible translation years, mission locations, and ethnic group characteristics.

Table 2 summarizes our final analytical sample after linking the individual DHS dataset with the other data sources. All of our 13 African countries are located in Central Africa except for Malawi. Column (1) shows the number of individuals surveyed by country, while column (2) shows the percentage of those individuals who are literate. Column (3) shows the number of languages per country. Notably, Chad has the highest number of languages within a country. Finally, column (4) displays the number of missions per country, ranging from only six missions in Chad to 189 missions in Nigeria.[18]

We restrict our main sample to language groups with observations in the DHS for a balanced panel cohorts born before and after the translation completed between 1979 and 1980.[19] Table 1 shows the languages that we use in our main specification and the years in which they received a Bible translation. In Table 1, we list the ten treated languages in Panel A and the 33 languages without a Bible in Panel B. Using languages that were translated between 1979 and 1988 allows us to compare cohorts born before and after the translation, as described in detail in our empirical approach.

3. Historical Missions and Selection into Translation

In this section, we describe the selection into translation across ethno-language groups. This section demonstrates that translation decisions are endogenous, raising concerns about simply comparing earlier-translated and later-translated languages and demonstrating the need for our empirical approach in the

[16] Figure B.3 shows the percentage of individuals in our DHS sample that speak those languages and have a Bible available in their mother tongue over time.

[17] It is possible that there are missions missing from our dataset, which could result in bias if the locations of the missing missions are correlated with translation timing.

[18] Figure B.3 shows the percentage of individuals in the DHS sample with access to a Bible translation in their mother tongue, by year of birth.

[19] For robustness, we also calculate the effects for the entire sample of languages, excluding only those that are always treated.

following section.[20]

In this approach, we exploit variation across ethno-language groups near a historical mission, comparing groups with and without access to mother-tongue Bible translations before 1970. Forty-five treated groups received a Bible translation before 1970, and 93 groups had no Bible translation available before 1970, as shown in Table B.2.[21] We estimate the following equation:

$$literate_{ilmcdt} = \beta Bible1970_l + \delta X_i + \alpha Z_l + \theta_m + \gamma_c + \phi_d + \alpha_t + \epsilon_{ilmcdt} \quad (1)$$

where indices indicate individual i, language l, mission m, country c, DHS cluster d, and cohort t. Our outcome, $literate_{ilmcdt}$, is either full or partial literacy. The treatment variable is $Bible1970_l$, equal to 1 if a Bible translation for language l is available before 1970. Individual characteristics, X_i, include male and rural indicators. Ethno-language group characteristics, Z_l, include gathering dependence, hunting dependence, fishing dependence, agriculture intensity, marital composition, major crop types, and settlement patterns. We include fixed effects for each mission θ_m, country γ_c, DHS cluster ϕ_d, and cohort α_t.[22] Finally, ϵ_{ilcmdt} is the error term. We are interested in β, which captures the correlation between of early Bible translations and literacy for individuals located near a mission. We cluster our standard errors at the language group level (Abadie et al., 2017).

We show our findings for specification (1) in Table B.3 and Table B.4. In column (1) of Panel A, access to a pre-1970 Bible translation is associated with an imprecisely estimated increase of 7.2 percentage points in the full literacy rate. In column (2), when we include indicators for male and rural and introduce fixed effects for DHS cluster, country, year of birth, and mission, the estimated coefficient becomes statistically significant but is smaller than in column (1). With these controls, access to an early Bible translation is associated with a three percentage point increase in literacy. Column (3) includes further controls for pre-colonial characteristics for each ethnic group, resulting in a less precise estimate of a three percentage point increase in literacy.

The results are consistent with positive selection into Bible translation.[23] The estimates in column (1) are likely to be upward-biased due to differential pre-colonial characteristics and are driven mainly by the largest ethnic group, as shown in Table 6. While Jedwab et al. (2018) and Michalopoulos and Papaioannou

[20] Previous papers on the impact of Christian missions find selection regarding mission locations (Becker and Woessmann, 2009; D. Kim, 2020; Bai and Kung, 2015; Waldinger, 2017; Wantchekon et al., 2015). Since access to a Bible translation is likely associated with distance from a mission, we reduce the potential for bias by narrowing our sample to respondents living within 10 km of a mission.

[21] This table summarizes the individual sample lost and the ethno-language groups from the DHS cluster located within 10 km of each mission.

[22] The mission fixed effects allow to capture the within-mission relationship between translations and literacy, accounting for time-invariant differences between missions.

[23] Another issue in the approaches that utilize mission locations comes from nonrandom measurement error of historical maps (Jedwab et al., 2018) (Fahs 1925).

(2013) argue that controlling for pre-colonial ethnic group characteristics as in column (3), corrects for bias, it is likely that this approach does not produce a causal estimate.[24]

4. Empirical Approach: Difference-in-Differences Strategy

In this section, we discuss our main empirical strategy for estimating the impact of access to a Bible translation on literacy and education-related outcomes. To identify the effects of access to a Bible translation on literacy and education, we make use of two dimensions of variation. First, we use time variation across birth cohorts within each language group. Second, we use cross-sectional variation at the language group level, comparing across languages that receive a Bible translation in different years. In this section, we explain how we use both the differential timing of Bible translations across languages and the differential treatment exposure across cohorts. We also present the specification, assumptions, and threats to identification.

4.1. Treatment Definition: Exposure to Bible Translation

Depending on when individuals are born and when the Bible is translated into their mother tongue, individuals will experience different levels of exposure to a Bible translation. We use two timing variables to define the treatment exposure of an individual to a Bible translation. First, we use the year of birth of each respondent by language group. Second, we use the year of translation of the Bible into the respondent's language.[25] We combine these two timing elements to create a variable that measures the exposure to a Bible translation. To calculate this exposure variable, which can take on both positive and negative values, we take the difference between the cohort's year of birth and the year of the Bible translation:

$$Exposure_{lt} = (BirthYear_{lt} - BibleYear_{l}) \tag{2}$$

In this equation, the indices indicate that the individual belongs to language group l and cohort (birth year) t. $Exposure_{lt}$ provides the number of years that the language group had access to a Bible translation at the cohort's year of birth. Negative values indicate that the language group had not yet received a Bible translation at the time of the cohort's birth (so negative values represent the length of time that passed after that cohort's birth and before translation).

In the event study framework, each of these possible values of $Exposure_{lt}$ represent a different time period

[24] These results are consistent with the findings by Jedwab et al. (2018) and Michalopoulos and Papaioannou (2013), who argue that the characteristics at the ethnic group level before the missionaries' arrival correct bias in the estimates of the impacts of missions on various outcome.

[25] One might be concerned that respondents from French and British colonies in Africa, speak French or English and had access to the Bible in those languages. In the analytical sample, only 0.3 percent of individuals chose to take the survey in English or French.

relative to the event of Bible translation. We restrict the event window to cover the 30-year time period from -15 to 14, meaning we include cohorts born from 15 years before the Bible translation up until 14 years after the translation. Older, untreated cohorts are those with negative values of $Exposure_{lt}$, while younger, treated cohorts are those with positive values of $Exposure_{lt}$. In other words, an individual is considered to be treated if she is born when there is already a translated Bible in his or her language, and the intensity of the treatment is based on how long the translation existed before the individual's year of birth. We estimate the treatment effect for each value of $Exposure_{lt}$, allowing us to identify differential impacts by level of exposure to a Bible translation.

For our main results, we restrict the sample to a balanced panel of treated ethno-language groups, which includes only ethnic groups that have individuals represented in the survey for all cohorts born 15 years before to 14 years after the Bible translation. This sample is balanced because we keep only languages with individual observations for cohorts at every level of exposure. We also include the never-treated groups without a Bible translation, which are not restricted to a certain time window since we cannot define an event time for those groups. Figure 3 illustrates the cohorts included for 3 languages, two of which were translated and one of which was never translated. The solid bars represent treated cohorts, while the hollow bars represent cohorts that are included in our analysis but are not treated (i.e., they have negative values of $Exposure_{lt}$). Our balanced sample consists of 43 languages in total.

4.2. Empirical Specification to Measure Effects

We evaluate the impacts of exposure to Bible translation across cohorts within language groups in a reduced form specification. Providing an event study framework, the two-way fixed effects (TWFE) equation is as follows:

$$y_{ilt} = \sum_{\alpha \neq -1} \beta_\alpha I(Exposure_{lt}) + \delta X_i + \mu_t + \lambda_l + \epsilon_{ilt} \tag{3}$$

Where the indices indicate individual i, language l, and birth cohort t. Our outcome of interest y_{ilt} is literacy and education outcomes (years of education and primary completion), and our secondary outcome is religious affiliation. In this specification, we include a fixed effect λ_l for the language group and a fixed effect μ_t for year of birth. X_i includes individual-level characteristics, namely indicators for males and for individuals living in urban areas.[26]

In our estimations, we use the indicator for $Exposure_{lt} = -1$ as the reference group, representing the group of individuals who were born one year before the Bible translation and were one year old at the time

[26] We do not include mission fixed effects, country fixed effects, or controls from the Ethnoatlas.

of the translation. The individuals included in the reference group likely already finished primary school around the time that the Bible translation was completed and distributed in their mother tongue. Thus, this group likely did not have access to a Bible translation when learning to read, since dispersion of the translation was not immediate.[27] As mentioned above, we also include never-treated ethno-language groups in our reference group. We rely on the assumption that literacy and educational attainment for individuals born before the translation is not influenced by access to a Bible translation. This assumption means that in the counterfactual with no Bible translations, individuals with all levels of treatment exposure would have had similar outcomes. The pre-trends in our event study model provides suggestive evidence that the underlying assumption for identification is not violated.

We estimate the β_α coefficients, which capture the effect of exposure to Bible translation separately for each level of exposure relative to the reference group. Point estimates for individuals born before the Bible translation will allow us to test for pre-trends in our sample, and point estimates for individuals born after the Bible translation will show the impacts of the translation on literacy and on educational attainment.

In summary, our reduced form difference-in-differences specification estimates the effect of exposure to a Bible translation in an individual's mother tongue based on their year of birth. Our identification comes from variation across cohorts and languages with different years of Bible translation. The individual's language, their year of birth, and the year of Bible translation into that language jointly determine their exposure to a Bible translation.

4.3. Assumptions and Threats to Identification

This approach uses an event-study estimation, and it is important to understand the assumptions that we are making when estimating the Equation 3. Since the event study is a generalized difference-in-differences design, the main assumption for identification is the parallel trends assumption. We assume that literacy trends over time between the treated and untreated languages are parallel. Intuitively, if there are compositional changes across cohorts, we assume that these appear in both the treatment and the control groups.

4.3.1. Staggered Treatment Timing

The model in Equation 3 is a naïve specification because it does not account for differential treatment timing, which is present in our model since the Bible was translated into different languages in different years. This differential treatment timing requires additional assumptions regarding parallel trends across cohort groups (Cunningham, 2021), namely that the treatment effects are constant for our language groups

[27] Note that if these individuals did have access to a translation while learning to read, this would tend to attenuate our estimates.

over time. In the context of differential treatment timing, TWFE estimates can be biased as a result of heterogeneity in treatment timing (Goodman-Bacon, 2021). Given these concerns regarding TWFE models, our main results use the approach found in (Sun and Abraham, 2021). These estimates correct for the differential timing of Bible translations across the treated ethno-language groups so that we are not improperly comparing earlier-treated and later-treated groups, essentially ensuring that post-treatment observations from earlier-treated groups are not included as a control.

Our results are robust to many different methods of estimation. In Appendix B, we show the TWFE estimates. Additionally, we calculate estimates using the semiparametric method by the Callaway and Sant'Anna (2021).[28] Finally, we calculate the effects using an imputation approach based from Borusyak et al. (2021). These results can all be found in Appendix B.

Despite the underlying assumptions necessary to this approach, using an event study design that exploits variation within and across language groups allows us to avoid the issues of endogenous selection into translation that would occur if we were simply comparing treated and untreated language groups.

4.3.2. Literacy Aquisition

An additional assumption in our model is that literacy acquisition occurs during the primary school years. In order for the Stable Unit Treatment Value Assumption (SUTVA) to hold, it must be true that individuals who were born before the translation do not change their literacy status as a result of access to the translation. This would happen if people attain literacy later in life. If, on the other hand, literacy is attained mainly in primary school, then SUTVA will be upheld.

There is evidence that most literacy acquisition in Sub-Saharan Africa takes place during primary school ages. Le Nestour et al. (2021) find that in the 1990s, 62.7 percent of women with 5 years of schooling are literate, while 76.4 percent of men with at least 5 years of education are literate in Sub-Saharan Africa. This suggests that individuals who attend primary school are gaining literacy. Additionally, an absence of differential pre-trends in our estimates suggests that individuals in the treatment group born before Bible translation were not impacted by access to the translation.

5. Main Results and Robustness

This section presents the results for the impact of exposure to a Bible translation on literacy and education outcomes and religious affiliation. Then, we present a series of robustness checks for these estimates.

[28]The Callaway and Sant-Anna (CS) estimates compare the treated and never-treated groups with semiparametric weights. However, these results are not very different from the naïve TWFE estimations, which are shown in Appendix B, indicating that differential treatment timing is not causing a large amount of bias.

5.1. Impact of Bible Translation on Literacy and Educational Attainment

Impacts on Adult Literacy

First, we find that access to a Bible translation in one's mother tongue increases adult literacy. We present the main results of the impact of exposure to a Bible translation on literacy in Figure 4. This figure shows the Sun and Abraham (2021) estimates for the impacts of access to a Bible translation on full literacy, defined as the ability to read a full sentence. The estimates account for differential timing of Bible translations and show no evidence of violation of the parallel trends assumption in the periods before the treatment. We find a positive effect of Bible translation on full literacy, which increases with exposure. These figures show that access to a Bible translation in one's mother tongue improves the probability of being literate by 12 percentage points for cohorts whose ethno-language group had access to a translation starting 10 years before the individual's birth.[29] By comparing across cohorts and across languages, these estimates account for the underlying trends of increasing literacy across all language groups in Africa (Le Nestour et al., 2021). This means that our results are not driven by increases in literacy over time across the entire sample, and we are able to identify the effect of access to a Bible translation separately.

We also calculate the TWFE, CS, and imputation estimates for the impact of access to a Bible translation on full literacy, and these are shown in Figure B.5. Table 3 shows the point estimates for the effect on full literacy for each of the four methods. The estimates are consistent with those from the Sun and Abraham (2021) specification, with increases in literacy ranging from 11 to 13 percentage points for individuals born 10 years after the translation.

This result of the impact of Bible translation on literacy is consistent with the literature on the effects of Christianity and missionary activity on human capital accumulation. Previous literature finds long-term effects of missionaries on literacy of around 3 to 10% (Calvi, Hoehn-Velasco, et al., 2022; Waldinger, 2017; Valencia Caicedo, 2018) in India and South America.[30] Meier Zu Selhausen and Weisdorf (2016) finds an increase of over 40% in literacy in Uganda when missionaries first arrived in the country. Although our estimation strategy and treatment are quite different from the previous papers, the estimated effects all indicate a positive effect of missionary activity on literacy. Additionally, the fact that our estimated treatment effect increases with exposure is not surprising, given that it would likely take some time for Bibles to become widely available after the translation.

Impacts on Adult Educational Attainment

[29] The results for partial literacy in Figure B.4 are similar, with flat pre-trends and a treatment effect that flattens out at around a 14 percentage-point increase.

[30] These results are similar to the long-term effects we find in Table B.2.

We also find that access to a Bible translation in one's mother tongue increases years of schooling and primary school completion. Figure 5 presents the impact of exposure to a Bible translation on years of schooling and primary school completion. The impacts on years of schooling are shown in Panel A and the impacts on primary school completion in Panel B. The effects on both outcomes are similar to the effects on literacy. For cohorts whose ethno-language group has been exposed to the translation for 10 years at the time of the individual's birth, the Bible translation increases schooling by around 1.3 years and the probability of completing primary school by 17 percentage points, as shown in Table 4. The positive effects of access to a Bible are also similar when we estimate the effect on an indicator for any education and on an indicator for completing secondary school or higher, as shown in Figure B.6.[31]

The impacts of exposure to a Bible translation on educational attainment are larger than the long-term effect of Christian missions on education found in previous papers. When using the ethnic group variation to estimate the effect of missions on years of education, Nunn (2014) finds an increase of 0.10 years, and similarly Okoye and Pongou (2023) finds an additional 2.1 years of education in areas more exposed to missionary activity in Nigeria. In contrast, as Jedwab et al. (2022) shows, effects estimated based on geographic variation of mission locations are smaller and not statistically significant once the selection into location is corrected for. Our results have the advantage of relying on the ethnic and cohort variation, not on the mission locations, to estimate the effect of Bible translation on literacy and education.

Impacts on Adult Religious Affiliation

This section discusses the impacts of access to a Bible translation on religious affiliation. Since Bible translations in Africa were completed with the purpose of increasing and spreading Christian beliefs, we also examine the direct effect of exposure to a Bible translation on the probability of holding Christian beliefs at the time of the survey. In Panel A of Figure 6, we present the effect of exposure to a Bible translation on current Christian affiliation. We find no effect of Bible translation on the probability of holding Christian beliefs for individuals born 10 years after the Bible translation. Similarly, we find no effect on the probability of holding Muslim beliefs, as shown in Panel B of Figure 6. The point estimates for these effects are shown in Table 5. Similar to Nunn (2010), who investigates the impacts of Bible translation on Christianity in Africa, the results suggest that access to a mother-tongue translation is not a key channel for the diffusion or adoption of Christian belief in Sub-Saharan Africa. As shown in the previous section, exposure to the Bible has a positive relationship with the human capital; these two main findings in Jedwab et al. (2022)

[31] One should be cautious when interpreting the impacts of completion of secondary school or higher. The younger cohorts in the sample are less likely to have finished secondary school, so the coefficient might be downward biased. Similar to the literacy outcome, the effect of access to a Bible translation on educational attainment is robust to various specifications. Figure B.7 and Figure B.8 show estimates for the impacts on years of education and primary completion using the TWFE, CS, and imputation approaches.

imply that the adoption of Christian belief is tied to economic development.

Reading the Bible is one of the cornerstones of Protestantism, promoted since the Reformation by Martin Luther (Becker and Woessmann, 2009). Additionally, recent translations into African languages in the 80s are Protestant Bibles.[32] One hypothesis is that the supply of reading material from Protestant missions increased the probability of Protestantism. We examine the relationship between current affiliation to Protestant Christianity versus Catholic Christianity and the access to Bible translations. Figure 7 shows the impacts on Catholicism in Panel A and Protestantism in Panel B and the point estimates are also shown in Table 5. While the estimates show evidence of pre-trends before the Bible translation for both affiliations, we document suggestive evidence of a slight decline in Catholic beliefs ten years after a Bible translation and a slight increase in Protestant beliefs. These results might imply substitution away from Catholic beliefs and towards Protestant beliefs, yet the pre-trend estimates for both Christian beliefs do not allow us to interpret them casually.

One explanation for these pre-trends is potential spillovers from Protestant to Catholic mission. While some historical reports suggest potential cooperation between both mission types to join efforts on translation (Vilhanova, 2006). Gallego and R. Woodberry (2010) documents competition between them when Catholic colonies restricted Protestant missionary activity. Thus, the effects on Protestantism and Catholicism might reflect the impacts of access to the religious text from the Protestant missions.

One concern regarding our main results is that Ethiopia and Mali are different from the rest of countries in our sample due to having earlier access to written languages and a long history of Christianity. We estimate the effects on literacy without the observations for these two countries in Figure B.9. This figure shows that the effects are nearly identical to our main results, meaning that observations from Ethiopia and Mali are not driving our results. The effects on years of schooling and primary school completion without Ethiopia and Mali are shown in Figure B.10. All of these estimates are generally similar to those in our main results.

6. Additional Estimations

This section provides information on the potential mechanisms by looking at the heterogeneity based on proximity to the nearest mission and based on gender.

[32] We briefly summarize the process of translation for our 12 treated language in the supplemental Appendix, available upon request.

6.1. Proximity to Missions

One potential issue regarding our estimates of the impacts of Bible translation is that missionary arrival and work around the time of the translation (such as providing increased access to schooling, literacy programs, etc.) may actually be driving the increase in literacy and education. These types of inputs and investments would likely be centered around a mission location. If these inputs were the primary drivers of the results, we would expect to see larger effects nearer to missions. To examine this, we estimate the effects of exposure to Bible translations separately for individuals living closer and further away from a mission location. Figure B.11 shows the impacts on literacy for individuals who live near (within 10 km) of a mission in Panel A and individuals who live farther (more than 10 km away) from a mission in Panel B. The pattern of effects are similar for individuals nearer and farther from missions, though the point estimates are slightly larger for individuals closer to a mission. The same is true for years of education in Figure B.12 and primary school completion in Figure B.13. Overall, the pre-trends for those closer to our mission locations are noisier around zero, yet the individuals far away from missions show very precise parallel pre-trends and small positive effects for those born ten years after a translation. This pattern suggests that access to the services and programs that the mission has to offer, such as schooling and literacy programs, is not the main driver of the increases in education outcomes.[33]

Theses results by mission proximity rule out the concern regarding whether the presence of missionaries in a region – required to complete the translations – is driving the results. In that case, we would expect to see effects *prior* to the completion of the translation; typically, it takes 7 years for the translation of the New Testament and 16 years for the translation of the entire Bible, on average. Since our effects are immediately following exposure to a Bible translation, missionary presence alone is unlikely to be driving our effects.

Additionally, while there is some anecdotal evidence of cooperation between Catholic and Protestant missions to translate the Bible together in Africa (Vilhanova, 2006), previous papers suggest a religious competition between the denominations (Gallego and R. Woodberry, 2010; Henn et al., 2021). In Figure B.16, we present estimates of the impact of Bible translation on literacy for only individuals whose DHS cluster was nearest to a Protestant mission. The results are similar to the main estimates, indicating that the results are not driven by a single type of mission. The same is true for impacts on years of schooling and primary school completion in Figure B.17 and Figure B.18.[34]

[33]The effect of a translation does not happen immediately but seems to be delayed until cohorts born around 10 years after the translation, which is possibly connected to the printing press and availability and distribution of the text, as indicated by (Cagé and Rueda, 2016). While the missionaries translate the Bible in the year indicated, it may take time before those translations can be reproduced at a low cost for use in schools. In Figure B.14, we show that the sample of DHS observation close to missions without printing press also results in positive effects on literacy. The same is true for years of schooling and primary school completion in Figure B.15. The effects are similar across those with and without access to the printing press, though perhaps slightly large for individuals located closer to missions.

[34]Note that we cannot directly estimate the effects for the subsample of individuals whose DHS cluster is closest to a Catholic

The similarities in effects across areas near and far from all missions and only Protestant missions are consistent with other studies that suggest larger effects of Protestantism on development and education (Bai and Kung, 2015; Cagé and Rueda, 2020; Dittmar, 2011).

6.2. Heterogeneity by Regional Religion

It is possible that the impacts of Bible translation are coming through an interaction with the impacts of Christianity. To illustrate, if access to a Bible translation increases the probability of being Christian, and those Christian beliefs direct individuals to value education more highly, then we should expect larger effect of exposure to translation in more Christian areas. To determine whether this is the case, we split the sample into geographical regions with high numbers of Muslims (>50% of individuals in the DHS in that region) and regions with low numbers of Muslims (<50% of individuals in the DHS in that region). We then estimate the effects separately for the two types of regions.

We find that the effects of access to a Bible translation on literacy are similar in both types of regions, as shown in Figure B.19. Additionally, the effects on years of education in Figure B.20 and the effects on primary school completion in Figure B.21 are similar for regions that are and are not predominantly Muslim.

This suggests that the effects are not solely driven by interactions with Christianity. Even in areas where Christianity is less prevalent, we still find large effects on literacy and education.

6.3. Heterogeneity by Gender

We show the effect of access to the Bible on literacy by gender in Figure 8. Panels A and B show the impacts on women and men, respectively. The pre-trends for women are more precise than for men because women represent around two-thirds of the DHS sample. More importantly, the effect of exposure to Bible translation for individuals born ten years after the translation is similar for both groups. Similarly, Figure 9 and Figure 10 show that the increases in years of education and primary school completion are also similar for both men and women. Although the sample for men is limited and shows imprecise pre-trends, the effect of exposure to a Bible translation on primary school completion seems more immediate for men than for women, and the effects are larger after ten years of exposure to a translation. These results contrast with the documented evidence of the larger effect of Protestant missions on females found in Nunn (2014) Jedwab et al. (2022) and Calvi, Hoehn-Velasco, et al. (2022). Thus, access to a text in one's mother tongue seems to produce a similar effect on education investment regardless of gender.

mission, because that subsample is too small due to fewer Catholic missions.

7. Conclusion

Using data on adult literacy and educational outcomes in Africa and historical data on mission locations and Bible translation into vernacular languages, we explore the long-term effects of access to a Bible translation in one's mother tongue on literacy and education. First, we document the endogeneity of language translation, showing that language groups with early (pre-1970) Bible translations are less dependent on gathering and nomadic behavior and more dependent on hunting and agricultural activities. Second, we use a novel approach of comparing individuals across cohorts and languages, using variation in the timing of the Bible translation and individuals' years of birth. We show that access to a Bible in one's mother tongue increases literacy by 12 percentage points, years of schooling by 1.3 years, and primary school completion by 17 percentage points for individuals born 10 years after the first translation.

Amidst the challenges of discord, social division, and the search for common ground, this study finds that Bible translations have played a pivotal role in improving literacy and education in sub-Saharan Africa. This finding highlights that efforts to translate religious texts into mother-tongue languages resulted in long-lasting positive effects on education, pointing to the role of missionary and religious institutions' involvement in development activities and lasting effects in improving education for all. These effects on literacy and education persist regardless of proximity to a mission location, regardless of gender, and regardless of whether the region is or is not predominantly Muslim.

Our findings align with evidence from other fields, such as history, religion, and education. Existing work finds that human capital accumulation and higher educational achievement for individuals from different ethnic groups around the missions depend on the original language structures. For instance, Chen (2013), Galor et al. (2021), and J. Kim et al. (2017) find that languages with periphrastic future tense are associated with higher educational attainment, more saving, more wealth, better health outcomes, and higher prevalence of earnings management. Our analysis considers the precolonial characteristics, because the ethnic group features are determinants of missionary contact and later survival of languages (Buzasi, 2015; Jedwab et al., 2022). Unfortunately, we do not have information in our dataset collection on the structure of the languages that would allow us to predict the probability that a language is translated in the first place, let alone predicting whether Bible translation increases the likelihood of being literate. However, future research can help understand which ethnic groups are more predisposed to learn to read from the Bible depending on their language structure and proximity between languages, as shown by Ginsburgh and Weber (2020).

Finally, ethnic groups may have incentives to adopt either the dominant language from the European colonial power or other languages to increase economic interactions with the developed world. The labor and

migration literature shows that language distance, which quantifies the dissimilarity between two languages, can increase communication costs, lower workplace productivity, and decrease trade between the two societies (Dale-Olsen and Finseraas, 2020; Heller, 2014; Isphording and Otten, 2013). Thus, it is probable that African ethno-language groups had economic incentives to interact with the missionaries and learn from those Bible translations to improve their relationship with Europeans.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge (2017). "When should you adjust standard errors for clustering?" arXiv: 1710.02926.

Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou (2020). "Intergenerational Mobility in Africa". In: 25534, pp. 1–65.

Alston, Lee J, Marie Christine Duggan, and Julio Alberto Ramos Pastrana (July 2022). "The Spanish Mission Legacy on Native American Reservations".

Bahrami-Rad, Duman, Anke Becker, and Joseph Henrich (2021). "Tabulated nonsense? Testing the validity of the Ethnographic Atlas". In: *Economics Letters* 204, p. 109880. ISSN: 01651765.

Bai, Ying and James Kai Sing Kung (2015). "Diffusing knowledge while spreading god's message: Protestantism and economic prosperity in China, 1840-1920". In: *Journal of the European Economic Association* 13.4, pp. 669–698. ISSN: 15424774.

Beach, H. P. (1903). *A Geography and Atlas of Protestant Missions: their environment, forces, distribution, methods, problems, results and prospects at the opening of the twentieth century (Vol. 2)*. New York: Student Volunteer Movement for Foreign Missions.

Becker, Sascha O., Jared Rubin, and Ludger Woessmann (2021). *Religion in economic history: a survey*. October 2019. Elsevier Inc., pp. 585–639. ISBN: 9780128158746.

Becker, Sascha O. and Ludger Woessmann (2009). "Was weber wrong? A human capital theory of protestant economic history". In: *Quarterly Journal of Economics* 124.2, pp. 531–596. ISSN: 00335533.

Boppart, Timo, Josef Falkinger, and Volker Grossmann (2014). "Protestantism and education: Reading (the bible) and other skills". In: *Economic Inquiry* 52.2, pp. 874–895. ISSN: 00952583.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (Aug. 2021). "Revisiting Event Study Designs: Robust and Efficient Estimation". In: arXiv: 2108.12419 [econ.EM].

Buzasi, Katalin (2015). "The historical determinants of language status in Sub - Saharan Africa". In: May.

Cagé, Julia and Valeria Rueda (2020). "Sex and the mission: The conflicting effects of early Christian missions on HIV in sub-Saharan Africa". In: *Journal of Demographic Economics* 86.3, pp. 213–257. ISSN: 20540906.

— (July 2016). "The Long-Term Effects of the Printing Press in Sub-Saharan Africa". In: *American Economic Journal: Applied Economics* 8.3, pp. 1–31.

Callaway, Brantly and Pedro H C Sant'Anna (Dec. 2021). "Difference-in-Differences with multiple time periods". In: *J. Econom.* 225.2, pp. 200–230.

Calvi, Rossella, Lauren Hoehn-Velasco, and Federico G. Mantovanelli (2022). "The Protestant Legacy Missions, Gender, and Human Capital in India". In: *Journal of Human Resources* 57.6, pp. 1946–1980. ISSN: 15488004.

Calvi, Rossella and Federico G. Mantovanelli (2018). "Long-term effects of access to health care: Medical missions in colonial India". In: *Journal of Development Economics* 135.July, pp. 285–303. ISSN: 03043878.

Cappelli, Gabriele and Joerg Baten (2017). "European Trade, Colonialism, and Human Capital Accumulation in Senegal, Gambia and Western Mali, 1770-1900". In: *Journal of Economic History* 77.3, pp. 920–951. ISSN: 14716372.

Chen, M. Keith (2013). "The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets". In: *American Economic Review* 103.2, pp. 690–731. ISSN: 00028282.

Cunningham, Scott (2021). *Causal inference: The mixtape.* Yale University Press.

Dale-Olsen, Harald and Henning Finseraas (2020). "Linguistic diversity and workplace productivity". In: *Labour Economics* 64.March, p. 101813. ISSN: 09275371.

Dittmar, Jeremiah E (2011). "Information technology and economic change: the impact of the printing press". In: *The Quarterly Journal of Economics* 126.3, pp. 1133–1172.

Doyle, Shane, Felix Meier zu Selhausen, and Jacob Weisdorf (2020). "The Blessings of Medicine? Patient characteristics and health outcomes in a Ugandan Mission Hospital, 1908–1970". In: *Social History of Medicine* 33.3, pp. 946–980.

Eberhard, David M, Gary F Simons, and Charles D Fennig (2021). "Ethnologue : Languages of Uganda". In: pp. 1–55.

— (2023). *Ethnologue. Languages of the World. Twenty-sixth edition.*

Feld, Elliot (2022). *Logos and Literacy.*

Fernihough, Alan and Stuart Henderson (2015). "Human Capital and Protestantism : Micro Evidence from Early 20th Century Ireland".

Gallego, Francisco A. and Robert Woodberry (2010). "Christian missionaries and education in former african colonies: How competition mattered". In: *Journal of African Economies* 19.3, pp. 294–329. ISSN: 09638024.

Galor, Oded, Ömer Özak, and Assaf Sarid (2021). "Linguistic Traits and Human Capital Formation". In: *SSRN Electronic Journal*, pp. 309–313.

Ginsburgh, Victor and Shlomo Weber (2020). "The Economics of Language". In: *Journal of Economic Literature* 58.2, pp. 348–404. ISSN: 00220515.

Gómez-i-Aznar, Èric (May 2022). "*Ad maiorem Dei gloriam* : Numeracy levels in the Guarani Jesuit missions". en. In: *Econ. Hist. Rev.*

Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics* 225.2, pp. 254–277. ISSN: 18726895.

Guttormsson, Loftur (1990). "The development of popular religious literacy in the seventeenth and eighteenth centuries". In: *Scandinavian Journal of History* 15.1-2, pp. 7–35. ISSN: 03468755.

Heller, Yuval (2014). "Language, meaning, and games: A model of communication, coordination, and evolution: Comment". In: *American Economic Review* 104.6, pp. 1857–1863. ISSN: 00028282.

Henn, Soeren, Horacio Larreguy, and Carlos Schmidt-Padilla (2021). "Missionary Activity, Education, and Long-run Political Development: Evidence from Africa". In: *Unpublished manuscript, Harvard University*.

Huillery, Elise (2009). "History matters: The long-term impact of colonial public investments in French West Africa". In: *American Economic Journal: Applied Economics* 1.2, pp. 176–215. ISSN: 19457782.

Isphording, Ingo Eduard and Sebastian Otten (2013). "The costs of babylon-linguistic distance in applied economics". In: *Review of International Economics* 21.2, pp. 354–369. ISSN: 09657576.

Jedwab, Remi, Felix Meier zu Selhausen, and Alexander Moradi (2018). "The economics of missionary expansion: evidence from Africa and implications for development".

— (2022). *The economics of missionary expansion: evidence from Africa and implications for development*. Vol. 27. 2. Springer US, pp. 149–192. ISBN: 0123456789.

Kim, Donghui (2020). "Church and Human Capital : Evidence from Korea during 1884-1930 I . Introduction II . Background". PhD thesis, pp. 1–19.

Kim, Jaehyeon, Yongtae Kim, and Jian Zhou (2017). "Languages and earnings management". In: *Journal of Accounting and Economics* 63.2-3, pp. 288–306. ISSN: 01654101.

Kirby, Kathryn R., Russell D. Gray, Simon J. Greenhill, Fiona M. Jordan, Stephanie Gomes-Ng, Hans Jörg Bibiko, Damián E. Blasi, Carlos A. Botero, Claire Bowern, Carol R. Ember, Dan Leehr, Bobbi S. Low, Joe McCarter, William Divale, and Michael C. Gavin (2016). "D-PLACE: A global database of cultural, linguistic and environmental diversity". In: *PLoS ONE* 11.7, pp. 1–14. ISSN: 19326203.

Le Nestour, Alexis, Laura Moscoviz, and Justin Sandefur (2021). "The long-term decline of school quality in the developing World". In: *Article submitted to a journal*, pp. 1–60.

Lowes, Sara (2020). "Ethnographic and Field Data in Historical Economics".

Ma, Chicheng (2021). "Knowledge Diffusion and Intellectual Change: When Chinese Literati Met European Jesuits". In: *J. Econ. Hist.*, pp. 1–46.

Meier Zu Selhausen, Felix and Jacob Weisdorf (2016). "A colonial legacy of african gender inequality? Evidence from Christian Kampala, 1895-2011". In: *Economic History Review* 69.1, pp. 229–257. ISSN: 14680289.

Michalopoulos, Stelios and Elias Papaioannou (2013). "Pre-Colonial Ethnic Institutions and Contemporary African Development". In: *Econometrica* 81.1, pp. 113–152. ISSN: 0012-9682.

Moilanen, Mikko and Hilde Leikny Sommerseth (2021). "'I will learn from it for as long as I live'–religious reading and functional literacy skills". In: *Scandinavian Economic History Review* 69.2, pp. 124–139. ISSN: 17502837.

Mosher, James S. (2016). "The Protestant Reading Ethic and Variation in Its Effects". In: *Sociological Forum* 31.2, pp. 397–418. ISSN: 15737861.

Müller-Crepon, Carl, Yannick Pengl, and Nils Christian Bormann (2021). "Linking Ethnic Data from Africa (LEDA)". In: *Journal of Peace Research*, pp. 1–11. ISSN: 14603578.

Murdock, George Peter (1959). *Africa: Its Peoples and Their Culture History*. New York: McGraw-Hill.

Mutibwa, Phares Mukasa. (2016). *A history of Uganda : the first 100 years, 1894-1995*. Kampala: Fountain Publishers. ISBN: 9789970253401.

Nunn, Nathan (2014). "Gender and missionary influence in Colonial Africa". In: *Africa's Development in Historical Perspective* February, pp. 489–512.

— (2010). "Religious conversion in Colonial Africa". In: *American Economic Review* 100.2, pp. 147–152. ISSN: 00028282.

Odén, B. (1975). "Läskunnighet och samhällsförändring [literacy and social change]". In: *Forskning om Utbildning* 1.

Okoye, Dozie and Roland Pongou (Sept. 2023). "Missions, fertility transition, and the reversal of fortunes: evidence from border discontinuities in the emirates of Nigeria". In: *J. Econ. Growth*.

Ricart-Huguet, Joan (2022). "The Origins of Colonial Investments in Former British and French Africa". In: *British Journal of Political Science* 52.2, pp. 736–757. ISSN: 14692112.

Roome, William R. M (1924). *Ethnographic Survey of Africa: Showing the Tribes and Languages; also the Stations of Missionary Societies [map]*.

Rubin, Jared (2014). "PRINTING AND PROTESTANTS: AN EMPIRICAL TEST OF THE ROLE OF PRINTING IN THE REFORMATION". In: *Rev. Econ. Stat.* 96.2, pp. 270–286.

Streit, K. (1913). *Atlas Hierarchicus: descriptio geographica et statistica Sanctae Romanae Ecclesiae tum occidentis tum orientis juxta statum praesentum accedunt nonnullae notae historicae necnon ethno- graphicae.* Paderborn: Typographiae Bonifacianae.

Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2, pp. 175–199. ISSN: 18726895. arXiv: 1804.05785.

Tuma, A D Tom and Phares Mukasa Mutibwa (1978). *A Century of Christianity in Uganda, 1877-1977: a historical appraisal of the development of the Uganda church over the last one hundred years*. Nairobi: Afropress.

Valencia Caicedo, Felipe (Oct. 2018). "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America*". en. In: *Q. J. Econ.* 134.1, pp. 507–556.

Vilhanova, Viera (2006). "Biblical Translations of Early Missionaries in East and Central Africa. I. Translations into Swahili." In: *Asian and African Studies* 15.1, pp. 80–89.

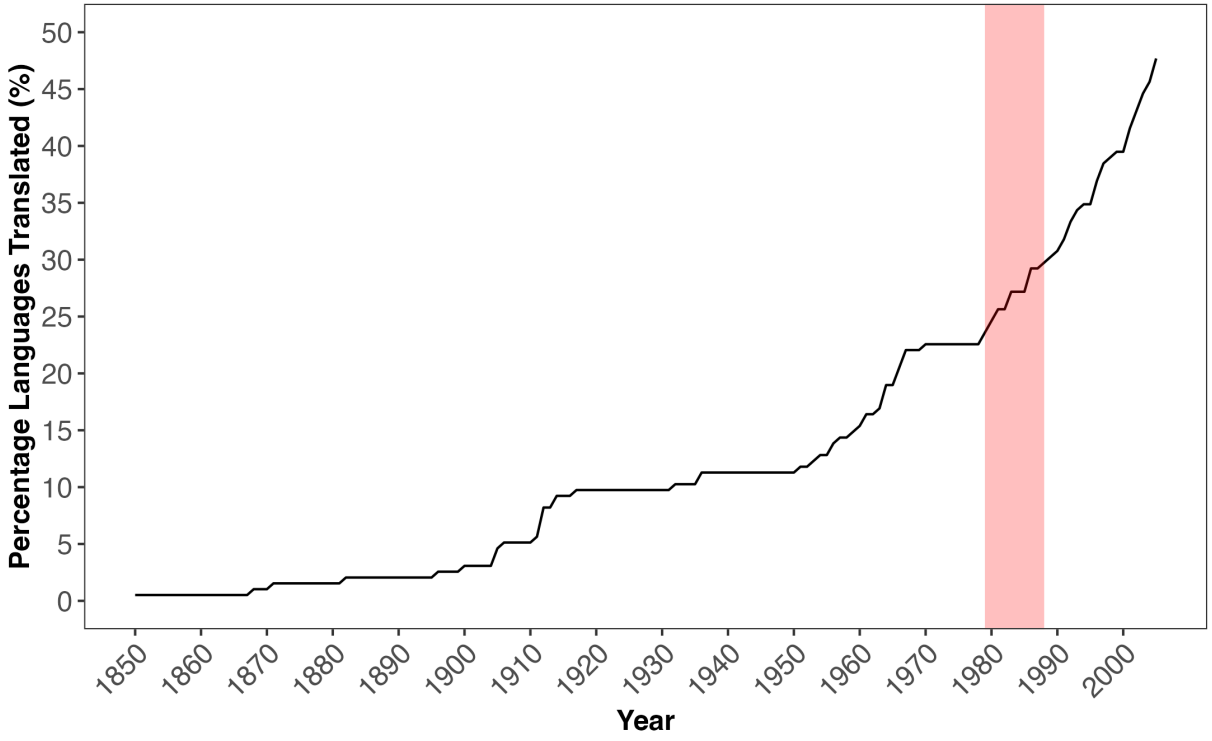
Waldinger, Maria (2017). "The long-run effects of missionary orders in Mexico". In: *Journal of Development Economics* 127.December 2015, pp. 355–378. ISSN: 03043878.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta (2015). "Education and human capital externalities: Evidence from colonial benin". In: *Quarterly Journal of Economics* 130.2, pp. 703–757. ISSN: 15314650.

Woodberry, Robert D. (2012). "The missionary roots of liberal democracy". In: *American Political Science Review* 106.2, pp. 244–274. ISSN: 00030554.

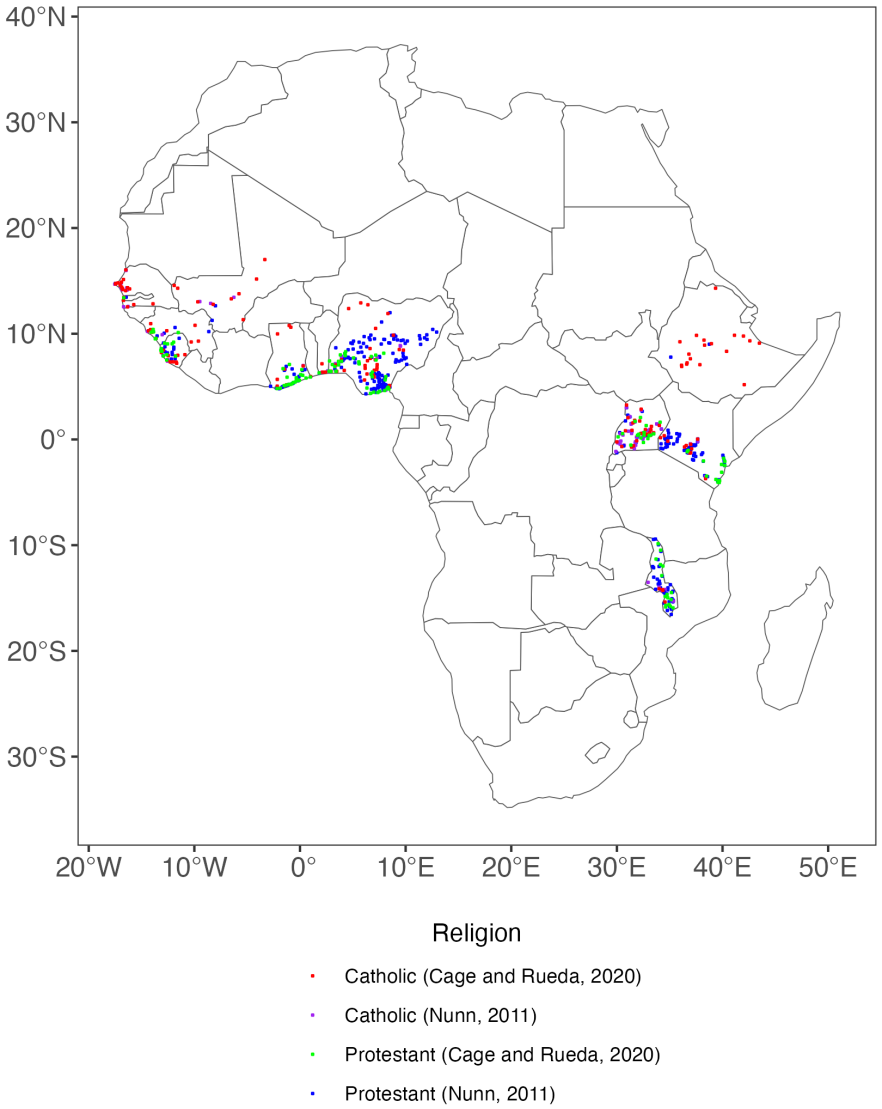
Figures

Figure 1: Distribution of Bible Translations in Sub-Saharan Africa



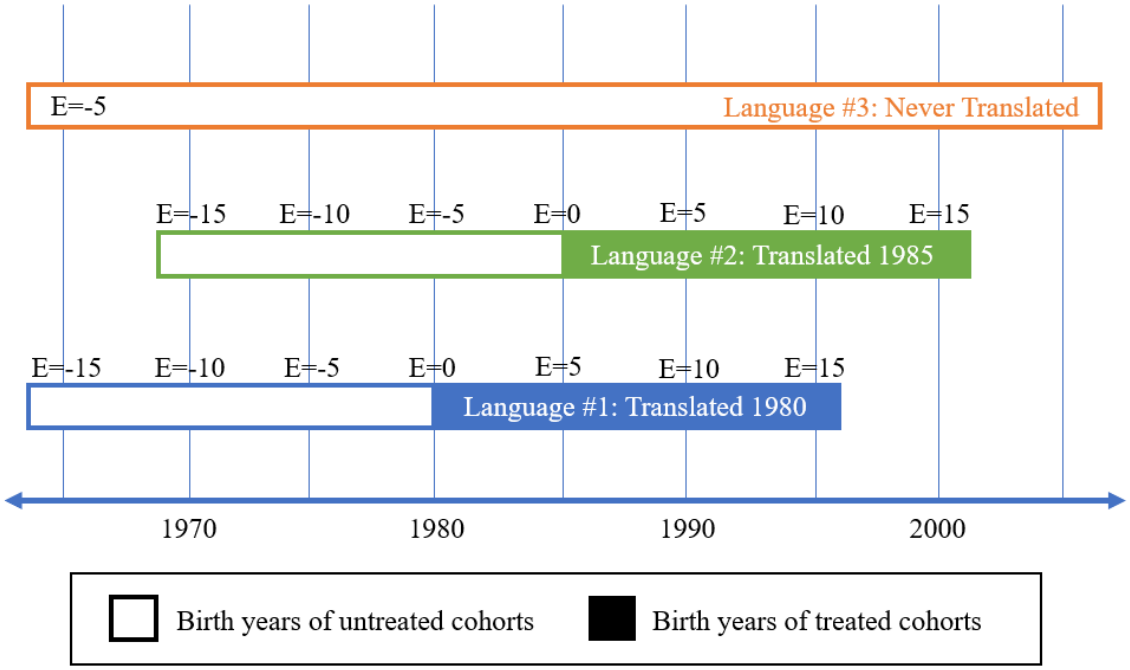
Notes: Distribution of Bible Translations across each decade and percentage of the sample that has a Bible in their own language. The red rectangle represents the years in which the sample in our analysis had a Bible translated.

Figure 2: Historical Missions Geographical Distribution



Notes: This figure shows the location of missions by religion and source of information. There are 224 total Catholic missions, from those 126 missions are from Cagé and Rueda (2020) and 78 missions from Nunn (2010). Similarly, there are 441 Protestant missions in total, where 146 missions are from Cagé and Rueda (2020) and 316 missions from Nunn (2010)

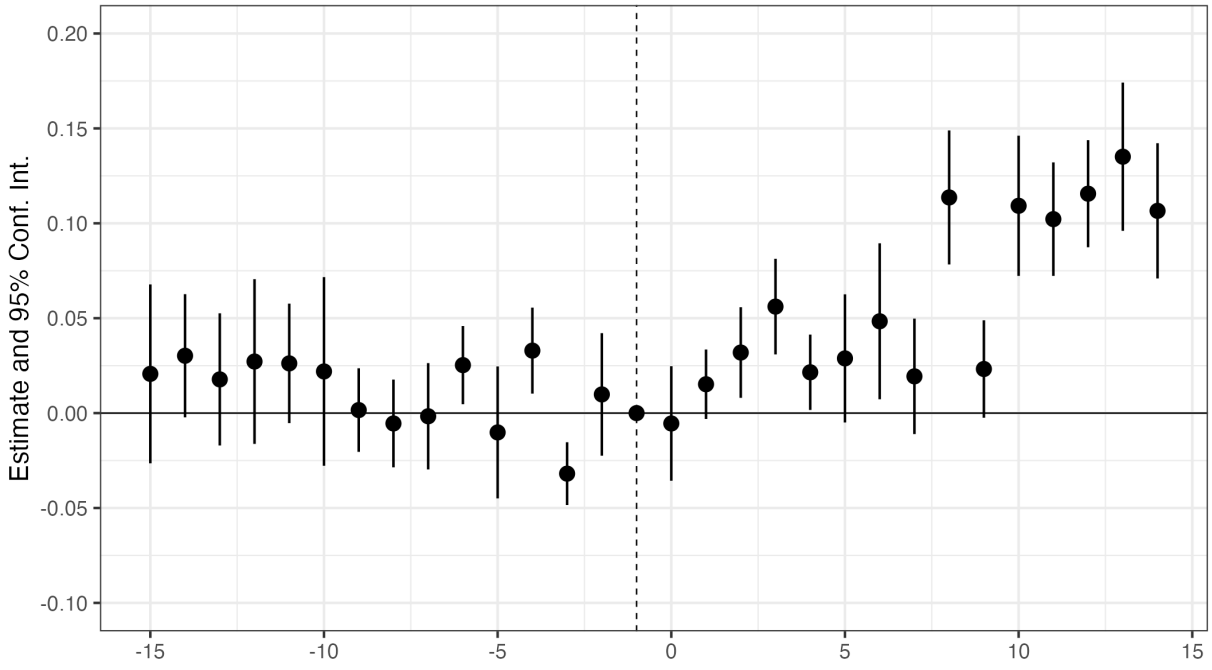
Figure 3: Description of Treatment



Notes: This figure shows the treated and untreated cohorts for three languages, one of which was translated in 1980, one of which was translated in 1985, and one of which was never translated. The solid bars represent cohorts with positive exposure to treatment, while the hollow bars represent cohorts that were not exposed to treatment. Note that for the language that was never translated, all cohorts become part of the reference group.

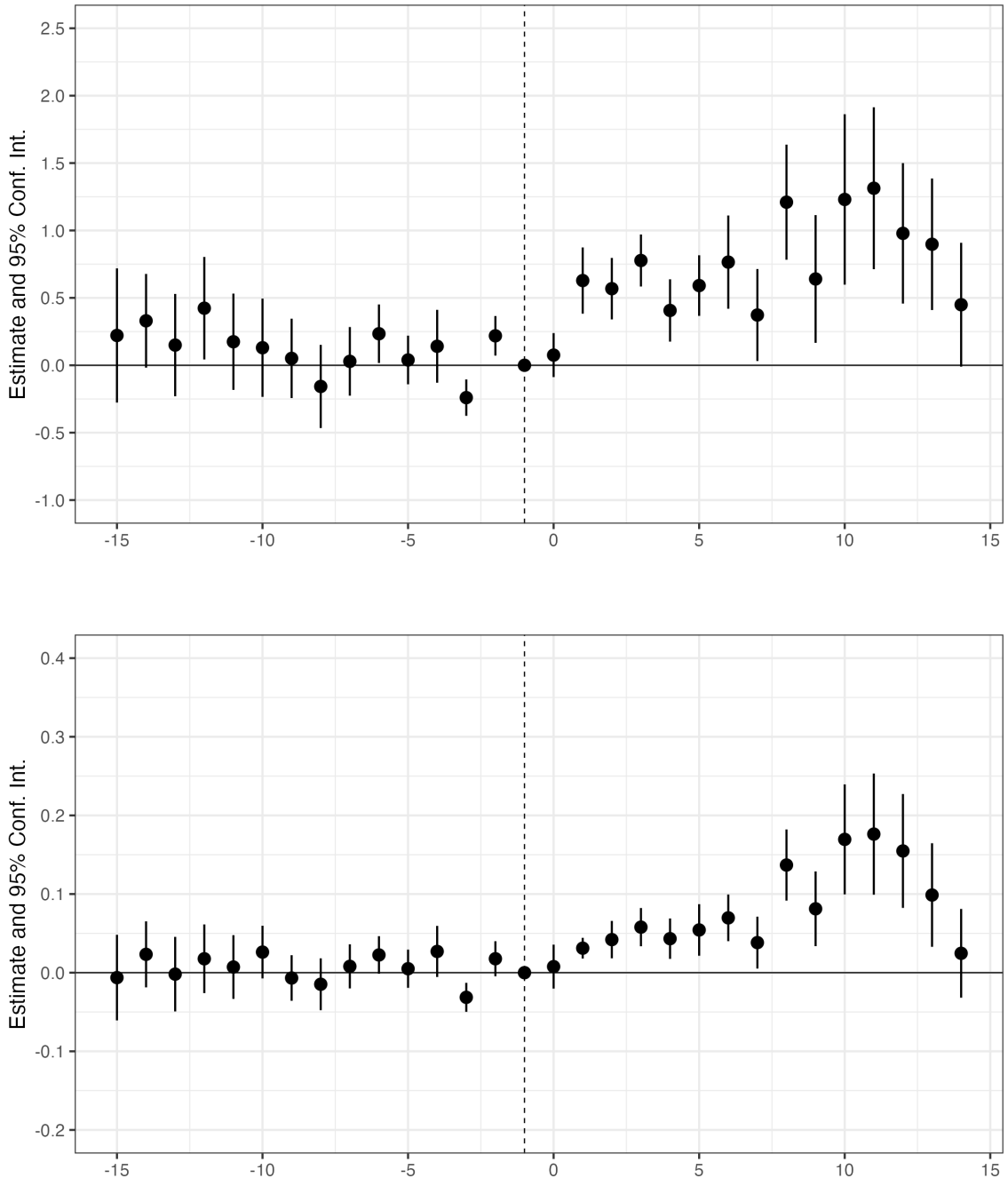
7.0.1. Results

Figure 4: Effects on Full Literacy



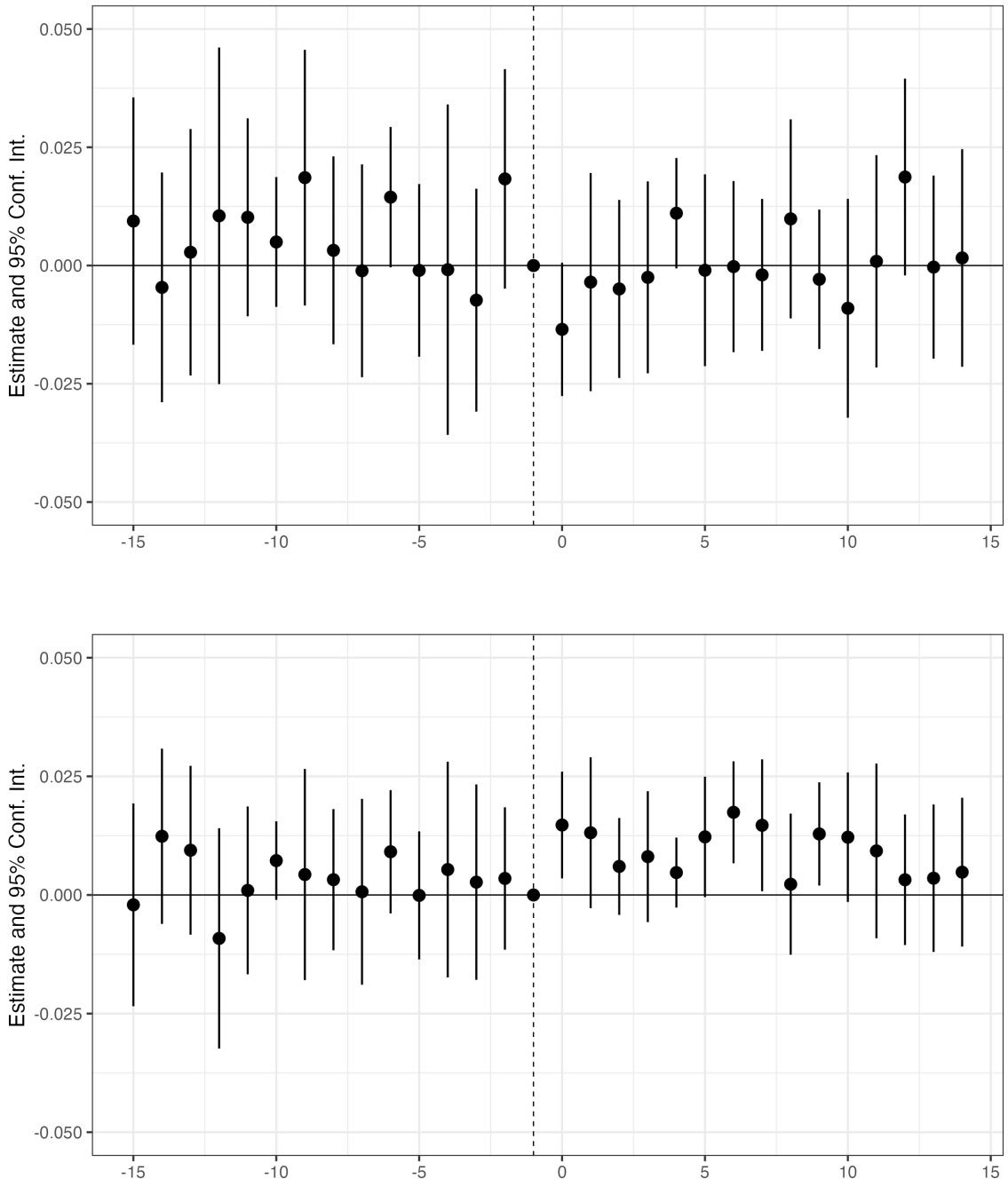
Notes: This figure shows the effect of exposure to Bible translation on literacy when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure 5: Effects on Years of Education and Primary School Completion



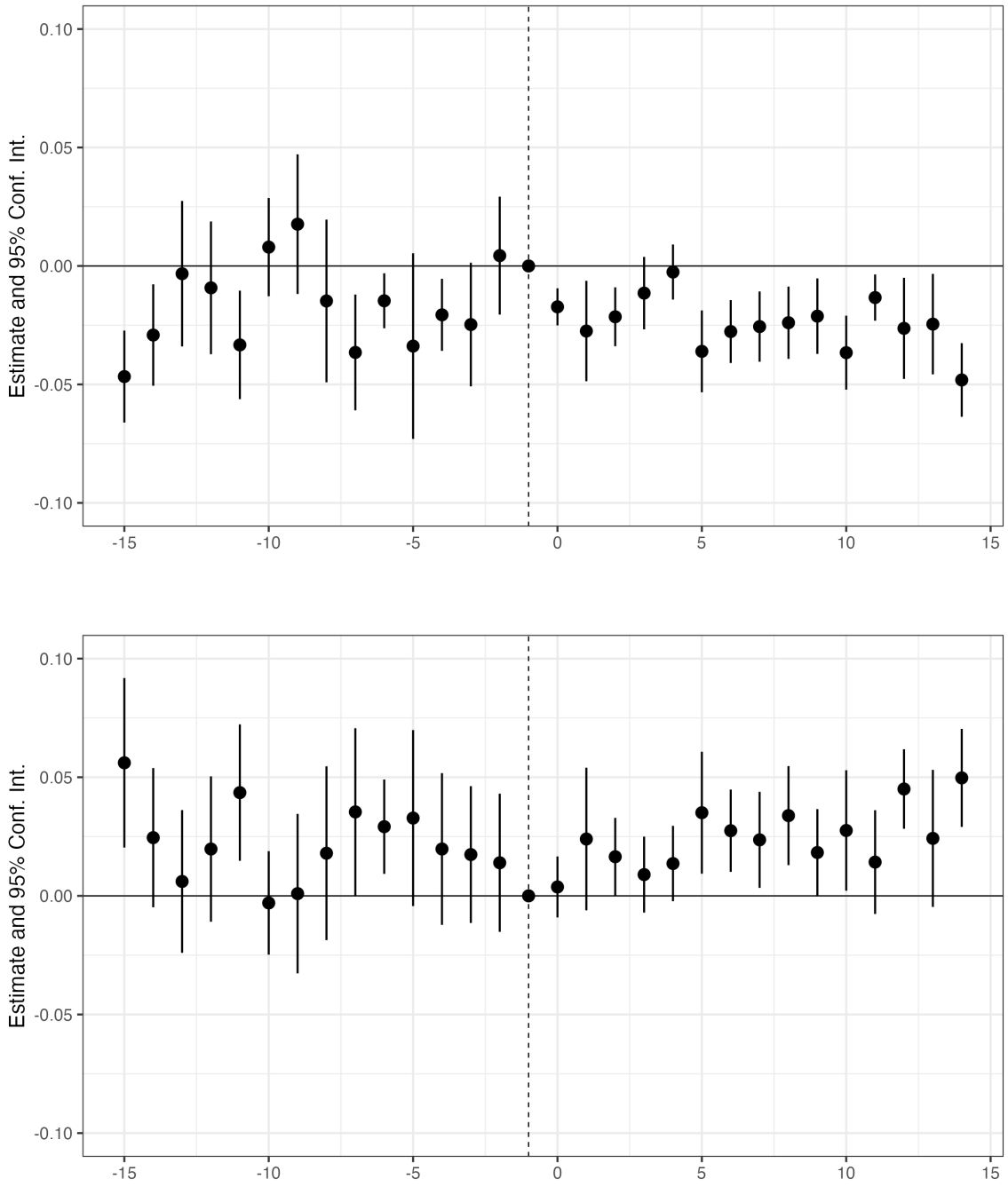
Notes: This figure shows the effect of exposure to Bible translation on years of education (Panel A) and primary school completion (Panel B) when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure 6: Effects on Christianity and Islam



Notes: This figure shows the effect of exposure to Bible translation on Christianity (Panel A) and Islam (Panel B) when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

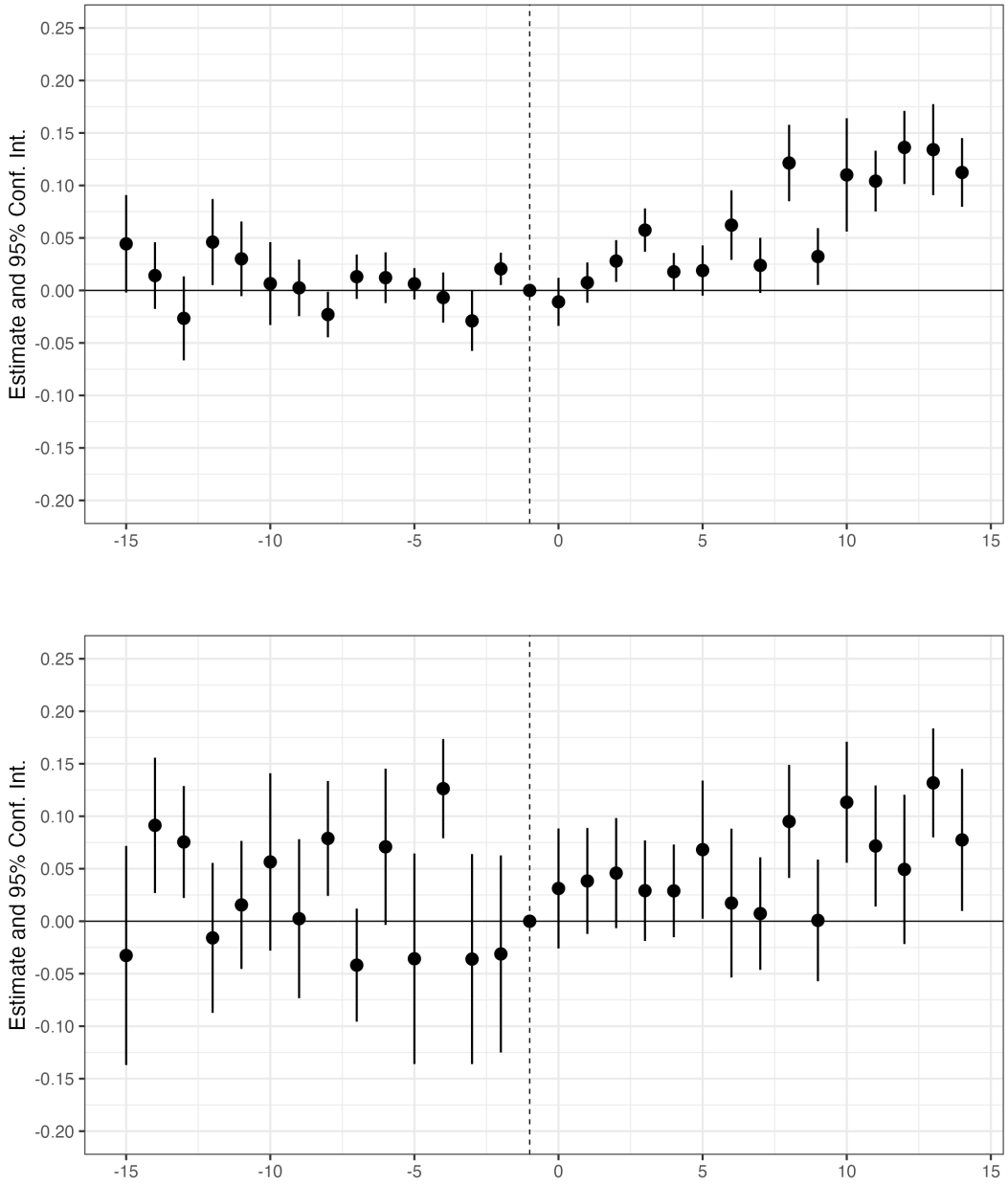
Figure 7: Effects on Catholicism and Protestantism



Notes: This figure shows the effect of exposure to Bible translation on Catholicism (Panel A) and Protestantism (Panel B) when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

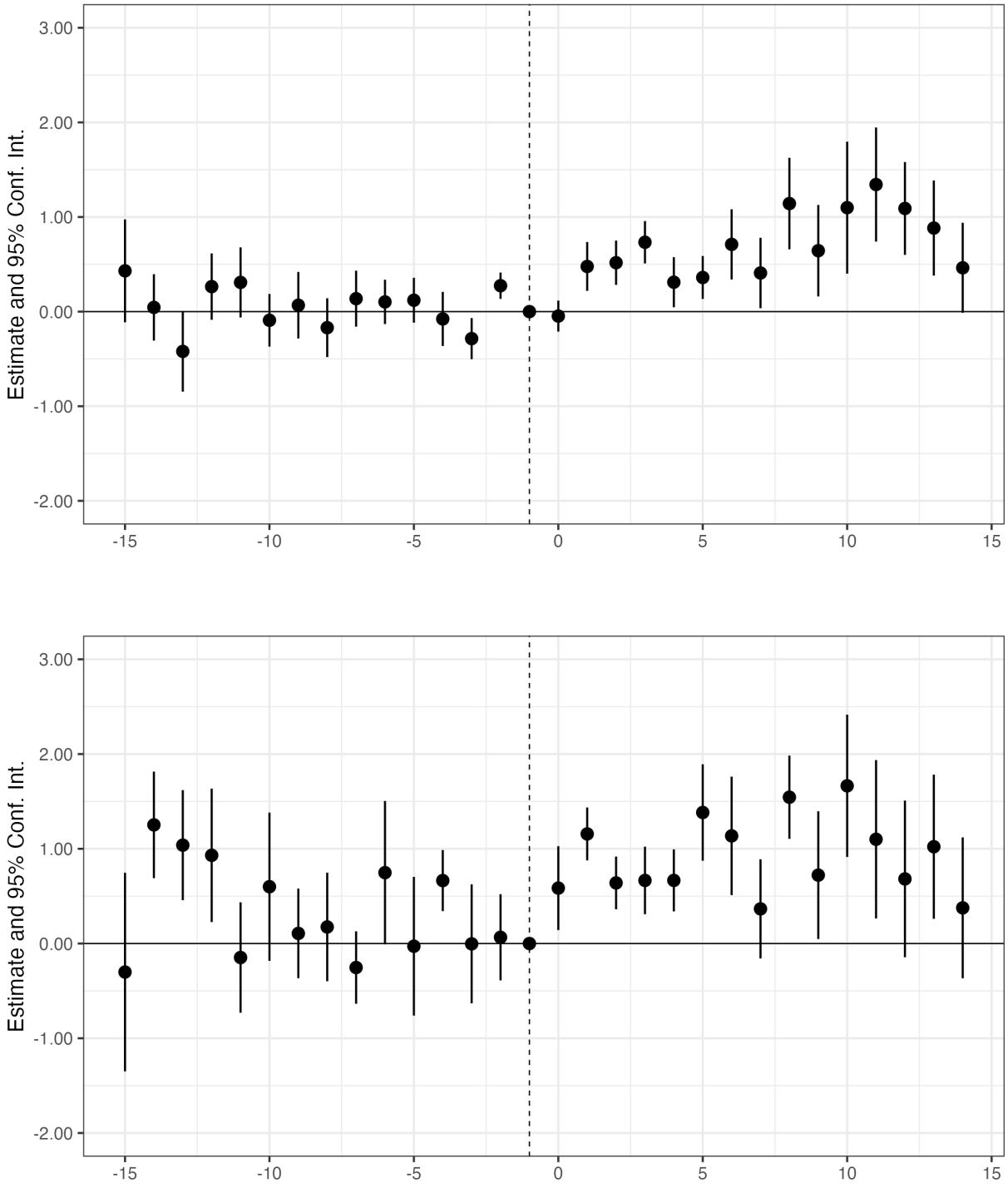
7.1. Heterogeneity

Figure 8: Effects on Full Literacy by Gender



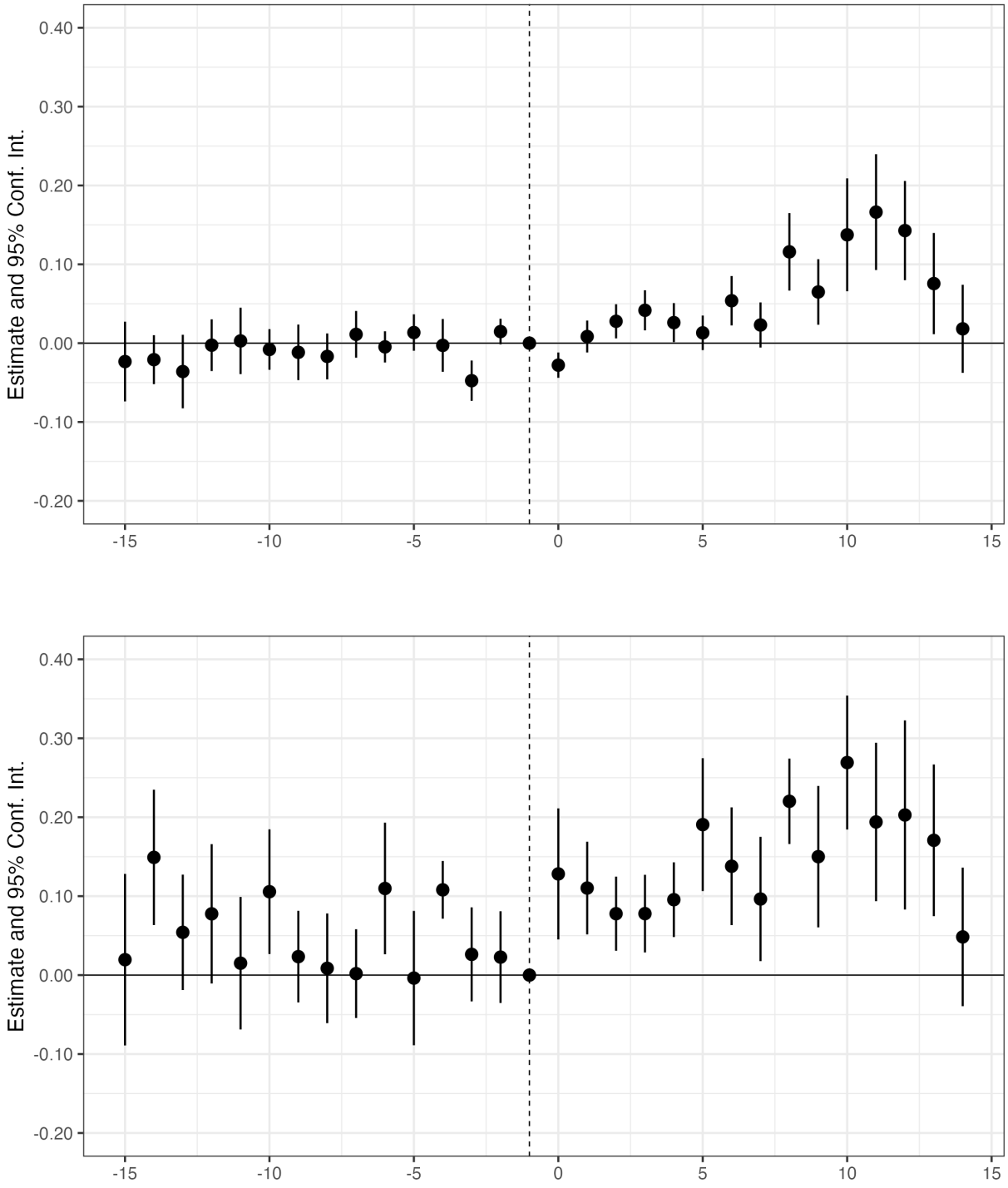
Notes: This figure shows the effect of exposure to Bible translation on full literacy split by the gender of the respondent when using the estimation method proposed by Sun and Abraham (2021). Panel A shows the results for the women's subsample, while Panel B shows the men's. For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure 9: Effects on Years of Education by Gender



Notes: This figure shows the effect of exposure to Bible translation on years of education split by the gender of the respondent when using the estimation method proposed by Sun and Abraham (2021). Panel A shows the results for the women's subsample, while Panel B shows the men's. For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure 10: Effects on Primary School Completion by Gender



Notes: This figure shows the effect of exposure to Bible translation on primary school completion split by the gender of the respondent when using the estimation method proposed by Sun and Abraham (2021). Panel A shows the results for the women's subsample, while Panel B shows the men's. For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Tables

Table 1: Treated and Untreated Languages in the Sample

DHS Ethnic Group	Year of Bible Translation
LANGO	1979
SOMALI	1979
SOMALIE	1979
KARO / ZIME / PEVE	1980
SARA (NGAMBAYE/SARA MADJIN-GAYE/MBAYE)	1980
LOKO	1983
MOUNDANG	1983
ACHOLI	1986
GABRI / KABALAYE / NANGTCHERE / SOUMRAYE	1986
WOLOF	1988
ALABA	ARAB
ARGOBA	BAKENYI
BOULALA / MEDEGO / KOUKA	DADAJO / KIBET / MOURO
DENDI AND RELATED	DOGON
FULA/TUKULUR/LOROBO	FULANI
FULLAH	GRUSI
HARARI	KALENJIN
KUKU	LUHYA
MANDINKA/JAHANKA	MAO
MESMEDJE / MASSALAT / KADJAKSE	MIJIKENDA/ SWAHILI
NGONI	NUBI
OROMO	OUADAI / MABA / MASSALIT / MIMI
PEUL / FOULBE / BODORE	PEULH
PEULH AND RELATED	SHEKO
SHERBRO	SHINASHA
SO (TEPETH)	TAMA / ASSONGORI / MARARIT
ZAGHAWA / BIDEYAT / KOBE	-

Note:
Treated ethnic groups based on the DHS classification in the balanced and binned sample, alongside with the year of Bible translation. Untreated languages are in the panel below.

Table 2: Data Description

DHS Country	Obs. by Country	Mean Literacy (%)	Total Languages	Total Untreated Languages	Number of Missions
Benin	2954	11.30	2	2	17
Chad	12051	19.43	12	8	5
Ethiopia	8690	39.13	8	7	25
Gambia	8699	34.50	2	2	8
Ghana	608	38.61	1	1	31
Guinea	5680	15.75	1	1	25
Kenya	15496	71.56	4	3	96
Malawi	4044	71.63	1	1	63
Mali	2709	20.02	2	2	16
Nigeria	3773	11.80	1	1	60
Senegal	6766	37.17	1	0	29
Sierra Leone	1735	27.47	3	2	49
Uganda	2431	47.38	6	4	51
Total	**75636**	**38.82**	**44**	**34**	**475**

Note:
Descriptive statistics of the balanced and unbinned sample

Table 3: Estimates for Full Literacy

	Sun and Abraham	TWFE	Callaway and Sant'Anna	Imputation Method
-15	0.025	0.027		0.008
	(0.024)	(0.036)		(0.023)
-14	0.031***	0.032	0.060	0.017
	(0.016)	(0.023)	(0.094)	(0.013)
-13	0.018	0.013	−0.040	−0.001
	(0.017)	(0.030)	(0.050)	(0.019)
-12	0.029	0.046	0.011	0.029
	(0.022)	(0.034)	(0.122)	(0.023)
-11	0.022	0.028	0.034	0.007
	(0.016)	(0.021)	(0.062)	(0.022)
-10	0.031	0.043	0.008	0.019
	(0.023)	(0.026)	(0.082)	(0.020)
-9	0.000	0.004	0.019	−0.007
	(0.012)	(0.025)	(0.089)	(0.018)
-8	−0.002	−0.003	−0.021	−0.015
	(0.016)	(0.026)	(0.077)	(0.015)
-7	0.004	0.018	0.005	−0.001
	(0.013)	(0.026)	(0.050)	(0.017)
-6	0.026**	0.031**	0.003	0.009
	(0.011)	(0.012)	(0.058)	(0.012)
-5	0.005	0.013	−0.015	−0.013
	(0.014)	(0.019)	(0.037)	(0.008)
-4	0.039**	0.048**	0.043	0.028***
	(0.014)	(0.019)	(0.038)	(0.017)
-3	−0.024**	−0.017	−0.086	−0.035***
	(0.011)	(0.024)	(0.063)	(0.020)
-2	0.015	0.031	0.061	0.009
	(0.018)	(0.035)	(0.075)	(0.029)
-1			−0.023	
			(0.070)	
0	0.003	0.020	0.012	−0.004
	(0.021)	(0.029)	(0.063)	(0.020)
1	0.016***	0.027	0.034	0.010
	(0.009)	(0.018)	(0.050)	(0.025)
2	0.034**	0.037**	0.041	0.019
	(0.012)	(0.016)	(0.057)	(0.021)
3	0.062**	0.077**	0.063	0.053**
	(0.014)	(0.018)	(0.054)	(0.019)
4	0.023**	0.028***	0.036	0.004
	(0.010)	(0.016)	(0.099)	(0.019)
5	0.041	0.046	0.060	0.022
	(0.028)	(0.030)	(0.061)	(0.031)
6	0.055**	0.062**	0.063	0.042
	(0.019)	(0.026)	(0.072)	(0.031)
7	0.028	0.038**	0.029	0.019
	(0.017)	(0.017)	(0.066)	(0.018)
8	0.120**	0.129**	0.134	0.112**
	(0.020)	(0.026)	(0.059)	(0.022)
9	0.027***	0.036***	0.059	0.012
	(0.015)	(0.018)	(0.096)	(0.019)
10	0.118**	0.128**	0.135	0.105**
	(0.023)	(0.031)	(0.066)	(0.026)
11	0.103**	0.108**	0.131	0.096**
	(0.017)	(0.027)	(0.085)	(0.026)
12	0.117**	0.123**	0.181	0.107**
	(0.014)	(0.026)	(0.094)	(0.028)
13	0.139**	0.149**	0.238	0.134**
	(0.020)	(0.023)	(0.121)	(0.022)
14	0.110**	0.116**	0.222	0.101**
	(0.018)	(0.020)	(0.061)	(0.023)
Num.Obs.	75066	75 066	75 066	75 066
R2	0.354	0.352		0.005

* p < 0.1, *** p < 0.1, ** p < 0.05

[a] Controls include: gender, urban or rural. Balanced and binned sample. Standard errors are clustered at the ethnic group level

Table 4: Sun and Abraham Estimates for Years of Education and Primary School Completion

	Years of Education	Primary School Completion
-15	0.266	−0.005
	(0.257)	(0.027)
-14	0.335***	0.024
	(0.175)	(0.021)
-13	0.149	−0.002
	(0.191)	(0.024)
-12	0.447**	0.019
	(0.208)	(0.023)
-11	0.132	0.006
	(0.178)	(0.022)
-10	0.232	0.030***
	(0.188)	(0.017)
-9	0.031	−0.008
	(0.140)	(0.015)
-8	−0.123	−0.013
	(0.163)	(0.016)
-7	0.087	0.010
	(0.131)	(0.014)
-6	0.240**	0.023***
	(0.107)	(0.012)
-5	0.217**	0.012
	(0.094)	(0.011)
-4	0.207	0.029***
	(0.136)	(0.017)
-3	−0.154***	−0.028**
	(0.085)	(0.009)
-2	0.275**	0.020***
	(0.102)	(0.012)
0	0.178	0.011
	(0.197)	(0.019)
1	0.639**	0.032**
	(0.140)	(0.007)
2	0.590**	0.043**
	(0.119)	(0.012)
3	0.839**	0.060**
	(0.165)	(0.013)
4	0.427**	0.044**
	(0.167)	(0.016)
5	0.723**	0.059**
	(0.315)	(0.024)
6	0.835**	0.072**
	(0.179)	(0.015)
7	0.472**	0.042**
	(0.183)	(0.017)
8	1.282**	0.140**
	(0.255)	(0.026)
9	0.675**	0.083**
	(0.268)	(0.025)
10	1.323**	0.173**
	(0.319)	(0.035)
11	1.318**	0.176**
	(0.334)	(0.041)
12	0.989**	0.155**
	(0.267)	(0.037)
13	0.937**	0.100**
	(0.263)	(0.039)
14	0.485***	0.026
	(0.244)	(0.034)
Num.Obs.	75609	75636
R2	0.393	0.265

* $p < 0.1$, *** $p < 0.1$, ** $p < 0.05$

[a] Controls include: gender, urban or rural. Balanced and binned sample. Standard errors are clustered at the ethnic group level

Table 5: Sun and Abraham Estimates for Religious Affiliation

	Christianity	Islam	Catholic	Protestant
-15	0.010	−0.002	−0.047**	0.057**
	(0.013)	(0.011)	(0.010)	(0.018)
-14	−0.005	0.012	−0.029**	0.025
	(0.012)	(0.009)	(0.011)	(0.015)
-13	0.003	0.009	−0.003	0.006
	(0.013)	(0.009)	(0.016)	(0.015)
-12	0.011	−0.009	−0.009	0.020
	(0.018)	(0.012)	(0.014)	(0.015)
-11	0.010	0.001	−0.033**	0.043**
	(0.011)	(0.009)	(0.012)	(0.015)
-10	0.006	0.007	0.008	−0.002
	(0.007)	(0.005)	(0.010)	(0.011)
-9	0.018	0.004	0.018	0.001
	(0.014)	(0.011)	(0.014)	(0.016)
-8	0.003	0.003	−0.015	0.018
	(0.010)	(0.008)	(0.016)	(0.017)
-7	−0.001	0.001	−0.037**	0.036**
	(0.011)	(0.010)	(0.012)	(0.017)
-6	0.014***	0.009	−0.015**	0.029**
	(0.008)	(0.007)	(0.006)	(0.010)
-5	0.000	0.000	−0.035***	0.035***
	(0.010)	(0.007)	(0.019)	(0.017)
-4	0.000	0.005	−0.021**	0.020
	(0.018)	(0.012)	(0.008)	(0.016)
-3	−0.007	0.003	−0.025**	0.018
	(0.012)	(0.010)	(0.012)	(0.013)
-2	0.019	0.003	0.004	0.015
	(0.012)	(0.008)	(0.012)	(0.014)
0	−0.013***	0.015**	−0.018**	0.005
	(0.007)	(0.006)	(0.004)	(0.006)
1	−0.003	0.013	−0.028**	0.024
	(0.012)	(0.008)	(0.011)	(0.015)
2	−0.005	0.006	−0.022**	0.017***
	(0.010)	(0.005)	(0.006)	(0.008)
3	−0.002	0.008	−0.012	0.010
	(0.010)	(0.007)	(0.007)	(0.008)
4	0.011***	0.005	−0.003	0.014***
	(0.006)	(0.004)	(0.006)	(0.008)
5	0.000	0.012***	−0.037**	0.036**
	(0.010)	(0.006)	(0.009)	(0.013)
6	0.000	0.017**	−0.028**	0.028**
	(0.009)	(0.005)	(0.007)	(0.009)
7	−0.001	0.015**	−0.026**	0.025**
	(0.008)	(0.007)	(0.007)	(0.010)
8	0.010	0.002	−0.024**	0.035**
	(0.011)	(0.008)	(0.008)	(0.010)
9	−0.003	0.013**	−0.021**	0.019**
	(0.007)	(0.005)	(0.007)	(0.009)
10	−0.008	0.012***	−0.037**	0.029**
	(0.012)	(0.007)	(0.008)	(0.013)
11	0.001	0.009	−0.013**	0.014
	(0.011)	(0.009)	(0.005)	(0.011)
12	0.019***	0.003	−0.026**	0.045**
	(0.011)	(0.007)	(0.011)	(0.009)
13	0.000	0.003	−0.025**	0.025***
	(0.010)	(0.008)	(0.010)	(0.014)
14	0.002	0.005	−0.048**	0.050**
	(0.012)	(0.008)	(0.008)	(0.011)
Num.Obs.	75636	75636	75636	75636
R2	0.741	0.807	0.313	0.491

* $p < 0.1$, *** $p < 0.1$, ** $p < 0.05$

[a] Controls include: gender, urban or rural. Balanced and binned sample. Standard errors are clustered at the ethnic group level

A. Effects of Earlier Translations on Literacy

In this appendix, we show the relationship between earlier mother-tongue Bible translations and adult literacy, presenting the results for ethnic group-level treatment.

Previous papers on the impact of Christian missions find selection regarding mission locations (Becker and Woessmann, 2009; D. Kim, 2020; Bai and Kung, 2015; Waldinger, 2017; Wantchekon et al., 2015). Since access to a Bible translation is likely associated with distance from a mission, we reduce the potential for bias by narrowing our sample to respondents living within 10 km of a mission.

In this approach, we exploit variation across ethno-language groups near a mission, comparing groups with and without access to mother-tongue Bible translations before 1970. Forty-five treated groups received a Bible translation before 1970, and 93 groups had no Bible translation available before 1970, as shown in Table B.2.[35] We estimate the following equation:

$$literate_{ilmcdt} = \beta Bible1970_l + \delta X_i + \alpha Z_l + \theta_m + \gamma_c + \phi_d + \alpha_t + \epsilon_{ilmcdt} \tag{4}$$

where indices indicate individual i, language l, mission m, country c, DHS cluster d, and cohort t. Our outcome, $literate_{ilmcdt}$, is either full or partial literacy. The treatment variable is $Bible1970_l$, equal to 1 if a Bible translation for language l is available before 1970. Individual characteristics, X_i, include male and rural indicators. Ethno-language group characteristics, Z_l, include gathering dependence, hunting dependence, fishing dependence, agriculture intensity, marital composition, major crop types, and settlement patterns. We include fixed effects for each mission θ_m, country γ_c, DHS cluster ϕ_d, and cohort α_t. The mission fixed effects allow to capture the within-mission relationship between translations and literacy, accounting for time-invariant differences between missions. Finally, ϵ_{ilcmdt} is the error term. We are interested in β, which captures the correlation between of early Bible translations and literacy for individuals located near a mission. Finally, for correct statistical inference of estimates, we cluster our standard errors at the language group level (Abadie et al., 2017).

We show our findings for specification (1) in Table B.3 and Table B.4. In column (1) of Panel A, access to a pre-1970 Bible translation is associated with an imprecisely estimated increase of 7.2 percentage points in the full literacy rate. In column (2), when we include indicators for male and rural and introduce fixed effects for DHS cluster, country, year of birth, and mission, the estimated coefficient becomes statistically significant but is smaller than in column (1). With these controls, access to an early Bible translation is associated with a three percentage point increase in literacy. Column (3) includes further controls for pre-

[35] This table summarizes the individual sample lost and the ethno-language groups from the DHS cluster located within 10 km of each mission.

colonial characteristics for each ethnic group, resulting in a less precise estimate of a three percentage point increase in literacy.

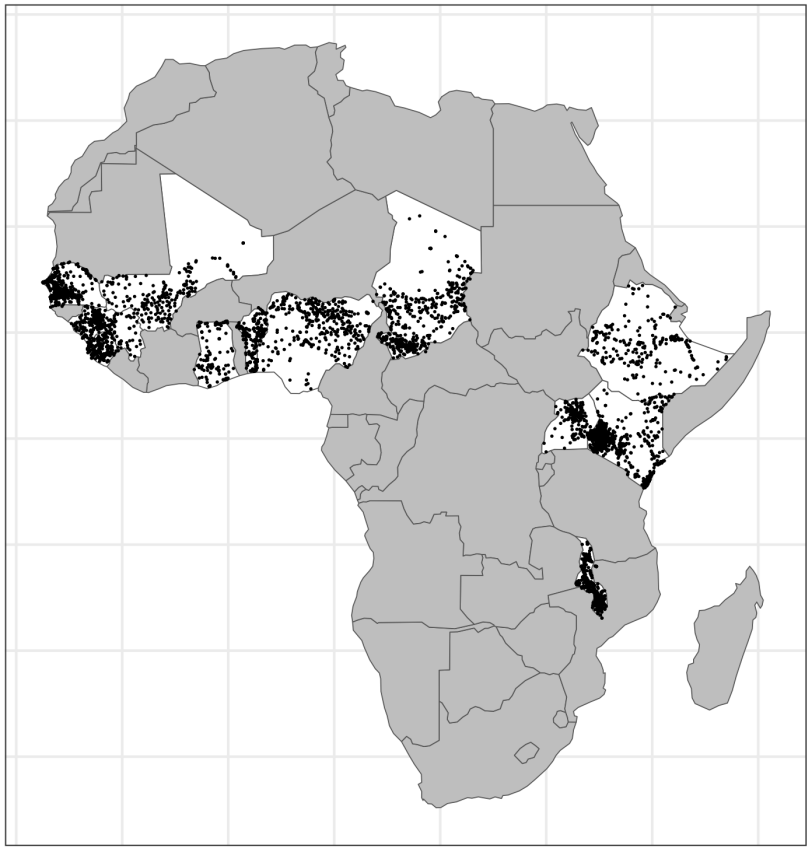
The results are consistent with positive selection into Bible translation.[36] The estimates in column (1) are likely to be upward-biased due to differential pre-colonial characteristics and are driven mainly by the largest ethnic group, as shown in Table 6. While Jedwab et al. (2018) and Michalopoulos and Papaioannou (2013) argue that controlling for pre-colonial ethnic group characteristics as in column (3), corrects for bias, it is likely that this approach does not produce a causal estimate.[37]

[36] Another issue in the approaches that utilize mission locations comes from nonrandom measurement error of historical maps (Jedwab et al., 2018) (Fahs 1925).

[37] These results are consistent with the findings by Jedwab et al. (2018) and Michalopoulos and Papaioannou (2013), who argue that the characteristics at the ethnic group level before the missionaries' arrival correct bias in the estimates of the impacts of missions on various outcome.

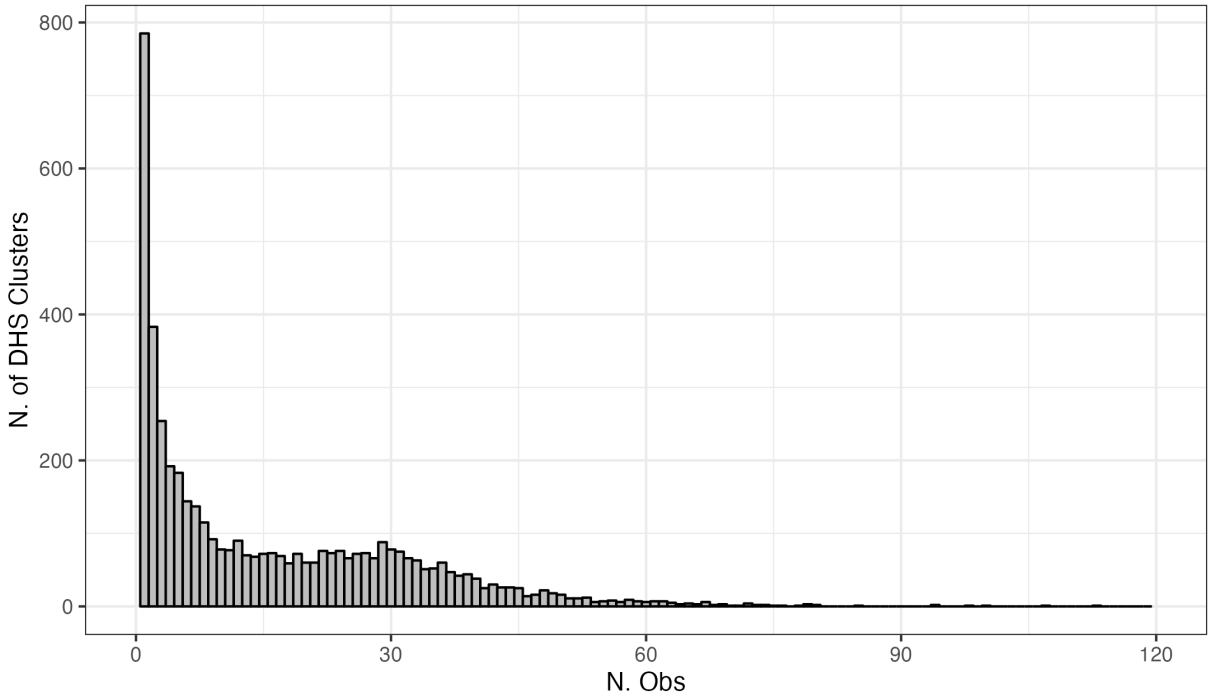
B. Figures and Tables

Figure B.1: Distribution DHS Cluster Across Selected Countries



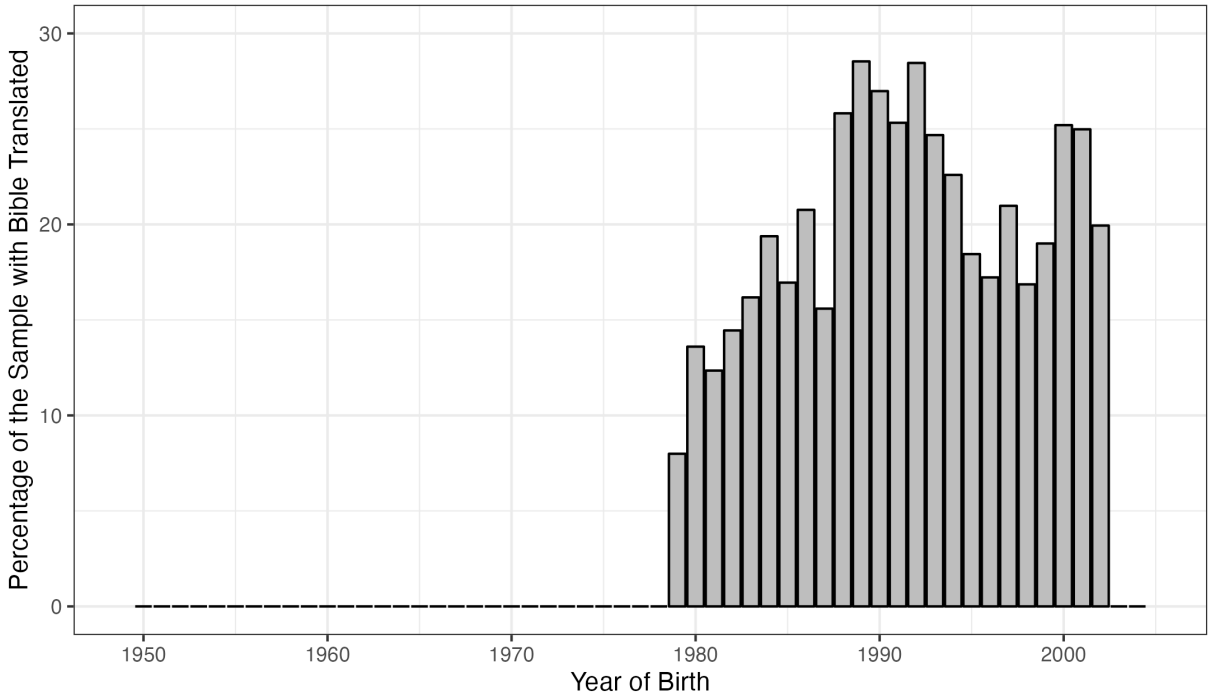
Notes: This figure shows in white the selected countries in our sample, and the black dots represent a DHS 7 cluster.

Figure B.2: Distribution of Observations per DHS cluster in the Sample Used for the Main Specification.



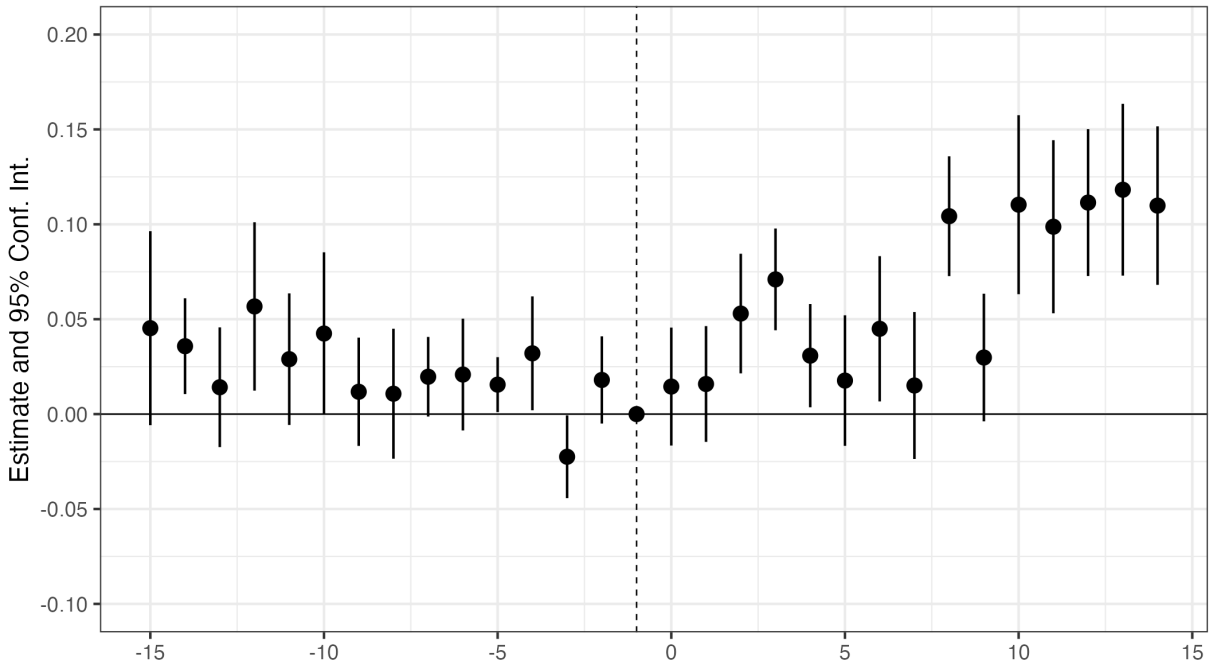
Notes: This figure shows the number of observations in our main analytical sample for each DHS 7 cluster.

Figure B.3: Percentage of the Sample with a Bible Translation by the Year They Were Born



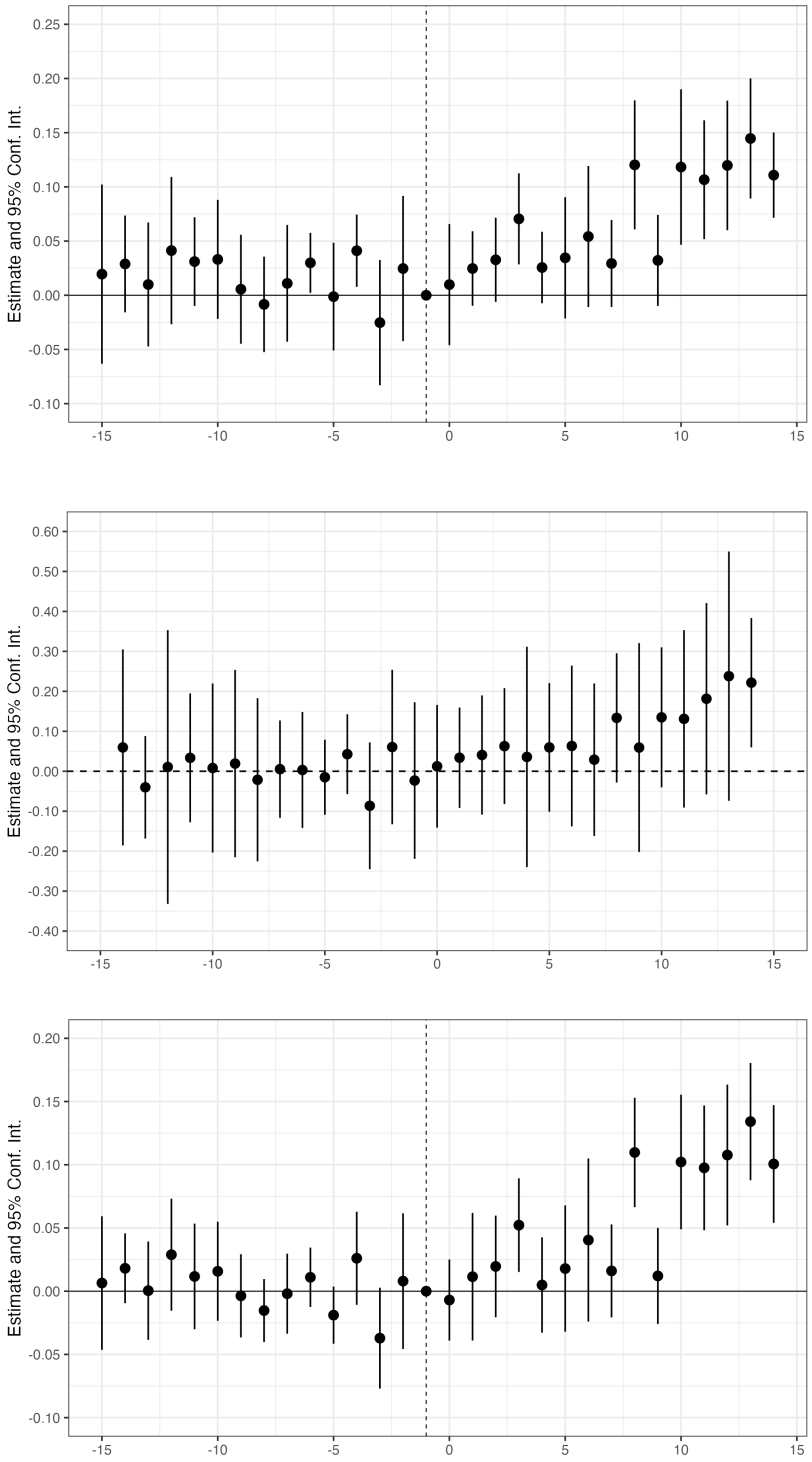
Notes: This figure describes the percentage of the sample that is treated by the year of birth of the respondent. Since the first translation in the sample was in 1979, the first time the respondents were treated began in 1979.

Figure B.4: Effects on Partial Literacy



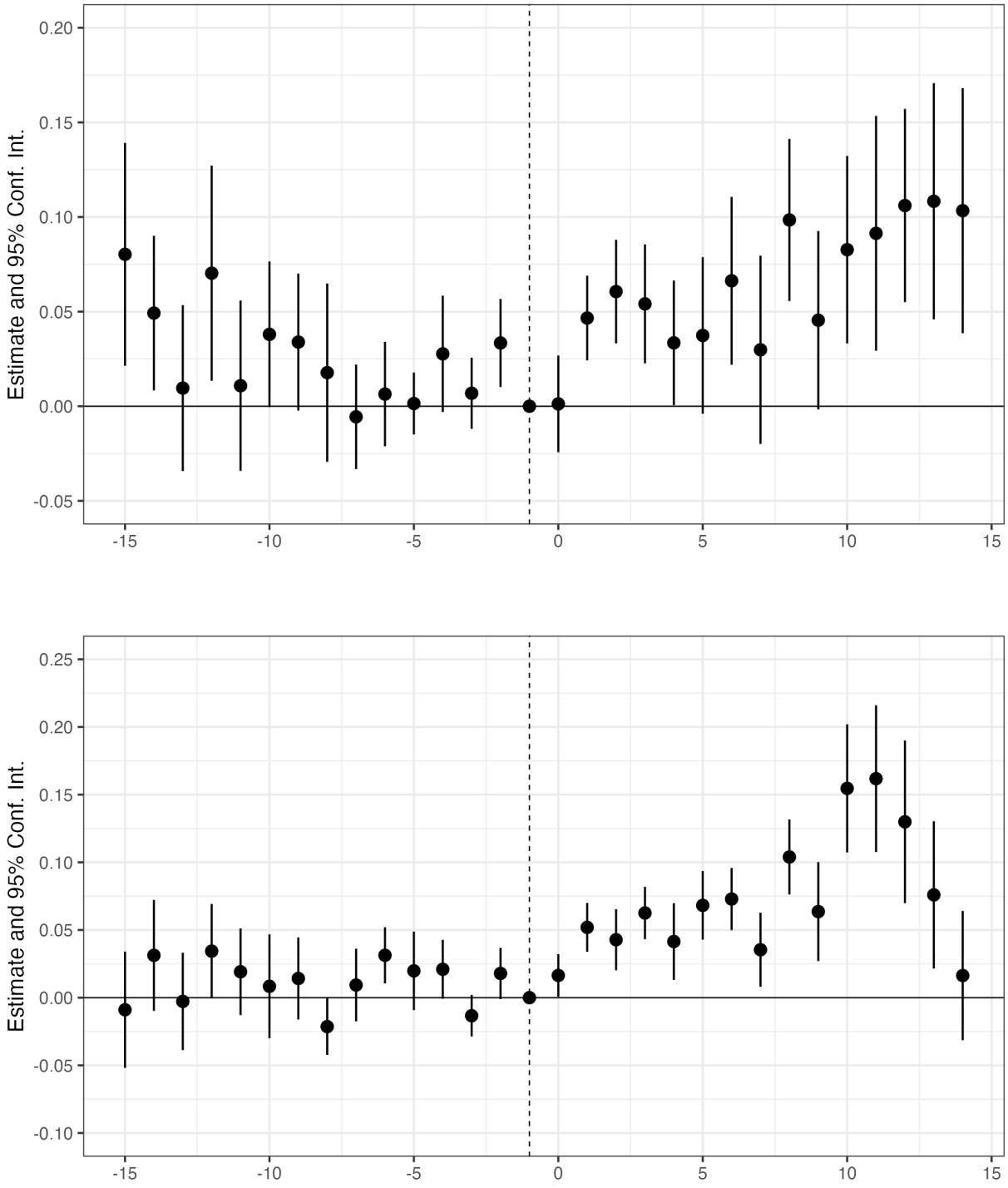
Notes: This figure shows the effect of exposure to Bible translation on partial literacy when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.5: Other Estimation Methods on Full Literacy



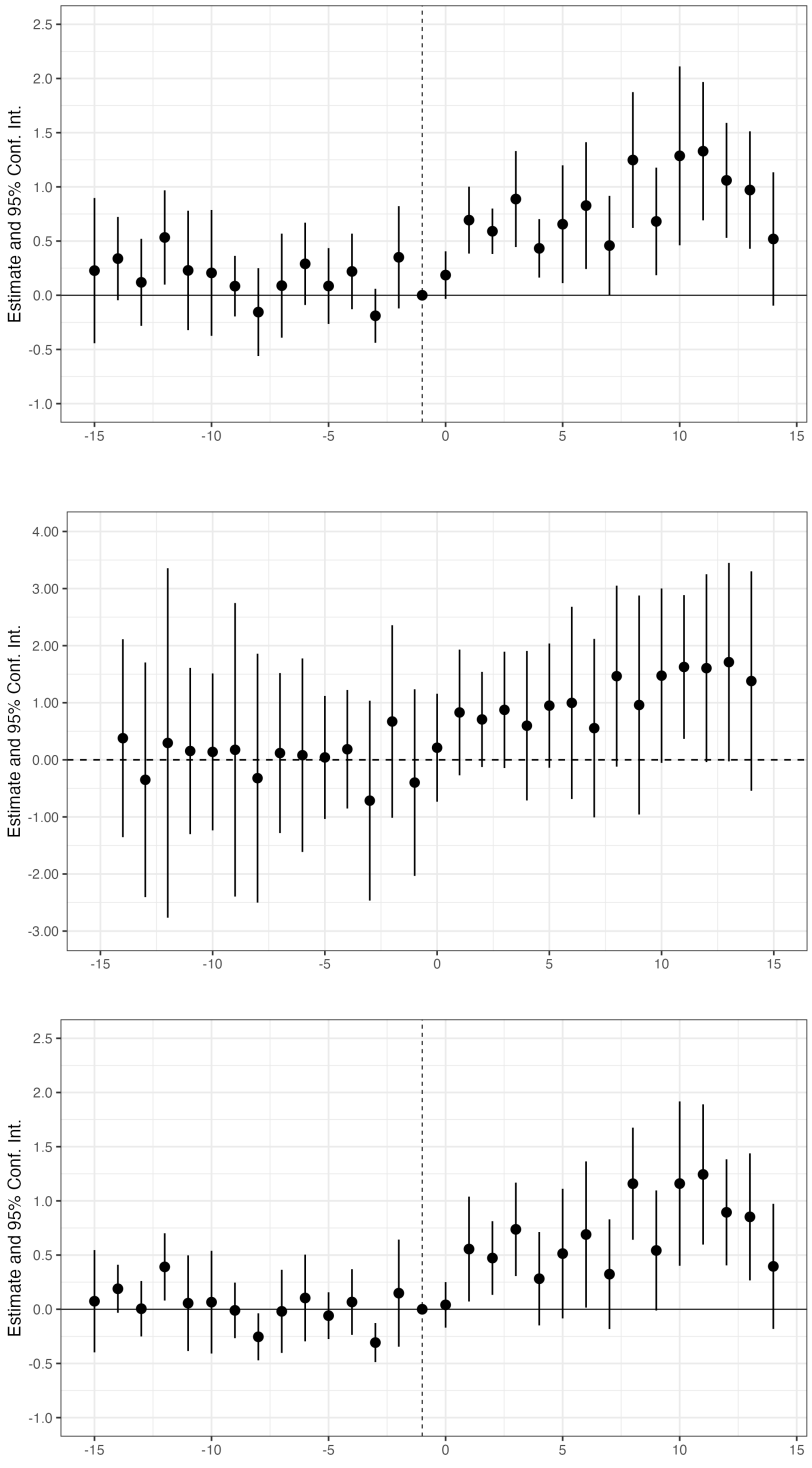
Notes: Effects of full literacy using three other estimation methods. In the first panel, a traditional two-way fixed effects model; in the second panel the estimation method proposed by Callaway and Sant'Anna (2021) and in the third panel, the method proposed by Borusyak et al. (2021).

Figure B.6: Effects on Any Education and Secondary Education



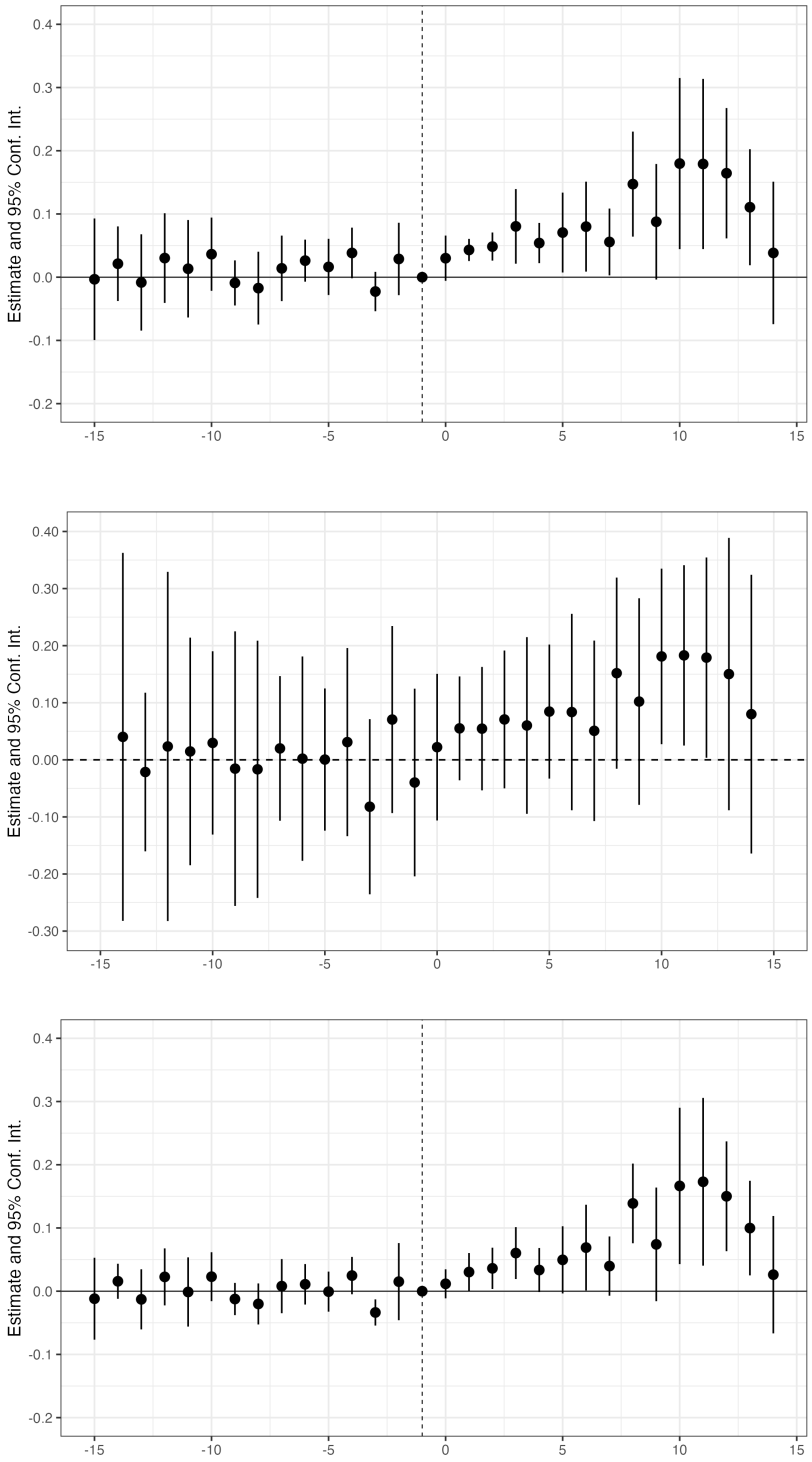
Notes: This figure shows the effect of exposure to Bible translation on any education (Panel A) or any secondary education (Panel B) when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.7: Other Estimation Methods on Years of Education



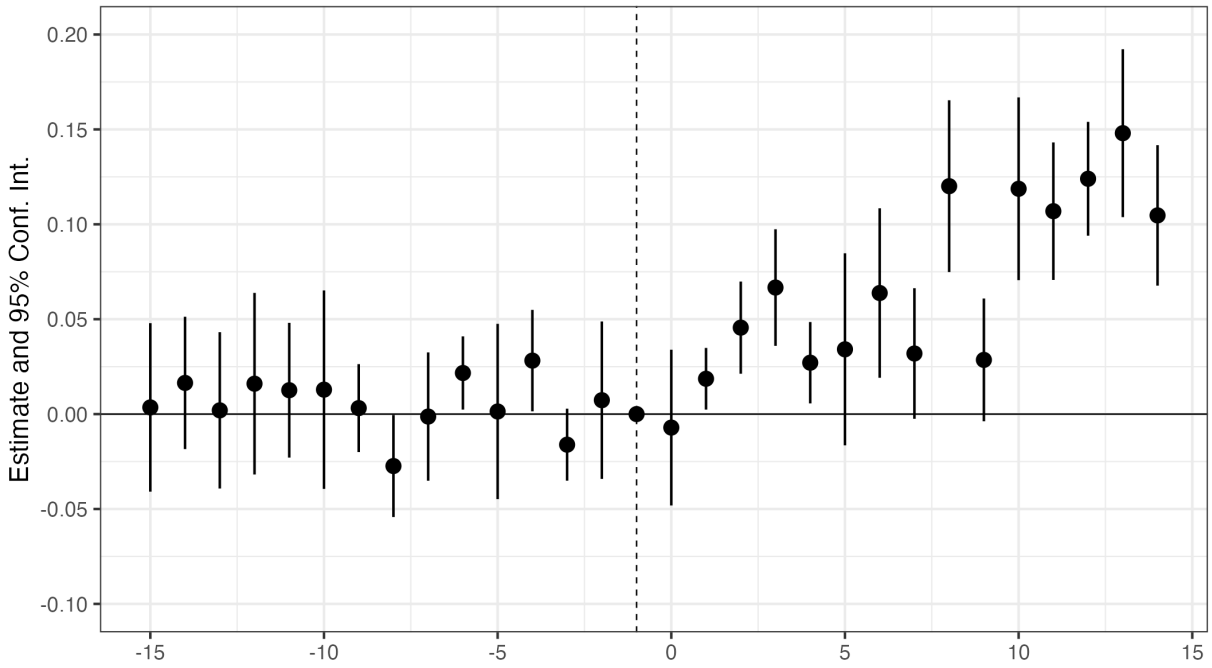
Notes: Effects on years of education using three other estimation methods. In the first panel, a traditional two-way fixed effects model; in the second panel the estimation method proposed by Callaway and Sant'Anna (2021) and in the third panel, the method proposed by Borusyak et al. (2021).

Figure B.8: Other Estimation Methods on Primary School Completion



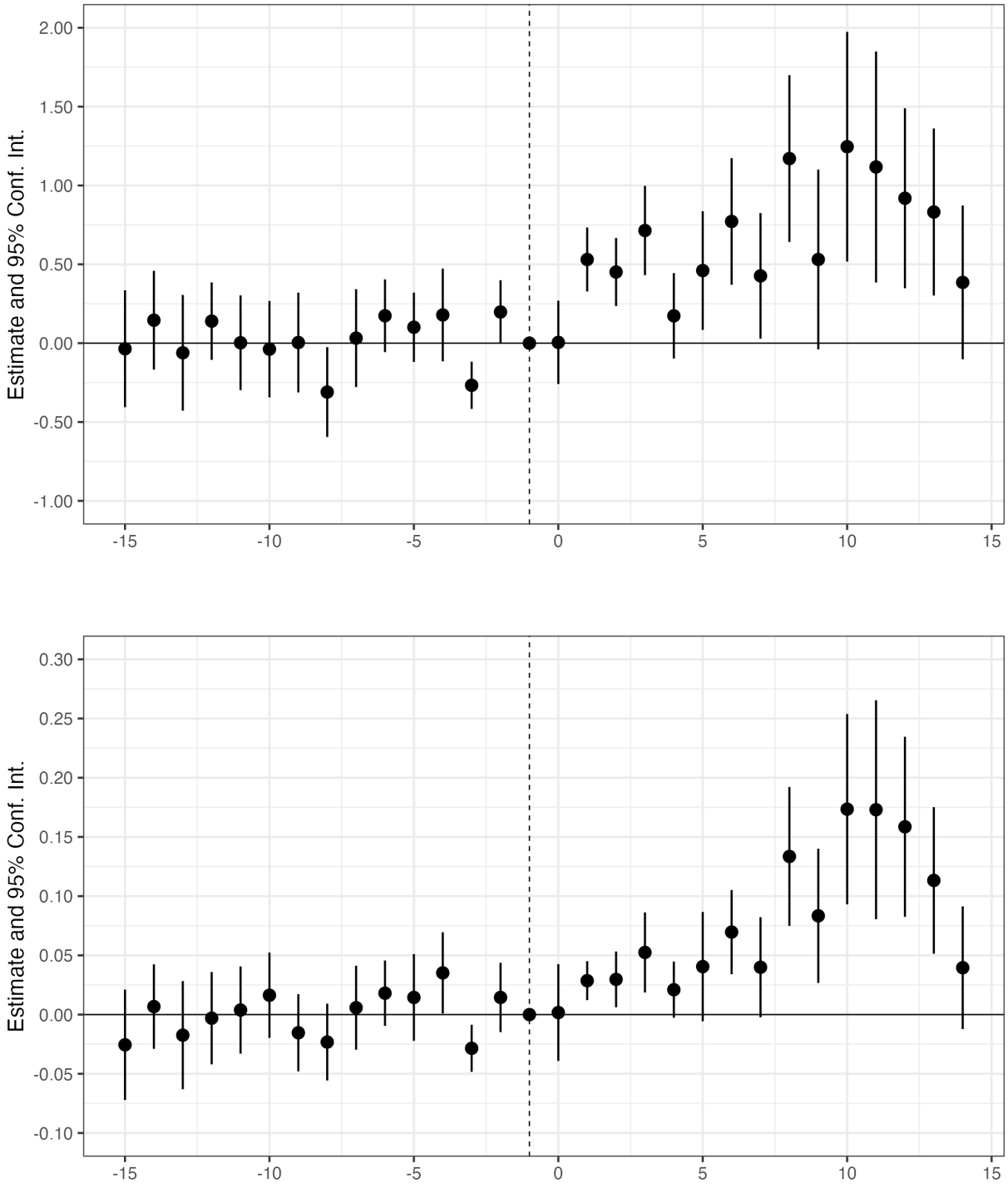
Effects of primary school completion using three other estimation methods. In the first panel, a traditional two-way fixed effects model; in the second panel the estimation method proposed by Callaway and Sant'Anna (2021) and in the third panel, the method proposed by Borusyak et al. (2021).

Figure B.9: Effects on Full Literacy - No Ethiopia or Mali Subsample



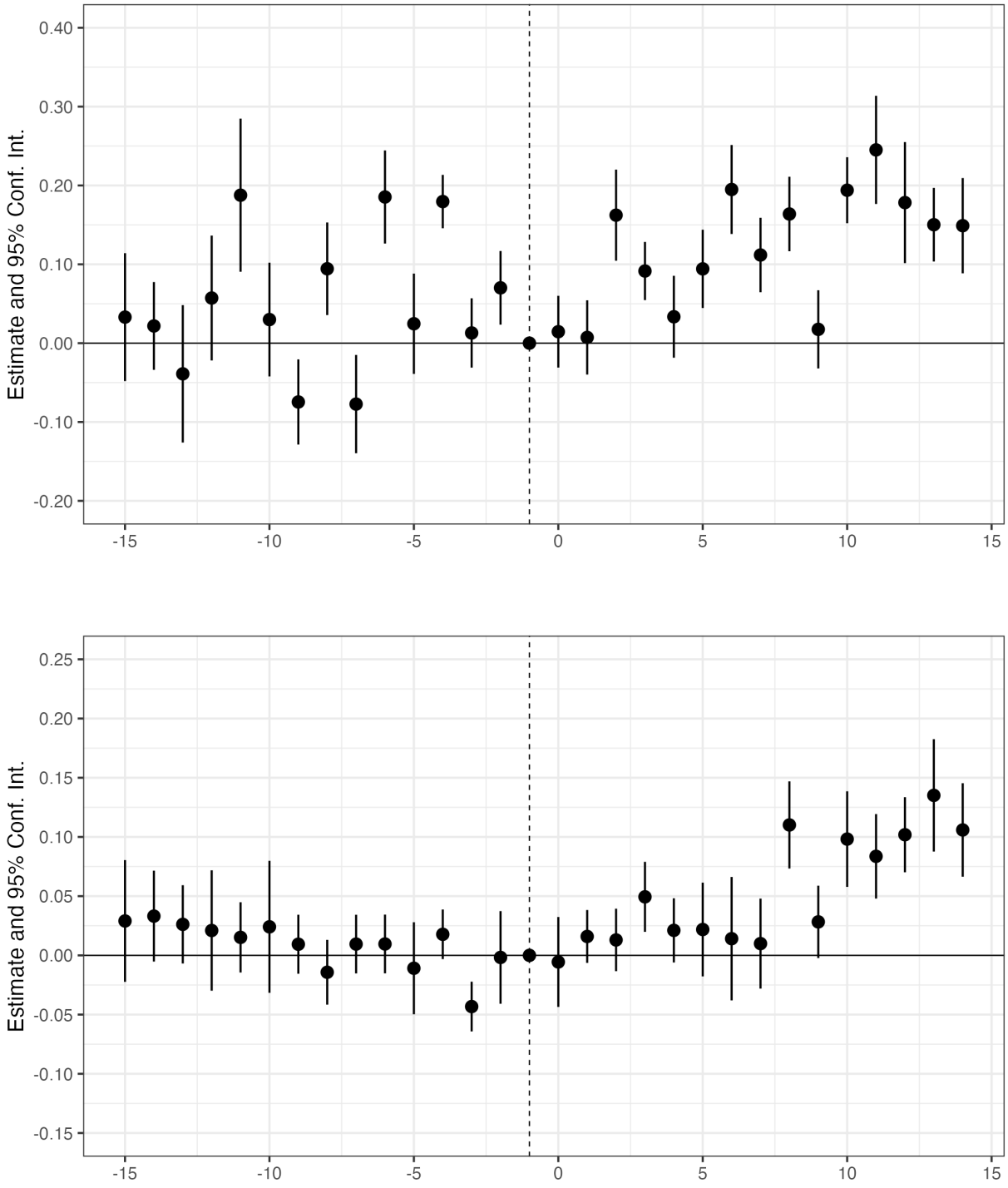
Notes: This figure shows the effect of exposure to Bible translation on full literacy when using the estimation method proposed by Sun and Abraham (2021) excluding Ethiopia and Mali. For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.10: Effects on Years of Education and Primary School Completion - No Ethiopia or Mali Subsample



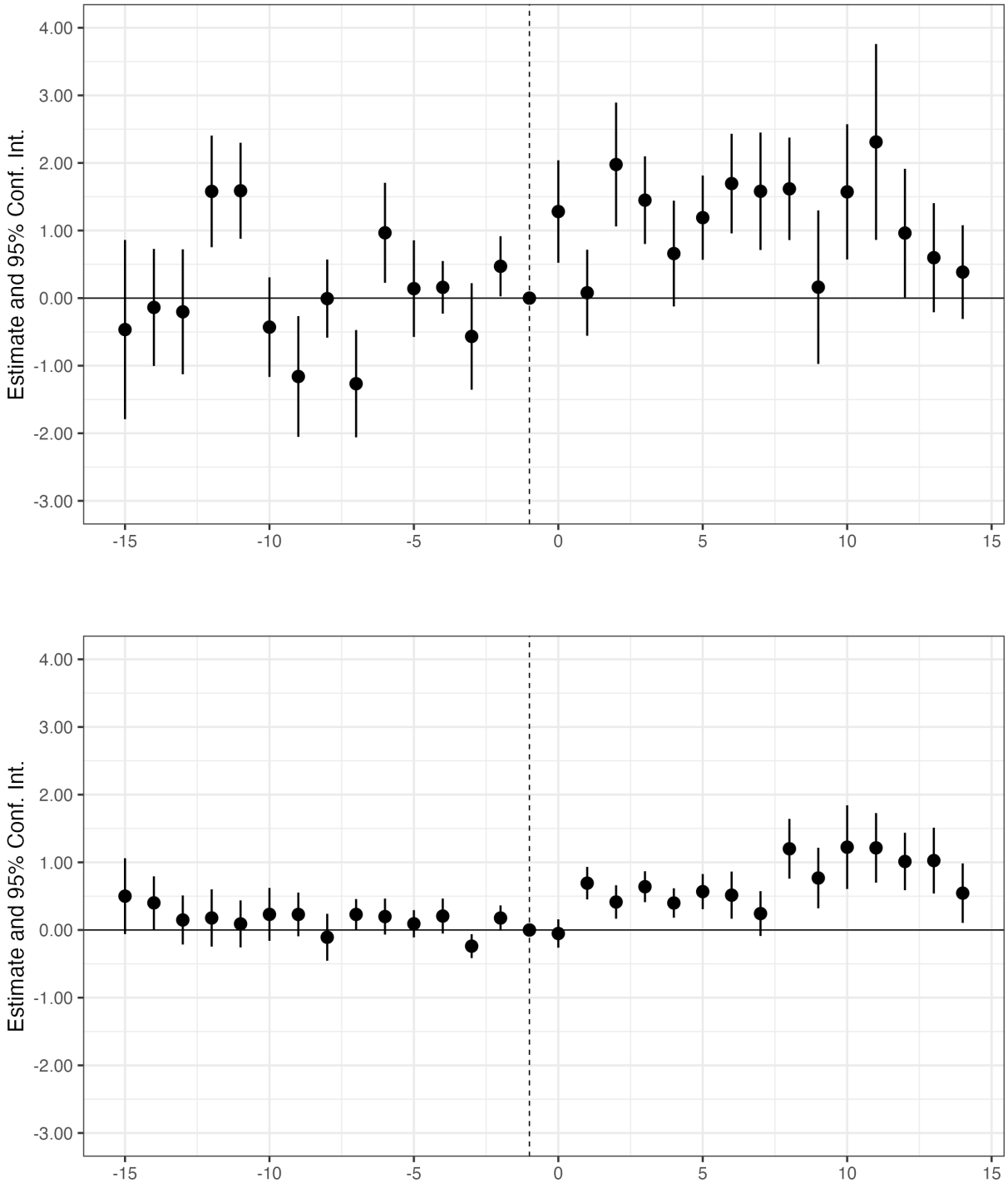
Notes: This figure shows the effect of exposure to Bible translation on years of education (Panel A) and primary school completion (Panel B) when using the estimation method proposed by Sun and Abraham (2021) excluding Ethiopia and Mali. For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.11: Effects on Full Literacy - Distance to Closest Historical Mission



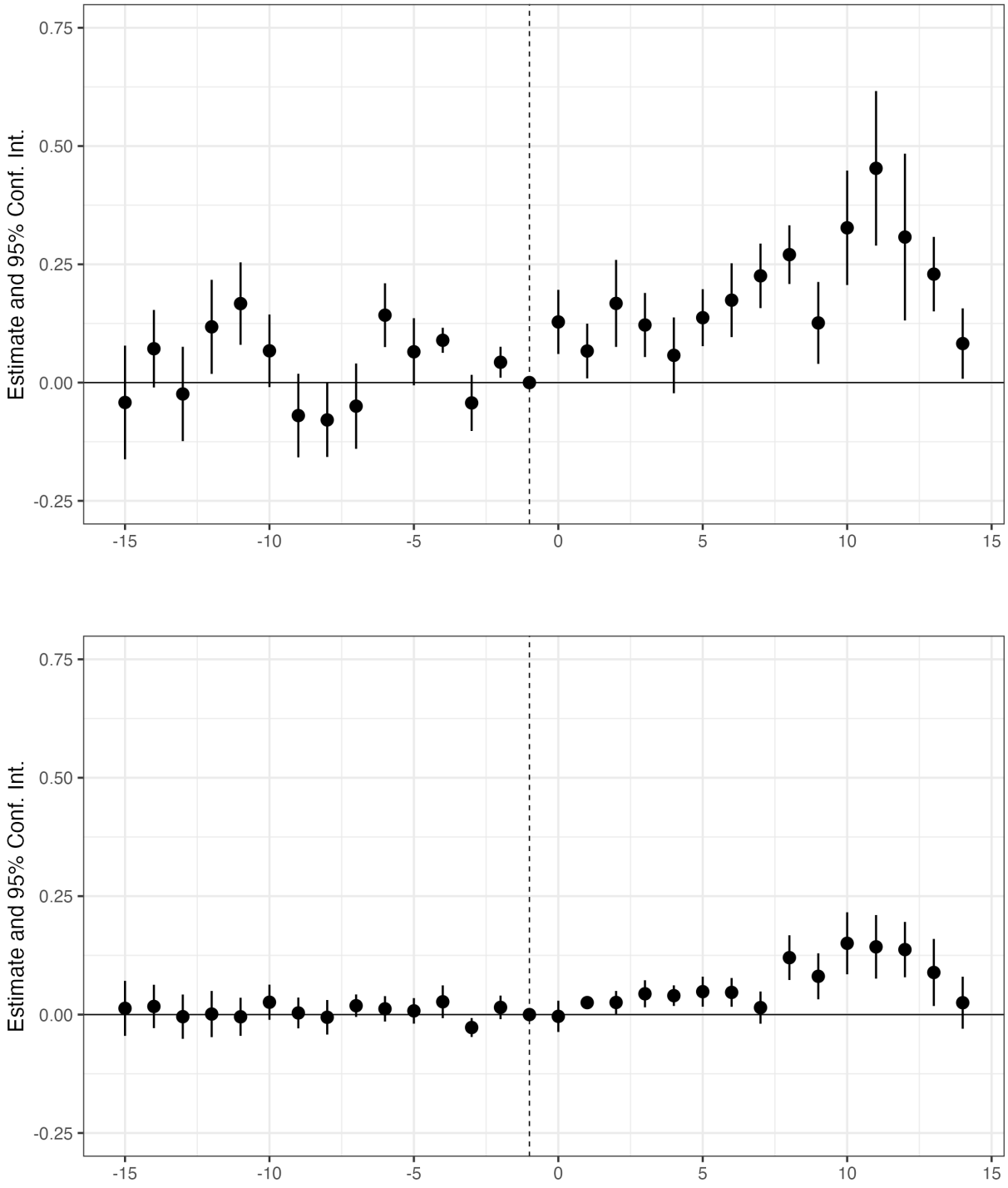
Notes: This figure shows the effect of exposure to Bible translation on full literacy split between DHS clusters less than (Panel A) or more than 10km (Panel B) from a historical mission when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.12: Effects on Years of Education - Distance to Closest Historical Mission



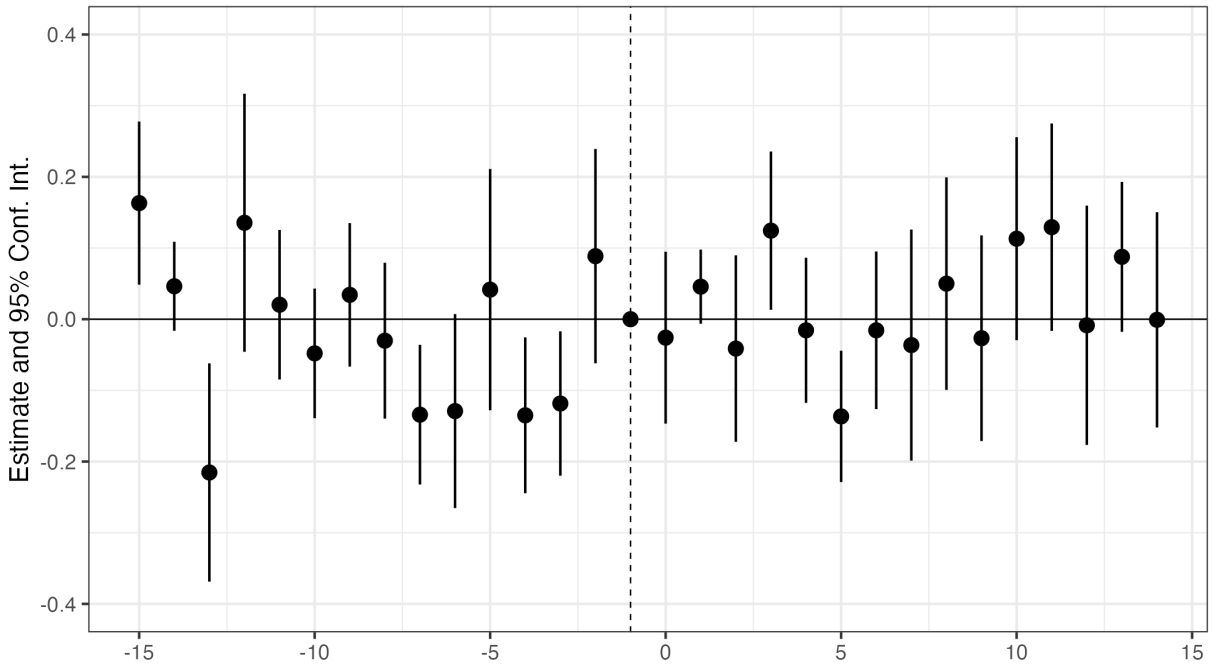
Notes: This figure shows the effect of exposure to Bible translation on years of education split between DHS clusters less than (Panel A) or more than 10km (Panel B) from a historical mission when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.13: Effects on Primary School Completion - Distance to Closest Historical Mission



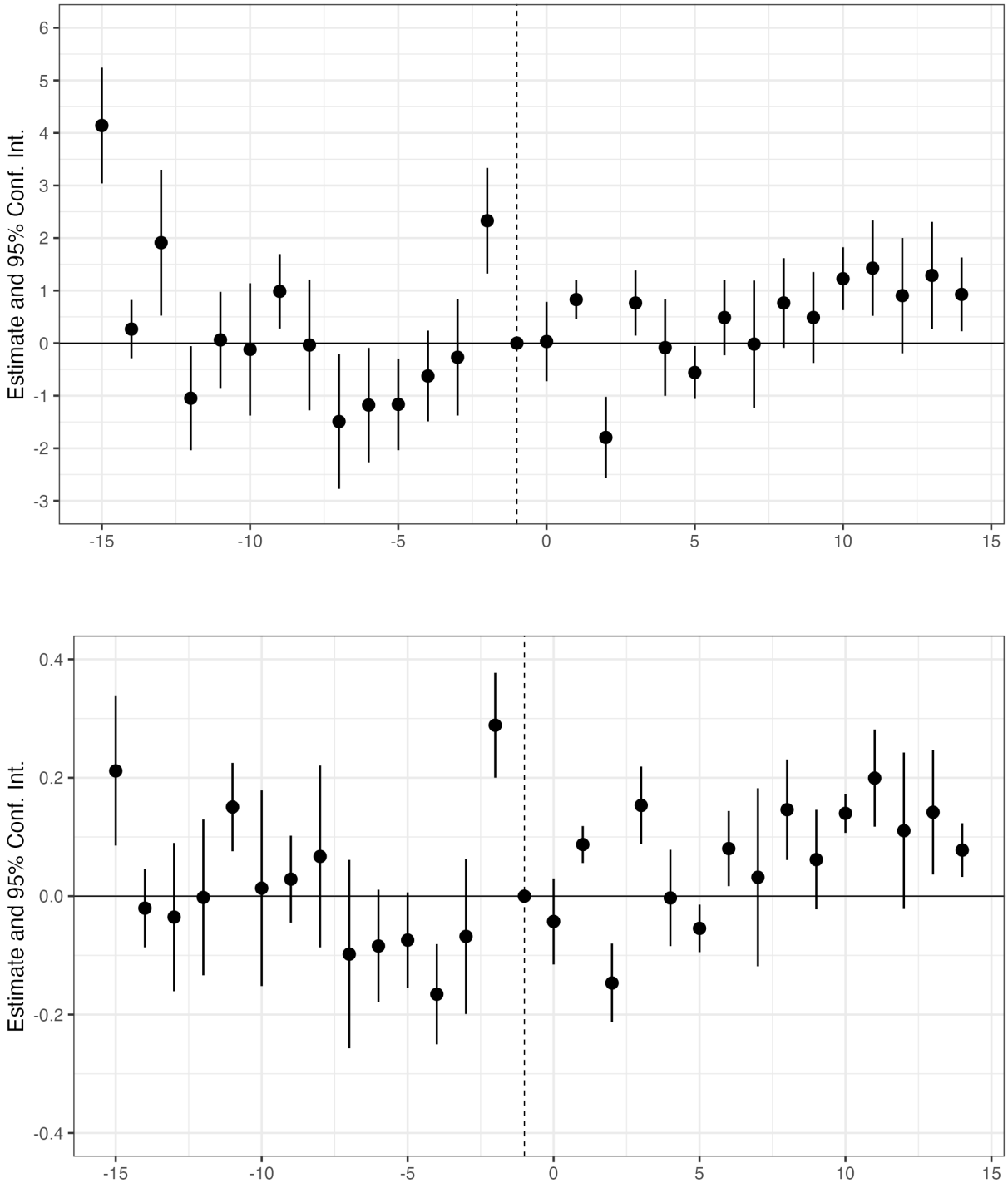
Notes: This figure shows the effect of exposure to Bible translation on primary school completion split between DHS clusters less than (Panel A) or more than 10km (Panel B) from a historical mission when using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.14: Effects on Full Literacy - No Printing Press Subsample



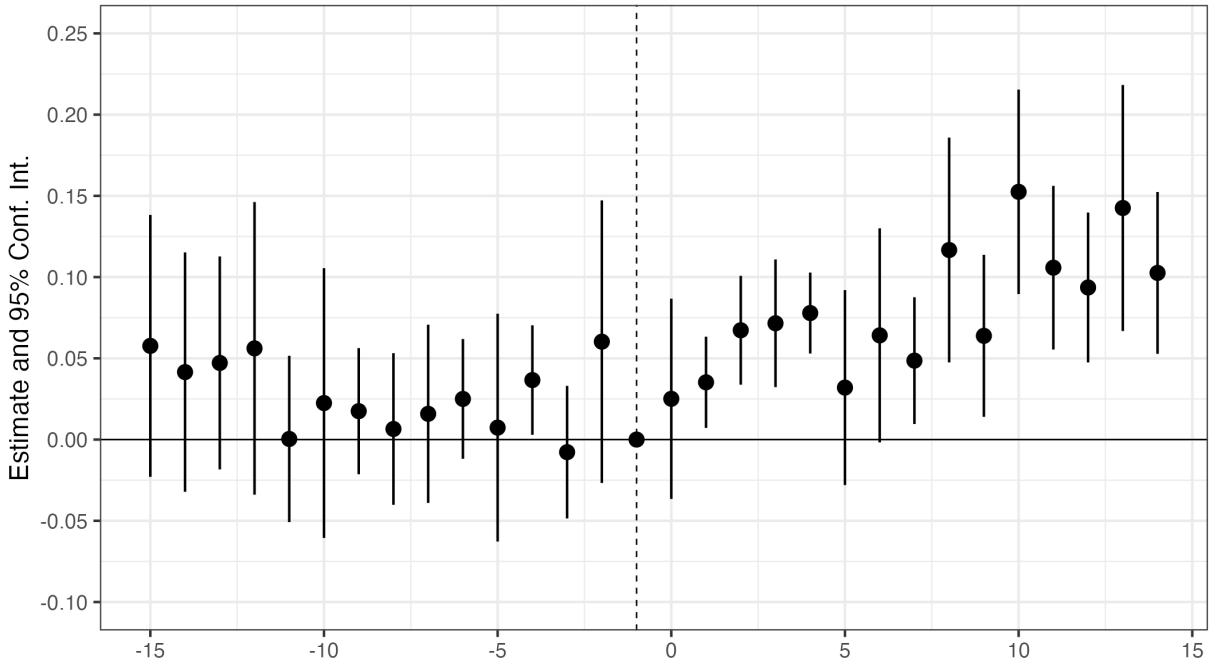
Notes: This figure shows the effect of exposure to Bible translation on full literacy on DHS clusters whose closest historical mission did not have a printing press using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.15: Effects on Years of Education and Primary School Completion - No Printing Press Subsample



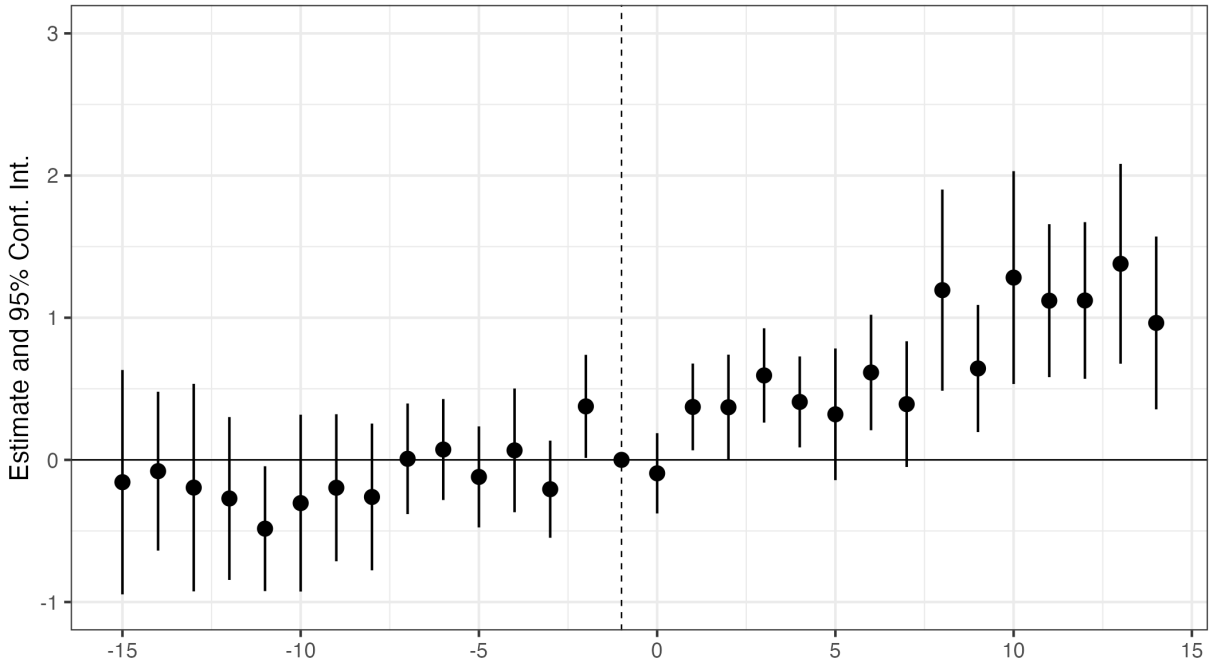
Notes: This figure shows the effect of exposure to Bible translation on years of education and primary school completion on DHS clusters whose closest historical mission did not have a printing press using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.16: Effects of Full Literacy - Protestant Mission Subsample



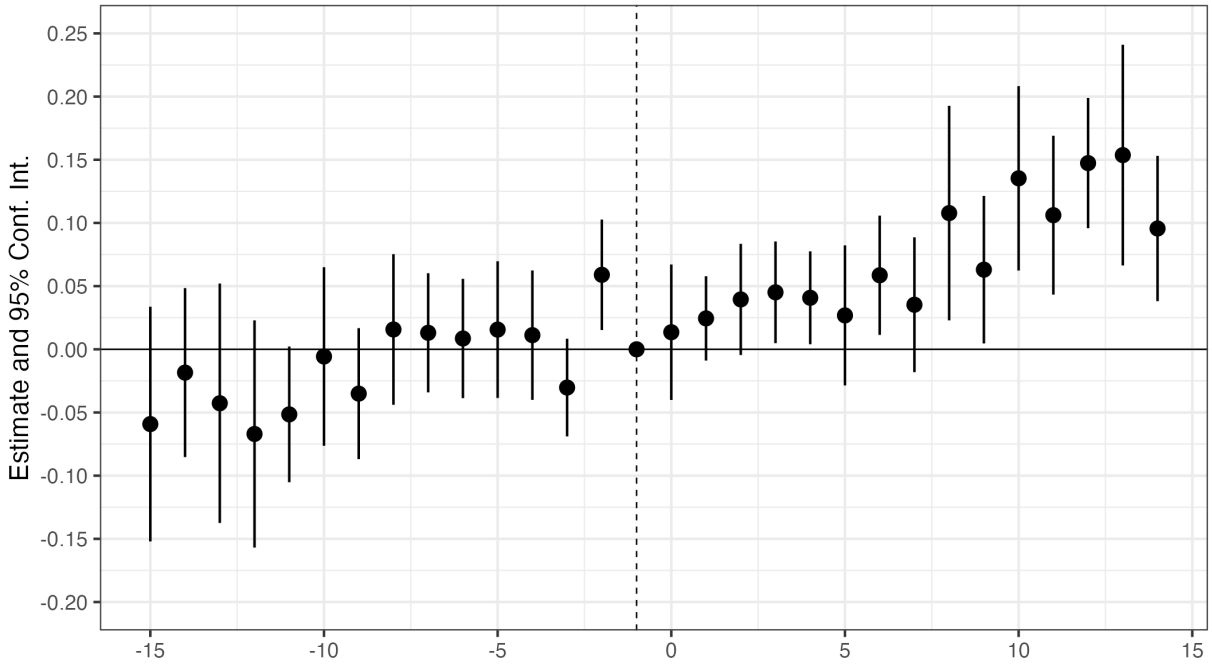
Notes: This figure shows the effect of exposure to Bible translation on full literacy on DHS clusters whose closest historical mission was from a Protestant denomination using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.17: Effects of Years of Education - Protestant Mission Subsample



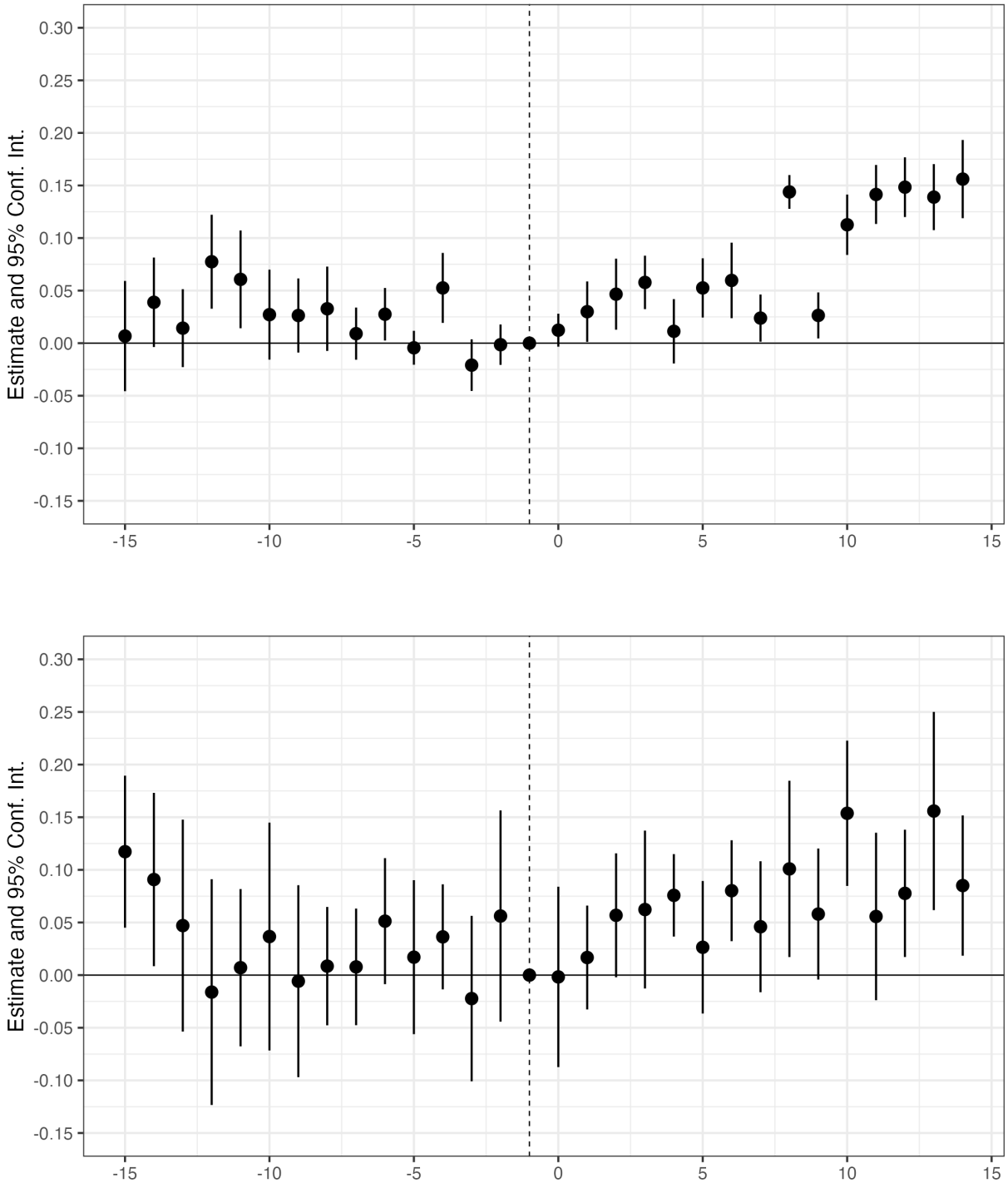
Notes: This figure shows the effect of exposure to Bible translation on years of education on DHS clusters whose closest historical mission was from a Protestant denomination using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.18: Effects of Primary School Completion - Protestant Mission Subsample



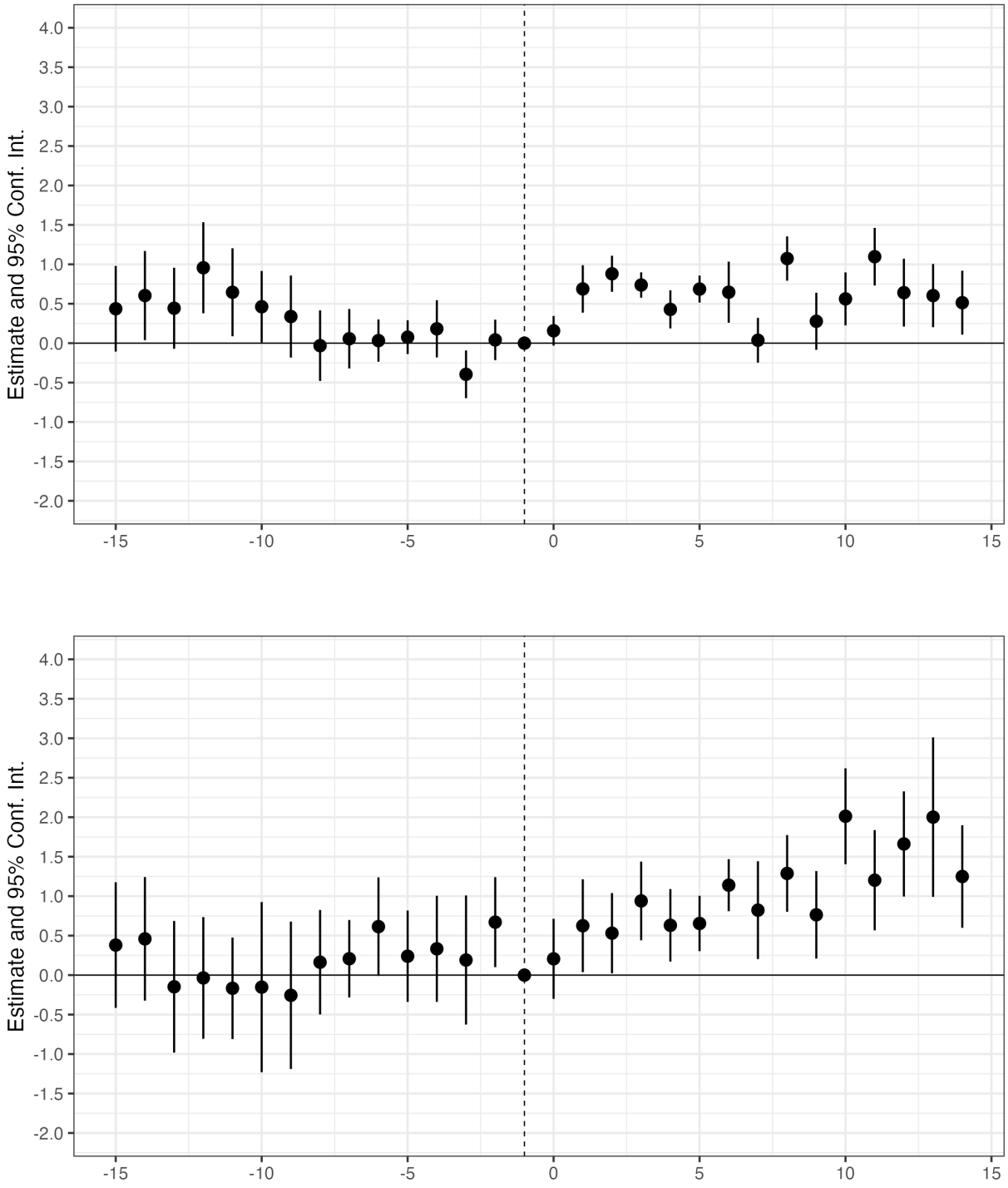
Notes: This figure shows the effect of exposure to Bible translation on primary school completion on DHS clusters whose closest historical mission was from a Protestant denomination using the estimation method proposed by Sun and Abraham (2021). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.19: Effects on Full Literacy By Muslim Regions



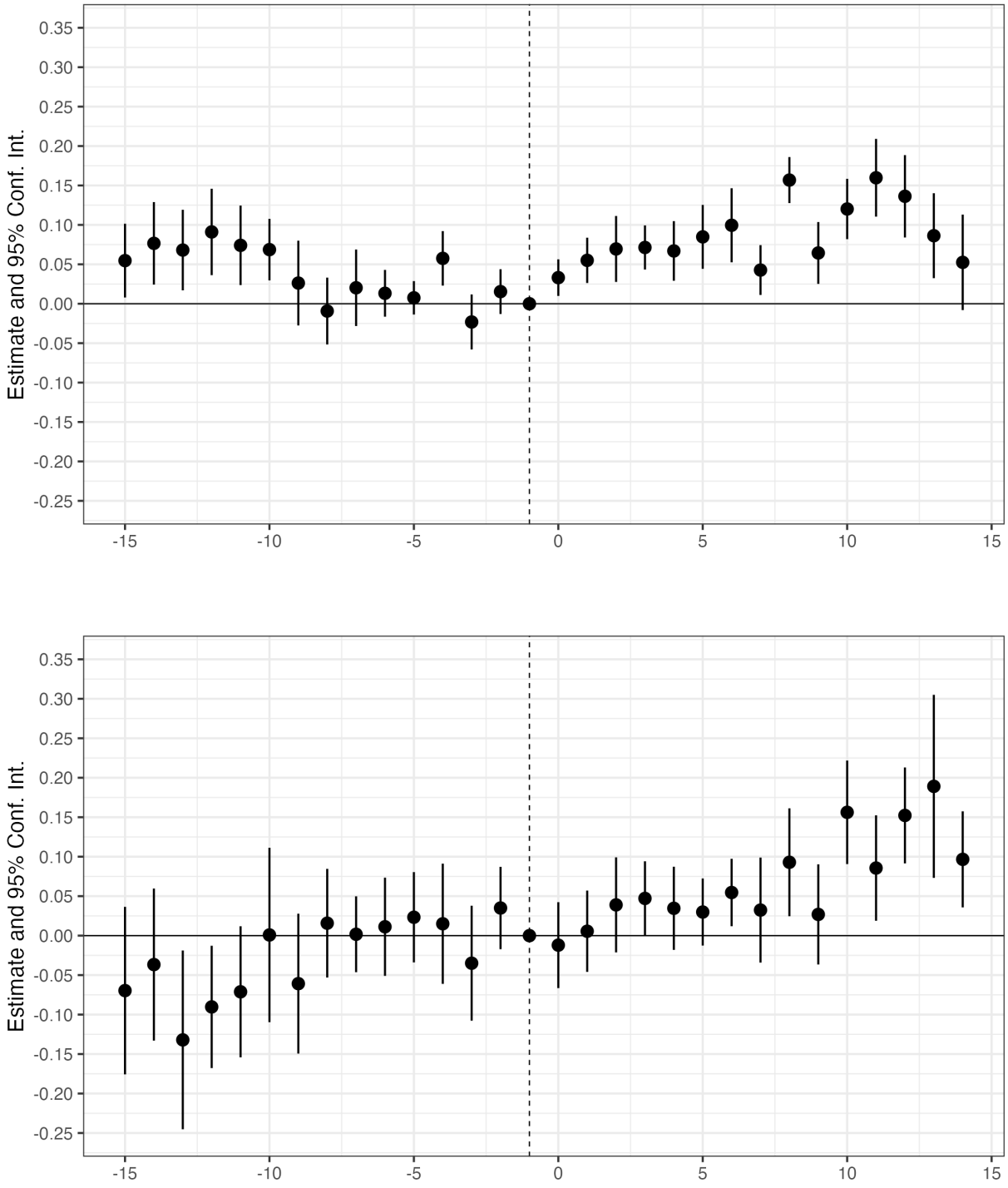
Notes: This figure shows the effect of exposure to Bible translation on full literacy on DHS clusters whose region is either minority (Panel A) or majority Muslim (Panel B). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.20: Effects on Years of Education By Muslim Regions



Notes: This figure shows the effect of exposure to Bible translation on years of education on DHS clusters whose region is either minority (Panel A) or majority Muslim (Panel B). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Figure B.21: Effects on Primary School Completion By Muslim Regions



Notes: This figure shows the effect of exposure to Bible translation on primary school completion on DHS clusters whose region is either minority (Panel A) or majority Muslim (Panel B). For further details on the methodology, see the empirical approach section. Confidence intervals are calculated using clustered standard errors at the ethnic group level and 95% confidence intervals.

Table B.1: Distribution on Pre-Colonial Measures

	Translated Before 1970 (N=45)		Translated After 1970 (N=86)		Never Translated (N=7)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Gathering Dependence:						
Gathering 0-5%	77.8	42.0	72.1	45.1	57.1	53.5
Gathering 6-15%	20.0	40.5	25.6	43.9	42.9	53.5
Gathering 16-25%	2.2	14.9	2.3	15.2	0.0	0.0
Hunting Dependence:						
Hunting 0-5%	33.3	47.7	43.0	49.8	85.7	37.8
Hunting 6-15%	53.3	50.5	52.3	50.2	14.3	37.8
Hunting 16-25%	13.3	34.4	4.7	21.2	0.0	0.0
Fishing Dependence:						
Fishing 0-5%	35.6	48.4	48.8	50.3	14.3	37.8
Fishing 6-15%	35.6	48.4	36.0	48.3	57.1	53.5
Fishing 16-25%	22.2	42.0	11.6	32.2	28.6	48.8
Fishing 26-35%	4.4	20.8	3.5	18.5	0.0	0.0
Agricultural Dependence:						
Agriculture 26-35%	0.0	0.0	0.0	0.0	14.3	37.8
Agriculture 36-45%	6.7	25.2	7.0	25.6	0.0	0.0
Agriculture 46-55%	24.4	43.5	16.3	37.1	28.6	48.8
Agriculture 56-65%	42.2	49.9	44.2	50.0	57.1	53.5
Agriculture 66-75%	15.6	36.7	15.1	36.0	0.0	0.0
Agriculture 76-85%	11.1	31.8	9.3	29.2	0.0	0.0
Intensity of Cultivation:						
Extensive/shifting	66.7	47.7	59.3	49.4	28.6	48.8
Horticulture	4.4	20.8	3.5	18.5	14.3	37.8
Intensive	28.9	45.8	31.4	46.7	57.1	53.5
Principal type of crop cultivated:						
Tree-fruit	4.4	20.8	9.3	29.2	14.3	37.8
Roots/tubers	22.2	42.0	16.3	37.1	14.3	37.8
Cereals	73.3	44.7	72.1	45.1	71.4	48.8
Marital composition of family:						
Limited polygny	0.0	0.0	10.5	30.8	14.3	37.8
Polygyny, sororal separate	4.4	20.8	3.5	18.5	0.0	0.0
Polygyny, non-sororal cohabit	73.3	44.7	65.1	47.9	14.3	37.8
Polygyny, non-sororal separate	17.8	38.7	17.4	38.2	42.9	53.5
Prevailing type of settlement pattern:						
Seminomadic	0.0	0.0	2.3	15.2	14.3	37.8
Semisedentary	4.4	20.8	4.7	21.2	0.0	0.0
Dispersed homesteads	35.6	48.4	31.4	46.7	28.6	48.8
Hamlets	20.0	40.5	10.5	30.8	28.6	48.8
Villages/towns	31.1	46.8	41.9	49.6	28.6	48.8

Table B.2: Summary of ethnic groups in the analytical sample for the simple approach to estimate the association between adult literacy and Bible translations before 1970

Percentage of observations from DHS individual data drop with the sample restriction < 10km	81.3
Total ethnic groups before matching with missions	214
Number of ethnic groups dropped with the sample restriction < 10km	80
Final number of ethnic groups in DHS after the sample restriction < 10km	138
Number of treated ethno-language groups with the sample restriction < 10km	45
Number of control ethno-language groups with the sample restriction < 10km	93

[a] This table summarizes the final analytical sample of ethnic groups for Approach I. The sample restriction "<10km" refers to the distance from the DHS cluster centroid to the mission geolocation.

Table B.3: Effects on Bible Translation Before 1970 on Full Literacy Rate

	Literate		
	(1)	(2)	(3)
Translation	0.072	0.030**	0.030**
	(0.047)	(0.012)	(0.015)
Mean Control	58.4	58.4	58.4
Std. Errors Cluster	Ethnic group	Ethnic group	Ethnic group
DHS Controls		✓	✓
Fixed Effects		✓	✓
Ethnolatlas Controls			✓
N	57 337	57 337	57 337
R^2	0.01	0.26	0.26

* p < 0.1, ** p < 0.05, *** p < 0.01

[a] Controls include: gender, urban or rural. Fixed effects include: DHS cluster by country, country , year of birth, and closest mission. Ethnoatlas control include societal characteristics such as reliance on hunting, gathering, fishing, and agriculture; marital composition, agricultural intensity, major crop types, and settlement patterns. Sample includes DHS Clusters within 10km of a former mission. Mean literacy for the sample is 62.7%

Table B.4: Effects on Bible Translation Before 1970 on Partial Literacy Rate

	Literate		
	(1)	(2)	(3)
Translation	0.110***	0.026**	0.026*
	(0.038)	(0.011)	(0.013)
Mean Control	69.2	69.2	69.2
Std. Errors Cluster	Ethnic group	Ethnic group	Ethnic group
DHS Controls		✓	✓
Fixed Effects		✓	✓
Ethnolatlas Controls			✓
N	57 337	57 337	57 337
R^2	0.02	0.27	0.27

* p < 0.1, ** p < 0.05, *** p < 0.01

[a] Controls include: gender, urban or rural. Fixed effects include: DHS cluster by country, country , year of birth, and closest mission. Ethnoatlas control include societal characteristics such as reliance on hunting, gathering, fishing, and agriculture; marital composition, agricultural intensity, major crop types, and settlement patterns. Sample includes DHS Clusters within 10km of a former mission. Mean literacy for the sample is 75.9%