French and British Colonial Legacies in Education: Evidence from the Partition of Cameroon  
  
Online Appendix

Yannick Dupraz

# A Data appendix

## A.1 Population census data

**Source and data quality.** I obtained the full 1976 population census from the Cameroonian National Statistical Institute via the MIMADEM project.[[1]](#footnote-1) For 2005, I use a representative 10% extract made available by IPUMS International.[[2]](#footnote-2) For the 1976 census, I had to identify and eliminate a lot of duplicate observations, all located in the Anglophone part. I do not know whether these duplicates were the result of an error during the coding process or an attempt to inflate population figures in certain districts.

**Geolocation.** I recovered the geographic coordinates of each village from its name using the website of the Cameroonian Ministry of Energy and Water (http://www.[mng-cameroon.org/SIG/](http://www.mng-cameroon.org/SIG/)). When I could not locate a village, I inferred its coordinates by taking the mean of villages in the same canton (a group of about 10 villages). In the 1976 census, in part of the Francophone Ouest region, I could not geolocate villages.[[3]](#footnote-3) 2005 census data cannot be geolocated more precisely than the district level.

**School attainment variables.** Census data give information on the last grade attended, from which I construct a measure of effective years of schooling (the theoretical number of years needed to attain the last grade attended and a standard measure of human capital). This measure takes into account the fact that, in 2005 as in 1976, the length of the primary cycle was 7 years in the Anglophone system versus 6 years in the Francophone system.[[4]](#footnote-4) However, primary-cycle length varied through time within French and British Cameroon, something the population census does not reflect.[[5]](#footnote-5) For that reason, I also consider the intensive margins on school participation (whether an individual has ever been to school) and school attainment variables (completion of the primary and secondary cycles).[[6]](#footnote-6) Because of differences in repetition rates (higher in the Francophone system), effective years of schooling need not translate into the same number of years actually spent in school. Using 2002–2003 data on the percentage of repeaters in each grade by region, gender, system (Francophone versus Anglophone) and order of schooling (general versus vocational), I construct a measure of years spent in school by adding to each effective year of schooling the region-grade specific repetition rate. If the repetition rate in grade 1 is 25%, pupils will spend on average 1.25 years in grade 1. Since I consider this measure when studying the cohort born between 1971 and 1980, I assume that repetition rates did not vary much between the 1980s and 2002, and that they were similar across districts of the same region (there were 10 regions in Cameroon in 2005).

**Skill content of occupation.** In 2005, I build a binary for having a high-skilled occupation (given one is working). High-skilled occupations are executive, manager, scientist/researcher, teacher/professor, engineer, physician, senior nurse or midwife, legal professional, writer/artist, or a member of the clergy. These are occupations having codes between 100 and 299 in the IPUMS dataset.

**Grade repetition.** In 2005, I construct a measure of grade repetition for students still in school (primary and secondary). A student is deemed to have repeated a grade if they are older than what their normal progression in the system would predict, taking into account across-district differences in school entry age. For each individual in district , I observe age and effective years of education . I compute, for each district , the average primary school starting age as the average age of individuals who are still in their first year of primary school. An individual is then considered to have repeated a grade if .

**Wealth index and public utilities.** I use information on house quality to build a wealth index for each household using the first component of a principal component analysis. The variables used are wall, roof and floor material, number of rooms per resident, type of toilet, light and drinking water sources, and, in 2005 only, type of lodging, cooking energy source, waste and waste water disposal systems, and accessibility. The wealth index is normalized to have zero mean and unit variance. I also decompose the index into a public and a private wealth index. The public wealth index uses housing quality variables which are strongly dependent on public utilities (drinking water, light and cooking energy sources, type of toilet, waste water disposal system, and accessibility), while the private wealth index uses all other variables. I also build two public utilities binary variables: a binary equal to one if the household’s source of drinking water is piped water (in the dwelling or via a public fountain), and a binary equal to one if the household’s source of light is electricity.

## A.2 Geographic data

**Temperature and precipitation.** I use gridded temperature and precipitation data from WorldClim ([http://www.worldclim.org](http://www.worldclim.org/)). Temperature is average temperature over 1950–2000, precipitation is average monthly precipitation over 1950–2000.

**Elevation and slope.** I use gridded elevation data from NASA Shuttle Radar Topography Mission, available on the CIAT-CSI SRTM website (http://srtm.csi.cgiar.org). I compute slope from the elevation data using the slope tool in ArcGIS.

**Malaria stability index.** I use the malaria stability index of Kiszweski et al. (2004) to proxy for the prevalence of malaria. The gridded data was downloaded from https://sites.google.com/site/gordoncmccord//datasets.

**Agricultural suitability.** I use gridded data on suitability for rainfed crops excluding forest ecosystems from Gobal Agro-Ecological Zones (http://webarchive.iiasa.ac. at/Research/LUC/GAEZ/index.htm). It is an index going from 0 (low suitability) to 11(high suitability).

**Night-time light.** I use gridded data on light intensity at night in 2005 from Version 4 DMSP-OLS Nighttime Lights Time Series: https://ngdc.noaa.gov/eog/dmsp/ downloadV4composites.html. I use ArcGIS to compute average pixel night-time light per district, and then multiply by the number of pixels in the district and divide by 2005 population to obtain night-time light per capita.

## A.3 Historical data

**German schools.** Schlunk (1914) gives the full list of missionary and government primary and secondary schools in Cameroon in 1911 (p. 85-113). From the name of the locality (Schulort), I recovered the latitude and longitude of each school using historical maps available on the Basel Mission Archives website (http://www.bmarchives.org) and the map available on the website of the Cameroonian Ministry of Energy and Water (http://www.mng-cameroon.org/SIG/). I was able to geolocate 84% of schools in Schlunk (1914) (92% if we exclude the district or Bezirk of Edea, far from the border). Non-located schools were given the geographic coordinates of their mission stations (all government schools were located).

**Public finance data.** I built series on colonial public expenditure (total expenditure and expenditure on education, discerning subsidies to the private sector) from the annual reports sent by both colonial powers to the League of Nations — the United Nations after World War II (France, Ministère des Colonies, 1921–1938, 1947–1957; Great Britain, Colonial Office, 1922–1938, 1949–1959), as well as French Cameroon’s budget estimates (Cameroun, various dates).

## A.4 Measures of historical school supply per district

To build measures of primary and secondary school supply per district at various dates, I use recent administrative school databases (from 2003 for primary schools, 2015 for secondary schools) giving the date of opening and status of every school in Cameroon. Assuming that few schools closed between their opening and today, I compute the number of schools per district at each date before using census data to divide by the corresponding school age population (6–11 for primary, 12–18 for secondary).

# B Additional maps and figures

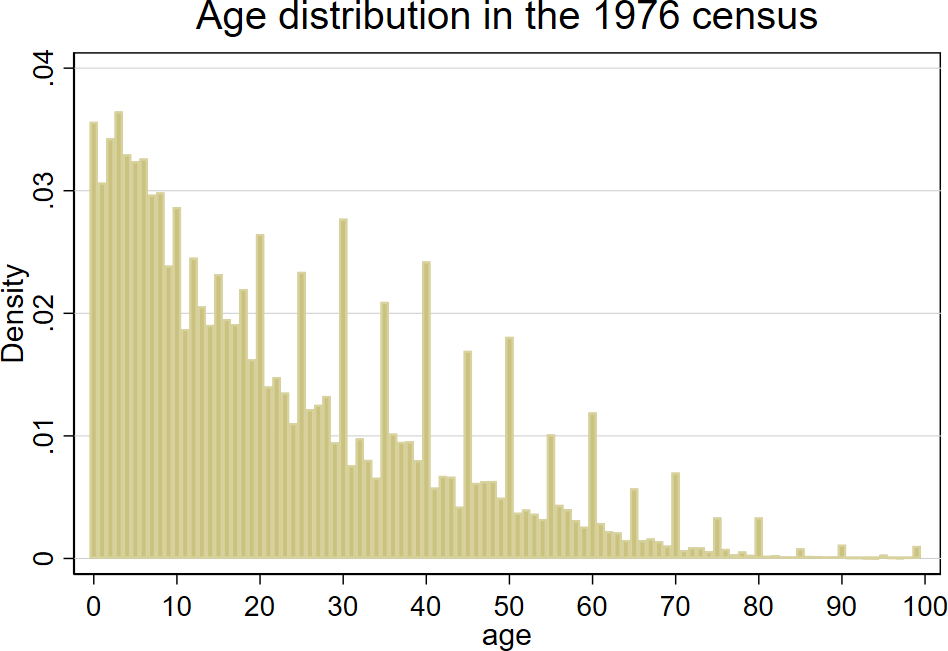
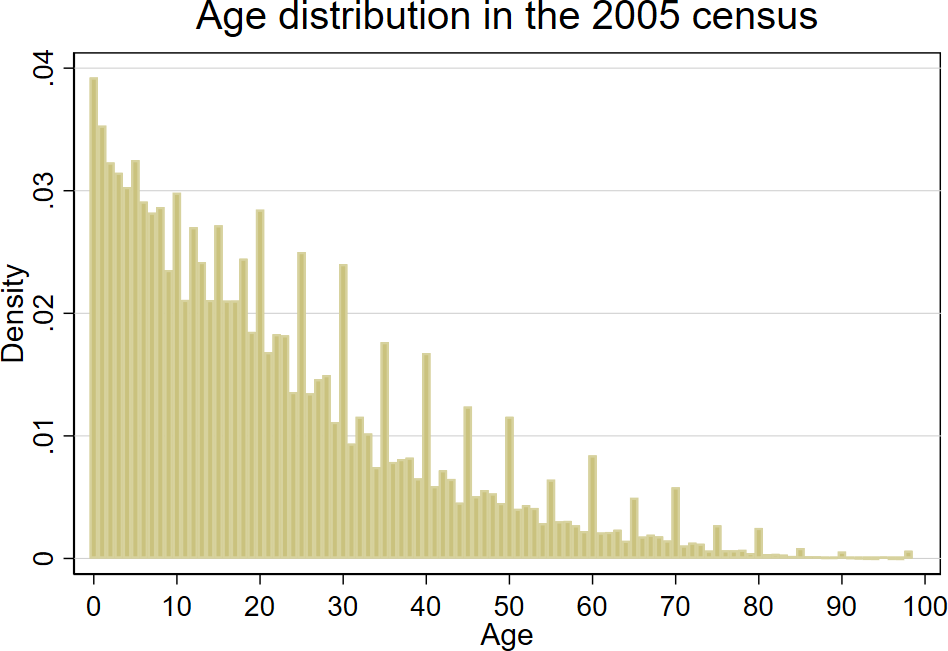
 

Figure B.1 — Evidence of age heaping in the 1976 and 2005 censuses

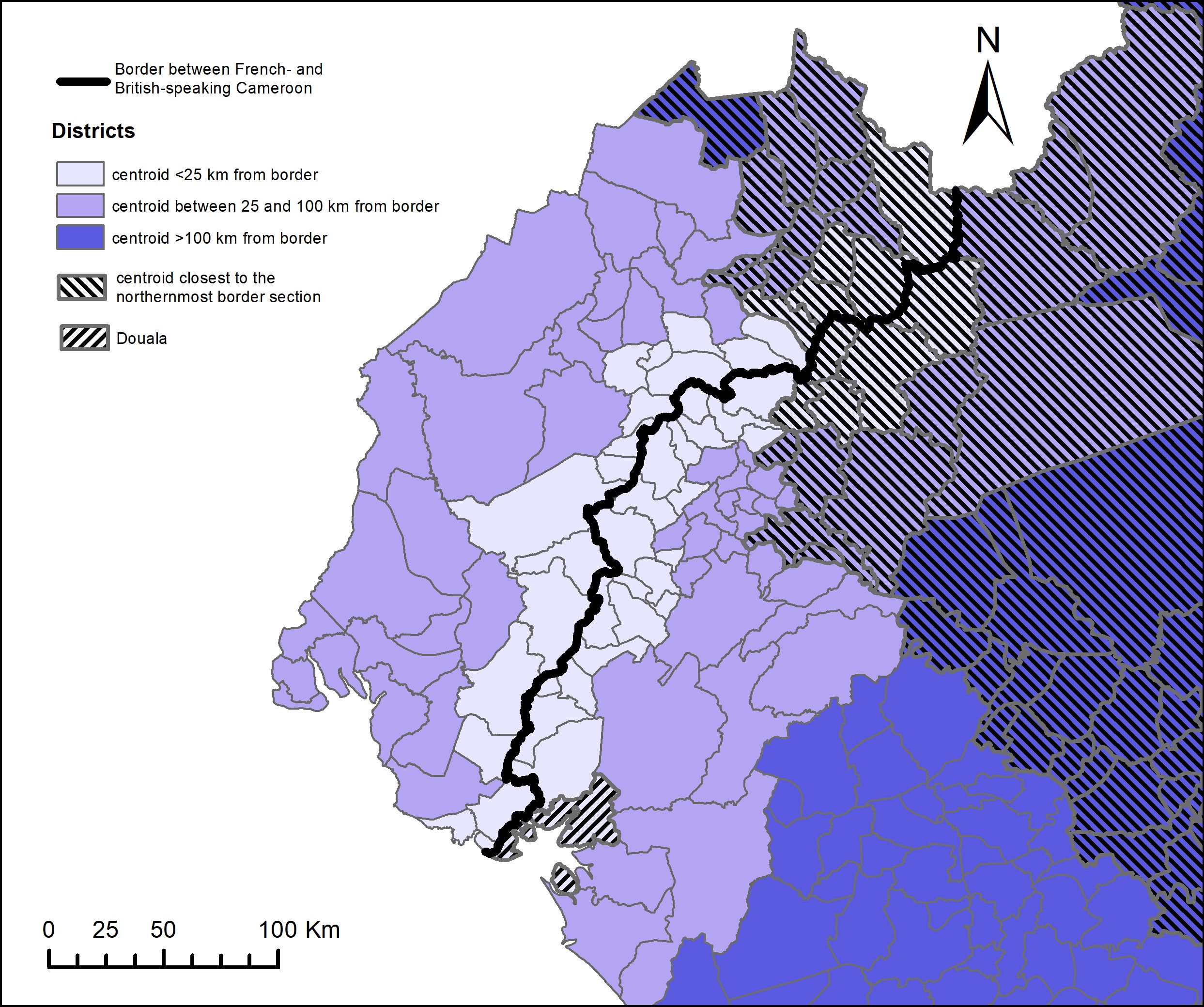
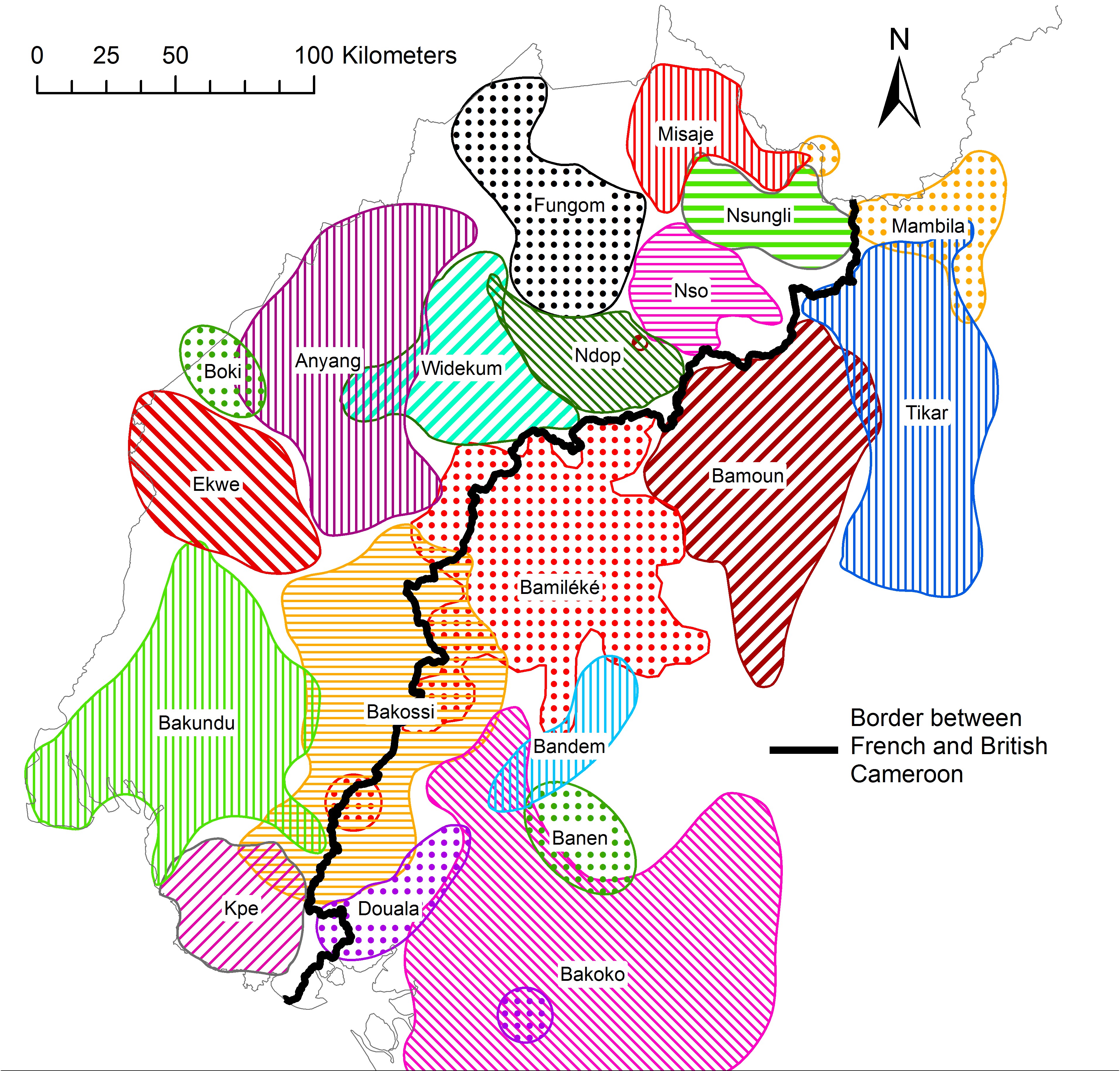


Figure B.2 — Districts in the 2005 population census

Figure B.3 — Ethnic groups in West Cameroon



Sources: Location of groups from population maps established by the Office de la recherche scientifique et technique outre-mer (ORSTOM) in the 1970s (ORSTOM, Atlas Régional du Cameroun). Definition of ethnic groups (“groups of essentially identical languages and culture”) from [Murdock (1959).](#_bookmark97) The “tribes’’ represented on the ORSTOM maps are alternatively tribes, chiefdoms, or ethnolinguistic groups — for example, the Bamilékés are represented as a single tribe, when this group is composed of several chiefdoms; the chiefdoms of Bangwa and Mundani, which are part of the Bamiléké group according to Murdock (1959) and Ngoh (1987), are represented as separate “tribes.”

# C Additional balance tests

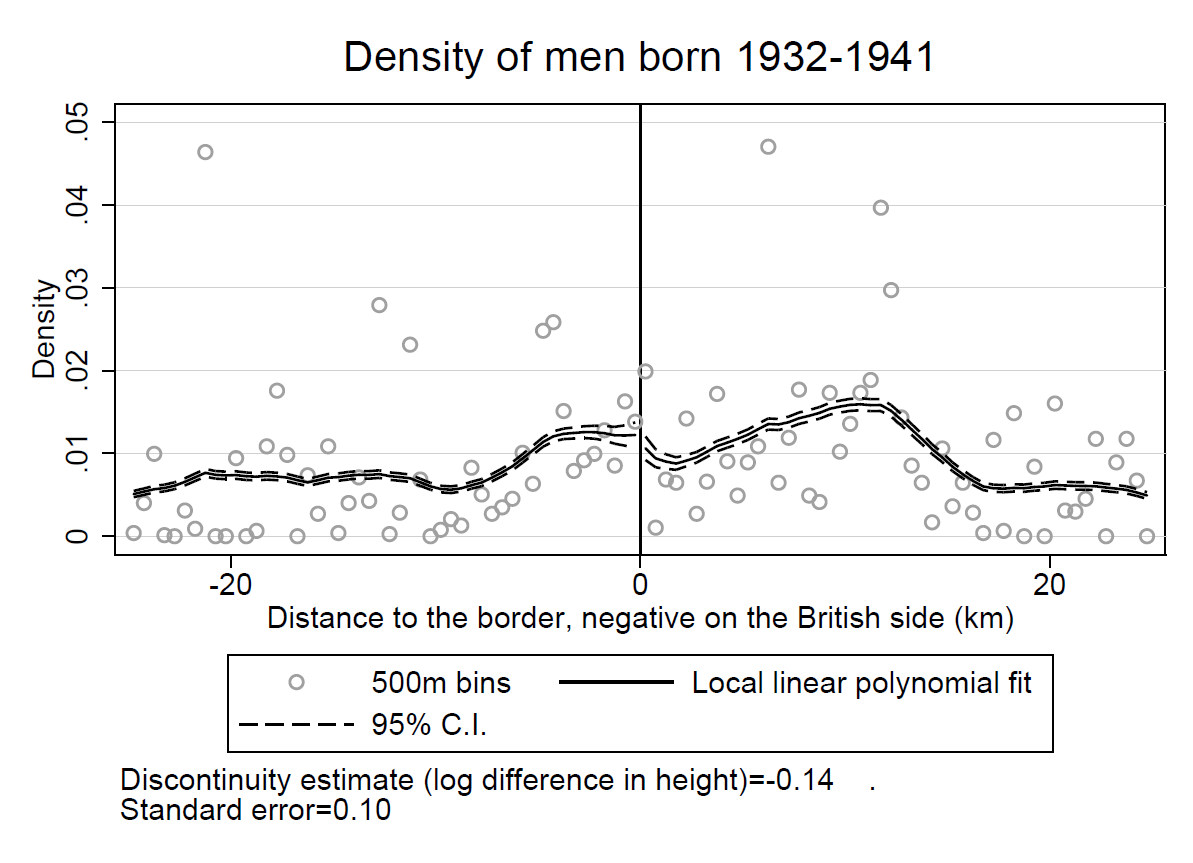
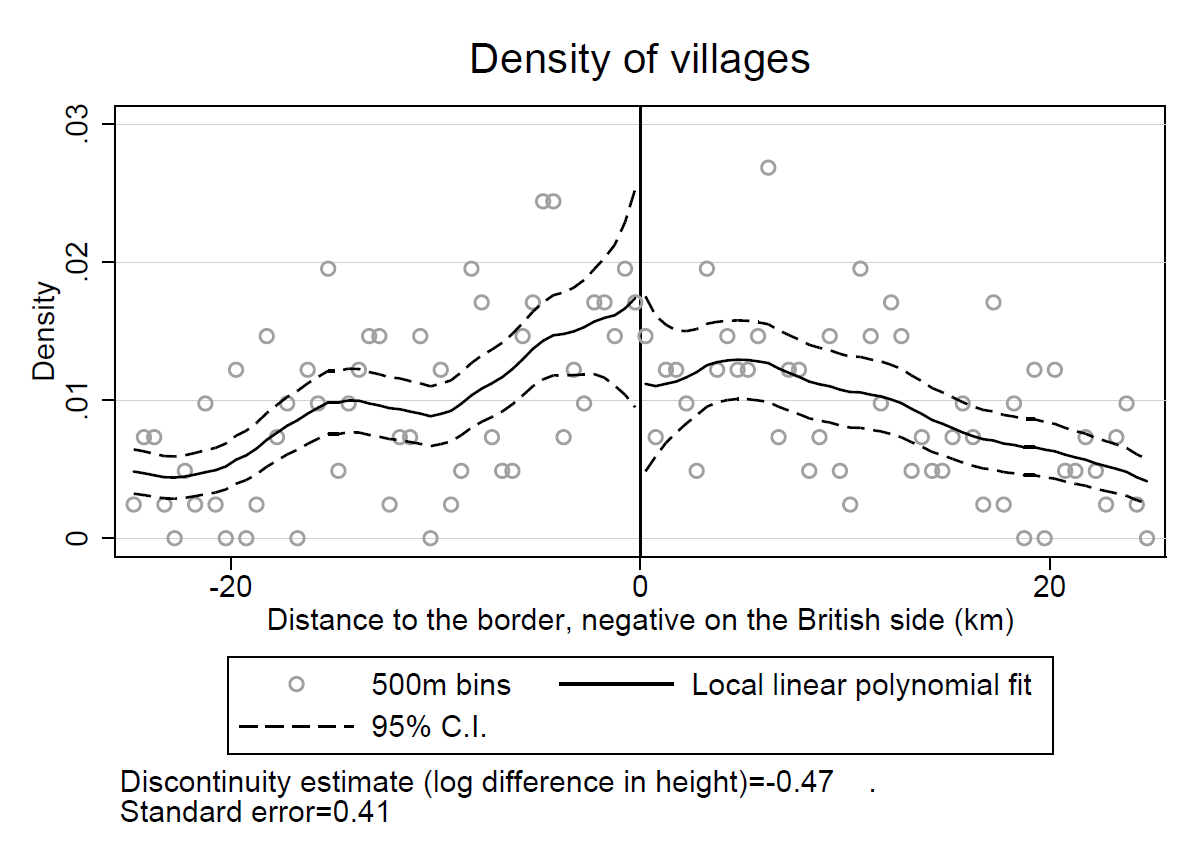
Table C.1 — Absence of discontinuity in the density of 1911 German schools

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|  | 10x10 km pixels | | | 1976 village-level sample | | | 2005 district-level sample | | |
|  | Mean within 30 km | |  | Mean within  15 km | |  | Mean within  20 km | |  |
|  | En. | Fr. | disc. | En. | Fr. | disc. | En. | Fr. | disc. |
| # German schools | 0,44 | 0.98 | -0.11 | 1.46 | 1.81 | 0.43 | 5.81 | 2.88 | 0.75 |
|  |  |  | (0.39) |  |  | (0.51) |  |  | (5.31) |
|  |  |  | [-0.86, 1.00] |  |  | [-0.63, 1.68] |  |  | [-12.40, 14.82] |
| Observations | 128 | 127 | 4,611 | 19,443 | 36,687 | 129,633 | 50,259 | 75,143 | 438,59 |

Notes: in columns (1) to (3), the unit of observation is the pixel. Robust standard errors in parentheses. The dependent variable is the number of 1911 German schools per pixel. In columns (4) to (6), the sample is the same as for results using 1976 census data: men older than 15 born in villages within 100 km of the southern section of the border (section 2). The dependent variable is the number of 1911 German schools within 5 km of the village. Standard errors clustered by village in parentheses. In columns (7) to (9), the sample is the same as for results using 2005 census data: all individuals older than 15 excluding those born in Douala and in districts closest to the northernmost section of the border (section 4). The dependent variable is the number of 1911 German schools per 2005 district. Standard errors clustered by district in parentheses. \*, \*\*, \*\*\*. Discontinuities obtained employing a local linear non-parametric estimation with triangular kernel and optimal mean squared error bandwidth. Robust Calonico et al. (2014) confidence interval in [ ]. Sources: Schlunk (1914).

*×*

Figure C.1 — McCrary tests for absence of discontinuity in density



*Notes*: The sample is the same as for the main results of the paper (figure 5 and table 2), which means discontinuities are estimated on border section 2 on figure 2. Bin size: 500 meters. Local linear fit with a triangular kernel and a 5 km bandwidth.

# D 1976 results: graphical representation and robustness

## D.1 1976 results: graphical representation

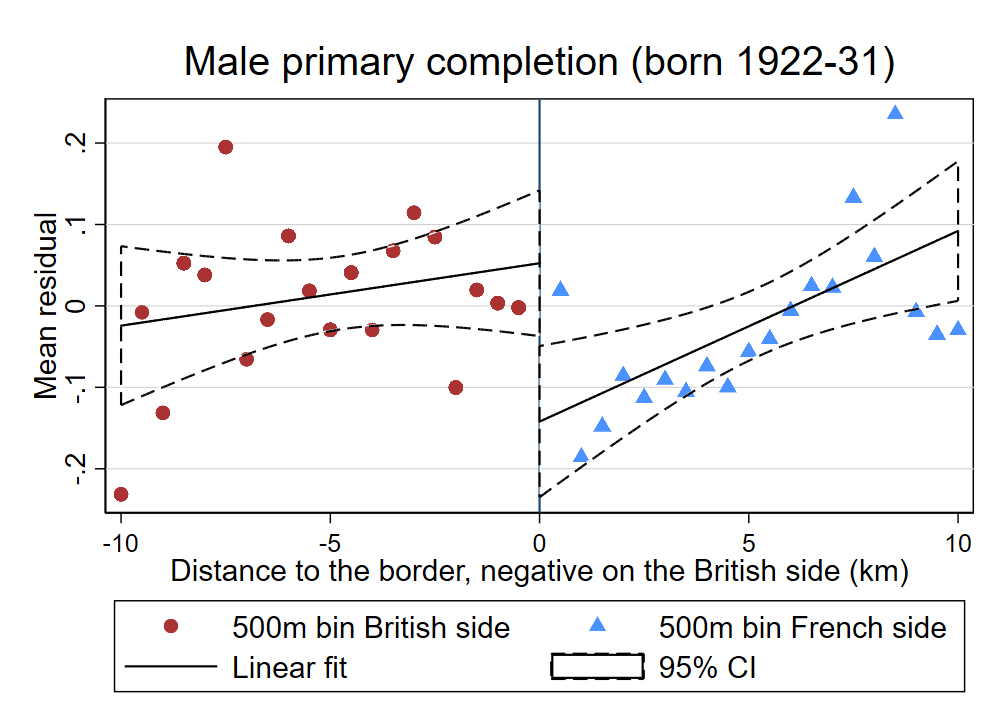
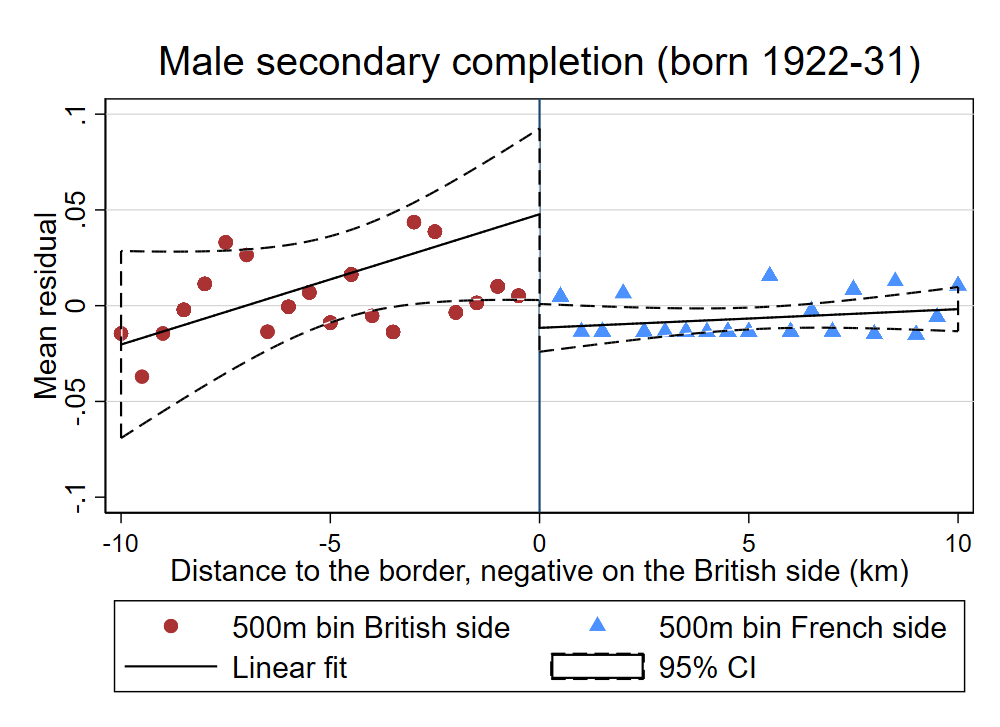
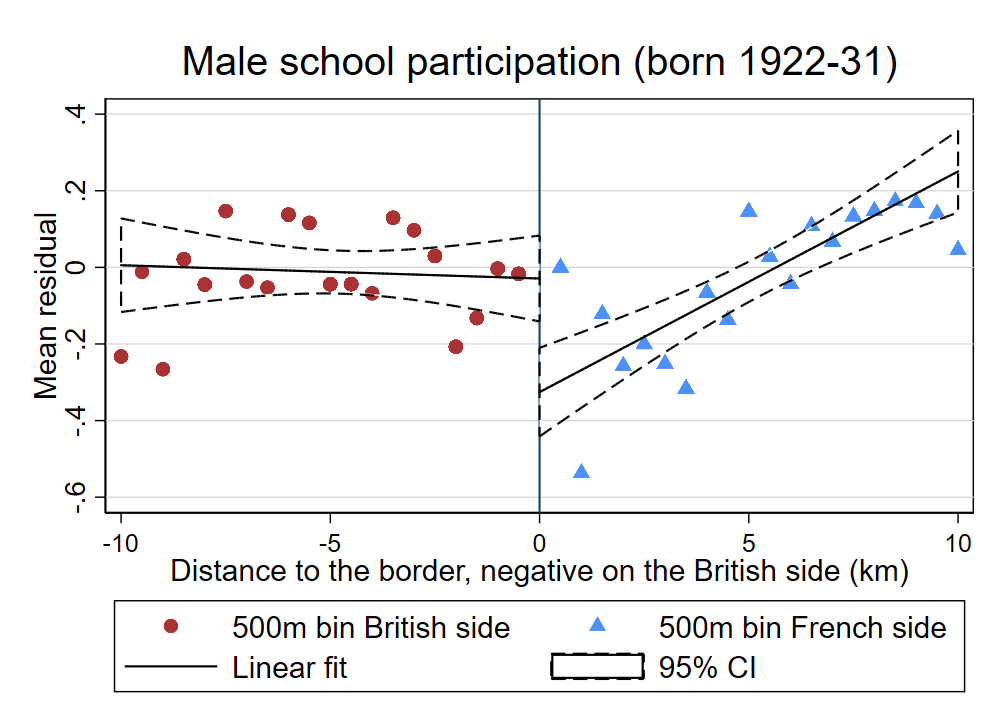
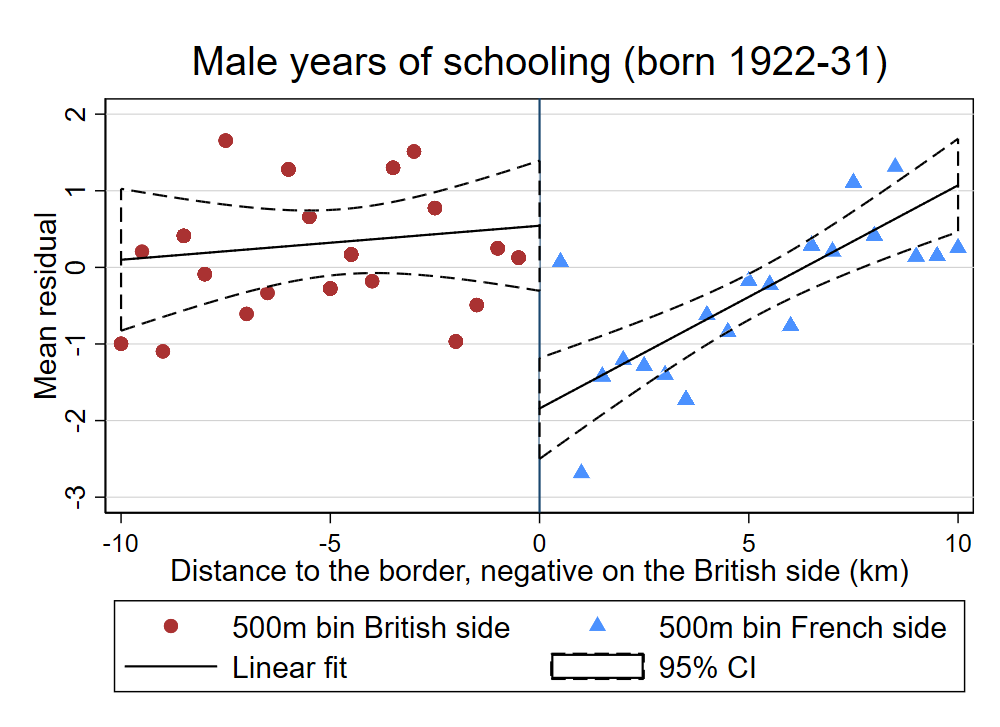


Figure D.1 — Graphical representation of 1976 discontinuities

Notes: The outcome variable is first regressed on the covariates (ethnic homeland fixed effects). Then each panel graphs the average value of the residuals for each 0.5 km bin on both sides of the border. The line on either side is fitted on observations collapsed by village (to take clustering into account). Distances are negative on the British side and positive on the French side and the border corresponds to a distance of zero. These graphs do not correspond directly to the nonparametric RD results, which use an optimal bandwidth and a triangular kernel.

## D.2 1976 results: robustness to other specifications

Table D.1 — Discontinuities in education in 1976: robustness to latitude-longitude specifications

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  | means on sample | | estimated discontinuities | | | |
|  | within <5 km of border | | <5 km from border | | <10 km from border | |
|  | British side | French side | (x,y) poly of degree 1 | (x,y) poly of degree 2 | (x,y) poly of degree 1 | (x,y) poly of degree 2 |
|  | **Men born 1922–1931** | | | | | |
| Years of education | 2.89 | 1.98 | 1.13\*\*\* | 1.19\*\*\* | 1.24\*\*\* | 1.25\*\*\* |
|  |  |  | (0.22) | (0.19) | (0.20) | (0.20) |
|  |  |  | [0.23] | [0.19] | [0.24] | [0.23] |
| Observations | 1,110 | 686 | 1,796 | 1,796 | 3,205 | 3,205 |
|  |  |  |  |  |  |  |
|  | **Men born 1942–1951** | | | | | |
| Years of education | 6.50 | 6.89 | 0.35\*\* | 0.21 | 0.26\* | 0.20 |
|  |  |  | (0.17) | (0.16) | (0.14) | (0.15) |
|  |  |  | [0.17] | [0.14] | [0.18] | [0.17] |
| Observations | 1,780 | 1,216 | 2,996 | 2,996 | 6,400 | 6,400 |
|  |  |  |  |  |  |  |
| # of villages | 65 | 45 | 110 | 110 | 195 | 195 |

Notes: sample of men born close to the southern section of the border. All regressions control for ethnic homeland fixed effects. Standard errors clustered by village in parentheses. \*, \*\*, \*\*\*. Conley standard errors in [ ] (cut-off window of 20 kms).

## D.3 1976 results: placebo borders

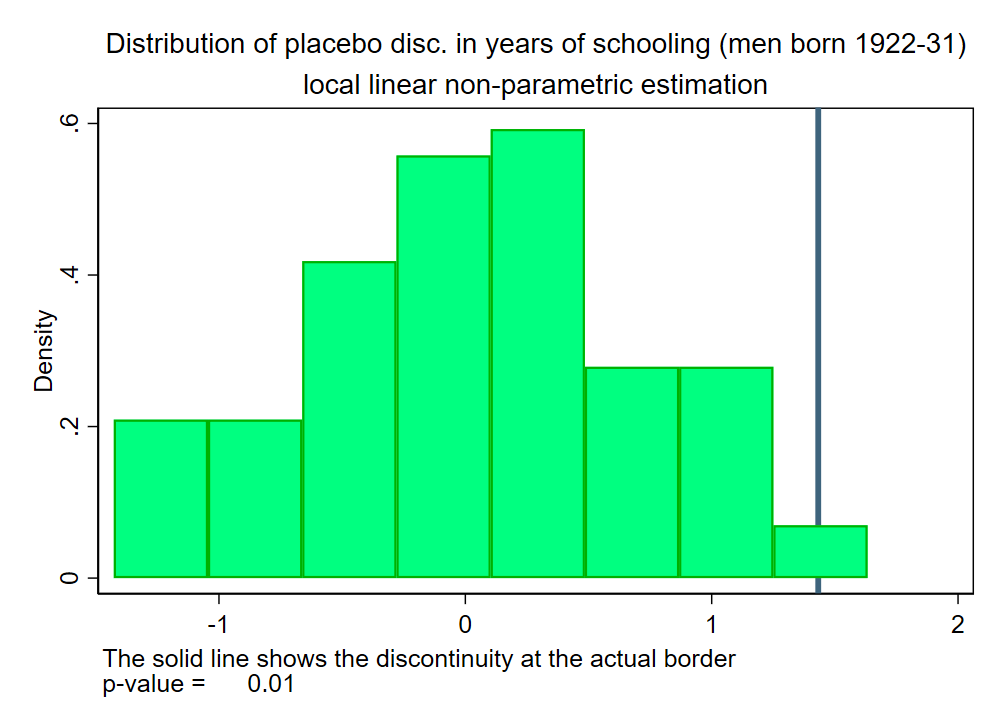


Figure D.2. — Placebo borders for 1976 results

Notes: this graph displays the distribution of discontinuities estimated on placebo borders running parallel to the actual border, but shifted East or West by up to 50 km in steps of 1 km. The solid line shows the discontinuity estimated at the actual border. The p-value for a two-sided test is simply the percentage of placebo discontinuities whose absolute value is larger than the absolute value of the actual discontinuity.

## D.4 Selection by mortality

Estimating results on different age groups observed at the same point in time creates the concern of potential selection by mortality, a problem similar to selective attrition. This is mostly a concern for the emergence of a British advantage in education after the partition of Cameroon, a result obtained on cohorts born between 1912 and 1931. In the census, I only observe the members of this cohort who survived up to 1976, when they were between 45 and 65 years old. We have to consider two types of mortality: peacetime mortality and mortality during the independence war. Peacetime mortality likely selects the poorest and least educated: it could explain the positive British effect for older cohorts if mortality was higher in British Cameroon than in French Cameroon. However, the only study comparing health systems in colonial Cameroon concludes it was better in the British part (Nzima Nzima, 2014).[[7]](#footnote-7) Moreover, if life expectancy was lower in British Cameroon, we would expect the age pyramid to be thinner at the top on the British side. To check that it is not the case, I estimate discontinuities in the share of each cohort in total population. Figure D.3 shows that, though age pyramids are roughly similar on either side of the border, the share of men aged 45–54 (born 1922–1931) is about 6 percentage points larger on the British side, while the share of men aged 35–44 (born 1932–1941) about 8 percentage points smaller on the British side. However, I estimate a large and positive discontinuity in years of schooling for both of these cohorts, as well as for the cohort born 1912–1921, which is the same size on each side. It is therefore unlikely that my results are explained by higher mortality in British Cameroon.

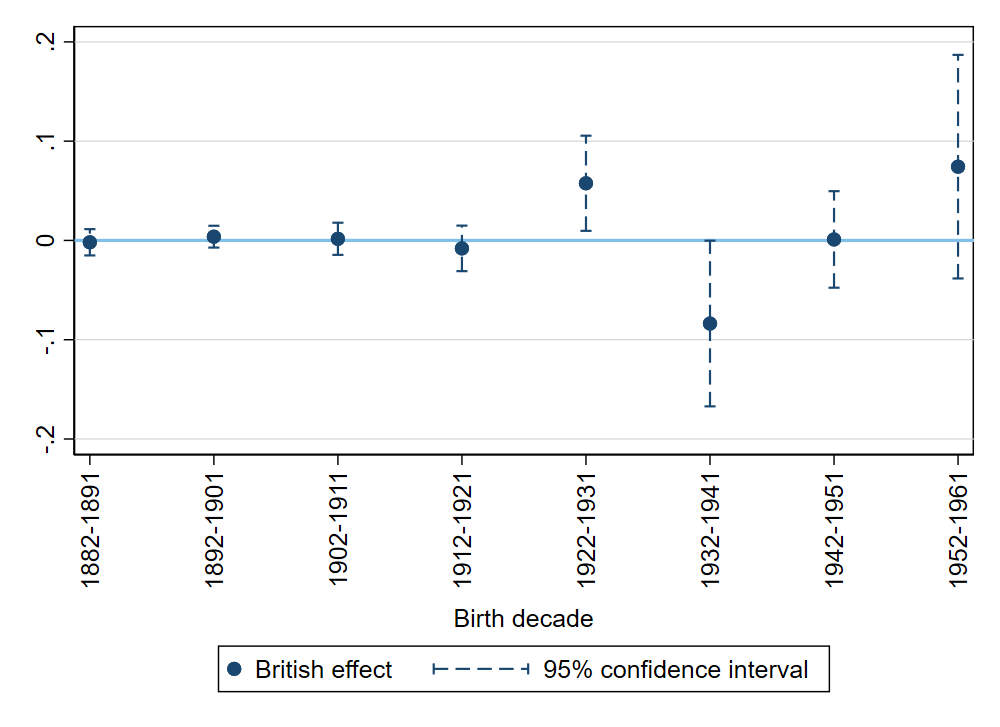


Figure D.3 — Border discontinuities in cohort share of total population

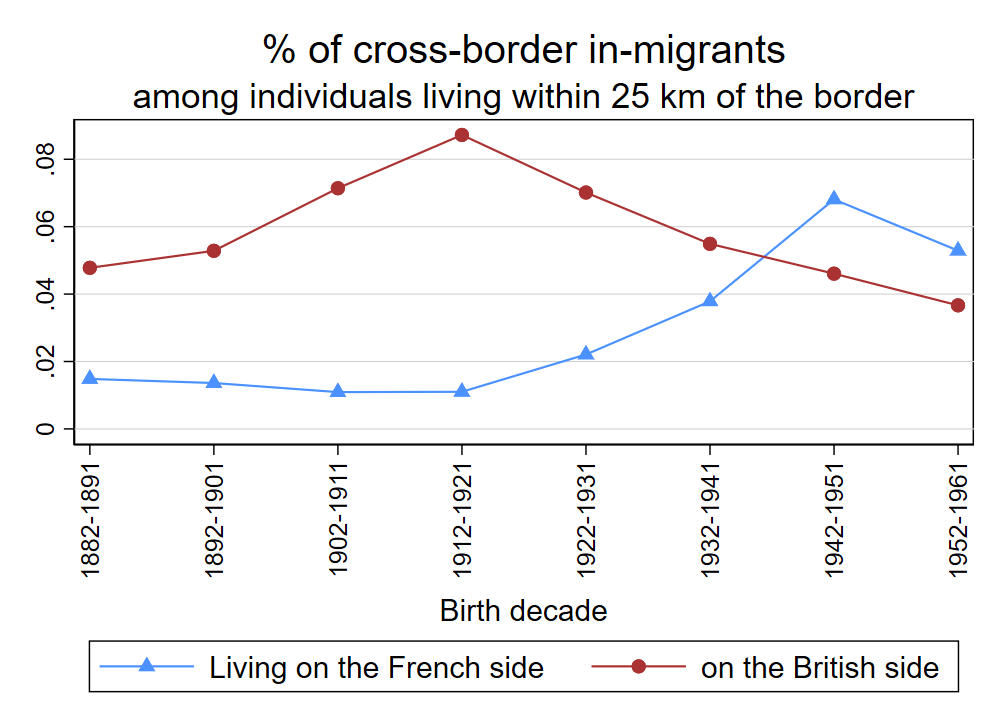
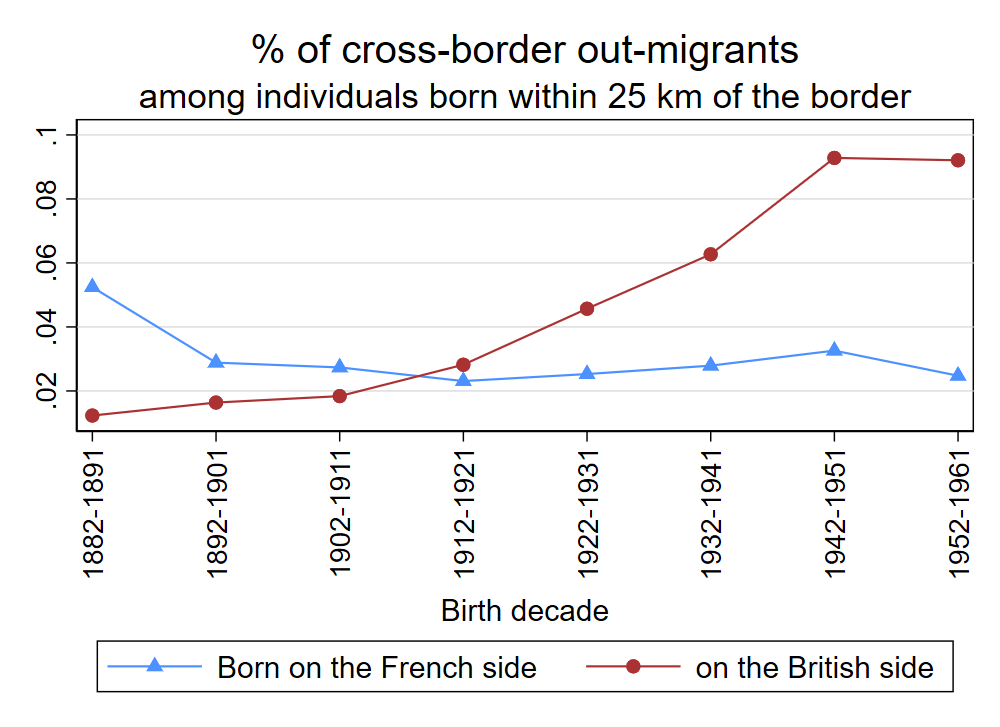
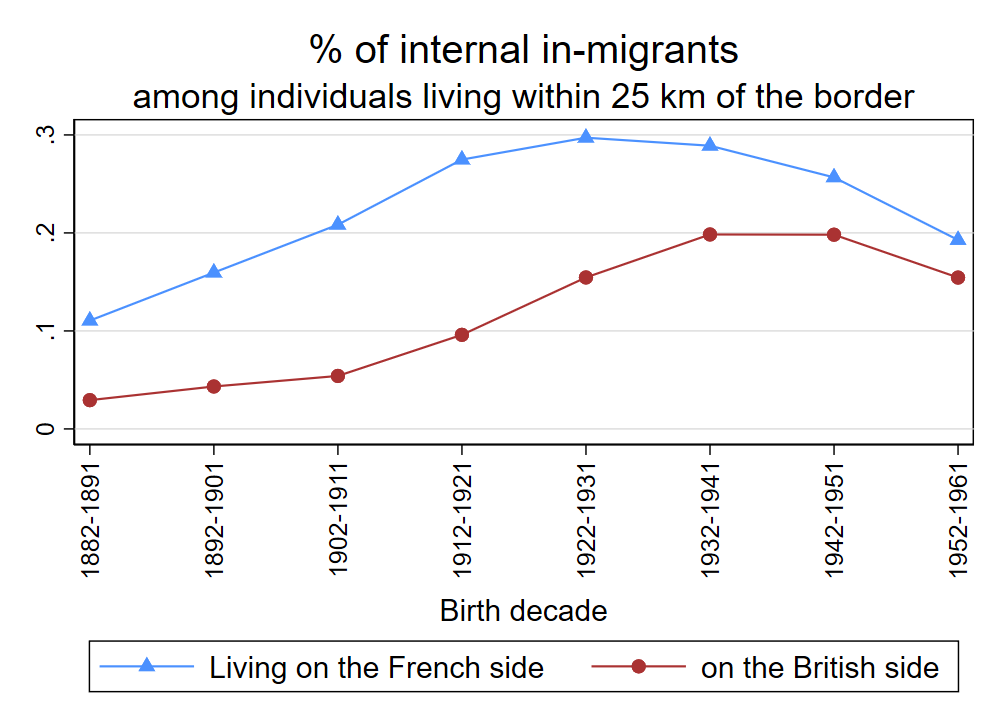
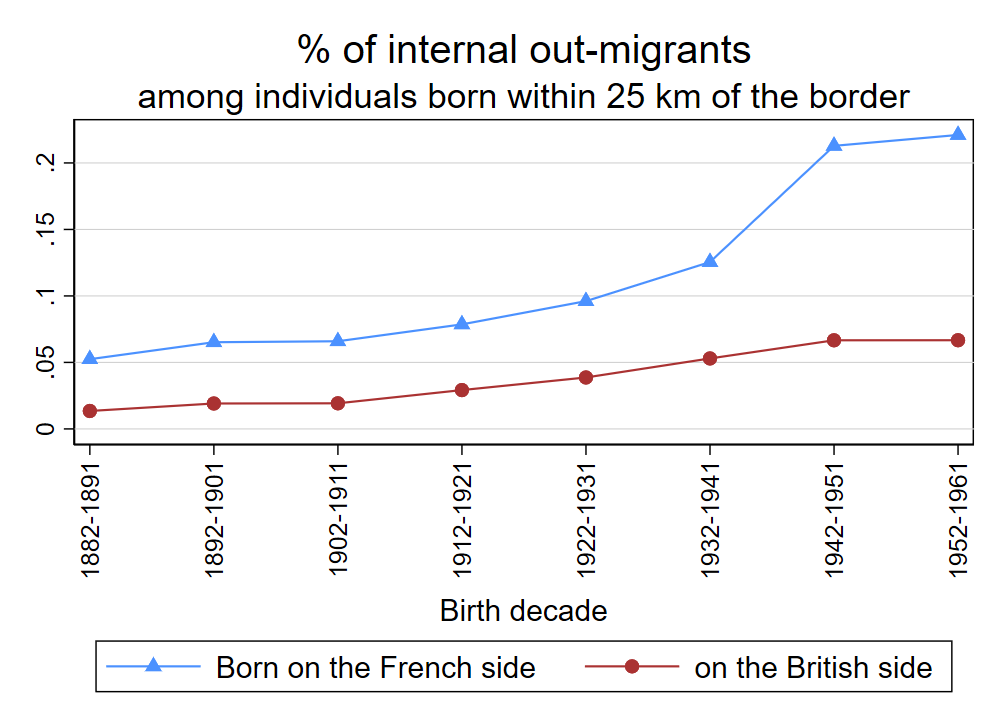
Notes: The dependent variable is the share of each cohort in the total male population of a village. Sample of villages within 100 km of the southern section of the border. Discontinuities obtained using a local linear non-parametric estimation with triangular kernel and optimal MSE bandwidth. Added covariates: ethnic homeland fixed effects.

What about wartime mortality? An estimated 100,000 to 200,000 people died during the independence war (Deltombe et al., 2011), which took place mostly in the French part. The conflict was most violent in the Bamiléké region, which is not part of my sample.[[8]](#footnote-8) For wartime mortality, the direction of selection is not obvious a priori. If wartime mortality selected the most educated, it could explain the positive British effect for older cohorts. It would be a concern mostly for the cohort born between 1932 and 1941, whose members were of fighting age during the most violent part of the conflict around 1960, but for this cohort, in my sample, the age pyramid is actually thicker on the French side (figure D.3). If wartime mortality selected the least educated, it could explain the convergence in education, but this convergence is also estimated for cohorts born after 1950, who did not take part in the conflict.

## D.5 Migration

**Within Cameroon migration.** Using 1976 census data, figure D.4 gives a better idea of migration patterns in the border region (defined as districts within 25 km of the border). The first panel displays, on both sides of the border and in each cohort, the percentage of internal out-migrants from the border region: people who were born close to the border region and moved away, but did not cross the border between French and British Cameroon. The second panel displays the percentage of internal in-migrants to the border region: people who were born away from the border but moved to the border region, without crossing the border. The third panel displays the percentage of cross-border out-migrants from the border region, and the fourth panel the percentage of cross-border in-migrants to the border region.[[9]](#footnote-9)

Figure D.4 – Internal and cross-border migration in the 1976 population census



Notes: migrants are people whose district of residence in 1976 is different from their district of birth. Internal migrants still live on the side of the border they were born in. Internal out-migrants were born within 25 km of the border (Douala excluded) and live more than 25 km away from it in 1976. Internal in-migrants live within 25 km of the border in 1976 (Douala excluded) and were born more than 25 km away from it.

Internal migration was overall more widespread than cross-border migration, which offers evidence that the border put in place by the colonizers had some reality for Cameroonians. Internal migration away from and to the border region was more widespread in French Cameroon, reflecting the fact that the French part was much larger than the British part, and offered more opportunity for migration.

Cross-border migration was overall more limited, but there was migration from French Cameroon to British Cameroon, peaking for the cohort 1912–1921 in which about 9% of individuals living close to the border in British Cameroon had migrated from French Cameroon (last panel of figure D.4). Such migration perhaps reflects the higher wages offered by German-owned plantations in British Cameroon (Le Vine, 1964). Migration from the British side of the border to the French side was particularly important for people who became adults after reunification: almost 10% of individuals born close to the border in British Cameroon between 1942 and 1961 lived in the former French part in 1976 (third panel of figure D.4).

Because I conduct a border discontinuity on the place of birth of individuals, these migration patterns matter only insofar as they affect the parents’ generation. Higher internal migration in French Cameroon could explain the emergence of a positive British effect if, among the parents of those for whom I estimate a discontinuity, the ones with a high preference for education migrated away from the border region before their child was born, or the ones with a low preference for education migrated to the border regions. Internal migrants were on average more educated than those who stayed, making the second scenario unrealistic.[[10]](#footnote-10) As for the first one, differences in education preferences would need to be very large to explain the divergence in education. Even if out-migrants were giving their children 6 more years of education than those who stayed (an extreme assumption) the 4% difference in internal out-migration rates between French and British Cameroon for older cohorts (first panel of figure 6) would account for a difference of years of education only.

Migration from French to British Cameroon could explain the emergence of a positive British effect if migrants from French Cameroon had a higher preference for education than locals, but this does not seem to have been the case: in the sample of men living in British Cameroon within 25 km of the border and born before 1912, regressing years of schooling on district of residence fixed effects and a dummy equal to one if an individual was born in French Cameroon yields a coefficient of (standard error clustered at the district level of ).

When using 1976 census data, I repatriate migrants to their village of origin assuming constant migration rates across villages of the same district. Figure D.5 shows what happens to the discontinuities in years of schooling when we exclude out-of-district migrants completely rather than repatriating them: for the first 5 cohorts, results are virtually unchanged, but for the cohorts born after 1932, rather than falling to zero, the British effect remains positive and significant, which confirms that the higher migration rates in French Cameroon in these cohorts (first panel of figure D.4) selected away the more educated individuals.

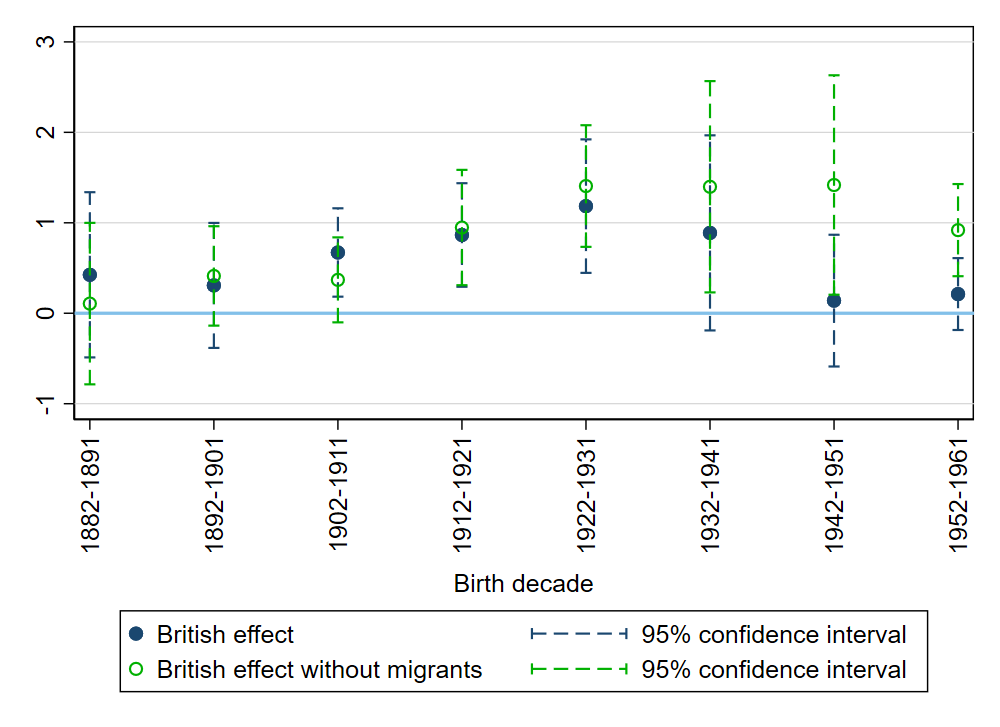


Figure D.5 — Discontinuities in 1976 with and without migrants

Notes: this graph compares discontinuities obtained on the sample of all individuals (including out-of-district migrants reallocated in their village of birth), with border discontinuities estimated on the sample of those who remained in their district of birth. Sample of men born in villages within 100 km of the southern section of the border. Discontinuities obtained employing a local linear non-parametric estimation with triangular kernel and optimal MSE bandwidth. Added covariates: ethnic homeland fixed effects. Standard errors clustered at the village level.

**Migration from/to Nigeria.** During the mandate period, a number of Nigerians moved to British Cameroon, which was administered as part of Southeast Nigeria, to work in road construction, on plantations, and in the civil service. In the 1952/1953 census of Nigeria, Nigerian “tribes” represent 11% of the population of Southern British Cameroon (Great Britain, Colonial Office, 1953).[[11]](#footnote-11) In the 1976 census, individuals born in Nigeria represent 6.5% of the population of the former British part. Again, individuals born in Nigeria are not actually the problem here as I conduct a border discontinuity on the place of birth of individuals, but the concern is that Nigerian migrants had a specific preference for educating their children born in Cameroon. In the 1976 census, I find that Nigerian migrants of the generation of the parents (born before 1912) were more educated that locals.[[12]](#footnote-12) To have an idea of the potential bias, we need to assume the percentage of Cameroon-born Nigerian-origins individuals in the relevant cohorts of the 1976 census. Assuming a figure of 5%, and assuming that Nigerian-born parents were giving their children 6 more years of education than locals (again, an extreme assumption), Nigerian migration would account for a British effect of years of education.[[13]](#footnote-13)

Another problem arises because a sizable number of Nigerians left Cameroon around reunification, as British Cameroonians pushed for an indigenization of the civil service and threatened to deport the Igbo migrants from southeastern Nigeria (Amaazee, 1990; Chem-Langhee, 1976). Individuals born in Cameroon who went back to Nigeria after 1960 would not be observed in the 1976 census. In 1953, Nigerian “tribes” represented 84,200 individuals in Southern British Cameroon. In the 1976 census, in the same area, individuals born in Nigeria before 1953 represented 45,265 individuals. It does not mean that 40,000 Nigerians left Cameroon, as the 1953 census counts people of Nigerian origins born in Cameroon while the 1976 does not, and as we need to take mortality into account. To construct an upper bound estimate of the bias, we can assume that the Nigerian-origins Cameroon-born individuals who left represented 5% of their cohort and that they had 6 more years of schooling than locals. The British effect would be underestimated by years of education.

## D.6 Replication of 1976 results at the district level

A concern is that 2005 district-level results are not directly comparable to 1976 village- level results. One first problem is that they are not obtained on exactly the same border section (because some data cannot be geolocated at the village level in 1976, see figure 2). A second problem is that the geographical variation is not the same (village-level versus district-level variation). On figure D.6, I replicate 1976 results at the 1976 district of birth level. Because of district splitting between 1976 and 2005, the number of districts within 100 km of the border is much lower in 1976 than in 2005 (56 versus 115). To retain enough statistical power, I consider all districts within 100 km of the border (excluding Douala), and control for geographic location using a linear function of distance to the border whose slope is allowed to vary on either side. Figure D.6 replicates the village-level results of figure 5, though the positive British effect for cohorts born between 1912 and 1941 is smaller and imprecisely estimated. On figure D.7, to increase the number of districts, I assign each 1976 village to its 2005 district and run the RD analysis as if I only had geographical data at the level of the 2005 district. This allows me to estimate 1976 discontinuities using the same specification and sample (of districts) as for 2005 results, excluding Douala and the northernmost border section.[[14]](#footnote-14) Figure D.7 replicates the village-level results of figure 5: men born between 1922 and 1941 have one more year of education if they were born on the British side of the border. The only difference is that I also estimate a positive border discontinuity of about 1 year for the cohort born 1942–1951, whereas I estimate no discontinuity when focusing on the southern border segment and using village-level variation.

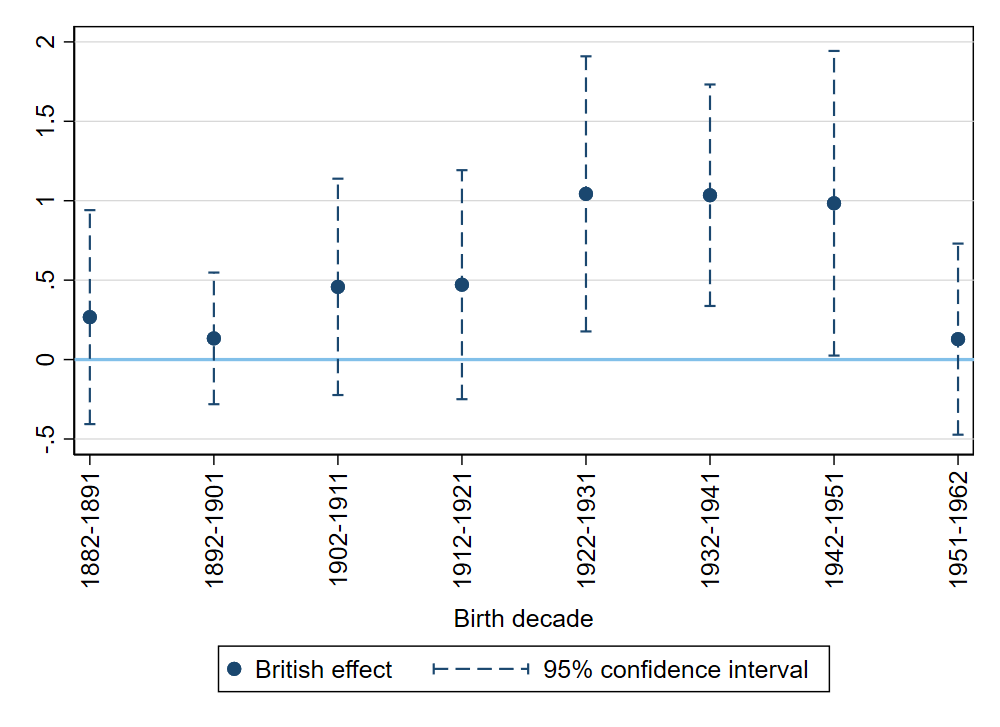
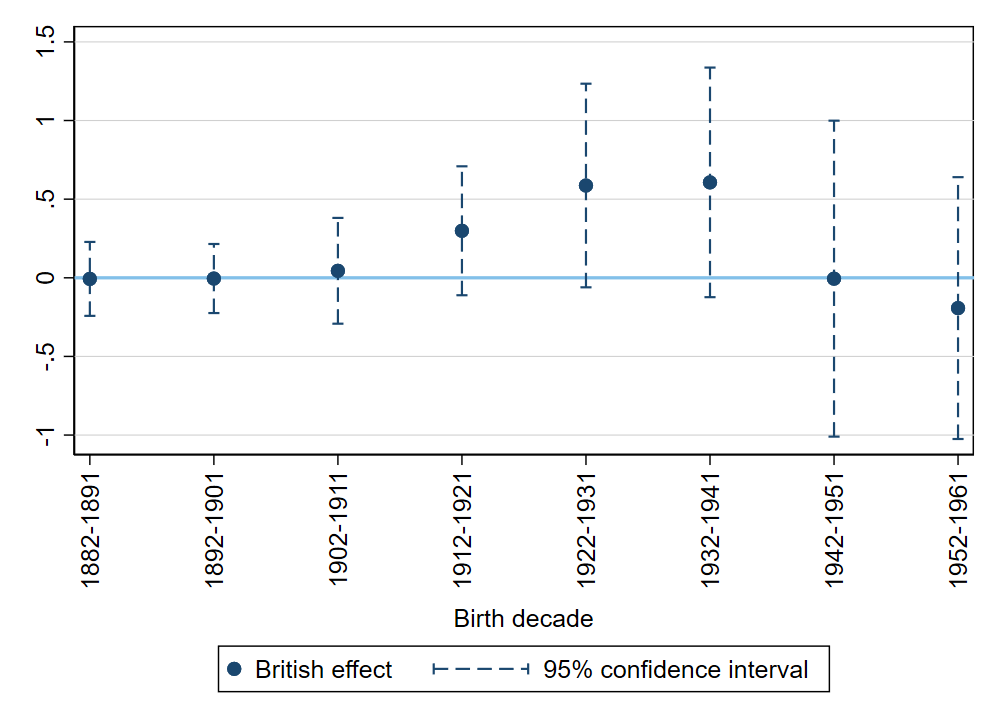


Figure D.6 — Discontinuities in male years of schooling in 1976, 1976-district-level specification

Notes: Sample of men born in districts located within 100 km of the border. Discontinuities estimated by controlling for a linear function of distance to the border with a different slope on each side. Added covariates: border section dummies. Standard errors clustered at the district level.

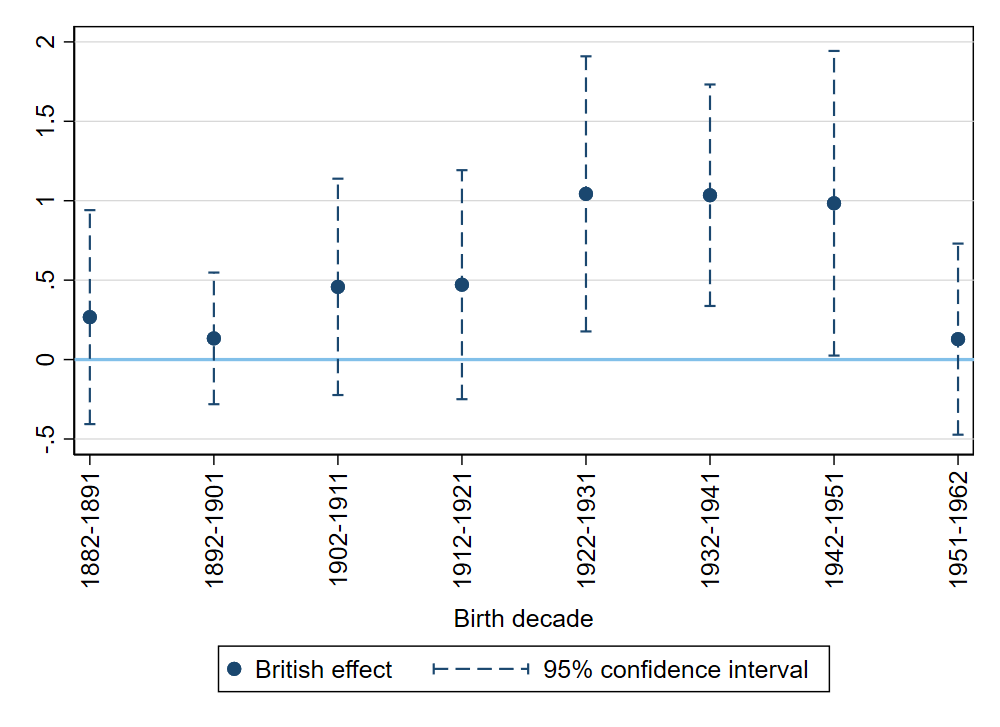


Figure D.7 — Discontinuities in male years of schooling in 1976, 2005-district-level specification

Notes: the sample excludes individuals born in Douala and in districts closest to the northernmost section of the border. Discontinuities obtained employing a local linear non parametric estimation with triangular kernel and a 20 km bandwidth (the bandwidth used to obtain the results of table [3](#_bookmark123) varies between 14 and 27 km). Added covariates: border section dummies. Standard errors clustered at the district level.

# E 2005 results: robustness and graphical representation of results

## E.1 2005 results: graphical representation

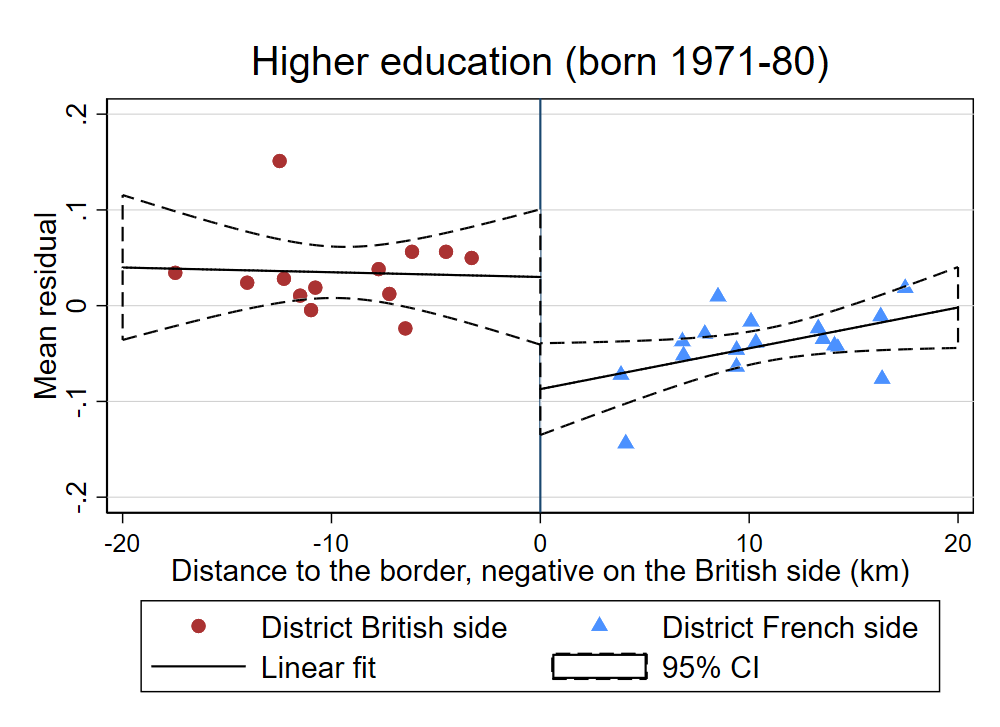
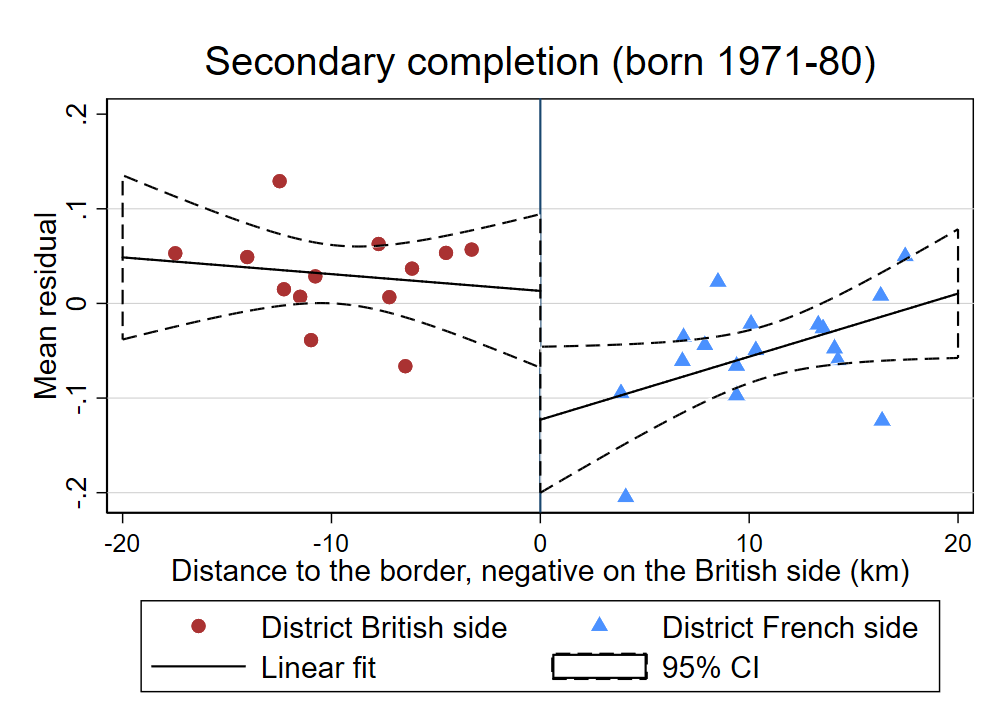


Figure E.1 — 2D visual representation of discontinuities in 2005

Notes: The outcome variable is first regressed on the covariates (a border section dummy). Then each panel graphs the average value of the residuals for each 0.5 km bin on both sides of the border. The line on either side is fitted on observations collapsed by village (to take clustering into account). Distances are negative on the English-speaking side and positive on the French-speaking side and the border corresponds to a distance of zero. These graphs do not correspond directly to the nonparametric RD results, which use an optimal bandwidth and a triangular kernel.

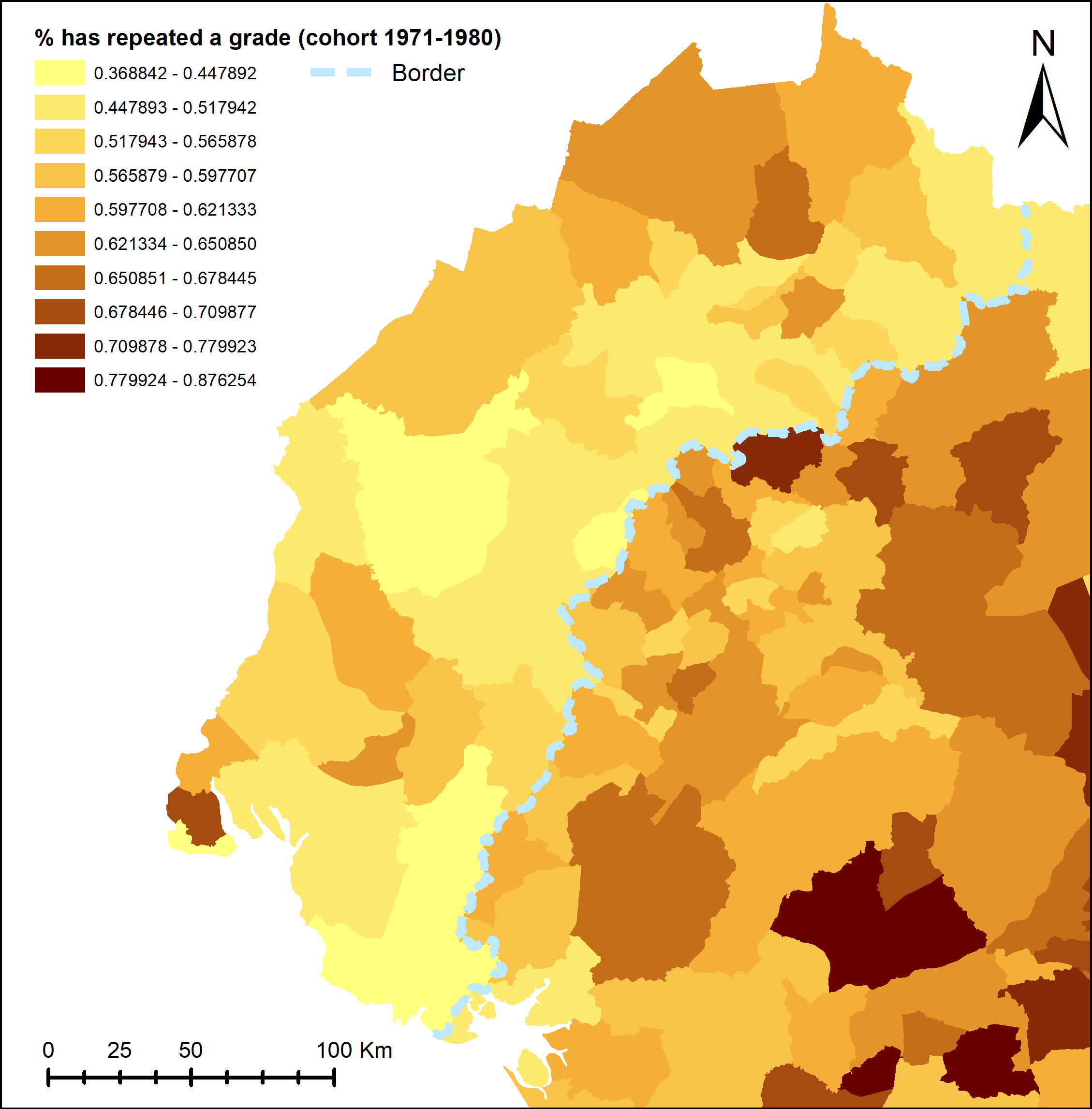
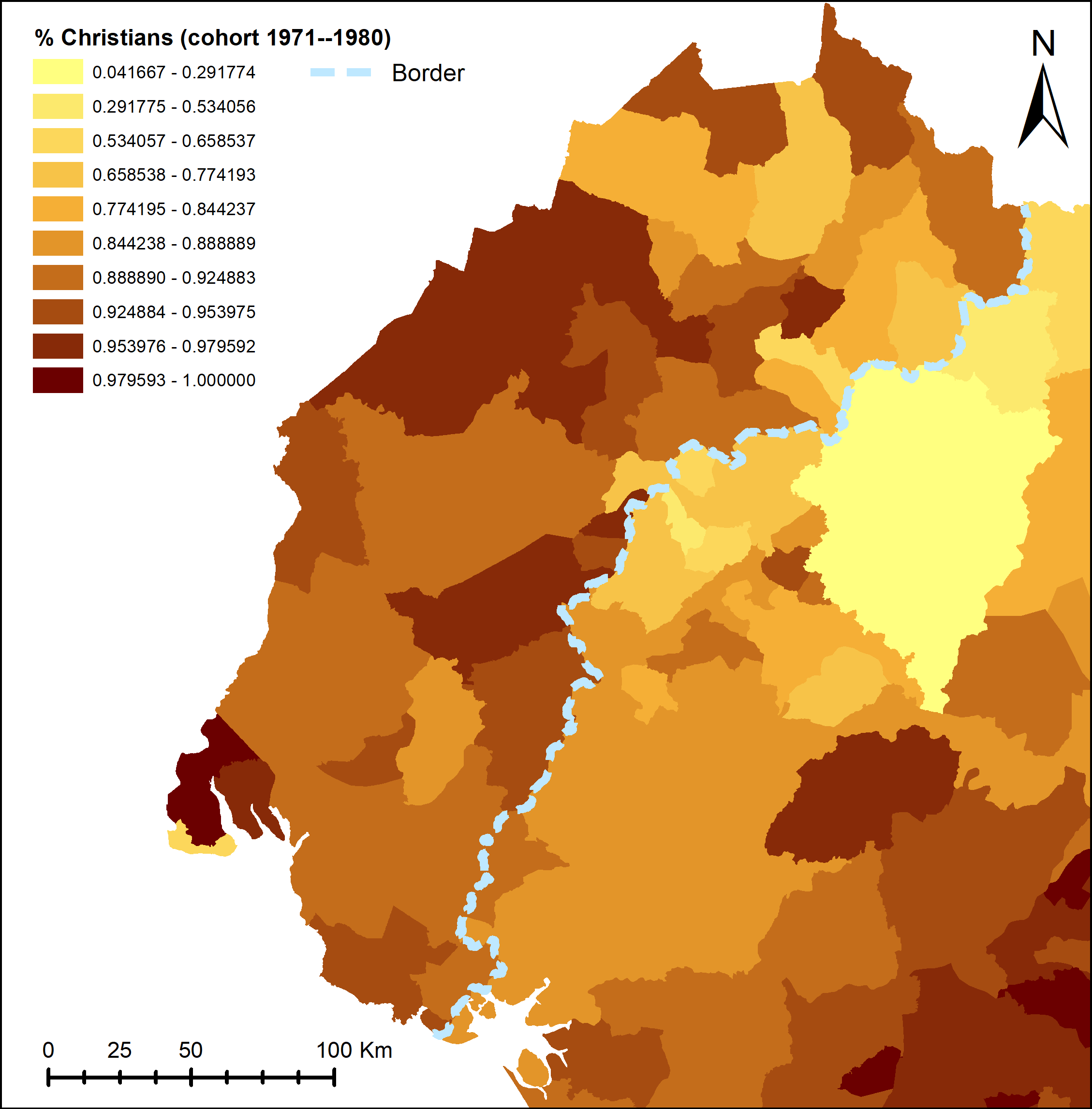
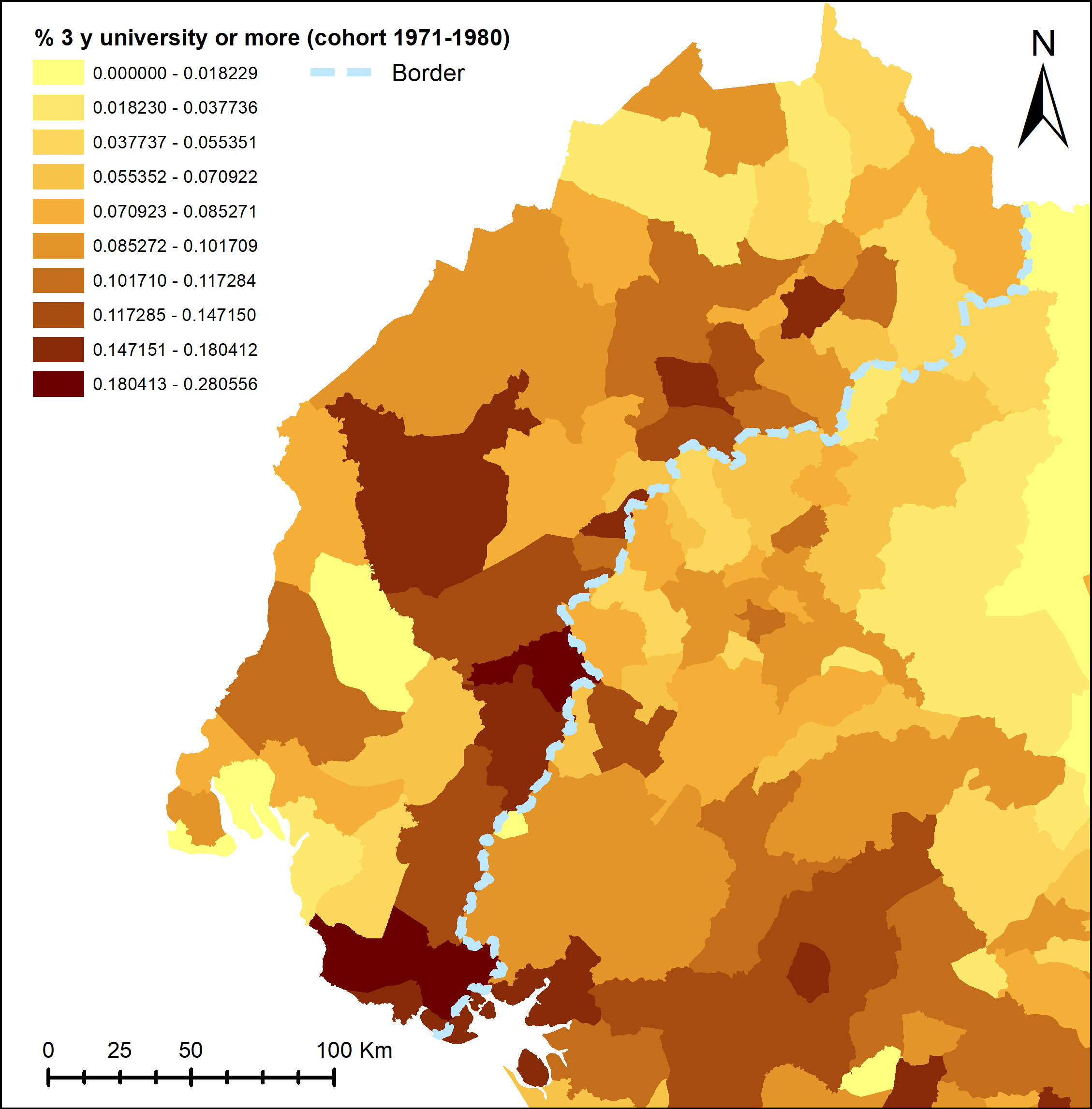
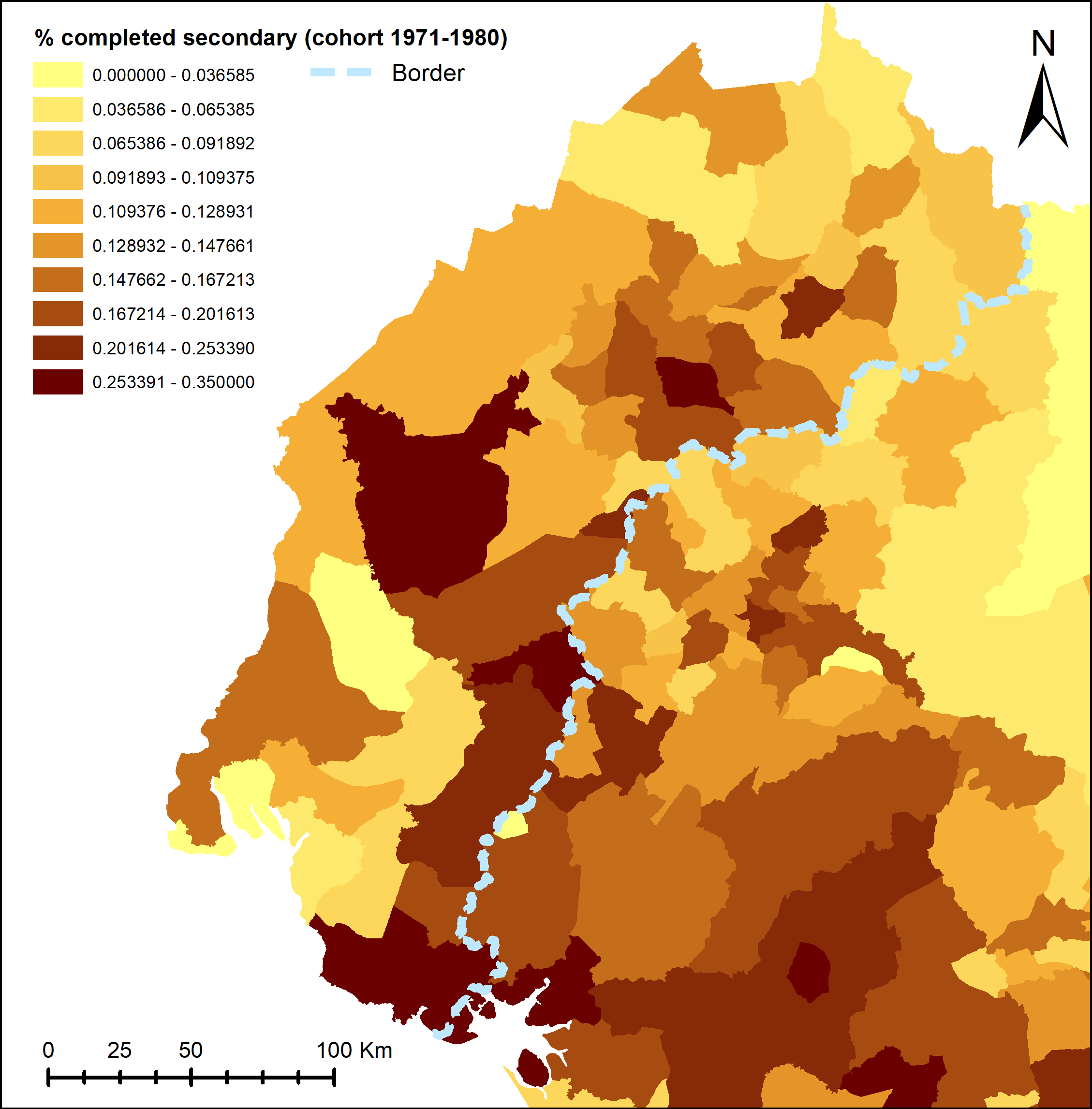


Figure E.2 — Maps of variables in the 2005 census

Notes: the first three maps display the average of the variable of interest among individuals born in the district between 1971 and 1980. The last map displays the percentage of students who have repeated at least a grade among students in primary and secondary school in the district. A student is deemed to have repeated a grade if they are older than what their normal progression in the system would predict, taking into account across-district differences in school entry age.

## E.2 2005 results: robustness to other specifications and samples

Table E.1 — Discontinuities in education in 2005: robustness to latitude-longitude specifications

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|  | means on sample | | estimated discontinuities | | | | | |
|  | within <25 km of border | | <25 km of border | | <50 km of border | | <100 km of border | |
|  | English-sp. side | French-sp. side | (x,y) poly of degree 1 | (x,y) poly of degree 2 | (x,y) poly of degree 1 | (x,y) poly of degree 2 | (x,y) poly of degree 1 | (x,y) poly of degree 2 |
|  | **Born 1971–1980** | | | | | | | |
| Ever been to school | 0.88 | 0.93 | -0.07\*\*\* | -0.13\*\*\* | -0.06\*\* | -0.08\*\*\* | -0.04 | -0.04 |
|  |  |  | (0.02) | (0.03) | (0.02) | (0.03) | (0.03) | (0.03) |
|  |  |  | [0.02] | [0.03] | [0.03] | [0.03] | [0.03] | [0.03] |
| Completed primary | 0.82 | 0.85 | -0.08\*\*\* | -0.12\*\*\* | -0.05 | -0.06\* | -0.03 | -0.03 |
|  |  |  | (0.03) | (0.04) | (0.03) | (0.03) | (0.04) | (0.04) |
|  |  |  | [0.03] | [0.04] | [0.04] | [0.04] | [0.04] | [0.04] |
| Completed secondary | 0.24 | 0.17 | 0.07\*\*\* | 0.08\*\* | 0.09\*\*\* | 0.09\*\*\* | 0.09\*\*\* | 0.09\*\*\* |
|  |  |  | (0.02) | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) |
|  |  |  | [0.03] | [0.03] | [0.03] | [0.03] | [0.03] | [0.03] |
| ≥3 years of higher ed. | 0.16 | 0.09 | 0.06\*\*\* | 0.08\*\*\* | 0.08\*\*\* | 0.08\*\*\* | 0.07\*\*\* | 0.07\*\*\* |
|  |  |  | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
|  |  |  | [0.02] | [0.02] | [0.02] | [0.02] | [0.02] | [0.02] |
| ≥15 years spent in school | 0.25 | 0.17 | 0.08\*\*\* | 0.10\*\*\* | 0.10\*\*\* | 0.11\*\*\* | 0.09\*\*\* | 0.09\*\*\* |
|  |  |  | (0.02) | (0.03) | (0.03) | (0.03) | (0.02) | (0.02) |
|  |  |  | [0.02] | [0.03] | [0.03] | [0.03] | [0.02] | [0.02] |
| High-skilled | 0.10 | 0.06 | 0.03\*\* | 0.04\*\* | 0.04\*\*\* | 0.06\*\*\* | 0.04\*\*\* | 0.04\*\*\* |
|  |  |  | (0.01) | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) |
|  |  |  | [0.01] | [0.02] | [0.01] | [0.01] | [0.01] | [0.01] |
|  |  |  |  |  |  |  |  |  |
|  | **Born 1981–1990** | | | | | | | |
| Ever been to school | 0.92 | 0.94 | -0.04\*\*\* | -0.06\*\*\* | -0.03\* | -0.04\*\*\* | -0.01 | -0.01 |
|  |  |  | (0.01) | (0.02) | (0.01) | (0.01) | (0.02) | (0.02) |
|  |  |  | [0.01] | [0.02] | [0.02] | [0.01] | [0.02] | [0.02] |
| Completed primary | 0.86 | 0.87 | -0.04\*\* | -0.04 | -0.02 | -0.02 | -0.01 | -0.01 |
|  |  |  | (0.02) | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) |
|  |  |  | [0.02] | [0.02] | [0.02] | [0.02] | [0.03] | [0.03] |
| Completed secondary | 0.21 | 0.15 | 0.05\*\*\* | 0.06\*\*\* | 0.07\*\*\* | 0.06\*\*\* | 0.06\*\*\* | 0.06\*\*\* |
|  |  |  | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
|  |  |  | [0.02] | [0.02] | [0.02] | [0.02] | [0.02] | [0.02] |
| ≥3 years of higher ed. | 0.04 | 0.03 | 0.01\*\*\* | 0.01\*\*\* | 0.02\*\*\* | 0.01\*\* | 0.01\*\*\* | 0.01\*\*\* |
|  |  |  | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
|  |  |  | [0.00] | [0.00] | [0.00] | [0.00] | [0.00] | [0.00] |
| ≥15 years spent in school | 0.24 | 0.15 | 0.07\*\*\* | 0.09\*\*\* | 0.09\*\*\* | 0.10\*\*\* | 0.08\*\*\* | 0.08\*\*\* |
|  |  |  | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
|  |  |  | [0.02] | [0.02] | [0.02] | [0.02] | [0.02] | [0.02] |
| High-skilled | 0.02 | 0.01 | 0.01\*\*\* | 0.01 | 0.01\*\*\* | 0.01\*\* | 0.01\*\*\* | 0.01\*\*\* |
|  |  |  | (0.00) | (0.01) | (0.00) | (0.00) | (0.00) | (0.00) |
|  |  |  | [0.00] | [0.01] | [0.00] | [0.00] | [0.00] | [0.00] |
|  |  |  |  |  |  |  |  |  |
| # of districts | 16 | 20 | 36 | 36 | 60 | 60 | 86 | 86 |

Notes: the sample excludes individuals born in Douala and born in districts closest to the northernmost section of the border. Added covariates: border section dummies. Standard errors clustered at the district level in parentheses. \*, \*\*, \*\*\*.01. Conley standard errors in [ ] (cut-off window of 20 kms).

Table E.2 — Discontinuities in education in the 2005 census estimated on the whole border

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  | Ever been to school | Completed primary | Completed secondary | ≥3 years of higher ed. | ≥15 years in school | High-skilled occupation |
|  | **Born 1971–1980** | | | | | |
| Discontinuity | -0.006 | -0.018 | 0.100\*\* | 0.087\*\*\* | 0.116\*\*\* | 0.026 |
|  | (0.068) | (0.079) | (0.040) | (0.031) | (0.042) | (0.020) |
|  | [-0.18, 0.19] | [-0.24, 0.18] | [0.04, 0.21] | [0.01, 0.17] | [0.06, 0.23] | [-0.03, 0.08] |
| Mean dep. var. | 0.850 | 0.758 | 0.182 | 0.111 | 0.186 | 0.074 |
| Observations (effective) | 33,104 | 39,594 | 33,104 | 39,594 | 33,104 | 20,310 |
| Observations (total) | 161,323 | 161,323 | 161,323 | 161,323 | 161,323 | 83,047 |
|  |  |  |  |  |  |  |
|  | **Born 1981–1990** | | | | | |
| Discontinuity | 0.069\*\* | 0.098\*\* | 0.081\*\*\* | 0.015\*\* | 0.100\*\* | 0.017\*\* |
|  | (0.032) | (0.050) | (0.031) | (0.007) | (0.044) | (0.008) |
|  | [0.01, 0.17] | [0.02, 0.25] | [0.01, 0.17] | [-0.00, 0.03] | [-0.02, 0.21] | [0.00, 0.04] |
| Mean dep. var. | 0.886 | 0.797 | 0.165 | 0.029 | 0.172 | 0.013 |
| Observations (effective) | 32,044 | 32,044 | 41,542 | 53,696 | 41,542 | 14,880 |
| Observations (total) | 234,767 | 234,767 | 234,767 | 234,767 | 234,767 | 59,048 |

Notes: the sample excludes individuals born in Douala and born in districts closest to the northernmost section of the border. Discontinuities obtained employing a local linear non-parametric estimation with triangular kernel and optimal mean squared error bandwidth. Added covariates: border section dummies. Standard errors clustered at the district level in parentheses. \*, \*\*, \*\*\*.01. Calonico et al. (2014) confidence interval in [ ]. The effective number of observations is the number of observations within the optimal bandwidth.

## E.3 Placebo borders

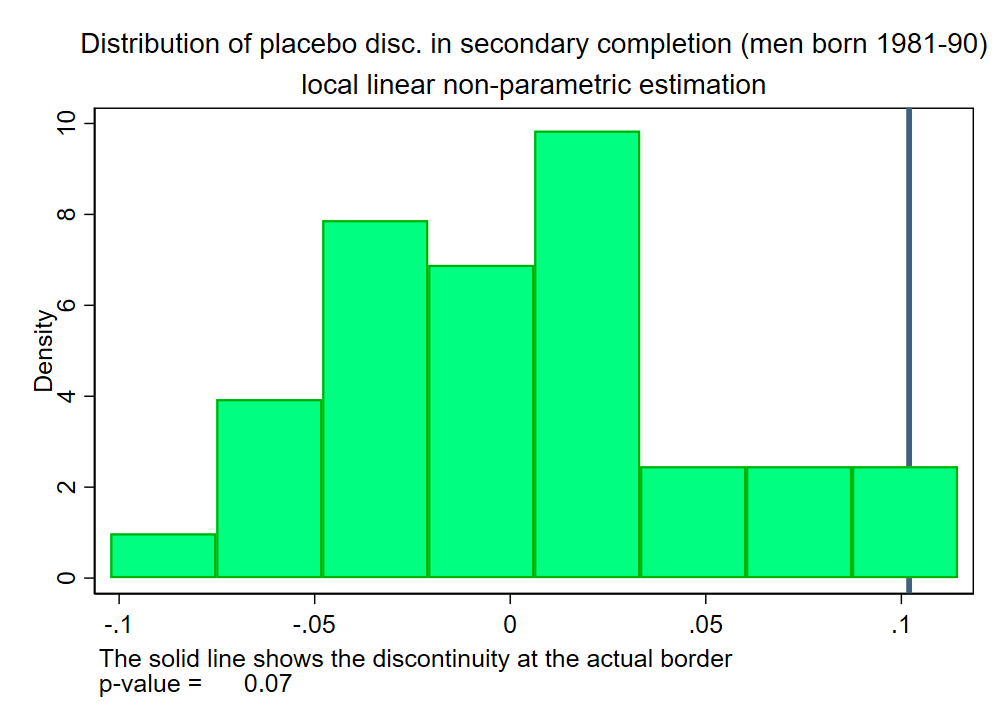


Figure E.3 — Placebo borders for 2005 results

Notes: This graph displays the distribution of discontinuities estimated on placebo borders running parallel to the actual border, but shifted East by up to 50 km and West by up to 25 km by steps of 1 km (the number of clusters quickly becomes small when the border is shifted West towards Anglophone Cameroon, which is why I limit the shift to 25 km). The solid line shows the discontinuity estimated at the actual border. The p-value for a two-sided test is simply the percentage of placebo discontinuities whose absolute value is larger than the absolute value of the actual discontinuity.

# F Additional results on mechanisms

Table F.1 — Discontinuities in male years of schooling with controls, cohort born 1922–1931

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Years of schooling | Years of schooling | Years of schooling | Years of schooling |
| Discontinuity | 1.18\*\*\* | 1.17\*\*\* | 1.15\*\*\* | 1.14\*\*\* |
|  | (0.38) | (0.37) | (0.35) | (0.35) |
|  | [0.48, 2.20] | [0.51, 2.19] | [0.55, 2.15] | [0.55,2.14] |
| Within 5 km of a plantation |  | 1.52\*\*\* | -0.15 | -0.16 |
|  |  | (0.57) | (0.67) | (0.65) |
| Within 10 km of a plantation |  |  | 1.92\*\*\* | 0.84 |
|  |  |  | (0.53) | (0.53) |
| Within 15 km of a plantation |  |  |  | 1.17\*\*\* |
|  |  |  |  | (0.33) |
| Observations (effective) | 4,981 | 4,981 | 4,981 | 4,981 |
| Observations (total) | 15,368 | 15,368 | 15,368 | 15,368 |

Notes: discontinuities obtained employing a local linear non-parametric estimation with triangular kernel and optimal mean squared error bandwidth. Added covariates: ethnic homeland fixed effects. Standard errors clustered at the village level in parentheses. \*, \*\*, \*\*\*.01. Calonico et al. (2014) confidence interval in [ ]. The effective number of observations is the number of observations within the optimal band- width. The controls are dummies equal to one if an individual was born in a village within x km of a German-owned plantation on the British side (the dummy is equal to zero on the French side). The list of German-owned plantations comes from Le Vine (1964).

Table F.2 — Discontinuities in religion in the 2005 census

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Christian | Protestant(a) | Catholic | Muslim | Animist | No religion |
| Discontinuity | 0.184\*\*\* | 0.281 | -0.115 | -0.044\* | -0.044 | -0.146\*\*\* |
|  | (0.052) | (0.200) | (0.193) | (0.026) | (0.051) | (0.055) |
|  | [-0.01, 0.43] | [-0.13, 0.89] | [-0.65, 0.29] | [-0.12, 0.01] | [-0.14, 0.08] | [-0.28, -0.03] |
| Mean dep. var. | 0.893 | 0.316 | 0.515 | 0.027 | 0.022 | 0.047 |
| Observations (effective) | 142,700 | 193,922 | 225,072 | 170,888 | 250,302 | 289,980 |
| Observations (total) | 713,814 | 713,814 | 713,814 | 713,814 | 713,814 | 713,814 |

Notes: the sample excludes individuals born in Douala and born in districts closest to the northernmost section of the border. Discontinuities obtained employing a local linear non-parametric estimation with triangular kernel and optimal mean squared error bandwidth. Added covariates: border section dummies. Standard errors clustered at the district level in parentheses. \*, \*\*, \*\*\*. Calonico et al. (2014) confidence interval in [ ]. (a) Protestants comprise other Christians, who are neither Protestant, nor Catholics, nor Orthodox.

1. Migrations, Marché du Travail et Dynamiques Démographiques en Afrique Subsaharienne, financed by the Hewlett Foundation, the AIRD (Agence Inter-établissement de la Recherche pour le Développement) and the AFD (Agence Française pour le Développement) [↑](#footnote-ref-1)
2. Integrated Public Use Microdata Series: https://international.ipums.org/ international/ [↑](#footnote-ref-2)
3. Codes given in the raw census data and in the village file do not match (census village codes very likely correspond to the chiefdom rather than the village). [↑](#footnote-ref-3)
4. The census asks whether the last grade attended was in the Francophone or Anglophone system, and I assume that individuals did their entire education within the same system. For instance, an individual who stopped school in the second grade of secondary school will be assigned 8 (6+2) years of schooling in the Francophone system and 9 (7+2) years of schooling in the Anglophone system. [↑](#footnote-ref-4)
5. The census gives a choice between only 6 grades in the Francophone system and 7 in the Anglophone one, however, in the Anglophone system, the primary cycle was 9 years long up to 1931, 8 years long from 1931 to 1967 and 7 years long from 1967 onwards. In the Francophone system, the primary cycle has always been 6 years long, but a lot of mission schools used to lengthen the cycle by offering beginners’ classes. [↑](#footnote-ref-5)
6. School participation includes Koranic schools, but omitting them does not affect results, especially because they were well developed only across the northernmost part of the border. [↑](#footnote-ref-6)
7. [Nzima Nzima](#_bookmark104) (2014) writes that infant mortality rates were higher in French Cameroon than in British Cameroon in the 1920s. [↑](#footnote-ref-7)
8. I could not geolocate villages there, see figure 2. [↑](#footnote-ref-8)
9. The census does not give an individual’s full migration history: a migrant is someone who lives in 1976 in a different district than the one they were born in. [↑](#footnote-ref-9)
10. In the sample of men living in French Cameroon within 25 km of the border and born before 1912, regressing years of schooling on district of residence fixed effects and a dummy equal to one if an individual was born away from the border region yields a coefficient of 1.75 (standard error clustered at the district level of 0.80). [↑](#footnote-ref-10)
11. Nigerian “tribes” comprise both people born in Nigeria and their children. They are 84,200 for a total population of 752,700 in the 1953 census of the Southern British Cameroon (the part of British Cameroon that was then reunited with Francophone Cameroon). [↑](#footnote-ref-11)
12. In the sample of men living in British Cameroon within 25 km of the border and born before 1912, regressing years of education on district of residence fixed effects and a dummy equal to one if an individual was born in Nigeria yields a coefficient of (standard error clustered at the district level of). [↑](#footnote-ref-12)
13. In the 1976 census, the percentage born in Nigeria in the generation of the parents (born before 1912) is about 2%. Assuming birth rates are roughly homogenous, the percentage of their Cameroon- born children in total population should not be too different. 5% seems like a reasonable upper-bound. [↑](#footnote-ref-13)
14. The number of districts is still smaller though because in the Ouest region I cannot geolocate villages beyond the 1976 district level (see figure 2). [↑](#footnote-ref-14)