Religious Institutions, Minority Incorporation, and Human Capital

Evidence from the Ottoman-Habsburg Frontier

Bogdan G. Popescu [*]

Word count: 8,833 (excluding abstract, references, and appendices)

Abstract

States often expand schooling and administration yet fail to erase the inequalities they inherit. I argue that state capacity converts into convergence only where the state incorporates a group into its national project; where a group is excluded, households withhold the compliance public institutions require, and inherited gaps persist. I test this argument along the 1739 Ottoman–Habsburg frontier, which Romania and Yugoslavia inherited jointly: Romania closed the Ottoman literacy disadvantage by 1991 while Yugoslavia did not. The central evidence is internal to Yugoslavia, where, under the same federal socialist state and increasingly equal school supply, Orthodox-majority Serbian districts converged at Romanian rates while Muslim-majority districts, whose population had lost its imperial-era standing, did not, with the residual concentrated on women's literacy. A within-Muslim discontinuity at the 1929 sharia-court closure is consistent with the mechanism: three decades later, women's illiteracy was higher where districts had lost their courts; men's was not.

[*]Assistant Professor of Political Science, Tec de Monterrey.

Email: bgpopescu@tec.mx

1 Introduction

A large body of research demonstrates that imperial and colonial institutions cast long shadows over contemporary development, generating inequalities that persist well beyond the regimes that created them (Acemoglu *et al.*, 2001; Becker *et al.*, 2016; Dell, 2010; Michalopoulos & Papaioannou, 2013; Nunn & Wantchekon, 2011; Valencia Caicedo, 2019). These legacies operate through inherited state capacity (Vogler, 2019), human capital (Becker & Woessmann, 2009), and political attitudes (Grosfeld & Zhuravskaya, 2015; Pop-Eleches & Tucker, 2017). What varies across them is not only how they transmit but how vulnerable they are to reversal, and by whom.

A parallel literature shows that centralized states can overcome inherited gaps through mass schooling (Darden & Grzymała-Busse, 2006; Paglayan, 2022), but not when such campaigns succeed and when they fail. Becker *et al.* (2016) show that Habsburg administrative inheritance outlasted the empire by a century; when inherited disadvantage erodes remains unspecified. I identify the conditioning variable on the Ottoman–Habsburg frontier in the Balkans, the same setting as their canonical persistence finding, where two successor states inherited the discontinuity but produced divergent literacy trajectories.

State capacity enables convergence but does not guarantee it. Incorporation, I argue, conditions whether convergence occurs. I define incorporation as the degree to which a successor state formally admits a given identity group into its national project through: (a) representation in central state institutions (e.g. parliament, civil service), (b) recognized institutional autonomy for communal life (e.g. religious-legal jurisdiction, confessional schooling), and (c) absence of targeted dispossession and dissolution of communal institutions. Not every institutional dismantling is hostile; the diagnostic is whether treatment is group-differential. A state that dismantles every community's institutions alike excludes no one; a state that strips one group's institutions while leaving others' standing, or pairs the stripping with exclusion from office and land, excludes that group. Interwar Yugoslavia was the second kind, withdrawing Muslim institutions while keeping Muslims out of office and dispossessing them (Section 3). Exclusion typically follows a status

reversal that turned an imperial-era ruling group into a post-imperial minority associated with a rival national project, and it erodes the household trust on which school uptake depends.

The evidence proceeds in three tiers. First, the cross-country contrast supplies the descriptive macro-puzzle: the imperial literacy gap of 14.3 percentage points in Romania and 43.5 in Yugoslavia (Figure 1) closed substantially in Romania by 1991 (83%) but far less in Yugoslavia (27%). The within-Yugoslavia comparison provides the central contrast: with the central socialist state and Ottoman inheritance held constant, Orthodox-majority Serbian districts converged at Romanian rates while Muslim-majority districts did not; the causal identification comes from the within-Muslim discontinuity of the third tier.

Second, the data reveal a distinctive empirical pattern of exclusion: the residual gap in Muslim-majority Yugoslav districts falls disproportionately on women. Because boys and girls within the same district face identical school infrastructure, the within-district gender asymmetry points to a household-compliance mechanism at the excluded margin rather than a district-level capacity gap.

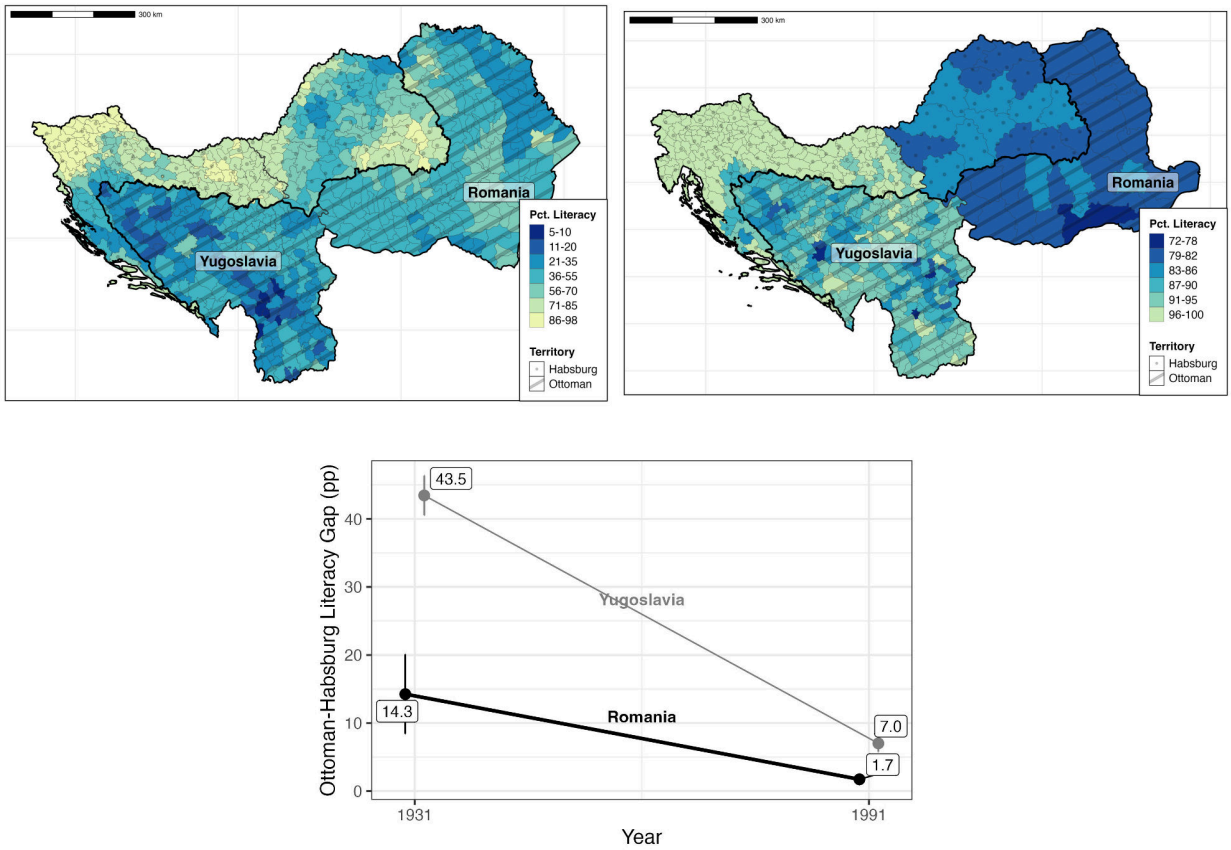
Third, the within-Muslim sharia-court RD probes the dissolution of communal institutions through one example: the 1929 closure stripped districts with fewer than 5,000 Muslims of local religious-legal jurisdiction. Where Muslim communities were present, the closure is associated with higher women's illiteracy, concentrated among the cohorts schooled after 1929, while men's illiteracy is barely affected; where Muslims were absent, it raised none, consistent with a channel operating only through Muslim communities.

2 Theory

The Persistence Consensus and Its Limits

Studies of institutional persistence trace imperial legacies through channels that fall into three broad categories: inherited state capacity (Becker *et al.*, 2016; Vogler, 2019), durable human capital differences (Lankina & Getachew, 2013; Valencia Caicedo, 2019), and persistent attitudes, norms,

Figure 1. *The Ottoman Literacy Gap in Romania and Yugoslavia: Levels and Convergence*



Notes: Panels (a) and (b): district-level literacy in 1931 and at the end of communism (1991/1992), with the Ottoman–Habsburg frontier; darker is higher. Panel (c): the estimated Ottoman–Habsburg gap by census year, OLS coefficients on Ottoman-side location (positive = Ottoman disadvantage) with 95% confidence intervals, declining from 43.5 to 7.0 pp in Yugoslavia and 14.3 to 1.7 in Romania.

and political behavior (Grosfeld & Zhuravskaya, 2015; Nunn & Wantchekon, 2011; Pop-Eleches & Tucker, 2017; Wittenberg, 2006). These channels share a common claim: imperial institutions generate inequalities that outlast the regimes that created them. What they do not specify is when state-led erasure of those inequalities succeeds and when it fails.

Recent work has begun to introduce conditionality into this framework. Simpser *et al.* (2018) distinguish institutional, attitudinal, and behavioral legacies according to their mechanisms of intergenerational transmission and argue that durability varies systematically with whether successor regimes reinforce or erode those mechanisms. Abad & Maurer (2021) push from the other side, arguing that many persistence claims rest on weak identification: observed correlations

may reflect selection or correlated unobservables rather than causal transmission across regimes. Both moves leave the conditionality question unresolved. The erasure literature shows that modern states can reshape educational outcomes through mass schooling (Darden & Grzymała-Busse, 2006; Paglayan, 2022, 2024), but does not specify why similar campaigns produce convergence in some settings and persistent gaps in others. The literature lacks a theory of the conditions that determine which outcome obtains.

A Theory of Reversal: Incorporation as the Binding Constraint

This article advances incorporation as the binding constraint on whether historical gaps close: state capacity supplies schooling, but incorporation determines uptake.

State capacity is the bureaucratic infrastructure and fiscal resources that regimes inherit or build. It determines what states can supply: schools, teachers, standardized curricula, and attendance enforcement. But supply is not uptake, and the persistence literature has little to say about what determines the gap between them.

Recent scholarship supplies the missing half by reconceiving mass schooling as a value-inculcation institution: states design it to mold identities and secure loyalty to a national project, not only to build human capital (Paglayan, 2024, 2026). Once schooling is understood as the transmission of a national project, uptake becomes conditional rather than automatic. A group incorporated into that project takes state schooling up as its own; a group the project excludes reads the same schooling as an instrument of assimilation and resists it at the margins households still control (Fouka, 2020; Bazzi *et al.*, 2025). Which response obtains depends on the state–minority relationship this article calls incorporation.

I argue that incorporation conditions whether centralized investments translate into schooling uptake. *Incorporation* is the degree to which the identities of governed populations are incorporated into the successor state's national project. Institutional access can be expanded within years through policy choices, but the mistrust that prior exclusionary policies generate persists across generations (Acharya *et al.*, 2018; Nunn & Wantchekon, 2011; Lupu & Peisakhin, 2017). The same

schooling investment therefore produces different outcomes depending on whether populations perceive the state as theirs. The argument links five steps: status reversal invites exclusionary policy; exclusion weakens incorporation; weak incorporation breeds distrust of state institutions; distrust lowers household compliance with state schooling; and withheld compliance falls hardest on daughters' education, the discretionary margin.

The argument predicts a heterogeneous pattern: within Yugoslavia, Orthodox-majority districts converge like Romania, while Muslim-majority districts lag, with the residual concentrated among women. Two channels operate together. First, traditional patriarchal norms (documented most extensively for Muslim societies (Fish, 2002; Norris & Inglehart, 2003; Spierings *et al.*, 2009) but also present in Orthodox Christian and other religiously traditional contexts) produce a baseline gender gap in literacy that household norms reproduce unless state-mediated schooling intervenes. Second, institutional exclusion determines whether that baseline closes or persists. Where incorporation holds, centralized schooling overcomes the baseline within two generations (Muslim-majority Albania, Orthodox Yugoslavia). Where it does not, the baseline persists. The interwar Yugoslav state's pattern of administrative exclusion, agrarian dispossession, and sharia-court contraction actively converted Muslim minority status into political exclusion. The nationalizing-states literature documents how successor states adopt such exclusionary projects (Brubaker, 1996; Wimmer, 2002, 2013) but does not trace their long-run consequences.

These interwar policies operated on inherited imperial legacies. Two are central to this account. First, the Habsburg and Ottoman empires left different *administrative infrastructures*: Habsburg rule generated professionalized bureaucracies with standardized schooling (Becker *et al.*, 2016; Ingrao, 2000; Judson, 2016; Vogler, 2019), whereas Ottoman governance relied on patrimonial arrangements and local intermediaries (Barkey, 1994; İnalcık & Quataert, 1994; İnalcık, 1994), leaving a thinner administrative inheritance in its Balkan provinces (Section 5.1 measures this difference directly). Second, Ottoman confessional governance through the millet system[1] created

[1]The millet system was the Ottoman Empire's confessional governance regime, under which legally recognized religious communities (e.g. Muslims, Orthodox Christians, Armenians, and Jews) exercised autonomy over matters such as family law, education, and communal taxation. It linked political status and administrative authority to

a specific *social structure*: religiously diverse populations in which Muslims were the ruling group whose identity was incorporated into the imperial state, the incorporation that the post-imperial Yugoslav state would later invert.

Under communism, institutional standardization narrowed but did not close educational gaps in Muslim-majority districts. The mistrust generated by interwar exclusion likely shaped household compliance with state schooling (Sorabji, 1989; Karčić, 2019; Fouka, 2020; Bazzi *et al.*, 2025): where state institutions were identified with a dominant ethnoreligious group, minorities often viewed schools as instruments of cultural domination (Alesina & La Ferrara, 2005; Darden & Grzymała-Busse, 2006; Ramet, 1992), and ethnographic work records older Bosnian Muslims perceiving state schools as culturally distant and morally threatening (Bringa, 1995, p. 204); (Hadžijahić *et al.*, 1977, pp. 122, 127). In Romania no comparable status reversal occurred: the dominant Orthodox population controlled both imperial-era communal institutions and post-independence state structures, and incorporation remained high.

Table 1 summarizes the article's main empirical claims and where each is tested.

Exclusion is different from coordination costs produced by ethnoreligious heterogeneity (Alesina & La Ferrara, 2005). Coordination costs take the form of disputes over language, script, and curricula that raise the cost of implementing uniform schooling. These frictions decline once a standard is selected and enforced (Charnysh, 2019). Exclusion is a relational condition: the state denies a group membership in its national project, and the denied group responds with durable mistrust of state institutions.[2]

Exclusion binds most tightly in domains that touch cultural reproduction, above all education. Within schooling, girls' education is the margin most elastic to household preferences and trust: in settings with conservative gender norms, parents treat sons' education as a default investment

religious identity.

[2]Educational standardization was thorough: by 1950 the last mektebs were closed, teaching in mosques was criminalized, and five of six historical medreses were shut (Mekić, 2017, p. 34); (Alibašić & Zubčević, 2009, pp. 50, 54); 756 mosques were destroyed or damaged postwar (Mekić, 2017, p. 196), exclusion rather than coordination friction. Yet literacy gaps persisted well beyond this standardization.

Table 1. *From Theoretical Claim to Empirical Test*

Claim	Observable Implication	Evidence	Result	Section
Ottoman inheritance produced an initial disadvantage	Lower literacy and administrative capacity, 1931	Frontier RD: Romania and Yugoslavia	Ottoman side lags in literacy and bureaucracy	5.1
Capacity alone does not explain convergence	Convergence varies within the same state	Within-Yugoslavia comparison	Orthodox districts converge like Romania; Muslim districts lag	5.2
Exclusion operates via household uptake	Residual gap concentrates on daughters' education	Gender asymmetry	Male penalty nearly closes; female penalty persists	5.3
Patriarchal baseline; exclusion blocks its erosion	Supply equalizes; cross-national Muslim gaps track incorporation, not religion	Supply test; cross-national comparison	Inputs equalize; gap closure tracks incorporation, not religion	5.3–5.4
Dissolving communal institutions reduces uptake	Gendered effect at the 1929 closure's 5,000-Muslim threshold	Within-Muslim RD at the 5,000-Muslim cutoff	Women's literacy falls where Muslims were present	6

in labor-market returns and daughters' as discretionary (Becker, 1981; Jayachandran, 2015), so distrustful families withdraw daughters first.

Hypotheses

H1 (Persistence as Default): *In the absence of countervailing state action, imperial legacies in human capital and administrative capacity persist. Localities historically governed by the Ottoman Empire will exhibit lower literacy rates and lower bureaucratic density than comparable Habsburg-administered localities.*

Section 5.1 tests this prediction, establishing the institutional substrate that Sections 5.2–5.4 trace forward.

H2 (Incorporation as the Binding Constraint): *Legacies fade where successor states govern populations incorporated into the national project; they endure where incorporation is absent, even when state capacity is comparable.*

Yugoslavia provides a within-country test. Orthodox-majority Serbian districts and Muslim-majority districts inherited the same Ottoman institutional bundle and were governed by the same central socialist regime; they differed in incorporation into the national project. If incorporation is the binding constraint, Serbian districts should converge like Romania while Muslim-majority districts lag. Incorporation is coded from institutional facts fixed before the outcome window (Section 3), not inferred from the convergence it is meant to explain, so two patterns would count against it: uniformly slow convergence across all Yugoslav republics, or Muslim-majority districts converging at the Orthodox rate. Two further observable implications follow. First, the two frictions should fade at different rates. Coordination costs decline once a standard is enforced, so the effect of religious fractionalization (the proxy for coordination costs) should fall toward zero; exclusion has no such endpoint, so the effect of Muslim population share (the contextual proxy for exclusion) should persist into late socialism. Second, the 1929 sharia-court closure (Section 6) separates conservative religious norms from institutional exclusion: districts below a 5,000-Muslim threshold lost their court, comparable districts above retained theirs. Because norms vary continuously with Muslim concentration and no cultural boundary coincides with an administrative count of 5,000, the institutional account predicts a gendered gap opening at the cutoff, while a norms-only account predicts none; the 1961 census measures it three decades later.

3 Historical Context

The Ottoman and Habsburg empires, which between them governed the territories that became Romania and Yugoslavia, followed different logics of rule and approaches to education. The Ottoman administrative order organized political authority by religion. Subjects were classified into recognized religious communities (millets), with communal authorities managing family law and fiscal obligations; political privileges (military service, senior administration) structurally favored Muslims (Barkey, 2008; Braude & Lewis, 1982; İnalcık, 1973). Education was correspondingly decentralized and confessional: Muslim instruction took place in Qur'anic schools (*mektebs*) attached to mosques or local endowments (*vakıf*), while non-Muslim communities governed their

own schools with minimal central oversight (Alibašić & Zubčević, 2009; Barkey, 2008; Fortna, 2011).

Habsburg administration moved in the opposite direction. Beginning with Maria Theresa's 1774 *Allgemeine Schulordnung*, the Habsburgs pursued centralized compulsory primary schooling through a professionalized civil service (Becker *et al.*, 2016; Ingrao, 2000; Judson, 2016; Vogler, 2019). This system transmitted two features to successor states: administrative infrastructure that could be repurposed for mass education, and populations accustomed to treating schooling as a state entitlement.

By 1931, decades after Ottoman rule had ended, the religious composition of the two countries' former Ottoman territories still differed sharply. The Ottoman side of Romania was predominantly Orthodox and homogeneous (mean fractionalization = 0.11), whereas the Ottoman side of Yugoslavia included Orthodox, Catholic, and Muslim communities (mean = 0.25; Appendix Figure A.14). Muslims were concentrated in Bosnia (31%), Kosovo (73%), Macedonia (28%), and Montenegro (16%), and largely absent from Croatia (0.15%), Romania (1%), and Serbia (2.8%).

Successor States and the Politics of Incorporation

In Romania, the Ottoman side's predominantly Orthodox population was incorporated into the Romanian national project from independence onward. No status reversal occurred for the population bearing the Ottoman legacy: Romanians, the titular group of the successor state, formed the majority on both sides of the imperial divide.

In Yugoslavia, the situation was different. The 1918 creation of the Kingdom of Serbs, Croats, and Slovenes transformed Muslims from an imperial ruling group into post-imperial minorities governed by a state organized around Christian national identities. This status reversal did not follow automatically from the change of states; successor-state policy produced it. Successor-state policies systematically excluded Muslim communities from institutions, administration, and economic opportunity.

Three patterns documented in contemporaneous sources establish the scale of this exclusion;

Appendix Table A.2 reports the full figures and sources. First, Muslim presence in state, economic, and educational institutions fell to roughly between 1 and 6% against an approximate 11% population share (Bougarel, 2018, pp. 43–44); (Kemura, 1986; Purivatra, 1974). Second, autonomous Muslim institutions were dismantled: mektebs and medreses contracted to a fraction of their 1918 numbers, sharia-court judgeships were cut under the 1929 Law on the Organization of Sharia Courts (*Zakon o uređenju šerijatskih sudova*, ZSS) (Karčić, 2005, p. 74), and the 1919 agrarian reform expropriated a landlord class that was roughly 91% Muslim (Dyker, 1972, pp. 246–247). Third, religious bloc voting was near-unanimous: the Yugoslav Muslim Organization captured 98–100% of the Muslim vote in every interwar election (Purivatra, 1974).

These patterns map onto the definition in the introduction: Muslim underrepresentation in state institutions is the failure of (a) representation; the dismantling of mektebs, medreses, and sharia courts is the withdrawal of (b) institutional autonomy; and the agrarian expropriation is (c) targeted dispossession. Bloc voting is not part of this coding; it is the group's revealed response to that treatment. Yugoslavia therefore excluded Muslims on all three dimensions, coded from facts that predate the outcomes (Appendix Table A.8 traces the temporal sequence on both sides of the frontier). Section 5 asks whether this historically measured exclusion predicts literacy trajectories; it does not infer exclusion from them.

4 Research Design

The research design exploits a distinctive feature of the Ottoman–Habsburg frontier: the same historical boundary separates imperial legacies in two successor states. The empirical strategy proceeds in three nested comparisons, each holding more fixed than the last. First, a geographic regression discontinuity along the frontier, estimated with the two-dimensional (RD2D) design of Cattaneo *et al.* (2025), documents the Ottoman–Habsburg literacy gap and its evolution in each country. Second, a within-Yugoslavia comparison exploits district-level variation in Muslim share, holding the successor state and its schooling campaigns fixed (Section 5.2). Third, because the republics administered schooling (Appendix Section I), the sharpest test holds even the republic

fixed: the within-Muslim regression discontinuity at the sharia-court threshold, entirely inside Bosnia (Section 6).

For the frontier discontinuity, treatment is inherited imperial exposure: a district is coded as treated if its centroid lies on the Ottoman side of the 1739 Treaty of Belgrade frontier, the most institutionally salient of the successive Ottoman–Habsburg demarcations (Appendix Section C).[3] The analysis draws on a newly harmonized panel of district-level literacy (by gender), religious composition, and demographics from interwar through late twentieth-century censuses of Romania and Yugoslavia's component units, spanning six decades.[4]

Re-estimating the discontinuity census by census then tests whether the Ottoman effect attenuated differently depending on incorporation: faster where incorporation was high and slower where it was absent, typically following status reversal.

The RD2D estimator treats treatment assignment as a function of two-dimensional geographic location $\mathbf{X}_i = (\text{lon}_i, \text{lat}_i)$ relative to the frontier $\mathcal{B}$. The estimand is the local boundary treatment effect $\tau(\mathbf{b}) \equiv \mathbb{E}[Y_i(1) - Y_i(0) \mid \mathbf{X}_i = \mathbf{b}]$, where $\mathbf{b} \in \mathcal{B}$ denotes a point on the frontier. The estimand $\tau(\mathbf{b})$ is not one number but a profile of treatment effects that varies along the border; I estimate it at 164 evaluation points spaced 10 km apart. Under standard continuity assumptions,[5] this is identified as the difference in conditional expectations from the Ottoman ($\mathcal{A}_1$) and Habsburg ($\mathcal{A}_0$) sides. I report the full profile and distinguish historically low-stakes from high-stakes segments, as classified below.

[3] Bosnia-Herzegovina, under Habsburg administration from 1878 and annexed in 1908, is coded "Ottoman side": its institutional inheritance (millet, *vakıf*, and mekteb networks) formed over three centuries of Ottoman rule and was not undone in four decades.

[4] The 1931 Yugoslav census measured literacy as the ability to read and write in Serbo-Croatian or Slovene (Latin or Cyrillic script); the 1930 Romanian census used Romanian. Arabic-script (*arebica*) literacy acquired through mekteb instruction is therefore not counted, which mechanically inflates the 1931 Ottoman-side gap; later censuses apply the same state-language definition, so the trajectories compare like with like. The final census in the panel is 1991 in Yugoslavia and 1992 in Romania; for readability, prose comparisons dated "1991" use each country's final census, and Table 2 labels the Romanian series 1930–1992.

[5] The potential-outcome surfaces $\mathbb{E}[Y_i(d) \mid \mathbf{X}_i = \mathbf{b}]$ are continuous in $\mathbf{b}$ at the frontier $\mathcal{B}$; see Cattaneo *et al.* (2025) for the formal statement.

Assignment to Treatment

The 1739 Treaty of Belgrade frontier was drawn by joint Ottoman–Habsburg commissions of engineers and cartographers, anchored to defensible rivers, marshes, and ridgelines (Gümüşçü, 2010; Kurtaran, 2018; Stoye, 1994).

The strategic salience of the boundary varied along its length. Where fortresses, trade corridors, or major rivers were at stake, border placement reflected military and economic considerations that could confound the RD. I classify these as "high-stakes" segments. Elsewhere, particularly along portions of the Una and Sava rivers (Appendix Figure A.15 maps the segments), the boundary tracked natural features for geographic convenience rather than strategic calculation. As Stoye observes, these stretches had "no fortresses, but no need for them" (Stoye, 1994, p. 186). I classify these as "low-stakes" segments and pre-specify them as the primary quasi-experimental settings. This strategy privileges historical plausibility over blanket frontier coverage, reporting high-stakes segments for transparency. The full historiographical analysis of border formation, including negotiator information, incentives, and capacity to manipulate the line, is summarized in Appendix Section C and included in the replication archive. This analysis follows the framework of Kocher & Monteiro (2016) and Paine *et al.* (2025).

Covariate Balance at the Frontier

Identification requires continuity of potential outcomes and predetermined covariates at the frontier. It also requires that units do not sort precisely around the boundary. I assess balance using five predetermined covariates (elevation, slope, precipitation, temperature, wheat suitability). At each of the 164 evaluation points, I estimate local discontinuities using the same RD2D procedure applied to literacy outcomes.

Covariate balance tracks how each border segment was drawn (Appendix Section D, Figure A.1). Where the frontier follows the Una and Sava rivers, drawn for geographic convenience, the two sides look alike: elevation discontinuities lie close to zero, and slope, precipitation, and temperature show the same pattern. Where the border was drawn strategically, as in parts of the Carpathians

and coastal Croatia, the two sides differ sharply. Wheat suitability is noisier overall but shows no systematic discontinuity along the Una–Sava corridor.

In the results that follow, I therefore classify each evaluation point as design-credible or not, using four criteria: (1) case-by-case historical evidence that the segment was drawn for reasons unrelated to the populations on either side (Kocher & Monteiro, 2016; Paine *et al.*, 2025); (2) balance on all five predetermined geographic covariates; (3) no bunching of population at the border in density tests; and (4) enough observations on both sides of the frontier. Statistical significance is reported for all estimates but plays no role in this classification.

Threats to Identification and Design Diagnostics

In practice the points satisfying all four criteria cluster tightly: on the Yugoslav frontier they lie along the Sava, on the Romanian frontier along the Danube and the land border, so the credible comparison is these corridors rather than the full frontier. Within them, three threats remain. If populations had sorted across the line, density would jump at the border: localized McCrary tests show smooth density at most locations, and the points that fail are excluded as non-credible (Appendix Section E). If selective migration had recomposed border populations, the estimates should move when areas of population decline are dropped or compositional shifts are controlled: they do not (Appendix Section H). And cross-border spillovers would make the two sides more alike, attenuating rather than generating discontinuities, so the estimates are conservative. Traditional distance-based RD estimates serve as secondary checks (Appendix Sections F and G).

Estimation

At each evaluation point, I estimate local treatment effects using bias-corrected local polynomial estimators (Cattaneo *et al.*, 2025): local linear estimation ($p = 1$), local quadratic inference ($q = 2$), triangular kernel, and MSE-optimal bandwidth selection.[6] Reported uncertainty bands are uniform confidence bands accounting for simultaneous inference across all evaluation points. Variance

[6]Because estimation and inference use different polynomial orders, point estimates may occasionally fall outside the robust confidence bands at locations with high curvature (Cattaneo *et al.*, 2025).

estimation uses HC1 standard errors clustered at the county (*judeţ*) level in Romania and the district level in Yugoslavia.[7] Bandwidth diagnostics are reported in Appendix Figure A.3, robustness to alternative specifications in Appendix Sections F and G, and full implementation details in the replication archive.

Main Results

Figure 2 summarizes the design and main estimates, for the total population (left) and for women (right). In each panel, the map plots the 164 evaluation points along the frontier over shading for observed 1931 literacy, and the strip below indexes the same points linearly along the border. Point color gives the estimated local Ottoman effect (red negative, blue positive); transparency flags local covariate imbalance, while the full four-criteria credibility classification enters the results below. Balance and density tests for every evaluation point are reported in Appendix Sections D and E.

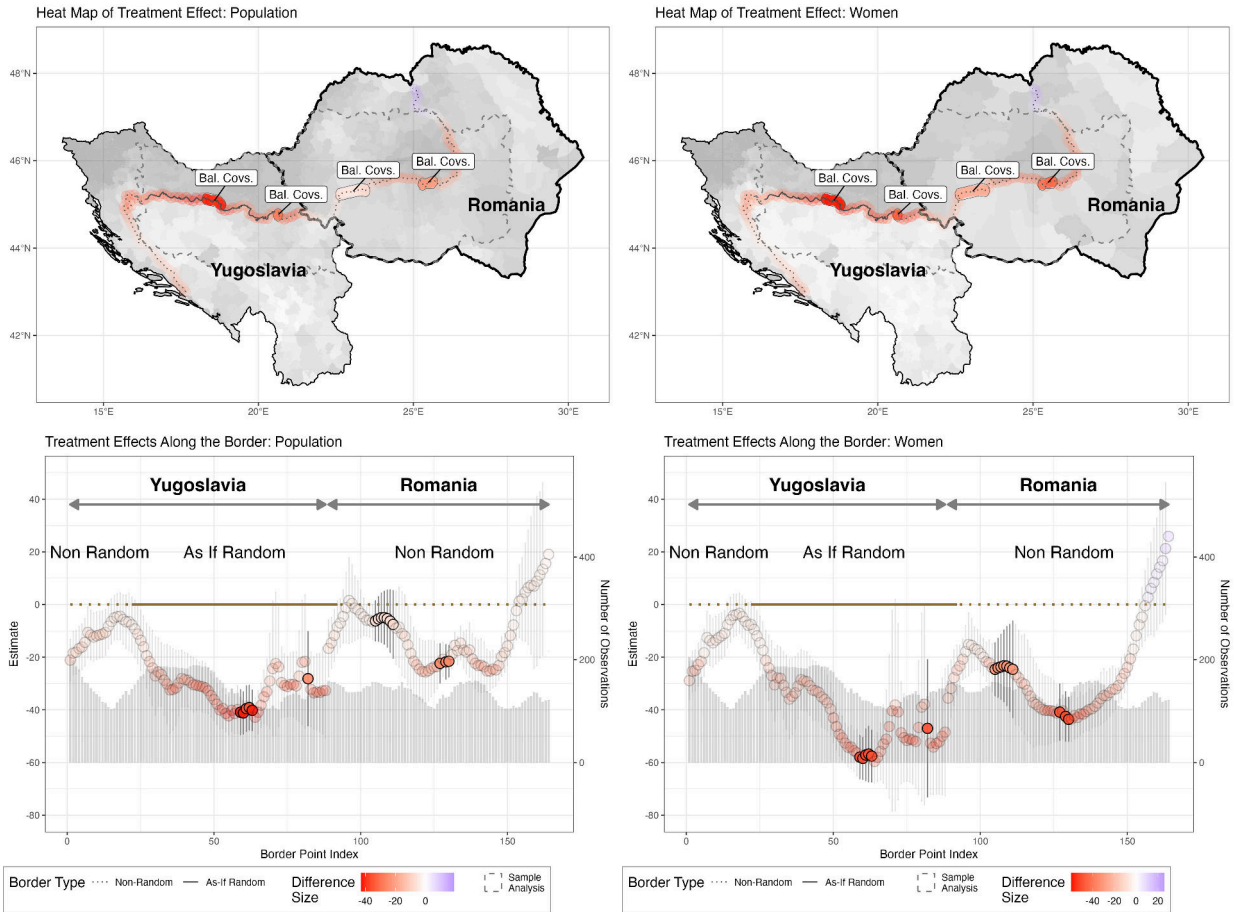
Across the frontier, estimated discontinuities are predominantly negative, indicating lower literacy in former Ottoman areas in 1931. Magnitudes are especially large along the Yugoslav riverine segments (Sava and Una), where local gaps typically range from -27 to -36 percentage points and reach -43 at some points; the Yugoslav land segment shows smaller effects around -11. Romania exhibits gaps that average around -11 percentage points overall, with the Danube segment showing the smallest effects (around -4).[8] Figure 2 visualizes this spatial heterogeneity.

Gender disaggregation reveals even larger discontinuities. Across the Yugoslav border, the Ottoman-side female literacy gap averages about -35 percentage points; across the Romanian border, about -25 points. Among points that satisfy all four credibility criteria, the Yugoslav effect

[7]The units are districts: Romanian *plăşi* grouped into counties (*judeţe*); Yugoslav *srezovi* in 1931, *opštine* grouped by srez thereafter. Errors cluster at the *judeţ* and the *srez*; in 1931 the Yugoslav srez is itself the unit, so clustering is district-robust, and from 1961 onward opštine are grouped under the harmonized 1961 srez classification, carried unchanged through 1991. Each regression is estimated within a single census, so boundary changes across years do not affect inference.

[8]The full distribution of local estimates, bandwidths, and uniform confidence bands is available in the replication archive.

Figure 2. *Romania and Yugoslavia Compared: RD2D Estimates of Population and Female Literacy in 1931*



Notes: Evaluation points along the frontier, mapped (top) and indexed linearly (bottom). Color: estimated local effect (red negative, blue positive); transparency: covariate imbalance. Map shading: observed literacy, total and female. Solid segments are low-stakes (as-if random), dotted are high-stakes; dashed outlines illustrate RD2D bandwidth neighborhoods. Grey bars: observations per local estimate.

is especially pronounced: the credible Yugoslav points all lie on the Sava River, where female estimates center near −56 points and reach −58. Credible Romanian points average around −30, ranging from about −24 along the Danube to −38 along the land border.

These magnitudes reflect extremely low baseline female literacy in former Ottoman areas in 1931: roughly 15% in Yugoslavia and 38% in Romania, compared to 71% and 61% respectively on the Habsburg side.

Figure 3 traces literacy differences in Romania and Yugoslavia from the interwar period to the end of communism. Each country block reports two sets of estimates. Nonparametric RD2D

estimates at border points with balanced covariates are shown in black; OLS comparisons using the full treatment and control regions are shown in grey. The upper row of each block displays the corresponding spatial samples (RD2D bandwidth strips vs. full-region OLS extent), and the lower row shows the literacy-gap trajectories.

In Romania, literacy differences were large in 1931: roughly -10 percentage points with the bivariate estimator and -14 with OLS. The gap was still about -10 points in 1948, at the onset of communism; by 1991 it had shrunk to -1.7 points (Figure 3, top row). The pattern is sharper in Yugoslavia (bottom row). Gaps of -38 to -43 points in the interwar period declined to -9 to -19 percentage points by 1961, after a decade and a half of communist reforms. By 1991, differences had narrowed further, to -7.0 points.

In levels, both countries closed a large share of the initial gap: Romania from 14.3 to 1.7 percentage points (88% closure) and Yugoslavia from 43.5 to 7.0 percentage points (84% closure). These rates appear similar but are misleading, because percentage-point gaps mechanically compress as literacy approaches universal levels. Three bounded-outcome transformations (log-odds, probit, arcsine) all show Romania converged more completely, with the closure-rate advantage ranging from 18 to 56 percentage points depending on functional form (Appendix Table A.1).[9] The nonparametric RD2D estimates (black series) track the OLS series closely and show the same divergence.

Gender differences further illuminate this decline. The right column in Figure 3 shows the evolution of female literacy gaps. In Romania, the Ottoman–Habsburg difference was about -23 points in 1930 (OLS), shrinking to -6 in 1977 and -3.1 in 1991. In Yugoslavia, the gap was far larger, around -56 points in 1931, but diminished to -11.7 by 1991. At near-universal literacy levels, a gap of this size is not trivial: roughly one in ten adults who would otherwise be literate remains illiterate. The persistence of such a gap after four decades of socialist schooling campaigns is therefore substantively important.

[9] Convergence rates are computed as $(\hat{\beta}_{1931} - \hat{\beta}_{1991})/\hat{\beta}_{1931}$, where $\hat{\beta}_t$ is the OLS coefficient on Ottoman treatment in year t. In log-odds: Romania closes 83% versus Yugoslavia's 27%. In probit: 84% versus 45%. In arcsine: 85% versus 67%.

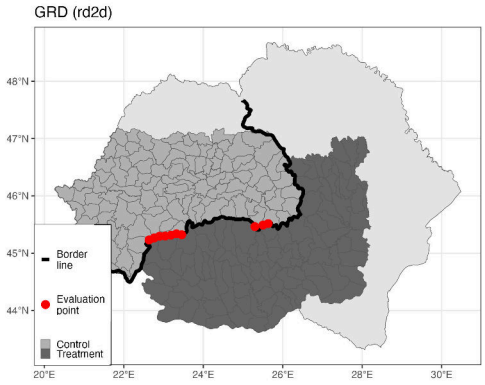
These patterns reflect targeted socialist campaigns: both countries made schooling compulsory for girls and undertook rural initiatives to expand female education (Brunnbauer, 2000; Poulton & Taji-Farouki, 1997).

A natural alternative attributes Romania's faster convergence to its comparatively more coercive communist regime. The evidence weighs against this: Serbia converged rapidly under Tito's comparatively less coercive regime (Appendix Section J).

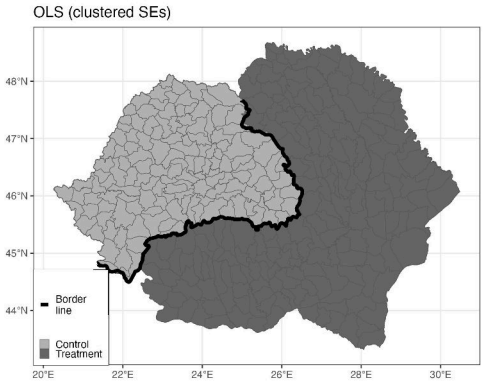
What explains the divergence between Romania and Yugoslavia? Figure 5 previews the answer, plotting male and female literacy trajectories in four incorporation contexts. In Romania both genders converge to near parity; in Yugoslavia the male gap closes while a female residual persists, and that residual is largest in the Muslim-majority districts. The same Ottoman inheritance, under the same socialist state, produced divergent outcomes depending on incorporation. The next section tests this pattern systematically.

Figure 3. *Evolution of Literacy Gaps: Romania (top block) vs. Yugoslavia (bottom block)*

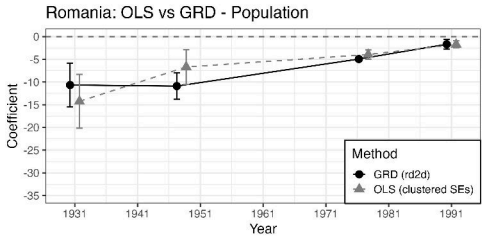
Romania: RD2D sample



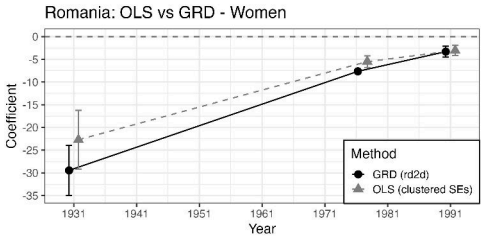
Romania: OLS sample



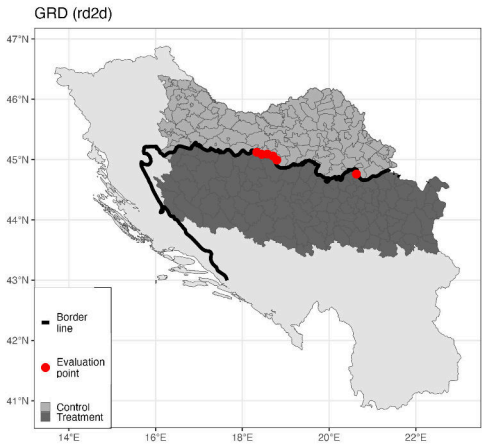
Romania: Total literacy gap over time



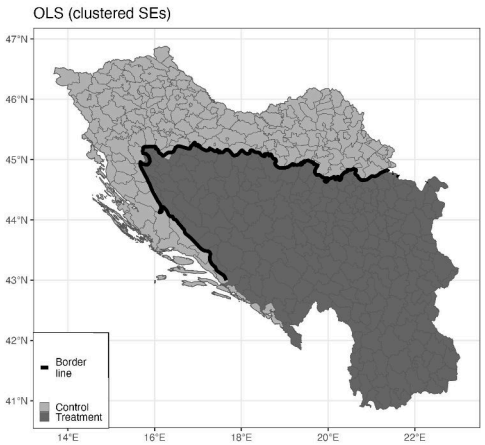
Romania: Female literacy gap over time



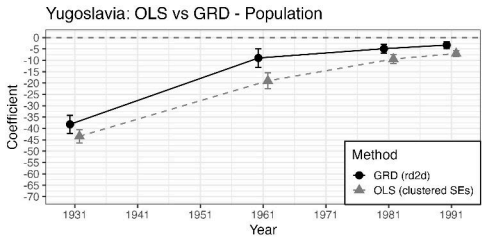
Yugoslavia: RD2D sample



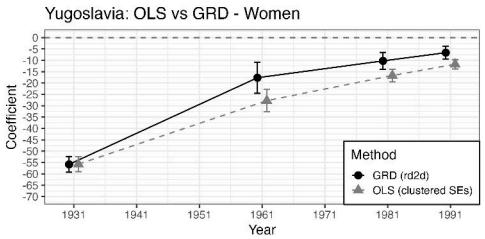
Yugoslavia: OLS sample



Yugoslavia: Total literacy gap over time



Yugoslavia: Female literacy gap over time



Notes: Top block: Romania; bottom: Yugoslavia. In each block, the upper row maps the estimators' samples (RD2D: districts within the frontier bandwidth; OLS: full treatment and control regions); the lower row plots the estimated Ottoman effect on literacy over time, total and female. Black: RD2D at design-credible points; gray: OLS with clustered standard errors; bars are 95% confidence intervals.

5 Testing the Incorporation Mechanism

The preceding section established that Ottoman legacies attenuated more completely in Romania than in Yugoslavia. I test the incorporation mechanism in four steps. Section 5.1 establishes the starting point: an inherited Ottoman deficit in both literacy and administrative capacity. Section 5.2 shows that the deficit's fate depended on incorporation: under the same successor state, Orthodox-majority districts converged like Romania while Muslim-majority districts lagged. Section 5.3 shows the lag was not a supply problem: school inputs equalized, and the residual concentrated on female literacy, where household compliance matters most. Section 5.4 asks whether patriarchal norms alone can explain that pattern. Section 6 then provides a sharper within-Muslim test, tracing the educational consequences of the 1929 sharia-court closure as an example of a state's withdrawal of a group's communal institutions.

5.1 Initial Ottoman Disadvantage: Literacy and State Capacity

H1 predicts that areas governed by the Ottomans entered the post-imperial period at a deficit. The data show two such deficits in 1931: lower literacy and thinner administrative capacity. The literacy gap is established in Section 4 (Figure 2). The administrative capacity gap is measured using the nearest state-employment category in each census: Yugoslavia's 1931 public-service share and Romania's 1930 public-administration share. The two differ in exact occupational coverage but both capture state-administered employment. Averaging RD2D estimates across balanced boundary points yields an Ottoman disadvantage of approximately 3.7 percentage points (mean 2.2%) in Yugoslavia and 0.23 points (mean 0.67%) in Romania (Appendix Figure A.16). These magnitudes are consistent with weaker administrative penetration under the late Ottoman system (Findley, 1996; İnalcık, 1994). Inherited capacity thus establishes the starting disadvantage but on its own cannot account for the variation in later convergence. Within Yugoslavia, Orthodox-majority districts inherited the same Ottoman institutional bundle as Muslim-majority Bosnian and Macedonian districts, yet only the latter failed to converge. Capacity is a contextual constraint

whose consequences depended on local incorporation into the successor state. Section 5.2 develops that contrast.

5.2 Incorporation and Convergence within Yugoslavia

The within-Yugoslavia comparison provides the central contrast. If capacity were the only constraint, convergence should be uniformly slow across all Yugoslav republics that inherited weak Ottoman administration. The incorporation mechanism predicts the opposite: Orthodox-majority Serbian districts, governed under the same federal constitution as Bosnia and Macedonia, should converge at Romanian rates despite sharing the Ottoman institutional bundle. Estimated on the same non-parametric frontier sample in each country, Serbia's Ottoman literacy gap declines from -39.1 percentage points in 1931 to -3.24 points in 1991, mirroring Romania's trajectory (-14.4 to -2.0 points) and ending well below Yugoslavia as a whole (-5.59 points; full results in Appendix Section I and Figure A.8). The same central state thus produced rapid convergence where incorporation was high and slow convergence where it was low, which capacity alone cannot explain. Two qualifications constrain this comparison. Serbia was the politically dominant republic, and Yugoslav federalism gave republics some control over education administration, so a comparison across republics does not hold every institution constant. The district-level analysis that follows therefore corroborates the pattern at finer resolution, and the within-Bosnia sharia-court discontinuity (Section 6), estimated entirely inside one republic, supplies the identification.

Figure 4 plots the treatment-by-religion dynamics in Yugoslavia for the two dominant Ottoman-side religious groups: Muslims and Orthodox Christians. The Catholic share is too small to support a separate interaction (6.7% on the Ottoman side in 1931). Each panel reports two specifications: an uncontrolled estimate (black) and one with geographic controls, elevation and population density (grey); district-level 1931 income is unavailable.[10]

The first panel reports the association between district-level Muslim population share and

[10]Standard errors are clustered at the district level (the *srez* from 1961 onward; in 1931 districts are the units themselves, so errors are district-robust).

literacy.[11] Districts with higher Muslim shares have substantially lower literacy in 1931 (coefficient ≈ -0.37), narrowing to -0.04 by 1991 but remaining statistically significant with geographic controls. The second panel reports the analogous Orthodox-share coefficient. Both the 1931 gradient (-0.27 with controls) and the 1991 residual (-0.03) are smaller in magnitude than for Muslim share. Orthodox-share districts thus started with a smaller penalty and closed it nearly completely, despite sharing the same state and schooling system as Muslim-majority districts, consistent with incorporation, since Orthodox communities experienced no status reversal under the successor state (Donia & Fine, 1994; Velikonja, 2003).

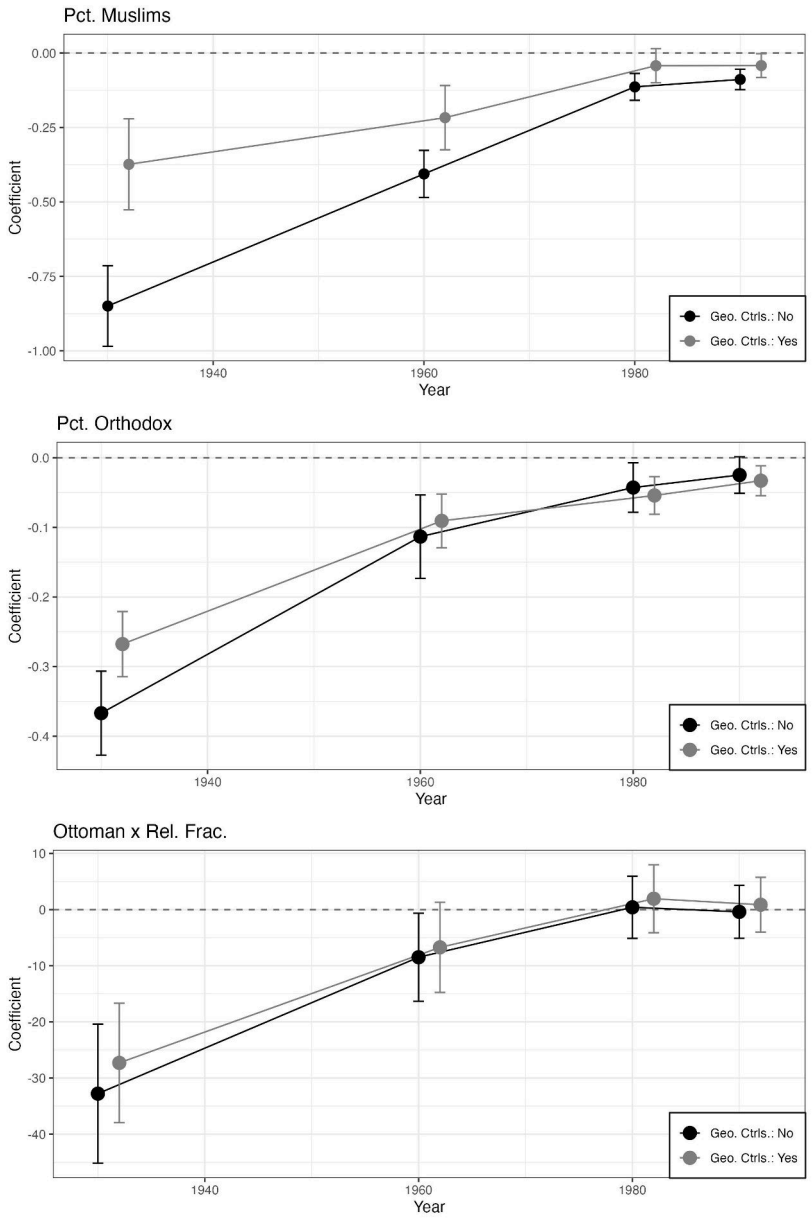
The third panel plots the fractionalization interaction. In 1931, higher religious fractionalization shows a stronger Ottoman literacy disadvantage, but by the 1980s this interaction attenuates to zero, consistent with coordination costs that decline as formal state institutions substitute for informal cooperation (Charnysh, 2019). The divergent temporal trajectories of these two predictors (Muslim share persisting, fractionalization vanishing) under converging supply match the theoretical framework: convergence is conditioned by incorporation, which capacity alone cannot supply.

Muslim share may proxy for correlated Ottoman legacies, poverty, or rural isolation. Three checks speak against this. First, Serbia's Ottoman territories, predominantly Orthodox, converged despite inheriting the same institutional bundle (Appendix Figure A.8). Second, railway density shows no association with Muslim share (Appendix Figure A.17, Panel C), and the gap persists among towns above the same population threshold on the same side of the frontier (Appendix Section K, Figure A.10). Third, these confounds are gender-neutral, while the residual gap concentrates on women within the same districts (Section 5.3). The gradient also predates communism: Muslim-share districts had fewer secondary schools in 1932, when state schooling was run by the interwar kingdom (Appendix Section M).

Selective migration between 1931 and 1961 could confound these estimates if populations with different literacy profiles moved differentially across the Ottoman–Habsburg boundary. The

[11] The independent variable is the percentage of Muslims, not the interaction with Ottoman treatment, since there are virtually no Muslims in the former Habsburg territory.

Figure 4. *Heterogeneous Effects on Literacy by Religion in Yugoslavia*



Notes: OLS coefficients on districts within the RD2D bandwidth; outcome is the literacy rate. Regressor: Muslim share (first panel), Orthodox share (second), Ottoman exposure × religious fractionalization (third). Black: no controls; gray: population density and elevation. Bars are 95% confidence intervals; errors clustered at the district level (the srez from 1961 onward; in 1931 districts are the units, so errors are district-robust).

Ottoman effect is essentially unchanged when reconstructing 1931–1961 population change on a harmonized grid, excluding areas with population decline, and controlling for compositional shifts. In Romania, the 1930 census *populația flotantă* (floating-population) mobility measure is uncorrelated with treatment (Appendix Section H).

5.3 Supply Equalized, Uptake Did Not

If supply-side constraints were binding, equalizing school inputs should have closed the literacy gap. H2 predicts otherwise: where incorporation is low, equalized inputs will not be enough. Appendix Figure A.17 reports regressions of school and teacher provision on Muslim share within the RD2D bandwidth (standard errors clustered by district). Schools per 1,000 exhibit no Muslim-share gradient in any census year. Teachers per capita were lower in Muslim-share areas in 1961 (-0.031), but the gradient had reversed by 1981 ($+0.009$, $p < 0.01$) and remained positive in 1991 ($+0.006$). The residual outlasted this compensatory expansion: whatever sustains it is not a shortage of schools or teachers.[12]

Gender disaggregation shifts the explanation from supply to demand: men and women in the same district face the same schools, so a persistent within-district gap cannot be a matter of district-level supply alone. The gap's behavior differs sharply by context. In Romania and in Orthodox-majority Yugoslav districts, the gender gap closed dramatically (35 to 4 pp and 45 to 12 pp); in Muslim-majority districts it barely narrowed (22.7 to 15.3 pp; Figure 5). By 1991 the male Ottoman penalty had nearly disappeared (2.4 pp), while the female penalty stood at 11.7 pp, against 3.1 pp in Romania (a 9 pp gendered residual). Two demand-side accounts could produce this pattern. Cultural norms about female schooling explain why the gap started large; the question is whether they also explain why it closed in some contexts and not others. The household-investment alternative (Becker, 1981, pp. 191–193) holds that daughters' schooling is the discretionary investment: families that distrust the state keep daughters home first, so state schooling closes the gap only under incorporation. Cross-national comparison already speaks against a uniform-Islam account (Section 5.4); the within-Muslim test of Section 6 separates the two directly.

[12]The teacher gradient had reversed by 1981, so by 1991 cohorts schooled entirely under equal or better provision had entered the adult population. A late-supply account faces two hurdles: supply that affected boys and girls alike cannot explain a female-specific residual, and a sons-first response to scarce places stopped operating once places stopped being scarce. Radio subscriptions, a non-schooling demand check, remained lower in districts with larger Muslim shares in 1981 and 1991 after geography and railway controls (Appendix Figure A.17, Panel B).

Figure 5 plots male and female literacy trajectories in the four incorporation contexts introduced above: Romania, Yugoslavia overall, and Yugoslavia's Muslim-majority and Orthodox-majority Ottoman-side districts, with majority status classified at the 1931 census and held fixed thereafter (later censuses' declared shares are endogenous to the identity politics under study). Three patterns emerge. First, in 1931 every context starts with a large gender gap (41 pp in Yugoslavia overall, 45 pp in Orthodox-majority districts). The Muslim-majority gap looks smaller (23 pp) only because there was almost no literacy to be unequal about: both men and women were near the floor. Second, as schooling arrived, the Muslim-majority gap *widened* (23 pp in 1931 to 34 pp in 1961) because men gained access faster than women, and it remained the largest residual through 1991 (15.3 pp, against 12.3 pp in Orthodox-majority districts). Part of that early widening is plausibly mechanical rather than incorporation-driven: teachers are themselves products of the schooling system, so staffing in previously unschooled areas lagged a generation behind demand, and scarce places went to boys first. The incorporation evidence therefore rests on the gap that remained after provision equalized. Third, by the early 1990s women had reached 96% of male literacy in Romania, 87% in Orthodox-majority districts, and 84% in Muslim-majority districts. Appendix Figure A.11 repeats the comparison with the continuous Muslim and Orthodox shares instead of majority splits, with the same qualitative pattern and sharper estimates; geographic controls leave the gender asymmetry intact (Appendix Table A.4).

How long did exclusion bind? The exclusion documented in Section 3 was interwar, and the state's posture eventually reversed: in 1968 the League of Communists recognized Muslims as a constituent nation (*narod*), with a corresponding category in the 1971 census (Bougarel, 2018, pp. 13–14). The gender gap follows a matching arc. Between 1931 and 1961, the Muslim-majority gap widened by about 12 points even as the Orthodox-majority gap narrowed by 14. After 1961 the Muslim-majority series reversed, narrowing slightly faster than the Orthodox one to 15.3 points by 1991. The residual still visible in 1991 is thus the durable component: the part of the gap that formal recognition did not erase.

Table 2 restates these trajectories as convergence rates. In percentage points the contexts

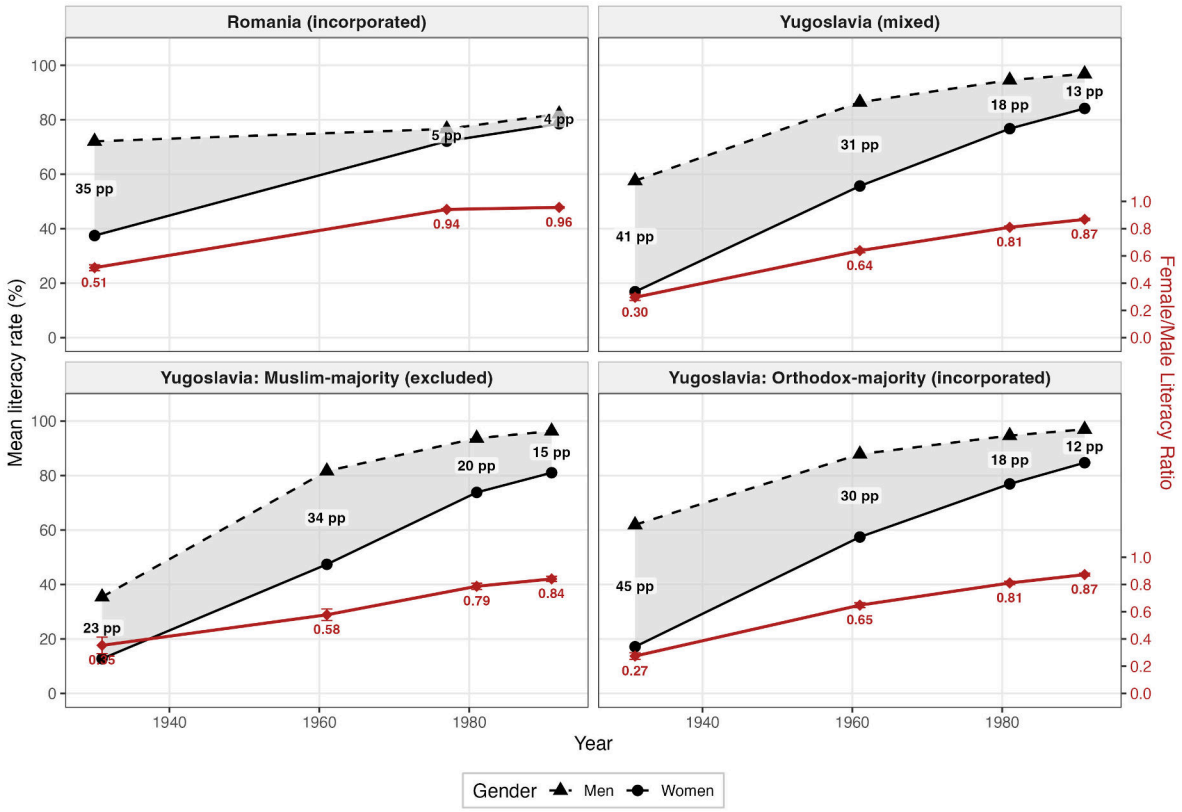
differ only in degree: Romania closes 89.6% of its 1931 gender gap, Muslim-majority districts 32.7%. But percentage points are compressed near the floor and ceiling, where literacy has little room to move (the problem noted above). The log-odds metric corrects this, and in it the contexts differ in kind: Romania closes 84.4% of its gap and Orthodox-majority districts 14.8%, while the Muslim-majority gap does not close at all but widens by 37.1%. The Muslim-majority series rests entirely on directly reported counts: all its districts lie in Bosnia-Herzegovina, whose censuses published gender-specific literacy; imputation touches only other republics' series (Appendix Section C).

Table 2. *Gender-Gap Closure by Incorporation Context*

	Gender Gap (pp)		% of 1931 Gap Closed by 1991	
Context	1931	1991	PP	Log-Odds
Romania (1930–1992)	34.6	3.6	89.6%	84.4%
Yugoslavia, all districts	40.7	12.7	68.9%	7.9%
Yugoslavia: Orthodox-majority districts	44.8	12.3	72.6%	14.8%
Yugoslavia: Muslim-majority districts	22.7	15.3	32.7%	-37.1%

Notes: The gender gap is male minus female literacy, computed from district-level means in the four contexts of Figure 5 (same sample; Romania uses the 1930 and 1992 censuses). "% Closed" is $(g_{1931} - g_{1991})/g_{1931}$ in each metric; log-odds is $\ln(p/(1-p))$, correcting floor and ceiling compression (Appendix Section A). Negative values indicate widening. Entries are descriptive means, not regression estimates; gender-specific literacy for the 1961–1991 Yugoslav censuses is directly reported for Bosnia-Herzegovina, home to the Muslim-majority panel, and partly imputed for Croatia, Macedonia, and Montenegro (Appendix Section C). Majority status is classified at the 1931 census and held fixed; later-year districts are assigned by centroid containment in the 1931 majority polygons (Muslim-majority: 8 districts in 1931, 16–19 successor units thereafter).

Figure 5. *Gender-Specific Literacy Trajectories by Incorporation Context*



Notes: Mean male (dashed) and female (solid) literacy rates per panel; grey ribbon: gender gap (pp); red diamonds: female-to-male ratio (right axis) with 95% CIs from district-level ratios. Top row: Romania | all Ottoman-side RD2D districts. Bottom row: Muslim-majority | Orthodox-majority districts, classified at the 1931 census and held fixed (later years assigned by centroid containment). Romania 1930/1977/1992; Yugoslavia 1931/1961/1981/1991.

Ethnographies of Bosnian Muslim communities consistently identify the Yugoslav state, not female education *per se*, as the object of household resistance. In the interwar kingdom, Muslim families viewed the new state schools as interfaith institutions whose "hidden agenda was to convert Muslim children to Christianity" (Giomi, 2021, p. 42). Under socialism, state schools were experienced as exposure to "state-sponsored areligious ideology" that families "saw as anti-Muslim" (Bringa, 1995, p. 110), and teachers humiliated students for mosque attendance (Henig, 2020, ch. 1).

5.4 Are Islam-Specific Gender Norms a Sufficient Explanation?

The most natural cultural interpretation of the persistent gender gap in Muslim-majority districts is that Islam-specific gender norms drove the outcome on their own, independent of incorporation.

The claim has a distinguished pedigree: in cross-country regressions, Fish (2002) finds female-to-male literacy ratios significantly lower in Muslim-majority societies and treats the subordination of women as one of their durable traits, and Norris & Inglehart (2003) locate the deepest value divide between Muslim and Western publics in attitudes toward gender equality rather than democracy. Under the joint-causation reading defended here, norms produce the baseline gap but cannot by themselves explain where it closes and where it persists.

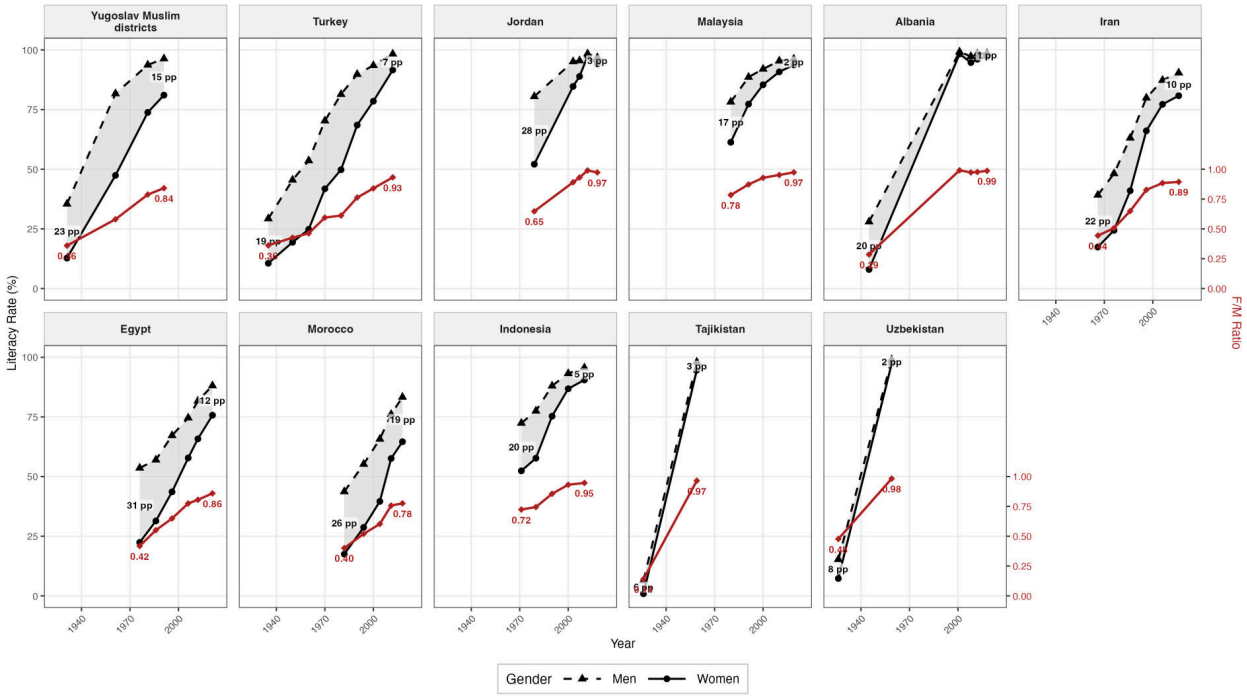
Figure 6 confronts the uniform-Islam reading with the cross-national record, plotting male and female literacy trajectories across eleven Muslim-majority populations. If Islam alone determined the gap, these populations should look alike. Instead, the latest female-to-male ratios range from 0.78 (Morocco) to 0.99 (Albania): some Muslim populations closed the gender gap entirely while others did not.[13] Yugoslavia's Muslim districts fall in the lower half at 0.84 (1991). The comparison is descriptive: these countries differ in development, state capacity, and schooling, so it cannot say why any particular gap persists. It shows only that a uniform-Islam account, by itself, does not fit the cross-national record.

6 A Within-Muslim Test: The 1929 Sharia-Court Closure

If Islam-specific norms were the whole story, institutional variation should not matter among Muslims. Testing that requires a design the previous comparisons cannot deliver. Cross-country comparison confounds religion with regime. The within-Yugoslavia comparison of Section 5 holds the state constant but varies religion, comparing Muslim with Orthodox districts, so a gendered residual there could still reflect Islam-specific norms rather than the state's treatment of Muslims. Separating the accounts requires the opposite variation: institutional treatment that differs *within* the Muslim population, with religion, language, and culture held constant.

[13] Albania also illustrates the group-differential criterion from the introduction: Hoxha's state dismantled the religious institutions of every faith, but the treatment was universal rather than aimed at one group, and Muslims remained the titular majority, so the definition codes Albania as incorporated. Its near-complete convergence is consistent with that coding.

Figure 6. *Male and Female Literacy Trajectories in Muslim-Majority Populations*



Notes: Male (dashed) and female (solid) literacy rates for selected Muslim-majority populations; grey shading is the gender gap, red diamonds the female-to-male ratio (right axis). Age coverage varies by source (15+ WDI, 10+ Yugoslav censuses, 9–49 Soviet). Yugoslav Muslim-majority districts: census data within the RD2D bandwidth (this paper). Soviet Central Asia: 1926 and 1959 censuses (Mickiewicz, 1973). Albania 1945 derived from 20% overall literacy and 92% female illiteracy. All other series from national censuses and World Bank WDI; latest years range from 2012 (Turkey) to 2021 (Egypt), with Morocco 2018 from the Haut-Commissariat au Plan.

This section provides that test: a regression discontinuity at the 5,000-Muslim threshold imposed by the 1929 ZSS (Section 3). Twelve of the fifty districts in the analysis sample (24%) fell below the threshold and lost their local sharia court; comparable districts just above retained theirs.

6.1 The 1929 Closure as an Institutional Shock

Sharia courts in Bosnia-Herzegovina were state-recognized tribunals administering Islamic personal-status law (marriage, divorce, inheritance, custody), operating as departments of the local district court (Karčić, 2005). The pre-closure baseline was uniform: since the Habsburg period, every district court housed a sharia department, "regardless of the number of Muslims" (Karčić, 2005, p. 67).

Under the January 1929 royal dictatorship, the ZSS (21 March 1929) stipulated that district-level (*srez*) sharia courts "shall be established at every district court whose area has at least 5,000 Muslims"; districts below the threshold had their jurisdiction folded into a neighbor's catchment. The official rationale was efficiency (Karčić, 2005, p. 67), but the same regime dissolved parliament, banned confessional parties, and partitioned BiH across four *banovine* in which Muslims were a minority (Djokić, 2007, pp. 68–83), part of broader pressure on Muslim institutional autonomy (Bougarel, 2018, pp. 35, 39–40).

6.2 Design

The design compares settlements on the two sides of the 5,000-Muslim line: those whose parent district lost its sharia court in 1929 against those whose district kept one, with the district's 1921 Muslim count as the running variable. Treatment is assigned at the district level because that is the unit on which the ZSS imposed the threshold, and it uses the 1921 count because that was the census available when the law was drafted (the next census, in 1931, came two years after the closure). Settlements whose parent district had a combined district and city Muslim count below 5,000 are coded *lost-court*; the rest are *retained-court*. Each 1961 settlement is matched to its 1921 parent district by spatial join.

Outcomes come from the 1961 census, which reports illiteracy by gender and four age cohorts (10–19, 20–34, 35–64, 65+). The cohorts recover people schooled before, during, and after the 1929–1946 window in which court access differed.[14] The gender outcomes are male and female illiterates per 100 of the total 10+ population (both sexes in the denominator), so they decompose total illiteracy by sex; the cohort outcomes are age-specific rates, illiterates per 100 of each band's population.

[14]With schooling at roughly ages 7 to 14, only the middle cohorts sat in school while court access differed: the 65+ cohort finished by about 1910, and the 10–19 cohort started after 1946, when both district groups were courtless. Direct exposure is concentrated in the 35–64 cohort and partial in the 20–34. The cohorts thus separate two channels: effects of the court's absence during schooling should sit in the 35–64 and 20–34, effects of durable community change in the 20–34 and 10–19, and the 65+ should show nothing under either, making it the shared placebo (Section 6.4).

Settlement coordinates and pre-treatment covariates come from the Habsburg settlement cross-section of Neubert & Nikolić (2026). Each 1961 *naselje* (settlement) is located by matching it, within its opština, to its 1885 counterpart, which also supplies elevation, an 1885 road indicator, and the 1910 confessional counts. Of the 6,052 Bosnian naselja in the 1961 census, 3,447 match an 1885 settlement and form the analysis frame; the rest, typically places founded, split, or renamed after 1885, cannot be assigned a coordinate or covariates. Because the match criterion, existence in 1885, predates the closure by four decades, it cannot itself respond to treatment; the balance tests below are computed on this matched frame.

Two design diagnostics probe the RD. The first raises no concern: a McCrary density test on the 50 districts shows no bunching at the cutoff ($T = 0.50$, $p = 0.62$), as expected since the 1921 Muslim count was fixed years before the law was drafted. The second, balance across eleven pre-treatment covariates and one post-treatment check, gives a mixed result. The five settlement-level covariates (elevation, the 1910 Orthodox and Catholic shares, 1910 fractionalization, and an 1885 road indicator) are imbalanced. This is the expected consequence of a threshold defined on a population count: districts above it are larger and more central, and Bosnian Muslims historically concentrated in lower-lying administrative centers with denser Habsburg state presence. The design absorbs the imbalance in two ways: the five covariates enter the headline specification as linear controls, and the zero-Muslim tercile introduced below serves as a placebo for whatever district-level confound they might still proxy. The seven district-level covariates are balanced, and these carry the substantive reassurance: confessional-school counts (1895, predating the closure), illiteracy from the 1931 census (schooling completed before the closure), and a WWII-victim share (a post-treatment check) do not jump at the cutoff, so neither pre-closure schooling infrastructure nor differential wartime losses can account for the 1961 gap. Full tests, including standard-error robustness, are in Appendix Section N.

Estimation uses local-linear RD with a triangular kernel and MSE-optimal bandwidth. Because the religious channel operates only through the Muslim community, the headline specification estimates the RD separately within *terciles* of the 1910 Muslim share: 37% of settlements had no

Muslims at all, so T1 collects the zero-share settlements, and T2 and T3 split the positive shares at their median (64%). The split is set by this zero mass point and the median, not searched; a continuous-share specification appears in Appendix Table A.7. Inference is the robust bias-corrected estimator of Calonico *et al.* (2014) with standard errors clustered by opština (municipality), reported as 95% confidence intervals; because the interval is centered on the bias-corrected estimate, it need not contain the local-linear point.

The design implies a three-part signature that any institutional effect of the closure must carry; Table 3 is laid out to test it. (i) *Muslim-conditional.* The religious channel has families to act on only where Muslims lived, which makes T1 a placebo: its settlements sit in the same districts and went through the same 1929 restructuring, but contain no Muslim families, so a district-level confound should appear there while the religious channel cannot. T1 screens confounds but cannot separate institutions from culture, since a norms account also predicts a null there; the culture test is the threshold itself, where norms are continuous. (ii) *Gendered.* Daughters' schooling was the discretionary margin (Section 2), so the harm should fall on women rather than men. (iii) *Cohort-timed.* The 65+ cohort, born by 1896, finished school decades before the closure, so it should show no harm. A confound would need to reproduce all three margins jointly; the results below are read against this signature.

One caution before the results: treatment varies at the district level and only twelve districts fall below the cutoff. That is too few treated clusters for district-level inference to be informative: the wild-cluster bootstrap in Appendix Table A.6 does not reject, but at this design size it under-rejects severely and cannot adjudicate either way (MacKinnon & Webb, 2018).

6.3 Main Estimates and Placebo Tests

Table 3 reports the closure's effect on 1961 illiteracy across the terciles. The first three columns are illiterates per 100 of the 10+ population (total, male, female); the last four are age-cohort illiteracy rates. Treatment is coded as lost-court, so a positive τ means the closure raised illiteracy (harm).

The effect appears only where Muslims were present. The pooled sample, mixing settlements

Table 3. *Naselje-level RD at the 5,000-Muslim Cutoff (1929 ZSS): 1961 Illiteracy by 1910 Muslim-Share Tercile*

	Illiterates per 100 (*svega*)			Age-cohort illiteracy			
	Illit. per 100	Male illit. per 100	Female illit. per 100	Illit. 10–19	Illit. 20–34	Illit. 35–64	Illit. 65+
Pooled (all retained-court + lost-court naselja)							
τ	3.76	-2.72^{***}	2.24	-11.82	-4.29^{***}	4.75	1.91
95% CI	$[-3.7, 2.5]$	$[-5.0, -1.7]$	$[-1.7, 2.8]$	$[-32.6, 1.4]$	$[-11.2, -2.9]$	$[-3.0, 6.8]$	$[-5.1, 5.7]$
p	0.712	0.000	0.654	0.072	0.001	0.445	0.910
Bandwidth	2442	1630	2571	985	1771	2483	2489
N	673	476	796	252	567	682	659
T1 (1910 Muslim share $= 0$*)*							
τ	-0.07	2.28	-2.59	7.54	2.35	-5.10^{**}	-0.61
95% CI	$[-10.7, 3.8]$	$[-2.4, 3.3]$	$[-8.9, 0.5]$	$[-0.9, 6.1]$	$[-10.0, 4.8]$	$[-27.4, -0.7]$	$[-8.4, 11.6]$
p	0.346	0.750	0.081	0.152	0.492	0.039	0.755
Bandwidth	2154	2655	1917	2138	2250	1293	1717
N	360	442	371	285	356	247	296
T2 (1910 Muslim share $\in (0, 64\%]$*)*							
τ	8.77^{***}	2.49	5.46^{***}	10.12^{**}	10.91^{***}	4.56	2.00
95% CI	$[2.0, 11.4]$	$[-1.9, 1.2]$	$[1.8, 8.2]$	$[0.6, 11.3]$	$[1.1, 7.2]$	$[-8.4, 6.7]$	$[-9.1, 6.2]$
p	0.005	0.623	0.002	0.030	0.007	0.819	0.714
Bandwidth	2926	2256	3055	2566	2821	2574	2961
N	264	201	266	233	258	253	254
T3 (1910 Muslim share $> 64\%$*)*							
τ	2.23	0.06	3.06	-16.11	4.14	5.04	2.62
95% CI	$[-9.3, 7.1]$	$[-6.7, 3.6]$	$[-1.6, 4.6]$	$[-57.4, 18.6]$	$[-9.9, 12.9]$	$[-4.1, 9.4]$	$[-10.6, 12.4]$
p	0.798	0.551	0.356	0.317	0.793	0.440	0.878
Bandwidth	3264	3203	3791	2319	3551	3131	3793
N	119	119	169	76	143	117	153
DV mean	39.6	9.9	29.6	15.4	31.7	64.3	77.5
DV SD	13.5	6.1	8.8	16.0	17.3	17.2	17.8

Notes: Each cell is a separate controlled local-linear RD (triangular kernel, MSE-optimal bandwidth, opština-clustered inference). Running variable: parent district's 1921 Muslim count; cutoff 5,000 Muslims (12 of 50 districts fall below). Terciles are of the 1910 Muslim share: T1 is the zero-share group; T2 and T3 split the positive shares at their median (64%). The five imbalanced pre-treatment covariates enter as linear controls (Appendix Table A.5). The estimate is the local-linear point (lost − retained); the 95% CI and p are the robust bias-corrected estimator of Calonico, Cattaneo & Titiunik (2014), so the interval is not generally centered on the estimate. The first three columns are illiterates per 100 of the total 10+ population (*svega*); the four age columns are age-specific illiteracy rates. N is the effective sample inside the bandwidth. Positive τ = closure raised illiteracy (harm). Treatment is assigned at the district; assignment-level inference (wild-cluster bootstrap and randomization inference) is reported in Appendix Table A.6 and cannot reject at conventional levels, so stars here reflect opština-clustered inference only. *** $p < 0.01$, ** $p < 0.05$.

the channel can reach with settlements it cannot, shows no harm anywhere: its only significant cells (men and the two youngest cohorts) are negative. In T1, the zero-Muslim tercile, losing the court likewise raises illiteracy for no one; every coefficient is null or negative (the 35–64 cohort's significantly so). Where cells are significant they favor the lost-court side, so any residual district-level imbalance that survives the controls runs against the harm the theory predicts; the T2 estimates below emerge despite that tilt, not because of it. In these settlements the religious channel has no Muslim families to act on, and the broader 1929 restructuring leaves no gendered residual on its own. The effect is concentrated in T2, the moderate-Muslim tercile. In T3, the high-share tercile, the female coefficient points the same way (about 3 points) but is imprecise, because few high-share districts sit near the cutoff. A continuous-share specification confirms the pattern is not an artifact of the split: pooled across all positive-share settlements the female effect is 5.25, varying little with the local share (Appendix Table A.7).

Within T2 the signature is gendered and confined to the post-1929 cohorts. Losing the court raises female illiteracy by about 5 points from a base near 30, more than twice the effect on men (about 2 points, not significant), and population illiteracy by about 9 (each outcome is estimated at its own MSE-optimal bandwidth, so the gender components need not sum to the population effect). The harm falls on the cohorts schooled after the closure: illiteracy rises about 10 points each among those aged 10–19 and 20–34 in 1961, and not at all in the two older cohorts. The 65+ cohort, schooled at least fifteen years before the closure, doubles as a timing placebo: the harm pattern should not appear there, and it does not. That pattern is what a religious-institutional channel predicts and what a pure cultural-norms account does not: norms travel with the household and cannot explain why the gap opens at the institutional threshold and only where Muslim communities were dense.

What did the court do, and why did its removal fall on daughters? The sharia court was the state-recognized Muslim institution that tied communal life to the modern state in a form households could accept, and its judges lent state schooling a credibility it did not otherwise have. Karčić (2019, p. 113) records that the qadis presented themselves as "champions of cultural

progress amongst Muslims," and that their own children's attendance helped "break down the suspicions that existed amongst the Muslim masses against modern education." Removing the local court withdrew that endorsement.

When the local Muslim institution was withdrawn, families uncertain about the state's intentions pulled back from state-mediated schooling where household discretion was greatest. Sons' education was a default investment in expected labor-market returns; daughters' was discretionary, and it touched modesty and family honor. On reaching puberty a daughter was expected to keep apart from men outside her family, which a secular school among strangers made impossible (Giomi, 2021, p. 53), and interwar Muslim reform associations were still contesting the widespread view that sending girls to school was "equivalent to destroying their faith and morality" (Giomi, 2021, p. 234). Under distrust, daughters were withdrawn first (Becker, 1981; Jayachandran, 2015). The channel therefore lands on women in Muslim-present districts only, producing exactly the gendered, T2-concentrated signature in the table.

6.4 Persistence and Interpretation

The treatment differential did not last. On 5 March 1946 the communist regime abolished the remaining sharia courts in Bosnia and Herzegovina, leaving lost-court and retained-court districts equally courtless (Karčić, 2019, p. 173). Treatment was therefore a seventeen-year window of difference, from 1929 to 1946, not a permanent divide. That timing clarifies what Table 3 shows. The largest effects appear in the 10–19 cohort, schooled entirely after 1946, and the 20–34 cohort, schooled partly after it, while the 35–64 cohort, the one actually in school while court access differed, shows no significant effect. Children schooled after 1946 never experienced the differential themselves; their deficit must reflect something the window left behind.

Two mechanisms could carry the window's effects forward. First, intergenerational transmission: mothers schooled or kept home during the window passed the resulting norm to their daughters. Second, community lock-in: once female schooling was de-normalized in a district, staying home remained the local default long after the courts were equalized. The design cannot

tell these apart. What it does show is that seventeen years of institutional difference left a deficit still visible in 1961, fifteen years after the difference had been erased.

The evidence in this section commits to one claim: institutional variation produced large gender-asymmetric effects *within* a Muslim population that was culturally, linguistically, and religiously homogeneous. Because the cultural baseline does not vary at the 5,000-Muslim threshold, a culture-only account cannot generate the T2 discontinuity. The result instead shows an institutional shock acting on a cultural baseline: patriarchal norms produced the gender gap, and they dissolved slowly everywhere; but where the local court was lost, the closure weakened households' engagement with the state schooling that was narrowing that gap, so there it narrowed more slowly still. The closure identifies one mechanism of interwar exclusion; the geographic RD of Section 4 captures the broader imperial-boundary inheritance.

This within-Muslim RD has two limits. First, the treated group is small: only twelve of fifty districts lost their courts. District-level inference therefore carries little information: the wild-cluster bootstrap does not reject but has little power with twelve treated clusters, and randomization inference at the district level yields p-values between 0.19 and 0.63 (Appendix Table A.6). The T2 estimates are suggestive local evidence, not a decisive test. Second, because the threshold is a population count, districts above it are larger and more penetrated by the state, and controls for the five imbalanced covariates cannot fully rule out correlated characteristics such as urbanization. A stronger version of this worry is that districts above the cutoff had denser Muslim institutional life across the board (more mektebs, mosques, and endowments), so the discontinuity may capture the loss of a bundle of Muslim institutions rather than of the court alone. That blurs which institution produced the effect; it does not rescue the cultural account, because a bundle of institutions is still institutions.

The three-part signature specified before estimation weighs against these limits. A rival account must explain three facts at once: why the harm falls on women rather than men; why it appears only in the cohorts schooled after 1929, and not at all among the 65+; and why it exists only where Muslims lived, vanishing in the zero-Muslim settlements that went through the same

1929 restructuring. Sampling noise or a generic size-and-urbanization confound might produce any one of these; producing all three together, at the institutional threshold, is predicted only by the sharia-court channel; no culture-only account or level confound does so.

7 Discussion

Imperial legacies are not erased by state capacity alone. In the cases examined, they faded where successor states governed populations incorporated into the national project, and persisted where incorporation was absent. Persistence is not a law but a default, one that obtains when the conditions for erasure are not met.

Within Yugoslavia, Orthodox-majority Serbian districts converged at rates comparable to Romania while Muslim-majority districts did not. Combined with evidence weighing against migration, federalism, and differential repression, this pattern is most consistent with the incorporation mechanism and difficult to reconcile with capacity or coordination costs alone. The two-country comparison reinforces this: convergence was nearly complete only in Romania, where incorporation was high. A within-Muslim regression discontinuity at the 5,000-Muslim cutoff of the 1929 sharia-court closure (Section 6) provides a stricter test. Holding religion, language, and culture constant, the exclusion shock raised women's illiteracy by about five points in the moderate-Muslim tercile, left men's largely unaffected, and produced no comparable disadvantage for women where there were no Muslims.

The argument does not deny that conservative gender norms produce a baseline gender gap in Muslim societies (Fish, 2002; Norris & Inglehart, 2003; Spierings *et al.*, 2009); it specifies the condition that governs how far state schooling closes it. Within Yugoslavia, Orthodox-majority districts began with the largest 1931 gender gap of any context (45 points) and narrowed it to 12 within the same federal socialist state. Conservative norms are present in converging and persistent cases alike; the persistent cases lack the incorporation that translates state schooling into household compliance.

The article makes two contributions. First, it adds the missing conditionality to the persistence

consensus: the literature documents that imperial institutions generate durable inequalities, but not when state-led efforts to erase them succeed. The incorporation framework supplies one answer: convergence requires not only that states can build schools but that the populations they govern are willing to use them, and incorporation governs that willingness. Second, it locates where the constraint binds: in household compliance, with effects concentrated on daughters' schooling, the decision most dependent on trust in the state.

Two cases beyond the Balkans illustrate the framework's reach. In Kemalist Turkey, the framework is consistent with the slower convergence Kurdish-majority provinces show despite comparable state investment (Tezcür & Gurses, 2017). In the Soviet successor states, convergence accelerated during *korenizatsiya* (the policy of minority nativization) and slowed after its withdrawal (Fitzpatrick, 1979). The contemporary implication follows directly: expanding schooling for a group the state simultaneously excludes should produce the pattern documented here, supply outrunning uptake in the choices left to households. Whether incorporation functions as a policy variable or a slow-moving inheritance, and how quickly it can be rebuilt once eroded, are open empirical questions this paper raises but does not resolve.

AI Disclosure

Data collection: ABBYY FineReader (ABBYY), Amazon Textract (AWS), Google Gemini 1.5 Pro (Google), and Claude 3.5 Sonnet (Anthropic) were used in 2025 for OCR and table extraction from scanned census volumes and statistical yearbooks (1931–1991). No data were generated beyond transcription of the printed sources, and all outputs were verified by the author against those sources.

Analysis code: Claude Opus 4.8 and Claude Fable 5 (Anthropic, via Claude Code, 2026) assisted in drafting and refactoring the R replication code; all code was reviewed and executed by the author.

Writing support: The same Claude models assisted revision of individual passages of author-written text; ChatGPT GPT-4o (OpenAI, 2025) was used for copyediting, with all edits individually accepted or rejected by the author. No section of the manuscript, including the literature review,

theoretical framework, and interpretation of results, was drafted by AI; the author assumes sole responsibility for all content. Years give approximate dates of use; precise model versions were not recorded.

Bibliography

Abad, Leticia Arroyo, & Maurer, Noel. 2021. History Never Really Says Goodbye: A Critical Review of the Persistence Literature. *Journal of Historical Political Economy*, **1**(1), 31–68.

Acemoglu, Daron, Johnson, Simon, & Robinson, James A. 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, **91**(5), 1369–1401.

Acharya, Avidit, Blackwell, Matthew, & Sen, Maya. 2018. *Deep Roots: How Slavery Still Shapes Southern Politics.* Princeton, NJ: Princeton University Press.

Alesina, Alberto, & La Ferrara, Eliana. 2005. Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, **43**(3), 762–800.

Alibašić, Ahmet, & Zubčević, Asim. 2009. Islamic Education in Bosnia and Herzegovina. *Pages 43–57 of:* Aslan, Ednan (ed), *Islamische Erziehung in Europa / Islamic Education in Europe.* Wiener islamisch-religionspädagogische Studien, vol. 1. Wien and Köln and Weimar: Böhlau Verlag.

Ballinger, Pamela. 2003. *History in Exile: Memory and Identity at the Borders of the Balkans.* Princeton, NJ: Princeton University Press.

Ballinger, Pamela, & Ghodsee, Kristen. 2011. Socialist Secularism: Religion, Modernity, and Muslim Women's Emancipation in Bulgaria and Yugoslavia, 1945–1991. *Aspasia*, **5**, 6–27.

Barkey, Karen. 1994. *Bandits and Bureaucrats: The Ottoman Route to State Centralization.* Ithaca, NY: Cornell University Press.

Barkey, Karen. 2008. *Empire of Difference: The Ottomans in Comparative Perspective.* New York: Cambridge University Press.

Bazzi, Samuel, Hilmy, Masyhur, & Marx, Benjamin. 2025. Religion, education, and the state. *The review of economic studies.* Advance article.

Becker, Gary S. 1981. *A Treatise on the Family.* Cambridge, MA: Harvard University Press.

Becker, Sascha, Boeckh, Katrin, Hainz, Christa, & Woessmann, Ludger. 2016. The Empire Is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy. *Economic Journal*, **126**(590), 40–74.

Becker, Sascha O., & Woessmann, Ludger. 2009. Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *The Quarterly Journal of Economics*, **124**(2), 531–596.

Blanc, Guillaume, & Kubo, Masahiro. 2026 (February). *The Making of France*. Working paper. Available at https://hal.science/hal-04292485.

Bougarel, Xavier. 2018. *Islam and Nationhood in Bosnia-Herzegovina: Surviving Empires*. London: Bloomsbury Academic.

Braude, Benjamin, & Lewis, Bernard (eds). 1982. *Christians and Jews in the Ottoman Empire. The Functioning of a Plural Society*. New York: Holmes and Meier.

Bringa, Tone. 1995. *Being Muslim the Bosnian Way: Identity and Community in a Central Bosnian Village*. Princeton, NJ: Princeton University Press.

Brubaker, Rogers. 1996. *Nationalism Reframed: Nationhood and the National Question in the New Europe*. Cambridge: Cambridge University Press.

Brunnbauer, Ulf. 2000. From Equality without Democracy to Democracy without Equality? Women and Transition in Southeast Europe. *Journal for Labour and Social Affairs in Eastern Europe*, **3**(3), 151–168.

Calonico, Sebastian, Cattaneo, Matias D., & Titiunik, Rocio. 2014. Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, **14**(4), 909–946.

Cattaneo, Matias D., Jansson, Michael, & Ma, Xinwei. 2018. Manipulation Testing Based on Density Discontinuity. *The Stata Journal*, **18**(1), 234–261.

Cattaneo, Matias D., Titiunik, Rocio, & Yu, Ruiqi Rae. 2025. rd2d: Causal Inference in Boundary Discontinuity Designs. *Working Paper. arXiv Preprint.* Available at https://doi.org/10.48550/arXiv.2505.07989.

Charnysh, Volha. 2019. Diversity, Institutions, and Economic Outcomes: Post-WWII Displacement in Poland. *American Political Science Review*, **113**(2), 423–441.

Darden, Keith, & Grzymała-Busse, Anna M. 2006. The Great Divide: Literacy, Nationalism, and the Communist Collapse. *World Politics*, **59**(1), 83–115.

Dell, Melissa. 2010. The Persistent Effects of Peru's Mining Mita. *Econometrica*, **78**(6), 1863–1903.

Djokić, Dejan. 2007. *Elusive Compromise: A History of Interwar Yugoslavia.* London: Hurst & Company.

Donia, Robert, & Fine, John. 1994. *Bosnia and Hercegovina: A Tradition Betrayed.* New York: Columbia University Press.

Douglas, R. M. 2012. *Orderly and Humane: The Expulsion of the Germans after the Second World War.* New Haven, CT: Yale University Press.

Dyker, D. A. 1972. The Ethnic Muslims of Bosnia—Some Basic Socio-Economic Data. *The Slavonic and East European Review*, **50**(119), 238–256.

Findley, Carter Vaughn. 1996. The Ottoman Administrative Legacy and the Modern Middle East. *In:* Brown, L. Carl (ed), *Imperial Legacy: The Ottoman Imprint on the Balkans and the Middle East.* Columbia University Press.

Fish, M. Steven. 2002. Islam and Authoritarianism. *World Politics*, **55**(1), 4–37.

Fitzpatrick, Sheila. 1979. *Education and Social Mobility in the Soviet Union, 1921–1934.* Cambridge: Cambridge University Press.

Fortna, Benjamin C. 2011. *Learning to Read in the Late Ottoman Empire and the Early Turkish Republic.* London: Palgrave Macmillan.

Fouka, Vasiliki. 2020. Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I. *The Review of Economic Studies*, **87**(1), 204–239.

Giomi, Fabio. 2021. *Making Muslim Women European: Voluntary Associations, Gender, and Islam in Post-Ottoman Bosnia and Yugoslavia (1878-1941).* Budapest: Central European University Press. xviii + 401 pp.

Grosfeld, Irena, & Zhuravskaya, Ekaterina. 2015. Cultural vs. economic legacies of empires: Evidence from the partition of Poland. *Journal of Comparative Economics*, **43**(1), 55–75.

Gümüşçü, Osman. 2010. Siyasi coğrafya açısından sınırlar ve tarihi süreç İçinde türkiye'de sınır kavramı. *Bilig: Türk dünyası sosyal bilimler dergisi*, **52**(Kış), 79–104.

Hadžijahić, Muhamed, Traljić, Mahmud, & Šukrić, Nijaz. 1977. *Islam i muslimani u Bosni i Hercegovini.* Sarajevo: Starješinstvo Islamske zajednice u SR Bosni i Hercegovini.

Henig, David. 2020. *Remaking Muslim Lives: Everyday Islam in Postwar Bosnia and Herzegovina.* Interpretations of Culture in the New Millennium. Urbana: University of Illinois Press.

İnalcık, Halil. 1973. *The Ottoman Empire: The Classical Age, 1300-1600.* London: Weidenfeld and Nicolson.

İnalcık, Halil. 1994. The Ottoman State: Economy and Society, 1300–1600. *Pages 9–409 of:* İnalcık, Halil, & Quataert, Donald (eds), *An Economic and Social History of the Ottoman Empire, 1300–1914.* Cambridge: Cambridge University Press.

İnalcık, Halil, & Quataert, Donald (eds). 1994. *An Economic and Social History of the Ottoman Empire, 1300-1914.* Cambridge: Cambridge University Press.

Ingrao, Charles W. 2000. *The Habsburg Monarchy, 1618–1815.* New York: Cambridge University Press.

Institutul Central de Statistică. 1930. *Recensământul general al populaţiei României din 29 decemvrie 1930 [The 1930 General Census of the Population of Romania].* Bucharest: Editura Institutului Central de Statistică.

Jayachandran, Seema. 2015. The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, **7**, 63–88.

Judson, Pieter M. 2016. *The Habsburg Empire: A New History.* Cambridge, MA: Belknap Press of Harvard University Press.

Karčić, Fikret. 2019. *Shari'a courts in Yugoslavia 1918–1941.* Sarajevo: Center for Advanced Studies and El-Kalem. Translated by Desmond Maurer; originally published in Bosnian, Sarajevo: Vrhovno islamsko starješinstvo, 1986.

Karčić, Fikret. 2005. *Šerijatski sudovi u Jugoslaviji 1918–1941.* 2 edn. Sarajevo: Fakultet islamskih nauka u Sarajevu and el-Kalem.

Keele, Luke, Titiunik, Rocío, & Zubizarreta, José R. 2015. Enhancing a Geographic Regression Discontinuity Design Through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout. *Journal of the Royal Statistical Society Series A: Statistics in Society*, **178**(1), 223–239.

Kemura, Ibrahim. 1986. *Uloga "Gajreta" u društvenom životu Muslimana Bosne i Hercegovine (1903–1941).* Sarajevo: Veselin Masleša.

Kocher, Matthew A., & Monteiro, Nuno P. 2016. Lines of Demarcation: Causation, Design-Based Inference, and Historical Research. *Perspectives on Politics*, **14**(4), 952–975.

Kurtaran, Uğur. 2018. Ottoman–Austria Border Determination Works and Newly Determined Borders According to the Treaty of Belgrade of 1739. *Pages 169–193 of:* Amedoski, Dragana (ed), *Belgrade 1521–1867.* Belgrade: Institute of History & Yunus Emre Enstitüsü, Turkish Cultural Centre Belgrade.

Lankina, Tomila, & Getachew, Lullit. 2013. Competitive Religious Entrepreneurs: Christian Missionaries and Female Education in Colonial and Post-Colonial India. *British Journal of Political Science*, **43**(1), 103–131.

Lupu, Noam, & Peisakhin, Leonid. 2017. The Legacy of Political Violence Across Generations. *American Journal of Political Science*, **61**(4), 836–851.

MacKinnon, James G., & Webb, Matthew D. 2018. The Wild Bootstrap for Few (Treated) Clusters. *The Econometrics Journal*, **21**(2), 114–135.

McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142**(2), 698–714.

Mekić, Sejad. 2017. *A Muslim Reformist in Communist Yugoslavia: The Life and Thought of Husein Đozo.* Contemporary Thought in the Islamic World, no. 5. London and New York: Routledge.

Michalopoulos, Stelios, & Papaioannou, Elias. 2013. Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, **81**(1), 113–152.

Mickiewicz, Ellen (ed). 1973. *Handbook of Soviet social science data.* New York: Free Press.

Neubert, Magnus, & Nikolić, Stefan. 2026. Why Railways Fail: Colonial Railways and Economic Development in Habsburg Bosnia-Herzegovina. *Journal of Development Economics*, **181**, 103747.

Norris, Pippa, & Inglehart, Ronald. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World.* Cambridge: Cambridge University Press.

Nunn, Nathan, & Wantchekon, Leonard. 2011. The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, **101**(7), 3221–3252.

Paglayan, Agustina. 2022. Education or Indoctrination? The Violent Origins of Public School Systems in an Era of State-Building. *American Political Science Review*, **116**(4), 1242–1257.

Paglayan, Agustina. 2024. *Raised to Obey: The Rise and Spread of Mass Education.* Princeton: Princeton University Press.

Paglayan, Agustina S. 2026. Education as a political institution. *Annual review of political science*, **29**, 525–546.

Paine, Jack, Qiu, Xiaoyan, & Ricart-Huguet, Joan. 2025. Endogenous Colonial Borders: Precolonial States and Geography in the Partition of Africa. *American Political Science Review*, **119**(1), 1–20.

Pop-Eleches, Grigore, & Tucker, Joshua A. 2017. *Communism's Shadow: Historical Legacies and Contemporary Political Attitudes*. Princeton, NJ: Princeton University Press.

Poulton, Hugh, & Taji-Farouki, Suha (eds). 1997. *Muslim Identity and the Balkan State*. New York: New York University Press.

Purivatra, Atif. 1974. *Jugoslavenska muslimanska organizacija u političkom životu Kraljevine Srba, Hrvata i Slovenaca*. Sarajevo: Svjetlost.

Ramet, Sabrina P. 1992. *Nationalism and Federalism in Yugoslavia, 1962–1991*. Bloomington: Indiana University Press.

Shoup, Paul. 1963. Yugoslavia's National Minorities under Communism. *Slavic Review*, **22**(1), 64–81.

Simpser, Alberto, Slater, Dan, & Wittenberg, Jason. 2018. Dead But Not Gone: Contemporary Legacies of Communism, Imperialism, and Authoritarianism. *Annual Review of Political Science*, **21**, 419–439.

Sorabji, Cornelia. 1989. *Muslim Identity and Islamic Faith in Sarajevo*. Ph.D. thesis, University of Cambridge.

Spierings, Niels, Smits, Jeroen, & Verloo, Mieke. 2009. On the Compatibility of Islam and Gender Equality: Effects of Modernization, State Islamization, and Democracy on Women's Labor Market Participation in 45 Muslim Countries. *Social Indicators Research*, **90**(3), 503–522.

Stoye, John. 1994. *Marsigli's Europe, 1680-1730: The Life and Times of Luigi Ferdinando Marsigli, Soldier and Virtuoso*. New Haven: Yale University Press.

Tezcür, Güneş Murat, & Gurses, Mehmet. 2017. Ethnic exclusion and mobilization: the Kurdish conflict in Turkey. *Comparative Politics*, **49**(2), 213–234.

Valencia Caicedo, Felipe. 2019. The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America. *The Quarterly Journal of Economics*, **134**(1), 507–556.

Velikonja, Mitja. 2003. *Religious Separation and Political Intolerance in Bosnia-Herzegovina.* College Station, TX: Texas A&M University Press.

Vogler, Jan P. 2019. Imperial Rule, the Imposition of Bureaucratic Institutions, and Their Long-Term Legacies. *World Politics*, **71**(4), 806–863.

Wimmer, Andreas. 2002. *Nationalist Exclusion and Ethnic Conflict: Shadows of Modernity.* Cambridge: Cambridge University Press.

Wimmer, Andreas. 2013. *Waves of War: Nationalism, State Formation, and Ethnic Exclusion in the Modern World.* Cambridge Studies in Comparative Politics. Cambridge: Cambridge University Press.

Wittenberg, Jason. 2006. *Crucibles of Political Loyalty: Church Institutions and Electoral Continuity in Hungary.* New York: Cambridge University Press.

Supplementary Information

Contents

A Bounded-Outcome Transformations **A2**

B Quantitative Indicators of Interwar Yugoslav Exclusion **A2**

C Data Collection, Gender Imputation, and Border Construction **A2**

D Covariate Balance along the Frontier **A2**

E Tests for Manipulation around the Threshold **A6**

F Multi-Score RD Design: Over-Time Estimates and Distance Comparison **A6**

G Traditional GRD Benchmark **A6**

H Migration **A9**

I Federalism **A11**

J Differential Levels of Repression **A12**

K Muslim Share and Poverty: A Descriptive Check **A12**

L Gender-Disaggregated Muslim-Share Regressions: Robustness **A12**

M Mechanism Diagnostics: Gender Residual, Adoption, and Pre-Communist Baseline **A13**

N Sharia-Court RD: Design Validation and Robustness **A15**

O Additional Figures and Tables Referenced in Main Text **A20**

Roadmap. This supplement is organized as follows. Section A reports bounded-outcome convergence transformations. Section B documents the quantitative indicators of interwar Yugoslav exclusion. Section C provides pointers to data construction materials. Sections D–I provide identification-defending evidence: covariate balance (D), manipulation tests (E), multi-score RD (F), a traditional GRD benchmark (G), migration controls (H), and federalism (I). Section J reports the differential-repression test. Sections K–N test alternative mechanisms and report robustness: Muslim share and poverty (K), gender-disaggregated Muslim-share regressions (L), mechanism diagnostics (M), and sharia-court RD design validation (N). Section O collects additional figures and tables.

A Bounded-Outcome Transformations

A potential concern with the convergence analysis is that percentage-point gaps mechanically compress as literacy approaches universal levels: one cannot have more than 100% literacy. A skeptical reader might worry that the similar percentage-point convergence rates in Romania (88%) and Yugoslavia (84%) genuinely reflect comparable progress, with the remaining gaps being small in absolute terms. To address this concern, I re-estimate all four gap regressions (two countries, total and female literacy) using three standard transformations for bounded outcomes: (1) the log-odds (logit) transformation, $\ln(p/(1-p))$; (2) the probit (inverse-normal) transformation, $\Phi^{-1}(p)$; and (3) the arcsine square-root transformation, $\arcsin(\sqrt{p})$.[15] Each maps the bounded [0,1] interval onto a wider scale that penalizes gaps near the ceiling more heavily than raw percentage points, but each does so with a different functional form. If the convergence difference between Romania and Yugoslavia is robust across all three, rather than an artifact of the specific curvature of the logit function, the finding is on solid ground.

Table A.1 presents the results. The pattern is consistent across all three transformations. While Romania and Yugoslavia appear nearly identical in percentage-point terms (88% vs. 84% of the 1931 gap closed), every bounded-outcome transformation reveals a large divergence. Romania's convergence rate is stable across specifications: 88% in percentage points, 83% in log-odds, 84% in probit, and 85% in arcsine. Yugoslavia, by contrast, drops sharply: from 84% in percentage points to just 27% in log-odds, 45% in probit, and 67% in arcsine. The difference in convergence between the two countries ranges from 18 percentage points (arcsine, the most conservative) to 56 percentage points (log-odds, the most aggressive), with probit intermediate at 38 percentage points. All three bounded-outcome transformations thus agree that Romania's convergence was substantially more complete than Yugoslavia's, a conclusion invisible in raw percentage points.

When both countries approach high literacy levels, percentage-point gaps mechanically shrink. Yugoslavia's residual gap, though only 7 percentage points in levels, remains large in proportional terms, equivalent to roughly 8 percentage points when evaluated at a 90% baseline.[16] Romania, by contrast, achieved near-complete equalization regardless of the metric used. The fact that this conclusion holds across three transformations with different functional forms, each penalizing ceiling compression differently, demonstrates that it is not an artifact of any particular modeling choice.

B Quantitative Indicators of Interwar Yugoslav Exclusion

Section 3 summarizes three patterns of interwar Muslim exclusion (underrepresentation, institutional dismantlement, and bloc voting). Table A.2 reports the underlying quantitative indicators.

C Data Collection, Gender Imputation, and Border Construction

A separate data appendix documents the GIS dataset of Ottoman borders (152 georeferenced maps, 1326–1920), the gender-specific literacy imputation procedure for 1961–1991 Yugoslav censuses, and the historiographical analysis of border formation underlying the as-if random assignment claim. These materials, including the complete map bibliography, imputation code with national benchmarks, and the border-drawing narrative with all supporting figures, are included in the replication archive.

D Covariate Balance along the Frontier

Figure A.1 reports segment-level covariate balance for elevation, slope, precipitation, temperature, and wheat suitability. The pattern is heterogeneous: elevation differences are sign-reversing along the border; slope is largely balanced in Yugoslavia; precipitation is higher on the Ottoman side in western Yugoslavia but converges northward. This heterogeneity motivates the

[15] For the log-odds and probit transformations, I apply a small boundary adjustment $p_{\text{adj}} = \max(\min(p, 0.9999), 0.0001)$ before computing the transformation to avoid undefined values at exactly 0% or 100%. This has negligible impact on the results since observed literacy rates in the data are well within these bounds. The arcsine transformation is defined on [0,1] and requires no boundary adjustment.

[16] To interpret the log-odds gap: if Ottoman-side literacy is approximately 90% (log-odds $= \ln(0.9/0.1) = 2.20$), adding the estimated gap of 1.59 yields Habsburg-side log-odds of 3.79, corresponding to 97.8% literacy, a difference of 7.8 percentage points. Romania's residual log-odds gap of 0.11 corresponds to approximately 1 percentage point at the same baseline.

Table A.1. *Convergence in Ottoman–Habsburg Literacy Gaps across Alternative Transformations*

Country	Percentage Points		% of 1931 Gap Closed by 1991			
	1931 Gap	1991 Gap	PP	Log-Odds (logit)	Probit (Φ^{-1})	Arcsine ($\arcsin\sqrt{p}$)
Romania	14.3***	1.7***	87.9%	82.7%	83.9%	85.5%
	(2.97)	(0.39)				
Yugoslavia	43.5***	7.0***	83.9%	26.9%	45.4%	67.3%
	(1.48)	(0.64)				
Difference			4.0pp	55.8pp	38.5pp	18.2pp

Notes: Standard errors in parentheses, clustered at county level (Romania) and *srez* level (Yugoslavia 1991). The "Gap" columns show the estimated Ottoman disadvantage in percentage points (positive values indicate lower literacy on the Ottoman side). "% of 1931 Gap Closed" shows the proportion of the initial gap eliminated by 1991 under each transformation: percentage points, log-odds (logit: $\ln(p/(1-p))$), probit ($\Phi^{-1}(p)$), and arcsine ($\arcsin\sqrt{p}$). All three bounded-outcome transformations account for mechanical compression of percentage-point differences near ceiling. *** $p < 0.01$, ** $p < 0.05$.

segment-specific RD2D approach, which restricts estimation to segments where local balance holds. Full estimation details (bandwidth diagnostics, bias-correction procedures) are available in the replication archive.

Table A.2. *Quantitative Indicators of Interwar Yugoslav Exclusion of Muslims*

Indicator	Muslims	Total	Share
Panel A. Underrepresentation in state and economic institutions			
Gymnasium students, Yugoslavia, 1939[a]	2,663	110,000	2.4%
High-ranking civil servants, Yugoslavia, 1939[a]	108	6,163	1.8%
Industrial enterprises, Yugoslavia, 1939[a]	36	3,054	1.2%
Senior officials in Bosnia, 18 August 1919[b]	17	273	6.2%
Panel B. Dismantlement of Muslim institutions			
Mektebs (Qur'anic schools)[c]	≈1,200 (1918) → 700 (1928)		
Medreses (religious secondary schools)[c]	43 (1918) → 12 (1938)		
Sharia-court judge positions[d]	85 → 60 (13 left unfilled)		
Landowners with *kmet* tenants[e]	91.2% of all such landowners in BiH expropriated 1919		
Panel C. Revealed-preference response: religious bloc voting			
JMO share of Muslim vote, 1920–1928 elections[b]	98–100%		

Notes: Population share of Muslims in Yugoslavia was approximately 11% during this period. Panel A: representation in state-administered institutions; Panel B: contraction of Muslim educational, legal, and economic institutions under successor-state policy; Panel C: revealed-preference indicator of perceived exclusion. None of these disparities can be reduced to Ottoman-era urbanization, *vakıf* infrastructure, or inherited land tenure; they were produced by interwar state policies. Source notes:

[a] Survey by Muslim cultural society *Gajret* (1939), cited in Bougarel (2018) and Kemura (1986).

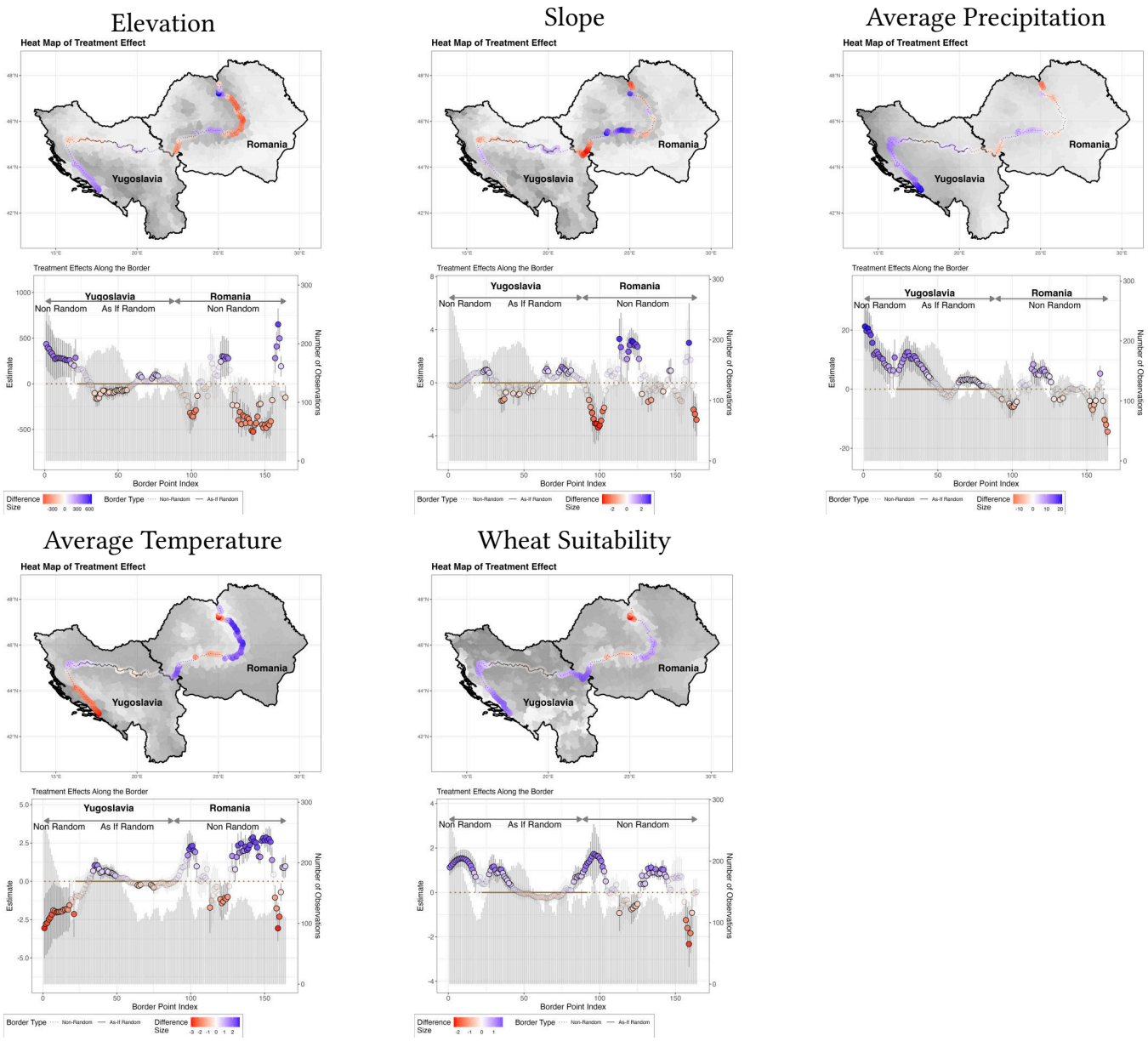
[b] Purivatra (1974, p. 44) for senior officials; JMO vote share elsewhere in Purivatra (1974).

[c] Bougarel (2018, p. 43).

[d] Karčić (2005). Requests from Muslim-majority districts for court establishment were denied on budget grounds; in some areas, courts mandated by the 1921 Vidovdan Constitution were not established until 1940, a nineteen-year lag.

[e] Dyker (1972); Purivatra (1974). The 1919 agrarian reform expropriated mostly Muslim landholders and transferred land to predominantly Orthodox tenant farmers with compensation widely regarded as inadequate (Bougarel, 2018, p. 34).

Figure A.1. *Border Segment Covariate Balance*



E Tests for Manipulation around the Threshold

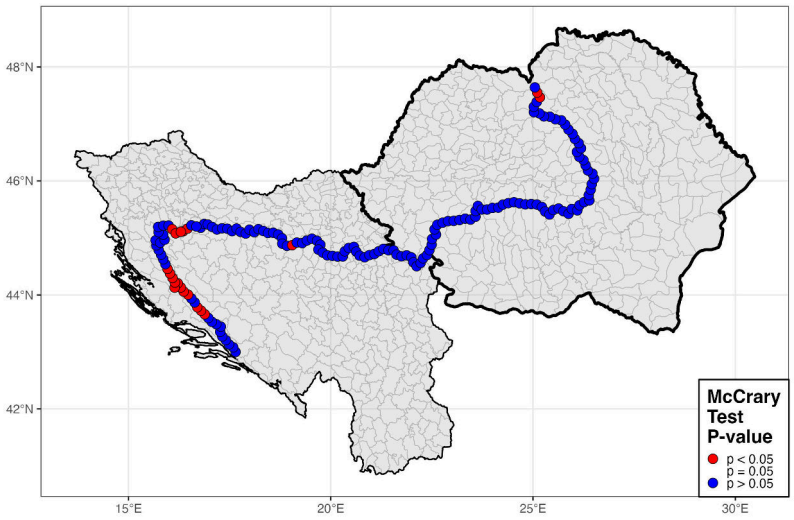
An additional identifying assumption is the absence of selective sorting across the treatment threshold. This assumption would be violated if people with different characteristics systematically moved from the Ottoman to the Habsburg side, thereby indirectly affecting literacy outcomes.

To test for such sorting, I employ the McCrary density test (McCrary, 2008), which detects discontinuities in the distribution of the running variable, in this case distance to the border. A discontinuity would indicate potential manipulation, such as a higher population density on the Habsburg side due to migration.

While the traditional McCrary test is typically estimated once over the full sample, this approach may miss localized sorting along particular border segments. To account for this, I compute the test at 164 points along the border using the `rdplotdensity` function of Cattaneo *et al.* (2018).

Figure A.2 reports the resulting p-values, revealing significant discontinuities in the western portion of the border, particularly near the Dalmatian coast, the Una river, and the Croatia–Serbia boundary, and at two points in northern Romania. Consequently, discontinuity estimates at those border points should be interpreted with caution.

Figure A.2. *Localized McCrary Density-Test P-values along the Frontier*



F Multi-Score RD Design: Over-Time Estimates and Distance Comparison

RD2D estimates at each census wave show the spatial evolution of the Ottoman effect from 1931 to 1991 (full maps in the replication archive). Comparing the bivariate location-score approach (Cattaneo *et al.*, 2025) with distance-based alternatives reveals substantial divergence along land sections of the border, consistent with the bias identified in Keele *et al.* (2015). Under the bivariate method, 111 of 164 evaluation points show statistically significant effects. Figure A.3 reports the comparison; over-time spatial maps for each census wave are in the replication archive.

G Traditional GRD Benchmark

For comparability with prior distance-to-border designs, I estimate a standard geographic RD (GRD) using Euclidean distance as the running variable, with parametric specifications across bandwidths from 10 to 150 km and non-parametric `rdrobust` estimates (Calonico *et al.*, 2014). Standard errors are clustered at the county (*județ*) level. I first verify covariate balance near the cutoff (Figure A.4). For elevation, slope, annual temperature, annual precipitation, and wheat suitability, the estimated treated–control difference is statistically indistinguishable from zero within narrow bandwidths (below roughly 65 km) and only diverges meaningfully beyond 80 km, particularly for elevation, where treated districts become significantly higher (50–67 m) at bandwidths of 80 km and above. The 1931 treatment effect on literacy is approximately -25 to -29 percentage points across parametric specifications (about -25 under the non-parametric variants) and stable across bandwidths from roughly 30 km to 150 km (Figure A.5), consistent with the RD2D results.

Figure A.3. *Bivariate Location vs. Distance-Based RD Estimates*

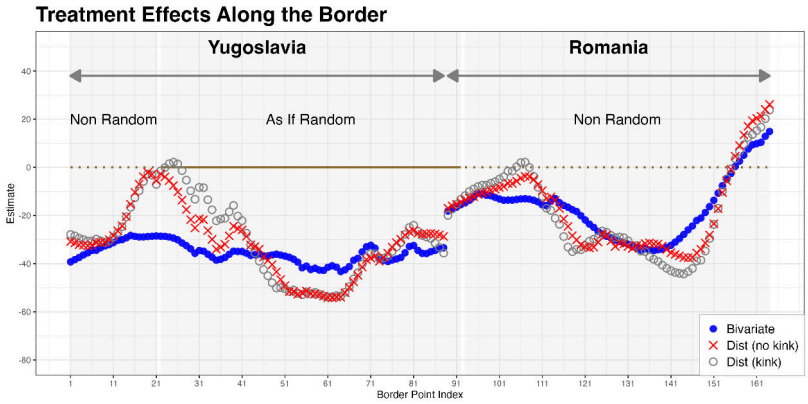
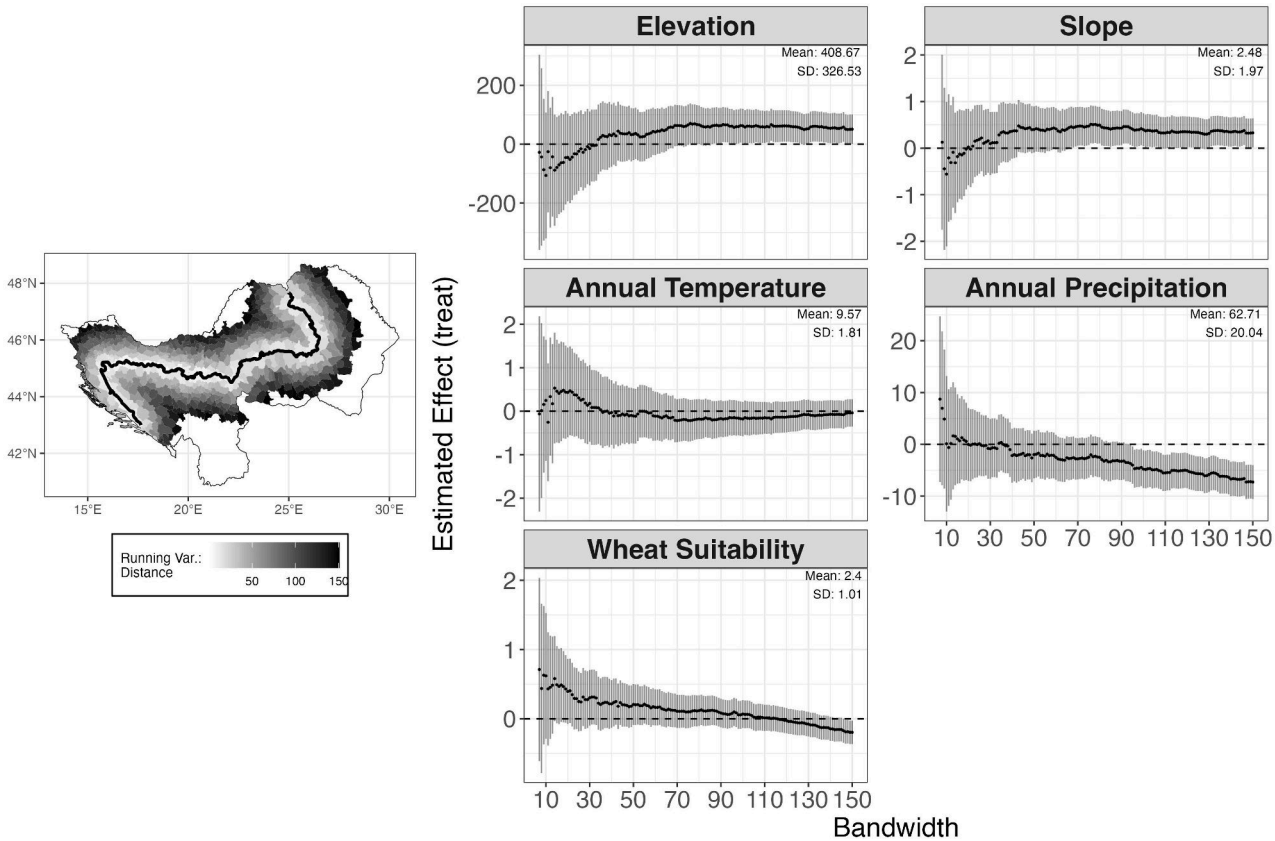
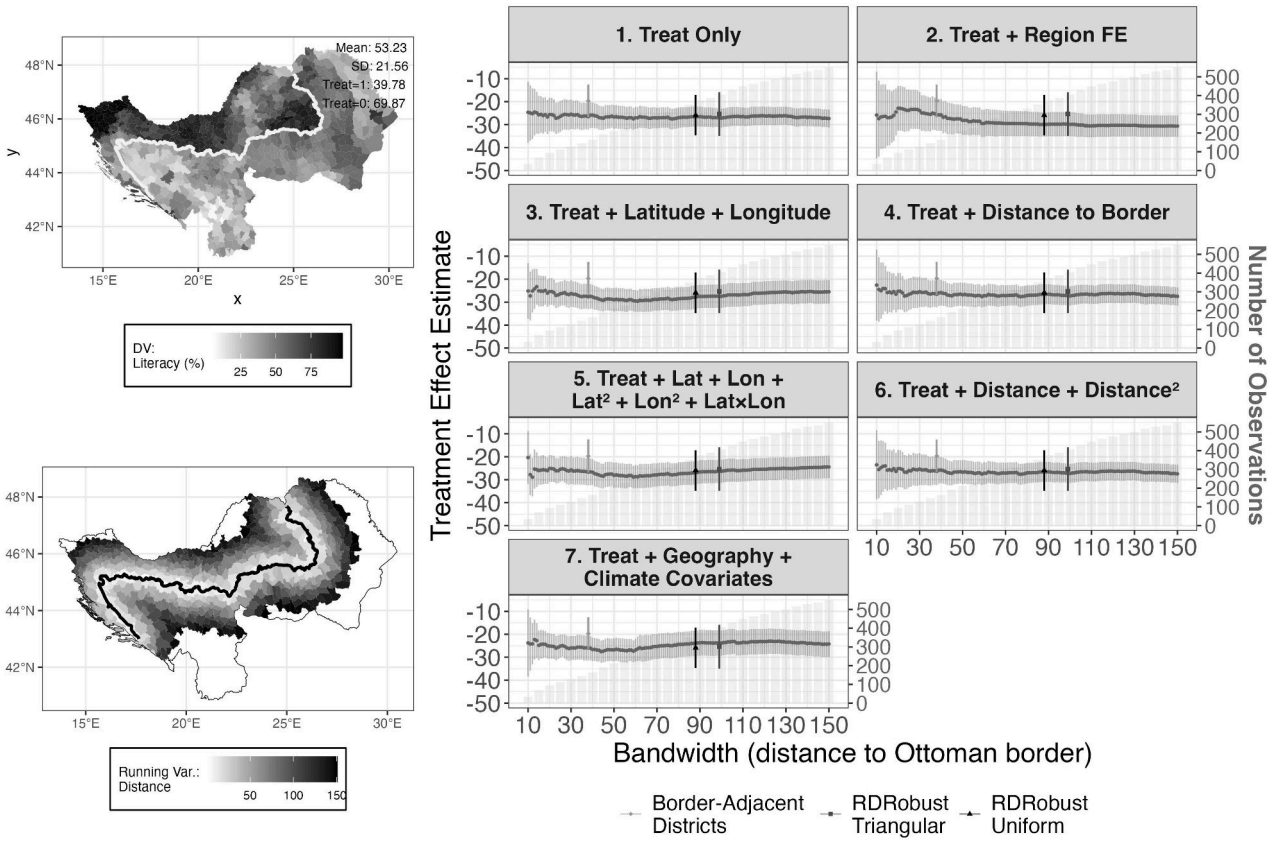


Figure A.4. *Traditional GRD: Covariate Balance by Bandwidth*



Notes: Estimated treated–control difference in each geographic/climate covariate (elevation in meters, slope in degrees, annual temperature in °C, annual precipitation in mm, wheat suitability index) across bandwidths from 10 to 150 km. Solid line: point estimate. Shaded band: 95% confidence interval. Annotations report the sample mean and SD for each covariate. Map (left): the running variable, distance to the Ottoman–Habsburg border (km), capped at 150 km. Standard errors clustered at the county level.

Figure A.5. *Traditional GRD: Bandwidth and Specification Sensitivity, 1931 (Yugoslavia + Romania)*



Notes: Treatment-effect estimates across bandwidths (10–150 km) for seven parametric specifications, non-parametric `rdrobust` *with triangular and uniform kernels, and a border-adjacent linear comparison. Standard errors clustered at the county level. Pooled Yugoslavia + Romania, 1931 census literacy.*

H Migration

Selective migration between 1931 and 1961 could confound the results if populations with different literacy profiles moved differentially across the Ottoman–Habsburg boundary. Yugoslavia experienced substantial minority outmigration: near-total German expulsion after WWII, Italian emigration from Istria (∼100,000), Turkish outmigration from Macedonia, and Albanian displacement from Kosovo (Shoup, 1963; Ballinger, 2003; Douglas, 2012). Because 1931 *srez* and 1961 *opština* boundaries differ, I construct a harmonized 0.25° population grid (approximately 20×28 km at this latitude) using area-weighted interpolation (methodology and intermediate maps in the replication archive).

Figure A.6 maps population changes. I then re-estimate the main model under four migration-robust specifications: (1) excluding grid cells where population declined; (2) controlling for absolute population change; (3) controlling for percentage change; and (4) controlling for religious composition change.

Figure A.6. *Population Changes 1931–1961*

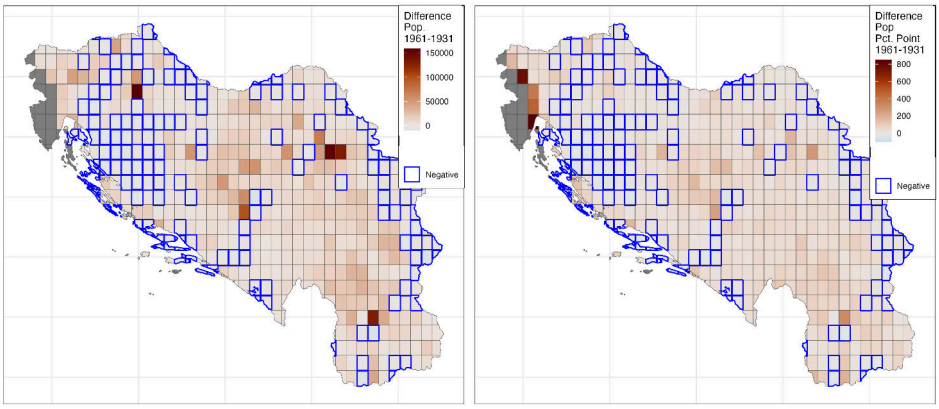


Figure A.7 compares the first two specifications to the baseline; the remaining two behave the same way (replication archive). The Ottoman effect is essentially unchanged, indicating that the documented migration episodes do not drive the findings.

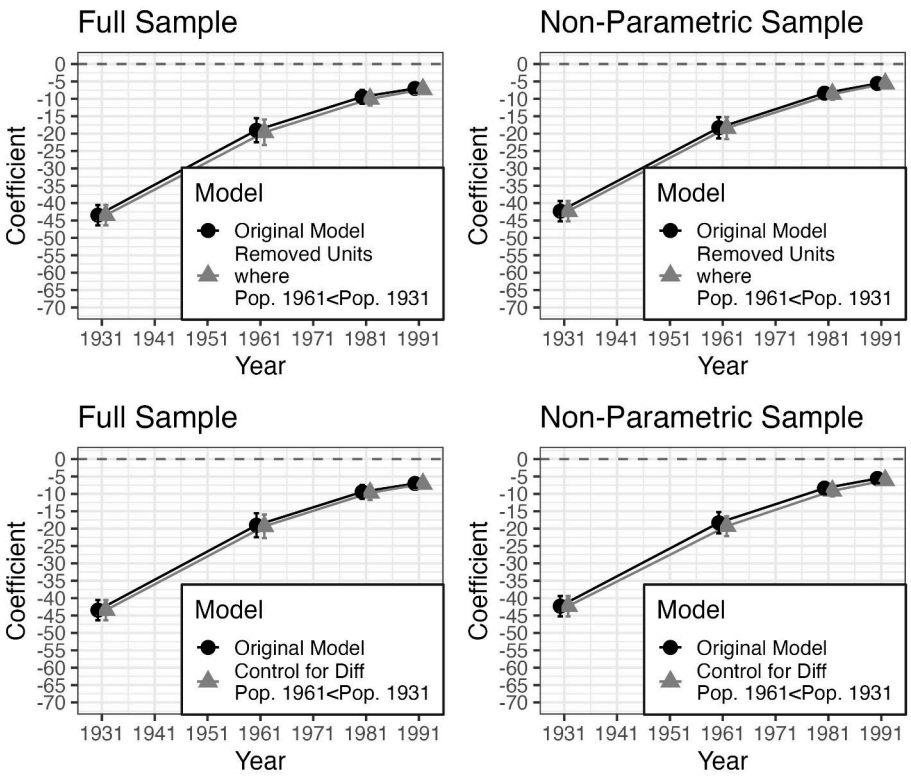
Migration in Romania

Romania did not experience comparable mass population movements. I assess pre-treatment mobility using the 1930 census *populația flotantă* (Institutul Central de Statistică, 1930). Regressing out-migration on Ottoman treatment yields a null coefficient (Table A.3), consistent with no differential pre-1931 mobility.

Table A.3. *Ottoman Treatment and Out-Migration, Romania (1930 Census)*

	Dependent variable:
	Pct. Outmigrants
Ottoman	0.027
	(0.048)
Observations	322
R^2	0.001
Adjusted R^2	−0.002
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure A.7. *Controlling for Migration*



I Federalism

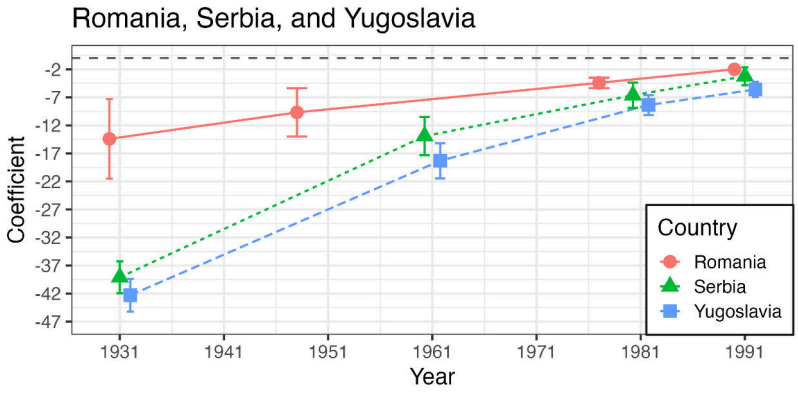
A prominent alternative explanation is that Yugoslavia's federal structure weakened the center's capacity to equalize human capital across regions. On this view, decentralization constrained the implementation of uniform education policy, producing slower convergence in Yugoslavia than in unitary Romania. If federalism were the primary mechanism, convergence should be systematically slower across Yugoslav republics relative to Romania, holding constant exposure to the Ottoman legacy. This prediction is not borne out by the data.

Timing cannot adjudicate the federalism claim: convergence slowed after the 1963 reforms expanded republican autonomy, but it slowed on the same schedule in Romania without any decentralization, because gains shrink mechanically as gaps close. The informative test is therefore cross-sectional, comparing groups held under the same constitutional arrangements in every period.

A within-Yugoslavia comparison provides that diagnostic. I examine Serbia alone using the same non-parametric frontier-based sample employed in the main analysis. Serbia was subject to the same federal framework and the same constitutional schedule of decentralization as Bosnia and Macedonia, and it operated under the same socialist education policies and national political economy as other Yugoslav republics. If federalism were the central impediment, Serbia should display convergence patterns similar to Yugoslavia as a whole. Instead, Serbia converges nearly as fully as unitary Romania. This test exploits the fact that federalism is constant across Serbian, Bosnian, and Macedonian districts, whereas ethnoreligious composition, and hence state–minority incorporation, varies substantially. I thus estimate the model $Outcome_v = \alpha + \beta\, Ottoman_v + \varepsilon_v$ by OLS on each country's non-parametric frontier sample (districts within the RD2D bandwidth), with standard errors clustered as in Section 4. Figure A.8 plots the point estimates and 95% confidence intervals over time.

The results weigh against a federalism-based account. In Romania, the estimated Ottoman disadvantage declines from -14.4 percentage points in 1931 to -2.0 points in 1991. In Serbia, which operated within the Yugoslav federation, the Ottoman disadvantage falls from -39.1 points in 1931 to -3.24 points in 1991. By contrast, for Yugoslavia as a whole the gap remains notably larger in 1991 (-5.59 points) and follows a slower convergence trajectory across the socialist period.

Figure A.8. *Federalism Test*

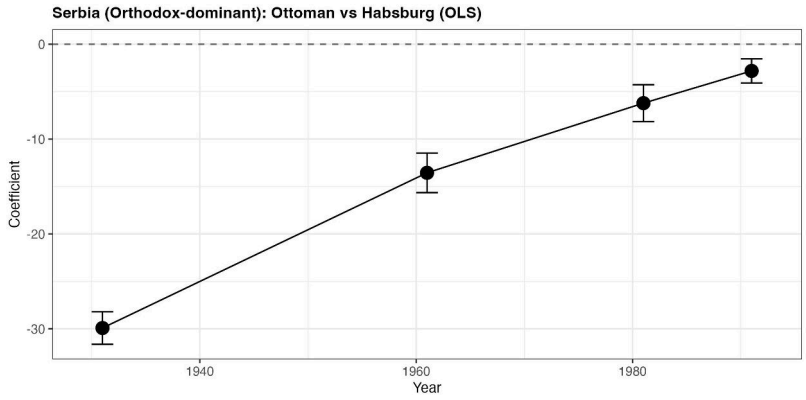


J Differential Levels of Repression

An alternative explanation is that Ceaușescu's more coercive regime, not incorporation, explains Romania's faster convergence. Yugoslavia's repression targeted Muslim religious institutions rather than school attendance (756 mosques destroyed or damaged postwar, *vakıf* nationalized, the veil banned in 1950, mosque-based teaching criminalized (Alibašić & Zubčević, 2009; Ballinger & Ghodsee, 2011; Mekić, 2017)): coercion of a kind the theory expects to deepen exclusion, not substitute for it.

Restricting to homogeneous Serbian districts (<1% Muslim, >85% Orthodox), which faced the same Yugoslav regime but high incorporation, shows rapid convergence (Figure A.9): Romania-level coercion is unnecessary where frictions are low.

Figure A.9. *The Ottoman Literacy Gap in Homogeneous Serbian Districts*



K Muslim Share and Poverty: A Descriptive Check

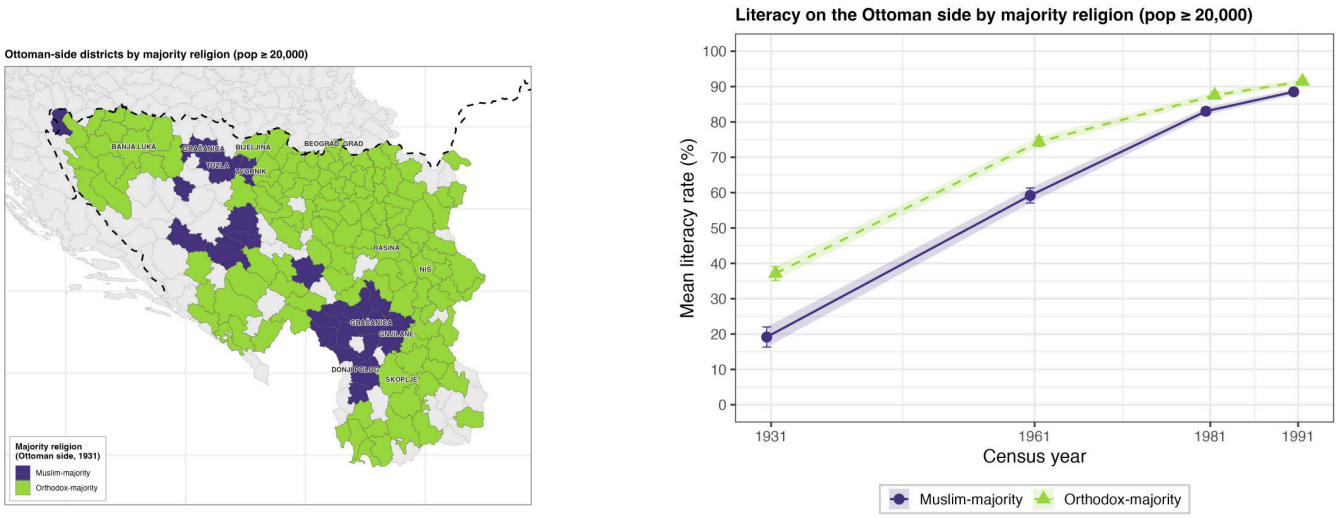
This section provides the descriptive evidence referenced in Section 5.2 to rule out the possibility that Muslim population share proxies for poverty or rural isolation. The test restricts the sample to Ottoman-side districts with populations above 20,000 and classifies each as Muslim-majority (>50% Muslim) or Orthodox-majority (>50% Orthodox), discarding mixed districts. This yields 24 Muslim-majority and 129 Orthodox-majority districts in 1931, changing to 61 and 128 respectively by 1981 as administrative boundaries were redrawn into smaller units.

Figure A.10 presents the results. The left panel maps the included districts by majority religion in 1931; the right panel plots mean literacy rates with 95% confidence intervals across the four Yugoslav censuses (1931, 1961, 1981, 1991). Muslim-majority towns start at 19.2% literacy versus 37.1% for Orthodox-majority towns, an 18 percentage-point gap among comparably sized places on the same side of the frontier. Both groups converge under communist schooling, but the gap never fully closes: 15.2 pp in 1961, 4.5 pp in 1981, and 3 pp in 1991. Since these are all districts above the same population threshold, located on the same side of the imperial frontier, the persistent gap cannot reflect district size or position relative to the frontier. Income differences among comparably sized districts remain possible; but a poverty account must then explain why the gap survives elevation and density controls (Table A.4), why the state's supply of schools and teachers overshot in Muslim-share districts by 1981 while households' voluntary uptake of state media, radio and television subscription, lagged (Section 5.3; Appendix Figure A.17), and, most sharply, why women's illiteracy jumps discontinuously at the 5,000-Muslim threshold of Section 6, where poverty varies smoothly. The pattern is instead consistent with the incorporation mechanism: among Ottoman-heritage towns of comparable size, those where the majority population experienced status reversal under the successor state lagged behind those where it did not.

L Gender-Disaggregated Muslim-Share Regressions: Robustness

Table A.4 reports the Muslim-share coefficient on male and female literacy across three specifications (baseline, geographic controls, region FE) within the RD2D bandwidth. The gender asymmetry (the female coefficient between roughly two and seven times the male in absolute value in most specifications) is present across every specification and census year from 1961 onward. Because men and women within the same district face identical infrastructure, any confounder that affects the sexes equally cannot explain this asymmetry. To survive, a confounder would have to change how the same family treats sons versus daughters, which narrows the alternatives to the household margin the theory is about.

Figure A.10. *Literacy on the Ottoman Side by Majority Religion (pop ≥ 20,000)*



Notes: Sample restricted to Ottoman-side districts (treat = 1) with total population ≥ 20,000. Districts classified as Muslim-majority if >50% Muslim, Orthodox-majority if >50% Orthodox. Left: geographic distribution. Right: mean literacy rate with 95% confidence intervals. Dashed line shows the Ottoman–Habsburg frontier.

Table A.4. *Effect of Muslim Population Share on Literacy by Gender: Robustness*

	1931		1961		1981		1991	
	Men	Women	Men	Women	Men	Women	Men	Women
Baseline	−0.916***	−0.826***	−0.226***	−0.579***	−0.043***	−0.197***	−0.027***	−0.157***
	(0.065)	(0.080)	(0.019)	(0.063)	(0.008)	(0.039)	(0.007)	(0.029)
+ Elev., pop. dens.	−0.612***	−0.192	−0.128***	−0.306***	−0.011	−0.081	−0.012	−0.078**
	(0.066)	(0.104)	(0.032)	(0.081)	(0.013)	(0.048)	(0.009)	(0.034)
+ Region FE	−0.518***	0.041	−0.112***	−0.193***	−0.005	−0.032	−0.008	−0.048
	(0.058)	(0.050)	(0.031)	(0.044)	(0.013)	(0.034)	(0.008)	(0.027)
N	193	193	357	357	243	243	243	243

Notes: Each cell reports the OLS coefficient on Muslim population share (%) from a regression of male or female literacy rate on Muslim share, with standard errors clustered by district (the *srez* from 1961 onward; in 1931 districts are the units themselves, so errors are district-robust) in parentheses. Geographic controls are elevation and population density. Region fixed effects absorb between-region (Ottoman/Habsburg) variation. *** $p < 0.01$, ** $p < 0.05$.

M Mechanism Diagnostics: Gender Residual, Adoption, and Pre-Communist Baseline

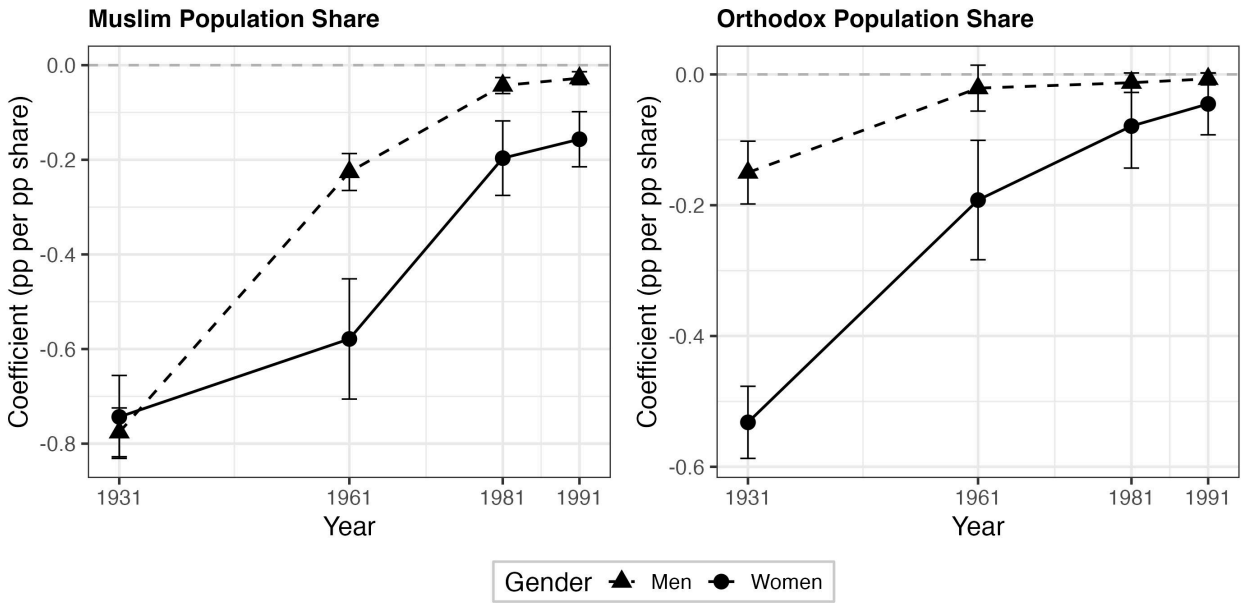
State provision vs. household adoption (Figure A.17). Figure A.17 contrasts the Muslim-share gradient in state-controlled educational inputs (Panel A: schools and teachers per 1,000 population) with household-level adoption of state media (Panel B: radio and television subscribers per 1,000 population), and reports an infrastructure-balance check using 1970 railway density (Panel C). The supply panels span 1961, 1981, and 1991; the adoption panels begin in 1981 because the 1961 Yugoslav census did not enumerate radio or television subscribers. The schools coefficient is statistically indistinguishable from zero in every census year. Teachers per capita were significantly lower in Muslim-share districts in 1961 (−0.031 per point of Muslim share); by 1981 the coefficient had flipped positive (+0.009, $p < 0.01$) and remained positive in 1991 (+0.006): the state did not merely equalize what it directly controlled, it overcompensated. Household adoption of state media, which requires voluntary engagement rather than coerced attendance, points the other way: radio subscription remained significantly lower in districts with larger Muslim shares in both 1981 and 1991, persisting after elevation and railway controls (red series),

while the television coefficients are negative but mostly insignificant once controls enter. The radio pattern is consistent with demand-side frictions rooted in exclusion. Panel C shows that 1970 railway density is uncorrelated with Muslim share, so differential rail access does not explain the adoption gap; household income differences remain an alternative demand-side channel, which the threshold evidence of Section 6 addresses.

Pre-communist secondary schooling. I regress district-level secondary schooling on Muslim population share in 1932, fourteen years after the creation of the Kingdom of Serbs, Croats, and Slovenes and before communist centralization. The Muslim-share coefficient is -0.00049 ($p < 0.05$) on secondary schools per 1,000 population and -0.131 ($p < 0.05$) on secondary students per 1,000 population, indicating that the educational gradient was already operative in the interwar period.

Within-Yugoslavia identification: continuous religion-share gradient (Figure A.11). Figure 5 in the main text uses majority-religion subsamples (Muslim-majority vs. Orthodox-majority Ottoman-side districts), which is intuitive but discards within-district intensity. The figure here reports the same identification using the continuous Muslim and Orthodox population shares: each point is the OLS coefficient on Muslim share (left panel) or Orthodox share (right panel) from a regression of male or female literacy on religious share, by census year, within the RD2D bandwidth. Standard errors are clustered at the district level (the *srez* from 1961 onward; in 1931 districts are the units themselves, so errors are district-robust). The 1991 Muslim-share gradient on women's literacy is roughly 3× the corresponding Orthodox-share gradient, sharper than the difference visible in the majority-cut version of Figure 5; continuous shares pick up intensity that a >50% threshold attenuates.

Figure A.11. *Effect of Religious Composition on Literacy by Gender, Yugoslavia 1931–1991 (Continuous Share Gradient)*



Notes: Each point is the OLS coefficient on Muslim (left) or Orthodox (right) population share from a regression of male or female literacy on religious share, with 95% CIs clustered at the district level (the srez from 1961 onward; in 1931 districts are the units themselves, so errors are district-robust). Sample restricted to the RD2D bandwidth. Dashed lines with triangles: men. Solid lines with circles: women. This is the continuous-share counterpart to the majority-cut version of Figure 5 in the main text.

N Sharia-Court RD: Design Validation and Robustness

This appendix validates the design of the 1929 sharia-court RD in Section 6: covariate balance at the cutoff, a density test for manipulation, and robustness to alternative standard errors. Two notes on data construction. First, each 1961 settlement is located by matching its name, within its opština, to the 1885 Habsburg settlement list of Neubert & Nikolić (2026); the match is one-to-one, so no 1885 settlement is assigned to more than one 1961 settlement, and the procedure is documented in the replication archive. Second, four 1921 census units were not districts of their own and are counted inside their parents: Doboj, a sub-office (*kotarska ispostava*) of the Tešanj district, and the towns of Velika Kladuša, Kozarac, and Bosanska Kostajnica, which belong to Cazin, Prijedor, and Bosanski Novi; their populations are included in the parent district's Muslim count.

S.1. Covariate Balance at the Cutoff

Table A.5 reports balance tests on twelve pre-treatment covariates. Each row treats the indicated covariate as the outcome of a local-linear RD with triangular kernel and MSE-optimal bandwidth (opština-clustered standard errors for the five settlement-level rows; each base district as its own cluster for the seven srez-level rows). Inference is the Conventional local-linear estimator.

Table A.5. *Sharia-Court RD: Covariate Balance at the 5,000-Muslim Cutoff*

Covariate	τ	SE	p	N	Mean (SD)
Elevation (m)	153.95^{***}	(17.12)	< 0.001	490	541.09 (304.23)
% Orthodox (1910)	-0.30^{***}	(0.07)	< 0.001	614	0.44 (0.43)
% Catholic (1910)	0.14^{***}	(0.01)	< 0.001	657	0.20 (0.34)
Ethnic fractionalization (1910)	-0.12^{***}	(0.02)	< 0.001	333	0.17 (0.20)
Road in 1885	0.06	(0.03)	0.067	874	0.19 (0.39)
Islamic schools in srez (1895)	−3.17	(3.35)	0.343	20	12.44 (12.55)
Orthodox schools in srez (1895)	−0.61	(0.80)	0.440	21	1.16 (1.31)
Catholic schools in srez (1895)	0.20	(0.41)	0.626	15	0.32 (0.77)
% illiterate (1931, total)	3.02	(6.76)	0.655	23	73.08 (8.38)
% illiterate (1931, men)	12.65	(14.06)	0.369	27	58.82 (12.59)
% illiterate (1931, women)	2.65	(6.35)	0.676	17	86.62 (7.32)
WWII victims, % of 1931 pop. (post-treatment)	7.18	(5.75)	0.211	14	9.62 (8.51)

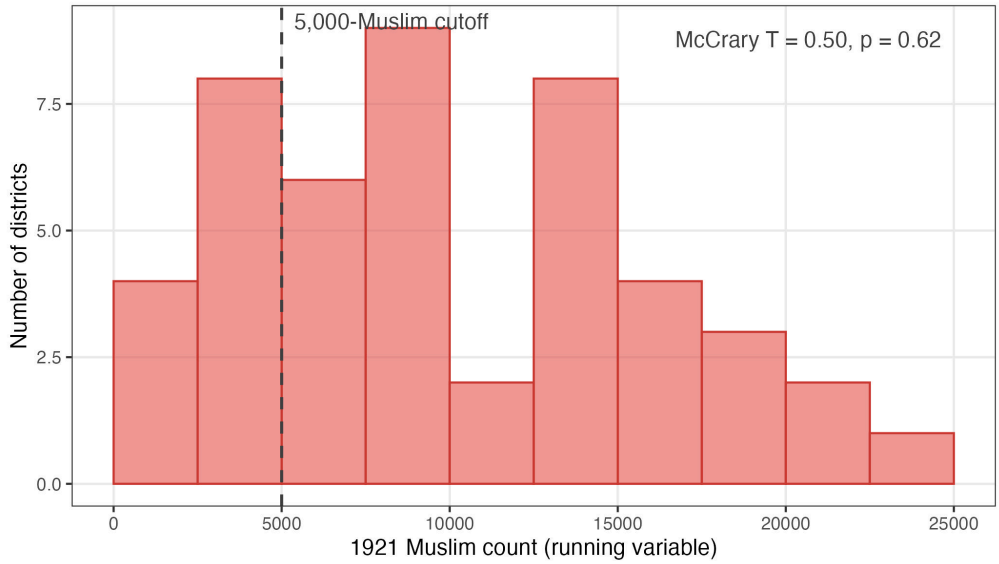
Notes: Each row is a separate local-linear RD (triangular kernel, MSE-optimal bandwidth) with the indicated covariate as the outcome; a significant τ signals imbalance. τ = (lost − retained), matching Table 3. The first five covariates are measured at the naselje (settlement) level and tested with opština-clustered standard errors; the remaining seven are measured at the district (srez) level and tested on the 50-district panel, each district its own cluster. The 1895 confessional-school counts predate the 1929 closure; the 1931 census illiteracy, taken two years after it, reflects schooling completed almost entirely before it; WWII victims (as % of 1931 population) is a post-treatment check that the 1961 gap does not reflect differential wartime losses. The five imbalanced naselje-level covariates (elevation, 1910 Orthodox and Catholic shares, 1910 fractionalization, 1885 road) enter the headline RD as linear controls, so its coefficients are conditional on them; the seven srez-level covariates are balanced. N is the effective sample inside the bandwidth. $^{***}\ p < 0.01$, $^{**}\ p < 0.05$.

Five of the twelve covariates show imbalance, all at the settlement level: elevation, the 1910 Orthodox and Catholic shares, and 1910 ethnic fractionalization at the 5% level, and the 1885 road indicator just outside it ($p = 0.067$), all measured in the Neubert & Nikolić (2026) cross-section. They reflect the 5,000-Muslim cutoff sorting districts along a state-penetration gradient as well as a Muslim-population one, and they enter the headline specification as linear controls (Table 3), so the reported coefficients are conditional on them. The seven srez-level covariates are balanced: the 1895 confessional-school counts (predating the closure), the 1931 census illiteracy (reflecting schooling completed almost entirely before the closure), and the WWII-victim share (a post-treatment check) do not jump at the cutoff, so neither pre-closure schooling infrastructure nor differential wartime losses can account for the 1961 gap.

S.2. McCrary Density Test at the Assignment Unit

Manipulation of the running variable would mean districts exercised discretion over their 1921 Muslim count. Timing alone makes this implausible: the count was fixed eight years before the 1929 law was drafted, and the next census came only in 1931, two years after the closure. The formal test agrees. At the assignment unit, the 50 base districts, the Cattaneo *et al.* (2018) density test finds no discontinuity at the cutoff ($T = 0.50$, $p = 0.62$; Figure A.12).

Figure A.12. *Distribution of the Running Variable across Districts*



Notes: Distribution of the 1921 Muslim count across the 50 analysis districts (capped at 25,000 for legibility). The dashed line marks the 5,000-Muslim cutoff. The annotation reports the McCrary density discontinuity test (Cattaneo–Jansson–Ma); a small p would indicate sorting across the threshold.

S.3. Robustness to Alternative Standard Errors

Table A.6 re-evaluates the T2 cell under three standard-error specifications: HC (no clustering), opština-clustered (the headline), and a wild-cluster bootstrap at the 1921 district, the assignment unit. The point estimates and bandwidths are those of Table 3; only the inference changes. The women's effect is significant under opština clustering and just short of the 0.05 threshold under HC ($p = 0.054$). The district bootstrap does not reject; with only twelve treated clusters, however, the null-imposed wild bootstrap is known to under-reject severely (MacKinnon & Webb, 2018), so a non-rejection at this design size is expected whether or not the effect is real and cannot adjudicate it. Randomization inference at the assignment level points the same way: collapsing T2 settlements to district means and permuting lost-court status within a ±3,000 window, the female difference is positive in every window but far from significance (Table A.6). At this design size no assignment-level test can adjudicate; both are reported for transparency, and the small number of treated districts is the design limit flagged before the estimates in Section 6.2.

S.4. Functional Form, Kernel, and Bandwidth

Table A.7 re-estimates the T2 cell under a local quadratic in place of the linear fit and under uniform and Epanechnikov kernels in place of the triangular default; its first panel reproduces the Table 3 headline. The women's estimate keeps its positive sign in every specification: 3.92 under the quadratic ($p = 0.08$, against the 5.46 linear headline), 6.08 under the Epanechnikov kernel (significant), and 1.42 under the uniform kernel ($p = 0.09$), which selects a much narrower window. The population and men's estimates are less steady, both turning slightly negative under the uniform kernel's narrow window. Figure A.13 varies the window directly, re-fitting each outcome at half to double its MSE-optimal bandwidth (the 1.00× point reproduces the headline). The women's estimate is positive across the entire range, significant from the optimal window outward, and weakens to non-significance only at the narrowest windows (1.02 at half the optimal); population illiteracy moves alongside it, while the men's estimate reverses sign at narrow windows, the few-mass-points instability visible elsewhere in this appendix. The pattern across all of these checks is consistent: the women's outcome is the one the design supports.

Table A.6. *Sharia-Court RD: Robustness to Alternative Standard Errors (T2 cell)*

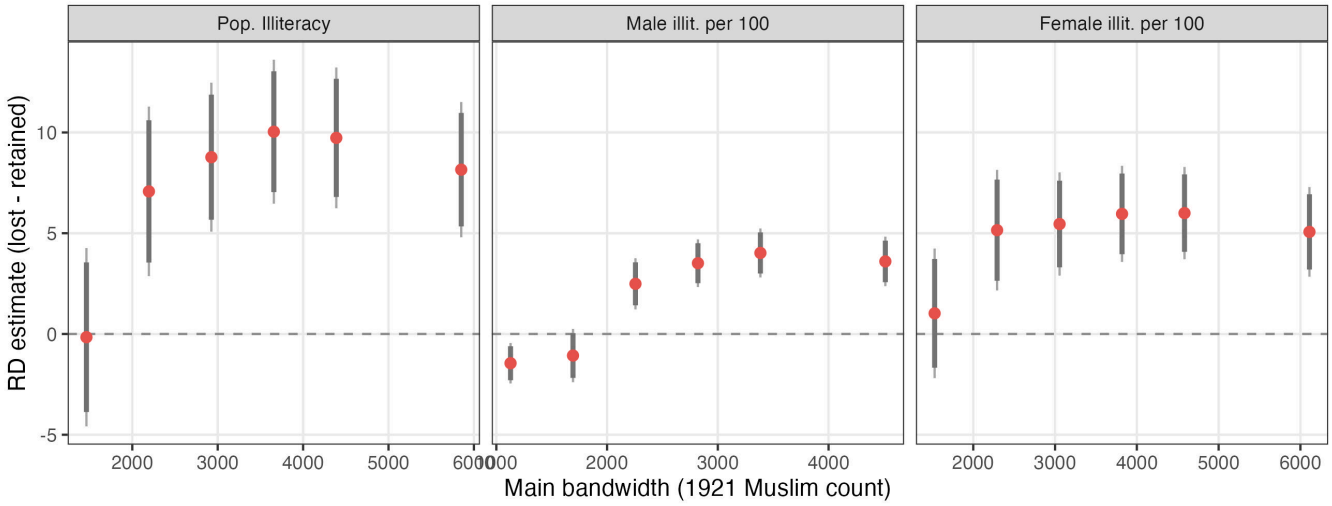
	Illit. per 100	Male illit. per 100	Female illit. per 100
T2 (moderate-Muslim tercile)			
τ	8.77	2.49	5.46
HC (no clustering) SE	(3.69)	(1.50)	(2.59)
opština-clustered SE	(2.42)***	(0.79)	(1.63)***
p (wild bootstrap, district)	0.421	0.565	0.406
District-level randomization inference (±3,000 window)			
District-mean difference	2.05	0.45	1.64
p (randomization inference)	0.552	0.752	0.459
Bandwidth	2926	2256	3055
N	264	201	266

Notes: Point estimates and bandwidths are those of the T2 cell in Table 3; only the inference changes. Three specifications: HC (no clustering, parenthesized), opština-clustered (the headline, parenthesized), and a wild-cluster bootstrap at the 1921 district (9,999 Rademacher draws), the small-cluster method appropriate to the assignment unit. Treatment is assigned at the district; with few treated clusters the null-imposed wild bootstrap under-rejects severely (MacKinnon and Webb 2018), so its p-values are conservative. The randomization-inference panel collapses T2 settlements to district means and permutes lost-court status among the 17 districts within ±3,000 of the cutoff (10 treated; 20,000 draws; two-sided p); narrower windows and a covariate-residualized female outcome (difference 2.48, $p = 0.19$) give the same picture. Stars on each row reflect that row's own inference. *** $p < 0.01$, ** $p < 0.05$.

The last two panels of Table A.7 ask whether the tercile split itself creates the result; they show it does not. Pooled across every settlement with any Muslim presence, without the split, the female effect is 5.25 ($p < 0.01$). Allowing the effect to vary smoothly with the 1910 Muslim share instead of by bins barely moves it: between 4.8 and 5.5 points from low shares to high, with no significant trend. What the data cannot say is whether the effect weakens at the very highest shares: the smooth fit gives 5.5 there while the imprecise T3 cell gives 3.1, and the two intervals are too wide to separate. Either way, the conclusion the design needs survives: the effect requires Muslim presence and does not depend on how the Muslim-present settlements are divided. Because the interaction model estimates more parameters from the same settlements, its individual estimates are less precise than the pooled one; the pooled male coefficient (-0.27) is a trivially small negative of the kind discussed with Table 3.

A scan over placebo cutoffs is uninformative in this setting. The running variable is the 1921 district Muslim count, which takes fifty distinct values, only twelve of them below the cutoff and seven within a thousand of it. Re-estimating the RD at arbitrary fake cutoffs fits the slope off a handful of district mass points, and the fit is singular across much of the upper range; such a scan reflects the coarseness of the running variable rather than the validity of the threshold, and is suited to the finely spaced running variables for which it was developed (Blanc & Kubo, 2026).

Figure A.13. *Bandwidth Sensitivity for the Sharia-Court RD, T2 Cell*



Notes: RD estimate (lost − retained) for the three substantive outcomes at multiples of each outcome's MSE-optimal main and bias bandwidths, from half to double; the estimate at the optimal bandwidth reproduces the Table 3 headline. Points are the local-linear estimate; thin and thick grey bars are the conventional 95% and 90% confidence intervals centered on the point (the robust bias-corrected interval recenters on the bias-corrected estimate, so significance in the tables uses the robust p). Same controlled specification as Table 3, opština-clustered.

Table A.7. *Sharia-Court RD: Functional-Form, Kernel, and Muslim-Share Robustness*

	Pop. Illiteracy	Male illit. per 100	Female illit. per 100
Triangular, local linear (headline)			
Estimate	8.77^{***}	2.49	5.46^{***}
95% CI	[2.0, 11.4]	[−1.9, 1.2]	[1.8, 8.2]
Bandwidth	2926	2256	3055
Effective N	264	201	266
Local quadratic (p = 2)			
Estimate	4.05	1.24	3.92
95% CI	[−3.6, 8.8]	[−3.1, 0.9]	[−0.4, 8.0]
Bandwidth	2876	3194	2840
Effective N	264	275	266
Uniform kernel			
Estimate	−1.16	−1.71	1.42
95% CI	[−5.4, 7.7]	[−2.9, 0.7]	[−0.5, 7.2]
Bandwidth	962	1077	1611
Effective N	65	82	144
Epanechnikov kernel			
Estimate	9.17^{***}	2.62	6.08^{***}
95% CI	[2.6, 12.2]	[−2.1, 1.3]	[3.1, 9.9]
Bandwidth	2284	1929	2147
Effective N	201	195	197
Pooled positive-share settlements (T2 + T3), local linear			
Estimate	5.63	-0.27^{**}	5.25^{***}
95% CI	[−2.9, 5.7]	[−6.1, −0.0]	[2.4, 5.7]
Bandwidth	2517	1961	2888
Effective N	290	283	384
Continuous 1910-share interaction (T2 + T3)			
Effect at share = 0.25	3.47	−1.33	4.80
Effect at share = 0.64	5.27	−0.44	5.23
Effect at share = 0.90	6.48	0.14	5.52
Lost-court × share	4.62	2.26	1.11

Notes: The first four panels re-estimate the T2 cell (1910 Muslim share >0 to 0.64), same controlled specification as Table 3 (five imbalanced covariates as linear controls, opština-clustered robust bias-corrected inference), re-estimated under a local quadratic and under uniform and Epanechnikov kernels, with the MSE-optimal bandwidth re-selected for each fit; the first panel reproduces the Table 3 headline. Estimate = (lost − retained). The 95% interval and stars are robust bias-corrected, so the interval is centered on the bias-corrected estimate and need not contain the local-linear point. The pooled panel estimates one local-linear RD over all positive-share settlements; the interaction panel adds a lost-court × 1910-share term inside the pooled bandwidth (triangular weights, opština-clustered), reporting the implied effect at low, median, and high shares; its saturated specification is less precise than the pooled RD, so the two panels are read together. Age-cohort outcomes are excluded: near the cutoff the district-level running variable has too few distinct values for narrow-window or higher-order fits to identify them stably. *** $p < 0.01$, ** $p < 0.05$.

O Additional Figures and Tables Referenced in Main Text

Table A.8. *Temporal Sequence: Legacies and Mechanisms*

Period	Ottoman Territory	Habsburg Territory
Pre-1878 (Imperial legacies)	Millet system; thin bureaucracy; Muslims as ruling group	Merit-based bureaucracy; compulsory standardized schooling
1878–1945 (Post-imperial)	Muslim status inverted; thin bureaucracy persists	Dense administration repurposed by successor states
1945–1960s (Early communist)	Centralized schooling; incorporation low; coordination costly	Centralized schooling builds on inherited capacity
1960s–1991 (Late communist)	Coordination resolved; incorporation frictions persist	Near-full convergence achieved

Figure A.14. *Muslims and Ethno-Religious Fractionalization*

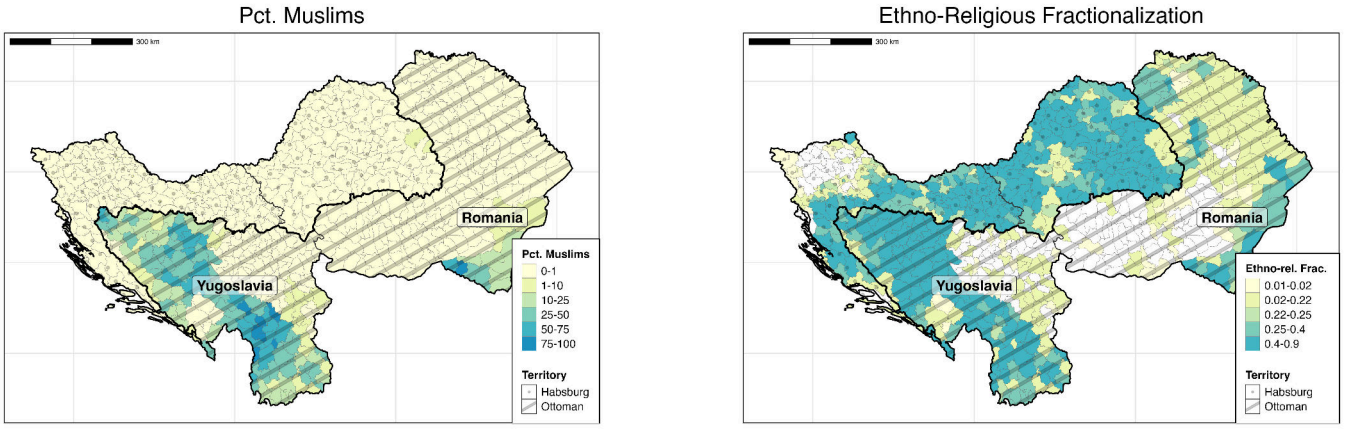


Figure A.15. *Border Types: High-Stakes and Low-Stakes Segments*

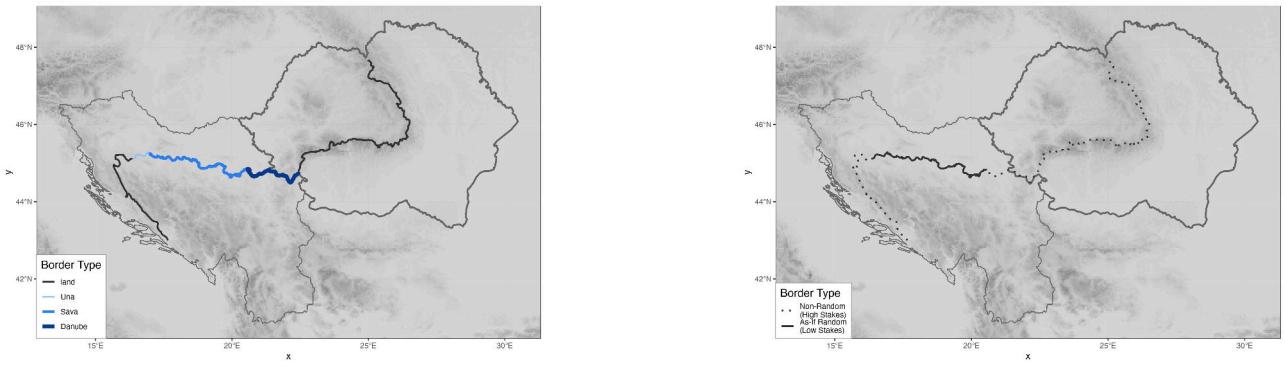


FIGURE A.16. *State Capacity in Yugoslavia and Romania, 1931*

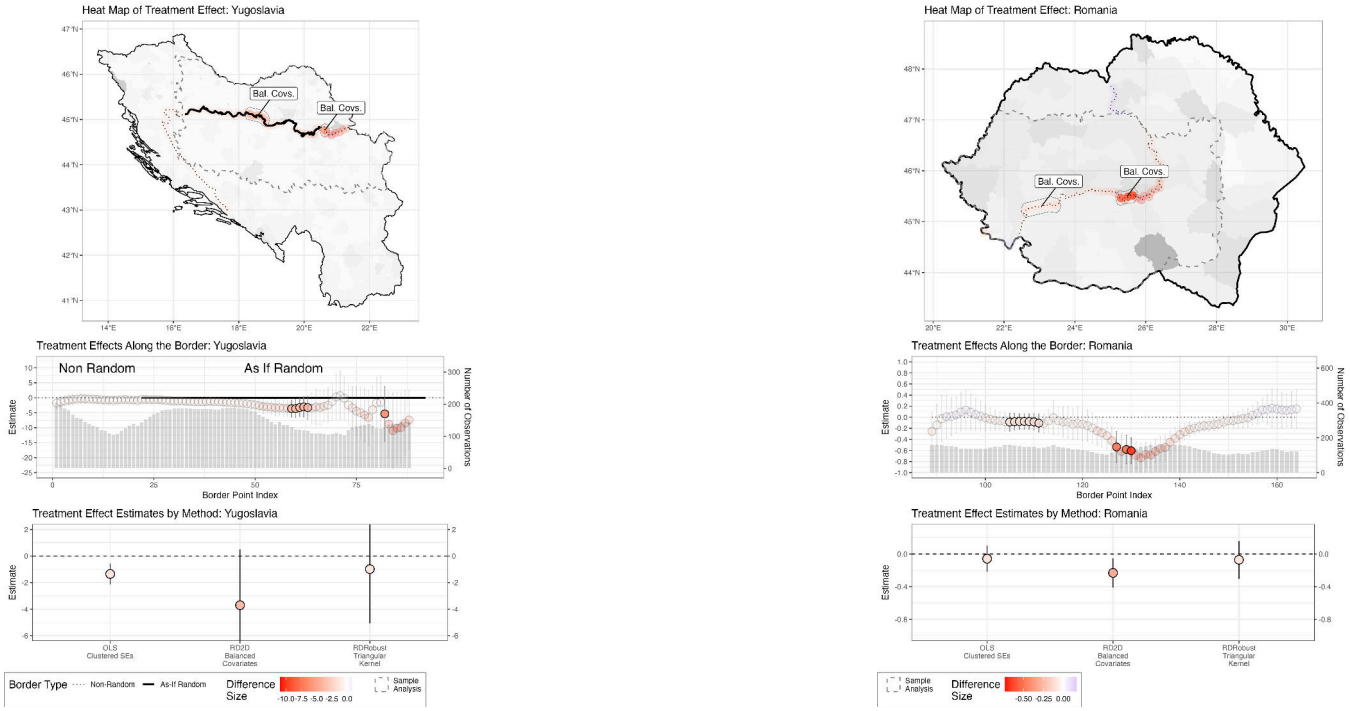


FIGURE A.17. *State Provision, Household Adoption, and Infrastructure Balance by Muslim Share*

