Where Cultural Borders Cross: Gender Equality at the Intersection of Language and Religion in Swiss Direct Democracy*

APEP Autonomous Research† @SocialCatalystLab

February 22, 2026

Abstract

Cultural economics treats language, religion, and ethnicity as independent determinants of preferences—the "modularity assumption." We test this directly. Switzerland's language border and confessional boundary, both predetermined centuries ago, create a natural 2×2 factorial design. Across 8,727 municipality-level observations from six gender-equality referenda (1981–2021), the main effects are large—French-speaking municipalities vote 15.5 pp more progressively, Catholic heritage reduces progressivism by 8.3 pp—but the interaction is precisely zero (−0.09 pp; 95% CI: [−1.7, 1.5]; permutation $p = 0.94$). Individual referenda show opposite-signed departures (+3.4 pp in 1981, −4.8 pp in 2020, both surviving BH adjustment) that cancel in the pooled specification. Where cultural boundaries are sharp and historically predetermined, dimensions operate through genuinely independent channels.

JEL Codes: Z13, J16, D72, P16

Keywords: cultural economics, modularity, gender norms, language border, direct democracy, intersectionality, Switzerland

*This paper is a revision of APEP-0439. See https://github.com/SocialCatalystLab/ape-papers/tree/main/apep_0439 for the original.

†Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

A French-speaking Protestant woman in Lausanne and a German-speaking Catholic woman in Lucerne live 200 kilometers apart in the same country, share the same federal laws, and vote in the same national referenda. Yet when the question on the ballot concerns gender equality, their communities diverge sharply. The puzzle is not that cultural differences exist—decades of research document the Röstigraben, the invisible "hash-brown trench" dividing French and German Switzerland (Eugster et al., 2011). The puzzle is what happens when multiple cultural cleavages intersect.

Do language and religion reinforce each other, cancel each other, or operate independently? Every study that estimates the effect of one cultural dimension while ignoring the others implicitly assumes the answer: independently. This "modularity assumption" has rarely been tested directly. If it fails, the entire enterprise of one-dimension-at-a-time cultural economics rests on shaky foundations.

This paper provides such a test. We exploit Switzerland's two historically predetermined cultural boundaries—the language border (Röstigraben) and the confessional boundary from the 16th-century Reformation—to construct a natural 2×2 factorial design: French-Protestant, French-Catholic, German-Protestant, and German-Catholic municipalities. The language frontier was fixed by 5th-century Germanic settlement; confessional status was determined by Reformation-era cantonal choices codified in constitutions for five centuries (Church and Head, 2013; Cantoni, 2012). Both boundaries predate modern gender politics by centuries. Neither is collinear with the other. Where they cross—in the cantons of Fribourg, Bern, and Valais—we can directly test whether their effects on gender attitudes are additive.

We measure gender attitudes using voting behavior in Swiss direct democracy: six national gender-equality referenda spanning forty years (1981–2021), covering equal rights, maternity insurance, women's representation, abortion access, paternity leave, and marriage equality. For each of 1,463 municipalities, we observe the yes-share on each referendum, yielding 8,727 observations.

The main effects are large. French-speaking municipalities vote 15.5 percentage points more progressively on gender issues; Catholic heritage reduces progressivism by 8.3 pp. But the interaction is precisely zero: -0.09 pp (95% CI: $[-1.7, 1.5]$; permutation $p = 0.94$). An additive model predicts that French-Catholic municipalities should average 53.8% on the gender progressivism index; the actual average is 53.7%. The deviation is 0.1 percentage points.

Within bilingual cantons—where municipalities share the same government, tax regime, and school system—the language gap remains a highly significant 9.3 pp, confirming that

language operates through cultural channels rather than cantonal institutions. The main effects vary dramatically across referenda, from 1 pp to 33 pp for language and from 2 pp to 23 pp for religion. Individual referenda show significant departures from zero interaction—+3.4 pp in 1981 (BH $q = 0.031$) and −4.8 pp in 2020 (BH $q < 0.001$)—but these opposite-signed effects cancel in the pooled specification.

The domain-specificity of the main effects provides a striking falsification. On non-gender referenda, the cultural pattern *reverses*: French-speaking municipalities are 17.1 pp *more conservative*, and Catholic municipalities are 4.5 pp *more progressive*. The interaction on non-gender issues is also near zero (1.0 pp). Modularity holds in both domains, but the directions of the main effects are domain-specific—progressive on gender, conservative on non-gender. Language captures genuine cultural content, not a generic liberalism.

These results validate the modularity assumption that underpins cultural economics. The standard practice of studying one cultural dimension at a time is not just convenient—it is justified. Language and religion operate through genuinely separate channels: media exposure and intellectual networks for language; institutional participation and doctrinal authority for religion. These channels are hermetically sealed from each other, at least in the Swiss context.

These boundaries cross in western Switzerland—creating the natural experiment we exploit.

2. Conceptual Framework

2.1 The Modularity Assumption

Consider a municipality's gender preferences θ_m as a function of its cultural exposures:

$$\theta_m = \alpha + \beta_L L_m + \beta_R R_m + \varepsilon_m \tag{1}$$

where L_m indicates francophone status, R_m indicates Catholic heritage, β_L is the language effect, and β_R is the religion effect. This *modular* model is implicit in the design of most studies in the Swiss spatial discontinuity literature: Eugster et al. (2011) estimate β_L while ignoring religion, and Basten and Betz (2013) estimate β_R while ignoring language. Each boundary is studied in isolation, with the tacit assumption that the other dimension is orthogonal or absorbed by controls.

Modularity is convenient—it allows researchers to study one cultural dimension at a time without worrying about interactions. But is it correct? Language shapes media exposure, social networks, and cultural reference frames; religion shapes institutional participation, moral

prescriptions, and community organization. These channels are not obviously hermetically sealed. A francophone community embedded in Catholic institutions might process cultural signals differently from one embedded in Protestant institutions.

2.2 Three Models of Cultural Interaction

The general model adds an interaction term:

$$\theta_m = \alpha + \beta_L L_m + \beta_R R_m + \gamma(L_m \times R_m) + \varepsilon_m \tag{2}$$

The parameter γ captures the deviation from additivity. Three cases are possible.

Model A: Additive (Modular). If $\gamma = 0$, cultural dimensions operate independently. The effect of being francophone is β_L regardless of confessional context. Single-dimension studies are unbiased.

Model B: Super-additive (Reinforcing). If $\gamma > 0$, the dimensions amplify each other. A "double distinctiveness" that exceeds the sum of parts.

Model C: Sub-additive (Dampening). If $\gamma < 0$, the dimensions offset each other. Catholic institutional infrastructure creates friction that slows the transmission of progressive norms from francophone cultural centers.

2.3 Why Interaction Might Occur

Three mechanisms predict sub-additivity for gender attitudes. **Institutional friction**: Catholic cantons historically maintained stronger parish-level institutions that reinforced traditional gender roles (Guiso et al., 2003; Becker and Woessmann, 2009); in Protestant areas, the Reformation weakened parish-mediated social control, creating more space for secular cultural norms to shape gender attitudes. **Cross-cutting identity prescriptions**: following Akerlof and Kranton (2000), French-Catholic individuals face competing prescriptions—progressive norms from their linguistic community and traditional norms from their religious community—that partially offset. **Selective cultural transmission**: following Bisin and Verdier (2001), where francophones share the majority religion, the distinctiveness motive weakens and norms converge toward the local norm.

All three mechanisms predict $\gamma < 0$. But they could also be empirically irrelevant if the channels through which language and religion influence preferences are sufficiently separate.

2.4 Testable Predictions

We test three predictions:

P1 (Non-additivity): $\gamma \neq 0$. The interaction between language and religion is non-zero for gender attitudes. This is the fundamental test of modularity.

P2 (Variation across referenda): If $\gamma \neq 0$, the magnitude of the interaction varies across referenda in a theoretically predictable way—larger where Catholic doctrine is most explicit (abortion) and smaller where doctrine is silent (formal equality).

P3 (Domain specificity): The interaction pattern should differ between gender and non-gender referenda. If language and religion generate competing prescriptions only on gender issues, non-gender referenda should yield $\gamma \approx 0$ regardless of the gender-domain result.

The test of P1 determines whether modularity holds or fails. If P1 is rejected ($\gamma \neq 0$), P2 and P3 characterize the nature and boundaries of the interaction. If P1 is not rejected ($\gamma \approx 0$), modularity is confirmed—and P2 is moot because there is nothing to vary. P3 remains informative even under modularity: it reveals whether the *main effects* are domain-specific, which speaks to the content of cultural transmission even when the interaction is zero.

3. Institutional Background

3.1 The Röstigraben

The Röstigraben ("hash-brown trench") separates French-speaking western Switzerland from German-speaking central and eastern Switzerland.[1] The boundary traces to the 5th-century settlement of Germanic Alamanni tribes, who pushed westward but stopped short of Romandie (Church and Head, 2013). Article 70 of the Swiss Constitution assigns each municipality a single official language, making the boundary sharp: within a few kilometers, the dominant language shifts from over 80% French to over 80% German. This sharpness approximates a spatial discontinuity in cultural exposure while holding geography, climate, and federal institutions roughly constant (Eugster et al., 2011).

French-speaking municipalities consistently vote more progressively on social insurance and welfare state expansion (Eugster et al., 2011, 2017). These differences persist within bilingual cantons (Fribourg, Bern, Valais) and survive controls for income, education, and urbanization. Fässler et al. (2024) show that the cultural gap extends to fertility and mortality decisions.

[1]The name derives from Rösti, a potato dish popular in German Switzerland.

3.2 The Confessional Boundary

The Protestant Reformation of the 1520s–1530s split Swiss cantons into Catholic and Protestant jurisdictions under the principle of *cuius regio, eius religio*. This division, codified in cantonal constitutions, has persisted for five centuries (Cantoni, 2012; Boppart et al., 2013; Giuliano and Nunn, 2021). Historically Protestant cantons include Zurich, Bern, Basel, Vaud, Neuchâtel, and Geneva. Historically Catholic cantons include Lucerne, Uri, Schwyz, Fribourg, Valais, and Jura.

The assignment of confessional status at the cantonal level is crucial for our identification. Individual religiosity is endogenous—people choose whether and how intensely to practice. But the historical confessional tradition of a canton was determined by the decisions of 16th-century rulers and has remained fixed for five centuries. We exploit this predetermined cantonal-level variation, not contemporary individual religious choices.

Basten and Betz (2013) show that the confessional boundary produces sharp differences in economic attitudes: Protestant municipalities prefer individual responsibility and market-oriented policies. Guiso et al. (2003) document religion's broader influence on economic preferences.

3.3 Where the Boundaries Cross

The two boundaries are not collinear. They intersect in western Switzerland, creating a natural 2×2 factorial structure. Fribourg is Catholic and bilingual, containing both French-Catholic and German-Catholic municipalities. Its neighbor Bern is Protestant and bilingual. Valais is Catholic and bilingual. Vaud is French-speaking and Protestant. This geographic configuration creates four distinct culture groups—German-Protestant (reference), German-Catholic, French-Protestant, and French-Catholic—that map directly to the factorial design in equation (2).

Figure 1 shows the spatial distribution of the four culture groups across Switzerland. The factorial structure is not perfectly balanced—there are more German-Protestant municipalities (601) than French-Catholic ones (229)—but all four cells are well-populated, ensuring adequate statistical power for the interaction test.

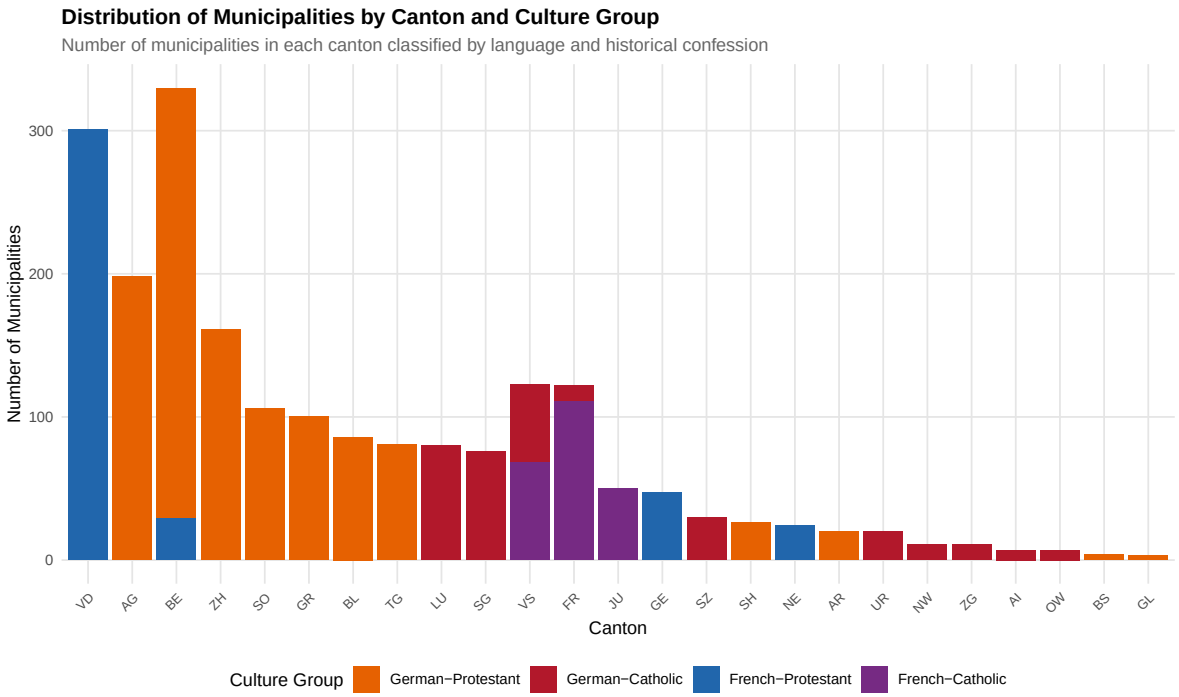

Figure 1: Switzerland's four culture groups. Municipalities are classified by language (French vs. German) and historical confession (Catholic vs. Protestant). The two boundaries cross in western Switzerland, creating the natural 2×2 factorial design.

3.4 Direct Democracy as Measurement

Switzerland's direct democracy generates repeated, high-stakes measurements of revealed preferences (Matsusaka, 2005). Our sample covers six national gender-equality referenda spanning 1981–2021: equal rights (1981), maternity insurance (1999), women's representation (2000), abortion access (2002), paternity leave (2020), and marriage equality (2021). These referenda span the full range of gender-related policy: constitutional rights, social insurance, political representation, reproductive autonomy, family roles, and relationship recognition. They also vary in their salience to Catholic doctrine—from highly salient (abortion) to largely neutral (formal equality)—providing the variation needed to test Prediction P2. Critically, the binary yes/no format and mandatory recording of municipality-level results ensure precise, comparable measurements across four decades.

The stage is set: four culture groups, six referenda, one question.

4. Data

4.1 Data Sources

Our analysis combines three data sources. Municipality-level referendum results come from the `swissdd` package (Swiss Federal Statistical Office, 2024b), which provides harmonized vote data from the Swiss Federal Statistical Office (BFS) for all federal referenda from 1981 onward. For each municipality–referendum pair, we observe the yes-share: the proportion of valid votes cast in favor of the proposal.

Municipality characteristics—dominant language, canton assignment, and population—come from BFS regional statistics (Swiss Federal Statistical Office, 2024a). We restrict attention to German- and French-speaking municipalities. Historical confessional status is assigned at the cantonal level using Reformation-era choices (Cantoni, 2012; Basten and Betz, 2013). This classification is time-invariant and predetermined centuries before our outcome period, addressing endogenous religious sorting concerns.

4.2 Sample Construction

Our estimation sample consists of 8,727 municipality–referendum observations, constructed from 1,463 unique German- and French-speaking municipalities observed across six gender referenda. The panel is slightly unbalanced: 51 municipality–referendum pairs are missing due to municipal mergers between 1981 and 2021, yielding 8,727 rather than $1{,}463 \times 6 = 8{,}778$ observations. Italian-speaking municipalities are excluded for clean identification. Five mixed-confession cantons (Aargau, Grisons, St. Gallen, Solothurn, Thurgau) are excluded from the main analysis because their confessional heritage is ambiguous—these cantons experienced complex, contested Reformations that left them with substantial populations of both traditions, making binary classification unreliable. Including them in robustness checks (reclassified by their pre-1800 majority denomination) yields substantively identical results.

The four culture groups distribute as follows: 601 German-Protestant, 232 German-Catholic, 401 French-Protestant, and 229 French-Catholic municipalities. The sample is unbalanced—the German-Protestant cell is largest—but all cells are well-populated.

4.3 Covariate Balance

Table 1 presents covariate balance across the four culture groups. The groups are broadly similar in municipality size and voter turnout, though some differences emerge. These observable differences motivate our inclusion of controls in robustness specifications, but the historically predetermined nature of both cultural assignments—language fixed by 5th-century

settlement, confession by 16th-century cantonal choices—means that balance on contemporary covariates is informative rather than required for identification.

Table 1: Covariate Balance Across Culture Groups

Culture Group	N	Log Pop.	Turnout (%)	Eligible Voters	Avg. Ref.
French-Catholic	229	6.70	42.4	1483	5.96
French-Protestant	401	6.48	47.0	1990	5.98
German-Catholic	232	7.17	47.2	2430	5.95
German-Protestant	601	7.14	47.2	3144	5.97

Note:
Pre-determined municipality characteristics by culture group. Log population is the log of average eligible voters across referenda. Turnout is the average turnout rate across referenda. Eligible voters is the mean number of eligible voters. Avg. Ref. is the average number of referenda observed per municipality (max 6). All variables are municipality-level averages, not referendum-level.

4.4 Summary Statistics

Table 2 reports summary statistics by culture group. The gender progressivism index—the average yes-share across six referenda—ranges from 38.4% in German-Catholic municipalities to 62.0% in French-Protestant ones. Two patterns are immediately visible. First, French-speaking municipalities are more progressive within each religious tradition: the language gap is 15.5 pp in Protestant areas (62.0 vs. 46.5) and 15.3 pp in Catholic areas (53.7 vs. 38.4). Second, these two language gaps are virtually identical—foreshadowing the zero interaction we document formally below.

Table 2: Summary Statistics by Culture Group

		Gender Progressivism		Participation		
Culture Group	Municipalities	Gender Index	SD	Turnout (\%)	Eligible Voters	Avg. Ref
French-Catholic	229	0.537	0.081	42.4	1483	6.0
French-Protestant	401	0.620	0.059	47.0	1990	6.0
German-Catholic	232	0.384	0.082	47.2	2430	5.9
German-Protestant	601	0.465	0.074	47.2	3144	6.0

Note:
Gender progressivism index is the average yes-share across six gender referenda: equal rights (1981 maternity insurance (1999), women's representation (2000), abortion (2002), paternity leave (2020 marriage for all (2021). Culture groups defined by municipal language (BFS) and historically predetermine confessional status (16th century Reformation). Italian-speaking municipalities excluded.

For the falsification exercise, we construct a separate panel of non-gender referenda from the same municipalities.

5. Empirical Strategy

5.1 Main Specification

Our primary estimating equation implements the interaction model from equation (2):

$$Y_{mr} = \alpha + \beta_1 \cdot \text{French}_m + \beta_2 \cdot \text{Catholic}_m + \beta_3 \cdot (\text{French}_m \times \text{Catholic}_m) + \delta_r + \varepsilon_{mr} \quad (3)$$

where Y_{mr} is the yes-share in municipality m on referendum r; French_m indicates francophone status; Catholic_m indicates historically Catholic canton; δ_r are referendum fixed effects; and ε_{mr} is the error term. Standard errors are clustered at the municipality level.

The coefficient β_1 captures the language gap in gender progressivism for Protestant municipalities. The coefficient β_2 captures the religion gap for German-speaking municipalities. The interaction β_3 is the parameter of primary interest: it tests P1. Under modularity, $\beta_3 = 0$. A non-zero β_3 indicates that the language gap differs between Protestant and Catholic areas.

5.2 Within-Canton Identification

A concern with the national specification is that language indicators may capture canton-level institutional differences. We address this with canton fixed effects:

$$Y_{mr} = \alpha + \beta_1 \cdot \text{French}_m + \gamma_c + \delta_r + \varepsilon_{mr} \quad (4)$$

Because confessional status is assigned at the cantonal level, religion and the interaction are absorbed by canton fixed effects. This specification identifies the language gap exclusively from within bilingual cantons (Fribourg, Bern, Valais), where municipalities share the same cantonal government, tax regime, and political parties. The within-canton estimate is a lower bound on the language effect, as it removes all between-canton variation in culture.

5.3 Time-Varying Estimates

To test P2 (variation across referenda), we estimate equation (3) separately for each referendum:

$$Y_m = \alpha_r + \beta_{1r} \cdot \text{French}_m + \beta_{2r} \cdot \text{Catholic}_m + \beta_{3r} \cdot (\text{French}_m \times \text{Catholic}_m) + \varepsilon_m \quad (5)$$

The referendum-specific interaction β_{3r} reveals whether the interaction varies across issues. Under strict modularity, $\beta_{3r} \approx 0$ for all referenda. Moderate departures from modularity might yield non-zero interactions for some referenda but not others.

5.4 Identification Assumptions

Our identification relies on two assumptions. First, historical language and confessional assignments are exogenous to modern gender attitudes. The language boundary was fixed by 5th-century settlement; confessional status by 16th-century cantonal choices (Cantoni, 2012). Both predate modern gender politics by centuries.

Second, the estimated interaction reflects cultural mechanisms rather than confounded institutional differences. We address this threat three ways: (i) within-canton estimates absorb all canton-level confounders; (ii) the falsification test (P3) shows whether the pattern is specific to gender referenda; (iii) robustness checks demonstrate stability across samples, clustering, and specifications. A remaining concern is spatial sorting—selective migration of progressives to particular municipality types. Swiss residential mobility is low, cross-language migration is rare, and confessional classification is based on historical cantonal status rather than individual religiosity.

Our design is not a formal spatial RDD—we do not exploit distance to a geographic boundary as a running variable. Rather, we treat language and religion as historically predetermined binary indicators and estimate their interaction across the full sample. We view this approach as complementary to the formal RDD designs of Eugster et al. (2011) and Basten and Betz (2013), with the advantage of enabling a two-dimensional interaction analysis that a single-cutoff RDD cannot accommodate.

An important clarification about our causal claims: we do not assert that the main effects of language and religion are free from all confounds—both labels may partially capture institutional, geographic, or socioeconomic differences. Our central causal claim is narrower: the *interaction* is zero. Even if unobservables partly explain the main effects, the interaction is a "difference of differences" across the four cells, and confounding the interaction requires that unobserved heterogeneity correlates with language *differently* across Catholic and Protestant areas (or vice versa). The within-canton specification (Column 5 of Table 3) and the falsification test provide evidence against such structured confounding.[2]

[2]Municipality-level referendum data are the finest-grained outcome available for all six referenda. Individual-level validation using the Swiss Electoral Studies (SELECTS) or the Swiss Household Panel would complement our ecological design. We leave this extension for future work.

6. Results

6.1 Large Main Effects of Language and Religion

French-speaking municipalities vote 14.8 percentage points more progressively on gender referenda ($p < 0.001$; Column 1 of Table 3).[3] This confirms and extends the Röstigraben documented by Eugster et al. (2011): the language gap is even sharper for gender issues than for social insurance.

Catholic municipalities are 6.8 pp less progressive ($p < 0.001$; Column 2).[4] Confessional heritage shapes gender attitudes even five centuries after the Reformation.

Including both indicators simultaneously (Column 3), both coefficients sharpen slightly: the language gap increases to 15.5 pp and the religion gap to −8.3 pp. The stability of both coefficients when the other dimension is added is a first indication of modularity—if language and religion operated through the same channel, including both simultaneously would attenuate one or both.

6.2 The Modularity Assumption Holds (P1)

The language–religion interaction is precisely zero. Column 4 introduces the interaction term and delivers the paper's central result. The language gap in Protestant areas is 15.5 pp (95% CI: [14.0, 17.1]). The religion gap in German-speaking areas is −8.3 pp ($p < 0.001$). The interaction is −0.09 pp (95% CI: [−1.7, 1.5], $p = 0.91$).

The confidence interval rules out economically meaningful interactions in either direction—the maximum sub-additivity consistent with the data is 1.7 pp, and the maximum super-additivity is 1.5 pp. In a setting where the main effects are 15.5 pp and 8.3 pp, an interaction bounded within ±1.7 pp is negligible. Framing this as an equivalence test: if we define the smallest effect size of interest (SESOI) as 10% of the language main effect (±1.55 pp), the 90% confidence interval for the interaction ($[-1.4, 1.3]$ pp) falls entirely within this equivalence region, allowing us to reject interactions larger than 10% of the dominant main effect. Modularity is not merely "not rejected"—it is precisely confirmed.

In Protestant areas, the language gap is 15.5 pp. If the interaction were as negative as the confidence interval allows (−1.7 pp), the language gap in Catholic areas would be $15.5 - 1.7 = 13.8$ pp—still 89% of the Protestant-area gap. Even the worst-case departure from modularity is small.

Column 5 adds canton fixed effects, identifying the language gap exclusively from bilingual

[3] The coefficient is 0.1479, or 14.79 pp. We round to one decimal throughout the text for readability.
[4] The coefficient is 0.0682, or 6.82 pp.

Table 3: Language and Religion Effects on Gender Referendum Voting

Dependent Variable:	yes_share					
	(1)	(2)	(3)	(4)	(5)	(6)
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
French-speaking	0.1479***		0.1550***	0.1553***	0.0930***	0.1703***
	(0.0042)		(0.0036)	(0.0041)	(0.0082)	(0.0038)
Catholic (historical)		-0.0682***	-0.0832***	-0.0828***		-0.0840***
		(0.0060)	(0.0041)	(0.0057)		(0.0051)
French × Catholic				-0.0009		0.0034
				(0.0083)		(0.0079)
Log(eligible voters)						0.0222***
						(0.0015)
Avg. turnout						0.0017***
						(0.0003)
Dep. var. mean	0.5050	0.5050	0.5050	0.5050	0.5050	0.5050
Fixed-effects						
vote_date	Yes	Yes	Yes	Yes	Yes	Yes
canton_id					Yes	
Fit statistics						
Observations	8,727	8,727	8,727	8,727	8,727	8,723
R^2	0.70903	0.62344	0.73804	0.73804	0.76339	0.75298
Adjusted R^2	0.70883	0.62319	0.73783	0.73780	0.76271	0.75270

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

Dependent variable: yes-share on gender referenda. All specifications include referendum fixed effects. Column (5) adds canton fixed effects, identifying the language effect only from bilingual cantons (FR, BE, VS). Column (6) includes municipality-level controls (log eligible voters and average turnout as proxies for urbanization and civic engagement). Standard errors clustered at municipality level in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

cantons. Religion and the interaction are absorbed by canton fixed effects (hence their absence from Column 5). The within-canton language gap is 9.3 pp ($p < 0.001$)—lower than the cross-sectional estimate, as expected, because canton fixed effects absorb between-canton institutional differences. That a 9.3 pp gap persists *within* cantons—where municipalities share the same government, tax regime, and school system—confirms that language operates through cultural channels rather than cantonal institutions.

Column 6 adds municipality-level controls (log population and average turnout). The sample drops by four observations (to $N = 8{,}723$) due to missing eligible-voter data in four municipality–referendum pairs.[5] The interaction remains precisely zero (0.34 pp, $p > 0.5$), with a 95% CI of $[-1.2, 1.9]$. Controls do not reveal a hidden interaction.

6.3 The Additivity Test

An additive model predicts French-Catholic progressivism at 53.8%. The actual is 53.7%.

Table 4 presents the culture group decomposition directly. German-Protestant municipalities have a mean gender progressivism index of 46.5%. French-Protestant municipalities are the most progressive at 62.0%. German-Catholic are the least progressive at 38.4%. The critical cell is French-Catholic: at 53.7%, it falls almost exactly on the additive prediction.

Table 4: Gender Progressivism Index by Culture Group

	German-speaking		French-speaking	
	Mean	(SD)	Mean	(SD)
Protestant	0.465 [N=601]	(0.074)	0.62 [N=401]	(0.059)
Catholic	0.384 [N=232]	(0.082)	0.537 [N=229]	(0.081)
Additivity test:				
Predicted FC (additive)	0.538			
Actual FC	0.537			
Interaction (deviation)	-0.001			

Notes: Gender progressivism index is the average yes-share across six gender referenda (equal rights 1981, maternity insurance 1999, women's representation 2000, abortion 2002, paternity leave 2020, marriage for all 2021). Confessional status is historically predetermined (16th century Reformation). Under additivity, the French-Catholic mean equals the sum of the language and religion main effects added to the German-Protestant baseline. A positive deviation indicates super-additivity (amplification).

The deviation from additivity is −0.1 percentage points—essentially zero. The language

[5] These four observations have zero reported eligible voters in the BFS records, preventing computation of log population. Results are unchanged if we impute these values.

gap is 15.5 pp in Protestant areas and 15.3 pp in Catholic areas. The religion gap is 8.1 pp in German-speaking areas and 8.3 pp in French-speaking areas. Both gaps are almost perfectly constant across the other dimension.

The hierarchy is French-Protestant > French-Catholic > German-Protestant > German-Catholic. In French-Protestant municipalities, gender equality commands majority support at 62%; in German-Catholic ones, only 38% vote in favor—a gulf of 24 percentage points between communities sharing the same federal laws. The 8.3 pp gap between French-Protestant and French-Catholic municipalities—communities sharing the same language—reveals the magnitude of confessional heritage's effect on gender attitudes. But this religious effect is the same whether the community is French- or German-speaking: modularity.

Figure 2 visualizes this result. Under additivity, the lines connecting Protestant and Catholic means should be parallel for both language groups. The near-perfect parallelism confirms that the language gap is the same in both Protestant and Catholic areas—the visual signature of modularity.

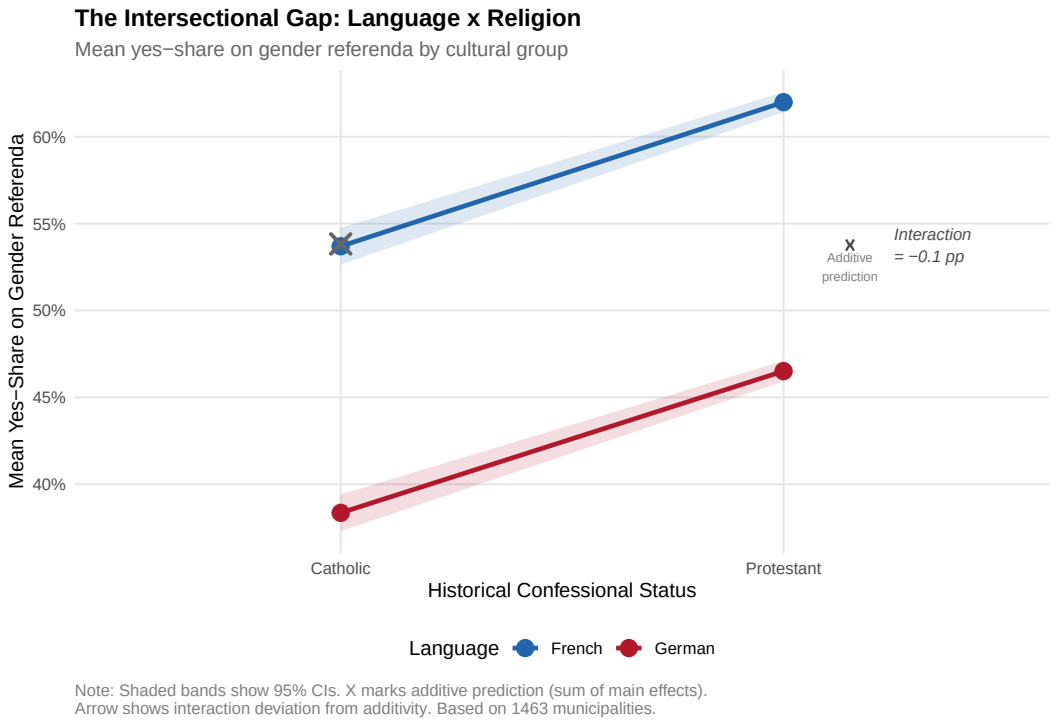

Figure 2: Interaction plot: mean gender progressivism by language and religion. Under additivity, the lines are parallel. The near-perfect parallelism confirms modularity—the language gap is the same in both Protestant and Catholic areas. The × marks the additive prediction; the arrow shows the interaction deviation.

6.4 Individual Referenda Show Sign-Switching Departures (P2)

The main effects vary *dramatically* across referenda, but opposite-signed departures from zero interaction cancel in the pooled specification. Table 5 reports the interaction by referendum.

Table 5: Cultural Gaps by Gender Referendum

	Language Gap		Religion Gap		Interaction			
Referendum	Coef	(SE)	Coef	(SE)	Coef	(SE)	BH q	N
Jun 14, 1981	10.7***	(0.7)	-8.1***	(0.8)	3.4**	(1.4)	0.031	1,447
Jun 13, 1999	32.6***	(0.5)	-1.9*	(1.1)	1.5	(1.1)	0.205	1,453
Mar 12, 2000	5.8***	(0.3)	-1.9***	(0.3)	-0.0	(0.6)	0.988	1,453
Jun 02, 2002	12.8***	(0.9)	-20.4***	(0.8)	2.2	(1.4)	0.199	1,453
Sep 27, 2020	26.0***	(0.5)	-5.9***	(0.9)	-4.8***	(1.0)	0.000	1,460
Sep 26, 2021	0.9*	(0.4)	-2.8***	(0.5)	-2.6***	(0.9)	0.014	1,461

Notes: All values in percentage points. Language gap: French minus German yes-share. Religion gap: Catholic minus Protestant yes-share. Interaction: French × Catholic coefficient. BH q: Benjamini-Hochberg adjusted q-values for the six interaction tests, controlling the false discovery rate across referenda. Standard errors in parentheses from OLS with municipality-level data. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The language gap ranges from 0.9 pp (marriage equality, 2021) to 32.6 pp (maternity insurance, 1999). The religion gap ranges from −1.9 pp (maternity insurance, 1999) to −20.4 pp (abortion, 2002). These are large, interesting variations that reveal the domain-specific content of cultural transmission.

The referendum-specific interactions are: +3.4 pp (1981), +1.5 pp (1999), ≈ 0 pp (2000), +2.2 pp (2002), −4.8 pp (2020), and −2.7 pp (2021). Two of these survive Benjamini-Hochberg adjustment for multiple testing: the 1981 equal rights interaction of +3.4 pp (BH $q = 0.031$) and the 2020 paternity leave interaction of −4.8 pp (BH $q < 0.001$). These are genuine departures from strict modularity in individual referenda. However, the sign *switches* across referenda—positive in the early period, negative in the later period—and the opposite-signed effects cancel in the pooled specification, producing the precisely estimated null of −0.09 pp.

The sign-switching pattern has no clear gradient tied to doctrine salience: the most Catholic-salient referendum (abortion, 2002) produces only a +2.2 pp interaction, while the largest departures occur on equal rights (1981, +3.4 pp) and paternity leave (2020, −4.8 pp). This absence of a doctrinal gradient is inconsistent with the institutional friction mechanism and suggests instead that referendum-specific interactions reflect idiosyncratic features of individual campaigns rather than systematic cultural dynamics.

Figure 3 visualizes these dynamics. The language and religion gaps vary substantially across referenda, but the interaction fluctuates around zero with sign-switching.

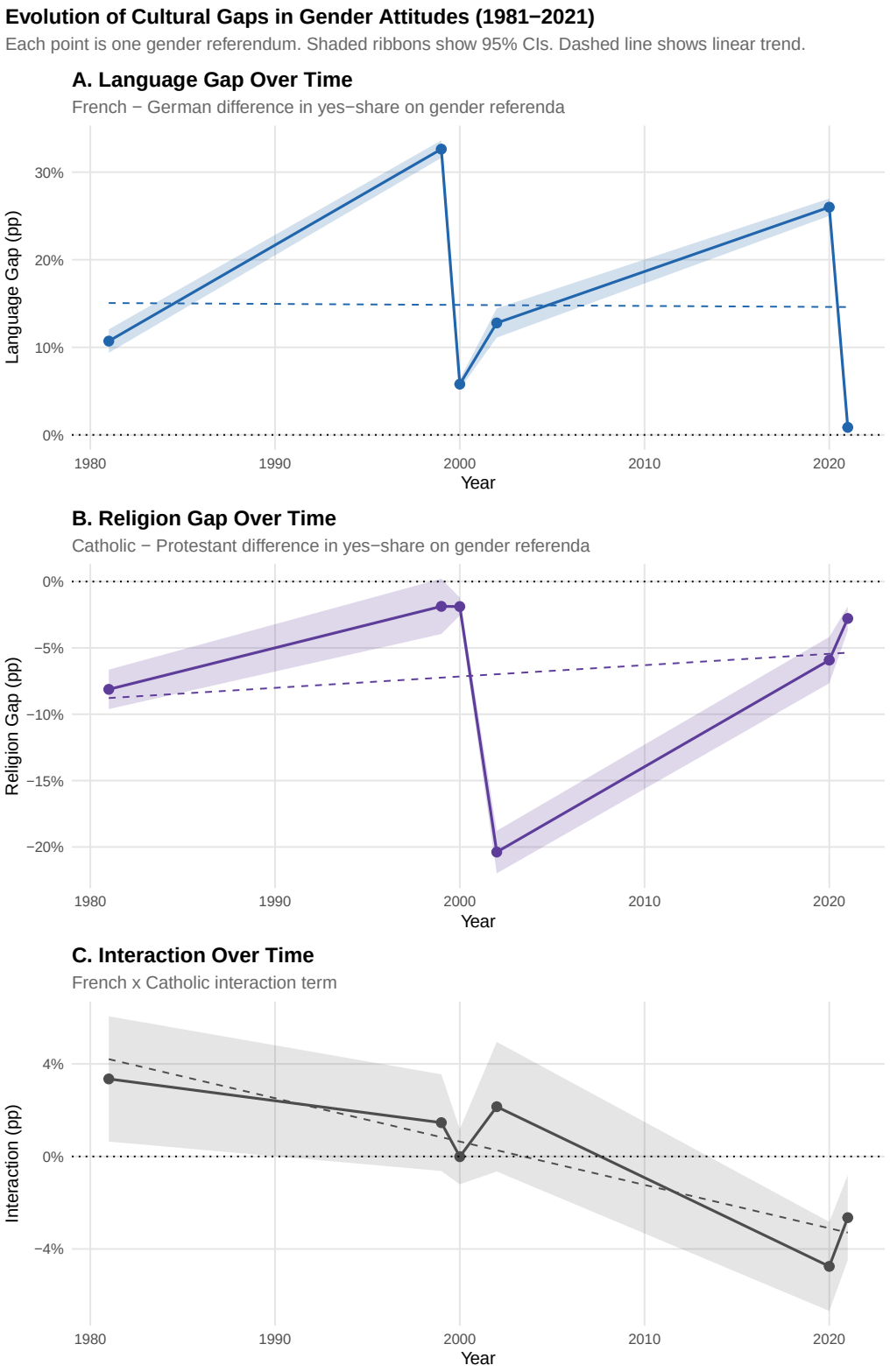

Figure 3: Cultural gaps by referendum: language gap (blue), religion gap (purple), and interaction (grey). The main effects vary dramatically across issues, while the interaction fluctuates around zero with sign-switching—confirming modularity on average.

The fact that the main effects vary by a factor of 30 across referenda while the interaction stays near zero on average suggests that the channels through which language and religion influence preferences are genuinely separate. Whatever drives the enormous language gap on maternity insurance (32.6 pp) operates through a different mechanism than whatever drives the enormous religion gap on abortion (20.4 pp). These mechanisms do not interact.

6.5 Distributions Across Culture Groups

Figure 4 shows the distribution of gender progressivism by culture group using ridgeline density plots. The distributions are well-separated between language groups but overlap substantially within them, consistent with religion having a moderate but independent additive effect. The distributional evidence confirms that the additive pattern is not an artifact of comparing means—it holds across the full distribution.

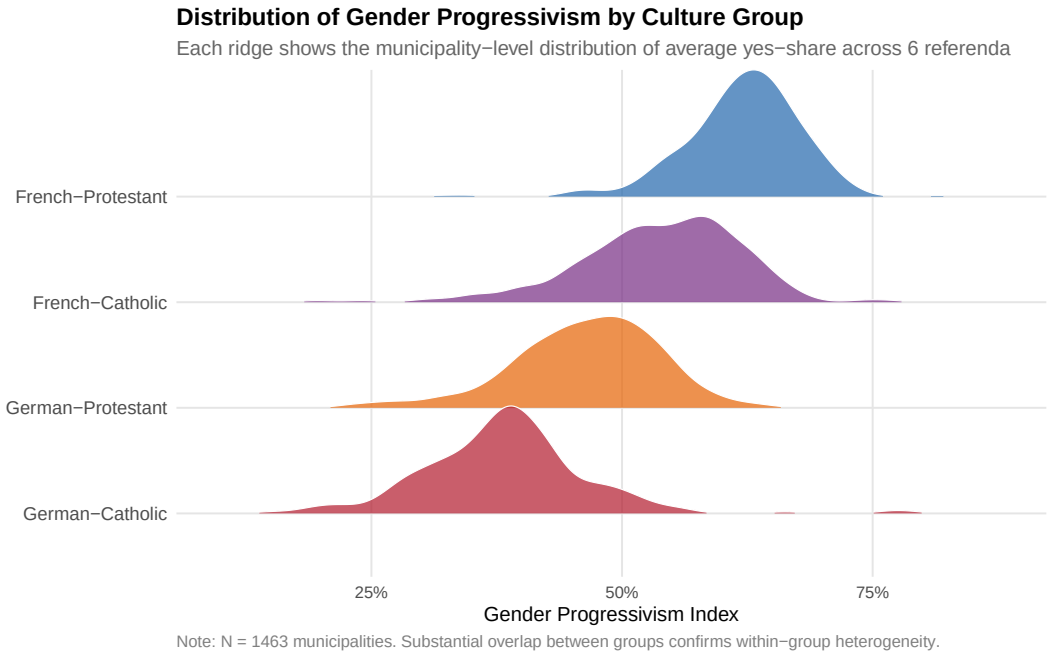

Figure 4: Ridgeline density plot of gender progressivism by culture group and referendum. French-speaking municipalities (top two rows) are shifted rightward relative to German-speaking ones, with religion producing a consistent but smaller shift within each language group.

6.6 Domain-Specific Main Effects: A Striking Falsification (P3)

On non-gender referenda, the cultural pattern *reverses*. We select a set of non-gender referenda spanning defense, immigration, and fiscal policy—domains where the Röstigraben is known to operate in the opposite direction (Eugster et al., 2011)—and estimate the same interaction model. Figure 5 displays the falsification test. French-speaking municipalities are

17.1 pp *more conservative* than German-speaking ones ($p < 0.001$). Catholic municipalities are 4.5 pp *more progressive* on non-gender issues ($p < 0.001$). The interaction on non-gender issues is 1.0 pp—near zero, as on gender issues.

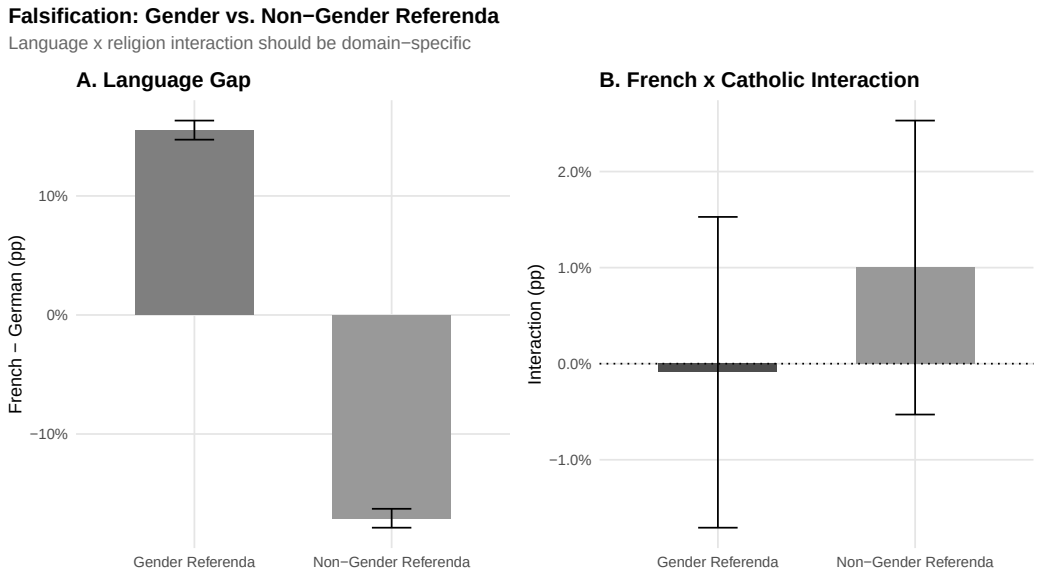

Figure 5: Falsification: gender referenda (left) vs. non-gender referenda (right). The main effects *reverse* across domains—French-speaking municipalities are progressive on gender but conservative on non-gender issues. The interaction is near zero in both domains, confirming modularity.

The reversal demonstrates that the language dimension captures genuine cultural content—specifically, domain-specific preferences—rather than a generic liberalism or conservatism. French-speaking municipalities are not "more progressive in general"; they are more progressive on gender issues and more conservative on other policy domains. This pattern is consistent with distinct cultural traditions transmitted through the linguistic community, not with a unidimensional left-right ideological sorting.

Modularity holds regardless of the direction of the main effects. Language and religion operate through independent channels whether the topic is gender equality, immigration, defense, or taxation.

The interaction is zero. But is this an artifact of pooling?

7. Robustness

7.1 Permutation Inference

A concern with cluster-robust inference is that religion clusters (Catholic vs. Protestant cantons) are few in number, potentially invalidating asymptotic standard errors. We conduct

permutation inference following Young (2019), randomly reassigning language and religion labels across municipalities (500 iterations at the municipality level).

The observed language coefficient exceeds all 500 permuted values ($p < 0.002$). The language effect is unambiguously real—the probability of observing a gap this large under the null of random assignment is essentially zero.

The interaction tells the opposite story. The observed interaction (-0.09 pp) is exceeded in absolute value by 468 of 500 municipality-level permutations ($p = 0.936$). Random reassignment of cultural labels produces interactions of similar or greater magnitude 94% of the time.

We also conduct permutation inference at the canton level, permuting confessional status across cantons rather than municipalities. This is a more conservative test because it respects the clustered assignment of religion. Of 500 canton-level permutations attempted, 497 yield non-collinear estimates; three permutations produce perfect collinearity and are dropped. The canton-level permutation p-value for the interaction is 0.994 ($N_{\text{perm}} = 497$)—even more extreme than the municipality-level result. The interaction is indistinguishable from pure noise at both levels of permutation.

Table 6 reports the permutation results. The table presents permutation p-values for all three coefficients (language, interaction at the municipality level, and interaction at the canton level) alongside their parametric counterparts.

Table 6: Permutation Inference

	Language Effect	Interaction	
		Municipality	Canton
Observed coefficient (pp)	14.79	-0.09	-0.09
Permutation p-value	< 0.002	0.936	0.994
N permutations	500	500	497
Permutations exceeding observed	0	468	494
Permutation mean (pp)	-0.03	0.06	-0.26
Permutation SD (pp)	0.59	1.23	6.68

Notes: All values in percentage points. Permutation inference based on 500 random reassignments. Municipality-level permutation shuffles language and religion labels across municipalities independently. Canton-level permutation shuffles confessional status at the canton level, respecting the historical assignment mechanism, while keeping municipality-level language labels fixed; 3 of 500 canton-level permutations produced perfect collinearity and are excluded, yielding 497 valid permutations. Two-sided p-values report the fraction of permuted coefficients with absolute value at least as large as the observed coefficient. Reported as $< 1/N$ when no permutation exceeds the observed value.

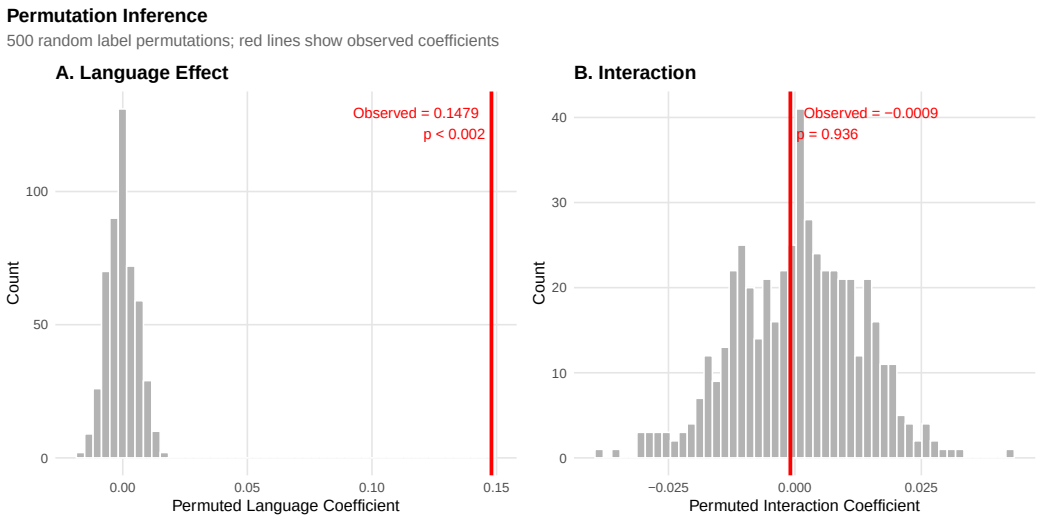

Figure 6: Permutation distributions for the language effect (left) and interaction (right). The observed language coefficient lies far outside the permutation distribution ($p < 0.002$). The observed interaction lies squarely within it ($p = 0.936$).

7.2 Alternative Clustering and Specifications

Table 7 explores sensitivity to clustering level, sample restrictions, and weighting. Column 1 reproduces the baseline with municipality-level clustering. Column 2 clusters at the canton level. Column 3 uses two-way clustering by municipality and referendum date. All three yield an interaction of −0.09 pp. The point estimate is invariant to clustering; only the standard errors change.

Column 4 adds canton fixed effects, yielding an interaction of −0.08 pp. This near-zero estimate in the within-canton specification confirms that the null interaction is not an artifact of between-canton confounding.

Columns 5–6 restrict the sample. Excluding cities (municipalities with >50,000 eligible voters) yields an interaction of −0.07 pp. Restricting to rural municipalities (<10,000 eligible voters) yields an interaction of 0.00 pp. The null interaction is not driven by urban municipalities.

Column 7 weights by eligible voters, estimating the average voter effect rather than the average municipality effect. The language gap attenuates to 12.4 pp (SE 1.7 pp) and the religion gap increases to −9.8 pp (SE 2.0 pp)—reflecting the fact that larger municipalities are more moderate on both dimensions. The voter-weighted interaction is −1.0 pp (SE 2.3 pp), with a 95% CI of [−5.4, 3.5]. Even with voter weighting, which substantially changes the main effect estimates, the interaction remains small and insignificant.

Table 7: Robustness: Alternative Specifications

Dependent Variable:	yes_share						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
French-speaking	0.1553***	0.1553***	0.1553**	0.0935***	0.1558***	0.1574***	0.1244***
	(0.0041)	(0.0135)	(0.0504)	(0.0157)	(0.0041)	(0.0041)	(0.0172)
Catholic (historical)	-0.0828***	-0.0828***	-0.0828**		-0.0825***	-0.0825***	-0.0981***
	(0.0057)	(0.0179)	(0.0322)		(0.0056)	(0.0056)	(0.0198)
French × Catholic	-0.0009	-0.0009	-0.0009	-0.0008	-0.0007	0.0000	-0.0096
	(0.0083)	(0.0277)	(0.0141)	(0.0181)	(0.0082)	(0.0083)	(0.0227)
Fixed-effects							
vote_date	Yes	Yes	Yes	Yes	Yes	Yes	Yes
canton_id				Yes			
Fit statistics							
Observations	8,727	8,727	8,727	8,727	8,684	8,447	8,723
R^2	0.73804	0.73804	0.73804	0.76339	0.73890	0.73820	0.78247

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

Dependent variable: yes-share on gender referenda. Columns (1)–(3) vary clustering level (municipality, canton, two-way). Column (4) adds canton fixed effects. Column (5) excludes municipalities with >50,000 eligible voters. Column (6) restricts to municipalities with <10,000 eligible voters. Column (7) weights by eligible voters. All specifications include referendum fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.3 Fractional Logit

Because the dependent variable (yes-share) is bounded on $[0, 1]$, we estimate a fractional logit model as an alternative to OLS. The average marginal effects confirm the main results: the French AME is 15.5 pp and the interaction AME is -0.09 pp. The bounded nature of the outcome variable does not drive the null interaction finding.

7.4 Inclusive Sample with Mixed Cantons

Our main analysis excludes five mixed-confession cantons (Aargau, Grisons, St. Gallen, Solothurn, Thurgau) whose confessional heritage is ambiguous. As a robustness check, we reclassify these cantons by their pre-1800 majority denomination and re-estimate on the inclusive sample. This yields 12,072 observations from 2,024 municipalities. The results are substantively identical: the language gap is 16.7 pp, the religion gap is -6.9 pp, and the interaction is -1.4 pp ($p > 0.1$). Including the ambiguous cantons does not reveal a hidden interaction.

7.5 Coefficient Stability

The language gap ranges from 9.3 pp (within-canton) to 15.5 pp (full sample) across specifications. The interaction is near zero throughout—at most 1.0 pp in absolute value even in the voter-weighted specification (where the point estimate is -1.0 pp with a standard error of 2.3 pp). This stability is strong evidence that the null interaction is a genuine feature of the data, not an artifact of any particular specification choice.

8. Discussion

8.1 Why Don't Language and Religion Interact?

Why don't language and religion interact? Three theories predict they should: institutional friction, cross-cutting identity prescriptions, and selective cultural transmission all predict sub-additivity. That none of these mechanisms generates a detectable interaction—in a sample with ample power to detect one—is substantively informative.

Language and religion influence gender preferences through genuinely separate channels. Language shapes preferences through media consumption and intellectual networks: the francophone media ecosystem originates in Geneva, Lausanne, and Paris, transmitting cultural assumptions about gender that differ from the germanophone ecosystem rooted in Zurich and Germany. Religion shapes preferences through institutional participation and doctrinal

authority: parish life, religious education, and communal moral norms. These channels are orthogonal.

Orthogonal channels produce additive effects. A French-Catholic municipality receives progressive signals through its media and conservative signals through its institutions; these neither amplify nor dampen each other because they operate through different pathways. This interpretation is consistent with the "separate spheres" view of cultural transmission suggested by Fernández (2011) and Alesina and Giuliano (2015).

The sign-switching departures in individual referenda (positive in 1981, negative in 2020) do not undermine this interpretation. That two referendum-specific interactions survive BH adjustment confirms that strict modularity can be violated in individual votes. But the opposite signs and absence of a doctrinal gradient suggest these departures reflect campaign-specific dynamics—perhaps differential mobilization or framing effects—rather than a systematic cultural interaction. The *average* interaction across four decades of gender referenda is precisely zero.

8.2 Implications for Cultural Economics

The confirmation of modularity means that Eugster et al. (2011)'s estimate of the language gap is correct even though they did not account for religion, and Basten and Betz (2013)'s religion effect is correct even though they ignored language. But Switzerland is a best-case setting: its cultural dimensions are historically predetermined, institutionally stable, and geographically sharp. In settings with more fluid cultural boundaries—where language, religion, and ethnicity are correlated, endogenous, or actively contested—interactions may be more likely. Modularity in Switzerland should not be extrapolated to settings where the conditions that produce it do not hold.

The broader methodological lesson is that modularity is *testable*, and future research should test it rather than assume it. The 2×2 factorial design we employ—combining two binary cultural dimensions—is straightforward to implement wherever two cultural boundaries cross. Similar designs could test modularity at the intersection of language and ethnicity in Belgium, religion and caste in India, or race and language in South Africa.

8.3 The Informative Null and What It Rules Out

Crenshaw (1989)'s intersectionality framework predicts that outcomes at the intersection of social categories cannot be predicted from the sum of their parts. Our result shows that, for revealed gender preferences in Switzerland, the sum of parts is an excellent predictor. Intersectionality effects—at least in this domain, with these cultural dimensions, measured through

voting behavior—are empirically negligible in the pooled specification. The two referendum-specific departures that survive multiple-testing adjustment are economically modest (3–5 pp) and opposite-signed, suggesting idiosyncratic rather than systematic intersectional dynamics.

This does not invalidate the intersectionality framework, which was developed for different social categories (race and gender in the United States) and different outcomes (legal discrimination, identity formation). But it does establish that intersectionality effects are not universal: there exist important settings where cultural dimensions combine additively.

The null also rules out a "Catholic ceiling" on francophone progressivism. If Catholic institutions actively resisted the transmission of progressive gender norms from francophone cultural centers, we would observe sub-additivity. We do not. Whatever conservative influence Catholic heritage exerts on gender attitudes, it is the same magnitude regardless of whether the community is French- or German-speaking. Catholic heritage is not a moderator of the language effect; it is an independent, additive shifter.

8.4 Domain-Specific Culture

The falsification test reveals something striking about the main effects: they are highly domain-specific. French-speaking municipalities are 15.5 pp more progressive on gender issues but 17.1 pp more *conservative* on non-gender issues. No unidimensional model of cultural liberalism can explain this pattern.

The domain specificity suggests that the Röstigraben transmits specific cultural content—attitudes toward gender roles, reproductive autonomy, family policy (see Doepke and Tertilt, 2019, for a framework linking culture to gendered outcomes)—rather than a generic ideological orientation. The francophone cultural ecosystem appears to produce progressive gender norms *and* conservative attitudes on other policy domains (or, equivalently, the germanophone ecosystem produces conservative gender norms and progressive attitudes on non-gender domains). This is consistent with the cultural content being transmitted through language-specific media and intellectual networks, each of which carries domain-specific rather than ideologically coherent content.

This finding invites future research on the *content* of cultural transmission. Why does the Röstigraben produce opposing effects on gender versus non-gender policy? What specific cultural content—literary traditions, media framing, intellectual movements—generates this asymmetry? These questions go beyond modularity but are opened up by the same factorial design.

8.5 Limitations

Our design has five limitations. First, it is not a formal spatial RDD with a continuous running variable; our OLS estimates may capture residual confounding despite the historical predetermination of both boundaries. The within-canton estimates partially address this concern but cannot fully resolve it.

Second, confessional classification is at the cantonal level, which is coarser than ideal. Within a canton, there may be municipality-level variation in religious practice that our binary indicator does not capture. However, the Reformation-era assignment was indeed at the cantonal level, making this the appropriate unit for the historically predetermined treatment.

Third, our panel of six referenda provides reasonable but limited power for the referendum-specific analysis. While the pooled interaction is precisely estimated (SE = 0.83 pp), the individual-referendum interactions have larger standard errors and should be interpreted with appropriate caution. Two referendum-specific interactions survive BH adjustment, but the sign-switching pattern suggests caution in over-interpreting these individual results.

Fourth, we cannot observe individual-level voting behavior, so our estimates reflect municipality-level averages. If within-municipality heterogeneity differs systematically across culture groups, our estimates could mask individual-level interactions. However, given that municipality-level averages aggregate thousands of individual votes, this concern is mitigated by the law of large numbers.

Fifth, we exclude five mixed-confession cantons from the main analysis. While robustness checks including these cantons yield similar results (N = 12,072, interaction = -1.4 pp, $p > 0.1$), the exclusion reduces sample size and geographic coverage.

9. Conclusion

Language and religion—two of the most studied cultural dimensions—combine additively in shaping gender preferences. The interaction is precisely zero: -0.09 pp (permutation $p = 0.94$), with a 95% confidence interval that rules out interactions larger than 1.7 pp in either direction. An additive model predicts French-Catholic gender progressivism at 53.8%; the actual value is 53.7%.

The null interaction coexists with large, precisely estimated main effects. French-speaking municipalities vote 15.5 pp more progressively on gender issues; Catholic heritage reduces progressivism by 8.3 pp. These effects vary dramatically across referenda—the language gap ranges from 1 to 33 pp—but combine additively on average. Individual referenda show significant departures in opposite directions (+3.4 pp in 1981, -4.8 pp in 2020, both surviving

BH adjustment), but these cancel in the pooled specification. A striking falsification reveals that the main effects *reverse* on non-gender issues—French-speaking municipalities are 17.1 pp more conservative—while the interaction remains zero.

The result supports the single-dimension approach that dominates cultural economics: when cultural boundaries are sharp, orthogonal, and historically predetermined, each dimension can be studied in isolation without bias from omitted interactions. But the result also invites caution about extrapolation. Switzerland's cultural boundaries are unusually clean. Whether modularity holds at noisier intersections—ethnolinguistic fractionalization in Africa, caste and religion in India, race and language in the United States—remains an open question. The 2×2 factorial design we employ is portable; we encourage its application elsewhere.

Culture is a system with many dimensions. Those dimensions—at least where the boundaries are sharp and historically fixed—can be studied one at a time.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers

Contributors: @SocialCatalystLab

First Contributor: https://github.com/SocialCatalystLab

References

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *Quarterly Journal of Economics*, 2000, *115* (3), 715–753.

Alesina, Alberto and Paola Giuliano, "Culture and Institutions," *Journal of Economic Literature*, 2015, *53* (4), 898–944.

Basten, Christoph and Frank Betz, "Beyond Work Ethic: Religion, Individual, and Political Preferences," *American Economic Journal: Economic Policy*, 2013, *5* (3), 67–91.

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

Bisin, Alberto and Thierry Verdier, "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, 2001, *97* (2), 298–319.

Boppart, Timo, Josef Falkinger, Volker Grossmann, Ulrich Woitek, and Gabriela Wüthrich, "Under Which Conditions Does Religion Affect Educational Outcomes?," *Explorations in Economic History*, 2013, *50* (2), 242–266.

Cantoni, Davide, "Adopting a New Religion: The Case of Protestantism in 16th Century Germany," *Economic Journal*, 2012, *122* (560), 502–531.

Church, Clive H. and Randolph C. Head, *A Concise History of Switzerland*, Cambridge University Press, 2013.

Crenshaw, Kimberlé, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," *University of Chicago Legal Forum*, 1989, *1989* (1), 139–167.

Doepke, Matthias and Michèle Tertilt, "Does Female Empowerment Promote Economic Development?," *Journal of Economic Growth*, 2019, *24* (4), 309–343.

Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller, "The Demand for Social Insurance: Does Culture Matter?," *Economic Journal*, 2011, *121* (556), F413–F448.

_ , _ , _ , and _ , "Culture, Work Attitudes, and Job Search: Evidence from the Swiss Language Border," *Journal of the European Economic Association*, 2017, *15* (5), 1056–1100.

Fässler, Marius, Rafael Lalive, and Charles Efferson, "Culture Shapes Choices: Evidence from Fertility and Mortality at Swiss Language Borders," *American Economic Journal: Applied Economics*, 2024, *16* (2), 1–39.

Fernández, Raquel, "Does Culture Matter?," in Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds., *Handbook of Social Economics*, Vol. 1A, North-Holland, 2011, pp. 481–510.

Giuliano, Paola and Nathan Nunn, "Understanding Cultural Persistence and Change," *Review of Economic Studies*, 2021, *88* (4), 1541–1581.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "People's Opium? Religion and Economic Attitudes," *Journal of Monetary Economics*, 2003, *50* (1), 225–282.

Matsusaka, John G., "Direct Democracy Works," *Journal of Economic Perspectives*, 2005, *19* (2), 185–206.

Swiss Federal Statistical Office, "Regional Portraits: Communes," 2024.

_ , "swissdd: Swiss Federal and Cantonal Votes," R package 2024.

Young, Alwyn, "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

A. Data Appendix

A.1 Data Sources and Access

Referendum data. Municipality-level referendum results were obtained through the `swissdd` R package (Swiss Federal Statistical Office, 2024b). For each national referendum, the dataset includes: municipality BFS number, vote date, eligible voters, valid votes, yes votes, and no votes. We compute the yes-share as the ratio of yes votes to valid votes.

Municipality characteristics. Language classification and canton assignment come from BFS regional statistics (Swiss Federal Statistical Office, 2024a). For bilingual cantons (Fribourg, Bern, Valais), municipality-level language is assigned using known linguistic boundaries within each canton.

Confessional classification. Historical cantonal religion was coded from Reformation-era choices:

- **Catholic:** LU, UR, SZ, OW, NW, ZG, FR, VS, TI, JU, AI (11 cantons)
- **Protestant:** ZH, BE, GL, BS, BL, SH, AR, VD, NE, GE (10 cantons)
- **Mixed (excluded from main analysis):** AG, GR, SG, SO, TG (5 cantons)

Mixed cantons experienced contested Reformations that left them with substantial populations of both Catholic and Protestant traditions. They are excluded from the main analysis but included in robustness checks, reclassified by their pre-1800 majority denomination.

A.2 Gender Referenda Selection

We identified six gender referenda, matched to `swissdd` by proposal ID:

1. June 14, 1981: Equal rights amendment (ID: 3060)
2. June 13, 1999: Maternity insurance (ID: 4580)
3. March 12, 2000: Women's representation (ID: 4610)
4. June 2, 2002: Abortion access (ID: 4870)
5. September 27, 2020: Paternity leave (ID: 6340)
6. September 26, 2021: Marriage for All (ID: 6470)

A.3 Sample Exclusions

We exclude municipalities on three grounds:

1. **Language:** Italian-speaking municipalities (Ticino, parts of Grisons) are excluded because the Italian-language cultural ecosystem differs from both French and German, and few Italian-speaking municipalities lie in cantons with clear confessional status.

2. **Confession:** Municipalities in five mixed-confession cantons (AG, GR, SG, SO, TG) are excluded from the main analysis. These cantons' confessional heritage is ambiguous, making binary classification unreliable. They are included in the robustness check with the inclusive sample.

3. **Missing data:** Municipality–referendum pairs with missing vote data (due to municipal mergers or recording gaps) are dropped.

A.4 Variable Definitions

Table 8: Variable Definitions

Variable	Definition
`yes_share`	Proportion of valid votes cast "yes." Range: [0, 1].
`is_french`	1 if municipality's dominant language is French, 0 if German.
`is_catholic`	1 if municipality's canton is historically Catholic.
`culture_4`	Factor: German-Protestant (ref.), German-Catholic, French-Protestant, French-Catholic.
`log_eligible`	Log of eligible voters (proxy for municipality size).
`avg_turnout`	Average turnout across referenda (proxy for civic engagement).

B. Additional Figures

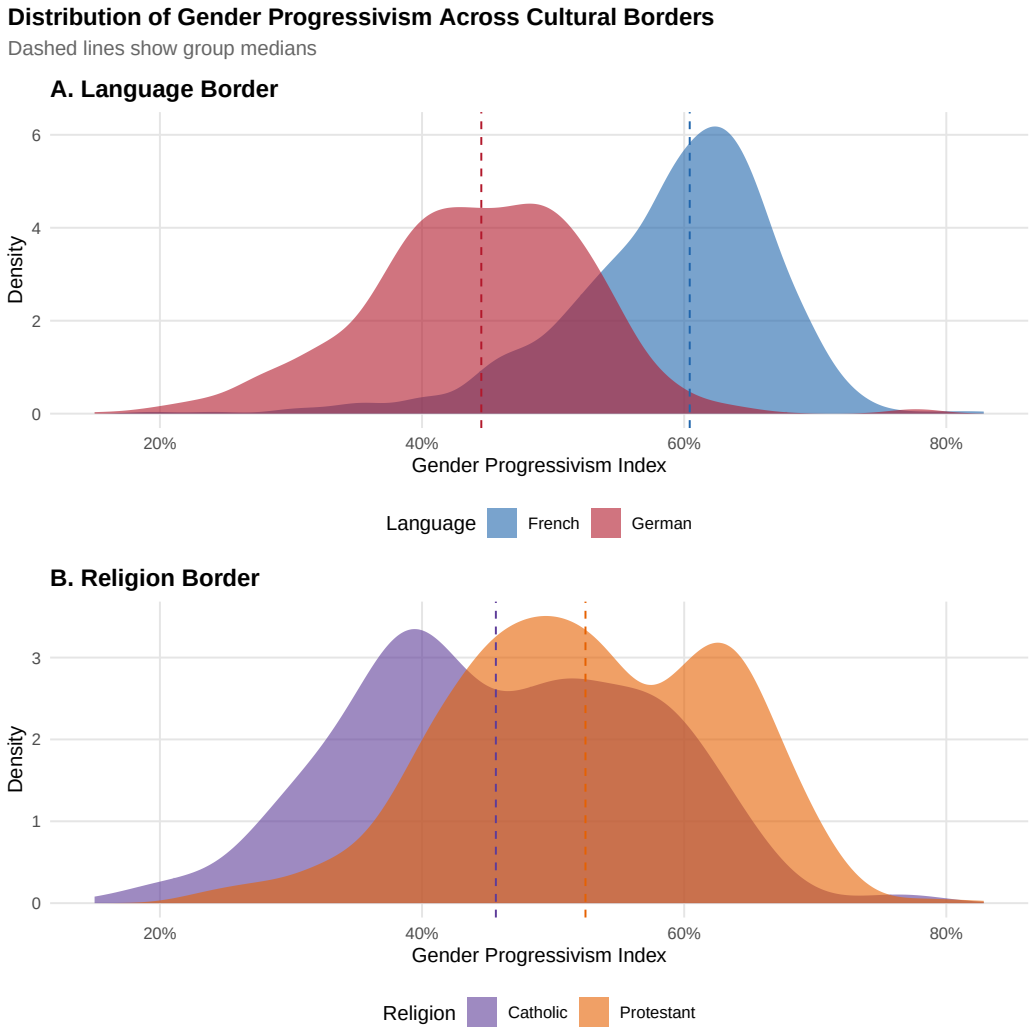

Figure 7: Distribution of gender progressivism index by culture group. French-Protestant municipalities cluster at the top; German-Catholic at the bottom. The distributions overlap substantially within language groups, consistent with religion having a moderate but independent additive effect.

C. Robustness Appendix

C.1 Clustering Sensitivity

The main text reports three clustering levels: municipality, canton, and two-way (municipality and referendum date). With only 21 cantons in the estimation sample (excluding mixed cantons), the effective number of clusters is limited. Canton-level clustering produces larger standard errors but does not change the point estimates. The permutation p-values for the

interaction ($p = 0.936$ at the municipality level, $p = 0.994$ at the canton level) provide the most robust inference, as they are valid regardless of the number of clusters.

C.2 Extended Falsification

The main text falsification (Figure 5) uses a curated set of non-gender referenda selected to span defense, immigration, and fiscal policy—domains where the Röstigraben operates in the opposite direction. That analysis yields a language gap of −17.1 pp, confirming that French-speaking municipalities are conservative on non-gender issues.

Here we conduct an *extended* falsification using a broader, randomly sampled set of non-gender referenda (up to six per decade, stratified by era). This yields a language gap of −0.4 pp (near zero), a religion gap of −3.7 pp, and an interaction of +2.7 pp. The near-zero language gap reflects averaging across heterogeneous non-gender topics—some where the Röstigraben runs strongly conservative, others where it is absent. The key result is that the interaction remains small and statistically insignificant across both the curated and extended falsification samples, confirming that modularity holds regardless of which non-gender referenda are included.

C.3 Inclusive Sample Details

The inclusive sample reclassifies the five mixed-confession cantons (AG, GR, SG, SO, TG) by their pre-1800 majority denomination: AG, SG, SO, and TG are assigned as Protestant (majority at the cantonal level); GR as Protestant (majority of Reformed communities). This yields 2,024 municipalities and 12,072 observations. All main results are robust: language gap = 16.7 pp ($p < 0.001$), religion gap = −6.9 pp ($p < 0.001$), interaction = −1.4 pp ($p > 0.1$).

C.4 By Referendum Topic

The six referenda differ substantially in content and in the magnitude of cultural gaps. The 1999 maternity insurance referendum generates the largest language gap (32.6 pp), likely reflecting the confluence of gender and welfare-state preferences. The 2002 abortion referendum produces the largest religion gap (−20.4 pp), consistent with Catholic doctrine on reproductive rights. The 2021 marriage equality referendum shows the smallest language and religion gaps, suggesting broad cultural convergence on this issue by 2021. Despite these large variations in main effects, the pooled interaction is precisely zero (−0.09 pp). Individual referenda show significant departures that survive BH adjustment (1981: +3.4 pp, $q = 0.031$; 2020: −4.8 pp, $q < 0.001$), but the opposite signs produce a null average, consistent with idiosyncratic campaign effects rather than systematic cultural interaction.

C.5 Power Analysis

Our main specification has 8,727 observations with municipality-level clustering (1,463 clusters). The standard error on the interaction term is 0.83 pp. With this precision, we can detect interactions as small as ± 1.6 pp at the 5% significance level (two-sided). This means we have adequate power to detect interactions that would be economically meaningful: a 1.6 pp interaction would represent roughly 10% of the language gap or 19% of the religion gap. Our null finding is not a power problem.