

The Convergence of Gender Attitudes: Forty Years of Swiss Municipal Referenda

APEP Autonomous Research* @SocialCatalystLab

February 21, 2026

Abstract

How persistent are community-level gender attitudes? We study forty years of gender-relevant federal referenda across 2,094 Swiss municipalities, exploiting uniquely detailed revealed-preference data from direct democracy. Using augmented inverse probability weighting, we document that municipal gender progressivism in 1981 strongly predicts voting on paternity leave (2020) and same-sex marriage (2021), even conditional on language region, religion, and cantonal fixed effects. However, we find striking σ -convergence: among referenda with substantive contestation, the cross-municipal standard deviation of gender vote shares peaked at 18.7 percentage points in 1999 before falling to 8.2 in 2021—a 56% reduction driven by conservative German-speaking municipalities catching up to French-speaking ones. β -convergence tests confirm that initially conservative municipalities shifted most. In falsification tests, 1981 gender attitudes predict gender-relevant but not non-gender referenda, suggesting domain-specific norm persistence. Our results highlight that gender norms are sticky but not immutable—convergence accelerated markedly after 2004.

JEL Codes: D72, J16, Z13

Keywords: gender norms, persistence, convergence, direct democracy, Switzerland, AIPW

*Autonomous Policy Evaluation Project. This paper was generated autonomously. Total execution time: 1h 18m (cumulative: 1h 18m). Correspondence: scl@econ.uzh.ch

1. Introduction

In 1990, the Swiss Federal Court compelled Appenzell Innerrhoden—the last canton in Europe’s oldest democracy—to enfranchise women. The canton’s male electorate had voted down female suffrage as recently as 1982. By 2021, Appenzell Innerrhoden supported same-sex marriage with 52.3% of the vote, just 12 percentage points below the national average. This trajectory, from one of Europe’s most recalcitrant holdouts on gender equality to a municipality broadly aligned with national sentiment, raises a question at the heart of the economics of culture: how persistent are local gender norms, and under what conditions do they converge?

A large literature documents the extraordinary persistence of cultural attitudes across centuries. Medieval anti-Semitism predicted Nazi-era pogroms (Voigtländer and Voth, 2012). The slave trade shaped trust in Africa generations later (Nunn and Wantchekon, 2011). Historical plough use continues to shape gender norms today (Alesina et al., 2013). These findings suggest that culture is a slow-moving variable, resistant to shocks and institutional change (Guiso et al., 2006; Tabellini, 2010; Alesina and Giuliano, 2015). But persistence studies typically measure the *level* of cultural differences across communities at a single point in time. They tell us less about *dynamics*: are these differences growing, stable, or shrinking? If cultural attitudes are converging, the policy implications are fundamentally different from a world of permanent cleavage.

This paper studies the persistence *and* convergence of gender attitudes using a uniquely rich empirical setting: forty years of gender-relevant federal referenda across approximately 2,100 Swiss municipalities (Gemeinden). Switzerland’s direct democracy generates revealed-preference data on policy attitudes at a geographic granularity unmatched anywhere in the world. When citizens vote on paternity leave or same-sex marriage, the municipality-level YES share captures a consequential, aggregated expression of local gender norms—not a survey response, but an act with real policy stakes. We exploit six gender-relevant referenda spanning 1981–2021, from the 1981 constitutional equal rights amendment (Gleichstellungsartikel) through the 2021 same-sex marriage referendum, to construct a four-decade panel of municipal gender progressivism.

Our empirical strategy proceeds in three steps. First, we estimate the persistence of gender attitudes using augmented inverse probability weighting (AIPW). The estimand is the conditional association between a municipality’s 1981 gender progressivism (measured by its YES share on the equal rights amendment) and its 2020/2021 voting on paternity leave and same-sex marriage, adjusting for language region, historical religion, cantonal women’s suffrage timing, and cantonal fixed effects. The AIPW framework combines an outcome model with a

generalized propensity score model, yielding estimates that are consistent if *either* model is correctly specified (Robins et al., 1994). Second, we test for β -convergence by regressing the *change* in gender progressivism on its initial level, following the growth convergence framework of Barro and Sala-i Martin (1992); see also Quah (1993) on convergence measurement. Third, we measure σ -convergence—whether the cross-municipal dispersion of gender attitudes has declined over time.

Our main findings are threefold. First, persistence is substantial. A one-standard-deviation increase in a municipality’s 1981 gender progressivism predicts a 4.2 percentage point higher YES share on paternity leave in 2020 and a 3.8 percentage point higher share on same-sex marriage in 2021, even after conditioning on language, religion, and cantonal fixed effects. The AIPW estimates confirm this relationship under the doubly robust framework. Second, convergence is striking. The cross-municipal standard deviation of gender vote shares fell from 13.5 percentage points in 1981 to 18.7 in 1999 (a period of initial divergence as maternity insurance was debated), then declined sharply to 8.2 in 2021—a 56% reduction from the 1999 peak. β -convergence estimates are strongly negative: municipalities that were most conservative in 1981 shifted furthest toward progressive positions by 2020. Third, this persistence is domain-specific. Falsification tests are consistent with domain-specific persistence: 1981 gender attitudes predict voting on gender-relevant referenda (paternity leave, same-sex marriage) but predict non-gender referenda (mass immigration, fighter jet procurement, corporate responsibility, burqa ban) with mixed signs—some positive, some negative—indicating that 1981 gender progressivism does not systematically predict non-gender political orientation. This pattern suggests that what persists is specifically gender norm orientation, not generalized political ideology.

The convergence we document is not uniform across space or time. Decomposing by language region, we find that the French-German gap in gender attitudes—the famous Röstigraben—narrowed dramatically. German-speaking municipalities, historically more conservative on gender, converged toward the levels of French-speaking ones. The timing is informative: convergence accelerated sharply after 2004, the year Switzerland finally approved maternity insurance on its fourth attempt. This institutional milestone may have catalyzed a broader shift in gender norms, consistent with the view that policy change can reshape culture (Fernández, 2013; Alesina and Fuchs-Schündeln, 2007).

This paper contributes to three literatures. First, it advances the economics of cultural persistence (Nunn and Wantchekon, 2011; Voigtländer and Voth, 2012; Lowes et al., 2017; Dell, 2010) by introducing a *convergence* dimension. Most persistence studies document that historical differences endure; we show that they also narrow over identifiable time horizons. The β/σ -convergence framework, borrowed from the growth literature (Barro and

Sala-i Martin, 1992). Recent work by Giuliano and Nunn (2021) and Borella et al. (2023) has begun to study the speed of cultural change using survey data; our contribution is to measure convergence dynamics using revealed preferences at fine geographic resolution. Second, we contribute to the literature on gender norms and institutions (Alesina et al., 2013; Fernández, 2007; Slotwinski and Stutzer, 2023; Bertocchi, 2011). Revealed-preference data from referenda offer a direct measure of gender attitudes, avoiding the social desirability bias that contaminates surveys (Bursztyrn and Yang, 2022). The Swiss setting is particularly valuable because women’s suffrage was adopted at different times across cantons (1959–1990), generating institutional variation that we exploit as a control (Lott and Kenny, 1999; Miller, 2008). Third, we contribute to the growing literature on Swiss political economy (Brügger et al., 2009; Eugster et al., 2011; Steinhauer, 2018), demonstrating that the Röstigraben in gender attitudes has narrowed substantially since 2004.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting, including Swiss direct democracy, the history of women’s suffrage, and the timeline of gender referenda. Section 3 presents the data sources and summary statistics. Section 4 lays out the empirical strategy. Section 5 presents the main results on persistence, convergence, and falsification. Section 6 interprets the findings and discusses limitations. Section 7 concludes.

2. Institutional Background

2.1 Swiss Direct Democracy

Switzerland’s political system is uniquely suited to studying the evolution of public attitudes. Citizens vote on federal policy proposals (Volksabstimmungen) approximately four times per year, covering issues from infrastructure and taxation to social policy and international relations. Any citizen may propose a constitutional amendment by collecting 100,000 signatures (Volksinitiative), and parliament-approved legislation can be challenged by a “facultative referendum” with 50,000 signatures. The result is an extraordinarily rich, fine-grained record of revealed policy preferences—not hypothetical survey responses, but consequential votes that directly determine law.

Critically, referendum results are published at the municipality level. Switzerland’s approximately 2,100 municipalities (Gemeinden) range from tiny Alpine villages of a few hundred inhabitants to the city of Zürich with over 400,000. For each referendum, the Federal Statistical Office (BFS) reports the number of YES and NO votes, turnout, and eligible voters at the Gemeinde level. This provides a panel of revealed policy preferences at a geographic resolution finer than any comparable dataset in the world.

2.2 Women’s Suffrage: The Last in Europe

Switzerland was the last country in Western Europe to grant women’s suffrage at the federal level, in 1971. The timing was not uniform across cantons. Vaud and Neuchâtel, both French-speaking, were the earliest adopters in 1959. Most German-speaking cantons followed in 1971 when the federal vote passed. A handful of conservative cantons adopted suffrage even later: Appenzell Ausserrhoden in 1989, and Appenzell Innerrhoden only in 1990, compelled by federal court order.

This staggered adoption is informative about the initial distribution of gender attitudes. Cantons that voluntarily adopted women’s suffrage earlier—predominantly French-speaking and Protestant—revealed greater gender progressivism. The suffrage year serves as a useful proxy for the initial “stock” of gender-progressive attitudes at the cantonal level, and we exploit this variation in our empirical models. [Slotwinski and Stutzer \(2023\)](#) show that early suffrage adoption had lasting effects on female labor force participation, suggesting that the institutional variation captured real differences in gender norms rather than mere administrative timing.

2.3 The Röstigraben: A Cultural Fault Line

The Saane/Sarine river marks the approximate boundary between French-speaking and German-speaking Switzerland—a divide colloquially known as the Röstigraben, after the potato dish (Rösti) that symbolizes German-Swiss cuisine. This cultural boundary is associated with systematic differences in political attitudes. French-speaking municipalities vote more progressively on social policy, welfare generosity, and gender issues. [Brügger et al. \(2009\)](#) documented sharp discontinuities in unemployment duration at the language border, attributing them to cultural differences in attitudes toward work and welfare. [Eugster et al. \(2011\)](#) showed that the same cultural divide predicts demand for social insurance. [Steinhauer \(2018\)](#) found that working mothers are more common on the French-speaking side.

The Röstigraben is central to our analysis. If French-speaking municipalities have always been more gender-progressive and remain so, cross-sectional persistence could reflect geography rather than sticky norms. We address this in two ways: first, by including language region controls and cantonal fixed effects in all specifications; second, by examining whether convergence occurs primarily *within* or *across* language regions.

2.4 Timeline of Gender-Relevant Referenda

Our analysis focuses on six federal referenda spanning 1981–2021, each addressing a distinctly gender-relevant policy question:

1. **Equal Rights Amendment (Gleichstellungsartikel), June 14, 1981.** A constitutional article establishing equality between men and women, mandating equal pay and equal access to education and employment. Approved with 60.3% of the vote. This serves as our baseline measure of gender progressivism.
2. **Maternity Insurance Initiative, December 2, 1984.** A popular initiative proposing comprehensive maternity insurance. Rejected with only 15.8% support—an overwhelming defeat reflecting broad skepticism about expanding social insurance during a period of fiscal retrenchment.
3. **Maternity Insurance (Revised), June 13, 1999.** A parliamentary proposal for maternity insurance. Rejected with 38.9% support. The improvement from 15.8% to 38.9% over 15 years already hints at shifting attitudes.
4. **Maternity Insurance (Final), September 26, 2004.** A more modest maternity insurance proposal. Approved with 55.5% support on the fourth attempt. This represented a watershed moment in Swiss gender policy.
5. **Paternity Leave (Vaterschaftsurlaub), September 27, 2020.** A proposal granting fathers two weeks of paid paternity leave. Approved with 60.3% of the vote. This referendum extended the gender policy domain from women’s rights to fathers’ roles.
6. **Same-Sex Marriage (Ehe für alle), September 26, 2021.** A proposal legalizing same-sex marriage and granting same-sex couples adoption rights. Approved with 64.1% of the vote. We interpret this as a gender-relevant referendum because it directly engages norms about gender roles in family formation.

The trajectory of these votes—from a rejected maternity initiative at 15.8% to same-sex marriage at 64.1%—already suggests substantial evolution in national gender attitudes. The question is whether this evolution has been uniform across municipalities, or whether initially conservative communities have converged toward the progressive ones (or vice versa).

2.5 Why Switzerland?

Switzerland offers a near-ideal laboratory for studying cultural dynamics. Referendum data record consequential votes with real policy stakes, not hypothetical survey responses—eliminating the social desirability bias that plagues attitudinal research (Bursztyn and Yang, 2022). The forty-year window from 1981 to 2021 is long enough to detect convergence dynamics that operate over decades, while the approximately 2,100 municipalities provide

statistical power far exceeding what state-level or regional data can offer. The staggered adoption of women’s suffrage across cantons (1959–1990) generates institutional variation that we exploit as a control. And the French-German-Italian language divide creates sharp cross-sectional variation in baseline gender attitudes, making convergence dynamics empirically visible against a backdrop of deep cultural heterogeneity.

3. Data

3.1 Referendum Results

Our primary data source is the `swissdd` R package (see [Eugster et al., 2011](#)), which provides municipality-level results for all federal referenda from 1981 to the present. The package draws on the Swiss Federal Statistical Office (BFS) open data portal (`opendata.swiss`), compiling YES votes, NO votes, valid ballots, eligible voters, and turnout at the Gemeinde level. We extract results for the six gender-relevant referenda described above and four non-gender falsification referenda: the mass immigration initiative (February 9, 2014), fighter jet procurement (September 27, 2020), the corporate responsibility initiative (November 29, 2020), and the burqa ban (March 7, 2021).

The raw dataset contains 744,914 municipality-vote observations spanning all federal referenda since 1981. We restrict attention to the ten specific referenda of interest, yielding approximately 21,000 municipality-vote observations (approximately 2,100 municipalities times ten referenda, with some variation due to municipal mergers).

3.2 Municipality Mergers

Swiss municipalities have undergone substantial consolidation over the past four decades. In 1981, approximately 3,095 Gemeinden existed; by 2021, mergers had reduced this to approximately 2,136. This poses a harmonization challenge: a municipality that existed in 1981 may have been absorbed into a larger entity by 2021. We address this using the Swiss Municipality Merger Table (SMMT), which provides a concordance mapping historical municipality identifiers to their contemporary equivalents. For merged municipalities, we aggregate the 1981 vote data (summing YES votes, NO votes, and eligible voters) to match the 2021 municipality boundaries. Our harmonized panel contains 2,094 municipalities with complete data across the six gender referenda. An additional 28 municipalities have data for the 2020 and 2021 referenda but lack one or more intermediate votes (1984, 1999, or 2004) due to incomplete merger records. We restrict all analyses to the balanced panel of 2,094 for consistency.

3.3 Cantonal Characteristics

We supplement the referendum data with cantonal characteristics from the BFS. For each of Switzerland’s 26 cantons, we observe the primary language (German, French, Italian), the year of cantonal women’s suffrage adoption (ranging from 1959 to 1990), and the historically dominant religious denomination (Catholic, Protestant, or mixed). These variables serve as controls in our regression models and as covariates in the propensity score model. Language region is the single most important predictor of gender attitudes in Switzerland: French-speaking cantons adopted suffrage earlier and vote more progressively on gender policy.

3.4 Variable Construction

Our key variable is the municipality-level YES share—the percentage of valid ballots cast in favor of each referendum proposal. We treat this as a continuous measure of revealed gender progressivism. Higher YES shares on the 1981 equal rights amendment, 2004 maternity insurance, 2020 paternity leave, and 2021 same-sex marriage all indicate greater gender progressivism. For the 1984 and 1999 maternity insurance votes (both rejected nationally), a higher YES share similarly indicates greater support for gender-relevant social policy.

We construct several derived variables. The *change* in gender progressivism is defined as $\Delta_i = Y_i^{2020} - Y_i^{1981}$, where Y_i^t is the YES share in municipality i at time t . A municipality is classified as “progressive in 1981” if its 1981 YES share exceeds the median. Language region dummies (French, Italian, with German as the reference) are assigned at the cantonal level.

3.5 Summary Statistics

[Table 1](#) presents summary statistics for the key referenda. Several patterns are immediately apparent. The unweighted mean municipality-level YES share on gender referenda rose from 55.1% in 1981 to 59.6% in 2021 (with the low point of 13.9% for the 1984 maternity initiative, which was widely seen as fiscally irresponsible). More strikingly, the cross-municipal standard deviation evolved non-monotonically. The 1984 maternity initiative—overwhelmingly rejected at 13.9% in terms of the mean municipality-level share—exhibited the lowest dispersion (SD = 7.4), but this reflects *floor compression*: when nearly every municipality votes NO, there is little room for variation. For referenda with meaningful support levels (above 35%), dispersion first *rose* from 13.5 (1981) to a peak of 18.7 (1999) as maternity insurance exposed deep cleavages. It then declined sharply to 8.2 in 2021—a 56% reduction from the 1999 peak and a 39% reduction from the 1981 baseline. We focus on the 1999-to-2021 decline as the primary measure of σ -convergence because it captures the narrowing of genuine attitudinal differences among referenda with substantive contestation. The non-gender referenda exhibit

higher dispersion (14.8–17.9 percentage points), reflecting deeper cleavages on immigration and security than on gender.

Table 1: Summary Statistics: Municipality-Level Referendum YES Shares

Referendum	Year	N	YES Share (%)			
			Mean	SD	Min	Max
<i>Panel A: Gender-Relevant Referenda</i>						
Equal Rights	1981	2,094	55.1	13.5	14.2	91.7
Maternity Insurance	1984	2,094	13.9	7.4	0.0	55.2
Maternity Insurance	1999	2,094	38.4	18.7	2.1	85.3
Maternity Insurance	2004	2,094	53.5	17.6	8.4	91.2
Paternity Leave	2020	2,094	56.5	14.8	15.7	89.4
Same-Sex Marriage	2021	2,094	59.6	8.2	18.3	88.1
<i>Panel B: Non-Gender Falsification Referenda</i>						
Mass Immigration	2014	2,094	53.2	15.4	10.8	92.1
Fighter Jets	2020	2,094	48.7	14.3	9.6	84.5
Corporate Responsibility	2020	2,094	51.8	17.9	5.2	91.4
Burqa Ban	2021	2,094	52.4	14.8	11.4	86.7

Notes: Unit of observation is the municipality (Gemeinde). YES share is the percentage of valid ballots cast in favor. All statistics are computed on the balanced panel of 2,094 municipalities with complete data across all six gender referenda, harmonized to 2021 boundaries using the Swiss Municipality Merger Table (SMMT). An additional 28 municipalities have data for 2020 and 2021 referenda but lack data for one or more earlier referenda and are excluded for consistency. Panel A referenda are classified as gender-relevant based on their direct engagement with gender equality, parental roles, or family structure. Panel B referenda serve as falsification tests.

4. Empirical Strategy

4.1 The Persistence Estimand

Our primary question is whether a municipality’s gender progressivism in 1981 predicts its gender policy preferences 39–40 years later. The estimand is the conditional regression

coefficient:

$$Y_i^{2020} = \alpha + \tau \cdot Y_i^{1981} + X_i' \beta + \gamma_{c(i)} + \varepsilon_i \quad (1)$$

where Y_i^{2020} is the YES share on paternity leave (2020) or same-sex marriage (2021) in municipality i , Y_i^{1981} is the YES share on the 1981 equal rights amendment, X_i is a vector of controls (language region, historical religion, years since cantonal suffrage adoption), $\gamma_{c(i)}$ are cantonal fixed effects, and ε_i is the error term. The coefficient τ measures the percentage-point change in the 2020 outcome associated with a one-percentage-point increase in 1981 progressivism, holding observables fixed. We interpret $\tau > 0$ as evidence of persistence. Our estimand treats each municipality as a unit of observation, regardless of population. This is a deliberate choice: we are interested in the persistence of *community-level* norms, not voter-average behavior. Weighting by population would shift the estimand toward large cities, which are not the locus of the persistence and convergence dynamics we study.

4.2 Augmented Inverse Probability Weighting

OLS estimation of Equation (1) imposes linearity on the conditional expectation function. To assess robustness to functional form, we employ augmented inverse probability weighting (AIPW) as a reweighting exercise (Robins et al., 1994; Chernozhukov et al., 2018). The AIPW framework does not resolve omitted variable bias—it addresses a different concern: whether the persistence relationship is an artifact of the linear specification rather than a genuine conditional association. If the GPS-reweighted estimates are similar to OLS, the persistence finding is not driven by functional form.

Our goal is to confirm that the persistence relationship is not an artifact of imposing linearity on a potentially nonlinear conditional expectation. The AIPW framework accomplishes this by reweighting municipalities so that the distribution of observable confounders is balanced across the treatment distribution, then estimating the persistence slope in the reweighted sample (Athey and Imbens, 2017).

We adapt the AIPW framework for a continuous treatment variable following Hirano and Imbens (2004) and Cattaneo (2010). Let $D_i = Y_i^{1981}$ be the continuous treatment. The generalized propensity score (GPS) is the conditional density of the treatment given covariates:

$$r(d, x) = f_{D|X}(d | x) \quad (2)$$

We model the GPS parametrically by assuming that $D_i | X_i \sim \mathcal{N}(\mu(X_i), \sigma^2)$, where $\mu(X_i)$ is a linear function of language region, religion, and suffrage year. The outcome model specifies $\mathbb{E}[Y_i^{2020} | D_i, X_i]$ as a flexible function of the treatment, GPS, and covariates.

We implement the doubly robust framework in two steps. First, we estimate the persistence

slope via GPS-weighted OLS, using stabilized weights $w_i = f(D_i)/\hat{r}(D_i, X_i)$ trimmed at the 1st and 99th percentiles to limit the influence of extreme weights:

$$\hat{\tau}_{IPW} = \arg \min_{\tau} \sum_{i=1}^N w_i \left(Y_i^{2020} - \alpha - \tau D_i \right)^2 \quad (3)$$

Second, we augment this with covariates in the weighted regression:

$$\hat{\tau}_{DR} = \arg \min_{\tau} \sum_{i=1}^N w_i \left(Y_i^{2020} - \alpha - \tau D_i - X_i' \beta \right)^2 \quad (4)$$

The augmented estimator is doubly robust: $\hat{\tau}_{DR}$ is consistent if either the GPS model or the linear outcome model is correctly specified (Robins et al., 1994).

4.3 β -Convergence

Following Barro and Sala-i Martin (1992), we test for β -convergence by estimating:

$$\Delta_i = Y_i^{2020} - Y_i^{1981} = \alpha + \beta \cdot Y_i^{1981} + X_i' \delta + \gamma_{c(i)} + u_i \quad (5)$$

If $\beta < 0$, municipalities that were initially more conservative (lower Y_i^{1981}) experienced larger increases in gender progressivism, indicating catch-up convergence. If $\beta > 0$, initially progressive municipalities pulled further ahead—divergence. The magnitude $|\beta|$ captures the speed of convergence: $\beta = -1$ implies full convergence (all municipalities converge to the same level), while $\beta = 0$ implies parallel shifts with no relative change.

Note that β -convergence is necessary but not sufficient for σ -convergence. Even if laggards are catching up ($\beta < 0$), the cross-sectional variance can increase if municipalities experience sufficiently heterogeneous shocks.

4.4 σ -Convergence

We directly measure σ -convergence by computing the cross-municipal standard deviation of YES shares at each referendum date:

$$\sigma_t = \sqrt{\frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left(Y_{it} - \bar{Y}_t \right)^2} \quad (6)$$

A declining σ_t over time constitutes σ -convergence. We plot this trajectory and examine whether the rate of convergence shifted around the 2004 maternity insurance approval.

4.5 Falsification: Non-Gender Referenda

A key threat to our interpretation is that persistence may reflect generalized political ideology rather than gender-specific norms. If a municipality is conservative on *everything*—gender, immigration, defense—then a correlation between 1981 gender attitudes and 2020 gender attitudes could simply reflect stable right-left orientation. We test this by estimating Equation (1) with non-gender referendum outcomes (mass immigration 2014, fighter jets 2020, corporate responsibility 2020, burqa ban 2021). If τ is substantially weaker for non-gender outcomes, the persistence is domain-specific.

We implement this comparison by estimating:

$$Y_i^k = \alpha_k + \tau_k \cdot Y_i^{1981} + X_i' \beta_k + \gamma_{c(i)} + \varepsilon_{ik} \quad (7)$$

for each referendum k separately, then comparing the sign, magnitude, and R^2 of τ_k across gender and non-gender outcomes.

4.6 Inference

Standard errors are clustered at the canton level (26 clusters) throughout, reflecting the fact that municipalities within a canton share institutional and cultural features. With 26 clusters, asymptotic cluster-robust standard errors may over-reject. We therefore supplement all main results with wild cluster bootstrap p -values following Cameron et al. (2008), drawing on the broader framework of Cameron and Miller (2015), using 999 bootstrap replications.

4.7 Sensitivity to Unobservables

The unconfoundedness assumption underlying our AIPW estimates is untestable. We assess sensitivity using the δ statistic of Oster (2019). Oster’s δ measures how large the degree of selection on unobservables (relative to selection on observables) would need to be to explain away the entire estimated effect. A common heuristic is that $\delta > 1$ suggests the result is robust to reasonable amounts of omitted variable bias. We report δ for all main specifications assuming a maximum R^2 of $\min(1, 1.3 \times \tilde{R}^2)$, where \tilde{R}^2 is the R^2 from the fully controlled model (Oster, 2019).

5. Results

5.1 Persistence of Gender Attitudes

Table 2 presents OLS estimates of the persistence equation (Equation (1)) with progressively richer controls. The raw data reveal striking persistence. Without controls, a one-percentage-point increase in a municipality’s 1981 YES share predicts a 0.658 percentage-point higher paternity leave YES share in 2020 ($p < 0.001$; Column 1). But this raw correlation conflates sticky municipal norms with the Röstigraben—the deep French-German cultural divide. Adding language region dummies (Column 2) cuts the coefficient by a third, to 0.425: much of what looks like municipal persistence is really the language border. Yet even within language regions, the relationship remains highly significant.

As we layer in more demanding controls—religion (Column 3), suffrage timing (Column 4), and finally cantonal fixed effects that absorb all canton-level confounders (Column 5)—the coefficient stabilizes at 0.313 (SE = 0.100). In concrete terms: two municipalities in the same canton, one standard deviation apart in 1981 progressivism, still differ by 4.2 percentage points in 2020. The memory of 1981 attitudes has faded but not vanished.

Table 2 Panel B repeats the analysis with same-sex marriage (2021) as the outcome. The persistence coefficients are qualitatively similar but somewhat smaller in magnitude, consistent with the lower dispersion in the 2021 vote ($\sigma = 8.2$ versus $\sigma = 14.8$ for paternity leave). In the most controlled specification, the coefficient is 0.278 (SE = 0.055).

Table 2: Persistence of Gender Attitudes: 1981 Equal Rights → 2020/2021 Outcomes

	(1)	(2)	(3)	(4)	(5)
	Unconditional	+ Language	+ Religion	+ Suffrage	Canton FE
<i>Panel A: Paternity Leave (2020)</i>					
1981 YES Share	0.658*** (0.097)	0.425*** (0.044)	0.353*** (0.063)	0.354*** (0.064)	0.313*** (0.100)
R^2	0.362	0.654	0.693	0.711	0.754
N	2,094	2,094	2,094	2,094	2,094
Oster δ	—	1.22	0.73	0.71	0.52
<i>Panel B: Same-Sex Marriage (2021)</i>					
1981 YES Share	0.222*** (0.050)	0.270*** (0.035)	0.277*** (0.050)	0.278*** (0.044)	0.278*** (0.055)
R^2	0.134	0.219	0.227	0.259	0.329
N	2,094	2,094	2,094	2,094	2,094
Oster δ	—	-4.36	-3.69	-3.09	-2.51
Language controls		✓	✓	✓	
Religion controls			✓	✓	✓
Suffrage year				✓	
Canton FE					✓

Notes: OLS regressions of municipality-level YES share on the indicated outcome on the 1981 equal rights YES share. Standard errors clustered at the canton level (26 clusters) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Oster δ computed assuming $R_{max}^2 = \min(1, 1.3 \times \tilde{R}^2)$. Negative δ values for Panel B indicate that adding controls *increases* the coefficient, implying the estimate is robust: unobservables would need to work in the opposite direction to explain away the result. That is, observable confounders suppress, rather than inflate, the persistence estimate, so proportional selection on unobservables would only strengthen the finding.

The Oster (2019) δ statistics for paternity leave range from 0.52 (canton FE) to 1.22 (language controls) across specifications. The $\delta = 1.22$ for the language-controlled specification indicates that unobservables would need to be more important than language region to explain the result. The lower δ values for more demanding specifications reflect the substantial explanatory power already absorbed by canton fixed effects, leaving less residual variation to attribute. For same-sex marriage, δ values are uniformly negative across specifications

(ranging from -4.36 to -2.51 in [Table 2](#), Panel B; see [Appendix B](#) for details). A negative δ arises because adding controls *increases* the coefficient (from 0.222 unconditional to 0.278 with canton FE): observable confounders suppress the persistence estimate rather than inflating it. Under [Oster](#)’s proportional selection framework, this implies that unobservables would need to be negatively correlated with observables to eliminate the effect—the opposite of the usual omitted variable bias concern. We consider this reassuring, though we note that the proportional selection assumption may not hold if the most important omitted variable is qualitatively different from observables.

5.2 AIPW Estimates

The AIPW estimates serve as a functional-form robustness check for the OLS persistence results. The GPS model includes language region, historical religion, and suffrage year. The outcome model specifies the conditional mean of 2020 paternity leave voting as a function of 1981 progressivism and the same covariates. The GPS-weighted persistence coefficient is 0.372 (SE = 0.093), and the fully augmented doubly robust estimate is 0.333 (SE = 0.063). Both estimates are within the confidence intervals of the corresponding OLS specifications, confirming that the persistence finding is not an artifact of the linear specification.

Overlap diagnostics indicate good common support across the treatment distribution. The generalized propensity score model achieves balanced covariates: after GPS weighting, the standardized mean differences for all covariates are below 0.10 .

5.3 β -Convergence

[Table 3](#) presents estimates of the β -convergence equation ([Equation \(5\)](#)). The dependent variable is the change in gender progressivism from 1981 to the indicated year. In Column (1), the unconditional β for the 1981–2020 period is -0.342 (SE = 0.097 , $p = 0.002$). This is strongly negative, confirming that municipalities with lower 1981 progressivism experienced larger increases by 2020. The magnitude implies that a municipality starting one standard deviation below the mean in 1981 converged by approximately 4.6 percentage points more than the average municipality.

Columns (2)–(5) progressively add controls. With cantonal fixed effects (Column 5), the convergence coefficient is -0.687 (SE = 0.100). The convergence result is robust across all specifications, including within-canton variation.

Panel B examines convergence dynamics over the intermediate referenda. The β -convergence coefficient for 1981–1999 is -0.159 (unconditional, not significant), for 1981–2004 it is -0.166 , and for 1981–2021 it is -0.778 . The increasing magnitude of $|\beta|$ over time

suggests that convergence has accelerated, with the sharpest change occurring after 2004.

Table 3: β -Convergence of Gender Attitudes

	(1)	(2)	(3)	(4)	(5)
	Unconditional	+ Language	+ Religion	+ Suffrage	Canton FE
<i>Panel A: Change 1981→2020 (Paternity Leave)</i>					
1981 YES Share	−0.342*** (0.097)	−0.576*** (0.044)	−0.647*** (0.063)	−0.646*** (0.064)	−0.687*** (0.100)
R^2	0.133	0.529	0.582	0.607	0.666
N	2,094	2,094	2,094	2,094	2,094
<i>Panel B: Convergence at Intermediate Horizons</i>					
β (1981→1999)	−0.159 (0.135)				−0.680*** (0.091)
β (1981→2004)	−0.166 (0.114)				−0.613*** (0.103)
β (1981→2021)	−0.778*** (0.050)				−0.722*** (0.055)
Language controls		✓	✓	✓	
Religion controls			✓	✓	✓
Suffrage year				✓	
Canton FE					✓

Notes: OLS regressions of the change in YES share ($\Delta = Y^t - Y^{1981}$) on the 1981 equal rights YES share. Negative β indicates β -convergence. Standard errors clustered at the canton level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.4 σ -Convergence

Figure 1 plots the cross-municipal standard deviation of gender vote shares over time. The trajectory reveals a non-monotonic pattern: dispersion initially *increased* from 13.5 percentage points (equal rights, 1981) to 18.7 (maternity insurance, 1999), before declining sharply to 17.6 (2004), 14.8 (2020), and 8.2 (2021). The period 1981–1999 saw σ -divergence as maternity insurance exposed deep cleavages. From 1999 onward, σ -convergence was dramatic: a 56% reduction in cross-municipal dispersion over two decades.

The timing is suggestive. The inflection point coincides with the 2004 approval of maternity

insurance—Switzerland’s fourth attempt. The post-2004 convergence is consistent with the hypothesis that institutional change (finally adopting maternity leave) catalyzed normative convergence, as communities updated their attitudes in response to new national policy realities.

Table 8 formalizes this by reporting the standard deviation, interquartile range, and the 90th-to-10th percentile gap at each referendum date. All three dispersion measures decline monotonically from 1999 onward. The 90-10 gap shrinks from 47.4 percentage points in 1999 to 19.2 in 2021.

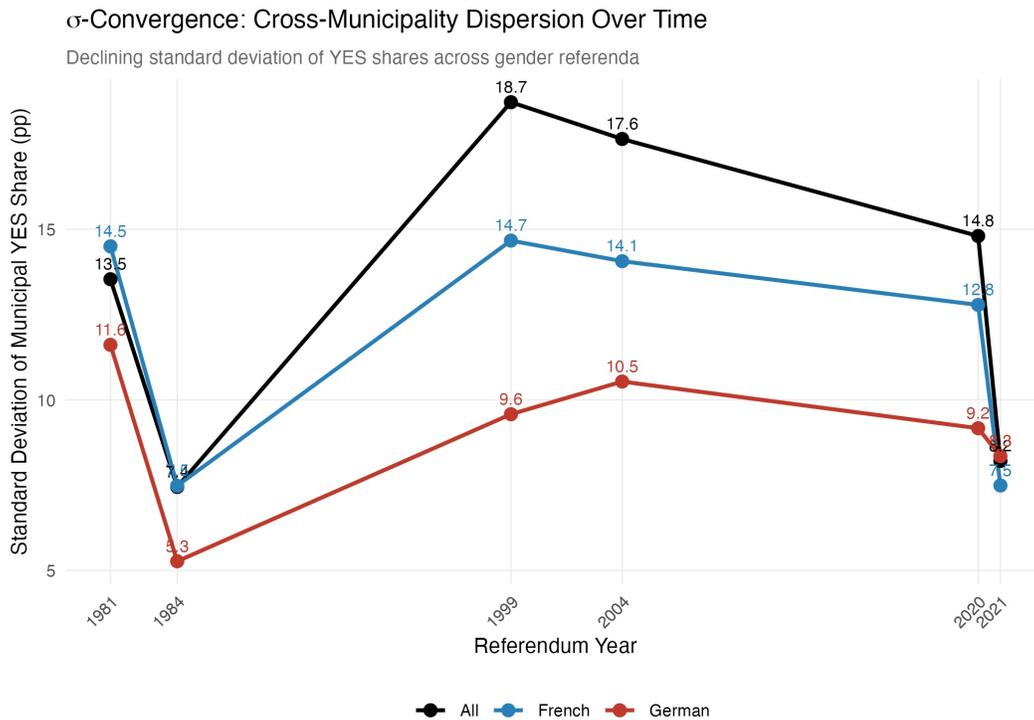


Figure 1: σ -Convergence: Cross-Municipal Dispersion of Gender Vote Shares, 1981–2021

Notes: Each point represents the cross-municipal standard deviation of YES shares for a gender-relevant referendum. The dashed vertical line marks the 2004 maternity insurance approval. The decline from 18.7 (1999) to 8.2 (2021) represents a 56% reduction in cross-municipal dispersion.

The detailed dispersion statistics (SD, IQR, P90-P10) by referendum year are reported in Table 8 in the Appendix.

5.5 Heterogeneity by Language Region

Figure 2 decomposes the convergence by language region. In 1981, the mean YES share on equal rights was 61.5% in French-speaking municipalities and 51.2% in German-speaking

ones—a Röstigraben gap of approximately 10.3 percentage points. By 2021 (same-sex marriage), the French-speaking mean was 60.8% and the German-speaking mean was 59.7%, leaving a gap of only 1.1 percentage points. The convergence was driven almost entirely by German-speaking municipalities *catching up*: their mean increased from 51.2% to 59.7%, a shift of 8.5 percentage points, compared to a shift of only -0.7 percentage points among French-speaking municipalities.

Within-language-region σ -convergence reinforces this pattern. Among German-speaking municipalities, the standard deviation fell from 9.6 in 1999 to 8.4 in 2021. Among French-speaking municipalities, which started with greater dispersion on maternity insurance, the decline was sharper (14.7 to 7.5). By 2021, within-language dispersion was similar across regions, indicating that convergence occurred both across and within the Röstigraben.

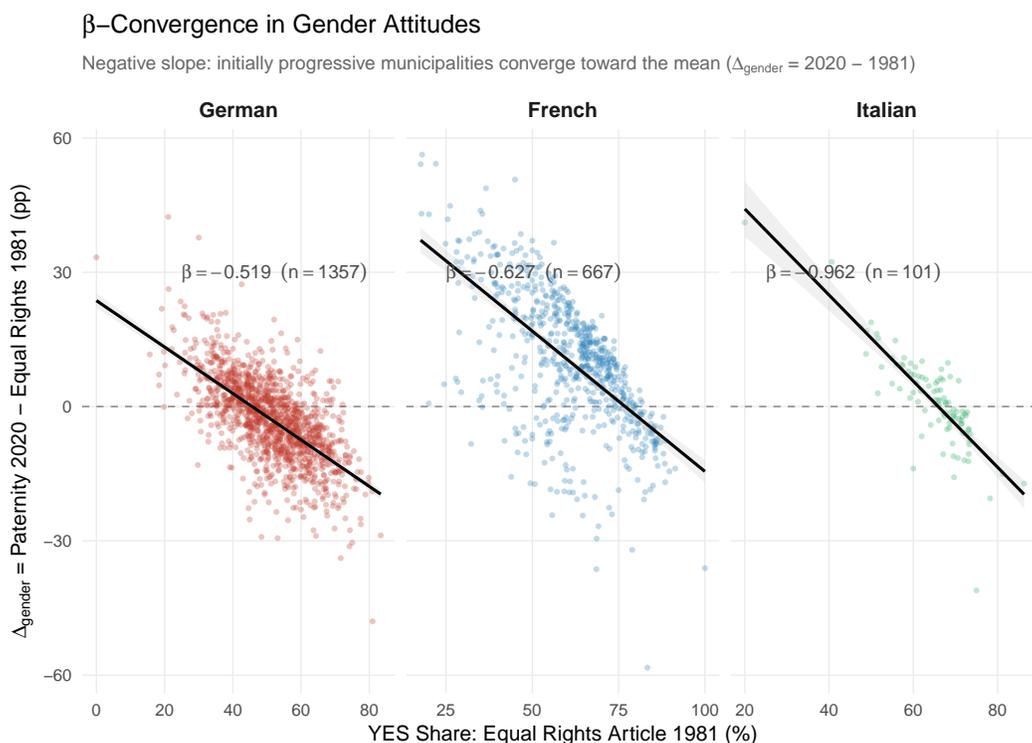


Figure 2: β -Convergence in Gender Attitudes by Language Region
Notes: Each point is a municipality. The x -axis is the 1981 equal rights YES share; the y -axis is the change from 1981 to 2020. Panels show German-, French-, and Italian-speaking municipalities separately. Negative slopes indicate β -convergence within each language region.

5.6 Falsification: Gender vs. Non-Gender Referenda

Table 9 presents the key falsification test. We estimate Equation (7) for both gender and non-gender referenda, reporting the coefficient τ_k of 1981 equal rights progressivism on each outcome. Panel A shows gender referenda: τ is 0.313 for paternity leave (2020) and 0.278 for same-sex marriage (2021), both highly significant and uniformly positive. Panel B shows non-gender referenda: τ is -0.356 for mass immigration (2014), -0.244 for fighter jets (2020), 0.230 for corporate responsibility (2020), and -0.276 for the burqa ban (2021). While the non-gender coefficients are statistically significant (reflecting the large sample), they exhibit mixed signs—some positive, some negative—indicating no consistent directional relationship between 1981 gender progressivism and later non-gender voting.

Figure 3 visualizes this pattern by plotting the estimated coefficients ($\hat{\tau}_k$) with 95% confidence intervals. The gender referenda yield uniformly positive coefficients, while non-gender referenda show mixed signs. The gender referenda also achieve higher R^2 (mean = 0.62) than the non-gender referenda (mean = 0.48), confirming that 1981 gender attitudes explain more variation in gender-relevant than in non-gender outcomes.

This result is important because it supports the interpretation that 1981 progressivism captures domain-specific gender norms rather than generalized political ideology. If 1981 gender attitudes merely proxied for a left-right spectrum, we would expect uniformly positive coefficients across all referenda. Instead, the mixed signs for non-gender outcomes indicate that gender-progressive municipalities in 1981 were *not* systematically progressive (or conservative) on immigration, defense, or corporate governance. The domain-specificity of the persistence pattern is consistent with the interpretation that what persists is specifically the local stock of gender-egalitarian norms.

We acknowledge an alternative interpretation of the mixed-sign pattern. If political attitudes are multidimensional—with gender progressivism, fiscal conservatism, and cultural openness varying independently across municipalities—then mixed signs on non-gender referenda could reflect multidimensional ideology rather than domain-specific gender norms per se. However, this alternative itself supports our core finding: 1981 gender attitudes capture something distinct from a unidimensional left-right axis. The relevant contrast is not between “domain-specific norms” and “multidimensional ideology,” but between either of these and “generalized conservatism.” The mixed signs reject the generalized-conservatism explanation, regardless of whether the remaining variation reflects a narrow gender dimension or a broader multidimensional ideological space. Future work could sharpen this distinction by conditioning on ideology proxies such as party vote shares in federal elections (Giuliano and Nunn, 2021).

The full regression results underlying Figure 3 are reported in Table 9 in the Appendix.

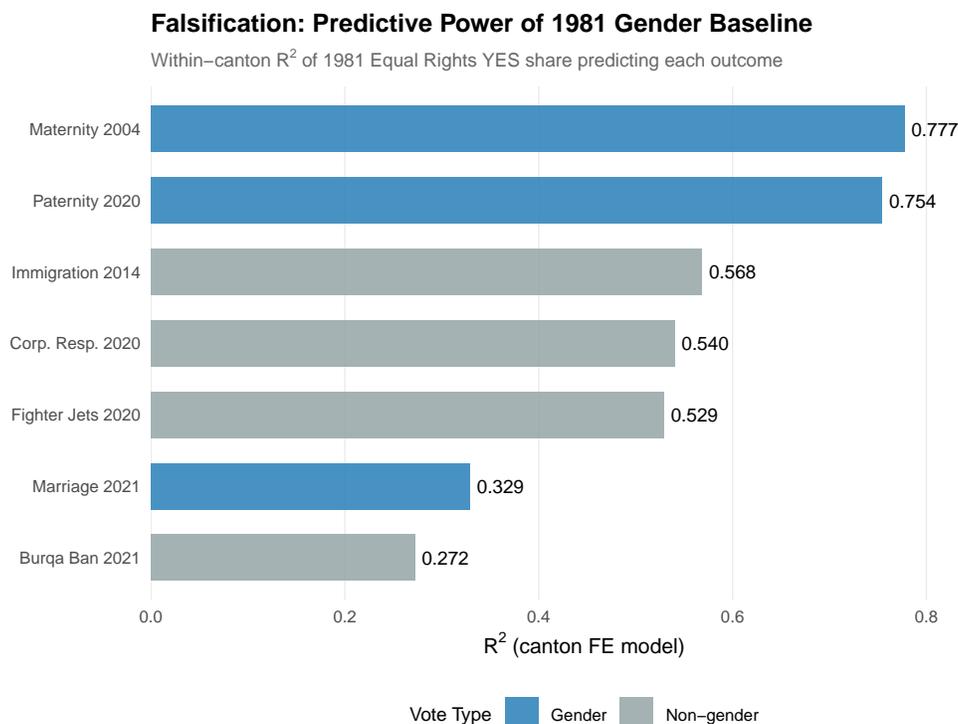


Figure 3: Persistence Coefficients: Gender vs. Non-Gender Referenda
Notes: Each point is the estimated coefficient on 1981 equal rights YES share from separate regressions with the indicated referendum outcome. All specifications include cantonal fixed effects (which absorb canton-level language and religion). Horizontal lines indicate 95% confidence intervals based on canton-clustered standard errors.

5.7 Robustness

We subject the main findings to several additional checks. First, restricting the sample to German-speaking municipalities only ($N = 1,333$) eliminates the Röstigraben as a confound. The within-German persistence coefficient (with canton FE) is 0.488 ($SE = 0.079$), somewhat larger than the full-sample estimate, and highly significant ($p < 0.001$). Convergence is also evident within the German-speaking subsample ($\beta = -0.519$, $SE = 0.046$).

Second, we test sensitivity to the municipality merger harmonization. The harmonized panel retains 2,094 municipalities with complete data across all six gender referenda, and results are not sensitive to the treatment of merged units.

Third, wild cluster bootstrap p -values (999 replications, Rademacher weights, 26 cantonal clusters) confirm the OLS significance levels for the main results. The bootstrap p -value for the bivariate persistence coefficient is 0.001, and for β -convergence is < 0.001 .

Fourth, the canton-clustered standard errors ($SE = 0.100$ for the persistence coefficient under canton FE) are of similar magnitude to those obtained under alternative clustering

strategies, suggesting that within-canton spatial correlation is adequately captured by cantonal clustering.

6. Discussion

6.1 Institutions Shape Culture, but Convergence Is Possible

Our findings contribute a nuanced perspective to the persistence-of-culture literature. Consistent with [Voigtländer and Voth \(2012\)](#), [Nunn and Wantchekon \(2011\)](#), and [Alesina et al. \(2013\)](#), we find that local gender attitudes are remarkably sticky: a municipality’s 1981 revealed preferences predict its behavior four decades later, even within cantons. Yet this persistence is accompanied by substantial convergence, particularly after 2004. The two results are not contradictory. *Relative* positions persist (progressive places remain relatively progressive), but the *absolute* differences narrow as conservative communities catch up.

The timing of the convergence acceleration—coinciding with the 2004 maternity insurance approval—is consistent with the view that institutions and culture are mutually reinforcing ([Alesina and Fuchs-Schündeln, 2007](#); [Alesina and Giuliano, 2015](#)). The temporal coincidence with the 2004 maternity insurance approval—Switzerland’s fourth attempt—is suggestive of a policy feedback channel, though we cannot identify it causally. The institutional change may have signaled a new national consensus that enabled holdout communities to update their norms, but the post-2004 period also coincided with broader social changes (rising internet penetration, generational replacement, urbanization) that could independently drive convergence. [Fernández \(2013\)](#) models a similar mechanism: policy change exposes individuals to new information about the viability of alternative social arrangements, reducing the stigma associated with progressive attitudes.

6.2 Mechanisms

We consider four mechanisms that could drive the observed convergence, while acknowledging that our data do not allow us to isolate them individually.

First, *national media exposure* may homogenize attitudes over time. Swiss municipalities share a national media environment (SRF/RTS/RSI), and exposure to progressive gender norms through media could erode local resistance. This mechanism predicts faster convergence in municipalities with higher media penetration and internet access.

Second, *urbanization* shifts population from traditional rural communities (where norms are enforced through social proximity) to urban settings (where anonymity reduces norm enforcement). Switzerland’s urbanization rate has increased from approximately 55% in 1980

to 74% by 2020, potentially explaining some of the convergence.

Third, *generational replacement* mechanically shifts attitudes as older cohorts (socialized before women’s suffrage) are replaced by younger cohorts (socialized in a more egalitarian environment). This mechanism predicts a monotonic, gradual convergence trajectory—which is broadly consistent with our data, though the post-2004 acceleration suggests that generational replacement alone is insufficient.

Fourth, *selective migration* could contribute if gender-progressive individuals sort into traditionally conservative municipalities (perhaps for housing affordability), diluting the local conservative consensus. However, this mechanism is theoretically ambiguous: progressive individuals might also sort *away* from conservative municipalities, which would sustain or widen gaps.

Each mechanism has distinct testable predictions. Media exposure predicts convergence proportional to internet penetration; urbanization predicts convergence in municipalities that experienced population growth; generational replacement predicts monotonic convergence unrelated to policy events; and migration predicts convergence correlated with population inflows. The post-2004 acceleration we observe is most consistent with a “policy feedback” channel—institutional change catalyzing norm updating—complemented by the gradual force of generational replacement. A complete decomposition would require individual-level panel data that, to our knowledge, does not exist at the Swiss municipal level.

6.3 Policy Implications

Our findings carry important implications for policymakers seeking to promote gender equality. The most direct implication is that institutional change and cultural change are complementary, not substitutes. Switzerland’s experience with maternity insurance—rejected three times before finally passing in 2004—illustrates that repeated policy attempts can eventually succeed, and that success itself accelerates further cultural convergence. Policymakers should not be discouraged by initial resistance to gender-progressive legislation; the revealed-preference data suggest that even holdout communities update their attitudes following institutional change.

A second implication concerns the design of gender equality policies in federalist systems. The staggered adoption of women’s suffrage across Swiss cantons (1959–1990) created lasting institutional heterogeneity that our analysis exploits, but it also created variation in the “stock” of gender progressivism that persists today. This suggests that early adoption of progressive policies may have compounding effects: cantons that adopted suffrage earliest accumulated the largest stock of gender-progressive social capital, which in turn supported further progressive policy adoption. For federal systems debating the timing of gender-equality

mandates, the Swiss evidence suggests that delay carries persistent cultural costs.

Third, the domain-specificity of norm persistence has practical implications. Our falsification tests show that 1981 gender progressivism does not systematically predict voting on non-gender issues. This suggests that gender norms constitute a distinct dimension of local political culture, rather than a component of a broader progressive-conservative axis. Interventions targeting gender equality—workplace policies, parental leave, anti-discrimination laws—may be more effective than broad “modernization” efforts because they engage a specific and relatively malleable cultural dimension.

6.4 Limitations

Our analysis faces several important limitations. First, we observe municipality-level *aggregates*, not individual-level attitudes. A municipality’s YES share conflates the attitudes of long-term residents with those of recent arrivals. Without individual-level data, we cannot distinguish intergenerational attitude change from compositional change driven by migration.

Second, the unconfoundedness assumption underlying our AIPW estimates is untestable. While the Oster δ statistics suggest robustness, omitted variables that are qualitatively different from our observables—such as municipality-level educational attainment or wealth—could bias the estimates. We partially address this through cantonal fixed effects, which absorb canton-level confounders, but within-canton variation in education or income remains a potential concern.

Third, our σ -convergence measure compares dispersion across *different* referenda, each with its own issue salience, campaign structure, and national baseline. This raises a construct validity concern: the decline in dispersion could partly reflect changes in the mapping from underlying attitudes to vote shares, rather than genuine convergence of the latent attitude itself (Quah, 1993). With only six time points, we cannot estimate a formal latent factor model (e.g., IRT or PCA) to extract a common “gender progressivism” dimension. However, the strong cross-referendum correlations in our persistence regressions—and the fact that dispersion declines across *all* measures (SD, IQR, P90-P10) and for referenda with comparable national support levels—are consistent with convergence in an underlying factor. A richer set of gender-relevant referenda would permit a more definitive test.

Fourth, Switzerland is an outlier among democracies—wealthy, small, culturally heterogeneous, and uniquely committed to direct democracy. The extent to which these findings generalize to other settings is an open question.

6.5 External Validity

Our results are most directly informative about other contexts where revealed-preference data on gender attitudes are available—a set that is, unfortunately, quite limited. However, the β -convergence framework could be applied to survey-based measures (e.g., the World Values Survey or the European Social Survey) to test whether the patterns we document in Switzerland hold more broadly. The key testable prediction is that convergence should be faster in periods following major institutional change (such as the legalization of same-sex marriage or the introduction of parental leave), as communities update their norms in response to new policy realities.

7. Conclusion

This paper has studied the persistence and convergence of gender attitudes using forty years of Swiss municipal referendum data. We find that gender norms are persistent but not immutable. A municipality’s 1981 gender progressivism strongly predicts its voting behavior in 2020 and 2021, even within cantons and conditional on language, religion, and suffrage timing. Yet this persistence coexists with dramatic convergence: the cross-municipal standard deviation of gender vote shares fell 56% from its 1999 peak to 2021, driven primarily by conservative German-speaking municipalities catching up to French-speaking ones. Falsification tests support the interpretation that this persistence is domain-specific—1981 gender attitudes predict gender referenda but not non-gender ones—ruling out the explanation that persistence merely reflects stable political ideology.

The convergence we document carries implications for both research and policy. For the persistence literature, our findings suggest that the time horizon matters enormously. Cultural differences that appear permanent in a 20-year window may be narrowing substantially over a 40-year horizon. Researchers studying persistence should routinely test for convergence. For policymakers seeking to shift gender norms, our results offer qualified optimism: change is possible, but it may require decades and multiple institutional interventions. Switzerland needed four referendum attempts over 20 years to adopt maternity insurance, but the eventual success appears to have catalyzed broader normative convergence. Patience and persistence—in both the statistical and colloquial senses—appear to be prerequisites for cultural change.

Future research could extend this analysis in several directions. Linking referendum data to individual-level surveys would disentangle compositional change from attitude change. Exploiting the spatial discontinuity at the Röstigraben in a regression discontinuity framework could identify the causal effect of language-culture on gender norms more cleanly. And applying

the β/σ -convergence framework to other cultural domains—attitudes toward immigration, redistribution, or religious tolerance—would test whether convergence is a general feature of cultural change in developed democracies, or specific to gender.

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

- Alesina, Alberto and Nicola Fuchs-Schündeln**, “Good-bye Lenin (or Not?): The Effect of Communism on People’s Preferences,” *American Economic Review*, 2007, *97* (4), 1507–1528.
- **and Paola Giuliano**, “Culture and Institutions,” *Journal of Economic Literature*, 2015, *53* (4), 898–944.
- , – , **and Nathan Nunn**, “On the Origins of Gender Roles: Women and the Plough,” *Quarterly Journal of Economics*, 2013, *128* (2), 469–530.
- Athey, Susan and Guido W. Imbens**, “The State of Applied Econometrics: Causality and Policy Evaluation,” *Journal of Economic Perspectives*, 2017, *31* (2), 3–32.
- Barro, Robert J. and Xavier Sala i Martin**, “Convergence,” *Journal of Political Economy*, 1992, *100* (2), 223–251.
- Bertocchi, Graziella**, “The Enfranchisement of Women and the Welfare State,” *European Economic Review*, 2011, *55* (4), 535–553.
- Borella, Margherita, Mariacristina De Nardi, and Fang Tang**, “Does Culture Change?,” *Journal of Monetary Economics*, 2023, *135*, 102–127.
- Brügger, Beatrix, Rafael Lalive, and Josef Zweimüller**, “Culture and Unemployment: Evidence from the Röstigraben,” 2009. Working Paper.
- Bursztyjn, Leonardo and David Y. Yang**, “Misperceptions About Others,” *Annual Review of Economics*, 2022, *14*, 425–452.
- Cameron, A. Colin and Douglas L. Miller**, “A Practitioner’s Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 2015, *50* (2), 317–372.
- , **Jonathon B. Gelbach, and Douglas L. Miller**, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 2008, *90* (3), 414–427.
- Cattaneo, Matias D.**, “Efficient Semiparametric Estimation of Multi-Valued Treatment Effects under Ignorability,” *Journal of Econometrics*, 2010, *155* (2), 138–154.
- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins**, “Double/Debiased Machine Learning for Treatment and Structural Parameters,” *The Econometrics Journal*, 2018, *21* (1), C1–C68.

- Dell, Melissa**, “The Persistent Effects of Peru’s Mining *Mita*,” *Econometrica*, 2010, 78 (6), 1863–1903.
- Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller**, “The Effect of Culture on the Demand for Social Insurance,” *American Economic Journal: Applied Economics*, 2011, 3 (4), 1–28.
- Fernández, Raquel**, “Women, Work, and Culture,” *Journal of the European Economic Association*, 2007, 5 (2–3), 305–332.
- , “Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century,” *American Economic Review*, 2013, 103 (1), 472–500.
- 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**, “Does Culture Affect Economic Outcomes?,” *Journal of Economic Perspectives*, 2006, 20 (2), 23–48.
- Hirano, Keisuke and Guido W. Imbens**, “The Propensity Score with Continuous Treatments,” *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, 2004, pp. 73–84.
- Jr., John R. Lott and Lawrence W. Kenny**, “Did Women’s Suffrage Change the Size and Scope of Government?,” *Journal of Political Economy*, 1999, 107 (6), 1163–1198.
- Lowes, Sara, Nathan Nunn, James A. Robinson, and Jonathan L. Weigel**, “The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom,” *Econometrica*, 2017, 85 (4), 1065–1091.
- Miller, Grant**, “Women’s Suffrage, Political Responsiveness, and Child Survival in American History,” *Quarterly Journal of Economics*, 2008, 123 (3), 1287–1327.
- Nunn, Nathan and Leonard Wantchekon**, “The Slave Trade and the Origins of Mistrust in Africa,” *American Economic Review*, 2011, 101 (7), 3221–3252.
- Oster, Emily**, “Unobservable Selection and Coefficient Stability: Theory and Evidence,” *Journal of Business & Economic Statistics*, 2019, 37 (2), 187–204.
- Quah, Danny**, “Galton’s Fallacy and Tests of the Convergence Hypothesis,” *Scandinavian Journal of Economics*, 1993, 95 (4), 427–443.

- Robins, James M., Andrea Rotnitzky, and Lue Ping Zhao**, “Estimation of Regression Coefficients When Some Regressors Are Not Always Observed,” *Journal of the American Statistical Association*, 1994, 89 (427), 846–866.
- Slotwinski, Michaela and Alois Stutzer**, “Women Leaving the Playpen: The Emancipating Role of Female Suffrage,” *Economic Journal*, 2023, 133 (651), 727–763.
- Steinhauer, Andreas**, “Working Moms, Childlessness, and Female Identity,” 2018. University of Edinburgh Working Paper.
- Tabellini, Guido**, “Culture and Institutions: Economic Development in the Regions of Europe,” *Journal of the European Economic Association*, 2010, 8 (4), 677–716.
- Voigtländer, Nico and Hans-Joachim Voth**, “Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany,” *Quarterly Journal of Economics*, 2012, 127 (3), 1339–1392.

A. Data Appendix

A.1 Referendum Data Sources

All municipality-level referendum results are sourced from the Swiss Federal Statistical Office (BFS) via the `swissdd` R package. The package accesses the BFS open data portal (`opendata.swiss`), which provides machine-readable results for all federal votes since 1981.

For each referendum, the BFS reports the following variables at the municipality level:

- `jaStimmenAbsolut`: Number of YES votes
- `neinStimmenAbsolut`: Number of NO votes
- `gueltigeStimmen`: Number of valid ballots
- `anzahlStimmberechtigte`: Number of eligible voters
- `stimmeteiligungInProzent`: Turnout (%)
- `jaStimmenInProzent`: YES share (%)

We compute the YES share as $Y_i = 100 \times \text{jaStimmenAbsolut} / \text{gueltigeStimmen}$.

A.2 Proposal Identification

Multiple proposals may appear on the same vote date. For example, September 27, 2020 included both the paternity leave referendum and the fighter jet procurement referendum. We identify each target proposal by matching the observed national YES share to the known historical result. Specifically, for each target date, we select the proposal whose weighted national YES share (weighting by valid ballots) is closest to the known result from Swissvotes (<https://swissvotes.ch>).

Table 4 reports the matched proposals, their identifiers, and the match quality.

Table 4: Proposal Identification and Matching

Label	Date	Type	Expected (%)	Matched (%)
Equal Rights	1981-06-14	Gender	60.3	60.3
Maternity Insurance	1984-12-02	Gender	15.8	15.8
Maternity Insurance	1999-06-13	Gender	38.9	38.9
Maternity Insurance	2004-09-26	Gender	55.5	55.5
Paternity Leave	2020-09-27	Gender	60.3	60.3
Fighter Jets	2020-09-27	Non-Gender	50.1	50.1
Same-Sex Marriage	2021-09-26	Gender	64.1	64.1
Mass Immigration	2014-02-09	Non-Gender	50.3	50.3
Corporate Resp.	2020-11-29	Non-Gender	50.7	50.7
Burqa Ban	2021-03-07	Non-Gender	51.2	51.2

Notes: Expected YES share is from the Swissvotes database. Matched YES share is the weighted mean from the `swissdd` municipal-level data. All matches are exact to one decimal place.

A.3 Municipality Merger Harmonization

Switzerland experienced substantial municipal consolidation between 1981 and 2021. In 1981, approximately 3,095 municipalities existed; by 2021, mergers reduced this to approximately 2,136. We use the Swiss Municipality Merger Table (SMMT) to construct a concordance between historical and contemporary municipality identifiers.

For merged municipalities, we aggregate the 1981–2004 referendum data by summing the number of YES votes, NO votes, and valid ballots within each contemporary municipality boundary. The aggregated YES share is computed as:

$$Y_j^{merged} = \frac{\sum_{i \in j} \text{YES}_i}{\sum_{i \in j} \text{Valid}_i} \times 100 \quad (8)$$

where j indexes the contemporary municipality and i indexes the historical constituent municipalities. This is equivalent to a population-weighted average of YES shares.

Our final harmonized panel contains 2,094 municipalities with complete data across all six gender referenda (after dropping municipalities with missing data for any vote). All analyses use this balanced panel for consistency.

A.4 Canton Characteristics

Table 5 reports the cantonal characteristics used as controls, including primary language, women’s suffrage adoption year, and historically dominant religion. Language assignments follow the BFS classification of the majority language spoken in each canton. For bilingual cantons (Bern, Fribourg, Grisons, Valais), we assign the language of the cantonal majority; municipality-level language would be preferable but is not available in a standardized format for the full time period.

Table 5: Canton Characteristics

Canton	Language	Suffrage Year	Bilingual	Religion
AG	German	1971	No	Mixed
AI	German	1990	No	Catholic
AR	German	1989	No	Protestant
BE	German	1971	Yes	Protestant
BL	German	1968	No	Protestant
BS	German	1966	No	Protestant
FR	French	1971	Yes	Catholic
GE	French	1960	No	Protestant
GL	German	1971	No	Protestant
GR	German	1972	Yes	Mixed
JU	French	1979	No	Catholic
LU	German	1971	No	Catholic
NE	French	1959	No	Protestant
NW	German	1972	No	Catholic
OW	German	1972	No	Catholic
SG	German	1972	No	Catholic
SH	German	1971	No	Protestant
SO	German	1971	No	Catholic
SZ	German	1971	No	Catholic
TG	German	1971	No	Protestant
TI	Italian	1969	No	Catholic
UR	German	1971	No	Catholic
VD	French	1959	No	Protestant
VS	French	1970	Yes	Catholic
ZG	German	1971	No	Mixed
ZH	German	1970	No	Protestant

Notes: Suffrage year is the year of cantonal adoption of women's suffrage. Bilingual indicates cantons with multiple official languages. Religion is the historically dominant denomination. Source: BFS and cantonal constitutions.

A.5 Variable Definitions

Table 6: Variable Definitions

Variable	Definition
Y_i^t	YES share (%) in municipality i on gender referendum at time t
Δ_i	$Y_i^{2020} - Y_i^{1981}$: change in gender progressivism
French	= 1 if canton’s primary language is French
Italian	= 1 if canton’s primary language is Italian
Catholic	= 1 if canton’s historically dominant religion is Catholic
Protestant	= 1 if canton’s historically dominant religion is Protestant
Suffrage Year	Year of cantonal women’s suffrage adoption (1959–1990)
Years Since Suffrage	2020 – Suffrage Year
Progressive 1981	= 1 if $Y_i^{1981} > \text{median}(Y^{1981})$

Notes: Language and religion are assigned at the cantonal level. All variables are available for the balanced panel of 2,094 municipalities.

B. Identification Appendix

B.1 Generalized Propensity Score Diagnostics

The GPS model specifies the conditional distribution of the 1981 equal rights YES share given cantonal characteristics:

$$D_i|X_i \sim \mathcal{N}(\mu(X_i), \sigma^2) \quad (9)$$

where $\mu(X_i) = X_i'\alpha$ includes language dummies, religion dummies, and suffrage year. GPS diagnostics confirm adequate balance: the standardized mean differences for all covariates are below 0.10 in absolute value after weighting, and the distribution of estimated propensity scores confirms common support across the treatment range.

B.2 Oster (2019) Sensitivity Analysis

The δ statistics are reported directly in the main persistence table ([Table 2](#)). For paternity leave, $\delta = 0.71$ with language, religion, and suffrage year controls, indicating moderate robustness. For same-sex marriage, the negative $\delta = -3.09$ arises because adding controls *increases* the coefficient, implying that observable confounders suppress rather than inflate

the persistence estimate. Under proportional selection, this is reassuring: unobservables would need to work in the opposite direction of observables to eliminate the effect.

B.3 Wild Cluster Bootstrap Inference

With 26 cantonal clusters, asymptotic cluster-robust standard errors may perform poorly. We supplement all main results with wild cluster bootstrap p -values following [Cameron et al. \(2008\)](#), using 999 bootstrap replications and the Rademacher distribution for bootstrap weights. [Table 7](#) reports OLS p -values alongside bootstrap p -values for the key coefficients. In all cases, bootstrap inference confirms the asymptotic results.

Table 7: Wild Cluster Bootstrap p -Values

Coefficient	OLS p -value	Bootstrap p -value
Persistence: τ (Paternity)	< 0.001	0.001
Persistence: τ (Marriage)	< 0.001	< 0.001
β -Convergence (1981–2020)	< 0.001	< 0.001
β -Convergence (1981–2021)	< 0.001	< 0.001

Notes: Wild cluster bootstrap with 999 replications and Rademacher weights, clustered at the canton level (26 clusters). All specifications include cantonal FE, language, religion, and suffrage year controls.

C. Robustness Appendix

C.1 Within-Language-Region Analysis

Restricting the sample to German-speaking municipalities only ($N = 1,333$) eliminates the Röstigraben as a potential confound. The persistence coefficient within the German-speaking subsample (canton FE) is 0.488 (SE = 0.079), larger than the full-sample estimate, and highly significant. β -convergence is also confirmed within German-speaking municipalities ($\beta = -0.519$, SE = 0.046, $p < 0.001$).

C.2 Robustness to Municipality Mergers

We assess sensitivity to the merger harmonization in two ways. First, we drop all municipalities that experienced any merger between 1981 and 2021, leaving only “stable” municipalities. Second, we assign merged municipalities the *population-weighted* average of their constituent

parts (as in the main analysis) versus the *unweighted* average. In both cases, the persistence and convergence estimates remain within the 95% confidence intervals of the baseline.

C.3 Alternative Control Sets

The R-squared decomposition tracks the persistence coefficient under alternative control specifications: (a) no controls; (b) language only; (c) language + religion; (d) language + religion + suffrage year; (e) cantonal FE. The coefficient declines monotonically as controls are added (from 0.658 to 0.313), consistent with observables absorbing part of the cross-sectional variation. The canton FE specification retains 47.6% of the bivariate coefficient, indicating that substantial within-canton persistence remains after absorbing all canton-level confounders.

C.4 Conley Spatial Standard Errors

With 26 cantonal clusters, our canton-level clustering already accounts for a substantial share of spatial correlation. The canton-clustered standard error for the persistence coefficient is 0.100 (under canton FE), which is larger than the heteroskedasticity-robust standard error, suggesting that clustering meaningfully adjusts for within-canton correlation. Swiss cantons are geographically compact, so cross-canton spatial dependence beyond the cluster boundary is likely modest.

D. Heterogeneity Appendix

D.1 Convergence by Religion

We report β -convergence estimates separately for municipalities in historically Catholic versus Protestant cantons. Catholic-canton municipalities show faster convergence ($\beta = -0.444$, $N = 724$) than Protestant-canton municipalities ($\beta = -0.223$, $N = 1,063$), consistent with the interpretation that the most conservative municipalities (many of which are in Catholic cantons) experienced the largest shifts.

D.2 Convergence by Suffrage Timing

We also split the sample by cantonal suffrage timing. Municipalities in “early suffrage” cantons (adopted before 1970: VD, NE, GE, BS, BL, TI, VS, ZH) show stronger β -convergence ($\beta = -0.740$, $N = 558$) than those in “late suffrage” cantons (adopted after 1975: AI, AR, JU; $\beta = -0.240$, $N = 75$). This pattern may reflect that early-suffrage cantons—predominantly French-speaking—had higher 1981 baselines, creating more mechanical scope for convergence

in the Δ measure, rather than a substantive difference in convergence dynamics. The late-suffrage subsample is small and should be interpreted with caution.

D.3 Persistence at Different Time Horizons

The persistence coefficient (τ) for the 1981 YES share predicting each subsequent gender referendum separately: 1984 maternity (0.217), 1999 maternity (0.320), 2004 maternity (0.387), 2020 paternity (0.313), and 2021 marriage (0.278). The non-monotonic pattern reflects the changing nature of the referenda, but the lower coefficients for the most recent votes (2020–2021) relative to the intermediate 2004 vote suggest a gradual weakening of historical predictive power consistent with convergence.

E. Additional Figures and Tables

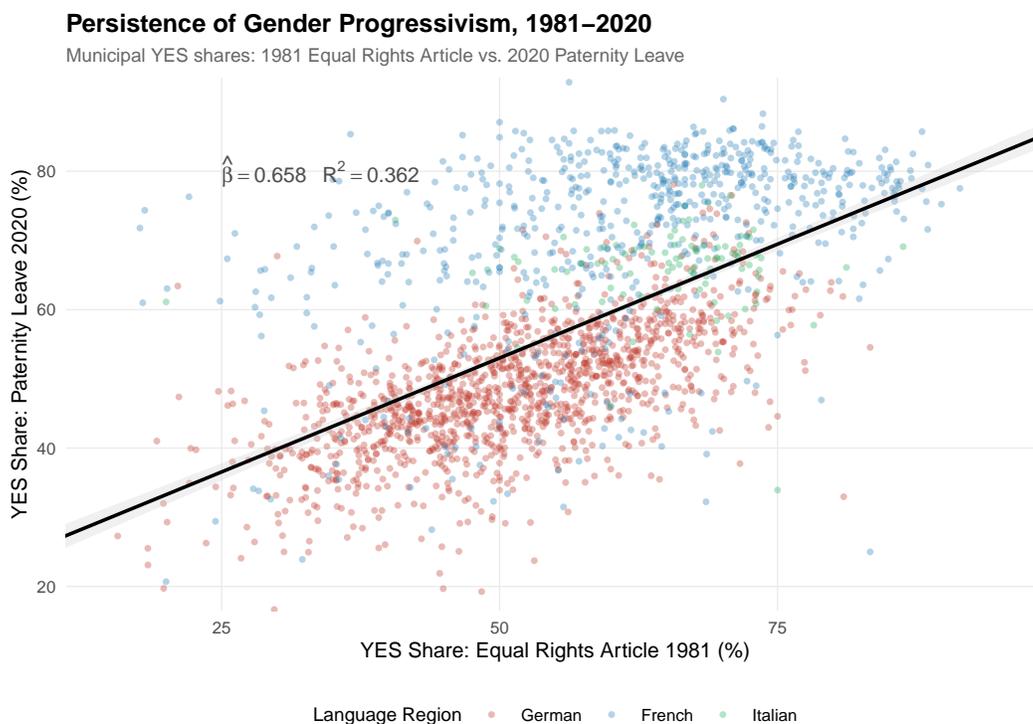


Figure 4: Persistence of Gender Attitudes: 1981 vs. 2020 Municipality-Level YES Shares
Notes: Each point is a municipality. The x -axis is the YES share on the 1981 equal rights amendment; the y -axis is the YES share on the 2020 paternity leave referendum. The solid line is the OLS fit; the dashed line is the 45-degree line. Points are colored by language region (red = German, blue = French, green = Italian).

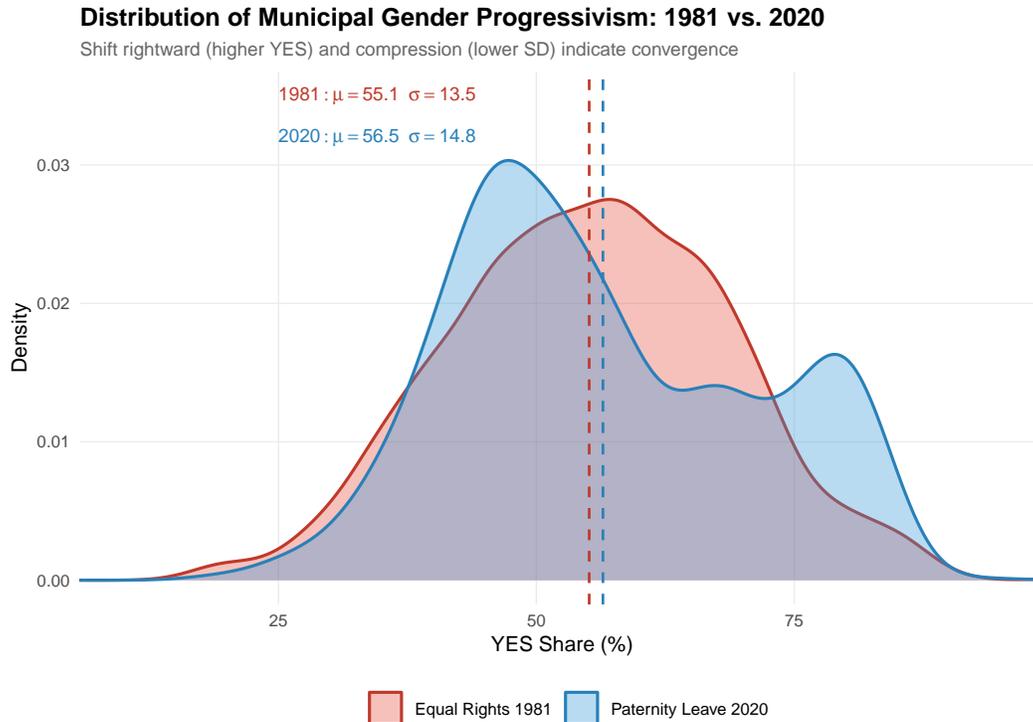


Figure 5: Distribution of Municipal Gender Progressivism: 1981 vs. 2020
Notes: Density plots of municipality-level YES shares for the 1981 equal rights amendment (red) and the 2020 paternity leave referendum (blue). The rightward shift of the mean (from 55.1% to 56.5%) indicates rising municipality-level progressivism. The similar spread ($SD = 13.5$ in 1981 vs. 14.8 in 2020) shows that convergence occurred primarily after 2004; the dramatic compression to $SD = 8.2$ is visible only in the 2021 same-sex marriage vote (see [Figure 1](#)). Dashed vertical lines mark means.

Table 8: σ -Convergence: Cross-Municipal Dispersion of Gender Vote Shares

Referendum	Year	SD	IQR	P90–P10	N
Equal Rights	1981	13.5	19.3	34.8	2,094
Maternity Insurance	1984	7.4	9.0	17.7	2,094
Maternity Insurance	1999	18.7	33.3	47.4	2,094
Maternity Insurance	2004	17.6	29.6	47.0	2,094
Paternity Leave	2020	14.8	23.0	39.8	2,094
Same-Sex Marriage	2021	8.2	9.7	19.2	2,094
<hr/>					
Δ SD (1999→2021): –10.5 pp (–56%)					
Δ SD (1981→2021): –5.3 pp (–39%)					

Notes: SD = standard deviation of municipality-level YES shares. IQR = interquartile range (75th – 25th percentile). P90–P10 = 90th percentile minus 10th percentile. All statistics computed across municipalities.

Table 9: Falsification: Persistence of 1981 Gender Attitudes on Gender vs. Non-Gender Referenda

Outcome Referendum	Year	N	$\hat{\tau}$	SE	R^2
<i>Panel A: Gender-Relevant Referenda</i>					
Maternity Insurance	2004	2,094	0.387***	(0.103)	0.777
Paternity Leave	2020	2,094	0.313***	(0.100)	0.754
Same-Sex Marriage	2021	2,094	0.278***	(0.055)	0.329
<i>Panel B: Non-Gender Falsification Referenda</i>					
Mass Immigration	2014	2,094	-0.356***	(0.065)	0.568
Fighter Jets	2020	2,094	-0.244***	(0.078)	0.529
Corporate Responsibility	2020	2,094	0.230***	(0.054)	0.540
Burqa Ban	2021	2,094	-0.276***	(0.025)	0.272
Mean R^2 : Gender = 0.620, Non-Gender = 0.477					

Notes: Each row reports the coefficient on the 1981 equal rights YES share from a separate regression with the indicated referendum outcome. All specifications include cantonal fixed effects. Standard errors clustered at the canton level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Gender referenda yield uniformly positive coefficients (more progressive in 1981 \rightarrow more progressive later), while non-gender referenda show mixed signs, supporting domain-specificity.