"A Pretty Blunt Approach":

Meta's Political Content Reduction Policy and Italian Parliamentarians' Facebook Visibility

Fabio Giglietto

University of Urbino

Working Paper Draft

January 2026

Abstract

In February 2021, Meta announced it would reduce the distribution of political content in users' News Feeds, implementing the policy globally by July 2022 before reversing course in January 2025. Despite the significance of these changes for democratic communication, no independent research has quantified their effects on elected officials. This study examines Meta's political content reduction policy using a dataset of over 2.5 million Facebook posts from Italian parliamentarians, prominent politicians, and political extremist accounts spanning 2021–2025, collected via the Meta Content Library API. Employing a discovery-validation design with structural breakpoint detection methods, the analysis identifies three cross-validated breakpoints that reveal policy effects emerged in Italy approximately ten months before Meta's announced global implementation. The findings document a 72% reduction in average reach for re-elected Members of Parliament, declining from approximately 53,000 views per post before the policy to 15,000 at the trough. Although Meta's January 2025 reversal produced a significant rebound, post-reversal reach recovered to only 65% of pre-policy levels. A robustness check reveals asymmetric effects: while mainstream politicians experienced substantial visibility losses, extremist accounts' increased posting frequency more than compensated for modest per-post declines, resulting in increased total weekly reach during the policy period. These findings expose significant transparency deficits in Meta's policy communication while also demonstrating the research value of the transparency infrastructure that made this study possible.

Keywords: platform governance, algorithmic content moderation, political communication, Meta, Facebook, transparency, Digital Services Act

Introduction

Social media platforms have become essential infrastructure for political communication in contemporary democracies. Elected representatives increasingly rely on platforms like Facebook to communicate with constituents, shape public discourse, and mobilize support. Yet these same platforms exercise considerable—and often opaque—control over whose voices reach which audiences. When Meta announced in February 2021 that it would begin reducing the distribution of political content in users' News Feeds (Stepanov & Gupta, 2021), the company initiated a policy experiment with profound implications for democratic communication. This study examines those implications empirically, asking: *When and to what extent did Meta's political content reduction policy affect Italian parliamentarians' reach on Facebook?*

Meta's approach to political content visibility evolved through multiple phases between 2021 and 2025. The company's stated rationale was user demand: as Mark Zuckerberg noted, "one common piece of feedback we hear is that people don't want political content to take over their News Feed" (Stepanov & Gupta, 2021). Beginning with tests in February 2021 in the United States, Canada, Brazil, and Indonesia, Meta experimented with reducing political content distribution before announcing global implementation in July 2022. The primary mechanism was de-emphasizing engagement signals, particularly shares and comments, when ranking posts classified as political. As Meta stated: "placing less emphasis on shares and comments for political content is an effective way to reduce the amount of political content people experience in their Feed" (Stepanov & Gupta, 2021, updated July 2022). In an April 2023 update, the company described "continuing to move away from ranking based on engagement, and towards giving more weight to signals based on the results of user surveys" (Stepanov & Gupta, 2021, updated April 2023). Industry research documents that such non-engagement approaches typically rely on predictive models trained on item-level surveys (e.g. asking users whether specific content is "worth your time," "informative," or "good for the world") though Meta has not disclosed which specific signals were incorporated into its political content ranking (Cunningham et al., 2024). By September 2024, Meta had introduced user-facing controls allowing individuals to adjust their political content settings, before announcing in January 2025 a shift toward a "more personalized approach" that would treat political content from followed accounts more like other content (Kaplan, 2025).

The January 2025 announcement marked a significant policy reversal. Meta's Chief Global Affairs Officer Joel Kaplan acknowledged that the company's content moderation systems had become overly restrictive, producing widespread errors and user dissatisfaction (Kaplan, 2025). He characterized the earlier policy as excessively blunt and announced a gradual return to a more personalized treatment of political content across Facebook, Instagram, and Threads, allowing greater visibility for users who wish to engage with such material. This reversal provides a natural endpoint for assessing the effects of the 2021–2024 policy period, while also highlighting the importance of understanding what that policy actually achieved.

In fact, despite the significance of these changes for political communication, fundamental questions about their implementation and effects remain unanswered. Meta's announcements provide limited detail about how political content is identified, what magnitude of reach reduction the policy produced, and whether effects varied across political contexts. The company noted that political content comprised approximately 3% of what users saw in their feeds—a figure updated in November 2022 (Stepanov & Gupta, 2021)—but offered no systematic data on how reach changed for political actors. More critically, the company failed to report on these effects in transparency documentation produced for institutions like the European Commission. The policy is conspicuously absent from official documents—including the EU Digital Services Act: Post-Election Reports—where an acknowledgment of mechanisms designed to drastically reduce political visibility would be expected. Consequently, no independent research has yet quantified the specific effects on elected officials whose democratic mandate depends on digital communication.

Italy provides an instructive case for examining these dynamics. Italian parliamentarians are highly active Facebook users, and the country's multiparty landscape allows for an examination of effects across a diverse ideological spectrum. Furthermore, Italy held parliamentary elections in September 2022—two months after Meta's global policy implementation—providing a unique opportunity to analyze the interplay between algorithmic suppression and intense political activity.

To address this research gap, this study leverages a dataset of over 2.5 million Facebook posts from Italian parliamentarians, prominent politicians, and political extremist accounts spanning 2021–2025. The analysis employs a discovery-validation design suited to the temporal ambiguity of Meta's rollout: because precise implementation dates for Italy were never publicly confirmed, structural breakpoint detection methods are first used to identify

significant discontinuities in the time series data, after which the magnitude of reach changes is assessed through interrupted time series analysis, controlling for secular trends and seasonal variations in political attention.

The remainder of this paper proceeds as follows. Section 2 describes our data collection strategy, the discovery-validation design that uses re-elected MPs as the discovery sample and other political actor groups for cross-validation, and our analytical methods for structural breakpoint detection. Section 3 presents findings from the breakpoint analysis, identifying three cross-validated breakpoints that define four policy phases—pre-policy baseline, initial policy implementation, adjusted policy (post-election), and post-reversal. The analysis reveals a distinctive pattern of sequential decline followed by partial recovery (DOWN → DOWN → UP) that validates across mainstream political actors but diverges for extremist accounts. Section 4 summarizes the magnitude of policy effects and presents a robustness check examining total weekly reach (rather than per-post averages), which reveals that extremist accounts' increased posting volume more than compensated for lower per-post reach, resulting in increased total visibility during the policy period. Section 5 discusses these findings in terms of both the transparency deficits they expose and the transparency progress that made this study possible, particularly the Meta Content Library infrastructure enabled by the Digital Services Act. The discussion concludes by arguing for a collaborative rather than adversarial orientation toward platform governance research. Section 6 addresses limitations and directions for future research, including cross-country replication and extension to Instagram.

Data and Methods

Data Structure

The study draws on data collected via the Meta Content Library API (Meta Platforms, Inc., 2024), accessed through Meta's Secure Research Environment. The Meta Content Library provides comprehensive access to publicly available content from Facebook and Instagram, including post-level engagement metrics and view counts. Data were queried programmatically using R within the cleanroom environment.

The dataset comprises Facebook post data from four distinct groups of Italian political actors: Members of Parliament elected in both 2018 and 2022 legislatures, Members of Parliament

elected only in 2022, Most-followed politicians not in Parliament, Italian pages/figures in alternative media ecosystem (Table 1).

The Prominent Politicians category comprises the most-followed Italian politicians on Facebook who do not hold seats in Parliament. The list was sourced from Politici italiani su Facebook (*https://politicasufacebook.it/politici/followers/TUTTI*), a tracking service that monitors Italian politicians' social media presence. These accounts represent high-profile political figures—including regional leaders, party officials, and former parliamentarians—who maintain significant public audiences but lack the institutional status of sitting MPs.

The 'Extremists' category comprises Italian Facebook pages and public figures characterized primarily by anti-establishment narratives, criticism of mainstream media and public figures, and vaccine/health policy skepticism. Operating within a non-mainstream media ecosystem, these accounts provide a curated feed of counter-information, news commentary, and lifestyle content. While Euroskepticism and criticism of Western foreign policy are present, the dominant themes center on distrust of institutions, COVID-19 vaccine criticism, and opposition to perceived elite narratives. The accounts span the political spectrum, with content leaning right, far-right, and far-left. This list was sourced from independent platform transparency research and was not compiled by the author for this study.

All lists are mutually exclusive. When an account appeared in multiple lists, we applied the following priority rules: (1) accounts from elected MPs appearing in the prominent politicians list were assigned to the appropriate MP category; (2) accounts appearing in both the prominent politicians and extremists lists were assigned to prominent politicians. This ensures each account contributes to only one analytical group.

The dataset consists of 2,529,933 posts across 901 accounts, spanning 256 complete weeks from January 3, 2021, to November 23, 2025. This represents approximately five years of Facebook activity. Table 2 presents the post counts and account coverage by group.

Table 1. Political Actor Groups.

Group	Source	Analytical Role
MPs (Re-elected)	Members of Parliament elected in both 2018 and 2022 legislatures	Discovery sample for breakpoint identification

Group	Source	Analytical Role
MPs (New)	Members of Parliament elected only in 2022	Validation sample (institutional actors)
Prominent Politicians	Most-followed politicians not in Parliament	Validation sample (non-institutional)
Extremists	Italian pages/figures in alternative media ecosystem	Validation sample (potentially targeted)

Table 2. Account and Post Counts by Group.

Group	N Accounts	Total Posts	Average Views
Extremists	146	1,311,247	10,563
MPs (New)	272	355,694	20,639
MPs (Re-elected)	281	499,164	26,343
Prominent Politicians	202	363,828	49,867

Table 3 presents weekly aggregated engagement statistics across the full observation period, revealing substantial variation in baseline reach and engagement patterns across groups. Several patterns are notable. First, Prominent Politicians consistently show the highest engagement metrics, with mean weekly views of 49,867 compared to 26,343 for re-elected MPs. Second, Extremists have substantially lower reach (mean 10,563 views) despite generating the most posts, suggesting either smaller audiences or pre-existing algorithmic limitations. Third, the high standard deviations across all groups (often exceeding the mean) indicate substantial week-to-week variation, motivating our use of breakpoint detection methods that can identify structural changes within noisy time series.

Table 3. Weekly Aggregated Engagement Statistics.

Views (Reach)	**Mean**	**SD**	**Median**	**Min**	**Max**
Extremists	10,562.8	2,830.1	10,535.8	4,491.0	18,299.5
MPs (New)	20,638.7	11,794.8	17,003.6	5,931.4	58,774.3
MPs (Re-elected)	26,342.9	15,432.6	21,277.8	7,529.1	78,420.0
Prominent Politicians	49,866.9	19,857.2	42,761.2	20,171.8	116,548.4
Reactions	**Mean**	**SD**	**Median**	**Min**	**Max**
Extremists	186.7	76.3	188.3	63.4	582.9
MPs (New)	405.2	239.7	310.6	143.3	1,103.5
MPs (Re-elected)	666.8	354.7	581.5	239.0	1,650.1
Prominent Politicians	1,051.5	429.9	888.2	475.1	2,211.8
Shares	**Mean**	**SD**	**Median**	**Min**	**Max**
Extremists	51.1	26.2	45.0	12.6	159.9
MPs (New)	52.3	35.6	38.4	16.3	184.1
MPs (Re-elected)	84.7	50.8	67.7	25.3	266.3
Prominent Politicians	171.7	90.3	134.4	53.7	435.4

Comments	Mean	SD	Median	Min	Max
Extremists	43.6	13.5	43.6	19.6	85.4
MPs (New)	90.9	71.1	59.8	30.1	342.7
MPs (Re-elected)	173.8	105.5	135.3	58.9	517.0
Prominent Politicians	226.0	104.5	182.0	90.8	591.5

Outcome Measure: Views

The primary outcome measure is views, as defined by Meta Content Library. Views represent the cumulative number of times a post appeared on screen for at least 250 milliseconds, excluding appearances on the post owner's screen. Several features of this metric warrant attention:

Impressions, not unique reach. Views count multiple appearances from the same account. If the same user sees a post three times, three views are recorded. This means views measure total impressions rather than unique audience reach.

Reshare amplification. When a reshared post appears on screen, view counts increment for both the original post and the reshare. For example, if User A creates a post and User B reshares it, a third user viewing B's reshare adds one view to both A's original post and B's reshare. This mechanism means that accounts whose content is frequently reshared may accumulate views at accelerated rates.

Minimum threshold and missing data handling. Only posts exceeding 100 views display a view count in the Meta Content Library; posts at or below this threshold return missing values (NA). Following clarification from Meta support, the analysis treated NA values as censored observations indicating ≤100 views. To preserve these observations in the analysis, missing view counts were imputed using a group-specific ratio-weighted approach derived from empirical analysis of each group's observed distribution.

For each political group, a power law model was fitted to bin counts in the 101–500 view range and extrapolated to estimate the distribution below the censoring threshold. This yielded group-specific parameters reflecting meaningful differences in how content falls below the visibility threshold. Extremists exhibited a left-skewed distribution ($\alpha = -0.36$, ratio = 1.65), indicating their censored posts are more likely to have very low view counts (expected imputed mean: 44). In contrast, Prominent Politicians showed a right-skewed distribution ($\alpha = 0.33$, ratio = 0.66), suggesting their censored posts tend to be closer to the

100-view threshold (expected imputed mean: 56). MPs fell between these extremes: re-elected MPs ($\alpha = 0.14$, ratio = 0.82, mean: 53) and newly elected MPs ($\alpha = 0.03$, ratio = 0.94, mean: 51) showed distributions closer to uniform. A fixed random seed (42) ensured reproducibility.

This censoring disproportionately affected different groups: 9.4% of Extremist posts had missing view counts, compared to 5.4% of re-elected MP posts, 4.3% of newly elected MP posts, and only 1.0% of Prominent Politicians posts.

These features have implications for interpretation. The reshare amplification mechanism could differentially affect groups whose content circulates through resharing networks. However, sensitivity analysis found a maximum difference of only 0.002% between imputation approaches, indicating that the censoring threshold does not materially affect the study's conclusions.

Unlike views, engagement metrics (reactions, shares, comments) have no documented censoring threshold; posts with missing values on these variables (N = 24,426) were therefore excluded from the dataset.

Policy Timeline and Expected Breakpoints

Meta's political content reduction policy followed a documented timeline with two critical inflection points that should produce detectable structural breaks in Italian political actors' reach.

The expected Breakpoint 1 is related to the policy implementation, between 2021 and 2022. Meta announced initial tests in February 2021 (US, Canada, Brazil, Indonesia), expanded to additional countries including European nations in August–October 2021, and implemented the policy globally in July 2022. The expected effect of this first breakpoint is a significant decrease in parliamentarian posts' reach.

The expected Breakpoint 2 regards instead the policy reversal in January 2025. On January 7, in fact, Meta announced a fundamental shift toward what the company calls a "more personalized approach" where political content from followed accounts would be treated like other categories of content in the feed. Expected effect: Significant increase in reach, returning toward pre-policy levels.

Discovery-Validation Design

Re-elected Members of Parliament constitute an optimal discovery sample for identifying the timing of policy implementation for three main reasons. First, these MPs maintained active Facebook pages throughout the entire study period (2021–2025), ensuring uninterrupted time series data that spans both the introduction of the policy and its subsequent reversal. This temporal continuity makes it possible to observe changes in reach before, during, and after the policy intervention. Second, because the sample consists of the same individuals over time, observed variations in reach cannot be attributed to compositional changes in the population of accounts. Any detected breakpoints therefore reflect shifts in platform-level content distribution rather than changes in who is posting. Finally, the policy reversal enables a quasi-experimental design based on the expectation of two breakpoints with opposite directional effects: a decline in reach at the time of implementation and a recovery following the reversal. When both breakpoints align with Meta's announcements, the causal attribution of observed effects to the policy intervention is substantially strengthened.

Analytical Methods

To identify and validate policy-related structural changes in engagement, the analysis followed a sequential procedure that moved from breakpoint detection to effect validation. The first stage focused on discovering potential structural shifts using weekly aggregated data from re-elected MPs. Two complementary breakpoint detection algorithms—Bai–Perron and PELT—were applied to four engagement metrics (views, reactions, shares, and comments), generating multiple independent signals of change. Rather than relying on single detections, temporally proximate breakpoints were consolidated into 30-day clusters, whose density of detections was used to assess the strength of consensus; only clusters supported by both algorithms were retained, ensuring that identified breakpoints reflected robust structural change. These validated breakpoints were then used to segment the time series into discrete policy phases, for which summary statistics were computed to characterize engagement levels across periods. Building on this phase structure, the magnitude of policy effects was assessed by estimating the reduction in engagement at policy onset, the increase associated with policy reversal, and the extent to which engagement returned to its pre-policy baseline. Finally, the external validity of these effects was examined by testing whether the same breakpoints produced significant and directionally consistent changes in engagement among additional actor groups, including newly elected MPs, prominent political figures, and extremist accounts.

Analysis and Results

This section presents findings from the breakpoint analysis addressing the RQ: *When and to what extent did Meta's political content reduction policy affect Italian parliamentarians' reach on Facebook?* The study employs a data-driven approach using Bai-Perron and PELT breakpoint detection methods, prioritizing empirical evidence over hypothesized policy dates. The analysis reveals both expected and unexpected patterns in the timing and magnitude of policy effects.

Breakpoint Detection Results

The consensus clustering procedure identified 19 distinct breakpoint clusters across the observation period. Of these, four meet the threshold for cross-algorithm validation (detected by both Bai-Perron and PELT): Cluster 4 (September 2021, STRONG), Cluster 5 (November 2021, MODERATE), Cluster 10 (January 2023, VERY STRONG), and Cluster 18 (March 2025, VERY STRONG).

The analysis consolidates Clusters 4 and 5 into a single implementation breakpoint (T_1) for three reasons. First, the two clusters are temporally proximate, separated by only eight weeks (September 19 to November 14, 2021). Second, both fall within Meta's documented expansion period, when the company announced it was "expanding the political content ranking tests to more countries around the world" (October 2021) and subsequently "put less emphasis on signals such as how likely someone is to comment on or share political content" (August 2021). Third, Cluster 5's MODERATE strength rating—based on only 3 detections compared to Cluster 4's 5 detections—suggests it may represent a secondary adjustment within the initial implementation phase rather than a distinct policy event. The consolidated T_1 breakpoint uses Cluster 4's point estimate (September 19, 2021) as the primary implementation date, with Cluster 5 interpreted as marking the completion of initial rollout. Table 4 summarizes the three resulting breakpoints.

Table 4. Cross-Validated Breakpoints (Detected by ≥2 Algorithms).

Breakpoint	Point Estimate	Detections	Strength	Detection Range
T_1 (Implementation)	2021-09-19	5	STRONG	Aug 29 – Oct 3, 2021
T_2 (Adjustment)	2023-01-01	13	VERY STRONG	Oct 30, 2022 – Feb 12, 2023
T_3 (Reversal)	2025-03-09	9	VERY STRONG	Mar 2 – May 4, 2025

Note: Point estimates are cluster medians. Detection range spans earliest to latest breakpoint date in each cluster; ranges exceeding 30 days occur when multiple detections chain together sequentially.

A critical finding emerges when comparing data-driven breakpoints with Meta's official policy announcements. The primary implementation breakpoint detected in our data (September 19, 2021) precedes Meta's announced "global implementation" date (July 19, 2022) by approximately 303 days. This substantial discrepancy suggests that Italian users experienced the policy effects considerably earlier than Meta's US-centric announcements would indicate.

This finding aligns with Meta's documented policy evolution. In October 2021, Meta announced it was "expanding the political content ranking tests to more countries around the world" (Meta Newsroom, October 13, 2021). Italy, as a major European market, likely received the policy during this expansion phase rather than waiting for the July 2022 "global" rollout.

The reversal breakpoint (March 9, 2025) follows Meta's January 7, 2025 announcement by approximately 61 days. This lag is consistent with the gradual rollout Meta described in its announcement, where the company indicated it would "start phasing this back into Facebook, Instagram and Threads with a more personalized approach" over time. The detected breakpoint thus captures when the policy reversal became sufficiently widespread to produce a measurable structural change in Italian parliamentarians' reach, rather than the initial announcement date.

Table 5 presents the magnitude of reach changes across the four policy phases defined by the detected breakpoints. Figure 1 visualizes these trends, displaying the weekly time series of average views per post for re-elected MPs with detected breakpoints (T_1, T_2, T_3) and election windows indicated.

Table 5. Reach Statistics by Policy Phase (Re-elected MPs).

Phase	N Weeks	Mean Views	Median Views	SD	Change
0: Pre-Policy	37	53,368	49,157	11,966	—
1: Policy-Active	67	26,079	25,025	8,420	−51.1%
2: Adjusted-Policy	114	14,869	13,573	4,857	−43.0%
3: Post-Reversal	38	34,918	35,185	9,589	+134.8%

Table 6. Engagement Metrics by Policy Phase (Re-elected MPs).

Phase	Reactions (Mean)	Shares (Mean)	Comments (Mean)
0: Pre-Policy	948.3	155.8	227.0
1: Policy-Active	669.1	81.7	139.8
2: Adjusted-Policy	396.5	44.3	111.0
3: Post-Reversal	1,199.7	142.4	370.3

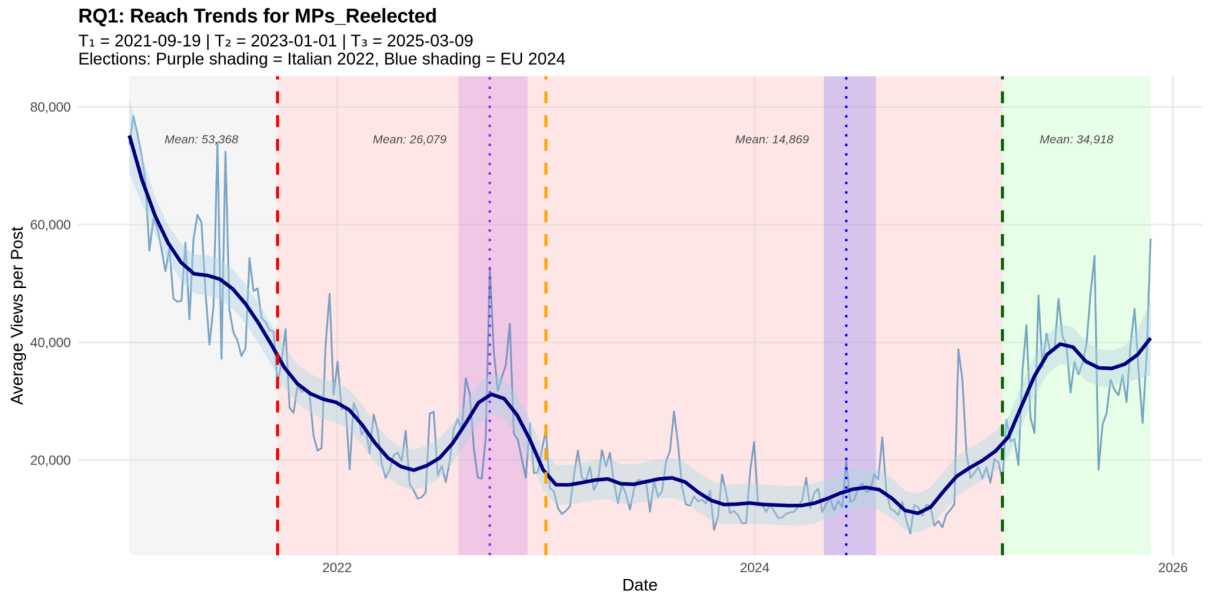

Figure 1. Reach trends for re-elected MPs with data-driven breakpoints. Vertical dashed lines indicate detected breakpoints: T_1 (September 2021, policy implementation), T_2 (January 2023, adjustment), T_3 (March 2025, reversal). Purple and blue shading indicate Italian 2022 and EU 2024 election windows respectively. Phase means are annotated.

The January 2023 Structural Break

A notable finding emerges from the breakpoint analysis: the January 1, 2023 breakpoint is detected with maximum consensus strength (13 different detections), making it one of the strongest structural breaks in the entire dataset. This breakpoint coincides with the conclusion of the Italian election period (September 2022 – December 2022).

The election window analysis reveals a significant temporary effect. As shown in Figure 1 (purple shading), during the election period (August 2022 – November 2022), mean views increased compared to non-election policy-active weeks. During the election period (August

2022 – November 2022), mean views increased compared to non-election policy-active weeks. This "election bounce" is statistically significant but transient—reach returned to suppressed levels after January 2023.

The abrupt January 2023 decline is consistent with either: (1) temporary policy suspension during the election followed by reactivation; (2) organic behavioral changes as heightened election-period interest subsided; or (3) coincidental algorithm updates. The strength of the breakpoint and its precise timing at the calendar year transition favors interpretation (1), though the current analysis cannot definitively distinguish between explanations.

Summary of Policy Effects

The magnitude analysis reveals a pronounced contraction in reach following policy implementation, followed by a statistically significant but incomplete recovery after the reversal. Relative to the pre-policy baseline, average views per post declined by 72.1%, falling from 53,368 in the initial phase to a trough of 14,869 during the adjusted-policy period. This contraction unfolded in two stages: an initial 51% decrease at the time of policy implementation (T_1), followed by a further 43% reduction associated with the January 2023 policy adjustment (T_2). Although the January 2025 reversal produced a substantial rebound, with average reach increasing by 134.8% from the trough to 34,918 views per post, post-reversal engagement remained markedly below baseline levels. Overall, recovered reach amounts to only 65.4% of pre-policy visibility, indicating that more than one-third of the initial loss persisted beyond the policy rollback. This pattern suggests that the effects of the intervention were not fully reversible, pointing either to continuing algorithmic constraints or to longer-term shifts in user behavior that emerged during the policy period.

Validation Across Political Actor Groups

To assess the generalizability of the detected breakpoints, the study examines whether the same directional pattern (DOWN → DOWN → UP) appears in the validation samples when applying the breakpoints discovered from re-elected MPs (Table 7).

Table 7. Breakpoint Validation Across Groups.

Group	Pattern	T_1 Valid?	T_3 Valid?	Kruskal-Wallis p
MPs (Re-elected)	DOWN → DOWN → UP	✓	✓	$< 2\times10^{-16}$
MPs (New)	DOWN → DOWN → UP	✓	✓	$< 2\times10^{-16}$

Prominent Politicians	DOWN → DOWN → UP	✓	✓	$< 2\times10^{-16}$
Extremists	DOWN → DOWN → DOWN	✓	✗	2.29×10^{-15}

The expected DOWN → DOWN → UP pattern is validated in 3 of 4 groups. The Extremists group shows an anomalous pattern (DOWN → DOWN → DOWN), with reach declining at T_1 and continuing to decline even after the policy reversal at T_3. Figure 2 visualizes these divergent trajectories across all four political actor groups, with the smoothed trend lines revealing the distinctive patterns identified through breakpoint detection.

All groups show highly significant phase differences ($p < 0.001$), confirming that the detected breakpoints identify genuine structural changes in reach dynamics regardless of the directional pattern.

Cross-Group Magnitude Comparison

Table 8 presents the magnitude of reach changes across all four political actor groups, enabling comparison of policy effects by group type.

Table 8. Cross-Group Magnitude Comparison (Views).

Group	Phase 0	Phase 1	Phase 2	Phase 3	Peak-to-Trough Δ
MPs (Re-elected)	53,368	26,079	14,869	34,918	−72.1%
MPs (New)	27,523	17,672	13,424	40,809	−51.2%
Prominent Politicians	84,670	49,262	36,153	58,187	−57.3%
Extremists	12,529	12,075	9,484	9,218	−24.3%

Note: Peak-to-Trough Δ = percentage change from Phase 0 (Pre-Policy) to Phase 2 (Adjusted-Policy trough).

The cross-group comparison reveals consistent patterns across mainstream political actors: MPs (both re-elected and new) and Prominent Politicians all show substantial reach declines of 51–72% from baseline to trough, followed by partial recovery in Phase 3. Note that for new MPs, this comparison uses a pre-MP baseline. The Extremists group presents a distinct pattern: their reach declined more modestly (−24.3%) and continued declining even after the policy reversal (Phase 2 → Phase 3: −2.8%).

To confirm that the observed phase differences are statistically significant, Kruskal-Wallis tests were conducted for each group, followed by Dunn's post-hoc pairwise comparisons with

Bonferroni correction. All groups show highly significant overall phase effects ($p < 0.001$) (Table 9).

Table 9. Pairwise Phase Comparisons (Dunn's Test, Bonferroni-adjusted).

Comparison	MPs (Re-el.)	MPs (New)	Prom. Pol.	Extremists
Phase 0 vs Phase 1	***	***	***	n.s.
Phase 0 vs Phase 2	***	***	***	***
Phase 1 vs Phase 2	***	***	***	***
Phase 2 vs Phase 3	***	***	***	n.s.
Phase 0 vs Phase 3	*	*	**	***

*Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; n.s. = not significant. Tests use Bonferroni correction for multiple comparisons.*

The pairwise comparisons reveal nuanced recovery patterns. For Extremists, both the initial implementation (Phase 0 to 1) and post-reversal transition (Phase 2 to 3) are non-significant, confirming their anomalous trajectory where neither the policy implementation nor reversal had immediate effects. Most notably, MPs (New) show marginally significant Phase 0 vs Phase 3 differences, indicating substantial but incomplete recovery to pre-policy baseline levels. Re-elected MPs show a similar pattern. In contrast, Prominent Politicians show significant Phase 0 vs Phase 3 differences at the $p < 0.01$ level, while Extremists show highly significant differences indicating continued decline rather than recovery.

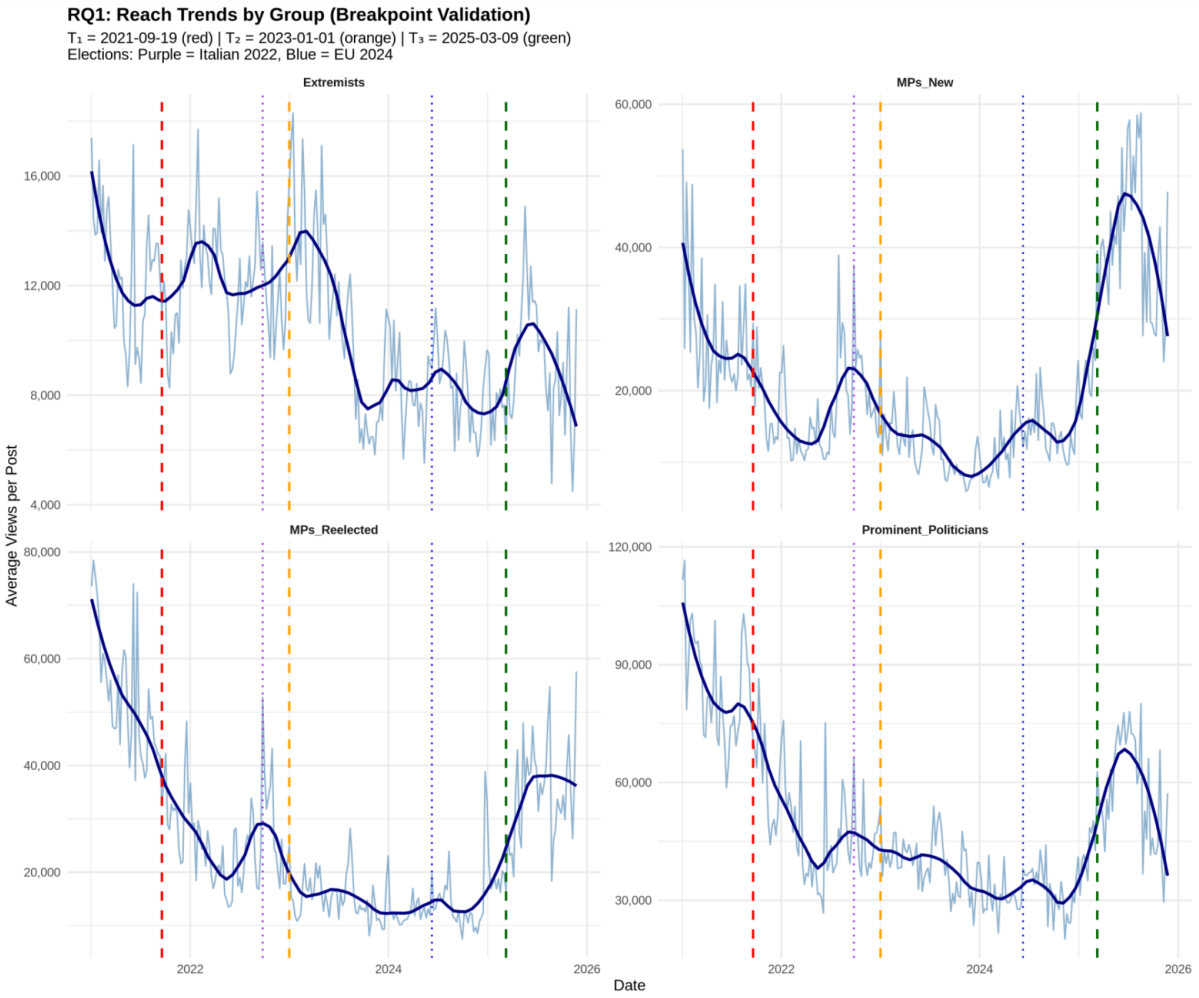

Figure 2. Individual group trends with detected breakpoints.

Robustness Check: Total Weekly Reach

The preceding analyses use per-post average reach as the primary outcome measure. However, groups differ substantially in posting frequency: Extremists post approximately three times more on average and up to five times more than some groups. This raises the question of whether high posting volume might compensate for lower per-post reach, producing different patterns when examining total weekly visibility (sum of all post views) rather than per-post averages.

Table 10. Per-Post vs Total Weekly Reach by Group.

Group	Per-Post Phase 0	Per-Post Phase 2	Δ	Total Phase 0	Total Phase 2	Δ
Extremists	12,529	9,484	−24.3%	43.4M	49.3M	+13.7%
MPs (New)	27,523	13,424	−51.2%	36.3M	18.5M	−49.1%

MPs (Re-elected)	53,368	14,869	−72.1%	127.5M	25.1M	−80.3%
Prominent Politicians	84,670	36,153	−57.3%	167.3M	42.2M	−74.8%

Note: Per-post figures show average views per post. Total figures show the sum of all post views per week (in millions). Δ = percentage change from Phase 0 to Phase 2.

The robustness check reveals a striking divergence. For mainstream political actors (MPs and Prominent Politicians), the directional pattern is consistent regardless of metric: both per-post and total reach show the DOWN → DOWN → UP pattern identified in the primary analysis. However, Extremists show fundamentally different patterns depending on the metric used:

- Per-post reach: DOWN → DOWN → DOWN (modest decline, continued decline after reversal)
- Total weekly reach: UP throughout all phases (continuous growth)

This divergence has substantive implications. While Extremists experienced modest per-post reach reductions during the policy period (−24.3%), their high posting volume more than compensated: total weekly visibility increased by 13.7% from pre-policy to adjusted-policy phases.

The ranking implications are particularly noteworthy. In the pre-policy period, Prominent Politicians held the highest total weekly reach (167.3M), followed by MPs Re-elected (127.5M), with Extremists a distant third (43.4M). By the adjusted-policy phase, Extremists had risen to first place in total weekly reach (49.3M), surpassing both Prominent Politicians (42.2M) and MPs (25.1M and 18.5M).

Table 11. Posting Frequency Changes from Pre-Policy Baseline.

Group	**Phase 0 (Baseline)**	**Phase 1**	**Phase 2**	**Phase 3**
Extremists	3,464	−1.5%	+61.5%	+140.5%
MPs (New)	1,322	+0.3%	+5.4%	+17.6%
MPs (Re-elected)	2,368	−6.3%	−26.4%	−28.4%
Prominent Politicians	1,995	−15.2%	−41.4%	−42.9%

Note: Phase 0 shows average weekly posts; subsequent columns show percentage change from baseline. Phase 0 = Pre-Policy (Jan–Sept 2021); Phase 1 = Policy-Active (Sept 2021–Dec 2022); Phase 2 = Adjusted-Policy (Jan 2023–Mar 2025); Phase 3 = Post-Reversal (Mar 2025–present).

These findings suggest that Meta's political content reduction policy had asymmetric effects on different types of political actors. Mainstream politicians—who typically post less frequently but achieve higher per-post engagement—experienced genuine reductions in total visibility. Extremist accounts, by contrast, appear to have maintained or increased total visibility through an increased posting volume. The timing of the volume increase aligns with the detected policy breakpoints: extremist posting remained stable through Phase 1, then increased sharply following the January 2023 adjustment (Table 11). Additionally, as noted in Section 2.2, Meta's view metric counts impressions on both original posts and reshares; if extremist content circulates extensively through resharing networks, this mechanism could amplify total view counts beyond what posting frequency alone would produce. Whether the observed pattern reflects strategic posting adaptation, differential resharing dynamics, or pre-existing behavioral differences requires further investigation.

Summary of the Findings

The analysis produces five central findings. First, the effects of the policy in Italy emerge considerably earlier than Meta's own timeline suggests: structural changes in reach are detectable from September 2021, roughly ten months before the company's July 2022 announcement of global implementation. Second, policy implementation is associated with a sharp contraction in visibility, reducing Italian parliamentarians' average post reach by approximately 72%, from around 53,000 views per post before the intervention to about 15,000 at the trough. Third, this decline does not occur as a single break but unfolds in two distinct stages, with an initial 51% drop at the time of implementation (T_1) followed by a further 43% reduction in January 2023 (T_2), a pattern that may reflect either a tightening of enforcement or a post-election normalization of engagement. Fourth, although the policy reversal produces a statistically significant rebound, recovery remains incomplete: post-reversal reach stabilizes at only about 65% of pre-policy levels. Finally, the timing of the detected reversal breakpoint follows Meta's policy announcement by approximately 61 days, consistent with the gradual rollout the company described and providing validation for the data-driven identification strategy's capacity to capture real-world platform interventions.

Discussion

The findings presented above warrant reflection on both what they reveal about platform governance and what they suggest about the conditions under which such findings become possible. This discussion addresses three interrelated themes: the transparency deficits that this study exposes, the transparency progress that made this study possible, and the implications for how researchers, policymakers, and platforms might productively engage with one another going forward.

Transparency Deficits

This opacity is particularly striking given the policy's historical positioning: the initial announcement in February 2021 was clearly situated in the aftermath of the January 6th Capitol invasion and the transition to the Biden administration, while the January 2025 reversal—announced just thirteen days before the second Trump inauguration—coincided with yet another American political transition. Both pivotal policy decisions were thus temporally anchored to US domestic politics, a distinctly American context that offers little guidance for understanding implementation timelines in other democracies.

The persistence of a US-centric approach to policy communication is notable. Meta's announcements referenced American user feedback, American political events, and American implementation timelines, even as the policy was being deployed across diverse political systems with different electoral calendars, media landscapes, and democratic traditions.

This US-centrism extends beyond the political content policy to Meta's broader approach to research transparency. The company's most prominent collaboration with academic researchers—the 2020 US Election study—produced a series of high-profile publications in Science and Nature examining platform effects on political attitudes and behavior. Yet these studies, conducted by predominantly US-based researchers, focused exclusively on the American electoral context. No comparable research collaboration has examined platform effects in European democracies, despite the significant regulatory attention the European Union has directed toward platform governance. Italian parliamentarians experienced a 51% reduction in reach coinciding with what appears to have been the policy's European expansion in autumn 2021, yet this occurred without the kind of rigorous academic scrutiny that accompanied Meta's US-focused research initiatives.

Equally concerning is the absence of this policy from Meta's official transparency documentation. The political content reduction mechanism—which this study estimates reduced Italian parliamentarians' reach by up to 72%—does not appear in the company's Digital Services Act compliance reports or other transparency disclosures to European institutions. This omission is difficult to reconcile with the DSA's emphasis on algorithmic transparency and the significant democratic implications of systematically reducing elected representatives' visibility.

The findings also reveal an unintended asymmetry in policy effects. While mainstream political actors experienced substantial reach reductions, extremist accounts responded by dramatically increasing their posting frequency (+61.5% by Phase 2, +140.5% by Phase 3)—effectively compensating for lower per-post visibility through sheer volume. This divergence suggests that algorithmic suppression may be least effective against the very actors it most plausibly targets. Parliamentarians and other institutional figures, whose communication patterns are constrained by their official roles, absorbed the full impact of reduced reach. Extremist accounts, unconstrained by such expectations, adapted their behavior to circumvent the policy's effects.

Transparency Progress

Yet the very existence of this study testifies to meaningful progress in platform transparency. The Meta Content Library, which provided the data for this analysis, represents a considerable investment by Meta in enabling independent research on its platforms. The infrastructure required to provide researchers with access to post-level engagement metrics, view counts, and account metadata across billions of pieces of content is substantial, and the company's ongoing maintenance and improvement of this system should not be taken for granted.

The Meta Content Library can be understood as a direct fruit of the Digital Services Act's Article 40.12 provisions requiring very large online platforms to provide vetted researchers with access to public data necessary to undertake research contributing to the detection, identification, and understanding of systemic risks. While the current implementation has limitations—the virtual clean room approach restricts data portability, the export process for aggregated results requires review, and the learning curve for the secure research environment is not trivial—these constraints reflect genuine tensions between research access and user privacy that do not admit easy solutions.

The progress should be measured carefully against what came before. CrowdTangle, which Meta acquired in 2016 and shut down in August 2024, provided researchers and journalists with valuable access to public engagement data on Facebook and Instagram for nearly a decade. Its closure was met with concern from the research community. Yet CrowdTangle had significant limitations. The Meta Content Library, while more restrictive in its clean room approach, offers capabilities that CrowdTangle lacked—including the view data that made it possible to detect the structural breaks documented in this paper. The transition represents a trade-off rather than unambiguous progress: greater depth of data access, but within a more controlled environment.

Toward Collaborative Platform Governance

These observations point toward a broader reframing of how stakeholders engage with questions of platform governance. There is a temptation—understandable given the power asymmetries involved—to use transparency provisions primarily as instruments for criticism, treating each new dataset as an opportunity to document platform failures. But this adversarial framing carries risks. Using transparency efforts against the platforms that provide them is neither fair nor strategically wise: it creates incentives for minimal compliance rather than genuine openness, and it may jeopardize the progress achieved thus far.

A more productive orientation would recognize that studies like this one, which illuminate the limits and unintended consequences of platform governance decisions, are not indictments but contributions. Governing socio-technical systems of Facebook's scale and complexity is genuinely difficult. No organization, regardless of its resources or intentions, can fully anticipate how algorithmic changes will interact with diverse user behaviors, political contexts, and information ecosystems across dozens of countries. The finding that extremist accounts maintained or increased total visibility despite per-post reach reductions—likely through volume compensation and resharing dynamics—exemplifies the kind of emergent complexity that only becomes visible through careful external analysis.

Companies like Meta are private entities that primarily pursue economic objectives. The DSA imposes transparency requirements, and external researchers can scrutinize platform activities. But the purpose of this scrutiny should be understood as helping platforms themselves better understand the complexity of their societal impact—a complexity that exceeds any single organization's capacity to fully comprehend. The more diverse and

wide-ranging the community of researchers analyzing social media, the better positioned platforms will be to steer these complex systems in less harmful directions.

It is worth acknowledging a structural reality that shapes all discussions of platform transparency: genuine transparency cannot truly be imposed on these platforms from the outside. In the final analysis, platforms and only platforms have access to the data warehouses where the evidence resides. Any tool that presents these data in a form that external researchers can analyze will necessarily be designed and developed by the platforms themselves. It will always be possible, in principle, for them to manipulate what researchers see. This is a structural condition we cannot change and must accept.

Yet the same complexity that makes governing these platforms so challenging also makes systematic deception difficult to execute. Hiding evidence by design would require the platform to possess a clear and comprehensive understanding of all its own shortcomings—precisely the kind of self-knowledge that this paper argues platforms lack. Furthermore, platform data must maintain internal consistency: metrics need to add up, time series must cohere, and cross-tabulations must align. Any attempt to selectively manipulate the data exposed to researchers risks generating inconsistencies that careful analysis could detect. Moreover, as the Knight-Georgetown Institute's Better Access framework emphasizes, independent data collection methods—including web scraping and crawling—serve as a crucial verification mechanism for platform-provided data (Chapman et al., 2025). The framework explicitly recommends that proactive data interfaces be "auditable, to verify the accuracy of data provided through the mechanism (including through Independent Data Collection)." This creates a practical check: researchers with appropriate technical capabilities can independently collect samples of public platform data and compare them against what platforms provide through official channels. Any systematic discrepancy between scraped data and API-provided data would be detectable, creating an additional layer of accountability beyond internal consistency checks alone. The potential backfire from discovering such a dishonest approach would be enormous—far exceeding whatever embarrassment the original findings might have caused. Regulatory consequences, reputational damage, and the collapse of researcher trust would follow any credible demonstration that a platform had deliberately falsified its transparency infrastructure.

This creates an imperfect but meaningful equilibrium. Platforms could theoretically deceive researchers, but the practical difficulty and catastrophic downside risks of doing so create

strong incentives for honest disclosure. External researchers cannot verify that they are seeing complete and accurate data, but they can—through careful analysis—identify patterns, inconsistencies, and unexpected findings that would be difficult to anticipate and preemptively sanitize. The collaborative orientation advocated here rests not on naive trust but on a realistic assessment of the incentive structures that shape platform behavior.

This does not mean abandoning critical analysis or excusing genuine failures of responsibility. The transparency deficits documented by this study are real and consequential. But it does mean recognizing that the path toward better platform governance runs through collaboration rather than pure confrontation. Policymakers, researchers, journalists, and platforms share an interest in understanding how algorithmic systems affect democratic discourse, individual well-being, and social cohesion. Adversarial dynamics—where each stakeholder group treats the others primarily as obstacles or targets—make this shared understanding harder to achieve.

The DSA's transparency provisions represent a policy framework oriented toward this collaborative vision. By requiring data access for vetted researchers, the regulation creates conditions for independent analysis without presuming that such analysis must be antagonistic. Studies like this one can identify problems, surface unintended consequences, and generate evidence that informs both platform decisions and regulatory responses. But they can only do so if the research infrastructure continues to develop—which in turn requires that platforms see transparency as serving their interests rather than merely exposing their vulnerabilities.

The findings presented in this paper—documenting substantial reach reductions for Italian parliamentarians, asymmetric effects across political actor types, and incomplete recovery following policy reversal—should be understood in this spirit. They contribute to a growing body of evidence about how algorithmic content moderation affects political communication, evidence that platforms, policymakers, and researchers alike need in order to navigate the difficult trade-offs that platform governance inevitably involves.

Limitations and Future Directions

Several limitations constrain the causal interpretation of these findings. The analysis relies on aggregate reach metrics without direct observation of Meta's algorithmic treatment. Alternative platform-wide changes—interface modifications, user behavior shifts,

competition from other platforms—may contribute to observed patterns. The election-period analysis cannot definitively distinguish policy modification from organic behavioral changes.

A further methodological constraint concerns the composition of reach metrics provided by the Meta Content Library. As confirmed by Meta's research support team, view counts aggregate both organic reach and paid/boosted views without distinction. If a post received 50,000 organic views and 200,000 paid/boosted views, the API returns only the combined total (250,000). The same aggregation applies to engagement metrics (reactions, comments, shares). This means the observed reach reductions could theoretically be confounded by changes in politicians' paid promotion strategies during the study period—for instance, if political actors reduced advertising spending concurrent with the policy implementation. However, this alternative explanation seems unlikely to account for the magnitude and timing of effects observed, particularly since the reach decline preceded Meta's announced policy date by ten months and affected groups with varying resources for paid promotion. Nevertheless, future research with access to advertising expenditure data could provide valuable triangulation of these findings.

A promising avenue involves constructing a model that predicts post reach from engagement signals (reactions, shares, comments). Meta's stated policy involved "de-emphasizing engagement signals" when ranking political content. If this description is accurate, the relationship between engagement metrics and reach should weaken during policy-active periods.

Additionally, the study is circumscribed to the Italian political context, providing depth but limiting generalizability. Replication studies in other countries are essential to determine whether the patterns observed here—particularly the ~72% reach reduction, the timing discrepancy with Meta's announced timeline, and the incomplete post-reversal recovery—represent general policy effects or Italy-specific phenomena. To facilitate such replication efforts, the complete analytical code, data processing pipelines, and methodological documentation are available in the study's online repository (see Replication Materials).

Priority should be given to countries where: (1) Meta announced specific policy rollout dates, enabling precise timeline comparisons; (2) parliamentary or electoral data is publicly accessible; (3) major elections occurred during the policy period, allowing examination of potential election-period modifications; and (4) the political media ecosystem differs

substantially from Italy's. Comparative analysis across multiple countries would strengthen causal inference and reveal whether Meta's policy implementation was uniform or adapted to local contexts.

Lastly, the study is limited to Facebook, yet Meta's political content reduction policy was announced as applying across its platforms, including Instagram and Threads. Replication on Instagram is particularly important given its distinct user demographics, content formats, and algorithmic architecture. Instagram's emphasis on visual content, Stories, and Reels may produce different policy effects than Facebook's text-and-link-heavy political discourse. Additionally, Instagram's younger user base may exhibit different engagement patterns with political content, potentially amplifying or attenuating policy effects. Comparative analysis across Facebook and Instagram for the same set of political actors would reveal whether Meta implemented its policy uniformly across platforms or adapted the approach to each platform's characteristics.

Replication Materials

To facilitate replication and extension of this study, all analysis code and documentation are available in a public GitHub repository. The complete analytical pipeline—from data download through breakpoint detection—is provided as a series of documented R notebooks that can be executed within the Meta Secure Research Environment (SRE) or other approved platforms with Meta Content Library API access.

Code Availability

The complete analytical pipeline is available at: https://github.com/fabiogiglietto/mcl-political-reach-study.

The repository includes seven core notebooks: (00) data download from MCL API, (01) dataset construction and standardization, (03) data cleaning and validation, (04) account metadata enrichment, (05) producer list verification, (06) breakpoint analysis, and (07) working paper outputs. Each notebook contains detailed documentation and is designed to run sequentially within the secure computing environment. The repository also provides configuration templates for adapting the methodology to other country contexts.

Data Access Requirements

Replication requires approved researcher access to the Meta Content Library API. The Meta Content Library provides comprehensive access to publicly available content from Facebook and Instagram, including post-level engagement metrics and view counts. Researchers must apply for access through Meta's official application process and conduct all data querying and analysis within Meta's secure computing environment. Raw post-level data cannot be downloaded for local analysis due to Meta's data protection policies.

Producer Lists

The original producer lists used in this study are publicly accessible to any researcher with Meta Content Library API access. These lists can be imported directly into new research projects, enabling exact replication of the sample. The following producer lists are available:

MPs (Re-elected): Members of Parliament elected in both 2018 and 2022 legislatures. https://www.facebook.com/transparency-tools/content-library/producer-lists/1508867967036191/

Parlamentari_ITA_Leg_XIX: All members of the XIX Italian legislature (2022–present). https://www.facebook.com/transparency-tools/content-library/producer-lists/1079401170932761/

Prominent Politicians: Most-followed Italian politicians not holding parliamentary seats. https://www.facebook.com/transparency-tools/content-library/producer-lists/2018160402336859/

Extremists (Cluster 1): Italian pages and public figures in the alternative media ecosystem. https://www.facebook.com/transparency-tools/content-library/producer-lists/1546733040084559/

Extremists (Cluster 2): Additional Italian pages and public figures in the alternative media ecosystem. https://www.facebook.com/transparency-tools/content-library/producer-lists/2543309726299 3192/

Cross-Country Replication

The repository includes a detailed replication guide (REPLICATION_GUIDE.md) for adapting the methodology to other national contexts. Researchers seeking to examine Meta's political content policy in other countries should: (1) identify analogous political actor

categories (elected representatives, prominent non-elected politicians, and comparison groups); (2) create producer lists in the MCL interface for each category; (3) adapt the configuration files to specify country-specific parameters such as election dates and policy timeline expectations; and (4) execute the pipeline notebooks in sequence. The modular design allows researchers to modify individual components while maintaining methodological consistency with the original study.

Researchers should note that while code and producer lists are publicly available, the raw post-level data cannot be shared outside the secure environment. Replication therefore requires independent data collection through the MCL API. Aggregated results and statistical outputs can be exported following Meta's review process for non-sensitive materials.

References

Chapman, P., et al. (2025). *Better Access: Data for the Common Good*. Knight-Georgetown Institute. https://kgi.georgetown.edu/research-and-commentary/better-access/

Cunningham, T., Pandey, S., Sigerson, L., Stray, J., Allen, J., Barrilleaux, B., Iyer, R., Milli, S., Kothari, M., & Rezaei, B. (2024). What we know about using non-engagement signals in content ranking. In arXiv [cs.SI]. arXiv. http://arxiv.org/abs/2402.06831

Meta Platforms, Inc. (2024). *Meta Content Library and Content Library API*. Developer Documentation. https://developers.facebook.com/docs/content-library-and-api

Kaplan, J. (2025, January 7). More speech and fewer mistakes. *Meta Newsroom*. https://web.archive.org/web/20251227094726/https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/

Stepanov, A., & Gupta, A. (2021, February 10). Political content in feeds [Updated May 28, 2025]. *Meta Newsroom*. https://web.archive.org/web/20251011124003/https://about.fb.com/news/2021/02/reducing-political-content-in-news-feed/