5 Main Results: Virality of Political Narratives

Access to information has reached unprecedented levels, nevertheless the impact of facts is often overshadowed by the power of narratives (Bursztyn et al. 2023). Narratives do not merely describe facts; they shape perceptions, opinions, and behaviors, often leading to measurable economic and social consequences (Andre et al. 2025). One of the primary mechanisms through which narratives gain traction and influence is virality, particularly on social media (Shiller 2017). Understanding what makes narratives go viral is therefore crucial to assessing their broader societal impact. In this section, we analyze the virality of narratives, first examining the fundamental nature of virality on Twitter, then exploring the relationship between narratives and virality, and finally identifying the key determinants that drive a narrative's virality.

5.1 Virality of Character-Roles

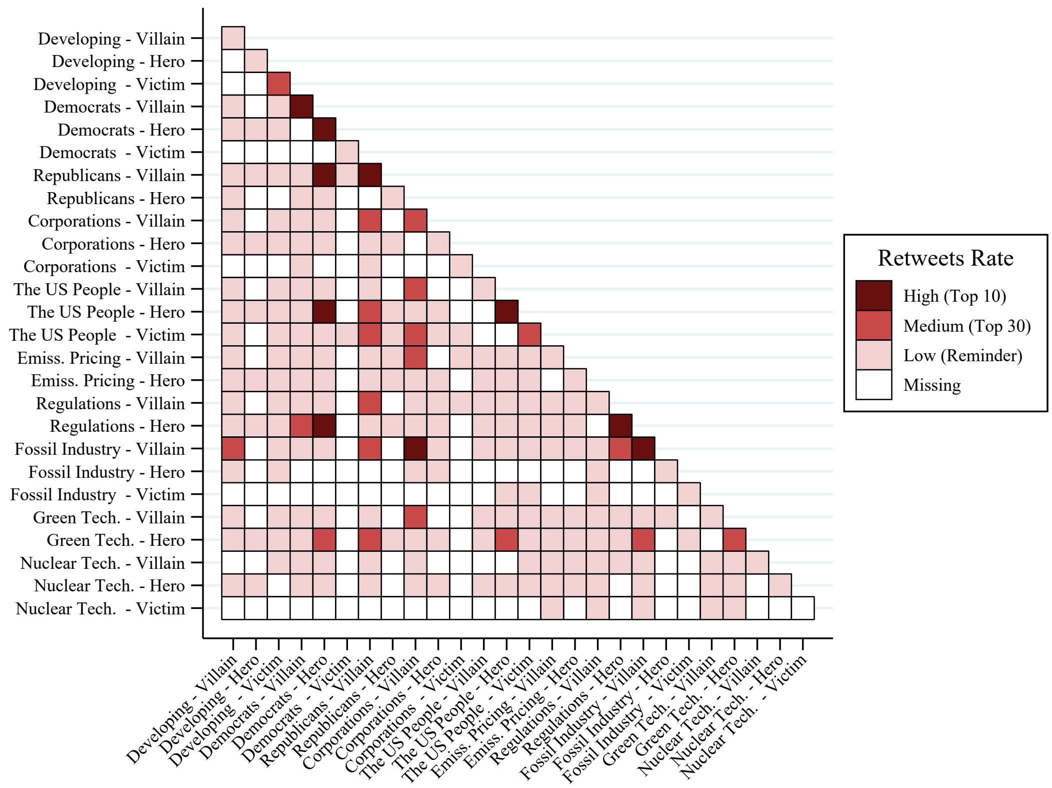
We define virality by the number of retweets a tweet receives. The first research question we address is straightforward: how viral are the *political narratives* in our dataset? We exploit the flexibility and capacity of our framework, and analyze virality for all combinations of one or two character-roles in Figure 5. The squares of the matrix capture combinations of character-roles, and the color of the squares indicates the level of retweets obtained by the specific character-role combination, out of the total amount of retweets obtained by any narrative containing one or two character-roles. *Political narratives* with single character-roles are shown on the diagonal, complex narrative combinations in the other entries of the matrices. The matrix is symmetrical.

Certain character-role combinations prove especially viral. Among the five most retweeted narratives, the most common pattern sees US DEMOCRATS as heroes and US REPUBLICANS as villains. Narratives that portray US PEOPLE as heroes likewise attract high retweet counts, whether they appear alone or alongside other roles. Depictions of the FOSSIL INDUSTRY as a villain also spread widely, particularly when CORPORATIONS are portrayed alongside as villains, a pairing that itself ranks in the top five. Although narratives featuring human actors tend to dominate, non-human roles such as REGULATIONS and GREEN TECH still break through, appearing in several of the thirty most retweeted narratives.

While Figure 5 provides a descriptive overview of the average virality of *political narratives*, it does not speak to the difference between the virality of *political narratives*, versus that of tweets only featuring characters in a neutral way. Figure 6 provides descriptive insights into this distribution. It displays the Log-Log Rank Distribution of retweets, distinguishing between tweets that feature at least one character-role from those that only feature characters in a neutral way.

In the figure, the x-axis represents the logarithm of the rank of tweets based on their retweets, respectively, with rank 1 corresponding to the most retweeted tweet in the dataset. The y-axis represents respectively the logarithm of the number of retweets, plus one. The distribution reveals key insights. First, for both neutral tweets and tweets containing a character-role, the distribution takes a linear shape. This suggests that virality follows a power-law rank distribution among the

Figure 5: *Virality of Character-Roles - Tweets with One or Two Character-Roles*



Notes: The figure shows the retweet rate of each character–role appearing alone or in combination with another role, among relevant tweets containing one or two roles. Retweet rates are computed as the share of total retweets received by a given role (or pair) relative to all retweets of tweets with one or two roles. The diagonal of the matrix shows the retweet rate when each character-role appears alone. Tweets with three or more character-roles are excluded. To avoid visual overload, we do not display exact rates. Instead, we use a color scheme to highlight the top 10 most frequently retweeted character-role combinations, the top 30 (which includes the top 10), and the remaining pairs. White indicates a pair that never appears together. The top 10 in order is: US Republicans-Villain (16.72%), US Democrats-Hero (12.98%), US People-Hero (7.92%), Fossil Industry-Villain + Corporations-Villain (7.89%), US Republicans-Villain + US Democrats-Hero (6.51%), US Democrats-Villain (4.79%), Fossil Industry-Villain (3.07%), US People-Hero + US Democrats-Hero (2.70%), Regulations-Hero (2.57%), Regulations-Hero + US Democrats-Hero (2.51%).

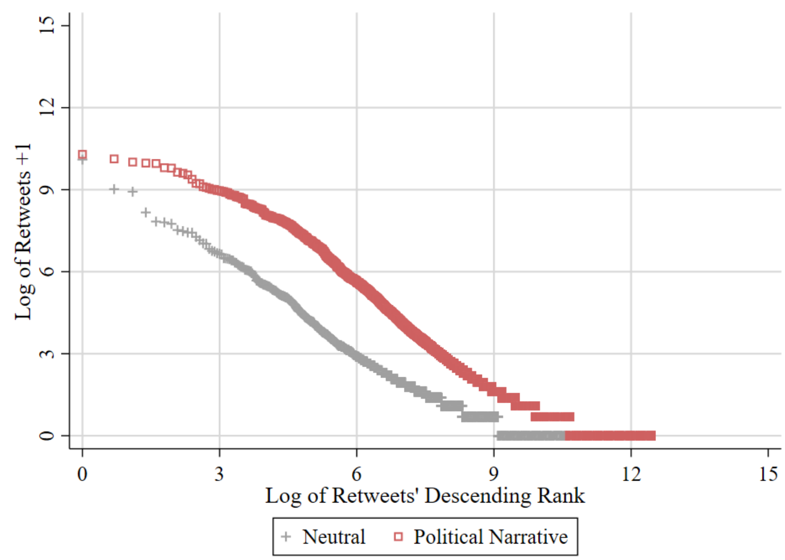
tweets in our dataset, which can be represented by the following functional form:

$$R(x) \propto x^{-\beta} \tag{1}$$

where $R(x)$ represents the expected number of retweets at rank x, and β is the scaling exponent that determines how quickly engagement declines as rank increases. A higher value of β results in a steeper decline, meaning engagement is concentrated in a few highly viral tweets, while most receive minimal interaction.

The exponent β directly influences the slope of the curves in the log-log distribution graphs. A steeper slope (β large) indicates a rapid drop from the most retweeted tweets to the least, suggesting

Figure 6: *Log-Log Rank Distribution of Retweets in Relevant Tweets (United States, 2010-2021)*



Notes: The figure shows the log–log rank distribution of retweets for relevant tweets, distinguishing between those with at least one character–role (red) and those with characters only in a neutral form (gray). The x-axis plots the logarithm of the rank, with rank 1 as the most retweeted tweet in the sample. The y-axis plots the logarithm of retweet counts (plus one). The slope of the curve indicates how quickly the distribution falls from the most viral to the least viral tweets: a steeper slope means engagement is clustered in a handful of viral tweets, whereas a flatter slope indicates that engagement is more evenly distributed. The vertical position at a given rank reflects relative virality: a higher curve means tweets at that rank receive more engagement than tweets at the same rank in a dataset with a lower curve.

that engagement is highly concentrated among a few viral tweets. Conversely, a flatter slope (β small) implies a slower decline, meaning engagement is more evenly distributed across tweets. The vertical position of the curves, at the same rank, provides additional information: at parity of rank, a higher-positioned curve signals a greater level of engagement, meaning that tweets in this subset receive more retweets or likes compared to those in another subset with a lower-positioned curve. This distinction is particularly relevant when comparing narrative tweets with neutral ones. The figures show how the distribution for tweets with narratives is vertically shifted, indicating that at parity of rank narrative tweets are generally more viral.

5.2 The Determinants of Virality

Beyond more descriptive results, we also use a regression framework to analyze the determinants of virality more generally. We use the following Poisson Pseudo-Maximum Likelihood regression equation:

$$E[Y_{i,s,t}|D_{i,s,t}] = exp[\alpha + \beta D_{i,s,t} + \theta X_{i,s,t} + \gamma_t + \delta_s] \quad \forall i \in C, \tag{2}$$

where $Y_{i,s,t}$ refers to the count of retweets for tweet i, originating from state s in year t. Each tweet i in the model is such that $i \in C$, where C is the sample of tweets located in the US with precision at least at the state level. α is a constant term, while γ_t and δ_s refer respectively to the year and state fixed effects. $D_{i,s,t}$ refers to a dummy variable reflecting the comparison of interest. E.g., when comparing villain and hero narratives it equals 1 if a villain narrative is present in the tweet and 0 for hero narratives. We exclude tweets that feature neither of the two, as well as all tweets that feature both simultaneously. $X_{i,s,t}$ collects the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by the user. We cluster standard errors at the week level.

Before turning to estimates, it is useful to clarify how we read $\widehat{\beta}$ in (2) in light of potential omitted drivers of virality on the platform. Let the true process be

$$Y_{i,s,t} = \exp\{\alpha + \beta D_{i,s,t} + \gamma A_{i,s,t} + \theta' X_{i,s,t} + \gamma_t + \delta_s + u_{i,s,t}\},$$

where $A_{i,s,t}$ denotes omitted variables, e.g. unobserved ranking/moderation shifters employed by Twitter. Estimating (2) without $A_{i,s,t}$ yields the standard omitted-variable term

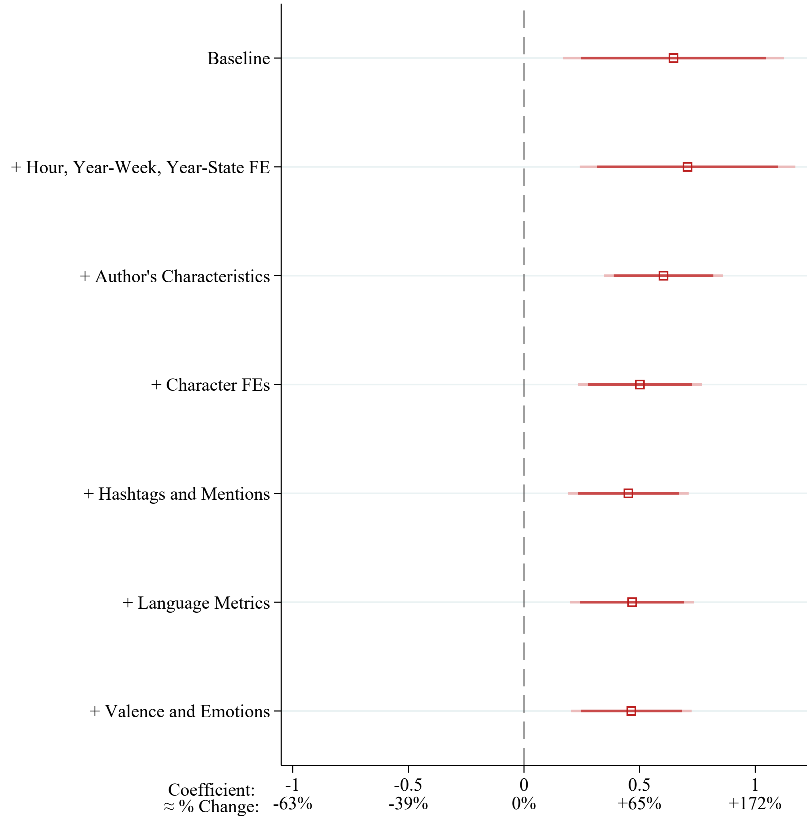
$$\gamma \cdot \frac{(D_{i,s,t}, A_{i,s,t} \mid X_{i,s,t}, \gamma_t, \delta_s)}{(D_{i,s,t} \mid X_{i,s,t}, \gamma_t, \delta_s)}$$

in $\widehat{\beta}$. If narrative tweets are systematically surfaced more ($\gamma > 0$ and $(D, A \mid \cdot) > 0$), $\widehat{\beta}$ is biased upward; if moderation disproportionately suppresses some role uses so that $(D, A \mid \cdot) < 0$, the bias is downward. This logic does not require time-variation in $A_{i,s,t}$; a time-invariant ranking rule can still generate bias whenever it differentially favors $D_{i,s,t} = 1$ tweets. In addition, misclassification in our role labels $D_{i,s,t}$ that is approximately mean-zero and independent of $Y_{i,s,t}$ conditional on covariates acts like classical measurement error, attenuating estimates toward zero under a linear approximation and, analogously, weakening PPML coefficients. Taken together, engagement-based surfacing can push $\widehat{\beta}$ upward, while moderation and measurement error plausibly pull it downward; the net bias is a priori ambiguous.

Our empirical design mitigates the most likely channels for $(D_{i,s,t}, A_{i,s,t} \mid \cdot)$ by conditioning on rich author and time effects and on text features – sentiment, emotions, length, hashtags, and mentions – that ranking systems can observe or proxy. Because the use of role language and emotional tone often co-occur within a tweet, conditioning on these features is appropriate rather than over-controlling. With these safeguards in place, we interpret $\widehat{\beta}$ as the association between character–role assignment and virality over and above observable author and text characteristics, and we place emphasis on within-sample contrasts across roles (hero, villain, victim versus neutral) and extensive robustness. We now turn to the estimates.

We begin our analysis by addressing the most basic question: Does using *political narratives* impact virality at all? Figure 7 provides insight into this question through a coefficient plot illustrating the impact of *political narratives* on virality compared to tweets where also at least one character appear, but only in neutral roles. The plot is structured top to bottom estimating a sequence of increasingly stringent models, beginning with a baseline specification without controls. The second

Figure 7: ***Regression Results - Impact of Political Narratives on Virality***



Notes: The figure shows the coefficients of Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring at least one character-role vs. featuring characters only in a neutral role on virality, measured as the count of retweets. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx e^{\beta} - 1$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour, week of the year, and year-state fixed effects. The third model controls for author characteristics: verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status. The fourth model adds character fixed effects. The fifth model controls for the tweet's number of hashtags and mentions. The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above.

model introduces hour and year-state fixed effects (FEs). The third model accounts for author characteristics, including whether the account is verified, number of followers and followings, total tweets created, and whether the author identifies as Democratic or Republican, religious, highly educated, or as a parent. The fourth model further controls for character fixed effects, and the fifth model incorporates tweet characteristics, such as word count, number of hashtags, and number of handles. The final column controls for emotions and valence.

Narratives have a persistent and positive impact on virality, an effect that remains significant even as we introduce increasingly stringent model specifications. One key factor influencing virality is the author's characteristics, which reduce the estimated effect size while simultaneously improving model precision by reducing noise. A natural concern is whether the observed effect is simply capturing the virality of the characters included in our analysis. However, even after introducing character fixed effects, the impact of narratives on virality remains substantial, suggesting that the effect is not merely driven by the presence of specific characters.

As discussed earlier, multiple stylistic text elements are able to construct *political narratives*, including tone, plot structure, character roles, and emotional framing. In previous sections, we examined how narratives differ from neutral tweets in terms of valence, emotional content, and text quality. While narratives exhibit clear distinctions along these dimensions, controlling for these factors does not eliminate their effect on virality. This suggests that narratives influence engagement through mechanisms beyond just emotion or linguistic style, reinforcing their distinct role in shaping public discourse.

Our definition and detailed, structured measurement of *political narratives* then allows us to refine our analysis by moving beyond the overall impact of featuring any narrative and addressing a more nuanced question: Which roles of the drama triangle drive virality most effectively, and do all roles contribute equally? We answer this question in Table 3, where present the effect of featuring specific character roles on virality.

To ensure a reliable and mutually exclusive comparison, the results focus solely on narratives containing a single character-role. We compare the effect of complex narrative in the next part of the analysis. Column 1 of the table reports the effect of featuring a hero narrative compared to neutral tweets, column 2 compares villain narratives to neutral tweets, and column 3 does the same for victim narratives. Since all samples are mutually exclusive, the estimated effects reflect the distinct impact of each role. Columns 4 through 7 introduce models that include all roles except for one, allowing for direct comparisons between them. In these models, the reference category changes across columns: column 4 uses neutral tweets as the baseline, column 5 hero narratives, 6 villain, and column 7 uses as baseline victim narratives.

Several noteworthy patterns emerge. Hero and villain narratives significantly enhance virality compared to tweets where characters appear in neutral roles. Column 1 shows that hero narratives increase retweets by approximately 55%, while villain narratives lead to a much stronger increase of 170% (computed as $\approx e^{\beta} - 1$ for Poisson models). In contrast, victim narratives do not exhibit a statistically significant effect on virality compared to neutral-character tweets, suggesting that they do not drive engagement in the same way. When considering the full models in columns 4 through 7, a more nuanced pattern emerges. Villain narratives consistently stand out, surpassing both hero and victim narratives in driving virality, reinforcing the idea that narratives centered around opposition and conflict tend to generate the highest levels of engagement. These findings suggest that negative framing, conflict-driven narratives, and oppositional storytelling play a crucial role in determining virality on social media.

Table 3: ***Regression Results - Impact of Individual Roles on Virality***

Dependent Variable	**Retweets' Count**						
	Coeff./SE/p-value						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hero	0.445			0.412		-0.552	0.239
	(0.193)			(0.211)		(0.210)	(0.279)
	[0.021]			[0.051]		[0.009]	[0.393]
Villain		0.923		0.964	0.552		0.791
		(0.135)		(0.198)	(0.210)		(0.291)
		[0.000]		[0.000]	[0.009]		[0.007]
Victim			0.090	0.173	-0.239	-0.791	
			(0.326)	(0.287)	(0.279)	(0.291)	
			[0.783]	[0.546]	[0.393]	[0.007]	
Neutral					-0.412	-0.964	-0.173
					(0.211)	(0.198)	(0.287)
					[0.051]	[0.000]	[0.546]
Sample: Neutral	✓	✓	✓	✓	✓	✓	✓
Sample: Hero	✓			✓	✓	✓	✓
Sample: Villain		✓		✓	✓	✓	✓
Sample: Victim			✓	✓	✓	✓	✓
Mean Outcome Reference Group	1.972	1.972	1.972	1.972	4.350	3.683	2.395
Pseudo R2	0.79	0.70	0.68	0.71	0.71	0.71	0.71
Obs	77047	90057	43952	137664	137664	137664	137664

Notes: The table reports coefficients from Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring Hero, Villain, and Victim roles on virality, measured by retweet counts. The sample includes relevant tweets featuring at most one character, either framed neutrally or in one of the drama-triangle roles, so that tweets are mutually exclusive (e.g., a tweet coded as Hero cannot also be coded as Neutral or Villain). Column 1 compares Hero tweets to the Neutral baseline; Column 2 compares Villain to Neutral; Column 3 compares Victim to Neutral. Columns 4–7 include all groups but vary the reference category: Neutral (Col. 4), Hero (Col. 5), Villain (Col. 6), and Victim (Col. 7). Coefficients are interpreted as percentage changes using the transformation $e^{\beta} - 1$. All regressions control for author characteristics (verified status, number of followers/followings, total tweets created, party affiliation, religiosity, higher education, and parenthood) and include character fixed effects. We also include hour, week-of-year, and year–state fixed effects. Standard errors are clustered at the week level, covering the full time frame.

5.3 Heterogeneous Effects

So far, our analysis has focused on the determinants of virality, examining both the overall impact of narratives and the effects of *political narratives* with different roles and combinations of character-roles. In this section, we present the final results of our empirical analysis, shifting the attention to the role of multiple character-roles and different types of characters.

We begin by addressing the following question: Does narrative complexity matter for virality? We define a narrative as simple if it contains only one of the three roles in the Drama Triangle – even if multiple characters are present within that role. Conversely, we define a narrative as complex if it features two or all three roles within the same tweet. This classification reflects the underlying structure of the message. When multiple characters appear but share the same role, the tone, emotional framing, and intended opinion remain consistent. However, when multiple roles are present, the narrative becomes more nuanced, incorporating multiple perspectives or conflicting viewpoints, thus increasing its complexity.

Figure 8 provides insights into the effect of narrative complexity on virality. We estimate models where each single-role narrative (e.g., hero) is compared to its three complex alternatives (e.g., hero-villain, hero-victim, and hero-villain-victim), as well as to a reinforced simple narrative, where multiple characters appear within the same role (e.g., multiple heroes). Panel A of the figure presents results for hero narratives, Panel B for villain narratives, and Panel C for victim narratives.

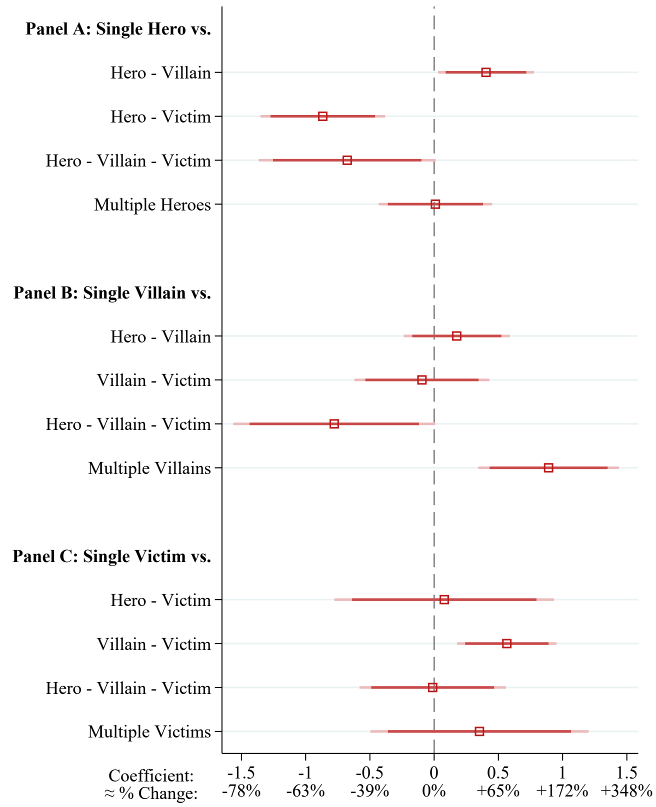
Compared to tweets featuring a single-character hero narrative, increasing complexity by introducing a hero-victim or hero-victim-villain structure reduces overall virality. However, when the hero role is combined exclusively with a villain, contagiousness increases, while popularity remains unchanged. A similar pattern emerges for villain narratives, where complexity tends to reduce virality, particularly when the full Drama Triangle is present. Notably, when villains are reinforced within a simple narrative – meaning multiple characters share the villain role – virality increases even further, suggesting that narratives centered around multiple villain tend to gain more traction. For victim narratives, no consistent pattern emerges, except when a victim is paired with a villain, which significantly boosts contagiousness.

Overall, these findings suggest a clear and consistent pattern: narrative complexity tends to reduce virality. When multiple character-roles interact within the same tweet, the added nuance appears to be detrimental to both retweets and likes. In contrast, when a message is reinforced by multiple characters of the same role, virality increases. This effect is particularly pronounced for villain narratives, which emerge as a strong driver of engagement. The contrast between villains and either heroes or victims appears to amplify a narrative's impact, making it more compelling. Most strikingly, tweets featuring multiple villains generate the highest levels of virality, suggesting that oppositional framing and conflict-driven narratives are particularly effective in capturing attention. Do people engage most with what they oppose?

In the final part of this section we ask whether certain kinds of narratives spread more easily than others. We exploit two dimensions of our narrative framework and classification pipeline. First, we test whether the type of characters matters for virality, comparing narratives built around human actors (e.g., US People or Corporations) with those centered on instrument characters such as regulations or technologies. Second, we test whether the number of character-roles in a tweet makes a difference, does adding more characters framed as heroes, villains, or victims make a difference for virality? Figure 9a and Figure 9b show the results of our heterogeneous-effects analysis.

Figure 9a reports Poisson Pseudo-Maximum Likelihood estimates, where the dependent variable is the number of retweets each tweet receives. Columns (1) and (2) compare tweets that feature only human characters with tweets that feature only instrument characters. To make the comparison clean, we restrict the sample to tweets that include at least one character-role and place each tweet in one of two mutually exclusive groups: all-human or all-instrument. Column (1) narrows the analysis to tweets with a single character-role, while Column (2) allows multiple roles. Columns (3) and (4) focus on the number of roles. Column (3) includes the raw count of character-roles as a regressor; Column (4) adds the squared term to test for non-linear effects. Figure 9b complements these regressions by plotting the marginal change in expected retweets associated with each additional

Figure 8: ***Regression Results - Heterogeneity in the Impact of Character-Role Combinations on Virality***



Notes: The figure shows coefficients from Poisson Pseudo-Maximum Likelihood regression models testing the effect of character-role combinations on virality, measured as the tweet-level count of retweets. The x-axis reports coefficient estimates together with the corresponding approximate percentage change, computed as $e^{\beta}-1$. Each panel corresponds to a separate regression model. Panel A examines the effects of hero combinations, with the comparison group being tweets featuring a single hero character only. The regressors capture cases where a hero is paired with one or more villains, with one or more victims, with both villains and victims, or with additional heroes. Panel B follows the same structure for villains, with the reference group being tweets that feature a single villain character, while Panel C does so for victims, using tweets with a single victim character as the baseline. All regressions include hour, week of the year, and year–state fixed effects. Standard errors are clustered at the week level, covering all weeks in the study period.

character-role, providing a more intuitive picture of how narrative complexity influences virality.

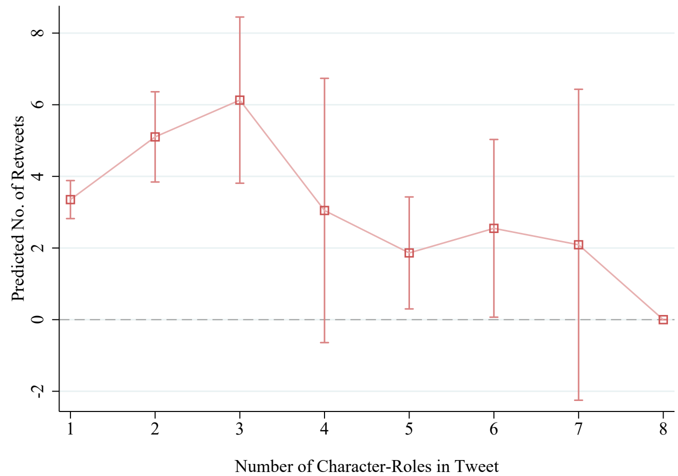
The results paint a consistent picture. Tweets built around human characters spread much farther than those built around instrument characters such as regulations or technologies. This holds whether a tweet features a single role or several, and the effect sizes are large: shifting from an instrument to a human focus raises expected retweets. In short, solutions-oriented or policy-focused messages attract far less engagement than stories that center on people. Turning to narrative complexity, adding characters boosts virality at first but gives diminishing returns. Moving from one to two character-roles delivers the biggest jump in retweets, and a third role still helps, but the incremental gain is small by the time a fourth role appears. Beyond that point the curve flattens, suggesting that audiences lose interest, or find the story harder to process, once the cast grows too

Figure 9: ***Regression Results – Heterogeneous Impact of Narratives on Virality***

(a) Heterogeneous Effects

Dependent Variable	**Retweets Count**			
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Human vs. Instrument	0.712	1.937		
(mut. excl.)	(0.341)	(0.861)		
	[0.037]	[0.024]		
Sum of CRs			0.214	0.659
			(0.057)	(0.150)
			[0.000]	[0.000]
Sum of CRs Squared				-0.125
				(0.045)
				[0.006]
Single Character-Role	✓			
Multiple Character-Roles		✓	✓	✓
Mean Outcome Reference Group	12.201	12.201	10.923	10.923
Pseudo R2	0.71	0.80	0.62	0.62
Obs	142611	42951	309531	309531

(b) Marginal Effect of Adding Character-Roles



Notes: The exhibit provides an overview of the heterogeneous effects of narratives on virality. The table reports estimates from Poisson Pseudo-Maximum Likelihood regression models where the dependent variable is virality, measured by the number of retweets a tweet receives. Columns (1)–(2) isolate character type: the sample is limited to tweets that contain at least one character-role and each observation is placed in a mutually exclusive group of either (i) tweets featuring only human characters or (ii) tweets featuring only instrument characters. Column (1) restricts the comparison to tweets with a single character-role; Column (2) allows multiple roles within the tweet. Columns (3)–(4) explore the cumulative effect of featuring additional character-roles in a tweet: Column (3) includes the total count of character-roles as a regressor, while Column (4) adds the squared term to test for non-linear effects. All regressions control for author characteristics (verified status, number of followers/followings, total tweets created, party affiliation, religiosity, higher education, and parenthood) and include character fixed effects. We also include hour, week-of-year, and year–state fixed effects. Standard errors are clustered at the week level, covering the full time frame. Figure 9b complements the regression results by plotting the marginal change in expected retweets associated with each additional character-role, providing an intuitive visualization of how increasing narrative complexity affects virality. Because of technical reasons working with Poisson, we modified the usual specification of regression to obtain these predictions: for this regression we exclude the year-state FEs.

crowded.

6 Experimental Results: Beliefs, Preferences and Memory

While repeated and increased exposure through higher virality alone makes *political narratives* attractive as a communication technology, their effectiveness could be further enhanced by their persuasive power and a potential effect on memory. Using a large-scale observational dataset of tweets, we find that narrative on average go more viral, with villain characters being even more efficient than hero characters. The well-documented mere-exposure effect in psychology shows that repeated exposure to the same message can already greatly enhance its effectiveness. However, is there also already an effect on beliefs, preferences and memory from single exposure to *political narratives*? To test this, we run three parallel survey experiments, each focused on *political narratives* with different character-role combinations that were also frequent and viral on social media.

We distinguish between beliefs, preferences and memory effects. Narratives may be effective by shaping beliefs and thus influencing expectations and future decisions. Regarding preferences, we distinguish preferences about support for specific policies from preferences about characters. Policy preferences are generally more stable, and potentially tied to a partisan identity, and thus harder to be influenced by a single exposure to a specific narrative, while preferences about a character are more directly tied to the portrayal of that character in a drama-triangle role in the political narrative. Changes in beliefs and in preferences about characters could be viewed as pre-requisites for changing concrete policy preferences.

The remainder of the section is organized as follows. First we provide an overview on the design of our experiments in Subsection 6.1, second we show our results on beliefs and preferences in Subsection 6.2, and thirdly, we explore the impact of *political narratives* on memory in Subsection 6.3.

6.1 Experimental Design

We conduct three pre-registered online experiments, each following an identical design. Our approach is in each case to test the effect of being exposed to a political narrative tweet compared to a control tweet that contains the same factual information and characters, but without portraying the characters in a drama-triangle role. In the first experiment, the narrative tweet frames Green Tech and the US People as heroes. In the second, Green Tech again appears as a hero, joined by Regulations as a second hero, and Fossil Industry as a villain. In the third, Fossil Industry and Corporations are framed as villains, with Green Tech as the hero. Given the structure of the narratives, from here on we refer to the first experiment as the Hero-hero (HH) experiment, the second as Hero-hero-villain (HHV) experiment, and the third as Villain-villain-hero (VVH) experiment.

In each experiment, participants are randomly assigned to either a treatment group – exposed to the narrative tweet – or a control group – exposed to the neutral version. We inform participants that the experiment investigates the effects of viewing a typical social media feed. All participants view a feed designed to resemble a Twitter/X timeline, with usernames blurred for anonymity. Each feed contains three posts. The first two – identical across both conditions – serve as obfuscation. The third post presents either the narrative tweet (in the treatment condition) or the neutral tweet (in the control condition). We recruit a representative sample of the US population via *Prolific*, randomly assigning participants to either the treatment or the control condition. Figure F.1 shows balance tests confirming no worrying differences between the treatment and control groups. To test memory effect, each experiment includes a follow-up survey administered one day later, with an overall attrition rate of roughly 22% (20% for the Hero-Hero experiment, 18% for the Hero-Hero-Villain, and 28% for the Villain-Villain-Hero). In this follow-up, all participants in both are shown the same feed as on the day before, but with the third (treatment or control) post blurred[7].

We use a combination of different type of questions, explained in more detail in the respective following sections. Our belief question asks participants to predict the share of renewable energy in the US by 2035. As this is unknown, there is no incentive to correctly guess; however, we also

[7]See the questionnaires of each experiment in Appendix F

Table 4: *Manipulation Check - Effectiveness of the Political Narrative Treatment*

Dependent Variable	**Mention of Character-Roles**							
	Coeff./SE/p-value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.293	0.281	0.254	0.271	0.358	0.322	0.280	0.316
	(0.032)	(0.032)	(0.031)	(0.018)	(0.034)	(0.035)	(0.032)	(0.019)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Outcome | Recall at least 1 Character					✓	✓	✓	✓
Experiment: Hero-Hero	✓			✓	✓			✓
Experiment: Hero-Hero-Villain		✓		✓		✓		✓
Experiment: Villain-Villain-Hero			✓	✓			✓	✓
Experiment FE				✓				✓
Mean Outcome Control Group	0.33	0.34	0.44	0.37	0.45	0.52	0.64	0.54
Observations	987	976	968	2931	742	686	695	2123

Notes: The table reports OLS estimates from manipulation checks of the narrative treatment. The dependent variable is a binary indicator equal to one if the participant recalled the role assigned to a character in the treatment tweet (e.g., Green Tech as hero in the Hero–Hero experiment, Fossil Industry as villain in the Hero–Hero–Villain experiment). Columns 1–4 report effects on the unconditional likelihood of recalling the character-role, while Columns 5–8 restrict the sample to participants who recalled at least one character from the tweet. Only the treatment group was exposed to characters explicitly framed in roles; control participants saw the same characters presented neutrally. Non-zero recall in the control group therefore reflects participants' prior beliefs or participants' interpretation of the control tweet, while the treatment effect captures the additional role attribution induced by framing. All regressions control for income, education, political preference, age, and sex, with standard errors clustered at the individual level. Appendix Table G.3 and Table G.9 report the corresponding models without controls and with randomization inference for p-values.

see no reason why the answers should be systematically biased. In the questions about policy preferences, the specific policy is adjusted to the specific treatment-control tweet combination. Given that changing preferences about the characters contained in narratives is a key channel of narrative persuasiveness that we highlight, we use an incentivized, real-stakes donation question for its measurement. For memory, we use a closed-form item-specific aided recall question with memory cues about numerical facts and an open, free recall question to measure character and role memory.

We begin by verifying that the treatment manipulation worked. To do so, we use the answers to the open, free recall question on the day of the experiment,: *"Please tell us anything you remember about the social media post that served as the basis of the previous questions. Describe your thoughts in the order they come to your mind; this can include full sentences, individual words, or attributes."* If the treatment manipulation was effective, treated participants should be more likely to mention the characters in the roles in which they were cast. Note than non-zero role assignment in the control group may reflect either prior beliefs or participants' subjective interpretation of the neutral character representation in the control tweet.

Table 4 shows results separately for each experiment, Hero-hero, Hero-Hero-Villain, and Villain-Villain-Hero, as well as for a pooled version with experiment fixed effects (FEs). The dependent variable is always 1 if the participant recalls at least one character-role correctly, and 0 otherwise. In columns (1) to (4), we show linear OLS models using this as an outcome without further restrictions. However, this bunches participants who remember the characters together with those that might not even remember the character. To examine specifically role perceptions more clearly, we thus also run specification in columns (5) to (8) that consider only participants who recall at least one character. The results are highly consistent across specifications, and slightly larger when conditioning on character recall. For instance, column 1 shows that compared to 33% in the control group, around 33+29=62% in the treatment group indicate at least one character in the correct role. The treatment manipulation work and those exposed to the political narrative version are approximately 30% more likely to mention the the characters in a role.

6.2 Experimental Results: Beliefs and Preferences

We turn to the effects of the three *political narratives* on beliefs and stated preferences. Figure 10a shows the impact of *political narratives* on beliefs, separately for each treatment-control pair: Hero–Hero (top panel), Hero–Hero–Villain (middle), and Villain–Villain–Hero (bottom). We show the coefficients for the belief question ("what percentage of US energy will come from green technologies in 2035") and the confidence in that estimate (ranging from 0 = not confident to 100 = very confident).

We find that the effects of the *political narratives* on beliefs are mirroring the roles and role combinations in the respective treatment. The Hero–Hero treatment with Green Tech and US People increases beliefs by about 2%, with the effect being statistically significant at the 10%-level. However, combining Green Tech and Regulations as heroes with Fossil Industry in a villain role changes this positive outlook, which shifts close to zero and statistically insignificant. We then further emphasize the villain roles by casting two characters, Fossil Industry and Corporations, against Green Tech as hero. This completely turns around the effect on beliefs, with a negative point estimate of about 5% that is also statistically significant at the 1%-level.

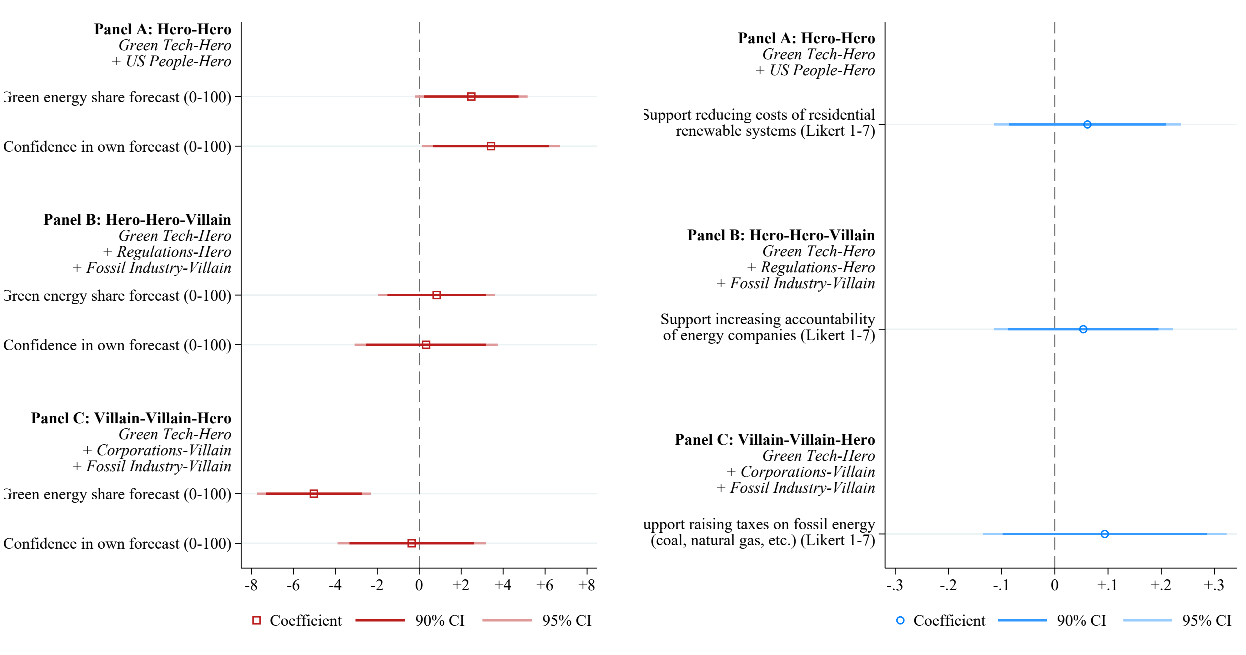
How should we interpret those effects on beliefs? First, *political narratives* clearly can shift beliefs even with single exposure. Second, the fact that different combinations affect beliefs in a plausible way demonstrate the value of our framework to understand narratives and their effect. Thirdly, the effect of using villains, especially a villain combination, is by far the strongest, in line with our results on virality. Fourth, the only combination that increased confidence in the participants estimates is the fully aligned, non-antagonistic hero-hero role combination, while confidence for the antagonistic combinations remains unaffected. Overall, we conclude that single exposure to different narrative combinations does affect beliefs linked to a character-role conditional on the combination of that character with other character roles.

Moving to the right part of the figure, Figure 10b shows effects on stated policy preferences, using the same vertical structure as for beliefs. To do so, we measure support for specific climate change policy proposals that fit the respective character combinations in the treatment-control pairs. For the Hero–Hero combination, the question is about support for higher government subsidies for residential renewable energy systems. For the Hero–Hero–Villain combination, the policy proposal is about increasing the accountability of energy companies for potential damages they cause. For the Villain–Villain–Hero experiment, the question refers to support for raising taxes on fossil energy

Figure 10: *Experimental Results - Impact of Political Narratives on Beliefs and Policy Preferences*

(a) Impact on Beliefs: Forecast for the Future

(b) Impact on Stated Policy Preferences: Policy Support



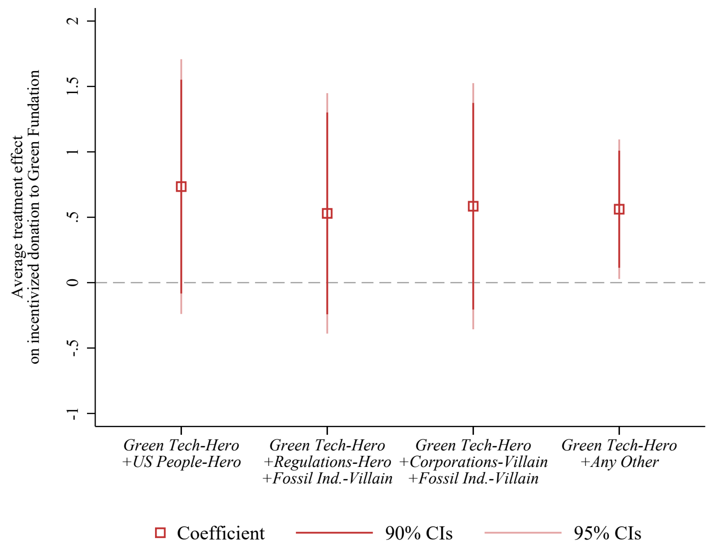
Notes: The figures show the coefficients from OLS regression models analyzing the impact of the narrative treatment on beliefs and stated preferences. In both figures, Panel A shows results for experiment Hero-Hero, Panel B for Hero-Hero-Villain and Panel C for Villain-Villain-Hero. In Figure 10a, we show the impact of the narrative treatment on two outcomes: the participants' forecast, as an answer to the question *'What percentage of US energy do you predict will come from renewable sources and green technology by the year 2035? Indicate a number between 0 and 100.'*, and their confidence in the forecast, as answer to *'Your response on the previous screen suggests that by 2035, [x]% of US energy will come from renewable sources and green technology. How certain are you that the actual share of renewable energy in 2035 will be between [x-5] and [x+5]%?'*. Appendix Table F.1, Table G.4, and Table G.10 show the correspondent regression models, with controls, without controls, and using randomization inference for p-values, respectively. In Figure 10b, we show the impact of the narrative treatment on the support or opposition for a policy or law that is in line with the content of the narrative. Policies questions are indicated in the graph and answers were collected as a 7-points Likert scale from 'Strongly Oppose' to 'Strongly Support'. All models include income, education, political preference, age, and sex as controls. We use robust standard errors. Appendix Table F.2, Table G.5, and Table G.11 show the correspondent regression models, with controls, without controls, and using randomization inference for p-values, respectively.

sources. We find no significant changes in policy support for any of the treatment-control combinations. This finding is consistent with existing research suggesting that policy preferences are less likely to be influenced by experimental interventions than beliefs (Berkebile-Weinberg et al. 2024).

Finally, we examine a crucial potential lever of *political narratives*. Characters, from individual politicians to institutions, policies and technologies, are at the core of our approach of defining and measuring narratives. Hence, to understand the functioning of *political narratives* as a communication technology, and their persuasive power, it is essential to analyze whether (single) exposure can change preferences about the characters featured in the narrative? To test this, we drafted our treatment-control pairs so that each contains Green Tech as a character, always in the hero role in the treatment condition. This enables us to consistently measure revealed preferences using a real-stakes, incentivized donation to a non-partisan organization supporting green technology as an outcome. Specifically, participants play a version of a dictator game in which they are asked to allocate $25 between themselves and the green technology organization. Donation outcomes like this tend to be statistically rather noisy relative to the expected effect size as many other factors influences the individual decision, hence it is important that our design allows us to also pool all three experiments to increase statistical power.

Figure 11 displays the coefficients together with 90% and 95% confidence intervals from each individual experiment, along with the pooled estimate for all experiments. The results consistently point in the same direction, indicating that the *political narratives* did shift the preferences about the character. Statistical power is limited in each individual experiment, because of a lot of idiosyncratic noise in the donation outcome. However, exposure to the political narrative, each with Green Tech as the hero, does always increase participants' willingness to donate to the institution promoting green technology. The pooled effect is positive and statistically significant at the 5%-level. On average, exposure to the Green Tech character in a hero role increases donations by around 0.56 dollars, with the overall average donation in the treatment condition being 7.02 dollars compared to 6.55 in the control condition. Accordingly, even a single exposure to a political narrative casting a characters in a specific drama-triangle role can be sufficient to change a real-stakes decision related to that character.

Figure 11: ***Experimental Results - Impact of Political Narratives on Revealed Preferences about Character GREEN TECH***



Notes: The figure displays the coefficients, 90% (dark red), and 95% confidence intervals (light red) from OLS models analyzing the impact of the *political narratives* on participants' revealed preferences. We measure revealed preferences with an incentivized decision to donate 25$ to themselves or to a foundation promoting green technology diffusion. Moving from the left of the graph, coefficients 1, 2, and 3 show results for each experiment (in order Hero-Hero, Hero-Hero-Villain, and Villain-Villain-Hero), while the last coefficient on the right shows results for the pooled sample with experiment fixed effects. All models include income, education, political preference, age, and sex as controls. We use robust standard errors. Appendix Table F.3, Table G.6, and Table G.12 show the full corresponding regression models, with controls, without controls, and also when using randomization inference for p-values.

6.3 Experimental Results: Memory

While our previous results indicate that a single exposure to a political narrative can shift beliefs and revealed preferences about a character, *political narratives* may also have an edge over other options to convey information with regard to information processing, storing and retrieval. Graeber, Roth, and Zimmermann (2024) show that linking quantitative with qualitative information can improve recall, especially if the memory retrieval question provides cues to the qualitative information. We investigate a related, but different question. We test (i.) whether the recall of numerical facts is improved when the facts are embedded in a political narrative, as well as (ii.) whether the recall of the characters is improved when they are framed in one of the drama-triangle roles.

A single exposure to a political narrative may not be enough to shift policy preferences on its own. However, narratives can influence preferences incrementally, as repeated exposure combines with belief and attitude shifts triggered by even one exposure. To assess the potential for such cumulative effects, it is crucial to examine what kind of information from a single exposure is stored in memory and how well it can be recalled. Do *political narratives* mainly serve as a vehicle to communicate and anchor specific factual information, such as numerical data? Or is their primary strength in shaping perceptions of characters and their roles, thereby shifting preferences and beliefs more indirectly?

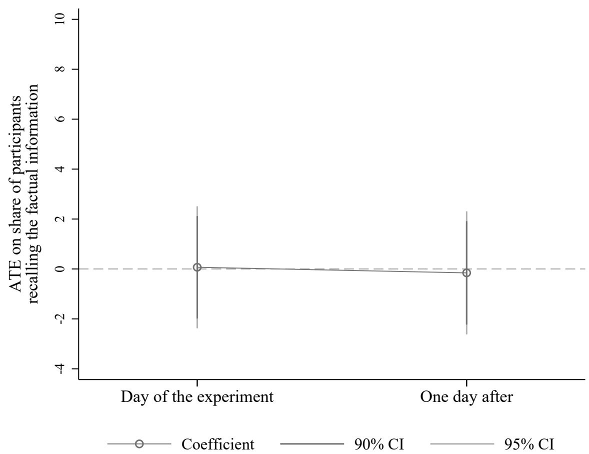
To investigate this, we included a memory recall task in the three experiments discussed above. We ask each respondent two information recall questions: first, a direct question with memory cues about the numerical fact contained in the texts, and second, a free-recall question about anything else they remember about the text they were exposed to. Both questions were posed both on the day of the experiment and again in a follow-up survey conducted a day later[8].

Figure 12a and Figure 12b show the results of the exposure to *political narratives* on memory, pooling all three treatment-control combinations. The left panel shows the coefficient measuring the effect of being exposed to the narrative treatment on remembering the factual information contained in the tweet, the right panel the effect on remembering characters contained in the tweets. The panels show that *political narratives* positively affect memory recall of at least one character, with the effect also being clearly statistically significant. There is only little memory decay between the two days (from around 6% to around 5%), and the effect remains significant. In contrast, *political narratives* have no effect on factual memory, with the effect on recall of numerical information being very close to zero and clearly insignificant. Those null effects are robust to using different target ranges instead of exact numerical recall as the outcome (see Appendix Table G.1 and Table G.2). Hence, it seems the main strength of *political narratives* as a communication technology with regard to memory is that the narrative characters, in their respective roles, are much better remembered by recipients.

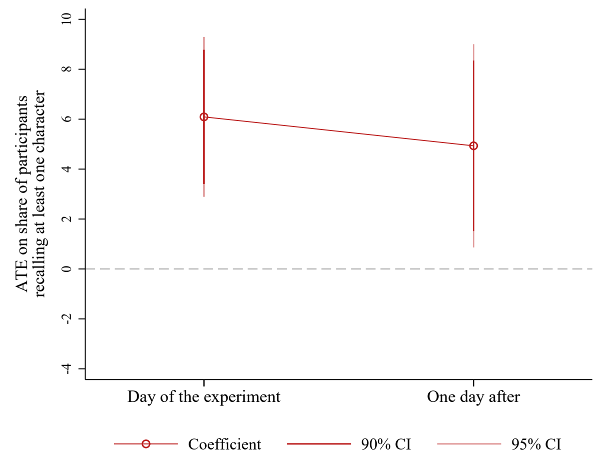
[8]The question posed on the day of the experiment serves as basis for the results of our manipulation check as well, in Table 4.

Figure 12: *Experimental Results - The Impact of Political Narratives on Memory (Pooled Sample)*

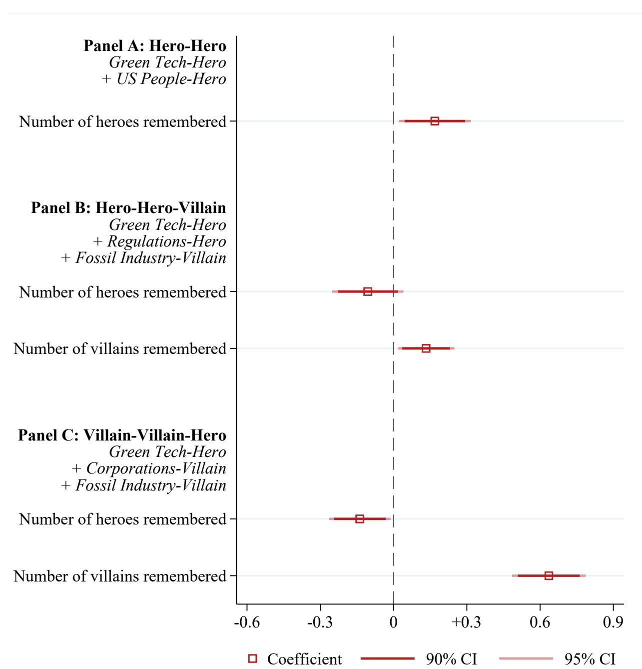
(a) Factual Memory:
Recall of Numerical Information



(b) Character Memory:
Recall of at least one character



(c) Character Memory:
Recall of characters by role



Notes: The figure reports OLS estimates of the treatment effect on participants' memory. Figure 12a shows the pooled impact on factual recall from the tweet on the day of the experiment (left) and one day later (right), expressed as percentage differences between treatment and control. Figure 12b shows the effect on recalling at least one character from the text, again same day (left) and next day (right), using an open-ended recall question coded into a binary outcome. Appendix Table F.4, Table G.7, and Table G.13 show the correspondent regression models, with controls, without controls, and using randomization inference for p-values, respectively. Figure 12c examines recall of specific characters: Panel A (Hero–Hero) tests recall of Green Tech and US People (heroes); Panel B (Hero–Hero–Villain) shows recall of Green Tech and Regulations (heroes, top) and Fossil Industry (villain, bottom); Panel C (Villain–Villain–Hero) shows recall of Green Tech (hero) and Fossil Industry/Corporations (villains). All models control for income, education, political orientation, age, and sex; SEs are clustered at the participant level. The first two figures also include experiment fixed effects. Appendix tables report the corresponding models with controls, without controls, and using randomization inference for p-values. Appendix Table F.5, Table G.8, and Table G.14 show the correspondent regression models, with controls, without controls, and using randomization inference for p-values, respectively.

Our framework allows us to go beyond this general memory effect and distinguish between the effect on characters cast as heroes compared to those cast as villains in the different treatment-control combinations. Figure 12c differentiates for each treatment-control combination between the effect on remembering those characters portrayed as hero from those portrayed as villain in the treatment condition. The first panel shows that solely using two characters in hero roles enhances memory of both characters, as expected. However, the second and third panel reveal an interesting relationship between hero and villain role memory. Once a character cast as villain is inserted, memory of the hero characters is crowded out and actually lower than in the control tweets. In contrast, recall of the villain characters is always enhanced, even more strongly if two characters are cast into the villain role.[9] Hence, casting characters as villains not only enhances virality and persuasiveness, but in addition even improves memory recall of that character at the expense of characters in other roles.

Discussion and Interpretation

This section discusses possible explanations specifically for the memory results, with some explanations also possibly relevant for the prior experimental results. We do not go into more detail here as this would go beyond the scope of the paper, but we are convinced that there is ample room for more research to investigate the channels and mechanisms in more detail.

a.) Information resonance. People are more likely to remember information that resonates with them and with their whole set of memories. Many of these memories are in the form of stories and narratives, and naturally feature characters in the three drama-triangle roles. Hence, the representation of characters in these roles plausibly has an inherent advantage for encoding information and storing it efficiently into memory. We can think of the roles in the treatment as adding meaning to the characters, which improve the encoding and retrieval of the characters in memory. The schema theory in psychology also highlights how using default, existing schemas to simplify complex information can be seen as a mechanism for efficient information storage and easier retrieval. Participants in the treatment condition also remember the roles (as we show in Table 4), in line with the idea that they store the character and its role jointly in their memory.

b.) Cue-similarity. A key insight in Graeber, Roth, and Zimmermann (2024) is that cue similarity between a story and the prompt/question use for memory retrieval plays a key role in further enhancing the better memory of stories. Our differential results between fact and character recall do not seem to be driven by cue similarity. In fact, our question about the fact ("How many billion kWhs were generated using solar energy in 2023 in the US? Please indicate your best guess.") has many more cues to the fact that our open ended question from which we infer characters ("Please tell us anything you remember about the social media post"). Hence, it is noteworthy that we find the improved character recall even with such a free-recall question.

c.) Type of recall. Our fact question is a verbatim numeric recall of a specific number,

[9] The character outcome is coded from open-ended free recall. We cannot exclude that item-specific cued recall of heroes would attenuate the crowding-out pattern. Accordingly, we put more emphasis on the difference between the effect on characters in hero versus villain roles than on the absolute effect size.

which are generally harder to remember for most people. This number in our context is linked to a character (Green Tech), but it is linked to that character in both treatment and control. While our manipulation adds qualitative role content to the text, similar to Graeber, Roth, and Zimmermann (2024), this role assignment is not directly number-focused. Hence, what our results show is that such verbatim recall of a specific number is not further improved by assigning a role to that character. The better memory of characters and roles can actually be regarded as in line with Graeber, Roth, and Zimmermann (2024), who also document better recall of information type and direction, not specific numbers.

d.) Attention. Attention influences what information people process, but also what and how they store it (Loewenstein and Wojtowicz 2025). Attention is drawn towards aspects of a text that are salient. The drama-triangle roles can be stimulating in that sense for a person because they reflect a salient category, e.g. a threat (Vuilleumier 2005). Similarly, when "searching" their memory for information retrieval, prior evidence indicates attention guides people how to search for information. If a character in memory is linked to a larger semantic cluster, like an archetypal role, that could help to retrieve it more easily.

e.) Emotions. Although it is plausible that emotions play an important role also for memory formation, a closer look at our results reveal that the character-role combinations of the *political narratives* capture more than just emotion. We can distinguish valence (how positive relative to negative a text is) from specific types of emotions like joy, fear or anger. Appendix Figure F.2 shows the valence difference of each treatment vs. control tweet comparison, using the same standard dictionaries as before, as well as the coefficient from the experimental results discussed above. Valence is in line with the changes in beliefs, with more negative treatments triggering a decline in beliefs. However, the donation outcomes related to the Green tech character is consistently positive and almost identical in size across the three treatment-control pairs, demonstrating that it is not the valence of the whole tweet that matters, but the role assignment of the relevant character. There is also no direct relationship between the level of anger and the experimental results on belief and preferences. The memory results further demonstrate that it is the specific character-role combinations that drive the effects, not simply valence or emotions.

f.) Strategic implication for political communication. Our experiments show that single-shot narrative exposure reliably imprints who the actors are and how they are cast, but does not improve recall of facts, even when those facts are tied to a key character. This possibly gives some leeway to the sender of such information in terms of spreading imprecise or even false facts. If most recipients of *political narratives* only remember the characters and their roles better, but not the facts, this lowers the costs of spreading false facts. Ex post fact-checking only partly solves this problem, as the original narrative - thanks to its virality - is usually read much more widely than possible corrections. New approaches like community notes are more interesting in that regard, as they tie the fact-checking directly to the political narrative. At least to the extent the narratives remains with the social media platform, are shared and displayed together with it.

7 Conclusion

In conclusion, our study demonstrates that the political narrative framework provides a powerful lens for understanding how narratives drive engagement and influence public opinion. The result is a numerical map of the narrative citizens share, one that economists can merge with behavioral and market data. By observing which characters and roles dominate, we gain predictive leverage over both the direction and the intensity of public debate.

By analyzing US climate change policy discussions on Twitter over a decade, we show what makes narratives go viral. *political narratives*, on average, are about sixty per cent more likely to be retweeted. This result holds controlling for a range of time and region fixed effects, for key authors characteristics like the number of followers, and when using character fixed effects. Negativity or emotions alone cannot explain that virality premium: when we hold eight discrete emotions and continuous valence constant, the effect on virality is only slightly decreased and remains clearly significant. A villain framing lifts retweets by roughly 170 per cent, hero framing by about 55 per cent, while victim framing has little effect. Pairing another role with a villain raises virality, whereas adding other roles has an ambiguous effect, indicating that complexity can impose an attention cost. Human characters generally tend to go more viral than instrument characters like technology or policies.

Across three preregistered surveys we embed single narrative tweets in an otherwise ordinary feed and compare the effect to a tweet with the same characters in neutral roles. A one-time exposure to the political narrative shifts beliefs: respondents adjust their expectations about the character in the direction implied by the narrative. It also nudges revealed preferences: when given an incentive-compatible choice, participants reallocate real money toward or away from the actor highlighted in the story, even though their stated policy support remains effectively unchanged from single exposure. One day later participants reliably remember which characters appeared more often, yet are not more likely to reproduce factual numbers that accompanied the text, showing that *political narratives* are more about characters than facts.

Economists value causal narratives for showing how one action leads to another, capturing an important aspect of narratives as a communication technology. Causal narratives trace sequences of events – taxes raise prices, prices curb emissions, monetary expansion causes inflation. *Political narratives* add an explicit assignment of agency, blame, and moral standing to relevant human (politicians, institutions) or instrument characters (technologies, policies). Corporations become villains, households victims, activists heroes, and the very same chain of events acquires a specific purpose and direction. Because these role labels can shift while the underlying causality stays fixed, *political narratives* add an orthogonal dimension that pure event-based stories cannot capture. They also reach far beyond raw emotion: a sentence may sound negative without naming a culprit or emphasize a hero without using strong emotion. Our evidence shows that the presence of a villain, not the amount of negativity, is what multiplies viral reach and shapes expectation. Distinguishing characters and roles from both causality and sentiment can therefore be essential for understanding how ideas travel and influence decision-making.

Our framework applied with the suggested pipeline outputs structured data with explicit character and role variables, enabling researchers to study narratives in any large text data set and easily combine it with other economic or political data. Because characters and the archetypal drama triangle roles are so fundamental to human story-telling, standard LLMs are really efficient in measuring well-defined character-roles, making this also a very cost-effective way of measuring narratives without manual coding. Possible applications beyond our context are widespread. For instance, macroeconomists could now test whether villain framing of central banks widens inflation-expectation tails; finance scholars can observe how firms move from hero to villain status during scandals and how this affects stock market evaluations; development economists can monitor shifts in donor narratives about recipient governments and the role of aid agencies. Measurement, once the bottleneck, no longer stands in the way of such inquiries.

References

Acemoglu, Daron, Tarek A Hassan, and Ahmed Tahoun (2018). "The Power of the Street: Evidence from Egypt's Arab Spring". In: *The Review of Financial Studies* 31.1, pp. 1–42.

Adena, Maja et al. (Mar. 2021). *Bombs, Broadcasts and Resistance: Allied Intervention and Domestic Opposition to the Nazi Regime During World War II.* SSRN Scholarly Paper. Rochester, NY. (Visited on 09/11/2025).

Akerlof, George A. and Dennis J. Snower (2016). "Bread and Bullets". In: *Journal of Economic Behavior & Organization* 126.Part B, pp. 58–71.

Andre, Peter et al. (2024). "Misperceived Social Norms and Willingness to Act Against Climate Change". In: *Review of Economics and Statistics*, pp. 1–46. (Visited on 05/14/2025).

Andre, Peter et al. (2025). "Narratives about the Macroeconomy". In: *The Review of Economic Studies*. (Visited on 01/28/2025).

Anker, Elisabeth (2005). "Villains, Victims and Heroes: Melodrama, Media, and September 11". In: *Journal of Communication* 55.1, pp. 22–37.

Aridor, Guy et al. (2024). "The Economics of Social Media". In: *Journal of Economic Literature* 62.4, pp. 1422–1474. (Visited on 08/05/2025).

Ash, Elliott and Sergio Galletta (Oct. 2023). "How Cable News Reshaped Local Government". In: *American Economic Journal: Applied Economics* 15.4, pp. 292–320. (Visited on 08/11/2025).

Ash, Elliott, Germain Gauthier, and Philine Widmer (2024). "RELATIO : Text Semantics Capture Political and Economic Narratives". In: *Political Analysis* 32.1, pp. 115–132. (Visited on 05/16/2025).

Ash, Elliott and Mikhail Poyker (2024). "Conservative News Media and Criminal Justice: Evidence from Exposure to Fox News Channel". In: *The Economic Journal* 134.660, pp. 1331–1355. (Visited on 08/05/2025).

Ash, Elliott et al. (2021). "Visual Representation and Stereotypes in News Media". In: *Center for Law & Economics Working Paper Series* 2021.15. Publisher: Center for Law & Economics, ETH Zurich.

Ash, Elliott et al. (2024a). "From viewers to voters: Tracing Fox News' impact on American democracy". In: *Journal of Public Economics* 240, p. 105256. (Visited on 08/05/2025).

Ash, Elliott et al. (2024b). "The Effect of Fox News on Health Behavior during COVID-19". In: *Political Analysis* 32.2, pp. 275–284. (Visited on 05/16/2025).

Barrera, Oscar et al. (2020). "Facts, Alternative Facts, and Fact Checking in Times of Post-Truth Politics". In: *Journal of Public Economics* 182, p. 104123.

Barron, Kai and Tilman Fries (2023). "Narrative Persuasion". In: *CESifo Working Paper No. 10206*.

Baylis, Patrick (2020). "Temperature and Temperament: Evidence from Twitter". In: *Journal of Public Economics* 184, p. 104161.

Beach, Brian and W Walker Hanlon (2023). "Historical Newspaper Data: A Researcher's Guide and Toolkit". In: *Explorations of Economic History* 90, p. 101541.

Berger, Jonah (July 2011). “Arousal Increases Social Transmission of Information”. In: *Psychological Science* 22.7, pp. 891–893. (Visited on 01/29/2025).

— (2016). *Contagious: Why Things Catch On*. Simon and Schuster.

Berger, Jonah and Katherine L Milkman (2012). “What Makes Online Content Viral?” In: *Journal of Marketing Research* 49.2, pp. 192–205.

Bergstrand, Kelly and James M Jasper (2018). “Villains, Victims, and Heroes in Character Theory and Affect Control Theory”. In: *Social Psychology Quarterly* 81.3, pp. 228–247.

Berkebile-Weinberg, Michael et al. (May 2024). “The differential impact of climate interventions along the political divide in 60 countries”. In: *Nature Communications* 15.1, p. 3885. (Visited on 08/27/2025).

Braghieri, Luca et al. (2024). *Article-Level Slant and Polarization of News Consumption on Social Media*. (Visited on 08/05/2025).

Bursztyn, Leonardo et al. (July 2023). “Opinions as Facts”. In: *The Review of Economic Studies* 90.4, pp. 1832–1864. (Visited on 05/12/2025).

Bénabou, Roland, Armin Falk, and Jean Tirole (2020). “Narratives, Imperatives, and Moral Reasoning”. In: *NBER Working Paper No. 24798*.

Caesmann, Marcel et al. (2021). “Going Viral: Propaganda, Persuasion and Polarization in 1932 Hamburg”. In: *Discussion Paper DP16356*.

Cage, Julia et al. (2022). *Hosting Media Bias: Evidence from the Universe of French Broadcasts, 2002-2020*. SSRN Scholarly Paper. Rochester, NY. (Visited on 12/12/2024).

Cagé, Julia (May 2020). “Media competition, information provision and political participation: Evidence from French local newspapers and elections, 1944–2014”. In: *Journal of Public Economics* 185, p. 104077. (Visited on 09/16/2025).

Cagé, Julia, Nicolas Hervé, and Marie-Luce Viaud (2020). “The Production of Information in an Online World”. In: *The Review of Economic Studies* 87.5, pp. 2126–2164.

Cagé, Julia et al. (July 2023). “Heroes and Villains: The Effects of Heroism on Autocratic Values and Nazi Collaboration in France”. In: *American Economic Review* 113.7, pp. 1888–1932. (Visited on 05/16/2025).

Chen, Jiafeng and Jonathan Roth (2024). “Logs with Zeros? Some Problems and Solutions”. In: *The Quarterly Journal of Economics* 139.2, pp. 891–936. (Visited on 05/07/2025).

Chu, Zi et al. (2012). “Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg?” In: *IEEE Transactions on dependable and secure computing* 9.6, pp. 811–824.

Dechezleprêtre, Antoine et al. (Apr. 2025). “Fighting Climate Change: International Attitudes toward Climate Policies”. In: *American Economic Review* 115.4, pp. 1258–1300. (Visited on 05/16/2025).

Djourelova, Milena and Ruben Durante (2022). “Media Attention and Strategic Timing in Politics: Evidence from U.S. Presidential Executive Orders”. In: *American Journal of Political Science* 66.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12683, pp. 813–834. (Visited on 09/11/2025).

Djourelova, Milena, Ruben Durante, and Gregory J Martin (2025). “The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist”. In: *Review of Economic Studies* 92.3, pp. 1738–1772. (Visited on 09/11/2025).

Djourelova, Milena et al. (2024). “Experience, Narratives, and Climate Change Beliefs”. In: *SSRN Electronic Journal*. (Visited on 03/27/2025).

Durante, Ruben and Brian Knight (2012). “Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi’s Italy”. In: *Journal of the European Economic Association* 10.3, pp. 451–481. (Visited on 09/11/2025).

Durante, Ruben, Paolo Pinotti, and Andrea Tesei (2019). “The Political Legacy of Entertainment TV”. In: *American Economic Review* 109.7, pp. 2497–2530. (Visited on 09/11/2025).

Durante, Ruben and Ekaterina Zhuravskaya (June 2018). “Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict”. In: *Journal of Political Economy* 126.3. Publisher: The University of Chicago Press, pp. 1085–1133. (Visited on 09/11/2025).

Eliaz, Kfir and Ran Spiegler (Dec. 2020). “A Model of Competing Narratives”. In: *American Economic Review* 110.12, pp. 3786–3816. (Visited on 02/25/2025).

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova (2020). “Social Media and Protest Participation: Evidence from Russia”. In: *Econometrica* 88.4, pp. 1479–1514.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya (2011). “Media and Political Persuasion: Evidence from Russia”. In: *American Economic Review* 101.7, pp. 3253–85.

Esposito, Elena et al. (June 2023). “Reconciliation Narratives: *The Birth of a Nation* after the US Civil War”. In: *American Economic Review* 113.6, pp. 1461–1504. (Visited on 05/12/2025).

Fog, Klaus et al. (2010). “The Four Elements of Storytelling”. In: *Storytelling*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 31–46. (Visited on 08/11/2025).

Gehring, Kai, Joop Age Adema, and Panu Poutvaara (2022). “Immigrant Narratives”. In: *CESifo Working Paper No. 10026*. CESifo Working Paper No. 10026.

Gentzkow, Matthew and Jesse M Shapiro (2010). “What Drives Media Slant? Evidence from US Daily Newspapers”. In: *Econometrica* 78.1, pp. 35–71.

Gomez-Zara, Diego, Miriam Boon, and Larry Birnbaum (2018). “Who is the Hero, the Villain, and the Victim? Detection of Roles in News Articles Using Natural Language Techniques”. In: *23rd International Conference on Intelligent User Interfaces*, pp. 311–315.

Graeber, Thomas, Christopher Roth, and Florian Zimmermann (2024). “Stories, Statistics, and Memory”. In: *The Quarterly Journal of Economics* 139.4, pp. 2181–2225. (Visited on 01/28/2025).

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart (Mar. 2023). “Designing Information Provision Experiments”. In: *Journal of Economic Literature* 61.1, pp. 3–40. (Visited on 03/21/2025).

Halberstam, Yosh and Brian Knight (2016). “Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter”. In: *Journal of Public Economics* 143, pp. 73–88.

Harari, Yuval Noah (2014). *Sapiens: A Brief History of Humankind*. Random House.

Jiangli, Su (2020). "European & American Think Tanks and the Reality of US-China Trade War: An NPF Application". In: *Journal of Social and Political Sciences* 3.2.

Jones, Michael D (2014). "Cultural Characters and Climate Change: How Heroes Shape our Perception of Climate Science". In: *Social Science Quarterly* 95.1. Publisher: Wiley Online Library, pp. 1–39.

Jones, Michael D. and Mark K. McBeth (2010). "A Narrative Policy Framework: Clear Enough to Be Wrong?" In: *Policy Studies Journal* 38.2, pp. 329–353.

Karpman, Stephen (1968). "Fairy Tales and Script Drama Analysis". In: *Transactional Analysis Bulletin* 7.26, pp. 39–43.

Kendall, Chad and Constantin Charles (Aug. 2022). "Causal Narratives". In: *NBER Workin Paper 30346*. (Visited on 05/12/2025).

Kirilenko, Andrei P and Svetlana O Stepchenkova (2014). "Public Microblogging on Climate Change: One Year of Twitter Worldwide". In: *Global Environmental Change* 26, pp. 171–182.

Lagakos, David, Stelios Michalopoulos, and Hans-Joachim Voth (2025). "American Life Histories". In: *NBER Working Paper 33373*.

Landis, J. Richard and Gary G. Koch (1977). "The Measurement of Observer Agreement for Categorical Data". In: *Biometrics* 33.1, p. 159. (Visited on 04/15/2025).

Loewenstein, George and Zachary Wojtowicz (Sept. 2025). "The Economics of Attention". In: *Journal of Economic Literature* 63.3, pp. 1038–1089. (Visited on 09/03/2025).

Macaulay, Alistair and Wenting Song (2022). *Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media*. Economics Series Working Papers 973. MPRA Paper No. 113620.

Merry, Melissa K (2016). "Constructing Policy Narratives in 140 Characters or Less: The Case of Gun Policy Organizations". In: *Policy Studies Journal* 44.4, pp. 373–395.

Michalopoulos, Stelios and Christopher Rauh (2024). "Movies". In: *NBER Working Paper 32220*. (Visited on 08/11/2025).

Michalopoulos, Stelios and Melanie Meng Xue (2021). "Folklore". In: *The Quarterly Journal of Economics* 136.4, pp. 1993–2046.

Mohammad, Saif M. and Peter D Turney (Nov. 2013). *NRC Emotion Lexicon*. Tech. rep. Artwork Size: 234 p. National Research Council of Canada, 234 p. (Visited on 02/28/2025).

Montoya, R. Matthew et al. (2017). "A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking". In: *Psychological Bulletin* 143.5, pp. 459–498.

Müller, Karsten and Carlo Schwarz (2021). "Fanning the Flames of Hate: Social Media and Hate Crime". In: *Journal of the European Economic Association* 19.4. Publisher: Oxford Academic, pp. 2131–2167.

— (2023). "From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment". In: *American Economic Journal: Applied Economics* 3, pp. 270–312.

Oehl, Bianca, Lena Maria Schaffer, and Thomas Bernauer (2017). "How to Measure Public Demand for Policies When There is no Appropriate Survey Data?" In: *Journal of Public Policy* 37.2, pp. 173–204.

O'Brien, Erin (2018). *Challenging the Human Trafficking Narrative: Victims, Villains and Heroes*. Routledge.

Polletta, Francesca et al. (2011). "The Sociology of Storytelling". In: *Annual Review of Sociology* 37.1, pp. 109–130.

Qian, Nancy and David Yanagizawa-Drott (2017). "Government Distortion in Independently Owned Media: Evidence from U.S. News Coverage of Human Rights". In: *Journal of the European Economic Association* 15.2, pp. 463–499. (Visited on 09/11/2025).

Serra-Garcia, Marta (2025). "The Attention–Information Tradeoff". In.

Shiller, Robert J. (Apr. 2017). "Narrative Economics". In: *American Economic Review* 107.4, pp. 967–1004. (Visited on 05/16/2025).

— (2020). *Narrative Economics: How stories go viral and drive major economic events*. Princeton University Press.

Tabassum, Fatima et al. (2023). "How Many Features Do We Need to Identify Bots on Twitter?" In: *Information for a Better World: Normality, Virtuality, Physicality, Inclusivity*. Ed. by Isaac Sserwanga et al. Cham: Springer Nature Switzerland, pp. 312–327.

Terry, Larry D (1997). "Public Administration and the Theater Metaphor: The Public Administrator as Villain, Hero, and Innocent Victim". In: *Public Administration Review*, pp. 53–61.

Voth, Hans-Joachim and David Yanagizawa-Drott (2025). "Image(s)". In: *CEPR Discussion Paper* DP19219.

Vuilleumier, Patrik (Dec. 2005). "How brains beware: neural mechanisms of emotional attention". In: *Trends in Cognitive Sciences* 9.12, pp. 585–594. (Visited on 09/03/2025).

Yanagizawa-Drott, David (2014). "Propaganda and Conflict: Evidence from the Rwandan Genocide". In: *The Quarterly Journal of Economics* 129.4, pp. 1947–1994. (Visited on 09/11/2025).

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov (2020). "Political Effects of the Internet and Social Media". In: *Annual Review of Economics* 12, pp. 415–438.

E Robustness Checks: Observational Data

E.1 Impact of Major Twitter Algorithm Changes

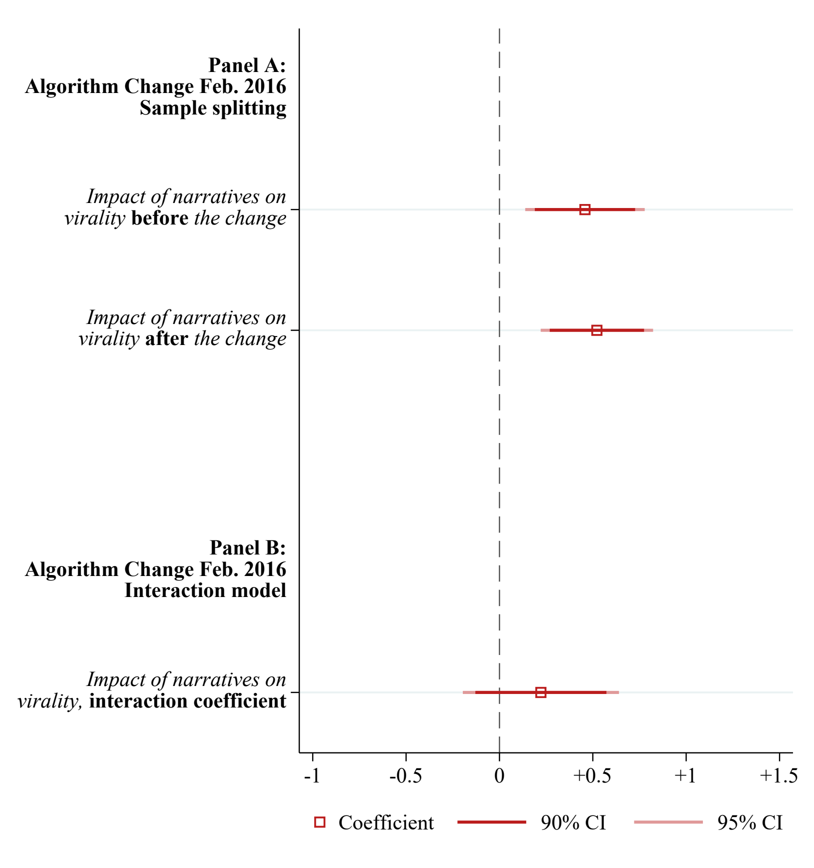
In the main sections of the paper, we show that featuring characters in *political narratives* generates a virality premium compared to presenting the same characters in a neutral way. This effect remains strong and consistent even after introducing strict controls for time, user profiles, and fixed effects. Our analysis so far has concentrated on the difference between narrative and neutral content, but we have not assessed how much of the overall variation in virality can be attributed to content itself. Since we study a social media platform, another crucial mechanism shaping virality is the algorithm that structures content exposure by curating each user's timeline. In this section, we investigate the role of Twitter's algorithm, exploiting a major change that occurred during our study period.

A social media platform's timeline algorithm can be understood as a set of rules and mechanisms to interpret signals, that determine what content a user sees when using the platform. The term timeline reflects the original mechanism used to regulate this flow in most of the social media platforms: posts were ranked chronologically, with users shown the most recent content from the accounts they followed. Over time, however, this purely time-based ranking was phased out and replaced with what is commonly called the algorithmic timeline. The idea is straightforward: an AI-powered algorithm curates each user's feed, tailoring it to individual tastes and past behavior, and prioritizing the content the system predicts the user is most likely to engage with.

Twitter adopted this shift in February 2016, when it introduced its own timeline algorithm. As BuzzFeed reported at the time, the algorithm was designed as "a way for Twitter to elevate popular content, and could solve some of Twitter's signal-to-noise problems. [...] The timeline will reorder tweets based on what Twitter's algorithm thinks people most want to see, a departure from the current feed's reverse chronological order" (BuzzFeed). In many ways, this marked a before-and-after moment for the platform, fundamentally reshaping how information and content were consumed. Some described it as "arguably the most fundamental change it has ever made: a major tweak to the timeline" (WIRED). Others were more skeptical. As VICE put it, "2016 was the year of politicians telling us what we should believe, but it was also the year of machines telling us what we should want" (VICE).

In the context of our study, the transition from a chronological to an algorithmic timeline could be a decisive factor shaping how narratives spread online. We do not have a clear prior on the potential effects of this change. Algorithms are designed to maximize engagement by serving users content they are most likely to interact with, based on past behavior and users' characteristics. This mechanism could amplify the virality of narratives if their emotional and dramatic structure makes them especially engaging. At the same time, algorithmic curation may work in the opposite direction: by reducing the disproportionate visibility of a few highly prolific accounts, it could dilute the dominance of narrative-heavy users and favor a more balanced and maybe neutral flow of information. Finally, it is possible that the algorithm change had little effect on our outcome of interest. If narratives are intrinsically more viral than neutral content, their relative advantage might persist regardless of how the platform ranks posts. Since the internal workings of the algorithm are not clear to us, the net effect is ultimately an empirical question, which we explore in Figure E.1.

Figure E.1: ***Impact of the Main Algorithm Change in Twitter History***



Notes: The figure shows coefficients from Poisson Pseudo-Maximum Likelihood regressions testing the effect of featuring at least one character-role, compared to featuring only neutral characters, on virality (measured as the number of retweets). The analysis focuses on Twitter's major algorithmic change in February 2016, when the platform moved from a chronological timeline to an algorithm-based feed, where an algorithm ranked and prioritized content for each user. The two panels correspond to different specifications: the top panel shows the impact of *political narratives* on virality before (top coefficient) the algorithm change and the same after the change (second coefficient); and the bottom panel reports the interaction effect between narratives and a post-change indicator variable. The x-axis reports coefficient estimates and the corresponding approximate percentage change, computed as $e^{\beta} - 1$ and rounded to the nearest unit. All regressions control for author characteristics (verified status, followers, followings, total tweets created, party affiliation, religiosity, higher education, parenthood status) and include character, hour, week-of-year, and year–state fixed effects. Standard errors are clustered at the week level.

In Figure E.1, we examine the impact of Twitter's algorithmic change in two ways. Panel A compares the relationship between narratives and virality before and after the change, using our standard specification: a Poisson Pseudo-Maximum Likelihood regression with user controls, character fixed effects, and a full set of time fixed effects. Panel B shows results from a specification that includes an interaction term between the narrative indicator and a post-change dummy (equal to 1 after the algorithm switch, 0 before). While we are aware that interaction terms in Poisson models require cautious interpretation, our goal here is simply to test whether the interaction effect is significantly different from zero.

The figure provides a clear message. Splitting the sample shows consistent evidence of the impact of narratives on virality: both before and after the algorithm change, the effect is positive, statistically significant, and comparable in size to the estimates over the full time period presented in the main paper. The interaction term, while positive, is not statistically significant. Taken together, the results suggest that the algorithm change had little effect on the relationship between narratives and virality. Although this exercise has clear limitations and should be interpreted with caution, it is nonetheless striking that the narrative premium remains virtually unchanged, point to a dynamic that may go beyond the algorithmic structure of the platform.

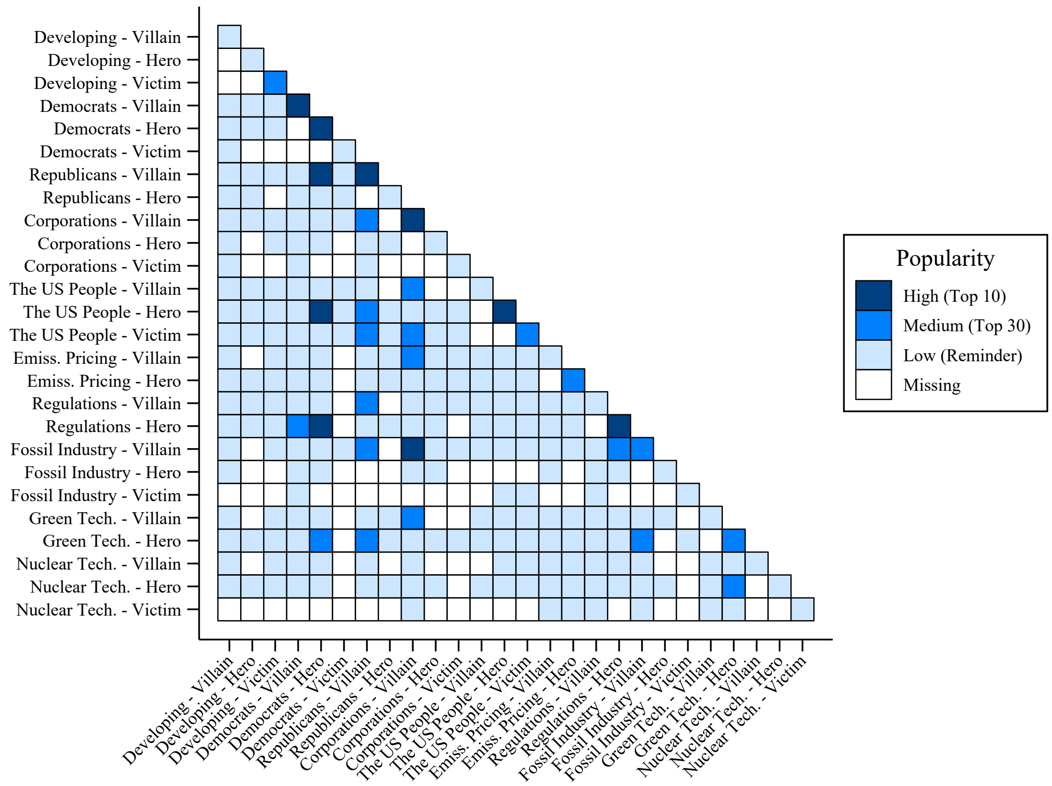
E.2 Impact of Narratives on Popularity

In this section of Appendix E, we check the robustness of our results by looking at how narratives affect tweet popularity, measured by the number of likes. We leave these results out of the main paper for two reasons. First, the findings for popularity are very similar to those for virality, which we measure using retweet counts. Second, we believe retweets are a better measure of virality because they capture not just approval and endorsement – as likes might – but also the actual spread of content.

We begin our analysis on popularity by replicating the same descriptive exercise from Subsection 4.3, now using the rate of likes instead of retweets. Figure E.2 shows a heatmap of like rates for all narratives that include one or two character-roles. Each square in the matrix reports the rate of likes for a specific pair of character-roles, or for a single character-role when it appears alone (along the diagonal). The rate is computed by summing all likes received by tweets featuring that specific combination (or single character-role), and dividing by the total number of likes across all tweets with one or two character-roles. The matrix is symmetrical. To keep the figure readable, we do not display the numerical values directly; instead, we highlight the top 10 and top 30 most frequent combinations using a color scale.

Figure E.2 shows that popularity rates follow patterns similar to those observed for virality (retweets), though with some notable differences. While the most polarized and politicized narratives – those on Democrats and Republicans – still drive engagement, US People-Hero narratives stand out even more strongly in this dimension, ranking among the top five both independently and when combined with US Democrats-Hero narratives. For both virality and popularity, human characters tend to be more viral, yet Regulations and Green Tech also feature prominently, appearing in several of the top 30 most popular narratives.

Figure E.2: ***Popularity (Likes) of Character-Roles - Tweets with One or Two Character-Roles***



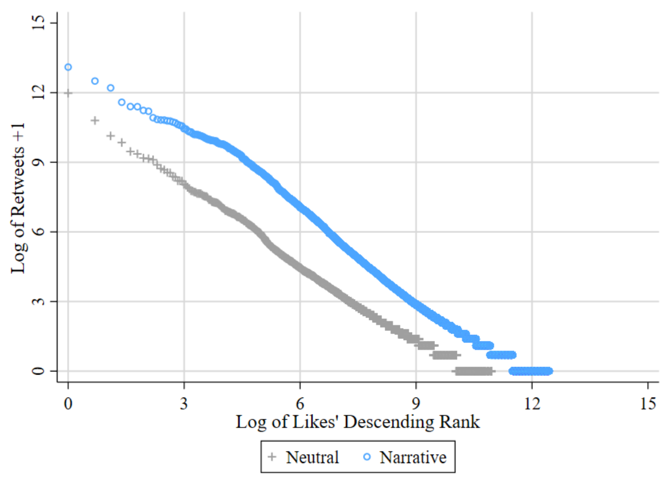
Notes: The figure shows the like rate of each character–role appearing alone or in combination with another role, among relevant tweets containing one or two roles. Like rates are computed as the share of total likes received by a given role (or pair) relative to all likes of tweets with one or two roles. The diagonal of the matrix shows the likes rate when each character-role appears alone. Tweets with three or more character-roles are excluded. To avoid visual overload, we do not display exact rates. Instead, we use a color scheme to highlight the top 10 most frequently retweeted character-role combinations, the top 30 (which includes the top 10), and the remaining pairs. White indicates a pair that never appears together. The top 10 in order is: US Democrats-Hero (13.90%), US Republicans-Villain (13.90%), US People-Hero (7.00%), US Republicans + US Democrats-Hero (6.13%), US People-Hero + US Democrats-Hero (5.32%), Fossil Industry-Villain + Corporations-Villain (5.20%), US Democrats-Villain (3.85%), Regulations-Hero + US Democrats-Hero (3.21%), Regulations-Hero (2.29%), Corporations-Villain (2.18%). Paper Figure 5 shows the same for virality (retweets) of tweets.

Building on the analysis in the main paper, we take a step further and examine how narratives influence the distribution of likes. Specifically, we compare tweets that feature a Political Narrative with those that do not. Figure E.3 shows the log-log rank distribution of likes, separating tweets that include at least one character-role from those that only present characters in a neutral form. The x-axis reports the logarithm of the tweet's rank, where rank 1 corresponds to the most liked or retweeted tweet in the dataset. The y-axis reports the logarithm of the number of retweets or likes, plus one.

The figure shows that the distribution of popularity also follows a power law, similarly to the case of retweets. Moreover, at every point along the rank distribution, tweets containing *Political*

Narratives tend to receive more likes – mirroring the pattern observed for retweets. Compared to the retweet distribution presented in the main paper, the curve for likes appears even steeper. This suggests that the most liked tweets receive more likes than the most retweeted tweets receive retweets, and that the drop-off from the most to the least liked tweets is even sharper. Overall, this descriptive evidence indicates that popularity follows similar patterns to virality, and that tweets featuring narratives consistently attract more likes than neutral ones.

Figure E.3: ***Popularity of Political Narratives in Relevant Tweets (United States, 2010-2021)***

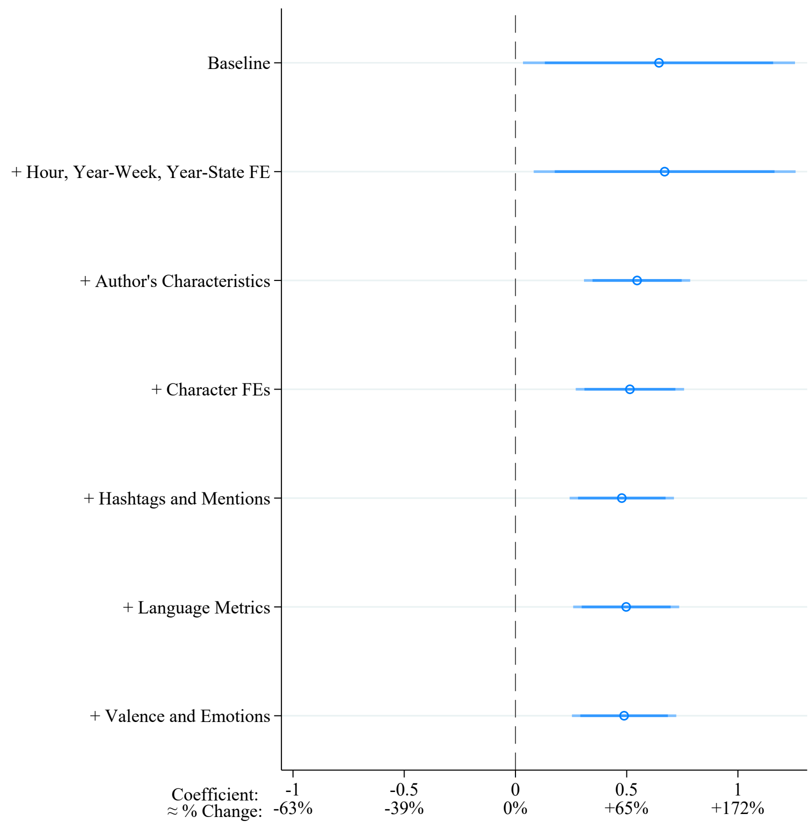


Notes: The figure shows the log–log rank distribution of likes for relevant tweets, distinguishing between those with at least one character–role (blue) and those with characters only in a neutral form (gray). The x-axis plots the logarithm of the rank, with rank 1 as the most liked tweet in the sample. The y-axis plots the logarithm of like counts (plus one). The slope of the curve indicates how quickly the distribution falls from the most popular to the least popular tweets: a steeper slope means engagement is clustered in a handful of viral tweets, whereas a flatter slope indicates that engagement is more evenly distributed. The vertical position at a given rank reflects relative popularity: a higher curve means tweets at that rank receive more engagement than tweets at the same rank in a dataset with a lower curve. The Paper Figure 5 shows the same for virality (retweets) of tweets.

We turn next to regression analysis, extending the descriptive findings by estimating the same models used in Figure 7, this time applied to popularity. Figure E.4 presents a coefficient plot based on Poisson Pseudo-Maximum Likelihood estimates. The models follow a sequential design, becoming progressively more restrictive from the top to the bottom panel. Each panel builds on the previous one by adding further controls, as noted in the titles. As for the paper figure, the coefficients can be interpreted as approximate percentage changes using the transformation: $\approx e^{\beta} - 1$.

The results align closely with those presented in the main paper for tweets' virality. Consistent with the findings on retweets, *political narratives* have a clear positive effect on popularity. Across all model specifications, tweets that include at least one character-role receive more likes than those that present characters in a neutral way. As in the main analysis, user characteristics emerge as an important control: accounting for them sharpens the results and helps reduce noise in the estimation.

Figure E.4: ***Regression Results - Impact of Political Narratives on Popularity (likes)***



Notes: The figure shows the coefficients of Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring at least one character-role vs. featuring characters only in a neutral role on popularity, measured as the count of likes. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx e^{\beta} - 1$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour, week of the year, and year-state fixed effects. The third model controls for author characteristics: verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status. The fourth model adds character fixed effects. The fifth model controls for the tweet's number of hashtags and mentions. The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above. The Paper Figure 7 shows the same for virality (retweets) of tweets.

In the final part of this robustness check, we take a closer look at *political narratives* by examining the individual impact of the three drama triangle roles: Hero, Villain, and Victim. Table E.1 reports Poisson Pseudo-Maximum Likelihood estimates of how each role affects the number of likes received by tweets. The analysis is restricted to tweets where roles are mutually exclusive – that is, if a tweet features a Hero, it does not also include a Villain or a Victim. Columns 1 to 3 compare Hero, Villain, and Victim narratives separately to tweets with only neutral framing. Columns 4 to 7 include all roles simultaneously, changing the reference category in each column: first neutral (column 4), then Hero, Villain, and Victim, respectively.

When comparing these results to those found for virality, several meaningful patterns emerge, highlighting both parallels and distinctions between the effects on retweets and likes. In both cases, Hero and Villain narratives significantly boost engagement relative to tweets that feature characters in neutral roles. Column 1 of Table E.1 shows that Hero narratives increase the number of likes by approximately 53% (computed as $\approx e^{\beta} - 1$ for Poisson models). Villain narratives, however, produce an even stronger effect, with a coefficient nearly twice as large. These effect sizes closely match those observed for retweets, indicating that narratives centered around Heroes and Villains consistently draw higher engagement. By contrast, Victim narratives do not show a statistically significant effect on either metric, suggesting that this role is less effective at driving user interaction.

When looking at the full models in columns 4 through 7, a more nuanced picture emerges. Villain narratives consistently stand out, outperforming both Hero and Victim narratives in driving engagement. This supports the idea that narratives built around conflict and opposition tend to be the most effective at capturing attention. The effect of Hero narratives remains statistically significant only when compared to neutral tweets. This suggests that while Heroes can boost engagement, their impact is less pronounced when set against the stronger pull of Villain narratives.

Lastly, we reproduce the results on the interactions of roles in Figure E.5. Each panel presents the results of a single Poisson Pseudo-Maximum Likelihood regression model. In each model, the comparison group consists of tweets featuring only one character-role (for example, a single hero narrative). The regressors capture all combinations where that role appears alongside other roles (e.g., hero + villain, hero + victim, hero + villain + victim) or where the same role applies to multiple characters (e.g., two different hero character-roles within the same tweet).

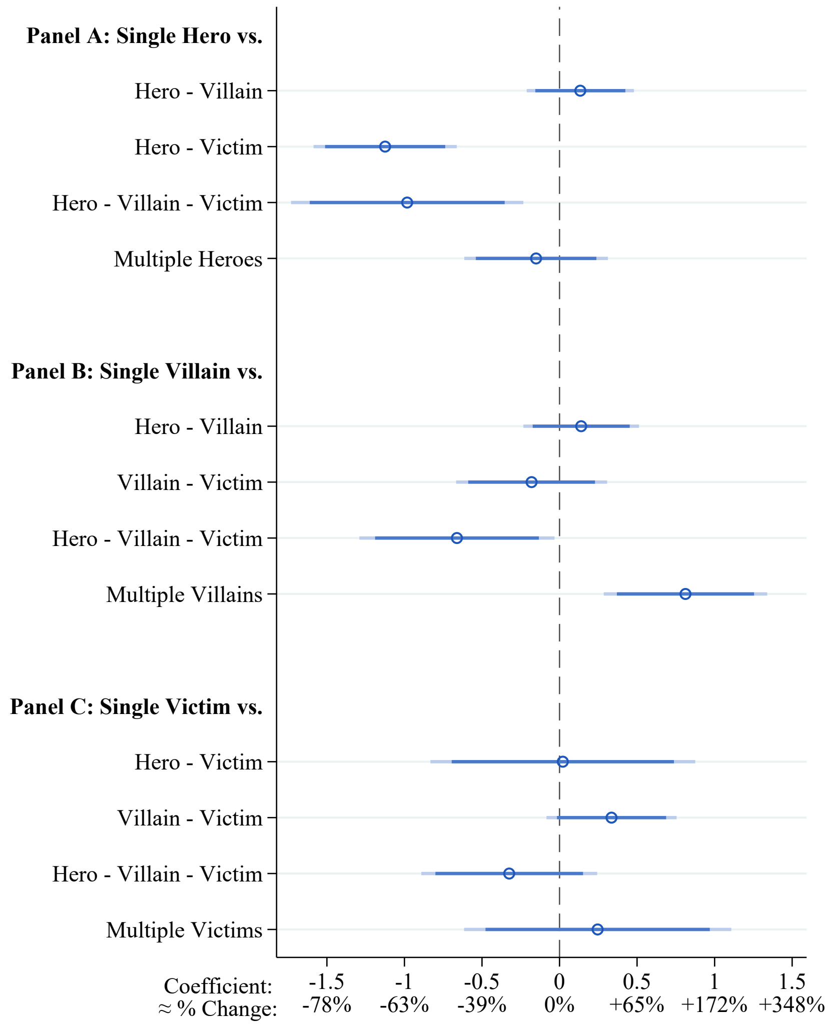
The results for virality (retweets) and popularity (likes) as a robustness check are strikingly similar. Compared to tweets featuring a single-character hero narrative, increasing complexity by introducing a hero-victim or hero-victim-villain structure reduces overall virality. However, when the hero role is combined exclusively with a villain, virality increases, while popularity remains unchanged. A similar pattern emerges for villain narratives, where complexity tends to reduce virality, particularly when the full Drama Triangle is present. Notably, when villains are reinforced within a simple narrative – meaning multiple characters share the villain role – virality increases even further, suggesting that narratives centered around multiple antagonists tend to gain more traction. For victim narratives, no consistent pattern emerges, except when a victim is paired with a villain, which significantly boosts virality.

Table E.1: ***Regression Results - Impact of Individual Roles on Popularity (likes)***

Dependent Variable	**Likes' Count**						
	Coeff./SE/p-value						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hero	0.430			0.434		-0.518	0.100
	(0.211)			(0.214)		(0.280)	(0.288)
	[0.042]			[0.042]		[0.064]	[0.728]
Villain		0.823		0.952	0.518		0.618
		(0.123)		(0.250)	(0.280)		(0.321)
		[0.000]		[0.000]	[0.064]		[0.054]
Victim			0.073	0.334	-0.100	-0.618	
			(0.306)	(0.267)	(0.288)	(0.321)	
			[0.811]	[0.210]	[0.728]	[0.054]	
Neutral					-0.434	-0.952	-0.334
					(0.214)	(0.250)	(0.267)
					[0.042]	[0.000]	[0.210]
Sample: Neutral	✓	✓	✓	✓	✓	✓	✓
Sample: Hero	✓			✓	✓	✓	✓
Sample: Villain		✓		✓	✓	✓	✓
Sample: Victim			✓	✓	✓	✓	✓
Mean Outcome Reference Group	9.523	9.523	9.523	9.523	34.136	16.535	13.269
Pseudo R2	0.88	0.72	0.70	0.80	0.80	0.80	0.80
Obs	74739	88417	43084	134948	134948	134948	134948

Notes: The table reports coefficients from Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring Hero, Villain, and Victim roles on popularity, measured by like counts. The sample includes relevant tweets featuring at most one character, either framed neutrally or in one of the drama-triangle roles, so that tweets are mutually exclusive (e.g., a tweet coded as Hero cannot also be coded as Neutral or Villain). Column 1 compares Hero tweets to the Neutral baseline; Column 2 compares Villain to Neutral; Column 3 compares Victim to Neutral. Columns 4–7 include all groups but vary the reference category: Neutral (Col. 4), Hero (Col. 5), Villain (Col. 6), and Victim (Col. 7). Coefficients are interpreted as percentage changes using the transformation $e^{\beta} - 1$. All regressions control for author characteristics (verified status, number of followers/followings, total tweets created, party affiliation, religiosity, higher education, and parenthood) and include character fixed effects. We also include hour, week-of-year, and year–state fixed effects. Standard errors are clustered at the week level, covering the full time frame. Paper Table 3 shows the same for virality (retweets) of tweets.

Figure E.5: ***Regression Results - Heterogeneity in the Impact of Character-Role Combinations on Popularity (likes)***



Notes: The figure shows coefficients from Poisson Pseudo-Maximum Likelihood regression models testing the effect of character-role combinations on popularity, measured as the tweet-level count of likes. The x-axis reports coefficient estimates together with the corresponding approximate percentage change, computed as $e^{\beta}-1$. Each panel corresponds to a separate regression model. Panel A examines the effects of hero combinations, with the comparison group being tweets featuring a single hero character only. The regressors capture cases where a hero is paired with one or more villains, with one or more victims, with both villains and victims, or with additional heroes. Panel B follows the same structure for villains, with the reference group being tweets that feature a single villain character, while Panel C does so for victims, using tweets with a single victim character as the baseline. All regressions include hour, week of the year, and year–state fixed effects. Standard errors are clustered at the week level, covering all weeks in the study period. The Paper Figure 8 shows the same for virality (retweets) of tweets.

E.3 Impact of Narratives on Conversation

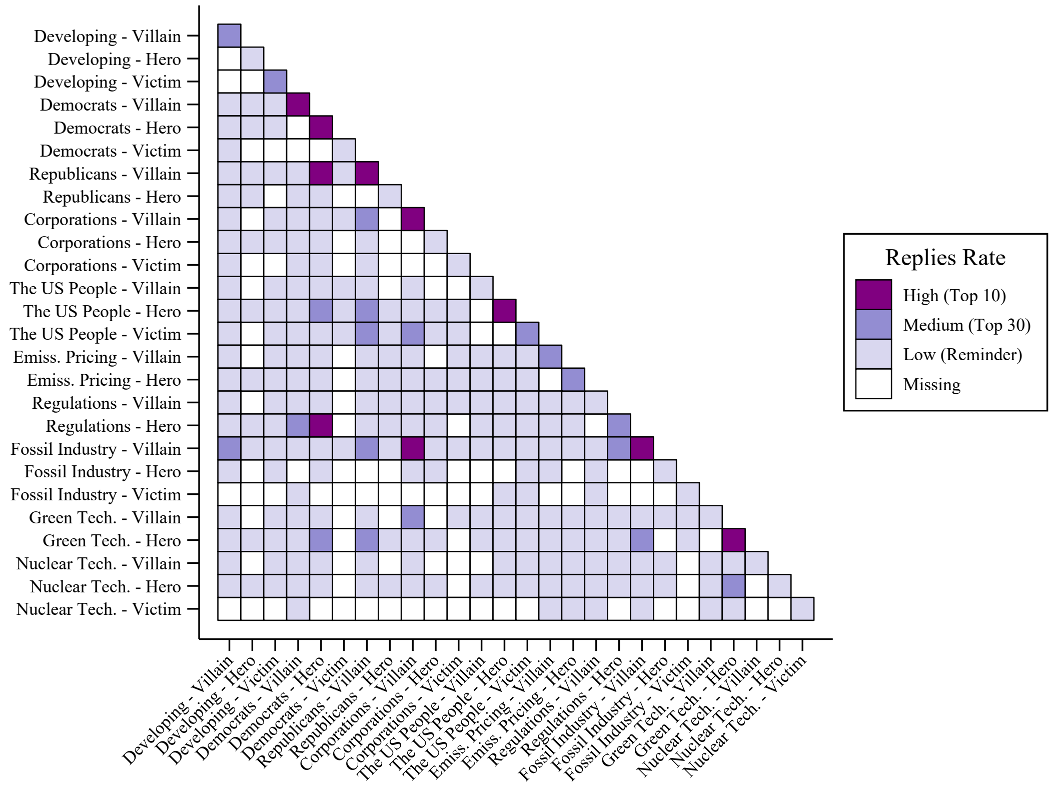
In this section of Appendix E.3, we provide a robustness check on the impact of narratives by examining their effect on conversation. We measure conversation using the count of comments a post receives. We do not include this outcome in the main analysis, as it captures a slightly different concept than virality. While a high number of comments may reflect popularity and engagement, it can also result from a small number of users exchanging opinions or arguing in the comment section. Despite being a less precise measure of virality, conversation provides a valuable robustness check.

We begin this exercise by exploring descriptive statistics on the rate of conversation for narratives featuring one or at most two character-roles, in line with the main results of the paper. Figure E.6 shows a heatmap of conversation rates for all combinations of one or two character-roles. The color of each square reflects the conversation rate for that combination, calculated by dividing the total number of comments received by tweets with that combination by the total number of comments across all tweets featuring one or two character-roles. The diagonal indicates the conversation rate for each character-role occurring alone. The matrices are symmetrical. Numerical values are not displayed to avoid visual overload, but the top 10 and top 30 most frequent combinations are highlighted using a color scheme.

Figure E.6 suggests a pattern in conversation rates that closely mirrors the patterns observed for virality and the robustness specification on popularity (likes). As with retweets and likes, the highest rates of conversation cluster around the more politicized narratives featuring characters US Democrats and US Republicans. High conversation rates also appear around some of the instrument characters: Regulations, the fossil industry, and Green Tech prompt many comments, especially in tweets framing the fossil industry as the Villain of the climate change debate.

One striking difference from the virality patterns is the prominence of Emission Pricing in the conversation. Whether framed as hero or villain, Emission Pricing generates more discussion than retweets or likes. This suggests that when attention shifts to policy instruments – such as carbon pricing – users are more inclined to engage in extended debates rather than simply signaling support for one political side. The contrast is less pronounced for popularity, however, as Emission Pricing framed as hero also ranks among the top 30 in terms of likes.

Figure E.6: ***Conversation (Replies) on Character-Roles - Tweets with One or Two Character-Roles***

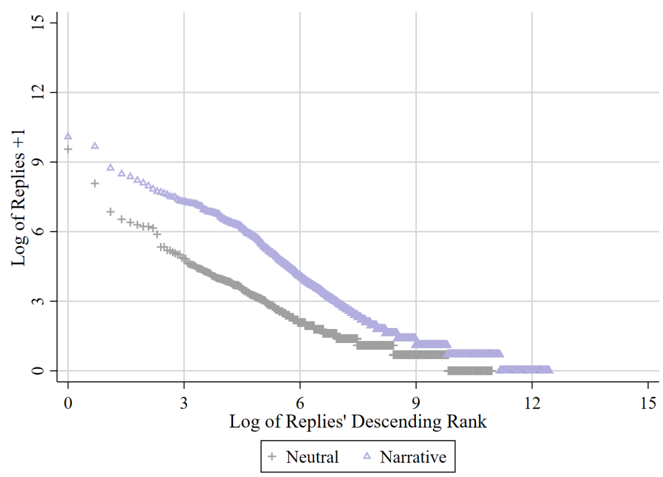


Notes: The figure shows the like rate of each character–role appearing alone or in combination with another role, among relevant tweets containing one or two roles. Replies rates are computed as the share of total comments received by a given role (or pair) relative to all comments of tweets with one or two roles. The diagonal of the matrix shows the replies rate when each character-role appears alone. Tweets with three or more character-roles are excluded. To avoid visual overload, we do not display exact rates. Instead, we use a color scheme to highlight the top 10 most frequently commented character-role combinations, the top 30 (which includes the top 10), and the remaining pairs. White indicates a pair that never appears together. The top 10 in order is: US Democrats-Hero (22.33%), US Republicans-Villain (11.64%), US Democrats-Hero + US Republicans-Villain (4.99%), US People-Hero (6.24%), US Democrats-Villain (4.45%), Fossil Industry-Villain (4.07%), Green Tech-Hero (4.03%), Corporations-Villain (3.01%), Fossil Industry-Villain + Corporations-Villain (2.85%), Regulations-Hero + US Democrats-Hero (2.64%). Paper Figure 5 shows the same for virality (retweets) of tweets.

We take a step forward into the analysis of the impact of narratives on conversation by exploring the distribution of comments in tweets featuring a narrative and tweets featuring characters only in neutral framing. Similar to what is presented in Subsection 4.3 of the paper for virality, Figure E.7 shows the log-log rank distribution of the count of comments, in tweets featuring a narrative (triangle shape) and tweets featuring no narrative (cross shape). The x-axis represents the logarithm of the rank of tweets based on their comment count, with rank 1 corresponding to the tweet that received the most comments. The y-axis represents the logarithm of the number of comments, plus one.

The distribution reveals several interesting insights. First, the distribution of comments appears to follow a power law, consistent with what we observe for retweets. The linear shape of the log-log rank distribution suggests that the data generation process follows an exponential structure, with most tweets receiving few comments and a small number receiving many. Second, compared to retweets, the curves are flatter and shifted downward, reflecting the smaller number of comments overall. Lastly, the figure provides an initial indication of the effect of narratives. The curve for narrative tweets lies above that of neutral tweets. At each rank, tweets featuring a narrative receive more comments, suggesting that *political narratives* may not only spark more engagement but also encourage greater participation through conversation.

Figure E.7: ***Conversation on Political Narratives in Relevant Tweets (United States, 2010-2021)***

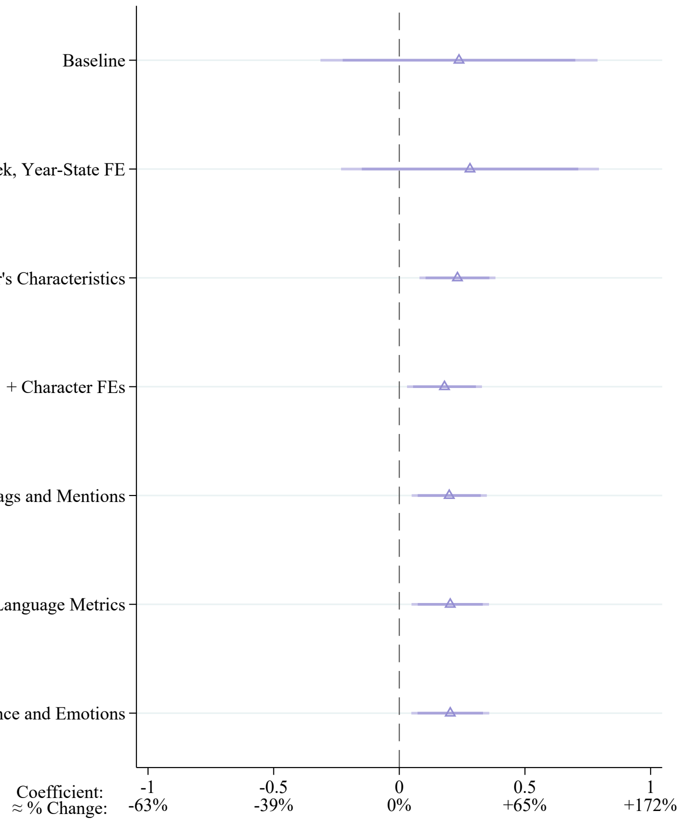


Notes: The figure shows the log–log rank distribution of replies for relevant tweets, distinguishing between those with at least one character–role (blue) and those with characters only in a neutral form (gray). The x-axis plots the logarithm of the rank, with rank 1 as the tweet with most replies in the sample. The y-axis plots the logarithm of replies count (plus one). The slope of the curve indicates how quickly the distribution falls from the most popular to the least popular tweets: a steeper slope means engagement is clustered in a handful of viral tweets, whereas a flatter slope indicates that engagement is more evenly distributed. The vertical position at a given rank reflects relative popularity: a higher curve means tweets at that rank receive more engagement than tweets at the same rank in a dataset with a lower curve. The Paper Figure 5 shows the same for virality (retweets) of tweets.

We move from descriptive to regression analysis by reproducing the results of Figure 7 for conversation. Figure E.8 presents a coefficient plot showing the results of models that mirror those used for virality as a robustness check. The models are estimated using Poisson Pseudo-Maximum Likelihood and apply increasingly restrictive specifications, moving from the top panel to the bottom panel. The additions to each specification are indicated in the panel titles. Coefficients can be interpreted as percentage changes using the following transformation: $\approx e^{\beta} - 1$.

The results are mostly consistent with those found for virality in the paper, and popularity in the appendix. Although the first two models have large confidence intervals – suggesting noisier measures compared to retweets and likes – the overall effect indicates that narratives spark more conversation than tweets where characters appear only in neutral framing. The author fixed effects play an important role by reducing excess noise and stabilizing the coefficients around 0.2, which corresponds to roughly a 22% increase in the number of comments on tweets featuring *political narratives*. The size of this coefficient effect is about half that observed for retweets and likes, but narratives still appear to play an important role in driving conversation.

Figure E.8: ***Regression Results - Impact of Political Narratives on Conversation***



Notes: The figure shows the coefficients of Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring at least one character-role vs. featuring characters only in a neutral role on conversation, measured as the count of replies. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx e^{\beta} - 1$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour, week of the year, and year-state fixed effects. The third model controls for author characteristics: verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status. The fourth model adds character fixed effects. The fifth model controls for the tweet's number of hashtags and mentions. The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above. The Paper Figure 7 shows the same for virality (retweets) of tweets.

In the final part of this exercise, we use the flexibility of our framework to move beyond the overall effect of *political narratives*. We ask which roles drive the increase in conversation and how the drama triangle roles interact to shape this effect. Table E.2 presents Poisson Pseudo-Maximum Likelihood regression models estimating the impact of the drama triangle roles on the number of comments received by tweets. All models include only tweets where roles are mutually exclusive, that is, e.g. if a hero appears in a tweet, that tweet does not contain a villain or a victim. The first three columns compare, respectively, hero, villain, and victim narratives to tweets featuring only neutral framing. Columns 4 to 7 include all roles at once, changing the comparison group for each column: starting with neutral (column 4), then hero, villain, and victim, respectively.

The results of these models provide a clear picture. Unlike what we observed for virality (retweets) and popularity (likes) as a robustness check, when focusing on mutually exclusive roles, only villain narratives have a meaningful impact on the amount of conversation. Whether directly compared to neutral tweets or in the full models, hero and victim narratives do not significantly increase the number of comments. The prominence of villain narratives may have several explanations, though it is important to remember that this is only correlational evidence. These results support the intuition proposed earlier: conversation may be driven more by argument and confrontation, which villain narratives are more likely to provoke, rather than by simple expressions of support or empathy.

Lastly, we reproduce the results on the interactions of roles in Figure E.9. Each panel presents the results of a single Poisson Pseudo-Maximum Likelihood regression model. In each model, the comparison group consists of tweets featuring only one role and only one character-role (for example, a single hero narrative). The regressors capture all combinations where that role appears alongside other roles (e.g., hero + villain, hero + victim, hero + villain + victim) or where the same role applies to multiple characters (e.g., two different hero character-roles within the same tweet).

The models reveal patterns similar to those observed for virality in the paper, and popularity as a robustness check. Although not always statistically significant, the results suggest that adding complexity to narratives often reduces the effect size. For all three roles – hero, villain, and victim – when all roles are featured together, forming a more nuanced narrative, the number of comments decreases significantly. Despite smaller effect sizes, the impact of multiple villains aligns with the results for retweets and likes: more villains tend to spark more comments. The only exception to this pattern is the victim role. Unlike for retweets and likes, tweets featuring multiple victims seem to increase conversation compared to those with a single victim.

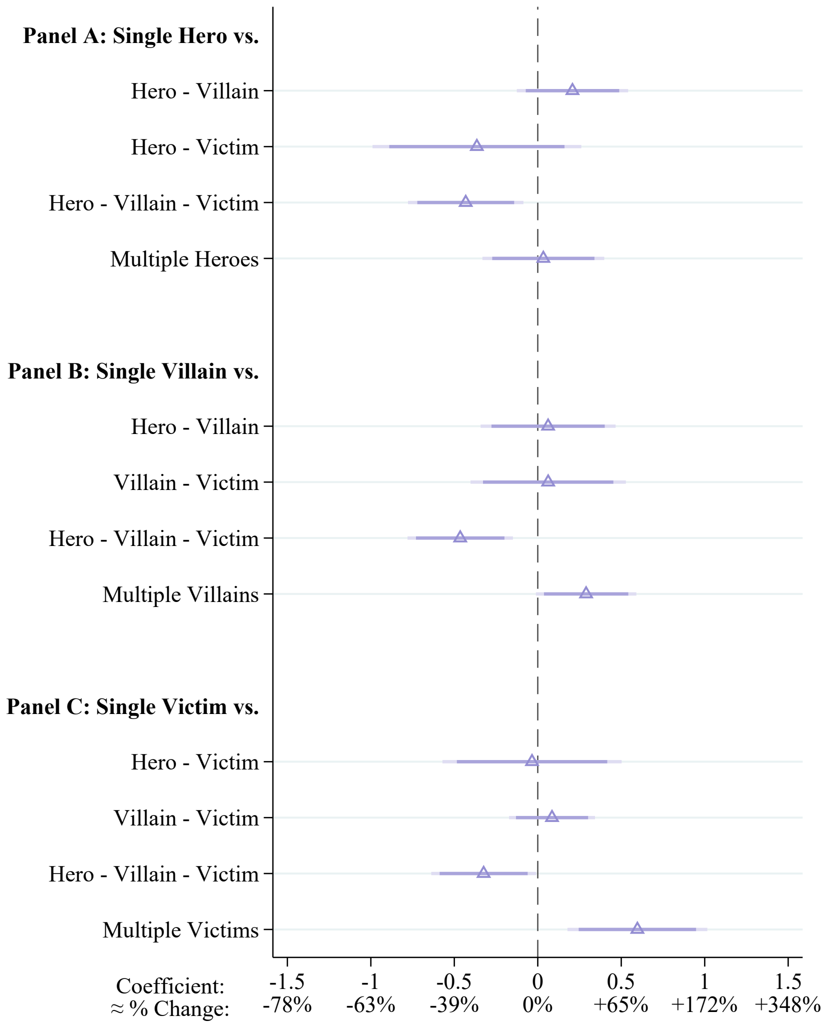
Overall, the results on conversation are strongly consistent with those for virality (retweets) and popularity (likes) as a robustness check. Villain narratives appear even more important in driving virality through conversation. However, featuring multiple victims also has a positive effect on the number of comments.

Table E.2: ***Regression Results - Impact of Individual Roles on Conversation***

Dependent Variable	**Replies' Count**						
	Coeff./SE/p-value						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hero	0.076			0.019		-0.352	-0.081
	(0.118)			(0.132)		(0.142)	(0.203)
	[0.521]			[0.887]		[0.013]	[0.690]
Villain		0.243		0.371	0.352		0.271
		(0.071)		(0.116)	(0.142)		(0.187)
		[0.001]		[0.001]	[0.013]		[0.147]
Victim			0.130	0.100	0.081	-0.271	
			(0.165)	(0.163)	(0.203)	(0.187)	
			[0.431]	[0.540]	[0.690]	[0.147]	
Neutral					-0.019	-0.371	-0.100
					(0.132)	(0.116)	(0.163)
					[0.887]	[0.001]	[0.540]
Sample: Neutral	✓	✓	✓	✓	✓	✓	✓
Sample: Hero	✓			✓	✓	✓	✓
Sample: Villain		✓		✓	✓	✓	✓
Sample: Victim			✓	✓	✓	✓	✓
Mean Outcome Reference Group	1.972	1.972	1.972	1.972	4.350	3.683	2.395
Pseudo R2	0.79	0.54	0.49	0.70	0.70	0.70	0.70
Obs	74554	88592	43191	134041	134041	134041	134041

Notes: The table reports coefficients from Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring Hero, Villain, and Victim roles on conversation, measured by replies count. The sample includes relevant tweets featuring at most one character, either framed neutrally or in one of the drama-triangle roles, so that tweets are mutually exclusive (e.g., a tweet coded as Hero cannot also be coded as Neutral or Villain). Column 1 compares Hero tweets to the Neutral baseline; Column 2 compares Villain to Neutral; Column 3 compares Victim to Neutral. Columns 4–7 include all groups but vary the reference category: Neutral (Col. 4), Hero (Col. 5), Villain (Col. 6), and Victim (Col. 7). Coefficients are interpreted as percentage changes using the transformation $e^{\beta} - 1$. All regressions control for author characteristics (verified status, number of followers/followings, total tweets created, party affiliation, religiosity, higher education, and parenthood) and include character fixed effects. We also include hour, week-of-year, and year–state fixed effects. Standard errors are clustered at the week level, covering the full time frame. Paper Table 3 shows the same for virality (retweets) of tweets.

Figure E.9: ***Regression Results - Heterogeneity in the Impact of Character-Role Combinations on Conversation***



Notes: The figure shows coefficients from Poisson Pseudo-Maximum Likelihood regression models testing the effect of character-role combinations on conversation, measured as the tweet-level count of replies. The x-axis reports coefficient estimates together with the corresponding approximate percentage change, computed as $e^{\beta} - 1$. Each panel corresponds to a separate regression model. Panel A examines the effects of hero combinations, with the comparison group being tweets featuring a single hero character only. The regressors capture cases where a hero is paired with one or more villains, with one or more victims, with both villains and victims, or with additional heroes. Panel B follows the same structure for villains, with the reference group being tweets that feature a single villain character, while Panel C does so for victims, using tweets with a single victim character as the baseline. All regressions include hour, week of the year, and year–state fixed effects. Standard errors are clustered at the week level, covering all weeks in the study period. The Paper Figure 8 show the same for virality (retweets).

E.4 Output Excluding Potential Bots

In this section of Appendix E we provide an additional robustness check on the main results of the paper about the virality of *political narratives*. In particular, we address the important issue of bot activity on the social media platform Twitter/X. In this context, a bot can be defined as an account operated by an algorithm, programmed to automate actions such as generating content, liking, retweeting, and commenting on other users' posts. Due to the automated nature of their behavior, bots are capable of interacting with a large number of users and can, at times, achieve considerable visibility. Given this potential influence, we test whether our results on the determinants of virality are affected by the presence and activity of bots.

Identifying bots on social media is challenging, as there is no universally accepted definition or detection method. Tools such as Botometer offer sophisticated ways to classify accounts as bots or humans. However, these tools require a substantial number of tweets per user to be effective, which is a limitation in our case due to data constraints. Furthermore, the Twitter/X API is no longer accessible, preventing us from employing such tools. We therefore adopt a simpler and more practical approach tailored to our dataset.

We implement two complementary strategies to detect potential bots, based on both tweet content and user characteristics. Specifically, we define tweets as originating from potential bots if they meet either one or both of the following two conditions:

1. **Repeated identical text**: Within our dataset, we observe instances where identical tweets appear multiple times. These are exact text duplicates, yet each is assigned a distinct tweet ID by the Twitter/X API, confirming that they are separate posts. We classify a tweet as bot-generated if its text appears at least five times, whether posted by the same author or by different authors. We interpret such repetitions as attempts to amplify specific narratives or content.

2. **Authors' activity patterns**: Following Chu et al. (2012), we use the *reputation* metric, calculated as the number of followers divided by the sum of followers and followees of a user. Bots typically have a low reputation, as they tend to follow many accounts indiscriminately. Additionally, we incorporate insights from Tabassum et al. (2023), who highlight the unusually high activity levels of bots. Specifically, bots tend to produce an exceptionally large number of tweets. Combining these two indicators, we flag tweets as bot-generated if their authors fall in the bottom 25% of the reputation distribution and simultaneously in the top 25% of the distribution of total tweets produced. The threshold of 25% is purely discretional but we argue it is a pretty conservative choice.

There is little overlap between the two definitions, hopefully indicating that we capture different kinds of bots successfully. Among our relevant tweets, a total of 16,700 tweets were identified through definition 1., a total of 5,748 through definition 2., while only 115 overlap. As a reminder we define relevant tweets as those featuring at least one character from our list. We include only character-roles that appear at least 100 times, thus excluding 'US Republicans-victim', 'Emission Pricing-victim', 'Regulations-victim', and 'Green Tech-victim'.

The first concern is that bots may have inflated the presence of certain narratives in the dataset. Therefore, our first check examines the distribution of narratives. Table E.3 corresponds to paper Table 2, but accounts for potential bot activity by dropping tweets potentially posted by bots. It reports the share of character-roles and neutral instances for each character of interest, excluding those character-role that did not reach at least 100 occurrences in the full period, in line with the paper's table.

The results suggest that bots did not disproportionately promote specific narratives. Removing all (potentially) bot-generated tweets does not substantially change the composition of narratives. Fossil Industry–Villain and Green Tech–Hero remain the most common narratives. The share of Green Tech–Hero drops slightly once bot activity is excluded, which may suggest some degree of coordinated posting. However, all other shares remain virtually unchanged.

We take a step further to strengthen these descriptive findings by reproducing Figure 7, which presents the main results of the paper. Figure E.10 plots the coefficients from the same models used in the paper, but excludes tweets potentially posted by bots. The models apply increasingly restrictive specifications, moving from the baseline in the top panel to the most controlled model in the bottom panel. The results remain virtually unchanged, suggesting that bots do not play a central role in shaping the discussion on climate change policy, at least within the scope of our dataset.

Table E.3: ***Share of Character-Roles in Relevant Tweets, Excluding Potential Bots (United States, 2010-2021)***

Panel A: Human Characters

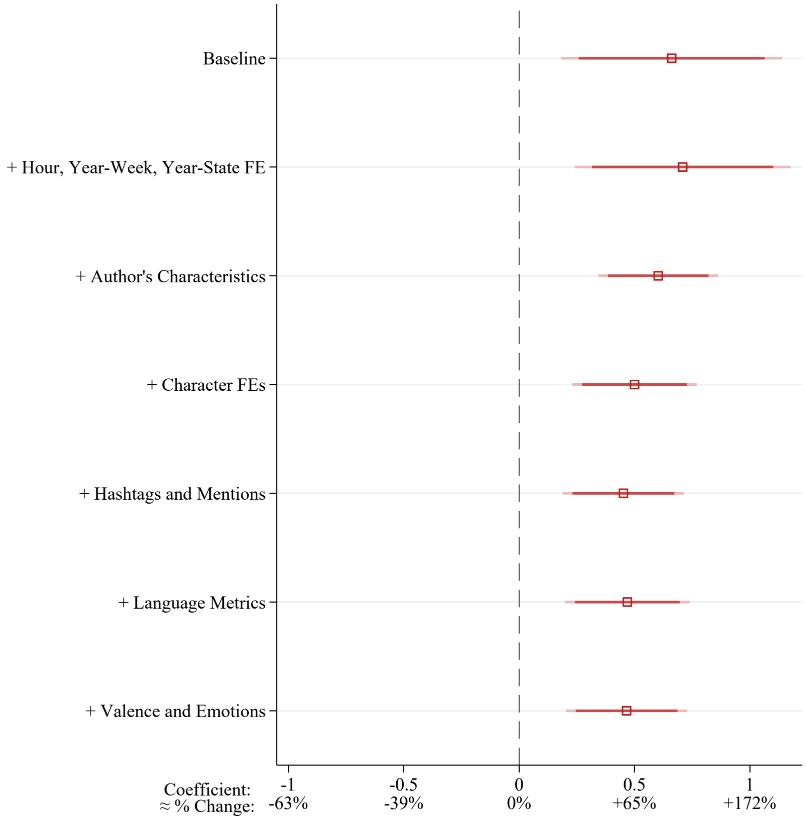
	Hero	Villain	Victim	Neutral	Total
Developing Economies	0.13	1.25	0.88	0.20	2.46
US Democrats	5.83	1.90	0.06	1.42	9.21
US Republicans	0.12	9.91	.	1.66	11.69
Corporations	0.96	8.10	0.06	7.78	16.90
US People	4.04	0.58	3.25	9.68	17.55

Panel B: Instrument Characters

	Hero	Villain	Victim	Neutral	Total
Emission Pricing	2.11	1.30	.	1.24	4.65
Regulations	3.51	1.44	.	3.14	8.09
Fossil Industry	0.07	11.48	0.04	1.67	13.26
Green Tech	10.41	0.88	.	3.16	14.45
Nuclear Tech	0.92	0.33	0.02	0.47	1.75

Notes: The table displays the frequency of character-roles in the classified data as shares of the total occurrences of each character. Shares are computed considering only the dataset of relevant tweets used in our analysis. For the analysis we exclude tweets potentially produced by bots, where bots are defined as explained in Appendix Subsection E.4. We define a tweet as relevant if it features at least one character from our list. We include in the computation of shares only character roles that appear at least 100 times, thus excluding 'US Republicans-victim', 'Emission Pricing-victim', 'Regulations-victim', and 'Green Tech-victim', indicated by a dot in the tables. Panel (a) displays the shares for characters of the human type, while Panel (b) displays the same for characters of the instrument type. The column Neutral in both panels reports cases where the character is present in the tweet but is not depicted in one of the three specific roles. The occurrence of character-roles is not mutually exclusive, meaning multiple roles may appear in the same tweet. For this analysis we exclude tweets produced by potential bots, following the definitions provided.

Figure E.10: ***Regression Results - Impact of Political Narratives on Virality, Excluding Potential Bots***



Notes: The figure shows the coefficients of Poisson Pseudo-Maximum Likelihood regression models testing the effect of featuring at least one character-role vs. featuring characters only in a neutral role on virality, measured as the count of retweets. For the analysis we exclude tweets potentially produced by bots, where bots are defined as explained in Appendix Subsection E.4. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx e^{\beta} - 1$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour, week of the year, and year-state fixed effects. The third model controls for author characteristics: verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status. The fourth model adds character fixed effects. The fifth model controls for the tweet's number of hashtags and mentions. The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above.

In conclusion, this section provides evidence that bots have limited influence on the conversation about climate change policy. However, these results should be interpreted with caution and should not be taken to mean that bots are unimportant in shaping social media discussions more broadly. First, in recent years, significant improvements in bot programming may have enhanced their performance in ways not captured during our study period. Second, while climate change policy does not appear to be heavily affected, other topics – such as war, abortion, or general politics – may be much more vulnerable to bot activity. Finally, as noted above, there is no exact method for identifying bots, and the approach we use may have limitations in accurately detecting automated accounts.

E.5 Alternative Regression Models

In this section, we want to dive deeper into the choice of model specifications used in our analysis. We explain the motivations and reasoning behind our decisions, while also considering potential alternative solutions. Our modeling choices are driven by the particular nature of the data and the research questions we address. In particular, the outcome variables of our analysis – retweets in the paper, likes and replies in the appendix – follow a power-law distribution: many observations take low values (including zeros), while a few take extremely high values. Such distributions are common in social media engagement and in other domains shaped by self-reinforcing processes. They pose unique challenges and require careful thinking when selecting an appropriate modeling strategy. In similar cases, researchers often apply a log transformation to ease interpretation and reduce skewness before estimating models using Ordinary Least Squares (OLS). This would be a sensible solution if the dependent variables were strictly positive. However, our measures contain a large proportion of zeros, which makes log transformations problematic.

Recent work by Chen and Roth (2024) highlights how applying log transformation, can introduce important biases. Specifically, adding a constant to handle zeros - e.g., log(y + 1) - effectively translates into rescaling the outcome variable in a way that can arbitrarily inflate or deflate estimated effects. Formally, any treatment effect estimate becomes a function of a scaling factor (denoted *"a"* in Chen and Roth (2024)), which depends on both the chosen constant and the distribution of the dependent variable. In the worst case, researchers could manipulate estimated effects simply by adjusting this scaling factor.

There are many contexts in economic research where data might seem suitable for such transformations, creating potential sources of faulty results. For example, if the dependent variable is e.g. hours worked, changing the unit from hours to days or weeks would improperly affect the coefficient after transformation – a clear violation of sound econometric practice. While such rescaling concerns are less pronounced for count data (like retweets or likes), the broader problem remains: log transformations with zeros can produce misleading or unstable estimates.

In light of these issues, for this study, we adopt Poisson Pseudo-Maximum Likelihood (PPML) regression models as our preferred choice. Poisson models are particularly suitable for count data, especially when the underlying distribution follows a power-law structure, as is common in social media engagement metrics like retweets and likes. These models naturally handle zeros, avoiding the need for arbitrary transformations that could make results misleading. They also yield coefficients that can be interpreted similarly to semi-elasticities, maintaining a clear and meaningful link between model estimates and real-world effects.

In the remainder of this section, we present some alternative approaches. Although we argue above that Poisson represents the best method for our analysis, we provide these alternative models to offer comparison and to improve understanding of the main results. In particular, we examine three approaches: OLS without any transformation, OLS applied after log-transforming the dependent variables, and the piecewise model proposed by Chen and Roth (2024), which distinguishes between extensive and intensive margin effects.

OLS Models

In the first part of this exercise we reproduce the results using simple OLS models. We do so leaving the dependent variable unvaried. The particular distribution of retweets and likes, following a power law, is problematic when using OLS and leads to a violation of the OLS assumptions, with regard to the distribution and variance of errors. Nevertheless, we want to provide a comparison to OLS, to highlight the differences with what we argue is the most correct modeling choice.

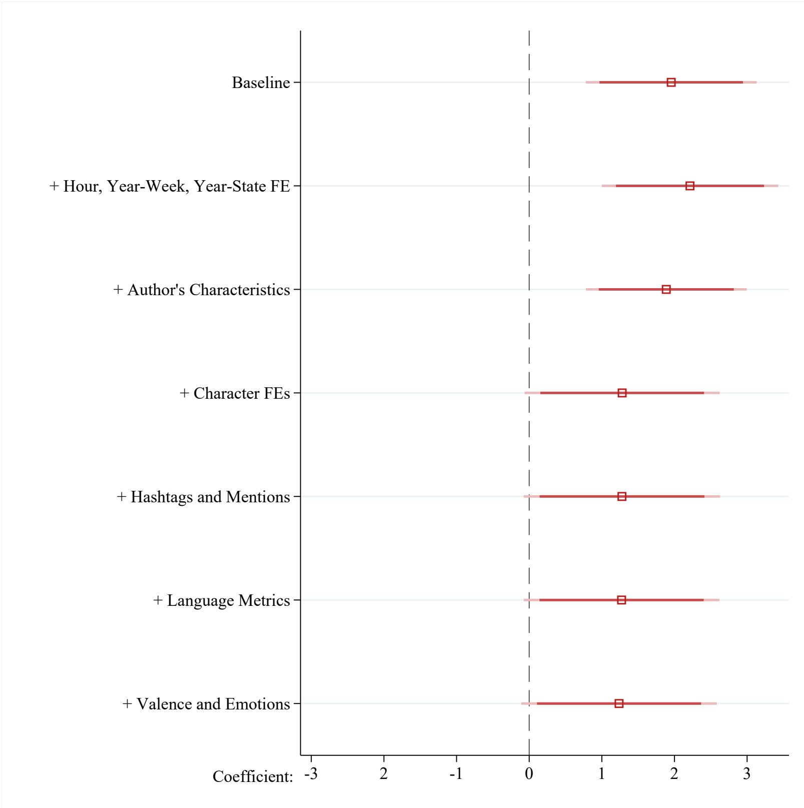
Figure E.11 reproduces the results of the paper Figure 7, where models are increasingly restrictive starting from the baseline model in the top panel, to the most restrictive one in the bottom panel. Overall the point estimates are in line with the main results of the paper. Featuring a Political Narrative increases the virality of tweets relative to tweets only featuring characters in a neutral framing. To add on what is said above, even if the point estimate can still be unbiased, the standard errors likely are not. That is why we should take with a grain of salt the significance levels of these coefficients.

In the second part of this exercise, we adopt an approach that is very common in economic research: applying a log transformation. Because our dependent variables contain many zeros, we first add a positive constant before taking the logarithm. While this practice is common in economics, as discussed above, it can lead to problematic results. Adding a constant to the dependent variable effectively rescales the coefficient by an arbitrary value. Changing the value of this constant alters the point estimates of the models, leading to potentially unreliable results. Figure E.13 serves two purposes. First, both sub-figures aim to reproduce the results of Figure 7 from the paper. Second, they illustrate the effects of adding different constants to the dependent variable before applying the log transformation to handle zeros. This allows us to demonstrate how changing the constant affects the estimated effects.

The first sub-figure on the left, Figure E.12a, applies the log transformation $Y + 1$. As in previous models, specifications become increasingly restrictive from the top to the bottom panel, with each panel's header indicating the additional controls included. Several key differences emerge. First, the percentage change implied by the logarithmic transformation is much smaller than the effects estimated using Poisson models. In the log-transformed models, the largest estimated increase is around 4%, compared to increases up to 160% in the Poisson models. The latter seems more consistent with the nature of the data. Nevertheless, the overall direction of the results remains in line with the main findings of the paper, providing an additional robustness check.

Figure E.12b also reproduces the results from Figure 7 in the paper, using log-transformed dependent variables. However, unlike Figure E.12a, this model adds 10 units to the dependent variable instead of 1 before applying the log transformation. If the transformation produced unbiased estimates, changing the constant should not affect the point estimates. Yet, in almost all models, the estimated effects are roughly half the size of those in the previous figure. Although the overall results still align with the main findings of the paper, this exercise highlights the weaknesses of log-transforming the dependent variable in this context. The instability of the estimates reinforces our argument that Poisson models provide a more reliable approach.

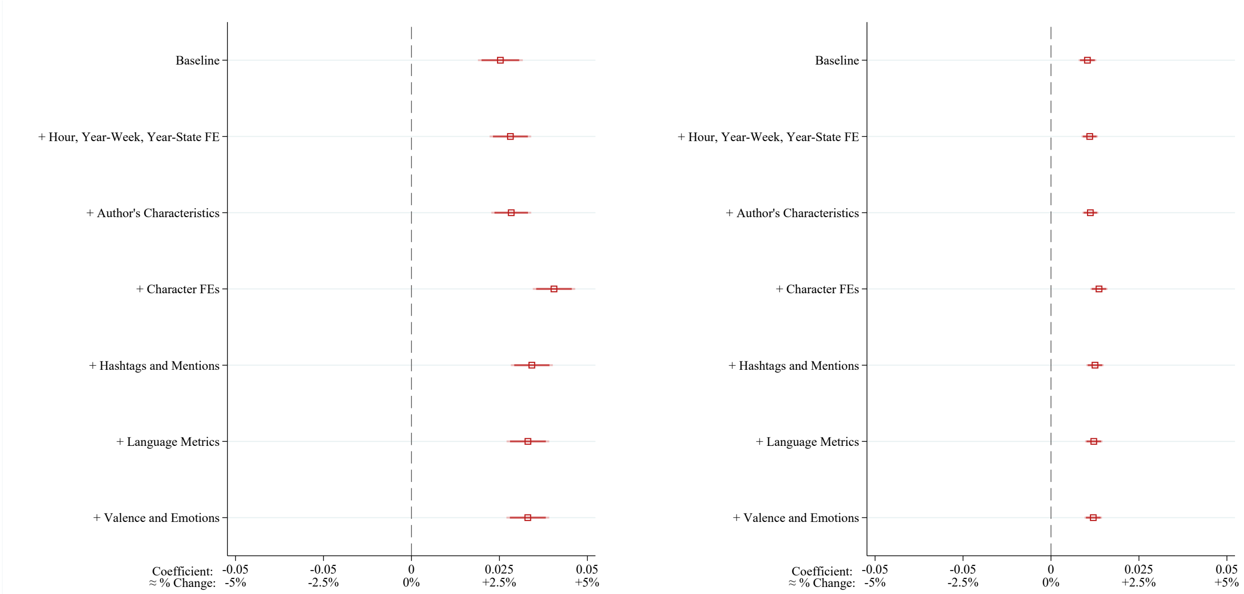
Figure E.11: ***Impact of Political Narratives on Virality - OLS Without Transformation***



Notes: The figure shows the coefficients of OLS regression models testing the effect of featuring at least one character-role vs. featuring characters only in a neutral role on virality, measured as the count of retweets. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx e^{\beta} - 1$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour, week of the year, and year-state fixed effects. The third model controls for author characteristics: verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status. The fourth model adds character fixed effects. The fifth model controls for the tweet's number of hashtags and mentions. The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above. The Paper Figure 7 is the reference plot.

Figure E.12: ***Impact of Political Narratives on Virality - OLS and Log Transformation***

(a) Transformation: $log(Y+1)$ **(b)** Transformation: $log(Y+10)$



Notes: The figure shows the coefficients of OLS regressions testing the effect of featuring at least one character-role, compared to featuring characters only in a neutral role, on virality, measured as the count of retweets. The dependent variable is log transformed after adding one unit, in Figure E.12a and 10 units, in Figure E.12b. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx \beta * 100$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour and year-state fixed effects. The third model accounts for author characteristics (verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status). The fourth model adds character fixed effects. The fifth model controls for tweet characteristics (number of hashtags and mentions). The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above.

Extensive and Intensive Margin Model

As a final step in this exercise, we adapt one of the suggestions proposed by Chen and Roth (2024). Specifically, we develop a piecewise regression that combines two models: one capturing the extensive margin effect and the other capturing the intensive margin effect of featuring narratives on virality. The dependent variable is transformed differently for each model. For the first model:

$$y^{*1} = \begin{cases} y = 1 & \text{if } y > 0 \\ y = 0 & \text{if } y = 0 \end{cases}$$

For the second model:

$$y^{*2} = \begin{cases} log(y) & \text{if } y > 0 \\ . & \text{if } y = 0 \end{cases}$$

The first model captures the extensive margin. It includes all observations but transforms any positive value of the dependent variable into 1, while leaving zeros unchanged. This approach captures whether featuring narratives increases the likelihood of any engagement with a tweet. By using OLS, we effectively estimate a Linear Probability Model, interpreting the results as the extensive margin of virality.

The second model captures the intensive margin. It retains only observations where the dependent variable is greater than zero. For this subset, we apply a log transformation to the dependent variable and estimate the model using OLS. This approach addresses a different question: Does featuring narratives increase or decrease the intensity of virality, conditional on receiving some engagement? In this model, the coefficients can be interpreted as semi-elasticities.

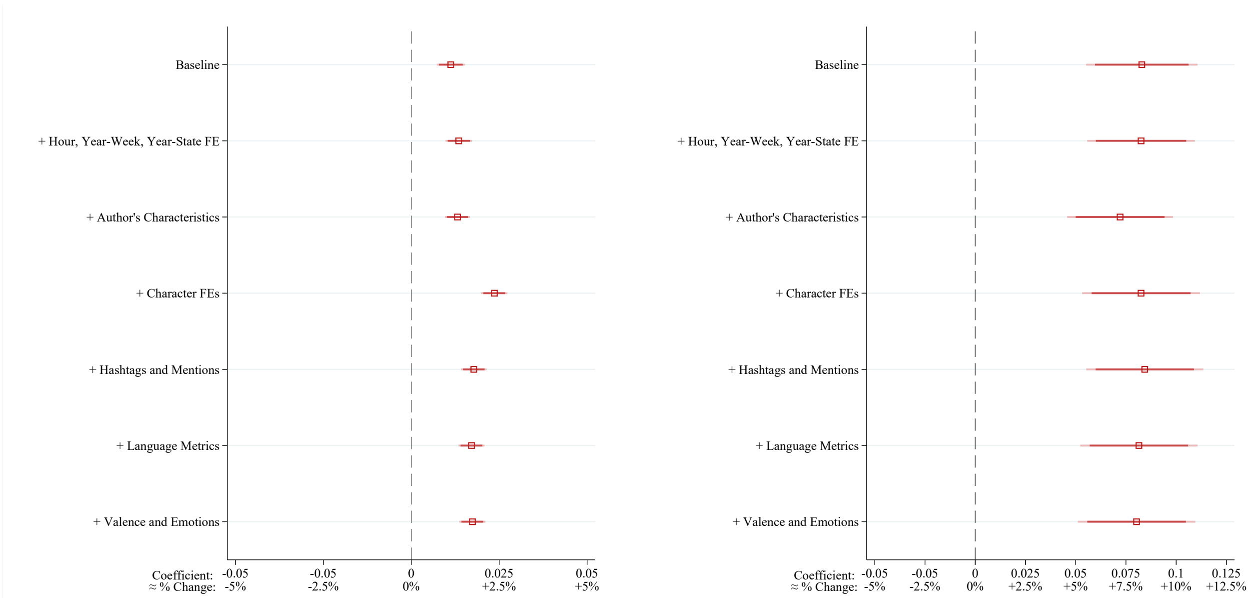
Figure E.13a and Figure E.13b show the extensive and intensive margin effects, respectively. The results point to a clear pattern. In both models, narratives have a positive impact on virality. Featuring a narrative increases both the likelihood of any user engagement and the degree of engagement, which we consistently refer to as virality throughout this study. The only exception appears with Popularity in the extensive margin model, where the effect is negative unless author characteristics are included. Once again, author fixed effects play an important role in reducing noise and producing more reliable estimates. We therefore consider the inclusion of author controls essential.

Summary

While the Poisson regression remains the theoretically preferred model for our data, the alternative specifications confirm that our core findings are robust across modeling choices. At the same time, these exercises highlight the potential pitfalls of common transformations and underscore the value of using models tailored to the data's structure and distribution.

Figure E.13: ***Impact of Political Narratives on Virality - Extensive and Intensive Margins***

(a) Extensive Margin **(b)** Intensive Margin



Notes: The figure shows the coefficients of OLS regressions testing the effect of featuring at least one character-role, compared to featuring characters only in a neutral role, on virality, measured as the count of retweets, and likes as a robustness check. In Figure E.13a the dependent variable is first transformed to have $Y = 1$ if $Y > 0$, then a Linear Probability Model is applied. In Figure E.13b we drop $Y = 0$ cases and then apply a log transformation. The regression models include relevant tweets, defined as those featuring at least one character from our list. We include only character-roles that appear at least 100 times, thus excluding 'US Republicans-victim', 'Emission Pricing-victim', 'Regulations-victim', and 'Green Tech-victim'. The x-axis reports coefficient estimates along with the corresponding percentage change rounded to the closest unit and computed as follows: $\approx e^{\beta} - 1$. Panels display results from increasingly restrictive models. The baseline model includes only the indicator variable for containing a character-role and clusters standard errors at the week level. The second model adds hour and year-state fixed effects. The third model accounts for author characteristics (verified status, number of followers and followings, total tweets created, party affiliation as Democrat or Republican, religiosity, higher education, and parenthood status). The fourth model adds character fixed effects. The fifth model controls for tweet characteristics (number of hashtags and mentions). The sixth and seventh models include, respectively, language metrics and valence/emotions, as used in the descriptive analysis above.

F Additional Output: Experimental Data

F.1 Descriptives for experiments

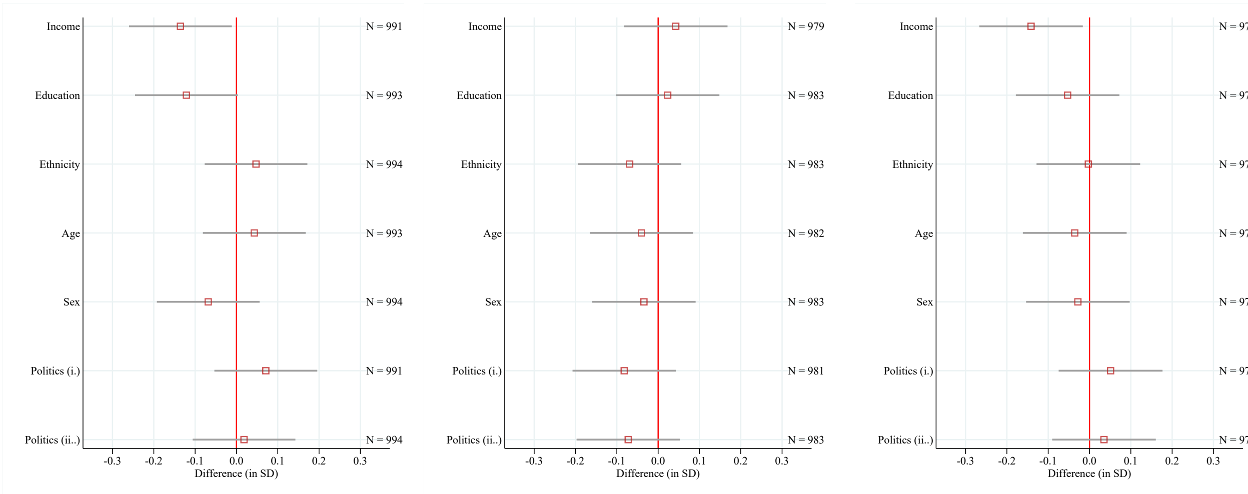
In this section of Appendix F we provide additional information on our online, pre-registered experiments. We conducted three separate studies, each with a representative sample of the US population based on age, ethnicity, gender, and political affiliation. Within each experiment, participants were randomly assigned to either a control or a treatment group. Both groups were told that the survey aimed to study the impact of exposure to social media, and both were shown a fictitious feed of three posts resembling Twitter or Facebook. Two posts were identical across groups and served as obfuscation. The third post differed: in the control group it conveyed a piece of factual information and depicted characters in a neutral way, while in the treatment group it presented the same information but framed the characters within a drama-triangle narrative.

In each experiment, the control–treatment pair of tweets featured different characters, with the exception of Green Tech, which appeared in all three designs. We label the experiments according to the characters included in the narrative treatment. The first is the Hero–Hero experiment, featuring Green Tech-Hero and US People-Hero. The second is the Hero–Hero–Villain experiment, with Green Tech-Hero, Regulations-Hero, and Fossil Industry-Villain. The third is the Villain–Villain–Hero experiment, featuring Green Tech-Hero, Corporations-Villain, and Fossil Industry-Villain. The day after exposure, participants were invited to complete a short follow-up survey, identical across experiments, to assess recall and other outcomes. In the main sections of the paper we provide further details on design choices; below we also provide links to all experimental questionnaires for transparency.

- Main Experiment - Hero-hero: Click here
- Main Experiment - Hero-hero-villain: Click here
- Main Experiment - Villain-villain-hero: Click here
- Follow-up Experiment (all conditions): Click here

At the core of our design is the assumption that randomization was successful, ensuring that the control and treatment groups in all three experiments are comparable on average. Figure F.1 presents balance tests for a range of participant characteristics. We obtain the coefficients by regressing each individual feature on an indicator for assignment to the treatment group. With only two exceptions, all tests confirm balance between treatment and control. Income appears slightly lower in the treatment group for both the Hero–Hero experiment and the Villain–Villain–Hero experiment. In principle, our main analyses already control for this set of personal characteristics, but for completeness we also report results without controls later in this appendix.

Figure F.1: ***Balance of Individual Characteristics by Experiment Wave***



(a) Hero-hero **(b)** Hero-hero-villain **(c)** Villain-villain-hero

Notes: The Figure shows control-treatment balance tests for several individual characteristics, among the participants of each experiment. Panel (a) shows result for the experiment Hero-Hero, Panel (b) shows result for the experiment Hero-Hero-Villain, and Panel (c) for Villain-Villain-Hero. Tests are performed by regressing each individual characteristics on a dummy variable that takes value 1 if the participant is in the treatment group. We use robust standard errors.

F.2 Additional Details on Experimental Output

In this section we provide additional details on the experimental output that we describe in the paper. In particular we provide the underlying models for each coefficient plot in Section 6. Table F.1 provides the models for the Paper Figure 10a, Table F.2 for the Paper Figure 10b, Table F.3 about the donation results in Figure 11. Table F.4 and Table F.5 provide additional details on our memory results, from the Paper Figure 12b and Figure 12c.

Table F.1: ***Experimental Results - The Impact of Political Narratives on Beliefs***

Dependent Variable	**Expectation for the Future:**					
	Forecast	**Confidence**	**Forecast**	**Confidence**	**Forecast**	**Confidence**
	Coeff./SE/p-value					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	2.487	3.426	0.831	0.331	-5.022	-0.358
	(1.368)	(1.680)	(1.425)	(1.740)	(1.387)	(1.803)
	[0.069]	[0.042]	[0.560]	[0.849]	[0.000]	[0.843]
Controls	✓	✓	✓	✓	✓	✓
Experiment: Hero-Hero	✓	✓				
Experiment: Hero-Hero-Villain			✓	✓		
Experiment: Villain-Villain-Hero					✓	✓
Mean Outcome Control Group	39.02	46.02	43.67	48.54	42.32	48.15
Observations	987	987	976	976	968	968

Notes: The table displays the coefficients of OLS regression models providing insights on two outcomes: first, the participants' forecast, as answer to the question *'What percentage of US energy do you predict will come from renewable sources and green technology by the year 2035? Indicate a number between 0 and 100.'* (in Columns 1, 3, 5), second, their confidence in the forecast, as answer to *'Your response on the previous screen suggests that by 2035, [x]% of US energy will come from renewable sources and green technology. How certain are you that the actual share of renewable energy in 2035 will be between [x-5] and [x+5]%?'* (in Columns 2, 4, 6). Columns 1 and 2 show results for the Hero-Hero experiment, columns 3 and 4 for the Hero-Hero-Villain experiment, and 5 and 6 for the Villain-Villain-Hero experiment. All models include income, education, political preference, age, and sex as controls. We use robust standard errors. The Paper Figure 10a is the reference plot.

Table F.2: ***Experimental Results - The Impact of Political Narratives on Stated Preferences***

Dependent Variable	**Stated Preferences: Policy Support**		
	Coeff./SE/p-value		
	(1)	(2)	(3)
Treatment	0.061	0.054	0.094
	(0.090)	(0.086)	(0.117)
	[0.495]	[0.533]	[0.421]
Controls	✓	✓	✓
Experiment: Hero-Hero	✓		
Experiment: Hero-Hero-Villain		✓	
Experiment: Villain-Villain-Hero			✓
Mean Outcome Control Group	5.70	5.72	4.19
Observations	987	976	968

Notes: The table displays the coefficients of OLS regression models providing insights on the support or opposition for a policy or law that is in line with the content of the narrative tweet in our experiments. In Column 1, we show results for the Hero-Hero experiment, where participants were asked whether they would support a policy that reduces the cost of residential renewable systems. In Column 2, we show results for the Hero-Hero-Villain experiment, where participants were asked whether they would support increasing transparency and accountability for energy companies. In column 3, we show results for the Villain-Villain-Hero experiment, where people were asked whether they would support raising taxes on fossil fuels. All models include income, education, political preference, age, and sex as controls. We use robust standard errors. The Paper Figure 10b is the reference plot.

Table F.3: *Experimental Results - Impact of Political Narratives on Revealed Preferences*

Dependent Variable	**Revealed Preference: Incentivized Donation**			
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Treatment	0.734	0.530	0.584	0.562
	(0.497)	(0.469)	(0.480)	(0.272)
	[0.140]	[0.259]	[0.224]	[0.039]
Controls	✓	✓	✓	✓
Experiment: Hero-Hero	✓			✓
Experiment: Hero-Hero-Villain		✓		✓
Experiment: Villain-Villain-Hero			✓	✓
Experiment FE				✓
Mean Outcome Control Group	6.52	6.40	6.73	6.55
Observations	987	976	968	2931

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' revealed preferences. We measure revealed preferences with the decision to donate to an association promoting sustainable development and local/national projects to support Green Tech diffusion. The decision is incentivized with a lottery: participants could be selected to win 25$ and had to allocate the amount between themselves and the association. No restrictions on the allocation were given. Columns 1, 2, and 3 show results for the single experiments, Column 4 shows results for the pooling together all experiments, and it includes experiment fixed effects. All models include income, education, political preference, age, and sex as controls. We use robust standard errors. The Paper Figure 11 is the reference plot.

Table F.4: *Experimental Results - The Impact of Political Narratives on Memory (Pooled Sample)*

Dependent Variable	**Information Retention:**			
	Facts		**Character**	
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Treatment	0.001	-0.002	0.061	0.049
	(0.012)	(0.013)	(0.016)	(0.021)
	[0.957]	[0.901]	[0.000]	[0.018]
Controls	✓	✓	✓	✓
Experiment FEs	✓	✓	✓	✓
Day of the Experiment	✓		✓	
Day After the Experiment		✓		✓
Mean Outcome Control Group	0.13	0.10	0.69	0.45
Observations	2931	2280	2931	2269

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' memory. Columns 1 and 2 show the effect on information retention, respectively in the day of the main experiment and a day later. Participants were asked to remember the factual information reported in the tweet. Column 3 and 4 show the effect on recalling the characters present in the text. Columns 5 and 6 show the effect on recalling the characters framed in their role. The dependent variables for columns from 3 to 6 are obtained encoding an open-ended question that asked participants to recall anything from the text they saw. All models include income, education, political preference, age, and sex as controls, and experiment FEs. We use robust standard errors. Paper Figure 12a and Figure 12b are the reference plots.

Table F.5: *Experimental Results - The Impact of Political Narratives on Memory of Roles*

Dependent Variable	**Information Retention:**				
	Hero	**Hero**	**Villain**	**Hero**	**Villain**
	Coeff./SE/p-value				
	(1)	(2)	(3)	(4)	(5)
Treatment	0.169	-0.106	0.133	-0.139	0.636
	(0.075)	(0.075)	(0.059)	(0.064)	(0.077)
	[0.025]	[0.156]	[0.025]	[0.032]	[0.000]
Controls	✓	✓	✓	✓	✓
Experiment: Hero-Hero	✓				
Experiment: Hero-Hero-Villain		✓	✓		
Experiment: Villain-Villain-Hero				✓	✓
Mean Outcome Control Group	1.35	1.25	0.67	1.01	0.81
Observations	784	792	792	693	693

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' memory of characters divided by role. Column 1 includes only the Hero–Hero experiment and tests recall of Green Tech and US People (both heroes). Columns 2 and 3 include the Hero–Hero–Villain experiment, showing effects on recall of Green Tech and Regulations (heroes, column 2) and Fossil Industry (villain, column 3). Columns 4 and 5 cover the Villain–Villain–Hero experiment, showing effects on recall of Green Tech (hero, column 4) and Fossil Industry/Corporations (villains, column 5). All models include income, education, political preference, age, and sex as controls. We use robust standard errors. Paper Figure 12c is the reference plot.

F.3 The Impact of Valence and Anger

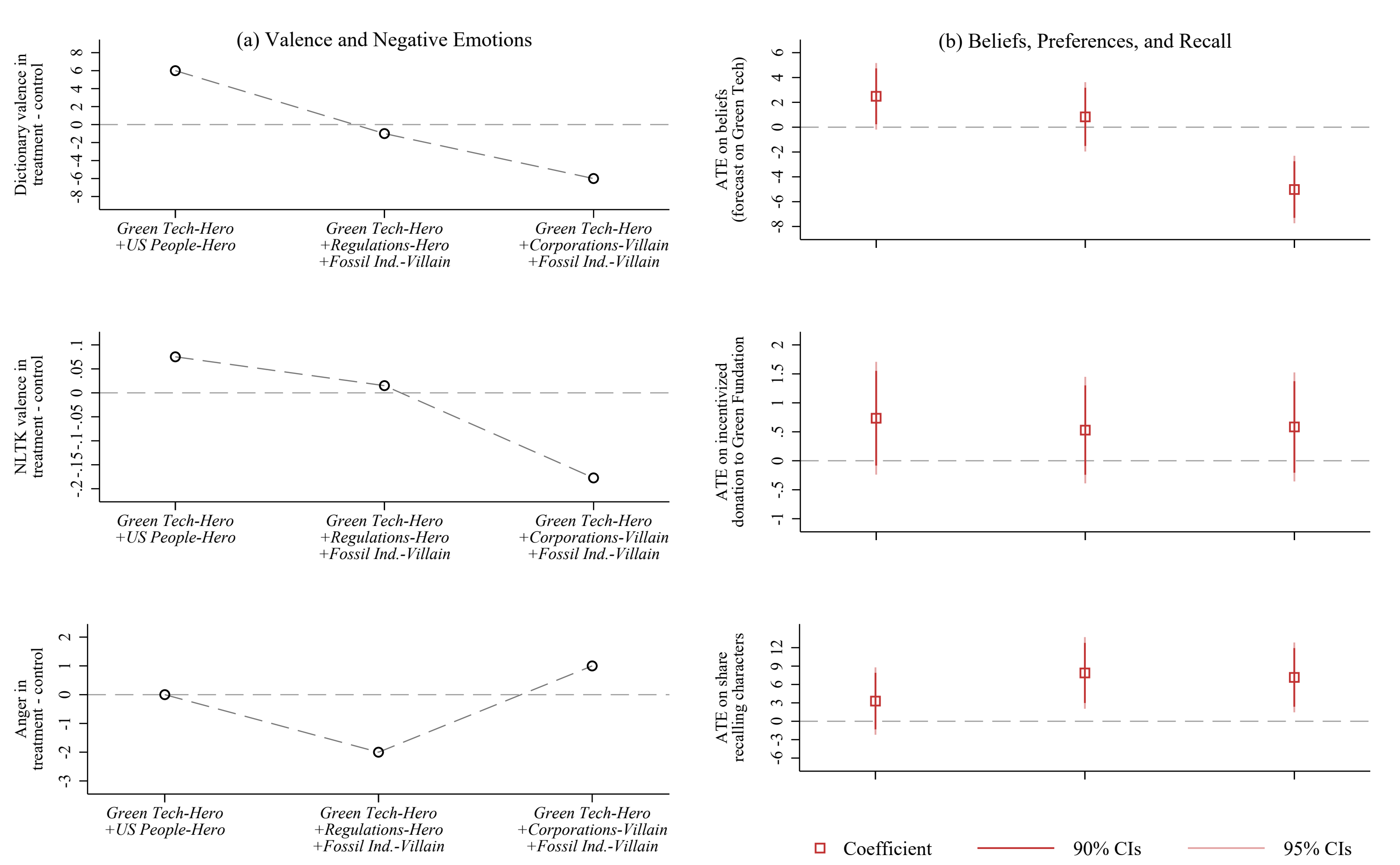
One potential concern with our experiments is that the results we describe in the main paper might be driven primarily by emotions rather than by the use of *political narratives*. The marketing literature in particular highlights the central role of emotions and arousal in determining the virality of content (Berger 2016; Berger 2011). For example, Berger (2011) argues that much of a message's contagiousness can be explained by whether it features emotions, such as disgust, that heighten the receiver's arousal. Guided by this evidence, we explore the correlation between emotions and our experimental results.

While we recognize that emotions are an important ingredient in building and communicating *political narratives*, our analysis throughout the paper shows that they are not sufficient to explain their effect. The roles of villain, hero, and victim capture deeper mental constructs that extend beyond emotional tone. To test this directly, recall that participants in our experiments were shown a social media feed containing tweets on climate change. In the control group, a tweet conveyed factual information with characters portrayed neutrally; in the treatment group, the same information was presented but with the characters framed in drama triangle roles. If emotions alone captured the essence of *political narratives*, then the emotional content of the tweets should fully explain the differences in outcomes we observe. Building on this idea, we test whether differences in valence – defined as the balance of positive emotions minus negative ones – between treatment and control tweets account for the direction and significance of our results.

Figure F.2 sheds light on this issue. The left side of the figure reports the difference in valence and anger between the treatment and control tweets of each experiment. In the top panel, we compute valence as the number of positive emotion words (joy, trust, anticipation, and surprise) minus the number of negative emotion words (anger, disgust, fear, and sadness), following the NRC dictionary (Mohammad and Turney 2013). We then calculate, for each experiment, the difference between the treatment tweet's valence and that of the control tweet. A positive value indicates that the treatment tweet carried a more positive tone, while a negative value signals greater negativity. The middle panel repeats this exercise using the NLTK sentiment package as an alternative measure of valence. Finally, the bottom panel focuses specifically on anger, showing the difference in the frequency of anger-related words between treatment and control. The right side of Figure F.2 provides a condensed summary of our experimental results. The top panel shows the effect of the narrative treatment on beliefs, measured by participants' forecasts of the future share of energy generated from green technologies. The middle panel presents the impact on revealed preferences, captured through donations to green technology institutions. The bottom panel reports our main results on memory, measured as the likelihood of recalling characters from the narratives. For a detailed discussion of these outcomes, we refer back to the main body of the paper.

Despite being purely correlational and descriptive, the figure delivers a clear message: there is no consistent link between the emotional tone of the treatment tweets and the experimental outcomes. For beliefs, the relationship with net valence appears plausible: tweets with a more negative tone seem to elicit a more pessimistic outlook on the future. However, for both donations and memory, the narrative treatment has a positive effect regardless of whether the tweet's valence is positive or negative. Finally, when focusing specifically on anger – the emotion most likely to heighten arousal – there is virtually no correlation with the experimental results. These findings should be interpreted with caution, but they suggest an important takeaway: while emotions, valence, and arousal certainly play a role, they cannot by themselves explain the effects we observe. Narratives structured around hero, villain, and victim roles capture a deeper mechanism of influence, one that goes beyond emotional tone alone.

Figure F.2: *Emotional Content in the Experimental Design*



Notes: Figure F.2 panel (a) displays the difference in the corresponding emotion score between the treatment tweet and the control tweet, for each of the three experimental designs (Hero-Hero, Hero-Hero-Villain, Hero-Villain-Villain). The emotion scores used are valence and anger. Valence is measured either as the difference in the number of positive words and negative words (top) or as a score using the NLTK Python package (middle). Anger is computed as a score using the NLTK Python package (bottom). Emotion and valence use the NRC Emotion Lexicon (Mohammad and Turney (2013)). Each graph of panel (a) represents the difference in the score between the treatment tweet and the control tweet. Figure F.2 panel (b) replicates the main results of the paper. Reference figures are Figure 10a, Figure 11, Figure 12c.

G Robustness Checks: Experimental Data

G.1 Recall of the Factual Information

One important insight from our experimental results concerns the nature of memory recall. On the one hand, participants in the treatment groups remembered the characters featured in the tweets much more – both on the day of the experiment and the day after. On the other hand, we find no significant difference between treatment and control groups in the recall of factual information. In the paper, factual recall is measured using a strict encoding: participants were asked to reproduce exactly the factual information reported in the tweets. A potential concern is that the absence of differences could simply reflect the narrowness of this measure. To address this, in this section we provide a robustness test. Even when we relax the encoding and accept answers within broader intervals around the correct value, no detectable treatment–control difference emerges.

Table G.1 and Table G.2 report robustness checks on the recall of factual information, measured on the day of the experiment and the day after, respectively. In both tables, Column 1 reproduces the baseline specification from the paper, while Columns 2, 3, and 4 progressively relax the coding of correct answers by accepting responses within ±10, ±20, and ±50 units of the true value. The results are clear: across all specifications, there is no statistically significant difference in factual recall between treatment and control groups.

Table G.1: ***Experimental Results - Different Encoding of Factual Recall on the Experiment Day (Pooled)***

Dependent Variable	**Interval of Recall: 163 +|-**			
	0	**10**	**20**	**50**
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Treatment	0.001	0.003	-0.005	-0.012
	(0.012)	(0.015)	(0.016)	(0.018)
	[0.957]	[0.845]	[0.777]	[0.499]
Mean Outcome Control Group	0.12	0.20	0.23	0.35
Observations	2931	2931	2931	2931

Notes: The table reports OLS regression results on the effect of narratives on participants' memory of the factual information embedded in both control and treatment tweets, measured on the day of the experiment. The dependent variable comes from the question: "How many billion kWhs were generated using solar energy in 2023 in the US? Please indicate your best guess." Column 1 codes the outcome as 1 if the answer is exactly correct; Column 2 as 1 if the answer falls within a 10-unit range; Column 3 within 20 units; and Column 4 within 50 units. All models include observations from all three experiments and experiment fixed effects. Additionally, all models control for income, education, political preference, age, and sex, and use robust standard errors. Paper Figure 12a is the reference plot.

Table G.2: *Experimental Results - Different Encoding of Factual Recall on the Follow-Up Day (Pooled)*

Dependent Variable	**Interval of Recall: 163 +|-**			
	0	**10**	**20**	**50**
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Treatment	-0.002	0.005	-0.006	-0.016
	(0.013)	(0.016)	(0.018)	(0.020)
	[0.901]	[0.756]	[0.748]	[0.436]
Mean Outcome Control Group	0.11	0.19	0.22	0.34
Observations	2280	2280	2280	2280

Notes: The table reports OLS regression results on the effect of narratives on participants' memory of the factual information embedded in both control and treatment tweets, measured on the day after of the experiment. The dependent variable comes from the question: "How many billion kWhs were generated using solar energy in 2023 in the US? Please indicate your best guess." Column 1 codes the outcome as 1 if the answer is exactly correct; Column 2 as 1 if the answer falls within a 10-unit range; Column 3 within 20 units; and Column 4 within 50 units. All models include observations from all three experiments and experiment fixed effects. Additionally, all models control for income, education, political preference, age, and sex, and use robust standard errors. Paper Figure 12a is the reference plot.

G.2 Output Excluding Controls

In this section we present robustness tests for the experimental results. In particular we reproduce all the results of the paper, exlcuding individual characteristics as controls in the models.

Table G.3: ***Manipulation Check - Effectiveness of the Political Narrative Treatment - Excluding Controls***

Dependent Variable	**Mention of Character-Role**							
	Coeff./SE/p-value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.296	0.276	0.237	0.270	0.359	0.330	0.254	0.315
	(0.030)	(0.031)	(0.031)	(0.018)	(0.033)	(0.034)	(0.031)	(0.019)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Outcome | Recall 1+ Characters					✓	✓	✓	✓
Experiment: Hero-Hero	✓			✓	✓			✓
Experiment: Hero-Hero-Villain		✓		✓		✓		✓
Experiment: Villain-Villain-Hero			✓	✓			✓	✓
Experiment FE				✓				✓
Mean Outcome Control Group	0.33	0.34	0.44	0.37	0.45	0.52	0.64	0.54
Observations	994	983	978	2955	744	690	702	2136

Notes: The table reports OLS estimates from manipulation checks of the narrative treatment. The dependent variable is a binary indicator equal to one if the participant recalled the role assigned to a character in the treatment tweet (e.g., Green Tech as hero in the Hero–Hero experiment, Fossil Industry as villain in the Hero–Hero–Villain experiment). Columns 1–4 report effects on the unconditional likelihood of recalling the character-role, while Columns 5–8 restrict the sample to participants who recalled at least one character from the tweet. Only the treatment group was exposed to characters explicitly framed in roles; control participants saw the same characters presented neutrally. Non-zero recall in the control group therefore reflects participants' prior beliefs or participants' interpretation of the control tweet, while the treatment effect captures the additional role attribution induced by framing. We use robust standard errors. Paper Table 4 is the reference table.

Table G.4: ***Experimental Results - The Impact of Political Narratives on Beliefs - Excluding Controls***

Dependent Variable	**Expectation for the Future:**					
	Forecast	**Confidence**	**Forecast**	**Confidence**	**Forecast**	**Confidence**
	Coeff./SE/p-value					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	1.838	3.018	0.755	0.037	-4.995	-0.856
	(1.321)	(1.694)	(1.360)	(1.716)	(1.311)	(1.717)
	[0.165]	[0.075]	[0.579]	[0.983]	[0.000]	[0.618]
Experiment: Hero-Hero	✓	✓				
Experiment: Hero-Hero-Villain			✓	✓		
Experiment: Villain-Villain-Hero					✓	✓
Mean Outcome Control Group	39.02	46.02	43.67	48.54	42.32	48.15
Observations	994	994	983	983	978	978

Notes: The table displays the coefficients of OLS regression models providing insights on two outcomes: the participants' forecast, as answer to the question *'What percentage of US energy do you predict will come from renewable sources and green technology by the year 2035? Indicate a number between 0 and 100.'* (in Columns 1, 3, 5), their confidence in the forecast, as answer to *'Your response on the previous screen suggests that by 2035, [x]% of US energy will come from renewable sources and green technology. How certain are you that the actual share of renewable energy in 2035 will be between [x-5] and [x+5]%?'* (in Columns 2, 4, 6). Columns 1 and 2 show results for the Hero-Hero experiment, columns 3 and 4 for the Hero-Hero-Villain experiment, and 5 and 6 for the Villain-Villain-Hero experiment. The models do not include personal characteristics as controls. We use robust standard errors. The Paper Figure 10a is the reference plot.

Table G.5: ***Experimental Results - The Impact of Political Narratives on Stated Preferences - Excluding Controls***

Dependent Variable	**Stated Preferences: Policy Support**		
	Coeff./SE/p-value		
	(1)	(2)	(3)
Treatment	0.040	0.052	-0.011
	(0.092)	(0.092)	(0.126)
	[0.664]	[0.568]	[0.930]
Experiment: Hero-Hero	✓		
Experiment: Hero-Hero-Villain		✓	
Experiment: Villain-Villain-Hero			✓
Mean Outcome Control Group	5.70	5.72	4.19
Observations	994	983	978

Notes: The table displays the coefficients of OLS regression models providing insights on the support or opposition for a policy or law that is in line with the content of the narrative. In Column 1, we show results for the Hero-Hero experiment, where participants were asked whether they would support a policy that reduces the cost of residential renewable systems. In Column 2 we show results for the Hero-Hero-Villain experiment, where participants were asked whether they would support increasing transparency and accountability for energy companies. In Column 3, we show results for the Villain-Villain-Hero experiment, where people were asked whether they would support raising taxes on fossil fuels. The models do not include personal characteristics as controls. We use robust standard errors. The Paper Figure 10b is the reference plot.

Table G.6: ***Experimental Results - Impact of Political Narratives on Revealed Preferences - Excluding Controls***

Dependent Variable	**Revealed Preference: Incentivized Donation**			
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Treatment	0.540	0.570	0.311	0.474
	(0.477)	(0.458)	(0.469)	(0.270)
	[0.257]	[0.214]	[0.507]	[0.079]
Experiment: Hero-Hero	✓			✓
Experiment: Hero-Hero-Villain		✓		✓
Experiment: Villain-Villain-Hero			✓	✓
Experiment FE				✓
Mean Outcome Control Group	6.52	6.40	6.73	6.55
Observations	994	983	978	2955

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' revealed preferences. We measure revealed preferences with the decision to donate to an association promoting sustainable development and local/national projects to support Green tech diffusion. The decision is incentivized with a lottery: participants could be selected to win 25$ and had to allocate the amount between themselves and the association. No restrictions on the allocation were given. Columns 1, 2, and 3 show results for the single experiments, Column 4 shows results for the pooling together all experiments, and it includes experiment fixed effects. The models do not include personal characteristics as controls. We use robust standard errors. The Paper Figure 11 is the reference plot.

Table G.7: *Experimental Results - The Impact of Political Narratives on Memory (Pooled Sample) - Excluding Controls*

Dependent Variable	**Information Retention:**			
	Facts		**Character**	
	Coeff./SE/p-value			
	(1)	(2)	(3)	(4)
Treatment	0.001	-0.002	0.059	0.053
	(0.012)	(0.012)	(0.016)	(0.021)
	[0.917]	[0.847]	[0.000]	[0.011]
Experiment FEs	✓	✓	✓	✓
Day of the Experiment	✓		✓	
Day After the Experiment		✓		✓
Mean Outcome Control Group	0.13	0.13	0.69	0.45
Observations	2955	2297	2955	2286

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' memory. Columns 1 and 2 show the effect on information retention, respectively in the day of the main experiment and a day later. Participants were asked to remember the factual information reported in the tweet. Column 3 and 4 show the effect on recalling the characters present in the text, respectively in the day of the main experiment and a day later. All models include experiment FEs. We use robust standard errors. Paper Figure 12a and Figure 12b are the reference plots.

Table G.8: *Experimental Results - The Impact of Political Narratives on Memory of Roles - Excluding Controls*

Dependent Variable	Information Retention:				
	Hero	Hero	Villain	Hero	Villain
	Coeff./SE/p-value				
	(1)	(2)	(3)	(4)	(5)
Treatment	0.194	-0.126	0.092	-0.157	0.620
	(0.073)	(0.072)	(0.058)	(0.062)	(0.077)
	[0.008]	[0.082]	[0.109]	[0.012]	[0.000]
Controls	✓	✓	✓	✓	✓
Experiment: Hero-Hero	✓				
Experiment: Hero-Hero-Villain		✓	✓		
Experiment: Villain-Villain-Hero				✓	✓
Mean Outcome Control Group	1.35	1.25	0.67	1.01	0.81
Observations	789	798	798	699	699

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' memory of characters divided by role. Column 1 includes only the Hero–Hero experiment and tests recall of Green Tech and US People (both heroes). Columns 2 and 3 include the Hero–Hero–Villain experiment, showing effects on recall of Green Tech and Regulations (heroes, column 2) and Fossil Industry (villain, column 3). Columns 4 and 5 cover the Villain–Villain–Hero experiment, showing effects on recall of Green Tech (hero, column 4) and Fossil Industry/Corporations (villains, column 5). We use robust standard errors. Paper Figure 12c is the reference plot.

G.3 Output with Randomization Inference

In this section we present robustness tests for the experimental results. In particular we reproduce all the results of the paper, computing the p-values via randomization inference, using the *ritest* Stata package.

Table G.9: *Manipulation Check - Effectiveness of the Political Narrative Treatment - Randomized Inference*

Dependent Variable	**Mention of Character-Role**							
	Coeff./SE/ *Randomized (n=1000)* p-value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.293	0.281	0.254	0.271	0.358	0.322	0.280	0.316
	(0.032)	(0.032)	(0.031)	(0.018)	(0.034)	(0.035)	(0.032)	(0.019)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Outcome | Character					✓	✓	✓	✓
Experiment: Hero-Hero	✓			✓	✓			✓
Experiment: Hero-Hero-Villain		✓		✓		✓		✓
Experiment: Villain-Villain-Hero			✓	✓			✓	✓
Experiment FE				✓				✓
Mean Outcome Control Group	0.33	0.34	0.44	0.37	0.45	0.52	0.64	0.54
Observations	987	976	968	2931	742	686	695	2123

Notes: The table reports OLS estimates from manipulation checks of the narrative treatment. The dependent variable is a binary indicator equal to one if the participant recalled the role assigned to a character in the treatment tweet (e.g., Green Tech as hero in the Hero–Hero experiment, Fossil Industry as villain in the Hero–Hero–Villain experiment). Columns 1–4 report effects on the unconditional likelihood of recalling the character-role, while Columns 5–8 restrict the sample to participants who recalled at least one character from the tweet. Only the treatment group was exposed to characters explicitly framed in roles; control participants saw the same characters presented neutrally. Non-zero recall in the control group therefore reflects participants' prior beliefs or participants' interpretation of the control tweet, while the treatment effect captures the additional role attribution induced by framing. All regressions control for income, education, political preference, age, and sex. Standard errors are clustered at the individual level. We compute the p-values using the STATA command *ritest* for randomized inference, using 1000 repetitions, with *strict* and *two-sided* specification. Paper Table 4 is the reference table.

Table G.10: *Experimental Results - The Impact of Political Narratives on Beliefs - Randomization Inference*

Dependent Variable	**Expectation for the Future:**					
	Forecast	**Confidence**	**Forecast**	**Confidence**	**Forecast**	**Confidence**
	Coeff./SE/*Randomized (n=1000)* p-value					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	2.487	3.426	0.831	0.331	-5.022	-0.358
	(1.368)	(1.680)	(1.425)	(1.740)	(1.387)	(1.803)
	[0.077]	[0.051]	[0.546]	[0.860]	[0.000]	[0.835]
Controls	✓	✓	✓	✓	✓	✓
Experiment: Hero-Hero	✓	✓				
Experiment: Hero-Hero-Villain			✓	✓		
Experiment: Villain-Villain-Hero					✓	✓
Mean Outcome Control Group	39.02	46.02	43.67	48.54	42.32	48.15
Observations	987	987	976	976	968	968

Notes: The table displays the coefficients of OLS regression models providing insights on two outcomes: the participants' forecast, as answer to the question *'What percentage of US energy do you predict will come from renewable sources and green technology by the year 2035? Indicate a number between 0 and 100.'* (in Columns 1, 3, 5), their confidence in the forecast, as answer to *'Your response on the previous screen suggests that by 2035, [x]% of US energy will come from renewable sources and green technology. How certain are you that the actual share of renewable energy in 2035 will be between [x-5] and [x+5]%?'* (in Columns 2, 4, 6). Columns 1 and 2 show results for the Hero-hero experiment, columns 3 and 4 for the Hero-hero-villain experiment, and 5 and 6 for the Villain-villain-hero experiment. All models include income, education, political preference, age, and sex as controls. We compute the p-value using the STATA command *ritest* for randomized inference, using 1000 repetitions, with *strict* and *two-sided* specification. The Paper Figure 10a is the reference plot.

Table G.11: *Experimental Results - The Impact of Political Narratives on Stated Preferences - Randomized Inference*

Dependent Variable	**Stated Preferences: Policy Support**		
	Coeff./SE/ *Randomized (n=1000)* p-value		
	(1)	(2)	(3)
Treatment	0.061	0.054	0.094
	(0.090)	(0.086)	(0.117)
	[0.501]	[0.558]	[0.474]
Controls	✓	✓	✓
Experiment: Hero-Hero	✓		
Experiment: Hero-Hero-Villain		✓	
Experiment: Villain-Villain-Hero			✓
Mean Outcome Control Group	5.70	5.72	4.19
Observations	987	976	968

Notes: The table displays the coefficients of OLS regression models providing insights on the support or opposition for a policy or law that is in line with the content of the narrative. In column 1, we show results for the Hero-Hero experiment, where participants were asked whether they would support a policy that reduces the cost of residential renewable systems. In Column 2 we show results for the Hero-Hero-Villain experiment, where participants were asked whether they would support increasing transparency and accountability for energy companies. In Column 3, we show results for the Villain-Villain-Hero experiment, where people were asked whether they would support raising taxes on fossil fuels. All models include income, education, political preference, age, and sex as controls. We compute the p-value using the STATA command *ritest* for randomized inference, using 1000 repetitions, with *strict* and *two-sided* specification. The Paper Figure 10b is the reference plot.

Table G.12: *Experimental Results - Impact of Political Narratives on Revealed Preferences - Randomized Inference*

Dependent Variable	**Revealed Preference: Incentivized Donation**			
	Coeff./SE/*Randomized (n=1000)* p-value			
	(1)	(2)	(3)	(4)
Treatment	0.734	0.530	0.584	0.562
	(0.497)	(0.469)	(0.480)	(0.272)
	[0.150]	[0.263]	[0.216]	[0.045]
Controls	✓	✓	✓	✓
Experiment: Hero-Hero	✓			✓
Experiment: Hero-Hero-Villain		✓		✓
Experiment: Villain-Villain-Hero			✓	✓
Experiment FE				✓
Mean Outcome Control Group	6.52	6.40	6.73	6.55
Observations	987	976	968	2931

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' revealed preferences. We measure revealed preferences with the decision to donate to an association promoting sustainable development and local/national projects to support Green tech diffusion. The decision is incentivized with a lottery: participants could be selected to win 25$ and had to allocate the amount between themselves and the association. No restrictions on the allocation were given. Columns 1, 2, and 3 show results for the single experiments, Column 4 shows results for the pooling together all experiments, and it includes experiment fixed effects. All models include income, education, political preference, age, and sex as controls. We compute the p-value using the STATA command *ritest* for randomized inference, using 1000 repetitions, with *strict* and *two-sided* specification. The Paper Figure 11 is the reference plot.

Table G.13: *Experimental Results - The Impact of Political Narratives on Memory (Pooled Sample) - Randomized Inference*

Dependent Variable	**Information Retention:**			
	Facts		**Character**	
	Coeff./SE/*Randomized (n=1000)* p-value			
	(1)	(2)	(3)	(4)
Treatment	0.001	-0.002	0.061	0.049
	(0.012)	(0.013)	(0.016)	(0.021)
	[0.969]	[0.906]	[0.000]	[0.014]
Controls	✓	✓	✓	✓
Experiment FEs	✓	✓	✓	✓
Day of the Experiment	✓		✓	
Day After the Experiment		✓		✓
Mean Outcome Control Group	0.13	0.10	0.69	0.45
Observations	2931	2280	2931	2269

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' memory. Columns 1 and 2 show the effect on information retention, respectively in the day of the main experiment and a day later. Participants were asked to remember the factual information reported in the tweet. Column 3 and 4 show the effect on recalling the characters present in the text. All regressions include experimental FEs and the following controls: income, income, education, political preference, age, and sex. We compute the p-value using the STATA command *ritest* for randomized inference, using 1000 repetitions, with *strict* and *two-sided* specification. Paper Figure 12a and Figure 12b are the reference plots.

Table G.14: *Experimental Results - The Impact of Political Narratives on Memory of Roles - Randomized Inference*

Dependent Variable	Information Retention:				
	Hero	Hero	Villain	Hero	Villain
	Coeff./SE/p-value				
	(1)	(2)	(3)	(4)	(5)
Treatment	0.169	-0.106	0.133	-0.139	0.636
	(0.075)	(0.075)	(0.059)	(0.064)	(0.077)
	[0.025]	[0.156]	[0.025]	[0.032]	[0.000]
Controls	✓	✓	✓	✓	✓
Experiment: Hero-Hero	✓				
Experiment: Hero-Hero-Villain		✓	✓		
Experiment: Villain-Villain-Hero				✓	✓
Mean Outcome Control Group	1.35	1.25	0.67	1.01	0.81
Observations	784	792	792	693	693

Notes: The table displays the results from OLS regression models analyzing the impact of the narratives on participants' memory of characters divided by role. Column 1 includes only the Hero–Hero experiment and tests recall of Green Tech and US People (both heroes). Columns 2 and 3 include the Hero–Hero–Villain experiment, showing effects on recall of Green Tech and Regulations (heroes, column 2) and Fossil Industry (villain, column 3). Columns 4 and 5 cover the Villain–Villain–Hero experiment, showing effects on recall of Green Tech (hero, column 4) and Fossil Industry/Corporations (villains, column 5). All models include income, education, political preference, age, and sex as controls. We compute the p-value using the STATA command *ritest* for randomized inference, using 1000 repetitions, with *strict* and *two-sided* specification. Paper Figure 12c is the reference plot.

H Political Narratives in Other Media

The main analysis in this paper investigates the virality of narratives on social media, focusing on Twitter/X. While these results provide detailed insights into the determinants of virality in the context of social media engagement, it is important to consider whether similar patterns in the use and distribution of narratives extend beyond the online environment. This appendix addresses this question by exploring the presence of *political narratives* – as defined in this study – in traditional media sources, specifically newspapers and television.

This exercise serves two purposes. First, it provides external validation by examining whether the types of narratives that drive engagement on social media are also present in other influential media platforms. Second, it assesses the potential scope extension of our findings, recognizing that while virality is a unique feature of social media, the content and prevalence of narratives may reflect broader patterns in public discourse. We present a simple descriptive analysis of narrative distribution in newspapers and television during the same period covered by our main study. Although we cannot reproduce the virality results in these settings, identifying similar narrative patterns across media contributes to a more comprehensive understanding of the role narratives play in shaping political communication.

H.1 Newspapers

In preparing and classifying the newspaper articles we follow as close as possible the procedure to prepare the tweets. We source a random and representative set of newspaper articles. To that end, we use the three most widely circulate newspapers in the US; The New York Times, The Wall Street Journal, and USA Today. We download the articles from Factiva. For each newspaper, we download 3000 articles for the period between 2010 and 2021. We download the articles in 4 year intervals. To ensure a balanced distribution of popular articles over time for each newspaper, we source the 333 and 334 most popular articles from 2010-2013, then 2014-2017, and 2018-2021. We use exactly the same list of keywords (Oehl, Schaffer, and Bernauer 2017)(see Section A.1 for the full list of keywords). We minimally adapt the prompts to classify the tweets, changing the references from tweets to newspaper articles. The adapted prompts can be found attached.

We find that the distribution of articles represents the overall distribution of articles from Twitter extremely well. The three most often recurring character-roles for human characters are US democrats heroes with 7.52%, corporations-villains with 6.33%, and US republicans with 6.14%. Similarly, the three most often recurring character-roles for instrument characters are US fossil industry-villain with 14.98%, green tech-hero with 9.92%, and regulations-hero with about 6%.

This resembles the ranking for instrument characters.

Table H.1: ***Share of Character-Roles in Relevant Newspaper Articles (United States, 2010-2021)***

Panel A: Human Characters

	Hero	Villain	Victim	Neutral	Total
Developing Economies	0.51	0.78	3.16	0.85	5.29
US Democrats	7.52	0.44	0.17	1.67	9.80
US Republicans	0.19	6.14	.	1.92	8.24
Corporations	1.92	6.33	0.17	10.64	19.05
US People	1.44	0.18	1.84	4.00	7.45

Panel B: Instrument Characters

	Hero	Villain	Victim	Neutral	Total
Emission Pricing	3.92	0.94	.	2.15	7.00
Regulations	5.96	1.23	.	3.71	10.90
Fossil Industry	0.08	14.98	0.06	2.18	17.29
Green Tech	9.92	0.36	.	2.80	13.08
Nuclear Tech	0.98	0.21	0.13	0.59	1.91

Notes: The table shows the frequencies of character-roles in the classified newspaper articles as a percentage of the total occurrences of each character. Shares are computed considering only the dataset of relevant snippets used in our analysis. We define a tweet as relevant if it features at least one character from our list. We include in the computation of shares only character roles that appear at least 100 times, thus excluding 'US Republicans-victim', 'Emission Pricing-victim', 'Regulations-victim', and 'Green Tech-victim', indicated by a dot in the tables. Panel (a) displays the shares for characters of the human type, while Panel (b) displays the same for characters of the instrument type. The column Neutral in both panels reports cases where the character is present in the tweet but is not depicted in one of the three specific roles. The occurrence of character-roles is not mutually exclusive, meaning multiple roles may appear in the same tweet. The appendix Table C.6 shows the frequency in absolute numbers, including also the categories excluded here, because not reaching the 100 instances in the time period.

H.2 Television

Table H.2: ***Share of Character-Roles in Relevant TV transcripts (Fox News and MSNBC, 2010-2021)***

Panel A: Human Characters

	Hero	Villain	Victim	Neutral	Total
Developing Economies	0.17	0.22	0.71	0.34	1.44
US Democrats	16.29	1.63	0.00	0.57	18.50
US Republicans	0.10	14.72	0.00	1.03	15.84
Corporations	1.51	7.97	0.00	1.74	11.23
US People	5.87	0.06	2.81	6.19	14.93

Panel B: Instrument Characters

	Hero	Villain	Victim	Neutral	Total
Emissions Pricing	2.01	2.23	0.00	3.09	7.33
Regulations	2.60	3.89	0.00	2.26	8.75
Fossil Industry	0.18	10.85	0.00	1.30	12.33
Green Tech	6.35	0.15	0.00	2.29	8.79
Nuclear Tech	0.20	0.07	0.00	3.42	3.68

Notes: The table shows the frequencies of character-roles in the classified tv transcripts as a percentage of the total occurrences of each character. Shares are computed considering only the dataset of relevant tv transcripts used in our analysis. We define a snippet as relevant if it features at least one character from our list. We include in the computation of shares only character roles that appear at least 100 times, thus excluding 'US Republicans-victim', 'Emission Pricing-victim', 'Regulations-victim', and 'Green Tech-victim', indicated by a dot in the tables. Panel (a) displays the shares for characters of the human type, while Panel (b) displays the same for characters of the instrument type. The column Neutral in both panels reports cases where the character is present in the tweet but is not depicted in one of the three specific roles. The occurrence of character-roles is not mutually exclusive, meaning multiple roles may appear in the same tweet. The appendix Table C.6 shows the frequency in absolute numbers, including also the categories excluded here, because not reaching the 100 instances in the time period.