Supplementary Information for "Socio-demographic and online behavioural drivers of misinformation engagement in Hungary: Evidence from linked survey and social media data"

1. Supplementary Information for Data and Methods

Final analytic sample size: 738

Demographic variable	**Mean**	**SD**
Age (years)	43.68	13.92
Female (%)	60.16%	—
Education		
	N	**Percent (%)**
Primary education	31	4.20
Lower secondary education	97	13.14
Upper secondary education	321	43.50
Higher education	289	39.16
Subjective wealth category		
	N	**Percent (%)**
Deprivation	26	3.52
Hardly getting by	77	10.43
Getting by	260	35.23
Living well	333	45.12
Living very well	42	5.69

Supplementary Table 1: Descriptive statistics of the analytic sample. Continuous variables are summarized using means and standard deviations, while categorical variables are reported as counts and percentages.

1.1. Facebook Interactions (PC1) Variable

For each participant, we aggregated the total number of such interactions across different engagement types (e.g., total page followings, total page reactions, total group memberships). To reduce these multiple correlated measures into a smaller set of summary dimensions, we applied principal component analysis (PCA). When including all interaction variables–pages liked, page reactions, page comments, groups liked, group reactions, group comments–the first principal component explained 39.1% of the total variance. The communalities of these variables with respect to the first principal component indicate that page-related activities (e.g., pages liked, page reactions, page comments) are relatively well represented (0.54–0.62), whereas group-related activities moderately captured (0.16–0.26). To address the low communalities, we iteratively removed variables with the weakest scores (those with communalities below 0.25) and reran the PCA. The final analysis retained pages liked (communality = 0.72), page reactions (0.51), page comments (0.64), and groups liked (0.30). This revised model explained 54.21% of the total variance and captured the dominant axis of variation in users' activity counts, serving as an overall index of engagement intensity across different parts of the platform.

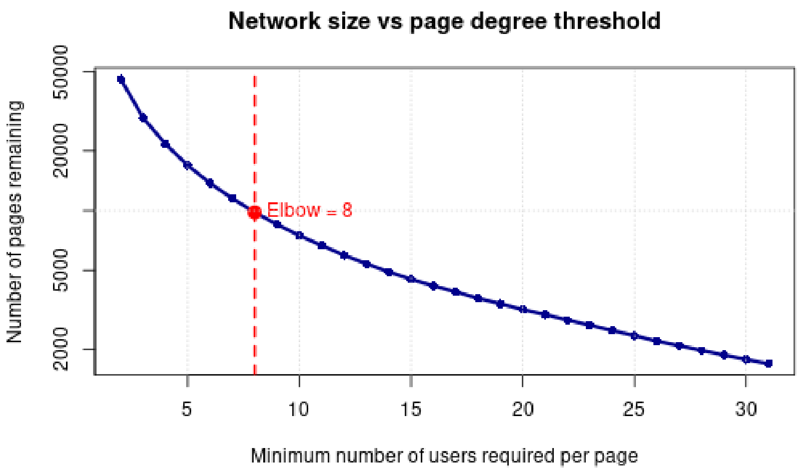
1.2. Facebook Page Network

We analyse patterns of content engagement by quantifying the similarity of page audiences.

To ensure that similarity estimates are based on sufficiently informative interaction patterns, we first restrict the analysis to pages with at least eight unique users. Pages with very small audiences can be dominated by a few highly active users, which may introduce noise into similarity estimates.

To determine an appropriate threshold, we examined how the number of retained pages varies as a function of the minimum number of unique users per page (Supplementary Figure 1). As shown in the figure, the number of pages declines sharply at low thresholds, reflecting the removal of sparsely engaged pages, and then transitions into a more gradual decrease. Using the Kneedle algorithm for automated knee detection [Satopaa et al., 2011], we identified an elbow in this curve at eight users, beyond which increasing the threshold yields diminishing returns in terms of noise reduction while substantially reducing network size. We therefore retain only pages with at least eight unique users, balancing the need to remove sparsely connected pages while preserving sufficient network size.

To quantify similarity between pages, for each page i, we define a user vector $\mathbf{u}_i$ of length U, where U denotes the total number of unique users who interact with pages in the dataset. The n-th element of $\mathbf{u}_i$ corresponds to the number of times user n has engaged with page i. For any pair of pages i and j, we then compute the cosine similarity between $\mathbf{u}_i$ and $\mathbf{u}_j$ and use this cosine similarity to create a network of pages, where an edge between two pages is weighted with the value of their cosine similarity. This approach captures not only the extent of audience overlap (i.e., the number of shared users) but also the relative intensity of their engagement with each page. In doing so, it preserves information that is typically lost when projecting page–user interactions into a lower-dimensional representation, thereby offering a more nuanced view of how audiences cluster around different content sources.

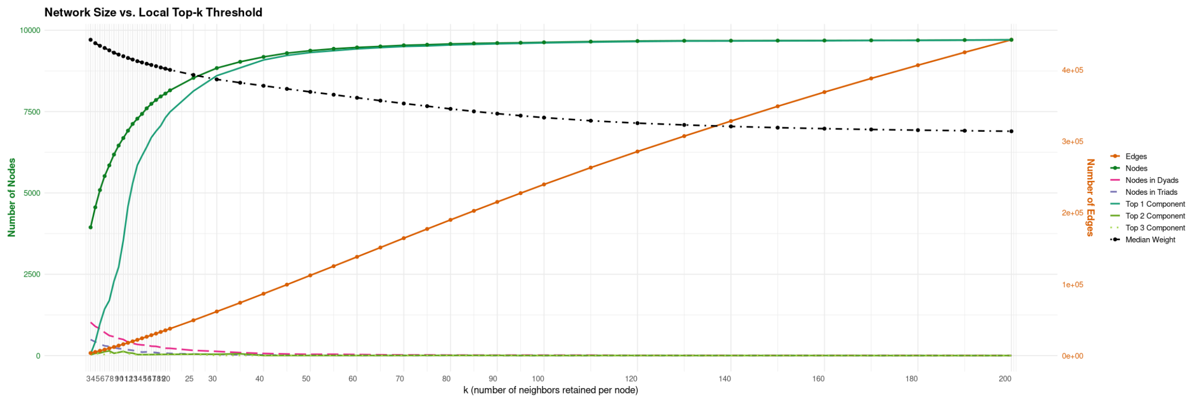


Supplementary Figure 1: Number of retained pages as a function of the minimum number of unique users per page. The curve exhibits a clear elbow at eight users, which we select as a threshold to balance noise reduction and network size.

To construct the page–page network, we adopt a local sparsification approach rather than relying on a single global cosine similarity threshold. Specifically, for each page i, we retain as neighbours only the top k most similar pages based on cosine similarity, ensuring that each node preserves its strongest connections. To obtain an undirected network, we apply an intersection rule, retaining an edge between two pages only when each appears among the other's top k most similar neighbours. In addition, we require that cosine similarity exceeds a minimum threshold of 0.35 to ensure that retained edges reflect meaningful co-engagement rather than incidental overlap.

We set $k = 97$, corresponding to approximately 1% of the total number of pages in the network (after removing disconnected nodes). This choice balances sparsity and connectivity, preserving local neighbourhood structure while avoiding overly dense networks.

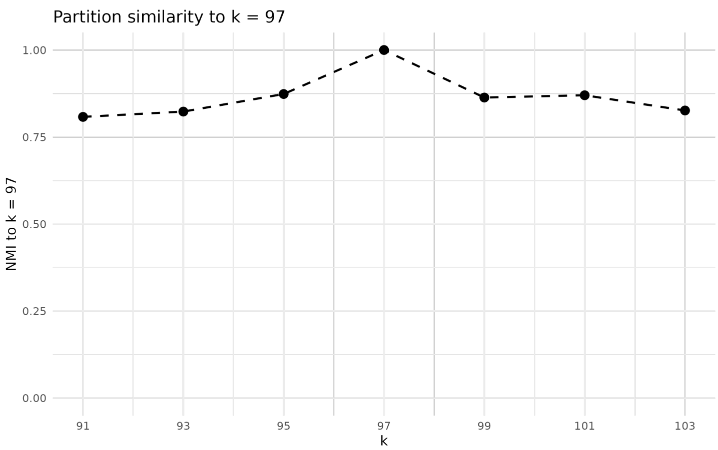
To evaluate the sensitivity of the network structure to the choice of k, we examined how key network properties vary across a range of k values (Supplementary Figure 2). The results show that the size of the largest connected component and overall edge density stabilise around this value, indicating that the network structure is robust to moderate changes in k.



Supplementary Figure 2: Network properties as a function of the k-nearest-neighbour parameter k. The size of the largest connected component stabilises around $k = 97$.

We then apply community detection to the page network to identify clusters of pages that are co-engaged by similar sets of users, using the Louvain algorithm [Blondel et al., 2008]. Since the Louvain algorithm can yield slightly different partitions across runs, we assess its stability by repeating the algorithm 100 times and computing normalized mutual information (NMI) [Manning et al., 2008] across results. To mitigate instability, we implement a consensus clustering approach [Nguyen and Caruana, 2007] that aggregates information across runs via a co-association matrix, yielding a final, robust set of communities.

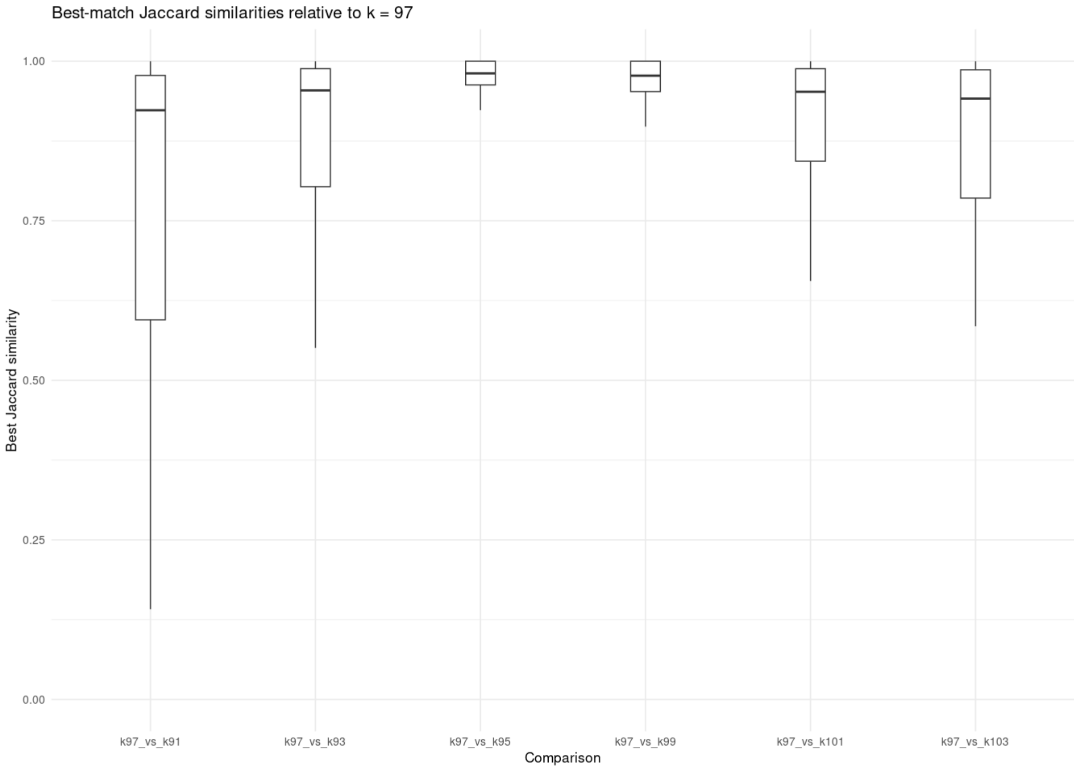
To further assess robustness, we repeated the community detection procedure across multiple nearby values of k and compared the resulting partitions using NMI and best-match Jaccard similarity. The resulting partitions exhibit high similarity across specifications, indicating that the identified communities are stable and not driven by a specific parameter choice (Supplementary Figures 3 and 4).



Supplementary Figure 3: Partition similarity (NMI) relative to the baseline partition at $k = 97$.

1.3. Alternative Outcome Variable Specifications for Testing Robustness

To facilitate transparency and robustness checks, we constructed four alternative specifications of the misinformation engagement indices, reflecting different assumptions about the relative



Supplementary Figure 4: Best-match Jaccard similarity of communities across alternative values of k.

weight of engagement behaviours. The "main" index is reported in the main text, while results based on the other specifications—emphasizing higher intent, greater exposure, or excluding potentially sceptical engagement—are provided below.

Supplementary Table 2 summarises the components and their associated multipliers for each outcome variable alternative.

Supplementary Table 2: Component Multipliers in Misinformation Engagement Indices

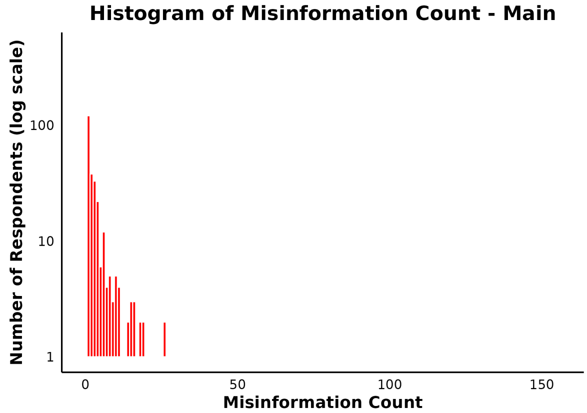
Component	Main Index	High-Intent Emphasis	Exposure Emphasis	Excl. Sceptical
Posts with Misinformation Links	1.0	1.2	1.0	1.0
Comments with Misinformation Links	1.0	1.2	1.0	1.0
Likes/Follows (Pages and Groups)	0.6	0.6	0.8	0.6
Reactions Pages/Groups	0.3	0.3	0.4	0.1
Comments on Pages/Groups	0.7	0.7	0.7	excluded

Each index was computed as the weighted sum of these components. For example, the main index was calculated as:

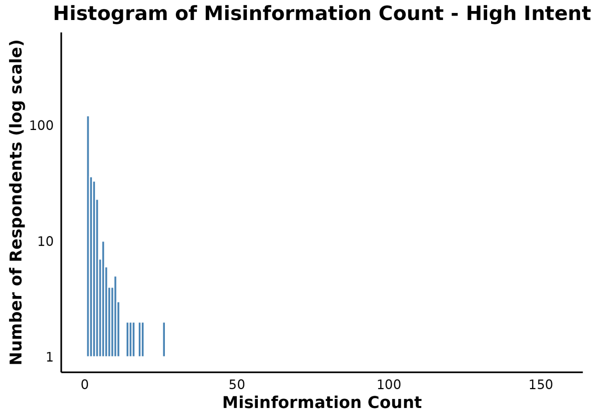
$$\begin{aligned}\texttt{misinfo\_count\_main} = {} & 1.0 \times \text{misinfo\_posts} \;+\; 1.0 \times \text{misinfo\_comments} \;+ \\ & 0.6 \times \big(\text{misinfo\_pages\_liked\_count} + \text{misinfo\_groups\_liked\_count}\big) \;+ \\ & 0.3 \times \big(\text{misinfo\_page\_reactions\_count} + \text{misinfo\_group\_reactions\_count}\big) \;+ \\ & 0.7 \times \big(\text{misinfo\_page\_comments\_count} + \text{misinfo\_group\_comments\_count}\big).\end{aligned}$$

Analogous formulas were applied to construct the other indices using the multipliers shown in Supplementary Table 2.

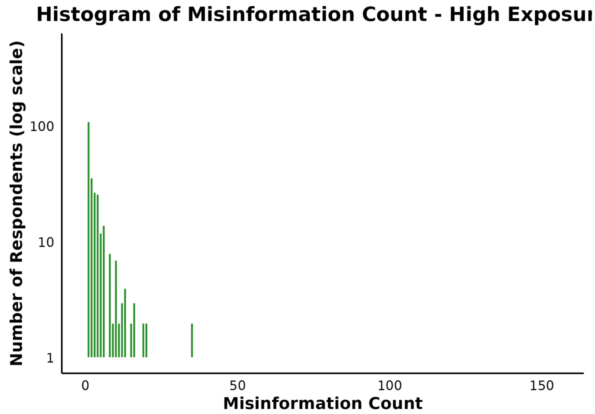
1.4. Distributions of the Four Outcome Variable Specifications



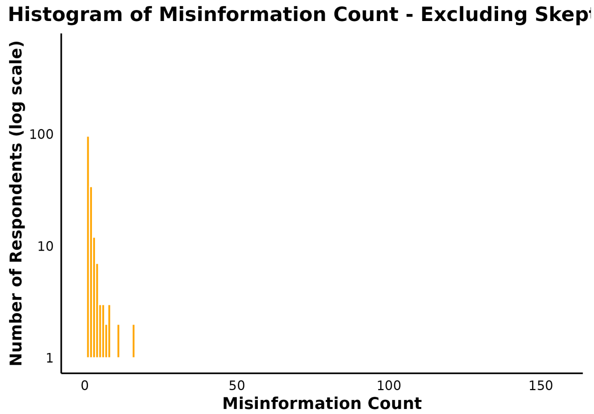
(a) Main



(b) Intent



(c) Exposure



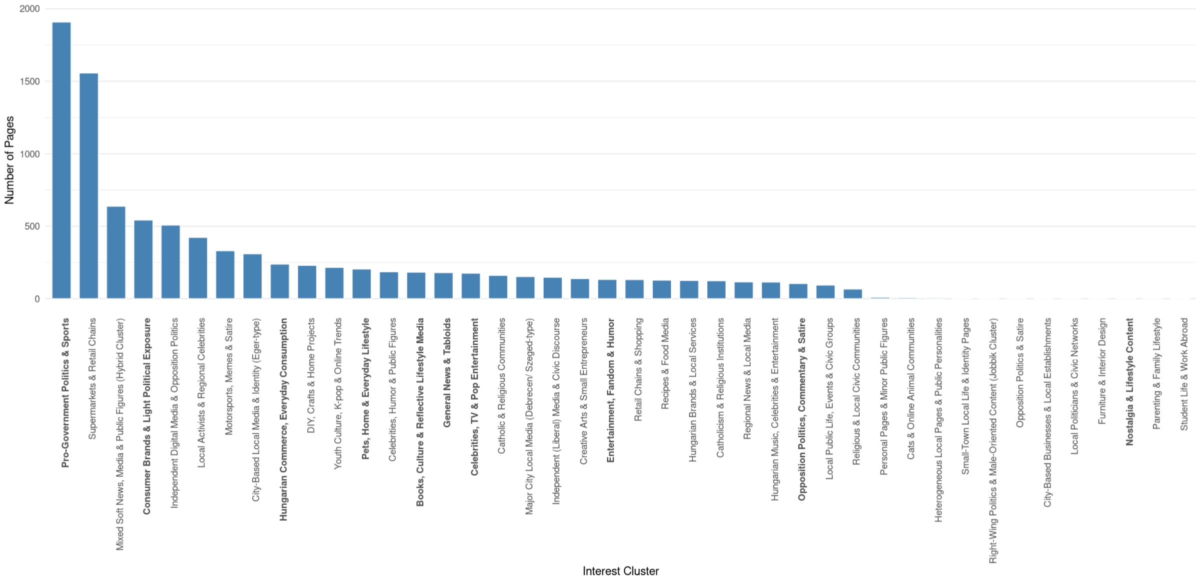
(d) Excluding sceptical

Supplementary Figure 5: Distribution of the four outcome variable specifications measuring misinformation engagement. A large share of observations take the value zero (no engagement), while the remainder form a long right-skewed tail. This distribution motivates the use of hurdle models, combining a logistic component for predicting any engagement and a Gamma regression with log link for modeling the positive values.

2. Supplementary Information for Results

2.1. Interest Clusters Identified in the Facebook Page Network

As described in the main text, we applied community detection to the Facebook page network to identify clusters of pages with similar patterns of follower overlap. (Supplementary Figure 6) shows the distribution of the number of pages in each of the identified interest clusters. The table below provides the full list of 46 clusters obtained through community detection on the Facebook page network. For each cluster, we assigned a descriptive label that captures its dominant theme, supplemented by a brief summary of the content and examples of representative pages. These examples are included for illustrative purposes and highlight the most central or recognizable nodes within each cluster. The descriptive labels are not intended as strict categorical definitions, but rather as interpretive guides to facilitate understanding of the structure of the network.



Supplementary Figure 6: Distribution of Interest Cluster Sizes. Clusters that are significant in any of the models are highlighted in bold.

Supplementary Table 3: Facebook Page Clusters

Name	Description	Example Pages	Page Count	User Count	Interaction Count
Pro-Government Politics & Sports	Government actors, pro-government media, and sports-related identity content.	Orbán Viktor (PM), Nemzeti Sport (sports daily), M4 Sport (sports TV), Trollfoci (football memes), Magyarország Kormánya (Government)	1909	705	357700
Supermarkets & Retail Chains	Large retail chains and supermarket brands focused on everyday consumption.	Lidl Magyarország, Aldi Magyarország, Spar Magyarország, Tesco Magyarország, Penny Market	1557	656	301116
Mixed Soft News, Media & Public Figures	Hybrid cluster of media pages, soft news, and public personalities.	Femina.hu (lifestyle media), NLC.hu (portal), Blikk Rúzs (tabloid), Promotions.hu (media), Contextus (culture)	638	651	72420
Consumer Brands & Light Political Exposure	Consumer brands with occasional exposure to right-leaning political content.	Tesco Magyarország, Regio Játék (toy store), Alfahír (right-wing news), Pizza Sprint (food chain), MediaMarkt Magyarország	543	623	71009
Independent Digital Media & Opposition Politics	Digital-native independent media with opposition-oriented political content.	Telex (news), 444.hu (liberal media), Partizán (political channel), Direkt36 (investigative), Átlátszó (watchdog)	508	631	79677

Name	Description	Example Pages	Page Count	Interaction Count	
Local Activists & Regional Celebrities	Local public figures, activists, and regionally known personalities.	Debrecenben Hallottam (local news), Szeged365 (city media), Influenszer Anita (local influencer), Szabolcsi Hírek (regional), Városi Pletykák (local gossip)	423	503	50301
Motorsports, Memes & Satire	Car culture, motorsports content, and humor-driven meme pages.	Totalcar (auto media), Vezess.hu (cars), Trollfoci (memes), Faszagyerek (humor), Autós Élet (car lifestyle)	331	519	41240
City-Based Local Media & Identity	Local media ecosystems centered around specific cities and identity.	Eger Hírek (Eger News), Egri Ügyek (local media), Egerben Hallottam (local gossip), Heves Megyei Hírek, Eger Város	310	590	70551
Hungarian Commerce, Everyday Consumption	Consumer goods, services, and everyday commercial activity.	Coop, Reál, Mayo Chix (fashion), Dr-PC (electronics), McSemege (grocery)	239	423	59435
DIY, Crafts & Home Projects	Do-it-yourself, crafting, and home improvement content.	Barkácsolás (DIY), Otthon Ötletek (home ideas), Kézműves Világ (crafts), Fa Műhely (woodwork), Praktiker Magyarország	230	553	64095
Youth Culture, K-pop & Online Trends	Youth-oriented content including K-pop, influencers, and online trends.	BTS Hungary (fan page), K-pop Hungary, TikTok Trendek, Influenszer Figyelő, Anime Világ	216	553	56804
Pets, Home & Everyday Lifestyle	Animal content, home-related media, and everyday lifestyle pages.	Sokszínű Vidék (rural lifestyle), Állatbarátok Klubja (animal lovers), Torta és Karamell (recipes), Német Juhász (dogs), Megaport (portal)	205	512	26310
Celebrities, Humor & Public Figures	Entertainment personalities, humor pages, and public figures.	Janklovics Péter (comedian), Bochkor (radio host), Celeb Hírek (celebrity news), Showder Klub, Humor Heroldok	186	547	57523
Books, Culture & Reflective Lifestyle Media	Books, psychology, and culturally oriented lifestyle content.	National Geographic, Könyvmolyok (book lovers), Tudatos Életmód (mindful living), Bihari Viktória (psychology), Bookline	183	511	34085
General News & Tabloids	Mass media outlets and tabloids focused on daily news and sensational stories.	24.hu, Blikk, Ripost, ATV, Időkép (weather)	180	595	68514

Name	Description	Example Pages	Page Count	Interaction Count	
Celebrities, TV & Pop Entertainment	Television personalities and mainstream entertainment content.	Till Attila (TV host), Kulcsár Edina (celebrity), TV2, RTL Klub, CelebVilág	176	494	60935
Catholic & Religious Communities	Religious pages focused on Catholic faith and community life.	Katolikus Egyház (Catholic Church), Mária Rádió (radio), Vatican News, Szentmise (mass), Bibliai Idézetek	161	483	16772
Major City Local Media	Media ecosystems of large Hungarian cities.	Debrecenben Hallottam, Szeged365, Pécs Aktuál, Győr Hírek, Miskolc Ma	153	486	31247
Independent (Liberal) Media & Civic Discourse	Liberal media and civic discussion platforms.	HVG, Mérce, Magyar Narancs, Klubrádió, Civil Hang	148	446	20302
Creative Arts & Small Entrepreneurs	Artists, creators, and small-scale businesses.	Kézműves Bolt (craft shop), Etsy Hungary, Illusztrátorok (illustrators), Handmade Hungary, Design Market	139	379	20596
Entertainment, Fandom & Humor	Mixed entertainment content including fandoms and humor pages.	Dombóvári István (comedian), Radio Face, Meme Oldalak, Stand-up Hungary, Popkultúra	133	498	40971
Retail Chains & Shopping	Retail-focused content including stores and shopping environments.	IKEA Hungary, Decathlon Magyarország, Pepco Hungary, Jysk Magyarország, Möbelix	132	498	25939
Recipes & Food Media	Cooking, recipes, and food-related media.	Street Kitchen, Nosalty, Mindmegette, Főzz Velünk (cook with us), Receptneked	128	481	24040
Hungarian Brands & Local Services	Domestic brands and service providers.	Benu Gyógyszertár (pharmacy), MOL (oil company), OTP Bank, Telekom Hungary, Posta (postal service)	126	554	65910
Catholicism & Religious Institutions	Institutional religious organizations and formal church structures.	Vatican, Katolikus Karitász (charity), Jezsuiták (Jesuits), Püspökség (bishopric), Egyházmegye (diocese)	124	494	24255
Regional News & Local Media	Regional media outlets and local news sources.	Borsod Online, Kisalföld, ZAOL (Zala news), HEOL (Heves news), VAOL (Vas news)	116	460	10373
Hungarian Music, Celebrities & Entertainment	Music artists and entertainment figures.	Azahriah (musician), Follow the Flow (band), Majka (rapper), Halott Pénz (band), Rádió 1	115	517	26791

Name	Description	Example Pages	Page Count	Interaction Count	
Opposition Politics, Commentary & Satire	Opposition politicians and satire-driven political discourse.	Tibi Atya (satire), Gyurcsány Ferenc (politician), Karácsony Gergely (mayor), Tudományos MémeK (memes), Kétfarkú Kutya Párt	105	500	47636
Local Public Life, Events & Civic Groups	Local institutions, events, and civic organizations.	Városi Programok (city events), Önkormányzat (municipality), Kulturális Központ, Helyi Egyesület (local association), Városi Napok	94	461	11713
Religious & Local Civic Communities	Small-scale religious and civic community groups.	Nyugdíjas Klub (retiree club), Helyi Egyházközség (parish), Közösségi Csoport, Ima Csoport (prayer group), Falusi Közösség	67	380	8029
Personal Pages & Minor Public Figures	Individual profiles and minor influencers with weak thematic cohesion.	Szőcs Renáta (person), Anett Tóth (person), Nemazalány (influencer), Best FM Debrecen (radio), Romanii au Talent (TV show)	10	90	243
Cats & Online Animal Communities	Cat-focused pages and small animal communities.	Cicás Képek (cat photos), Macska Világ (cat world), Cuki Cicák (cute cats), Macska Klub (cat club), Állatbarátok (animal lovers)	8	77	650
Heterogeneous Local Pages & Public Personalities	Mixed local content and personalities without a clear theme.	Helyi Pletykák (local gossip), Városi Arcok (local figures), Influenszerek, Közösségi Oldalak, Vegyes Tartalom	7	136	1212
Small-Town Local Life & Identity Pages	Hyper-local identity and community pages of small towns.	Salgótarján Hírek, Vidéki Élet (rural life), Falusi Közösség, Helyi Bolt, Kisvárosi Hírek	3	27	68
Right-Wing Politics & Male-Oriented Content	Right-wing political content combined with male-oriented interests.	Jobbik, Alfahír, Autós Oldalak (cars), Férfi Életmód (men's lifestyle), Katonai Tartalom (military)	2	21	70
Opposition Politics & Satire	Political satire and opposition discourse.	Kétfarkú Kutya Párt, Tibi Atya, Politikai MémeK, Ellenzéki Oldalak, Humor Politika	2	22	91
City-Based Businesses & Local Establishments	Urban businesses and local commercial establishments.	Étterem Debrecen (restaurant), Kávézó Szeged (café), Helyi Bolt, Városi Üzletek, Szolgáltatók	2	45	242

Name	Description	Example Pages	Page Count	Interaction Count	
Local Politicians & Civic Networks	Local political actors and civic networks.	Polgármester (mayor), Képviselő (representative), Önkormányzat (municipality), Helyi Politika, Városi Tanács	2	27	139
Furniture & Interior Design	Furniture stores and interior design content.	IKEA, Jysk, Möbelix, Lakberendezés (home decor), Bútor Világ (furniture world)	2	22	68
Nostalgia & Lifestyle Content	Nostalgia-driven content and reflective lifestyle pages.	Nosztalgia Klub (nostalgia club), Régi Idők (old times), Retro Hungary, Emlékek (memories), Régi Képek (old photos)	2	19	28
Parenting & Family Lifestyle	Family life, parenting advice, and child-related content.	Anyák Klubja (mothers club), Baba-Mama (baby-mother), Gyereknevelés (parenting), Családi Élet (family life), Szülők Oldala	2	43	118
Student Life & Work Abroad	Student communities and working abroad experiences.	Erasmus Hungary, Külföldi Munka (work abroad), Diákélet (student life), Work & Travel, Tanulás Külföldön (study abroad)	2	87	709

Note: Clusters vary in size and thematic cohesion. While larger clusters are organized around clearly identifiable domains such as political communication, news media, or consumer activity, smaller clusters often capture more narrowly defined or weakly structured content, including personal profiles, local communities, and niche online content. The naming of clusters is therefore partly heuristic, aiming to balance descriptive clarity with interpretive flexibility.

2.2. Full Results for the Main Outcome Variable Specification

In this subsection, we present the full regression results for our models. In the main paper, we only reported statistically significant coefficients for the Main outcome. Here, we include the complete coefficient estimates for both the zero models (estimating the likelihood of sharing no misinformation) and the positive models (estimating the volume of misinformation shared, conditional on sharing at least once).

2.2.1 Zero models

Supplementary Figure 7 presents the results of the zero models, which estimate the probability that a respondent engages with misinformation at all (versus never sharing any). These models therefore capture the baseline likelihood of misinformation engagement.

Supplementary Table 4 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.046 (Model 1), and 0.283 (Model 2), reflecting the significant contribution of interest clusters and behavioural controls to model fit.

Summary. These models show that age, particular Facebook page–based interest clusters, and Facebook activity are significantly associated with individual misinformation engagement.

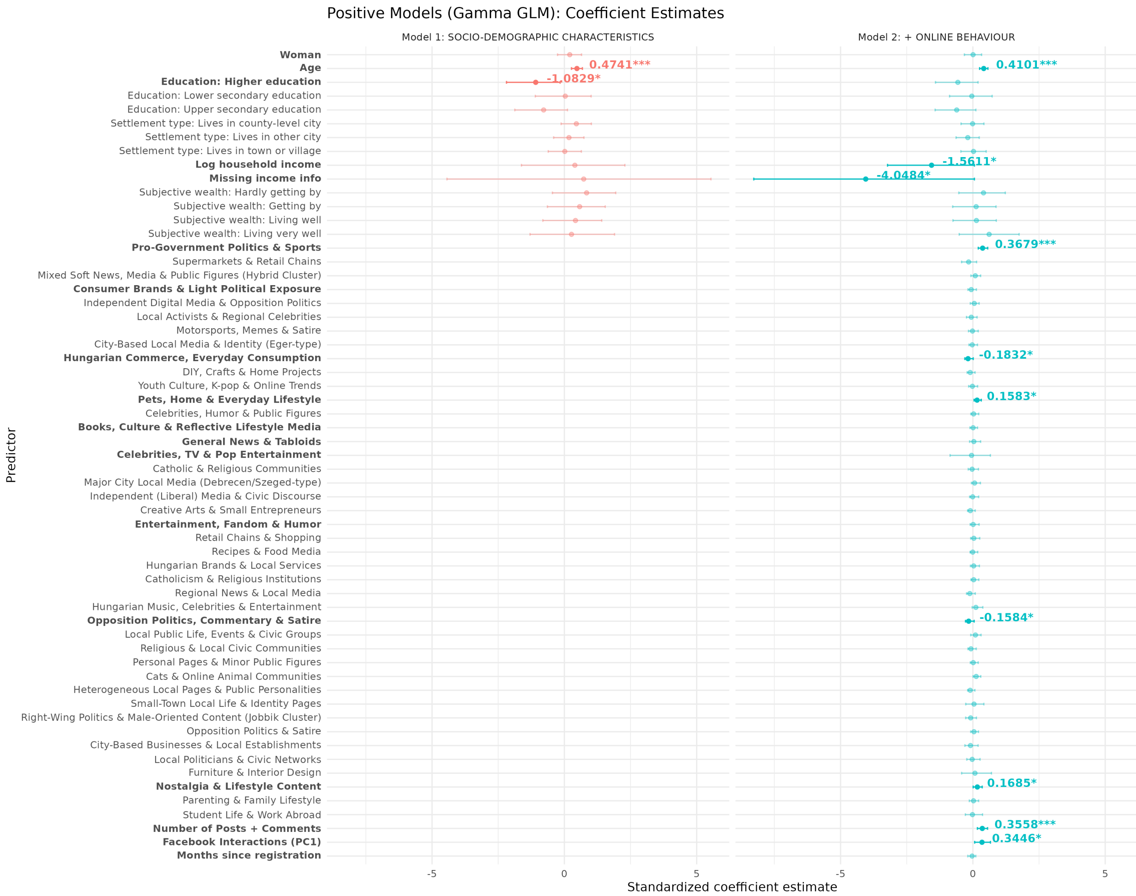
Supplementary Figure 7: Zero models predicting zero vs non-zero sharing (Main outcome).

2.2.2 Positive models

Supplementary Figure 8 shows the results of the positive models, which estimate the volume of misinformation shared among those who have engaged at least once. These models thus shed light on what drives the intensity of engagement.

Supplementary Table 4 presents that adjusted McFadden R^2 values for the positive models are 0.048 (Model 1), and 0.150 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. The positive models show that, conditional on engaging with misinformation, age, Facebook page–based interest clusters (some different from those affecting misinformation engagement likelihood), and platform activity are significantly associated with the volume of individual misinformation engagement.



Supplementary Figure 8: Positive models predicting volume of sharing (Main outcome).

Supplementary Table 4: Model Fit: Adjusted McFadden R^2 for Zero and Positive Misinformation Engagement Models with the Main Outcome Variable

Model	**Model 1**	**Model 2**
Zero models	0.046	0.283
Positive models	0.048	0.150

2.2.3 Combined summary of main outcome models

Taken together, the zero-inflation and positive models clarify the two stages of misinformation engagement. The zero models highlight the factors that determine whether individuals engage at all, showing robust roles for age, specific Facebook page–based interest clusters, and Facebook activity measures. The positive models then show that, among those who engage with misinformation, similarly to the zero models, age, Facebook page–based interest clusters–although some different from those in the zero models–, and platform activity are significantly associated with the volume of individual misinformation engagement. Together, these results provide the full picture underlying the significant-only results reported in the main text.

2.3. Robustness Checks with the Three Alternative Variable Specifications

In this section we present results for three alternative outcome variables: *High Intent*, *Exposure*, and *Excluding sceptical*. These serve as robustness checks to demonstrate that our findings are not dependent on the exact operationalization of misinformation sharing.

2.3.1 High Intent outcome

2.3.2 Zero models

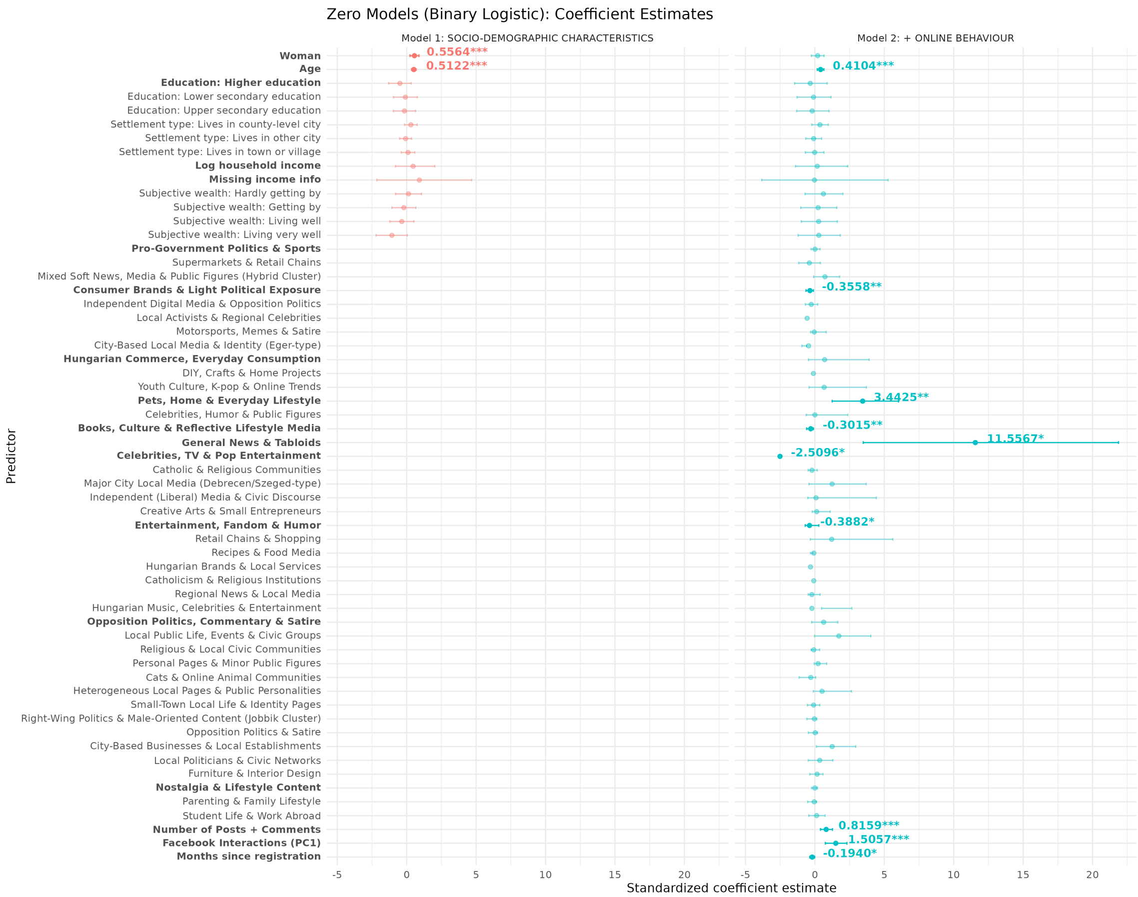
Supplementary Figure 9 presents the results of the zero models for the high-intent outcome, which predict whether a respondent engages with misinformation at all. The patterns closely resemble those reported for the Main outcome (Supplementary Figure 7), with age emerging as the most consistent predictor: older respondents are significantly more likely to engage with misinformation. As in the Main outcome models, the effect of gender is significant in the simplest specification, with women more likely to engage, but attenuates once interest clusters and behavioural controls are included. Education and income remain non-robust predictors.

The set of significant Facebook page–based interest clusters closely resembles the Main outcome. Engagement with *General News & Tabloids* and *Eger and Urban Culture & Media* predicts a higher likelihood of misinformation sharing, while affinity with clusters such as *Independent Liberal Media & Civic*, *Jobbik & Male Hobbies*, and *DIY, Woodworking & Car Enthusiasts* predicts lower engagement.

Finally, platform activity variables show the same pattern as in the Main outcome: the number of posts/comments and Facebook interaction intensity are positively associated with a greater likelihood of engagement, while account age shows a negative association.

Supplementary Table 5 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.046 (Model 1), and 0.283 (Model 2), reflecting the significant contribution of interest clusters and behavioural controls to model fit.

Summary. These results confirm the robustness of the Main outcome models: while higher activity levels increase the likelihood of misinformation engagement, age and certain Facebook page–based interest clusters robustly predict whether individuals engage with misinformation at all.



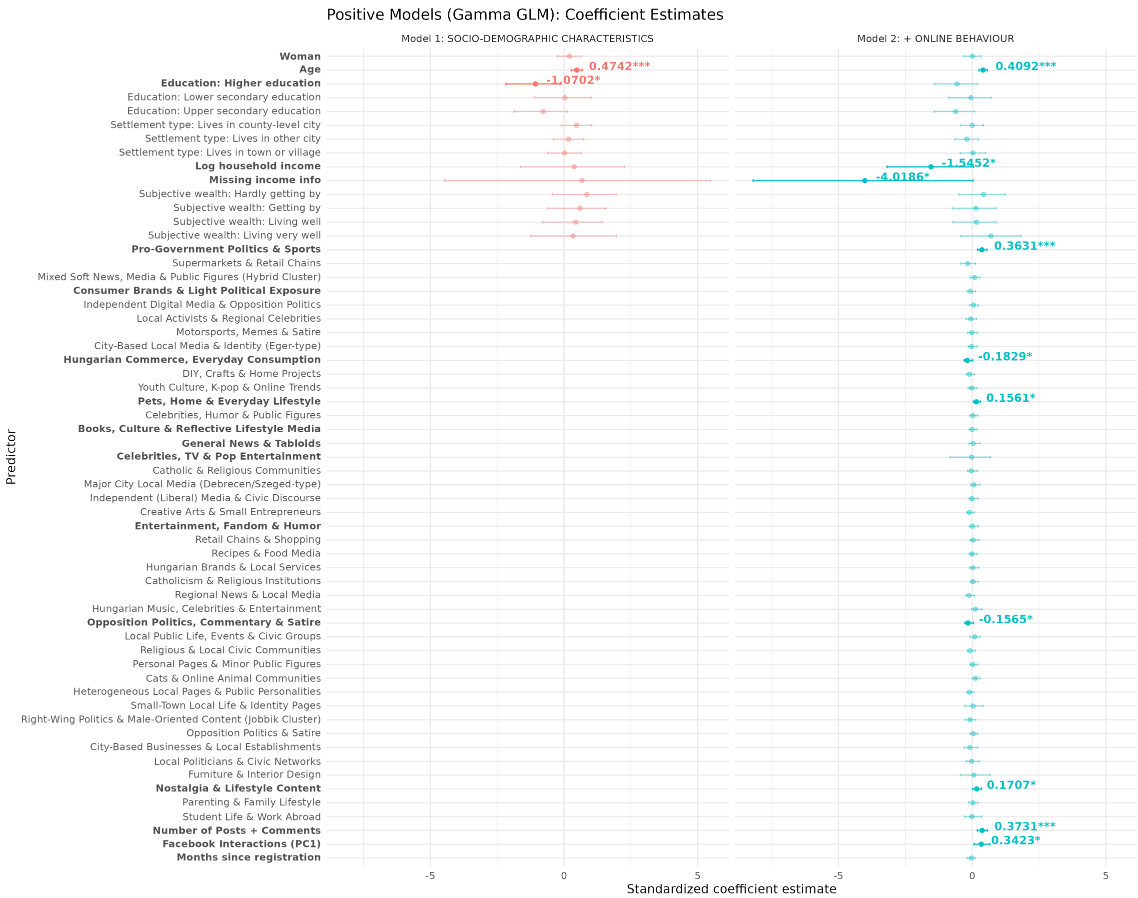
Supplementary Figure 9: Zero models predicting zero vs non-zero sharing (High Intent outcome).

2.3.3 Positive models

Supplementary Figure 10 shows the results of the positive models for the high-intent outcome, which estimate the volume of misinformation shared among those who have engaged at least once. As with the Main outcome models (Supplementary Figure 8), age remains the most consistent demographic predictor, with older respondents sharing larger volumes of misinformation. Gender is not significant once engagement is established, while education shows a negative effect in the simplest model but loses significance in the full model. The same Facebook page–based interest clusters remain significant predictors of volume as in the Main outcome. Behavioural predictors again mirror the Main outcome: the number of posts/comments is a significant positive predictor, with the effect size being somewhat stronger than in the Main outcome model. The effect is account age is not significant.

Supplementary Table 5 presents that adjusted McFadden R^2 values for the positive models are 0.048 (Model 1), and 0.149 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. The positive models for the high-intent outcome largely replicate the Main outcome findings: conditional on engaging with misinformation, age, Facebook page–based interest clusters, and platform activity are significantly associated with the volume of individual misinformation engagement.



Supplementary Figure 10: Positive models predicting volume of sharing (High Intent outcome)

Supplementary Table 5: Adjusted McFadden R^2 for Zero and Positive Intent Models

Model	Model 1	Model 2
Zero models	0.046	0.283
Positive models	0.048	0.149

2.3.4 Combined summary of high-intent models

The high-intent outcome models confirm that the Main outcome results are robust to this alternative specification. Both sets of models highlight the same structural patterns: age, certain interest clusters, and platform activity are consistently associated with misinformation engagement. The overlap across outcomes indicates that the Main results are not driven by the choice of outcome variable.

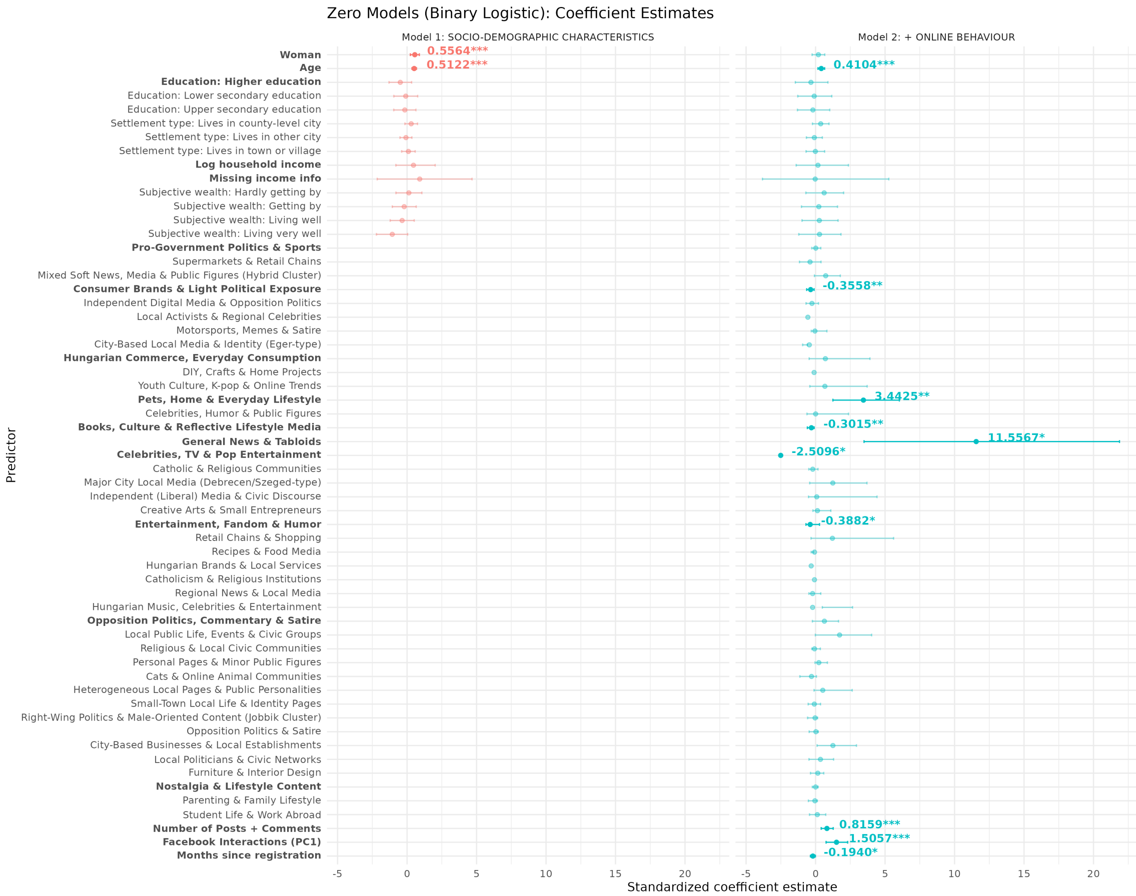
2.3.5 Exposure outcome

2.3.6 Zero models

Supplementary Figure 11 presents the zero models predicting whether respondents engage with misinformation at all under the Exposure specification. The demographic patterns largely mirror the Main outcome (Supplementary Figure 7) and the High Intent specification (Supplementary Figure 9).

Supplementary Table 6 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.046 (Model 1), and 0.283 (Model 2), reflecting the significant contribution of interest clusters and behavioural controls.

Summary. The zero model results in the Exposure specification largely replicate the Main and High Intent outcomes: while higher activity levels increase the likelihood of misinformation engagement, age and certain Facebook page–based interest clusters robustly predict whether individuals engage with misinformation at all.



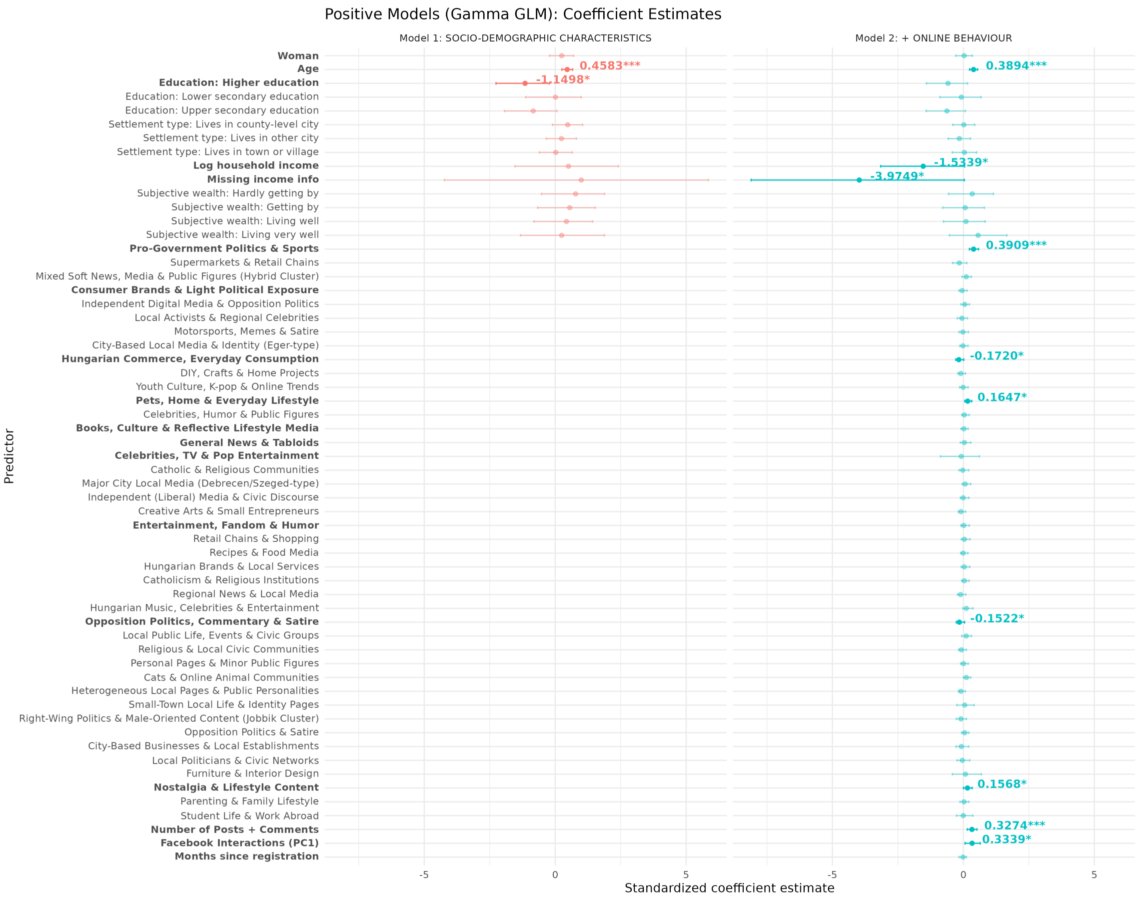
Supplementary Figure 11: Zero models predicting zero vs non-zero sharing (Exposure outcome).

2.3.7 Positive models

Supplementary Figure 12 shows the positive models estimating the volume of misinformation shared under the Exposure specification. Again, results closely resemble the Main outcome. Age is a consistent positive predictor of sharing volume. Higher education predicts lower sharing in the first model, but this effect weakens after adding behavioural variables. Interest cluster coefficients and the effect of Facebook activity replicate the Main and High Intent outcome patterns.

Supplementary Table 6 presents that adjusted McFadden R^2 values for the positive models are 0.045 (Model 1), and 0.149 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. Positive models with the exposure outcome variable reproduce the Main and High Intent outcome findings, confirming the robustness of the patterns across specifications.



Supplementary Figure 12: Positive models predicting volume of sharing (Exposure outcome).

Supplementary Table 6: Adjusted McFadden R^2 for Zero and Positive Exposure Models

Model	**Model 1**	**Model 2**
Zero models	0.046	0.283
Positive models	0.045	0.149

2.3.8 Combined summary of exposure models

Overall, the Exposure outcome variable specification confirms that the Main findings are robust to alternative modelling of engagement. Both the zero and positive model results show the same demographic and behavioural patterns as the Main and High Intent outcomes. This suggests that the observed predictors of misinformation engagement are not sensitive to minor differences in model specification or sampling.

2.3.9 Excluding sceptical respondents

2.3.10 Zero models

Supplementary Figure 13 shows the zero models predicting whether respondents engage with misinformation at all, excluding comments from the outcome variable specification, to avoid ambiguity due to potential scepticism. The main demographic patterns largely replicate the Main outcome Supplementary Figure 7), with age emerging again as the strongest predictor of engagement. Older individuals remain significantly more likely to engage. Women show a positive association in the simplest specification, which disappears when Facebook page–based interest clusters and behavioural controls are added. A difference that arises compared to the Main outcome is that Living in a county-level city is positively associated with misinformation engagement. The effect is however only weakly significant.

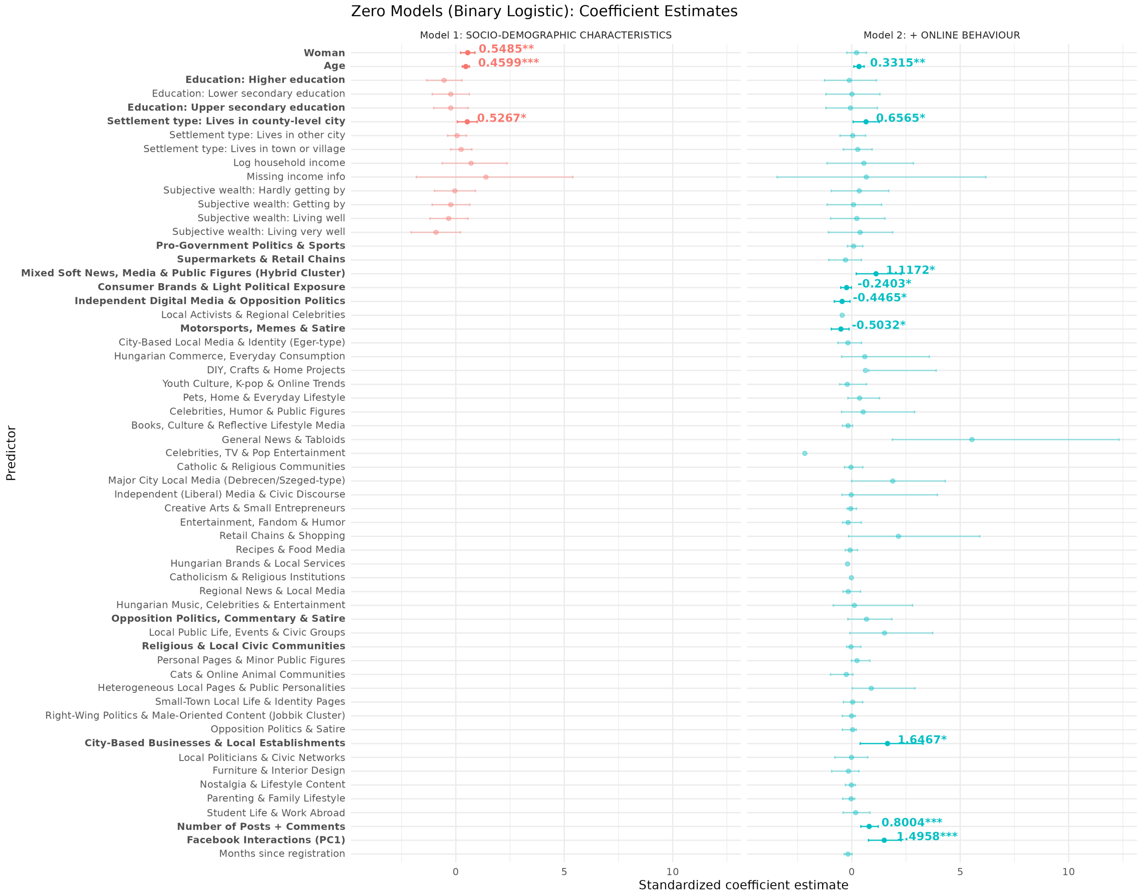
Regarding Facebook page–based interest clusters, some the same patterns re-emerge as for the other outcome specifications: *General News & Tabloids* and *Eger and Urban Culture & Media* are associated with higher engagement, while *Independent Liberal Media & Civic*, and *Recipes & Food Media* predict lower engagement.

Differently from the other outcome variable specifications, *Consumer Brands & Government Mix*, and *Local Fashion, Pets & Retail* emerge as negative associations with misinformation engagement likelihood, and none of the other interest clusters turn out to be significantly associated with misinformation engagement.

Finally, among behavioural variables, as in the case of the other outcome variable specifications, more posts/comments and higher Facebook interaction intensity strongly predict misinformation engagement, replicating the full-sample findings. However, the effect of account age is not significant.

Supplementary Table 7 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.038 (Model 1) and 0.263 (Model 2) reflecting the significant contribution of interest clusters and behavioural controls to model fit.

Summary. Excluding potentially sceptical engagements from the outcome specification leaves the zero model results largely unchanged: while controlling for activity, age and set of interest clusters remain key predictors.



Supplementary Figure 13: Zero models predicting zero vs non-zero sharing (Excluding sceptical outcome).

2.3.11 Positive models

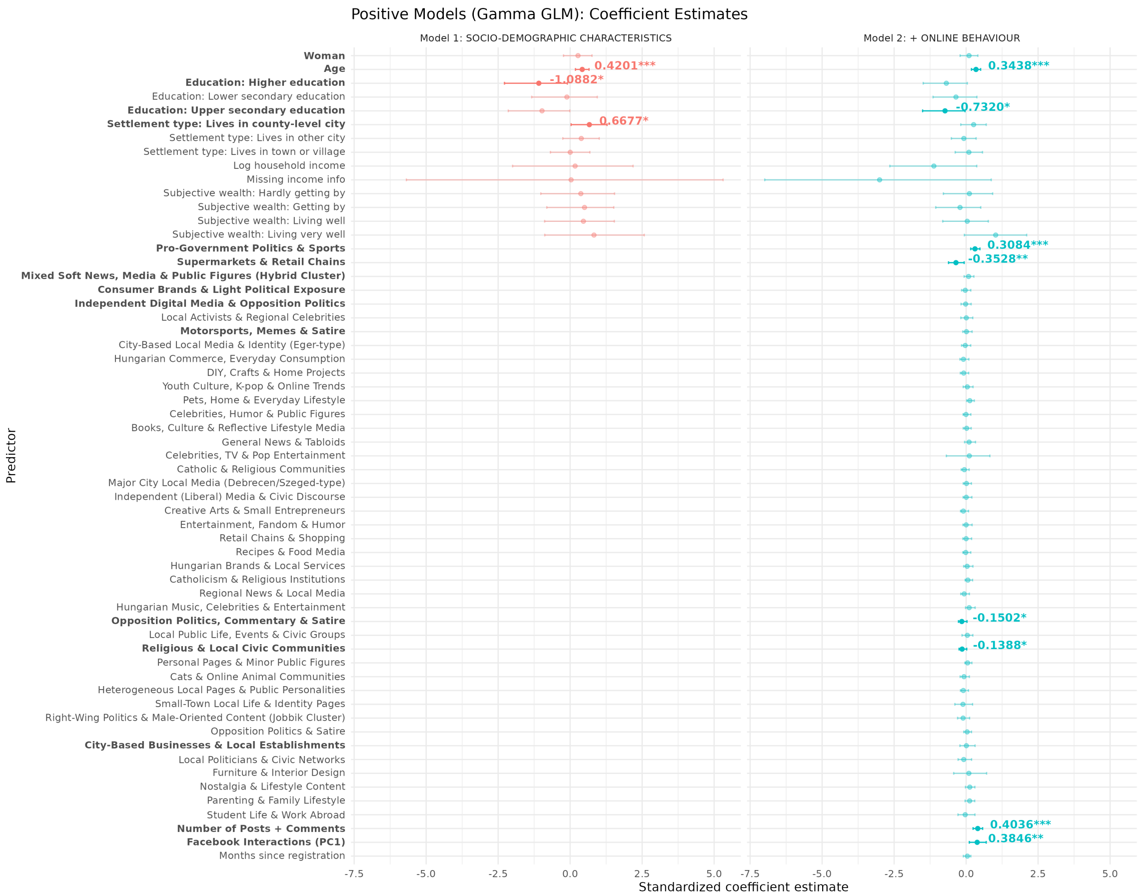
Supplementary Figure 14 presents the positive models, which estimate the volume of misinformation engagement among those who engage, excluding sceptical engagement. The results again mirror the Main outcome (Supplementary Figure 8), with age a consistent and strong predictor of sharing volume. As in the case of the other outcome variables, education shows a negative effect in the simplest specification (higher education associated with less sharing), but this attenuates once behavioural variables are added. As opposed to the other outcome specifications, higher subjective income is associated with higher volume of misinformation engagement in the simplest model. But this effect is explained away once behavioural variables are introduced.

Compared to the Main outcome, most interest cluster effects remain, while few shifts in interest cluster associations emerge. As seen for other outcome variable specifications, *Hungarian Brands & Commercial Services* predicts lower volumes, while *Pro-Government Politics & Sports* and *Hungarian Books, Personalities & Pop Entertainment* show positive associations with volume. Differently from the other outcome variable specifications, *Opposition Politics & Satire* appear negatively associated, and other interest clusters effects are not significant.

Behavioural predictors remain robust. The number of posts/comments and Facebook interaction intensity is a significant positive predictor of volume, consistent with the Main outcome.

Supplementary Table 7 presents that adjusted McFadden R^2 values for the positive models are 0.052 (Model 1), and 0.204 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. The positive models excluding sceptical engagement confirm the robustness of the Main findings: while controlling for activity age and some interest clusters remain the strongest predictors, with minor shifts in certain effects.



Supplementary Figure 14: Positive models predicting volume of sharing (Excluding sceptical outcome).

Supplementary Table 7: Adjusted McFadden R^2 for Zero and Positive Excluding Sceptical Models

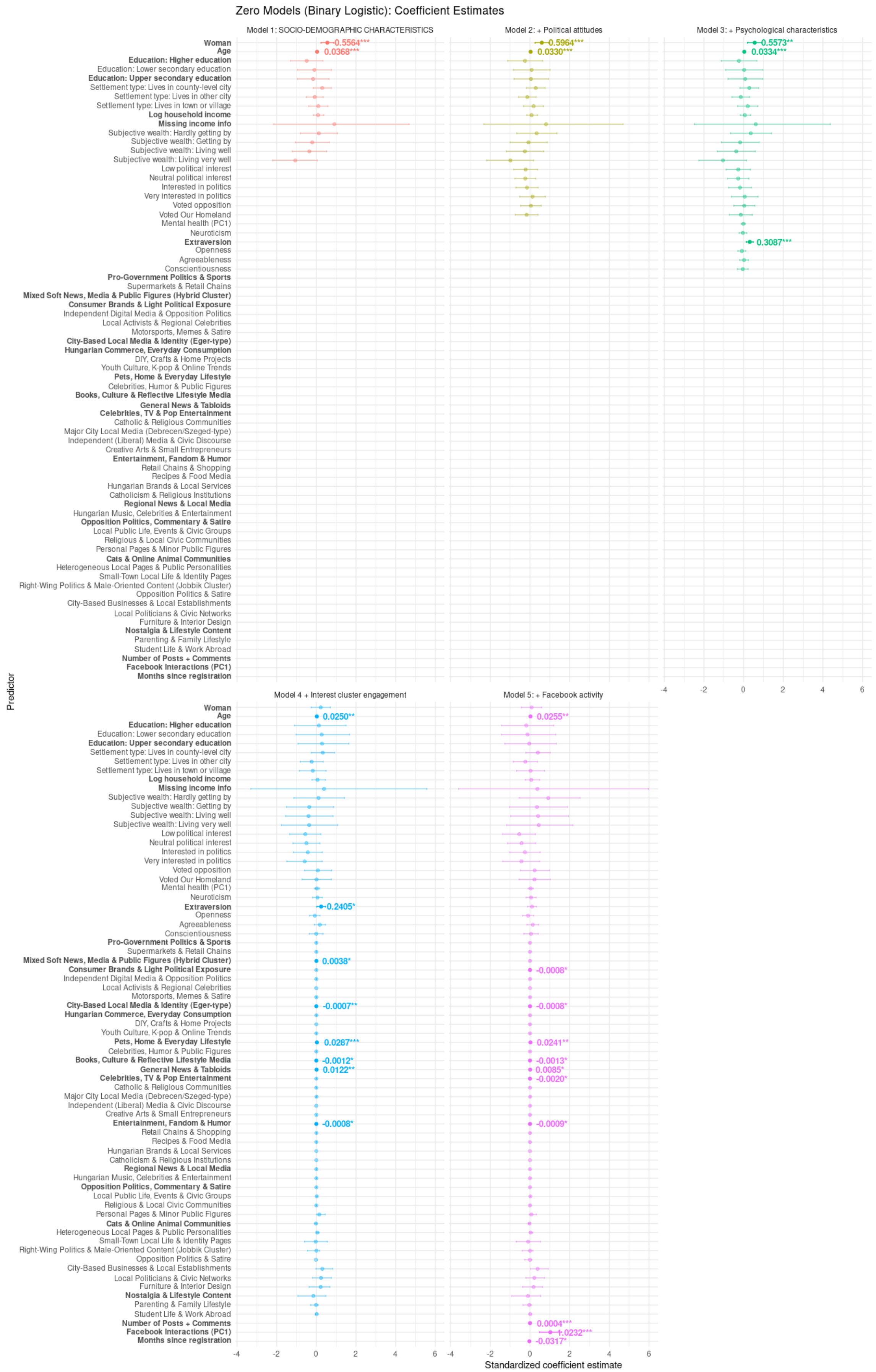
Model	**Model 1**	**Model 2**
Zero models	0.038	0.263
Positive models	0.052	0.204

2.3.12 Combined summary of exclusion models

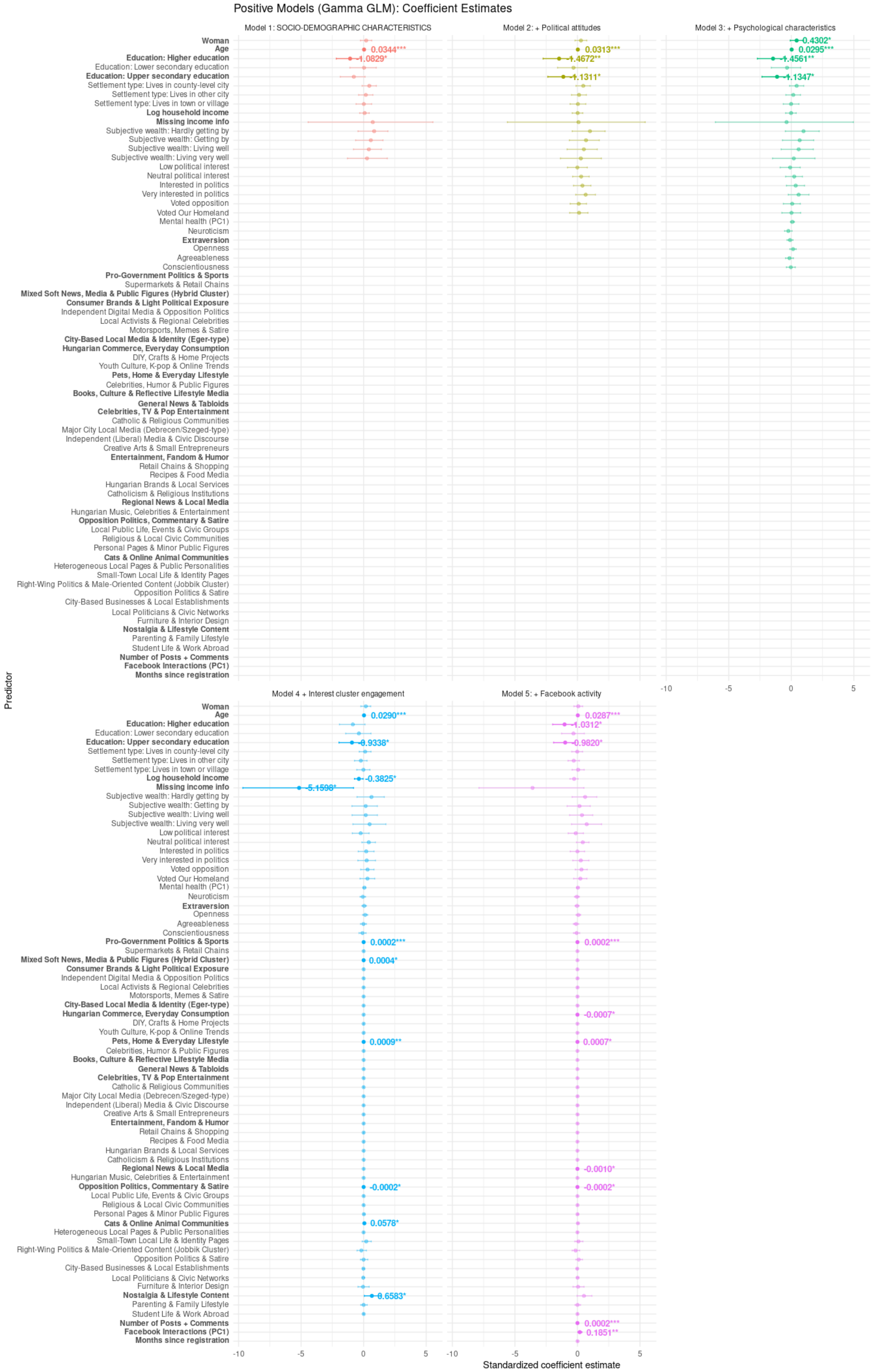
Overall, excluding sceptical respondents does not materially change the structure of the results. Across both model types, when controlling for Facebook activity, age and a set of Facebook page–based interest clusters remain central predictors of misinformation engagement and volume. A handful of additional or shifted associations emerge, but most cluster associations remain qualitatively similar. These findings indicate that the Main results are not driven by sceptical respondents, further strengthening robustness.

2.4. Robustness Checks with the Inclusion of Political Attitudes and Psychological Characteristics

The following figures present results from models that additionally include respondents' political attitudes and psychological characteristics. Supplementary Figure 15 shows the zero models, estimating the likelihood of engaging with misinformation, while Supplementary Figure 16 shows the positive models, estimating the volume of engagement among participants who shared at least once. These additions allow us to examine the potential role of political and psychological factors alongside socio-demographic and behavioural predictors, and the results indicate that these additional variables are generally not robustly significant, with the main patterns observed in the original models remaining largely unchanged. Supplementary Table 8 presents the adjusted McFadden R^2 values for each of the 5 zero and positive models.



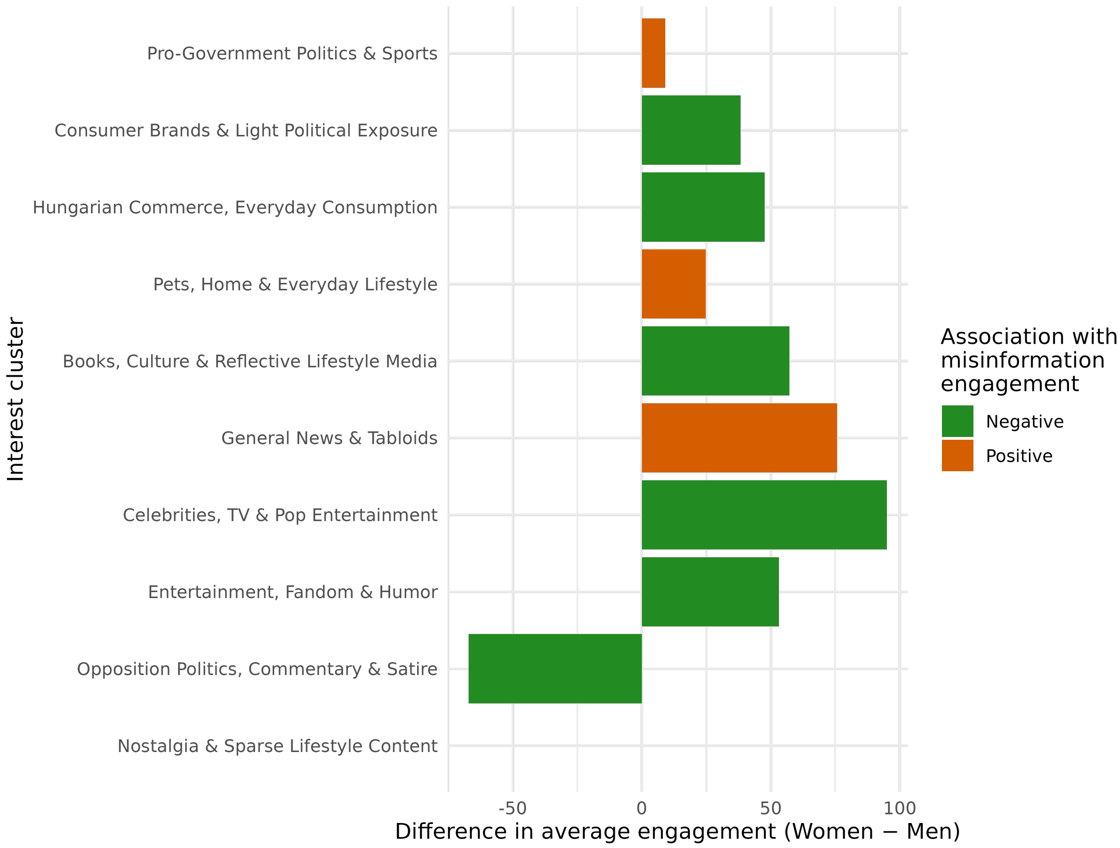
Supplementary Figure 15: Zero models predicting sharing, including Political Attitudes and Psychological Characteristics.



Supplementary Figure 16: Positive models predicting volume of sharing, including Political Attitudes and Psychological Characteristics.

Supplementary Table 8: Model Fit: Adjusted McFadden R^2 for Zero and Positive Misinformation Engagement Models with the Main Outcome Variable

Model	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Zero models	0.048	0.093	0.095	0.269	0.298
Positive models	0.048	0.095	0.094	0.170	0.203



Supplementary Figure 17: Gender differences in engagement across interest clusters significantly associated with misinformation engagement. The figure shows the difference in average engagement with each cluster between women and men (women - men), based on the user-level cluster engagement measures. Positive values indicate clusters with higher engagement among women, while negative values indicate clusters with higher engagement among men. Bar colors indicate the direction of association between each cluster and misinformation engagement based on hurdle model estimates: clusters positively associated with higher engagement are shown in orange, while those negatively associated are shown in green. Only clusters significant in the regression models are shown.

References

Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008. doi: 10.1088/1742-5468/2008/10/P10008. URL `https://doi.org/10.1088/1742-5468/2008/10/P10008`.

Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press, Cambridge, 2008.

Nam Nguyen and Rich Caruana. Consensus clusterings. *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, pages 607–612, 2007. URL `https://api.semanticscholar.org/CorpusID:15397162`.

Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In *Proceedings of the 31st International Conference on Distributed Computing Systems Workshops*, 2011.