

Unveiling Echo Chambers on YouTube: Analyzing Political Discourse and Social Dynamics
Through Advanced Quantitative Methods

Roman Nekrasov, Huub van de Voort, Andy Huang, Oumaima Lemhour, & Tom Teurlings

Jheronimus Academy of Data Science

Unveiling Echo Chambers on YouTube: Analyzing Political Discourse and Social Dynamics
Through Advanced Quantitative Methods

Executive Summary

This study explores the phenomenon of echo chambers on YouTube. It aims to understand how certain users or groups influence conversations and network structures, particularly in the context of politically charged content on smaller, politically focused YouTube channels. The study employs advanced quantitative methods, including Exponential Random Graph Models (ERGMs) and centrality measures, to analyze the presence and impact of prominent users in YouTube comment sections and the existence of echo chambers within these sections. The results contribute to a deeper understanding of social interactions on YouTube, addressing gaps and offering new insights into the platform's social dynamics, especially related to far-right content.

Introduction

Research on social media platforms, such as Twitter and Facebook, extensively explores echo chambers - environments where individuals connect with like-minded peers, reinforcing selective exposure to information aligning with their beliefs (Cinelli, De Francisci Morales, Galeazzi, Quattrociocchi, & Starnini, 2021). These principles, observed on social media platforms marked by informational homogeneity, apply to broader political discourse and policy debates (Jasny, Waggle, & Fisher, 2015). This suggests that the mechanisms of selective exposure observed in social media echo chambers may extend to diverse communication networks (Colleoni, Rozza, & Arvidsson, 2014). These tendencies contribute to polarization and extreme political positions (Colleoni et al., 2014). This harms social cohesion and trust, challenging finding common ground between political parties (McCoy & Somer, 2019) and shaping public discourse across diverse communication networks (Levy & Razin, 2019). Despite extensive research on platforms such as Twitter, the impact of echo chambers on YouTube, the second-largest social platform, remains understudied. YouTube's

unique structure and user interaction patterns, distinct from platforms like Twitter, may pose challenges in recognizing and understanding echo chambers on this platform. Recent research in social network analysis indicates a weaker presence of echo chambers, showing a tendency towards heterogeneity in discussions, even on controversial topics (Röchert, Neubaum, Ross, Brachten, & Stieglitz, 2020). However, prior research aggregates comment threads from numerous videos or focuses on singular, high-traffic videos, potentially overlooking the enclosed homogenous interactions typical of echo chambers. This methodological limitation risks underestimating their presence and influence on YouTube. Evidence from HosseiniMardi et al. (2020) challenges assumptions about algorithmic amplification in YouTube echo chambers, revealing them to be more subtle and insular than previously understood (HosseiniMardi et al., 2020). With the recent addition of user handles on YouTube, allowing individuals to tag each other, there has been a shift in user interactions (YouTube, 2022). This new feature holds the potential to offer insights into the formation of echo chambers, an aspect earlier research did not investigate. The identified gaps, coupled with recent changes in user interactions, necessitate a more targeted approach to precisely capture and analyze the evolving dynamics of social networks, revealing how echo chambers form and influence discourse. The concern about echo chamber effects in online communities, especially on social media platforms, is pronounced. Research by Cinelli (2021) and Du et al. (2017) emphasizes the contribution of echo chambers to heightened polarization and the formation of distinct opposing groups with online communities (Cinelli et al., 2021; Du & Gregory, 2017). Platforms like Twitter exemplify polarization attributed to echo chamber effects (Du & Gregory, 2017). This trend is alarming as social media platforms evolve from sources of entertainment to influential shapers of the social narrative. Echo chamber effects impact critical aspects of society, including policy-making, political communication and public discourse on controversial subjects (Cinelli et al., 2021). Grusauskaite, Carbone, Harambam, and Aupers (2023) underlines this concern, emphasizing social media's role in deepening societal divides and fueling extremist ideologies (Grusauskaite et al., 2023). This highlights

the importance of investigating echo chamber contexts on YouTube. The heightened significance arises from individuals perceiving videos as more credible than text, particularly in politics (Wittenberg, Tappin, Berinsky, & Rand, 2021). Given YouTube's status as the largest and most engaging online media consumption platform globally (HosseiniMardi et al., 2020), there is a pressing need to investigate echo chambers within its channels. To investigate YouTube echo chambers, our study proposes examining smaller, politically focused videos or channels with pronounced viewpoints. This approach aligns with research indicating a higher likelihood of echo chambers in political discussions (Villa, Pasi, & Viviani, 2021). Our focus on potentially more insular communities aims to uncover the true extent and nature of echo chambers on YouTube, contributing a nuanced understanding of the platform's social dynamics, especially related to far-right content (HosseiniMardi et al., 2020).

This will be done by conducting two studies:

RQ 1 *To what extent are prominent users present within YouTube channels' comment sections?* This study explores the presence of prominent users within YouTube channels' comment networks and their impact on the network structure. We hypothesize that YouTube comment sections feature influential users, drawing a parallel to the social network dynamics observed on Twitter (Sung, Moon, & Lee, 2013). This is based on the observed similarities in the structural dynamics of these platforms' social networks (Wattenhofer, Wattenhofer, & Zhu, 2012). We will use centrality measures to evaluate the extent of prominent users' presence in YouTube comment sections. These measures will provide quantitative insights into the influence and connectivity of users within the network. Conditional uniform graphs (CUGs) will be used to contextualize and validate these findings. CUGs will enable us to determine whether the centrality measures observed in the YouTube comment networks are exceptionally high or low, compared to what might be expected in a randomly structured network.

RQ 2 *To what extent are there echo chambers within YouTube channels' comment*

sections? The second study explores the existence of echo chambers within the comment sections of YouTube channels. Building upon our earlier research, which we discussed before, we hypothesize that echo chambers will be present in YouTube channels' comment sections which have a politically pronounced viewpoint. We use Exponential Random Graph Models (ERGMs) to test this hypothesis. This statistical framework allows us to examine the structural patterns of relationships in YouTube comments, providing insights into echo chambers' emergence and characteristics within these channels. By answering these research questions, our study contributes to understanding social interactions on YouTube, addressing gaps and exploring new dimensions. These insights are important to grasp how certain individuals or groups can influence conversations and network structures. Our second study investigates the presence and nature of echo chambers, thereby addressing concerns about social media's role in promoting ideological division and polarization. These findings will enhance our understanding of how certain users can initiate or reinforce echo chamber dynamics, leading to a more homogenous environment where similar viewpoints are echoed and amplified. Consequently, these individuals and dynamics can increase polarization and reduce exposure to diverse perspectives, affecting public discourse and social cohesion. In the remainder of the report, we will progress from discussing the dataset to outlining the rationale behind our chosen research methods. We then present and discuss the results of our studies and conclude with a summary of our findings, their implications and suggestions for future research.

Methodology

Dataset

A dataset was assembled by scraping the YouTube comment section. Ongehoord Nederland was selected for its unique nature as a YouTube channel, as it is known for broadcasting controversial and pronounced right-wing-oriented viewpoints. Having also compared Ongehoord Nederland to other right-wing-oriented channels in terms of channel subscribers, views and videos it is relatively a small channel. Together, this makes it an appropriate channel for our study's focus on political echo chambers, especially considering the channel's scale and content nature, which align with the methodological pivot outlined earlier. The larger the dataset, the more computationally expensive it is. Hence, we randomly selected a sequence of five videos from the channel. Rather than choosing videos randomly across the channel's history, this approach was intended to increase the chances of finding users who commented on multiple videos. Selecting sequentially adjacent videos, yet chosen randomly, allows for a more integrated analysis of user interactions and patterns, avoiding the disjointedness that might arise from sampling videos from widely separated time periods on the channel. In our dataset, nodes correspond to YouTube users engaged in the comment section using handles, which are unique and short identifiers and start with the "@" symbol. Directed edges in our network are formed based on explicit user interactions. This occurs when a user directly references another by their handle in a comment or is similarly mentioned by others. We restrict our analysis to interactions involving handles, as this clearly indicates directed communication. Comments without such specific mentions are excluded, as we cannot conclusively determine if they are responses to the main thread or to specific individuals, thus ensuring clarity and precision in our network analysis.

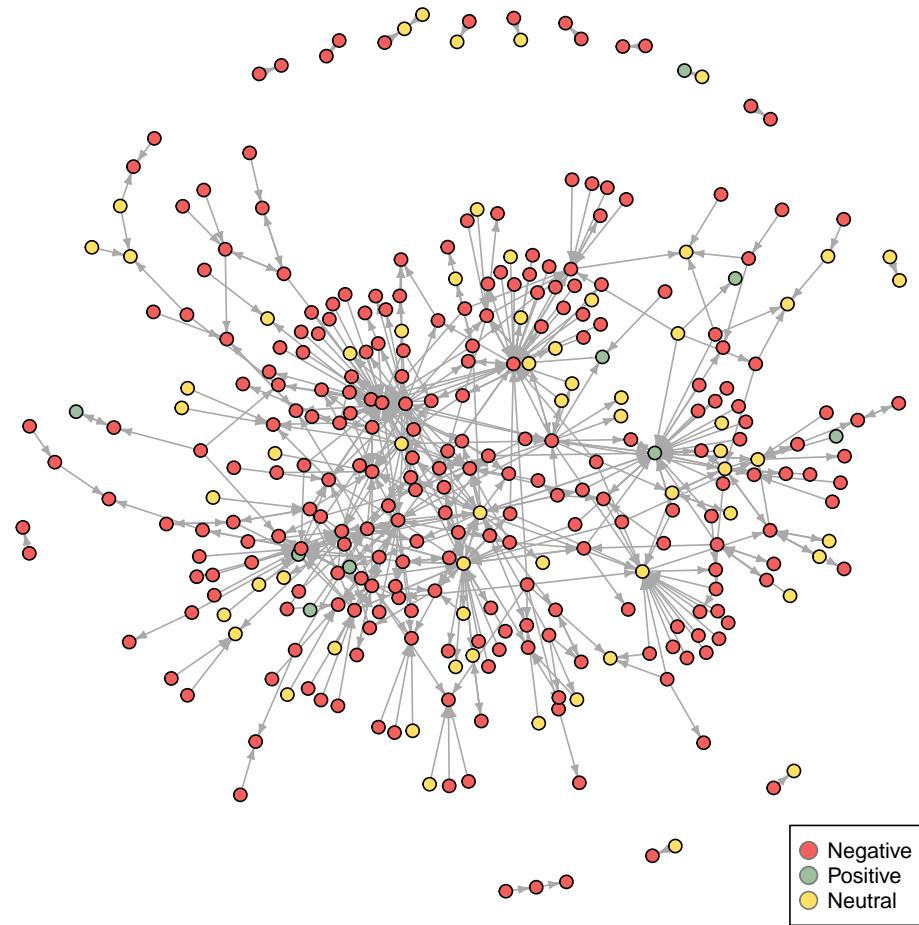


Figure 1. The plot of the network used for analysis. It shows the majority of the nodes labelled negative.

Furthermore, our data was initially a weighted network due to the potential for users to comment multiple times on each other's posts, adding complexity to our analysis.

However, to simplify and focus on the presence of interactions rather than their frequency, we convert this into an unweighted network. This means we consider whether an interaction occurred between users, regardless of how often it happened. Ultimately, we ended up with a network of 324 distinct actors interconnected through 553 ties, derived from comments posted between July 2023 and October 2023. This network is characterized by a notably low indegree and outdegree, as detailed in Figures 2, 3, pointing to its sparse connectivity.

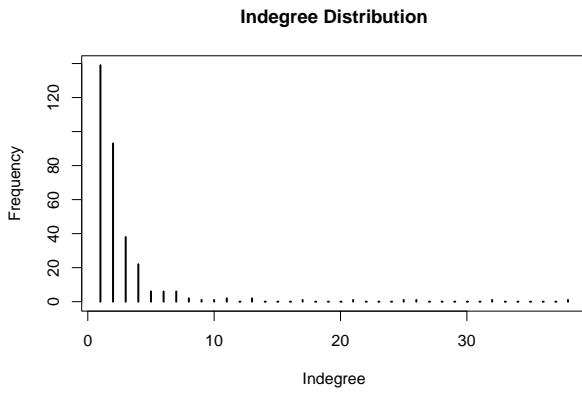


Figure 2. The plot above shows the indegree distribution for the actors in the network.

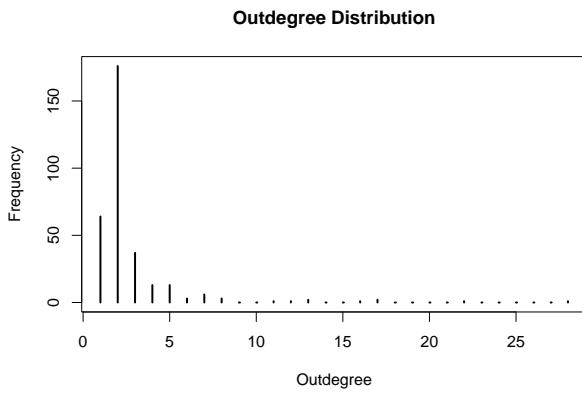


Figure 3. The plot above shows the outdegree distribution for the actors in our network.

Additionally, Table 1 reveals that reciprocal interactions within this network are infrequent, while unidirectional interactions are more prevalent. The visual representation of the network, depicted in Figure 1, further highlights the sparse connectivity as we see a lot of isolated dyads. In this Figure 1 a small subset of actors with significantly higher degrees of connectivity can also be found, an observation which underscores the relevance of our first study. The network descriptives can be found in Appendix A.

Table 1

*An overview of the reciprocal
dyads.*

Mutual	Asymmetric	Null
70	413	51843

Potential Bias

Our research methodology, while comprehensive, is not without potential biases that could impact the interpretation of our results. The primary source of bias stems from Ongehoord Nederland as the sole YouTube channel for analysis. This channel's distinct right-wing orientation and ties to specific political ideologies may not represent the broader YouTube community or other political spectrums. Therefore, our findings, while indicative of echo chambers within this particular context, may not be generalizable to other YouTube channels or political orientations. Another potential bias arises from our method of selecting videos. Although we chose a sequence of five videos randomly to increase the likelihood of finding users commenting on multiple videos, this approach might have inadvertently focused on a specific time frame or subset of topics, which may not encompass the full range of discussions typically found on the channel. This could limit the diversity of user interactions

and viewpoints captured in our study. It is crucial to acknowledge that these biases could influence the conclusions drawn from our study. The results should be interpreted with an understanding of these limitations and future research should aim to include a more diverse range of channels, video selection methods and user interaction types to provide a more comprehensive view of echo chambers on YouTube. This will help better understand online discourse's nuances and dynamics, especially in politically charged environments.

Research Rationale

A shared understanding of ‘echo chambers’ within social networks is essential to address our research questions effectively. Building upon Jasny et al. (2015) ‘s framework, we recognize echo chambers as social structures marked by significant homophily, where individuals sharing similar views are grouped together (Jasny et al., 2018, 2015; Jasny & Fisher, 2019). This concept is encapsulated in the ‘chamber’, represented as a network configuration known as a transitive triad. In this configuration, information circulates from one actor (A) to another (C) both directly and via an intermediary (B), illustrating multiple paths of information flow within the same source and endpoint. In this ‘chamber,’ a speaker, a mediating actor and a receiver are involved. Information in this configuration can travel directly from the speaker to the receiver or from the speaker through the mediating actor to the receiver. The similarity in content and sentiment of communications between actors with aligned opinions and outlooks identifies the ‘echo’ aspect. This phenomenon is visualized in Figure 4.

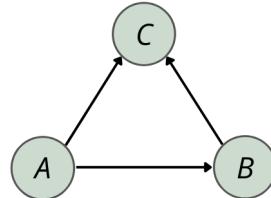


Figure 4. An echo chamber representation within the structural network configuration, also known as a transitive triad.

To assess the sentiment of a comment within a discussion section, we employed ChatGPT for zero-shot sentiment analysis. This analysis will categorize the comments of individual users as aligning with, being neutral toward, or opposing the political subject in question. ChatGPT will be used for sentiment analysis because it has shown remarkable zero-shot sentiment analysis proficiency, rivaling fine-tuned models such as BERT and

state-of-the-art models trained with labeled data within their respective domains (Wang, Xie, Ding, Feng, & Xia, 2023). The sentiment focuses on the users' attitudes towards specific political entities, parties, politicians, policy issues and broader socio-cultural themes. This sentiment was categorized as positive, negative, or neutral, incorporating a nuanced understanding of the complexities and diversities inherent in the topics discussed:

1. **Positive Sentiment:** This indicates approval, support, or positive feelings towards a subject. For instance, a comment expressing hope or confidence in a political movement is classified as positive.
2. **Negative Sentiment:** This reflects dissatisfaction, disapproval, or negative emotions. Comments critical of political parties, policy decisions, or societal issues such as gender identity and LGBTQ+ topics are examples of negative sentiment.
3. **Neutral Sentiment:** This denotes an objective, impartial, or moderate stance, devoid of strong positive or negative emotions, often informative in nature.

This approach enables us to develop a more detailed understanding of public perception and sentiment about the different topics discussed in the comments. This assists in understanding and analyzing the complexity of public opinions and emotions in relation to these diverse subjects. In assessing our hypothesis for our first research question we use CUGs because we cannot guarantee the normality assumptions required by many statistical tests. CUGs allows us to test our hypotheses despite the statistical complexities of the network representation. We use CUGs to determine whether centrality measures of our observed graph occur at levels exceeding what we would expect by chance. With centrality measures such as eigenvector centrality, degree centrality and betweenness centrality we aim to identify whether there are central players in our network. Eigenvector centrality is used to recognize influential nodes based not just on their direct connections, but more significantly on their connections to other highly connected and central nodes. Degree centrality is used

to highlight nodes that have more connections, providing a measure of popularity. Betweenness centrality is included to identify nodes that act as bridges within the network structure, controlling the flow of information between other nodes. For future research, methodologies such as the Multiple Regression Quadratic Assignment Procedure (MRQAP) may prove valuable for exploring how various attributes — like subscriber count, number of likes, or number of comments — are associated with centrality measures. Such an approach could contribute to our understanding of how certain attributes predict a node's likelihood of being a central figure. Echoing the approach of Jasny et al. (2015), which focused on operationalizing the components of echo chambers in the US climate policy network, our second study also employs ERGMs (Jasny et al., 2015). ERGMs are ideal for this type of analysis due to its ability to comprehensively model complex social networks. It integrates both endogenous network effects and exogenous variables into a unified statistical framework, accurately capturing the intricate dynamics of social interactions (Cranmer, Desmarais, & Morgan, 2020). This is important because this allows us to effectively analyze groups which are unified by shared sentiments. Additionally, ERGM's approach of treating the network as a single multivariate observation aligns well with our objective of empirically testing the structure and impact of echo chambers.

Results

Study 1

In our first study, we aimed to investigate if the centralization observed in YouTube comment networks — measured through eigenvector, degree and betweenness centrality — was higher than what would be expected in a random scenario. This was analyzed using Conditional Uniform Graphs (CUGs), conditioned on the network's number of edges.

Table 2

P-values resulting from CUG Tests.

centralization measure	$\Pr(X \geq \text{Obs})$
betweenness	0.97
indegree	0.00
outdegree	0.00
eigenvector	0.19

The CUG results, presented in Table 2, provide insight into the network structure in terms of these centrality measures. For both indegree and outdegree centrality, the findings were significant. A significant indegree centrality indicates that certain users in the network receive more interactions than others, highlighting the presence of particularly influential or attention-gathering users. Regarding outdegree centrality, the findings point to certain users being more active in initiating interactions or making connections within the network. However, the situation is different for eigenvector centrality. Despite its high empirical value of 0.92, the result was not statistically significant. A high empirical value typically indicates a tendency for influential users within the network to be interconnected, suggesting the presence of clusters where influential users are more likely to be connected with each other. However, the non-significant result in this case means that such a tendency of influential users being connected is not different from what could happen in a random network. The betweenness centrality also showed a non-significant result. This means that the network's

nodes do not play a central role as intermediaries in the information flow compared to a random network. This indicates that while there are nodes that help in spreading information, their role is similar to what might occur randomly.

Study 2

In this study, we analyse network formation based on the sentiment of user comments. Figure 5 presents the result of the ERGM from the best fitting model.

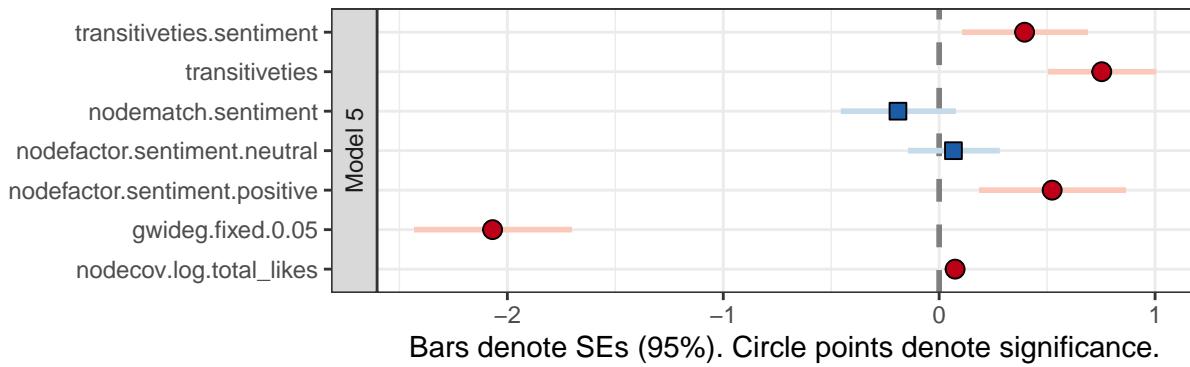


Figure 5. ERGM results for model 5.

In our initial attempts to fit the model, we encountered challenges with achieving convergence. Our approach replicated the methodology used by Jasny et al. (2015), who also faced similar convergence issues. They resolved this by imposing constraints on the indegree distribution, a strategy we adopted as well. Implementing this constraint enabled our models to converge more effectively. However, it's important to note that we excluded the edges term from the model specification, as it is unfeasible to estimate it while constraining the indegree distribution. This adjustment alters the interpretation of our results, as the ERGM now shows how network structures deviate from what would be expected given the fixed indegree distribution. We included a term for homophily, *nodematch.sentiment*, to represent our ‘echo’. However, with a negative coefficient of -0.19 translating to a log odds of approximately 0.82, this term suggests that sharing the same sentiment inversely influences the formation of ties

between two nodes. In other words, nodes with matching sentiments are slightly less likely to form ties than those with differing sentiments. We also included transitive triads as a term, with *transitivities*, to represent the ‘chamber’. The positive coefficient of 0.75 for *transitivities*, translating to log odds of about 2.11, indicates a strong influence of transitive relationships in tie formation, irrespective of sentiment. This suggests that the presence of a transitive triad significantly increases the likelihood of tie formation within the network. The full echo chamber mechanism is represented in the term *transitivities.sentiment*, with a positive coefficient of 0.4 and a corresponding log odds of approximately 1.49. The significance of both *transitivities.sentiment* and *transitivities* implies that the formation of transitive relationships is not due to chance and is influenced by the sentiment attribute. The positive coefficients here indicate a tendency for tie formation in scenarios where transitive relationships have similar sentiments. It is also worth mentioning that *nodefactor.sentiment.positive* is significant in our model, with a value of 0.52 translating to log odds of approximately 1.68. This finding highlights that comments characterized by positive sentiment are more inclined to establish connections within the network. Lastly, the term *nodecov.log.total_like* is also significant with a coefficient of 0.07, which translates to log odds of approximately 1.07. This indicates that the total likes of an user is a significant predictor for the formation of ties within the network. This finding underscores the role of engagement, as measured by likes, in facilitating connections between nodes in the network.

Discussion and Conclusion

Our study aimed to investigate the existence of echo chambers on YouTube, a platform less studied in this context compared to others like Twitter. It also focuses on if there are prominent users as these could initiate or reinforce the echo chamber effects, which in turn contributes to polarization and ideological homogeneity. Our research has uncovered significant insights that not only enhance our theoretical understanding but also offer practical guidance for various stakeholders. Our approach was centered around two key research questions:

RQ 1 *To what extent are prominent users present within YouTube channels' comment sections?*

RQ 2 *To what extent are there echo chambers within YouTube channels' comment sections?*

Study 1 revealed that while some users are more prominent or active in the network (as shown by the indegree and indegree centrality), these users are not more likely to be connected to each other (eigenvector centrality) and their role as information brokers (betweenness centrality) are not significantly different from what might be seen in a random network. This suggests that while there are prominent users in terms of activity, they do not form a cohesive group of influencers or dominate the information brokerage within the network. Looking at the in degree centrality plot (figure plot) it's noticeable there relative high amount of non-negative sentiment nodes . In the discourse environment of the comment section this could possibly indicate an hostile environment for having a different opinion. In Study 2, we found evidence supporting the existence of echo chambers on YouTube. The formation of transitive ties was also significant, indicating that these are not mere products of randomness. Important to note is that, contrary to the findings of Jasny et al. (2015), our study observed a lack of significant homophily. The lack of significant homophily within the

YouTube comment networks, particularly in contrast to the homogeneous environment dominated by negative sentiments, could be partially explained by the insights from Study 1. As the central nodes do not seem to have an influential role in the network and the sentiments of nodes with high degrees seem to. Another probable possibility could be due to the widespread negative sentiment in our network, making it difficult to observe homophily because most users share this sentiment. The lack of diverse sentiments in our network reduces the chances of forming distinct groups based on sentiment. Although this suggests our network is largely a single, homogeneous cluster, it does not rule out the possibility of smaller clusters forming based on other factors besides sentiment. It does suggest that one could describe our network as one big echo, in which chambers are being formed. The findings contribute significantly to the theoretical understanding of social media dynamics as our research extends the concept of echo chambers beyond traditional platforms like Twitter and Facebook, demonstrating their presence in the unique context of YouTube. This contributes to the broader discourse on how digital platforms facilitate ideological segregation and polarization. Understanding the role of prominent users and the dynamics of echo chambers is crucial for a range of stakeholders, including policymakers, educators and content creators. The study's findings can inform strategies to promote more balanced and diverse online conversations, thereby potentially mitigating the effects of polarization. These strategies could involve strengthening community guidelines, moderating content to prevent extremism, raising awareness about echo chambers and empowering users to diversify their content exposure. For policymakers, these insights underscore the need for policies that foster diversity in online discourse and prevent platforms like YouTube from becoming hotbeds of extremism or misinformation. Educators and content creators can leverage these findings to promote critical thinking and digital literacy, equipping audiences to engage constructively in online environments. The scope for future research is extensive and includes expanding the study to encompass a broader range of YouTube channels and political perspectives, conducting robustness checks for sentiment analysis and delving

deeper into the characteristics of prominent users. Such research could provide a more nuanced understanding of user interactions and opinions, as discussed in Chapter 2, further enriching our knowledge of the dynamics at play in online echo chambers. It is also interesting to further research users with a high indegree as these individuals are likely to be influenced by the incoming interactions from other users, making them pivotal in the spread and reinforcement of ideas within echo chambers. Utilizing Temporal Exponential Random Graph Models could offer a dynamic view of how sentiments change over time within these networks. This approach could help identify users who are more susceptible to influence, thereby enabling the development of strategies to foster a more balanced and diverse discourse. In conclusion, our research has revealed the complex dynamics of echo chambers on YouTube, highlighting the significant role of prominent users and the trend towards ideological homogeneity in the comment sections of politically oriented videos. These findings are critical in understanding the unique characteristics of social media platforms and their impact on public discourse and societal polarization. As digital platforms continue to evolve, the insights gained from our study can offer a roadmap for stakeholders in the digital space, aiding in the development of more informed and effective strategies to manage and comprehend the impact of social media on society. This study not only contributes to the academic discourse but also serves as a catalyst for action among policymakers, educators and content creators, steering us towards a more informed, balanced and cohesive digital world.

References

- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118.
- Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in twitter using big data. *Journal of Communication*, 64(2), 317–332.
- Cranmer, S. J., Desmarais, B. A., & Morgan, J. W. (2020). *Inferential network analysis*. Cambridge University Press.
- Du, S., & Gregory, S. (2017). The echo chamber effect in twitter: Does community polarization increase? *Complex Networks & Their Applications v: Proceedings of the 5th International Workshop on Complex Networks and Their Applications (COMPLEX NETWORKS 2016)*, 373–378. Springer.
- Grusauskaite, K., Carbone, L., Harambam, J., & Aupers, S. (2023). Debating (in) echo chambers: How culture shapes communication in conspiracy theory networks on YouTube. *New Media & Society*, 14614448231162585.
- HosseiniMardi, H., Ghasemian, A., Clauset, A., Rothschild, D. M., Mobius, M., & Watts, D. J. (2020). Evaluating the scale, growth, and origins of right-wing echo chambers on YouTube. *arXiv Preprint arXiv:2011.12843*.
- Jasny, L., Dewey, A. M., Robertson, A. G., Yagatich, W., Dubin, A. H., Waggle, J. M., & Fisher, D. R. (2018). Shifting echo chambers in US climate policy networks. *PloS One*, 13(9), e0203463.
- Jasny, L., & Fisher, D. R. (2019). Echo chambers in climate science. *Environmental Research Communications*, 1(10), 101003.
- Jasny, L., Waggle, J., & Fisher, D. R. (2015). An empirical examination of echo chambers in US climate policy networks. *Nature Climate Change*, 5(8), 782–786.
- Levy, G., & Razin, R. (2019). Echo chambers and their effects on economic and political

- outcomes. *Annual Review of Economics*, 11, 303–328.
- McCoy, J., & Somer, M. (2019). Toward a theory of pernicious polarization and how it harms democracies: Comparative evidence and possible remedies. *The Annals of the American Academy of Political and Social Science*, 681(1), 234–271.
- Röchert, D., Neubaum, G., Ross, B., Brachten, F., & Stieglitz, S. (2020). Opinion-based homogeneity on YouTube: Combining sentiment and social network analysis. *Computational Communication Research*, 2(1), 81–108.
- Sung, J., Moon, S., & Lee, J.-G. (2013). The influence in twitter: Are they really influenced? *International Workshop on Behavior and Social Informatics and Computing*, 95–105. Springer.
- Villa, G., Pasi, G., & Viviani, M. (2021). Echo chamber detection and analysis: A topology-and content-based approach in the COVID-19 scenario. *Social Network Analysis and Mining*, 11(1), 78.
- Wang, Z., Xie, Q., Ding, Z., Feng, Y., & Xia, R. (2023). Is ChatGPT a good sentiment analyzer? A preliminary study. *arXiv Preprint arXiv:2304.04339*.
- Wattenhofer, M., Wattenhofer, R., & Zhu, Z. (2012). The YouTube social network. *Proceedings of the International AAAI Conference on Web and Social Media*, 6, 354–361.
- Wittenberg, C., Tappin, B. M., Berinsky, A. J., & Rand, D. G. (2021). The (minimal) persuasive advantage of political video over text. *Proceedings of the National Academy of Sciences*, 118(47), e2114388118.
- YouTube. (2022). *Announcing YouTube handles: A unique identifier to discover and connect on YouTube*. <https://support.google.com/youtube/thread/183284252>.

Appendix A

This first appendix includes all of the R chunks of code that were hidden throughout the document to help with reproducibility readability and/or setup.

R set-up code

```
library("papaja")
r_refs("r-references.bib")

knitr::opts_chunk$set(fig.pos = "H", out.extra = "",
                      echo = FALSE, cache = TRUE)

setwd("../data_preparation/prepared_data")

graph <- igraph::read_graph(
  file = "network.graphml",
  format = "graphml"
)

# Loading environments with our data

setwd("../..../data_analysis")

load("model_results.RData")

# Seed for random number generation

set.seed(42)

knitr::opts_chunk$set(cache.extra = knitr::rand_seed)
```

R code for plots and tables

```
# Code for plotting the network

igraph::V(graph)$color <- "black"

igraph::V(graph) [sentiment == 'negative']$color <- "#F25F5C"
```

```

igraph::V(graph) [ sentiment == 'positive' ]$color <- "#9DBF9E"
igraph::V(graph) [ sentiment == 'neutral' ]$color <- "#FFE066"

plot(
  graph,
  vertex.label = NA,
  edge.arrow.size = .35,
  vertex.size = 3,
  vertex.color = igraph::V(graph)$color
)

colrs <- c("#F25F5C", "#9DBF9E", "#FFE066")
graphics::legend(x = 0.75, y = -.85, c("Negative", "Positive", "Neutral"),
                 pch = 21, col = "#777777", pt.bg = colrs, pt.cex = 1.5,
                 cex = .8, bty = "o", ncol = 1)

# Code for plotting the indegree distribution
degree_distribution_in <- snafun::g_degree_distribution(graph,
                                                       mode = "in",
                                                       type = "count"
)
# Create a frequency plot with lines
plot(1:length(degree_distribution_in), degree_distribution_in, type = "h",
      lwd = 2, main = "Indegree Distribution",
      xlab = "Indegree", ylab = "Frequency")

```

```
# Code for plotting the outdegree distribution

degree_distribution_out <- snafun::g_degree_distribution(graph,
                                                       mode = "out",
                                                       type = "count"
)

# Create a frequency plot with lines

plot(1:length(degree_distribution_out), degree_distribution_out, type = "h",
      lwd = 2, main = "Outdegree Distribution",
      xlab = "Outdegree", ylab = "Frequency")

# code for plotting the dyad count table

dyad_count <- snafun::count_dyads(graph, echo = FALSE)

apa_table(dyad_count,
          placement = "!h",
          caption = "(ref:dyad-count-table)")

# Reading file from figures directory

knitr::include_graphics(
  "SNA4DS_Report_files/figures/transitive_triad_2.png"
)

apa_table(cug_results,
          placement = "h",
          caption = "(ref:cug-results-table)")
```

Appendix B

```
centralities_plot <- snafun::plot_centralities(
  graph,
  measures = c("betweenness", "closeness", "degree", "eccentricity"),
  directed = TRUE,
  mode = "all",
  k = 3,
  rescaled = FALSE
)
```

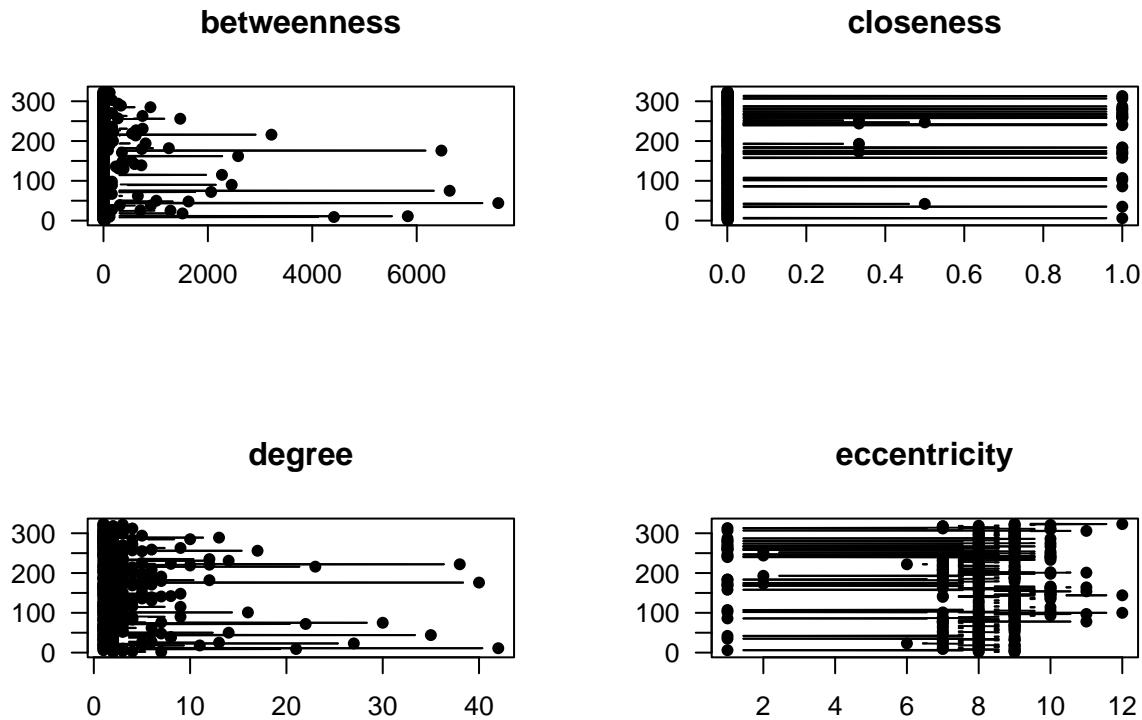


Figure 6. The centrality scores for each node in our network plotted.

```
# Plotting indegree for each node

# Coloring each node based on its sentiment

igraph::V(graph)$color <- "orange"

igraph::V(graph) [ sentiment == 'negative' ]$color <- "#F25F5C"
igraph::V(graph) [ sentiment == 'positive' ]$color <- "#9DBF9E"
igraph::V(graph) [ sentiment == 'neutral' ]$color <- "orange"

my_df <- data.frame(
  igraph::get.vertex.attribute(graph, name = "id"),
  igraph::get.vertex.attribute(graph, name = "color"),
  1:324
)

v_degree <- igraph::degree(graph, mode = "in")

my_df <- cbind(my_df, v_degree)

names(my_df) <- c("id", "color", "new_id", "degree")

plot(y = my_df$new_id,
      x = my_df$degree,
      xlab = "indegree",
      ylab ="",
      col = my_df$color,
      type = "b",
      pch = 16)
```

```
# Adding a legend

legend("topright", # Location of the legend. Change as needed.

       legend = c("Negative: 253",
                  "Positive: 6",
                  "Neutral: 62"), # Labels in the legend

       col = c("#F25F5C",
                  "#9DBF9E",
                  "orange"), # Colors corresponding to labels

       pch = 16)
```

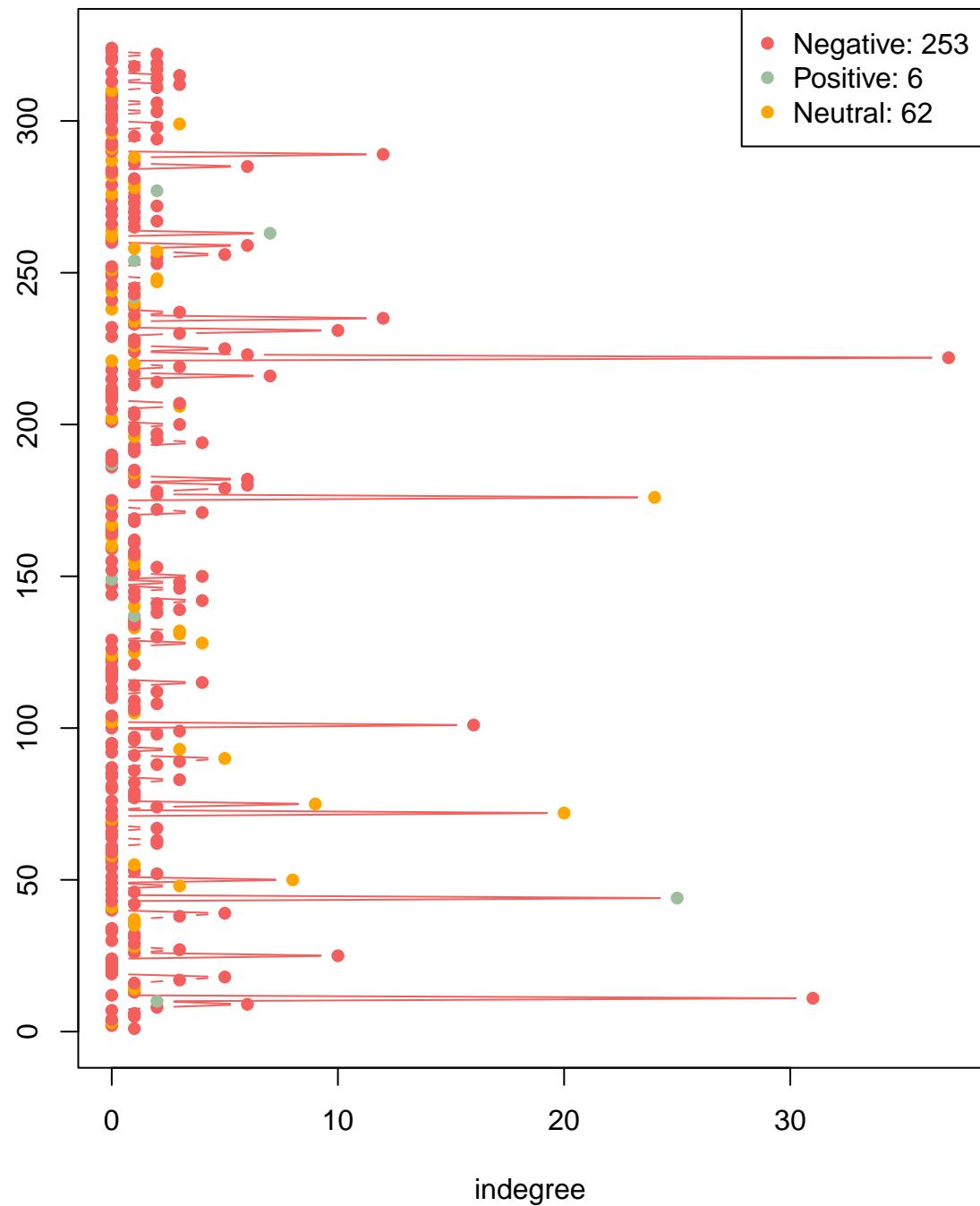


Figure 7. The indegree plotted for each node in the network

```
# Plotting outdegree for each node

# Coloring each node based on its sentiment

igraph::V(graph)$color <- "orange"

igraph::V(graph) [ sentiment == 'negative' ]$color <- "#F25F5C"

igraph::V(graph) [ sentiment == 'positive' ]$color <- "#9DBF9E"

igraph::V(graph) [ sentiment == 'neutral' ]$color <- "orange"

my_df <- data.frame(
  igraph::get.vertex.attribute(graph, name = "id"),
  igraph::get.vertex.attribute(graph, name = "color"),
  1:324
)

v_degree <- igraph::degree(graph, mode = "out")

my_df <- cbind(my_df, v_degree)

names(my_df) <- c("id", "color", "new_id", "degree")

plot(y = my_df$new_id,
      x = my_df$degree,
      xlab = "outdegree",
      ylab ="",
      col = my_df$color,
      type = "b",
      pch = 16)
```

```
# Add a legend  
legend("topright", # Location of the legend  
      legend = c("Negative: 253",  
                "Positive: 6",  
                "Neutral: 62"), # Labels in the legend  
      col = c("#F25F5C",  
              "#9DBF9E",  
              "orange"), # Colors corresponding to labels  
      pch = 16)
```

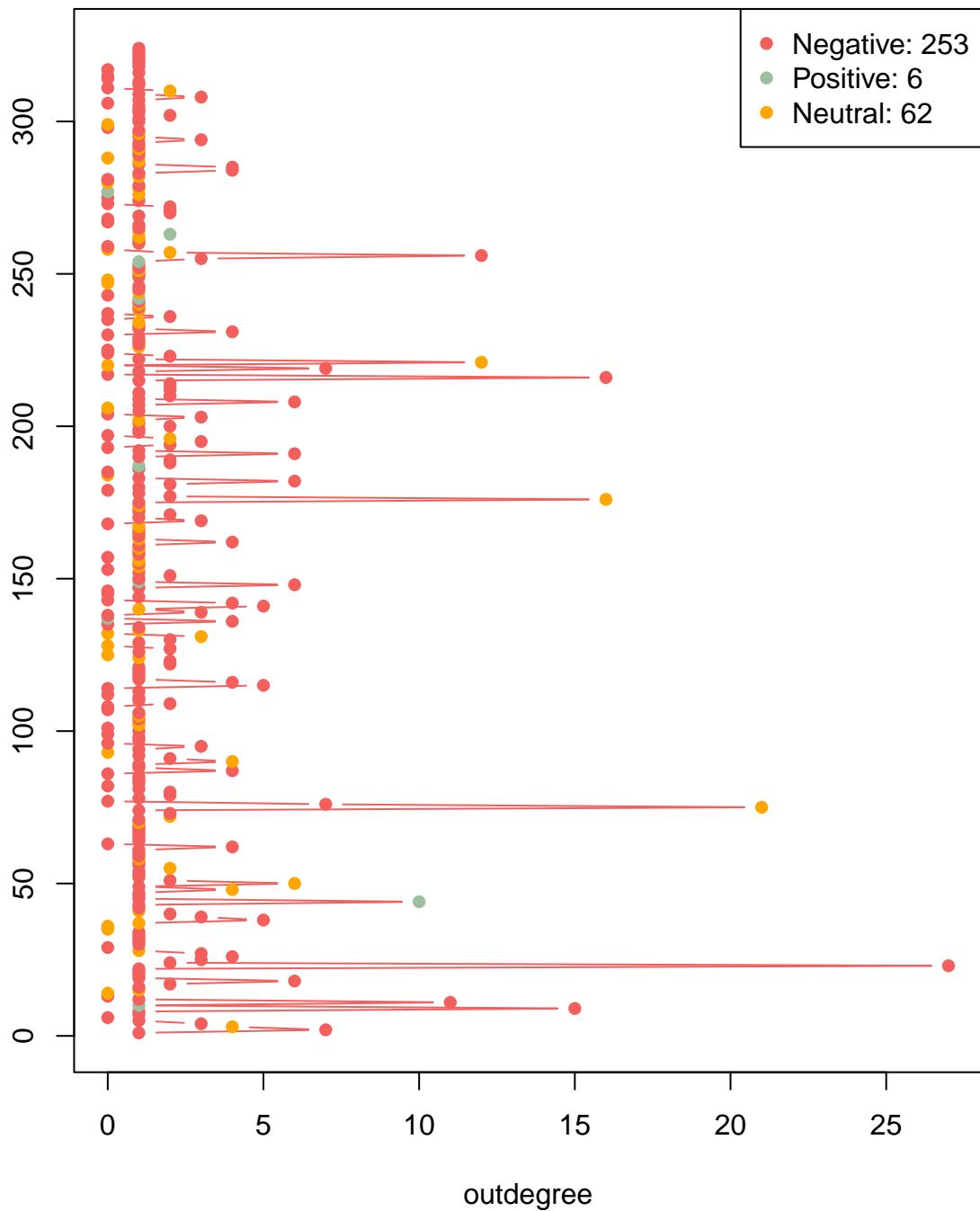


Figure 8. The outdegree plotted for each node in the network

ERGM Results

We display the results of the different models we ran as well as the diagnostics for model 5.

```
texreg:::plotreg(models)
```

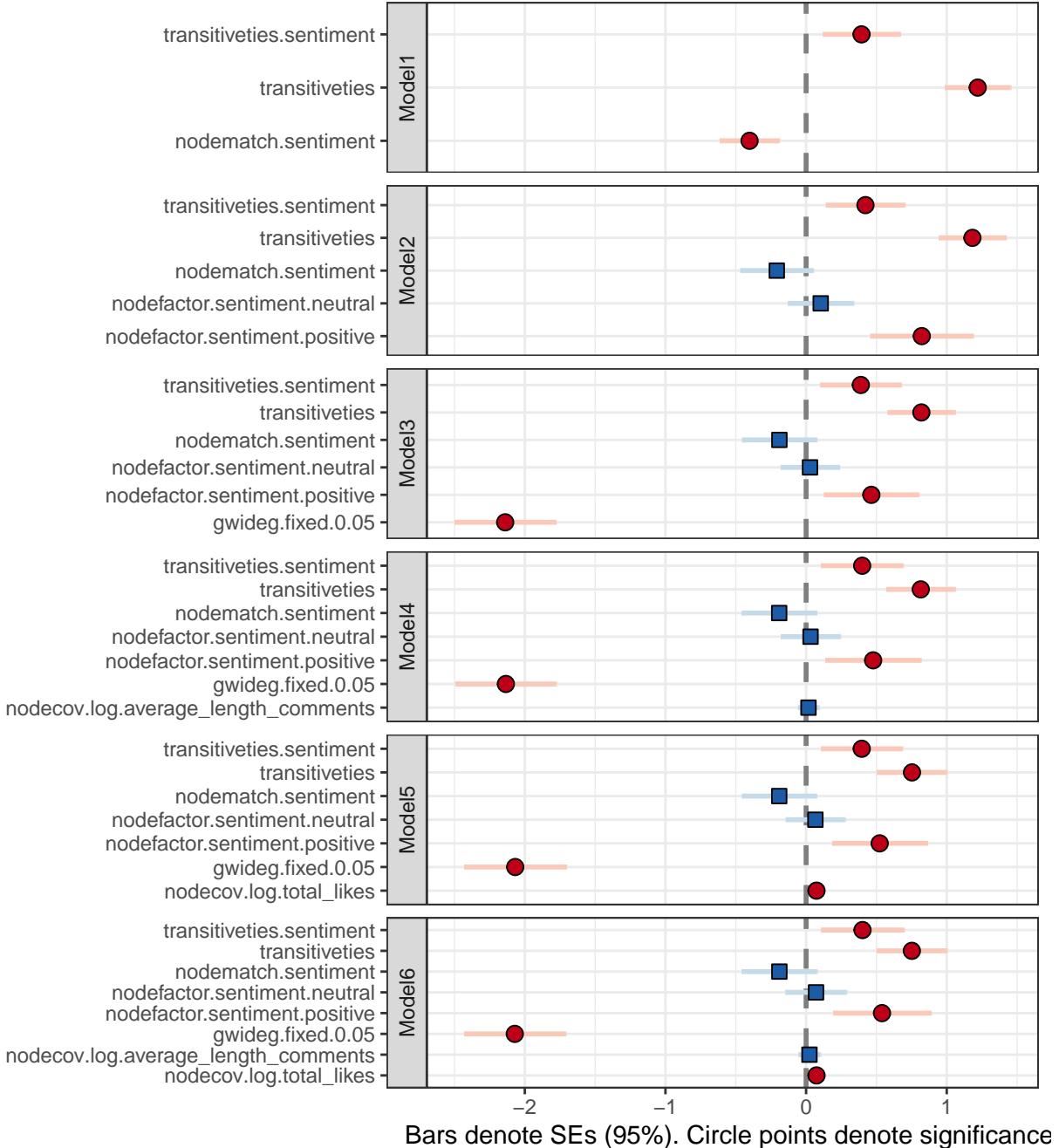


Figure 9

Model Diagnostics of model 5

```
# Plotting MCMC diagnostic
model_5 <- models$model_5
par(mai=c(.6,0.3,.4,0.3), omi=c(.1,.1,.1,.1)) # setting margins due to error

ergm::mcmc.diagnostics(
  model_5,
  which = c("plots")
)
```

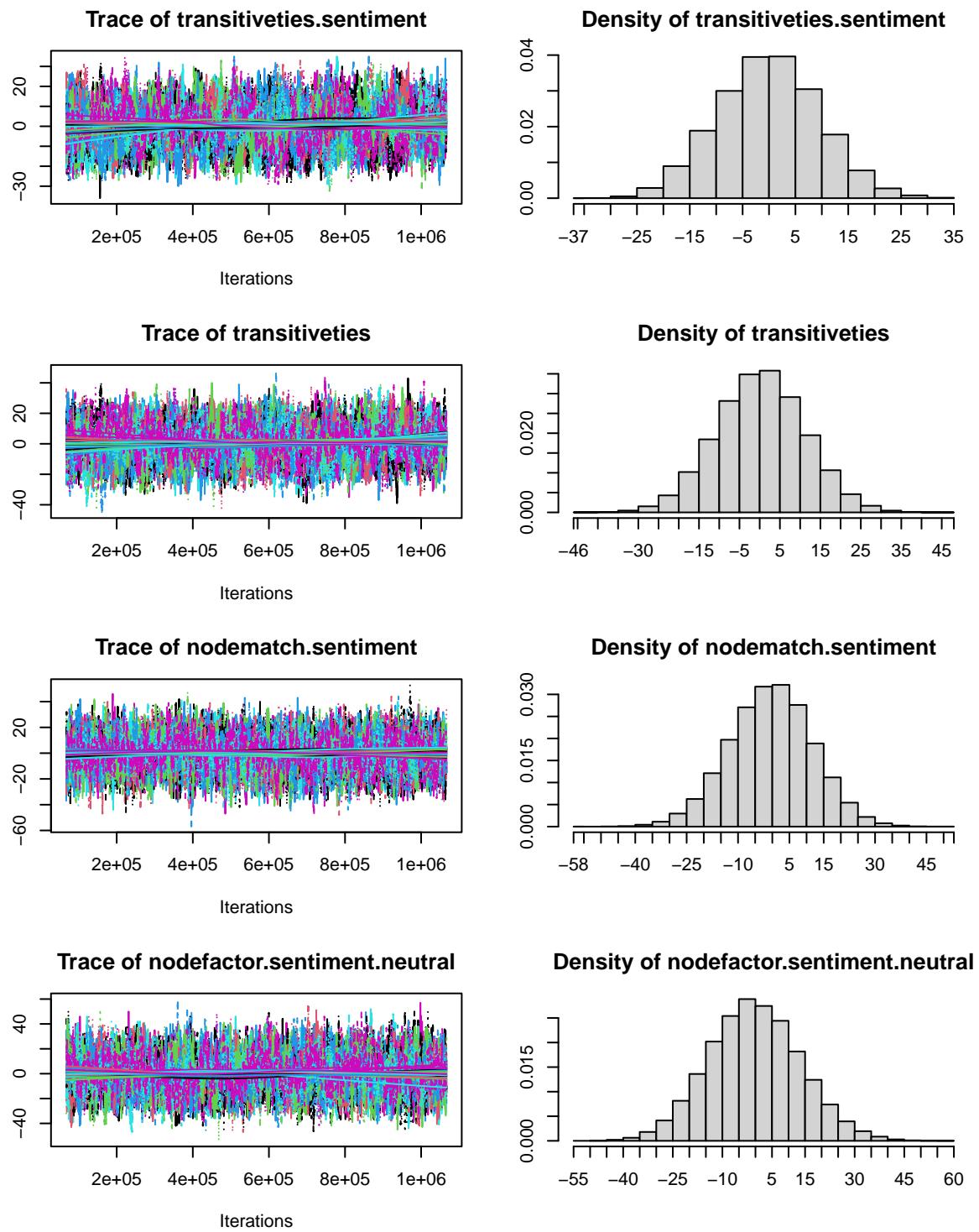


Figure 10. The MCMC diagnostics for model 5.

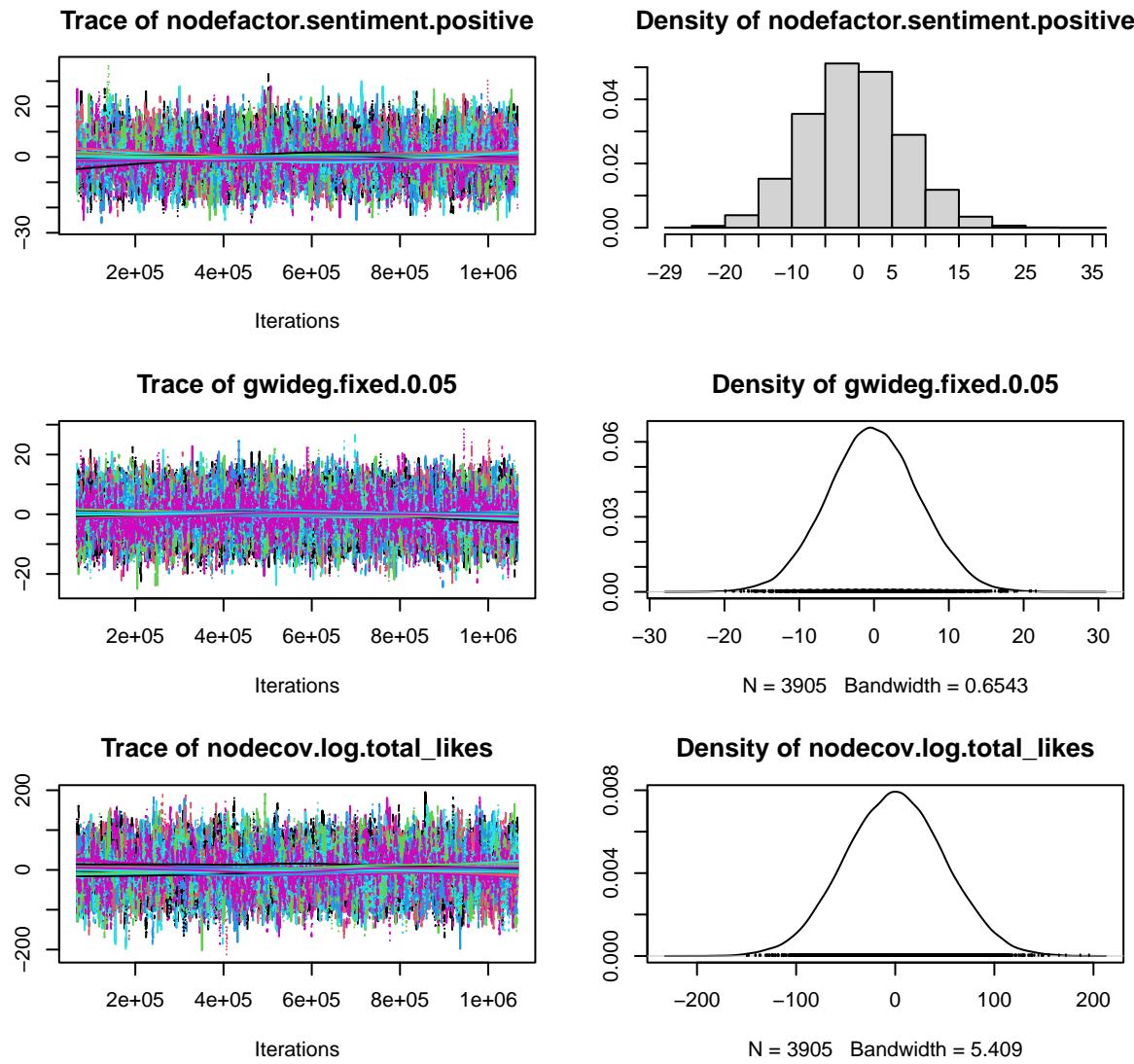


Figure 11. The MCMC diagnostics for model 5.

Note: MCMC diagnostics shown here are from the last round of simulation, prior to computation of final parameter estimates. Because the final estimates are refinements of those used for this simulation run, these diagnostics may understate model performance. To directly assess the performance of the final model on in-model statistics, please use the GOF command: `gof(ergmFitObject, GOF=~model)`.

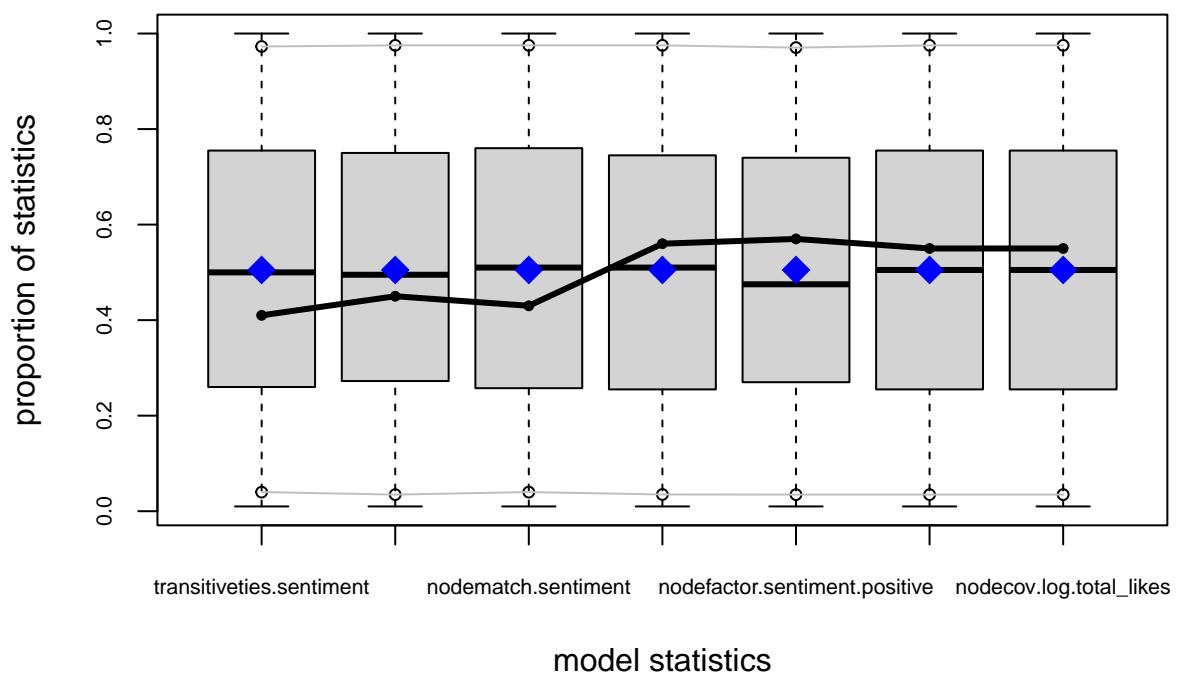
Goodness of Fit of model 5

Figure 12. The plot below shows the Goodness of Fit statistics for model 5

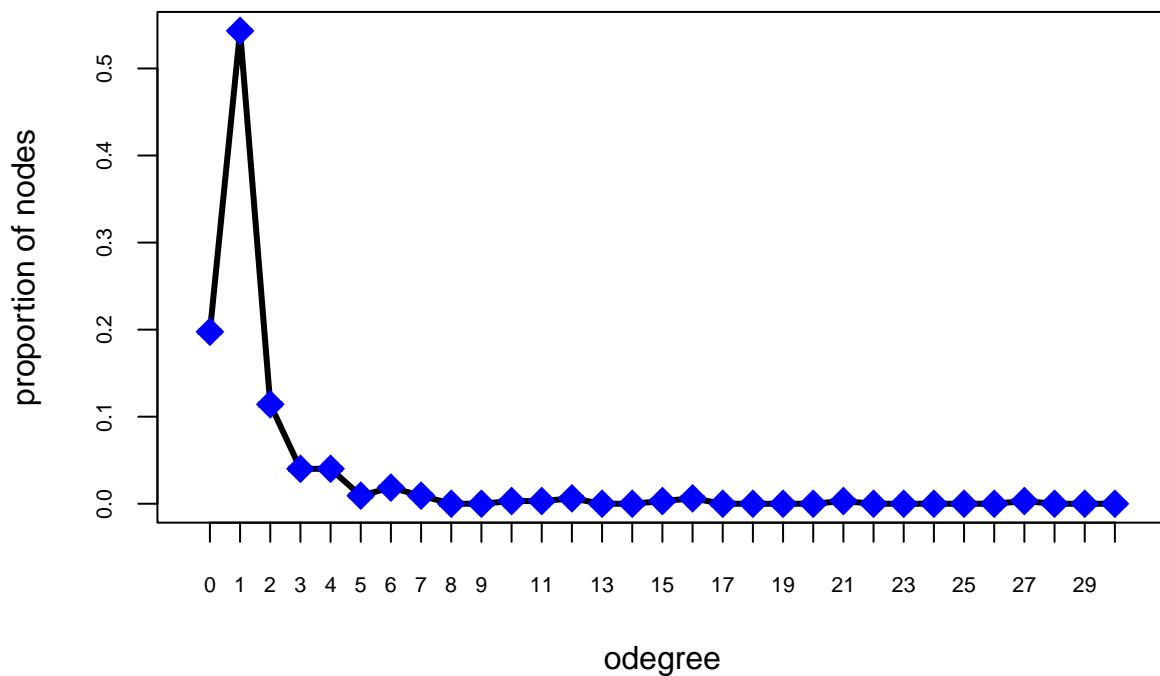


Figure 13. The plot below shows the Goodness of Fit statistics for model 5

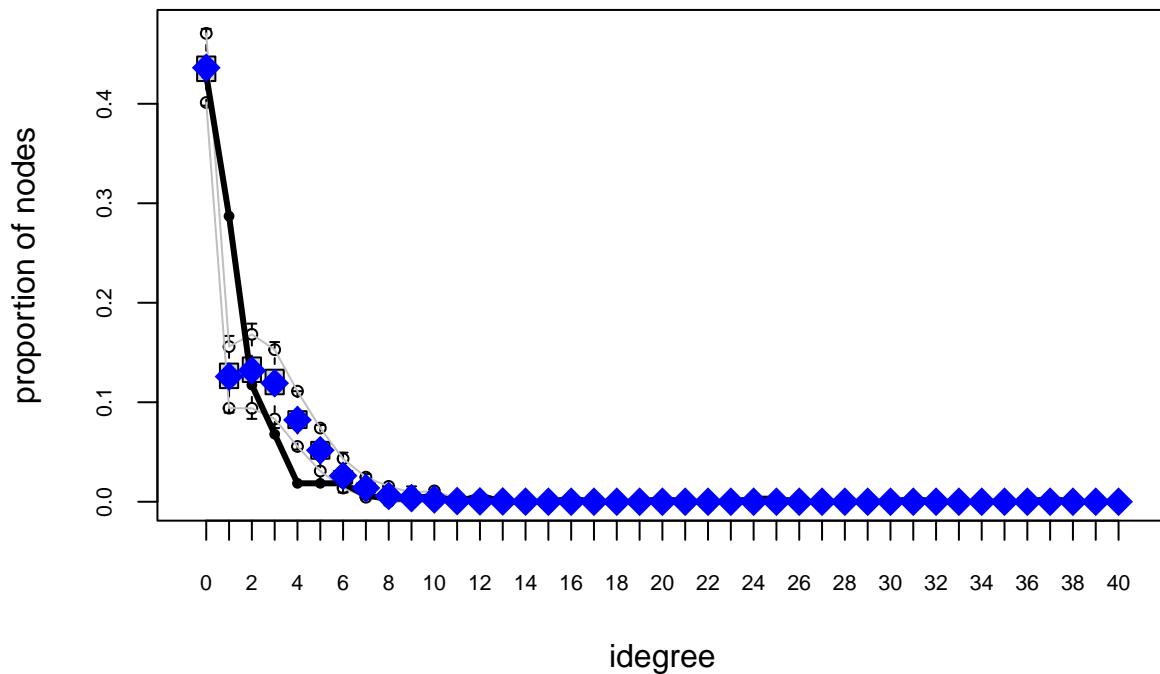


Figure 14. The plot below shows the Goodness of Fit statistics for model 5

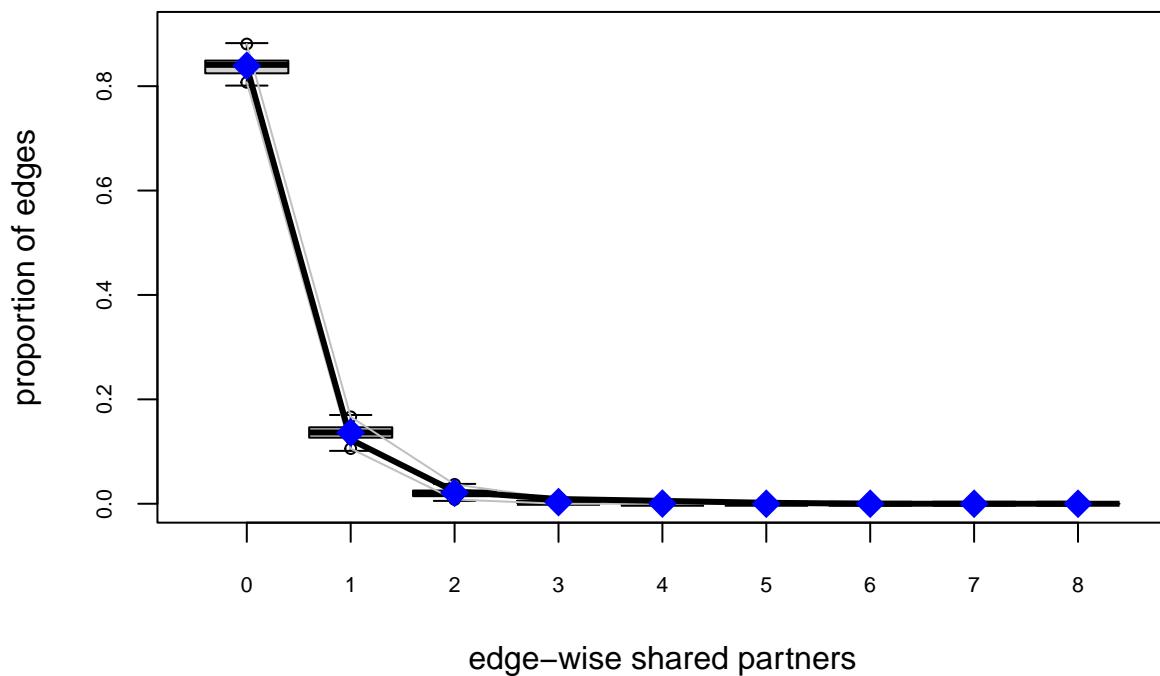


Figure 15. The plot below shows the Goodness of Fit statistics for model 5

Goodness-of-fit diagnostics

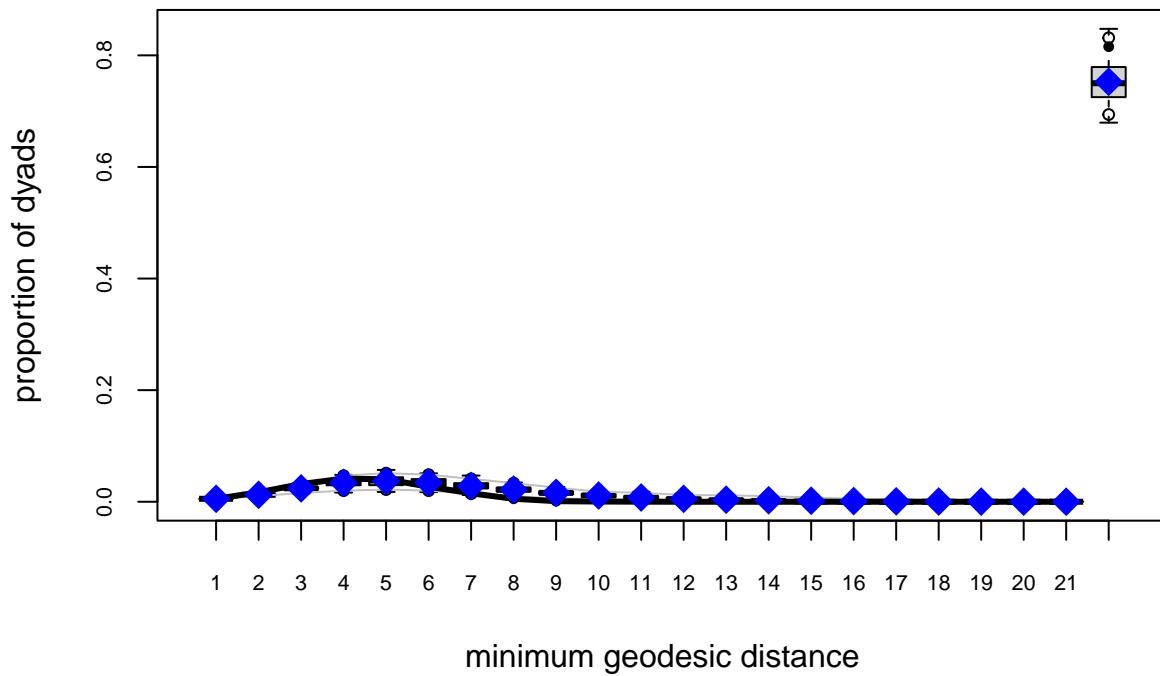


Figure 16. The plot below shows the Goodness of Fit statistics for model 5