Supplementary Material for *Sentiment and structure in word co-occurrence networks on Twitter*

Mikaela Irene Fudolig,[1,*] Thayer Alshaabi,[2,1] Michael V. Arnold,[1]
Christopher M. Danforth,[1,3] and Peter Sheridan Dodds[1,4]

[1] *Computational Story Lab, Vermont Complex Systems Center,*
MassMutual Center of Excellence for Complex Systems and Data Science,
Vermont Advanced Computing Core, University of Vermont, Burlington, VT, USA
[2] *Advanced Bioimaging Center, UC Berkeley, Berkeley, CA, USA*
[3] *Department of Mathematics & Statistics, University of Vermont, Burlington, VT, USA*
[4] *Department of Computer Science, University of Vermont, Burlington, VT, USA*

A. Removing hashtag hijacking

When we performed community detection on all the tweets, we found a community with words unrelated to Hillary Clinton or the elections, such as "santa" and "christmas". Upon checking the raw tweets with these words, we found that they all contained the hashtag `#smptweettest`. All tweets with this hashtag were removed in the dataset analyzed in the manuscript.

B. Most frequently used words removed in network backboning

Table S1 shows the words removed in the initial backboning step by comparing the words that consistently made it to the top 400 1-grams in Twitter for 100 random days and the top 200 case-insensitive 1-grams in each of the "favor", "against", and "all" corpora.

Table S1: Words from the intersection of the top 400 1-grams in Twitter for 100 random days, and whether or not they were also present in the most frequently occurring 200 case-insensitive words in each of the "favor", "against", and "all" corpora. Note that the top 400 1-grams from Twitter is case-sensitive and includes non-alphabet symbols, while the words in this table are case-insensitive and only contain characters from the English alphabet. Words with a check mark were removed in the initial backboning step.

word	FAVOR	AGAINST	ALL	**word**	FAVOR	AGAINST	ALL
a	✓	✓	✓	**many**		✓	✓
about	✓	✓	✓	**me**	✓	✓	✓
after	✓	✓	✓	**miss**			
again	✓	✓	✓	**more**	✓	✓	✓
all	✓	✓	✓	**morning**			
always				**most**	✓	✓	✓
am	✓		✓	**much**	✓	✓	✓
an	✓	✓	✓	**my**	✓	✓	✓
and	✓	✓	✓	**need**	✓	✓	✓
any		✓	✓	**never**	✓	✓	✓
are	✓	✓	✓	**new**	✓	✓	✓
as	✓	✓	✓	**next**	✓		
at	✓	✓	✓	**night**	✓		
back	✓	✓	✓	**no**	✓	✓	✓
bad		✓		**not**	✓	✓	✓

* mikaela.fudolig@uvm.edu

be	✓	✓	✓
because	✓	✓	✓
been	✓	✓	✓
before	✓		
being	✓	✓	✓
best	✓		
better	✓		
but	✓	✓	✓
by	✓	✓	✓
can	✓	✓	✓
cant			
come	✓		✓
could	✓	✓	✓
day	✓		✓
days			
did	✓	✓	✓
do	✓	✓	✓
doing			
done	✓	✓	✓
dont			
down	✓	✓	✓
even	✓	✓	✓
ever	✓	✓	✓
every		✓	✓
everyone	✓		
feel	✓		
find			
first	✓		✓
follow			
for	✓	✓	✓
friends			
from	✓	✓	✓
fuck			
get	✓	✓	✓
getting			
girl			
give			
go	✓	✓	✓
going	✓	✓	✓
gonna	✓		
good	✓	✓	✓
got	✓	✓	✓
great	✓		✓
guys			
had	✓	✓	✓
happy			
hard			
has	✓	✓	✓
hate	✓		
have	✓	✓	✓
he	✓	✓	✓
her	✓	✓	✓
here	✓	✓	✓
him	✓	✓	✓
his	✓	✓	✓
home			
hope	✓		✓
how	✓	✓	✓
i	✓	✓	✓
if	✓	✓	✓

now	✓	✓	✓
of	✓	✓	✓
off	✓	✓	✓
on	✓	✓	✓
one	✓	✓	✓
only	✓	✓	✓
or	✓	✓	✓
other			
our	✓	✓	✓
out	✓	✓	✓
over	✓	✓	✓
people	✓	✓	✓
please	✓		✓
real		✓	✓
really	✓	✓	✓
right	✓	✓	✓
rt		✓	
said	✓	✓	✓
same		✓	
say	✓	✓	✓
see	✓	✓	✓
she	✓	✓	✓
shit			
should	✓	✓	✓
so	✓	✓	✓
some	✓	✓	✓
someone			
something			
still	✓	✓	✓
stop		✓	✓
take	✓	✓	✓
tell			
than	✓	✓	✓
thanks			
that	✓	✓	✓
thats	✓		
the	✓	✓	✓
their	✓	✓	✓
them	✓	✓	✓
then	✓	✓	✓
there	✓	✓	✓
these	✓	✓	✓
they	✓	✓	✓
thing		✓	✓
things			
think	✓	✓	✓
this	✓	✓	✓
time	✓	✓	✓
to	✓	✓	✓
today	✓		✓
tomorrow	✓		✓
tonight	✓		✓
too	✓	✓	✓
tweet			
u	✓	✓	✓
up	✓	✓	✓
us	✓	✓	✓
very			
via	✓	✓	✓
wait			

ill				**wanna**			
im				**want**	✓	✓	✓
in	✓	✓	✓	**was**	✓	✓	✓
into		✓		**watch**	✓		✓
is	✓	✓	✓	**way**	✓	✓	✓
it	✓	✓	✓	**we**	✓	✓	✓
its				**well**	✓	✓	✓
ive				**were**	✓	✓	✓
just	✓	✓	✓	**what**	✓	✓	✓
keep	✓	✓	✓	**when**	✓	✓	✓
know	✓	✓	✓	**where**		✓	✓
last	✓		✓	**who**	✓	✓	✓
let	✓	✓	✓	**why**	✓	✓	✓
life				**will**	✓	✓	✓
like	✓	✓	✓	**with**	✓	✓	✓
little				**work**	✓		
lol				**world**	✓		✓
look		✓	✓	**would**	✓	✓	✓
love	✓		✓	**year**			
made	✓		✓	**you**	✓	✓	✓
make	✓	✓	✓	**your**	✓	✓	✓
man				**youre**	✓	✓	✓

C. Comparison with null models

We include here the histograms (Figures S1-S12) corresponding to Figures 2-4 for the null models using the configuration network model, the Erdos-Renyi model, the shuffled score model, and the uniform score model (Section 2.4).

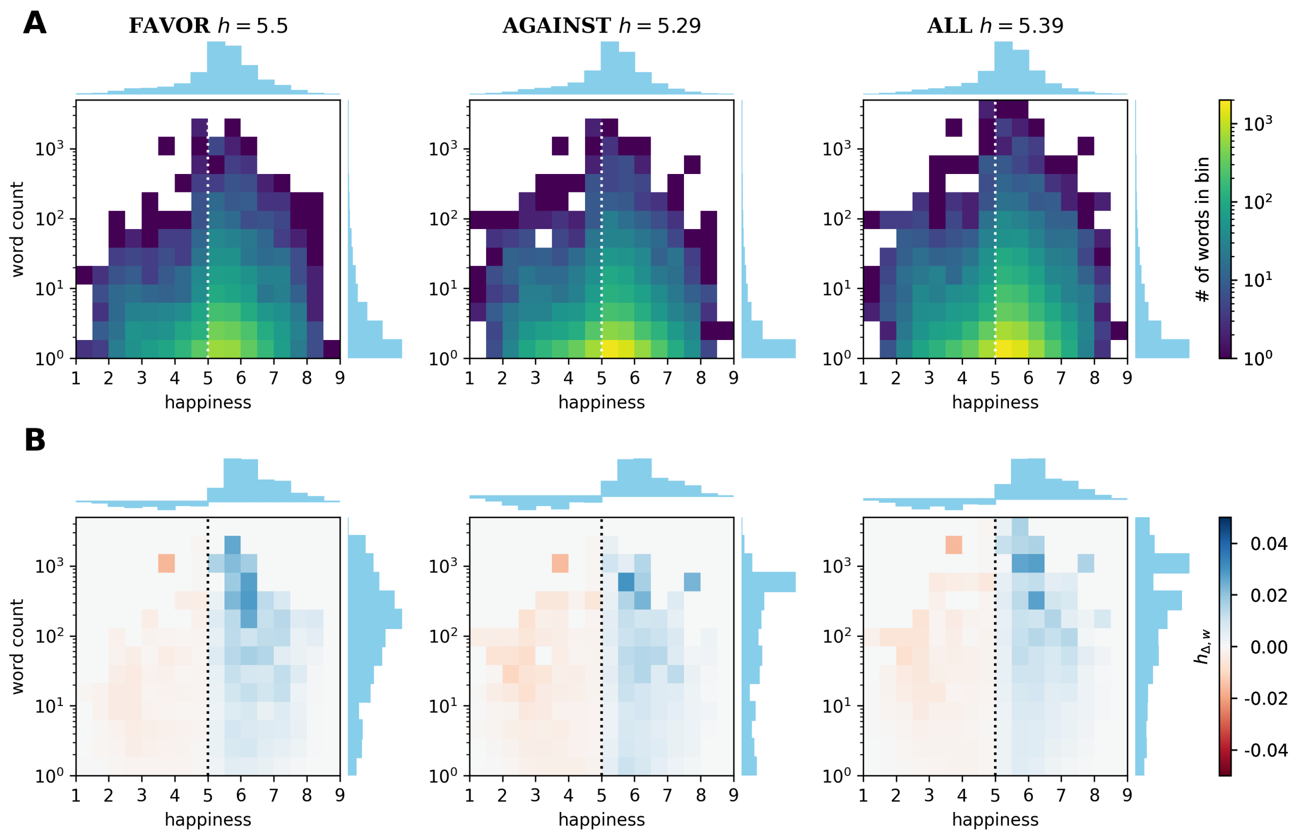

Figure S1. (a) The 2D histogram for word count vs. happiness score, with the corresponding marginal distributions shown (note that each word has a weight of 1) for the **configuration model**. (b) A 2D histogram of the contributions of words in word count-happiness space to the deviation from neutrality, $h_{\Delta,w} = (h_w - 5) * N_w / \sum_{w'} N_{w'}$, where h_w is the word's happiness score of word and N_w is the number of times the word appears in the corpus. The marginal distributions are also included. Vertical lines at $h = 5$ are added to guide the eye.

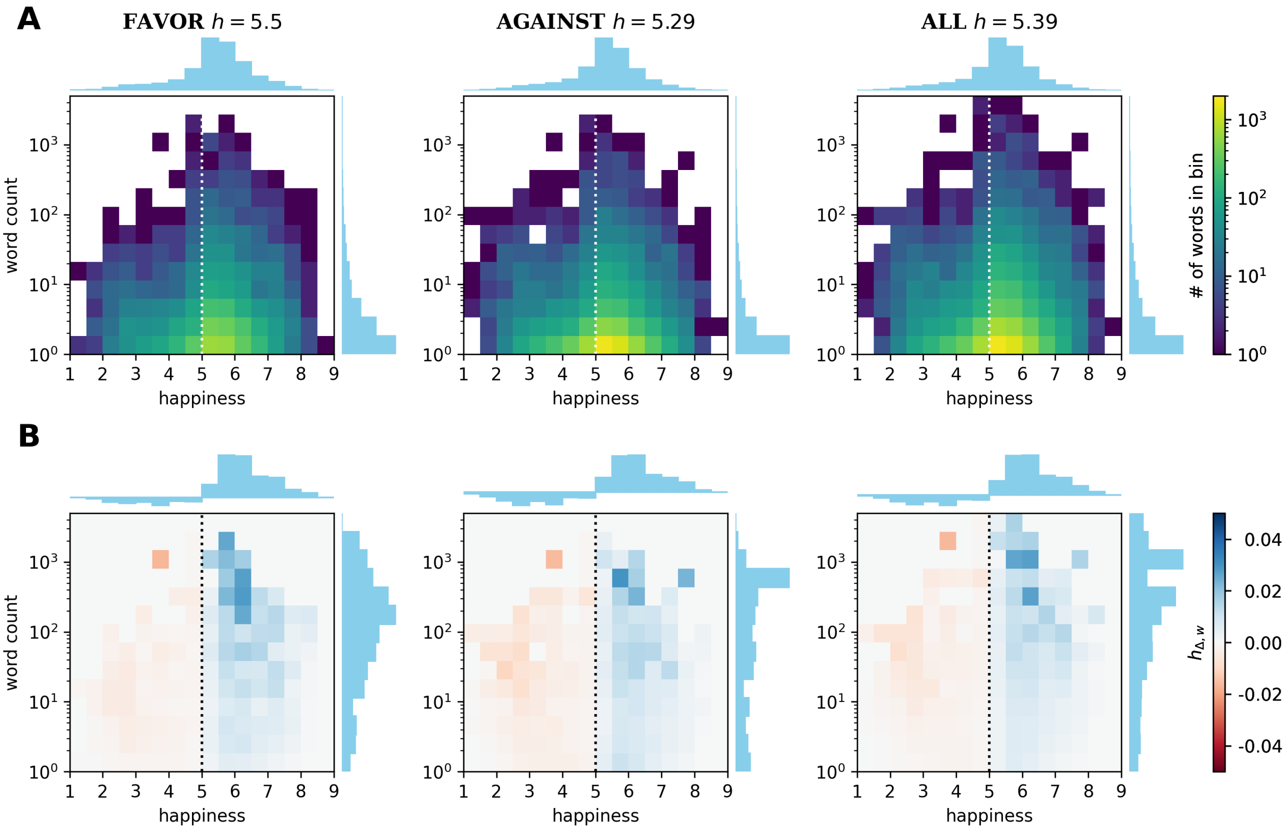

Figure S2. (a) The 2D histogram for word count vs. happiness score, with the corresponding marginal distributions shown (note that each word has a weight of 1) for the **Erdos-Renyi null model**. (b) A 2D histogram of the contributions of words in word count-happiness space to the deviation from neutrality, $h_{\Delta,w} = (h_w - 5) * N_w / \sum_{w'} N_{w'}$, where h_w is the word's happiness score of word and N_w is the number of times the word appears in the corpus. The marginal distributions are also included. Vertical lines at $h = 5$ are added to guide the eye.

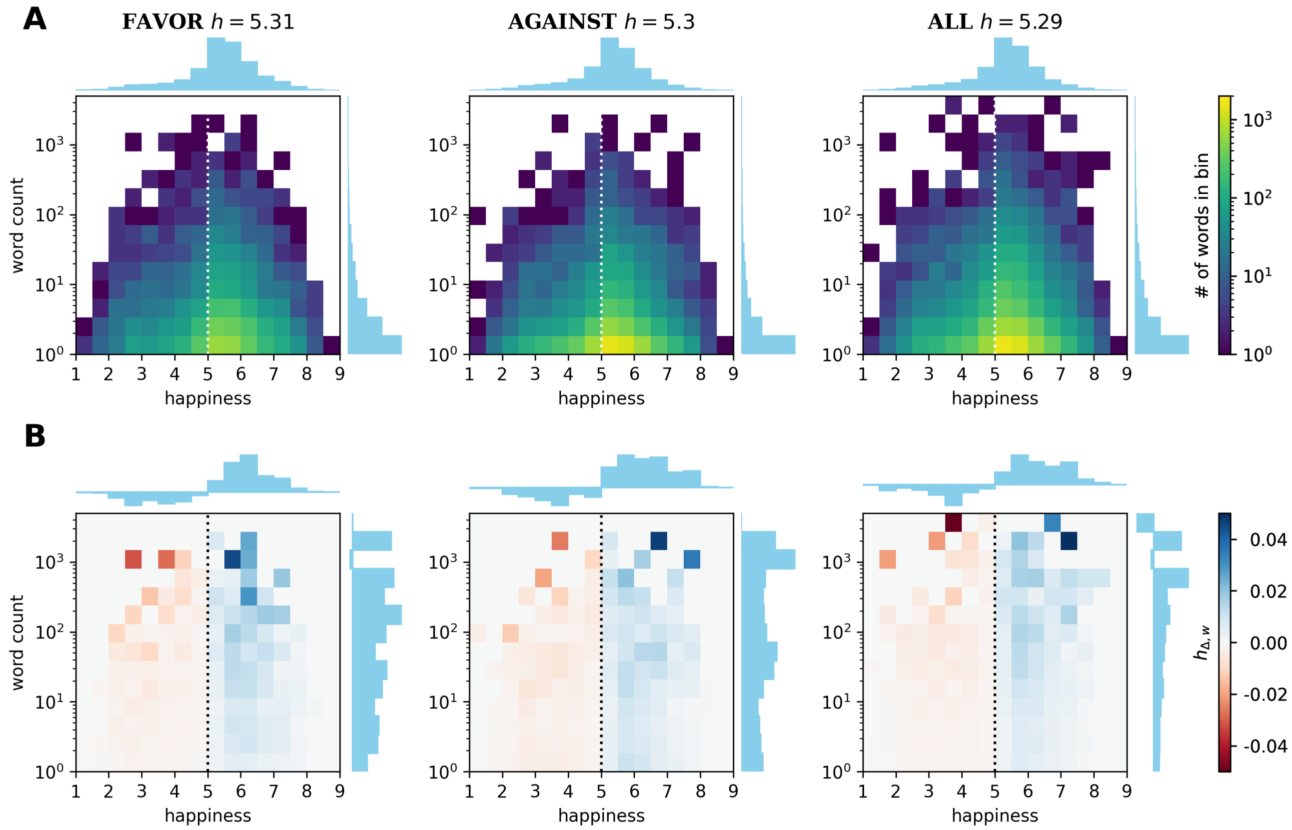

Figure S3. (a) The 2D histogram for word count vs. happiness score, with the corresponding marginal distributions shown (note that each word has a weight of 1) for the **shuffled score model**. (b) A 2D histogram of the contributions of words in word count-happiness space to the deviation from neutrality, $h_{\Delta,w} = (h_w - 5) * N_w / \sum_{w'} N_{w'}$, where h_w is the word's happiness score of word and N_w is the number of times the word appears in the corpus. The marginal distributions are also included. Vertical lines at $h = 5$ are added to guide the eye.

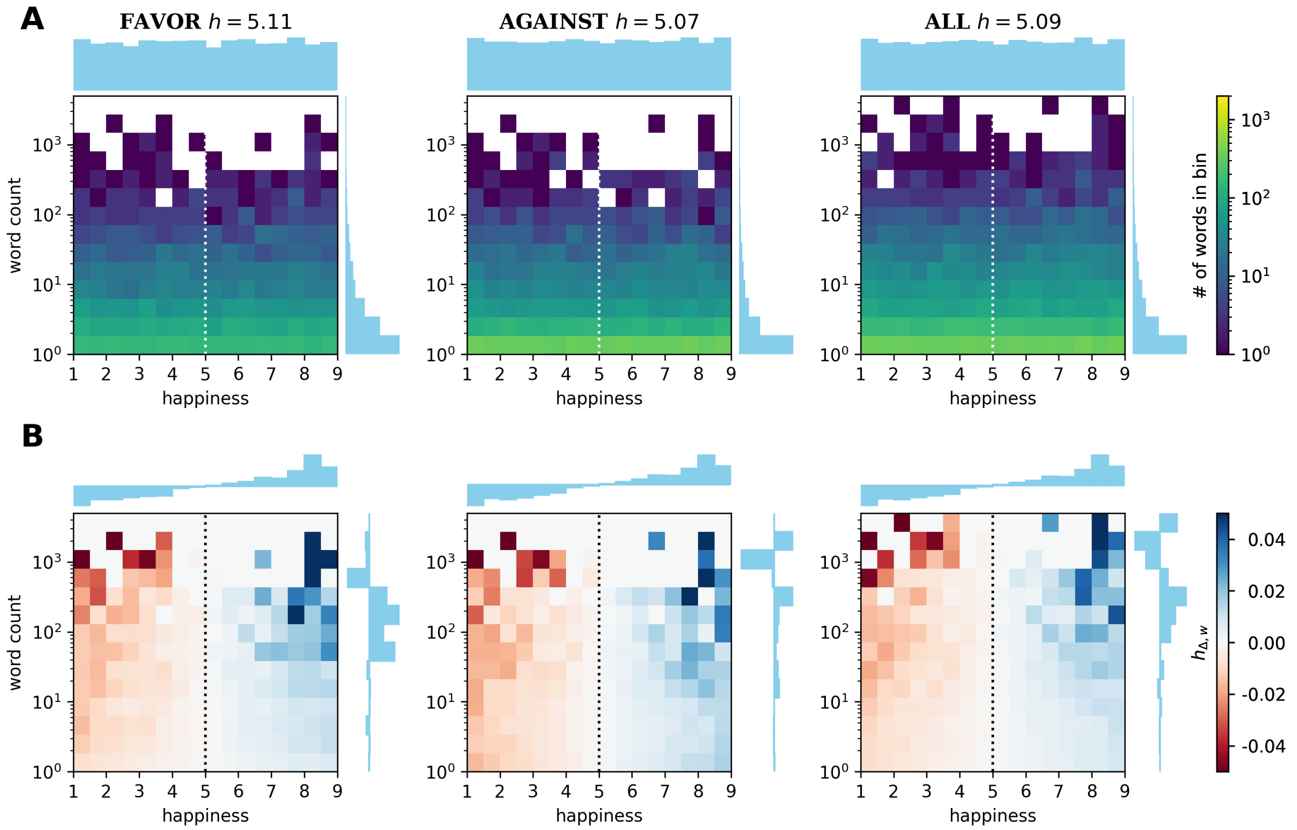

Figure S4. (a) The 2D histogram for word count vs. happiness score, with the corresponding marginal distributions shown (note that each word has a weight of 1) for the **uniform score model**. (b) A 2D histogram of the contributions of words in word count-happiness space to the deviation from neutrality, $h_{\Delta,w} = (h_w - 5) * N_w / \sum_{w'} N_{w'}$, where h_w is the word's happiness score of word and N_w is the number of times the word appears in the corpus. The marginal distributions are also included. Vertical lines at $h = 5$ are added to guide the eye.

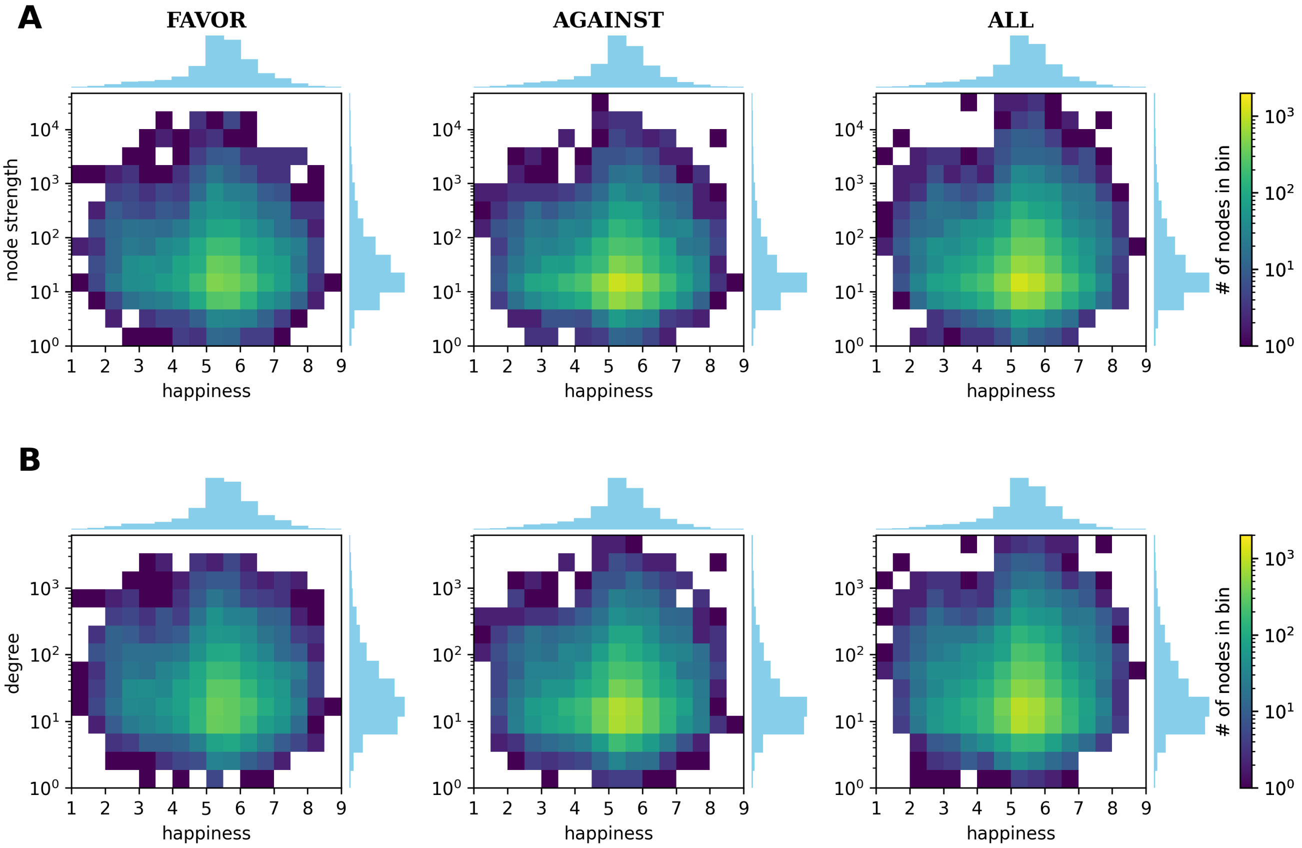

Figure S5. 2D histograms for both the node strengths (top row) and degree (bottom row) vs. happiness scores for the **configuration model**.

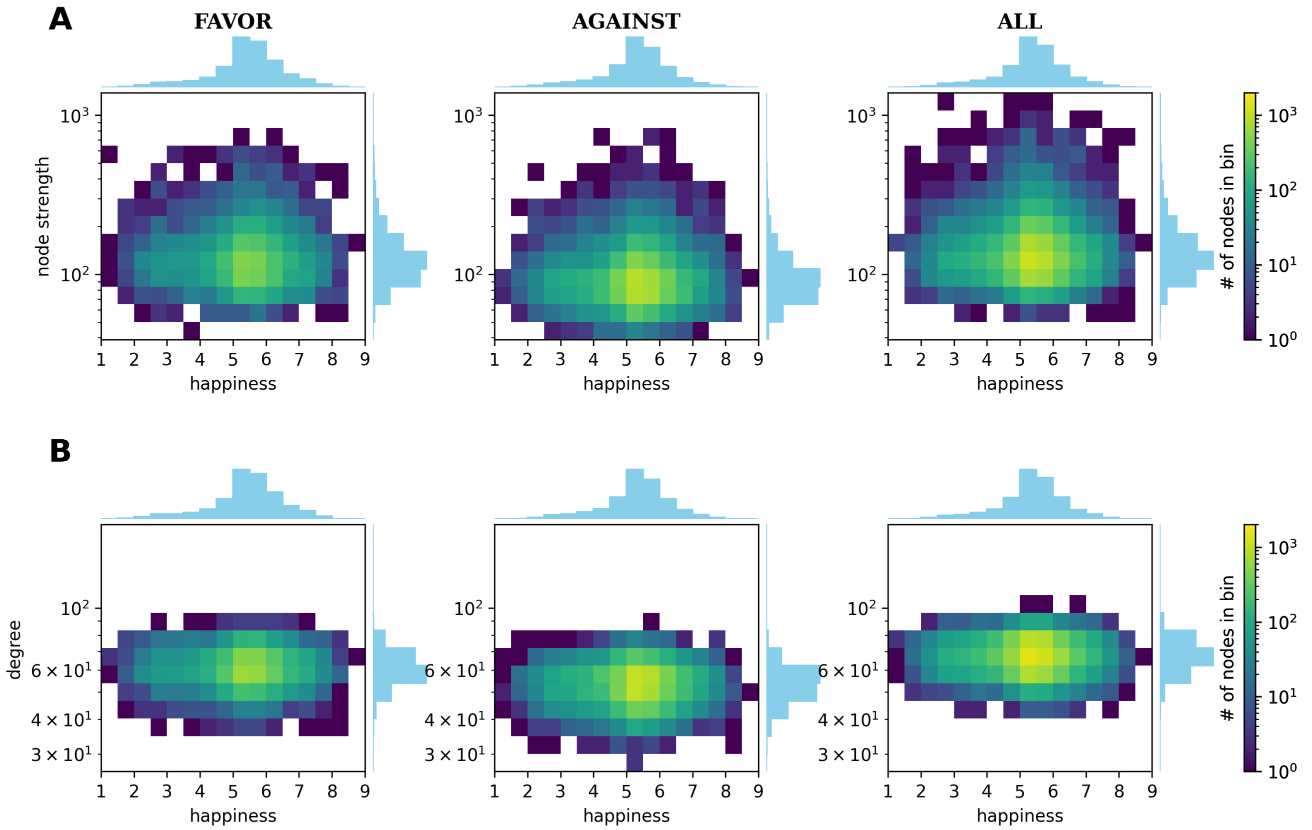

Figure S6. 2D histograms for both the node strengths (top row) and degree (bottom row) vs. happiness scores for the **Erdos-Renyi null model**.

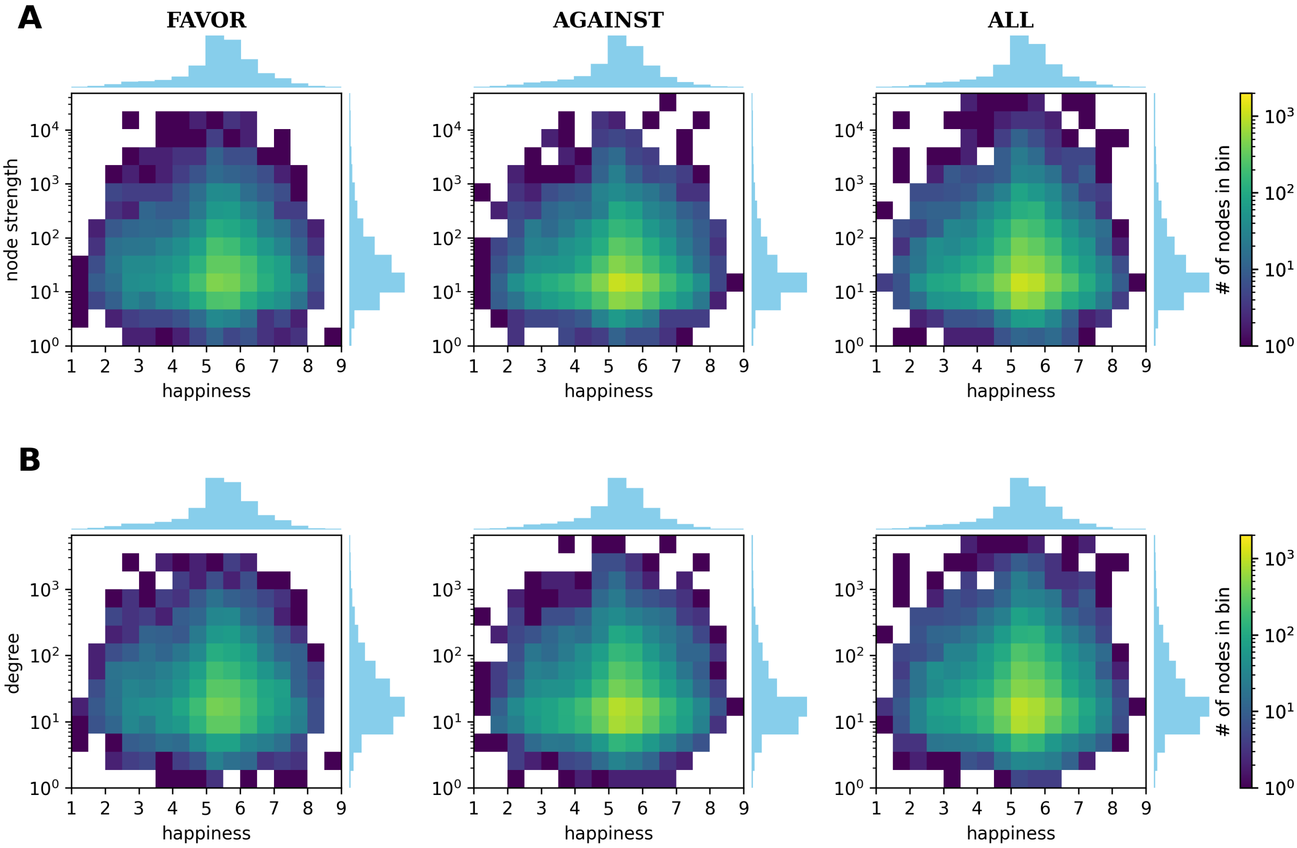

Figure S7. 2D histograms for both the node strengths (top row) and degree (bottom row) vs. happiness scores for the **shuffled score model**.

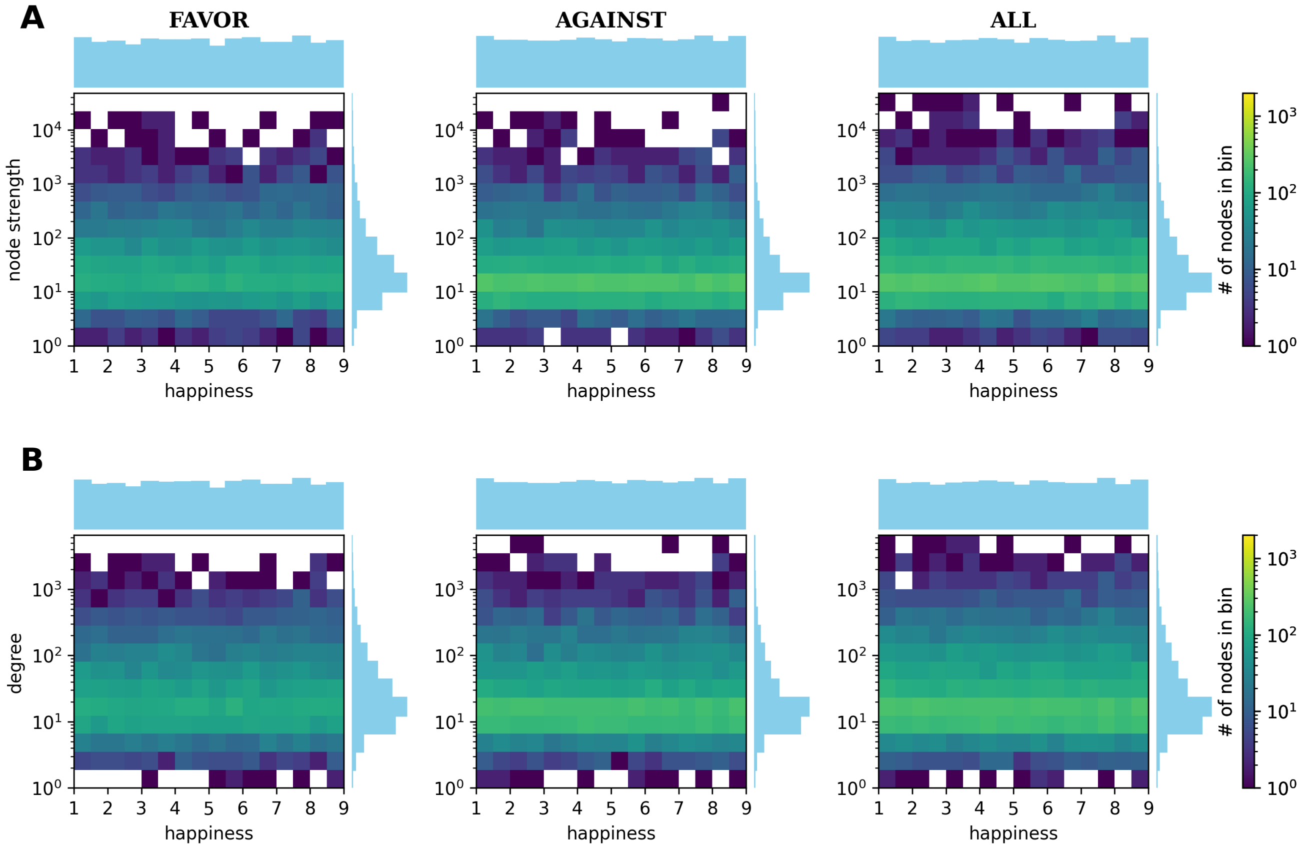

Figure S8. 2D histograms for both the node strengths (top row) and degree (bottom row) vs. happiness scores for the **uniform score model**.

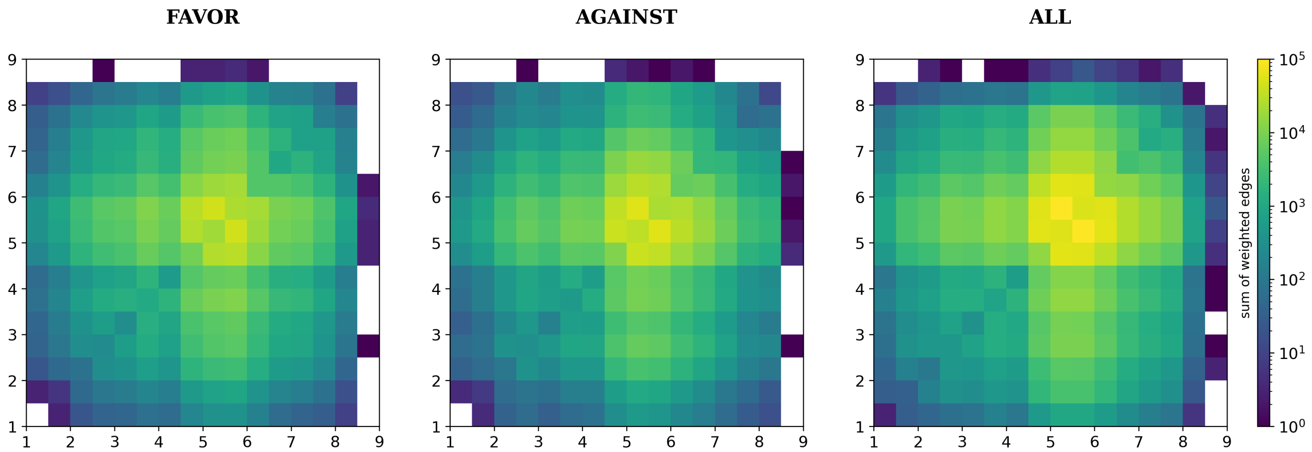

Figure S9. The happiness scores of each pair of nodes for the **configuration model** are plotted in this 2D histogram. Each pair of nodes is weighted by the weight of the edge connecting them. We made the histogram to be symmetric about the 45° line so that one can analyze it from either the horizontal or vertical direction.

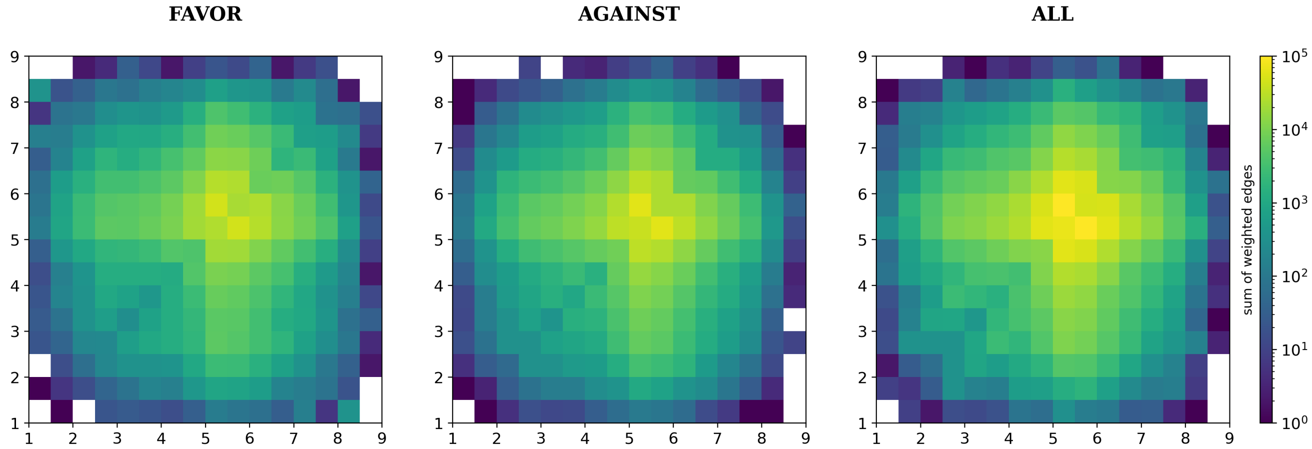

Figure S10. The happiness scores of each pair of nodes for the **Erdos-Renyi null model** are plotted in this 2D histogram. Each pair of nodes is weighted by the weight of the edge connecting them. We made the histogram to be symmetric about the 45° line so that one can analyze it from either the horizontal or vertical direction.

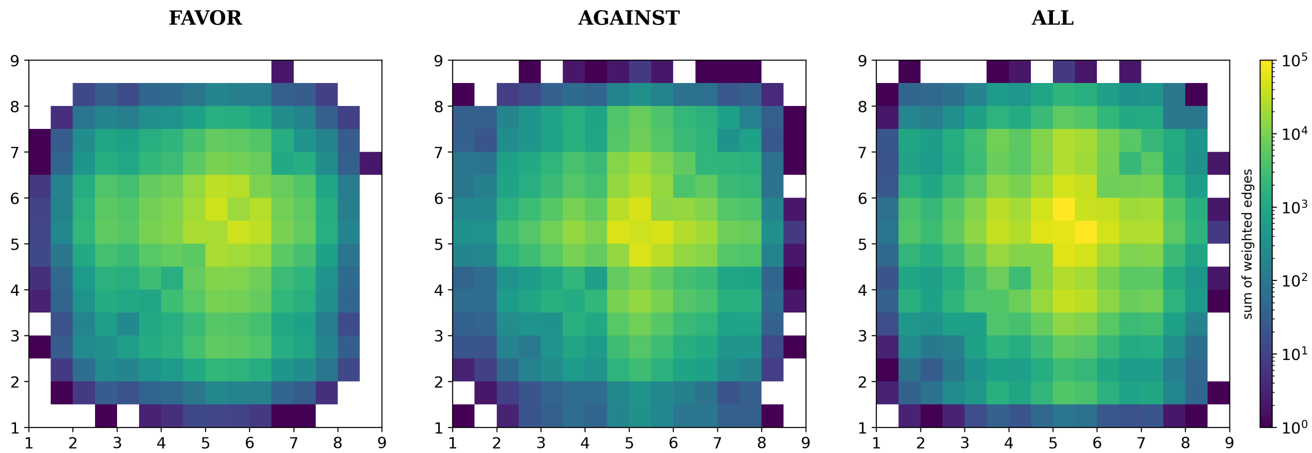

Figure S11. The happiness scores of each pair of nodes for the **shuffled score model** are plotted in this 2D histogram. Each pair of nodes is weighted by the weight of the edge connecting them. We made the histogram to be symmetric about the 45° line so that one can analyze it from either the horizontal or vertical direction.

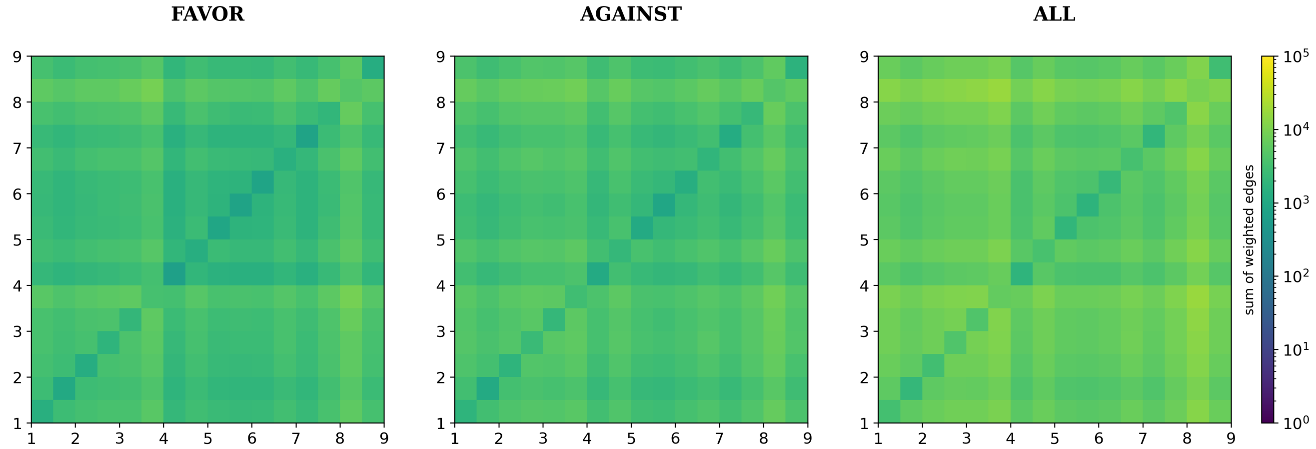

Figure S12. The happiness scores of each pair of nodes for the **uniform score model** are plotted in this 2D histogram. Each pair of nodes is weighted by the weight of the edge connecting them. We made the histogram to be symmetric about the 45° line so that one can analyze it from either the horizontal or vertical direction.

D. Disparity filter

Figure S13 shows how the size and order of the backbone vary depending on the significance level α chosen after removing the most commonly used words and applying the disparity filter, with $\alpha = 1$ corresponding to the network without the most common words on Twitter and the values on the y-axis corresponding to the complete network from which neither the most common words on Twitter nor any likely spurious edges were removed. The removal of words from Twitter eliminated several edges from the network, but few edges are removed by the disparity filter until the threshold value reaches $\alpha = 0.4$. As the value of α decreases, more edges are eliminated, but this also means more nodes are likely to be removed from the backbone. As a result, the number of components does not change monotonically with α (Figure S14). Note that since the disparity filter removes isolated nodes, there is a huge drop in the number of components from $\alpha = 1$ to $\alpha = 0.9$ even if the number of nodes does not change much. The giant component is also by far the largest of all the components, as can be deduced in the plot of the fraction n_2/n, where n is the number of nodes in the entire network and n_2 is the number of nodes in the second largest component (Figure S14).

The resulting score distributions, however, are very similar regardless of the threshold (Figure S15). As the most commonly used words on Twitter are also mostly neutral, removing these manually from the network increases the relative frequency of positive and negative words compared to the original network. Note that this change is due to the removal of the most commonly used words on Twitter; the disparity filter itself only minimally alters the score distribution except for low values of α, where a slight decrease in the relative frequency of negative words in the "favor" corpus is observed. This implies that in terms of the criteria set by the disparity filter, the relevance of nodes does not depend much on their happiness scores, as both neutral and non-neutral words get filtered out.

As the disparity filter becomes more restrictive, it removes more edges and alters the degree distribution (Figure S16), reducing the peak from around $k = 10$ to around $k = 1$. The filter also disproportionately targets edges with lower weights. Whereas the maximum degrees and edge weights remain similar across different values of α, edges with weights close to 1 have been completely eliminated with $\alpha = 0.05$.

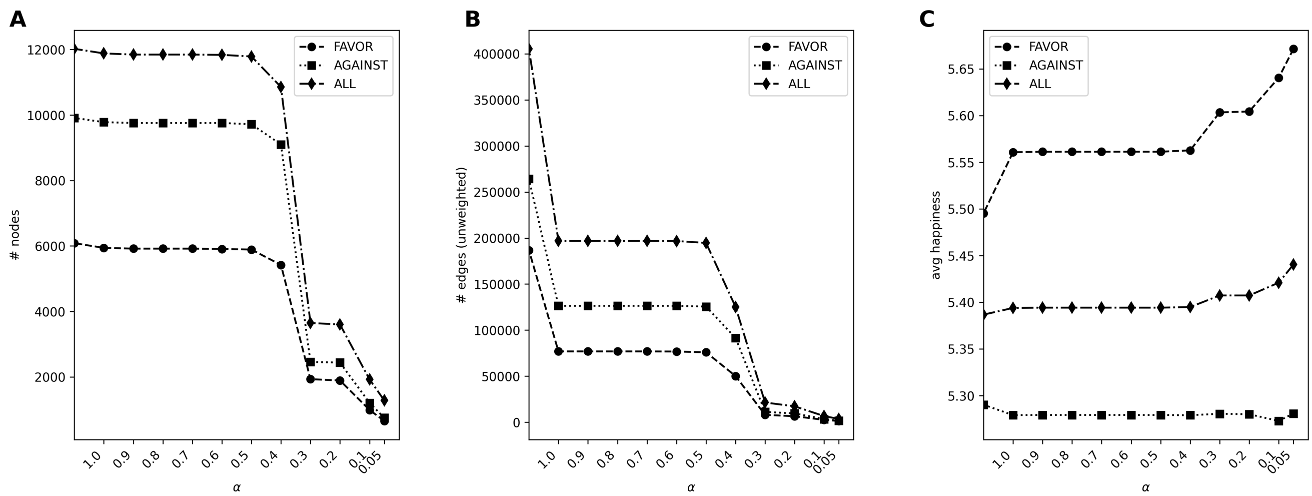

Figure S13. Graph (a) order, (b) size and (c) average happiness score weighted by word count as a function of the threshold used in the disparity filter. $\alpha = 1$ indicates that the top words from Twitter were removed, but the disparity filter was not applied, while the points on the y-axis correspond to the values from the complete network without any backboning.

[1] D. J. Griffiths, *Introduction to Electrodynamics* (Cambridge University Press, Cambridge, 2017).
[2] A. Bobrinha, Revista Brasileira de Lorem Ipsum **23**, 179 (2002).
[3] R. P. Feynman, R. B. Leighton and M. Sands, *Lições de Física de Feynman* (Editora Bookman, Porto Alegre, 2008).
[4] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, Danvers, 1999).

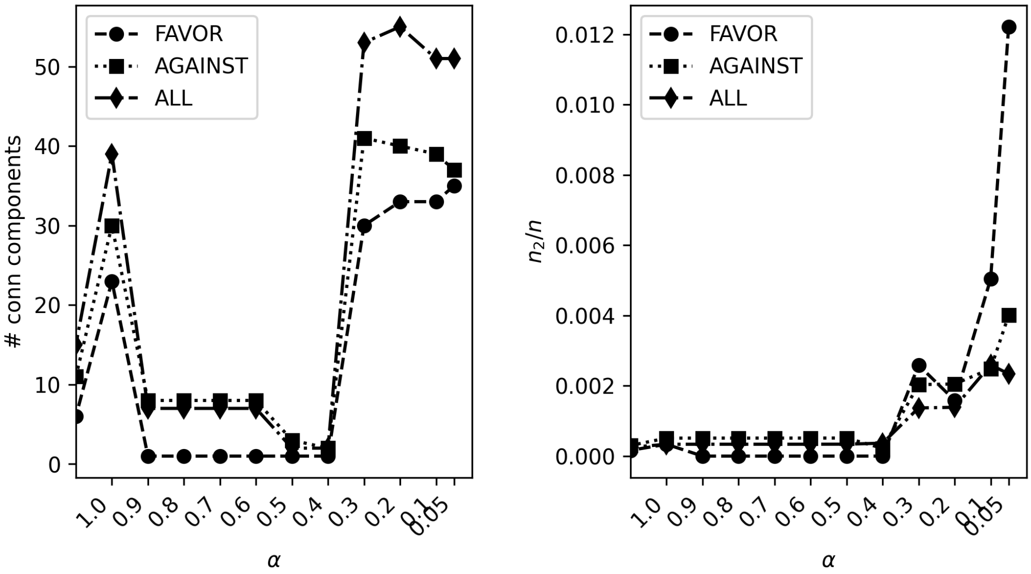

Figure S14. Number of connected components and relative size of the second-largest component to the largest component for all thresholds.

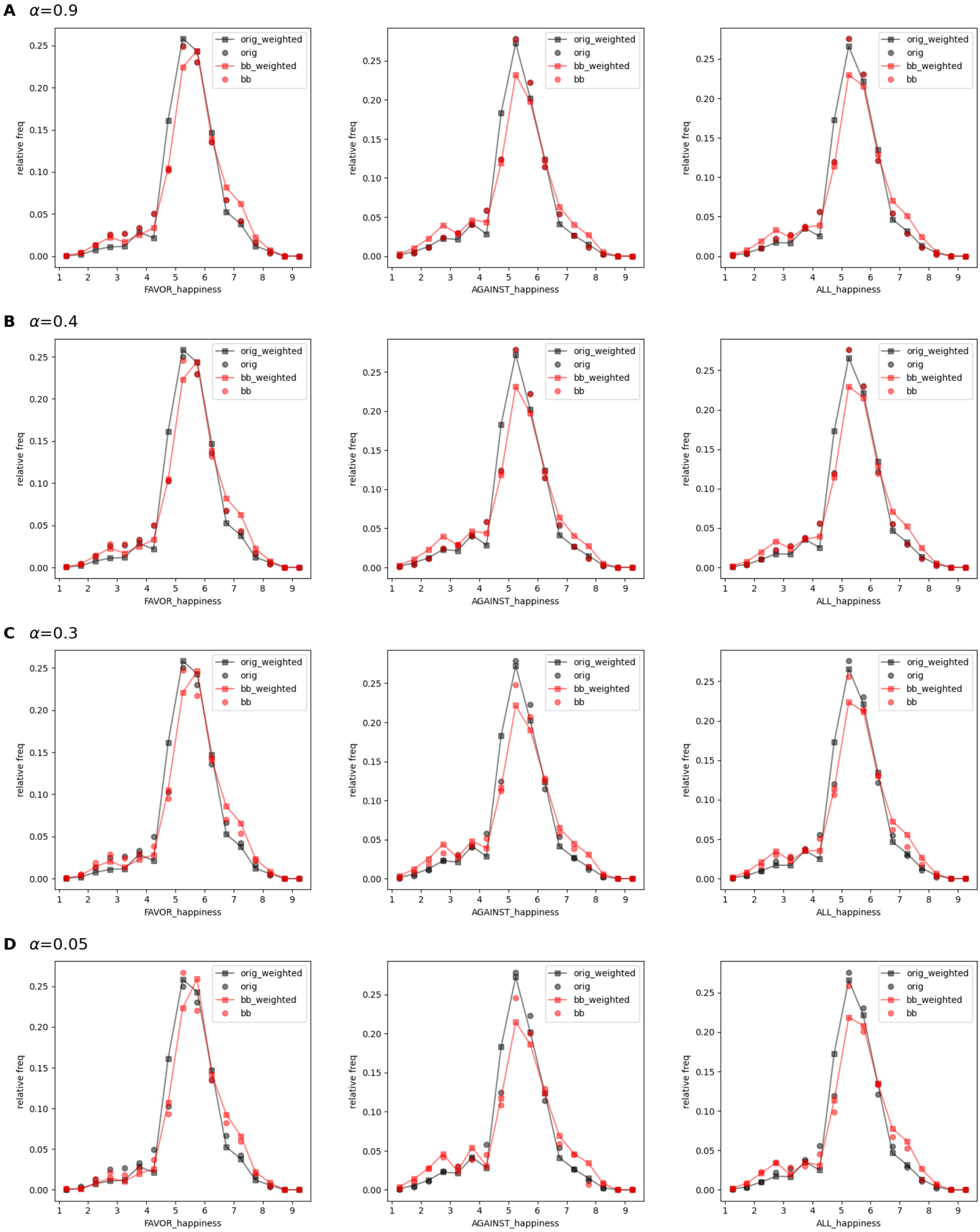

Figure S15. Relative frequencies of happiness scores in the backbone for significance levels $\alpha = 0.9, 0.4, 0.3, 0.05$. Curves marked “orig” refer to the original network, while those marked “bb” refer to the backbone. The suffix “_weighted” indicates that the words are weighted by their word counts.

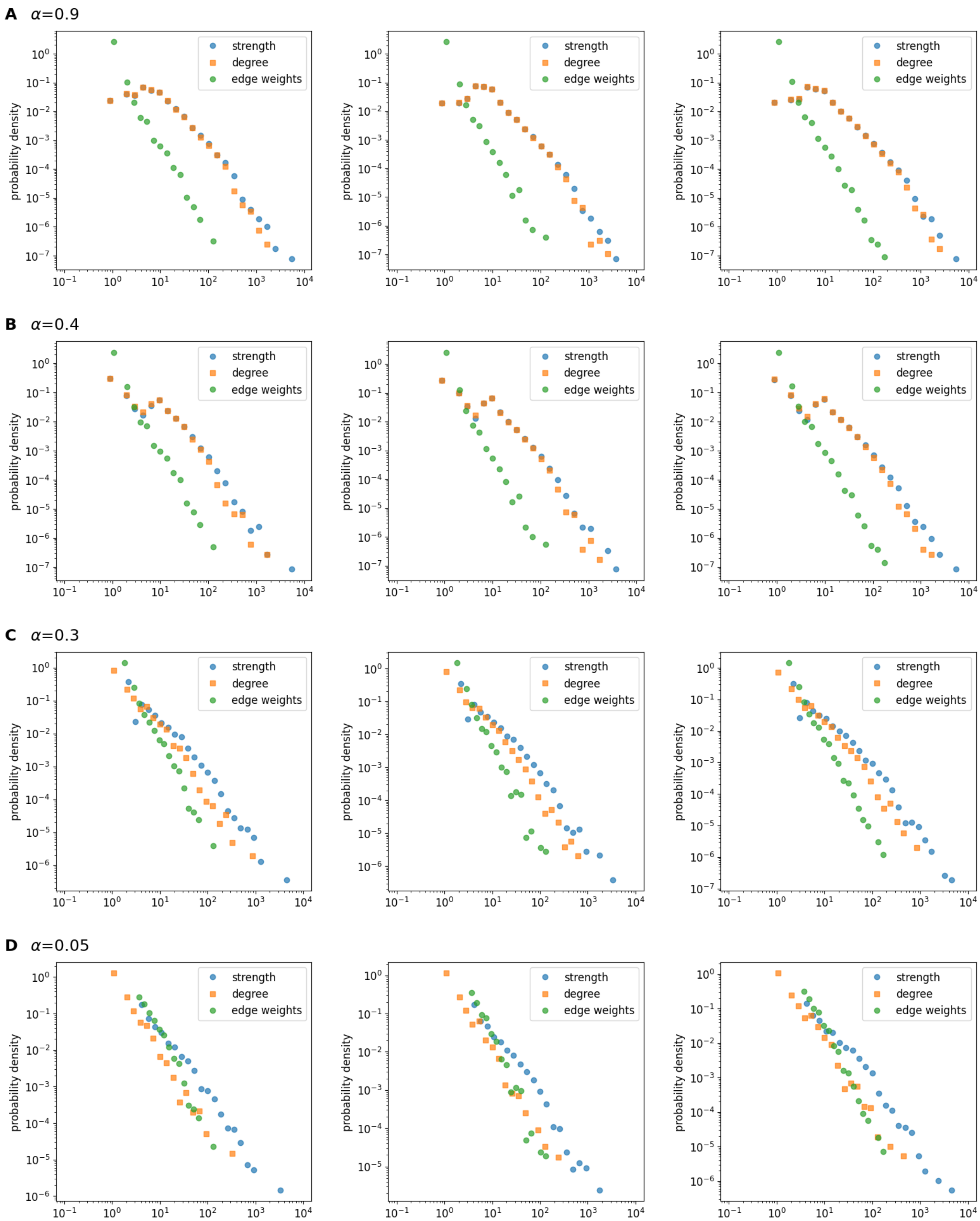

Figure S16. The distributions of the node degrees, node strengths, and edge weights for significance levels $\alpha = 0.9, 0.4, 0.3, 0.05$.

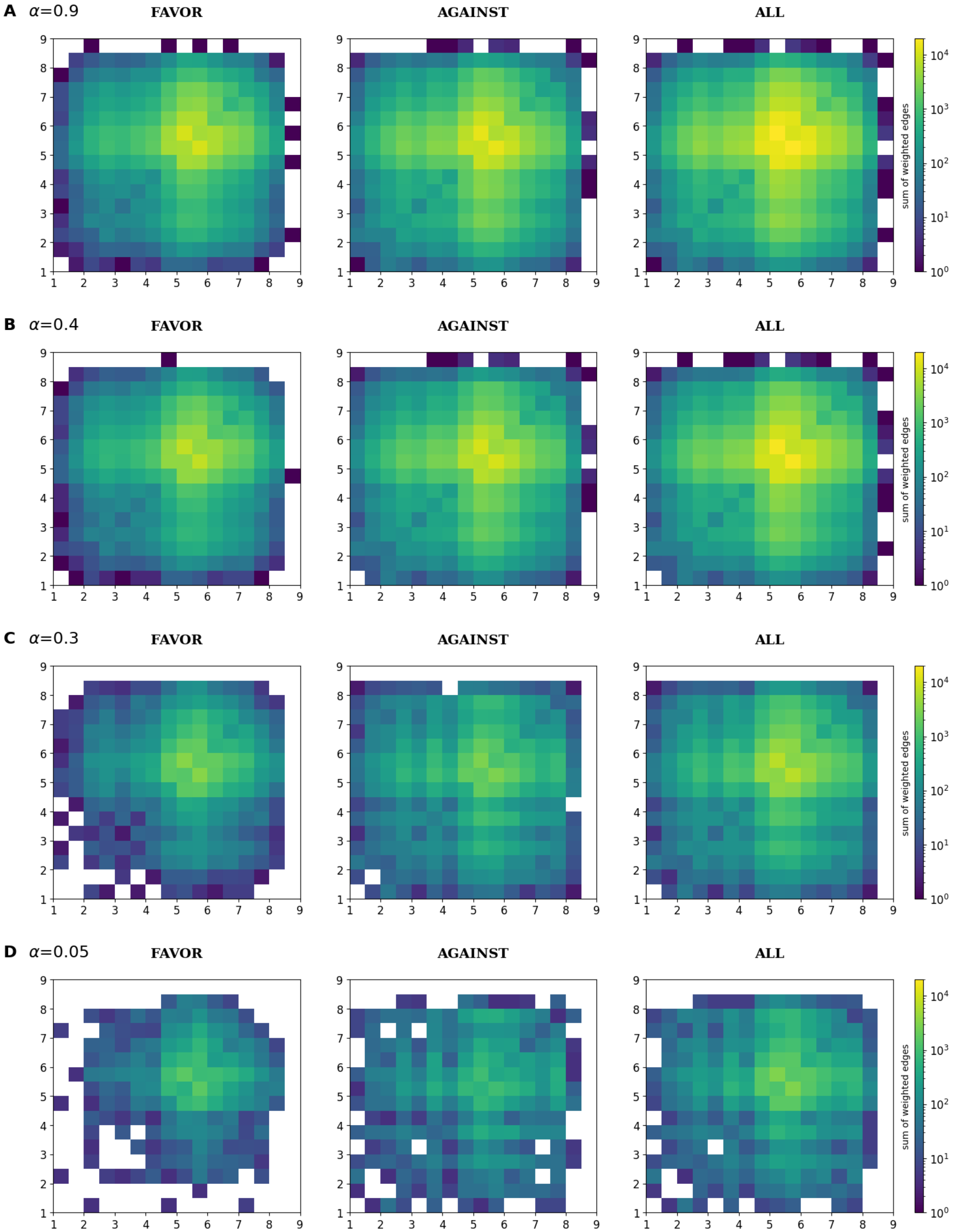

Figure S17. 2D histogram of scores of connected nodes weighted by the edge weights for significance levels $\alpha = 0.9, 0.4, 0.3, 0.05$. Note that the heat map is made to be symmetric about the 45° line so it can be analyzed from either the horizontal or vertical direction.