

Facebook Results Memo

ddd

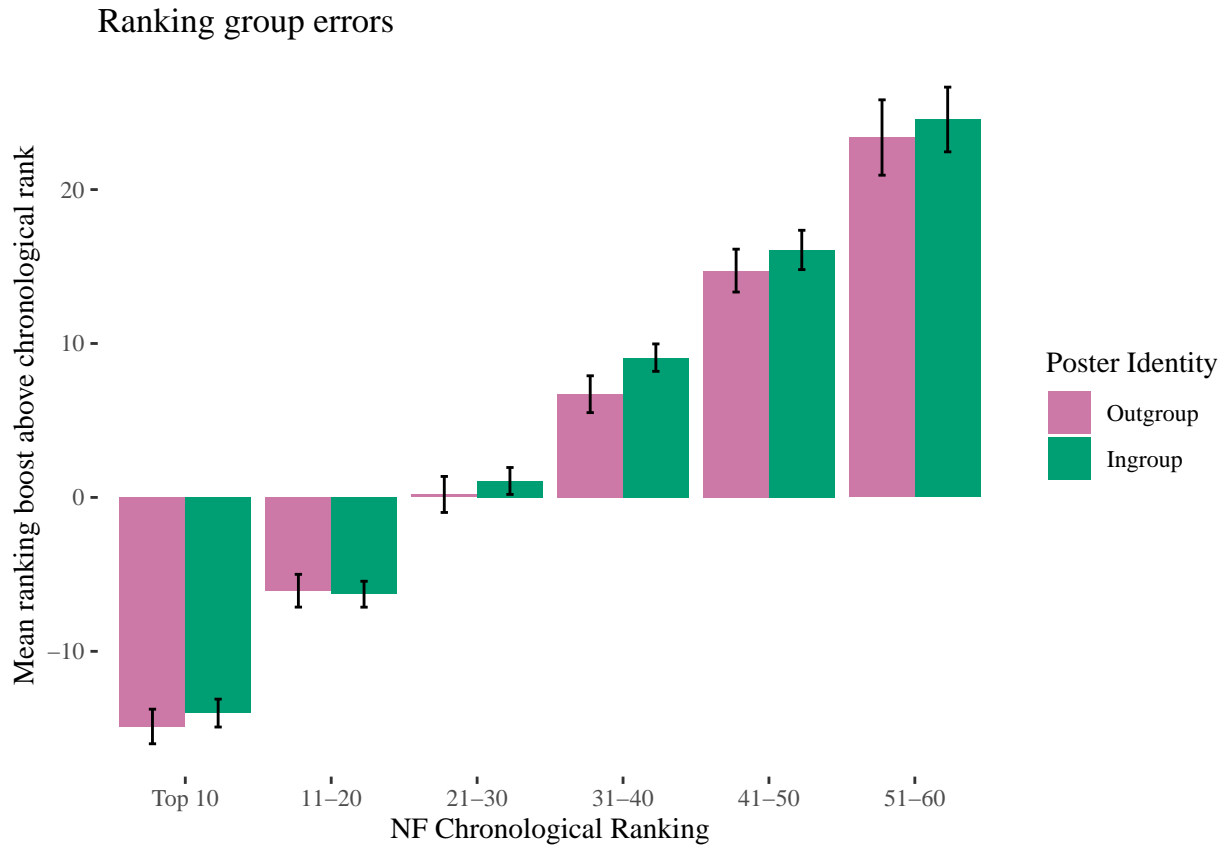
8/27/2020

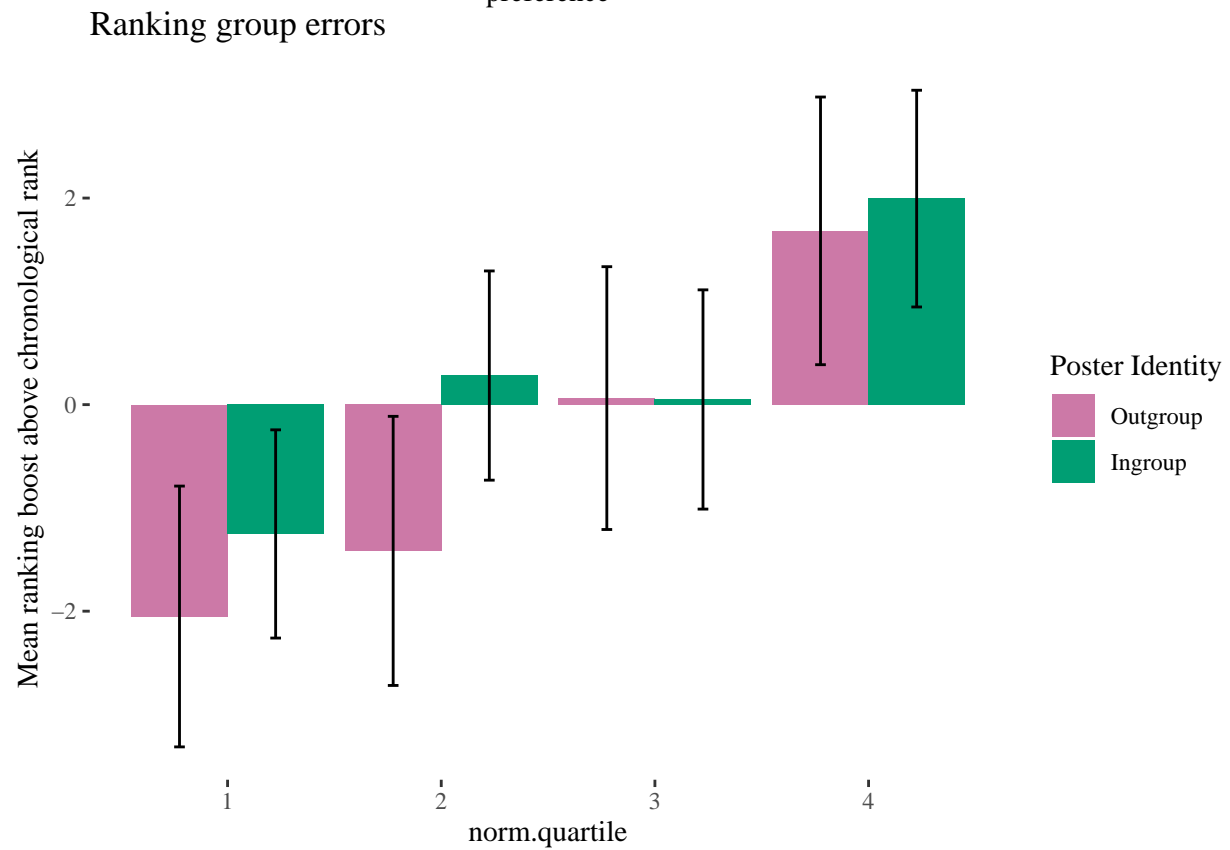
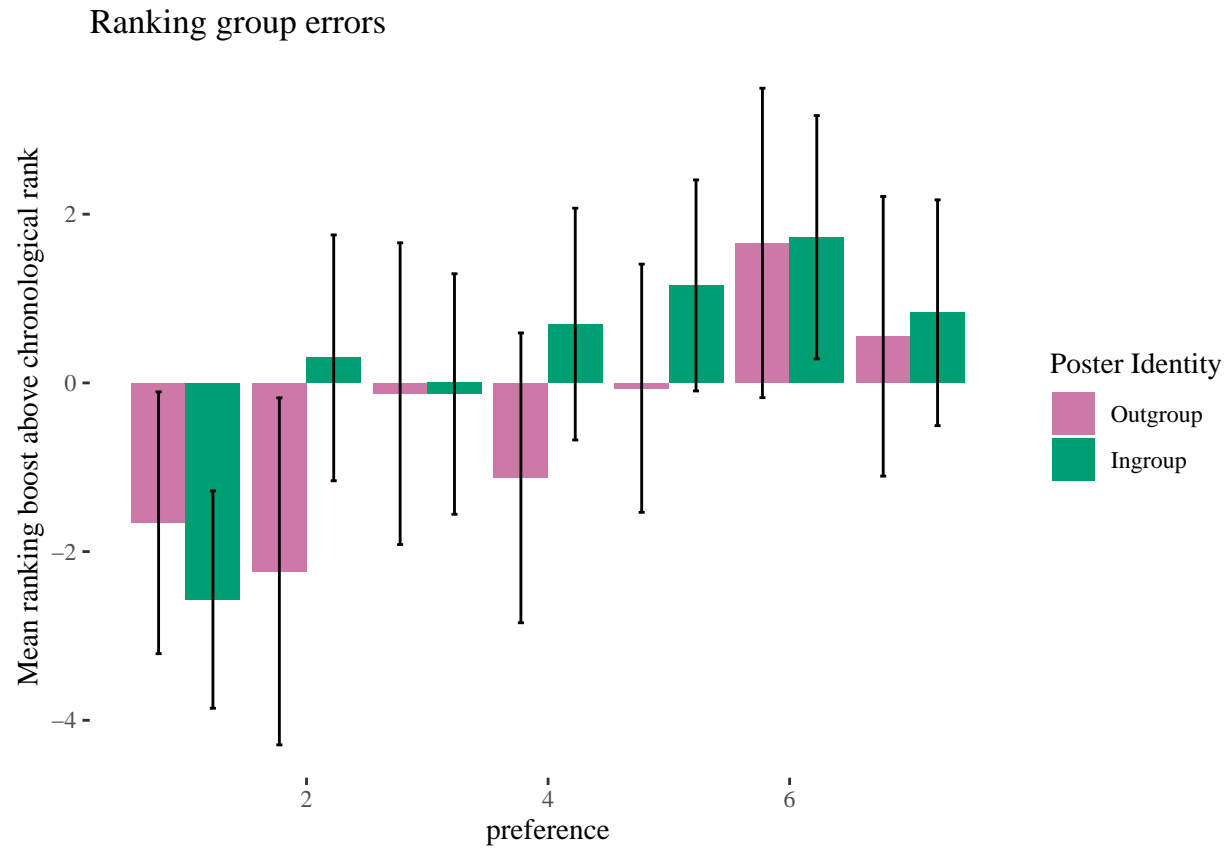
Executive Summary

Artificial intelligence has become an important component of how social media platforms try to achieve the goal of bringing people together, by helping prioritize what we see and consume online. These algorithms have the potential to expand people’s social networks, but – given evidence of bias with algorithms in other settings – also have the risk of narrowing the breadth of those with whom we interact online, and reinforcing or potentially even exacerbating the high levels of segregation that characterize ‘normal’ (real-life) interactions. To explore this possibility, we conduct an audit study in which each subject (along with an RA) records their first 60 news feed posts (NF) and the first 60 users recommended by the ‘People You May Know’ algorithm (PYMK).

We find evidence of significant discrimination in the NF sorting. When the author and subject are of the same race, the post receives a boost equivalent to 20 percentile points of stated preference; a same-race post in the 50th percentile of stated preference is ranked the same on average as an opposite-race post in the 70th percentile. We find no evidence of discrimination in the PYMK recommendations. We reconcile these findings by distinguishing between behaviors dominated by System 1 (driven by implicit/subconscious attitudes) vs System 2 (driven by explicit/conscious attitudes).

Paper Figures





% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:06:51

Table 1:

	<i>Dependent variable:</i>					
	new.rank					
	(1)	(2)	(3)	(4)	(5)	(6)
religion.in.group	−1.006296*** (0.339386)			−1.007704*** (0.338416)	−0.931093*** (0.330125)	
I(100 *norm.pctle)		−0.038949*** (0.005737)		−0.038960*** (0.005735)		−0.040080*** (0.005579)
time_rank			0.232724*** (0.010967)		0.232392*** (0.010963)	0.233469*** (0.010932)
Constant	23.903490*** (0.263650)	25.246370*** (0.331596)	17.874620*** (0.302280)	25.855030*** (0.389394)	18.444260*** (0.363434)	19.864020*** (0.409220)
Observations	7,866	7,866	7,866	7,866	7,866	7,866
R ²	0.001117	0.005826	0.054160	0.006946	0.055116	0.060329
Adjusted R ²	0.000990	0.005700	0.054040	0.006693	0.054876	0.060090

Note:

*p<0.1; **p<0.05; ***p<0.01

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:06:52

Table 2:

	<i>Dependent variable:</i>					
	new.rank					
	(1)	(2)	(3)	(4)	(5)	(6)
religion.in.group	−0.320043 (0.351422)			−0.173920 (0.348483)	−0.260766 (0.351182)	
I(100 *norm.pctle)		−0.079599*** (0.005687)		−0.079514*** (0.005690)		−0.077353*** (0.005939)
friend_rank			0.050286*** (0.009575)		0.050057*** (0.009580)	0.013012 (0.009923)
Constant	30.368040*** (0.287258)	34.141870*** (0.328728)	28.637850*** (0.332681)	34.253810*** (0.397966)	28.819020*** (0.412562)	33.636990*** (0.506265)
Observations	10,882	10,882	10,882	10,882	10,882	10,882
R ²	0.000076	0.017689	0.002529	0.017711	0.002579	0.017844
Adjusted R ²	−0.000016	0.017599	0.002437	0.017531	0.002396	0.017664

Note:

*p<0.1; **p<0.05; ***p<0.01

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:06:53

Table 3:

	<i>Dependent variable:</i>					
	new.rank					
	(1)	(2)	(3)	(4)	(5)	(6)
religion.in.group	-0.320043 (0.351422)			-0.173920 (0.348483)	-0.251979 (0.350972)	
I(100 *norm.pctle)		-0.079599*** (0.005687)		-0.079514*** (0.005690)		-0.075465*** (0.005937)
pct_friend_rank			0.060297*** (0.009570)		0.060085*** (0.009574)	0.023990** (0.009920)
Constant	30.368040*** (0.287258)	34.141870*** (0.328728)	28.335990*** (0.332496)	34.253810*** (0.397966)	28.510760*** (0.412088)	33.211360*** (0.506023)
Observations	10,882	10,882	10,882	10,882	10,882	10,882
R ²	0.000076	0.017689	0.003636	0.017711	0.003683	0.018217
Adjusted R ²	-0.000016	0.017599	0.003544	0.017531	0.003500	0.018036

Note:

*p<0.1; **p<0.05; ***p<0.01

control correlations

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:06:56

Table 4: Correlations 1

	Pearson	Spearman	Kendall
NF Rank, Time	0.233	0.238	0.165
PYMK Rank, Pct Friends	0.060	0.060	0.041

Correlation matrix for benchmarks

preference correlations

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

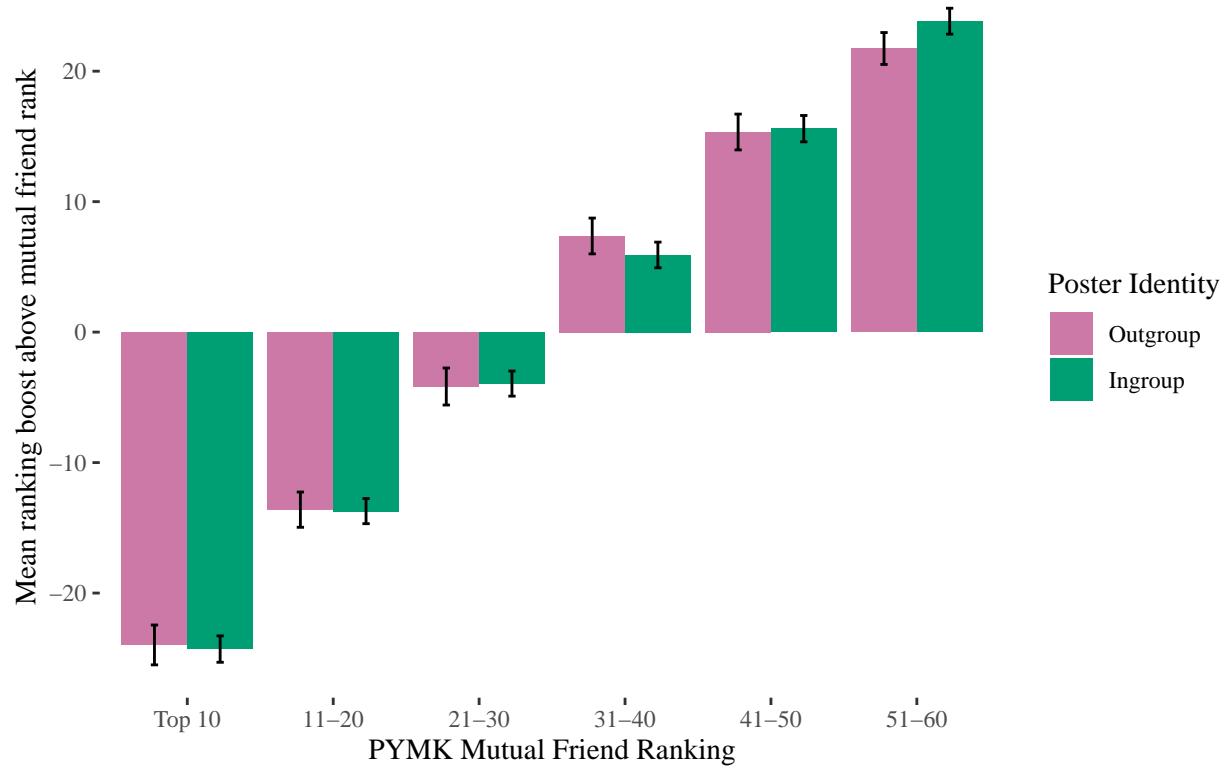
% Date and time: Tue, Jan 05, 2021 - 18:06:59

Table 5: Correlations 2

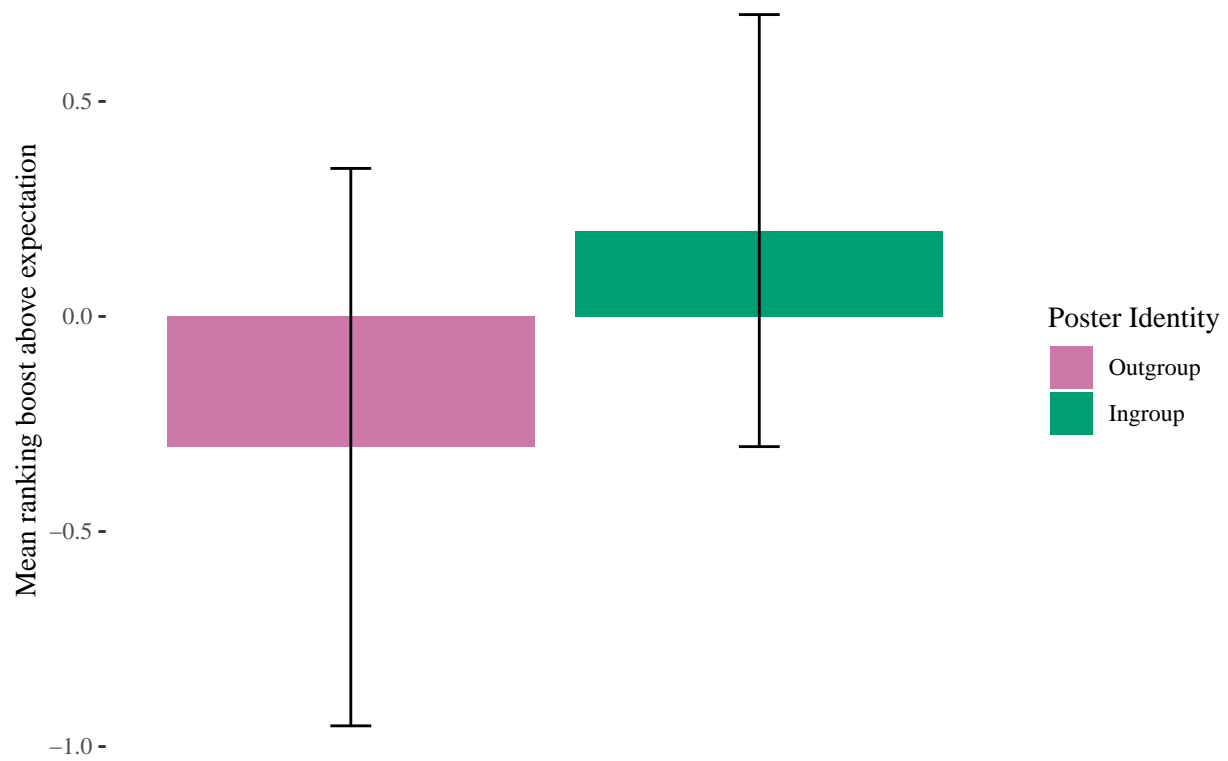
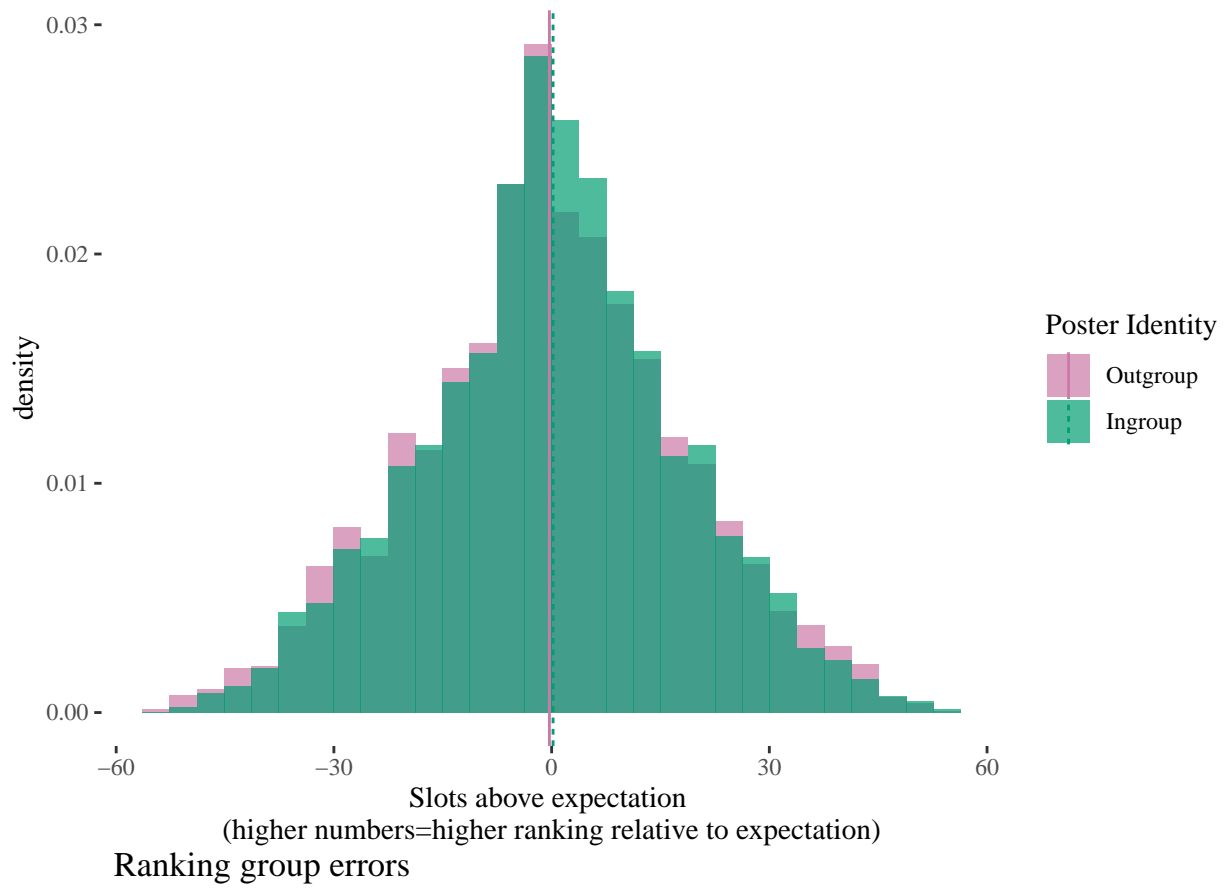
	Pearson	Spearman	Kendall
NF Rank, Preference	-0.078	-0.079	-0.053
PYMK Rank, Familiarity	-0.140	-0.134	-0.090

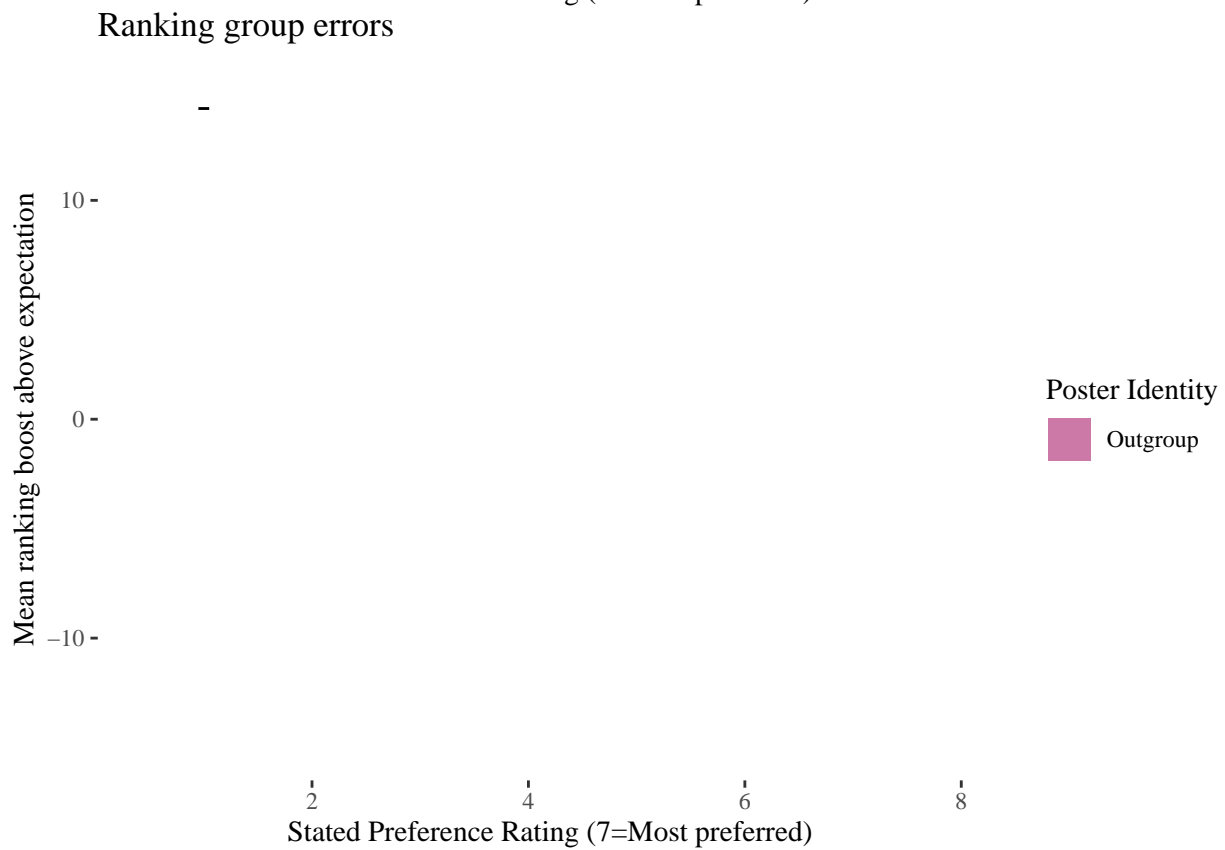
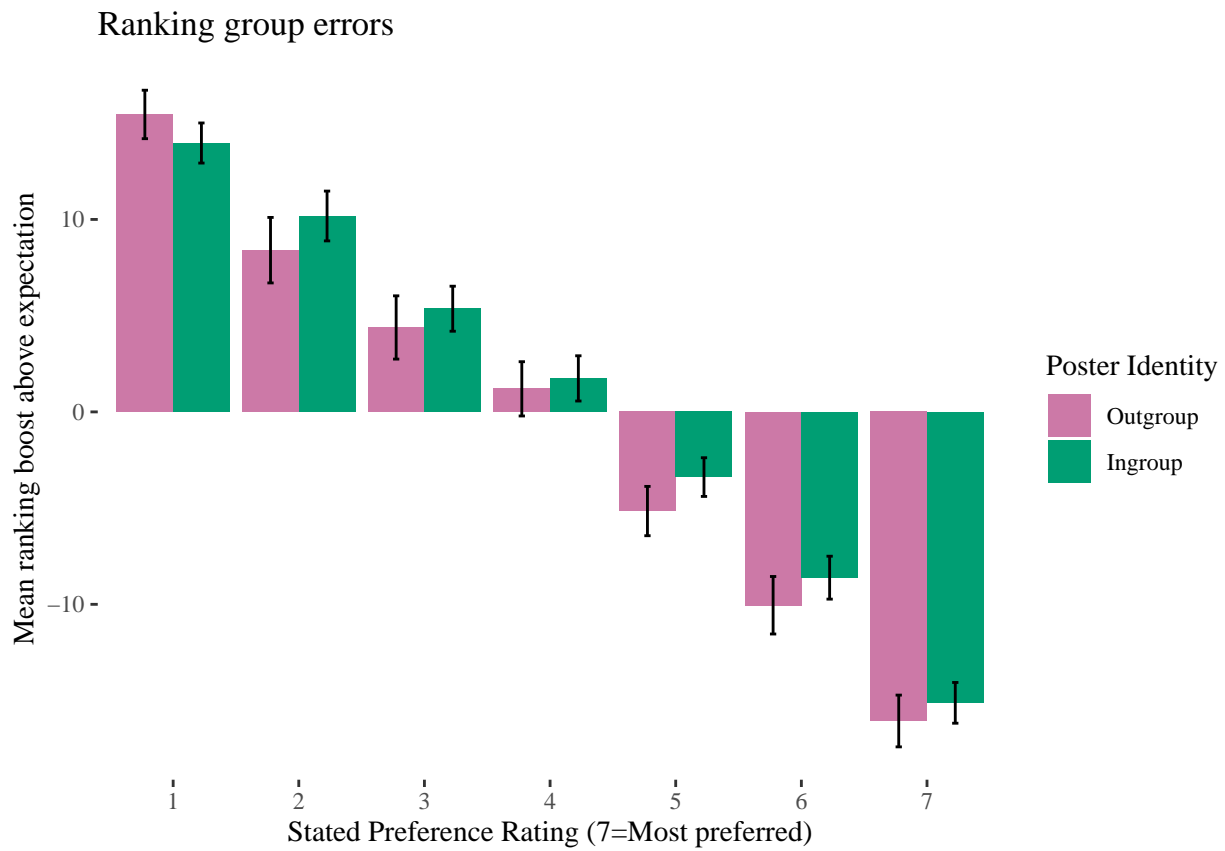
Correlation matrix for preferences

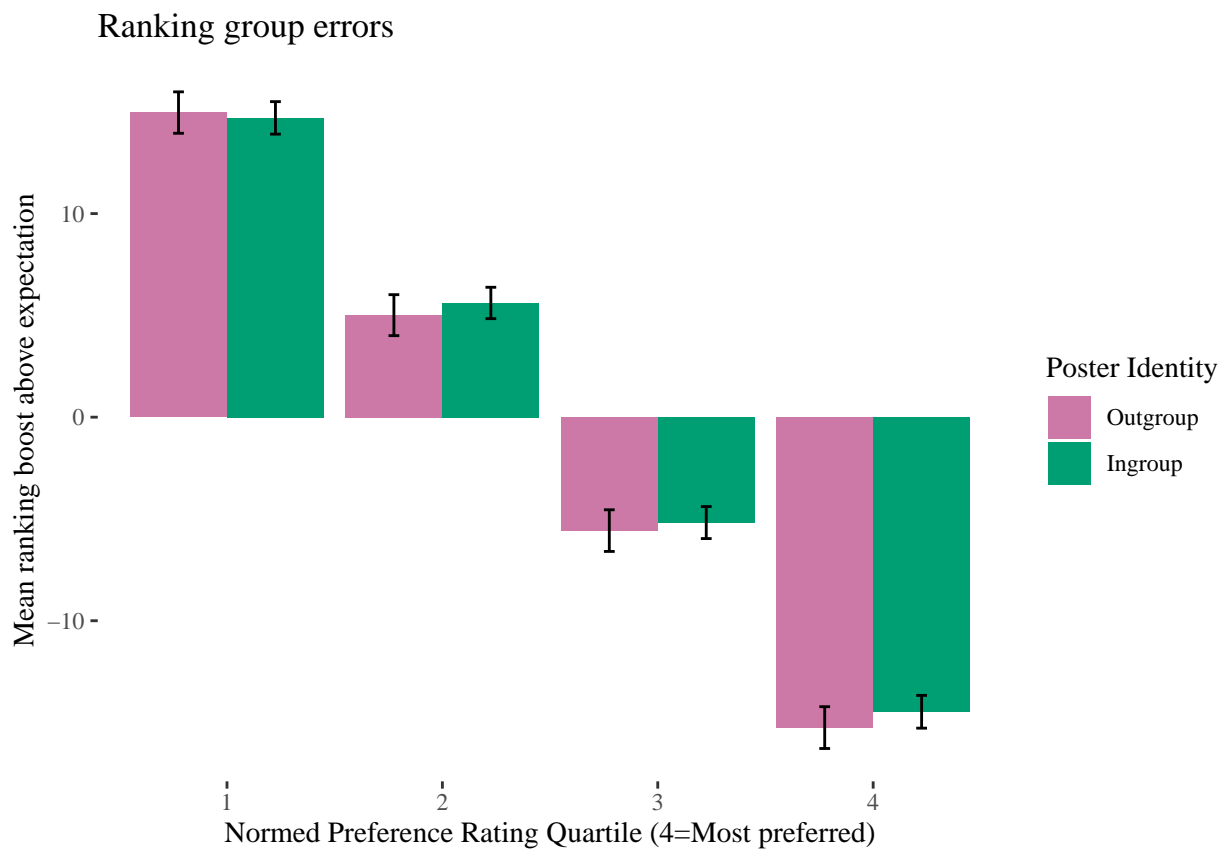
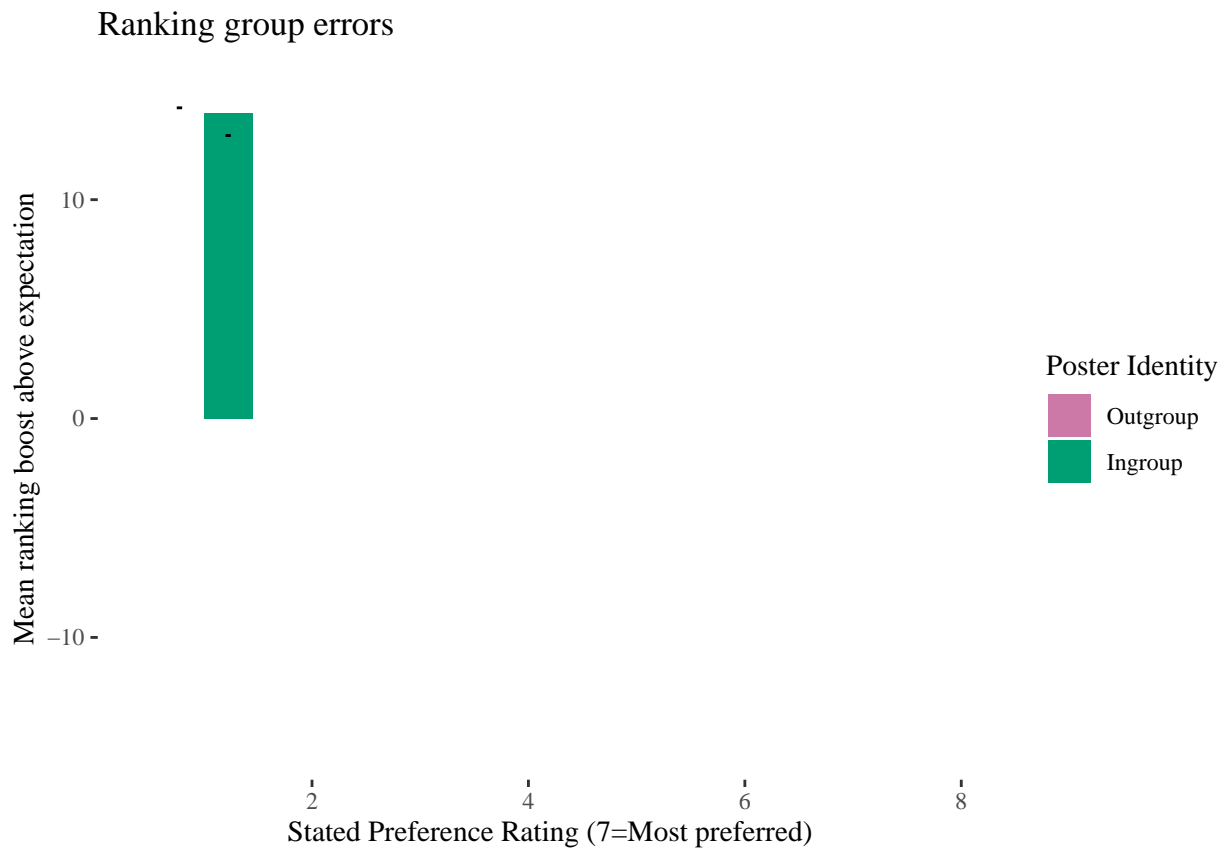
Ranking group errors



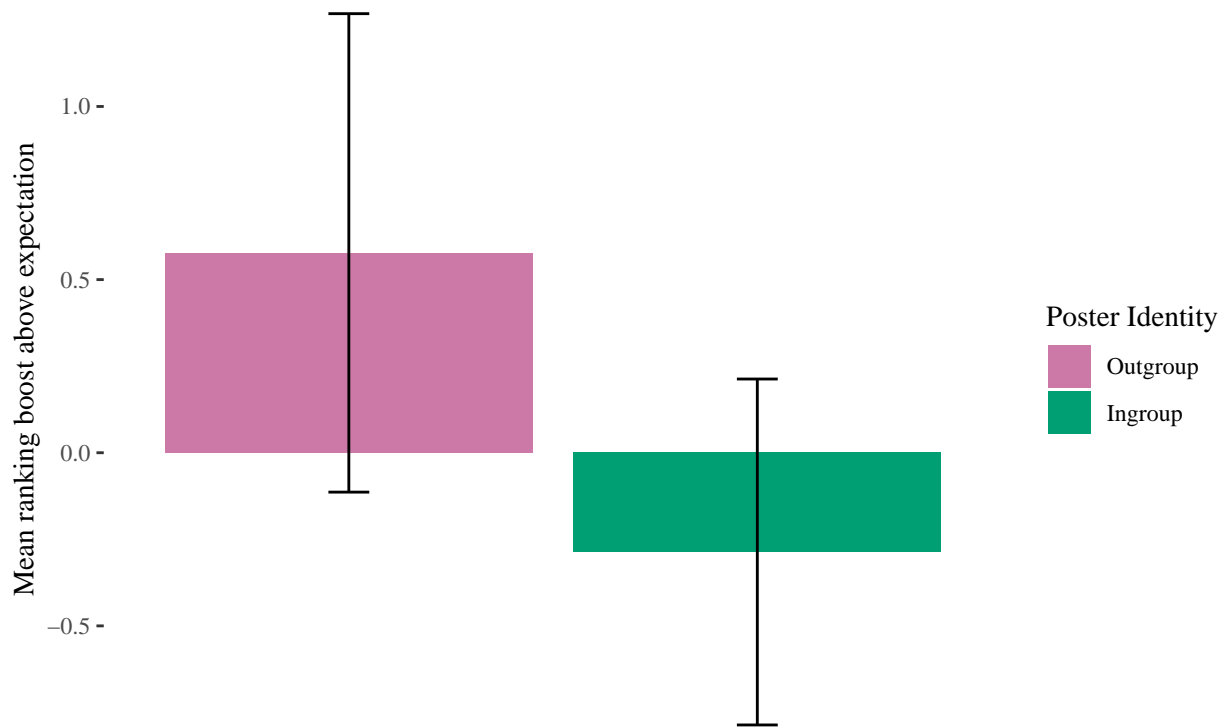
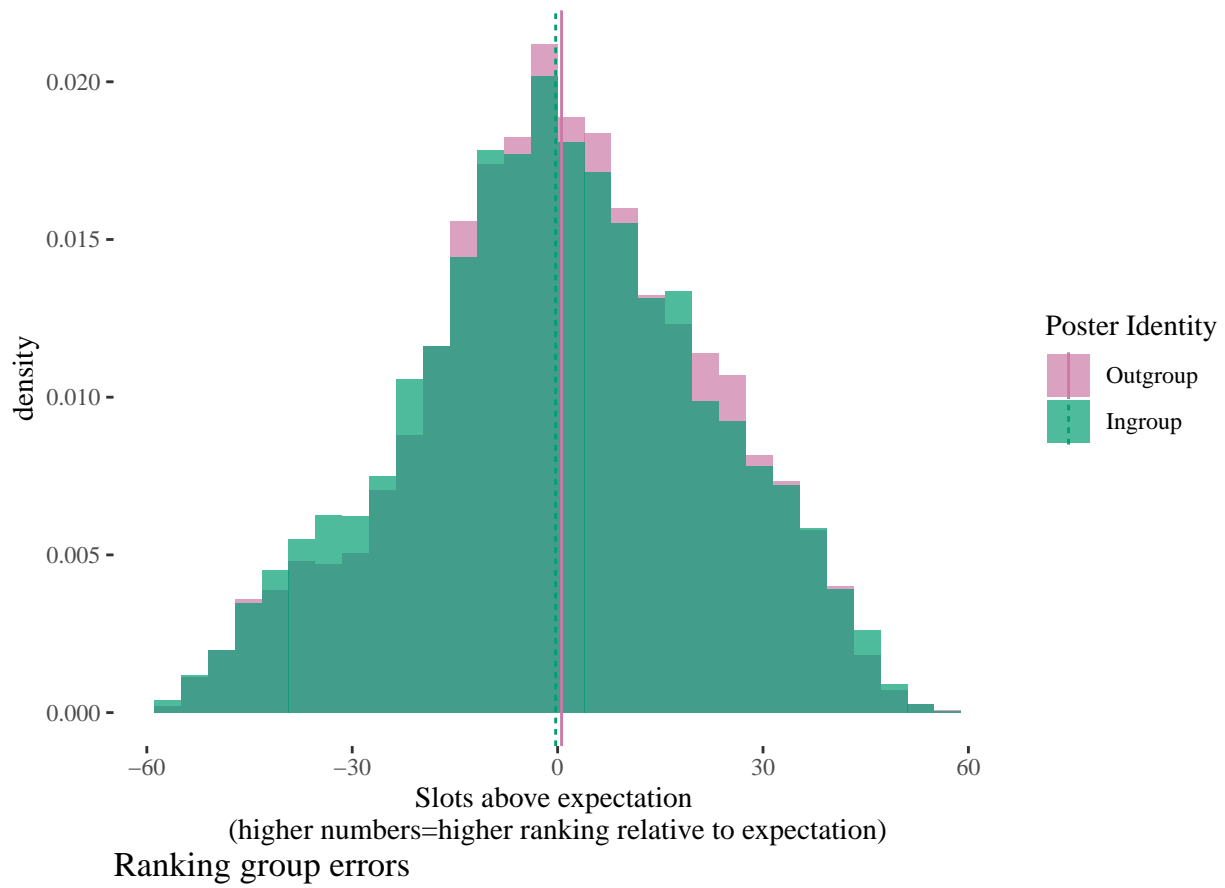
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

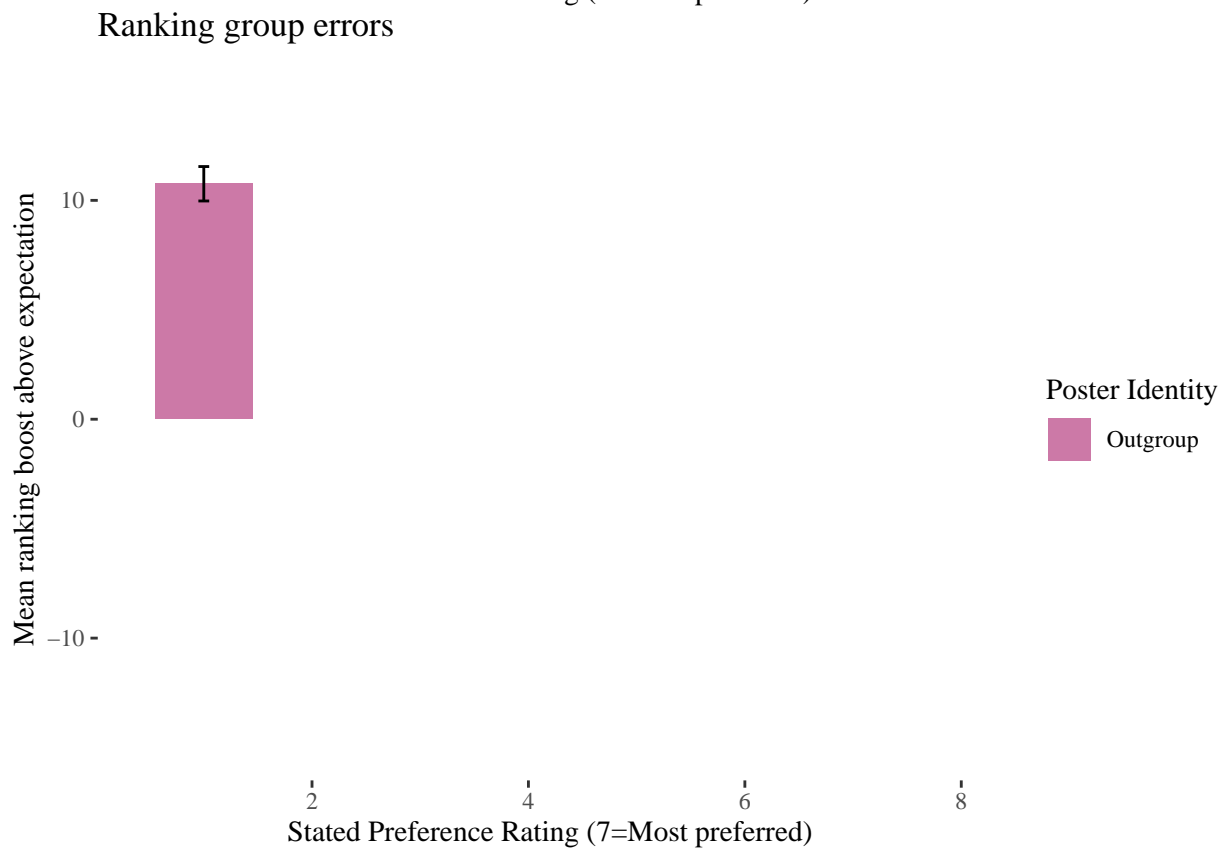
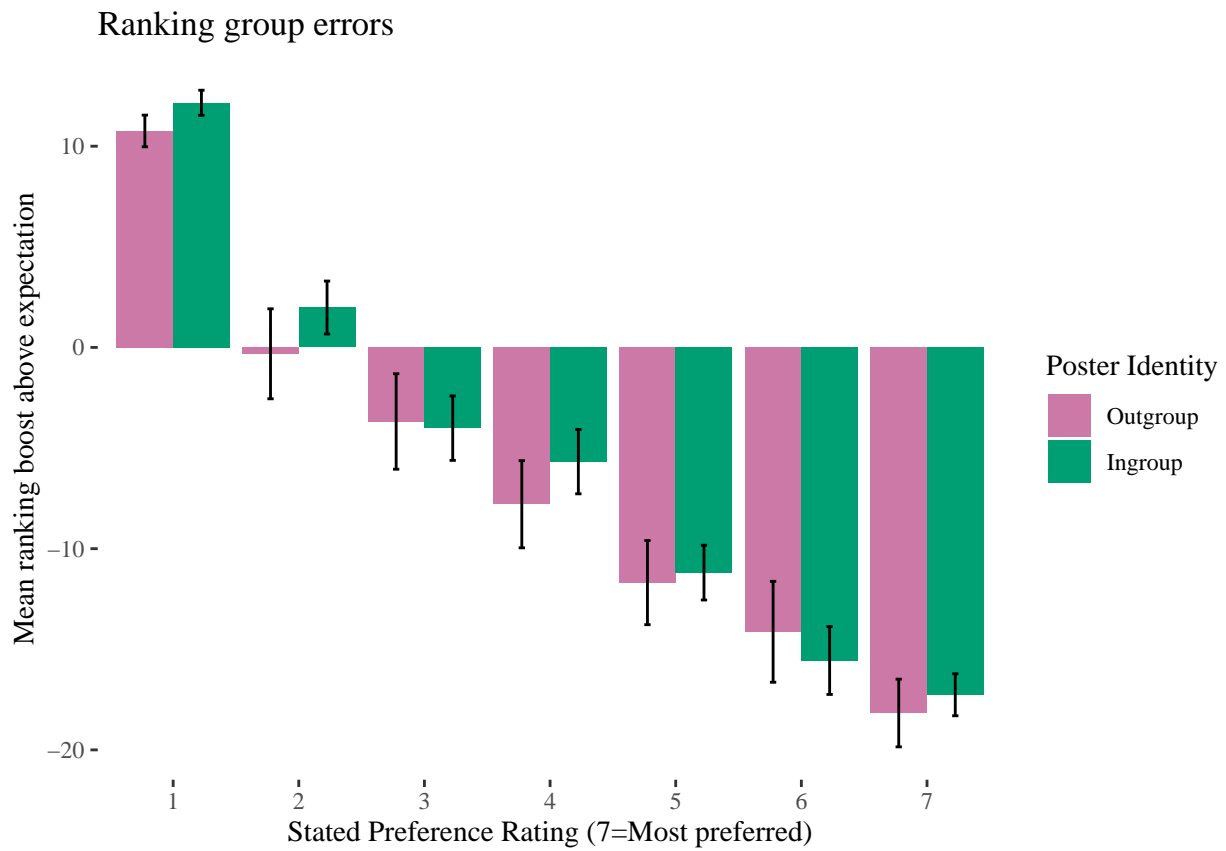


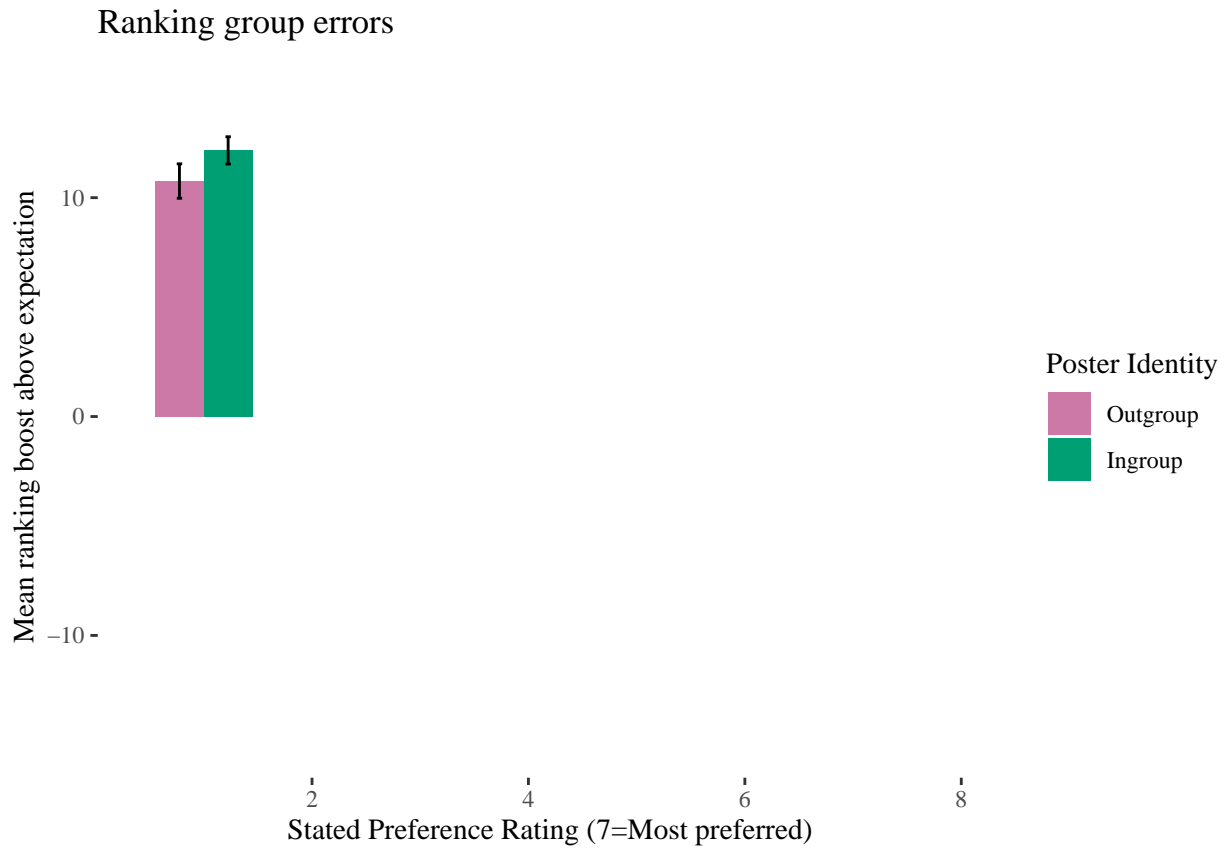




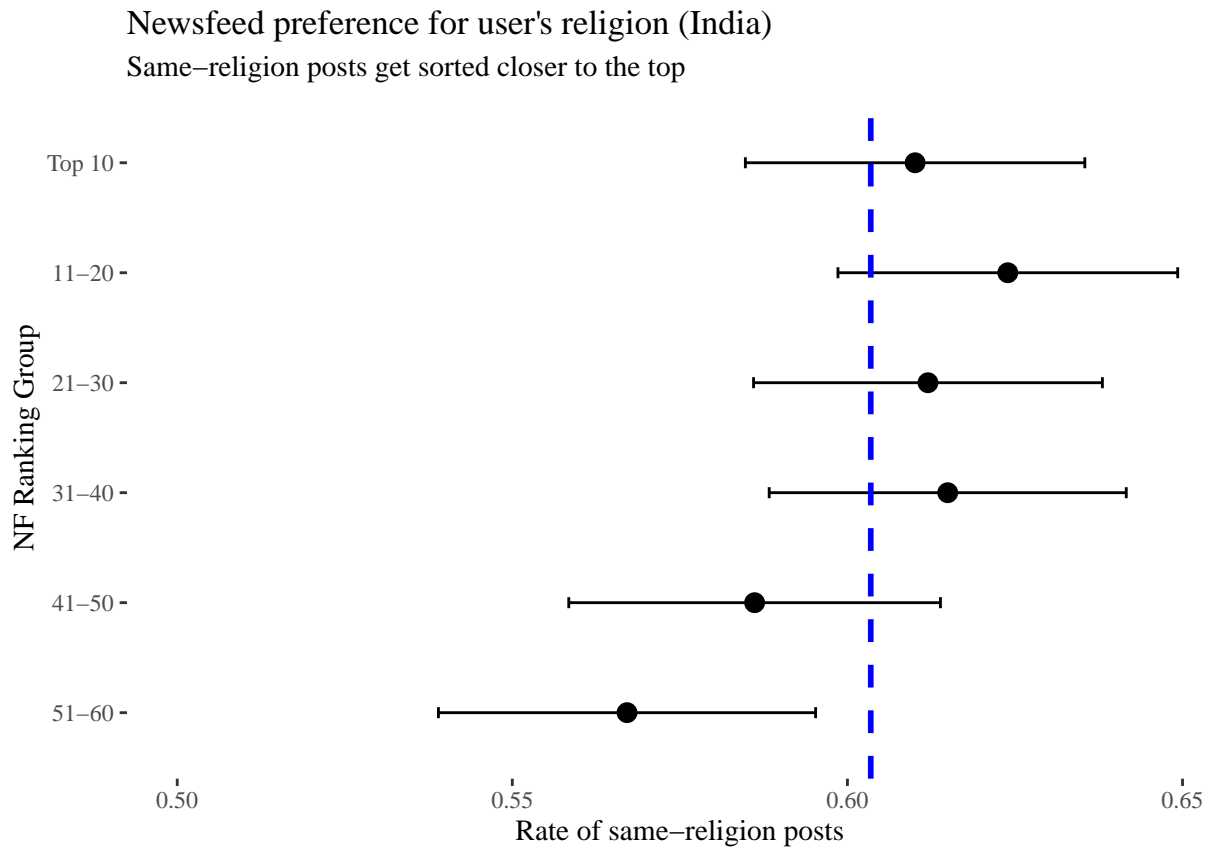
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```





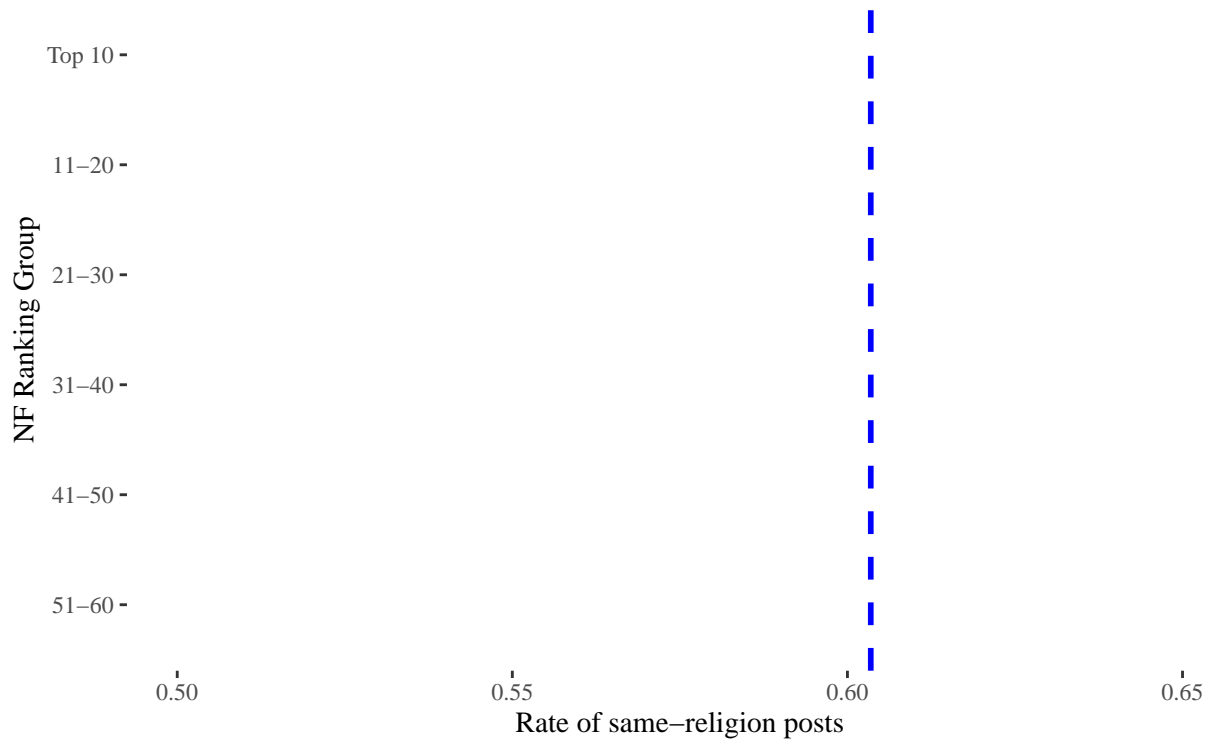


```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```



Newsfeed preference for user's religion (India)

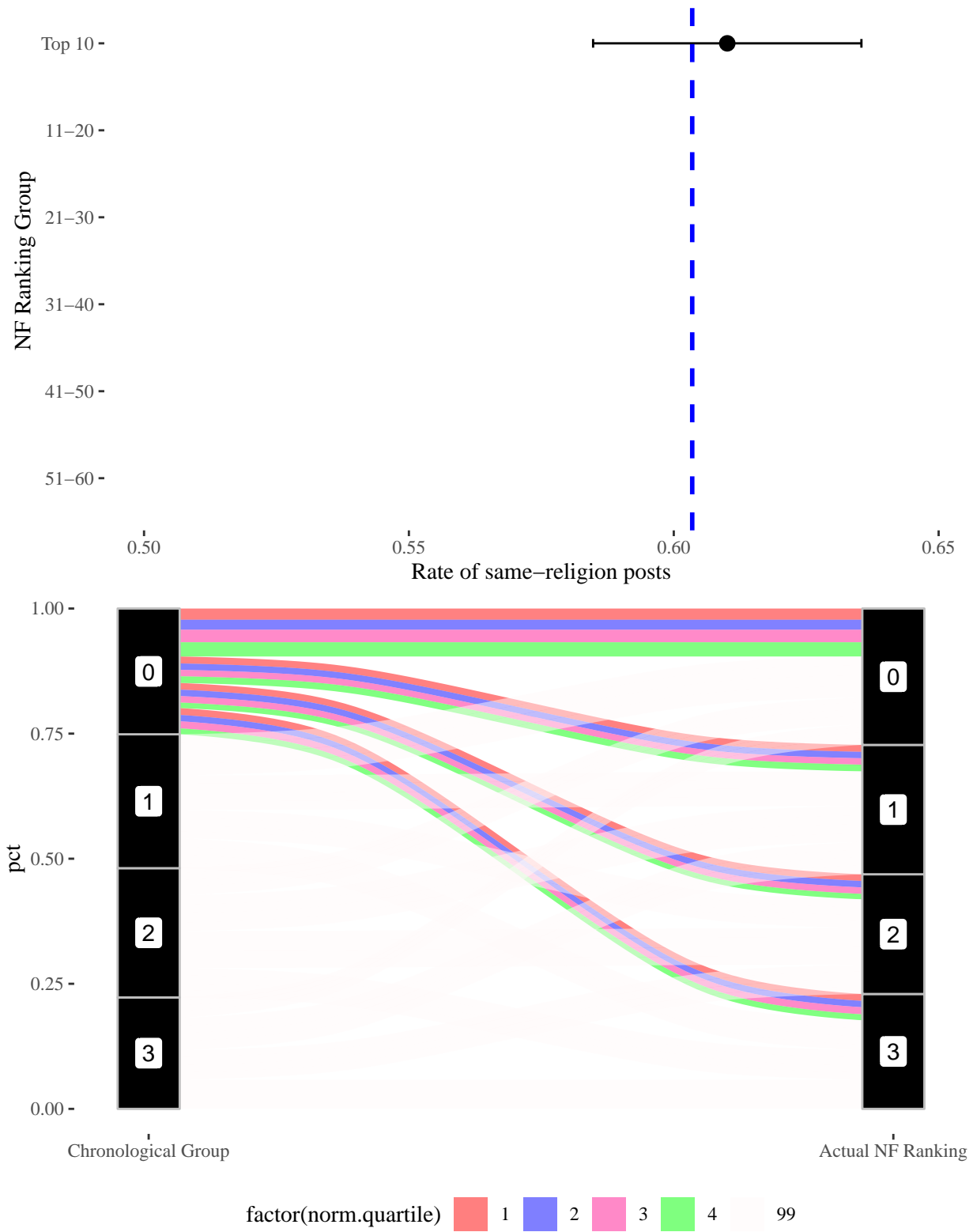
Same-religion posts get sorted closer to the top

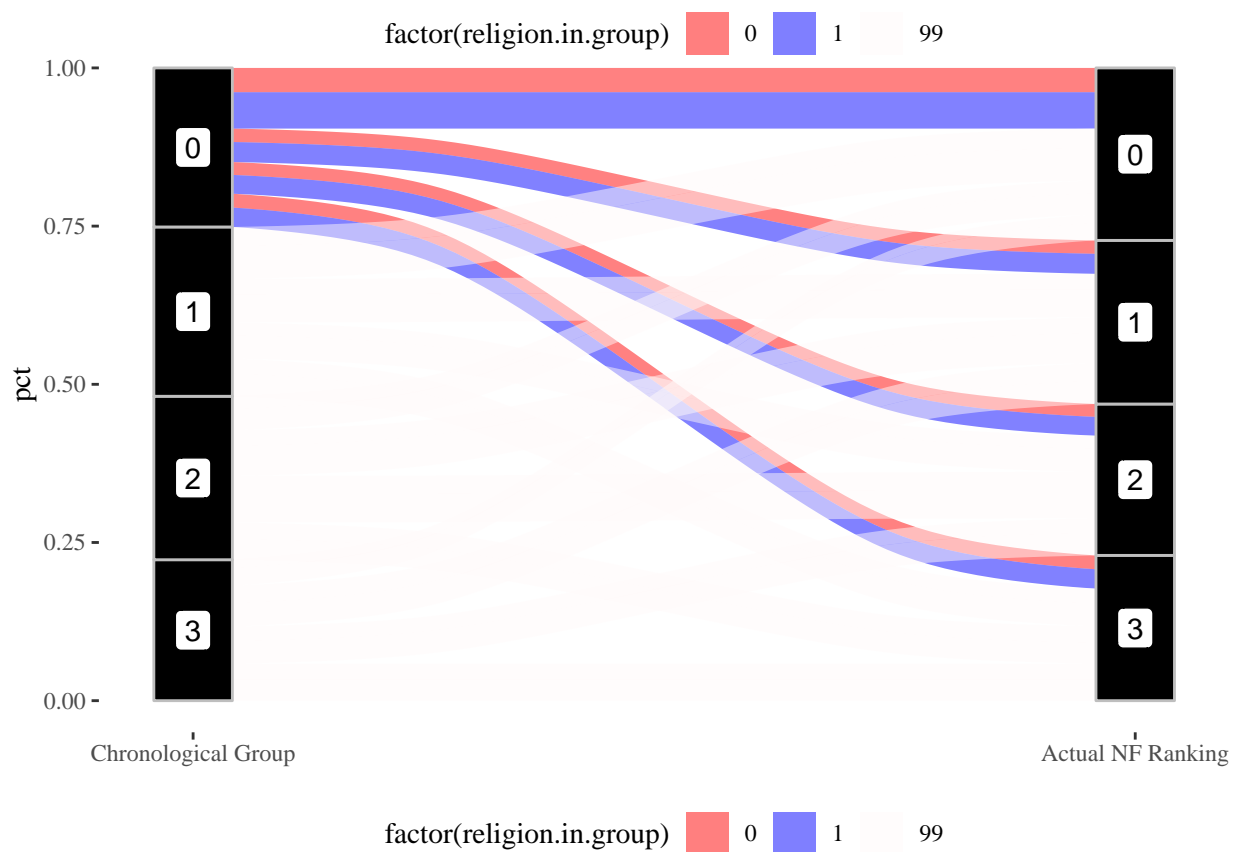
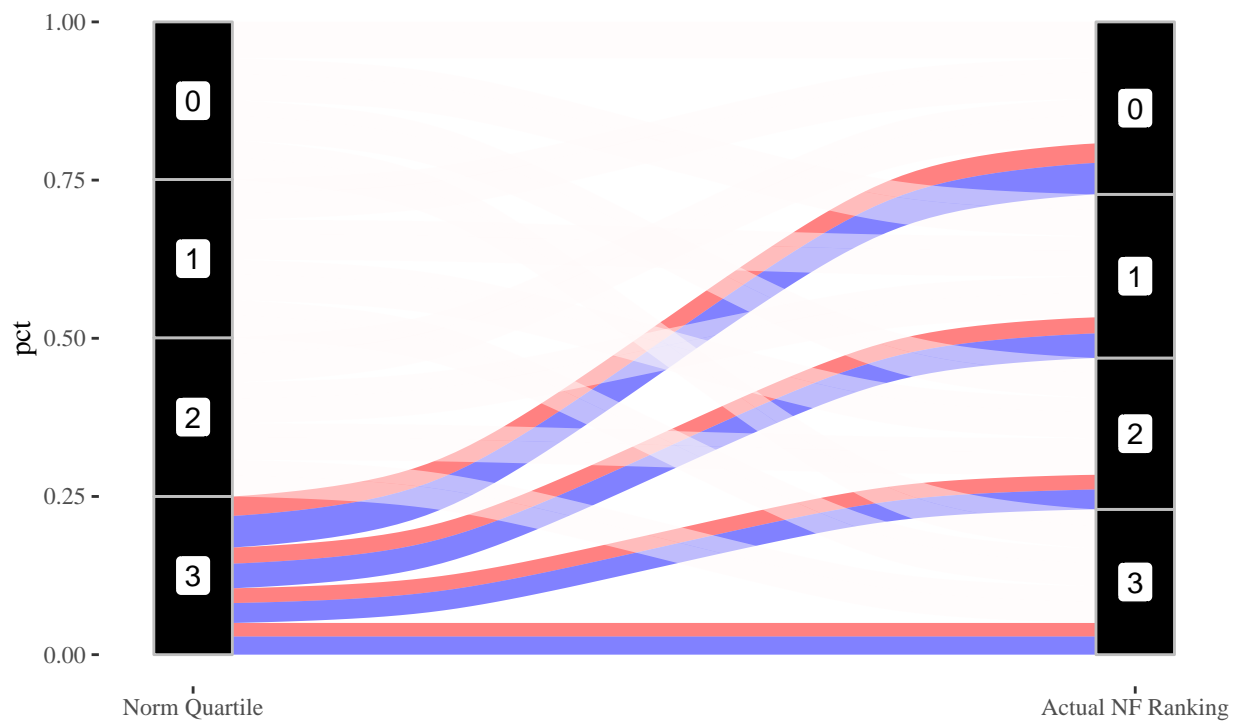


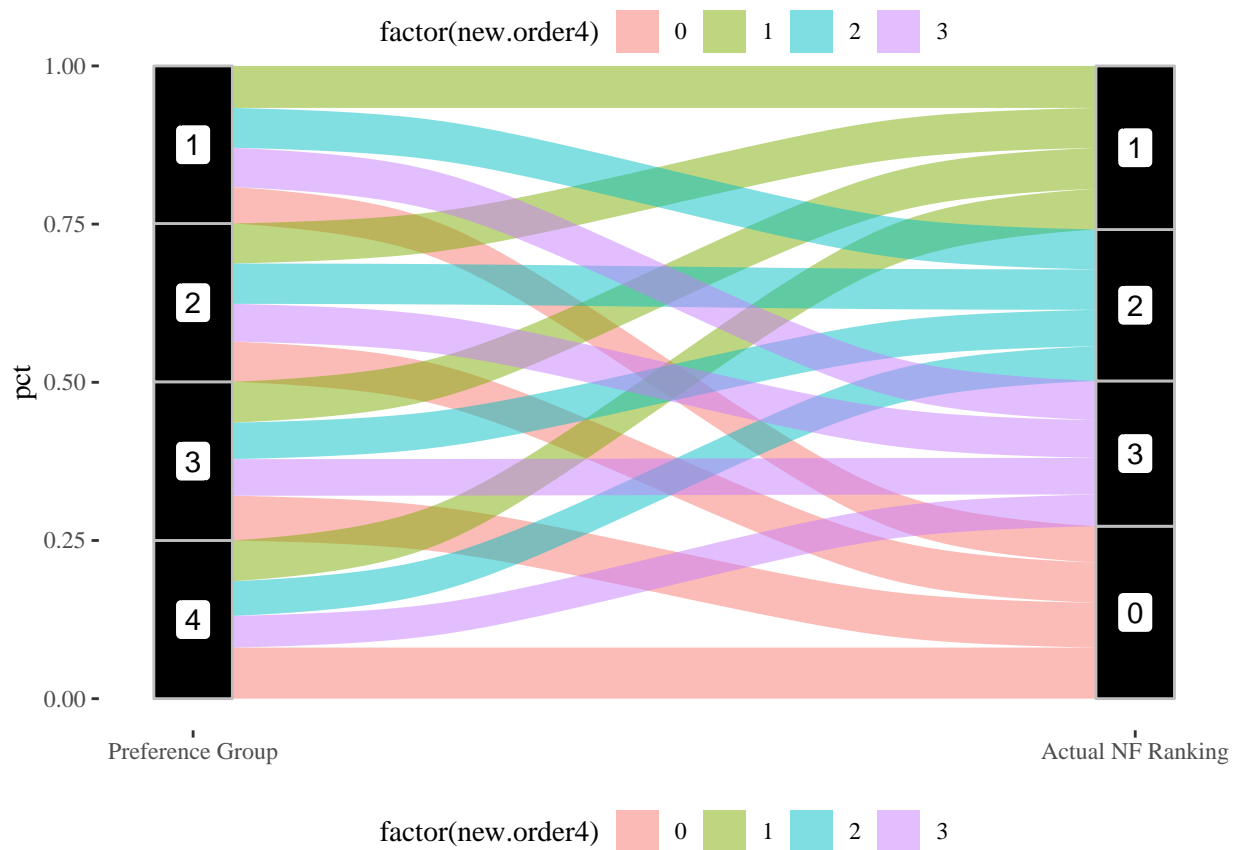
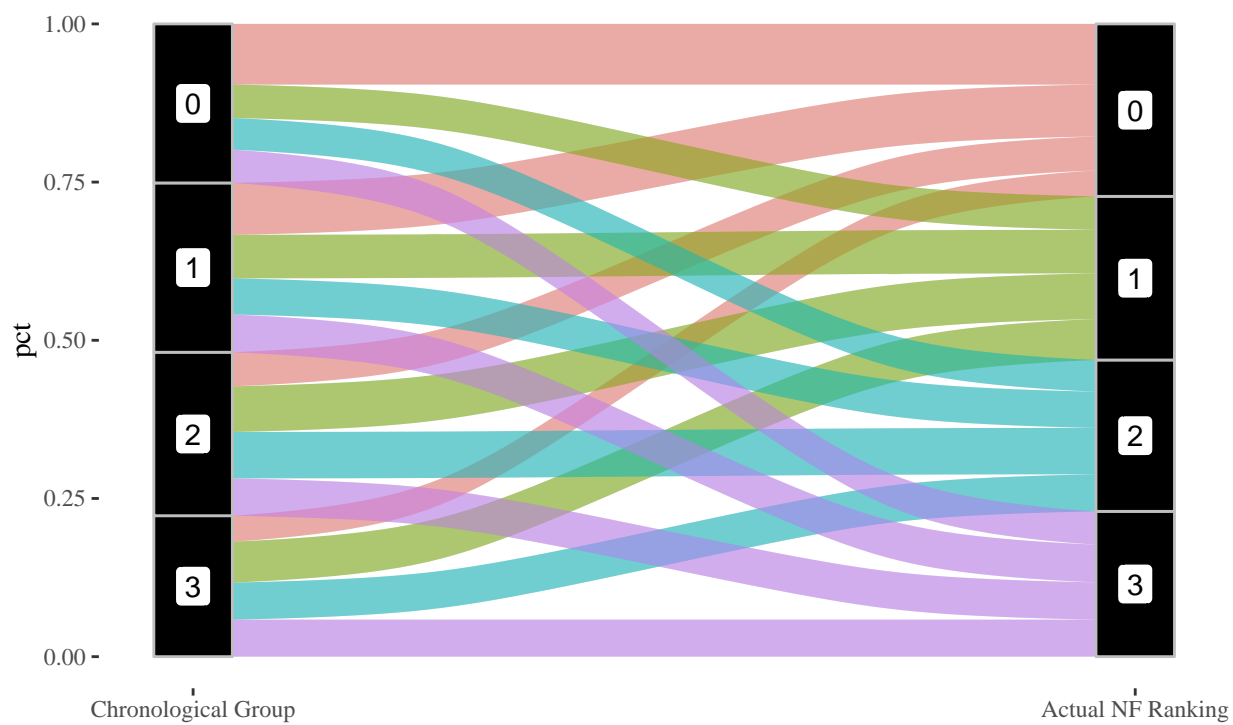
```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

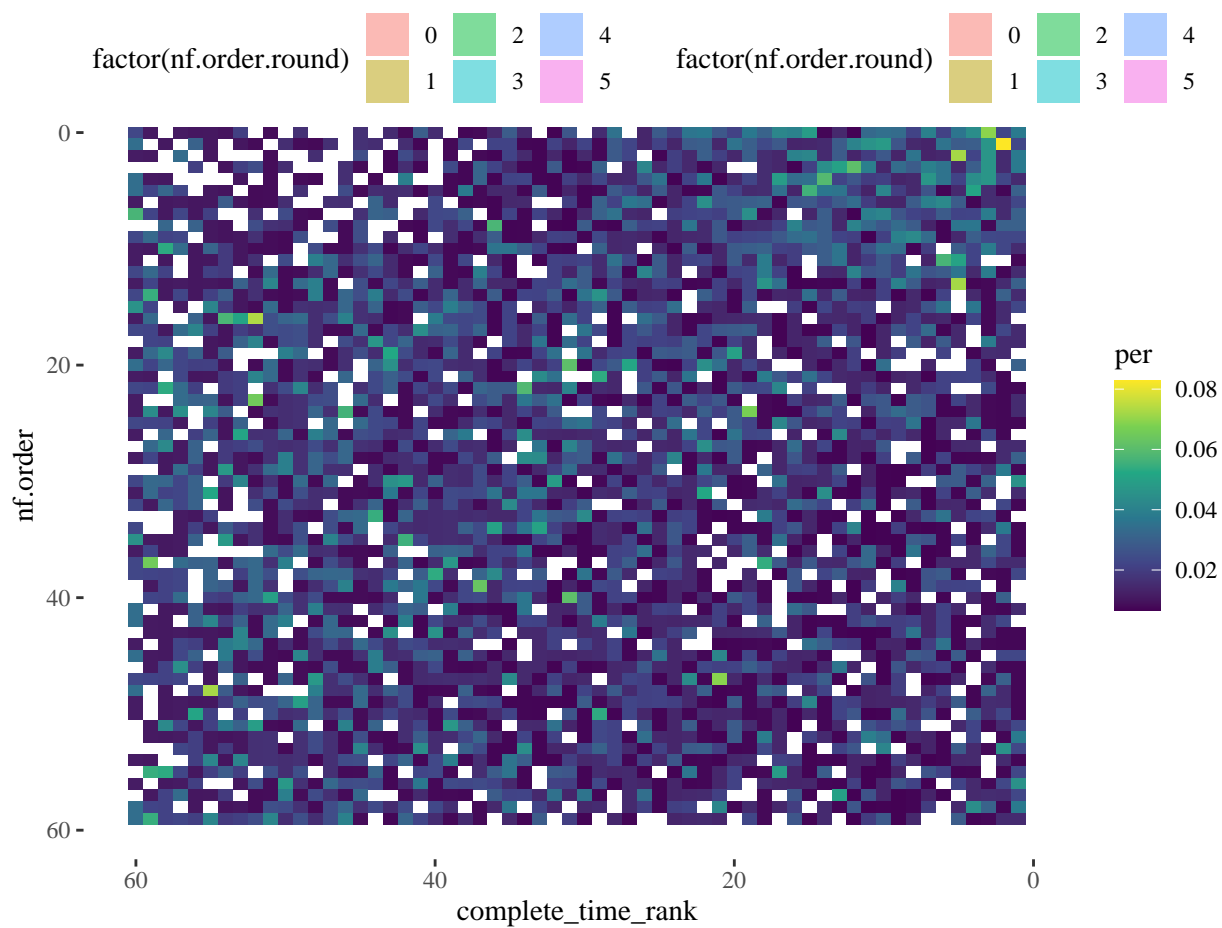
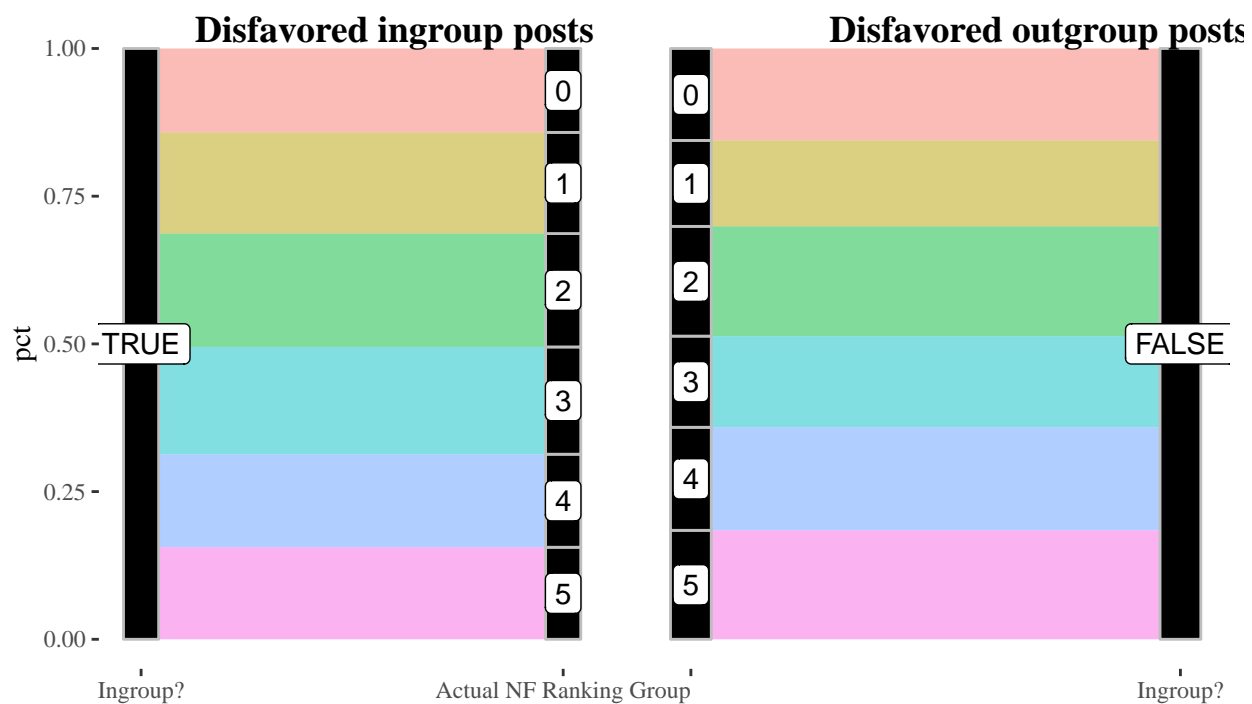
Newsfeed preference for user's religion (India)

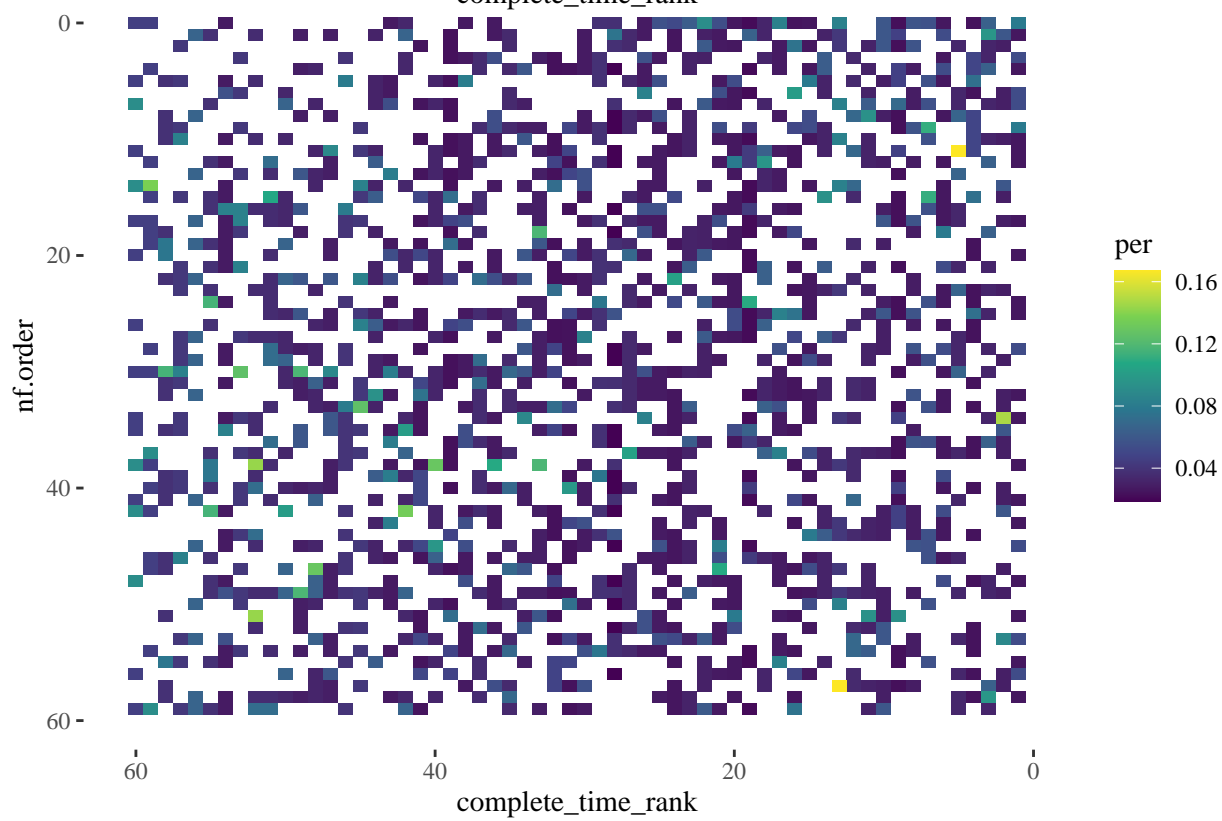
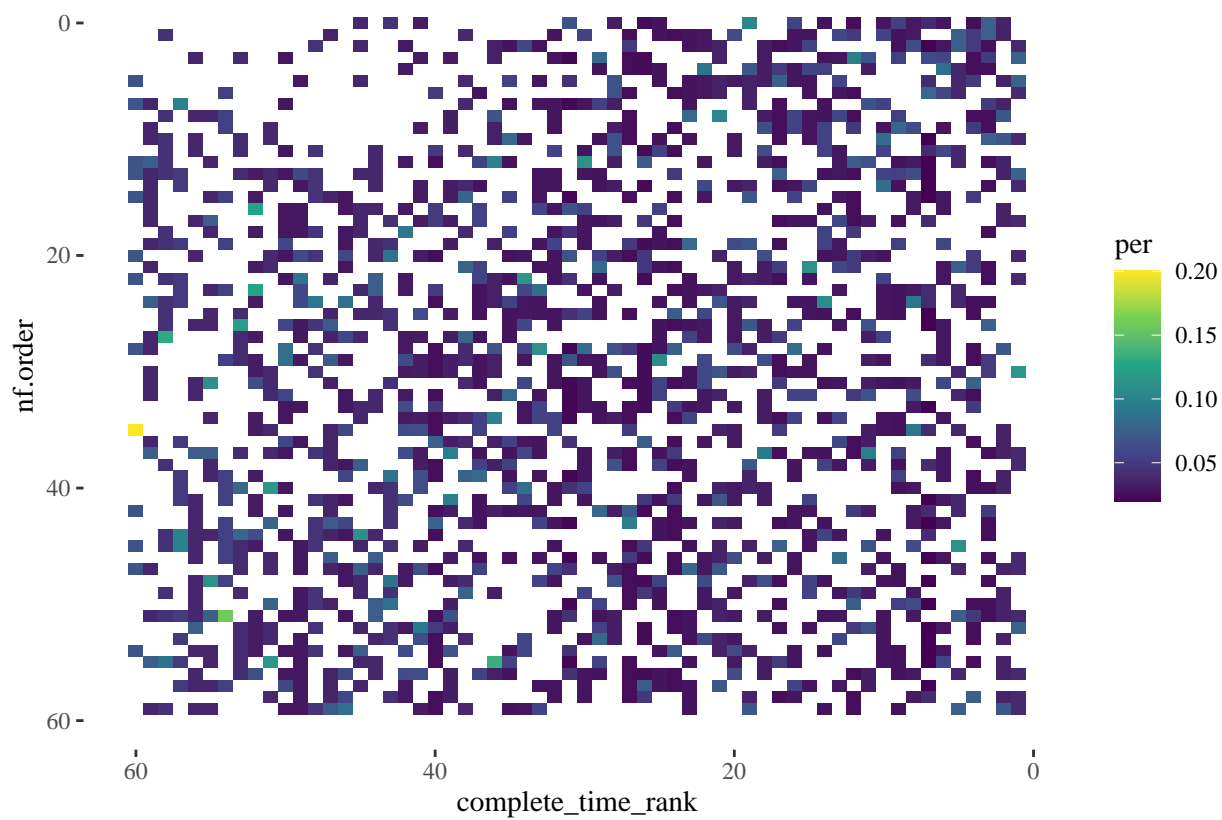
Same-religion posts get sorted closer to the top

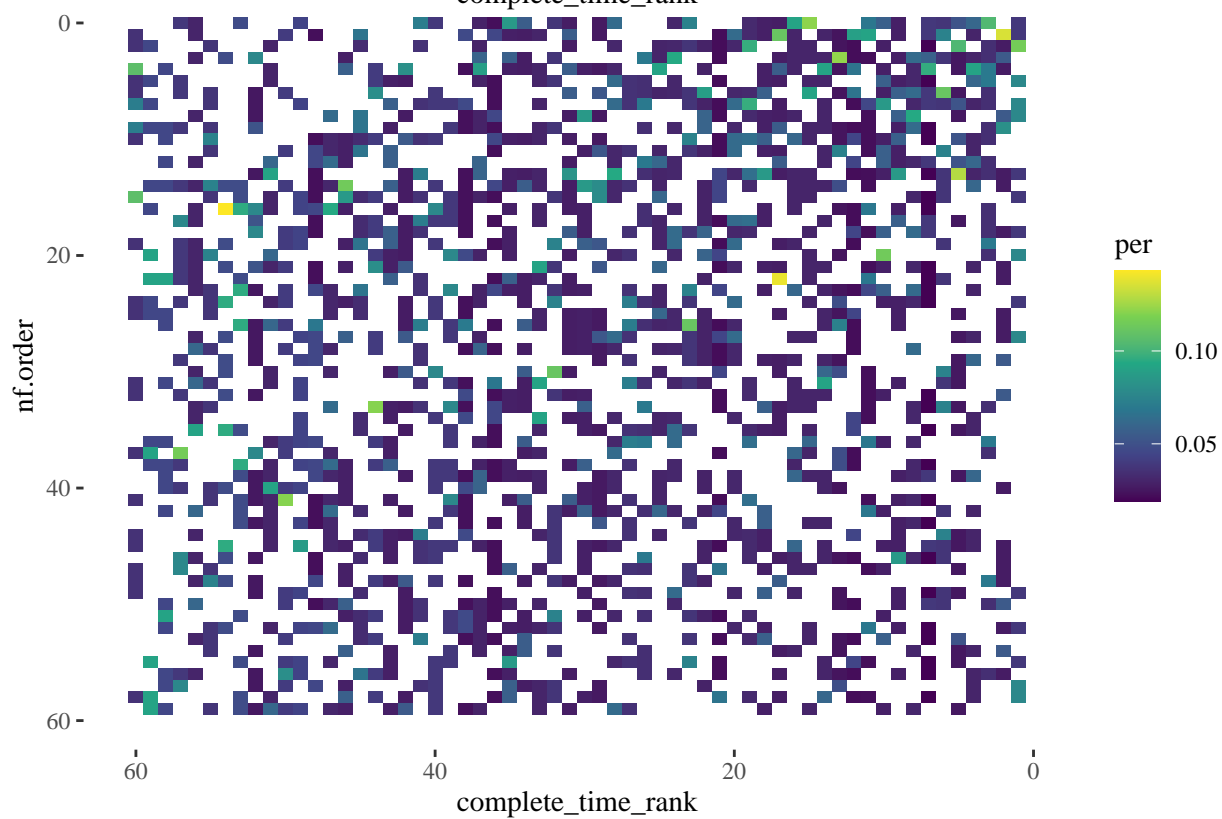
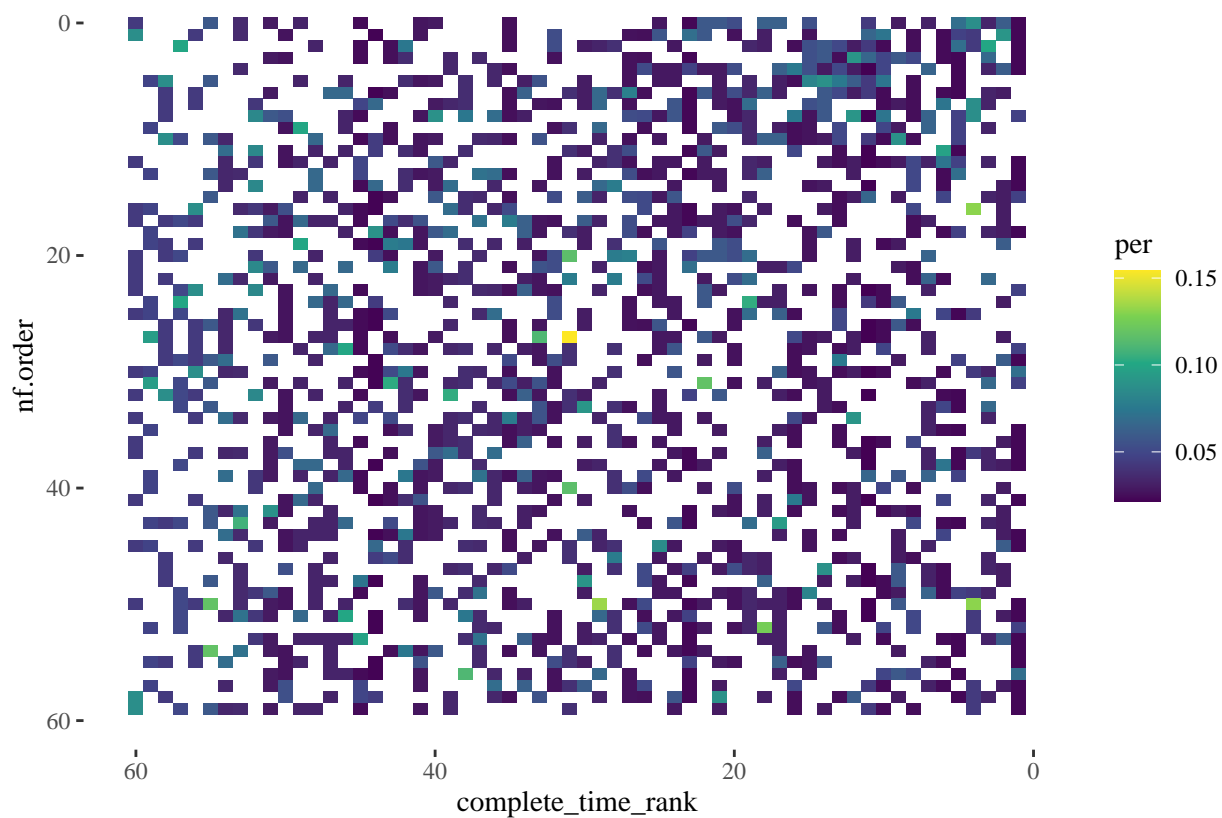


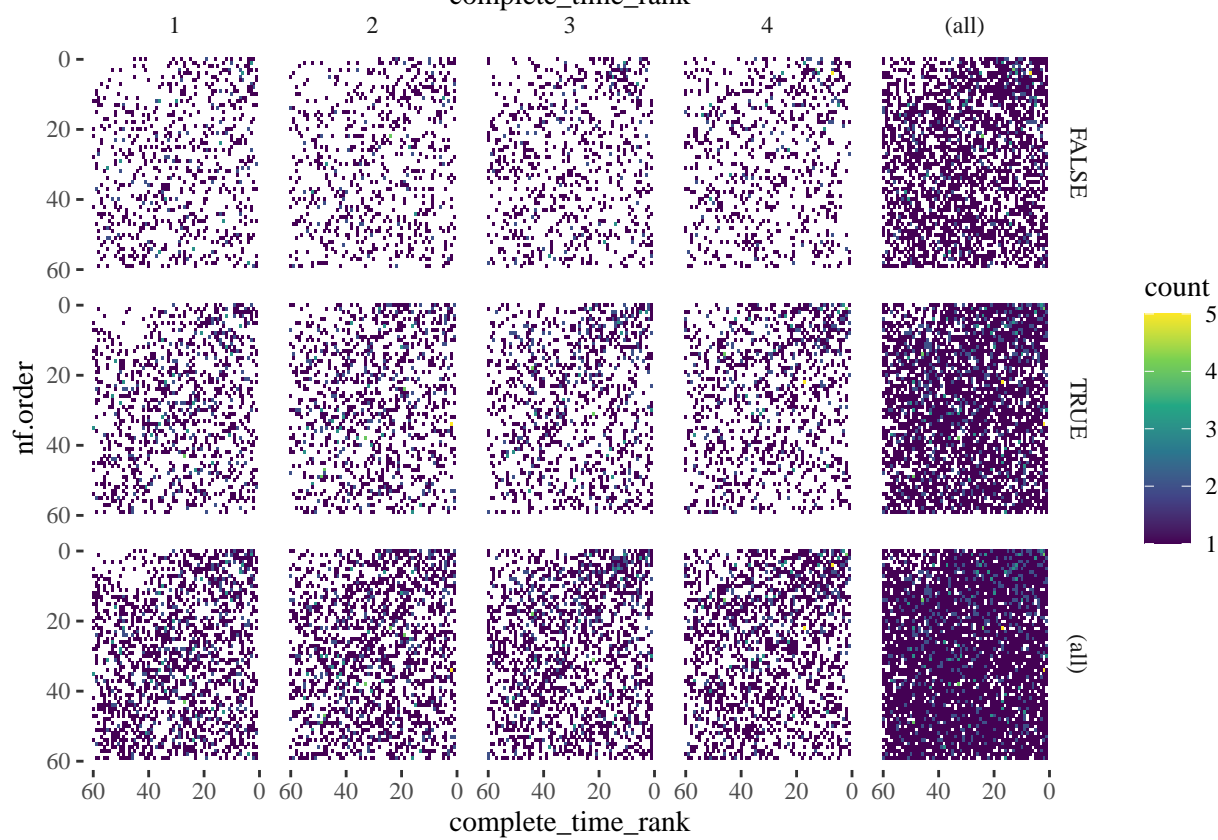
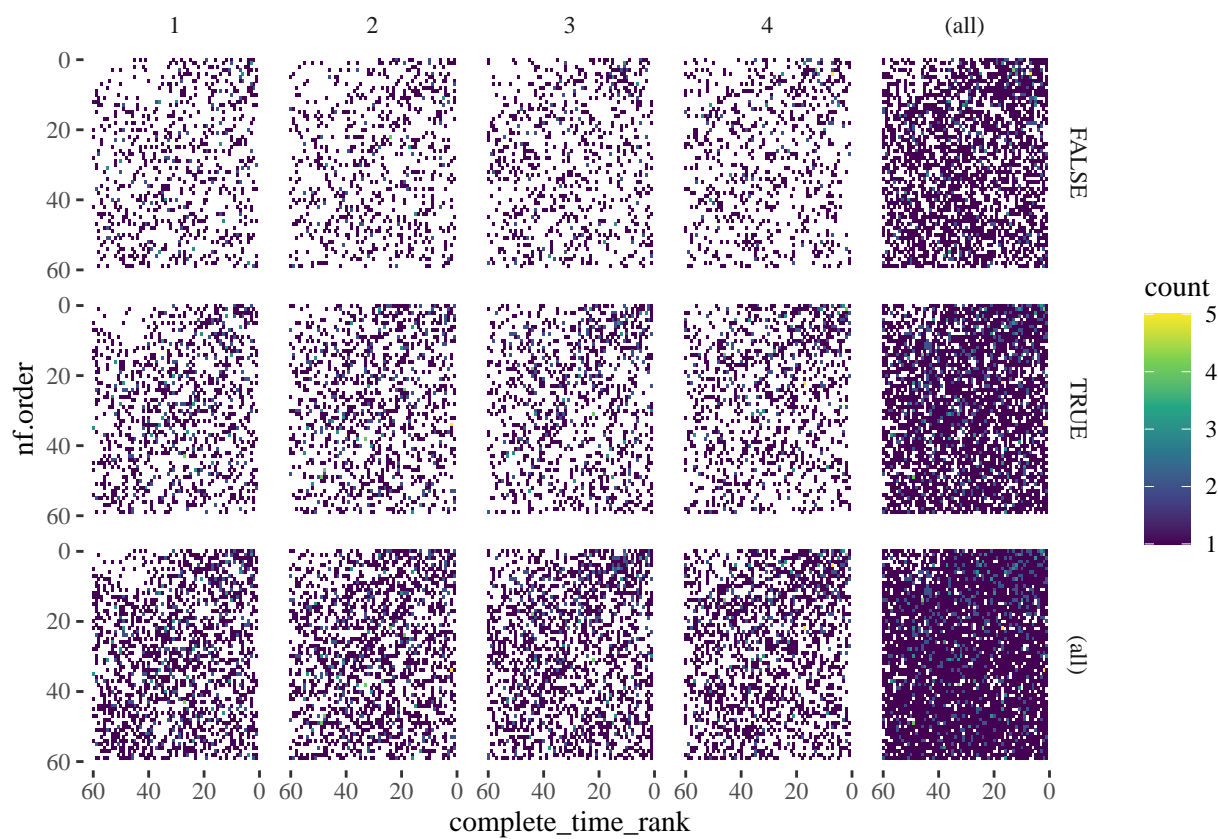


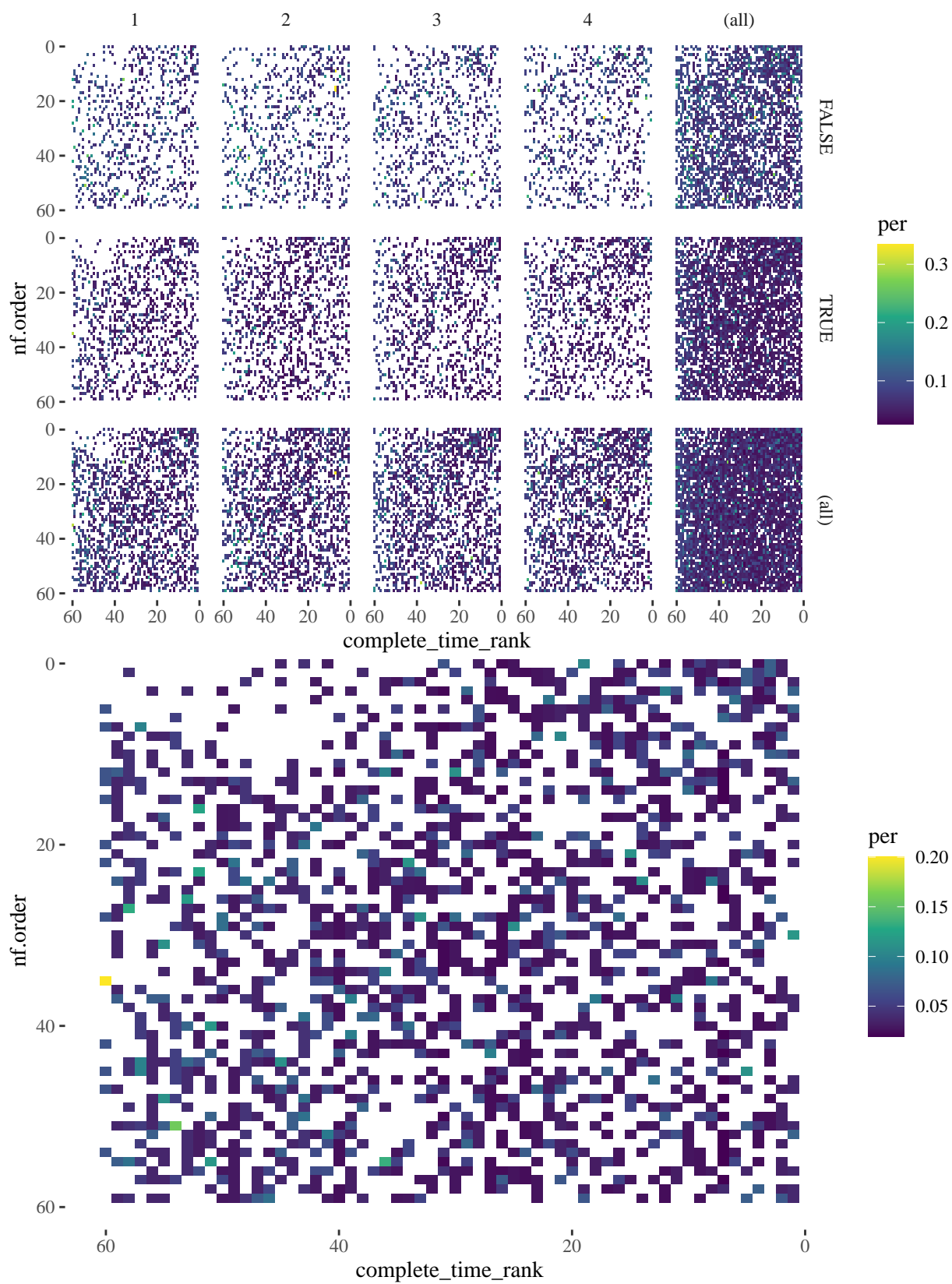


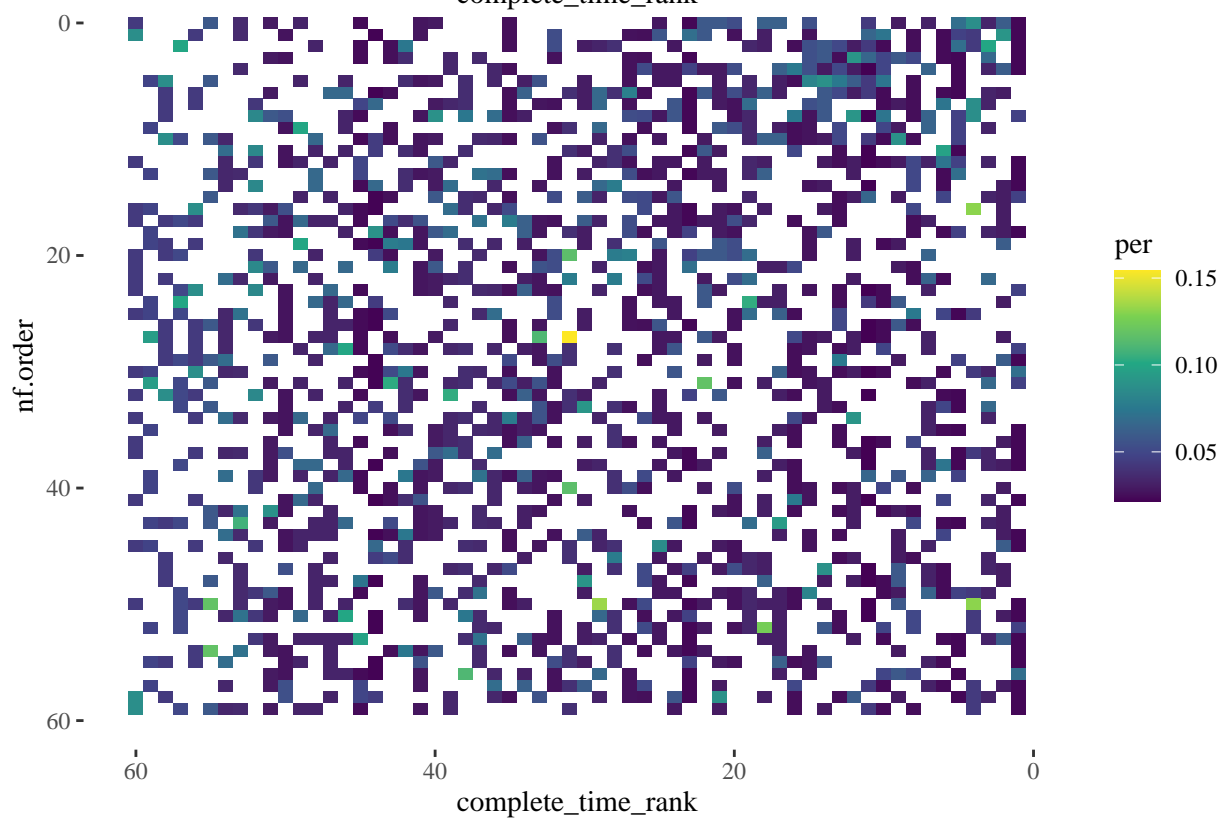
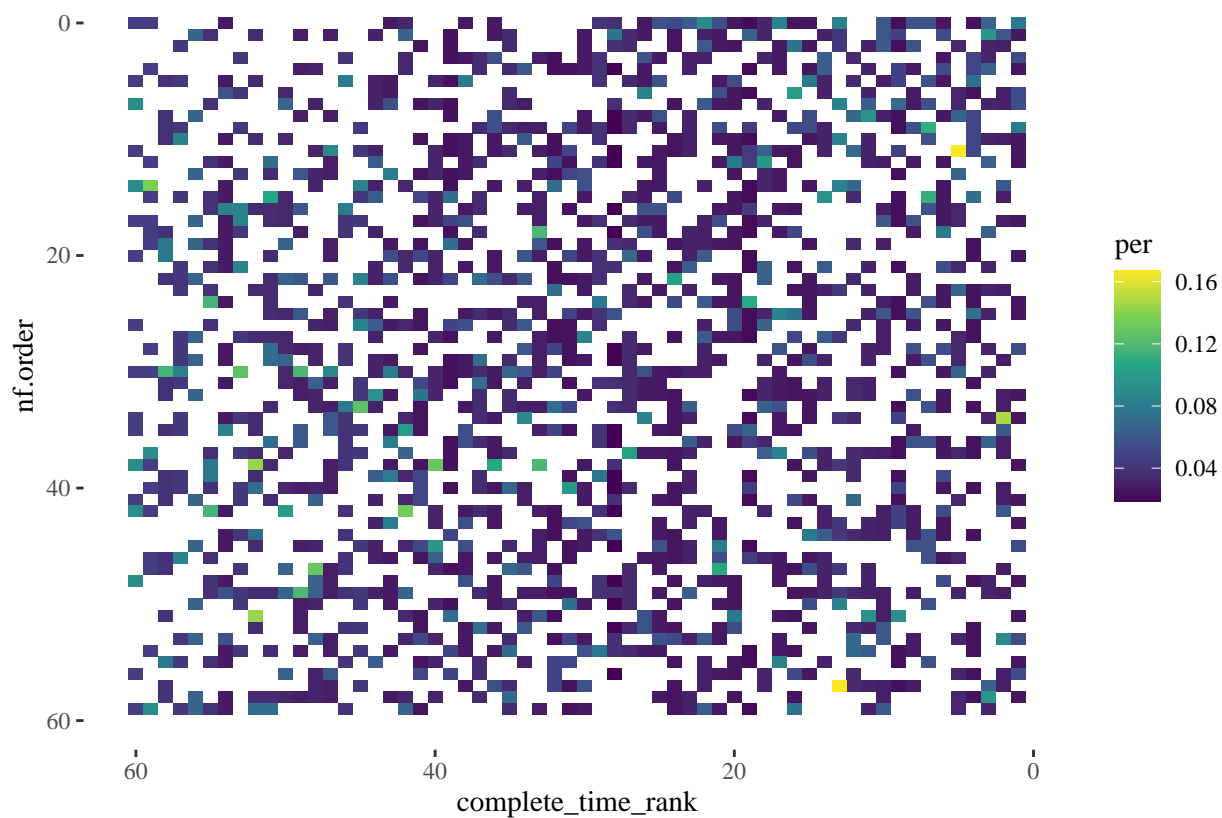


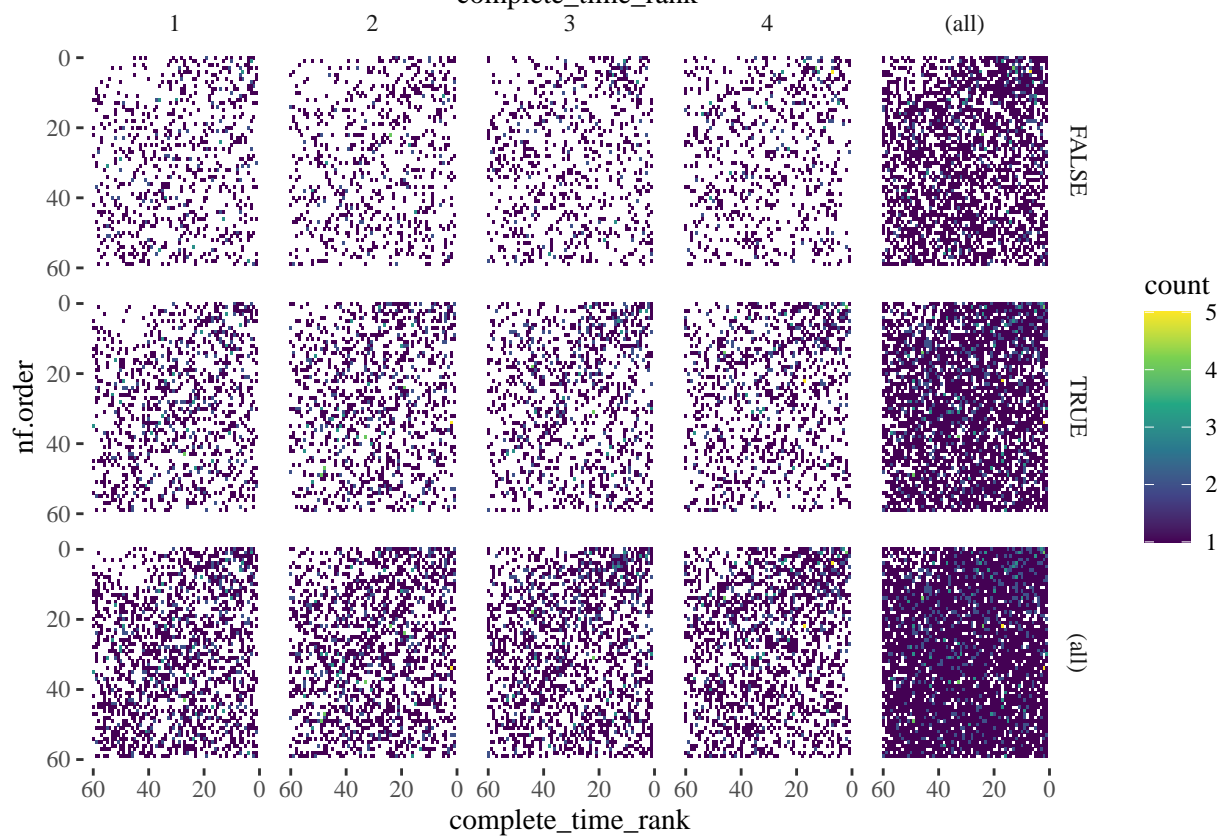
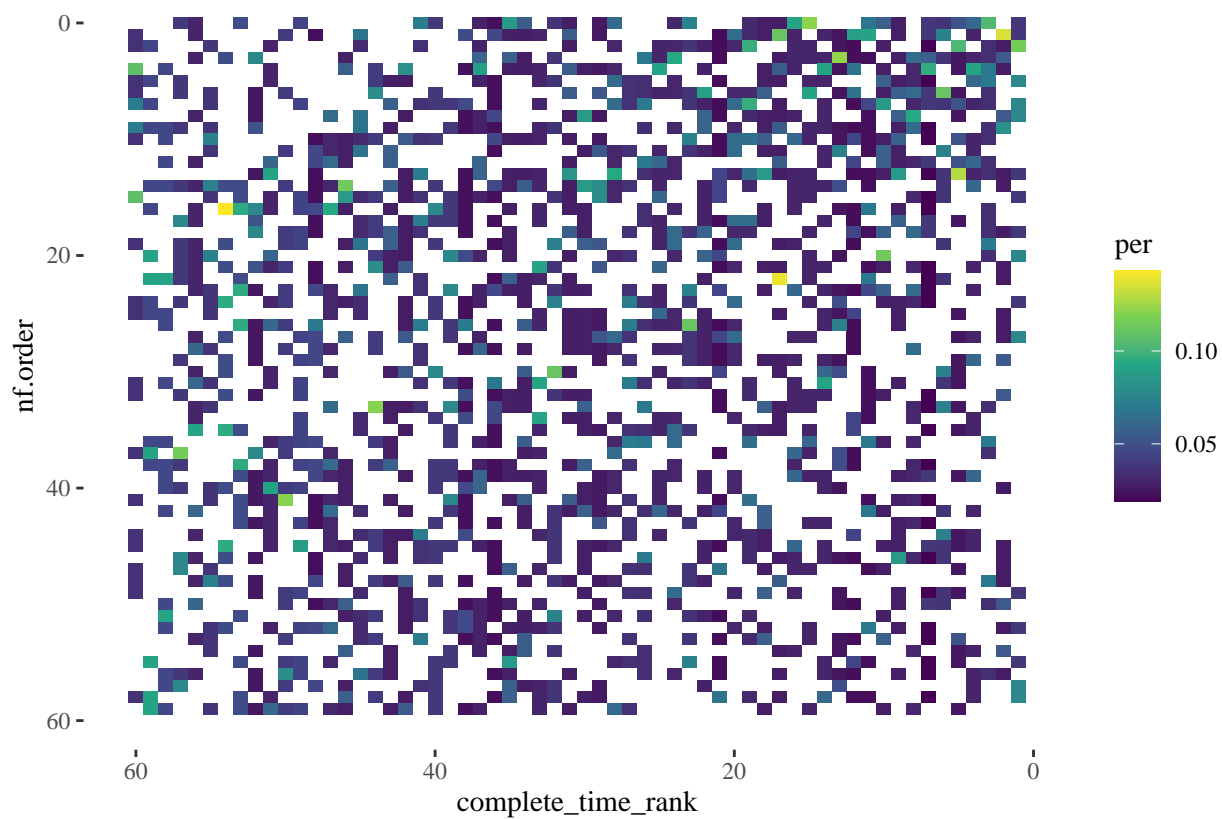


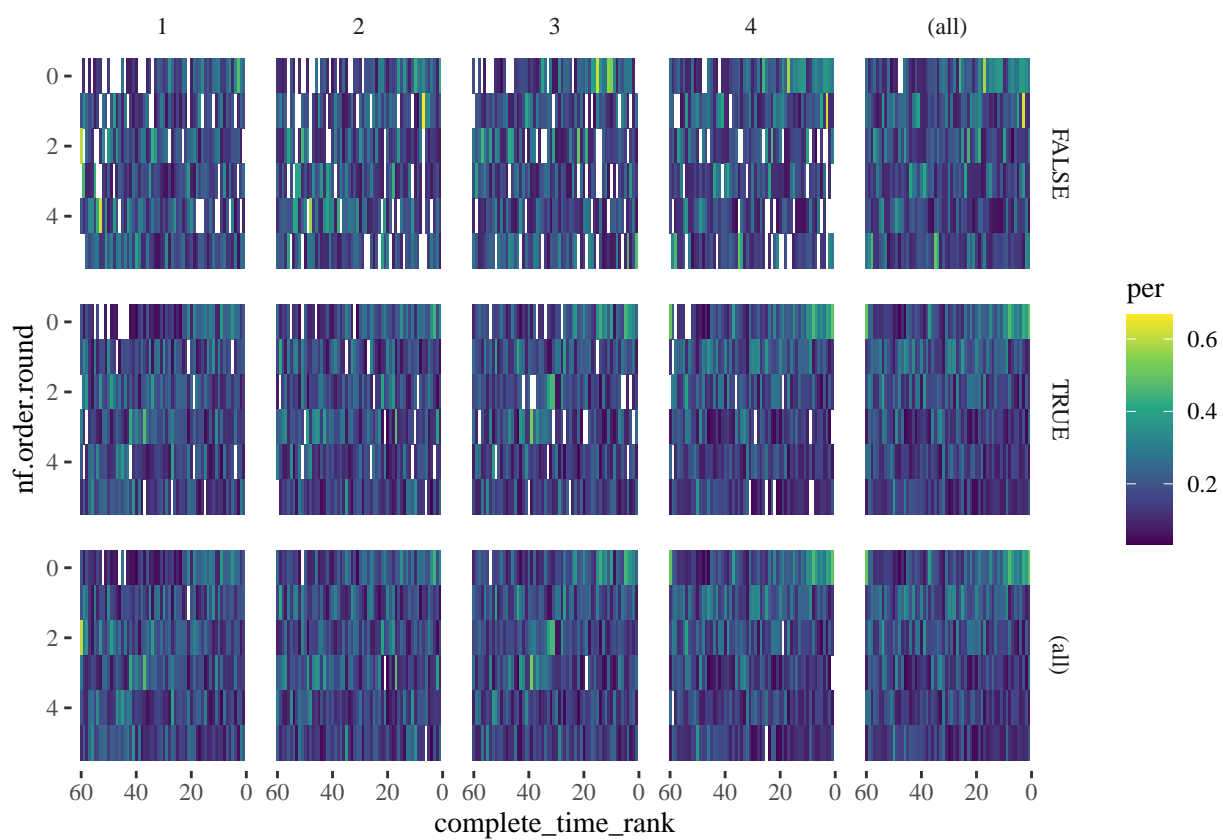
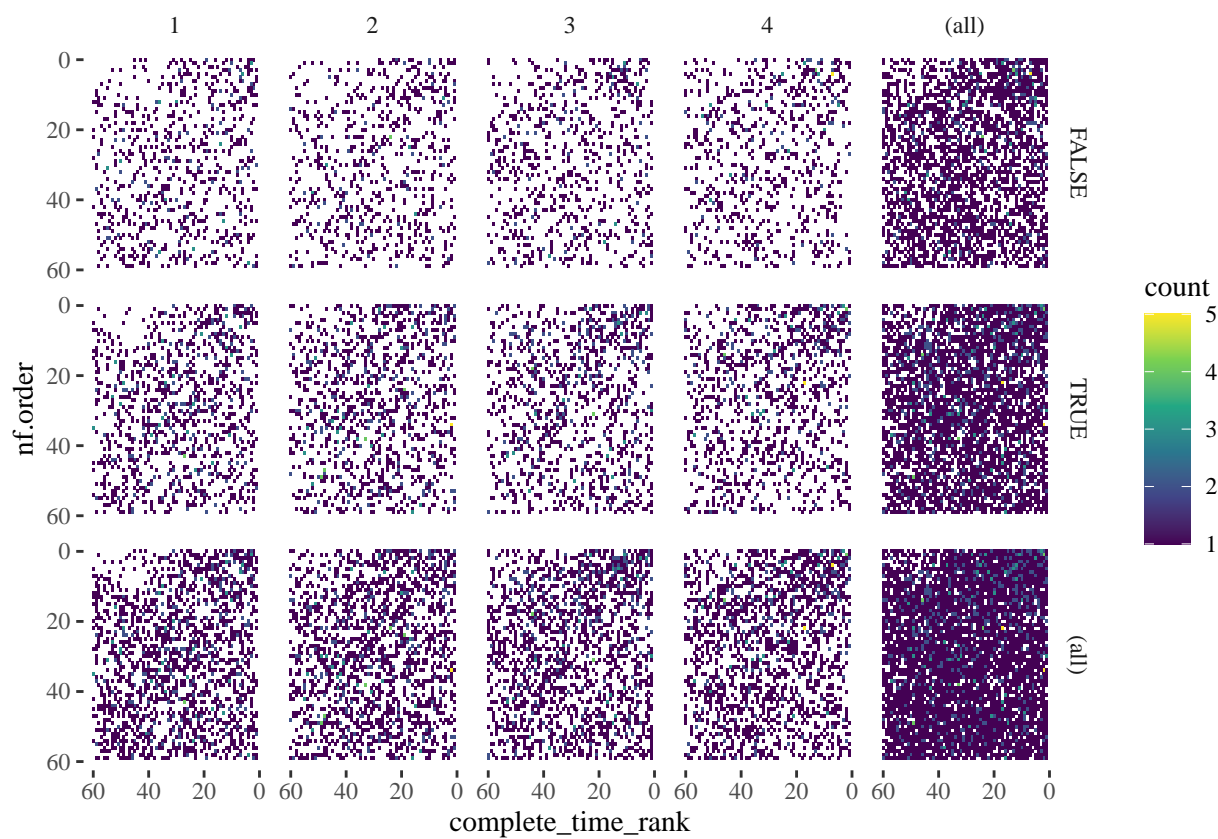


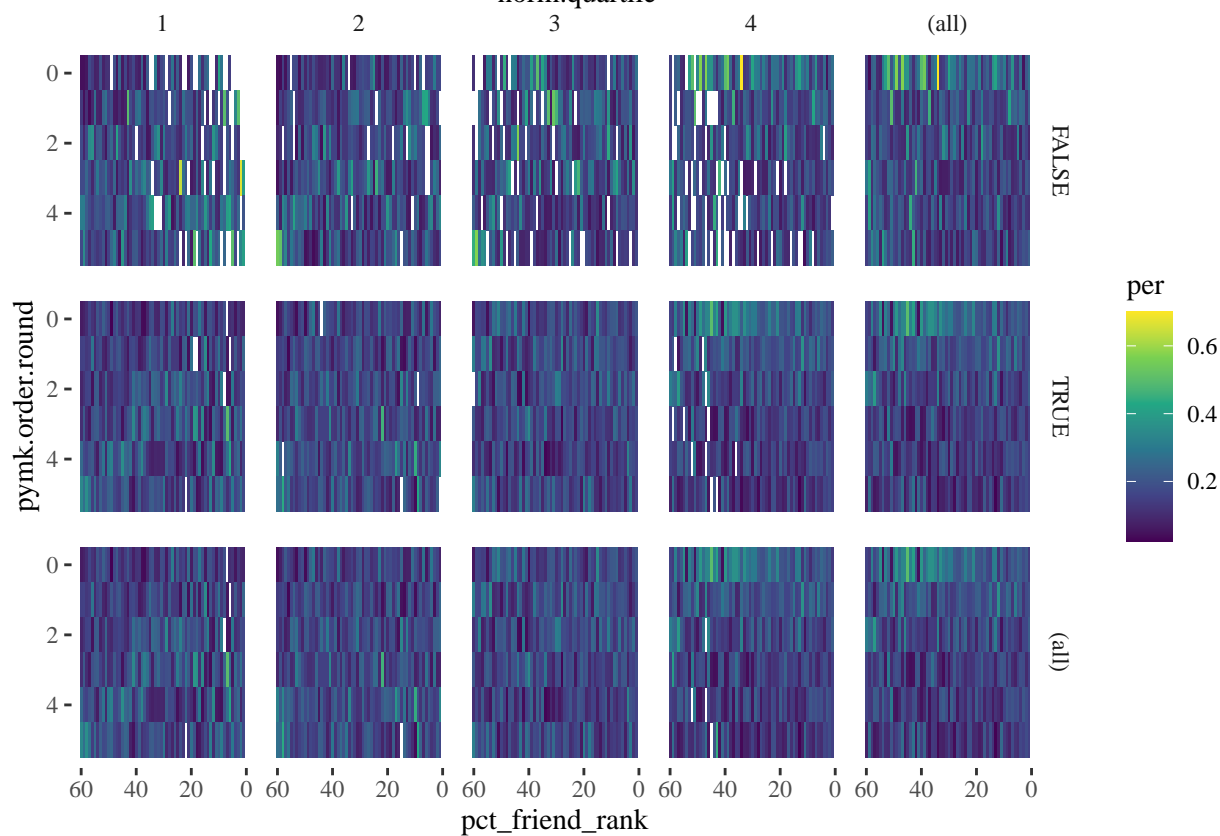
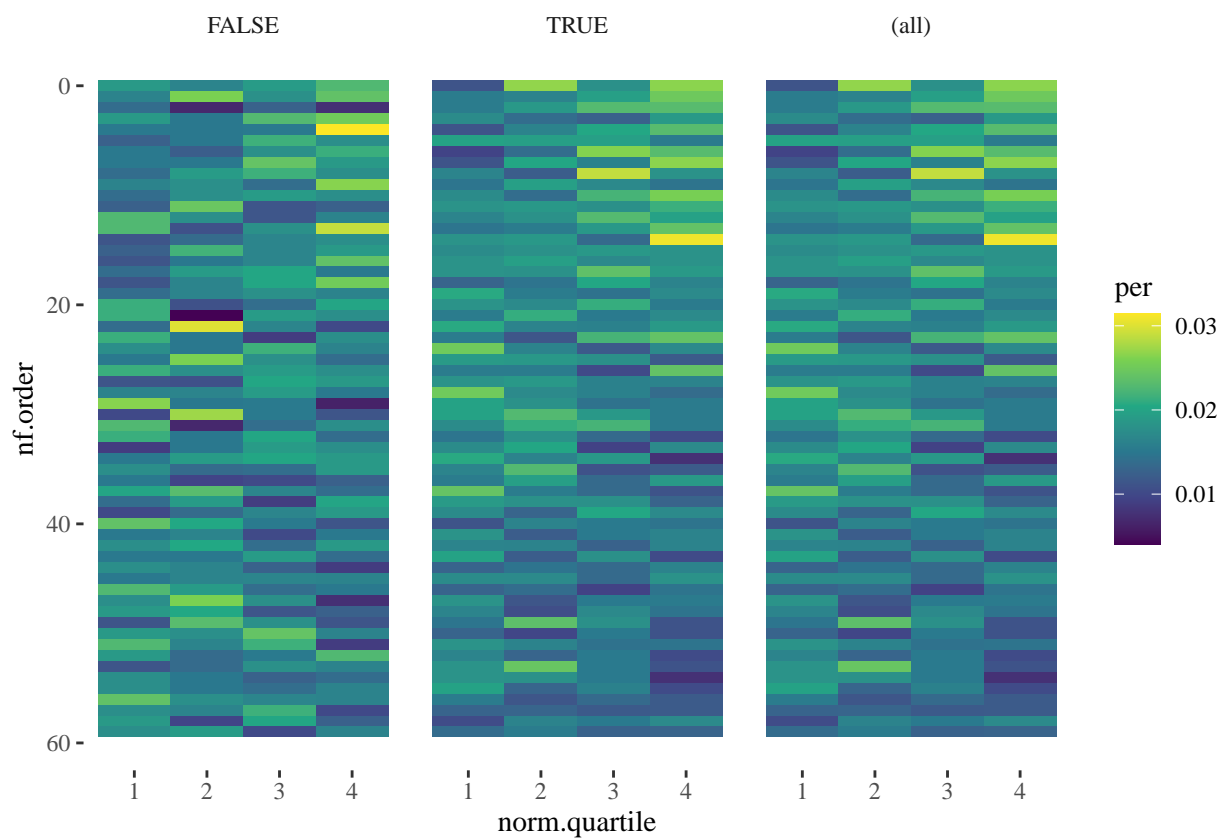


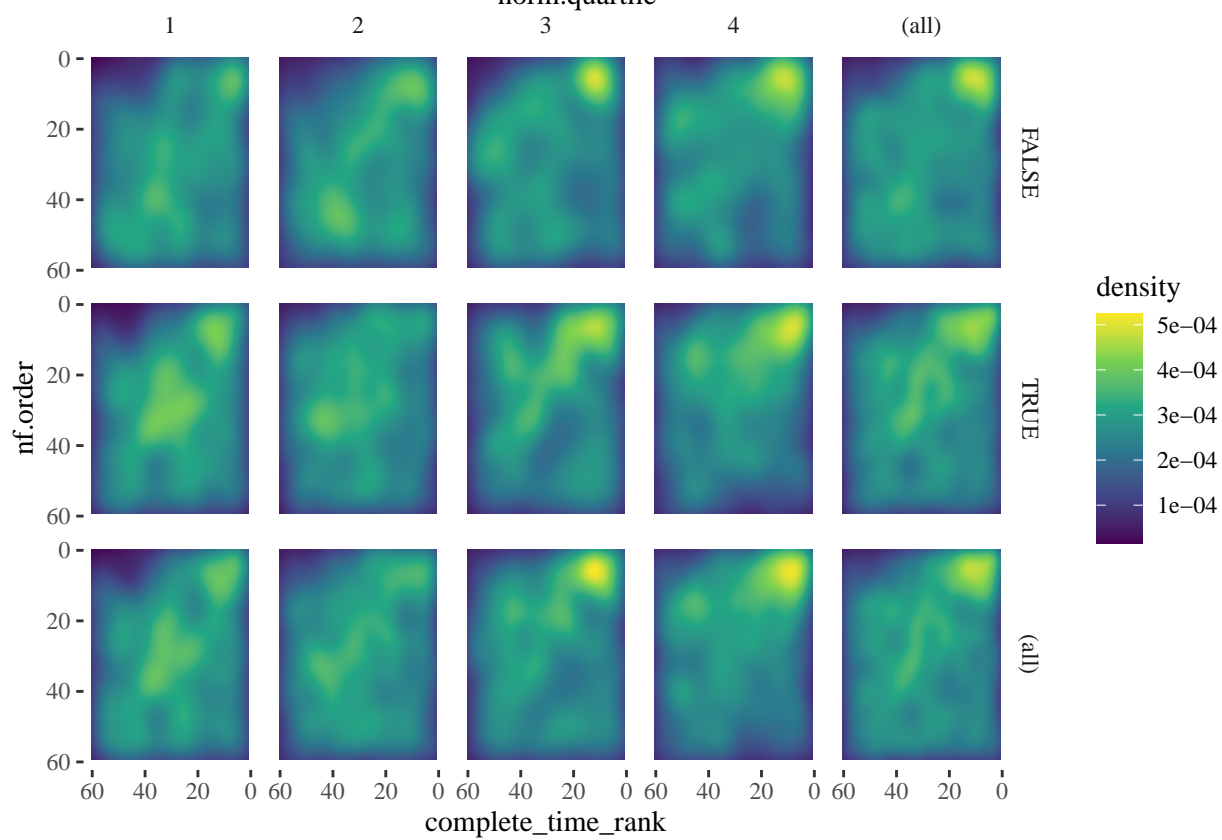
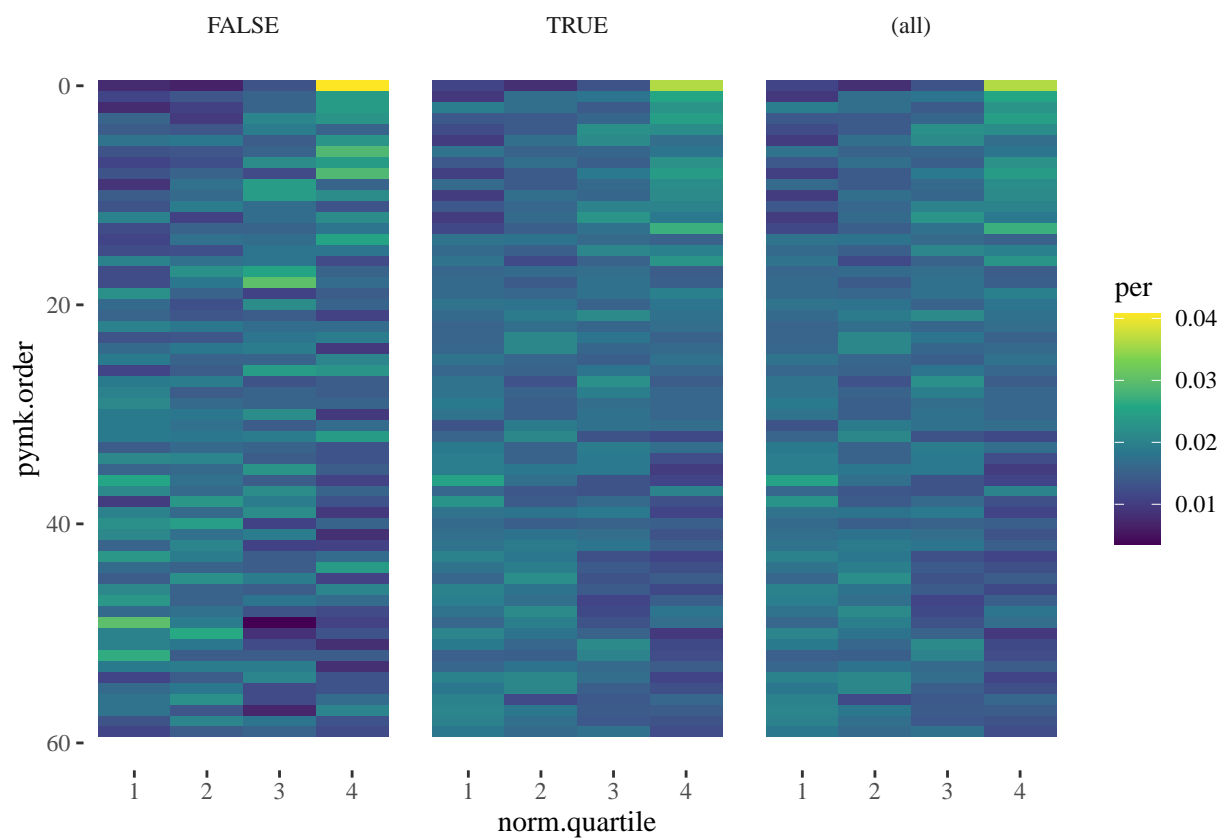


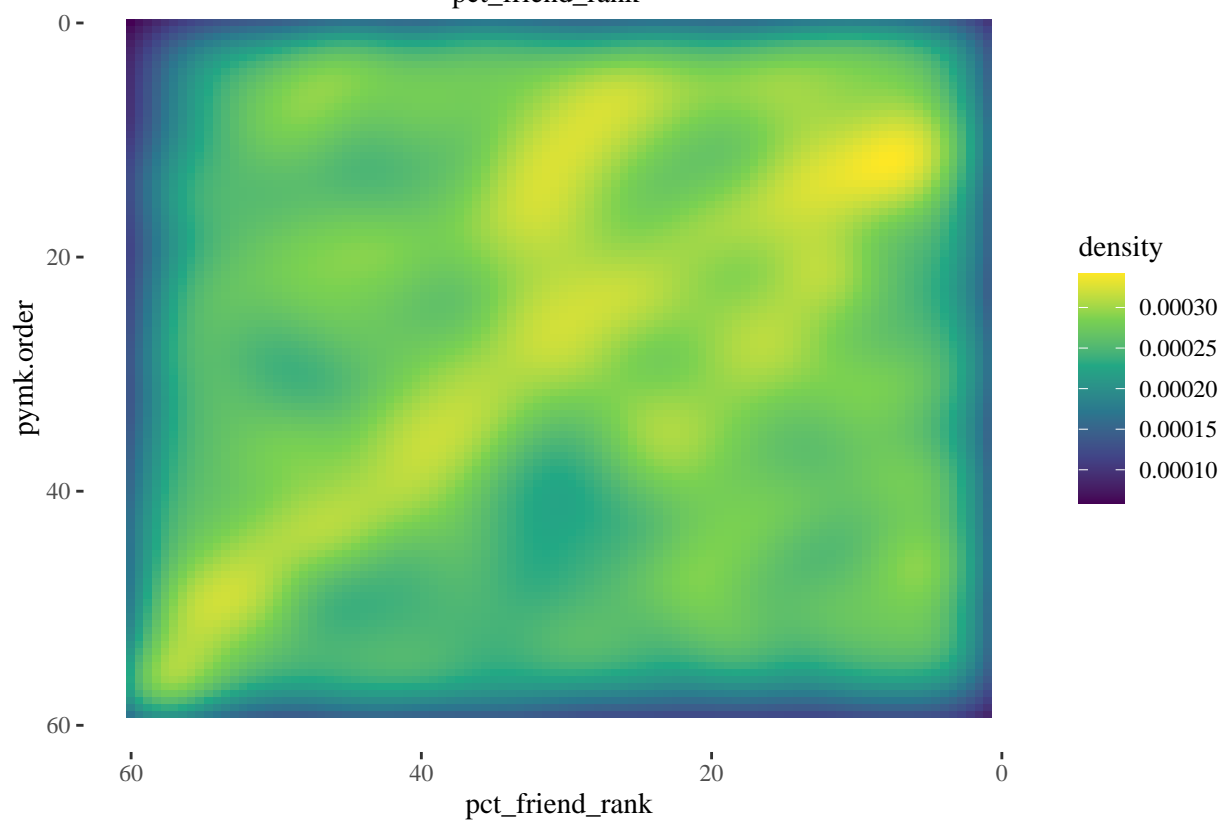
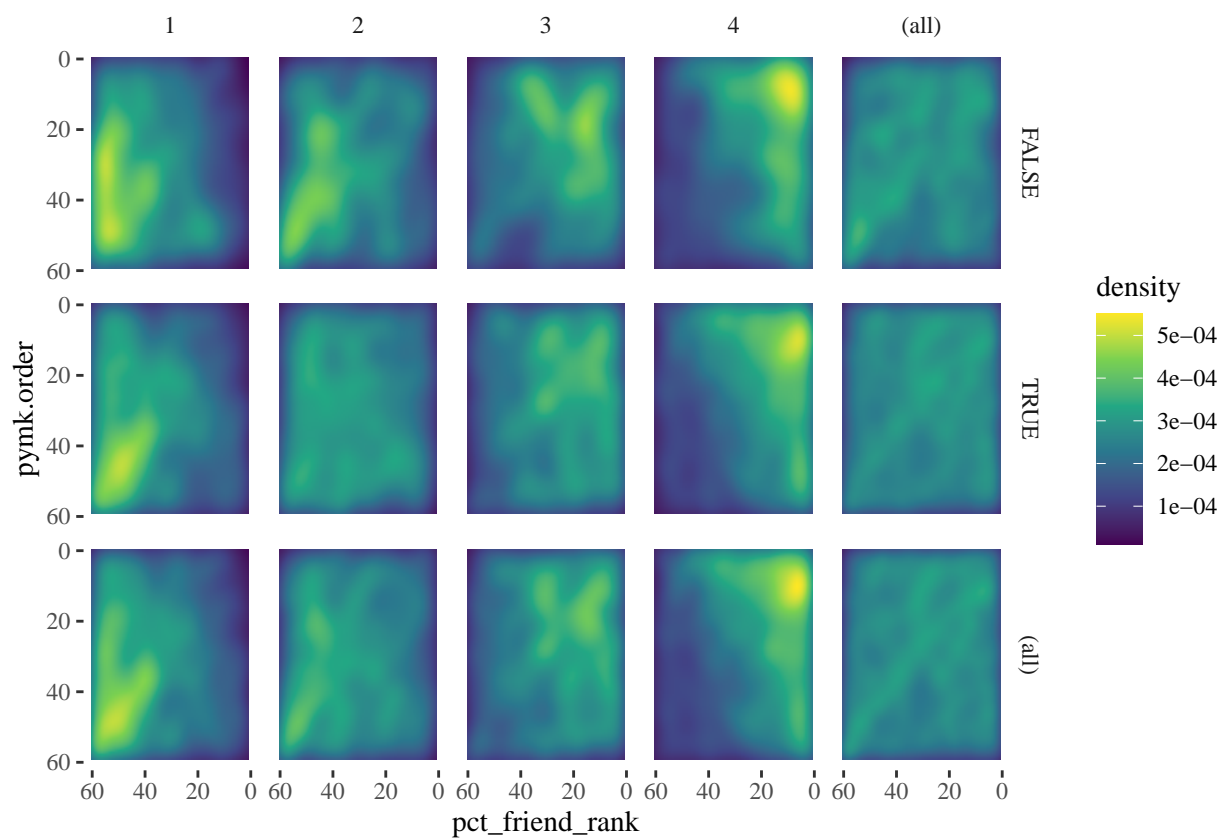




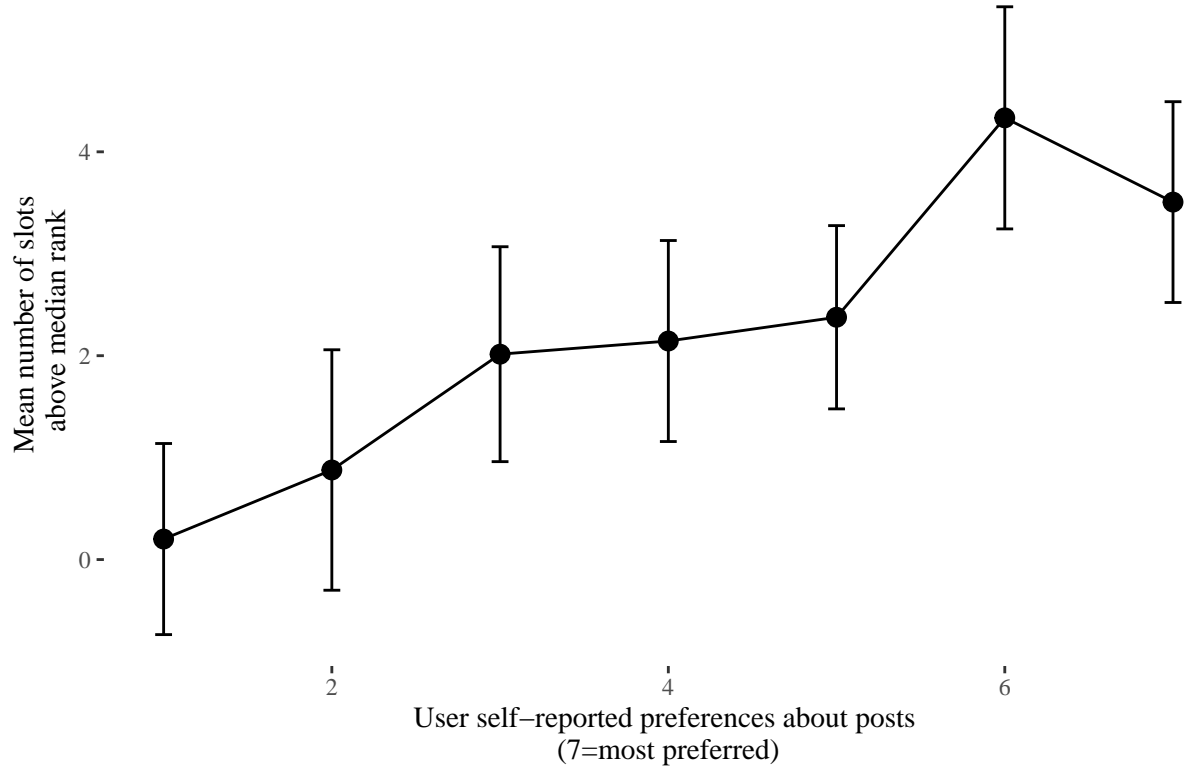




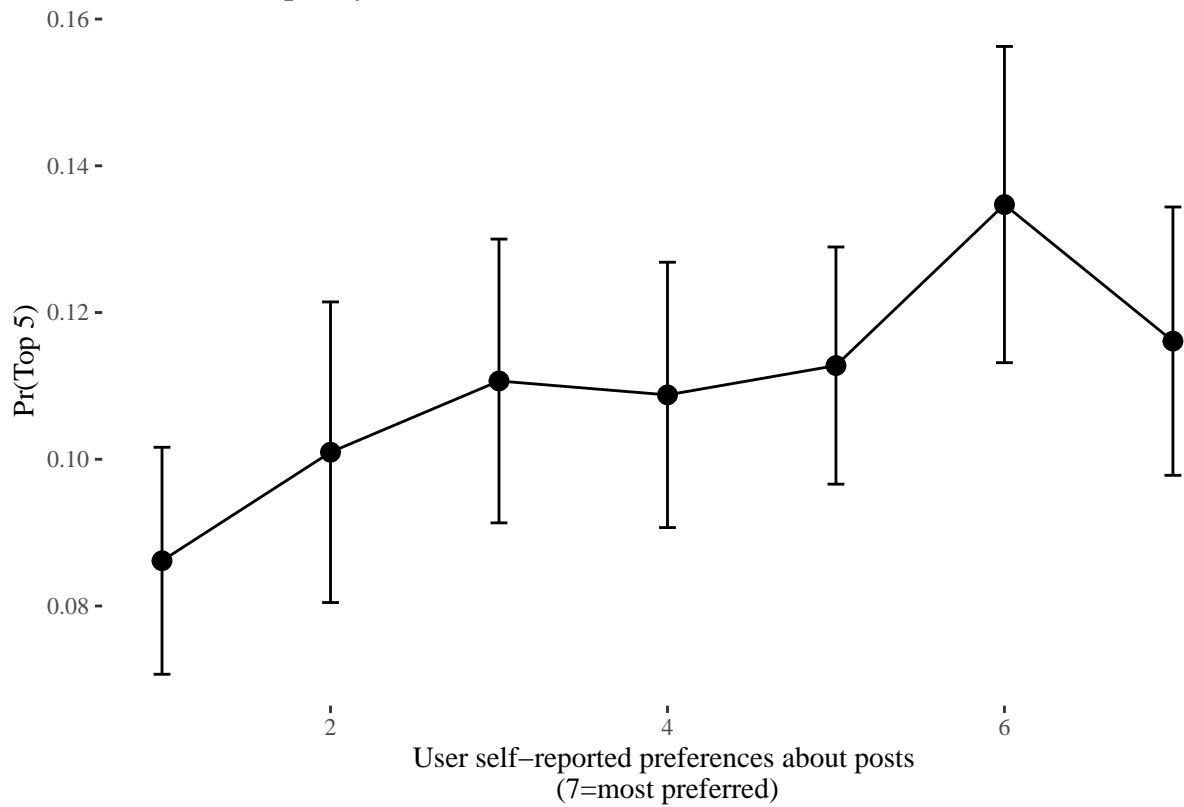




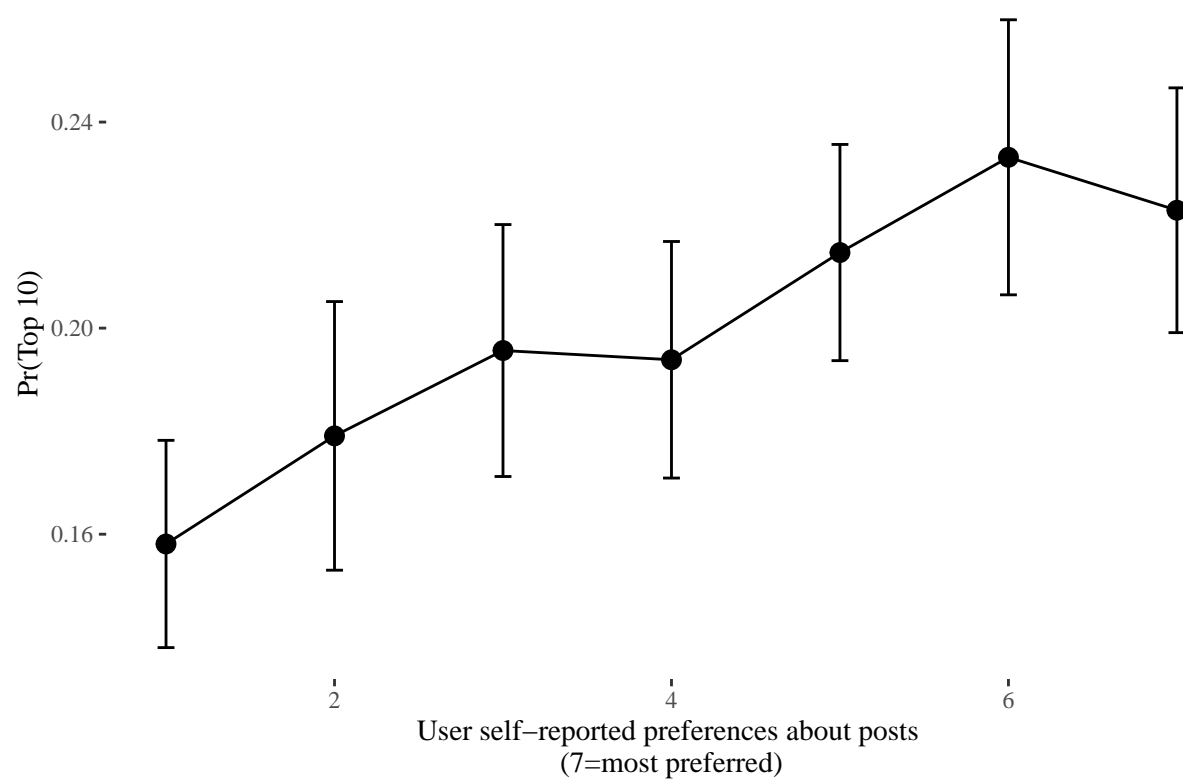
Newsfeed Ranking by Preference (India)



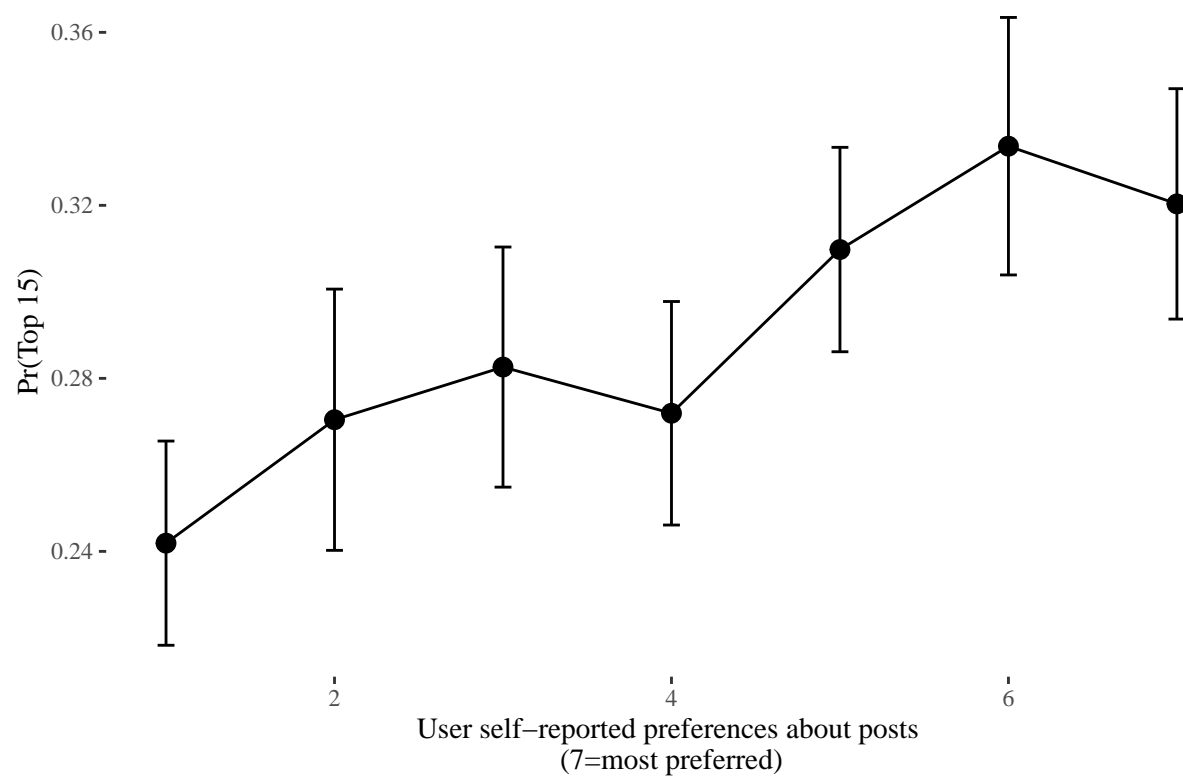
Newsfeed Top 5 by Preference (India)



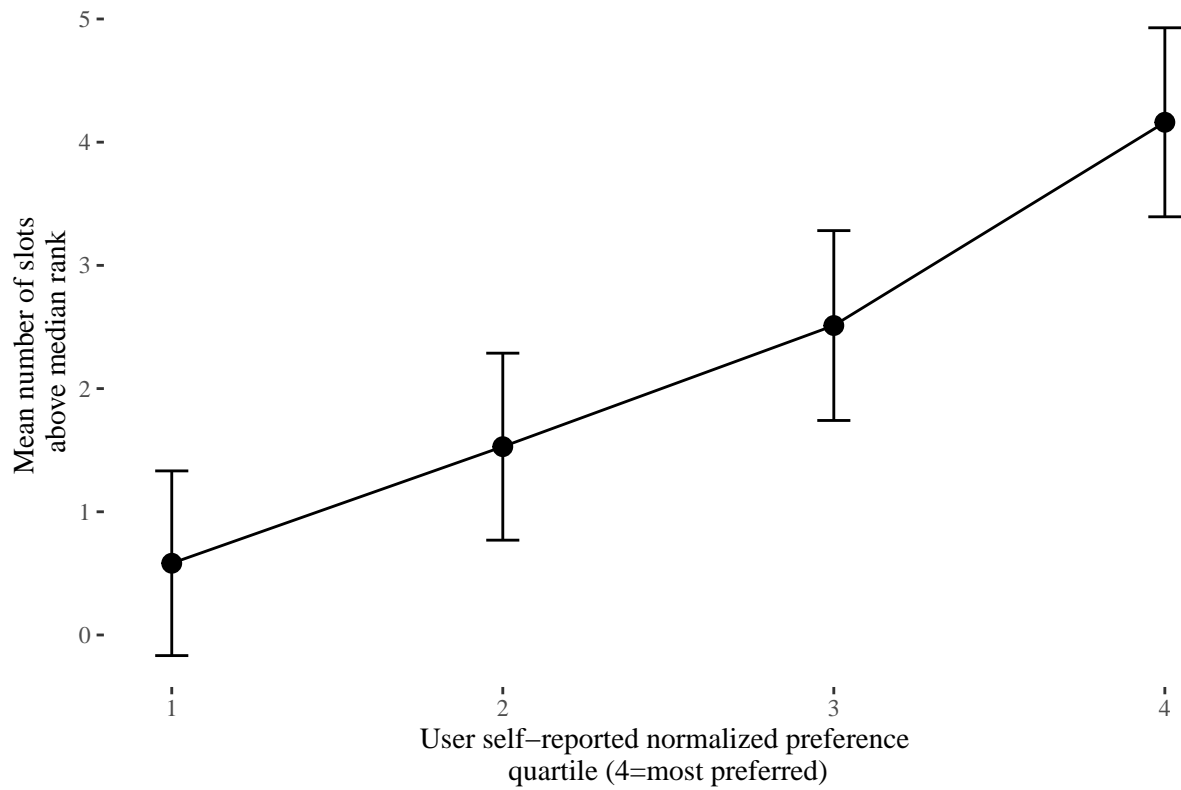
Newsfeed Top 10 by Preference (India)



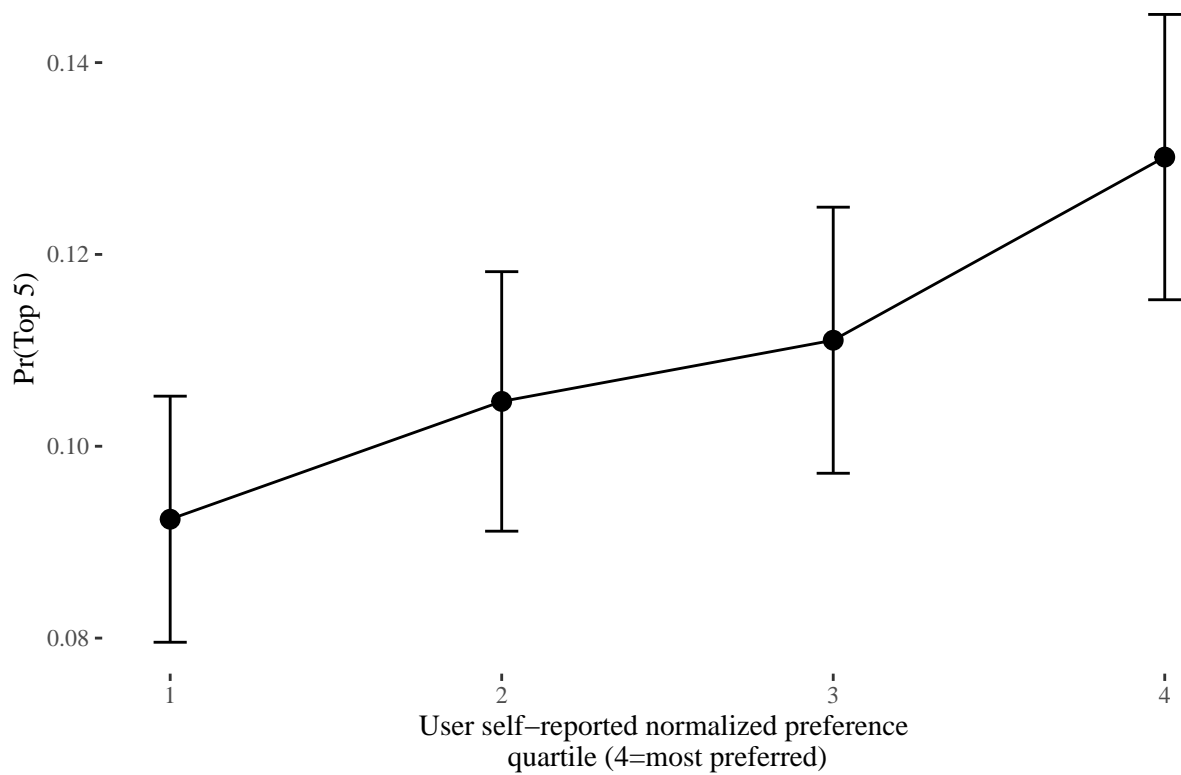
Newsfeed Top 15 by Preference (India)



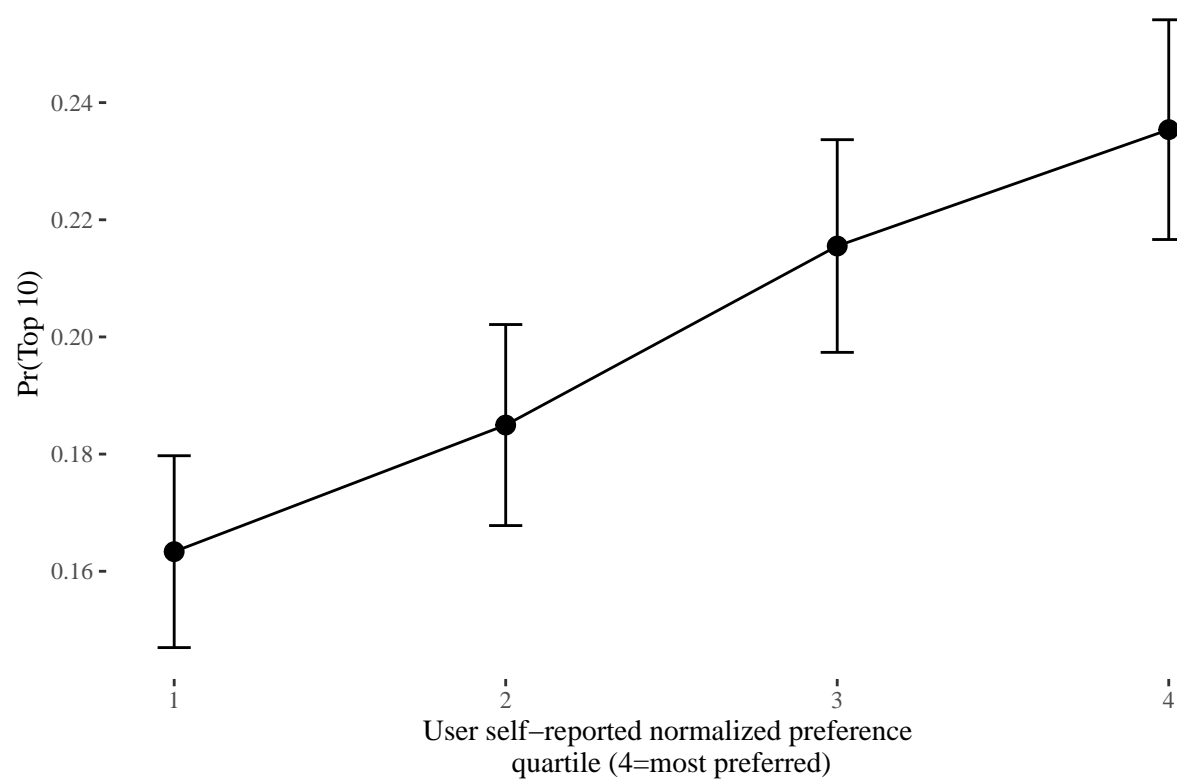
Newsfeed Ranking by Preference (India)



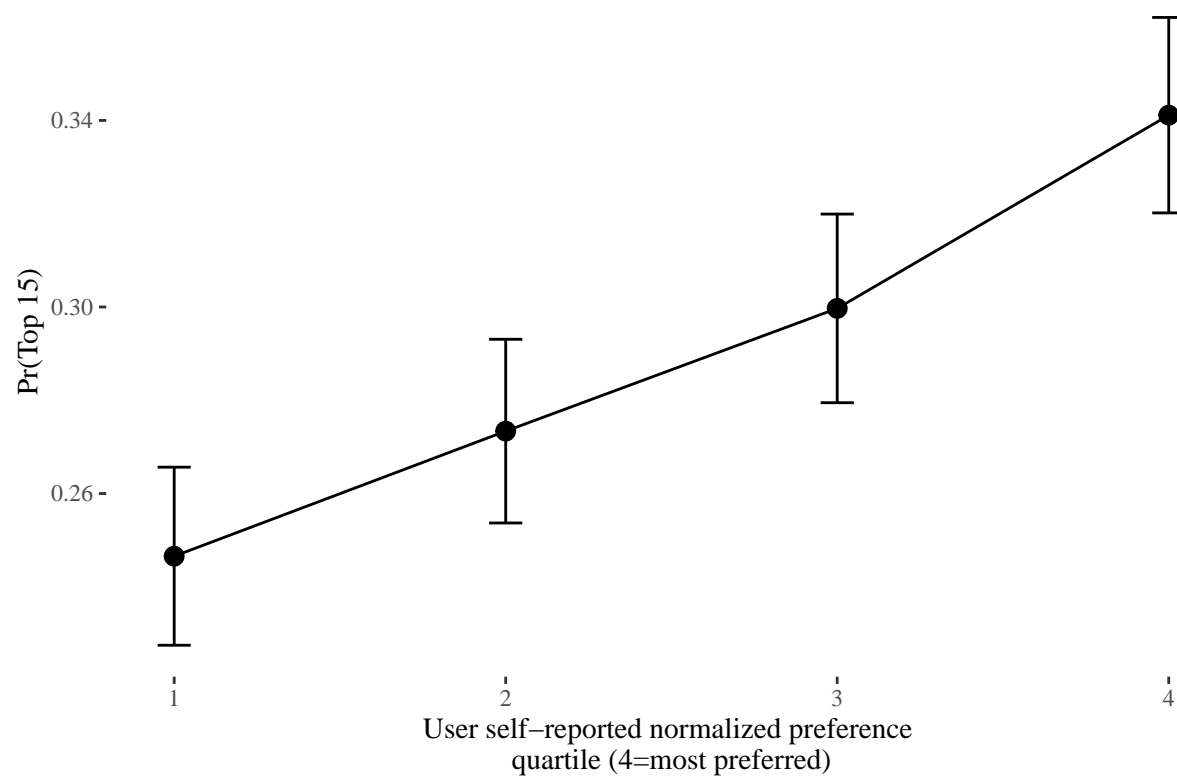
Newsfeed Top 5 by Preference (India)

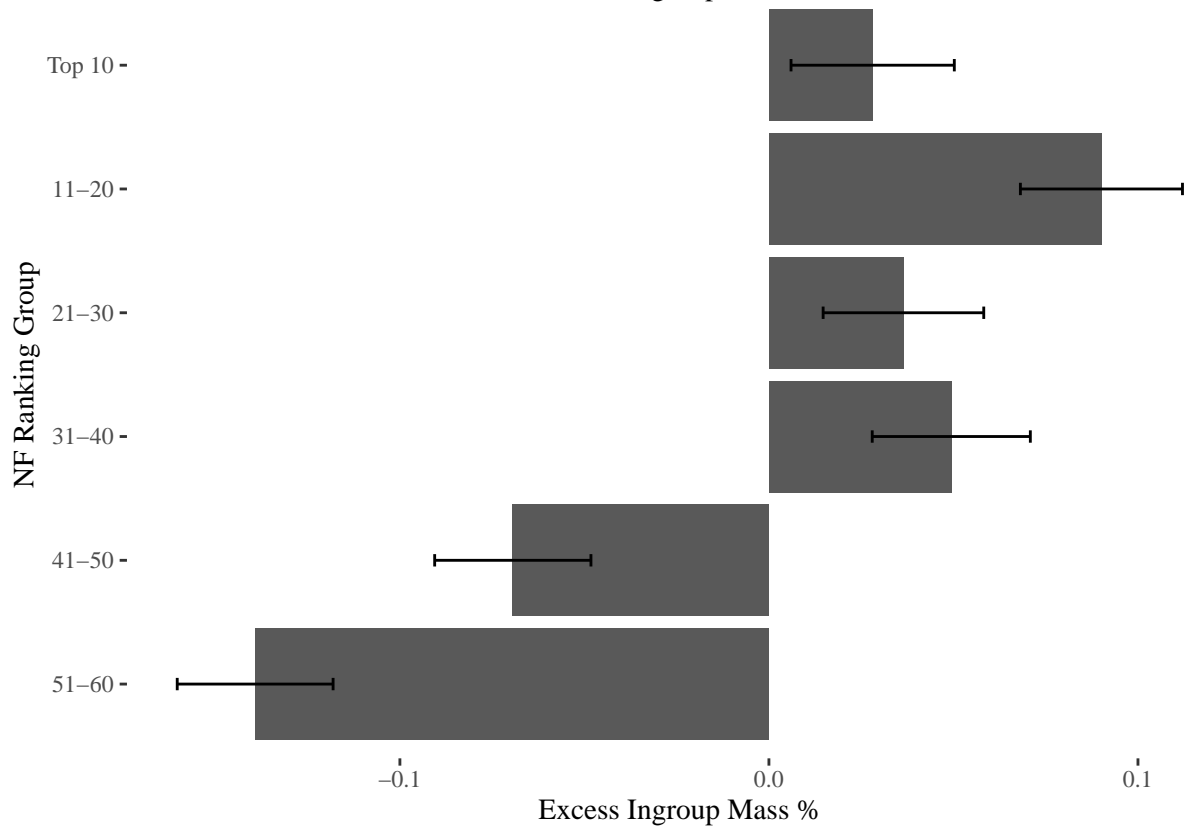
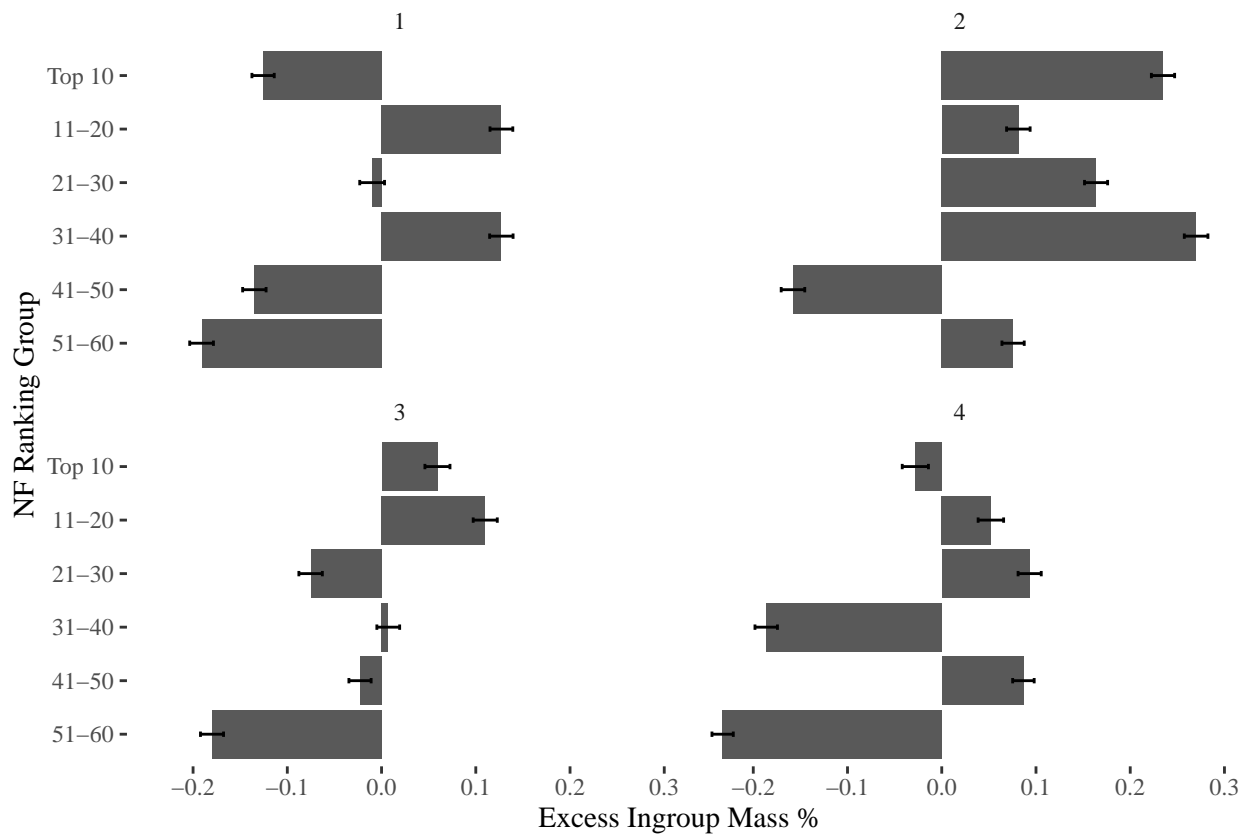


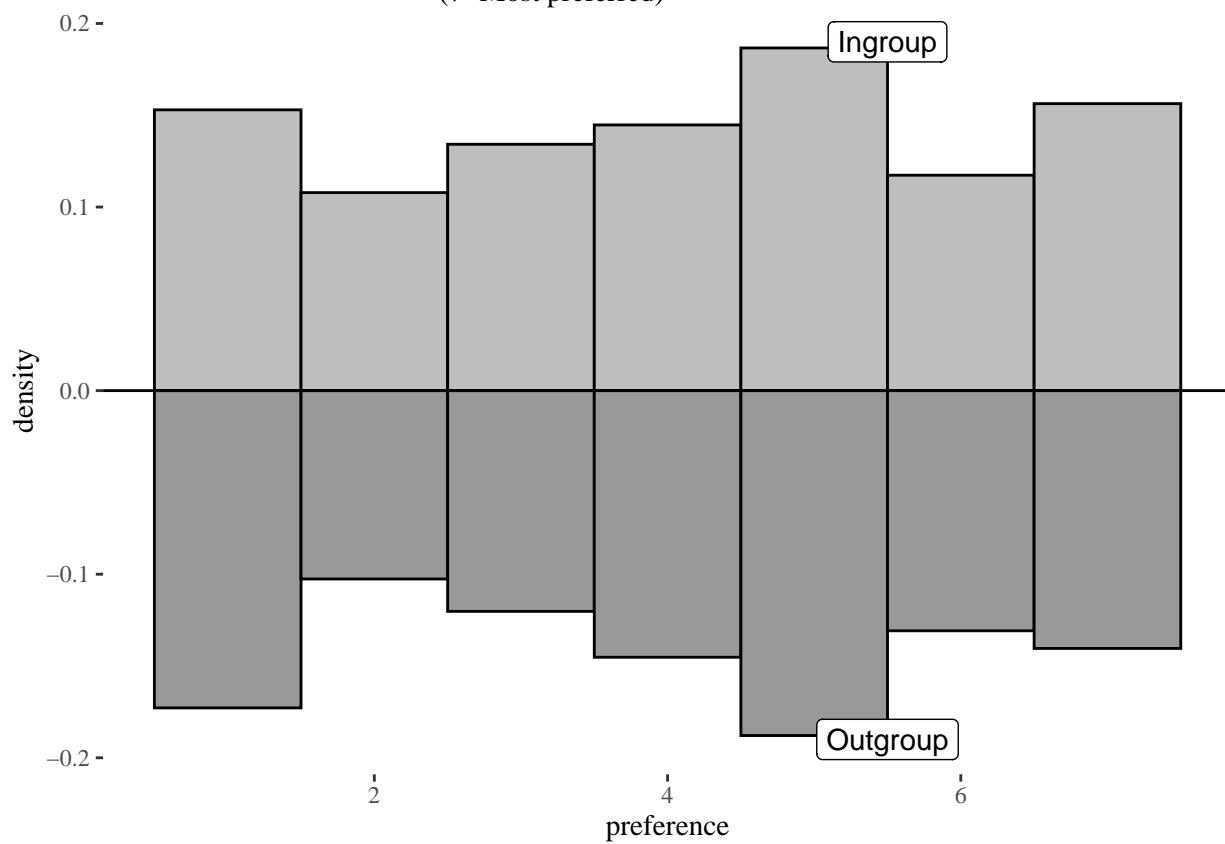
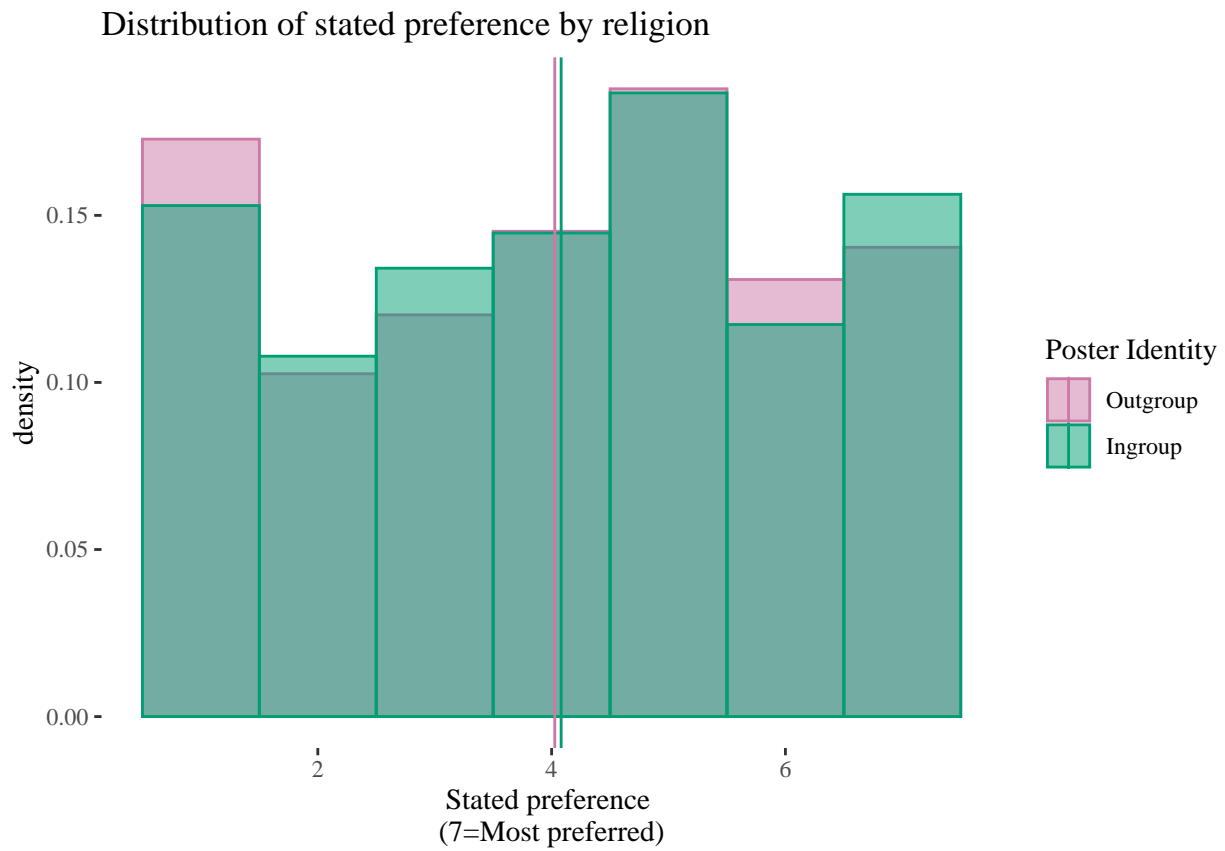
Newsfeed Top 10 by Preference (India)

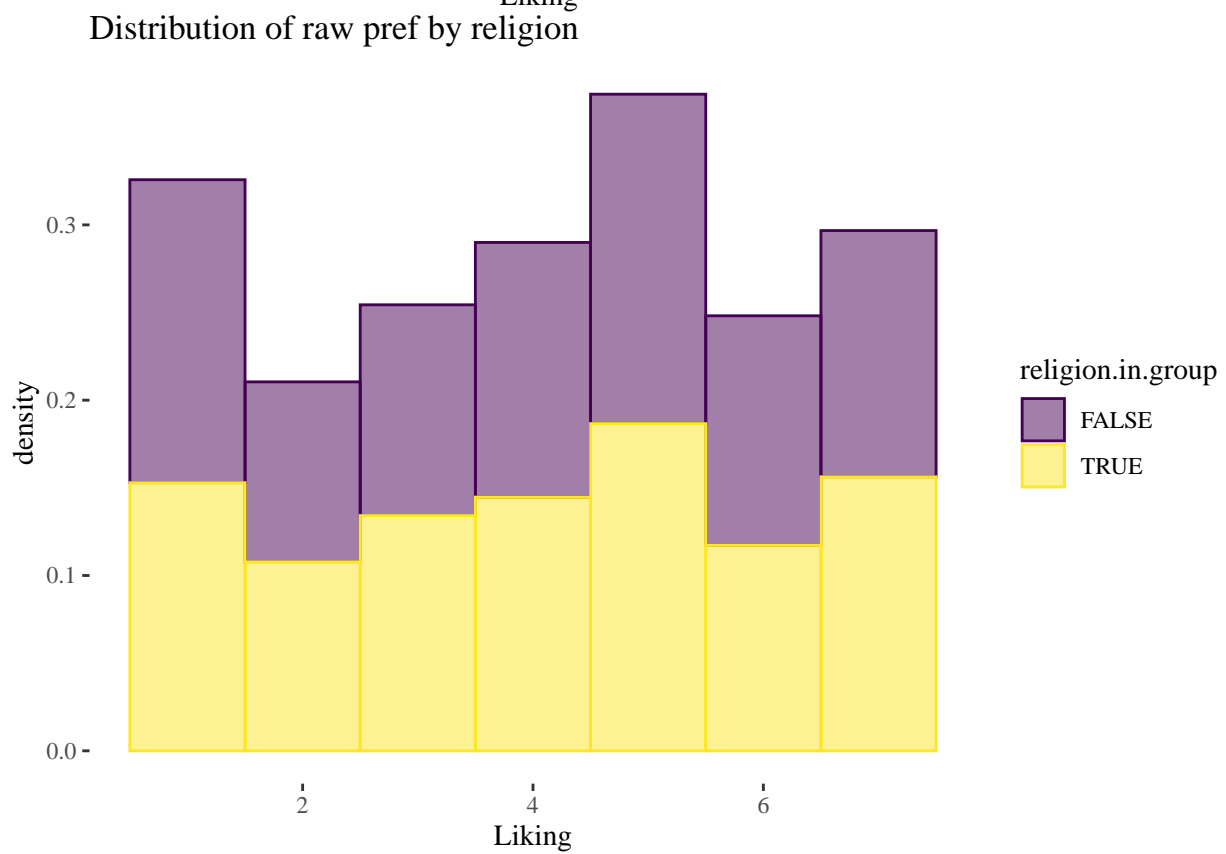
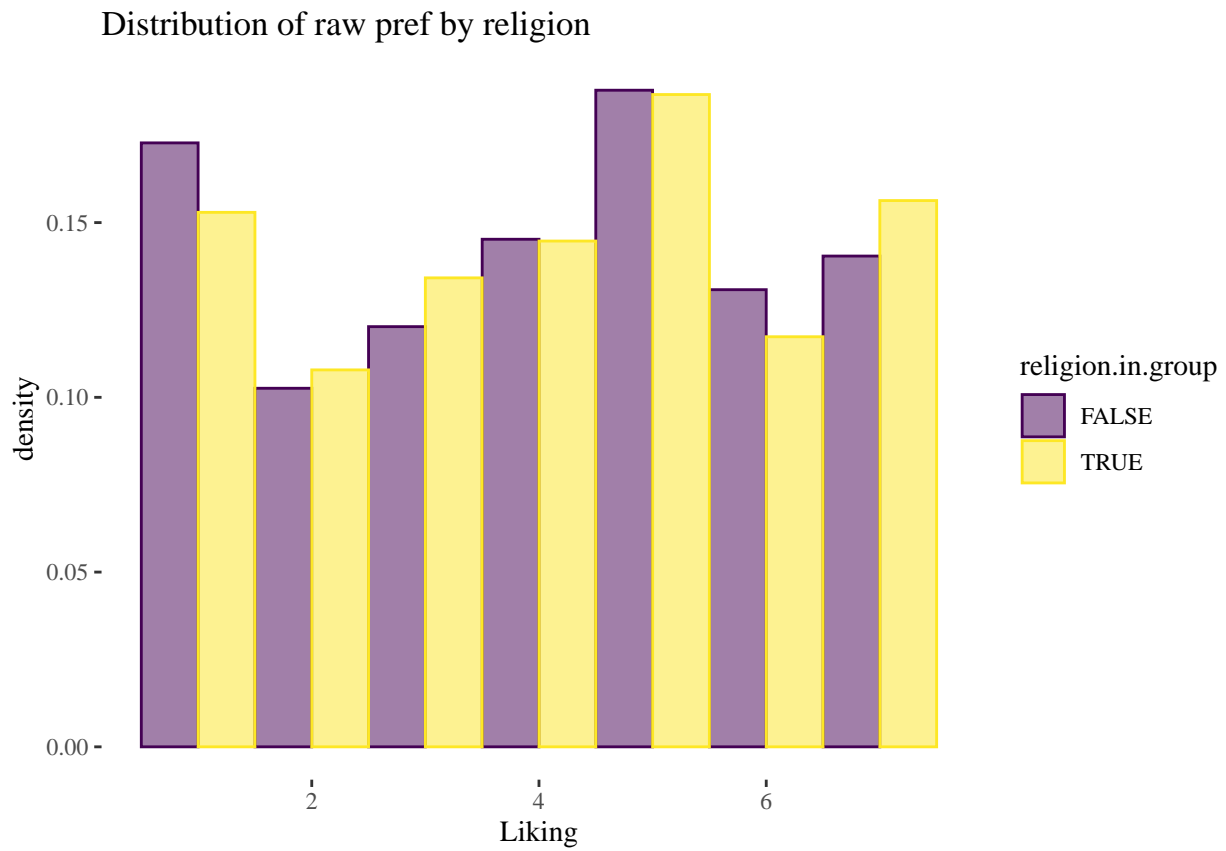


Newsfeed Top 15 by Preference (India)

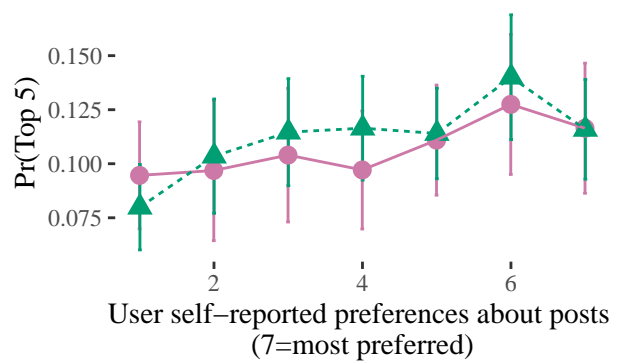
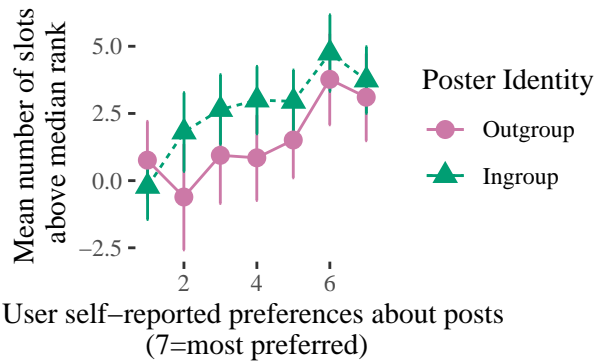




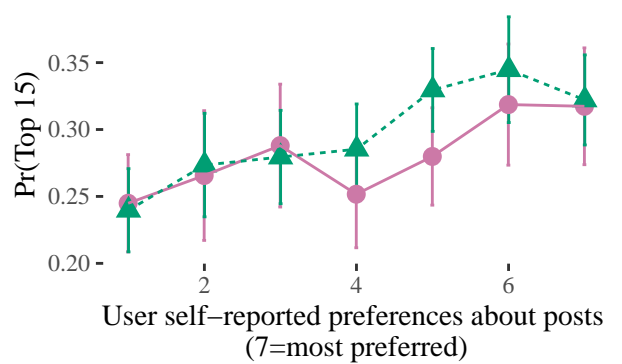
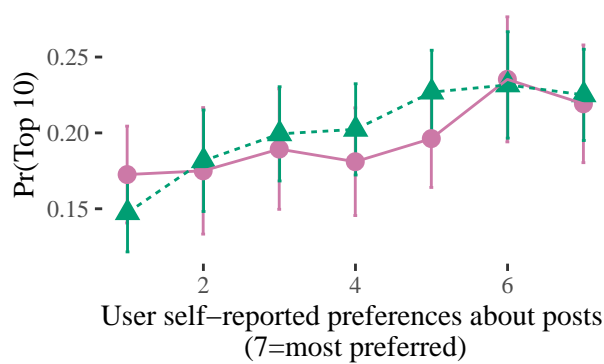




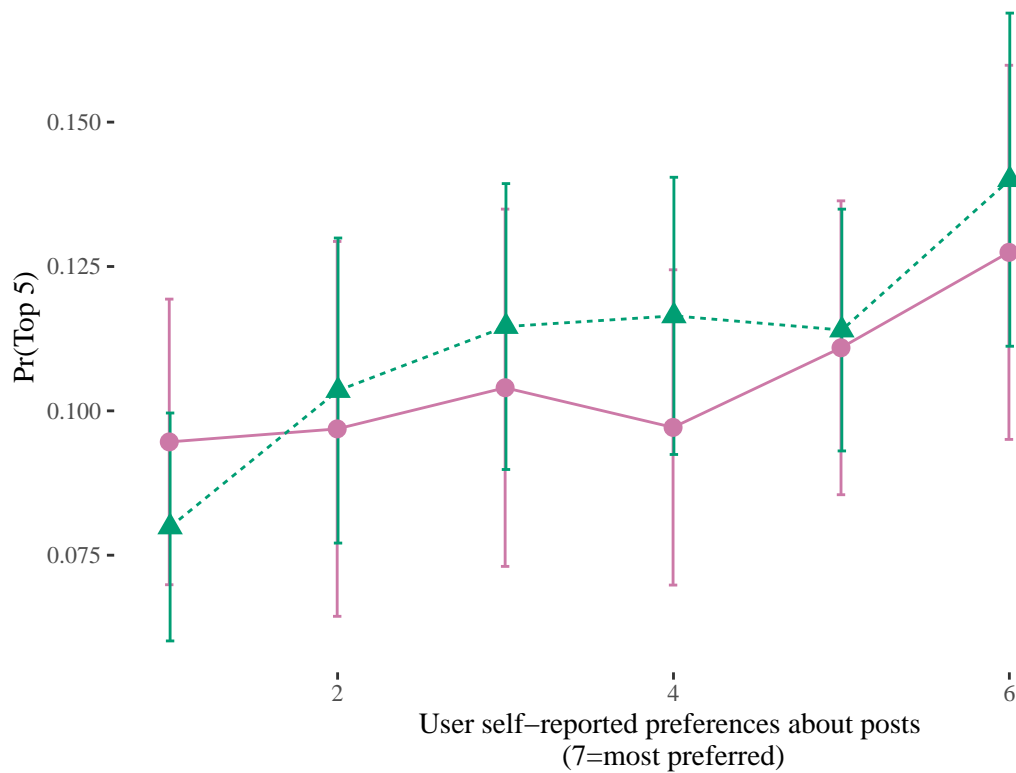
A Newsfeed Ranking by Preference (India) **B** Newsfeed Top 5 by Preference (India)

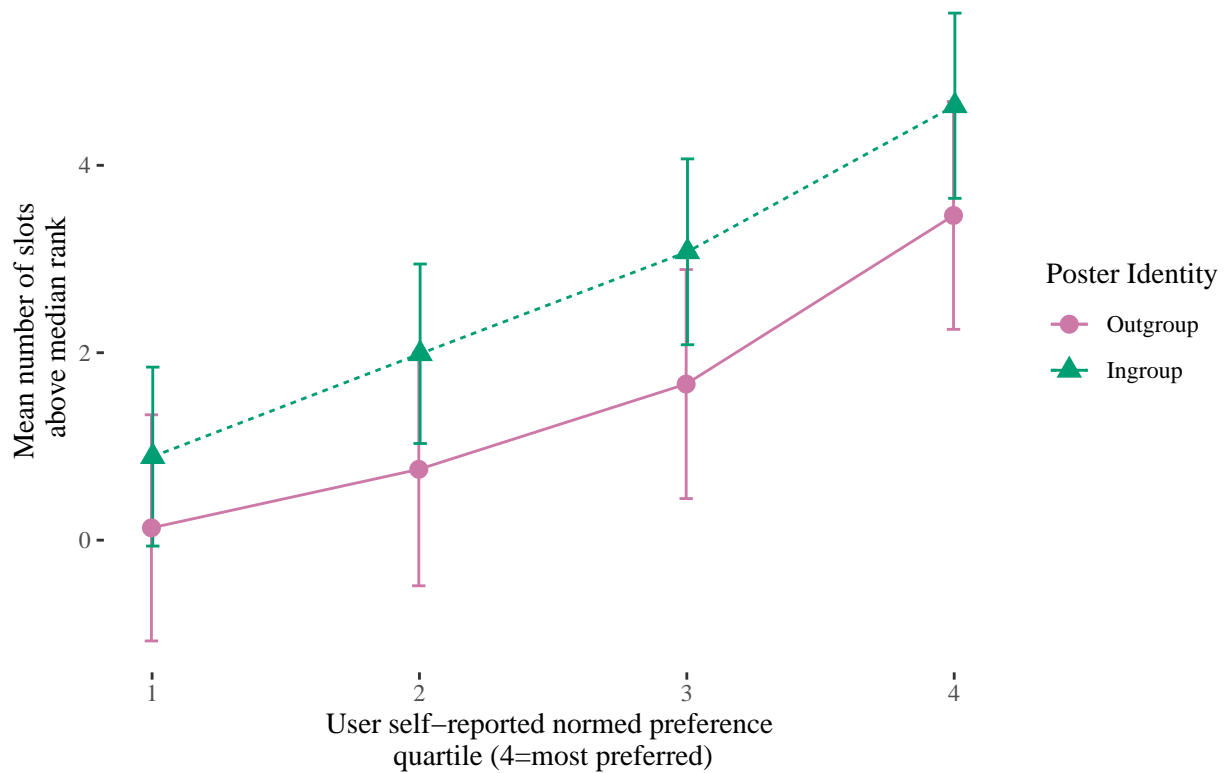
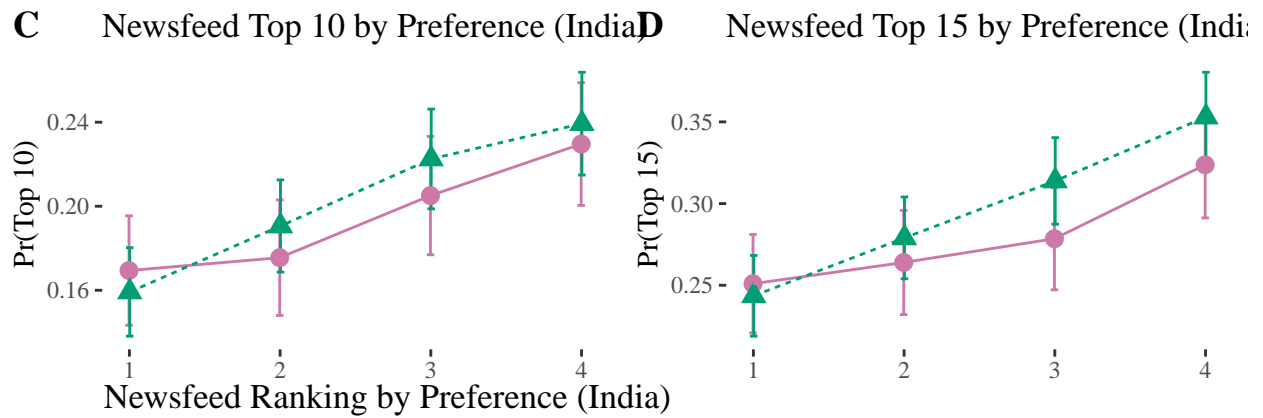
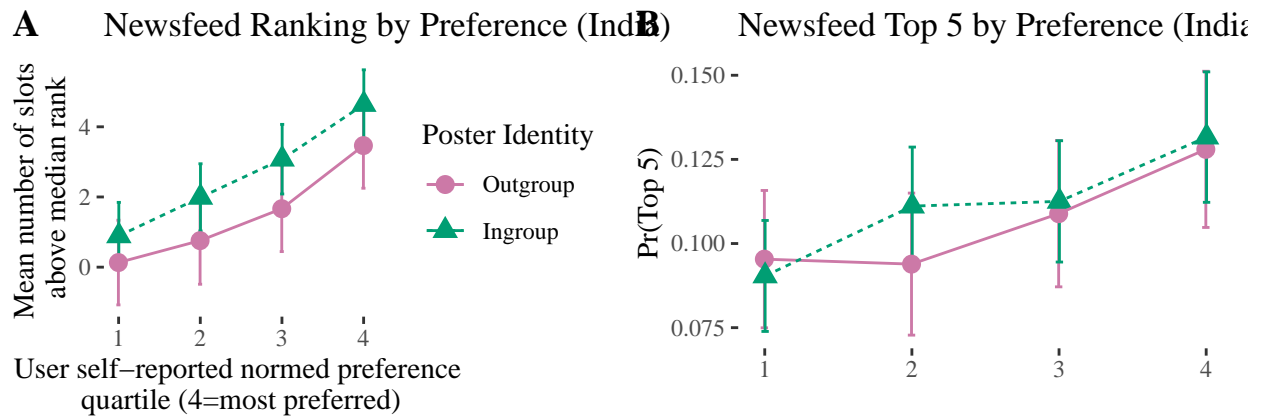


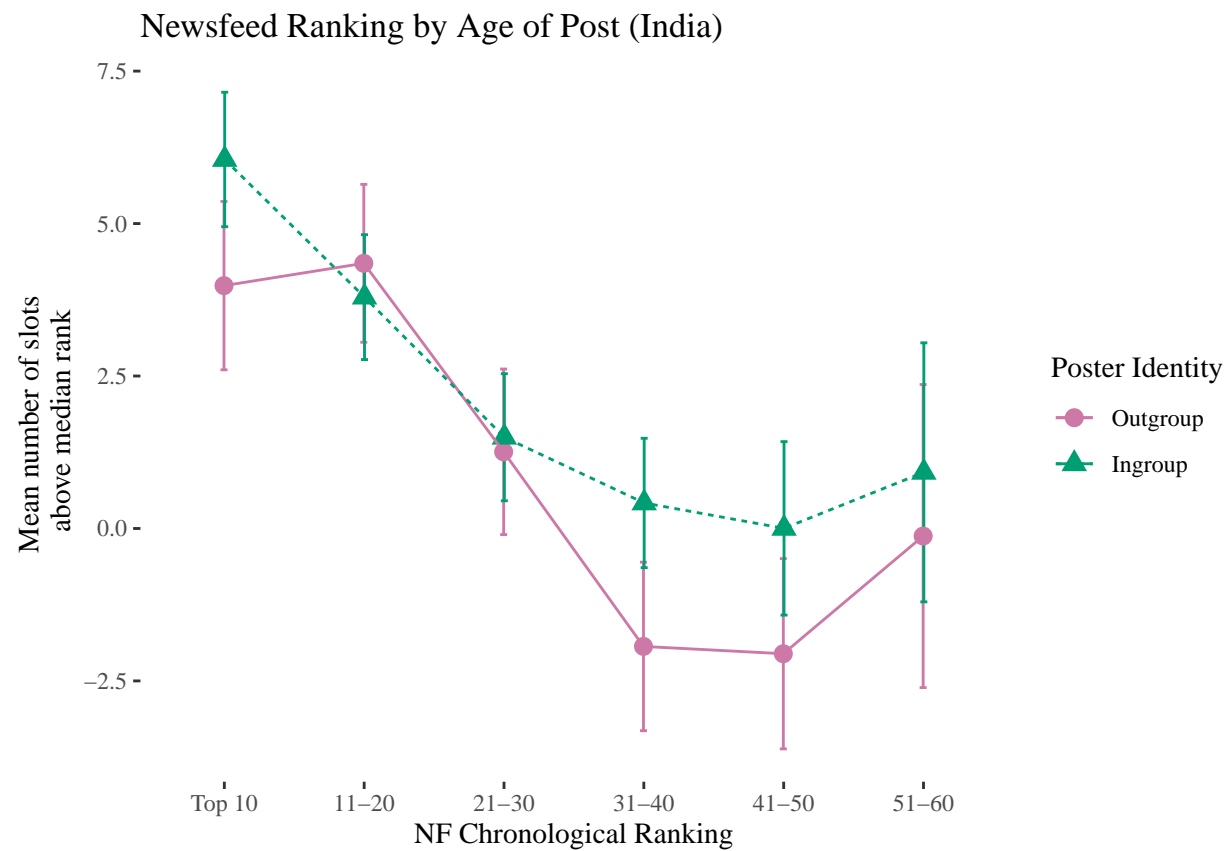
C Newsfeed Top 10 by Preference (India) **D** Newsfeed Top 15 by Preference (India)



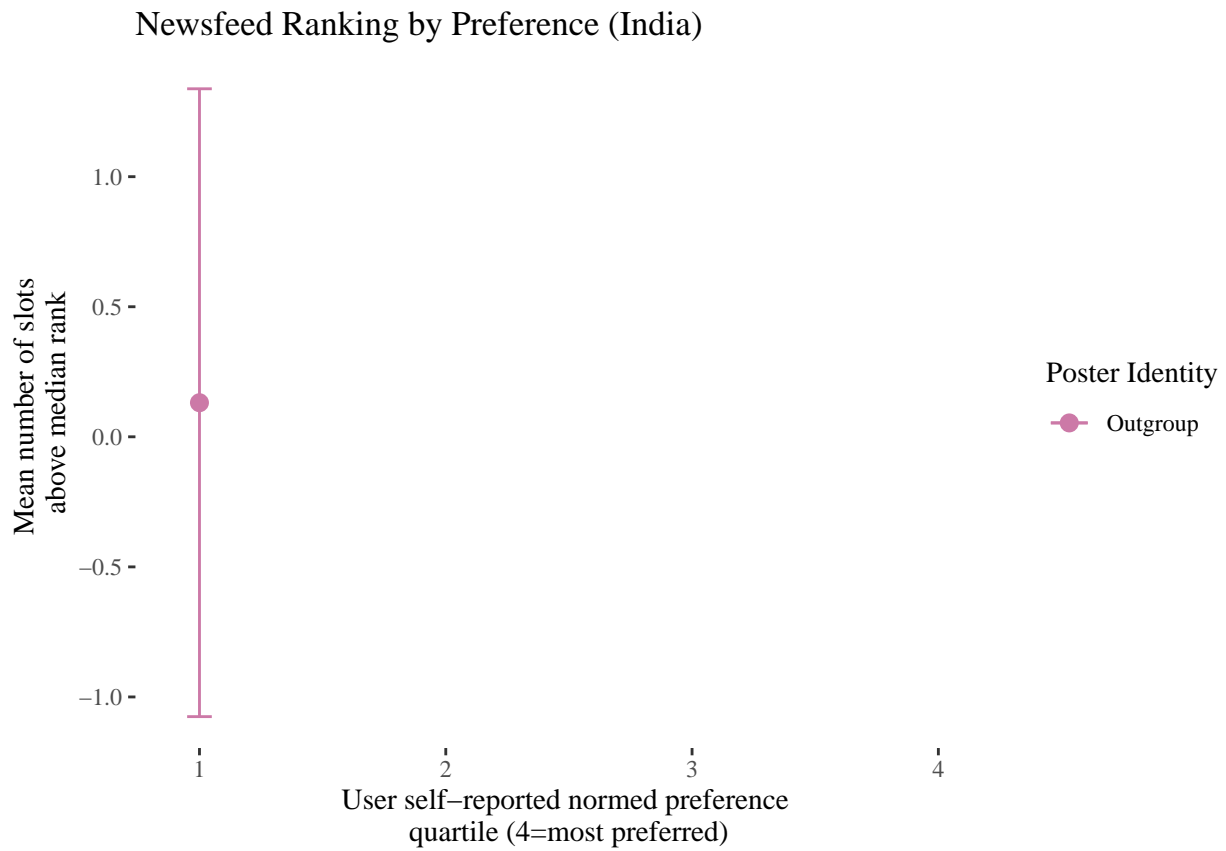
Newsfeed Top 5 by Preference (India)





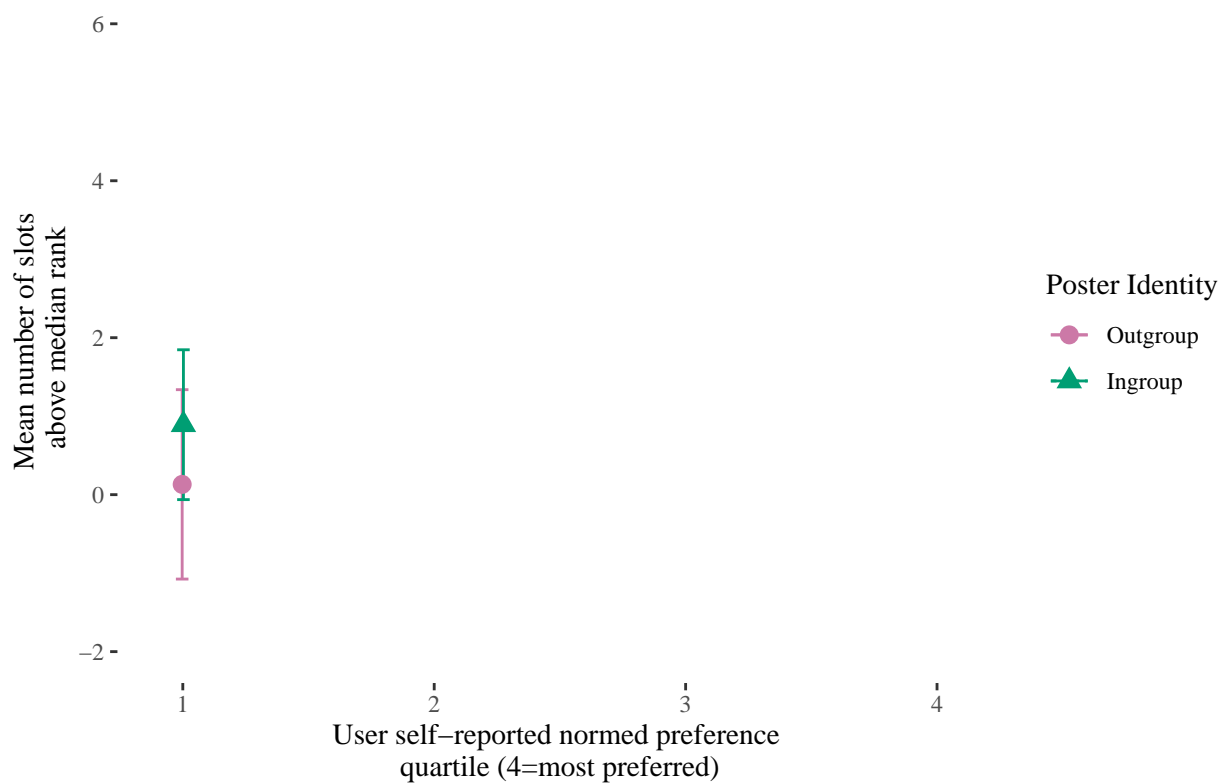


geom_path: Each group consists of only one observation. Do you need to adjust
 ## the group aesthetic?

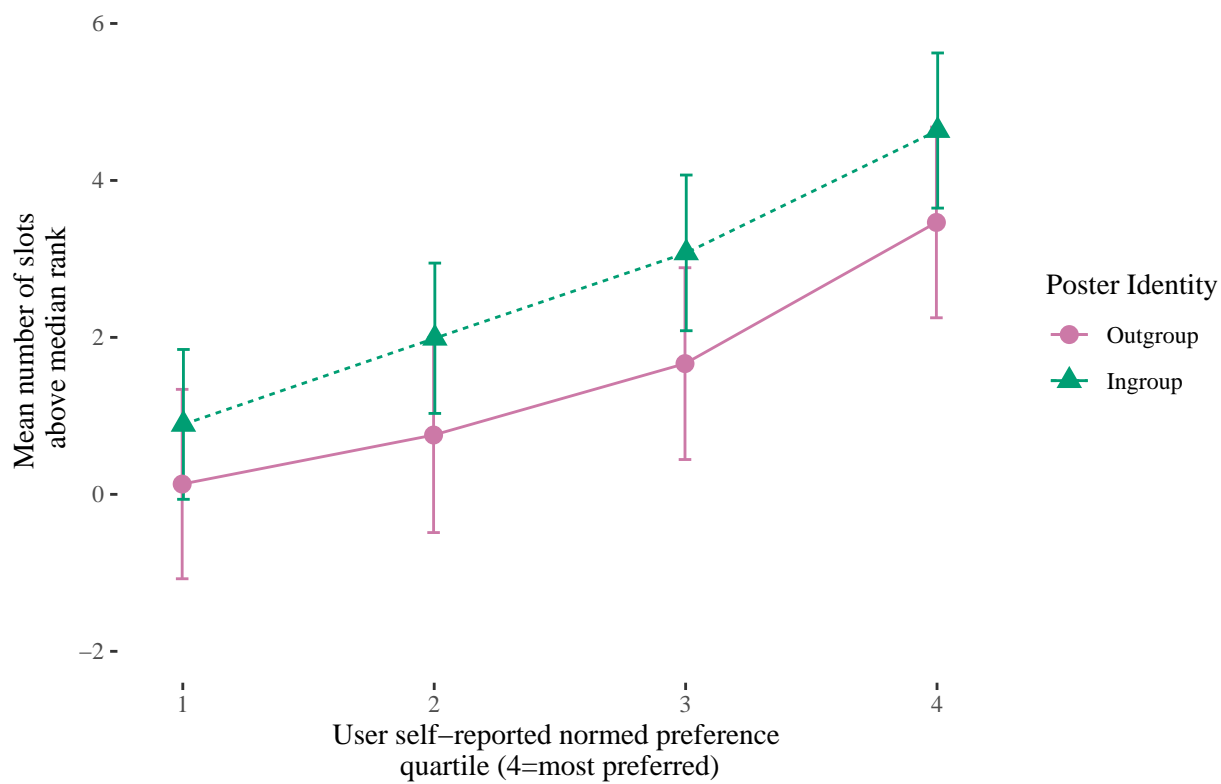


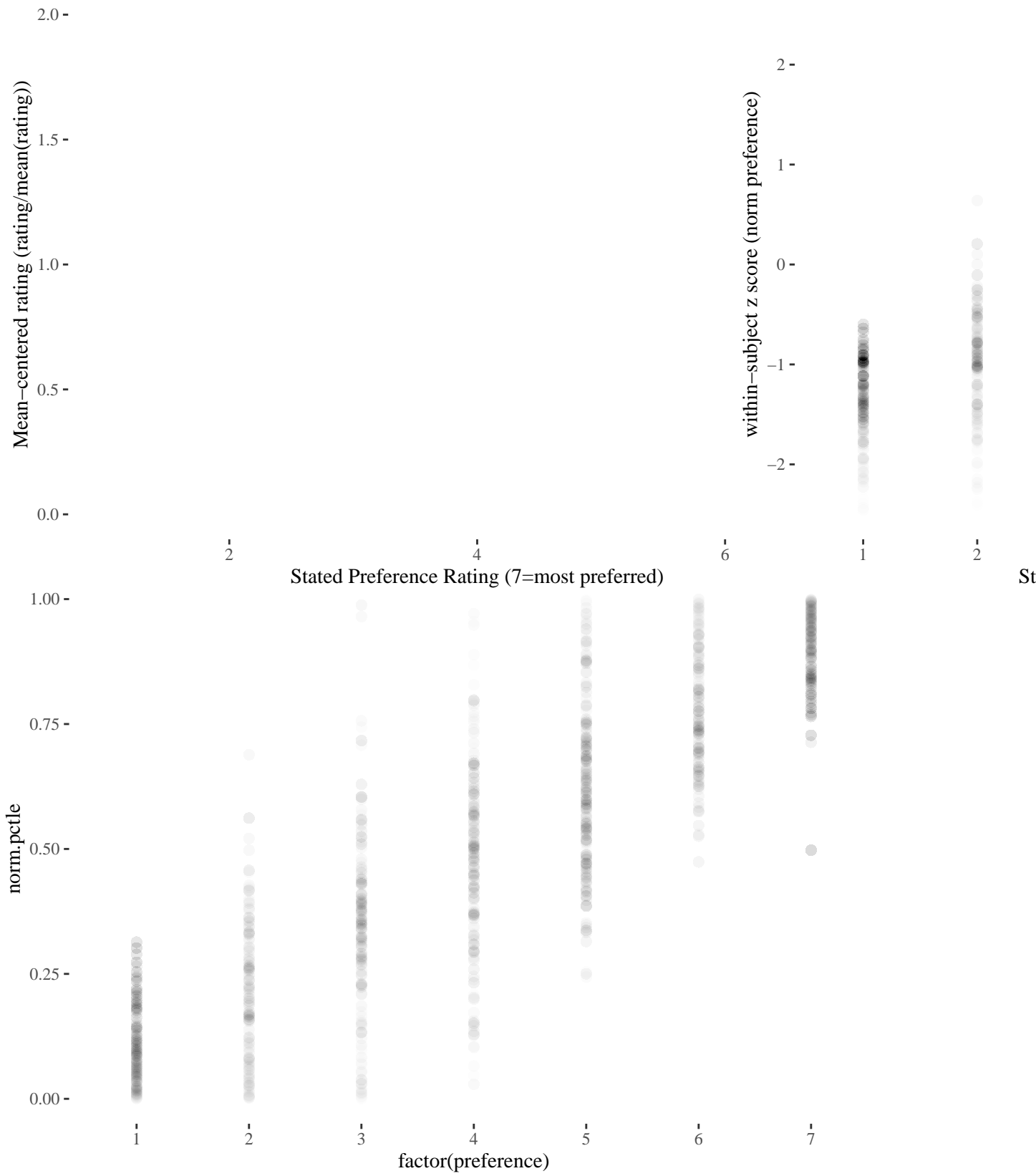
```
## geom_path: Each group consists of only one observation. Do you need to adjust  
## the group aesthetic?
```

Newsfeed Ranking by Preference (India)



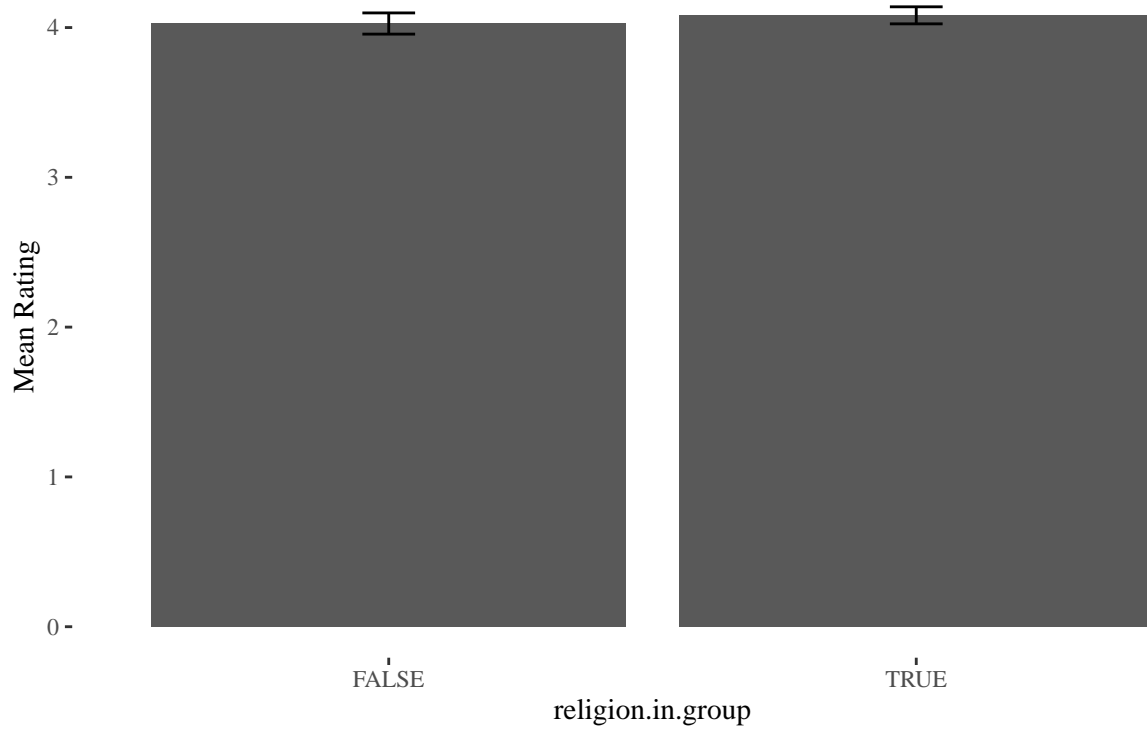
Newsfeed Ranking by Preference (India)



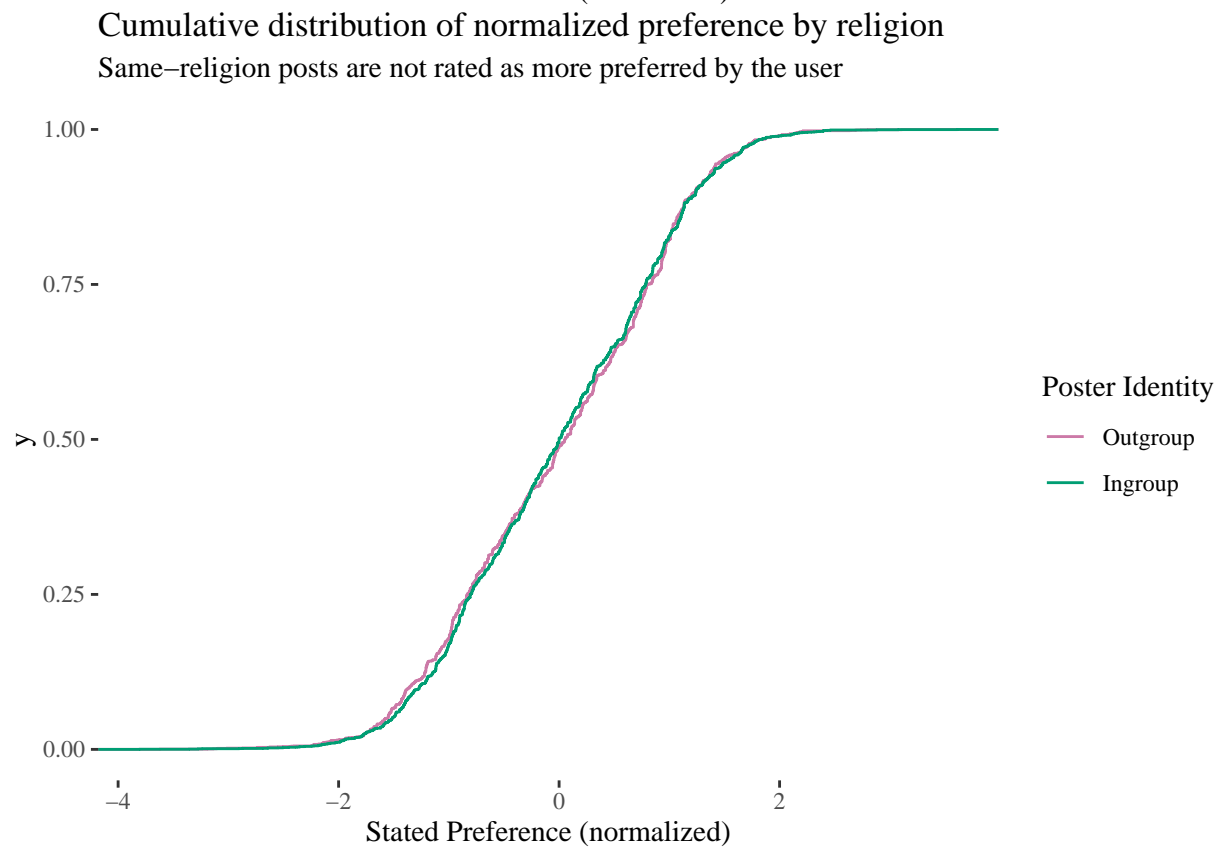
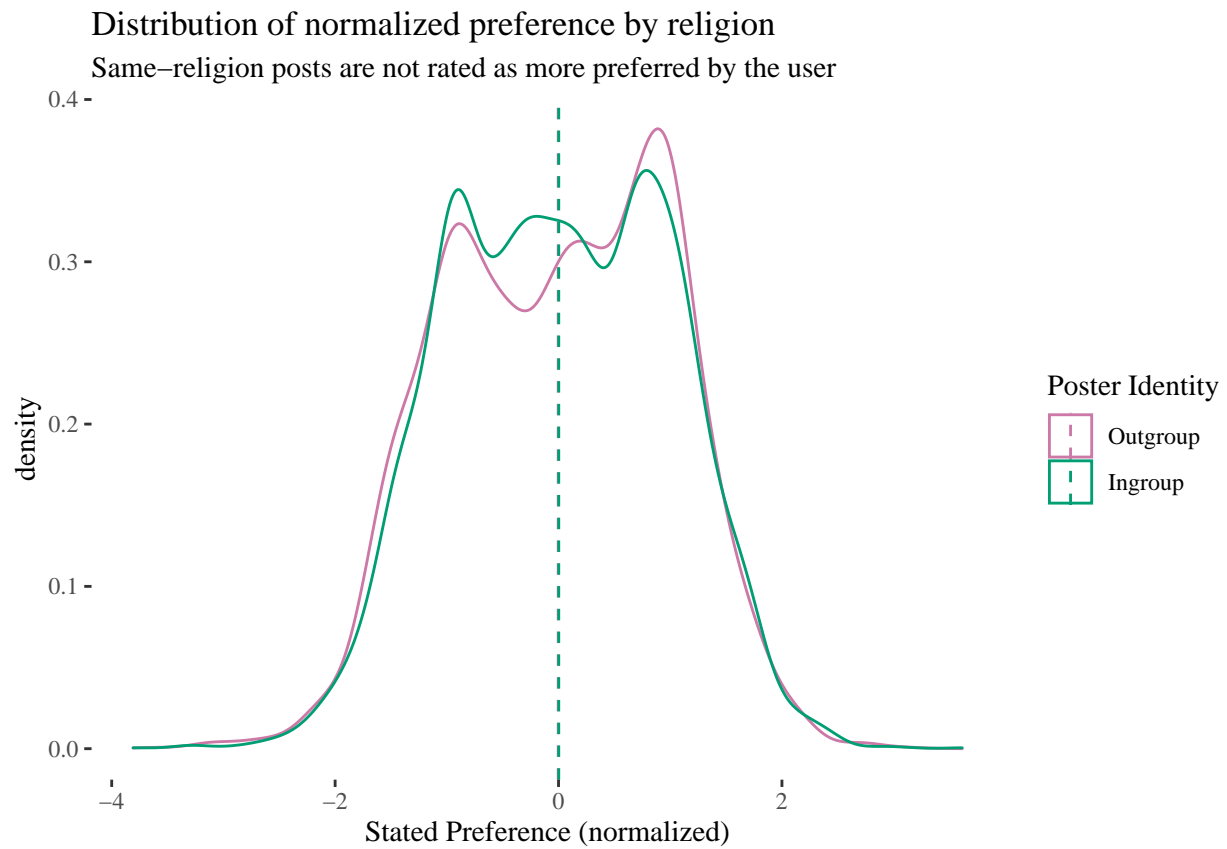


Raw Reported Preference by Race Group

Same-religion posts are slightly more preferred



Adding missing grouping variables: `dedupid`



Adding missing grouping variables: `dedupid`

Adding missing grouping variables: `dedupid`

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:08:20

Table 6: Time Results - Rank

	<i>Dependent variable:</i>		
	nf.order		
	(1)	(2)	(3)
	all	all2	base
I(100 *norm.pctle)	−0.048*** (0.007)	−0.048*** (0.007)	−0.047*** (0.007)
time_rank	0.164*** (0.021)	0.150*** (0.013)	
religion.in.group	−0.510 (0.736)	−1.078*** (0.394)	−1.126*** (0.397)
time_rank:religion.in.group	−0.024 (0.027)		
Constant	27.529*** (0.657)	27.864*** (0.545)	31.340*** (0.457)
Observations	7,866	7,866	7,866
R ²	0.024	0.023	0.007
Adjusted R ²	0.023	0.023	0.007

Note:

*p<0.1; **p<0.05; ***p<0.01

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:08:20

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Tue, Jan 05, 2021 - 18:08:20

Table 7: Time Results - Top 10

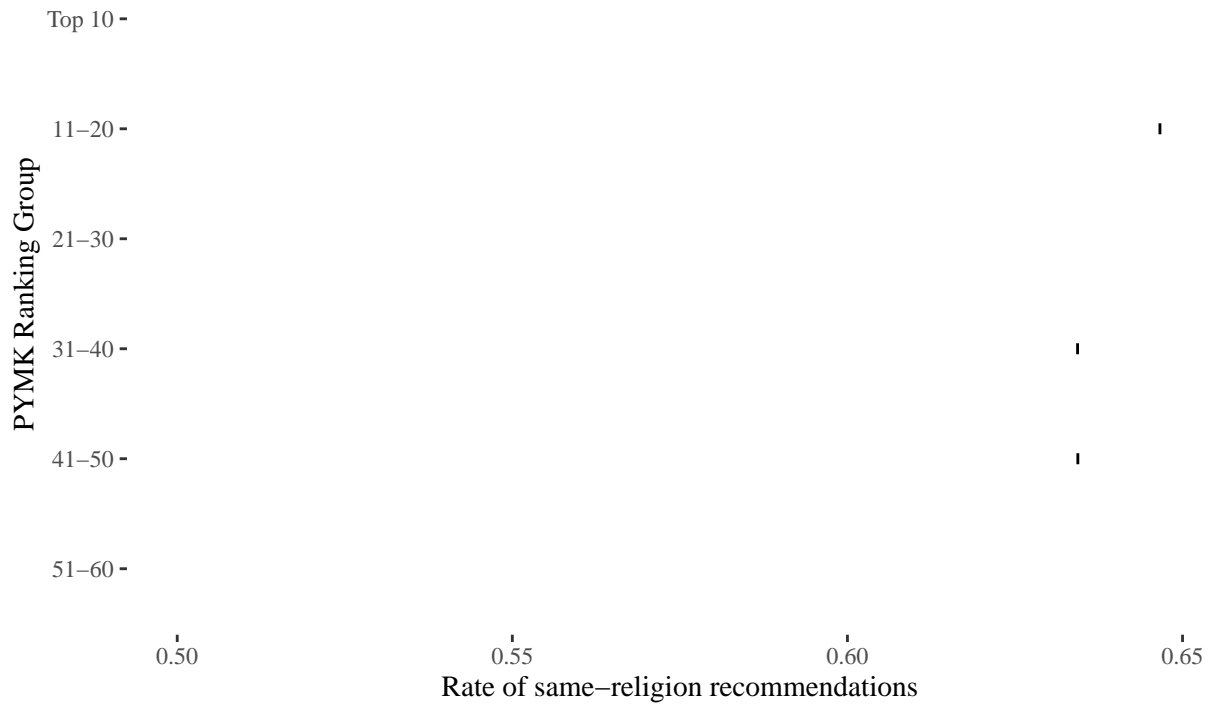
	<i>Dependent variable:</i>			
	nf.order			
	(1) all	(2) days	(3) hours	(4) mins
religion.in.group	−1.078*** (0.394)	−1.106*** (0.397)	−0.520 (0.617)	0.061 (1.635)
I(100 *norm.pctle)	−0.048*** (0.007)	−0.049*** (0.007)	−0.039*** (0.011)	−0.005 (0.028)
time_rank	0.150*** (0.013)	0.147*** (0.013)	0.147*** (0.034)	0.168 (0.225)
Constant	27.864*** (0.545)	28.015*** (0.550)	27.484*** (0.830)	27.718*** (2.090)
Observations	7,866	7,758	3,512	542
R ²	0.023	0.023	0.010	0.001
Adjusted R ²	0.023	0.023	0.009	−0.004
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01				

Table 8: Time Results - Top 10

	<i>Dependent variable:</i>			
	nf.order			
	(1) all	(2) recent10	(3) recent20	(4) recent30
religion.in.group	−1.078*** (0.394)	−1.476* (0.847)	−0.605 (0.598)	−0.615 (0.491)
I(100 *norm.pctle)	−0.048*** (0.007)	−0.047*** (0.014)	−0.043*** (0.010)	−0.044*** (0.008)
time_rank	0.150*** (0.013)	0.096 (0.144)	0.176*** (0.051)	0.187*** (0.028)
Constant	27.864*** (0.545)	27.974*** (1.246)	26.779*** (0.863)	26.796*** (0.704)
Observations	7,866	1,922	3,721	5,317
R ²	0.023	0.008	0.008	0.014
Adjusted R ²	0.023	0.006	0.008	0.014
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01				

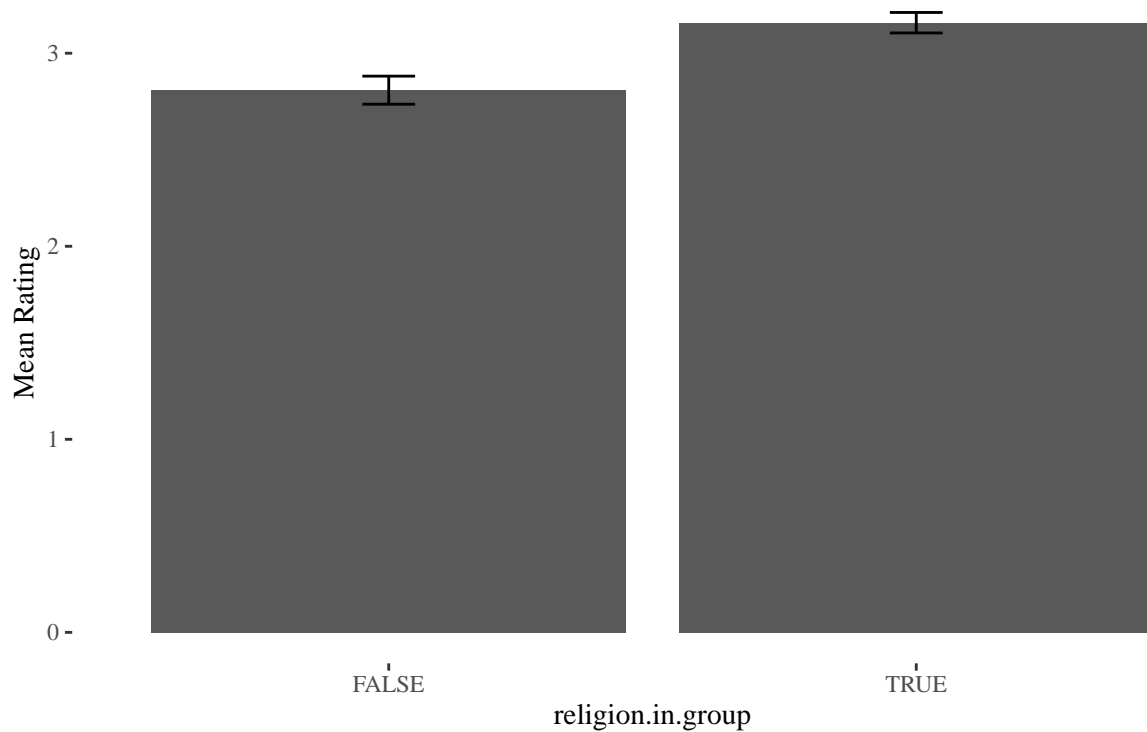
PYMK preference for user's religion (India)

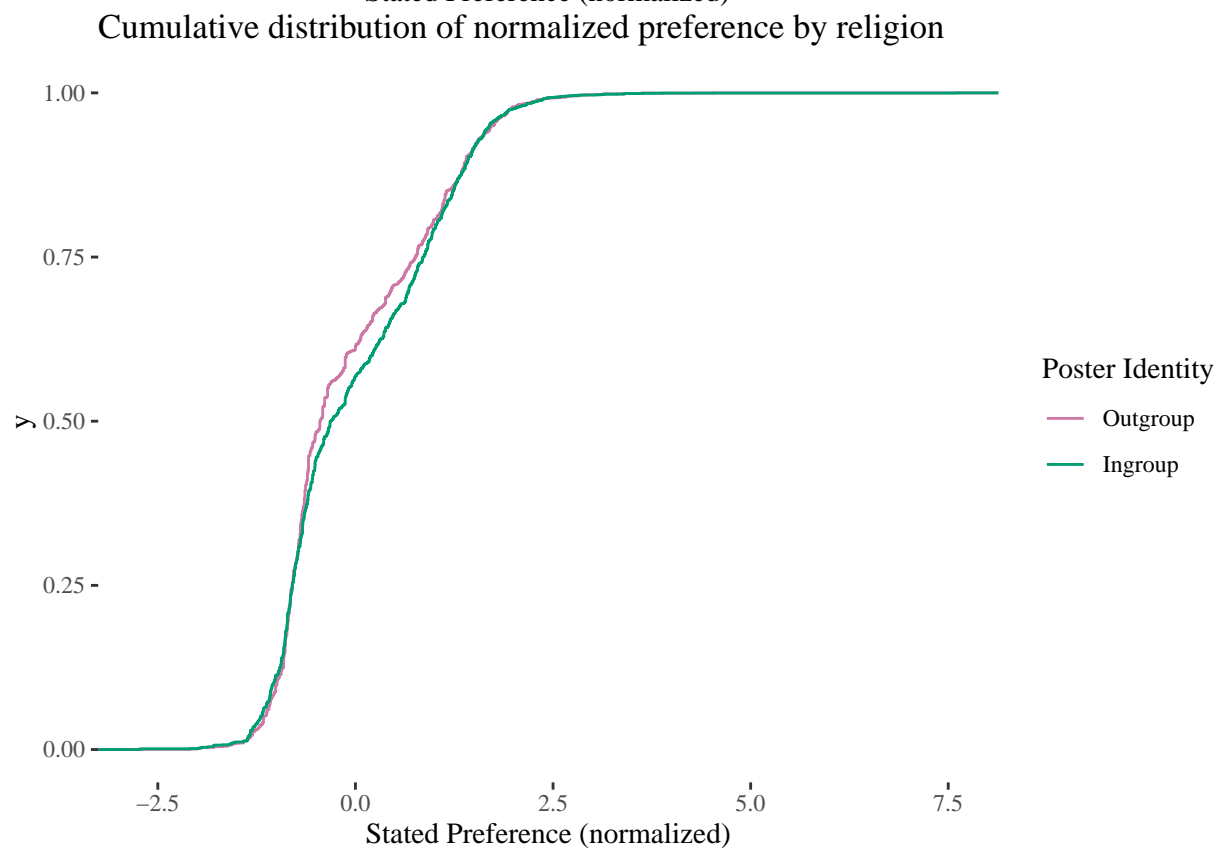
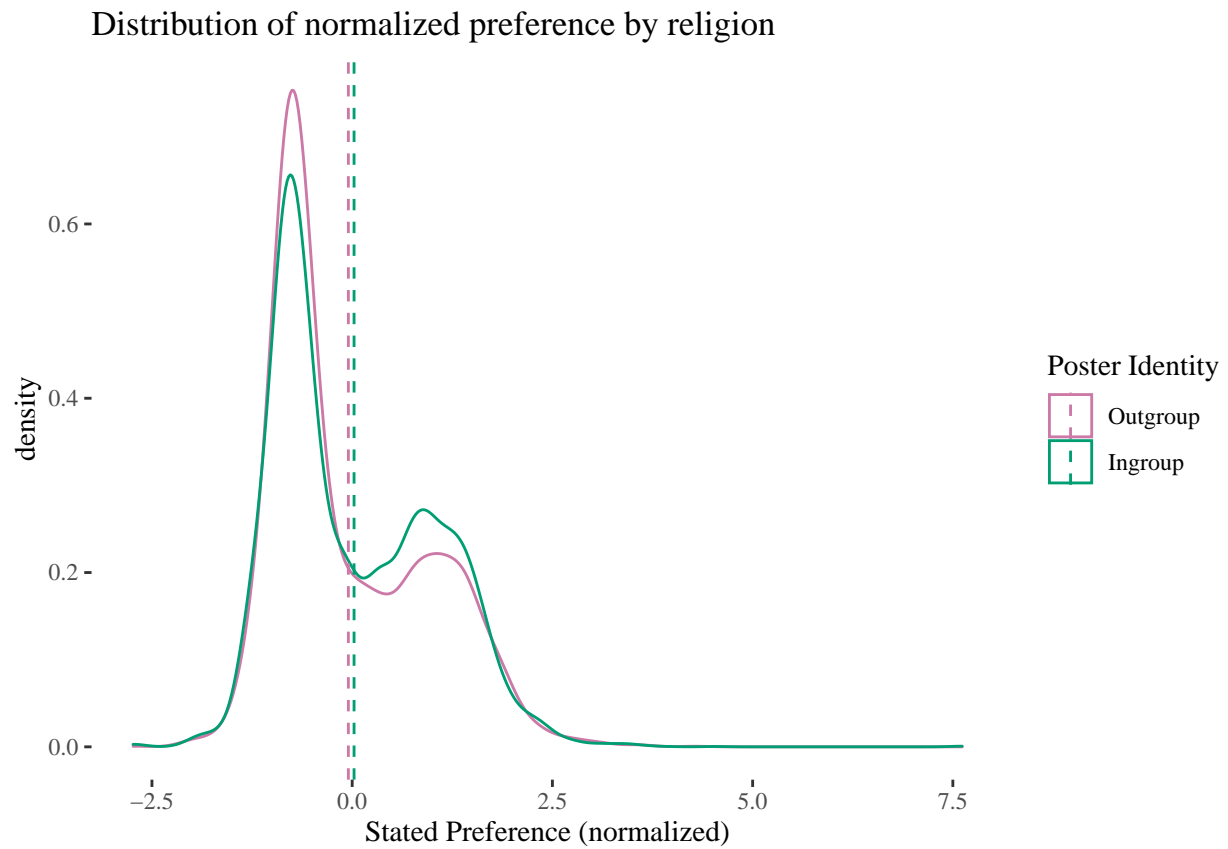
Same-religion recommendations are not sorted closer to the top



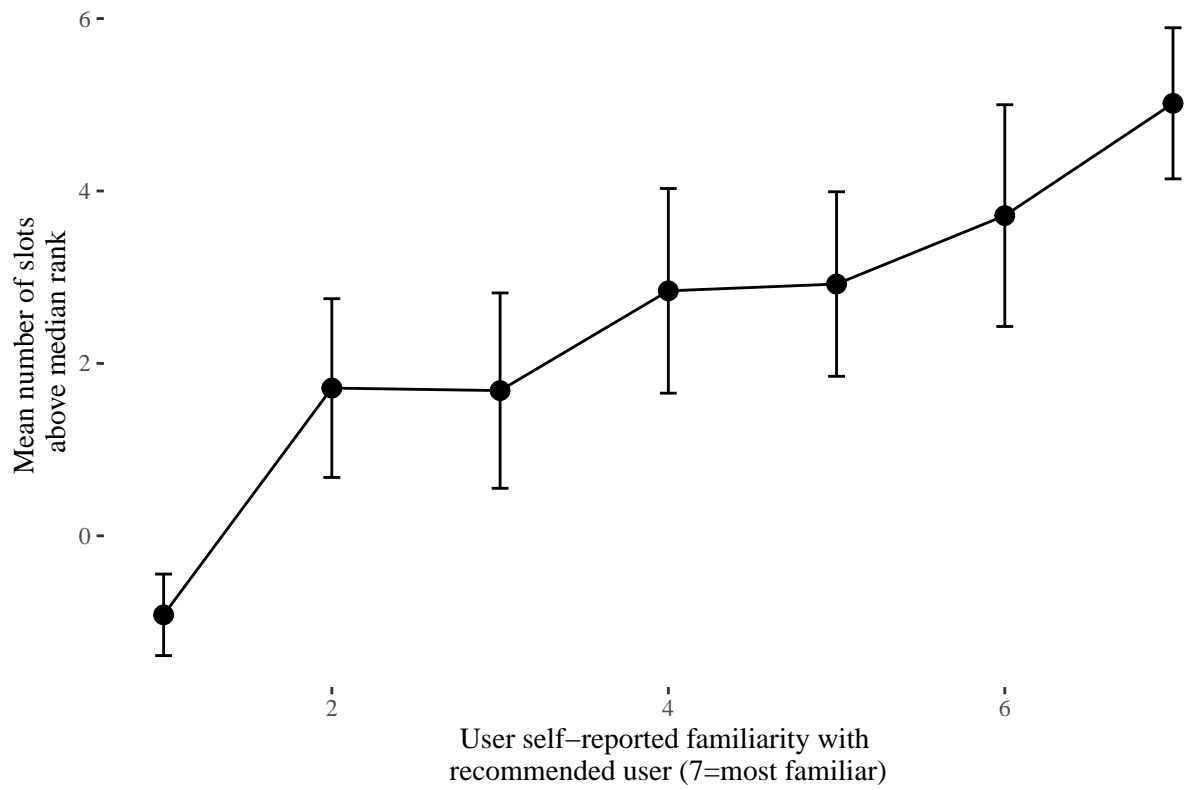
Raw Reported Preference by Race Group

Same-religion posts are slightly more preferred

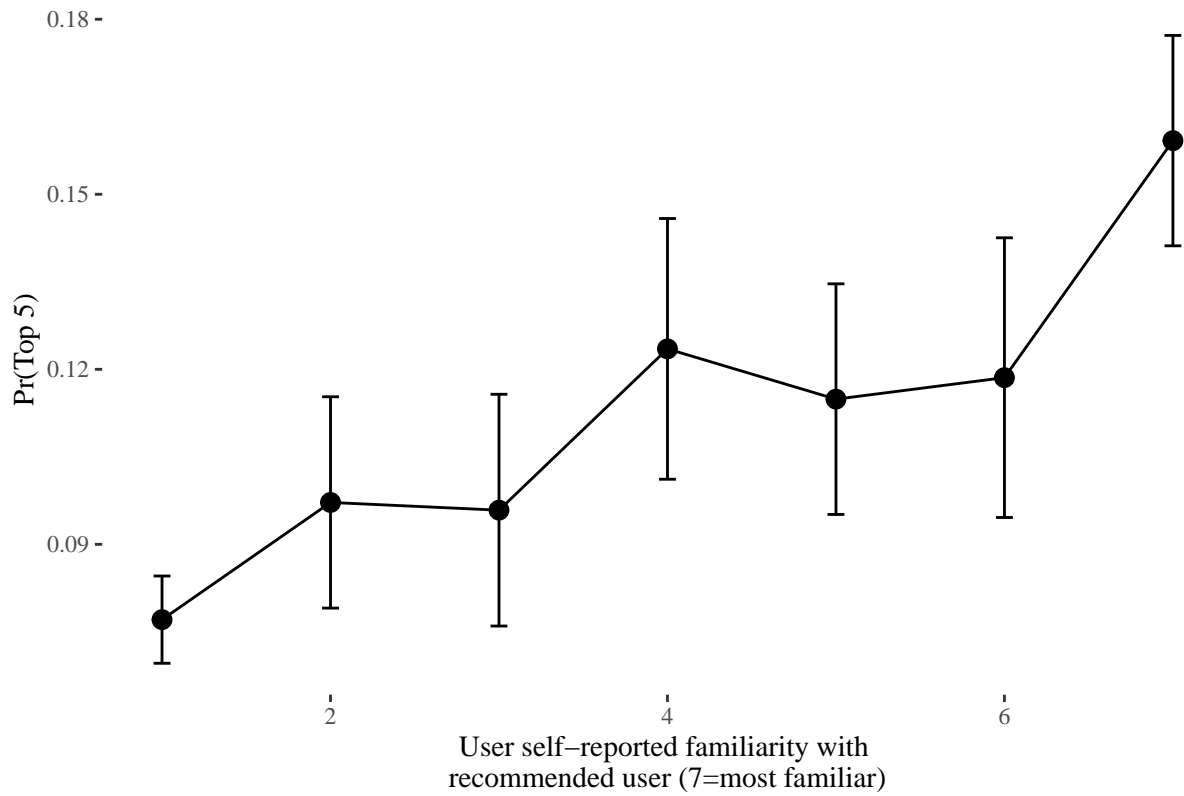


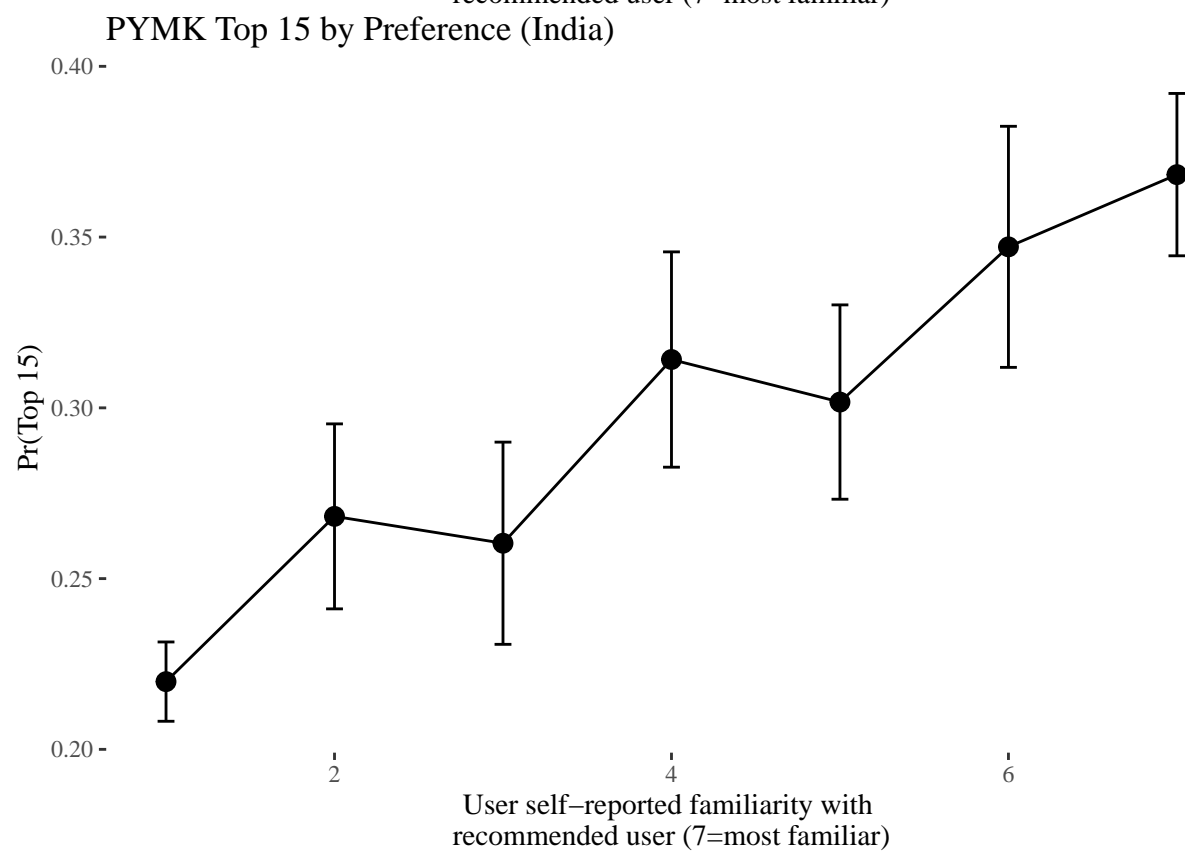
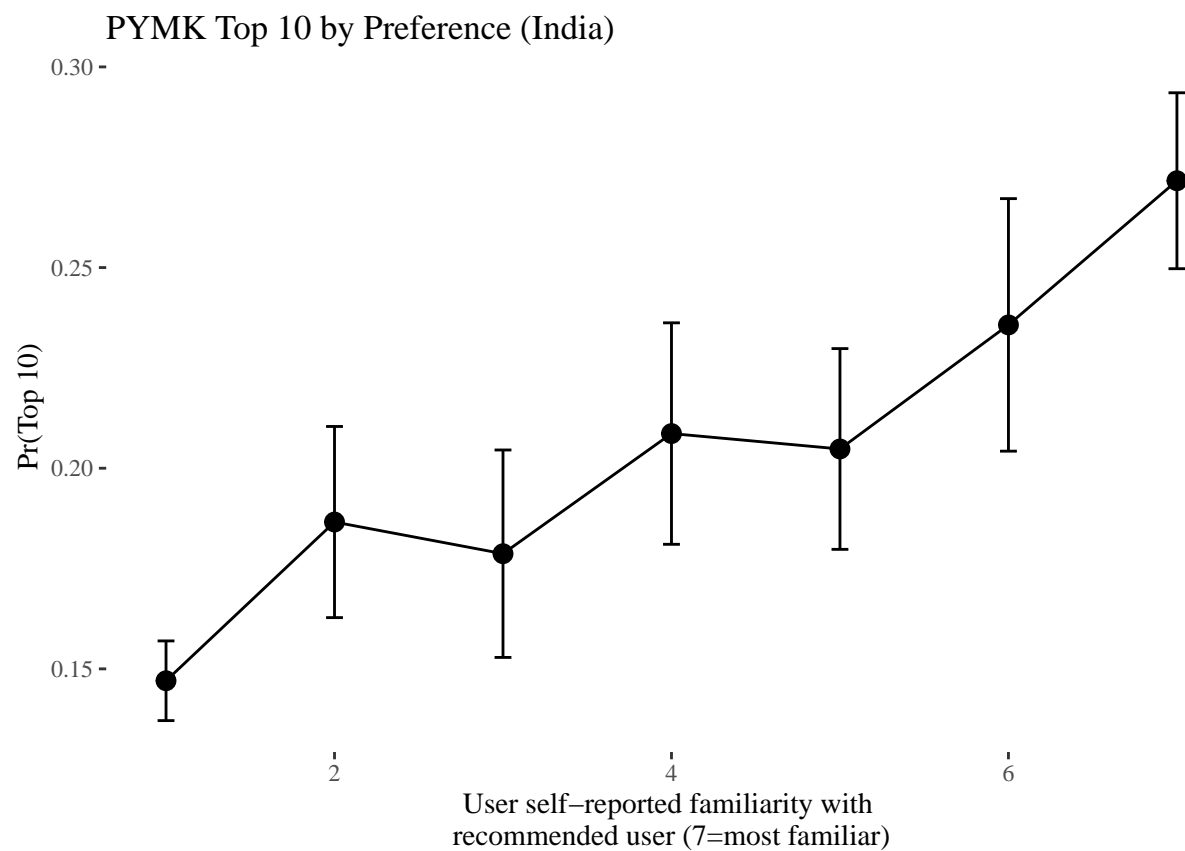


PYMK Ranking by Preference (India)



PYMK Top 5 by Preference (India)

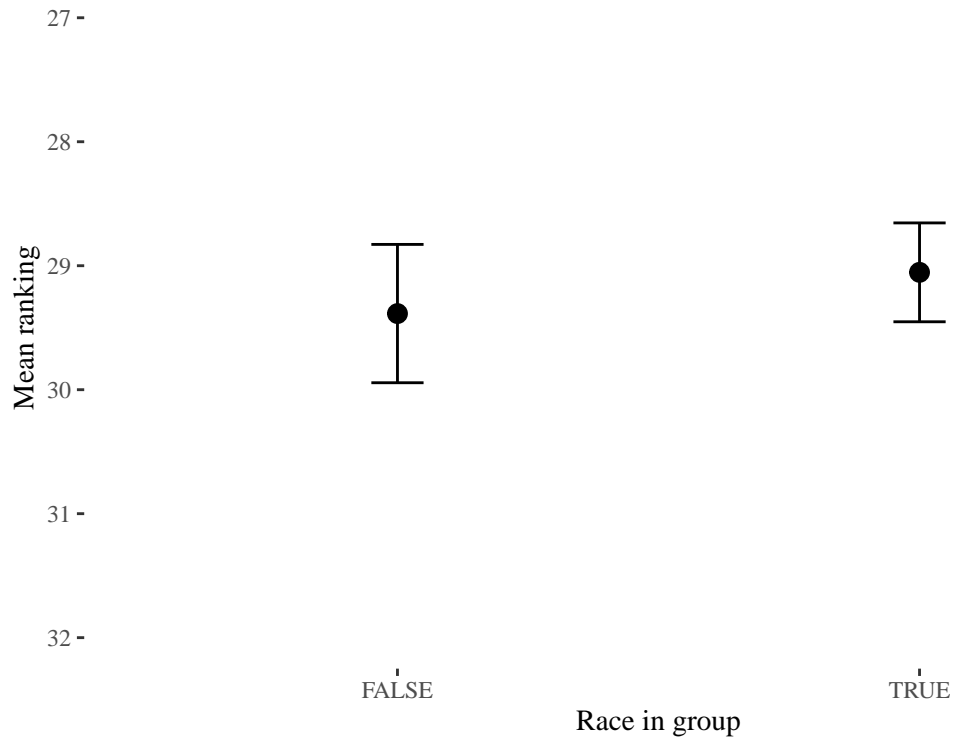




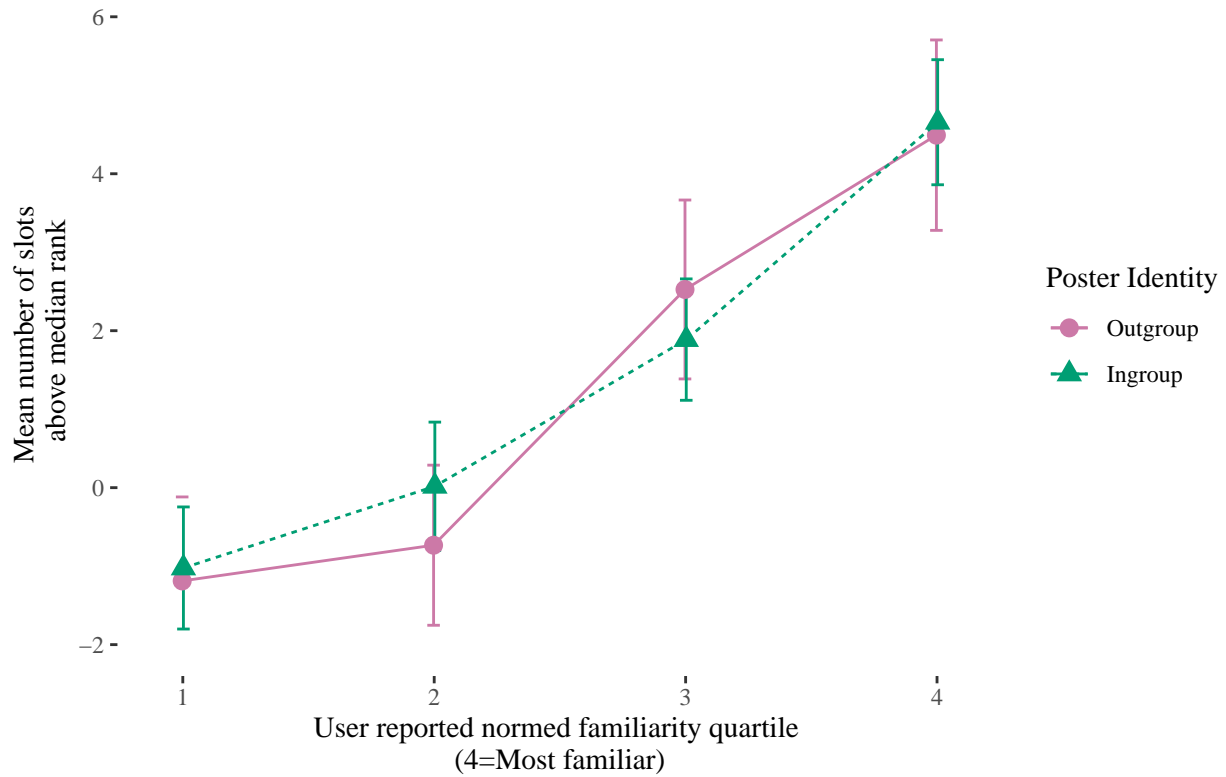
```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

PYMK Ranking by Identity

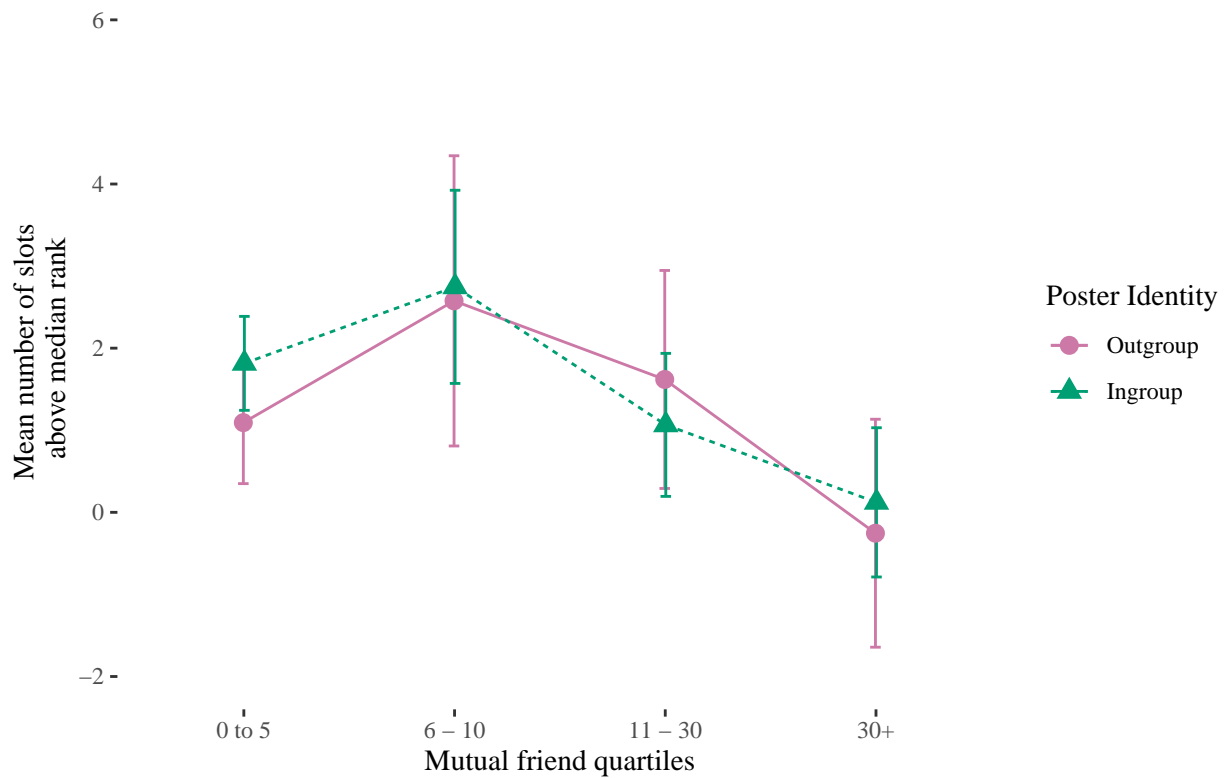
No evidence of differential treatments by religion



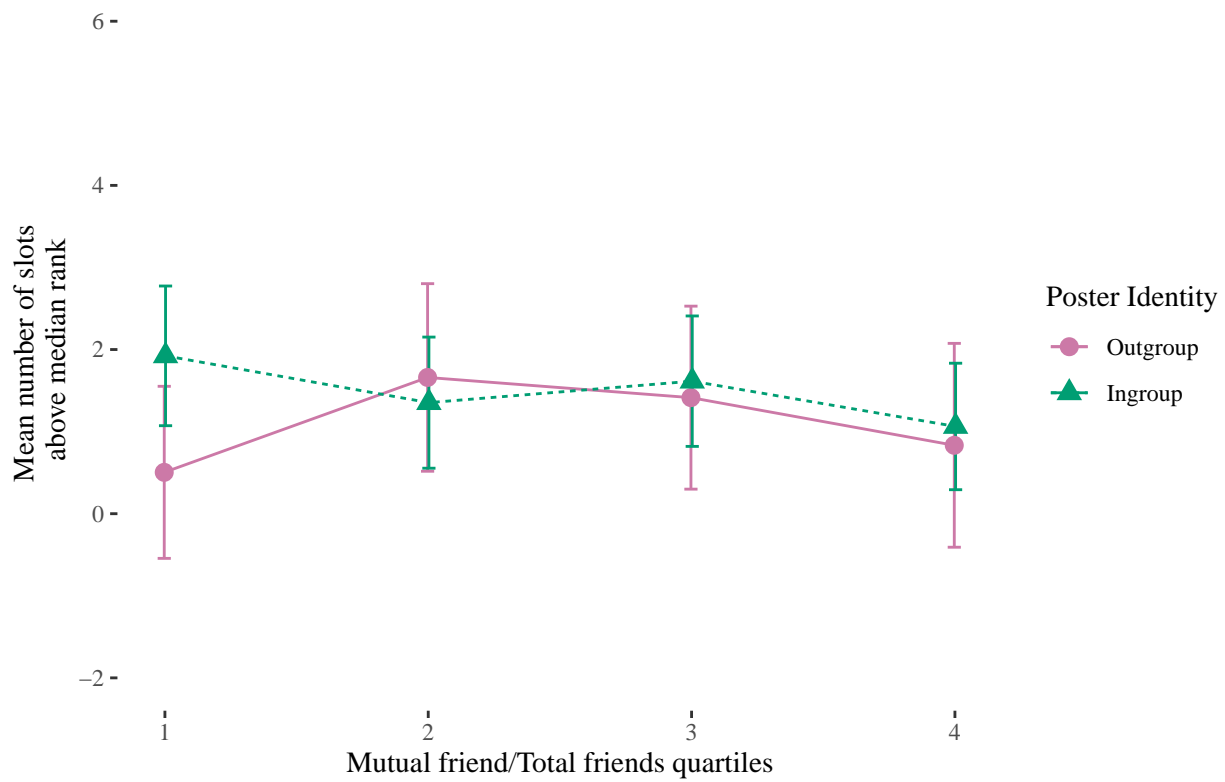
PYMK Ranking by Familiarity (India)



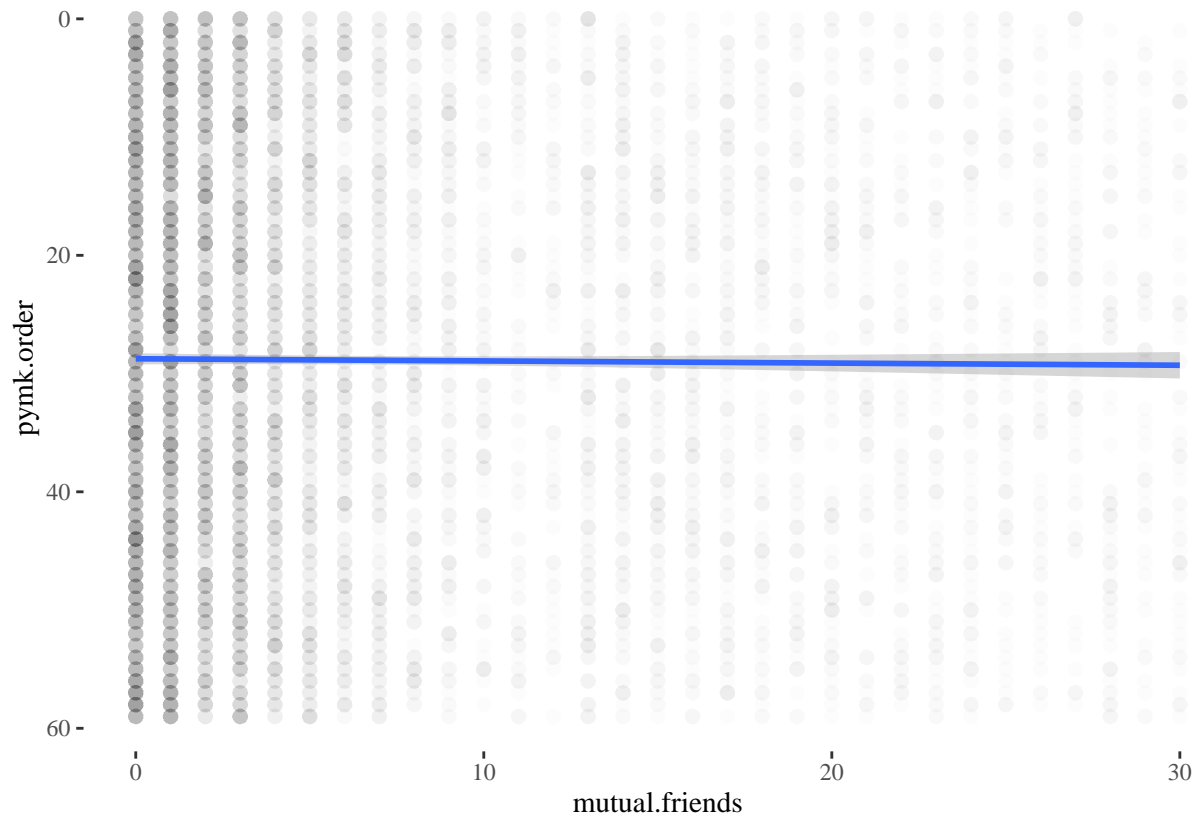
PYMK Ranking by Familiarity (India)

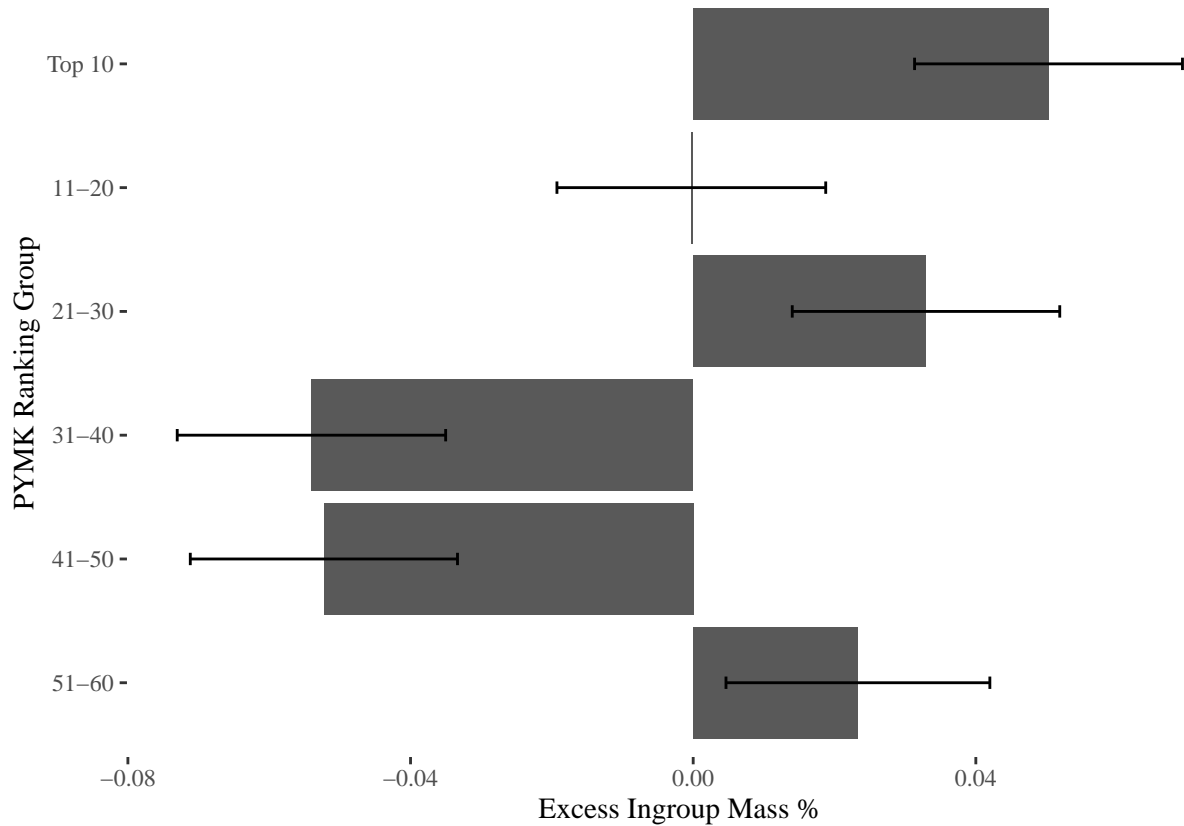


PYMK Ranking by Familiarity (India)



`geom_smooth()` using formula 'y ~ x'





Regression Tables

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Tue, Jan 05, 2021 - 18:08:35

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Tue, Jan 05, 2021 - 18:08:35

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Tue, Jan 05, 2021 - 18:08:36

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Tue, Jan 05, 2021 - 18:08:36

Table 9: Primary NF Rank Results

	<i>Dependent variable:</i>			
	nf.order			
	(1) all.mdl	(2) hindu.mdl	(3) muslim.mdl	(4) other.mdl
religion.in.group	1.126*** (0.397)	1.492*** (0.538)	0.958 (0.692)	2.420 (2.156)
I(100 *norm.pctle)	0.047*** (0.007)	0.046*** (0.009)	0.043*** (0.012)	0.107*** (0.033)
Constant	-31.340*** (0.457)	-31.614*** (0.627)	-30.949*** (0.745)	-34.107*** (2.003)
Observations	7,866	4,912	2,521	319
R ²	0.007	0.007	0.006	0.036
Adjusted R ²	0.007	0.007	0.005	0.030

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 10: Primary NF Top 10 Results

	<i>Dependent variable:</i>			
	top10			
	(1) all.mdl	(2) hindu.mdl	(3) muslim.mdl	(4) other.mdl
I(100 *norm.pctle)	0.001*** (0.0002)	0.001*** (0.0002)	0.001*** (0.0003)	0.002*** (0.001)
religion.in.group	0.005 (0.009)	0.002 (0.012)	0.015 (0.015)	0.060 (0.049)
Constant	0.134*** (0.010)	0.136*** (0.014)	0.138*** (0.017)	0.076* (0.046)
Observations	7,866	4,912	2,521	319
R ²	0.004	0.004	0.004	0.030
Adjusted R ²	0.004	0.003	0.003	0.023

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 11: Primary PYMK Rank Results

	<i>Dependent variable:</i>			
	pymk.order			
	(1) all.mdl	(2) hindu.mdl	(3) muslim.mdl	(4) other.mdl
religion.in.group	0.186 (0.349)	0.200 (0.540)	0.561 (0.584)	−2.191 (2.241)
I(100 *norm.pctle)	0.080*** (0.006)	0.077*** (0.007)	0.089*** (0.010)	0.041 (0.028)
Constant	−33.272*** (0.398)	−33.282*** (0.592)	−33.639*** (0.647)	−31.209*** (1.639)
Observations	10,882	6,833	3,449	480
R ²	0.018	0.017	0.022	0.007
Adjusted R ²	0.018	0.017	0.021	0.003
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01		

Table 12: Primary PYMK Top 10 Results

	<i>Dependent variable:</i>			
	top10			
	(1) all.mdl	(2) hindu.mdl	(3) muslim.mdl	(4) other.mdl
I(100 *norm.pctle)	0.002*** (0.0001)	0.001*** (0.0002)	0.002*** (0.0002)	0.001 (0.001)
religion.in.group	0.003 (0.008)	0.015 (0.012)	−0.007 (0.013)	−0.027 (0.050)
Constant	0.107*** (0.009)	0.100*** (0.013)	0.103*** (0.015)	0.139*** (0.037)
Observations	10,882	6,833	3,449	480
R ²	0.013	0.012	0.017	0.006
Adjusted R ²	0.013	0.012	0.016	0.002
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01		