Facebook Results Memo

ddd

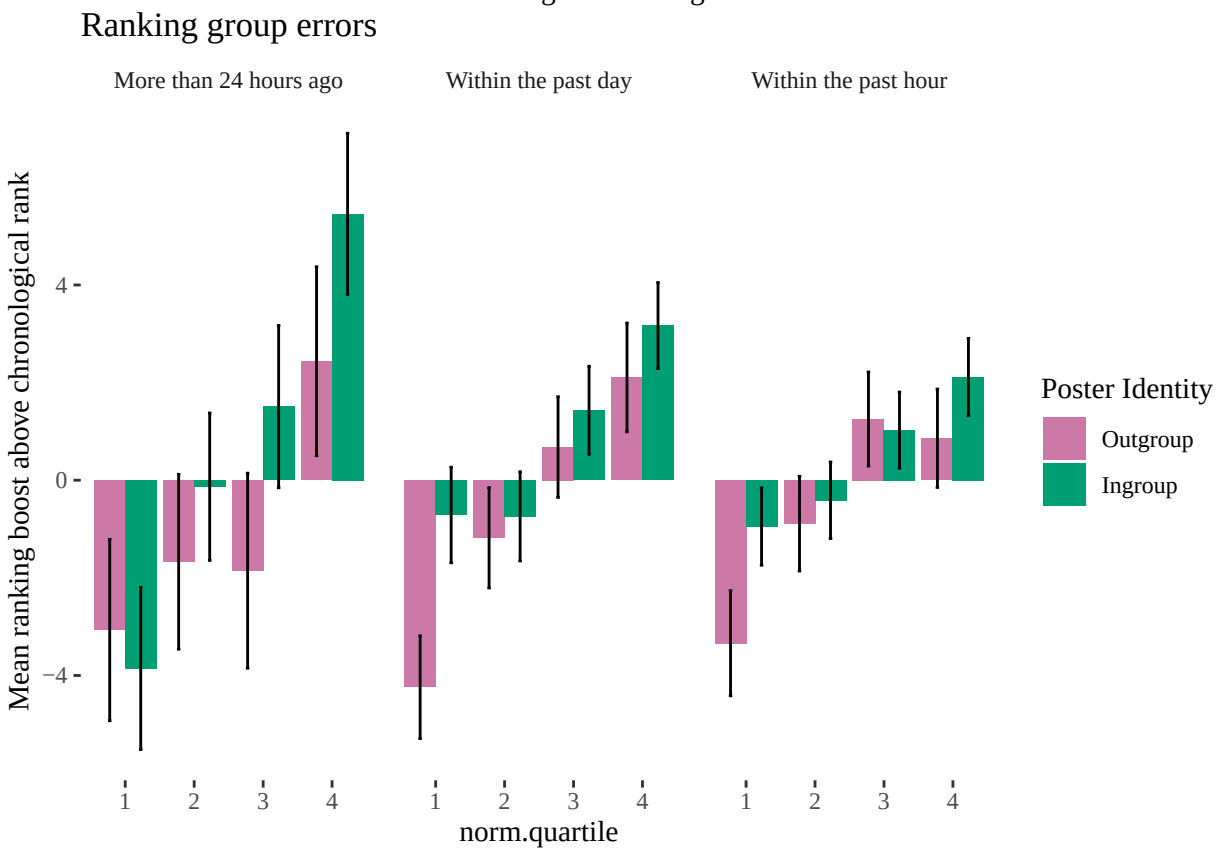
8/27/2020

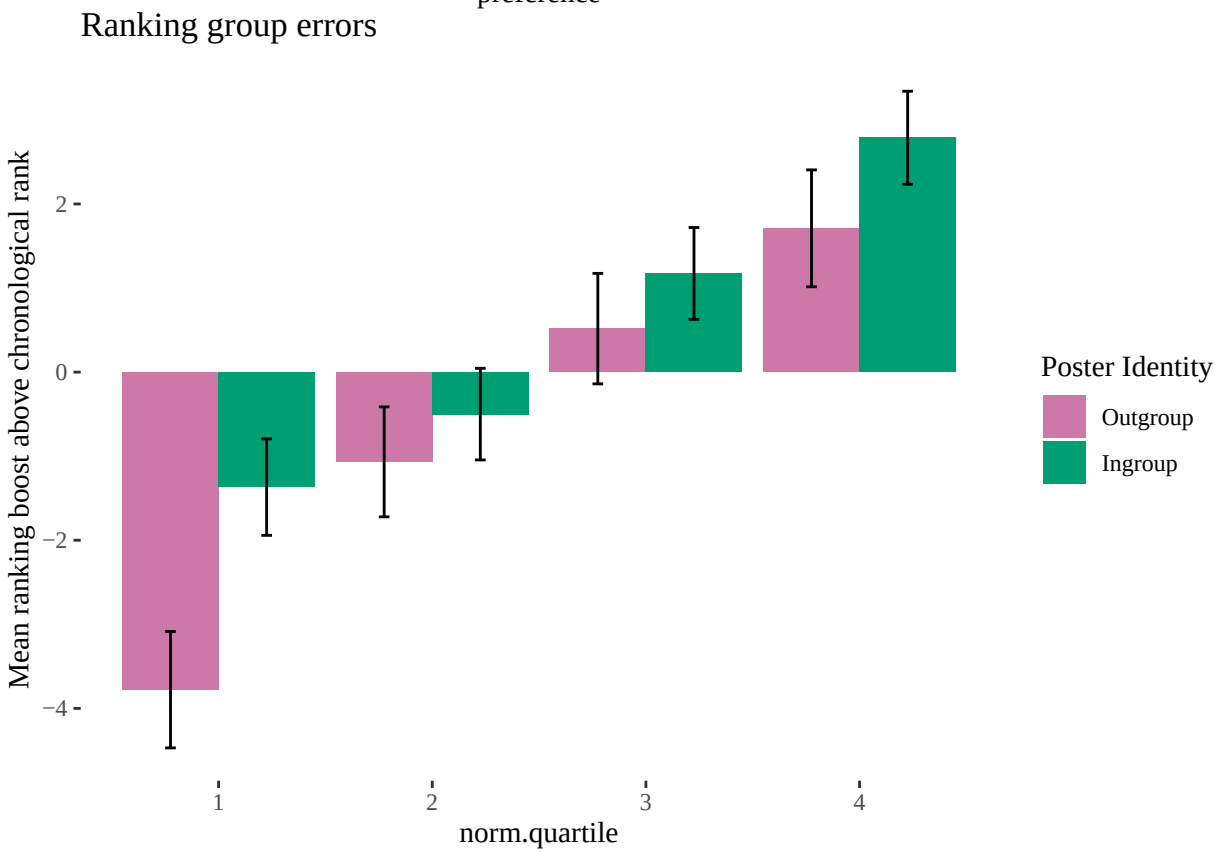
Executive Summary

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 26, 2021 - 14:21:56

Table 1:

	Dependent variable:					
	new.rank					
	(1)	(2)	(3)	(4)	(5)	(6)
race.in.group	−0.924533***			−0.897882***	−0.994101***	
	(0.181622)			(0.180885)	(0.176421)	
I(100 ∗norm.pctle)		−0.047141***		−0.046995***		−0.051191***
		(0.003068)		(0.003066)		(0.002978)
time_rank			0.237119***		0.237430***	0.240324***
			(0.005768)		(0.005765)	(0.005741)
Constant	24.534850***	26.333440***	18.278960***	26.867460***	18.870830***	20.760330***
	(0.141024)	(0.177040)	(0.163313)	(0.207102)	(0.194101)	(0.217346)
Observations	28,288	28,288	28,288	28,288	28,288	28,288
R^2	0.000915	0.008280	0.056385	0.009143	0.057443	0.066138
Adjusted R^2	0.000880	0.008245	0.056352	0.009073	0.057377	0.066072
Note:						*p<0.1; **p<0

% 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 26, 2021 - 14:21:56

Table 2:

	Dependent variable:					
	new.rank					
	(1)	(2)	(3)	(4)	(5)	(6)
race.in.group	0.055384			0.227486	0.018597	
	(0.219362)			(0.216976)	(0.219569)	
I(100 ∗norm.pctle)		−0.090238***		−0.090365***		−0.097824***
		(0.003704)		(0.003706)		(0.003790)
friend_rank			−0.019711***		−0.019688***	−0.052120***
			(0.005651)		(0.005658)	(0.005719)
Constant	30.301490***	34.851030***	30.954010***	34.724300***	30.942410***	36.870530***
	(0.167764)	(0.213997)	(0.208068)	(0.245773)	(0.249121)	(0.307819)
Observations	25,593	25,593	25,593	25,593	25,593	25,593
R^2	0.000002	0.022669	0.000475	0.022711	0.000475	0.025830
Adjusted R^2	−0.000037	0.022630	0.000436	0.022634	0.000397	0.025754
Note:						*p<0.1; **p<0

% 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 26, 2021 - 14:21:57

Table 3:

	Dependent variable:					
	new.rank					
	(1)	(2)	(3)	(4)	(5)	(6)
race.in.group	0.055384			0.227486	0.116784	
	(0.219362)			(0.216976)	(0.218006)	
I(100 ∗norm.pctle)		−0.090238***		−0.090365***		−0.088493***
		(0.003704)		(0.003706)		(0.003683)
pct_friend_rank			0.101516***		0.101563***	0.097926***
			(0.005617)		(0.005618)	(0.005557)
Constant	30.301490***	34.851030***	27.140020***	34.724300***	27.070240***	31.682750***
	(0.167764)	(0.213997)	(0.206802)	(0.245773)	(0.244410)	(0.278514)
Observations	25,593	25,593	25,593	25,593	25,593	25,593
R^2	0.000002	0.022669	0.012602	0.022711	0.012613	0.034387
Adjusted R^2	−0.000037	0.022630	0.012564	0.022634	0.012536	0.034312
Note:						*p<0.1; **p<0.0

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 26, 2021 - 14:22:16

Table 4: Correlations 1

	Pearson	Spearman	Kendall
NF Rank, Time	0.237	0.243	0.169
PYMK Rank, Pct Friends	0.112	0.125	0.086

Correlation matrix for benchmarks

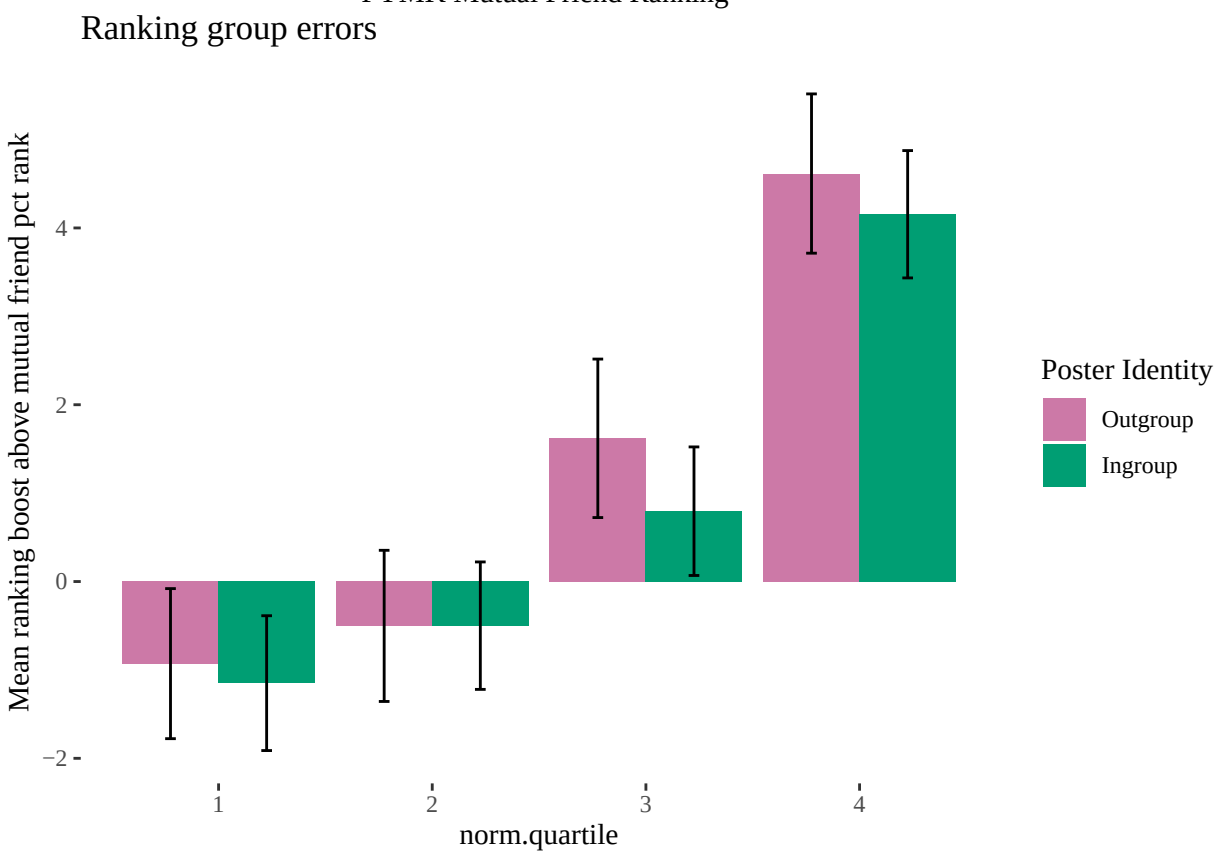
preference correlations

% 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 26, 2021 - 14:22:34

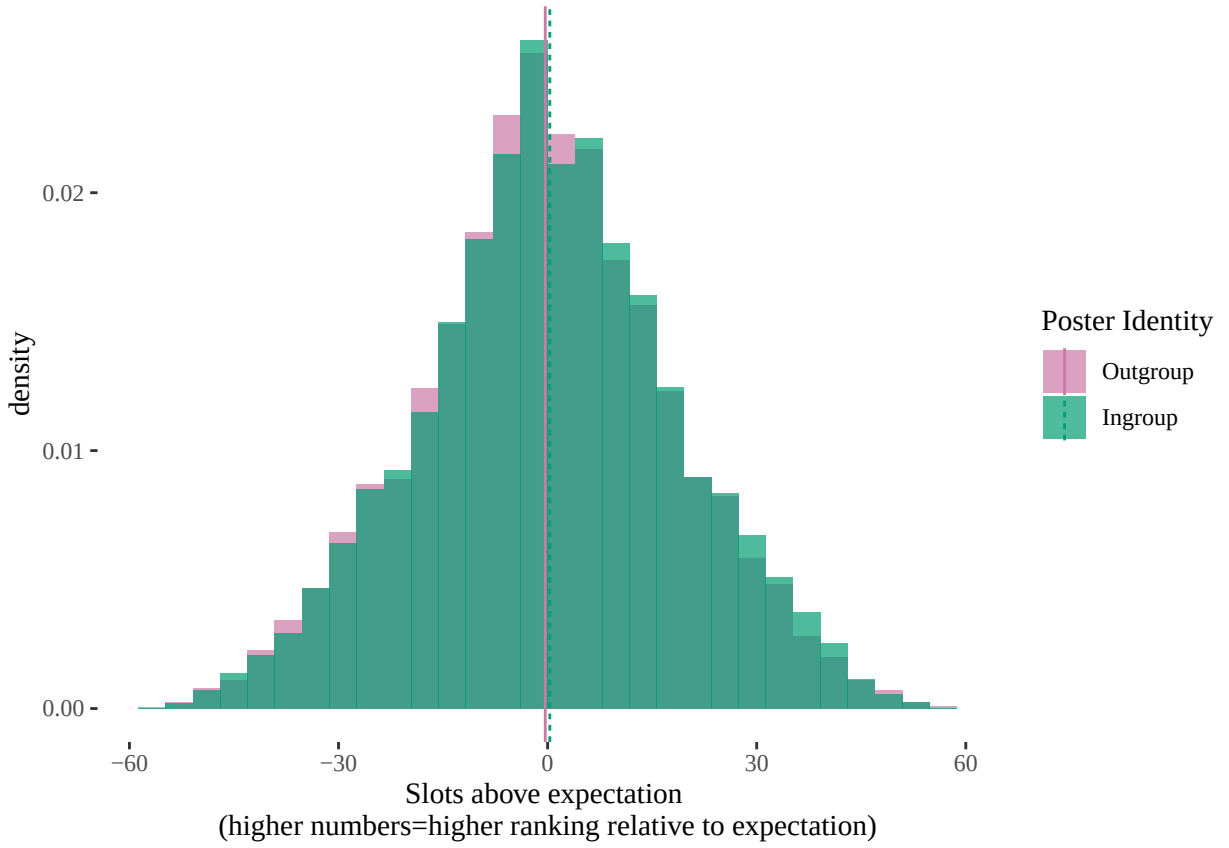
Table 5: Correlations 2

	Pearson	Spearman	Kendall
NF Rank, Preference	-0.095	-0.093	-0.063
PYMK Rank, Familiarity	-0.164	-0.151	-0.102

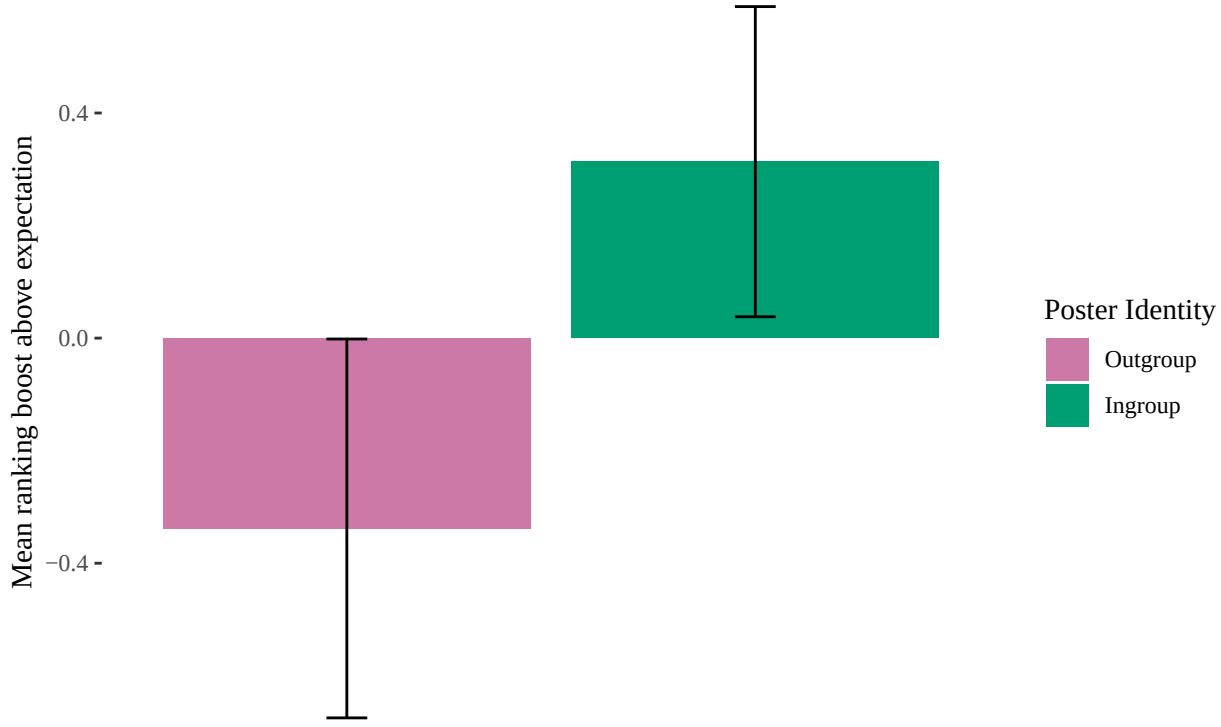
Correlation matrix for preferences

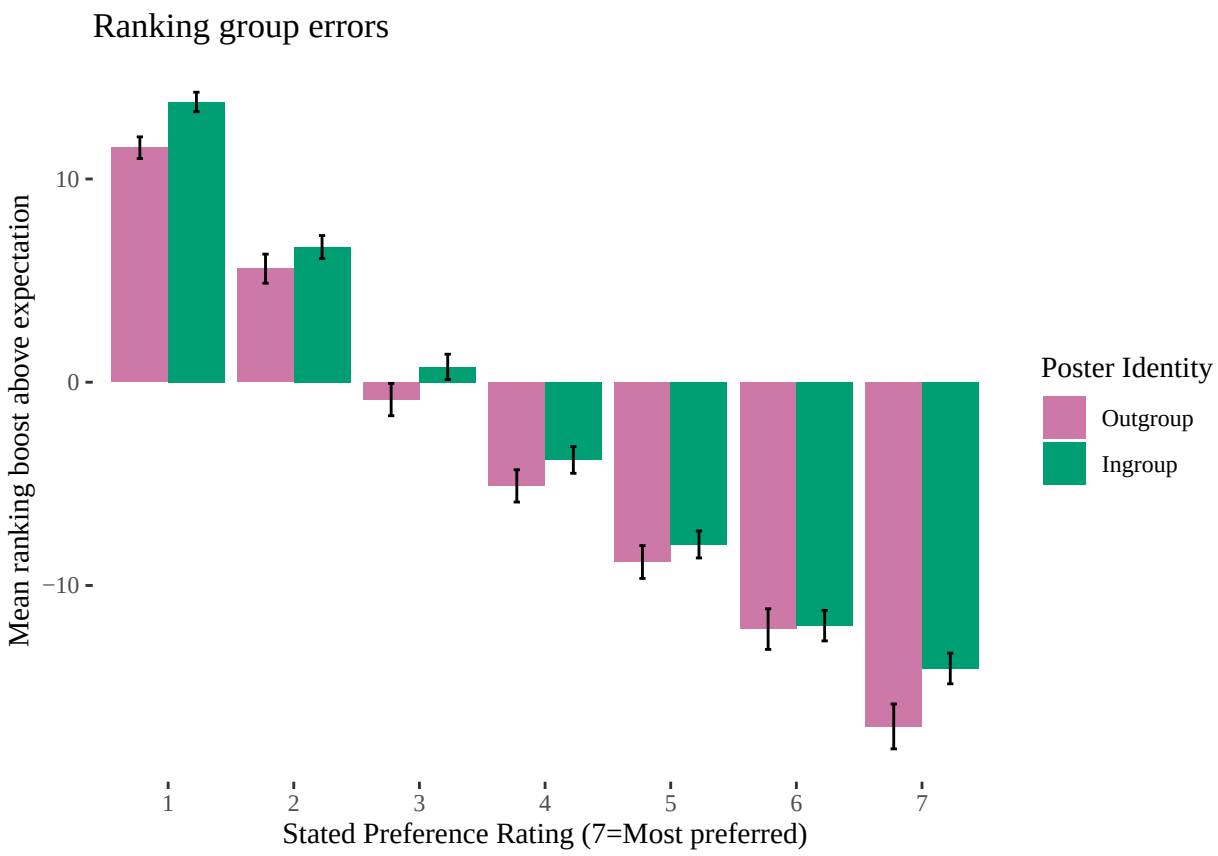


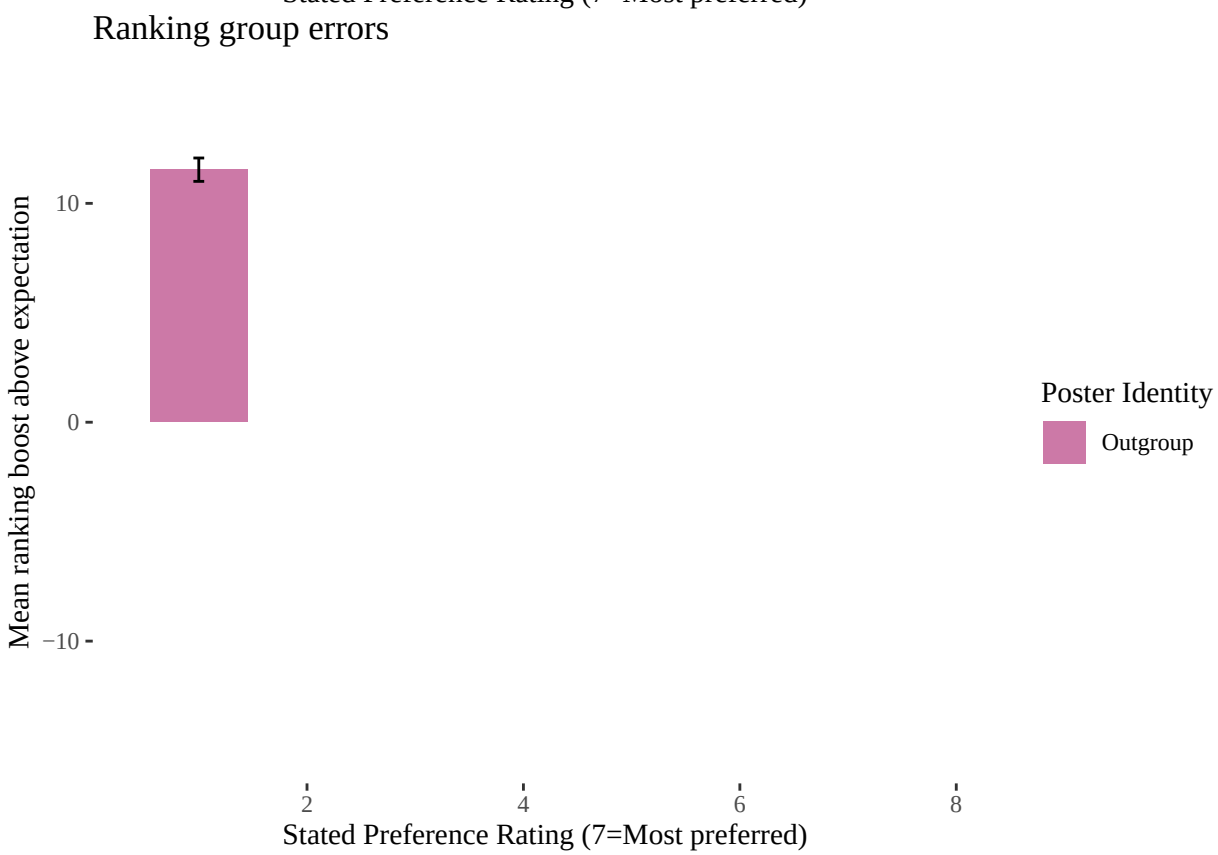

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

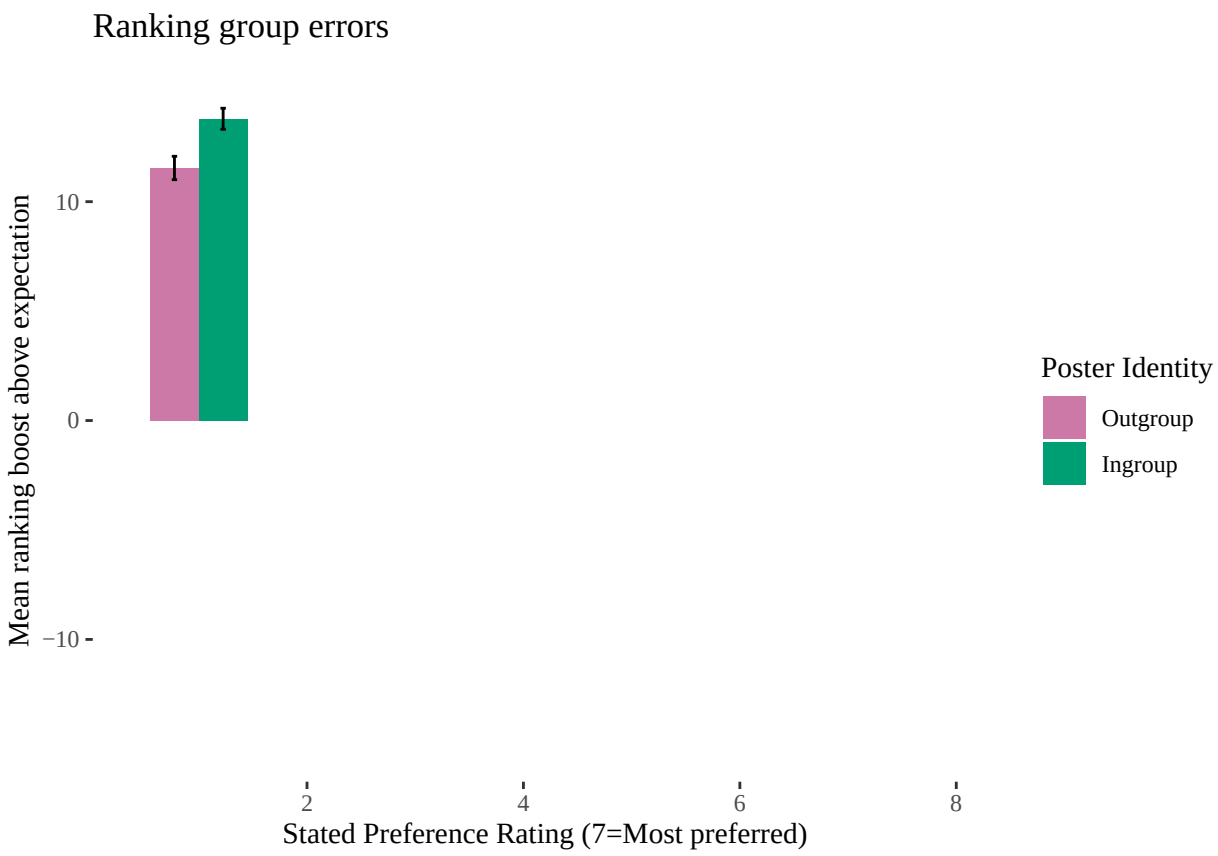


Ranking group errors



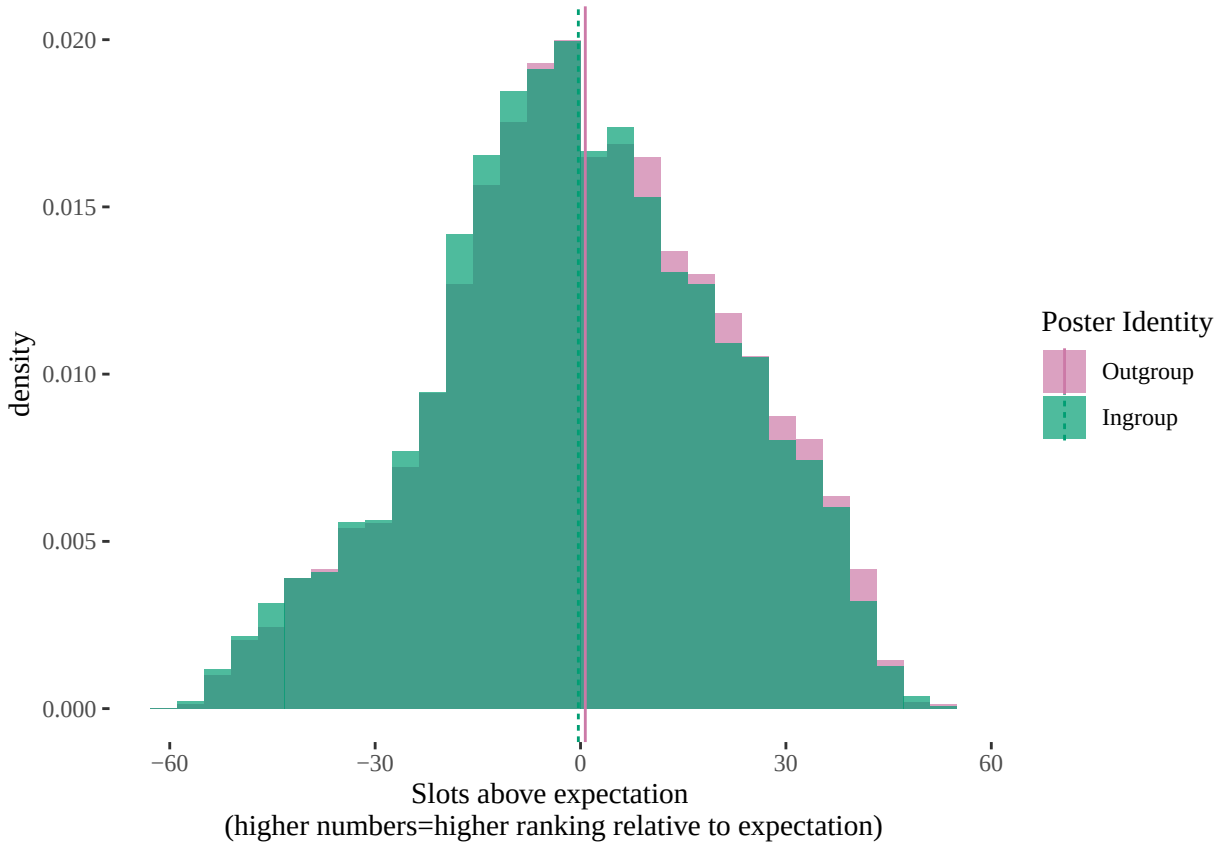

Ranking group errors
Mean ranking boost above expectation
10
0
−10
1
2
3
4
5
6
7
Stated Preference Rating (7=Most preferred)
Poster Identity
Outgroup
Ingroup


Ranking group errors
Mean ranking boost above expectation
10
0
−10
2
4
6
8
Stated Preference Rating (7=Most preferred)
Poster Identity
Outgroup

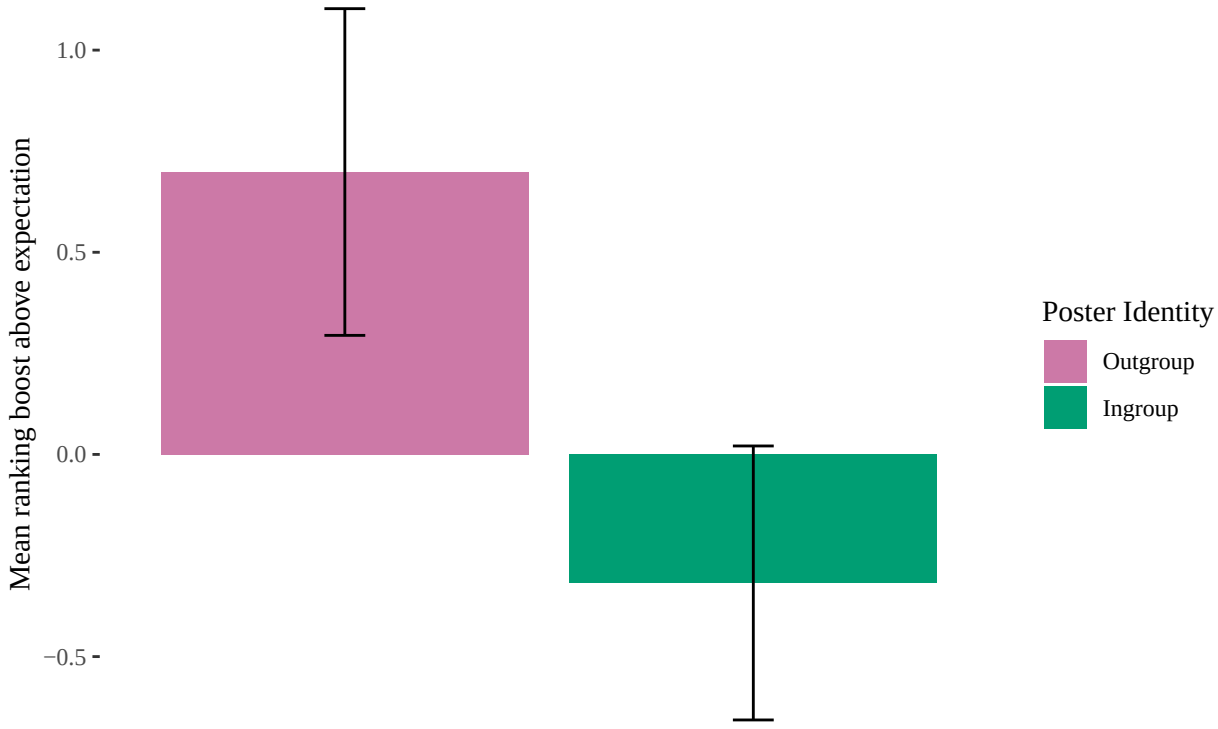

Ranking group errors
Mean ranking boost above expectation
10
0
−10
2
4
6
8
Stated Preference Rating (7=Most preferred)
Poster Identity
Outgroup
Ingroup

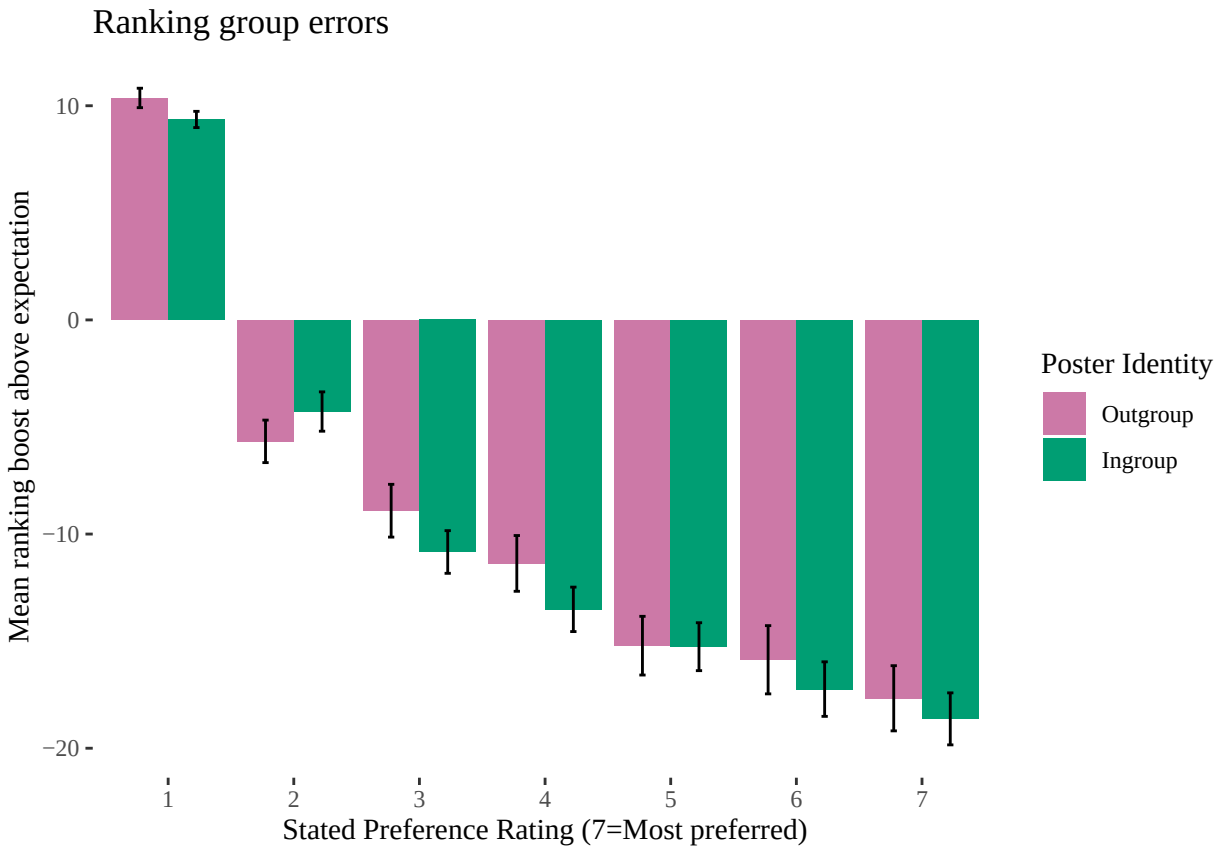
Ranking group errors
Mean ranking boost above expectation
10
0
−10
1
2
3
4
Normed Preference Rating Quartile (4=Most preferred)
Poster Identity
Outgroup
Ingroup

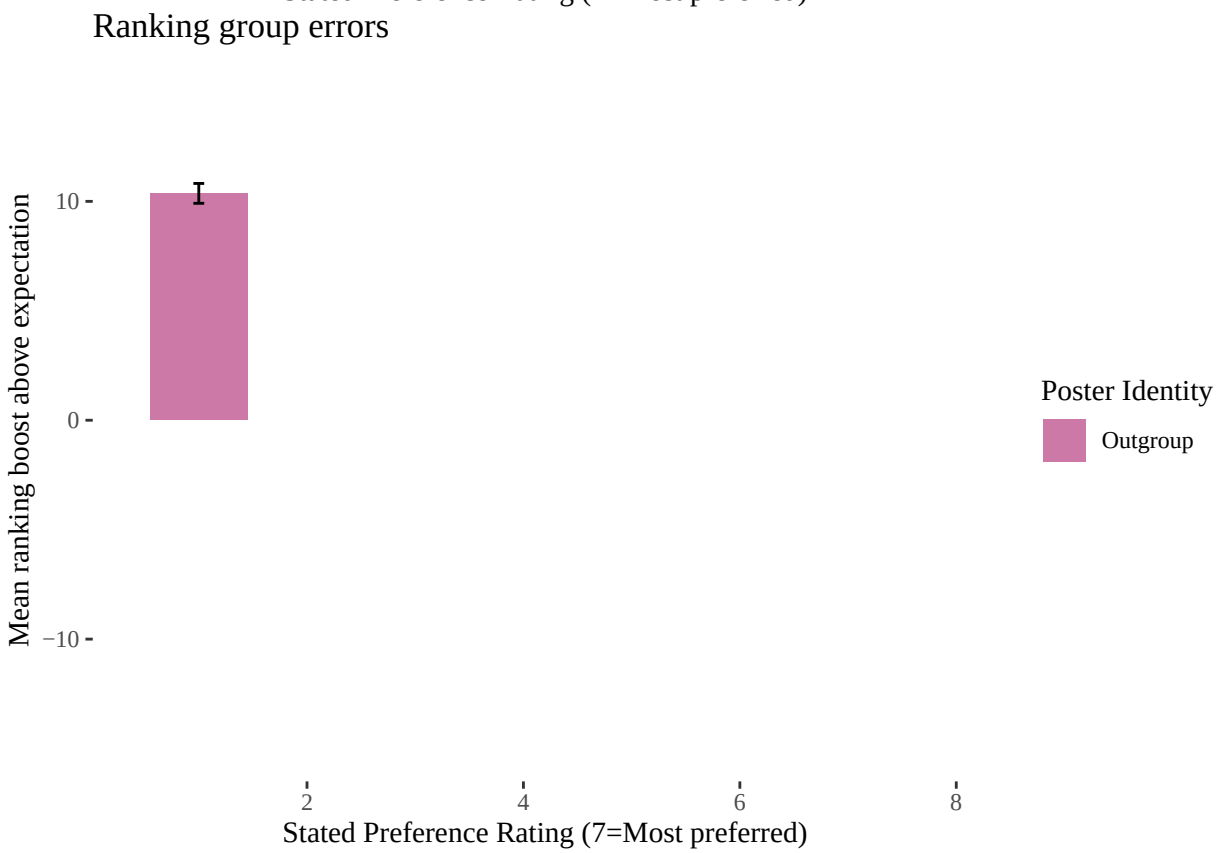
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

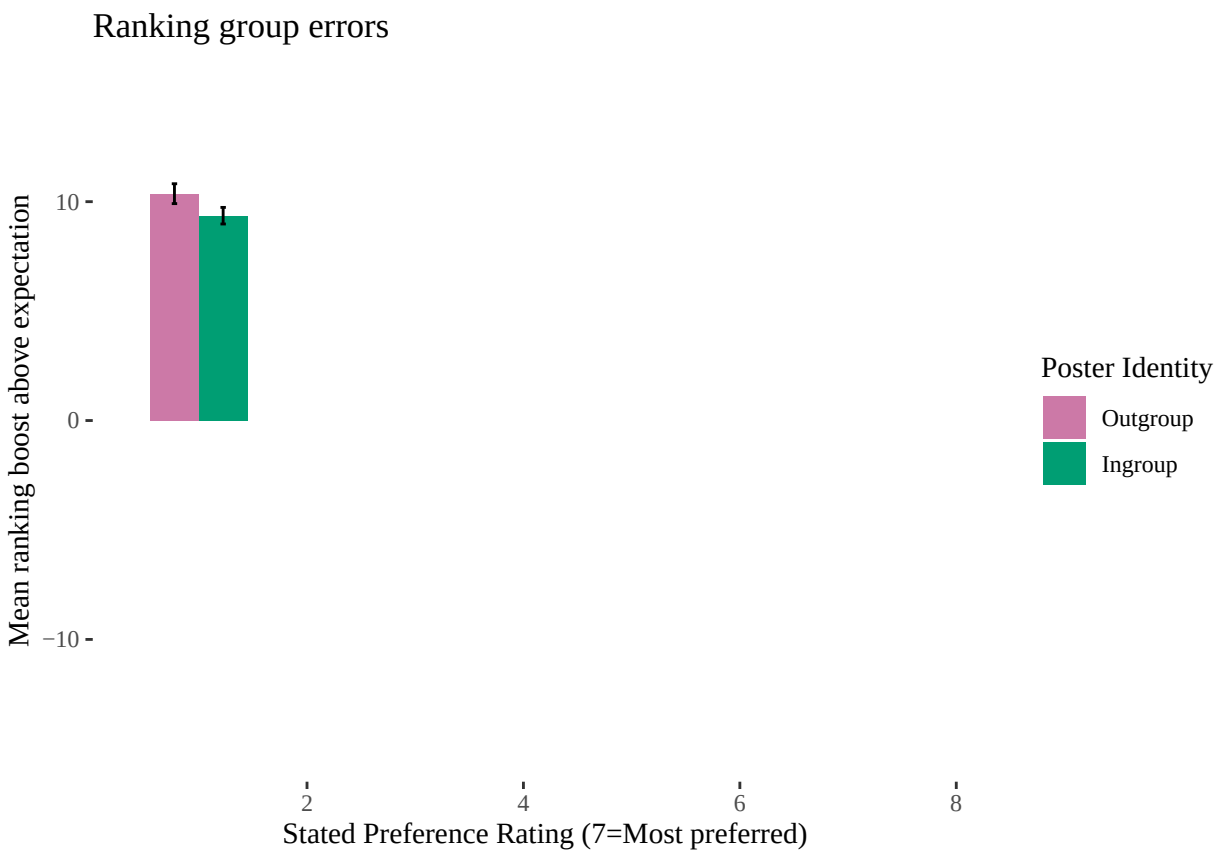


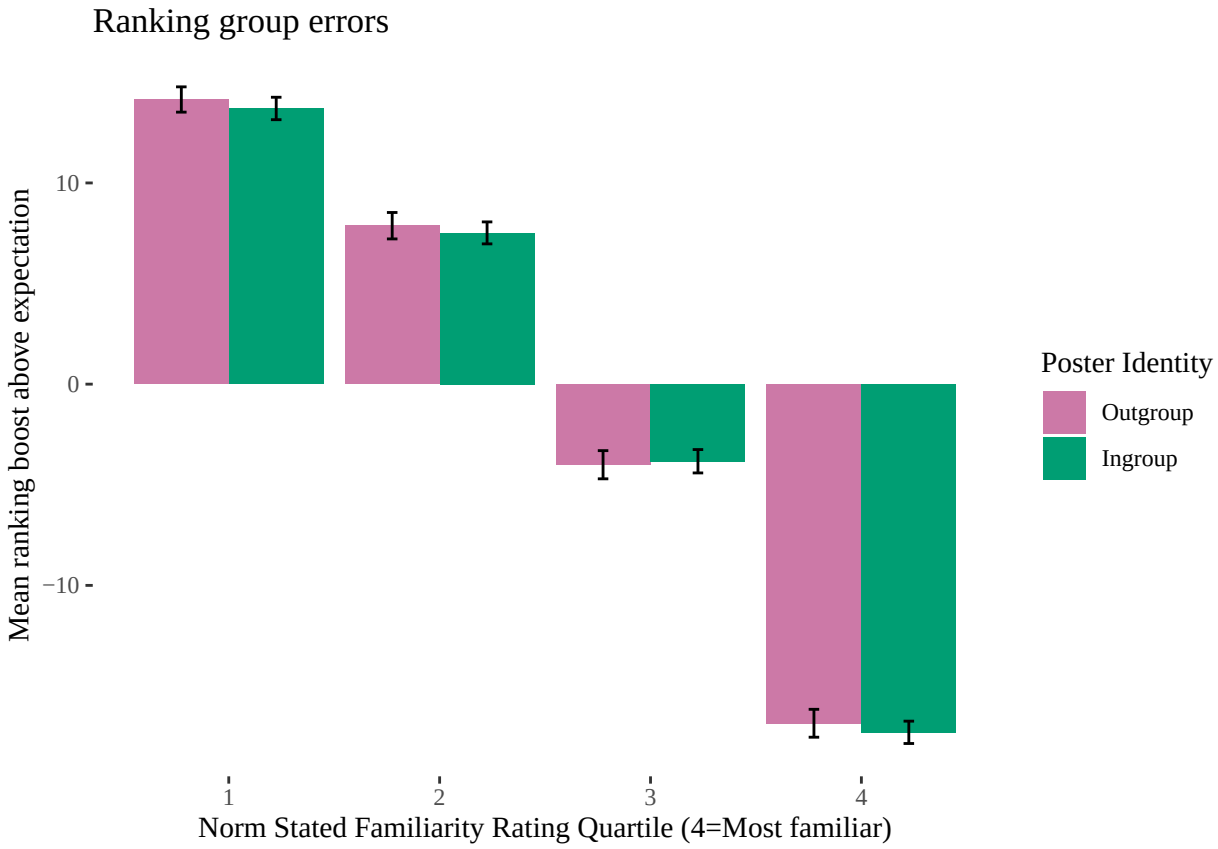
Ranking group errors



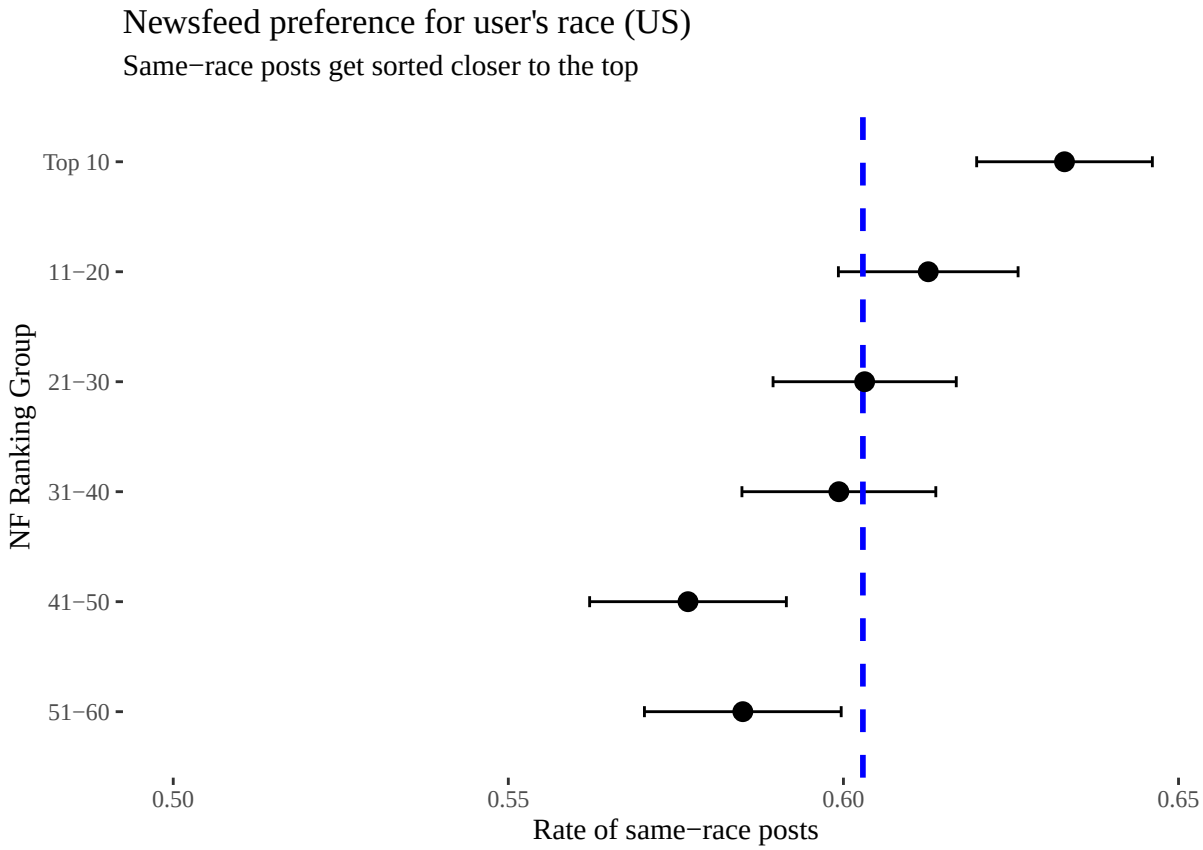

Ranking group errors
Mean ranking boost above expectation
10
0
−10
−20
1
2
3
4
5
6
7
Stated Preference Rating (7=Most preferred)
Poster Identity
Outgroup
Ingroup

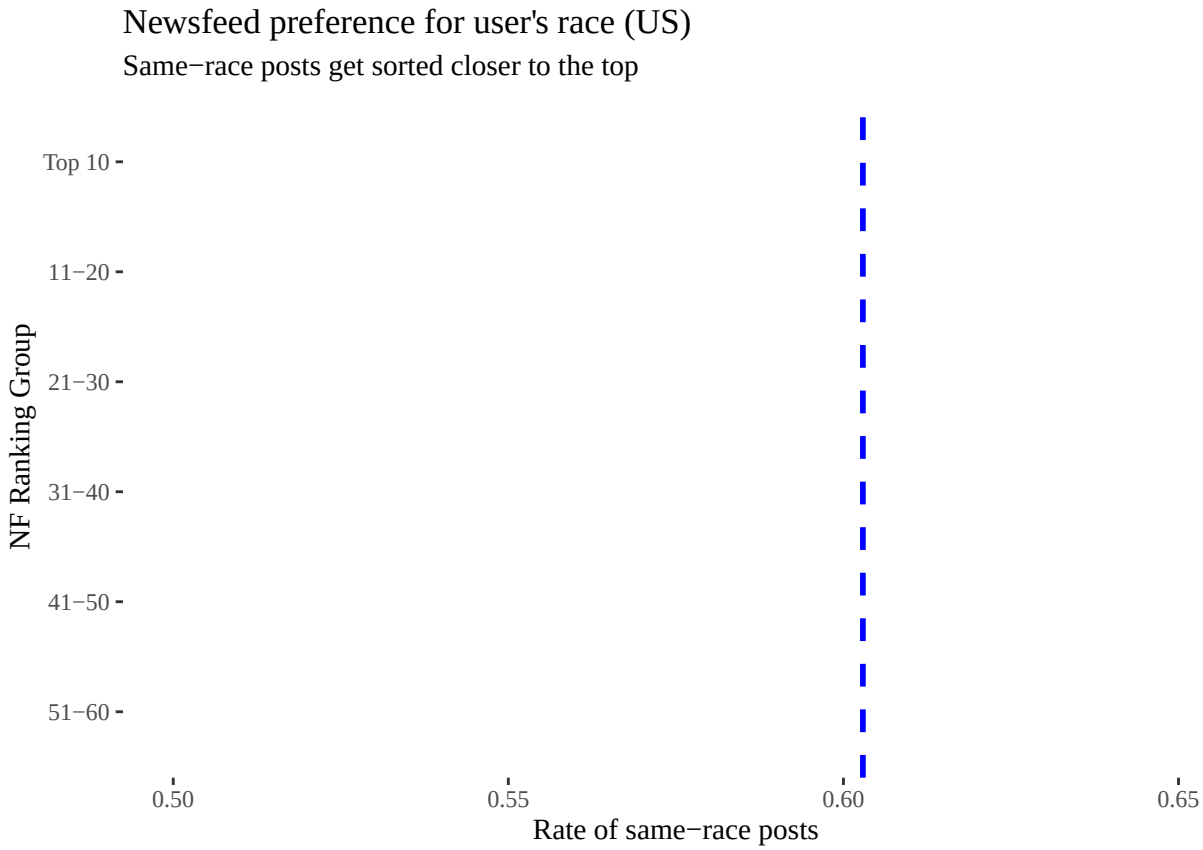

Ranking group errors
Mean ranking boost above expectation
10
0
−10
2
4
6
8
Stated Preference Rating (7=Most preferred)
Poster Identity
Outgroup


Ranking group errors
Mean ranking boost above expectation
10
0
−10
2
4
6
8
Stated Preference Rating (7=Most preferred)
Poster Identity
Outgroup
Ingroup

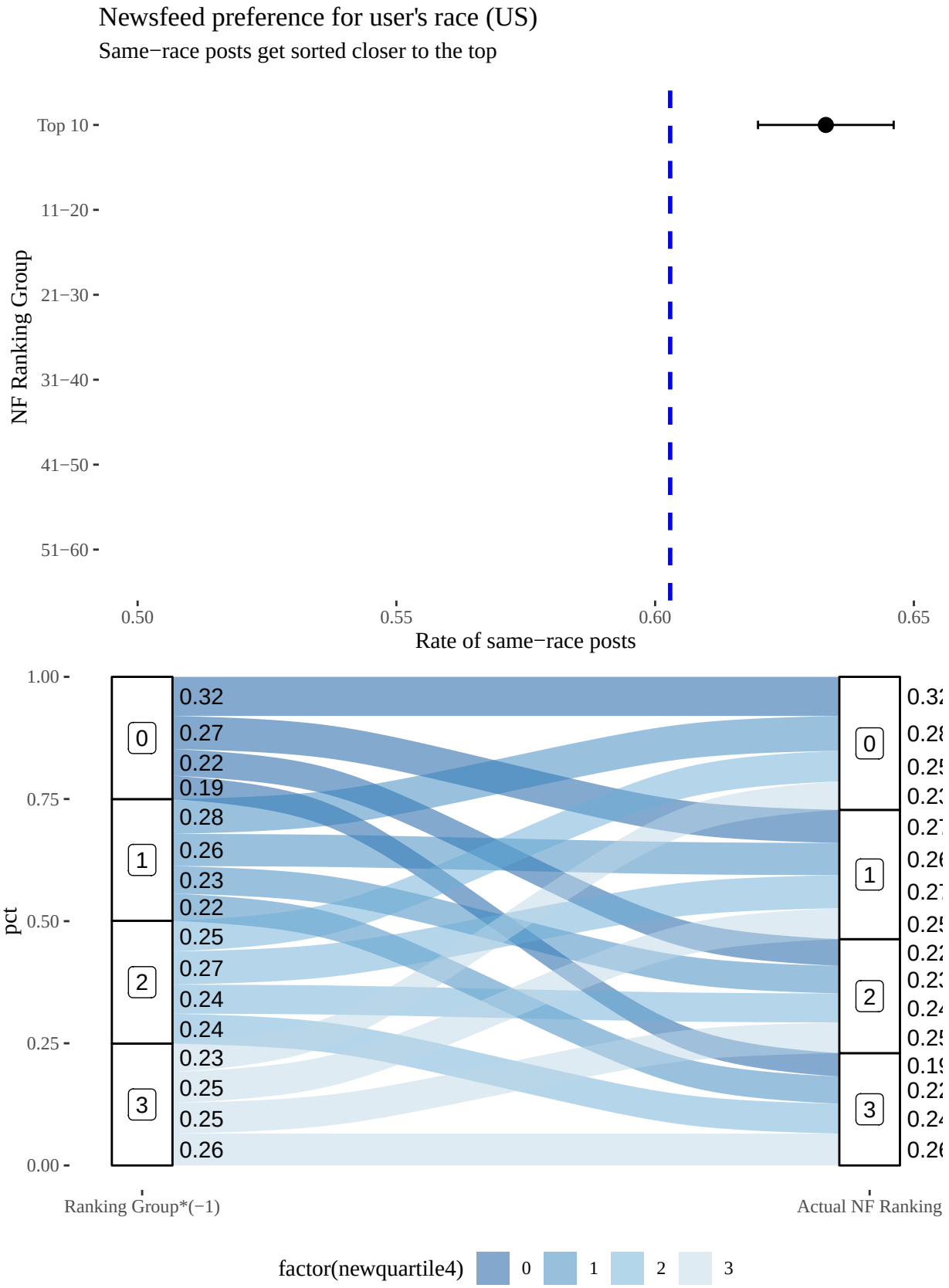

Ranking group errors
Mean ranking boost above expectation
10
0
−10
1
2
3
4
Norm Stated Familiarity Rating Quartile (4=Most familiar)
Poster Identity
Outgroup
Ingroup

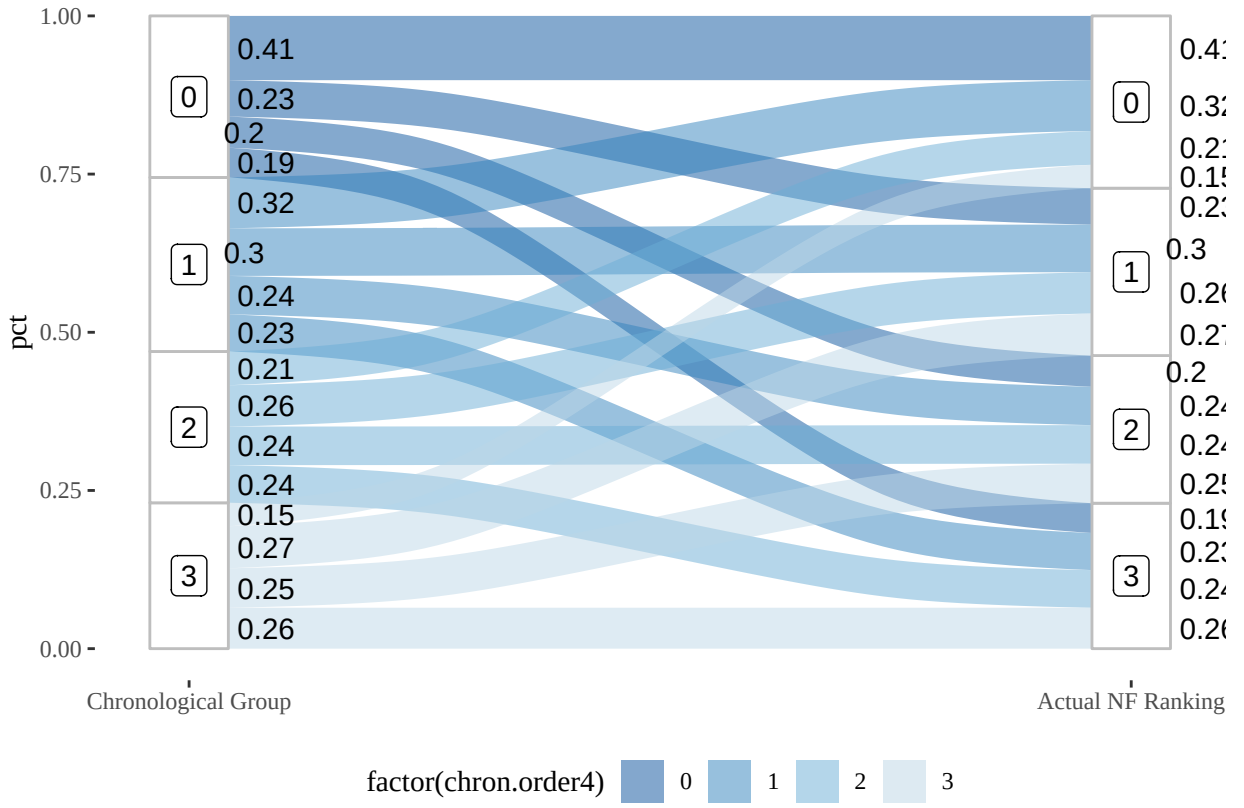
```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

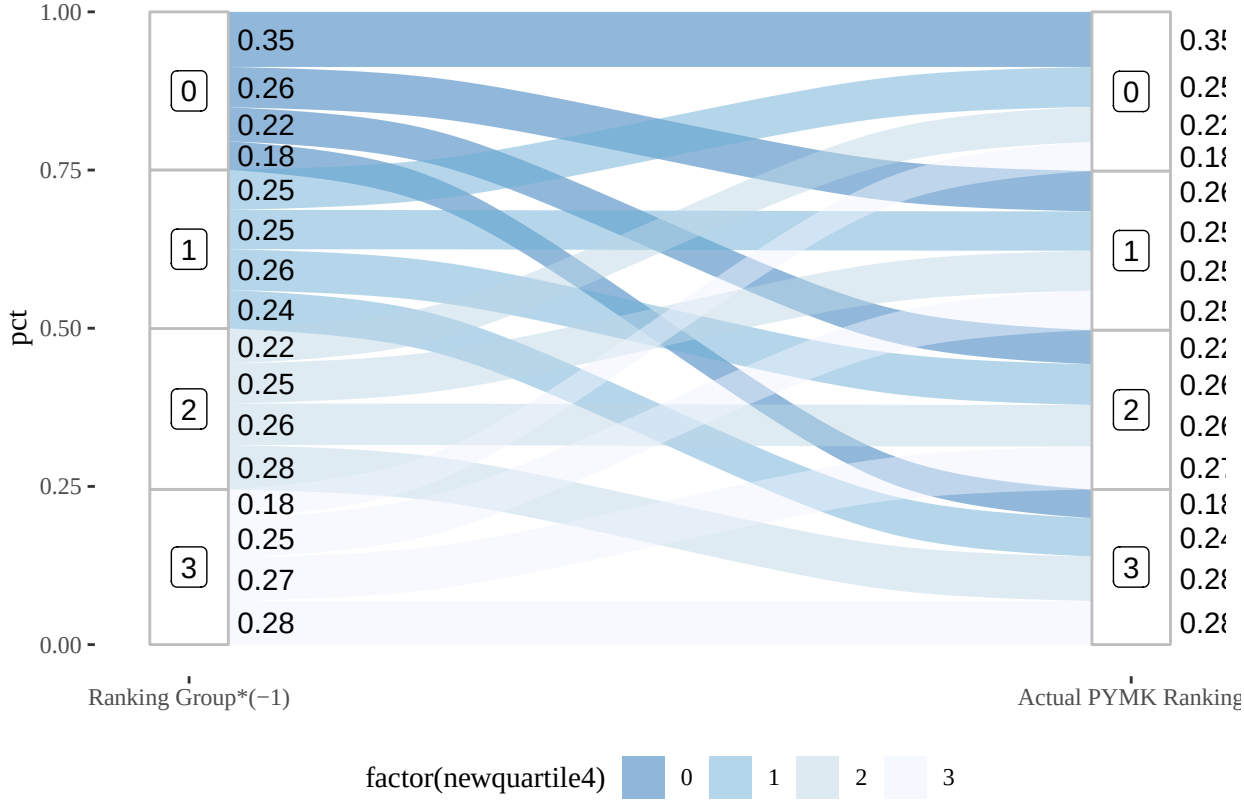


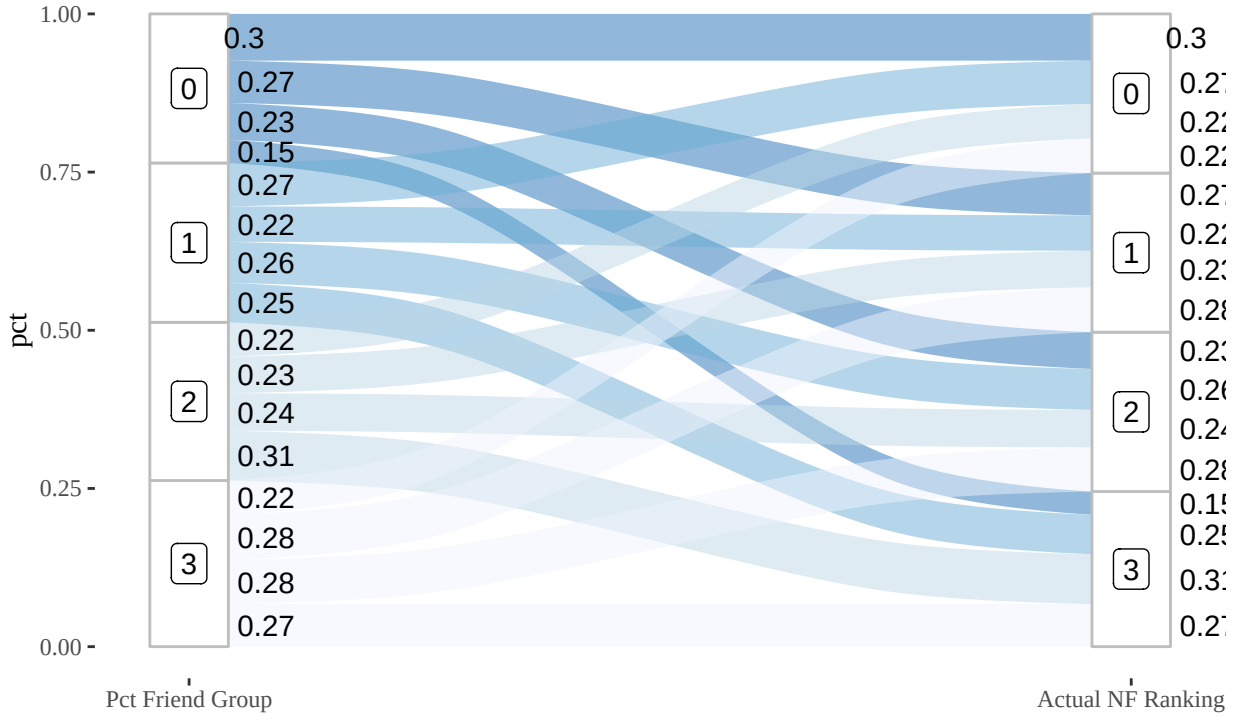


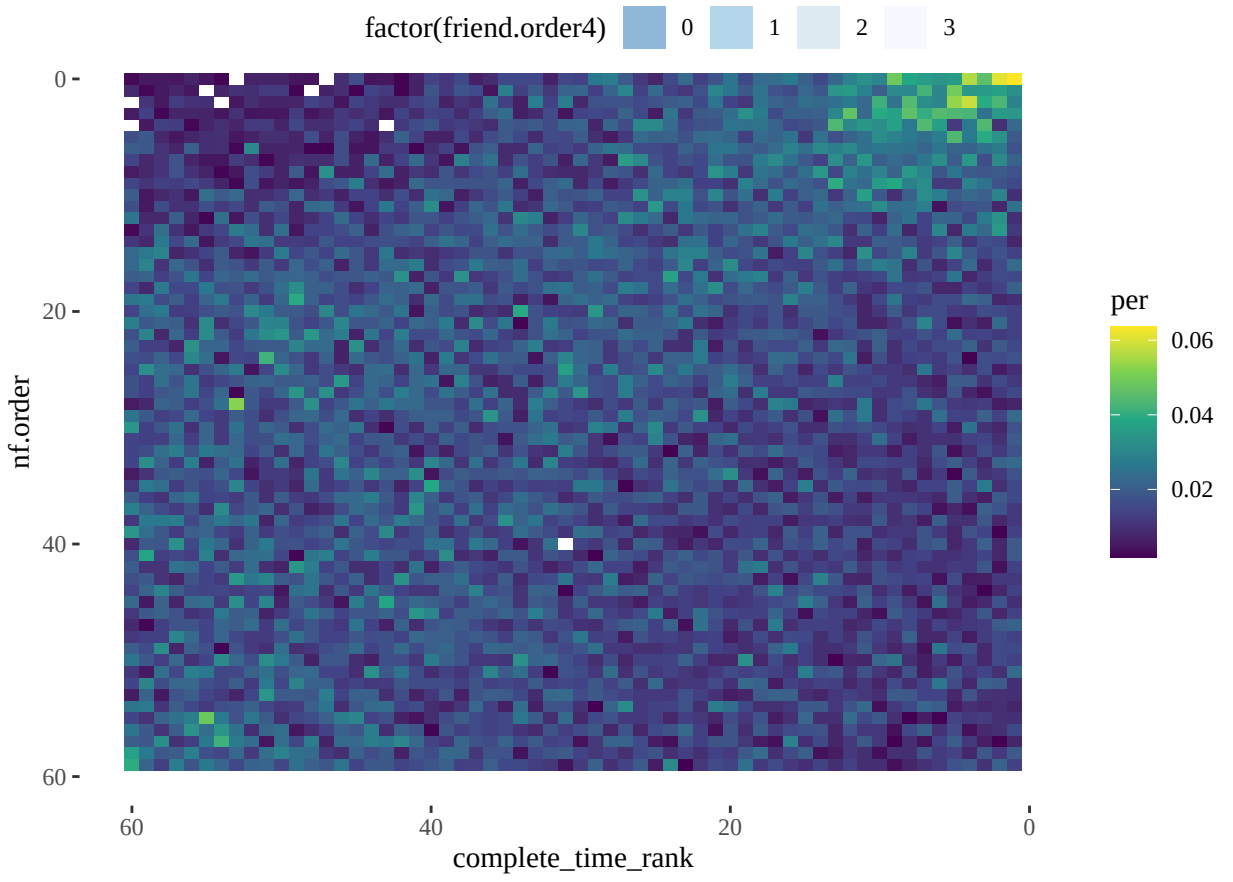

```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

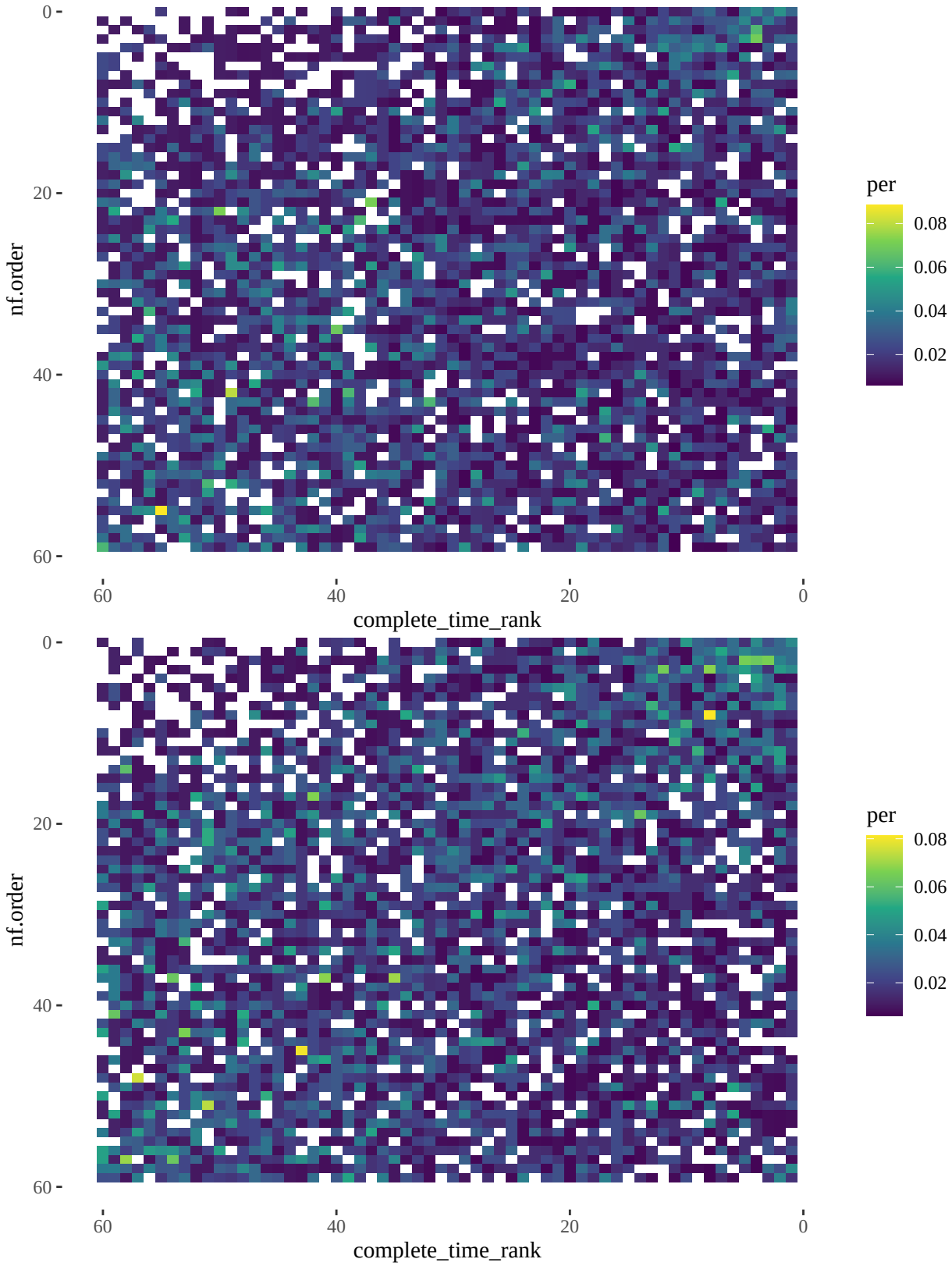

Newsfeed preference for user's race (US)
Same−race posts get sorted closer to the top
NF Ranking Group
Top 10
11−20
21−30
31−40
41−50
51−60
0.50
0.55
0.60
0.65
Rate of same−race posts
pct
1.00
0.75
0.50
0.25
0.00
0
1
2
3
0.32
0.27
0.22
0.19
0.28
0.26
0.23
0.22
0.25
0.27
0.24
0.24
0.23
0.25
0.25
0.26
Ranking Group*(−1)
Actual NF Ranking
factor(newquartile4)
0
1
2
3

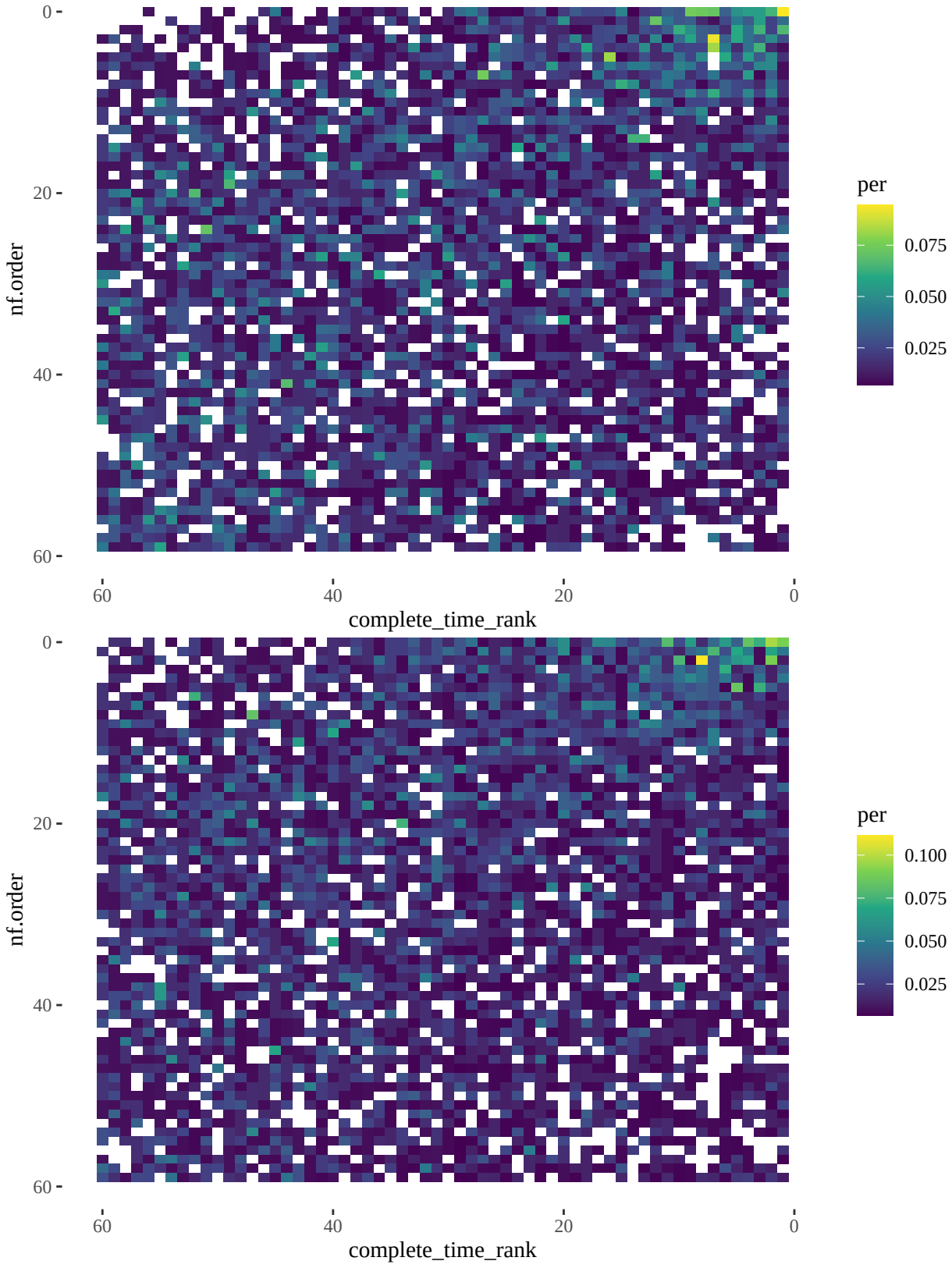

1.00
0.75
0.50
0.25
0.00
pct
0
1
2
3
0.41
0.23
0.2
0.19
0.32
0.3
0.24
0.23
0.21
0.26
0.24
0.24
0.15
0.27
0.25
0.26
Chronological Group
Actual NF Ranking
factor(chron.order4)
0
1
2
3

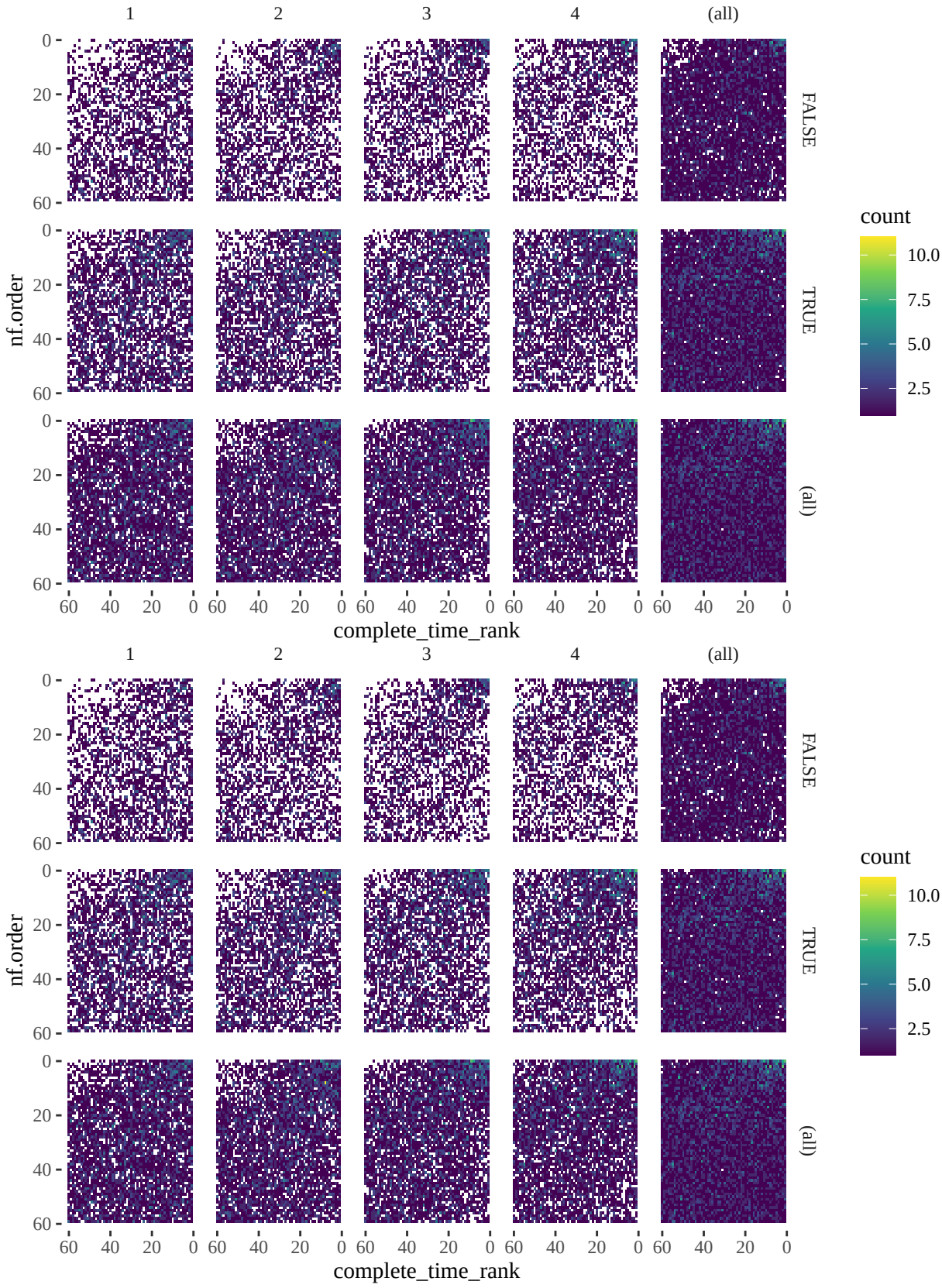

1.00
0.75
0.50
0.25
0.00
pct
0
1
2
3
0.35
0.26
0.22
0.18
0.25
0.25
0.26
0.24
0.22
0.25
0.26
0.28
0.18
0.25
0.27
0.28
Ranking Group*(−1)
Actual PYMK Ranking
factor(newquartile4)
0
1
2
3

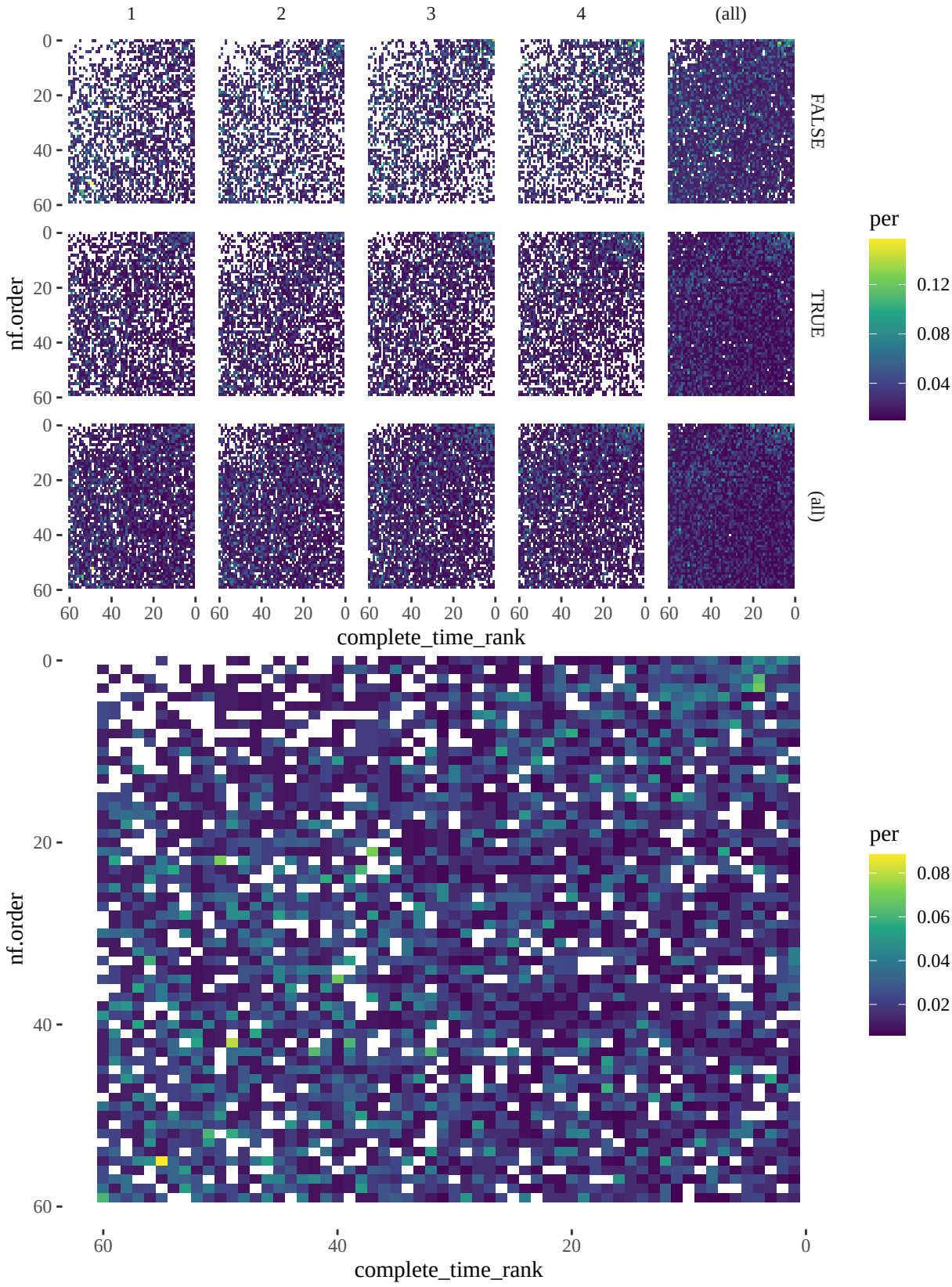

1.00
0.75
0.50
0.25
0.00
pct
0
1
2
3
0.3
0.27
0.23
0.15
0.27
0.22
0.26
0.25
0.22
0.23
0.24
0.31
0.22
0.28
0.28
0.27
Pct Friend Group
Actual NF Ranking

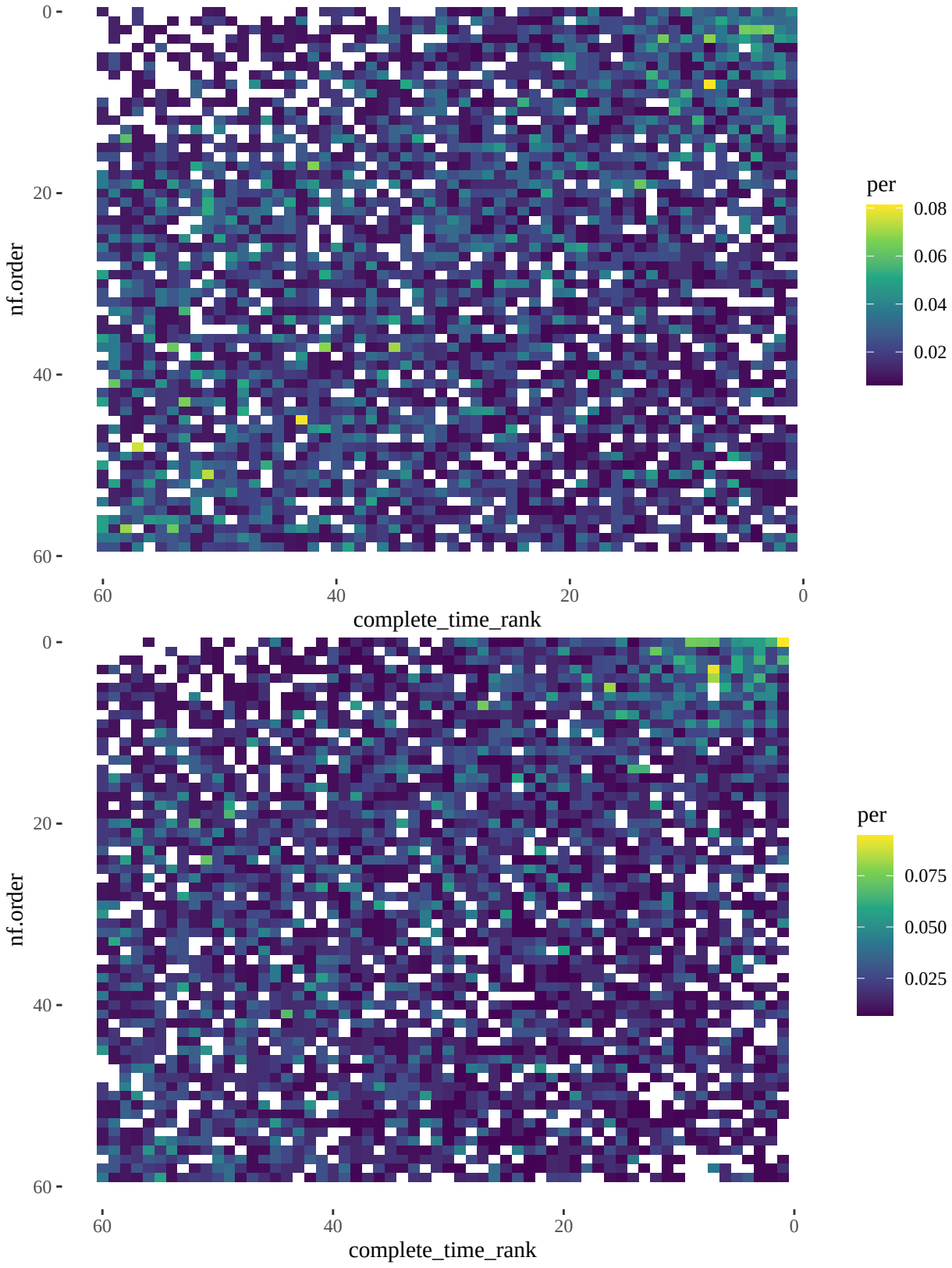

factor(friend.order4)
0
1
2
3
0
20
40
60
nf.order
60
40
20
0
complete_time_rank
per
0.06
0.04
0.02

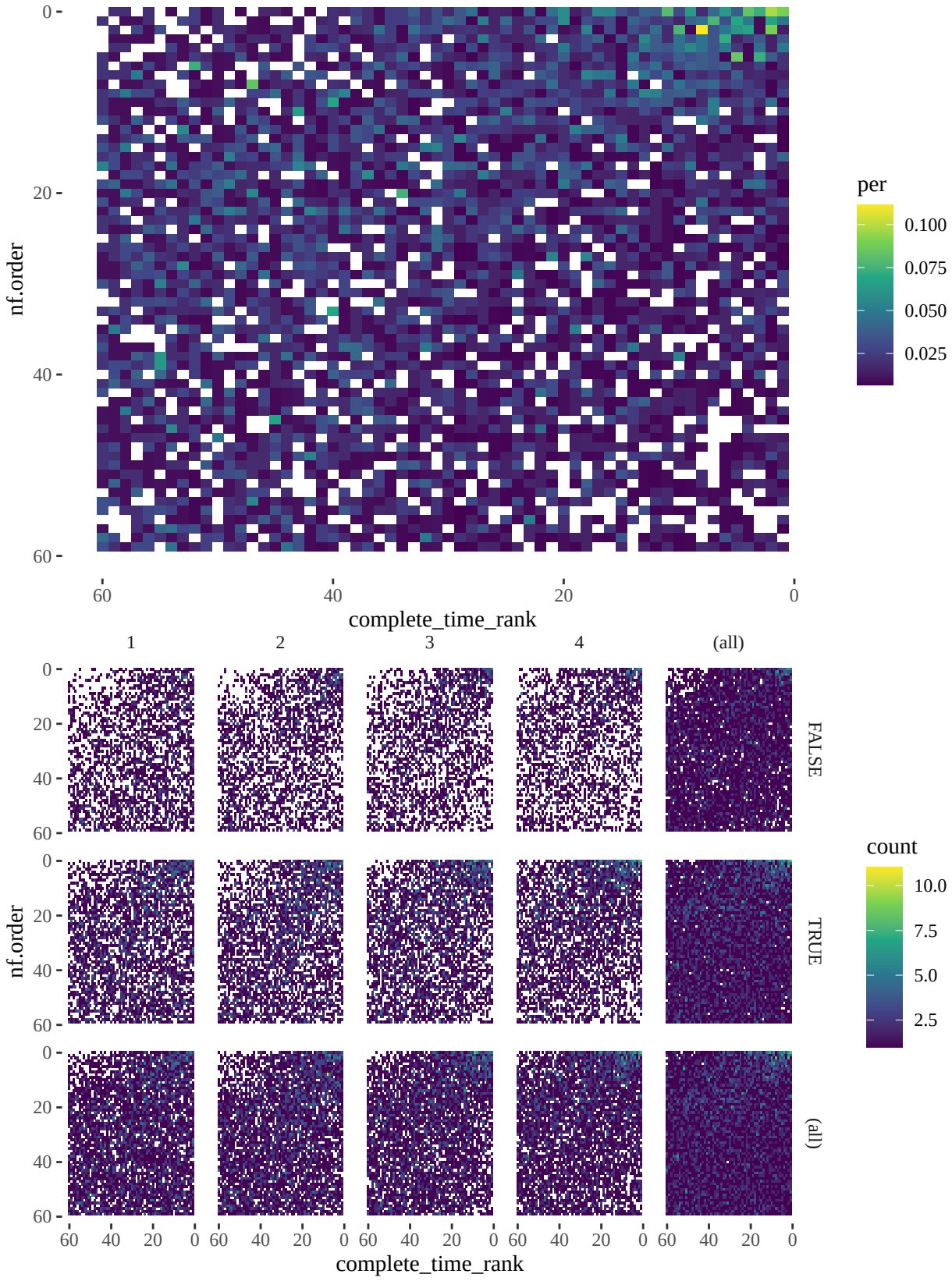

nf.order
0
20
40
60
60
40
20
0
complete_time_rank
per
0.08
0.06
0.04
0.02
nf.order
0
20
40
60
60
40
20
0
complete_time_rank
per
0.08
0.06
0.04
0.02

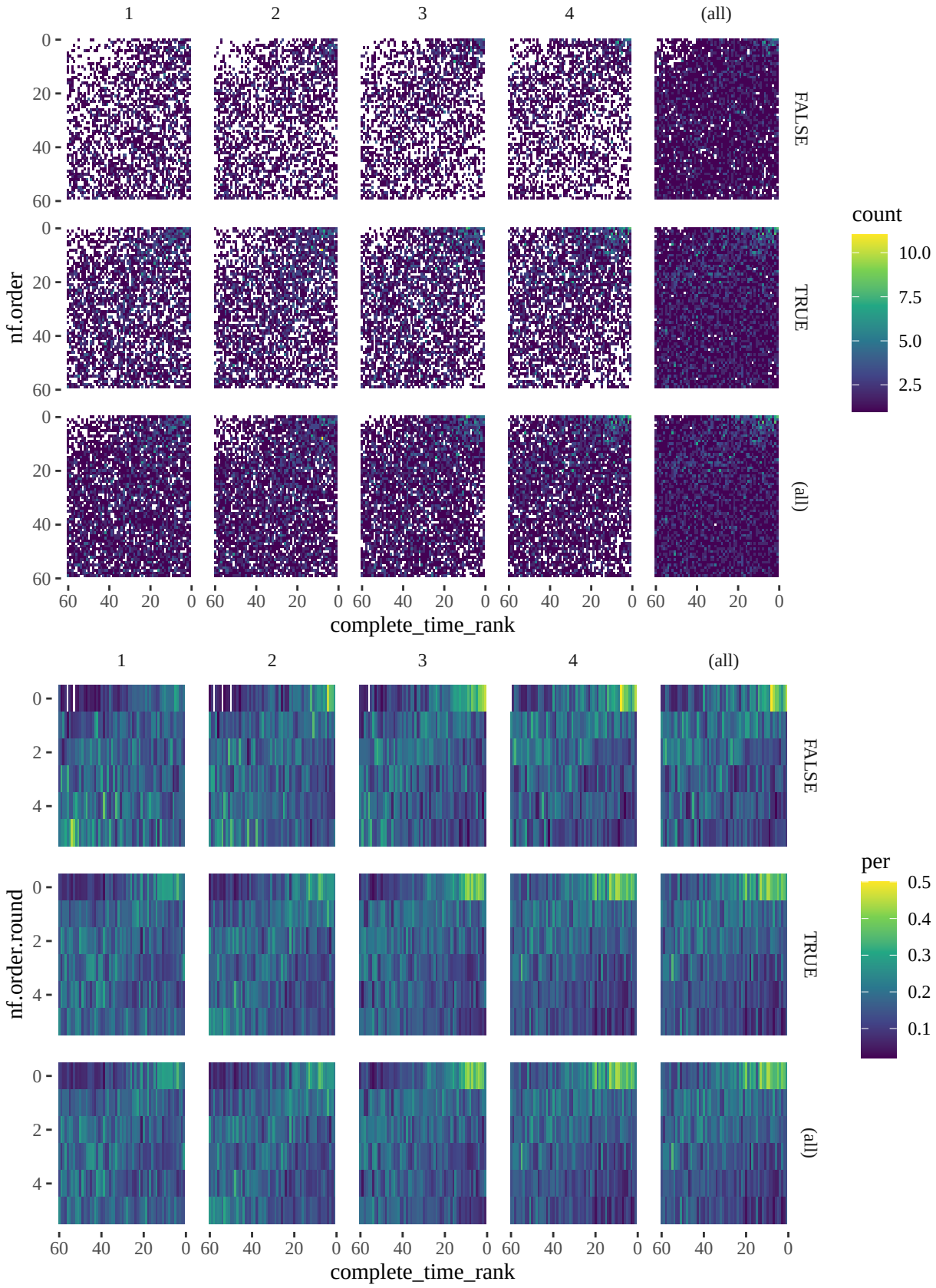

nf.order
0
20
40
60
60
40
20
0
complete_time_rank
per
0.075
0.050
0.025
nf.order
0
20
40
60
60
40
20
0
complete_time_rank
per
0.100
0.075
0.050
0.025

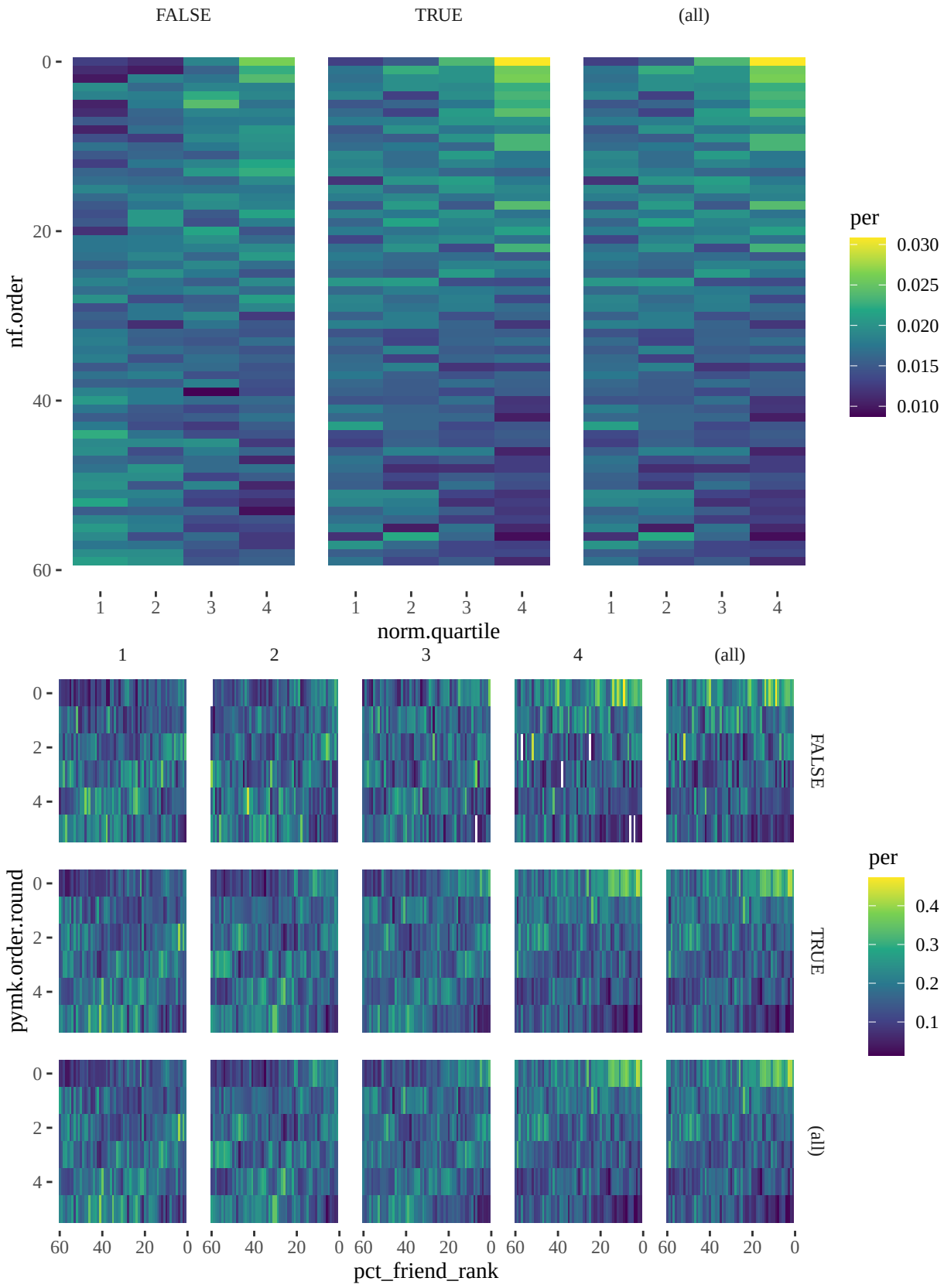

1
2
3
4
(all)
nf.order
FALSE
TRUE
(all)
0
20
40
60
complete_time_rank
count
10.0
7.5
5.0
2.5

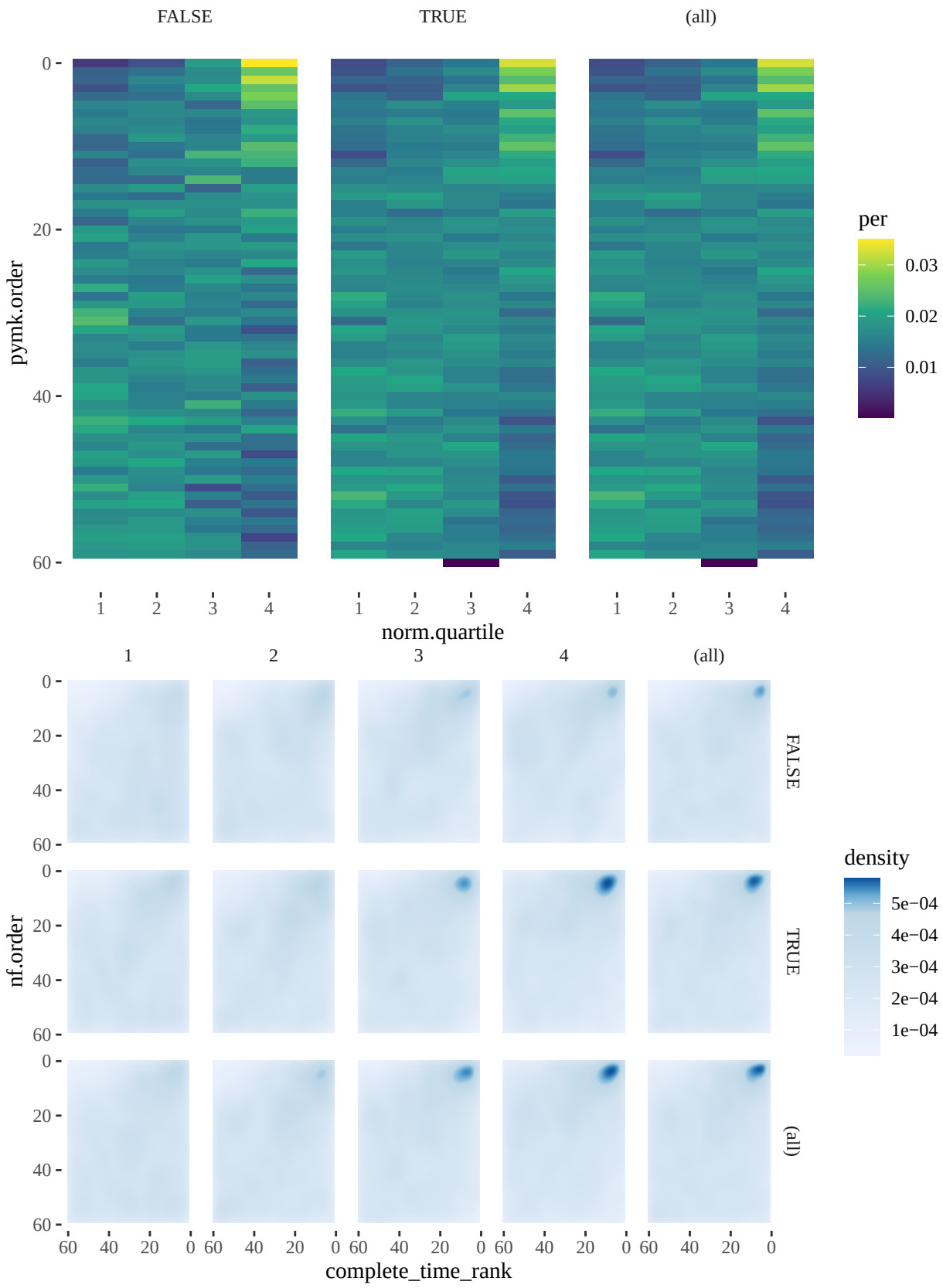

1
2
3
4
(all)
FALSE
TRUE
(all)
nf.order
0
20
40
60
complete_time_rank
per
0.12
0.08
0.04
0.06
0.02

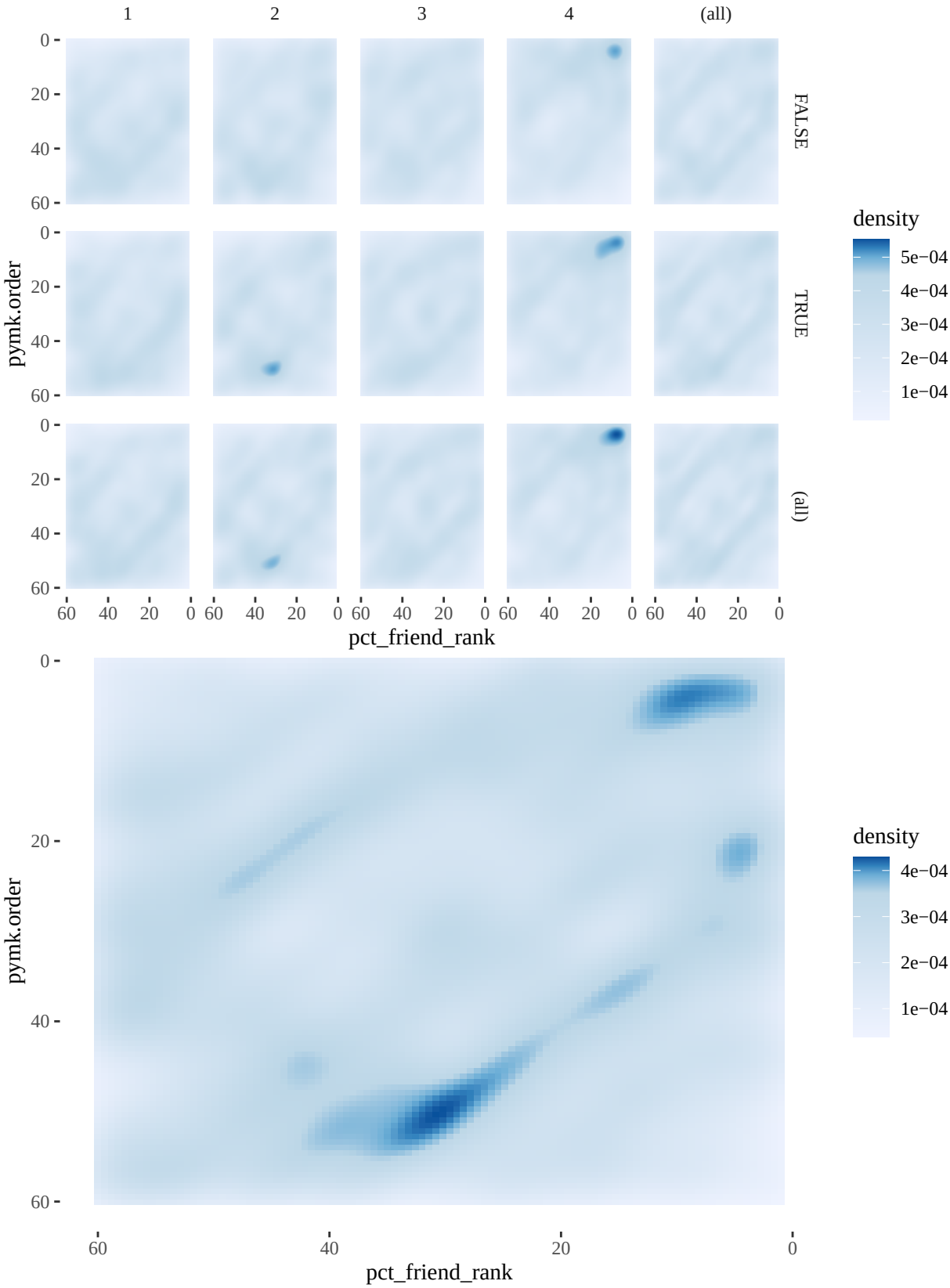

0
20
40
60
nf.order
60
40
20
0
complete_time_rank
per
0.08
0.06
0.04
0.02
0
20
40
60
nf.order
60
40
20
0
complete_time_rank
per
0.075
0.050
0.025

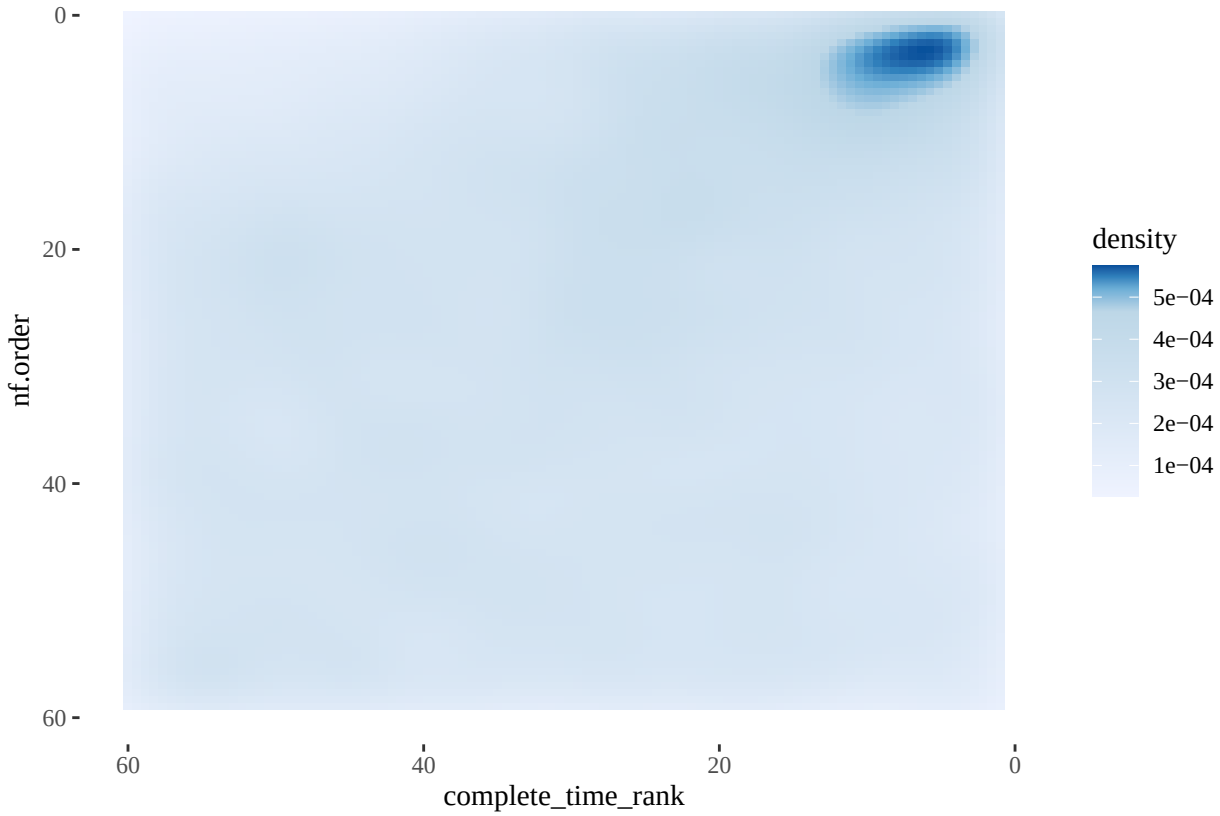

nf.order
0
20
40
60
60
40
20
0
complete_time_rank
per
0.100
0.075
0.050
0.025
1
2
3
4
(all)
FALSE
TRUE
(all)
count
10.0
7.5
5.0
2.5

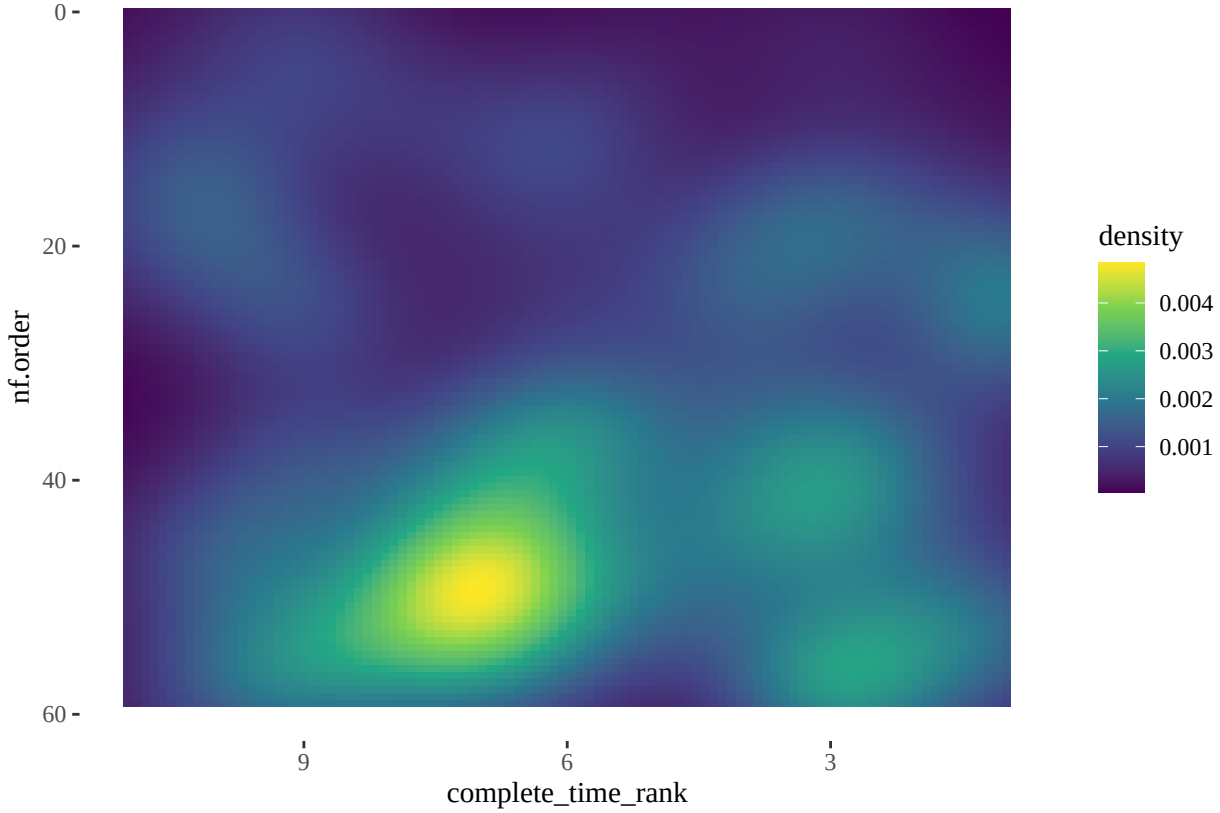

1
2
3
4
(all)
FALSE
TRUE
(all)
nf.order
0
20
40
60
count
10.0
7.5
5.0
2.5
60 40 20 0
complete_time_rank
nf.order.round
0
2
4
per
0.5
0.4
0.3
0.2
0.1

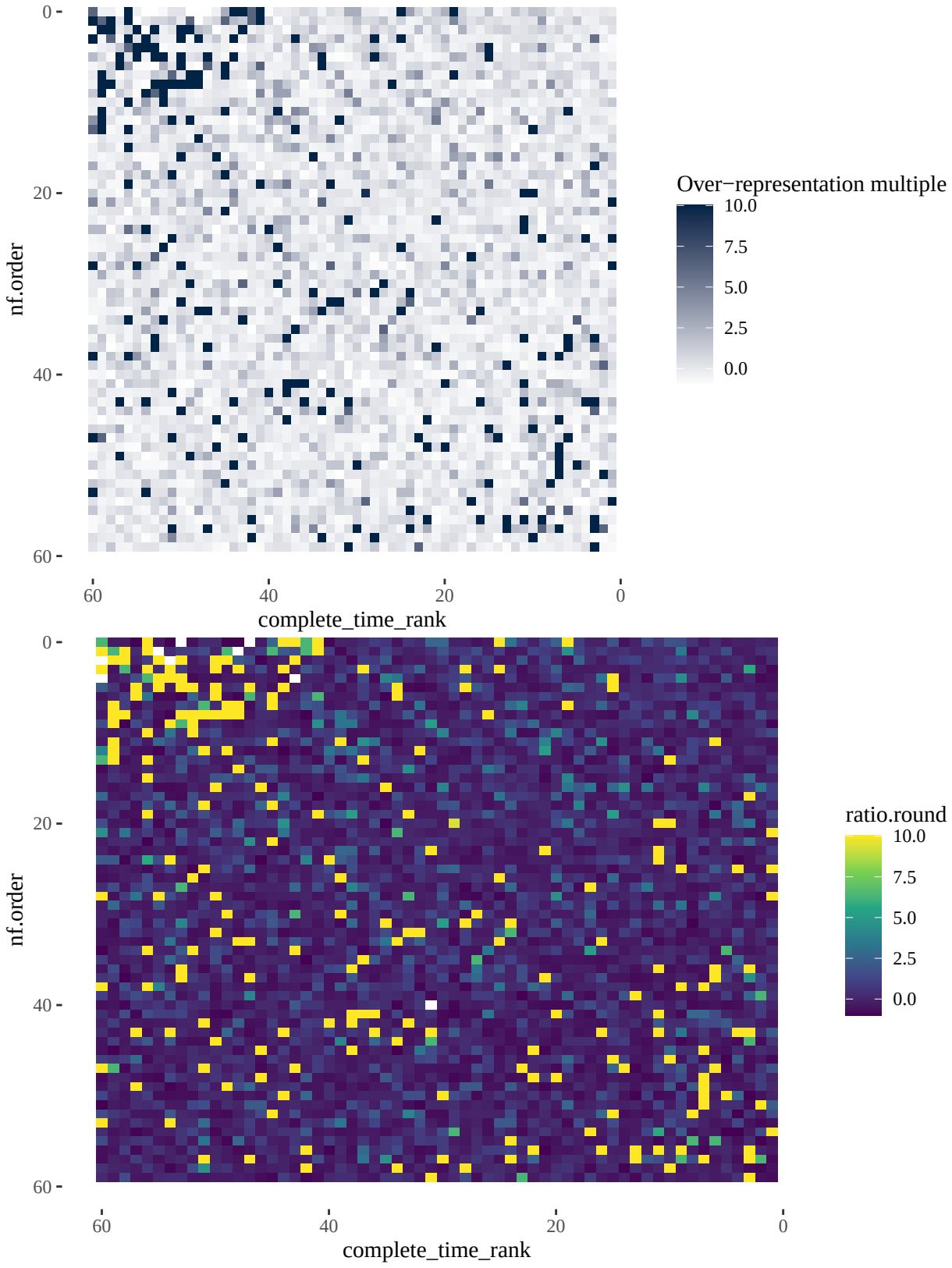

FALSE
TRUE
(all)
nf.order
0
20
40
60
1
2
3
4
norm.quartile
per
0.030
0.025
0.020
0.015
0.010
pymk.order.round
0
2
4
pct_friend_rank
60
40
20
0
0.4
0.3
0.2
0.1

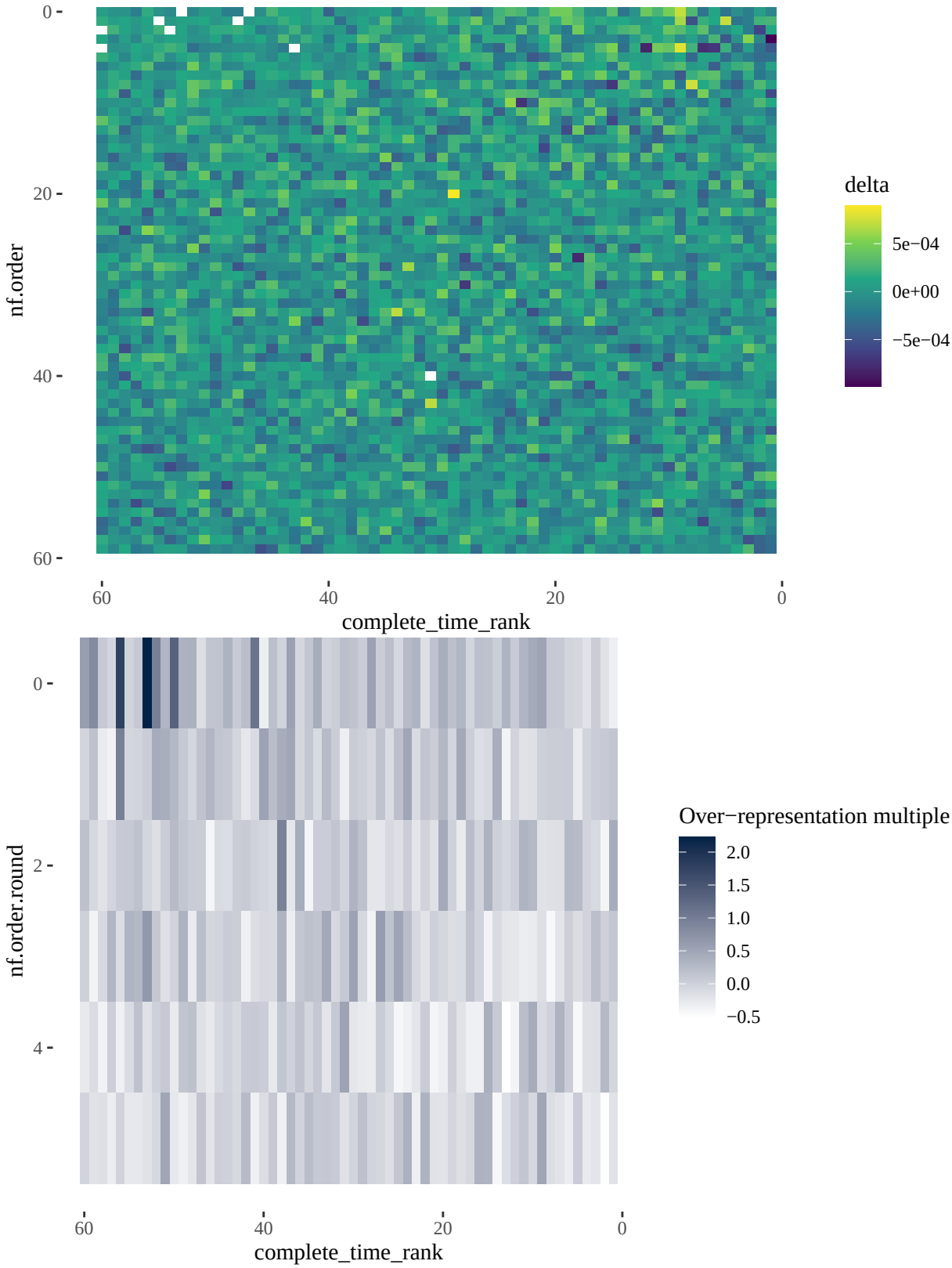

FALSE
TRUE
(all)
pymk.order
0
20
40
60
1
2
3
4
norm.quartile
per
0.03
0.02
0.01
nf.order
FALSE
TRUE
(all)
density
5e−04
4e−04
3e−04
2e−04
1e−04
complete_time_rank

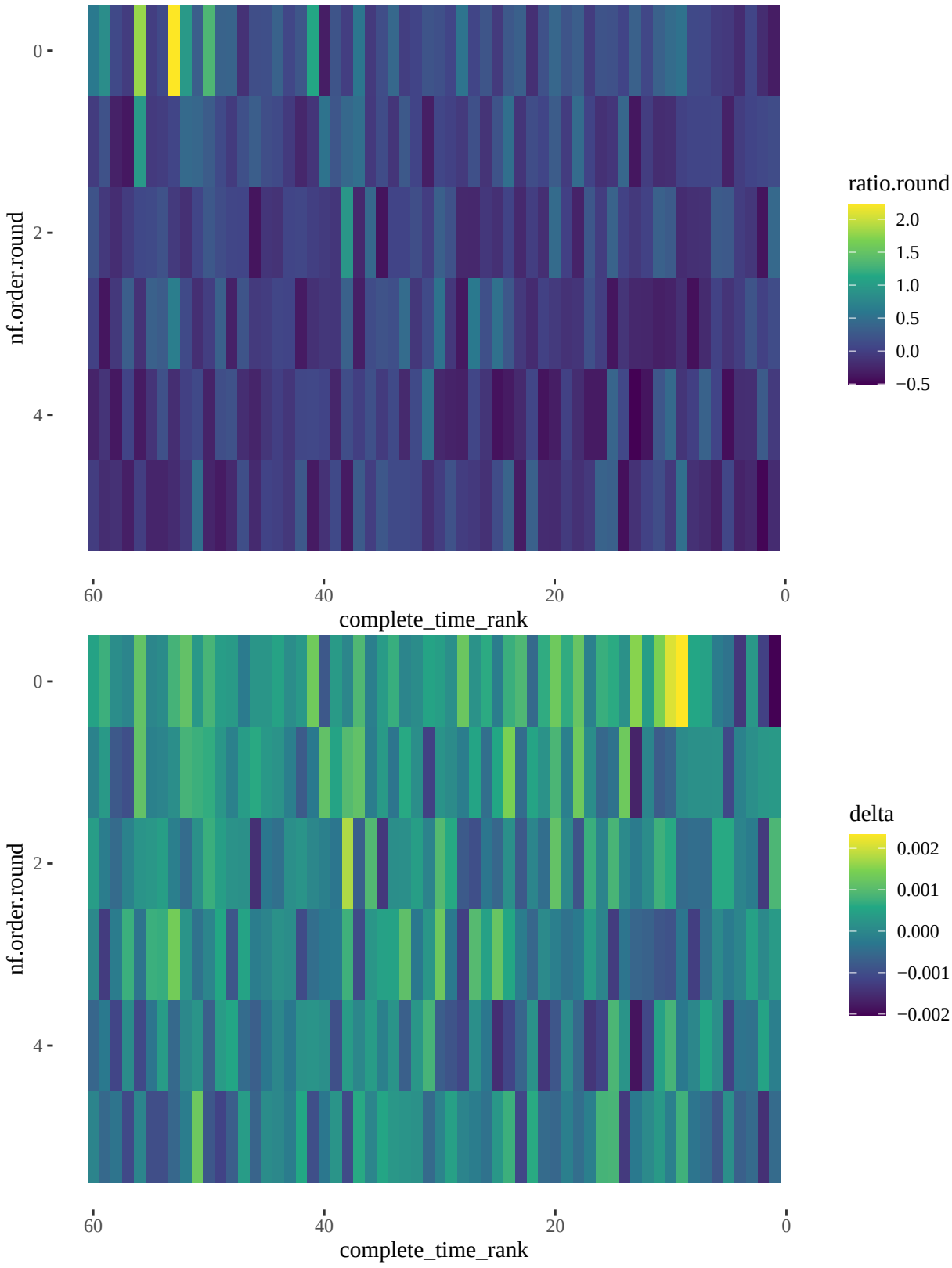

1
2
3
4
(all)
FALSE
TRUE
(all)
pymk.order
pct_friend_rank
density
5e−04
4e−04
3e−04
2e−04
1e−04
0
20
40
60

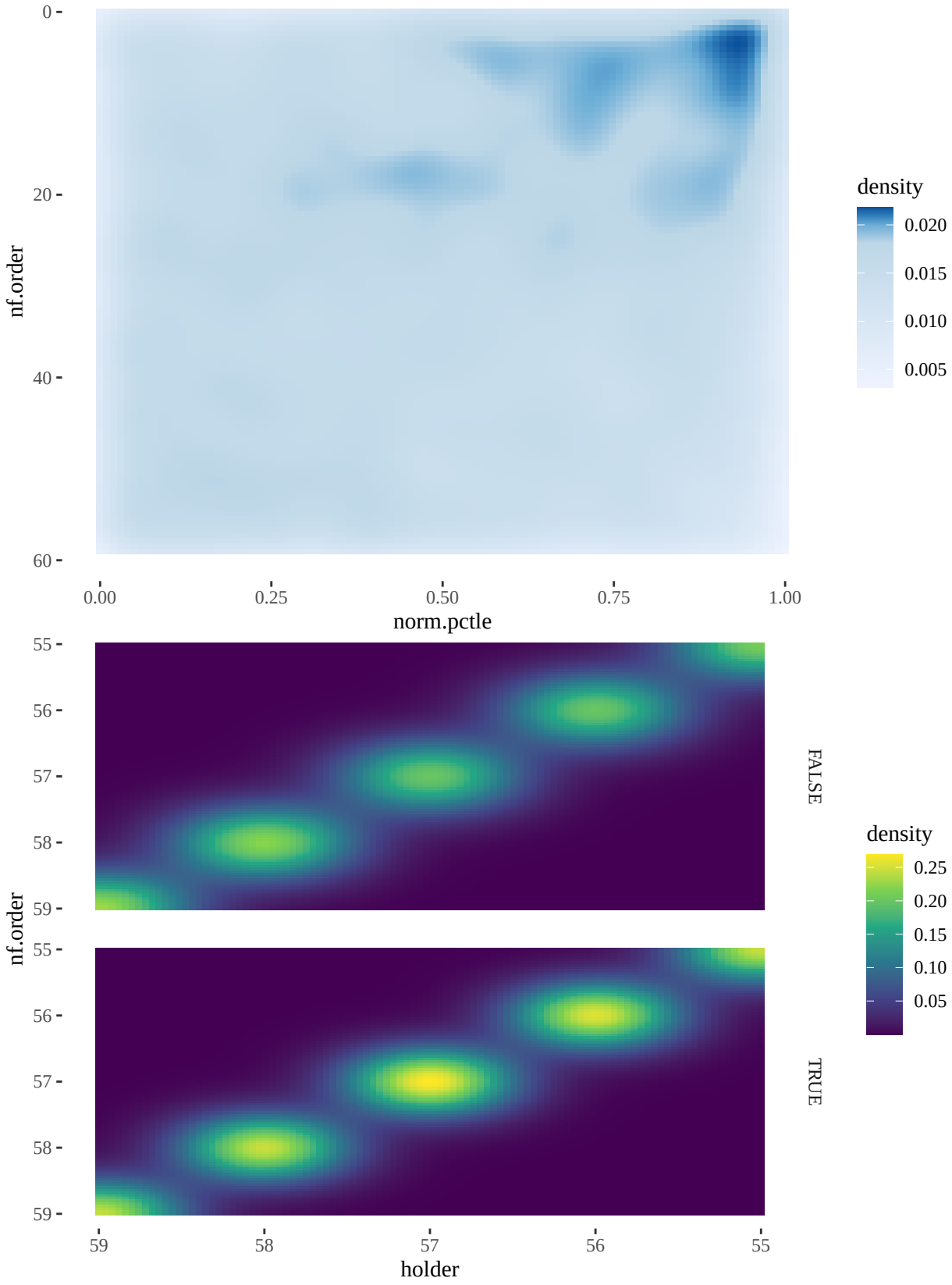

0
20
40
60
nf.order
60
40
20
0
complete_time_rank
density
5e−04
4e−04
3e−04
2e−04
1e−04

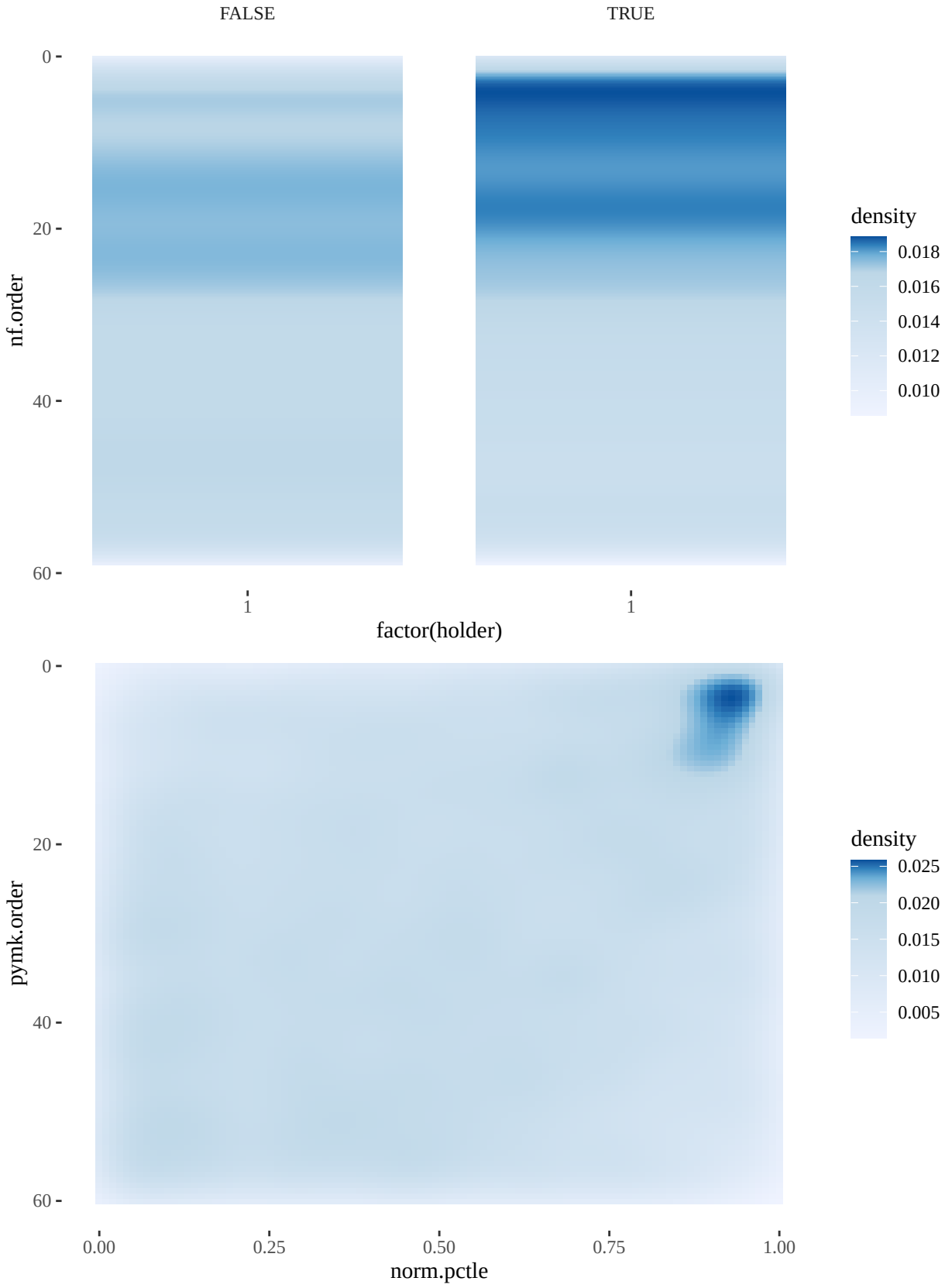

0
20
40
60
nf.order
9
6
3
complete_time_rank
density
0.004
0.003
0.002
0.001

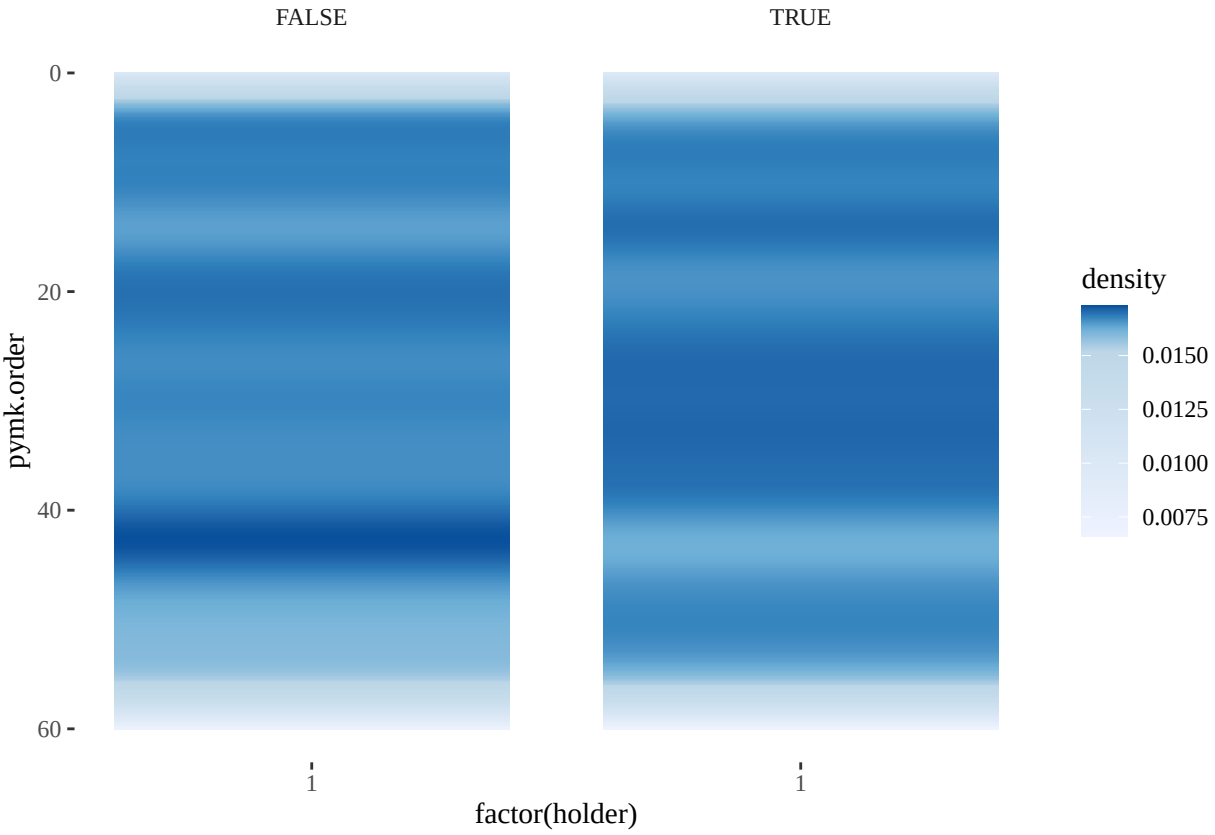

0
20
40
60
nf.order
60
40
20
0
complete_time_rank
Over−representation multiple
10.0
7.5
5.0
2.5
0.0
ratio.round
10.0
7.5
5.0
2.5
0.0


nf.order
0
20
40
60
60
40
20
0
complete_time_rank
delta
5e−04
0e+00
−5e−04
nf.order.round
0
2
4
60
40
20
0
complete_time_rank
Over−representation multiple
2.0
1.5
1.0
0.5
0.0
−0.5

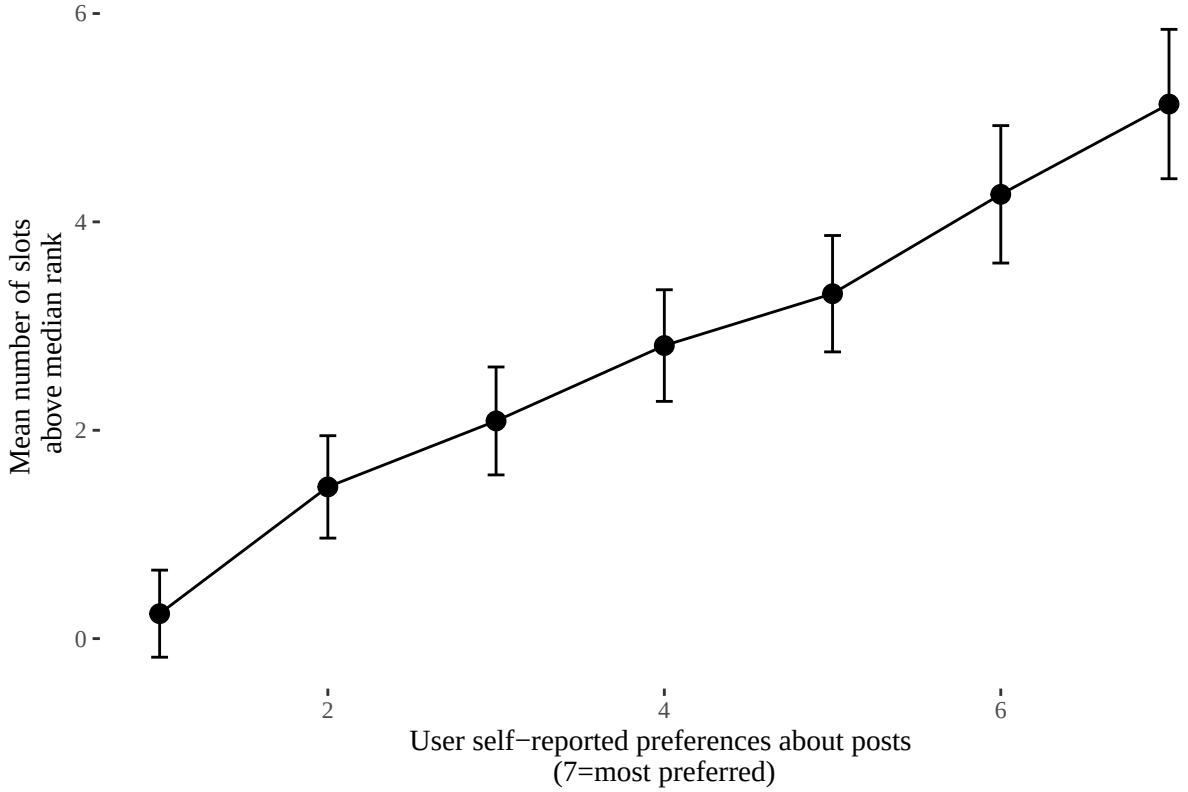

nf.order.round
0
2
4
60
40
20
0
complete_time_rank
ratio.round
2.0
1.5
1.0
0.5
0.0
−0.5
nf.order.round
0
2
4
60
40
20
0
complete_time_rank
delta
0.002
0.001
0.000
−0.001
−0.002


nf.order
0
20
40
60
0.00
0.25
0.50
0.75
1.00
norm.pctle
density
0.020
0.015
0.010
0.005
nf.order
55
56
57
58
59
FALSE
TRUE
density
0.25
0.20
0.15
0.10
0.05
holder

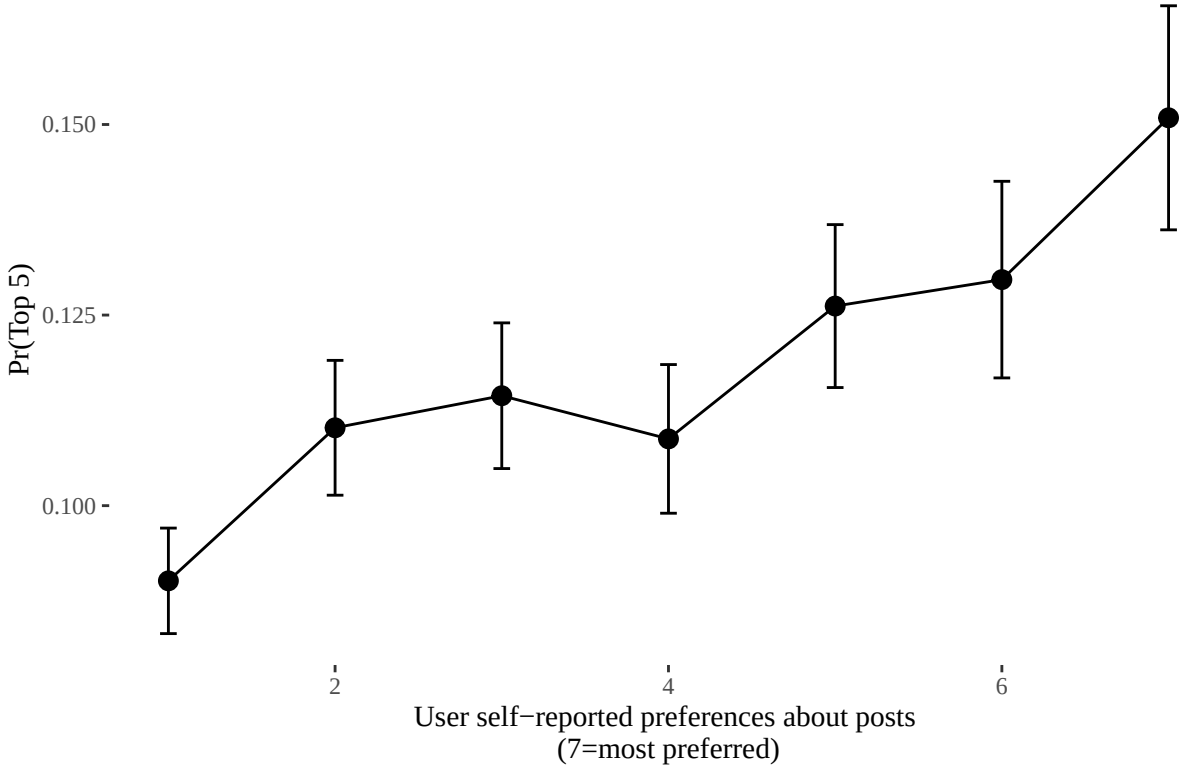

FALSE
TRUE
0
20
40
60
nf.order
1
1
factor(holder)
density
0.018
0.016
0.014
0.012
0.010
0
20
40
60
pymk.order
0.00
0.25
0.50
0.75
1.00
norm.pctle
density
0.025
0.020
0.015
0.010
0.005



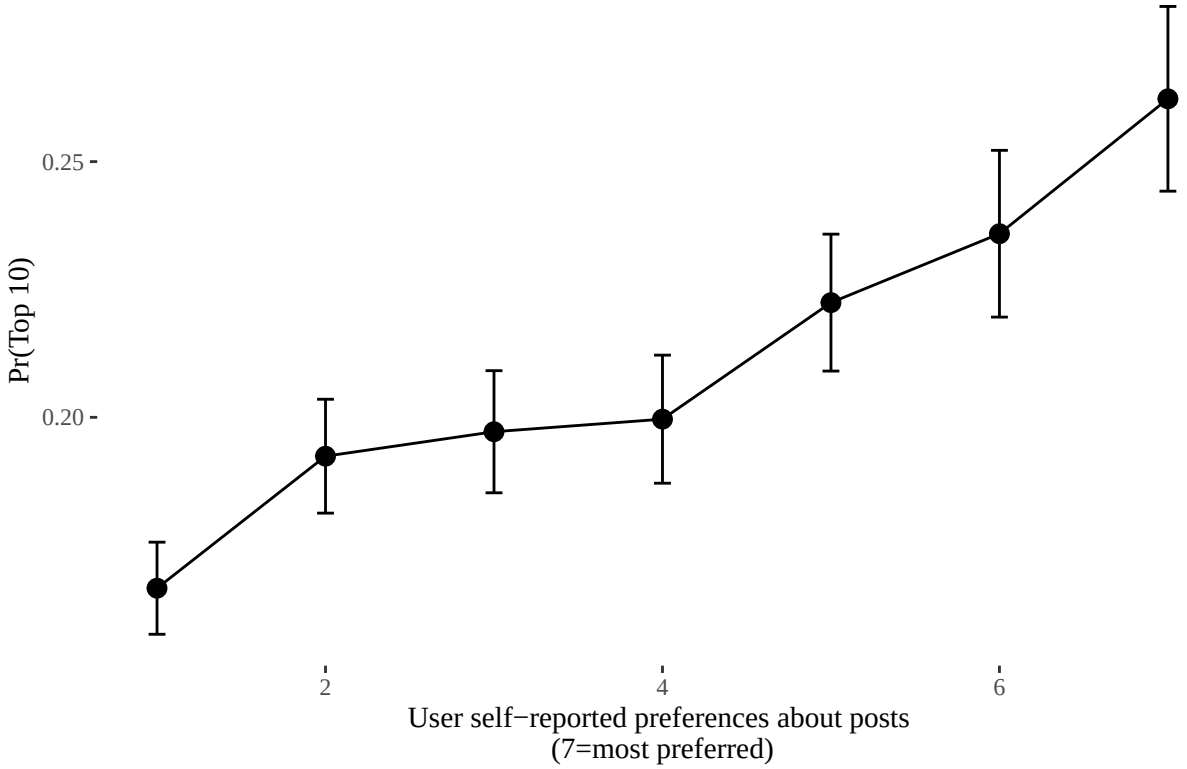
Newsfeed Ranking by Preference (US)

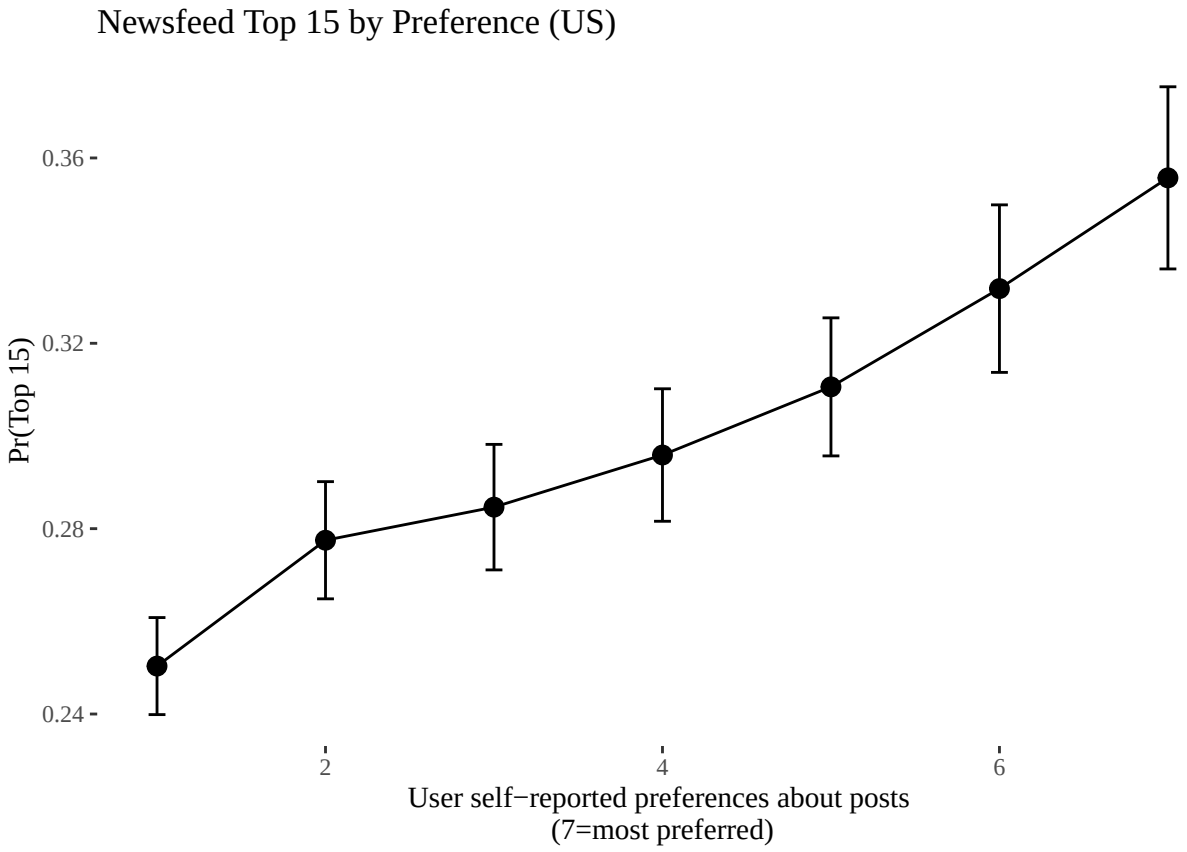


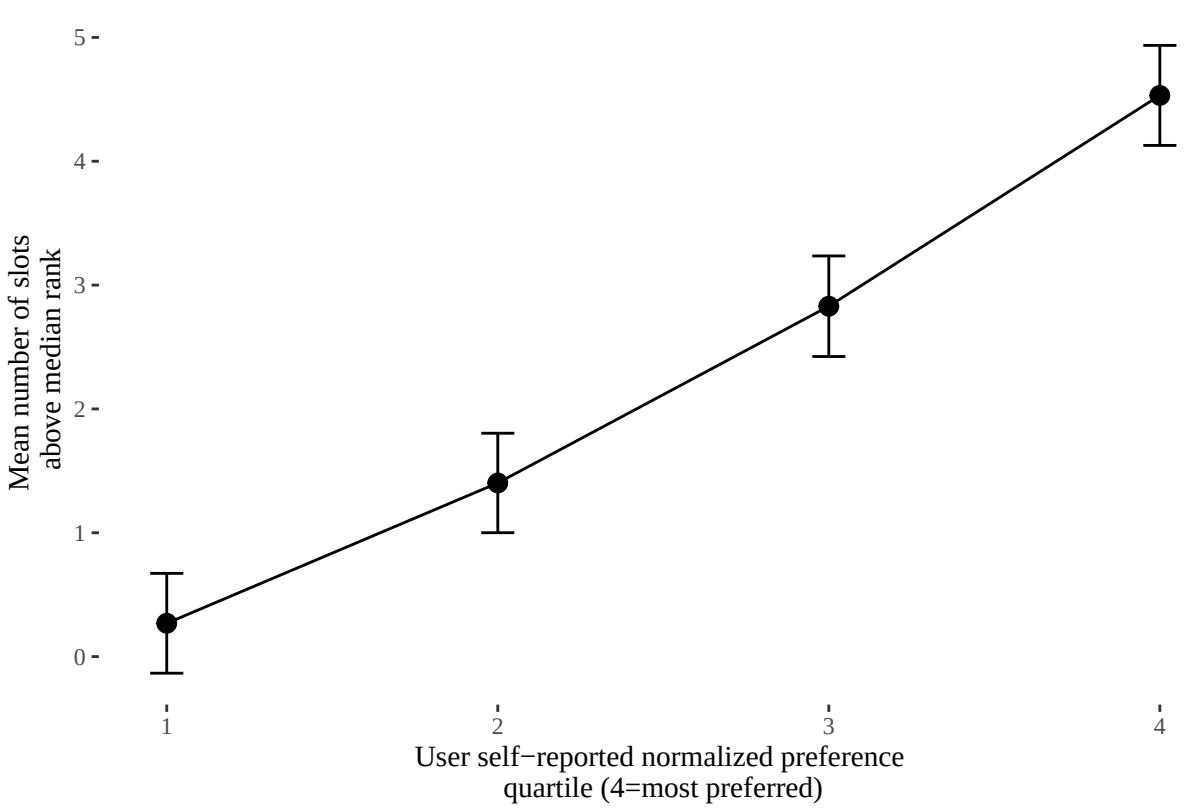
Newsfeed Top 5 by Preference (US)

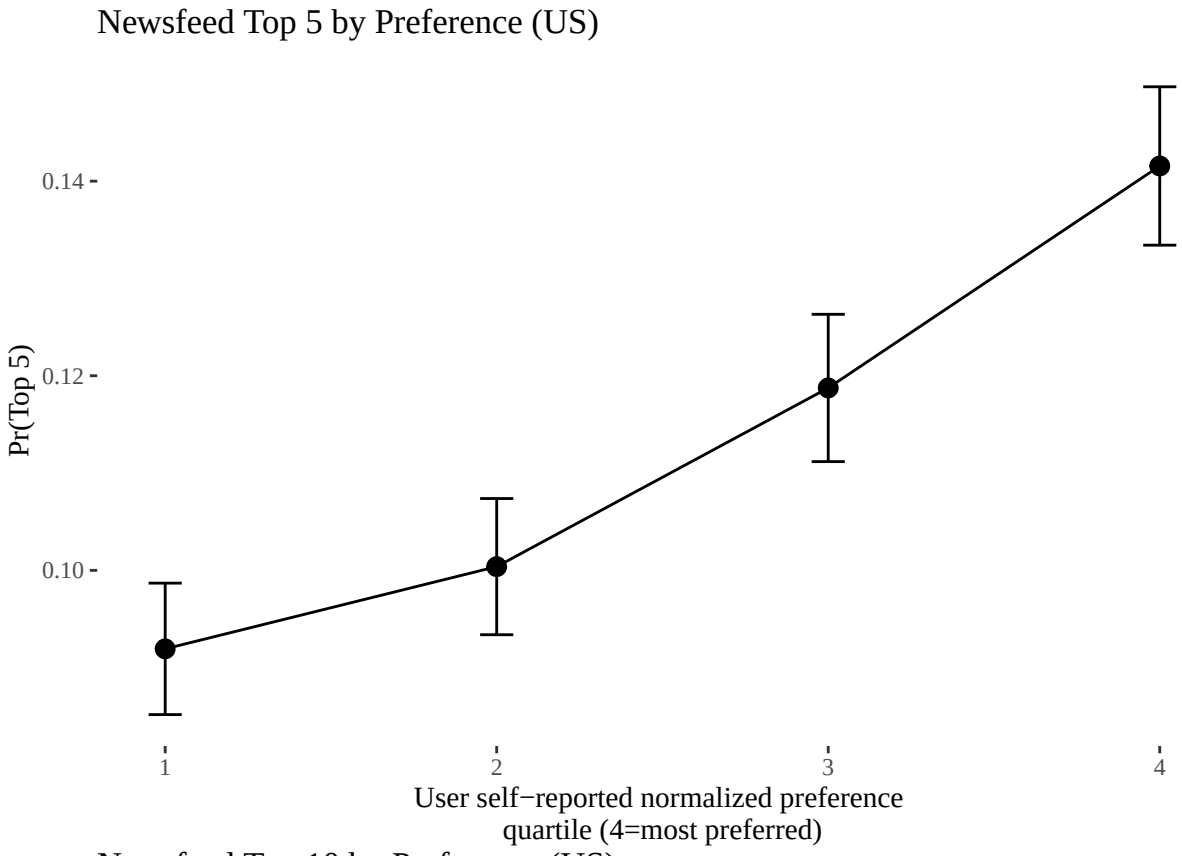


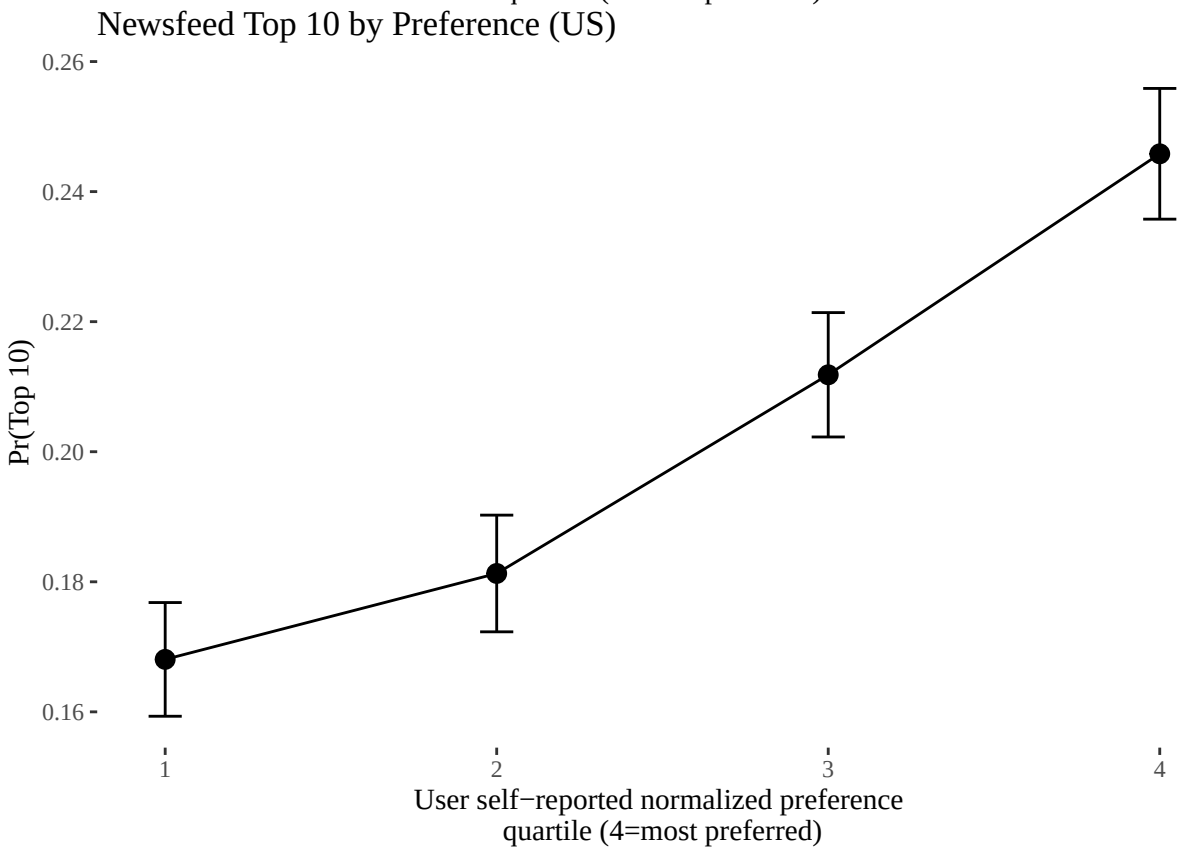
Newsfeed Top 10 by Preference (US)

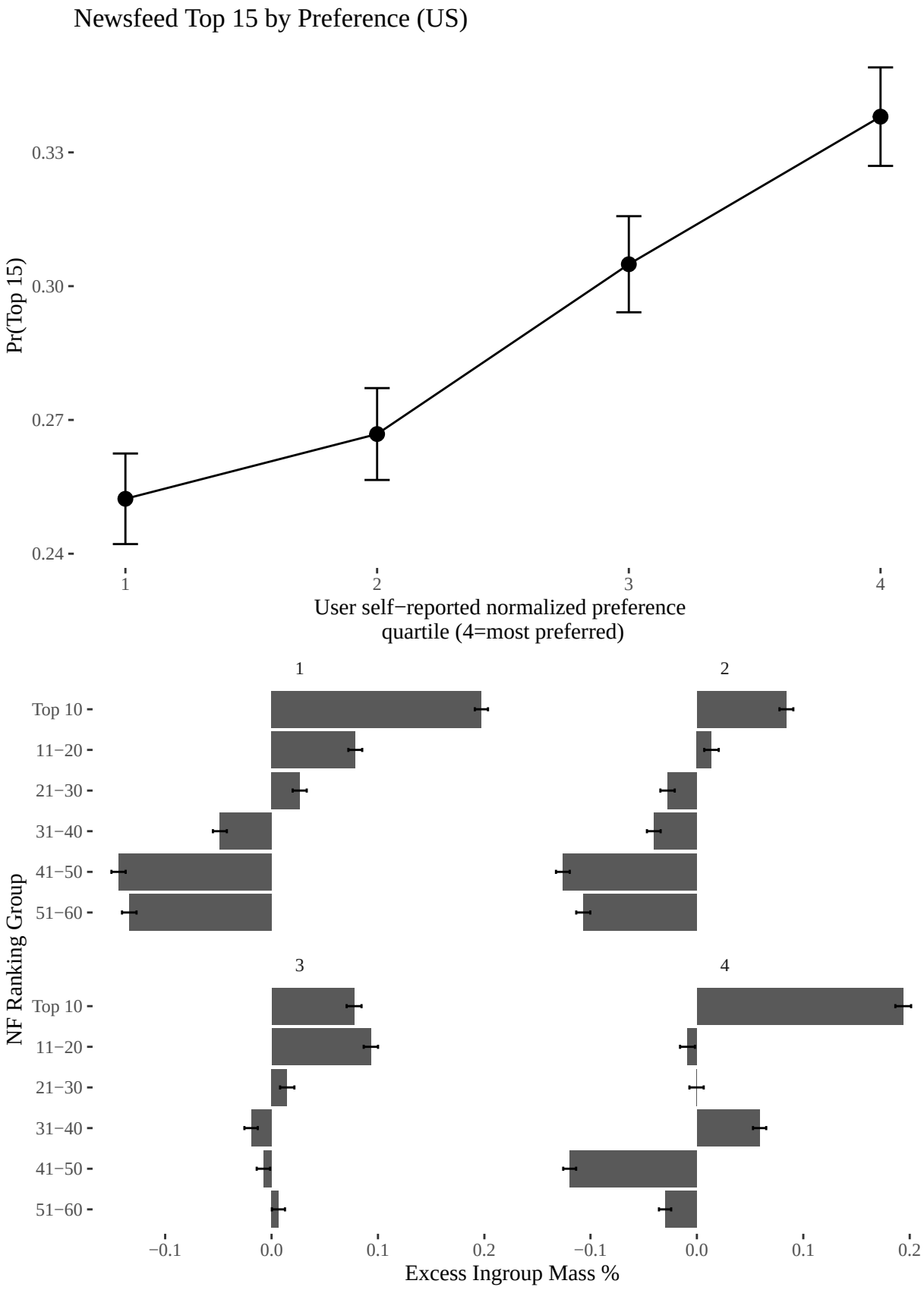


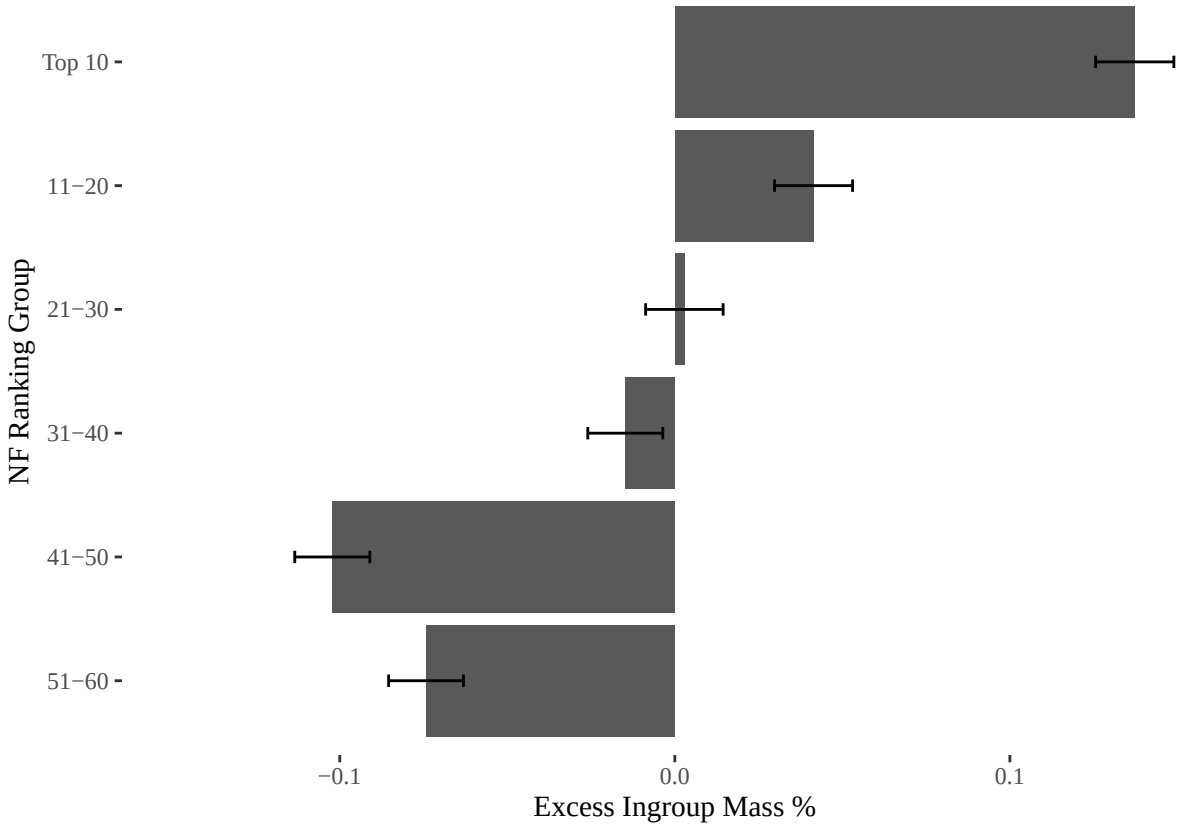

Newsfeed Top 15 by Preference (US)
0.36
0.32
0.28
0.24
Pr(Top 15)
2
4
6
User self−reported preferences about posts
(7=most preferred)

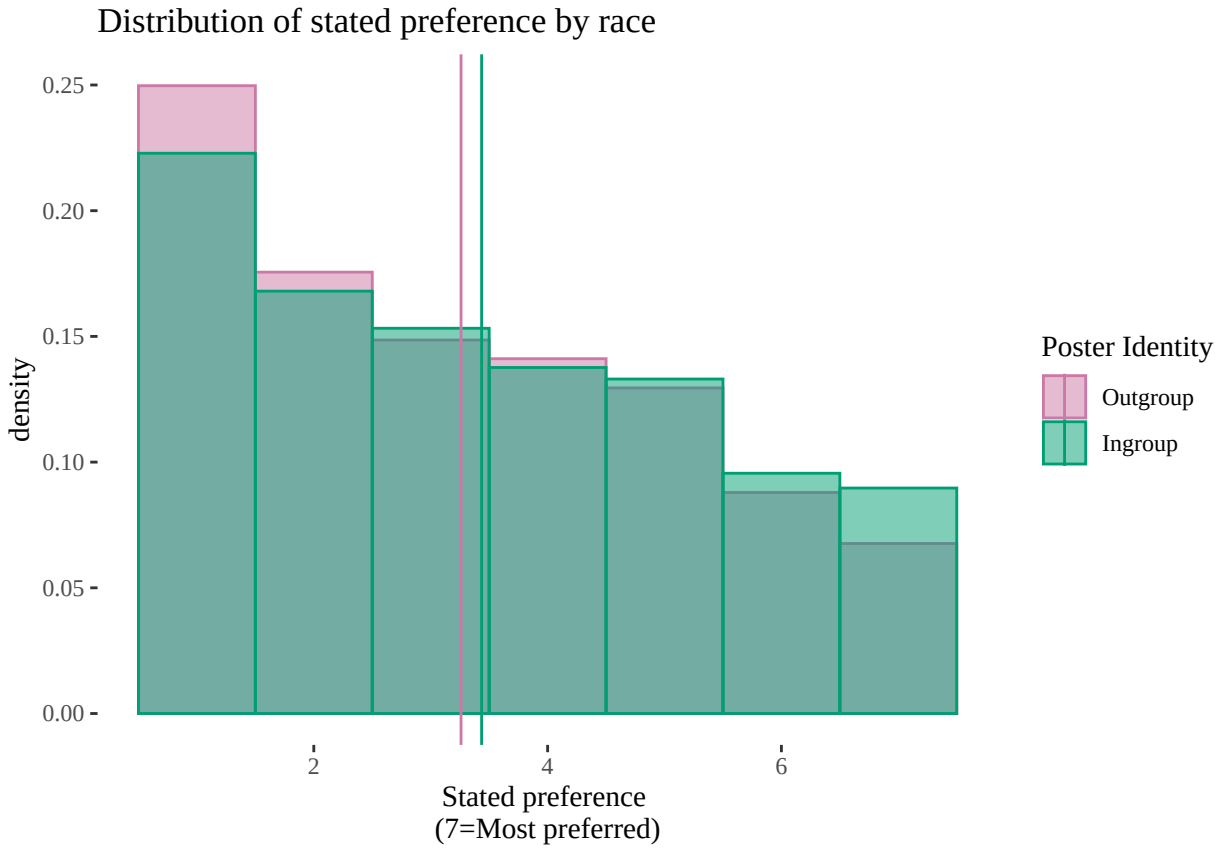

Newsfeed Ranking by Preference (US)
5
4
3
2
1
0
Mean number of slots
above median rank
1
2
3
4
User self−reported normalized preference
quartile (4=most preferred)

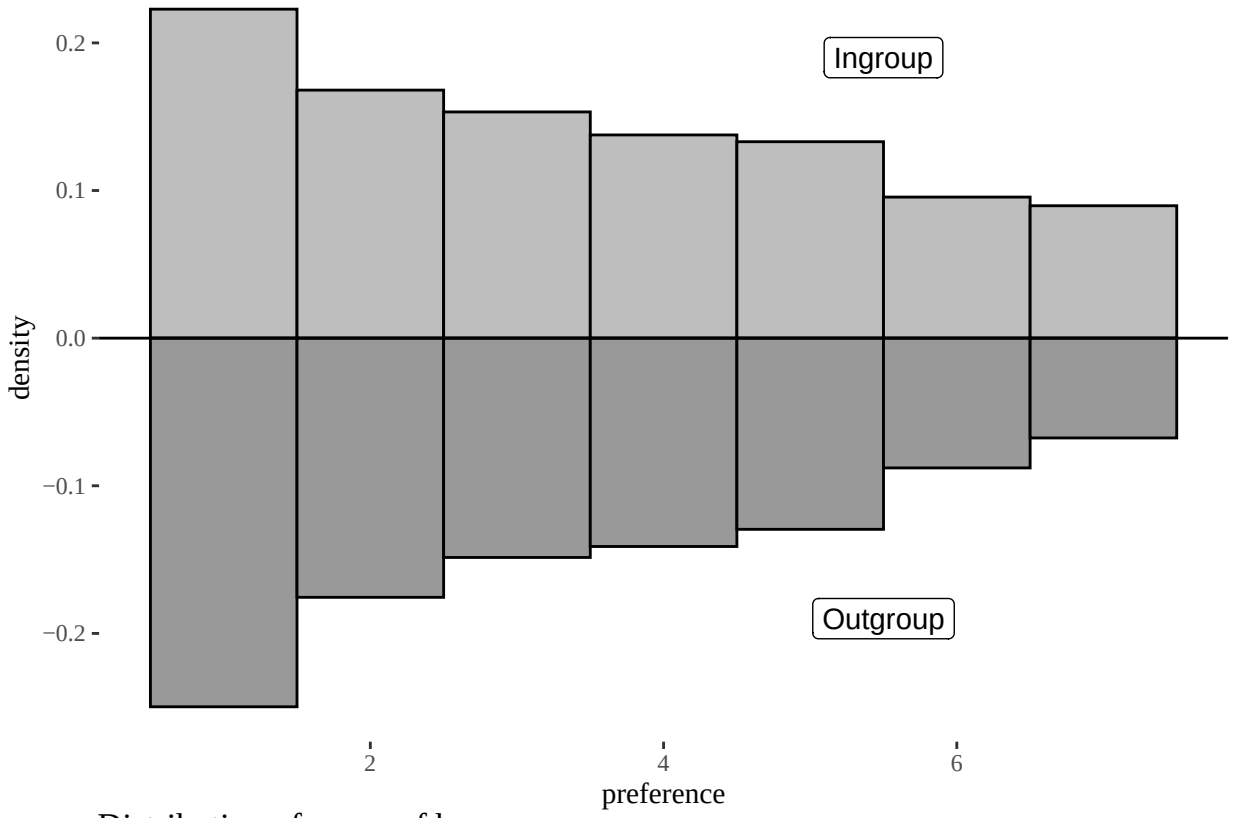

Newsfeed Top 5 by Preference (US)
0.14
0.12
0.10
Pr(Top 5)
1
2
3
4
User self−reported normalized preference
quartile (4=most preferred)


Newsfeed Top 10 by Preference (US)
0.26
0.24
0.22
0.20
0.18
0.16
Pr(Top 10)
1
2
3
4
User self−reported normalized preference
quartile (4=most preferred)

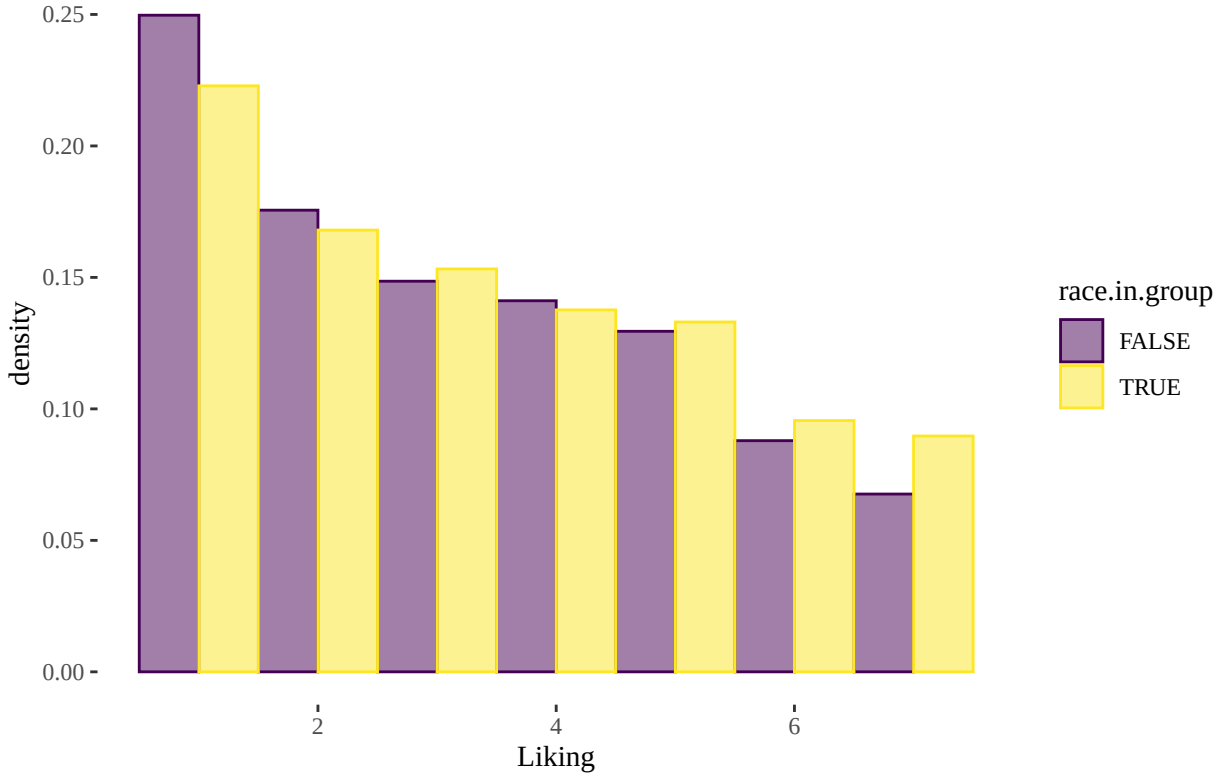

Newsfeed Top 15 by Preference (US)
Pr(Top 15)
0.33
0.30
0.27
0.24
1
2
3
4
User self−reported normalized preference
quartile (4=most preferred)
1
2
3
4
NF Ranking Group
Top 10
11−20
21−30
31−40
41−50
51−60
−0.1
0.0
0.1
0.2
Excess Ingroup Mass %

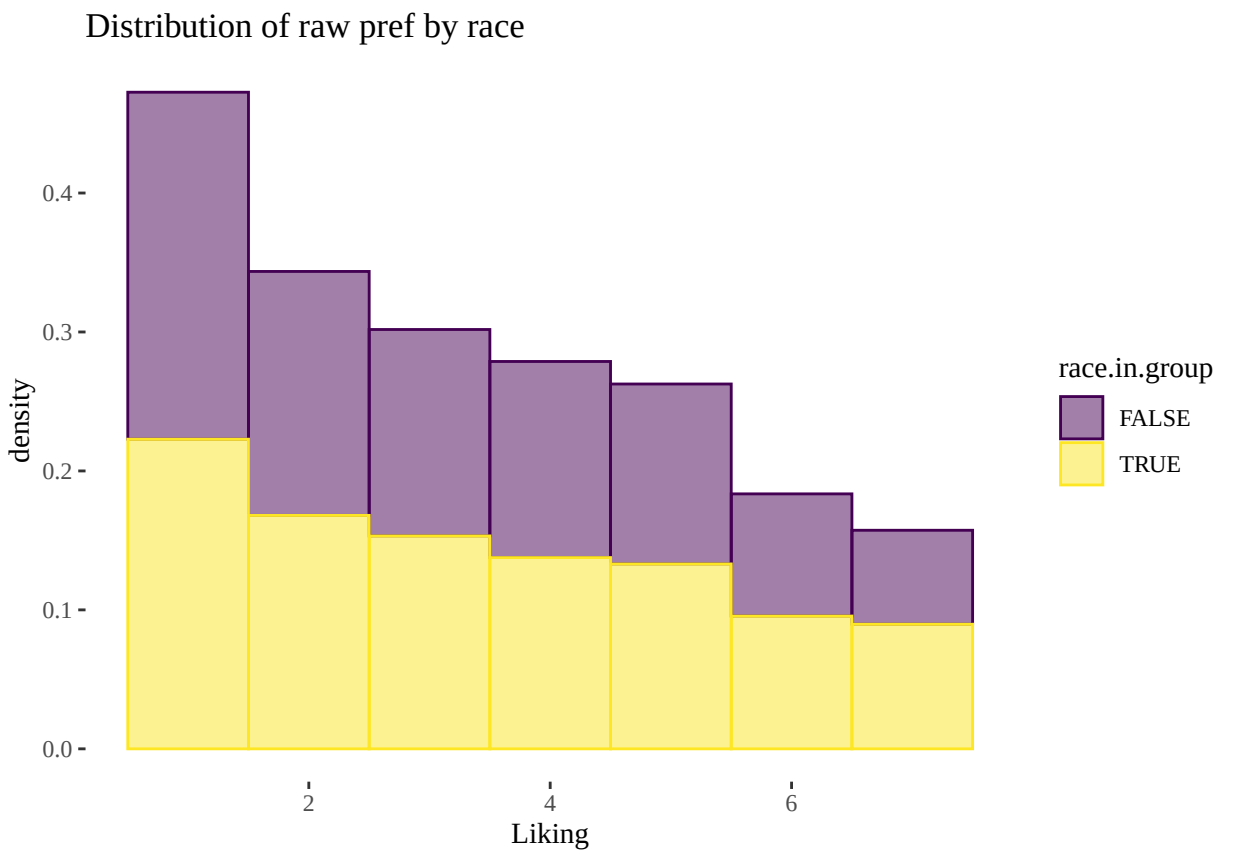

Top 10
11−20
21−30
31−40
41−50
51−60
NF Ranking Group
−0.1
0.0
0.1
Excess Ingroup Mass %

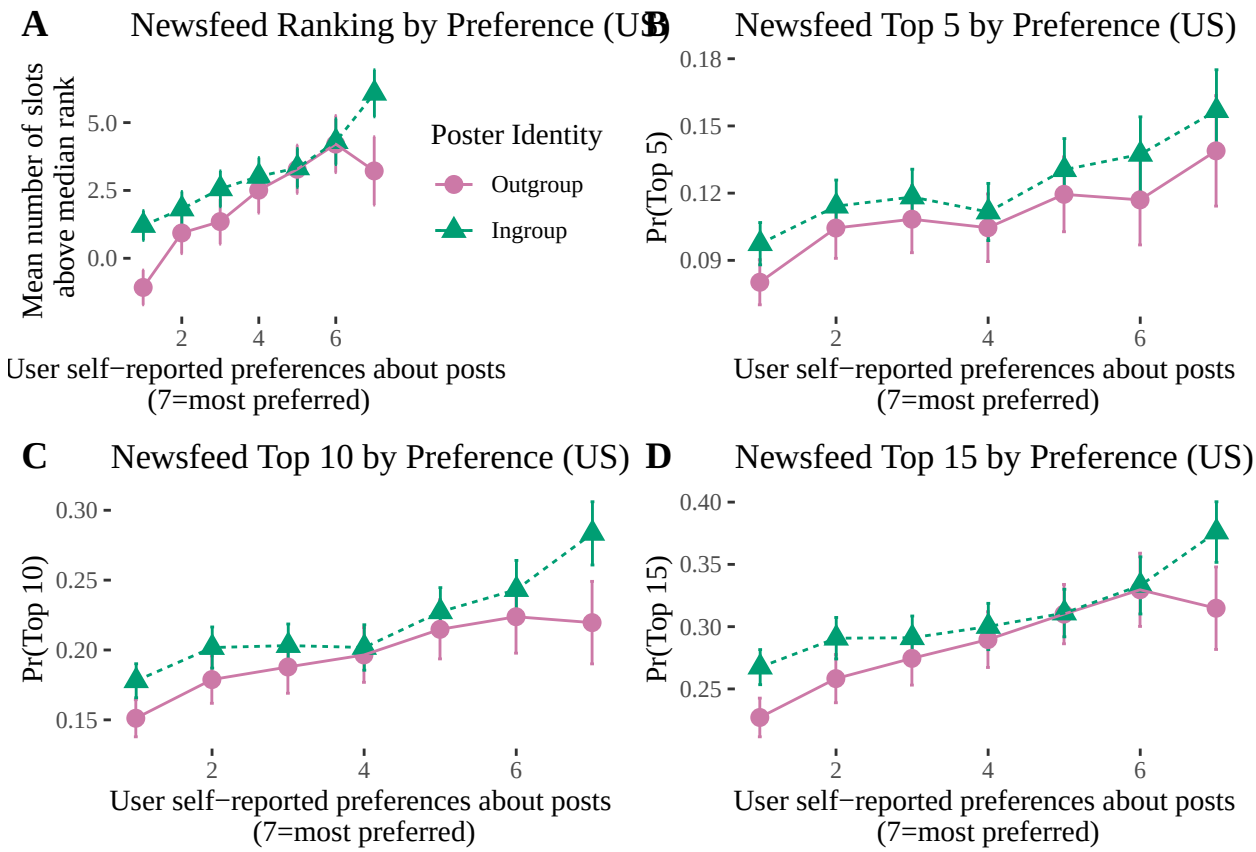

Distribution of stated preference by race
0.25
0.20
0.15
0.10
0.05
0.00
density
2
4
6
Stated preference
(7=Most preferred)
Poster Identity
Outgroup
Ingroup

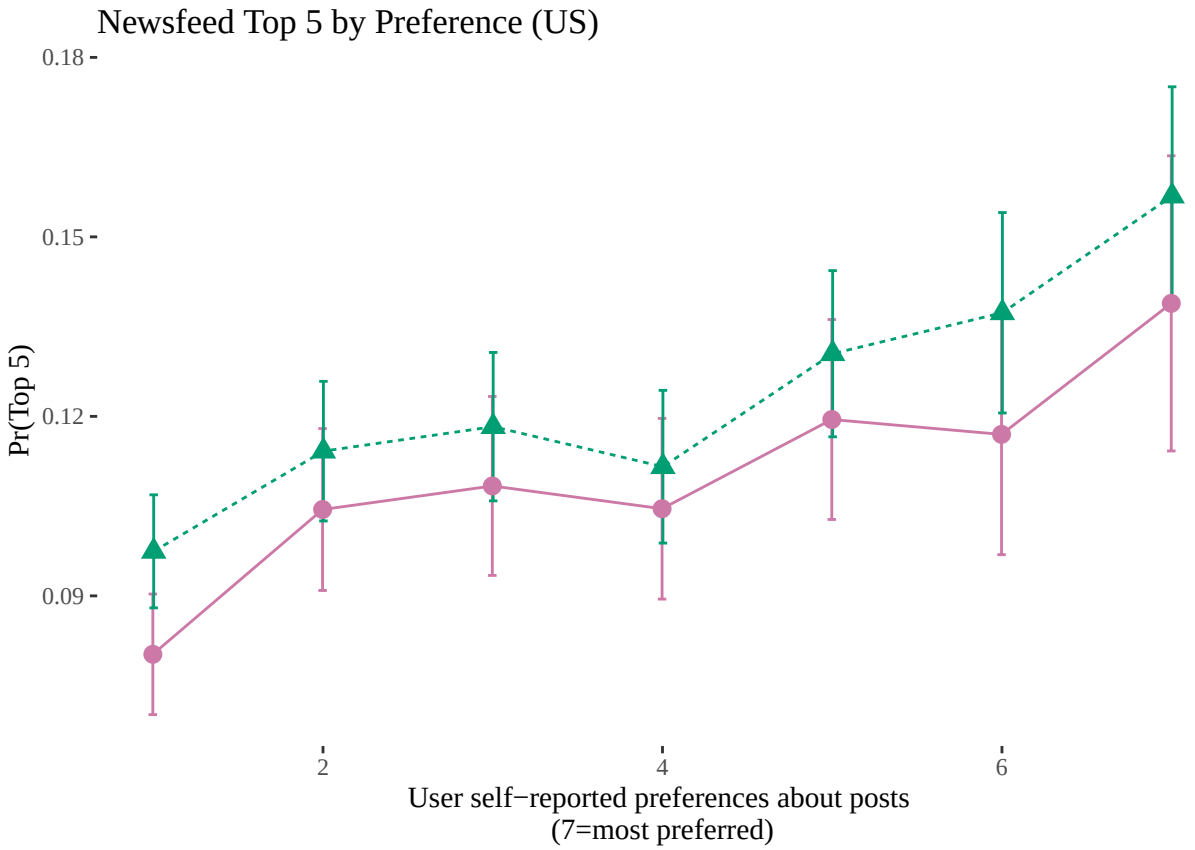


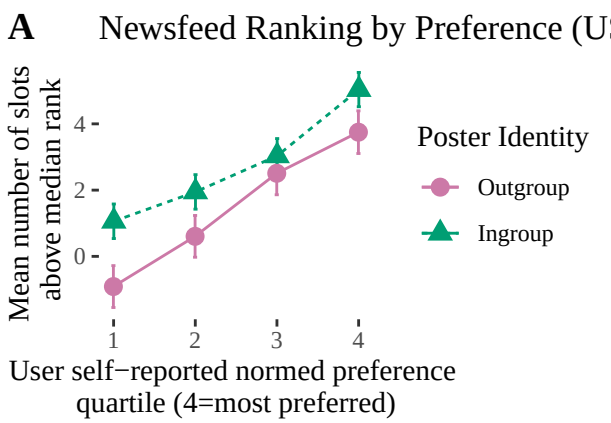
Distribution of raw pref by race

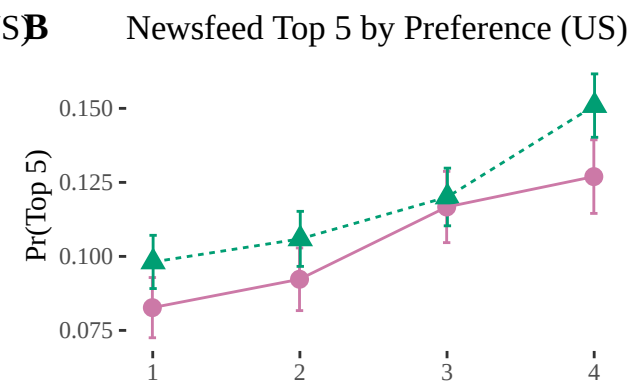


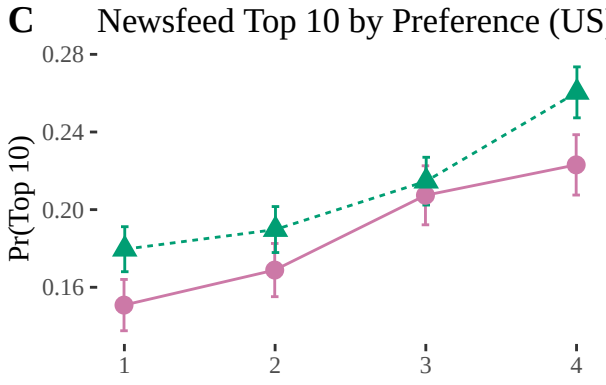

Distribution of raw pref by race
density
0.0
0.1
0.2
0.3
0.4
2
4
6
Liking
race.in.group
FALSE
TRUE

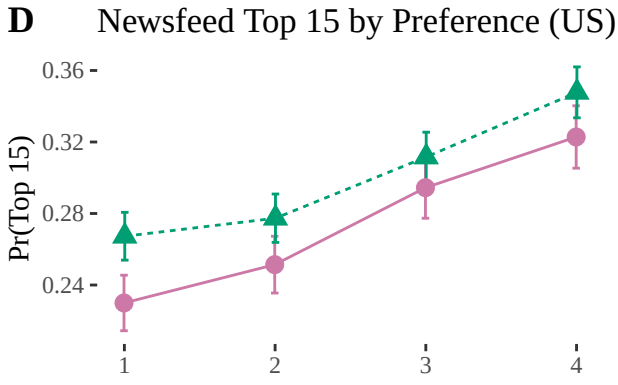

A
Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
0.0
2.5
5.0
2
4
6
User self−reported preferences about posts (7=most preferred)
Poster Identity
Outgroup
Ingroup
B
Newsfeed Top 5 by Preference (US)
Pr(Top 5)
0.09
0.12
0.15
0.18
2
4
6
User self−reported preferences about posts (7=most preferred)
C
Newsfeed Top 10 by Preference (US)
Pr(Top 10)
0.15
0.20
0.25
0.30
2
4
6
User self−reported preferences about posts (7=most preferred)
D
Newsfeed Top 15 by Preference (US)
Pr(Top 15)
0.25
0.30
0.35
0.40
2
4
6
User self−reported preferences about posts (7=most preferred)

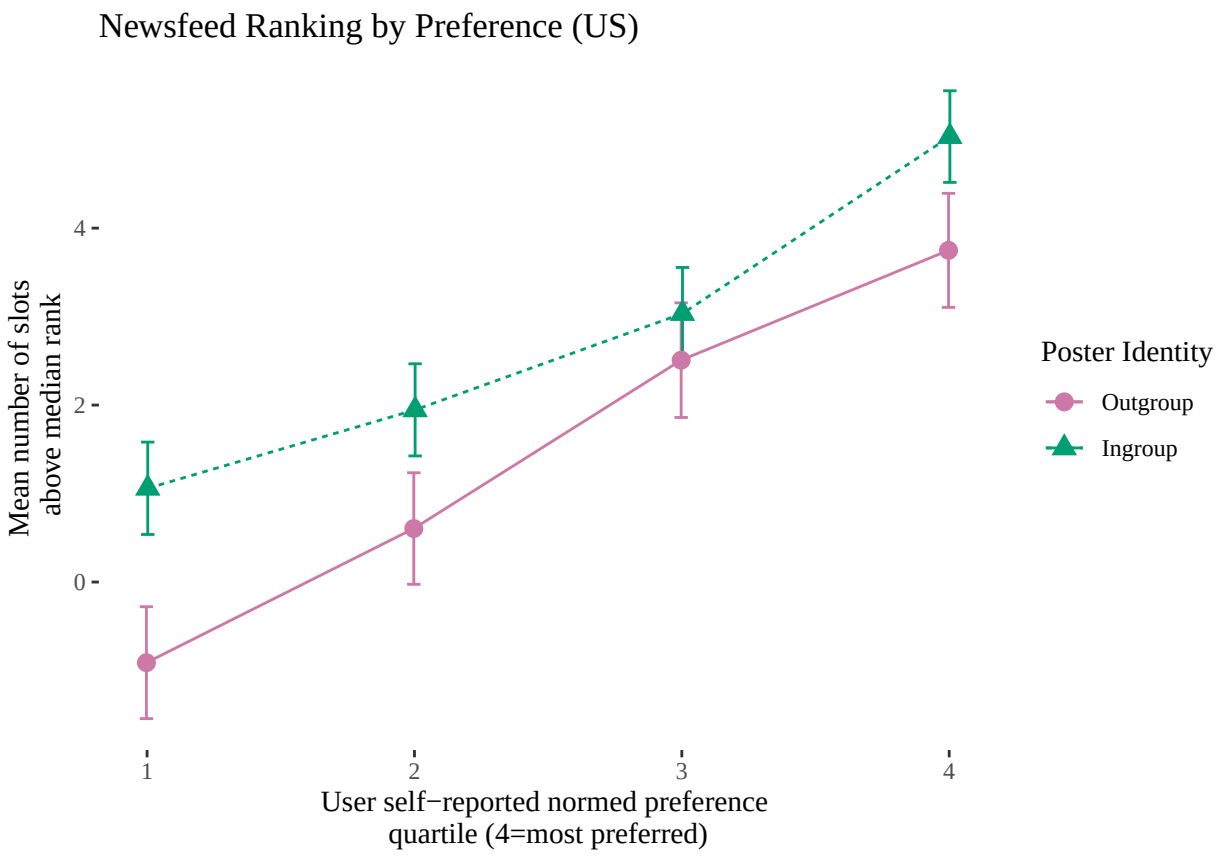

Newsfeed Top 5 by Preference (US)
Pr(Top 5)
0.18
0.15
0.12
0.09
2
4
6
User self−reported preferences about posts
(7=most preferred)

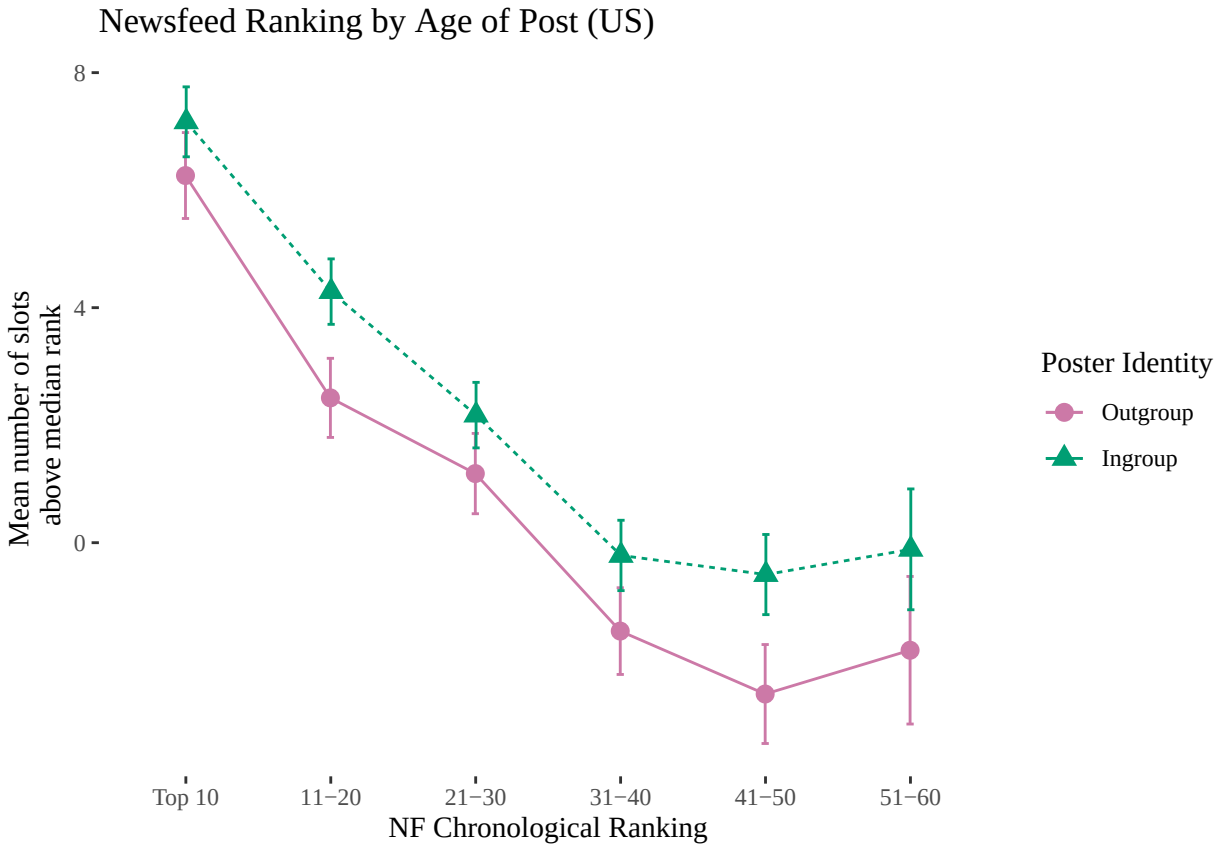

A
Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
4
2
0
1
2
3
4
User self−reported normed preference quartile (4=most preferred)
Poster Identity
Outgroup
Ingroup


B
Newsfeed Top 5 by Preference (US)
Pr(Top 5)
0.150
0.125
0.100
0.075
1
2
3
4

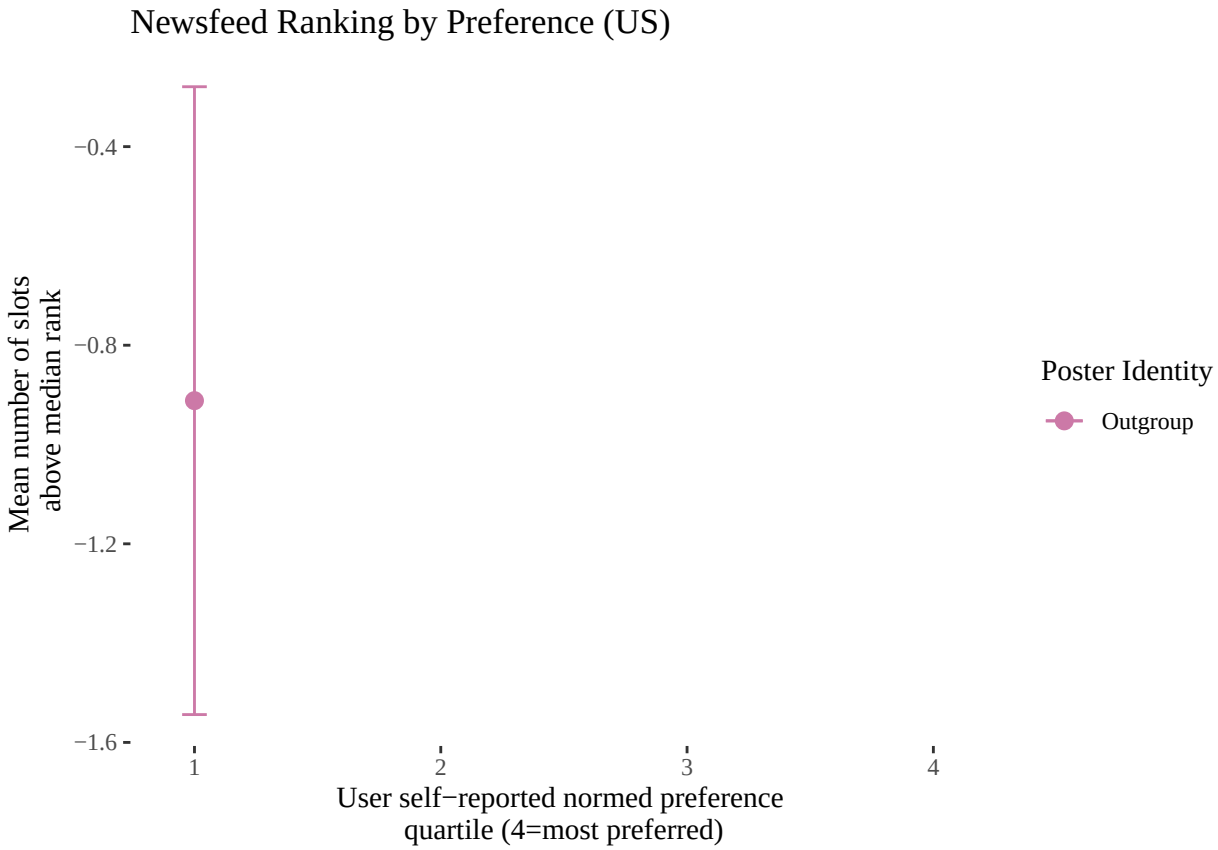

C
Newsfeed Top 10 by Preference (US)
Pr(Top 10)
0.28
0.24
0.20
0.16
1
2
3
4


D
Newsfeed Top 15 by Preference (US)
Pr(Top 15)
0.36
0.32
0.28
0.24
1
2
3
4

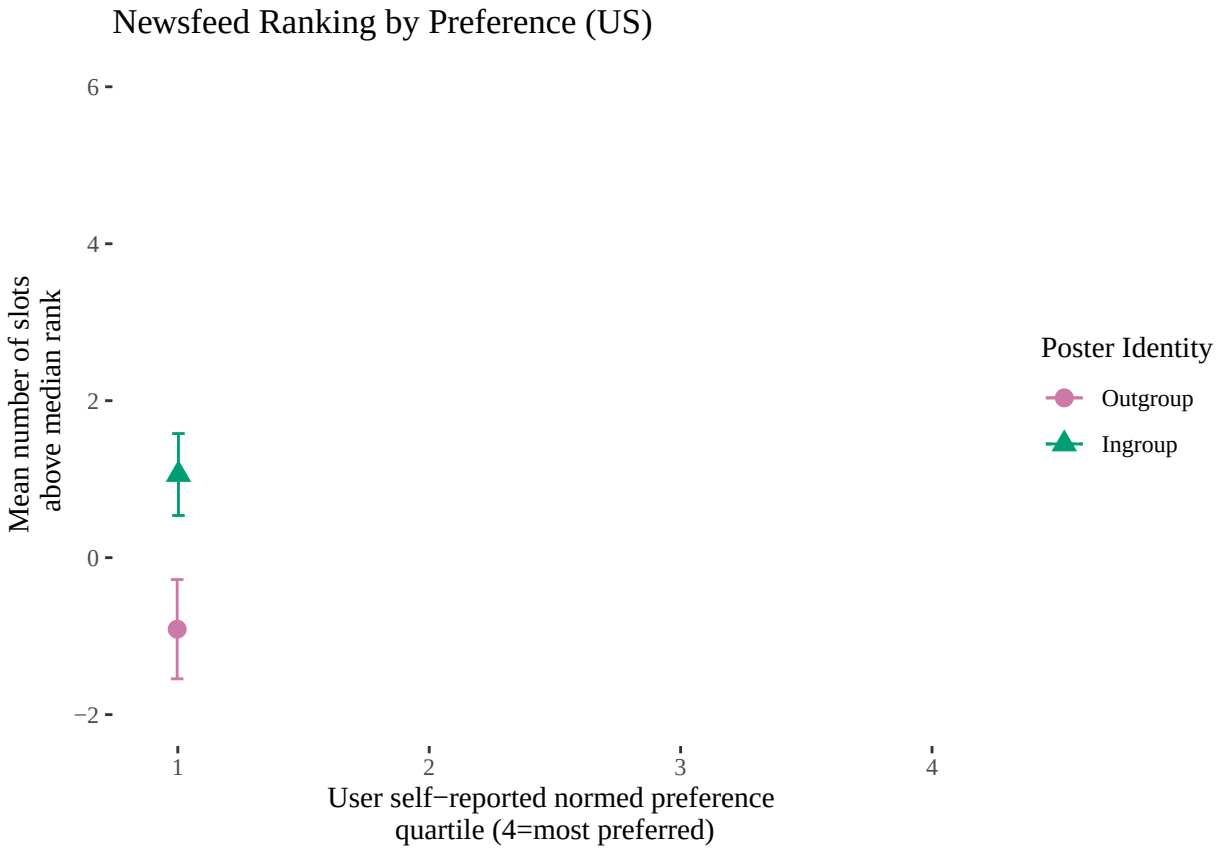

Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
4
2
0
1
2
3
4
User self−reported normed preference quartile (4=most preferred)
Poster Identity
Outgroup
Ingroup

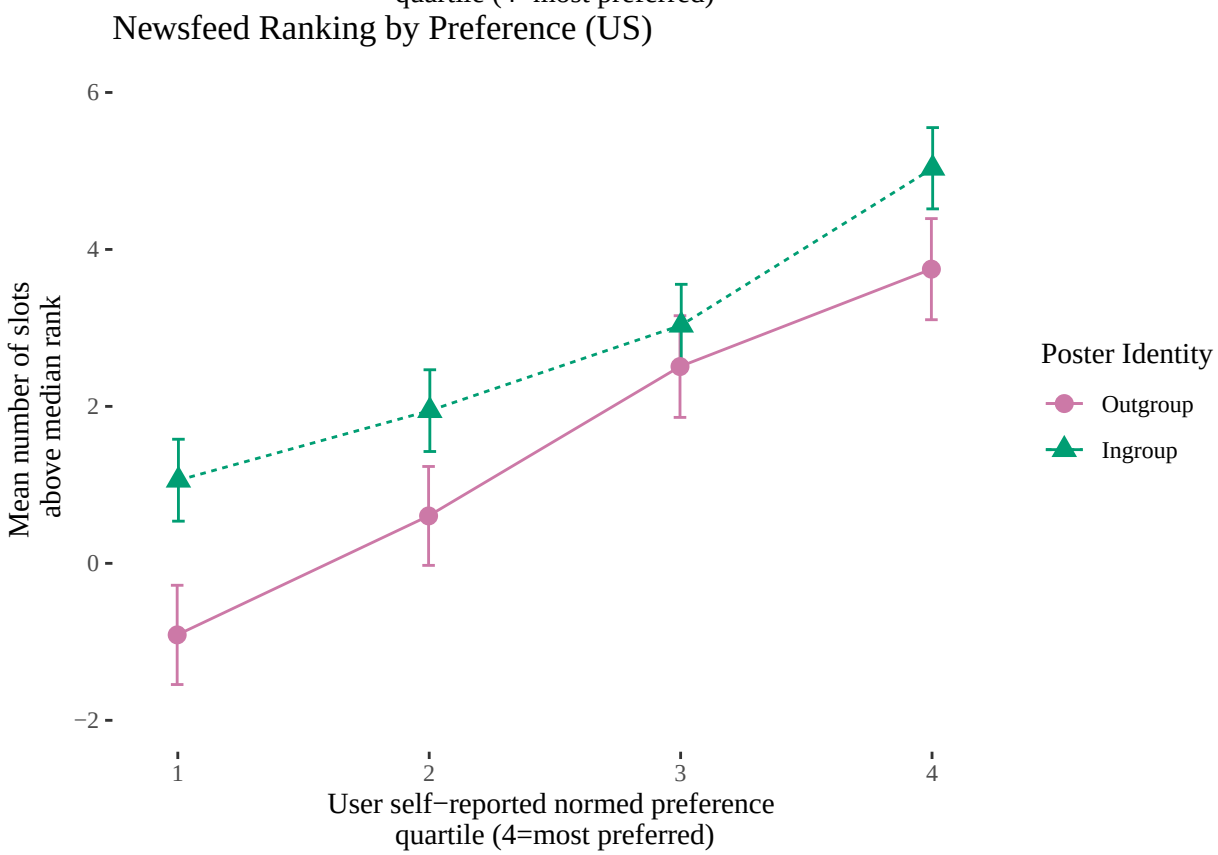

Newsfeed Ranking by Age of Post (US)
Mean number of slots above median rank
8
4
0
Top 10
11−20
21−30
31−40
41−50
51−60
NF Chronological Ranking
Poster Identity
Outgroup
Ingroup

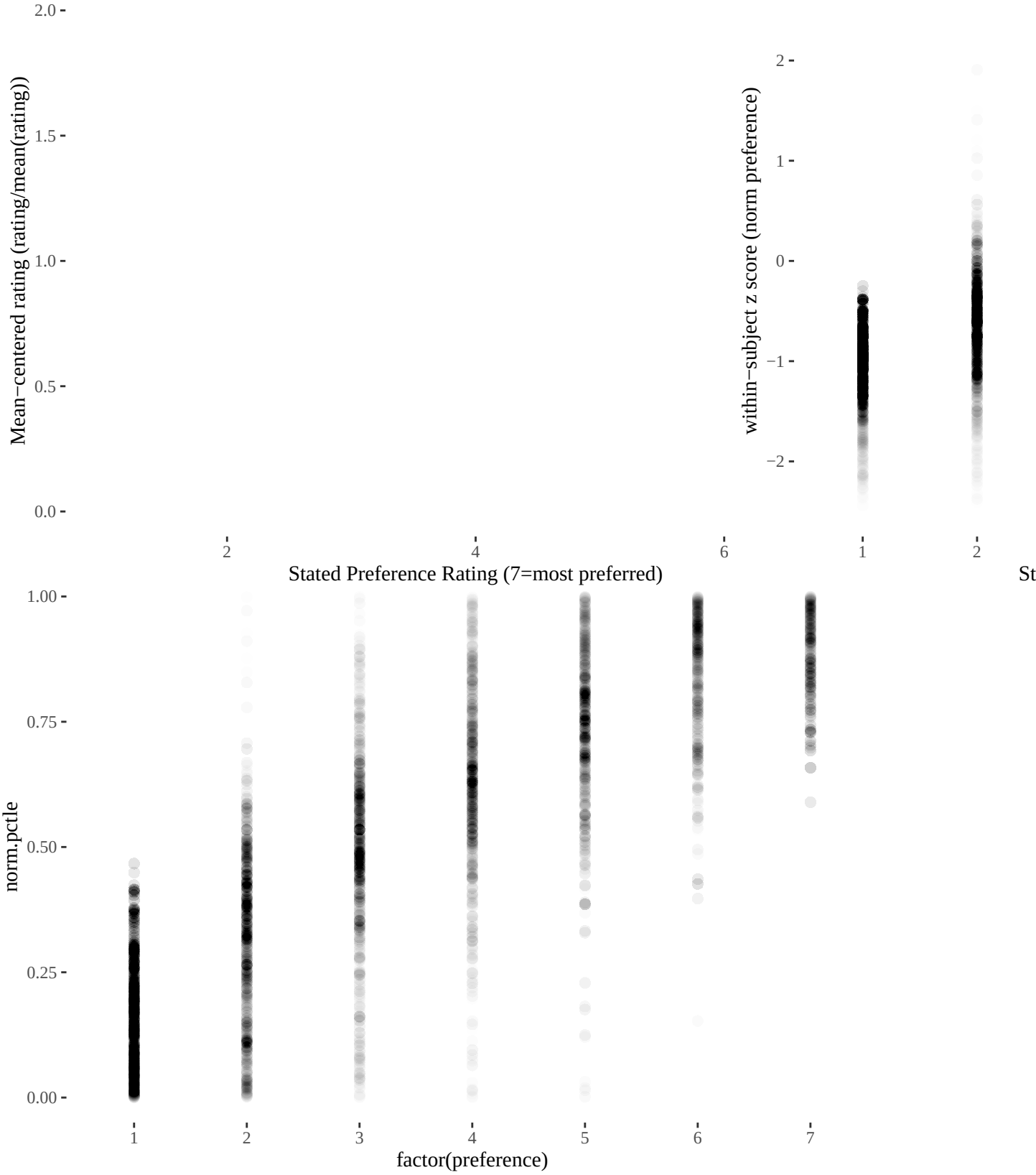
```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

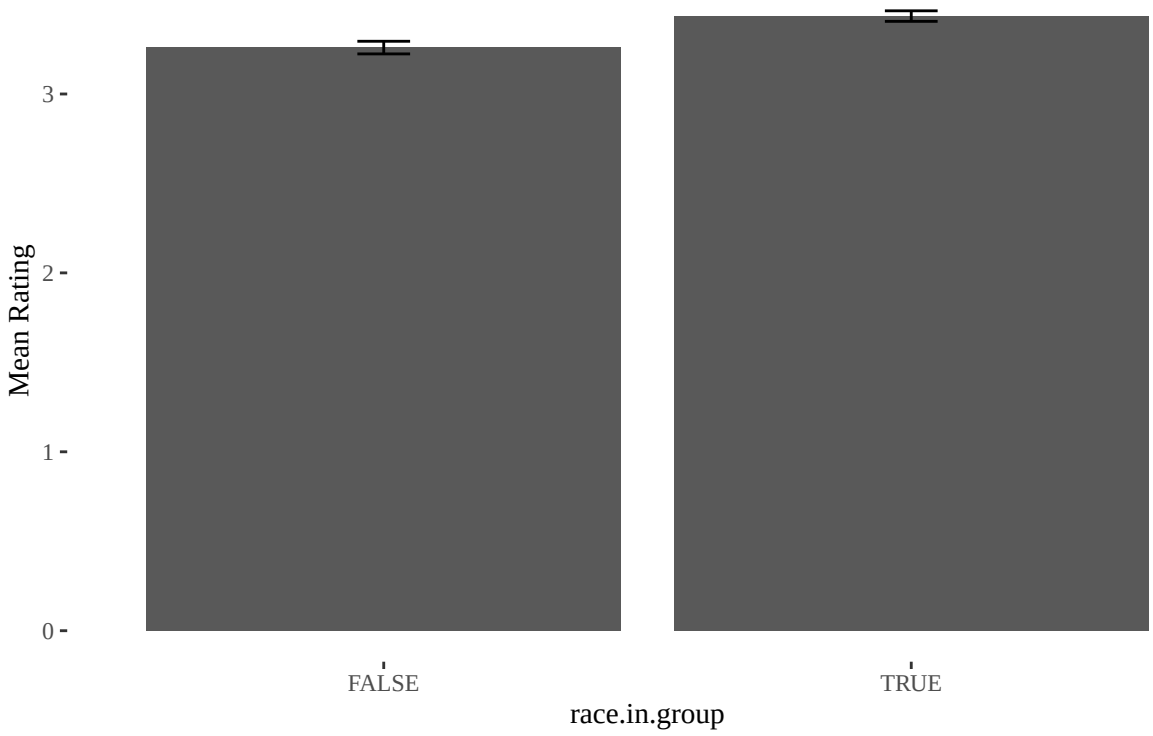



```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

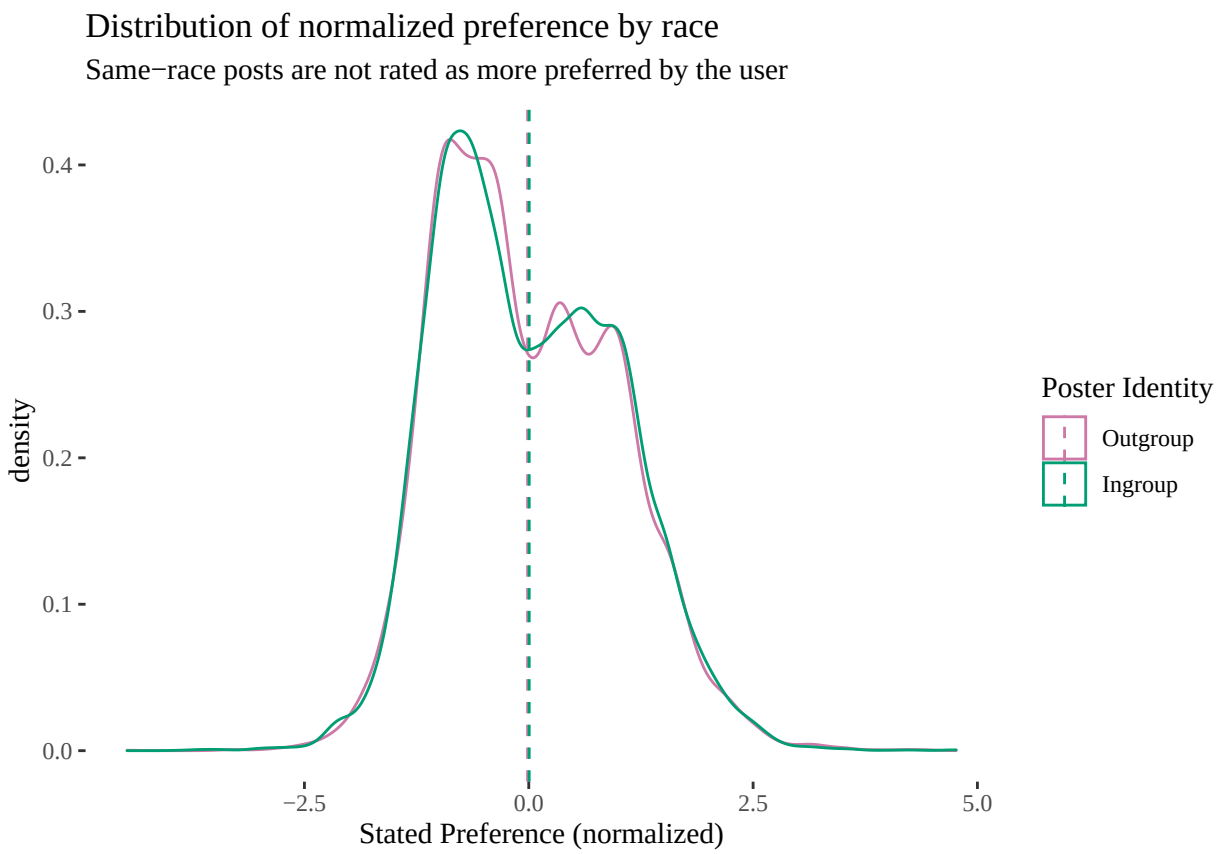

Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
6
4
2
0
−2
1
2
3
4
User self−reported normed preference quartile (4=most preferred)
Poster Identity
Outgroup
Ingroup

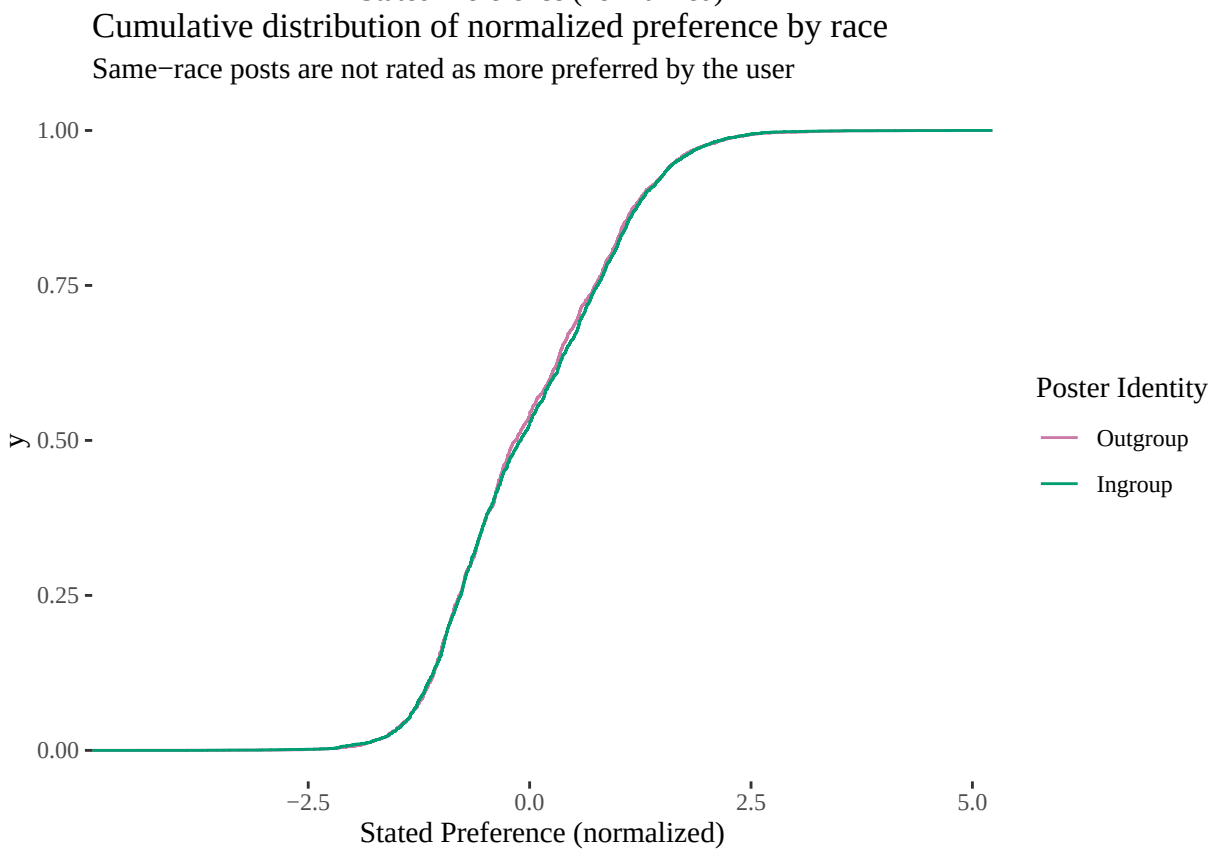

Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
6
4
2
0
−2
1
2
3
4
User self−reported normed preference quartile (4=most preferred)
Poster Identity
Outgroup
Ingroup


Mean−centered rating (rating/mean(rating))
2.0
1.5
1.0
0.5
0.0
2
4
6
Stated Preference Rating (7=most preferred)
within−subject z score (norm preference)
2
1
0
−1
−2
1
2
St
norm.pctle
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
6
7
factor(preference)

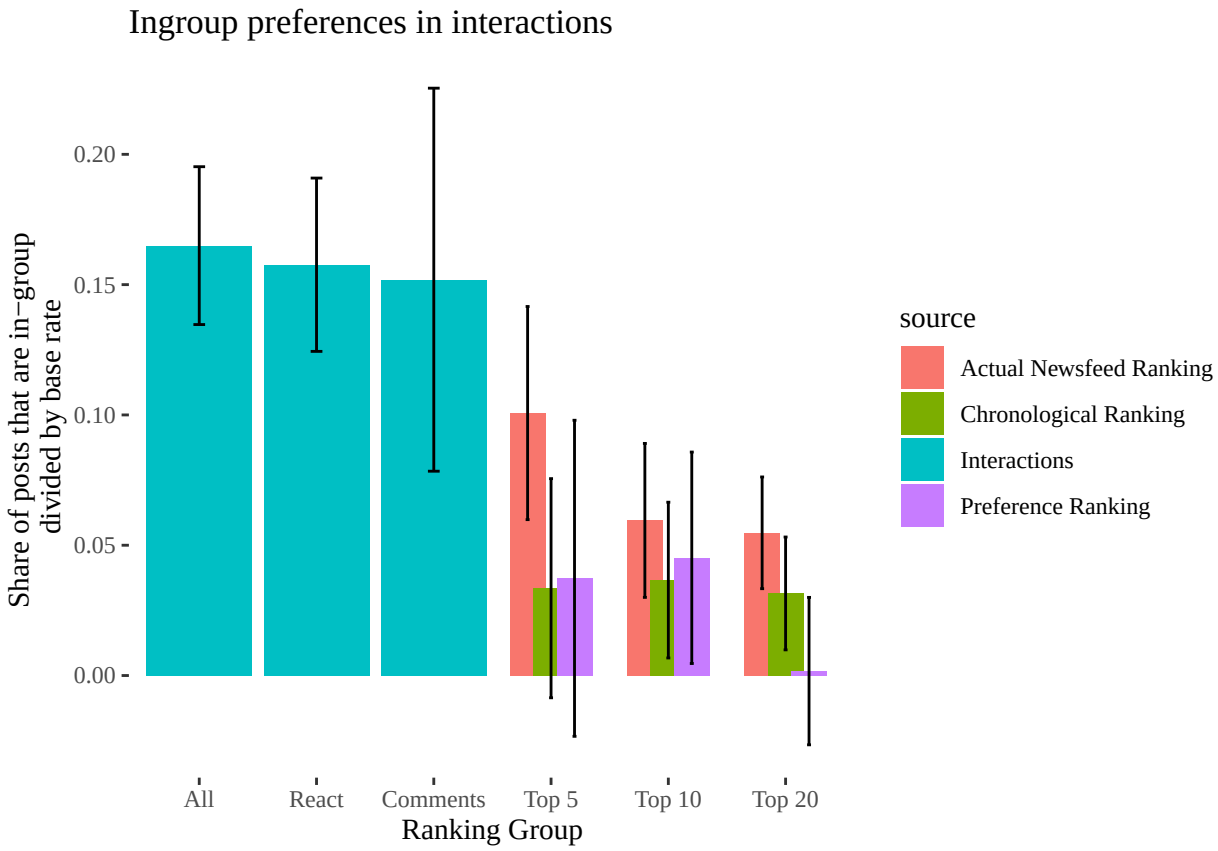


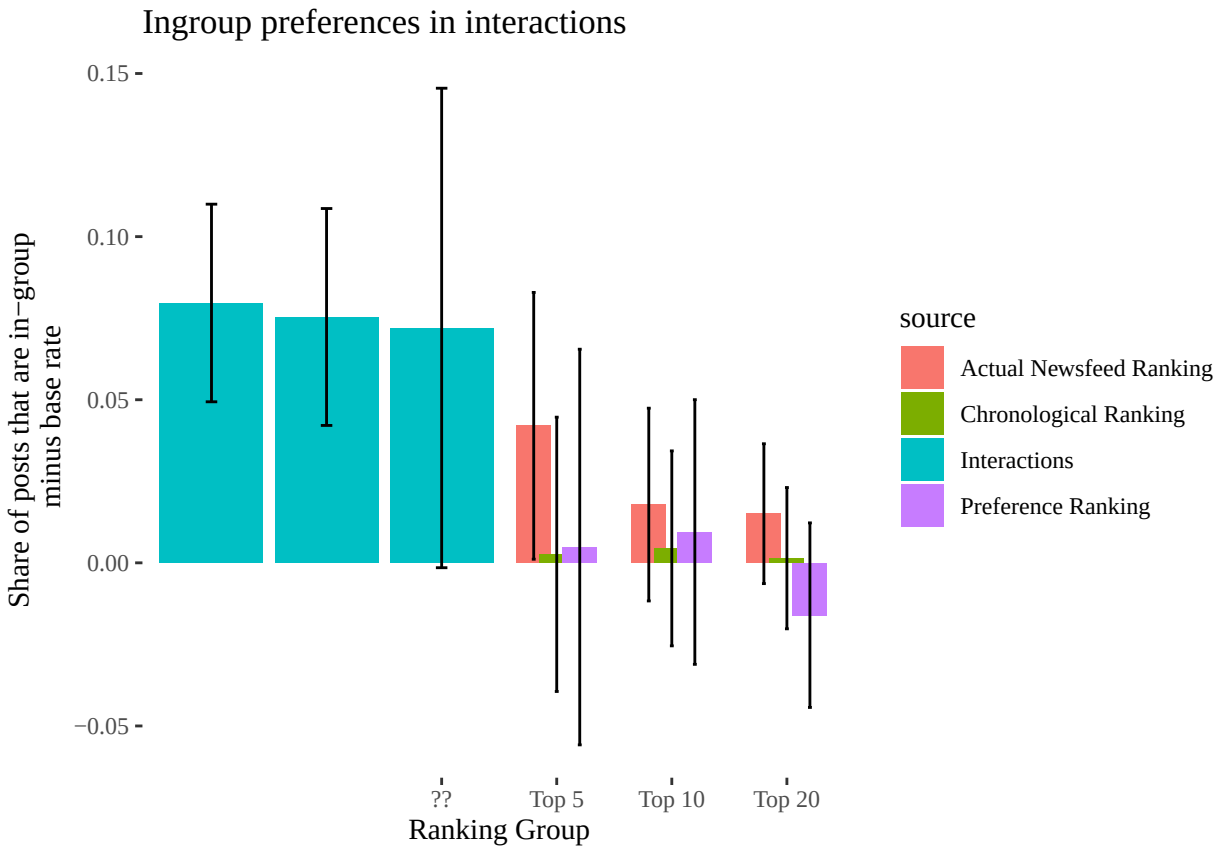

```
## Adding missing grouping variables: `dedupid`
```


Distribution of normalized preference by race
Same−race posts are not rated as more preferred by the user
density
0.4
0.3
0.2
0.1
0.0
−2.5
0.0
2.5
5.0
Stated Preference (normalized)
Poster Identity
Outgroup
Ingroup


Cumulative distribution of normalized preference by race
Same−race posts are not rated as more preferred by the user
y
1.00
0.75
0.50
0.25
0.00
−2.5
0.0
2.5
5.0
Stated Preference (normalized)
Poster Identity
Outgroup
Ingroup

```
## Adding missing grouping variables: `dedupid`
## Adding missing grouping variables: `dedupid`
```



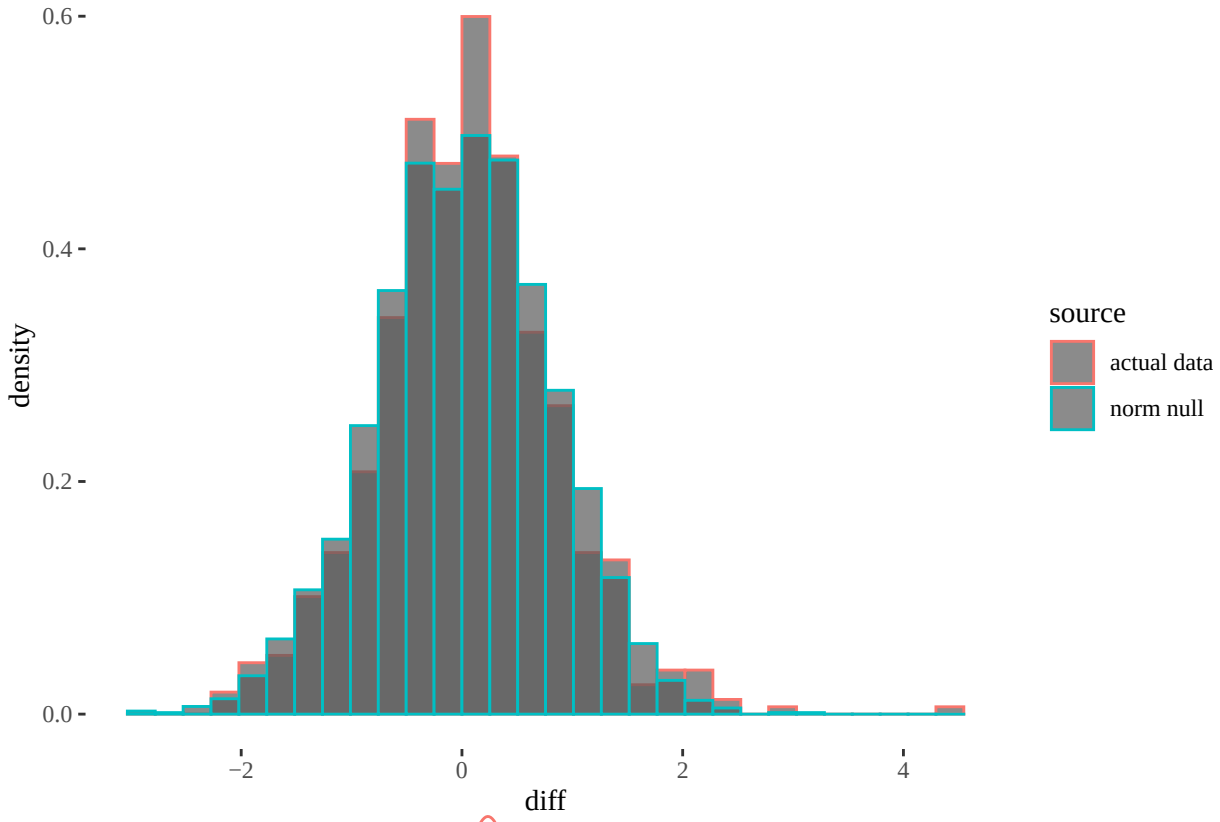


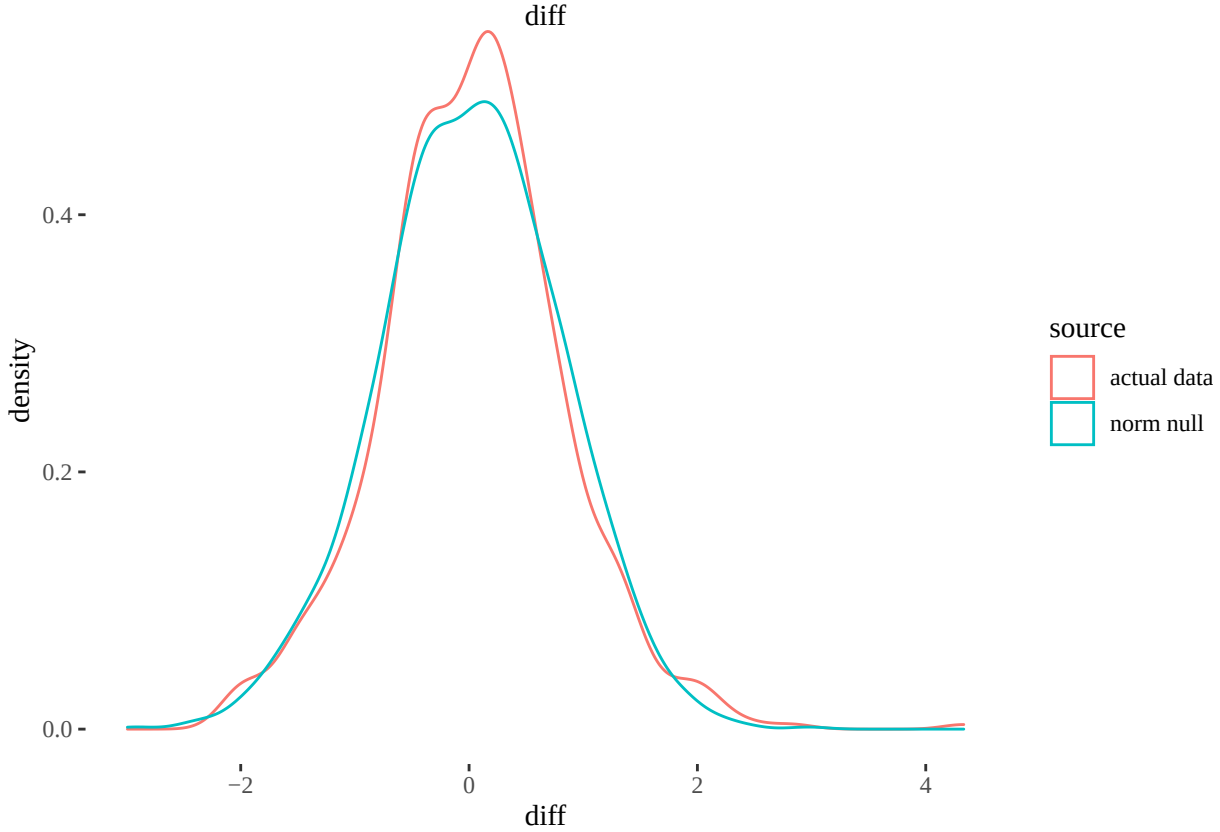
% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Jan 26, 2021 - 14:24:14

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

Table 6: In-group Baseline Heterogeneity Results - NF

	Dependent variable:		
	nf.order		
	(1)	(2)	(3)
	pctle	qrtle	med
race.in.group	−0.748*	−0.733	−1.108***
	(0.447)	(0.476)	(0.303)
I(100 ∗in.group.pctle)	0.017**		
	(0.007)		
factor(in.group.quartile)1		0.454	
		(0.389)	
factor(in.group.quartile)2		0.368	
		(0.463)	
factor(in.group.quartile)3		1.719***	
		(0.650)	
factor(in.group.med.split)1			0.590
			(0.380)
I(100 ∗norm.pctle)	−0.058***	−0.058***	−0.058***
	(0.004)	(0.004)	(0.004)
race.in.groupTRUE:I(100 ∗in.group.pctle)	−0.015*		
	(0.009)		
race.in.groupTRUE:factor(in.group.quartile)1		−0.747	
		(0.633)	
race.in.groupTRUE:factor(in.group.quartile)2		−0.779	
		(0.664)	
race.in.groupTRUE:factor(in.group.quartile)3		−1.662**	
		(0.799)	
race.in.groupTRUE:factor(in.group.med.split)1			−0.546
			(0.477)
Constant	31.469***	31.694***	31.868***
	(0.328)	(0.303)	(0.261)
Observations	27,267	27,267	27,267
R^2	0.011	0.011	0.011
Adjusted R^2	0.011	0.011	0.011
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 26, 2021 - 14:24:16

Table 7: Diff in mean in-out liking Heterogeneity Results - NF

	Dependent variable:		
	nf.order		
	(1)	(2)	(3)
	pctle	qrtle	med
race.in.group	−0.981**	−0.987**	−0.769**
	(0.441)	(0.447)	(0.306)
I(100 ∗jens.pctle)	0.006		
	(0.006)		
factor(jens.quartile)1		0.059	
		(0.485)	
factor(jens.quartile)2		1.226**	
		(0.483)	
factor(jens.quartile)3		0.285	
		(0.491)	
factor(jens.med.split)1			0.739**
			(0.332)
I(100 ∗norm.pctle)	−0.058***	−0.058***	−0.057***
	(0.004)	(0.004)	(0.004)
race.in.groupTRUE:I(100 ∗jens.pctle)	−0.005		
	(0.008)		
race.in.groupTRUE:factor(jens.quartile)1		0.483	
		(0.615)	
race.in.groupTRUE:factor(jens.quartile)2		−1.082*	
		(0.616)	
race.in.groupTRUE:factor(jens.quartile)3		−0.204	
		(0.620)	
race.in.groupTRUE:factor(jens.med.split)1			−0.878**
			(0.430)
Constant	31.708***	31.561***	31.576***
	(0.414)	(0.421)	(0.311)
Observations	27,267	27,267	27,267
R^2	0.011	0.011	0.011
Adjusted R^2	0.011	0.011	0.011
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 26, 2021 - 14:24:17

% 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 26, 2021 - 14:24:18

% 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 26, 2021 - 14:24:19

% 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 26, 2021 - 14:24:19

% 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 26, 2021 - 14:24:19

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

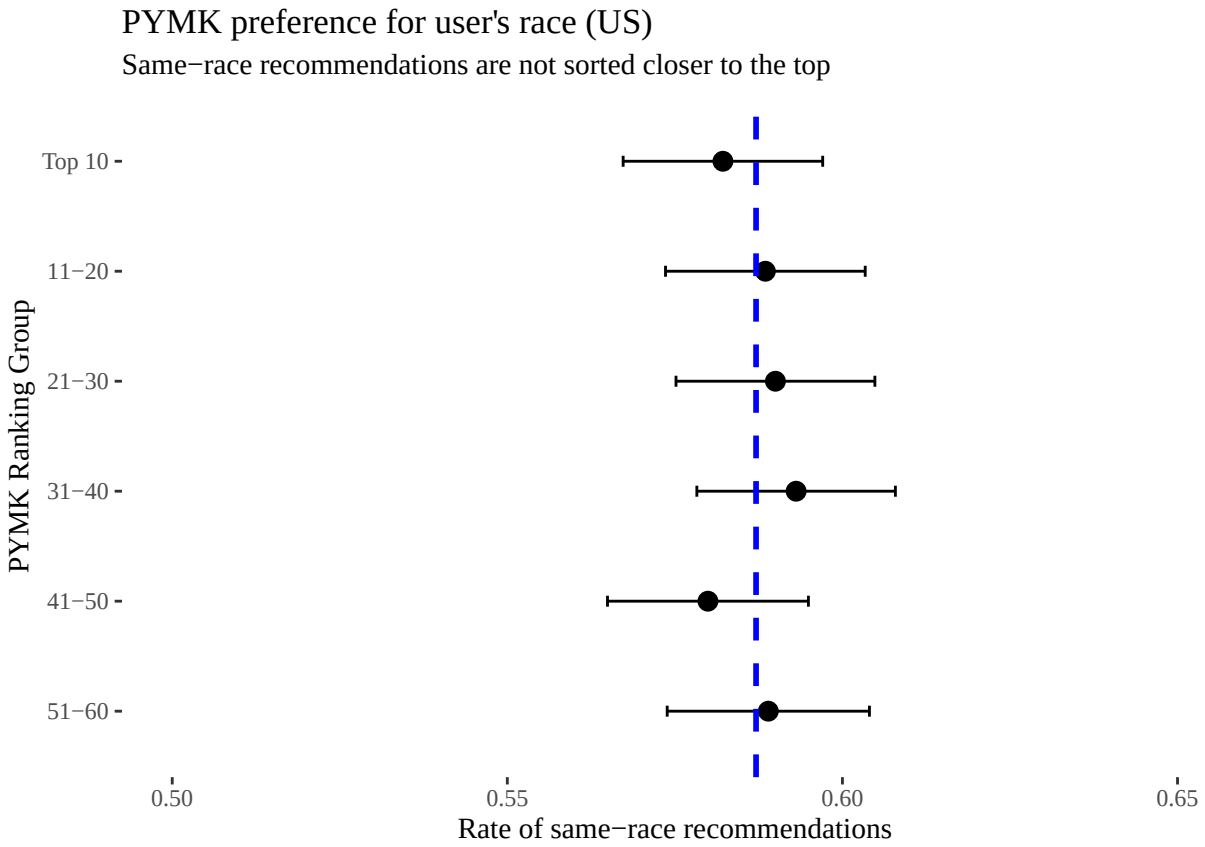


Table 8: In-group Baseline Heterogeneity Results - PYMK

	Dependent variable:		
	pymk.order		
	(1)	(2)	(3)
	pctle	qrtle	med
race.in.group	0.614	−0.008	0.594*
	(0.474)	(0.517)	(0.323)
I(100 ∗in.group.pctle)	−0.003		
	(0.007)		
factor(in.group.quartile)1		−0.793**	
		(0.403)	
factor(in.group.quartile)2		−0.017	
		(0.464)	
factor(in.group.quartile)3		0.046	
		(0.619)	
factor(in.group.med.split)1			0.325
			(0.375)
I(100 ∗norm.pctle)	−0.089***	−0.089***	−0.089***
	(0.004)	(0.004)	(0.004)
race.in.groupTRUE:I(100 ∗in.group.pctle)	−0.004		
	(0.009)		
race.in.groupTRUE:factor(in.group.quartile)1		1.208*	
		(0.679)	
race.in.groupTRUE:factor(in.group.quartile)2		0.124	
		(0.698)	
race.in.groupTRUE:factor(in.group.quartile)3		−0.256	
		(0.798)	
race.in.groupTRUE:factor(in.group.med.split)1			−0.680
			(0.486)
Constant	33.767***	33.892***	33.574***
	(0.344)	(0.317)	(0.273)
Observations	24,511	24,511	24,511
R^2	0.022	0.023	0.022
Adjusted R^2	0.022	0.022	0.022
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 9: Diff in mean in-out liking Heterogeneity Results - PYMK

	Dependent variable:		
	pymk.order		
	(1)	(2)	(3)
	pctle	qrtle	med
race.in.group	0.595	0.411	0.498
	(0.444)	(0.440)	(0.311)
I(100 ∗jens.pctle)	−0.001		
	(0.006)		
factor(jens.quartile)1		0.110	
		(0.463)	
factor(jens.quartile)2		0.037	
		(0.473)	
factor(jens.quartile)3		0.111	
		(0.482)	
factor(jens.med.split)1			0.019
			(0.338)
I(100 ∗norm.pctle)	−0.089***	−0.089***	−0.089***
	(0.004)	(0.004)	(0.004)
race.in.groupTRUE:I(100 ∗jens.pctle)	−0.006		
	(0.008)		
race.in.groupTRUE:factor(jens.quartile)1		0.175	
		(0.622)	
race.in.groupTRUE:factor(jens.quartile)2		−0.098	
		(0.626)	
race.in.groupTRUE:factor(jens.quartile)3		−0.635	
		(0.631)	
race.in.groupTRUE:factor(jens.med.split)1			−0.458
			(0.445)
Constant	33.715***	33.594***	33.647***
	(0.399)	(0.387)	(0.307)
Observations	24,511	24,511	24,511
R^2	0.022	0.022	0.022
Adjusted R^2	0.022	0.022	0.022
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 10: Time Results - Rank

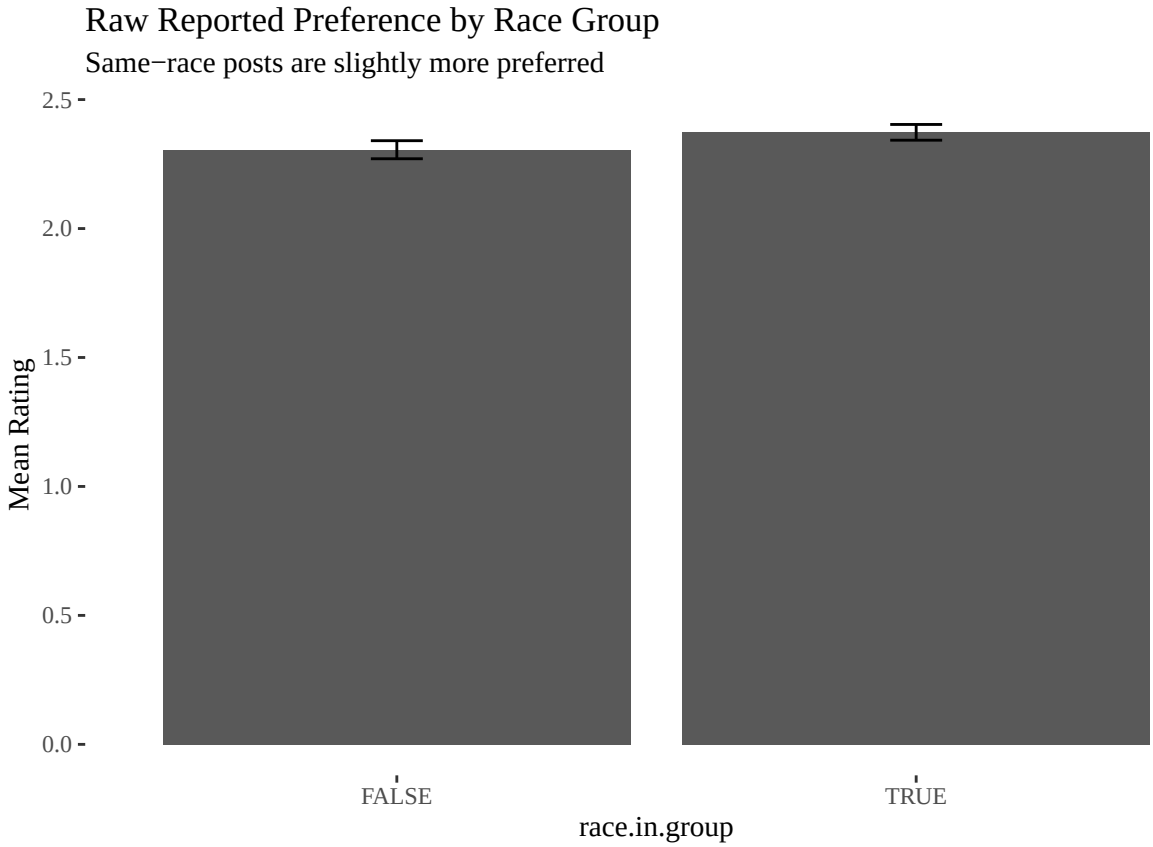
	Dependent variable:		
	nf.order		
	(1)	(2)	(3)
	all	all2	base
I(100 ∗norm.pctle)	−0.062***	−0.062***	−0.059***
	(0.004)	(0.004)	(0.004)
time_rank	0.201***	0.192***	
	(0.011)	(0.007)	
race.in.group	−0.958**	−1.344***	−1.285***
	(0.391)	(0.207)	(0.210)
time_rank:race.in.group	−0.016		
	(0.014)		
Constant	27.327***	27.556***	31.947***
	(0.346)	(0.284)	(0.241)
Observations	28,175	28,175	28,175
R^2	0.038	0.038	0.011
Adjusted R^2	0.038	0.038	0.011
Note:			*p<0.1; **p<0.05; ***p<0.01

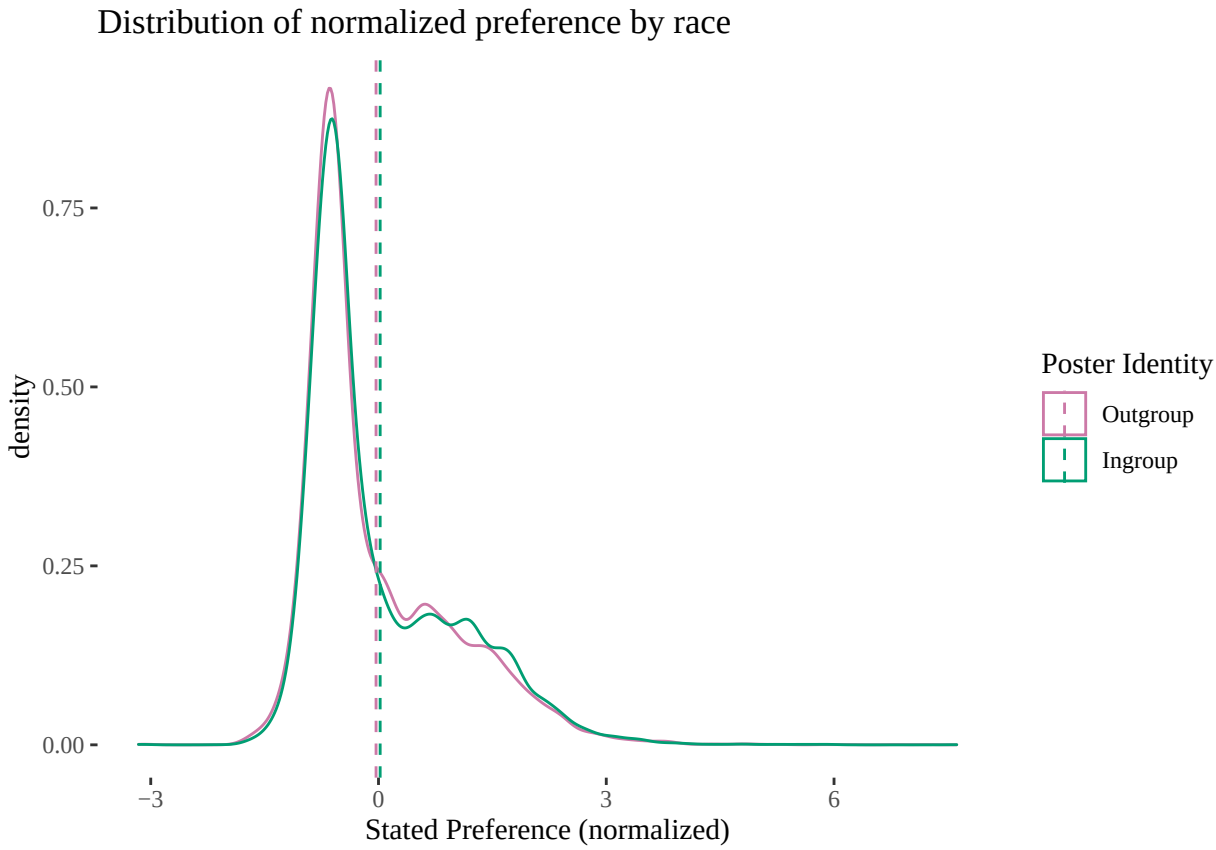
Table 11: Time Results - Top 10

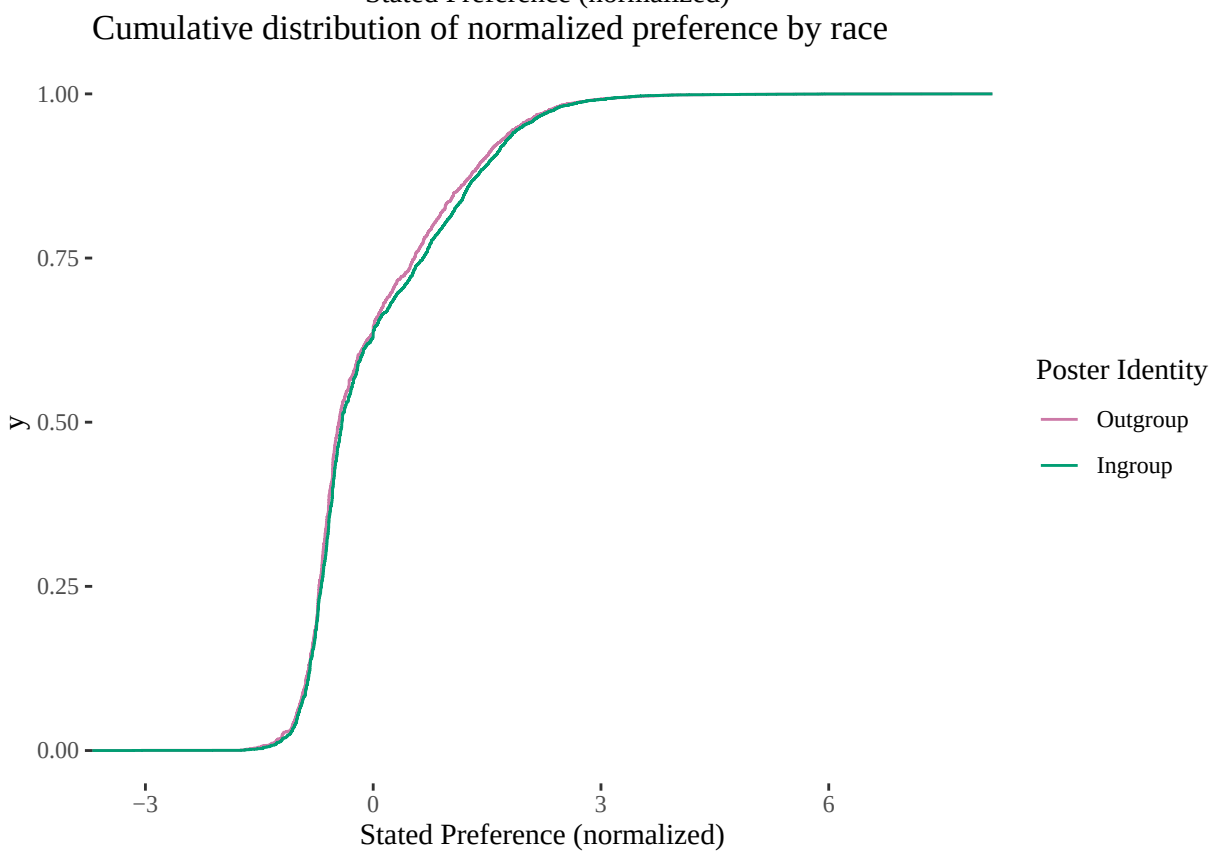
	Dependent variable:			
	nf.order			
	(1)	(2)	(3)	(4)
	all	days	hours	mins
race.in.group	−1.344***	−1.321***	−1.285***	−1.706***
	(0.207)	(0.213)	(0.272)	(0.648)
I(100 ∗norm.pctle)	−0.062***	−0.062***	−0.063***	−0.076***
	(0.004)	(0.004)	(0.005)	(0.011)
time_rank	0.192***	0.187***	0.195***	−0.012
	(0.007)	(0.007)	(0.012)	(0.069)
Constant	27.556***	27.565***	27.555***	28.772***
	(0.284)	(0.291)	(0.360)	(0.803)
Observations	28,175	26,724	17,557	3,323
R^2	0.038	0.037	0.027	0.017
Adjusted R^2	0.038	0.037	0.027	0.016
Note:				*p<0.1; **p<0.05; ***p<0.01

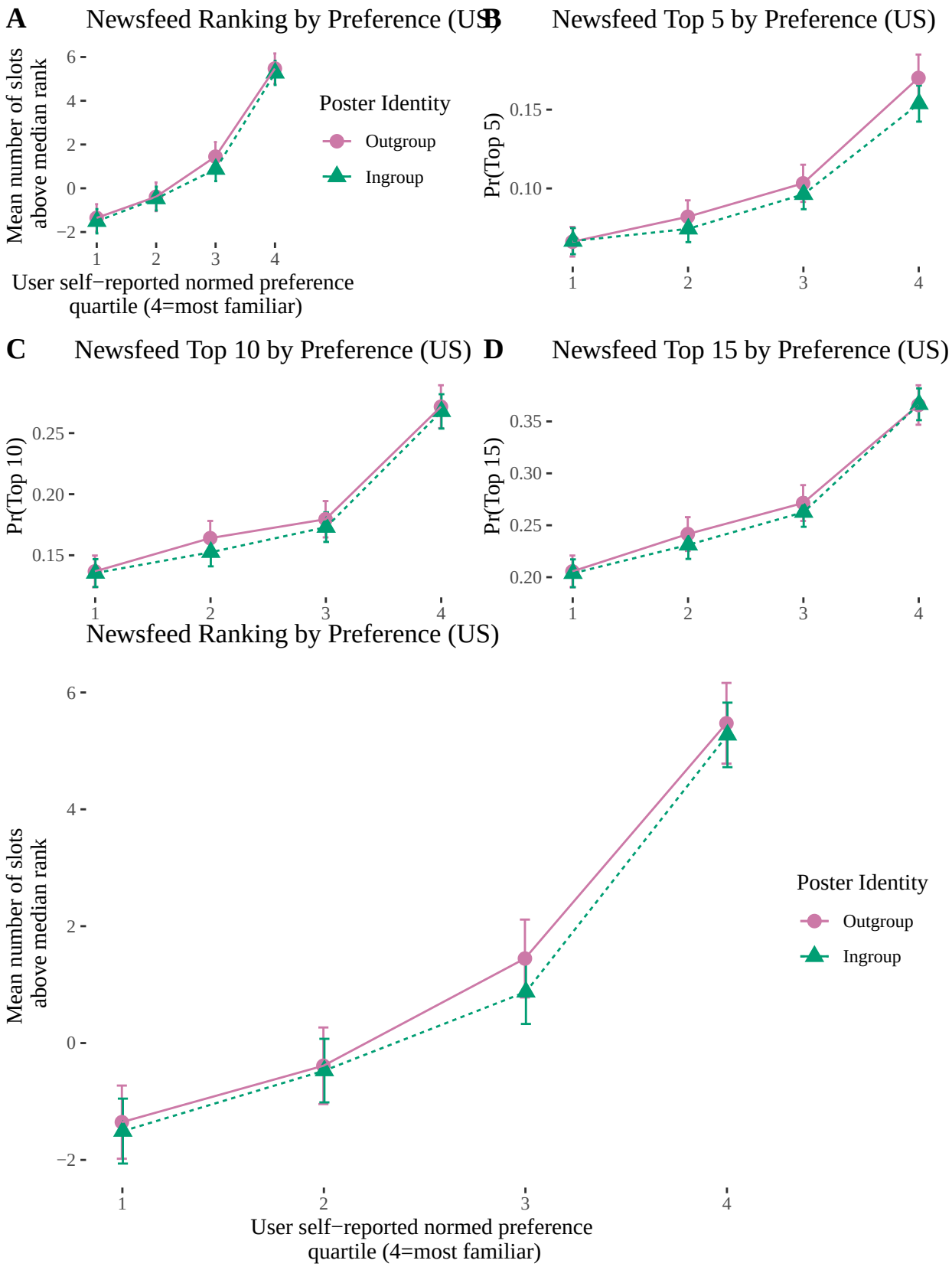
Table 12: Time Results - Top 10

	Dependent variable:			
	nf.order			
	(1)	(2)	(3)	(4)
	all	recent10	recent20	recent30
race.in.group	-1.344^{***}	-0.898^{**}	-1.416^{***}	-1.196^{***}
	(0.207)	(0.454)	(0.318)	(0.260)
I(100 $*$norm.pctle)	-0.062^{***}	-0.073^{***}	-0.064^{***}	-0.061^{***}
	(0.004)	(0.008)	(0.005)	(0.004)
time_rank	0.192^{***}	0.160^{**}	0.276^{***}	0.248^{***}
	(0.007)	(0.077)	(0.027)	(0.015)
Constant	27.556^{***}	27.157^{***}	26.611^{***}	26.579^{***}
	(0.284)	(0.655)	(0.454)	(0.367)
Observations	28,175	6,540	12,882	18,689
R^2	0.038	0.015	0.020	0.026
Adjusted R^2	0.038	0.015	0.020	0.026
Note:				*p<0.1; **p<0.05; ***p<0.01

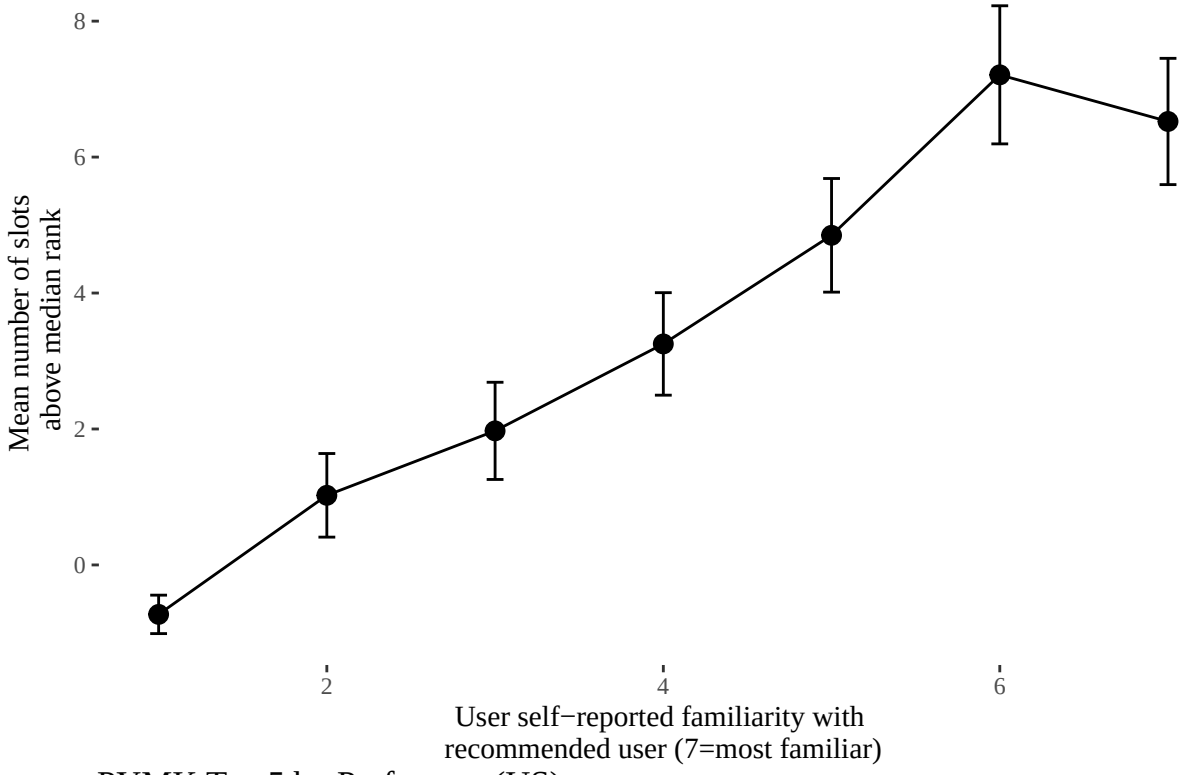



Distribution of normalized preference by race
density
0.75
0.50
0.25
0.00
−3
0
3
6
Stated Preference (normalized)
Poster Identity
Outgroup
Ingroup

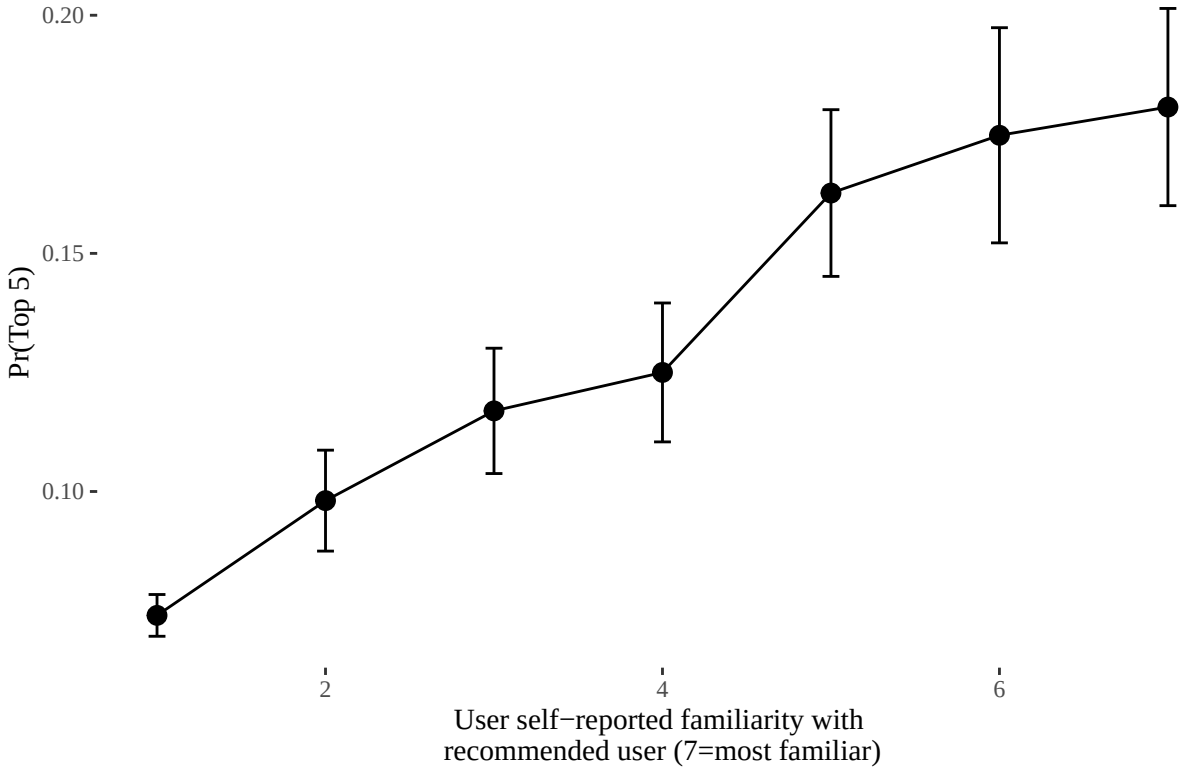

Cumulative distribution of normalized preference by race
y
1.00
0.75
0.50
0.25
0.00
−3
0
3
6
Stated Preference (normalized)
Poster Identity
Outgroup
Ingroup


A
Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
6
4
2
0
−2
1
2
3
4
User self−reported normed preference quartile (4=most familiar)
Poster Identity
Outgroup
Ingroup
B
Newsfeed Top 5 by Preference (US)
Pr(Top 5)
0.15
0.10
C
Newsfeed Top 10 by Preference (US)
Pr(Top 10)
0.25
0.20
0.15
D
Newsfeed Top 15 by Preference (US)
Pr(Top 15)
0.35
0.30
0.25
0.20
Newsfeed Ranking by Preference (US)
Mean number of slots above median rank
User self−reported normed preference quartile (4=most familiar)
Poster Identity
Outgroup
Ingroup

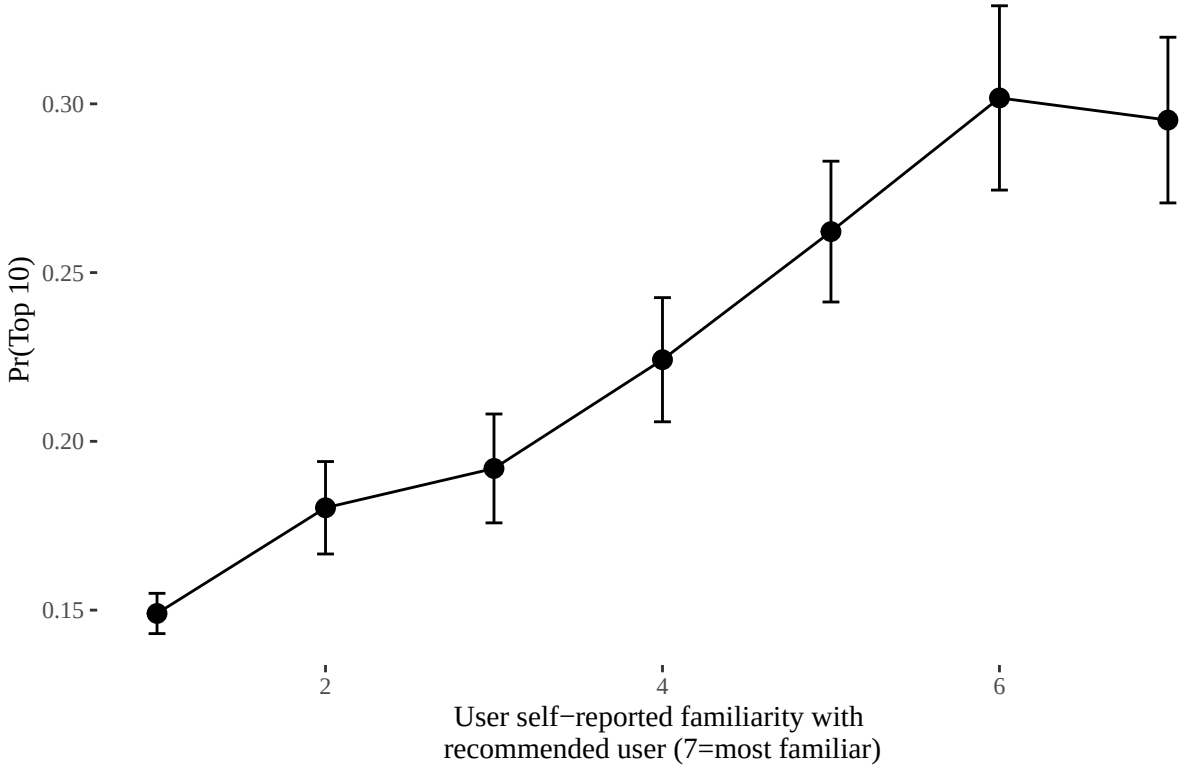
PYMK Ranking by Preference (US)



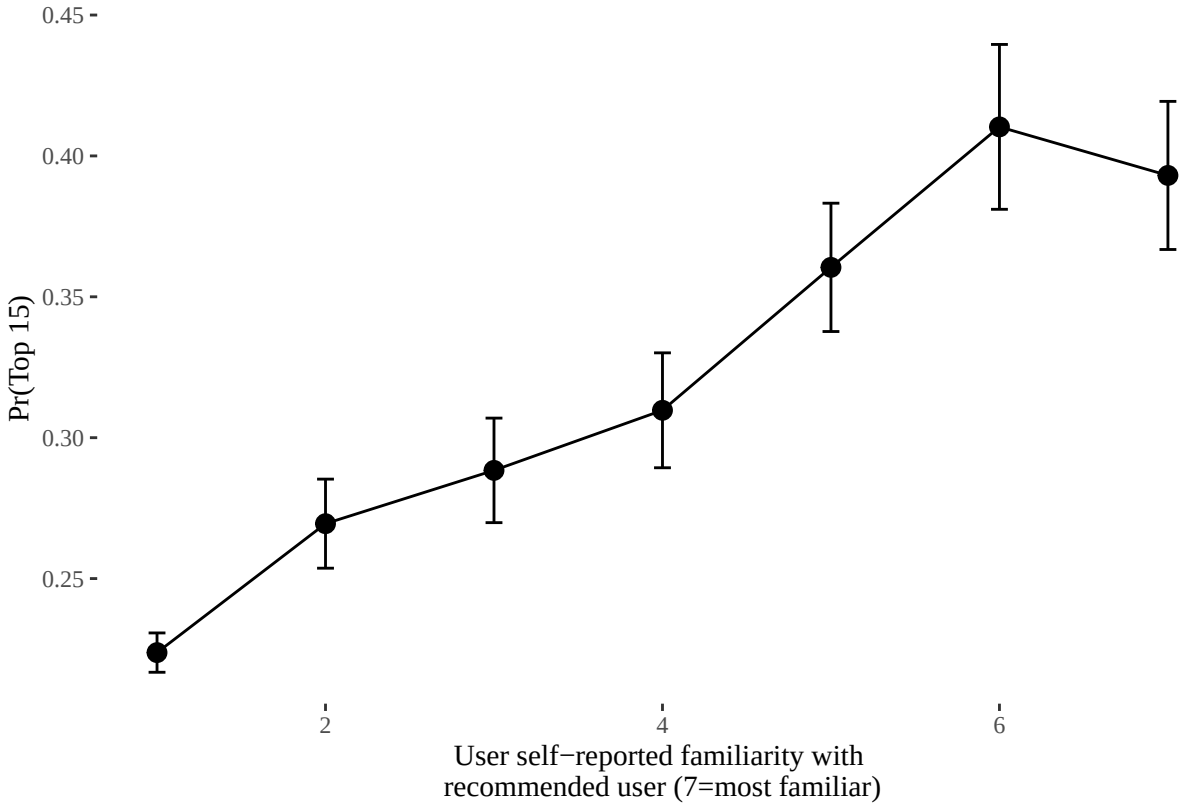
PYMK Top 5 by Preference (US)



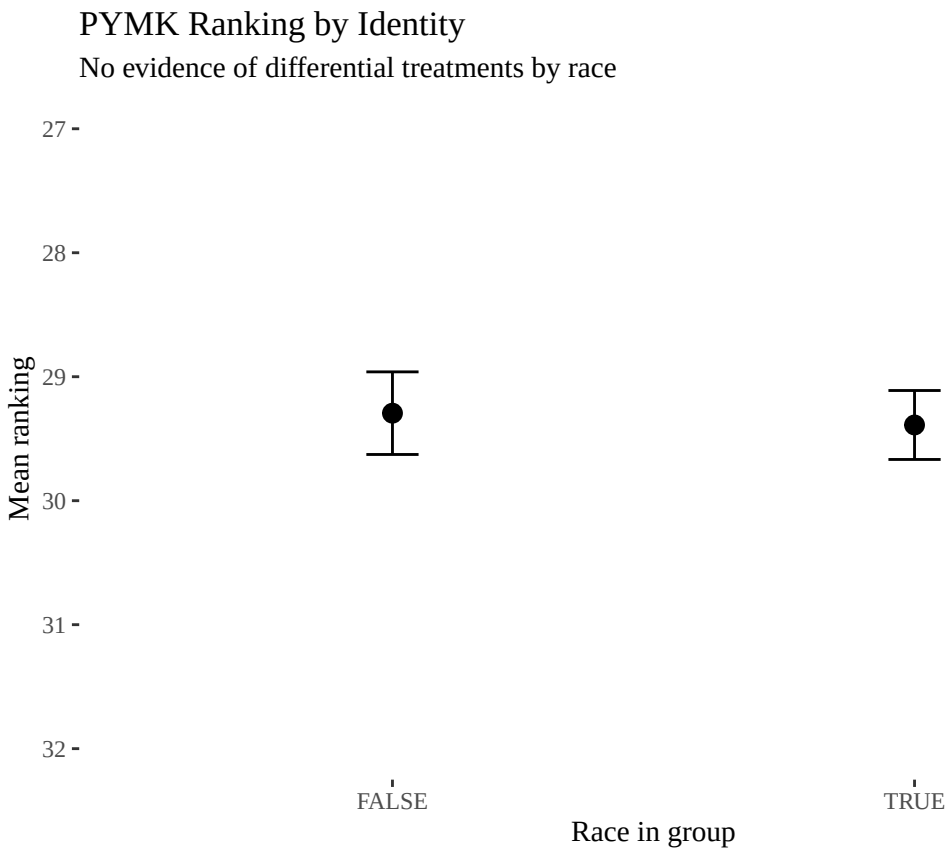
PYMK Top 10 by Preference (US)

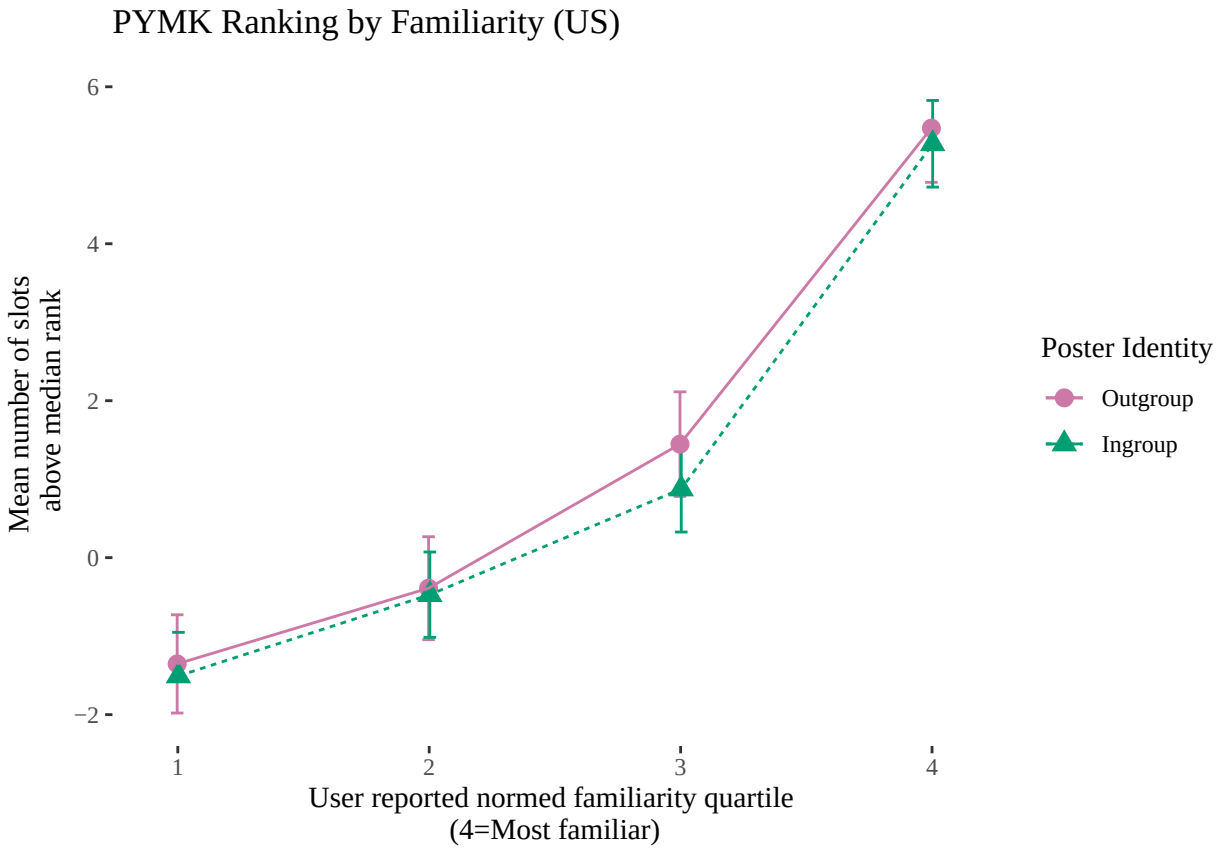


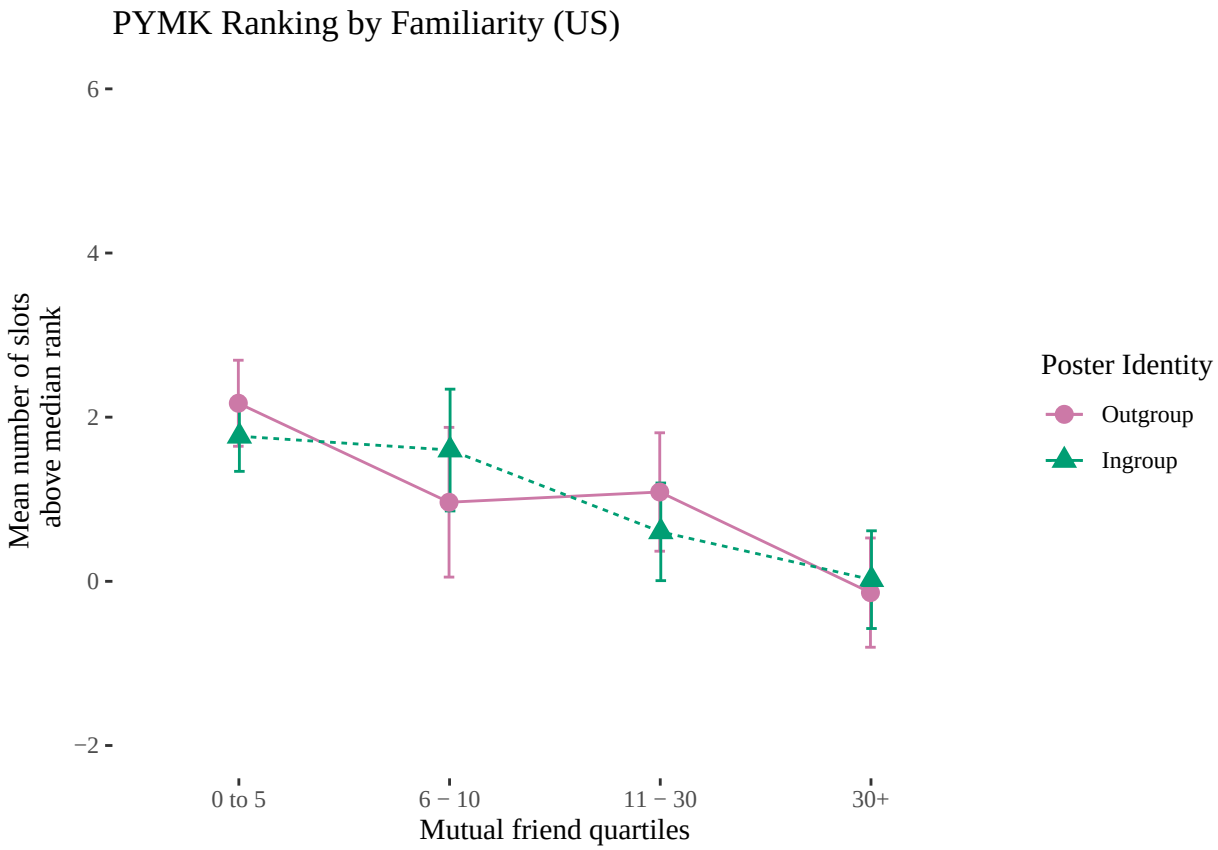
PYMK Top 15 by Preference (US)

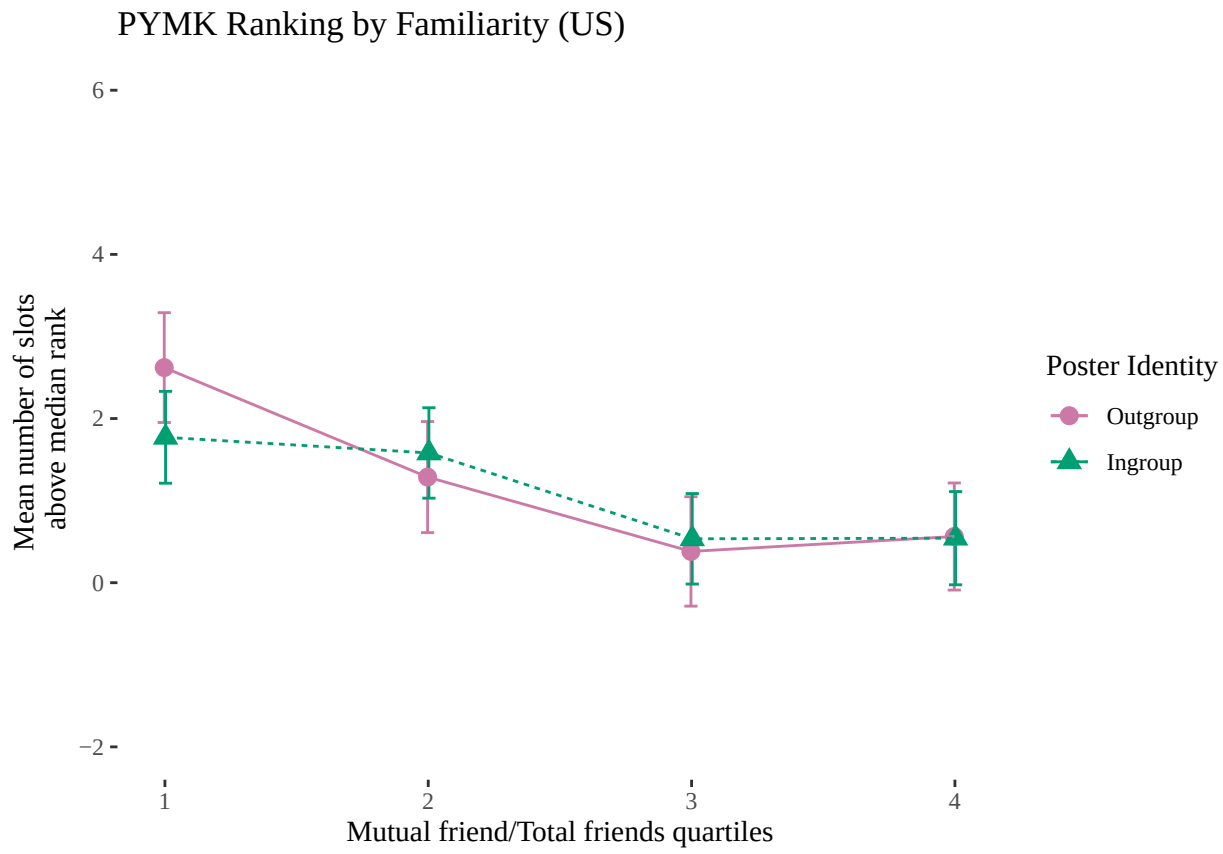


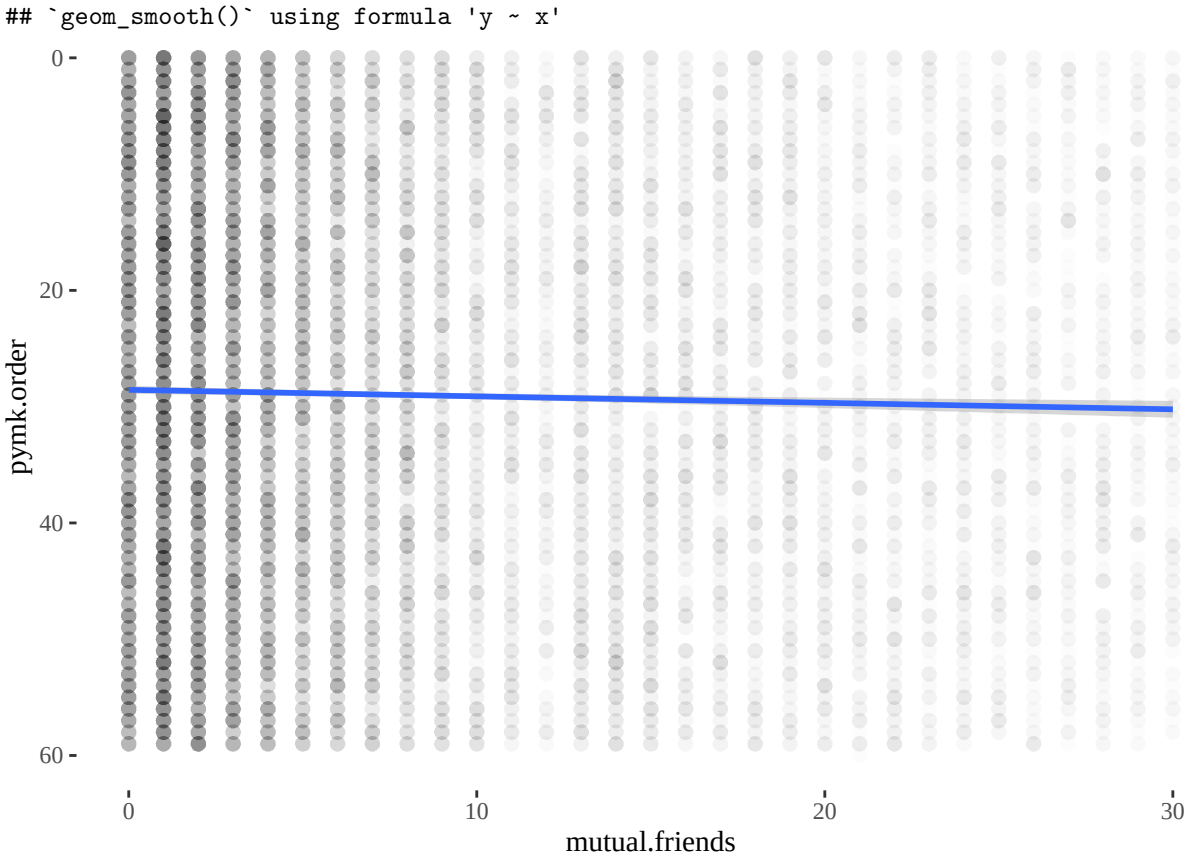
```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```

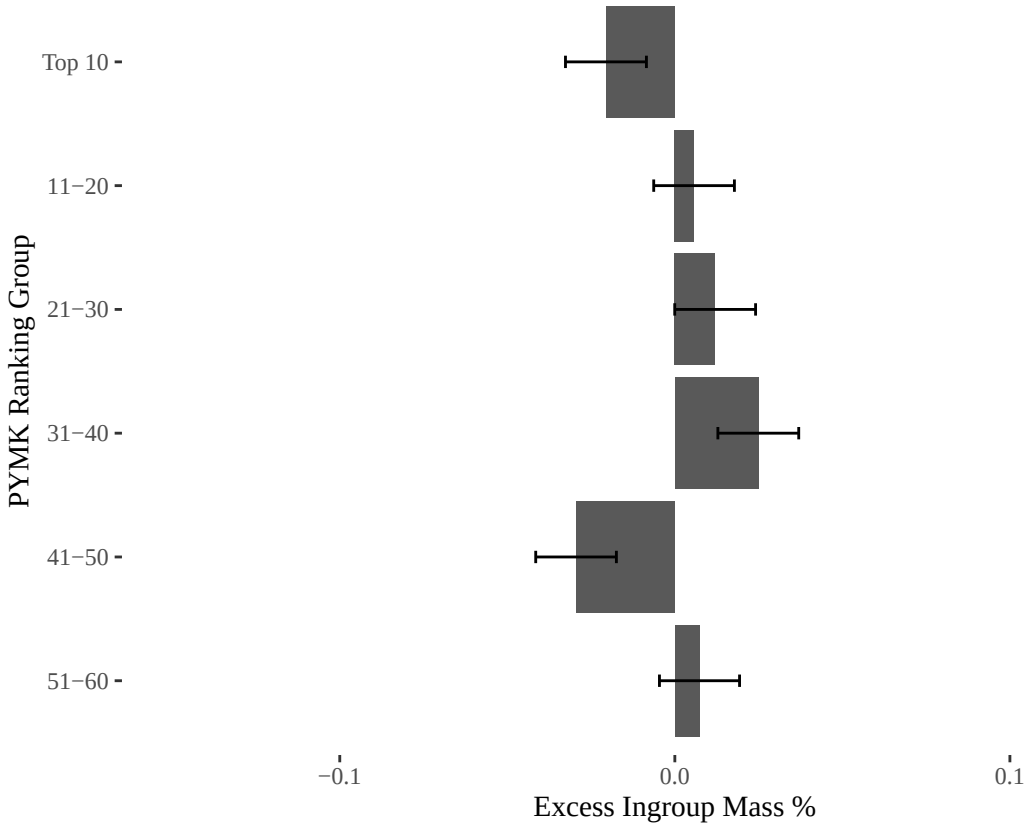



PYMK Ranking by Familiarity (US)
Mean number of slots above median rank
6
4
2
0
−2
1
2
3
4
User reported normed familiarity quartile
(4=Most familiar)
Poster Identity
Outgroup
Ingroup


PYMK Ranking by Familiarity (US)
Mean number of slots above median rank
6
4
2
0
−2
0 to 5
6 − 10
11 − 30
30+
Mutual friend quartiles
Poster Identity
Outgroup
Ingroup


PYMK Ranking by Familiarity (US)
Mean number of slots above median rank
6
4
2
0
−2
1
2
3
4
Mutual friend/Total friends quartiles
Poster Identity
Outgroup
Ingroup


`geom_smooth()` using formula 'y ~ x'
pymk.order
0
20
40
60
0
10
20
30
mutual.friends



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 26, 2021 - 14:24:37

% 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 26, 2021 - 14:24:37

% 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 26, 2021 - 14:24:38

% 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 26, 2021 - 14:24:38

% 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 26, 2021 - 14:24:39

% 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 26, 2021 - 14:24:40

Table 13: Primary NF Rank Results

	Dependent variable:				
	nf.order				
	(1)	(2)	(3)	(4)	(5)
	all.mdl	asian.mdl	black.mdl	hispanic.mdl	white.mdl
race.in.group	1.285***	1.088***	0.940	0.809	1.946***
	(0.210)	(0.319)	(0.681)	(0.691)	(0.383)
I(100 ∗norm.pctle)	0.059***	0.072***	0.039***	0.040***	0.051***
	(0.004)	(0.006)	(0.012)	(0.012)	(0.006)
Constant	−31.947***	−32.806***	−29.776***	−30.382***	−32.044***
	(0.241)	(0.355)	(0.766)	(0.770)	(0.438)
Observations	28,175	11,777	2,505	2,532	10,954
R^2	0.011	0.015	0.005	0.005	0.010
Adjusted R^2	0.011	0.015	0.004	0.004	0.009
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 14: Primary NF Top 10 Results

	Dependent variable:				
	top10				
	(1)	(2)	(3)	(4)	(5)
	all.mdl	asian.mdl	black.mdl	hispanic.mdl	white.mdl
I(100 ∗norm.pctle)	0.001***	0.001***	0.001***	0.001**	0.001***
	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0001)
race.in.group	0.023***	0.017**	0.037**	0.016	0.034***
	(0.005)	(0.007)	(0.016)	(0.015)	(0.009)
Constant	0.120***	0.109***	0.112***	0.152***	0.120***
	(0.005)	(0.008)	(0.018)	(0.017)	(0.010)
Observations	28,175	11,777	2,505	2,532	10,954
R^2	0.006	0.009	0.010	0.002	0.005
Adjusted R^2	0.006	0.009	0.009	0.001	0.005
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 15: Primary PYMK Rank Results

	Dependent variable:				
	pymk.order				
	(1)	(2)	(3)	(4)	(5)
	all.mdl	asian.mdl	black.mdl	hispanic.mdl	white.mdl
race.in.group	−0.258	−0.388	0.564	−0.237	−0.415
	(0.219)	(0.320)	(0.746)	(0.799)	(0.390)
I(100 ∗norm.pctle)	0.090***	0.084***	0.112***	0.075***	0.096***
	(0.004)	(0.006)	(0.013)	(0.014)	(0.006)
Constant	−33.698***	−33.359***	−35.319***	−33.202***	−33.792***
	(0.248)	(0.358)	(0.839)	(0.855)	(0.438)
Observations	25,111	11,405	2,085	1,920	9,401
R^2	0.022	0.020	0.033	0.015	0.026
Adjusted R^2	0.022	0.020	0.033	0.014	0.026
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 16: Primary PYMK Top 10 Results

	Dependent variable:				
	top10				
	(1)	(2)	(3)	(4)	(5)
	all.mdl	asian.mdl	black.mdl	hispanic.mdl	white.mdl
I(100 ∗norm.pctle)	0.002***	0.002***	0.002***	0.001***	0.002***
	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0001)
race.in.group	−0.006	−0.014*	−0.007	0.009	0.001
	(0.005)	(0.007)	(0.017)	(0.018)	(0.009)
Constant	0.101***	0.107***	0.094***	0.107***	0.095***
	(0.006)	(0.008)	(0.019)	(0.019)	(0.010)
Observations	25,111	11,405	2,085	1,920	9,401
R^2	0.017	0.017	0.018	0.011	0.018
Adjusted R^2	0.016	0.016	0.017	0.010	0.018
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 17: Secondary NF Rank Results (Robustness for nonlinear in preference)

	Dependent variable:				
	nf.order				
	(1)	(2)	(3)	(4)	(5)
	all.mdl	asian.mdl	black.mdl	hispanic.mdl	white.mdl
race.in.group	1.296***	1.077***	0.879	0.804	1.973***
	(0.210)	(0.319)	(0.682)	(0.694)	(0.383)
factor(norm.decile)0.2	0.166	0.536	2.575*	−2.033	0.024
	(0.460)	(0.718)	(1.464)	(1.582)	(0.734)
factor(norm.decile)0.3	0.439	0.179	2.237	−2.046	1.425**
	(0.458)	(0.721)	(1.670)	(1.608)	(0.716)
factor(norm.decile)0.4	0.783*	1.163	0.711	−2.068	1.075
	(0.459)	(0.725)	(1.439)	(1.457)	(0.749)
factor(norm.decile)0.5	1.767***	2.411***	2.661*	0.011	1.507**
	(0.459)	(0.723)	(1.573)	(1.545)	(0.722)
factor(norm.decile)0.6	2.427***	3.333***	2.972**	−1.647	2.328***
	(0.459)	(0.730)	(1.410)	(1.503)	(0.735)
factor(norm.decile)0.7	2.465***	2.752***	3.545***	−1.399	2.720***
	(0.458)	(0.736)	(1.347)	(1.496)	(0.734)
factor(norm.decile)0.8	2.732***	3.467***	4.324***	0.723	1.984***
	(0.461)	(0.750)	(1.389)	(1.621)	(0.711)
factor(norm.decile)0.9	3.486***	4.630***	0.926	1.032	3.638***
	(0.460)	(0.729)	(1.409)	(1.495)	(0.738)
factor(norm.decile)1	6.200***	7.426***	5.778***	3.921**	5.616***
	(0.460)	(0.721)	(1.544)	(1.553)	(0.731)
Constant	−31.067***	−31.786***	−30.289***	−27.987***	−31.523***
	(0.349)	(0.551)	(1.027)	(1.135)	(0.591)
Observations	28,175	11,777	2,505	2,532	10,954
R^2	0.012	0.017	0.010	0.011	0.011
Adjusted R^2	0.012	0.016	0.007	0.007	0.010
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 18: Secondary NF Rank Results (Robustness for white vs not white)

	Dependent variable:				
	nf.order				
	(1)	(2)	(3)	(4)	(5)
	all.mdl	asian.mdl	black.mdl	hispanic.mdl	white.mdl
same.race.alt	0.887***	0.303	0.396	0.563	1.914***
	(0.222)	(0.331)	(0.736)	(0.734)	(0.380)
I(100 ∗norm.pctle)	0.059***	0.073***	0.039***	0.040***	0.051***
	(0.004)	(0.006)	(0.012)	(0.012)	(0.006)
Constant	−31.786***	−32.461***	−29.558***	−30.380***	−32.008***
	(0.255)	(0.379)	(0.833)	(0.859)	(0.435)
Observations	28,175	11,777	2,505	2,532	10,954
R^2	0.010	0.014	0.005	0.005	0.010
Adjusted R^2	0.010	0.014	0.004	0.004	0.009
Note:					*p<0.1; **p<0.05; ***p<0.01