

Supplementary Material for “Modeling competitive and complementary ecological relationships between online communities”

ACM Reference Format:

. 2018. Supplementary Material for “Modeling competitive and complementary ecological relationships between online communities”. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/1122445.1122456>

1 CLUSTERING SELECTION

Here we provide additional details on our cluster selection procedure. We test three different algorithms: k -means, HDBSCAN, and affinity-propagation clustering. The table shows the set of parameters that we began testing for each algorithm. We set out to find a clustering that would have the best possible silhouette score given our choice of hyper parameters. For Affinity propagation and k -means we use the implementations in `sklearn`. For HDBSCAN we use the `hdbscan` python package and set the `minimum_cluster_size` parameter to 2 so that clusters of size 2 can be detected.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.

Manuscript submitted to ACM

Algorithm	Parameter	Values
All	LSI dimensions k	10,50,100,200,300,400,500,600,700,850
HDBSCAN	minimum_cluster_size	2
HDBSCAN	min_samples	2,3,4,5
HDBSCAN	cluster_selection_epsilon	0,0.01,0.05,0.1,0.15,0.2
HDBSCAN	cluster_selection_method	eom
Affinity Propagation	damping	0.5,0.6,0.7,0.8,0.95,0.97,0.99
Affinity Propagation	preference_quantile	0.1,0.3,0.5,0.7,0.9
Affinity Propagation	convergence_iters	15
k -means	max_iters	3000
k -means	n_inits	3000
k -means	n_clusters	100,500,1000,1250,1750,2000

2 FORECAST PLOTS

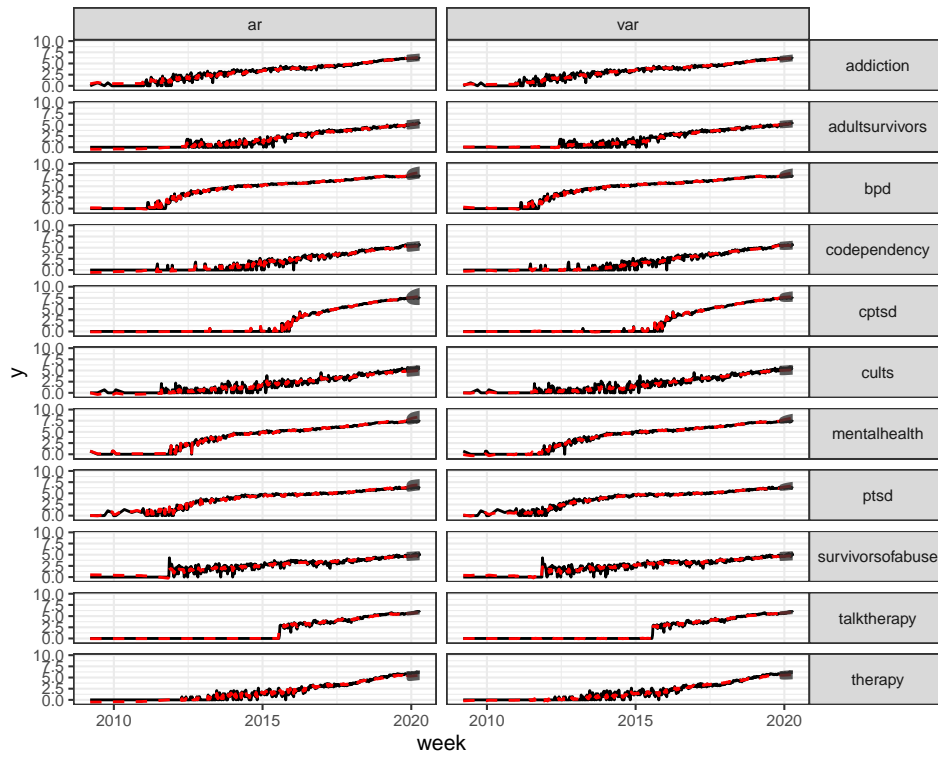


Fig. 1. Model fit and forecast from VAR model estimated on cluster of mental health subreddits

Woodstock '18, June 03–05, 2018, Woodstock, NY

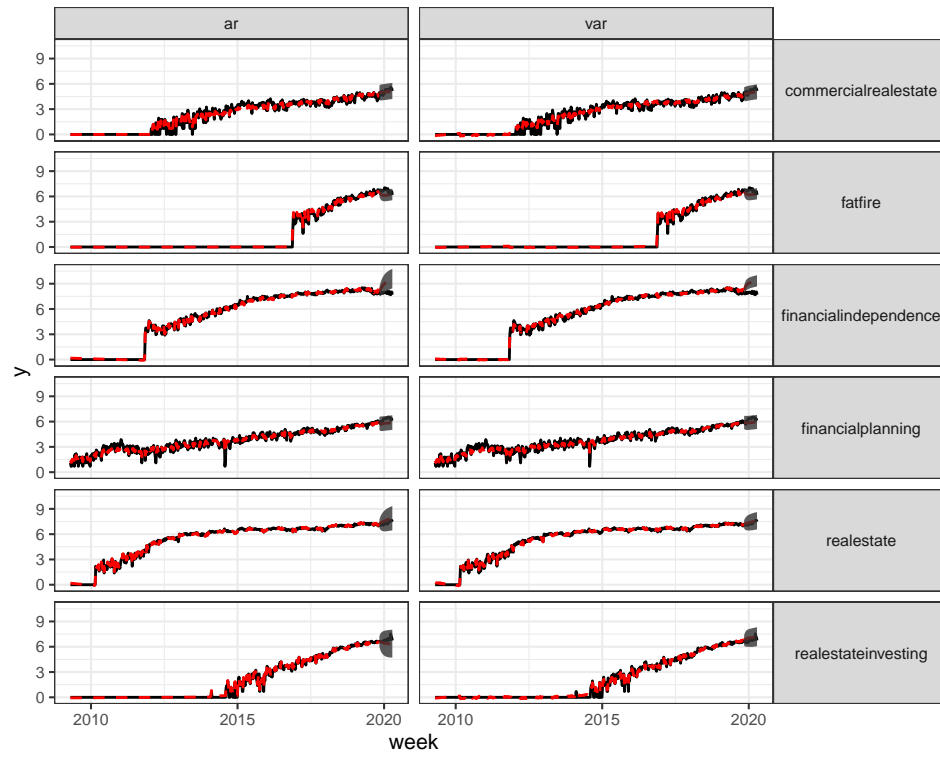


Fig. 2. Model fit and forecast from VAR model estimated on cluster of real estate and finance subreddits

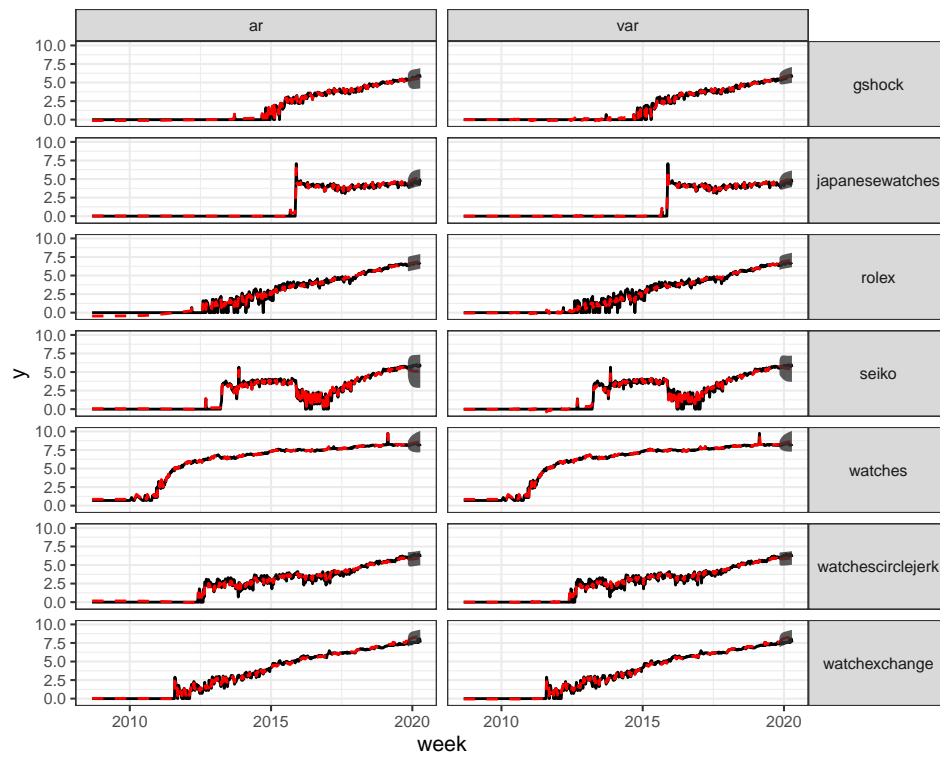


Fig. 3. Model fit and forecast from VAR model estimated on cluster of timepiece subreddits

Woodstock '18, June 03–05, 2018, Woodstock, NY

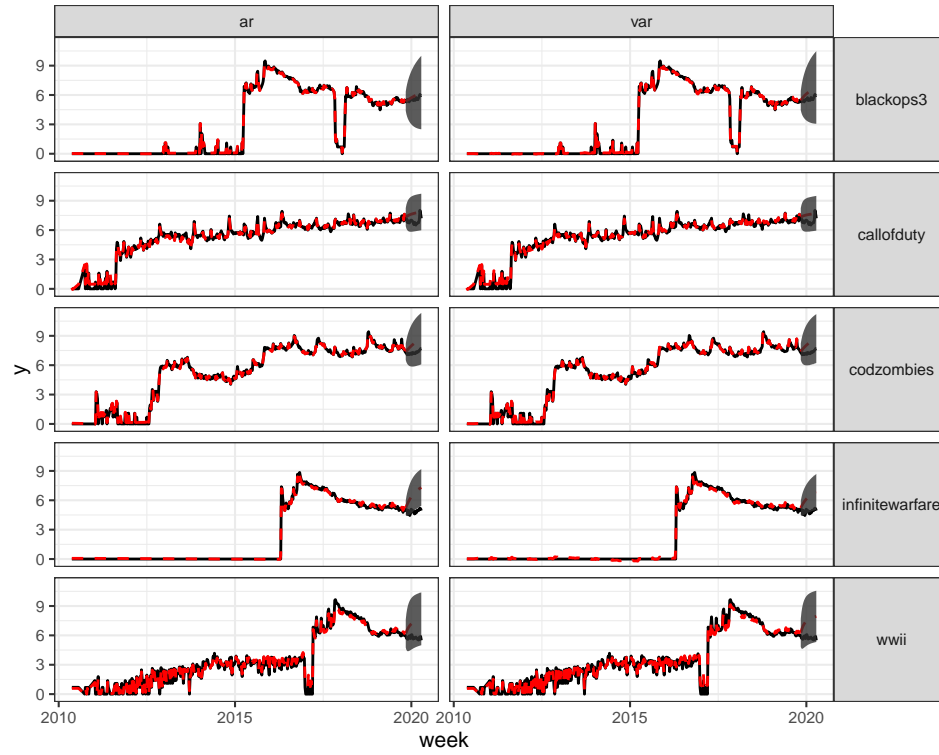


Fig. 4. Model fit and forecast from VAR model estimated on cluster of call-of-duty subreddits

3 VAR MODEL COEFFICIENTS

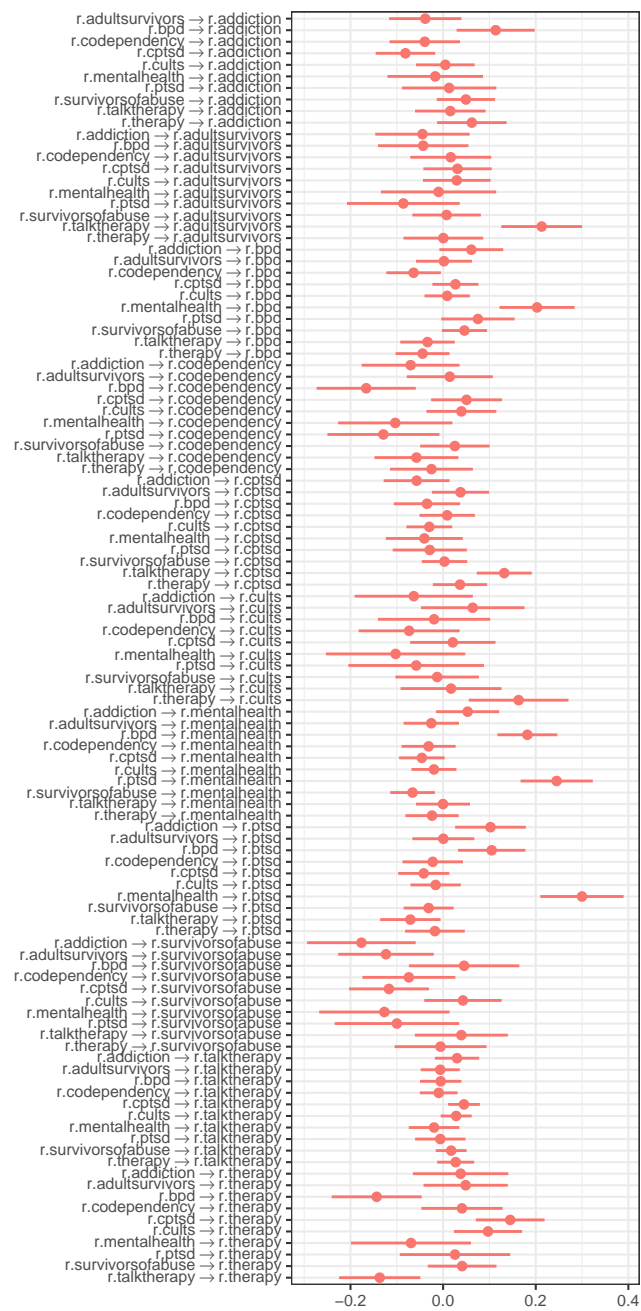


Fig. 5. Commensal coefficients from VAR model estimated on cluster of mental health subreddits.

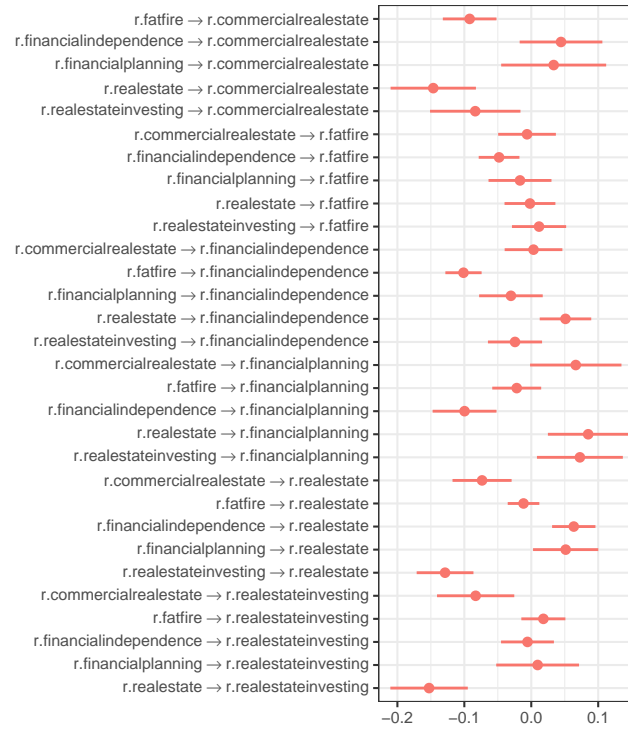


Fig. 6. Commensal coefficients from VAR model estimated on cluster of real estate and finance subreddits.

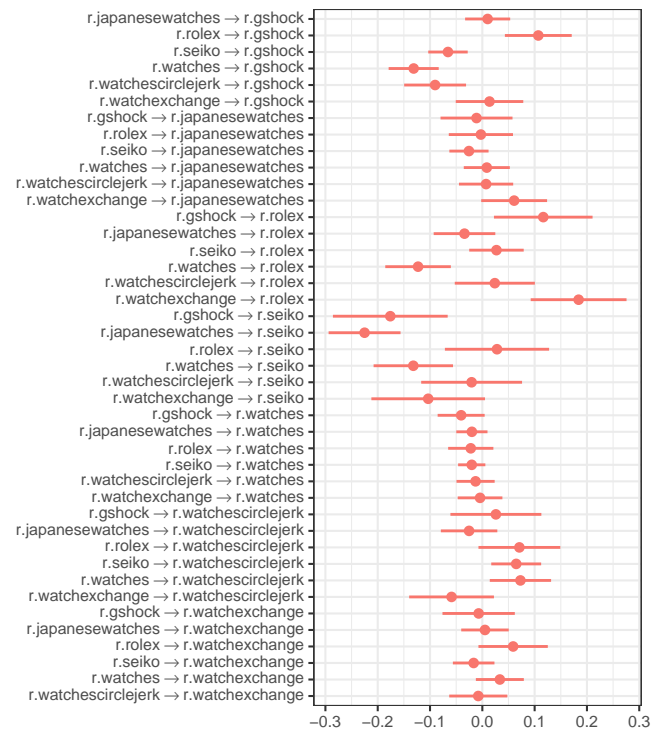


Fig. 7. Commensal coefficients from VAR model estimated on cluster of timepiece subreddits.

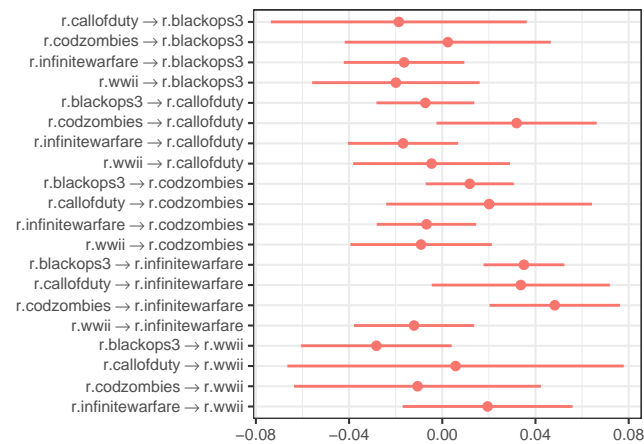


Fig. 8. Commensal coefficients from VAR model estimated on cluster of art call of duty subreddits.

Woodstock '18, June 03–05, 2018, Woodstock, NY

4 IMPULSE RESPONSE FUNCTIONS

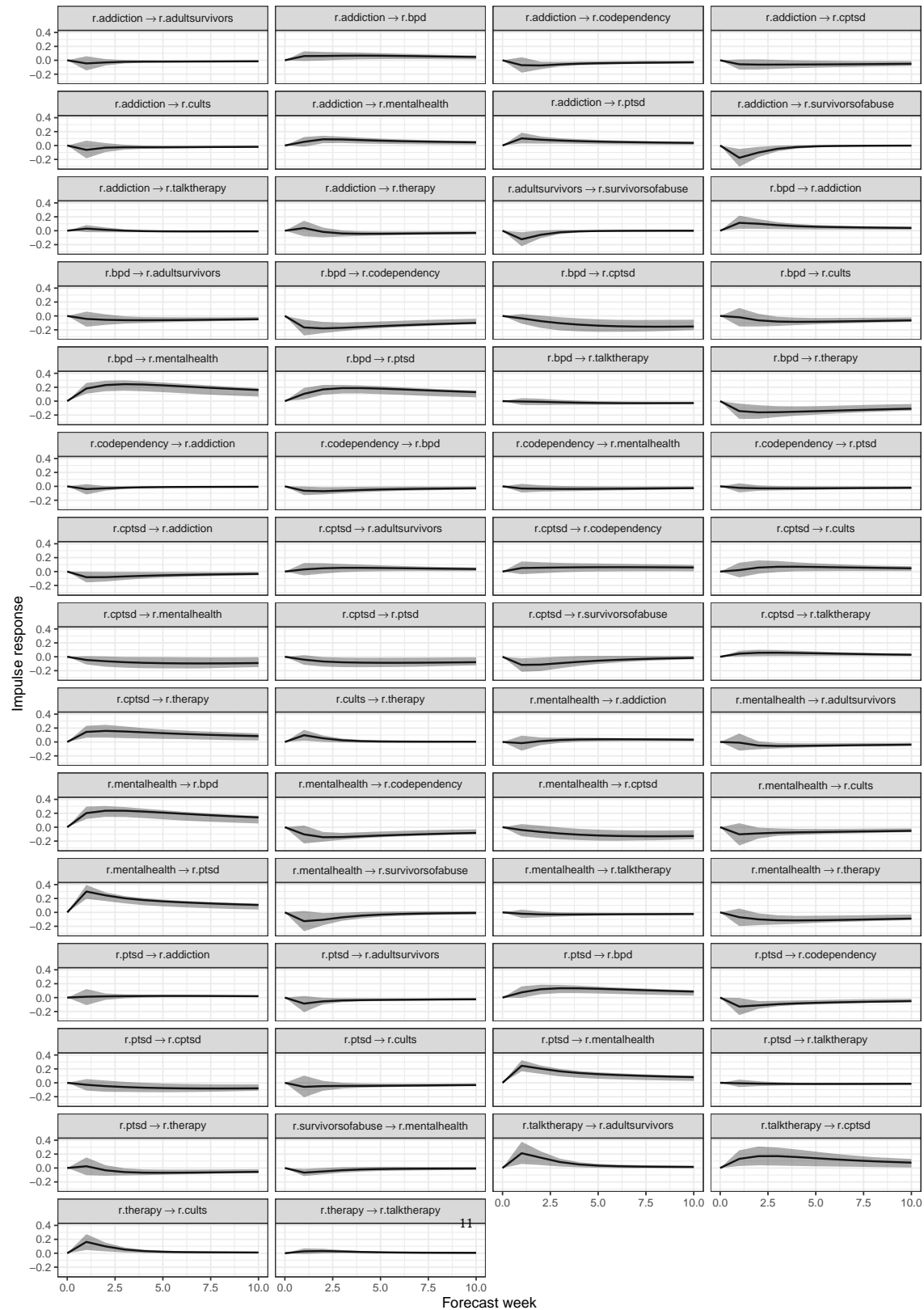


Fig. 9. Impulse response functions from VAR model estimated on cluster of mental health subreddits. Only impulses where the 95% confidence interval of the impulse response function does not include 0 are shown.

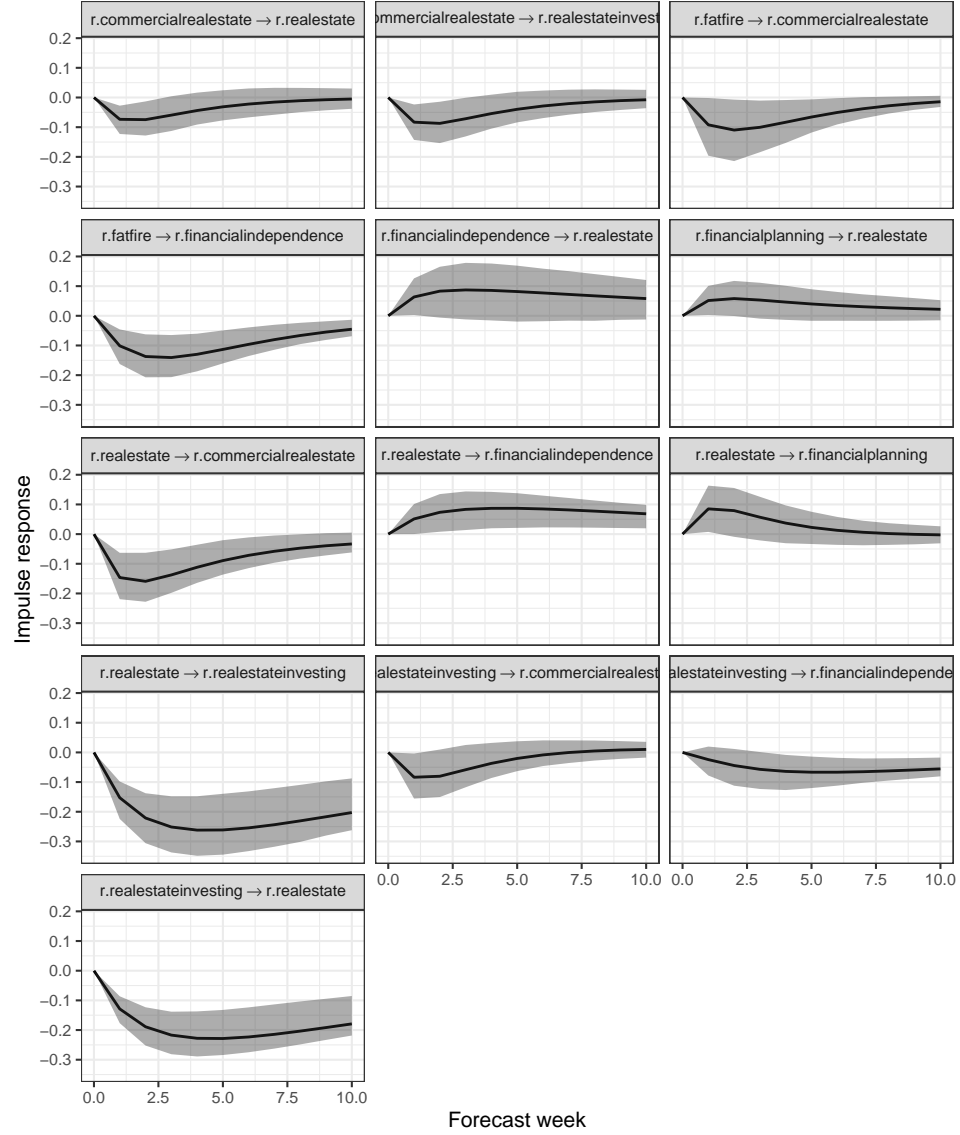


Fig. 10. Impulse response functions from VAR model estimated on cluster of real estate and finance subreddits. Only impulses where the 95% confidence interval of the impulse response function does not include 0 are shown.

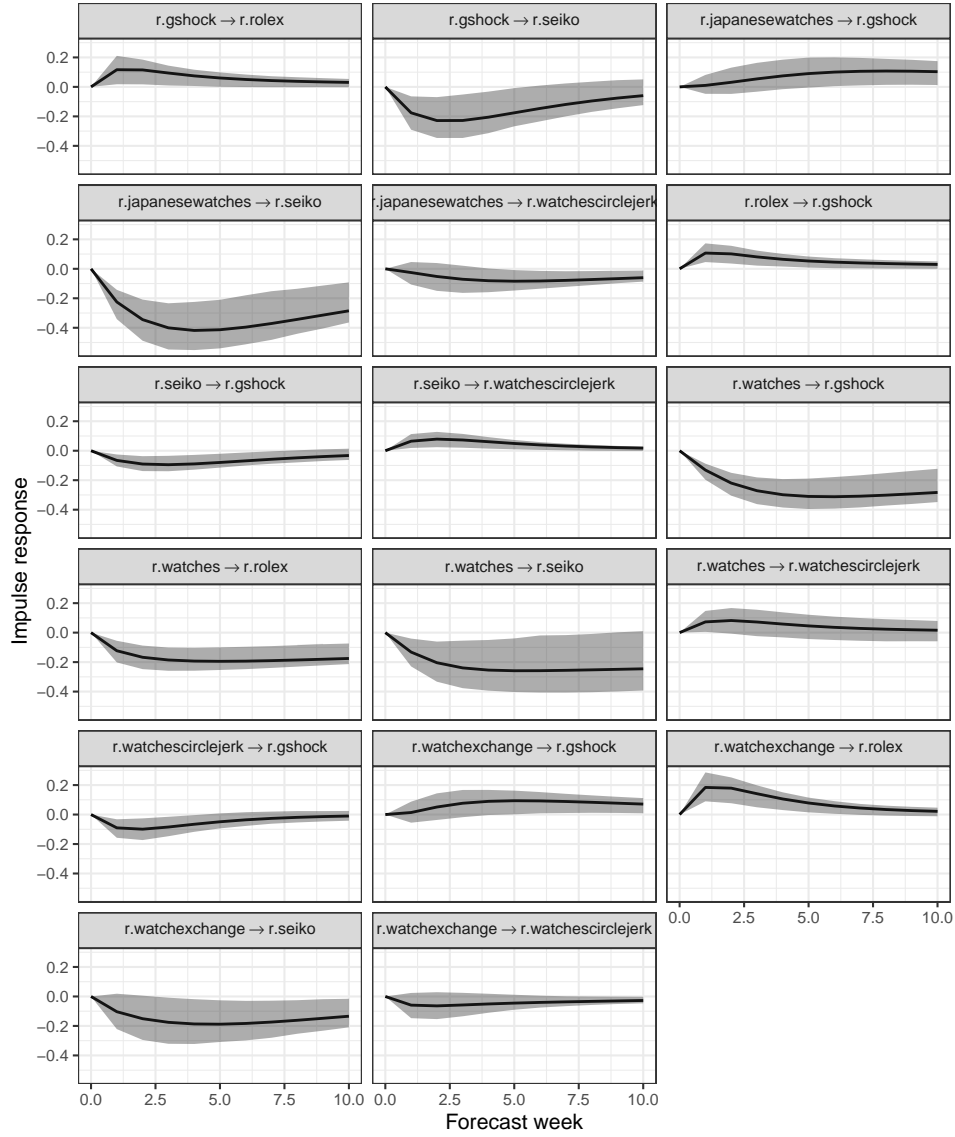


Fig. 11. Impulse response functions from VAR model estimated on cluster of timepiece subreddits. Only impulses where the 95% confidence interval of the impulse response function does not include 0 are shown.

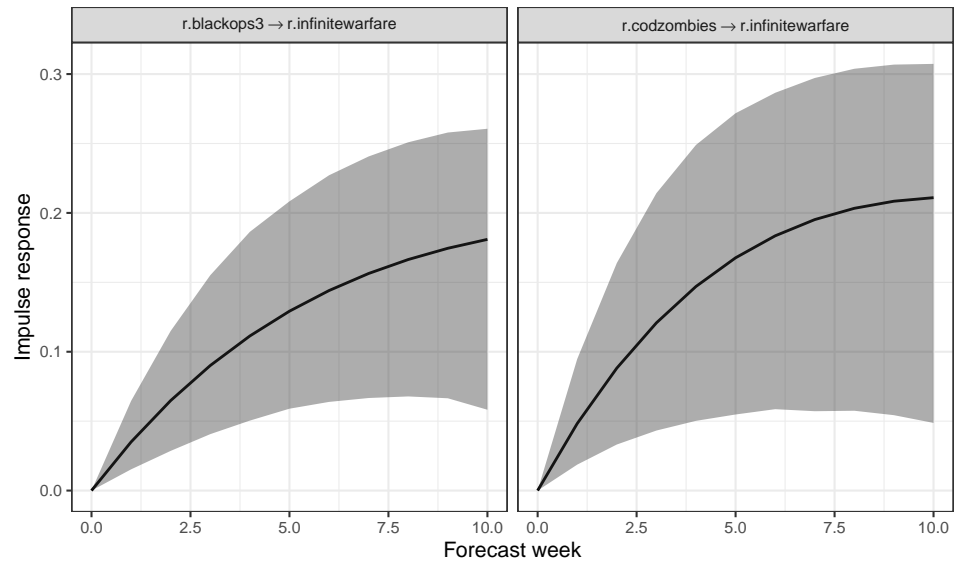


Fig. 12. Impulse response functions from VAR model estimated on cluster of call of duty subreddits. Only impulses where the 95% confidence interval of the impulse response function does not include 0 are shown.