

# Chapter 48 Outline

## Network Modeling: Estimation, Inference, Comparison, and Selection

John P. Schoeneman<sup>1</sup> and Bruce A. Desmarais<sup>1</sup>

<sup>1</sup>Penn State, Political Science, University Park, Pond Lab

**Introduction:** Most statistical models used in the social sciences rely on the assumption that the observations in a dataset (or at least clusters of observation) are drawn independently from a common data generating process. However, when analyzing networks it is often inappropriate to assume that observations are independent (Desmarais and Cranmer, 2017). Indeed, the network analyst is often interested in studying the ways in which observations depend upon each other (e.g., do friends influence each others' choices to vote (Bond et al., 2012), do legislators reciprocate support for legislation (Kirkland and Williams, 2014)).

**Overview:** In this chapter we will review three of the most popular statistical models that have been developed to analyze network data, along with their extensions. These include the exponential random graph model (ERGM) (Cranmer and Desmarais, 2011), the latent space model (LSM) (Dorff, Minhas and Ward, 2016), and the stochastic block model (SBM) (Latouche, Birmelé and Ambroise, 2011). We will review the application of each method to networks in which ties are binary (e.g., are two legislators on at least one committee together), count (e.g., on how many committees do two legislators both serve), and continuous (e.g., how much time do two legislators spend together in committee meetings each week). The complexity and analytical intractability of the mathematical forms of these models render exact approaches to estimation infeasible. As such, estimation in these models proceeds by sampling from a Bayesian posterior and/or approximate methods of maximum likelihood (Nowicki and Snijders, 2001; Raftery et al., 2012; Van Duijn, Gile and Handcock, 2009). We will review these estimation methods, and how they are applied to the ERGM, LSM, and SBM. One complication that arises from the variety of sampling-based and approximate approaches to inference with network models is that standard information criteria that would normally be used to select between different families of statistical models (e.g., AIC, BIC, DIC) cannot be calculated in a manner that is comparable across different families of network models. Recently, researchers have considered various methods of out-of-sample prediction in order to compare and select among network models in terms of generalization error (e.g., Desmarais and Cranmer, 2012; Dorff, Minhas and Ward, 2016), with some debate regarding whether this is an appropriate approach (Block et al., 2018). We will conduct a simulation study to evaluate the effectiveness of comparing the performance of models in terms of generalization error. We will consider and compare several measures of generalization error, and provide recommendations regarding which measures should be used with specific types of networks and models. This simulation study will be, to our knowledge, the first experiment designed to evaluate approaches to comparing network models based on generalization error.

**Organization:** This chapter will be divided into three sections. In the first section we will give an overview of the most common models used in network analysis. In the second section we will discuss the approaches to estimation that are commonly used with these models. In the last section we will conduct a simulation study with the goal of evaluating the effectiveness of measuring generalization error to assist practitioners in selecting among these models.

## 1 Introduction

Statistical inference with network data permits researchers to test hypotheses about network generation, make predictions regarding network structure, and simulate networks from probabilistic models.

## 2 Model description

LSM is common in polisci

ERGM is common in polisci

SBM is common elsewhere, and should be used more in polisci

## 3 Model comparison and selection

Issue is that we need Bayes for both SBM and LSM, which makes AIC/BIC difficult to calibrate. Would prefer to use held out likelihood.

Challenge with ERGM is that joint likelihood cannot really be evaluated, unlike with LSM or SBM  
[Wang et al. \(2016\)](#) provide an excellent example of doing both multiple imputation and held out edge prediction to evaluate ERGM predictive performance.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4743534/>

## 4 Replication

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Edges	−0.87*** (0.17)	−5.15*** (1.23)	−2.50*** (0.30)	−6.00*** (0.76)	−5.12*** (1.01)	−5.12*** (1.13)	−4.58*** (0.56)	−6.43*** (0.80)
Preference Similarity	0.18** (0.06)	0.11 (0.07)	−0.02 (0.11)	−0.02 (0.12)	0.09 (0.07)	0.11 (0.07)	−0.02 (0.11)	0.02 (0.13)
Government Alter		0.60** (0.19)		0.37 (0.33)	0.61** (0.19)	0.58** (0.19)	0.67* (0.30)	0.51 (0.34)
Scientific Ego		0.06 (0.22)		1.54*** (0.36)	0.35 (0.20)	0.07 (0.22)	1.59*** (0.34)	1.58*** (0.37)
Common Committee		0.30*** (0.05)		0.14* (0.06)	0.34*** (0.05)	0.29*** (0.05)	0.21*** (0.05)	0.16** (0.06)
Scientific Communication		2.88*** (0.60)				2.86*** (0.63)		
Political Communication				2.89*** (0.58)				2.94*** (0.59)
Interest Group Homophily		1.06*** (0.29)		1.22* (0.59)	1.10*** (0.28)	1.04*** (0.30)	1.55** (0.59)	1.28 (0.68)
Influence Attribution		0.94*** (0.18)		0.32 (0.31)	0.92*** (0.18)	0.88*** (0.20)	0.56 (0.30)	0.26 (0.32)
gwesp.fixed.0.1		2.52* (1.06)		0.49* (0.23)	2.48** (0.86)	2.47* (0.97)	0.61* (0.24)	0.34 (0.23)
gwdsp.fixed.0.1		−0.13** (0.04)		−0.17 (0.10)	−0.13** (0.04)	−0.13** (0.05)	−0.13 (0.09)	−0.22* (0.09)
Reciprocity		0.81** (0.25)		1.76*** (0.50)	0.86*** (0.25)	0.82** (0.25)	1.85*** (0.52)	1.69*** (0.51)
nodeicov.betweenness						0.00 (0.00)		0.02*** (0.00)
AIC	1161.59	849.59	456.09	313.09	887.52	850.39	352.91	303.07
BIC	1171.13	902.04	465.62	365.55	935.21	907.61	400.60	360.29
Log Likelihood	−578.80	−413.80	−226.04	−145.55	−433.76	−413.19	−166.46	−139.54

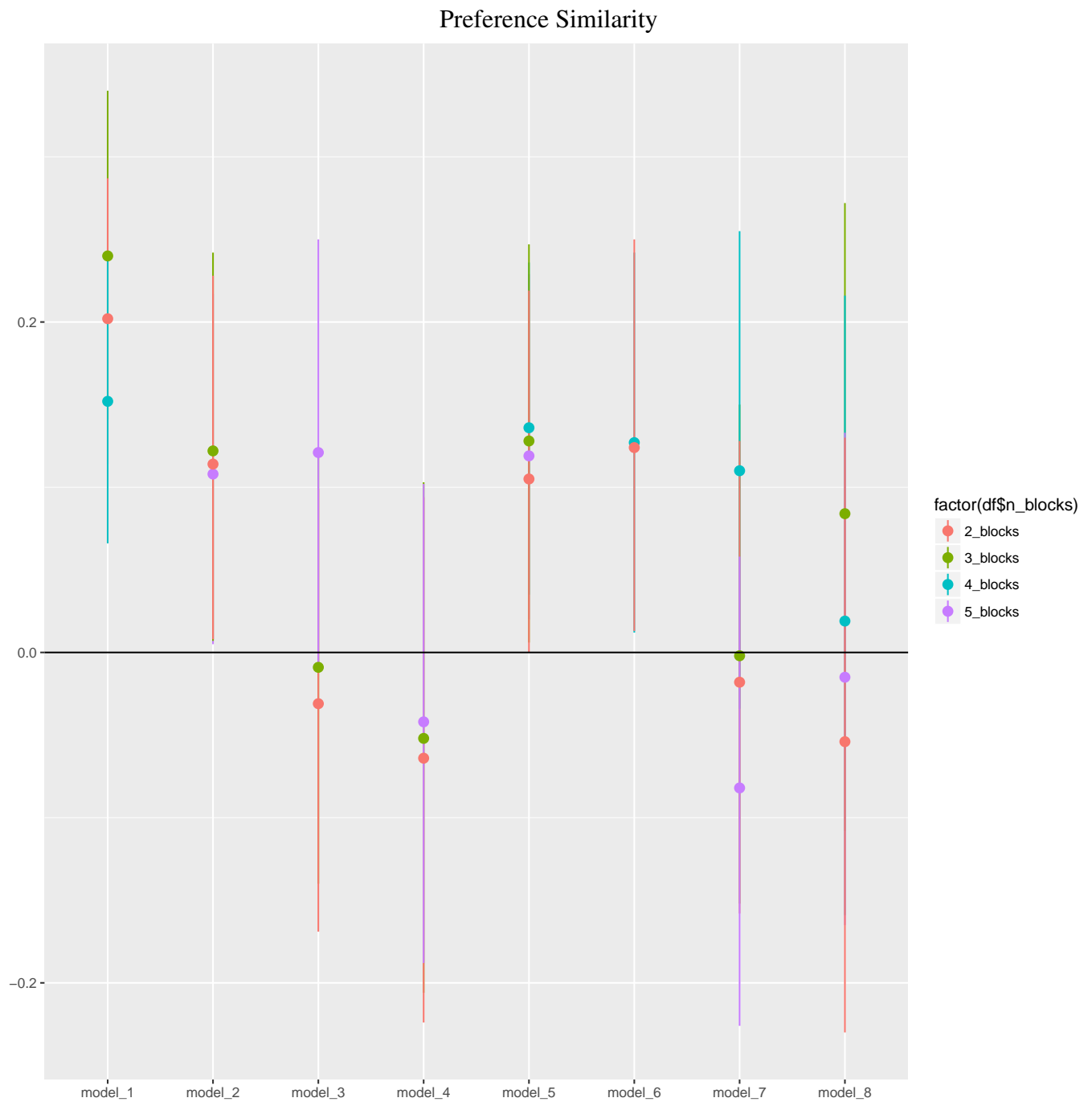
\*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$

**Table 1.** ERGM Replication

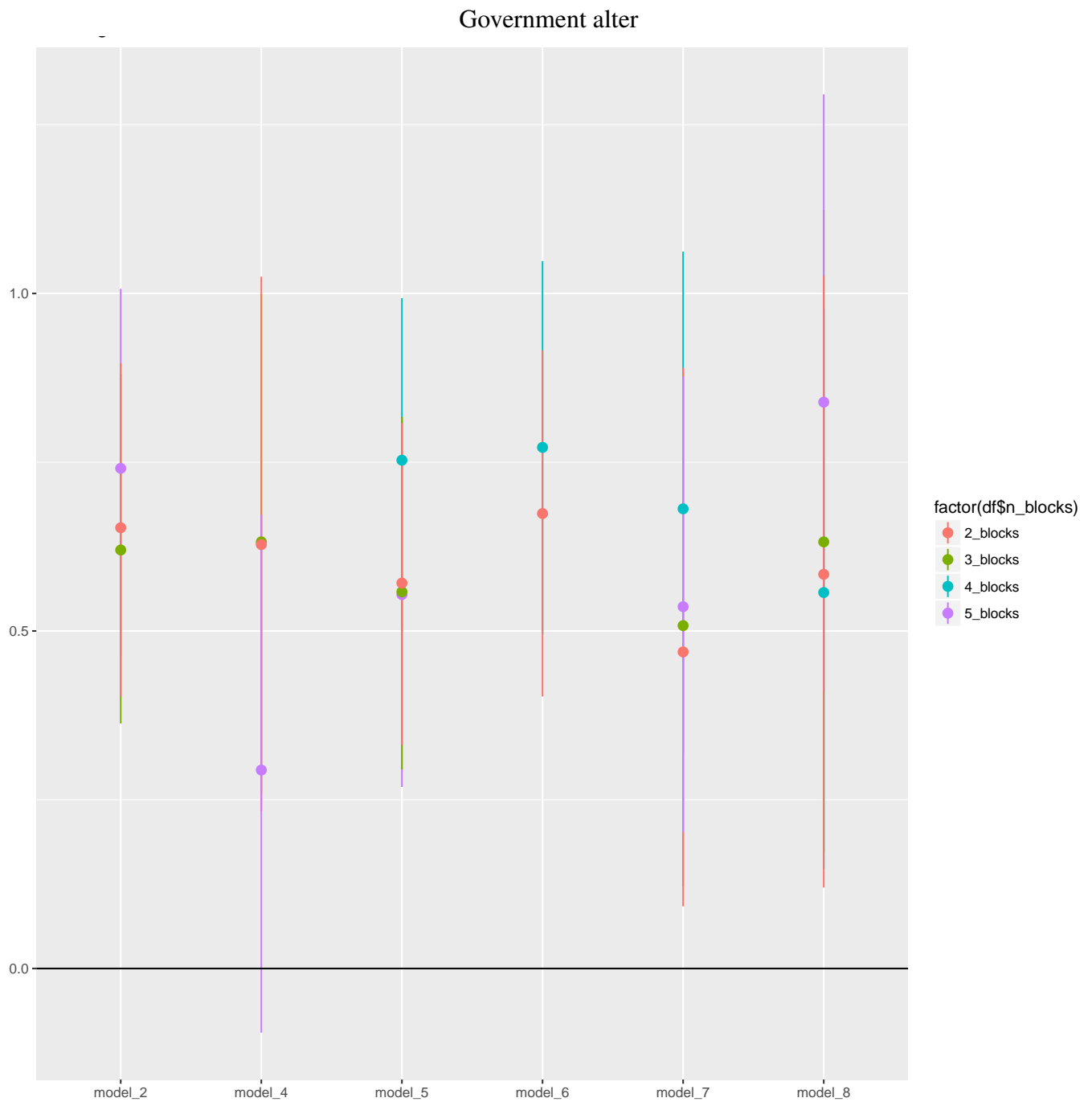
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
edges	2.30 [−0.44; 4.98]	1.45 [−1.42; 4.07]	−0.07 [−4.04; 3.66]	−2.40 [−6.31; 1.36]	1.54 [−1.29; 4.18]	2.13 [−1.23; 5.17]	−1.80 [−6.13; 2.09]	−3.10 [−6.90; 0.70]
Preference Similarity	0.20 [−2.40; 3.03]	−1.59 [−4.22; 1.21]	0.45 [−3.32; 4.44]	−1.77 [−5.33; 2.12]	−1.59 [−4.24; 1.26]	−2.30 [−5.14; 1.09]	−1.44 [−5.45; 2.91]	−0.50 [−5.33; 4.33]
Government Alter		0.66* [0.66; 0.66]		0.68* [0.42; 0.97]	0.85* [0.42; 1.30]	0.96* [0.55; 1.43]	0.82 [−0.02; 1.61]	0.70 [0.43; 0.97]
Scientific Ego		−0.01 [−0.45; 0.54]		0.95* [0.03; 1.67]	0.25 [−0.31; 0.83]	−0.26* [−0.27; −0.25]	1.33* [0.48; 2.19]	1.10 [0.88; 1.32]
Common Committe		1.59* [1.58; 1.60]		1.22 [−0.44; 2.82]	1.89* [1.36; 2.41]	1.74* [1.32; 2.35]	2.30* [0.84; 3.64]	1.10 [−0.50; 2.70]
Scientific Commuication		3.02* [1.73; 4.74]				3.35* [2.09; 4.43]		
Political Commuication				2.74* [1.47; 4.11]				3.10 [2.40; 3.80]
Interest Group Homophily		1.91* [1.86; 1.95]		1.10 [−0.75; 2.82]	2.13* [1.07; 3.24]	1.89* [0.98; 2.65]	1.86 [−0.09; 3.88]	1.10 [1.11; 1.09]
Influence Attribution		1.40* [1.40; 1.40]		0.87* [0.10; 1.68]	1.20* [0.81; 1.63]	1.28* [0.90; 1.91]	1.27* [0.56; 2.03]	0.70 [0.22; 1.18]
nodeicov.betweenness						−0.00* [−0.00; −0.00]		0.70 [0.00; 1.40]
BIC (Overall)	1057.64	954.68	487.17	506.51	971.64	961.00	501.48	520.10
BIC (Likelihood)	833.39	726.85	316.74	275.73	737.39	732.03	307.71	280.10
BIC (Latent Positions)	224.25	227.83	170.43	230.78	234.25	228.96	193.77	240.10

\* 0 outside the confidence interval

**Table 2.** LSM models

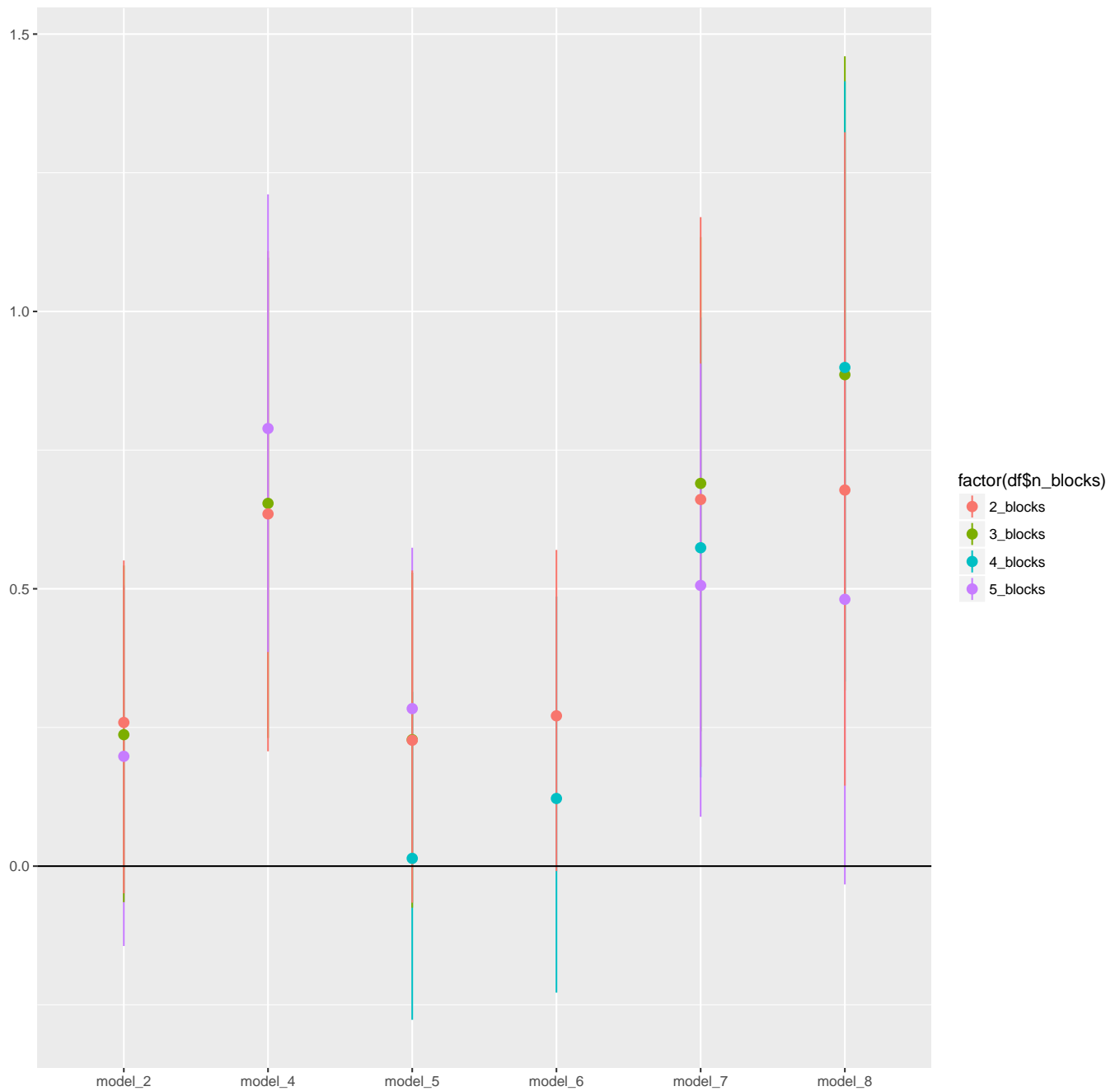


**Table 3.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.

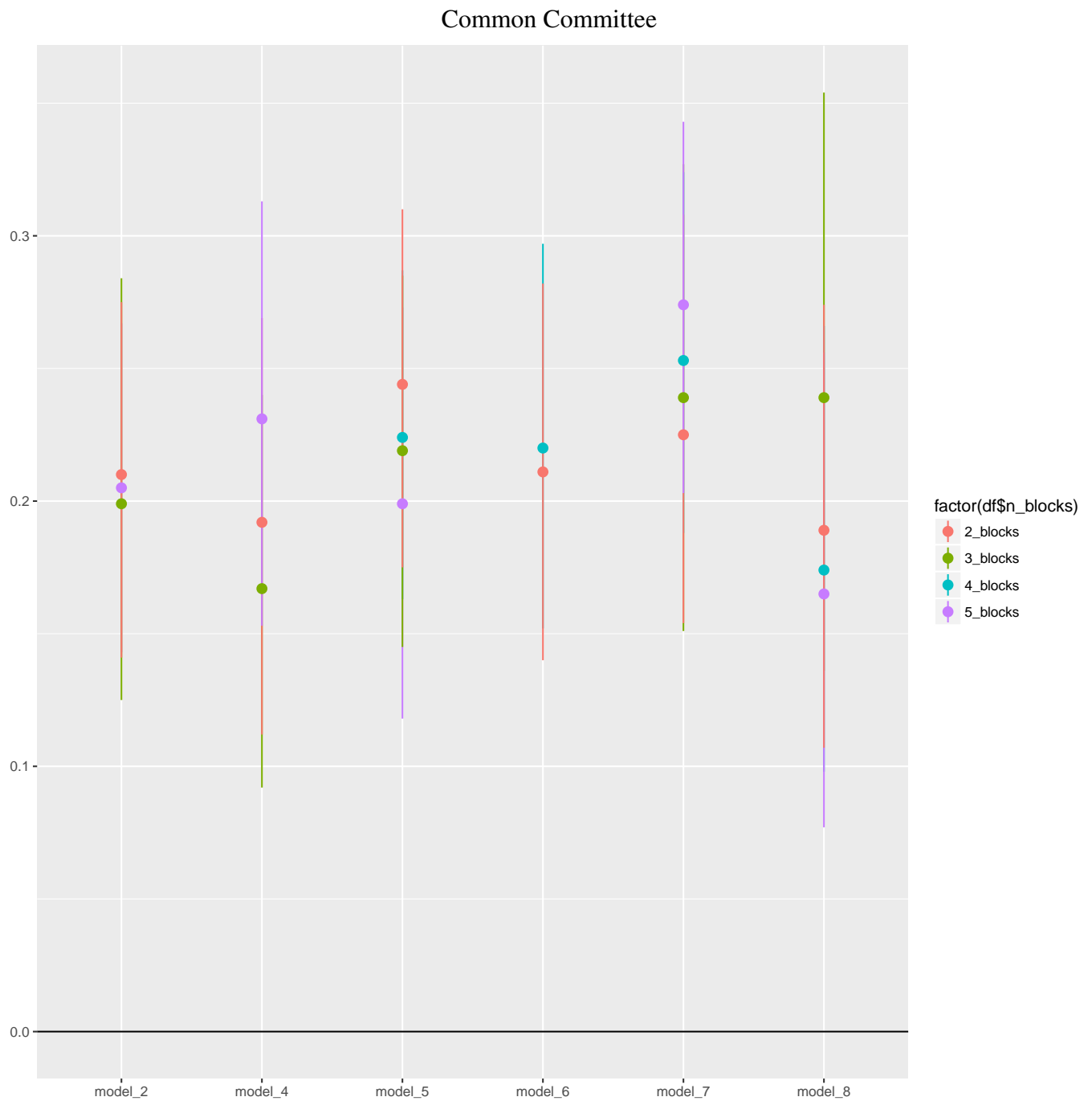


**Table 4.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.

## Scientific Ego

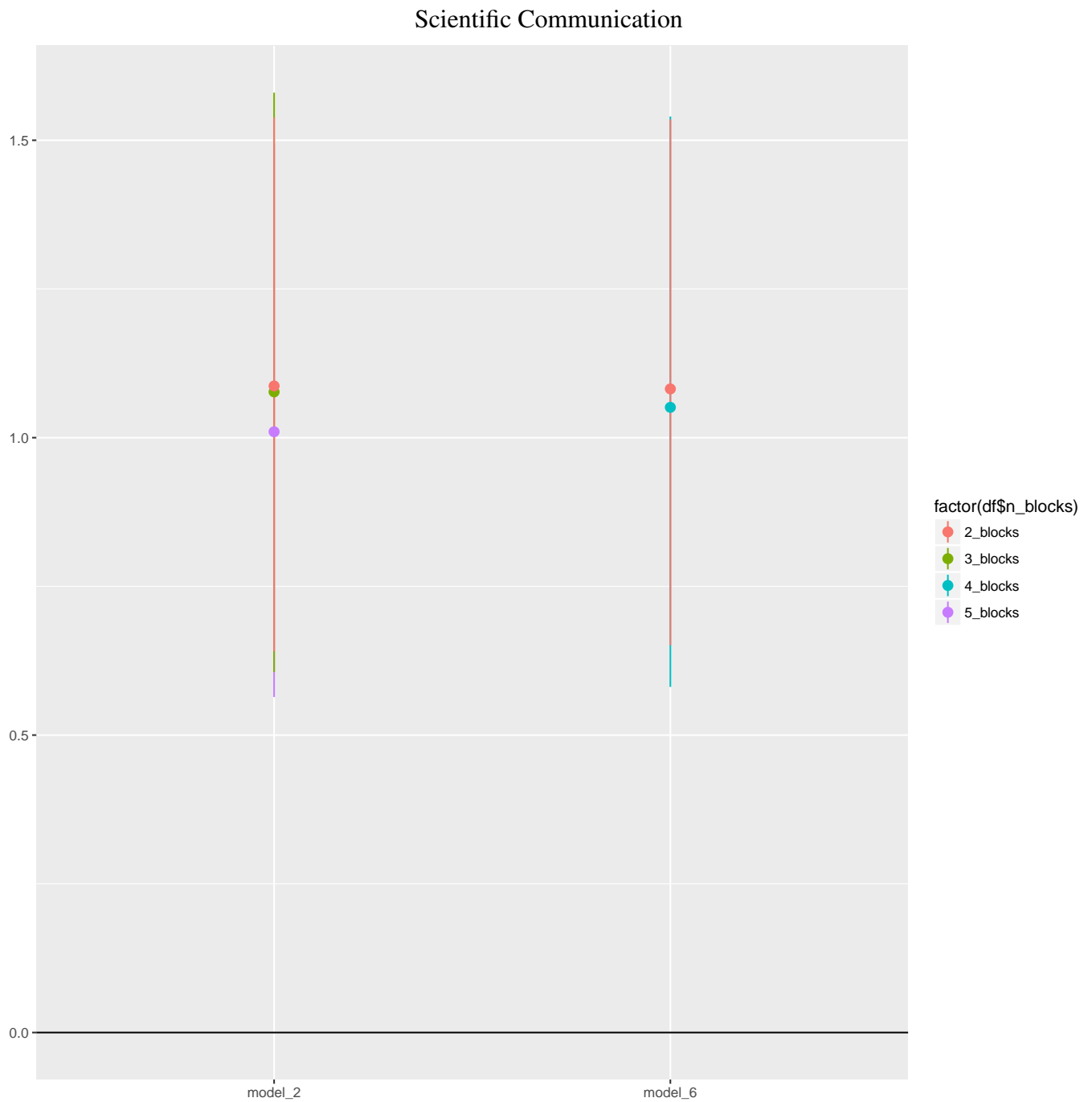


**Table 5.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.

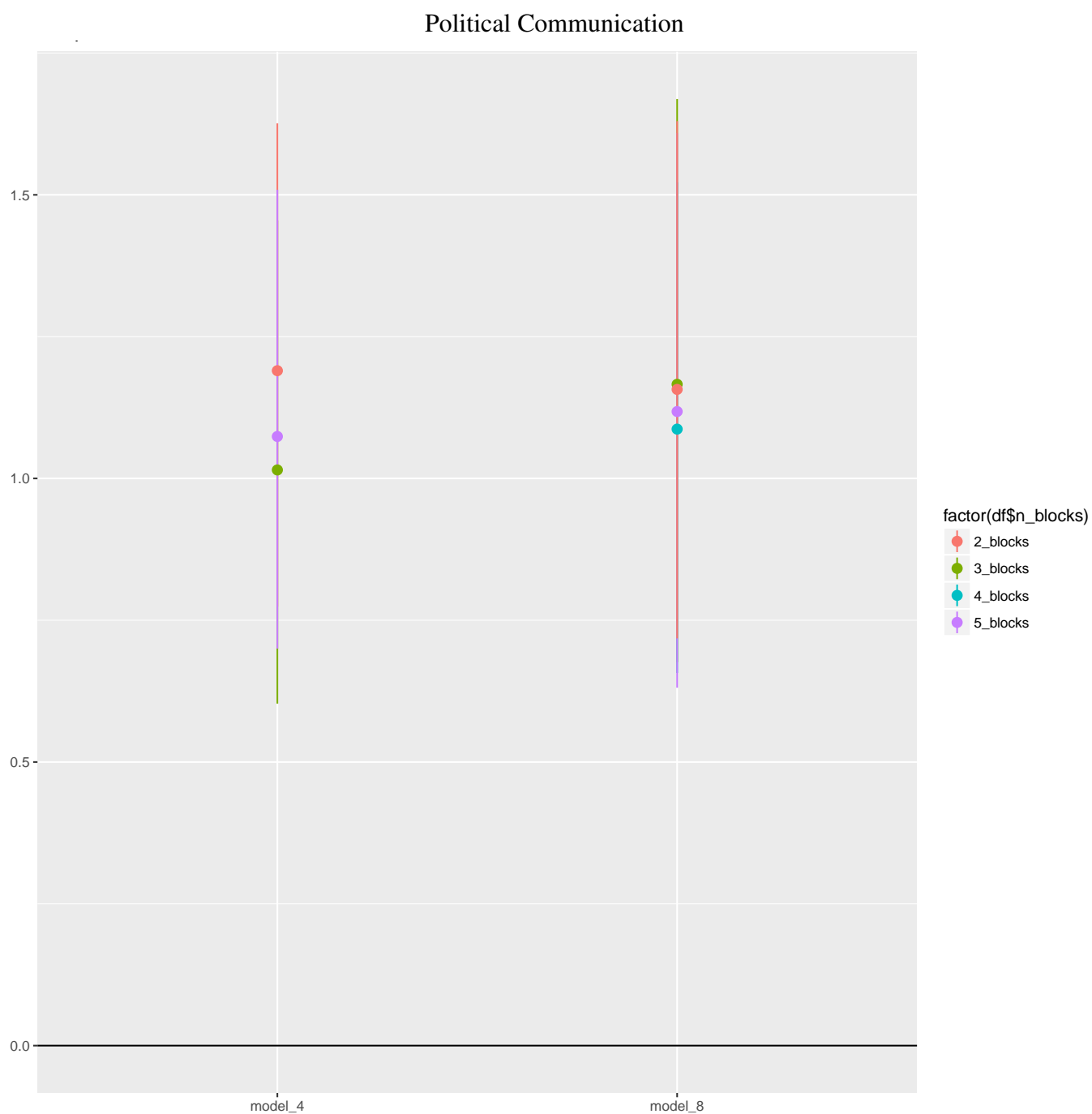


**Table 6.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.

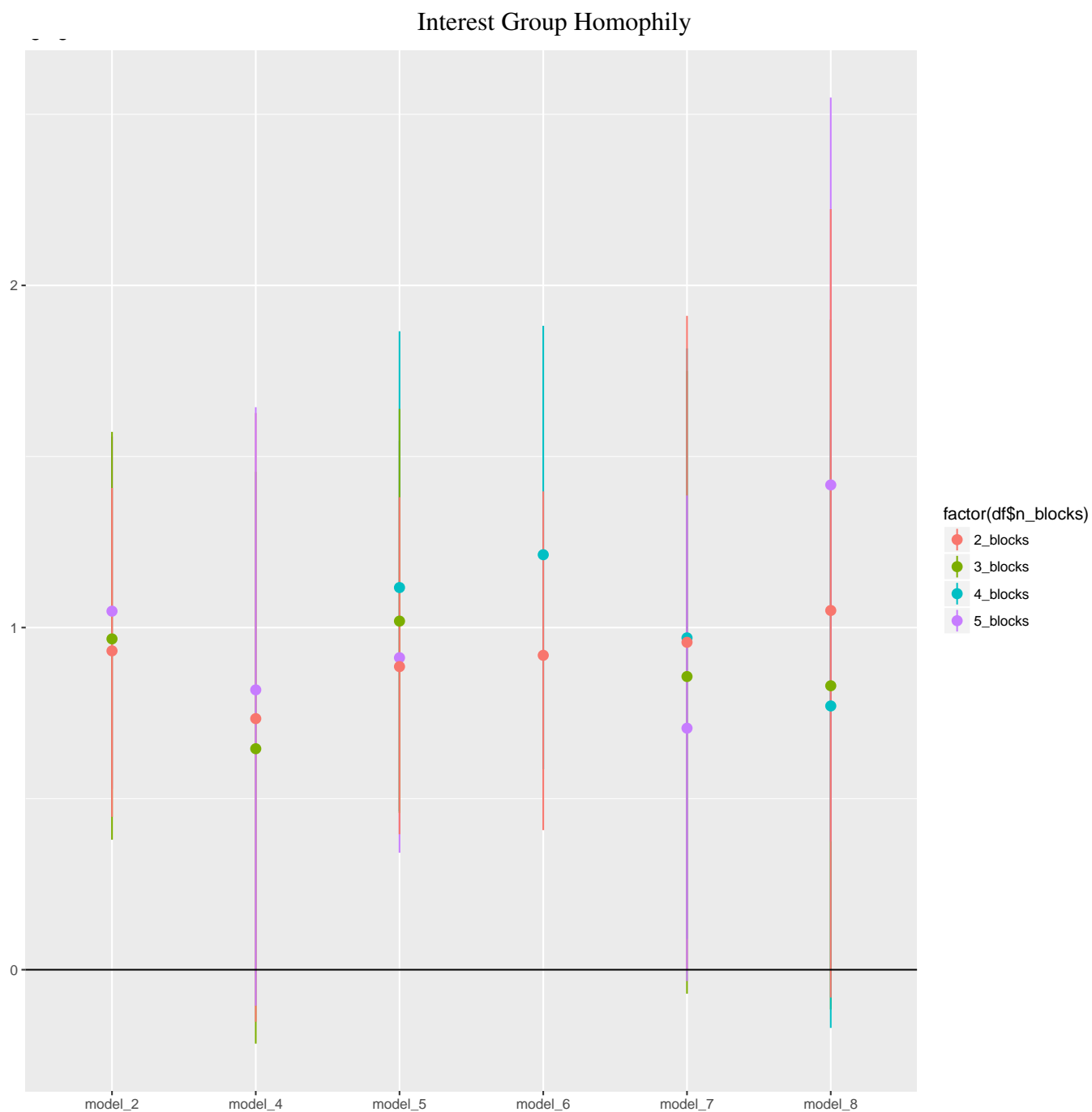




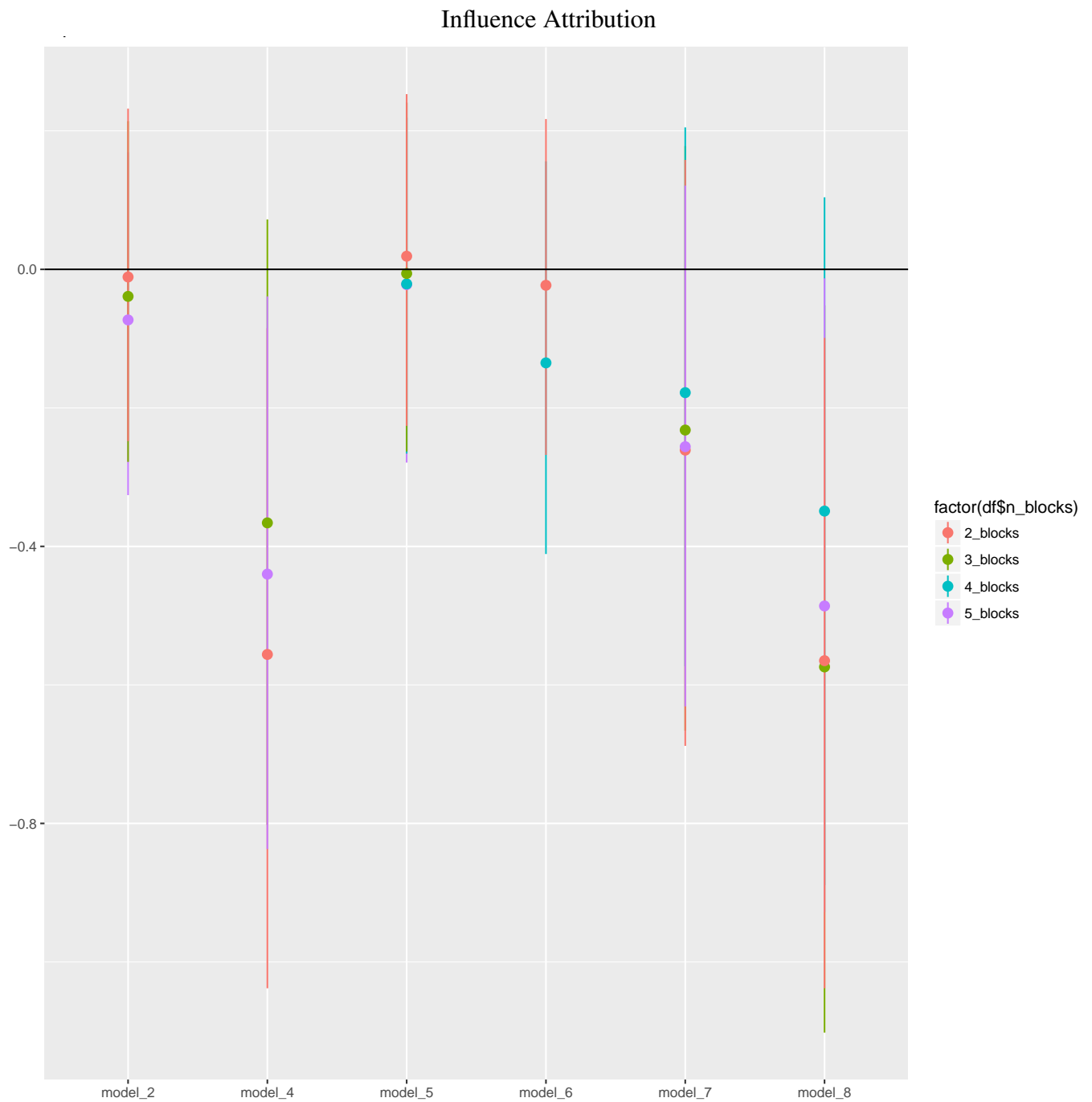
**Table 7.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.



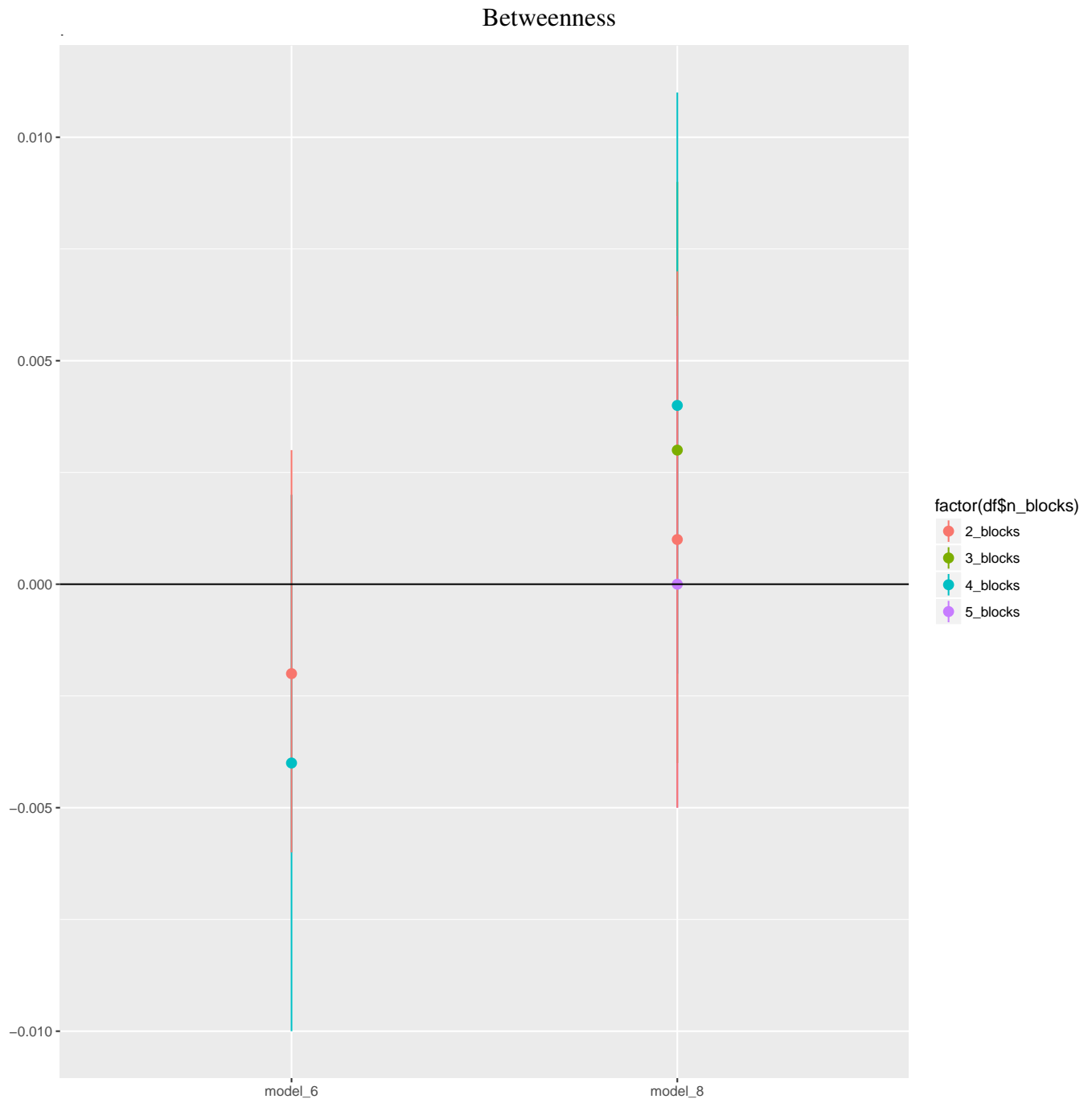
**Table 8.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.



**Table 9.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.



**Table 10.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.



**Table 11.** SBM covariate plot. Results from models for which DIC could not be calculated were discarded.

#### 4.1 Model Comparison

I only compare models one through four, as models five through eight are not reported in the main paper and are merely robustness tests. I included the results though above so that we could still review them and take them out later if we need too.

1. ERGM replication

- (a) Models two and four do not perfectly replicate. Most of the substantive findings remain, but in model four the significance level for common committee drops from 99.9% to 95%. There are a few other drops in significance for other covariates, but not as big. Signage remains the same for the replication across all models even when there are changes in the coefficient size or significance.

## 2. LSM

- (a) There are substantial differences between the ERGM results and the LSM results. For example, in model 1, LSM does not have significant correlation between preference similarity and political communication. In model 2 the point estimate for common committee is five times larger and is not significant in model 4. Influence attribution is almost two times larger in model 2 and 4. In the LSM models, intercept estimates are not significant in any model, but they are highly significant in the ERGM models. Interest group homophily is not significant in model 4.

## 3. SBM

- (a) Preference Similarity: In model 2, point estimates remain above 95% significance cutoff. The variables for model 2 do not absorb preference similarity measures as they do in the ERGM.
- (b) Government Alter: Point estimates are a little higher and in model 4 it is not 95% significant for five blocks.
- (c) Scientific Ego: Still statistically significant in model 4, but point estimates are less than half.
- (d) Common Committee: Point estimate is about two thirds for model 2's and slightly bigger than model 4's.
- (e) Scientific Communication: Point estimate in model 2 is about 1/3
- (f) Political Communication: Point estimate in model 4 is about 1/3
- (g) Interest Group Homophily: Point estimates are slightly smaller and in model 4 they are not significant at the 95% threshold.
- (h) Influence attribution: point estimates are negative and not statistically significant instead of being around one.
- (i) The number of blocks that results in the lowest DIC is not consistent across models. For the most part, models with four blocks have the lowest DIC for political communication but two blocks is best for scientific communication. Many of the institutional/relational structure variables are smaller not always significant at the same level.

## References

- Block, Per, Johan Koskinen, James Hollway, Christian Steglich and Christoph Stadtfeld. 2018. "Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power." *Social Networks* 52:180–191.
- Bond, Robert M, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle and James H Fowler. 2012. "A 61-million-person experiment in social influence and political mobilization." *Nature* 489(7415):295.

- Cranmer, Skyler J and Bruce A Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19(1):66–86.
- Desmarais, Bruce A and Skyler J Cranmer. 2012. "Micro-level interpretation of exponential random graph models with application to estuary networks." *Policy Studies Journal* 40(3):402–434.
- Desmarais, Bruce A and Skyler J Cranmer. 2017. "Statistical Inference in Political Networks Research." *The Oxford Handbook of Political Networks* p. 203.
- Dorff, Cassy, Shahryar Minhas and Michael D Ward. 2016. Latent Networks and Spatial Networks in Politics. In *The Oxford Handbook of Political Networks*.
- Kirkland, Justin H and R Lucas Williams. 2014. "Partisanship and reciprocity in cross-chamber legislative interactions." *The Journal of Politics* 76(3):754–769.
- Latouche, Pierre, Etienne Birmelé and Christophe Ambroise. 2011. "Overlapping stochastic block models with application to the french political blogosphere." *The Annals of Applied Statistics* pp. 309–336.
- Nowicki, Krzysztof and Tom A B Snijders. 2001. "Estimation and prediction for stochastic blockstructures." *Journal of the American statistical association* 96(455):1077–1087.
- Raftery, Adrian E, Xiaoyue Niu, Peter D Hoff and Ka Yee Yeung. 2012. "Fast inference for the latent space network model using a case-control approximate likelihood." *Journal of Computational and Graphical Statistics* 21(4):901–919.
- Van Duijn, Marijtje AJ, Krista J Gile and Mark S Handcock. 2009. "A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models." *Social Networks* 31(1):52–62.
- Wang, Cheng, Carter T Butts, John R Hipp, Rupa Jose and Cynthia M Lakon. 2016. "Multiple imputation for missing edge data: A predictive evaluation method with application to Add Health." *Social networks* 45:89–98.