The Sensitivity of Spatial Regression Models to Network Misspecification

Supplementary Materials

Sebastian Juhl
University of Mannheim
sebastian.juhl@gess.uni-mannheim.de

Overview

This document includes all supplementary materials referred to in the main article. For any further information, please contact the author.

A Comparison of the m-STAR and the BMA Parameter Estimates

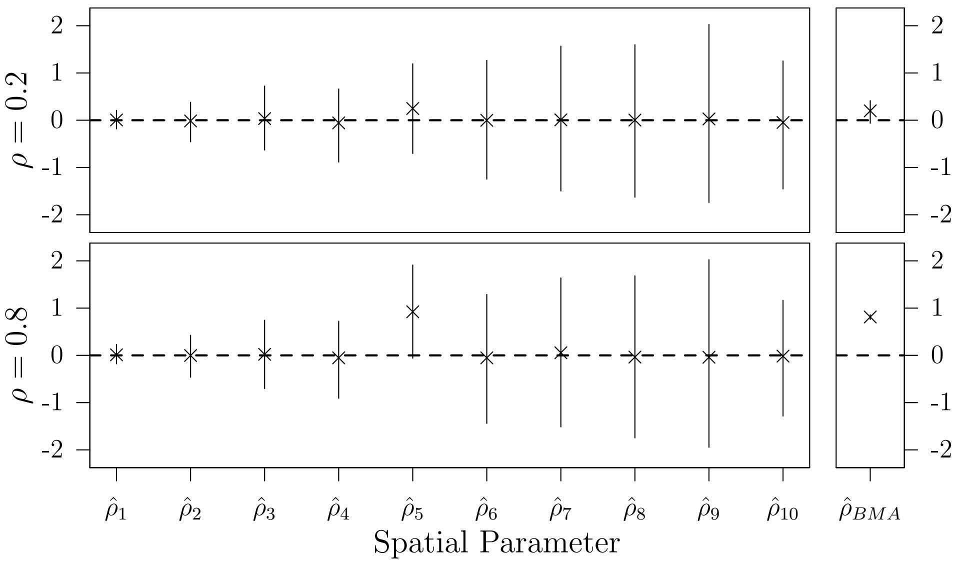
This section uses one simulation study and one application presented in the article in order to compare the performance of the m-STAR and the BMA approach. However, as mentioned in the main article, note that the parameter estimates have a slightly different interpretation. Therefore, the focus is on the inferences suggested by each approach.

In a first step, Figure 1 uses the setup of the first Monte Carlo study where the dependence structure is unidimensional by design. The left part depicts the parameter estimates obtained by the m-STAR model and the right part shows the BMA estimate. Due to multicollinearity in the spatial lags, the m-STAR estimates differ greatly across the 1,000 simulations of $\boldsymbol{y}$. As a consequence, the m-STAR approach is unable to detect non-random spatial clustering in the low-dependency scenario. In contrast, BMA provides accurate and precise parameter estimates.

In a second step, I investigate the performance of the m-STAR approach using the data from Plümper and Neumayer (2010), which is the second application in the article. Given that the models considered by the authors not only differ with respect to $\boldsymbol{W}$ but also on other dimensions, I only consider Models 2, 7, 8, 9, and 10 (see Supplementary Materials G for a summary of the differences between the models). To obtain parameter estimates, I restrict the model space in the BMA approach to these models and estimate the m-STAR model with the five different spatial lags.

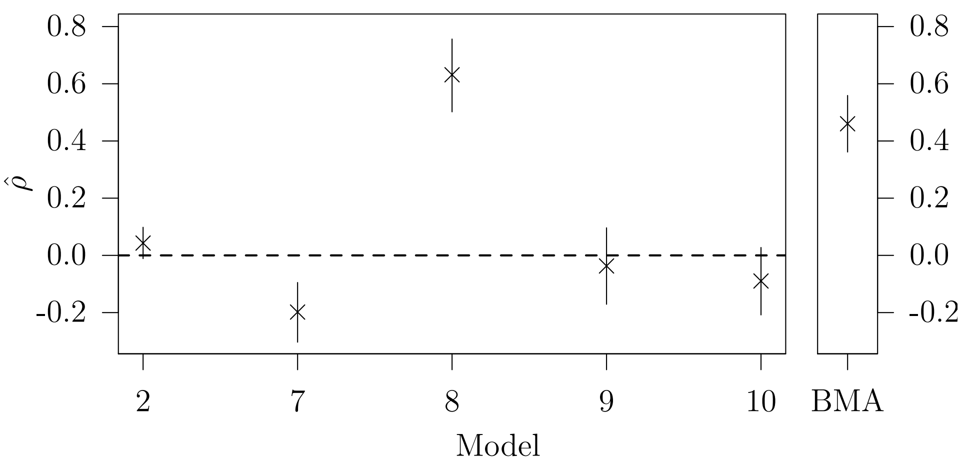
The horizontal axis in Figure 2 refers to the weighting scheme that belongs to the model with the respective number in the original article. Clearly, the parameter estimates are less affected by multicollinearity and identify two significant spatial parameters: a negative parameter for the spatial lag with an unstandardized connectivity matrix based

Figure 1: Monte Carlo Study 1: Spatial Parameter Estimates



on contiguity (Model 7) and a positive parameter for the spatial lag with the inverse distance specification of $\boldsymbol{W}$ (Model 8). However, given that contiguous countries also have the highest values in the network based on the inverse distance, the m-STAR model suggests that the countries' capital tax rates are positively correlated. At the same time, since the BMA estimate is a weighted average of the candidate models, it produces a single positive spatial parameter estimate. Taken together, both approaches come to similar substantive conclusions.

Figure 2: Plümper & Neumayer (2010): Spatial Parameter Estimates



B Monte Carlo Experiments: Estimated Impacts

B.1 Monte Carlo Study 1: Uncertainty in the Neighborhood Definition

	$\rho = 0.2$			$\rho = 0.8$		
	Direct	Indirect	Total	Direct	Indirect	Total
True	2.009	0.474	2.483	2.354	6.994	9.348
$\boldsymbol{W}_1$	2.004	0.116	2.120	2.874	4.103	6.976
	(0.089)	(0.142)	(0.174)	(0.154)	(0.636)	(0.759)
$\boldsymbol{W}_2$	2.015	0.255	2.270	2.761	6.308	9.069
	(0.090)	(0.221)	(0.255)	(0.140)	(0.922)	(1.027)
$\boldsymbol{W}_3$	2.023	0.355	2.379	2.621	6.927	9.548
	(0.091)	(0.260)	(0.296)	(0.127)	(1.021)	(1.109)
$\boldsymbol{W}_4$	2.024	0.429	2.453	2.511	7.391	9.902
	(0.091)	(0.296)	(0.331)	(0.121)	(1.147)	(1.228)
$\boldsymbol{W}_5$	2.012	0.517	2.530	2.381	7.673	10.055
	(0.090)	(0.329)	(0.358)	(0.115)	(1.305)	(1.381)
$\boldsymbol{W}_6$	2.009	0.504	2.513	2.333	7.758	10.091
	(0.089)	(0.342)	(0.367)	(0.115)	(1.468)	(1.542)
$\boldsymbol{W}_7$	2.001	0.485	2.486	2.277	7.785	10.062
	(0.089)	(0.355)	(0.377)	(0.114)	(1.657)	(1.729)
$\boldsymbol{W}_8$	2.005	0.488	2.493	2.277	7.907	10.184
	(0.089)	(0.365)	(0.388)	(0.116)	(1.877)	(1.951)
$\boldsymbol{W}_9$	2.007	0.484	2.491	2.277	8.038	10.315
	(0.090)	(0.372)	(0.397)	(0.118)	(2.128)	(2.203)
$\boldsymbol{W}_{10}$	2.004	0.483	2.487	2.254	8.113	10.367
	(0.089)	(0.382)	(0.405)	(0.120)	(2.416)	(2.492)

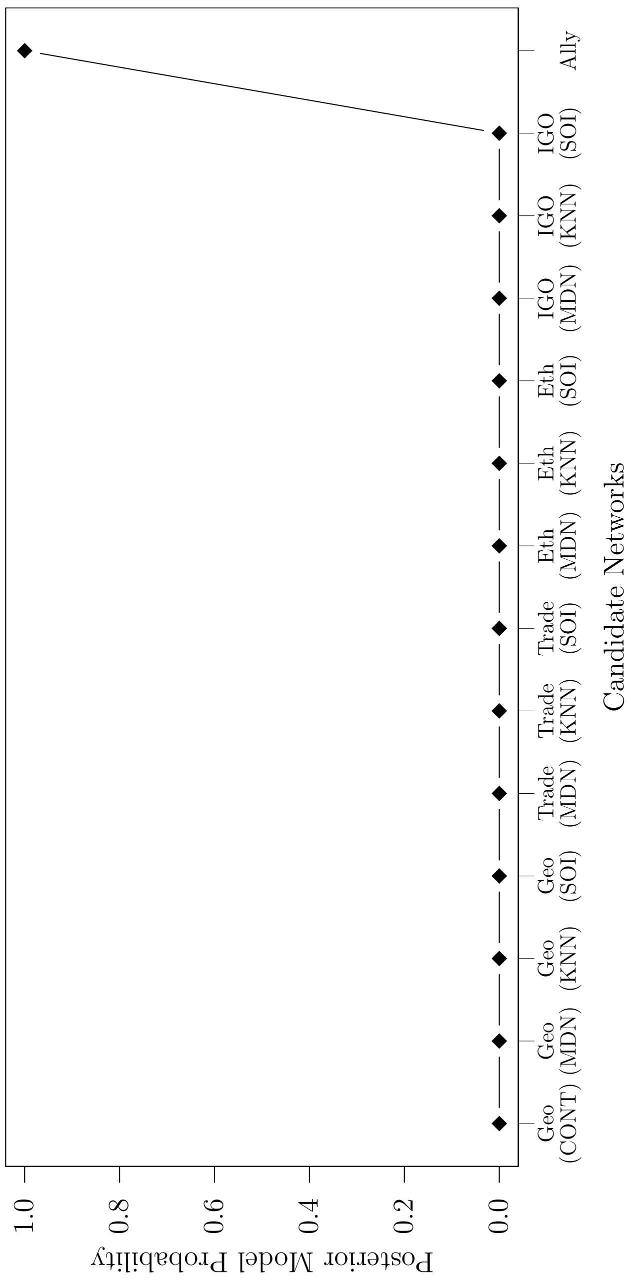
B.2 Monte Carlo Study 2: Uncertainty in the Weighting Scheme

	$\rho = 0.2$			$\rho = 0.8$		
	Direct	Indirect	Total	Direct	Indirect	Total
True	2.013	0.475	2.487	2.484	7.025	9.509
Binary	2.009	0.480	2.489	2.410	7.426	9.837
	(0.090)	(0.333)	(0.361)	(0.123)	(1.495)	(1.579)
Contiguity	2.006	0.226	2.232	2.865	4.611	7.477
	(0.089)	(0.150)	(0.186)	(0.159)	(0.712)	(0.842)
$1/d$	2.018	0.537	2.555	2.535	7.994	10.529
	(0.090)	(0.314)	(0.346)	(0.129)	(1.356)	(1.451)
$1/ln(d)$	2.013	0.533	2.545	2.442	7.806	10.249
	(0.090)	(0.338)	(0.368)	(0.125)	(1.492)	(1.581)
$1/(rev.\ d)$	2.021	0.531	2.552	2.527	7.861	10.389
	(0.091)	(0.321)	(0.355)	(0.129)	(1.360)	(1.453)
$1/d^2$	2.014	0.389	2.403	2.684	6.882	9.566
	(0.090)	(0.235)	(0.269)	(0.136)	(1.056)	(1.158)

C Replication of Model 1 in Gleditsch & Ward (2006)

	Mean	SD	2.5%	97.5%
Democracies				
Intercept	2.856	0.822	1.249	4.481
Logged GDP	-0.551	0.090	-0.732	- 0.381
Proportion Neighboring Democracies	-0.311	0.255	-0.812	0.181
Civil War	0.370	0.228	-0.103	0.789
Years of Peace at Territory	0.002	0.002	-0.003	0.006
Economic Growth	-0.024	0.012	-0.048	-0.001
Global Proportion of Democracies	-0.935	1.044	-3.071	1.044
Neighboring Transition to Democracy	-	-	-	-
Autocracies				
Intercept	4.088	0.553	3.000	5.157
Logged GDP	-0.093	0.061	-0.213	0.027
Proportion Neighboring Democracies	-0.502	0.227	-0.952	-0.045
Civil War	-0.011	0.153	-0.300	0.306
Years of Peace at Territory	-0.004	0.002	-0.008	-0.001
Economic Growth	0.003	0.007	-0.011	0.018
Global Proportion of Democracies	-2.753	0.681	-4.087	-1.442
Neighboring Transition to Democracy	-0.417	0.142	-0.688	-0.130

D Posterior Model Probabilities for the Candidate Models in Zhukov & Stewart (2013)

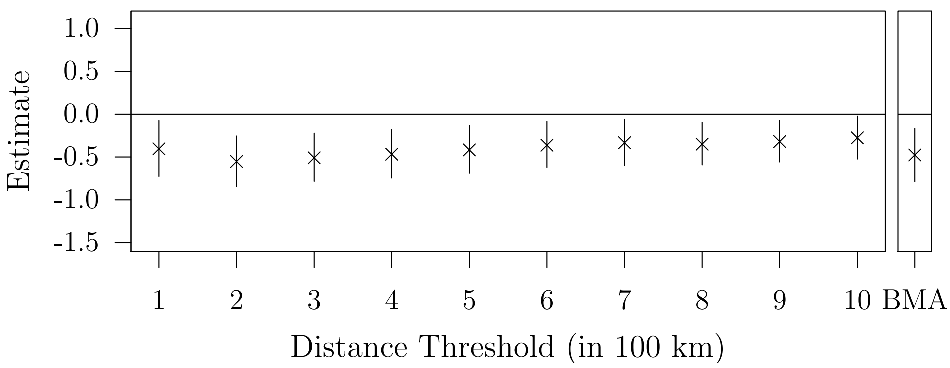


This figure shows the posterior model probabilities for the candidate networks analyzed by Zhukov and Stewart (2013) where all networks are equally likely *a priori*. The network specifications differ in their distance metric and in their connectivity criteria. The different distance metrics and their abbreviations in the figure are: geographical distance (Geo), ethnic proximity (Eth), trade proximity (Trade), joint membership in intergovernmental organizations (IGO), and joint membership in a military alliance (Ally). The connectivity criteria employed are: thresholding (CONT), minimum distance (MDN), k-nearest neighbor (KNN), and the sphere of influence (SOI). Like Zhukov and Stewart

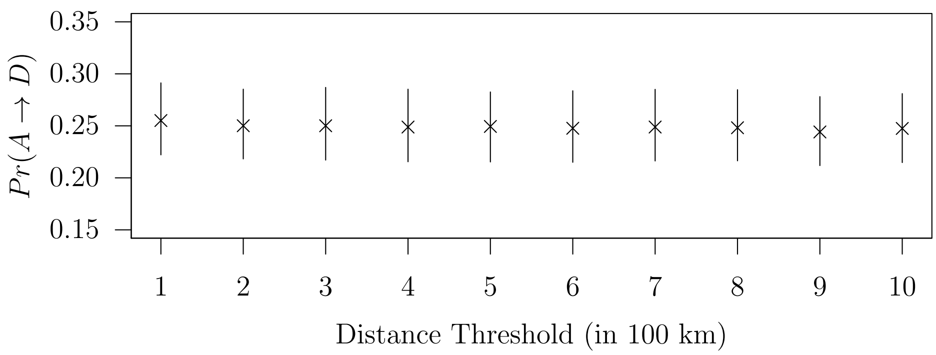
(2013, 279), I use the interborder distance of 500 km, the presence of one mutual defense pact, and $k = 4$ nearest neighbors as thresholds for connectivity. See Zhukov and Stewart (2013, 277ff) for a detailed description of the different metrics and connectivity criteria.

In this example, both BMA and the three-step procedure reach the same substantive conclusion: military alliance networks appear to be the most plausible channel for the spread of regime type, given the candidate pool (see also Zhukov and Stewart 2013, 282).

E Parameter Estimates of the Effect of Neighboring Transition to Democracy



F Predicted Transition Probability $Pr(A \to D)$ for Different Neighborhood Definitions



G Differences in Model Specification for the Nine Models Considered in Plümper & Neumayer 2010

Model	Temporal Lag	Unit FEs	Period FEs	Row-Standardized $\boldsymbol{W}$	Weighting Scheme
2	✓	✓	✓	✓	contiguity
3	-	✓	✓	✓	contiguity
4	✓	✓	-	✓	contiguity
5	-	✓	-	✓	contiguity
6	✓	-	✓	✓	contiguity
7	✓	✓	✓	-	contiguity
8	✓	✓	✓	-	$1/(distance)$
9	✓	✓	✓	-	$1/ln(distance)$
10	✓	✓	✓	✓	$1/(distance\ reversed)$