The Wald Test of Common Factors in Spatial Model Specification Search Strategies

Sebastian Juhl

University of Mannheim

Overview

This document contains supplementary information referred to in the article. For any further information, please feel free to contact the author.[1]

A The Delta Method

Based on the mean value theorem, the delta method uses a first-order Taylor series expansion around $\boldsymbol{\lambda}$ in order to approximate the asymptotic distribution of the non-linear function $g(\hat{\boldsymbol{\lambda}})$:

$$g(\hat{\boldsymbol{\lambda}}) = g(\boldsymbol{\lambda}) + \boldsymbol{G}(\bar{\boldsymbol{\lambda}})(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}). \tag{A1}$$

Let $\bar{\boldsymbol{\lambda}}$ take on values between $\hat{\boldsymbol{\lambda}}$ and $\boldsymbol{\lambda}$. By Slutsky's theorem and because of the assumed consistency which implies that $\text{plim}_{n\rightarrow\infty}\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}$, the vector of partial derivatives evaluated at $\bar{\boldsymbol{\lambda}}$, denoted $\boldsymbol{G}(\bar{\boldsymbol{\lambda}})$, converges in probability to $\boldsymbol{G}(\boldsymbol{\lambda})$. It follows that

$$\sqrt{n}\left(g(\hat{\boldsymbol{\lambda}}) - g(\boldsymbol{\lambda})\right) = \boldsymbol{G}(\bar{\boldsymbol{\lambda}})\sqrt{n}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}). \tag{A2}$$

Consequently, $g(\hat{\boldsymbol{\lambda}})$ has the same limiting distribution as $\boldsymbol{G}(\boldsymbol{\lambda})(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda})$.

The first expression on the right-hand side of Equation (A2), $\boldsymbol{G}(\bar{\boldsymbol{\lambda}})$, converges in probability to a constant and the second term, $\sqrt{n}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda})$, converges in distribution to a multivariate normal distribution with a mean vector of $\mathbf{0}$ and an asymptotic variance-covariance matrix given by $\boldsymbol{\Sigma}$. By a linear transformation of normal variables, the product of these terms is also normally distributed with a mean of 0 and a variance given by $\boldsymbol{G}(\hat{\boldsymbol{\lambda}})\boldsymbol{\Sigma}\boldsymbol{G}(\hat{\boldsymbol{\lambda}})'$:

$$\sqrt{n}(g(\hat{\boldsymbol{\lambda}}) - g(\boldsymbol{\lambda})) \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{G}(\boldsymbol{\lambda})\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\lambda})'\right). \tag{A3}$$

This result is given in Equation (7) in the article.

[1]Email: `sebastian.juhl@gess.uni-mannheim.de`.

B Correlation Between Included Regressor and Omitted Variable in the DGP

Table B.1 contains the median correlation between the regressor $\boldsymbol{x}$ and the omitted variable $\boldsymbol{z}$ across the $1,000$ simulation trials. As already indicated in the article, the setup of the Monte Carlo experiment assures that an increase in γ leads to a stronger correlation.

Table B.1: Median Correlation Between Included and Omitted Regressor

γ	$n = 49$	$n = 100$	$n = 225$	$n = 400$
0.0	0.00 [−0.28; 0.29]	0.00 [−0.20; 0.20]	0.00 [−0.13; 0.13]	0.00 [−0.09; 0.10]
0.2	0.22 [−0.06; 0.46]	0.22 [0.03; 0.40]	0.19 [0.06; 0.30]	0.20 [0.10; 0.29]
0.4	0.40 [0.14; 0.61]	0.42 [0.25; 0.57]	0.35 [0.24; 0.46]	0.38 [0.29; 0.45]
0.6	0.55 [0.35; 0.71]	0.57 [0.43; 0.68]	0.49 [0.39; 0.58]	0.52 [0.45; 0.58]
0.8	0.66 [0.49; 0.78]	0.68 [0.58; 0.76]	0.60 [0.52; 0.67]	0.63 [0.57; 0.68]
1.0	0.74 [0.61; 0.83]	0.76 [0.68; 0.82]	0.69 [0.62; 0.74]	0.71 [0.67; 0.75]

C Additional Simulation Results & Robustness Tests

Given the space constraints in the article, this section contains a number of additional simulation results as well as some robustness tests.

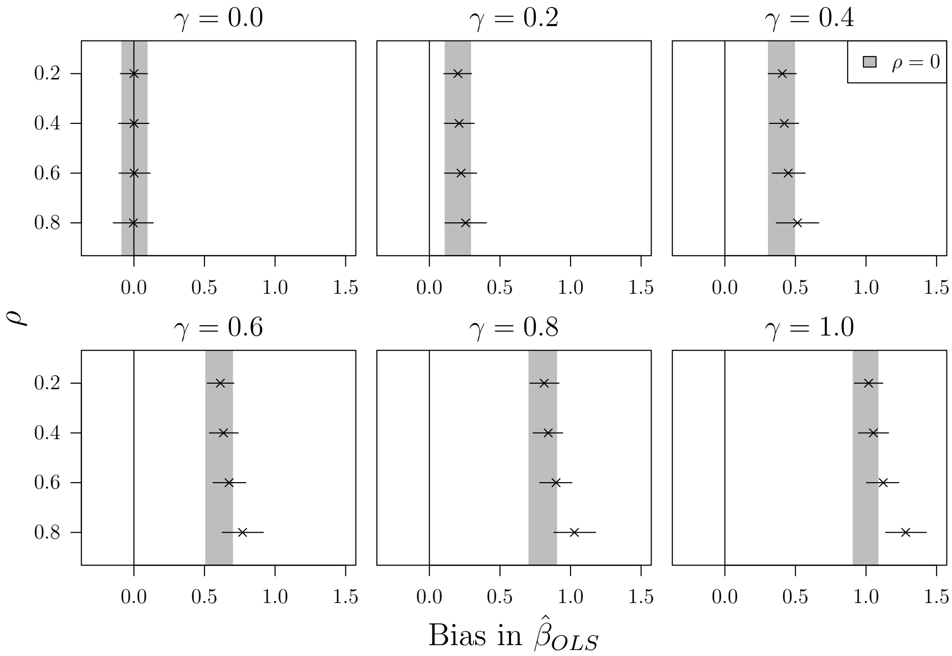
C.1 Biased Coefficient & Impact Estimates

C.1.1 Omitted Variables Bias in the Non-Spatial OLS Model

This section demonstrates that spatial dependence in an omitted variable which is correlated with an included regressor amplifies the standard omitted variables bias in non-spatial OLS models (see also Pace and LeSage, 2010). To this end, Figure C.1.1 reports the bias in the coefficient estimates from an OLS model across different parameter settings outlined in the main article using a sample size of $n = 400$. The gray areas in each of the six panels represent the estimates obtained in a scenario without spatial dependence ($\rho = 0$). This arrangement provides a baseline for the comparison with the omitted variables bias in scenarios with cross-sectional dependence. In the upper left panel, the true DGP resembles a SEM process while the other panels show SDM DGPs.

If an omitted variable $\boldsymbol{z}$ is correlated with the included regressor $\boldsymbol{x}$ but does not follow a spatial process, $\hat{\beta}_{OLS}$ is biased by γ. In addition to this textbook example of omitted

Figure C.1.1: Bias in the Non-Spatial OLS Coefficient Estimate



variables bias, Figure C.1.1 illustrates the effect of spatial dependence in $\boldsymbol{z}$ on the size of this bias. The upper left panel shows that non-random spatial clustering in an omitted regressor that is uncorrelated with $\boldsymbol{x}$ does not cause any bias in $\hat{\beta}_{OLS}$. At the same time, while the point estimate $\hat{\beta}_{SEM}$ would equal $\hat{\beta}_{OLS}$, the estimates derived from a correctly specified SEM model would be more efficient (e.g., Elhorst, 2010, 14).

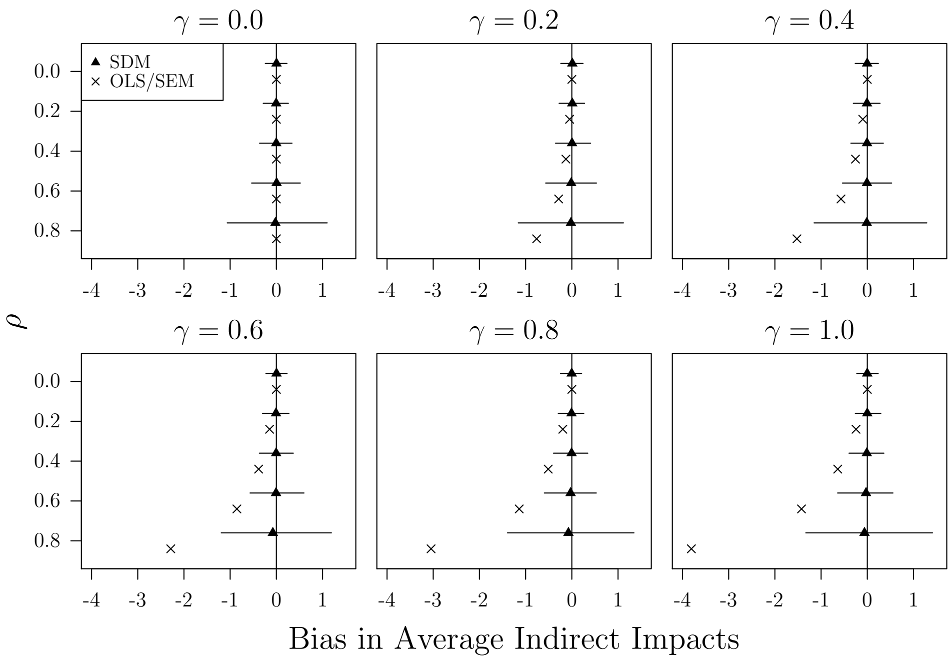
However, the situation is different when correlation exists between omitted and the included regressors. The five remaining panels show that spatial autocorrelation magnifies the standard omitted variables bias. This effect is most pronounced if there are strong interdependencies in the data and if the correlation increases. In a scenario without cross-sectional dependence and where $\gamma = 1$, for example, the bias in $\hat{\beta}_{OLS}$ is 1 with 95% of the empirical density within $[0.91; 1.09]$. However, it increases by almost 30% to 1.28 $[1.14; 1.43]$ if the spatial dependence in $\boldsymbol{z}$ is strong ($\rho = 0.8$). Therefore, while cross-sectional dependence in an omitted regressor does not cause bias in regression coefficients by itself, these results show that it amplifies the existing bias stemming from an omitted variable. For applied research, this implies that the threat arising from an omitted regressor for the validity of inferences is even more pronounced if there are reasons to expect that this variable is spatially clustered.

C.1.2 Biased Indirect Impact Estimates

Only looking at coefficient estimates in the context of spatial autocorrelation is insufficient (e.g., Whitten, Williams and Wimpy, 2019; Elhorst, 2010; LeSage and Pace, 2009).

Interpreting spatial models in terms of substantive effects requires additional considerations since the overall effect of a change in a regressor not only depends on its associated coefficient but also on the strength of the cross-sectional dependence and the spatial configuration. In general, it is informative to partition the overall effect into direct, indirect, and total impacts (e.g., Whitten, Williams and Wimpy, 2019). Direct impacts describe the change in one unit's outcome caused by a change in the same unit's regressor. Indirect impacts, also called spillover effects, quantify the shift in one unit's outcome induced by a change in another unit's covariate. The sum of these quantities constitutes the total impact for each unit. By implication and depending on the spatial configuration, the modification of one covariate in a single observation potentially affects all other units in the sample to a different degree. As a way to handle this wealth of information provided by spatial regression models, LeSage and Pace (2009) suggest to compute these impact estimates for each individual unit in the sample and report their averages as meaningful summary measures.

Figure C.1.2: Comparing the Bias in the Estimated Average Indirect Impact Derived from the SDM and OLS/SEM Models



By doing so, Figure C.1.2 reports the average indirect impact (AII) estimates in order to investigate the substantive effect of omitted spillovers for substantive inferences. Since both the SEM and the non-spatial OLS model assume no spillover effects, the estimated coefficient associated with $\boldsymbol{x}$ represent the regressor's average direct impact (ADI) which also equals the estimated average total impact (ATI) for these models. Consequently, Figure C.1.2 does not report confidence intervals for the bias in the AII estimates from the OLS/SEM model specification since they are not estimated but assumed to be zero. In contrast, the SDM model accounts for indirect spillovers and the estimate of a regressor's total impact is the sum of its average direct and indirect impact. For the SDM model, the the following partial derivatives matrix allows researchers to obtain substantive impact estimates:

$$(\boldsymbol{I}_n - \hat{\rho}\boldsymbol{W})^{-1}(\hat{\beta}\boldsymbol{I}_n + \boldsymbol{W}\hat{\theta}). \tag{C1}$$

In Equation (C1), $\boldsymbol{I}_n$ is the $n \times n$ dimensional identity matrix $\boldsymbol{W}$ is the exogenously given connectivity matrix of the same size and $\hat{\rho}$, $\hat{\beta}$, and $\hat{\theta}$ are the parameter estimates obtained from the unrestricted SDM model. The ADI can be computed by averaging over the while diagonal elements of the matrix shown in Equation (C1) the AII is the average of its cumulative off-diagonal elements (e.g., Lacombe and LeSage, 2015).

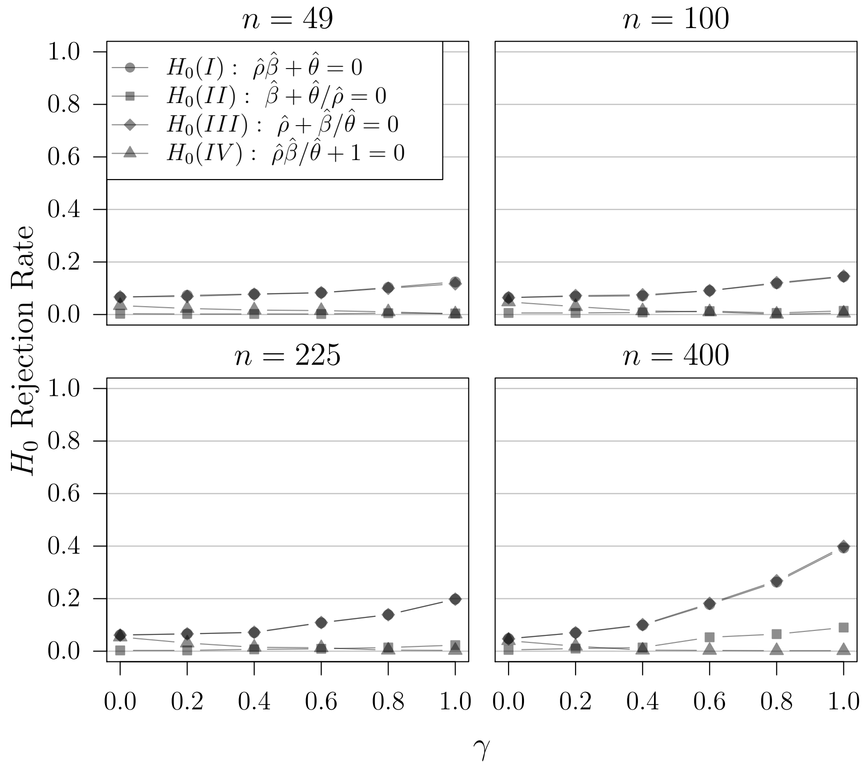
Across the different sample sizes and for all levels of spatial autocorrelation, Figure C.1.2 illustrates that all models – the non-spatial OLS, the SEM, and the SDM specification – yield unbiased estimates of the AII if the true DGP does not feature any indirect spillover effects. However, if the omitted variable is correlated with the included regressor, both models that rule out spillovers by assumption (OLS/SEM models) underestimate the regressor's indirect impact. The results also show that this bias increases (i) as the correlation between $\boldsymbol{x}$ and $\boldsymbol{z}$ increases and (ii) with the strength of the cross-sectional interdependence. Whereas moderate levels of correlation do not induce bias if ρ is small, higher levels of spatial dependence in the omitted variable induce stronger indirect spillover effects in the true DGP which, in turn, biases estimates from the OLS/SEM models. The size of the sample only affects the estimate's variability but has no effect on the bias. In line with the findings presented by LeSage and Pace (2009), the ability to derive unbiased effect estimates across a range of different spatial regimes is an important feature of the SDM model that makes it valuable in applied research settings (see also Mur and Angulo, 2006). At the same time, since the calculation of the impact estimates for the SDM model involves three estimated parameters ($\hat{\rho}$, $\hat{\beta}$, and $\hat{\theta}$) a major drawback of the SDM model is its inefficiency of the impact estimates.

C.2 Negative Spatial Autocorrelation

While positive spatial autocorrelation is considered to be more common in applied settings, this section investigates the performance of the four Wald tests in a scenario with negative autocorrelation. To this end, I rerun the simulation experiment described in the article for the different values for γ and across the different sample sizes and specify the true spatial parameter such that $\rho = -0.2$. This represents a scenario with a moderate level of negative spatial autocorrelation.

Figure C.2.1 depicts the results of this additional simulation exercise. As the results

Figure C.2.1: Performance of the Different Wald Tests in a Scenario with Negative Spatial Autocorrelation



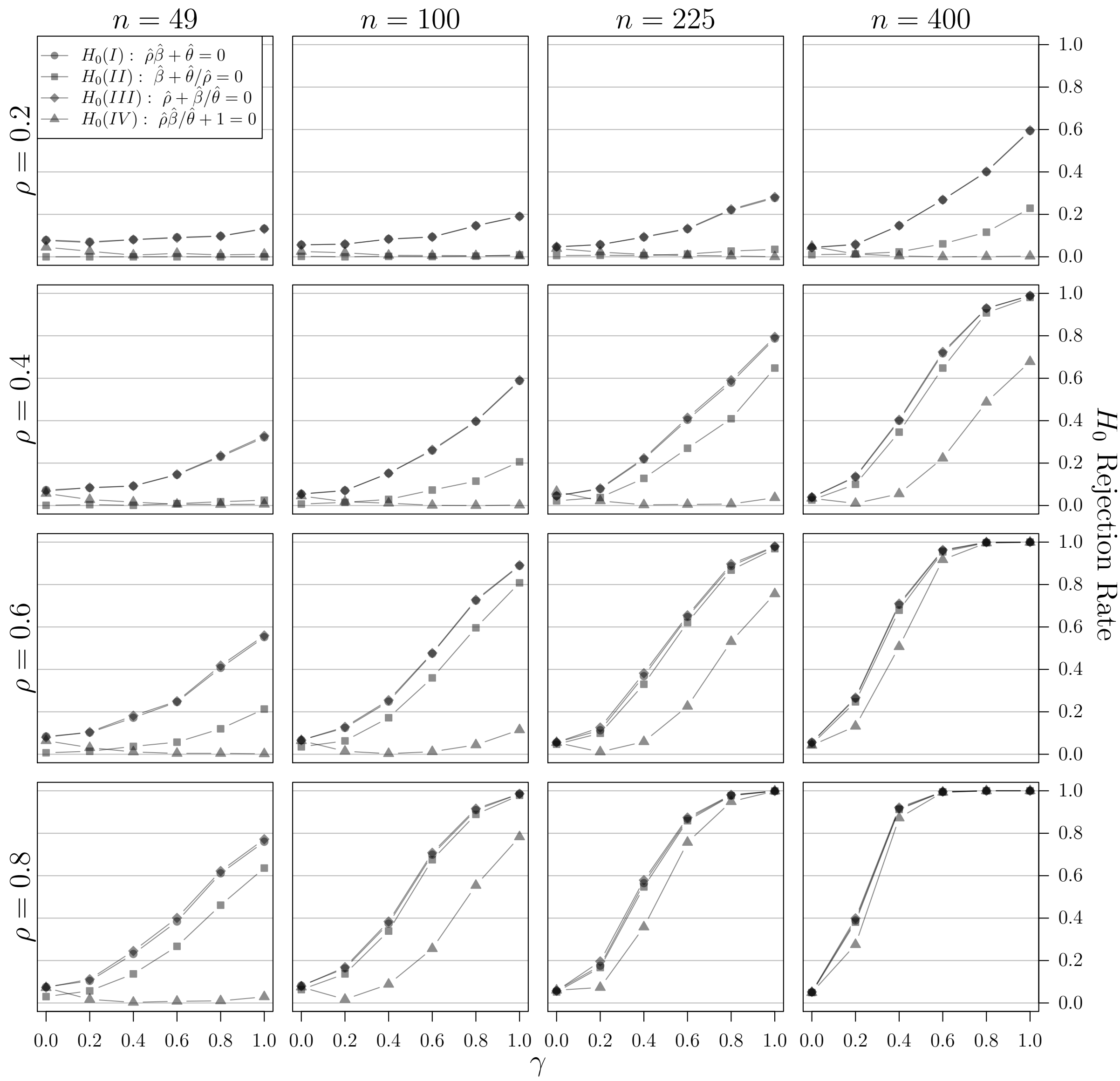
indicate, there is no difference between positive or negative spatial autocorrelation with respect to the Wald test's non-invariance to reparameterizations of the null hypothesis. Again, while $H_0(I)$ and $H_0(III)$ perform comparatively well, the power functions of the other two algebraically equivalent expressions of the common factor hypothesis are unsatisfactory. Based on these tests, researchers would frequently conclude that no meaningful spillovers exist although the true DGP features sizable indirect effects. Moreover, Figure C.2.1 clearly identifies notable differences between the Wald tests despite the fact that all tests not only use the same data and parameter estimates to calculate their respective test statistic. They also work with algebraically identical expressions of the null hypothesis.

C.3 Alternative Connectivity Scheme and Possible Edge Effects

In order to verify that the results presented in the main article are not solely driven by the queen contiguity scheme used to construct $\boldsymbol{W}$, Figure C.3.1 presents the results from the same simulations as conducted in the main part of the article with a rook instead of a queen connectivity scheme. To further avoid possible edge effects, the grid structure is mapped onto a torus so that there are no spatial units on the edges.

The results are in line with the ones already reported in the article. The different variants of the Wald test frequently come to conflicting conclusions regarding the null

Figure C.3.1: Share of Null Hypothesis Rejections (Rook Connectivity Mapped Onto a Torus)



hypothesis of common factors. This finding is also in line with the simulation results reported by Mur and Angulo (2006) who demonstrate that the connectivity scheme has a minor impact on the behavior of the LM, LR, and Wald test. Therefore, the results hold when controlling for possible edge effects and alternative connectivity structures.

C.4 Alternative Values for β

As the analytical results presented in the article suggest, the accuracy of the Taylor series approximation not only depends on ρ and γ but also on β. Consequently, this parameter also affects the performance of the different variants of the Wald test of common factors (see also Gregory and Veall, 1986). Against this background, this section contains additional simulation experiments in which the relevant parameters – ρ, γ, and β – vary while the sample size is fixed to $n = 100$.

Figure C.4.1: Share of Null Hypothesis Rejections Across Different Values of β

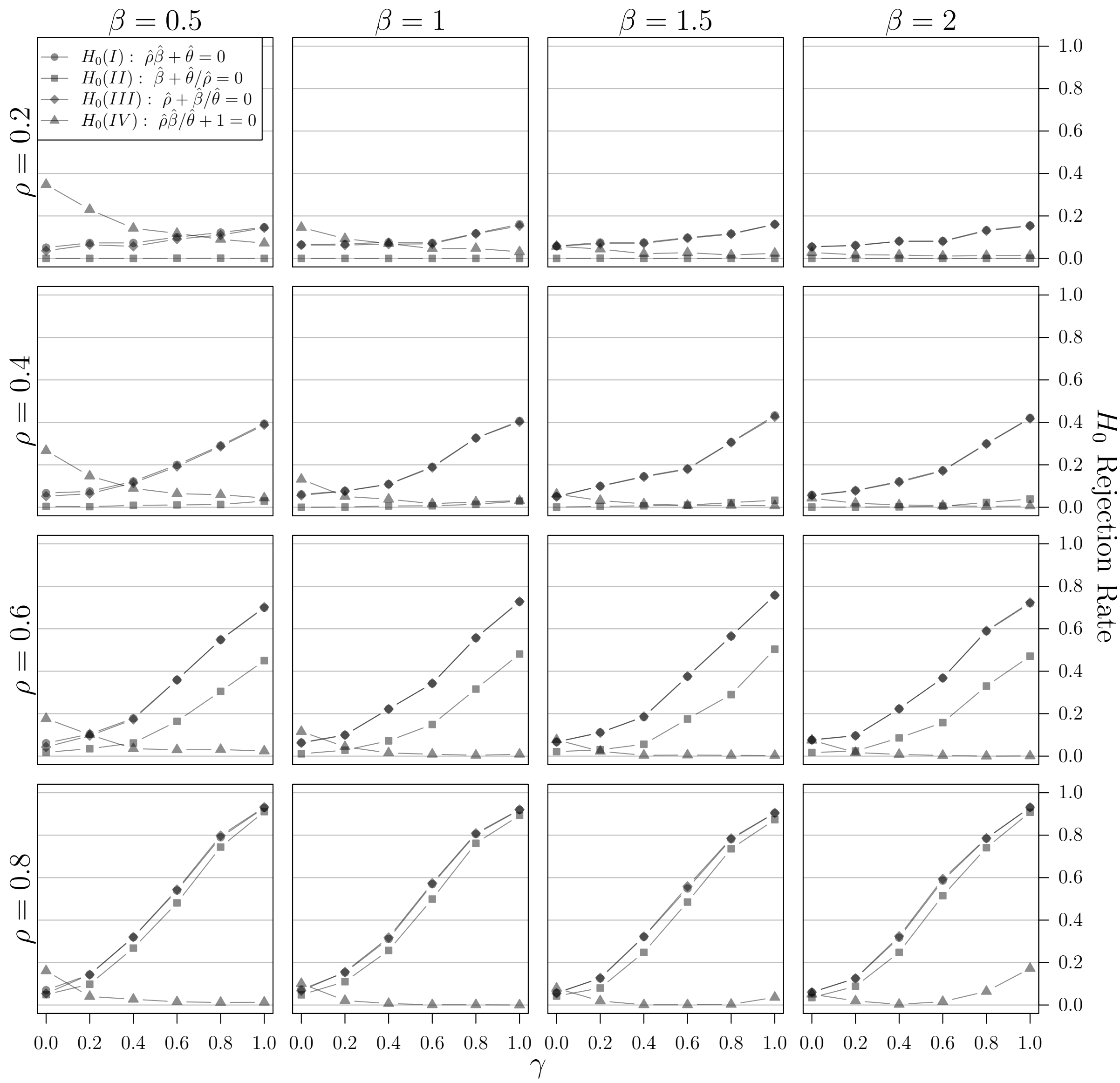


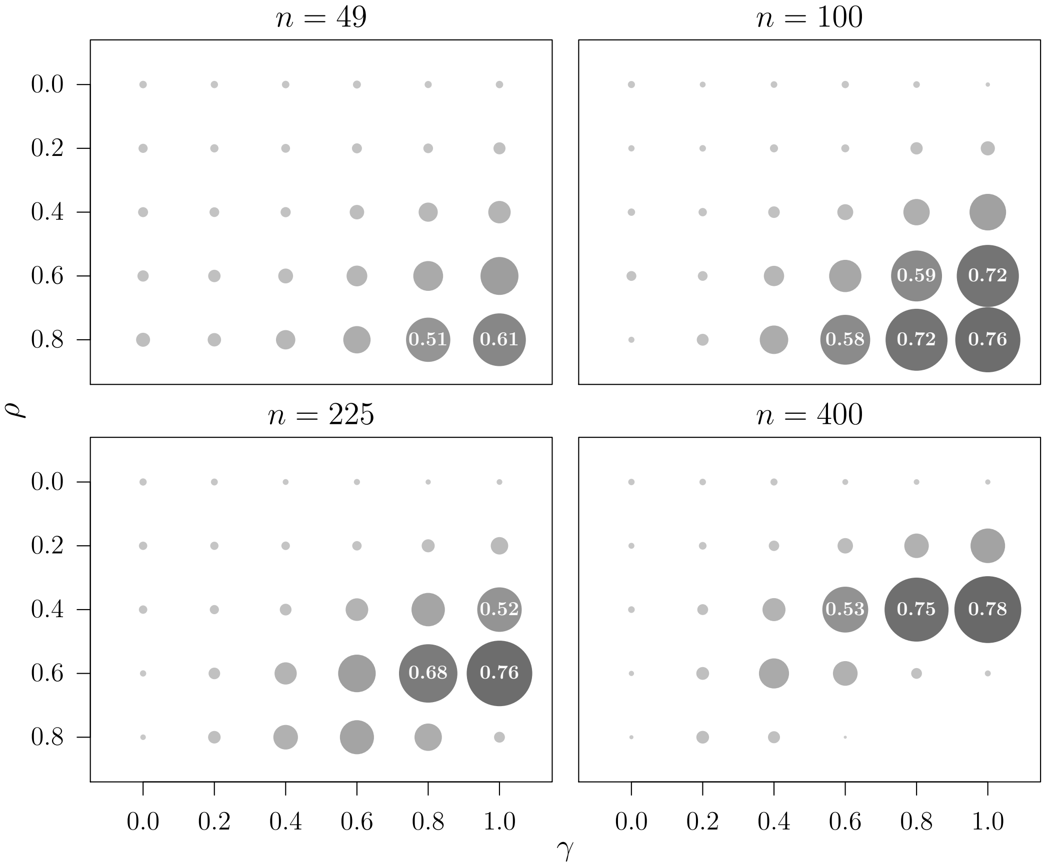
Figure C.4.1 displays the rejection rates of the null hypothesis of common factors for the different Wald tests based on asymptotic critical values across the specified parameter settings. The right column depicts the scenario considered in the main article where $\beta = 2$. Again, $H_0(I)$ and $H_0(III)$ perform equally well compared to $H_0(II)$ and especially $H_0(IV)$. $H_0(II)$ has almost no power unless there is a considerable level of spatial dependence in the DGP. As outlined in the article, this is because parameter values for ρ close to zero constitute an approximate violation of the assumed continuity of derivatives.

More strikingly, as Figure C.4.1 clearly shows, while the other three expressions of the null hypothesis remain relatively unaffected by the value of β in the true DGP, this parameter has serious consequences for the empirical size of $H_0(IV)$. Especially when there is only a trace of spatial autocorrelation in the DGP and β takes on smaller values, $H_0(IV)$ dramatically over-rejects the true null hypothesis of common factors.

C.5 Share of Conflicting Inferences from the Wald Test

As stated in the article, Wald tests based on different functional expressions of the common factor restriction can come to substantively different conclusions regarding the true DGP despite the fact that they use the same data and model estimates. To further investigate the severity of these conflicts, Figure C.5.1 reports the share of simulation iterations where at least one pair of tests suggests opposing conclusions regarding the existence of common factors using asymptotic critical values at the conventional α-level of 0.05. It identifies the regions of the parameter space where the tests diverge most frequently. In the four panels, the radius and the color of each circle signifies the share of inconsistent conclusions, where larger and darker circles indicate a higher share.

Figure C.5.1: Share of Inconsistent Inferences Across the 1,000 Simulation Trials

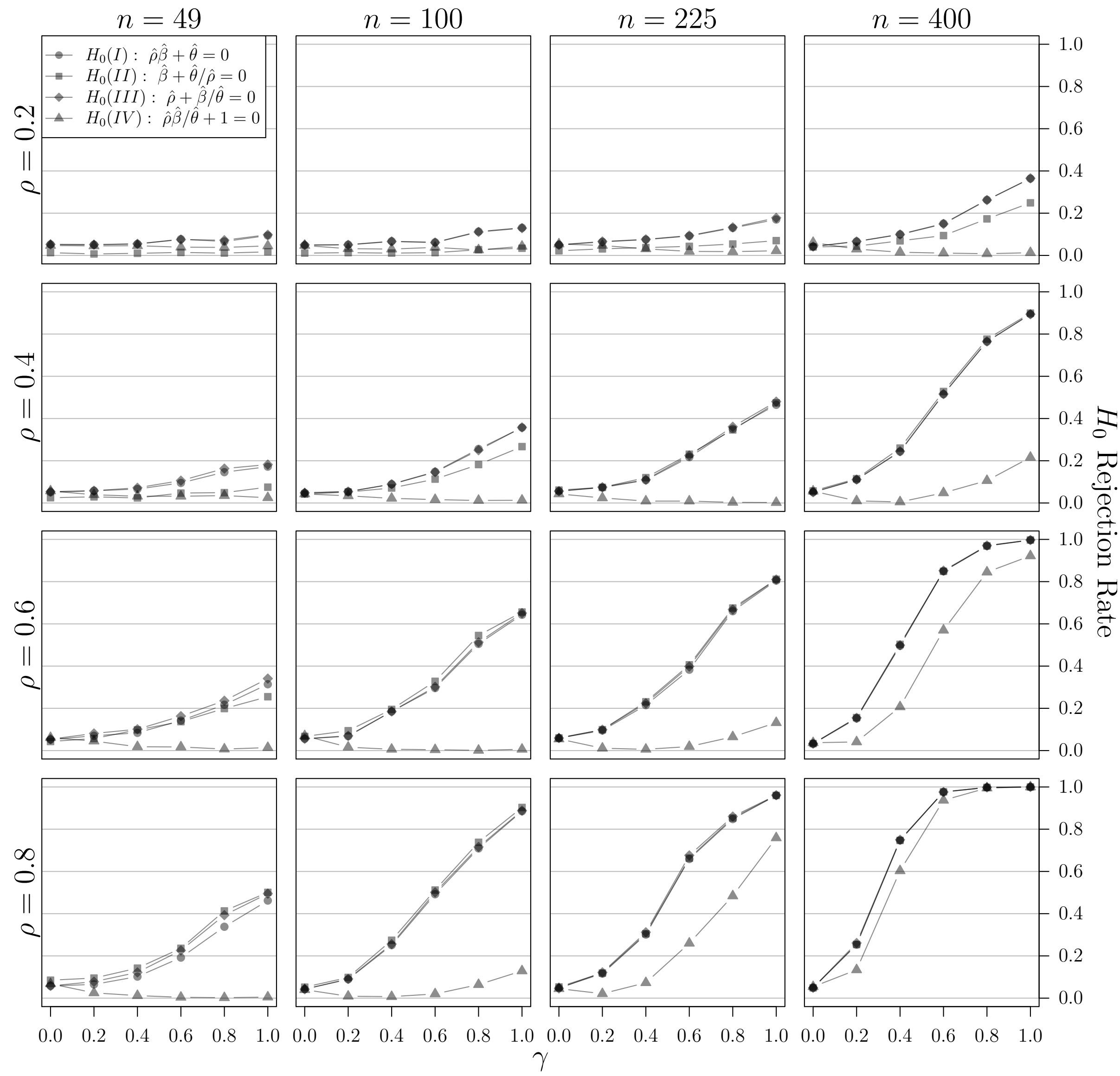


The area with the highest share of contradictory inferences is located at the lower-right part of the panels. Strong cross-sectional dependence in conjunction with high correlation between the regressor and the omitted variable constitutes a problem for the Wald tests as they regularly come to contradictory conclusions. While a larger sample size reduces the inconsistencies in a scenario where the spatial dependence in the omitted variable is strong, the Wald tests still contradict each other most of the time if the cross-sectional dependence in $\boldsymbol{z}$ is moderate. This illustrates the great potential for incorrect inferences caused by the Wald test's sensitivity to alternative expressions of the null hypothesis even in modestly sized samples.

C.6 Bootstrap Critical Values for the Wald Tests of Common Factors

While the main article already discusses the improvement of the Wald test's size properties when bootstrap critical values are used instead of the asymptotic critical value, this section evaluates the power of the modified Wald test. Figure C.6.1 compares the power functions of the four variants of the Wald test based on their estimated critical values across the different parameter settings specified in the article.

Figure C.6.1: Power of the Wald Test Based on Bootstrap Critical Values



The results illustrate that, although the bootstrap approach does not improve the performance of the Wald test based on $H_0(IV)$, it reduces the discrepancies in the power functions of the remaining three formulations of the common factor restriction. This is primarily a result of the improved power of $H_0(II)$. In contrast to the test's poor performance reported in the main article, basing inferences on the simulated null distribution as the reference distribution noticeably boosts its performance. At the same time, the problem with small values of ρ already discussed in the article still remains. Notwithstanding this, using bootstrap critical values reduces the conflict between the alternative expressions of the common factor restriction and better aligns the tests based on the four alternative variants of the common factor hypothesis. Therefore, this modification of the Wald test constitutes a superior assessment of the common factor restriction and should be used in empirical model search strategies as an alternative to the original Wald test that relies on the asymptotic χ^2 distribution.

C.7 Performance of the Likelihood Ratio Test

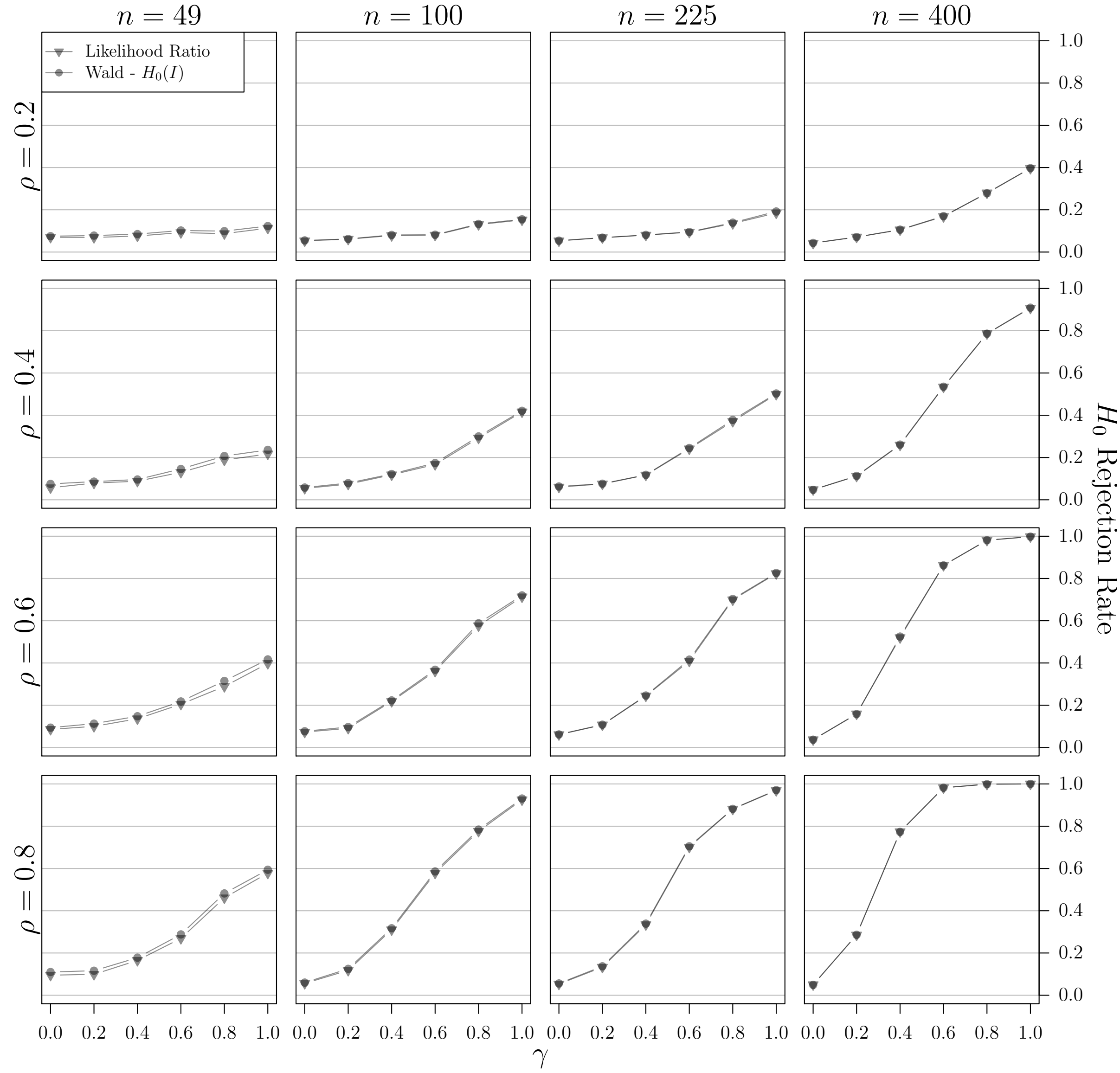
The LR test constitutes an alternative procedure to assess the common factor restriction in an unrestricted SDM model. In contrast to the Wald test, this statistic has the advantage that it is invariant to reparameterizations of the null hypothesis. Moreover, as Angulo and Mur (2011) show, the LR test of common factors performs well even under non-ideal conditions such as heteroscedasticity and non-linearity. However, it requires the estimation of both the unrestricted SDM and the restricted SEM model (e.g., Elhorst, 2014; Burridge, 1981). To investigate the performance of this alternative test, Figure C.7.1 compares the power functions of the LR test and the original Wald test based on asymptotic critical values. To this end, the figure shows the Wald specification with the restriction $H_0(I)$: $\hat{\rho}\hat{\beta} + \hat{\theta} = 0$ which was identified to be the best performing variant of the original Wald test in this simulation.

In contrast to the Wald test, the empirical size of the LR test does not notably differ across the sample sizes and the different values of ρ. In all scenarios, the share of simulation trials in which the test falsely rejects the null hypothesis is close to the specified α-level of 0.05. In line with previous research (e.g., Angulo and Mur, 2011; Mur and Angulo, 2009), the simulation experiment indicates that the LR test performs comparatively well across a wide range of parameter settings. As expected, all H_0 rejection rates increase in the sample size and as the strength of the spatial dependence increases.

Interestingly, the comparison shows that the original Wald test based on $H_0(I)$ performs equally well in these simulations. While this is an encouraging finding, it should be noted that no single formulation of the null hypothesis exists that outperforms the alternatives in all regions of the parameter space (e.g., Phillips and Park, 1988; Gregory and Veall, 1986). Consequently, despite the good performance of $H_0(I)$ in these simulations, there is no guarantee that the Wald test based on this particular parameterization always outperforms alternative formulations of the null hypothesis.

Taken together, this comparison suggest that, in small to medium sized samples, the LR test constitutes a powerful alternative to the Wald test in order to scrutinize the common factor restriction. The small disadvantage that the LR test requires more models to be estimated is negligible as compared to the considerable benefit that the test is not sensitive to arbitrary reparameterizations of the null hypothesis. In addition, the rapid increase in computational power steadily diminish concerns about the higher computational demand of the LR test in relation to the Wald test (see also Gibbons and Overman, 2012).

Figure C.7.1: Comparison of the LR Test and the Original Wald Test Based on $H_0(I)$



References

Angulo, Ana M. and Jesús Mur. 2011. "The Likelihood Ratio Test of Common Factors under Non-Ideal Conditions." *Investigaciones Regionales* 21:37–52.

Burridge, Peter. 1981. "Testing for a Common Factor in a Spatial Autoregression Model." *Environment and Planning A: Economy and Space* 13(7):795–800.

Elhorst, J. Paul. 2010. "Applied Spatial Econometrics: Raising the Bar." *Spatial Economic Analysis* 5(1):9–28.

Elhorst, J. Paul. 2014. "Matlab Software for Spatial Panels." *International Regional Science Review* 37(3):389–405.

Gibbons, Stephen and Henry G. Overman. 2012. "Mostly Pointless Spatial Econometrics?" *Journal of Regional Science* 52(2):172–191.

Gregory, Allan W. and Michael R. Veall. 1986. "Wald Tests of Common Factor Restrictions." *Economics Letters* 22(2-3):203–208.

Lacombe, Donald J. and James P. LeSage. 2015. "Using Bayesian Posterior Model Probabilities to Identify Omitted Variables in Spatial Regression Models." *Papers in Regional Science* 94(2):365–383.

LeSage, James P. and R. Kelley Pace. 2009. *Introduction to Spatial Econometrics.* Boca Raton, USA: CRC Press.

Mur, Jesús and Ana Angulo. 2006. "The Spatial Durbin Model and the Common Factor Tests." *Spatial Economic Analysis* 1(2):207–226.

Mur, Jesús and Ana Angulo. 2009. "Model Selection Strategies in a Spatial Setting: Some Additional Results." *Regional Science and Urban Economics* 39(2):200–213.

Pace, R. Kelley and James P. LeSage. 2010. Omitted Variables Biases of OLS and Spatial Lag Models. In *Progress In Spatial Analysis: Theory and Computation, and Thematic Applications*, ed. Antonio Páez, Julie Le Gallo, Ron N. Buliung and Sandy Dall'erba. Berlin: Springer pp. 17–28.

Phillips, P. C. B. and Joon Y. Park. 1988. "On the Formulation of Wald Tests of Nonlinear Restrictions." *Econometrica* 56(5):1065–1083.

Whitten, Guy D., Laron K. Williams and Cameron Wimpy. 2019. "Interpretation: The Final Spatial Frontier." *Political Science Research and Methods* pp. 1–17.