Spatial Regression Models: A Systematic Comparison of Different Model Specifications using Monte Carlo Experiments

Tobias Rüttenauer

University of Kaiserslautern
Department of Social Sciences
Erwin-Schrödinger-Str. 57
67663 Kaiserslautern, Germany
ruettenauer@sowi.uni-kl.de

Authors' preprint version. Please cite the final article:
Rüttenauer, T. (2019). Spatial Regression Models: A Systematic Comparison of Different Model Specifications using Monte Carlo Experiments. *Sociological Methods and Research*, Forthcoming.

LAST EDITED
June 9, 2019

ABSTRACT
Spatial regression models provide the opportunity to analyse spatial data and spatial processes. Yet, several model specifications can be used, all assuming different types of spatial dependence. This study summarises the most commonly used spatial regression models and offers a comparison of their performance by using Monte Carlo experiments. In contrast to previous simulations, this study evaluates the bias of the impacts rather than the regression coefficients and additionally provides results for situations with a non-spatial omitted variable bias. Results reveal that the most commonly used spatial autoregressive (SAR) and spatial error (SEM) specifications yield severe drawbacks. In contrast, spatial Durbin specifications (SDM and SDEM) as well as the simple SLX provide accurate estimates of direct impacts even in the case of misspecification. Regarding the indirect 'spillover' effects, several – quite realistic – situations exist in which the SLX outperforms the more complex SDM and SDEM specifications.

KEYWORDS
Spatial econometrics; Spatial regression; Monte Carlo experiments; Spillover effects; Spatial impacts; Simulations

---

Acknowledgements: I would like to thank Scott J. Cook for very helpful comments on an earlier version of this manuscript. Furthermore, I would like to thank two anonymous reviewers and the editors for helpful and constructive feedback. All remaining errors are my own.

Supplementary material: Supplementary results, the R script of the Monte Carlo program, and all other scripts necessary for reproduction of the results are available at the author's GitHub repository: `https://github.com/ruettenauer/Reproduction-Material-Spatial-Monte-Carlo-Experiments`.

1. Introduction

The increasing availability of spatially aggregated and georeferenced data has led to an increasing interest in spatial analyses amongst social scientists (Logan, 2012). For instance, social scientists have investigated the influence of contextual conditions (e.g. Crowder, Hall, & Tolnay, 2011; Friedrichs, Galster, & Musterd, 2003; Kling, Liebman, & Katz, 2007; Sampson, Morenoff, & Earls, 1999; Sampson, Morenoff, & Gannon-Rowley, 2002), or have dealt with explicitly spatial research questions like segregation, neighbourhood boundaries, or the exposure to environmental conditions (e.g. Dokshin, 2016; Downey, 2006; Legewie & Schaeffer, 2016; Lichter, Parisi, & Taquino, 2015; Reardon et al., 2008; Rüttenauer, 2018).

Still, researchers need to be aware of the fact that analysing spatial data bears new challenges regarding the applied methods. In many cases, the spatial processes are of specific interest, and thus require the use of spatial regression models. However, spatial methods may be necessary for consistent estimators albeit the spatial processes not being of explicit interest. As Tobler's first law of geography puts it: 'everything is related to everything else, but near things are more related than distant things' (Tobler, 1970, p. 236). For instance, it seems plausible to assume that the house prices in one district are correlated to or even influenced by the house prices in neighbouring districts. In consequence, the observations cannot be considered independent and identically distributed (i.i.d.), which violates a standard assumption of linear regression models: $\mathrm{E}(\varepsilon_i \varepsilon_j) = \mathrm{E}(\varepsilon_i)\mathrm{E}(\varepsilon_j) = 0$, where ε is the random error term for each observation i and j ($i \neq j$). Intuitively, this violation results in erroneous inferential statistics when using the conventional OLS estimator of the equation $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$. However, spatial autocorrelation can also lead to biased point estimates, depending on the spatial process underlying the spatial correlation in the data (e.g. Pace & LeSage, 2010).

Several spatial model specifications exist to deal with this issue by explicitly modelling the spatial dependence. Using the spatial weights matrix $\boldsymbol{W}$, the most common spatial model specification is the spatial autoregressive model (SAR), which adds an endogenous spatially lagged dependent variable $\boldsymbol{Wy}$ to the conventional regression formula. Alternatively, the spatial error model (SEM) models the spatial dependence among the error terms $\boldsymbol{u} = \boldsymbol{Wu} + \boldsymbol{\varepsilon}$, and the spatial lag of X model (SLX) comprises the spatial lags of the exogenous covariates $\boldsymbol{WX}$. Further specifications use a combination of these most basic specifications. The spatial autoregressive combined model (SAC) includes autocorrelation in the dependent variable and the error term ($\boldsymbol{Wy}$ and $\boldsymbol{Wu}$), the spatial Durbin model (SDM) combines an autoregressive dependent variable and spatially lagged covariates ($\boldsymbol{Wy}$ and $\boldsymbol{WX}$), and the spatial Durbin error model (SDEM) comprises a spatial error term combined with spatially lagged covariates ($\boldsymbol{Wu}$ and $\boldsymbol{WX}$). The specification containing all three spatial terms ($\boldsymbol{Wy}$, $\boldsymbol{Wu}$ and $\boldsymbol{WX}$) is called general nesting spatial model (GNS).

On the one hand, those models provide a way to obtain unbiased point estimates of the true parameters. On the other hand, spatial models offer the opportunity to estimate spatial spillover coefficients, thereby informing the researcher about processes of spatial correlation or influence. Still, researchers have to consider several options, as the different specifications rely on different assumptions regarding the spatial dependence. Unfortunately, empirical specification tests of spatial regression models yield severe drawbacks and there is no general rule for selecting the correct model specification in applied research. Therefore, it is of substantial interest to evaluate the performance of spatial model specifications under different scenarios of misspecification.

This study conducts a systematic comparison of different spatial model specifications in different scenarios of spatial dependence by using Monte Carlo experiments. It demonstrates under which conditions conventional linear models yield biased estimates and how spatial model specifications perform throughout different scenarios of spatial dependence. The study extends previous simulations in several ways. First, it evaluates the bias of the model specifications by relying on the impact estimates rather than the point estimates, as the model impacts are the measures of interest in applied research (LeSage & Pace, 2017). Second, it systematically evaluates the performance of the spatial model specifications in the absence and presence of a non-spatial omitted variable bias. Third, it incorporates multiple explanatory variables with distinct spatial effects, as this resembles the case in applied research.

The results of the Monte Carlo experiments reveal that the most commonly used spatial models – spatial autoregressive model (SAR) and spatial error model (SEM) – have severe drawbacks for applied research. In line with previous findings, those models are outperformed by more flexible Durbin specifications incorporating spatially lagged covariates. However, the results also reveal that, under highly realistic conditions, SLX offers a better performance than the Durbin specifications, especially regarding the indirect spillover effects.

2. Theoretical background

As a first step in spatial econometrics, the researcher is required to specify the spatial relationship between the units of observation, or more precisely, to define which units j are neighbours of unit i for all units $i = \{1, 2, ..., N\}$. This is done by setting up an $N \times N$ dimensional neighbours weights matrix $\boldsymbol{W}$, where all elements $w_{ij} > 0$ for all neighbouring units i and j ($i \neq j$), and 0 otherwise. This study relies on a row-normalised contiguity weights matrix, defining all units as neighbours that share at least one common border. Several specifications for $\boldsymbol{W}$ exist, such as for example k nearest neighbour or distance based approaches (see e.g. Dubin, 2009), and choosing the correct or incorrect specification may be vital for the results. However, these aspects have been discussed elsewhere (Corrado & Fingleton, 2012; Elhorst & Halleck Vega, 2017; LeSage & Pace, 2014; Neumayer & Plümper, 2016), and the focus of this

study remains on the model specifications, thereby assuming a correctly specified $\boldsymbol{W}$.

2.1. *Model specifications*

As mentioned above, spatial dependence can be modelled in various ways (for a comprehensive introduction see e.g. Elhorst, 2014; LeSage & Pace, 2009). The most popular spatial model specification is the spatial autoregressive model (SAR), which incorporates the spatially weighted dependent variable $\boldsymbol{y}$ as an endogenous regressor at the right-hand side of the equation. The SAR model is defined as:

$$\boldsymbol{y} = \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{y}$ is an $N \times 1$ vector of the dependent variable, $\boldsymbol{W}$ as defined above, $\boldsymbol{X}$ an $N \times K$ matrix of $k = \{1, 2, ...K\}$ covariates, and $\boldsymbol{\varepsilon}$ an $N \times 1$ vector of normally distributed disturbances. $\boldsymbol{\beta}$ is a $K \times 1$ vector of parameter estimates and ρ represents the autoregressive scalar parameter.[1] This SAR specification assumes that the dependent variable of unit i is directly influenced by the spatially weighted dependent variable of neighbouring units j. For illustration purposes, let's consider an example analysing the effect of air quality and available green spaces on the house prices of a spatial district (e.g. Anselin & Lozano-Gracia, 2008). The house price is the dependent variable $\boldsymbol{y}$, while air quality and the availability of green spaces constitute the covariates $\boldsymbol{X}$. According to the specification in the SAR model, we would assume that the house prices in one district directly influence the house prices in neighbouring districts. An intuitive interpretation for this process might be that sellers or estate agents determine the prices based on the prices they observe in neighbouring districts.

Another specification of spatial models is the spatial error model (SEM). In contrast to the SAR specification, the SEM explicitly models spatial autocorrelation between the disturbances $\boldsymbol{u}$, represented by the scalar parameter λ. The SEM is defined as:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda \boldsymbol{W} \boldsymbol{u} + \boldsymbol{\varepsilon}. \end{aligned} \tag{2}$$

In this specification we assume that the spatial correlation between our units is caused by unobserved characteristics, which are either spatially clustered or follow a spatial pattern, and which are independent of the included covariates. Using the example above, we could for instance assume that spatially clustered or diffusing crime rates influence the house prices in the affected areas (but are independent of air quality and available green space).

A third approach does not incorporate the spatial dependence as an autoregressive term of the dependent variable or the error term, but directly models so called spatial 'spillover' effects by including the spatially lagged covariates into the equation. This

spatial lag of X (SLX) specification is defined as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \boldsymbol{\varepsilon}, \tag{3}$$

where $\boldsymbol{\theta}$ is an $K \times 1$ vector of spatial spillover parameters. This model incorporates the direct effects $\boldsymbol{\beta}$ of the covariates as well as the indirect spillover effects $\boldsymbol{\theta}$ from the covariates of neighbouring units. An important property of the SLX model is that $\boldsymbol{\theta}$ constitutes a $K \times 1$ vector, thus including a distinct spatial effect for each covariate. Here we would assume that house prices in the focal unit are not only influenced by the characteristics in the focal unit but also by the air quality and green space availability in neighbouring districts. Moreover, we can hypothesise that the neighbouring districts matter only regarding the green spaces but not regarding the air quality, as we receive a distinct spatial parameter for each covariate.

The three specifications shown above represent the most basic specifications of spatial models. Yet, there are further specifications which combine the models mentioned above. The spatial autoregressive combined model (SAC) comprises an autocorrelated dependent variable and an autocorrelated disturbance, resulting in:

$$\begin{aligned} \boldsymbol{y} &= \rho\boldsymbol{Wy} + \boldsymbol{X\beta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda\boldsymbol{Wu} + \boldsymbol{\varepsilon}. \end{aligned} \tag{4}$$

The spatial Durbin model (SDM), in contrast, combines the spatial spillover specification of the covariates (SLX) with the spatial autoregressive term of the dependent variable, resulting in:

$$\boldsymbol{y} = \rho\boldsymbol{Wy} + \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \boldsymbol{\varepsilon}. \tag{5}$$

A third combined model is the spatial Durbin error model (SDEM), combining the specifications of SEM and SLX:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda\boldsymbol{Wu} + \boldsymbol{\varepsilon}, \end{aligned} \tag{6}$$

thereby comprising the spatial spillover effects of the covariates as well as an autocorrelated disturbance term.

Combining all three basic model specifications mentioned above leads to the general nesting spatial model (GNS):

$$\begin{aligned} \boldsymbol{y} &= \rho\boldsymbol{Wy} + \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda\boldsymbol{Wu} + \boldsymbol{\varepsilon}. \end{aligned} \tag{7}$$

Though the GNS specification combines all the spatial processes of the previous specifications, this model only plays a minor role in applied research, as this specification – analogous to Manski's neighbourhood effects model (Manski, 1993) – is not or only weakly identifiable (Cook, Hays, & Franzese, 2015; Gibbons & Overman, 2012).[2] Though Burridge, Elhorst, and Zigova (2016) show that the GNS model can be identified in case of a grouped (block-diagonal) weights matrix, they find that even in this situation the model might be overparameterised, thereby providing no additional information over SDM or SDEM.

Note that most of the spatial model specifications cannot be estimated by Least Squares (LS), as using (constrained) LS estimators for models containing a spatially lagged dependent variable or disturbance leads to inconsistent results (e.g. Anselin & Bera, 1998; Franzese & Hays, 2007). However, an extensive amount of econometric literature discusses different estimation methods based on (quasi-) maximum likelihood (e.g. Anselin, 1988; Lee, 2004; Ord, 1975) or instrumental variable approaches using generalized methods of moments (e.g. Drukker, Egger, & Prucha, 2013; Kelejian & Prucha, 1998, 2010), in which the endogenous lagged variables can be instrumented by q higher order lags of the exogenous regressors $(\boldsymbol{X}, \boldsymbol{WX}, \boldsymbol{W}^2\boldsymbol{X}, ..., \boldsymbol{W}^q\boldsymbol{X})$ (Kelejian & Prucha, 1998).

2.2. *Local and global spatial impacts*

At first glance, the specifications presented above seem relatively similar in the way of modelling spatial effects. Yet, they differ in very important aspects. First, models with an endogenous spatial term (SAR, SAC, and SDM) assume a very different spatial dependence structure than models with only exogenous spatial terms as SLX and SDEM specifications. While the first three assume global spatial dependence, the second two assume local spatial dependence (Anselin, 2003; Halleck Vega & Elhorst, 2015; LeSage & Pace, 2009). Second, the interpretation of the coefficients differs greatly between models with and without endogenous effects. This becomes apparent when considering the reduced form of the equations above. Exemplary using the SAR model of (1), the reduced form is given by:

$$\begin{aligned} \boldsymbol{y} - \rho\boldsymbol{W}\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \\ (\boldsymbol{I}_N - \rho\boldsymbol{W})\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \\ \boldsymbol{y} &= (\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}). \end{aligned} \tag{8}$$

When subsequently taking the first derivative of the explanatory variable $\boldsymbol{x}_k$ from the reduced form in (8) to interpret the partial effect of a unit change in variable $\boldsymbol{x}_k$ on

$\boldsymbol{y}$, we receive

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}_k} = (\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1} \beta_k, \tag{9}$$

for each covariate $k = \{1, 2, ..., K\}$. As can be seen from (9), the partial derivative with respect to $\boldsymbol{x}_k$ produces an $N \times N$ matrix, thereby representing the partial effect of each unit i onto the focal unit i itself and all other units $j = \{1, 2, ..., i-1, i+1, ..., N\}$. The diagonal elements of (9) indicate how each unit i influences itself (change of x_i on change of y_i), and each off-diagonal elements in column i represents the effect of i on each other unit j (change of x_i on change of y_j). Since reporting the individual partial effects is usually not of interest, LeSage and Pace (2009) proposed to average over these effect matrices. While the average diagonal elements of the effects matrix resulting from (9) represent the so called direct[3] impacts of variable $\boldsymbol{x}_k$, the average column-sums of the off-diagonal elements represent the so called indirect impacts (or spatial spillover effects). The direct impacts refer to an average effect of a unit change in x_i on y_i, and the indirect (spillover) impacts exhibit how a change in x_i, on average, influences all neighbouring units y_j.

Table 1. Direct and indirect impacts, adopted from Halleck Vega and Elhorst (2015)

	Direct Impacts	Spatial Spillovers	Type
OLS/SEM	β_k	0	none
SAR/SAC	Diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \beta_k$	Off-diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \beta_k$	global
SLX/SDEM	β_k	θ_k	local
SDM/GNS	Diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1} [\beta_k + \boldsymbol{W} \theta_k]$	Off-diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1} [\beta_k + \boldsymbol{W} \theta_k]$	global

Table 1 summarises the direct and indirect impacts for all model specifications outlined above (adopted from Halleck Vega & Elhorst, 2015, p. 345). For OLS, SEM, SLX, and SDEM, the point estimates obtained in the regression models can be interpreted as partial (direct and indirect) impacts. However, in case of SAR, SAC, and SDM, point estimates differ from the partial derivatives (or impacts). Two important consequences follow from the impact estimates presented in Table 1.

First, for SAR, SAC, and SDM even the direct impacts differ from the point estimates. This results from the fact that an endogenous term of the dependent variable $\boldsymbol{Wy}$ contains feedback loops through the system of neighbours (Betz, Cook, & Hollenbach, 2019; Franzese & Hays, 2007; Halleck Vega & Elhorst, 2015). A change of x_i in the focal unit i influences the focal unit i itself, but also the neighbouring unit j, which in turn influences the focal unit i in a feedback loop. This feedback loop is part of the direct impact.

Second, the kind of indirect spillover effects in SAR, SAC, and SDM models differs from the kind of indirect spillover effects in SLX and SDEM models: while the first three specifications represent global spillover effects, the latter three represent local

spillover effects (Anselin, 2003; LeSage, 2014; LeSage & Pace, 2009). In case of SLX and SDEM the spatial spillover effects can be interpreted as the effect of a one unit change of $\boldsymbol{x}_k$ in the spatially weighted neighbouring observations on the dependent variable of the focal unit; when using a row-normalised contiguity weights matrix, $\boldsymbol{W}\boldsymbol{x}_k$ is the average value of $\boldsymbol{x}_k$ in the neighbouring units. Thus, only direct neighbours – as defined in $\boldsymbol{W}$ – contribute to those local spillover effects. In contrast, spillover effects in SAR, SAC, and SDM models do not only include direct neighbours but also neighbours of neighbours (second order neighbours) and further higher-order neighbours. This can be seen by rewriting the inverse in (9) as power series:[4]

$$(\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}\beta_k = (\boldsymbol{I}_N + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + \rho^3\boldsymbol{W}^3 + ...)\beta_k = (\boldsymbol{I}_N + \sum_{h=1}^{\infty}\rho^h\boldsymbol{W}^h)\beta_k, \quad (10)$$

where the identity matrix represents the direct effects and the sum represents the first and higher order indirect effects and the above mentioned feedback loops. This implies that a change in one unit i does not only affect the direct neighbours but passes through the whole system towards higher-order neighbours, where the impact declines with distance within the neighbouring system. Global indirect impacts thus are 'multiplied' by influencing direct neighbours as specified in $\boldsymbol{W}$ and indirect neighbours not connected according to $\boldsymbol{W}$, with additional feedback loops between those neighbours.

Note furthermore, that all diagonal elements $\text{diag}(\boldsymbol{W}) = w_{ii} = 0$, whereas the diagonal elements $\text{diag}(\boldsymbol{W}^2) = \text{diag}(\boldsymbol{W}\boldsymbol{W}) \neq 0$. Intuitively, $\rho\boldsymbol{W}$ only represents the effects between direct neighbours (and the focal unit is not a neighbour of the focal unit itself), whereas $\rho^2\boldsymbol{W}^2$ contains the effects of second order neighbours, where the focal unit is a second order neighbour of the focal unit itself. Thus, equation (10) includes feedback effects from $\rho^2\boldsymbol{W}^2$ on, which are part of the direct impacts according to the measures in Table 1.

In consequence, local and global spillover effects represent two distinct kinds of spatial spillover effects (LeSage, 2014). The interpretation of local spillover effects is straightforward: it represents the effect of all neighbours as defined by $\boldsymbol{W}$ (the average over all neighbours in case of a row-normalised weights matrix). For instance, the environmental quality in the focal unit itself but also in neighbouring units could influence the attractiveness of a district and its house prices. In this example it seems reasonable to assume that we have local spillover effects: only the environmental quality in directly contiguous units (e.g. in walking distance) is relevant for estimating the house prices. In contrast, interpreting global spillover effects can be a bit more difficult. Intuitively, the global spillover effects can be seen as a kind of diffusion process. For example, an exogenous event might increase the house prices in one district of a city, thus leading to an adaptation of house prices in neighbouring districts, which then leads to further adaptations in other units (the neighbours of the neighbours), thereby globally diffusing the effect of the exogenous event due to the endogenous term. Yet,

those processes happen over time. In a cross-sectional framework, the global spillover effects are hard to interpret. Anselin (2003) proposes an interpretation as an equilibrium outcome, where the partial impact represents an estimate of how this long-run equilibrium would change due to a change in $\boldsymbol{x}_k$ (see also LeSage, 2014).

2.3. *Bias in non-spatial OLS*

2.3.1. With exogenous covariates

So far, this study has summarised different spatial model specifications and discussed the types of spatial effects defined by those specifications. However, even if spatial effects are not of specific interest, spatial dependence can influence the estimation results. Non-spatial OLS models may not only exhibit erroneous inference but also biased estimates in some cases of spatial correlation. Still, it is important to distinguish between two kinds of biases, resulting from the discussion of direct and indirect impacts above. First, one could say that the unbiased estimate is the non-spatial parameter β_k. Second, one could also say that the unbiased estimate is the direct impact of $\boldsymbol{x}_k$, which does not only include the non-spatial effect but also the feedback loops. As discussed elsewhere (Gibbons & Overman, 2012; Gibbons, Overman, & Patacchini, 2015), if a researcher is interested in the treatment effect (e.g. of a political intervention), the total direct impact including feedback loops might be of more interest than the non-spatial effect. The non-spatial parameter β_k does not include feedback effects, and thus actually underestimates the impact of a chance in x_{ik} on y_i. Note that the non-spatial parameter reflects the first derivative of the non-reduced spatial regression formula as in (1), while the direct impacts (with feedback effects) are given by the diagonal elements of the first derivative of the reduced form as in (8). Thus, the choice of which effect is the correct 'treatment' effect does also affect the discussion under which conditions a non-spatial OLS model produces biased estimates of a spatial data generating process (DGP).

Suppose the DGP follows a GNS as defined in (7), but we erroneously assume the DGP was $\boldsymbol{y} = \boldsymbol{x}\beta + \boldsymbol{v}$, and use the OLS estimator $\hat{\beta} = (\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal\boldsymbol{y}$ for estimation of the parameter β. For simplicity, we will consider the case with only a single explanatory variable in the following section. As shown by Franzese and Hays (2007), using the non-reduced form of (7) as DGP leads to the following estimate for β:

$$\begin{aligned}\hat{\beta} =& (\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal(\rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{x}\theta + \lambda\boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon}) \\ \operatorname{plim}\hat{\beta} =& \beta + \rho\frac{\operatorname{Cov}(\boldsymbol{x}, \boldsymbol{W}\boldsymbol{y})}{\operatorname{Var}(\boldsymbol{x})} + \theta\frac{\operatorname{Cov}(\boldsymbol{x}, \boldsymbol{W}\boldsymbol{x})}{\operatorname{Var}(\boldsymbol{x})} + \lambda\frac{\operatorname{Cov}(\boldsymbol{x}, \boldsymbol{W}\boldsymbol{u})}{\operatorname{Var}(\boldsymbol{x})} + \frac{\operatorname{Cov}(\boldsymbol{x}, \boldsymbol{\varepsilon})}{\operatorname{Var}(\boldsymbol{x})}, \qquad (11)\end{aligned}$$

where $\boldsymbol{W}\boldsymbol{y}$, $\boldsymbol{W}\boldsymbol{x}$, $\boldsymbol{W}\boldsymbol{u}$, and $\boldsymbol{\varepsilon}$ are omitted variables, producing a bias similar to the conventional (non-spatial) omitted variable bias resulting from $\operatorname{Cov}(\boldsymbol{x}, \boldsymbol{v}) \neq 0$. If we assume that we do not suffer from a non-spatial omitted variable bias and $\boldsymbol{\varepsilon}$ is indepen-

dent and randomly distributed, $\mathrm{E}(\boldsymbol{\varepsilon}|\boldsymbol{x}) = 0$, therefore also $\mathrm{E}(\boldsymbol{W}\boldsymbol{u}|\boldsymbol{x}) = 0$, we get an expectation of 0 for the last two terms in (11). As in the standard case of an omitted variable, the bias resulting from estimating a non-spatial OLS if spatial dependence is present depends on $\mathrm{Cov}(\boldsymbol{x}, \boldsymbol{z})$ and $\mathrm{Cov}(\boldsymbol{y}, \boldsymbol{z})$, where $\boldsymbol{z}$ is the omitted variable $\boldsymbol{W}\boldsymbol{y}$ or $\boldsymbol{W}\boldsymbol{x}$.

Still, if one argues that the unbiased causal effect of a change in $\boldsymbol{x}$ is given by the direct impacts as described in Table 1, we need to rewrite the DGP in reduced form:

$$\boldsymbol{y} = (\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}(\boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{x}\theta + \boldsymbol{L}(\lambda)\boldsymbol{\varepsilon}), \tag{12}$$

where $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$ are independent and randomly distributed $\mathcal{N}(0, \sigma_x^2)$ and $\mathcal{N}(0, \sigma_\varepsilon^2)$ with a mean of zero, and for simplicity we define $\boldsymbol{L}(\lambda) = (\boldsymbol{I}_N - \lambda\boldsymbol{W})^{-1}$. Consequently, we derive at a different expectation for the estimate in non-spatial OLS models:

$$\begin{aligned}
\hat{\beta} =& (\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal(\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}(\boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{x}\theta + \boldsymbol{L}(\lambda)\boldsymbol{\varepsilon}) \\
=& (\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal(\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}\boldsymbol{x}\beta \\
&+ (\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal(\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}\boldsymbol{W}\boldsymbol{x}\theta \\
&+ (\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal(\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}\boldsymbol{L}(\lambda)\boldsymbol{\varepsilon}.
\end{aligned} \tag{13}$$

For independent random vectors $\boldsymbol{x}$ and symmetric real matrices $\boldsymbol{A}$, Girard (1989); Pace and LeSage (2010) show that $\mathrm{E}((\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal\boldsymbol{A}\boldsymbol{x}) = N^{-1}\mathrm{tr}(\boldsymbol{A})$. Furthermore, we know that $\mathrm{E}((\boldsymbol{x}^\intercal\boldsymbol{x})^{-1}\boldsymbol{x}^\intercal\boldsymbol{L}(\lambda)\boldsymbol{\varepsilon}) = 0$, as $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$ are independent, and therefore $\mathrm{E}(\boldsymbol{\varepsilon}|\boldsymbol{x}) = 0$. Thus, (13) can be simplified to

$$\operatorname{plim}\hat{\beta} = \frac{1}{N}\mathrm{tr}((\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}(\beta + \boldsymbol{W}\theta)), \tag{14}$$

which equals the average diagonal elements of $(\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}(\beta + \boldsymbol{W}\theta)$. Note that this equals exactly the summary measure for the direct impacts in the GNS model as defined by LeSage and Pace (2009) and described in Table 1. Thus, the non-spatial OLS provides an unbiased estimate of the direct impacts, though it does not provide an unbiased estimate of the parameter β. Betz et al. (2019) show in detail why the parameter estimates of a non-spatial model are even biased if the covariates are randomly distributed but interdependence in the dependent variable ($\rho > 0$) is present.

The discussion above shows that non-spatial OLS provides unbiased estimates of the direct impacts if $\boldsymbol{x}$ is exogenous even in case that the true DGP follows a process with spatial autocorrelation and spatial spillover effects. Thus, if we are not particularly interested in spatial effects, a spatially autocorrelated dependent variable does not necessarily require the use of spatial regression models. However, this is only valid if a change in $\boldsymbol{x}$ is exogenously determined and values of the covariates do not follow a spatial pattern. The following section shows in detail under which circumstances the

non-spatial estimator is unbiased and which circumstances lead to biased estimates in the non-spatial OLS model.

2.3.2. With spatially dependent covariates

Though the case of a spatially autocorrelated dependent variable and randomly distributed covariates might be possible, it often seems more plausible that both variables of interest exhibit at least some kind of spatial dependence when using spatial data. For instance in the example above, it seems reasonable to assume that house prices as well as environmental quality are spatially correlated.

To see what happens to the estimator, we will now consider the case in which not only the dependent but also the covariate $\boldsymbol{x}$ is spatially autocorrelated. Therefore, we define $\boldsymbol{x} = \delta \boldsymbol{W}\boldsymbol{x} + \boldsymbol{\upsilon}$, or in reduced form $\boldsymbol{x} = (\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{\upsilon}$, where $\boldsymbol{\upsilon} \sim \mathcal{N}(0, \sigma_\upsilon^2)$, and δ denotes the autocorrelation in $\boldsymbol{x}$. Consequently, we can rewrite (13) as:

$$\begin{aligned}\hat{\beta} =& \frac{\boldsymbol{\upsilon}^\intercal((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{\upsilon}}{\boldsymbol{\upsilon}^\intercal((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{\upsilon}}\beta \\ &+ \frac{\boldsymbol{\upsilon}^\intercal((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{W}\boldsymbol{\upsilon}}{\boldsymbol{\upsilon}^\intercal((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{\upsilon}}\theta \\ &+ \frac{\boldsymbol{\upsilon}^\intercal((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{L}(\lambda)\boldsymbol{\varepsilon}}{\boldsymbol{\upsilon}^\intercal((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\boldsymbol{\upsilon}}.\end{aligned} \tag{15}$$

Multiplying both numerator and denominator by $(\boldsymbol{\upsilon}^\intercal\boldsymbol{\upsilon})^{-1}$ (Pace & LeSage, 2010) and using $\mathrm{E}(\boldsymbol{\varepsilon}|\boldsymbol{x}) = 0$ leads to

$$\begin{aligned}\operatorname{plim}\hat{\beta} =& \frac{\operatorname{tr}\big[(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}\big]}{\operatorname{tr}\big[((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\big]}\beta \\ &+ \frac{\operatorname{tr}\big[(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}\boldsymbol{W}\big]}{\operatorname{tr}\big[((\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1})^\intercal(\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}\big]}\theta.\end{aligned} \tag{16}$$

To see that terms in (16) exceed the direct impacts in (14) for positive values of ρ and δ, we can rewrite the traces in (16) as the sum of the elements of the Hadamard product

$$\begin{aligned}\operatorname{plim}\hat{\beta} =& \frac{\sum_{ij}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho))_{ij}}{\operatorname{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal)}\beta \\ &+ \frac{\sum_{ij}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho)\boldsymbol{W})_{ij}}{\operatorname{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal)}\theta,\end{aligned} \tag{17}$$

where $\circ$ denotes the Hadamard product, $\boldsymbol{M}(\delta) = (\boldsymbol{I}_N - \delta \boldsymbol{W})^{-1}$, and $\boldsymbol{M}(\rho) = (\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}$. Now recall that $\boldsymbol{M}(\delta) = \boldsymbol{I}_N + \delta \boldsymbol{W} + \delta^2 \boldsymbol{W}^2 + ...$, thus all diagonal elements of $\boldsymbol{M}(\delta) > 1$ and all off-diagonal elements of $\boldsymbol{M}(\delta) \geq 0$ for $\delta > 0$, therefore $\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal$ is a non-negative matrix with all diagonal elements > 1. Similarly, $\boldsymbol{M}(\rho)$ is non-

negative with diagonal elements > 1. It follows that

$$\sum_{ij}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho))_{ij} > \mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho)), \tag{18}$$

and using $\mathrm{E}(\boldsymbol{a} \circ \boldsymbol{b}) = \mathrm{E}(\boldsymbol{a})\mathrm{E}(\boldsymbol{b}) + \mathrm{Cov}(\boldsymbol{a}, \boldsymbol{b})$ shows that when taking only the traces of (17) instead of the total sum (and leaving aside the positive off-diagonal elements):

$$\begin{aligned}\mathrm{E}\left(\frac{\mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho))}{\mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal)}\right) =& \frac{1}{N}\mathrm{tr}(\boldsymbol{M}(\rho)) \\ &+ \frac{\mathrm{Cov}(\mathrm{diag}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal), \mathrm{diag}(\boldsymbol{M}(\rho))}{N^{-1}\mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal)}.\end{aligned} \tag{19}$$

As both $\boldsymbol{M}(\delta)$ and $\boldsymbol{M}(\rho)$ are constructed from the same weights matrix $\boldsymbol{W}$, $\mathrm{Cov}(\mathrm{diag}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal), \mathrm{diag}(\boldsymbol{M}(\rho)) > 0$ for positive ρ, δ. Thus, it follows from (18) and (19) that the first term of (17) exceeds the first term of the direct impacts $N^{-1}\mathrm{tr}(\boldsymbol{M}(\rho)\beta)$ for positive values of δ and ρ. Similarly, the second term of (17) exceeds the second term of the direst impacts $N^{-1}\mathrm{tr}(\boldsymbol{M}(\rho)\boldsymbol{W}\theta)$ for positive values of δ and ρ, which adds an additional bias when $\theta > 0$. Thus, $\hat{\beta}_{OLS}$ exceeds the direct impacts of $N^{-1}\mathrm{tr}(\boldsymbol{M}(\rho)[\beta + \boldsymbol{W}\theta])$, and is upwardly biased for positive values of ρ and δ. Furthermore, this bias in the impacts is a non-linear function of the parameter estimates, thereby giving a strong motivation to compare the biases in impacts rather than coefficients. Note that the first part of (17) goes to $N^{-1}\mathrm{tr}(\boldsymbol{M}(\rho)\beta)$ only if either $\rho = 0$ or $\delta = 0$ (leading to $\boldsymbol{M}(\rho) = \boldsymbol{I}_N$, or $\boldsymbol{M}(\delta) = \boldsymbol{I}_N$ respectively), and the second term goes to $N^{-1}\mathrm{tr}(\boldsymbol{M}(\rho)\boldsymbol{W}\theta)$ only if $\delta = 0$. Obviously, the latter part of the bias also disappears if $\theta = 0$.

Note that the bias in the impacts as described above is not related to autocorrelation in the disturbances λ. Respectively, erroneously omitting spatial autocorrelation in the disturbances if $\lambda > 0$ would only lead to a loss of efficiency. However, this is only true if the disturbances are independent of our covariates. LeSage and Pace (2009); Pace and LeSage (2010) show that in the presence of an omitted variable and $\mathrm{E}(\boldsymbol{\varepsilon}|\boldsymbol{x}) \neq 0$, spatial correlation in the disturbances leads to an amplification of the non-spatial omitted variable bias. Replacing the random disturbance in (15) by $\boldsymbol{\varepsilon} = \gamma\boldsymbol{x} + \boldsymbol{\eta}$, where γ defines the covariance between the error term (or an omitted variable) and the covariate $\boldsymbol{x}$, adds an additional bias of the form

$$+\frac{\mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal\boldsymbol{M}(\rho)\boldsymbol{L}(\lambda))}{\mathrm{tr}(\boldsymbol{M}(\delta)^\intercal\boldsymbol{M}(\delta))}\gamma \tag{20}$$

to equation (17). Following the same argument as above, (20) is positive and $> \gamma$ for positive parameters ρ, δ, $\lambda > 0$, but also in the case of ρ, $\delta = 0$ and $\lambda > 0$. The term (20) goes to γ only if both ρ, $\lambda = 0$.

In sum, the discussion above provides some important conclusions regarding the

bias in non-spatial OLS models. First, if we agree that the unbiased effect estimate of a change in $\boldsymbol{x}_{ik}$ on $\boldsymbol{y}_{ik}$ is given by the average direct impacts, OLS provides unbiased estimates of the true effect even if the dependent variable is spatially autocorrelated and the covariates are exogenous and randomly distributed. Second, several quite realistic constellations of spatial correlation exist, in which non-spatial OLS produces biased estimates of the direct impacts. In particular, the non-spatial OLS estimator $\hat{\beta}_{OLS}$ is biased in the presence of either:

(1) Spatial autocorrelation in the dependent variable ($\rho \neq 0$) and spatial autocorrelation in the covariate ($\delta \neq 0$). This bias increases with ρ, δ, and β.
(2) Local spatial spillover effects ($\theta \neq 0$) and spatial autocorrelation in the covariate ($\delta \neq 0$). This is analogous to the omitted variable bias resulting from the omission of $\boldsymbol{Wx}$. It increases with θ and δ, but additionally with ρ if $\theta \neq 0$ and $\delta \neq 0$.
(3) An omitted variable and $\mathrm{E}(\boldsymbol{\varepsilon}|\boldsymbol{x}) \neq 0$. This non-spatial omitted variable bias γ is amplified by spatial dependence in the disturbances (λ) and spatial autocorrelation in the dependent variable (ρ), but also increases with positive values of δ if either $\rho \neq 0$ or $\lambda \neq 0$. Obviously, it also increases with γ.

3. Model selection

By showing that the non-spatial OLS estimates are biased in some constellations of spatial dependence, the previous chapter gives a strong motivation for the use of spatial regression models. However, as described in Chapter 2.1, a variety of spatial model specifications exist that can be used to account for the spatial structure of the data. Thus, selecting the correct model specification remains a crucial task in applied research.

3.1. *Specification tests*

One way of selecting the model specification is the application of empirical specification tests. In general, there are two different strategies: a specific-to-general or a general-to-specific approach (Florax, Folmer, & Rey, 2003; Mur & Angulo, 2009).

The specific-to-general approach is more common in spatial econometrics. This approach starts with the most basic non-spatial model and tests for possible misspecifications due to omitted autocorrelation in the error term or the dependent variable. Therefore, Anselin, Bera, Florax, and Yoon (1996) proposed to use Lagrange multiplier (LM) tests for the hypotheses H_0: $\lambda = 0$ and H_0: $\rho = 0$, which are robust against the alternative source of spatial dependence. The specific-to-general approach based on the robust LM test offers a good performance in distinguishing between SAR, SEM, and non-spatial OLS (Florax et al., 2003). Still, in their original paper, Anselin et al. (1996) already note the declining power of the robust LM_λ test for

spatial error dependence with increasing autocorrelation in the dependent variable (indicating some uncertainty under a SAC-like DGP). Furthermore, Mur and Angulo (2009) demonstrate strong drawbacks of the specific-to-general approach under non-optimal conditions like heteroscedasticity or endogeneity. Another shortcoming of this approach is its disregard of spatial dependence from local spillover effects, as resulting from an SLX-like process. Cook et al. (2015), for instance, show theoretically that an SLX-like dependence structure leads to the rejection of both hypotheses H_0: $\lambda = 0$ and H_0: $\rho = 0$, though no autocorrelation is present (see also Elhorst & Halleck Vega, 2017). This is also validated by simulation results: the LM test is not helpful in distinguishing between SDM- and SDEM-like processes, and produces heavily biased results in case of a GNS-like structure (see Table C1 of the Appendix).

The general-to-specific approach depicts the opposite method of specification search. This approach starts with the most general model and stepwise imposes restrictions on the parameters of this general model. In theory, one would start with a GNS specification and subsequently restrict the model to simplified specifications based on the significance of parameters in the GNS. The problem with this strategy is that the GNS is only weakly identified and, thus, is of little help in selecting the correct restrictions (Burridge et al., 2016). The most intuitive alternative would be to start with one of the two-source models SDM, SDEM, or SAC. This, however, bears the risk of imposing the wrong restriction in the first place (Cook et al., 2015). Furthermore, Cook et al. (2015) show that more complicated restrictions are necessary to derive all single-source models from SDEM or SAC specifications.

Thus, both ways of specification testing suffer from different sources of uncertainty, thereby making it hard to develop a general guideline for empirically finding the correct model specification. Consequently, scholars have developed other strategies to select the spatial model specification.

3.2. *Alternative strategies*

Some argue that the best way of choosing the appropriate model specification is to exclude one or more sources of spatial dependence – autocorrelation in the dependent variable, autocorrelation in the disturbances, or spatial spillover effects of the covariates – by design (Gibbons & Overman, 2012; Gibbons et al., 2015). Natural experiments are probably the best way of making one or more sources of spatial dependence unlikely, thereby restricting the model alternatives to a subset of all available models. However, the opportunities to use natural experiments are restricted in social sciences, making it a favourable but often impractical way of model selection. Therefore, Cook et al. (2015) argue that theoretical considerations should guide the model selection. First, it might be possible to rule out some sources of spatial dependence by theory, and thus restricting the specification alternatives to a subset. Second, theoretical mechanisms might guide the choice of either global or local spillover effects.

Still, others (Elhorst, 2014; LeSage, 2014; LeSage & Pace, 2009) argue that there are strong analytical reasons to restrict the model specifications to a subset, as the SDM subsumes the SLX and SAR model, and the SDEM subsumes SLX and SEM. It is easily observed that SDM reduces to SLX if $\rho = 0$ and to SAR if $\boldsymbol{\theta} = 0$, while the SDEM reduces to SLX if $\lambda = 0$ and to SEM if $\boldsymbol{\theta} = 0$. Less intuitively, Anselin (1988) has also shown that the SDM subsumes the SEM. Therefore, we can express the reduced form of (2) and rearrange terms:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{I}_N - \lambda \boldsymbol{W})^{-1}\boldsymbol{\varepsilon} \\ (\boldsymbol{I}_N - \lambda \boldsymbol{W})\boldsymbol{y} &= (\boldsymbol{I}_N - \lambda \boldsymbol{W})\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ (\boldsymbol{I}_N - \lambda \boldsymbol{W})\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} - \lambda \boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{y} &= (\boldsymbol{I}_N - \lambda \boldsymbol{W})^{-1}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}). \end{aligned} \tag{21}$$

Thus, the SEM constitutes a special case of an SDM with the relative simple restriction $\boldsymbol{\theta} = -\lambda\boldsymbol{\beta}$, meaning direct and indirect effects are constrained to a common factor (Anselin, 1988, 2003). The fact that SDM subsumes SAR, SLX, and SEM leads to the conclusion that applied research should only consider SDM and SDEM as model specifications (LeSage, 2014). Especially in the case of a likely omitted variable bias, LeSage and Pace (2009, p. 68) argue in favour of using the SDM.

Nonetheless, others propose to use the SLX specification as point of departure (Gibbons & Overman, 2012; Halleck Vega & Elhorst, 2015). First, scholars have argued that SAC and SDM models are only weakly identified in practice (Gibbons & Overman, 2012; Pinkse & Slade, 2010). Second, the global spillover specification in SAR, SAC, and SDM often seems to be theoretically implausible. Recall, for instance, the example of 'diffusing' house prices, where the house price in one district influences house prices in neighbouring districts. Specifying a SAR-like process means that house prices directly influence each other. Yet, it might be more plausible that house prices are driven by the demand within the focal, but also by the demand in neighbouring units, which reflects an SLX-like structure. Third, the SLX is computationally efficient, as it can be estimated by using Least Squares. Fourth, it turns attention back to the question of whether $\boldsymbol{X}$ and $\boldsymbol{W}\boldsymbol{X}$ are exogenous, which should be the main focus when investigating the dependence between $\boldsymbol{X}$ and $\boldsymbol{y}$. Furthermore, SLX, SDM, and SDEM share the advantage that all three models estimate flexible indirect spillover effects, which are not bound to a common ratio between direct and indirect effects for all covariates (as in SAR and SAC).

4. Monte Carlo experiment

As outlined in the previous section, diverging recommendations exist for selecting the model in applied research. To further improve the discussion on the selection of spatial

model specifications, this study compares the performance of different spatial model specifications by using a Monte Carlo experiment. Following the discussion above, a Monte Carlo experiment should consider several aspects to provide implications for applied research. First, as proposed by LeSage and Pace (2017), it is important to evaluate the bias of the impacts rather than the point estimates. The impacts are the measures of interest in applied research and the bias in impacts follows a non-linear function of the bias in parameter estimates. Second, it is important to incorporate more than one covariate, as bias in one parameter β_k could be counterbalanced by a bias in ρ, θ_k or λ, thus producing unbiased impacts for covariate k. However, a counterbalancing bias in ρ or λ could, in turn, affect the impacts of other covariates, as the autoregressive parameters affect all covariates. Third, we should evaluate the performance in two different worlds, one without omitted variable (omv) bias and one with omv bias, as this is likely to occur in applied research.

The DGP of the Monte Carlo simulation follows a GNS and is shown in (22) to (24), where $\boldsymbol{\upsilon}_k$ and $\boldsymbol{\varepsilon}$ are independent and randomly distributed $\mathcal{N}(0, \sigma_\upsilon^2)$ and $\mathcal{N}(0, \sigma_\varepsilon^2)$ with a mean of zero, and $\boldsymbol{x}_k$ is the kth column-vector of $\boldsymbol{X}$ for $k = 1, ..., K$ covariates (K is fixed at 2 in the simulations). The parameter ρ represents the autocorrelation in the dependent variable, λ the autocorrelation in the disturbances, and δ_k the autocorrelation in covariate k.

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{u}, \tag{22}$$

$$\boldsymbol{u} = \lambda \boldsymbol{W}\boldsymbol{u} + \boldsymbol{X}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{23}$$

$$\boldsymbol{x}_k = \delta_k \boldsymbol{W}\boldsymbol{x}_k + \boldsymbol{\upsilon}_k. \tag{24}$$

The parameter-vector $\boldsymbol{\gamma}$ in (23) specifies the correlation between $\boldsymbol{x}$ and the disturbance vector $\boldsymbol{u}$, thereby defining the strength of an omitted variable bias. In reduced form, this DGP can be written as

$$\begin{aligned} \boldsymbol{y} =& (\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1} \big[(\boldsymbol{I}_N - \delta_k \boldsymbol{W})^{-1} \boldsymbol{\upsilon}_k \beta_k \\ &+ \boldsymbol{W} (\boldsymbol{I}_N - \delta_k \boldsymbol{W})^{-1} \boldsymbol{\upsilon}_k \theta_k \\ &+ (\boldsymbol{I}_N - \lambda \boldsymbol{W})^{-1} ((\boldsymbol{I}_N - \delta_k \boldsymbol{W})^{-1} \boldsymbol{\upsilon}_k \gamma_k + \boldsymbol{\varepsilon}) \big]. \end{aligned} \tag{25}$$

The parameter vector $\boldsymbol{\beta}$ was fixed at $\boldsymbol{\beta} = \begin{pmatrix} 0.2 & 0.5 \end{pmatrix}^\intercal$, and the noise parameters were fixed at σ_υ^2, $\sigma_\varepsilon^2 = 1$ for all trials. All other parameters vary between the following two options for each parameter (vector):

- $\rho \in \{0, 0.5\}$,
- $\lambda \in \{0, 0.5\}$,
- $\boldsymbol{\delta} \in \left\{ \begin{pmatrix} 0 & 0 \end{pmatrix}^\intercal, \begin{pmatrix} 0.4 & 0.7 \end{pmatrix}^\intercal \right\}$,
- $\boldsymbol{\theta} \in \left\{ \begin{pmatrix} 0 & 0 \end{pmatrix}^\intercal, \begin{pmatrix} 0.1 & 0.8 \end{pmatrix}^\intercal \right\}$,

- $\boldsymbol{\gamma} \in \left\{ \begin{pmatrix} 0 & 0 \end{pmatrix}^{\intercal}, \begin{pmatrix} 0.3 & 0 \end{pmatrix}^{\intercal} \right\}$,

leading to a total of 32 distinct combinations. Note that this selection of parameters intentionally violates the common ratio assumption between direct and indirect effects, as this should be a more common case in practical research. All combinations were simulated in 1000 trials, with the same starting seed for each combination. All spatial models were estimated using *R*'s package *spdep* (Bivand & Piras, 2015).

The simulations build on a square grid of 900 observations and use a row-normalised ('queen') contiguity weights matrix $\boldsymbol{W}$, where all units sharing a common border are defined as neighbours, thus $w_{ij} > 0$ for contiguous units and $w_{ij} = 0$ otherwise. This setting leads to an average of 7.6 neighbours. Note that many alternative specifications may be chosen and that the choice should be made based on the assumed theoretical mechanisms (Neumayer & Plümper, 2016). However, contiguity is the most commonly used weights specification (e.g. Elhorst & Halleck Vega, 2017), and has been shown to be the best choice when no theoretical or empirical reasons justify a specific form of connectivity (Stakhovych & Bijmolt, 2009). Furthermore, the contiguity weights matrix exhibits a high level of similarity with a wide range of k nearest neighbours matrices (LeSage & Pace, 2014), as contiguous neighbours are (most likely) also the nearest neighbours. In accordance, additional simulations using different specifications of $\boldsymbol{W}$ lead to identical conclusions as presented in the following chapter.[5]

4.1. *Results without omv*

Figure 1 shows the bias of the direct and indirect impacts for the simulations without a non-spatial omitted variable bias. Respective numbers and the root mean squared error (RMSE) are shown in Appendix A. Several findings can be observed in the plot.

First and less surprising, all models perform reasonably well when correctly specifying the DGP. For instance, when the DGP follows a SAR-like process (e.g. line 2 and 4), the SAR model yields very precise estimates of direct and indirect impact. Similarly, SDEM yields the lowest bias if the DGP contains positive error-correlation $\lambda > 0$ and positive local spillover effects $\boldsymbol{\theta} > 0$, but no correlation in the dependent variable $\rho = 0$ (see line 13 and 15). These findings hold throughout all model specifications (though SAC could be seen as an exception).

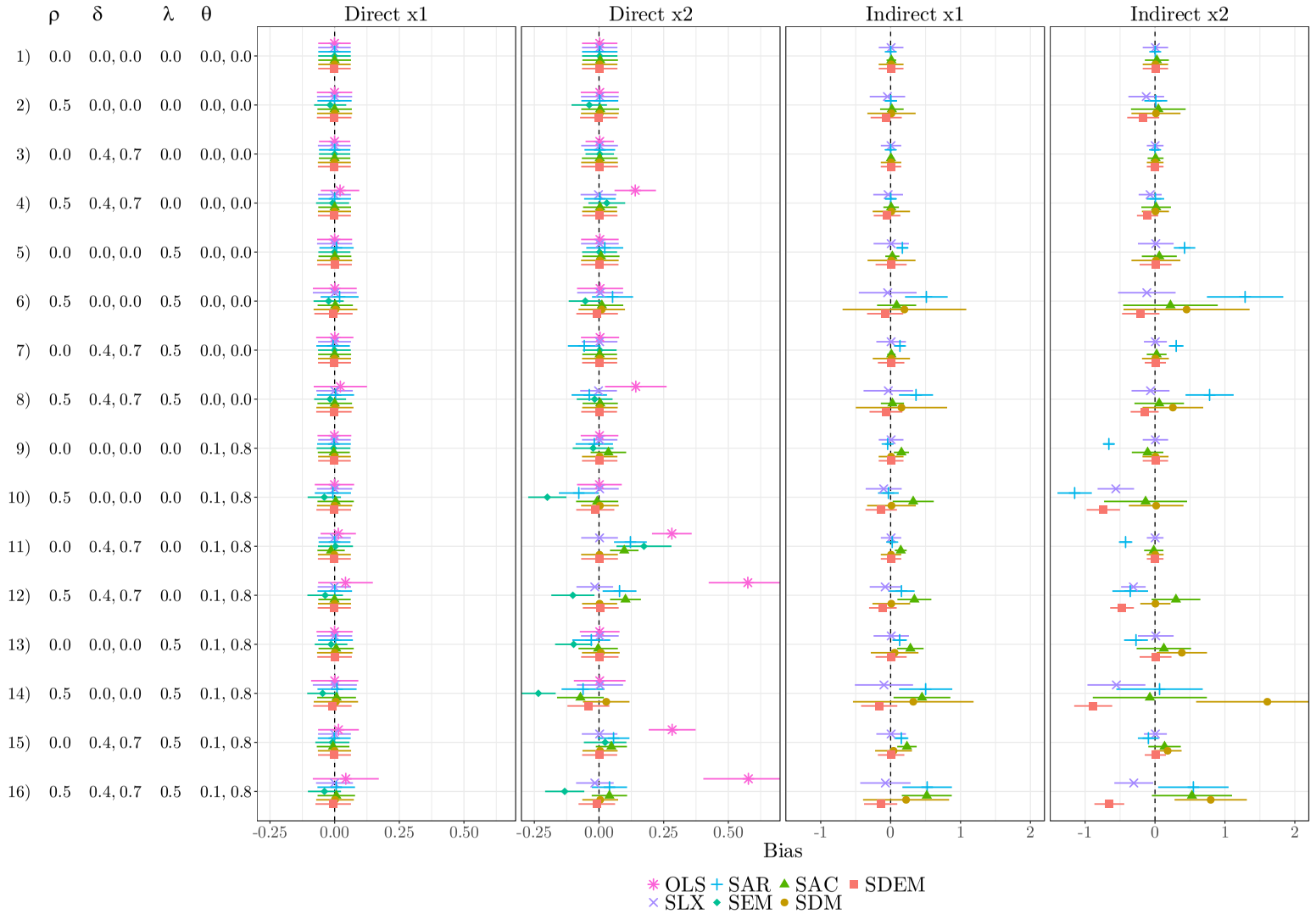

Figure 1. Bias of impacts and 95% confidence interval of empirical standard deviation without omv: $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$. ρ = autocorrelation in the dependent variable ($\boldsymbol{Wy}$); $\boldsymbol{\delta}$ = autocorrelation in the covariates ($\boldsymbol{x}_k = f(\boldsymbol{Wx}_k)$); λ = autocorrelation in the disturbances ($\boldsymbol{Wu}$); $\boldsymbol{\theta}$ = spatial spillover effects of covariates ($\boldsymbol{WX}$); $\boldsymbol{\gamma}$ = strength of omv.

Second, OLS yields an unbiased estimate of the direct impacts in many situations. The results confirm the theoretical predictions of Chapter 2.3: OLS estimates of the direct impacts are only biased in case of either positive autocorrelation in the dependent variable $\rho > 0$ *and* autocorrelation in the covariate $\delta_k > 0$ or local spillover effects $\theta_k > 0$ *and* autocorrelation in the covariate $\delta_k > 0$. Furthermore, this bias is rather moderate for low values of δ, θ, and β as can been seen in the first column of Figure 1. Note, however, that the bias is a conservative estimate, because the simulations use a relatively symmetric neighbours weights matrix. For instance, when increasing the variance in the number of neighbours per unit (thereby increasing the covariance between the diagonals of the inverted matrices $\boldsymbol{M}(\delta)$ and $\boldsymbol{M}(\rho)$), the bias in non-spatial OLS becomes more severe (see Equation 19 for the theoretical explanation).

Third, SLX, SDM, and SDEM all provide quite accurate estimates of the direct impacts (most visible in column 2). SAR, SEM, and SAC, in contrast, yield some drawbacks: especially in the presence of local spillover effects, these three specifications are biased (see lower part of Figure 1). Furthermore, SAR and SEM suffer from bias if autocorrelation in the disturbance and autocorrelation in the dependent variable are present simultaneously (see line 6 and 8). Though SLX is downwardly biased in case of autocorrelation in the dependent variable and the covariates (e.g. line 12 and 16), and SDM as well as SDEM yield some bias in case of a GNS-like process (line 14 and 16), those biases are rather moderate. This indicates that SLX, SDM, and SDEM are most robust against misspecification regarding the direct impacts.

Fourth, several differences exist regarding the indirect impacts. Most obviously, the often used SAR specification suffers from considerable bias: it overestimates indirect impacts in case of autocorrelation in the disturbances, and offers biased estimates if local spillover effects exist (which are not restricted to a common ratio). The latter also applies to SAC: though SAC offers relatively accurate estimates for $\boldsymbol{x}_2$, it overestimates indirect impacts for $\boldsymbol{x}_1$. Regarding the remaining three specifications – SLX, SDM, and SDEM – conclusions are less obvious. SDM and SDEM suffer from large bias for high values of $\boldsymbol{\theta}$ (see $\boldsymbol{x}_2$) if the DGP follows a GNS-like process (line 14 and 16): SDM overestimates the indirect impacts, while SDEM underestimates the indirect impacts. In addition, SDM performs badly if the true DGP is SDEM (line 13), and SDEM performs badly if the true DGP is SDM (line 10), whereas the bias increases with higher values of θ_k in both cases. Similar to SDEM, SLX underestimates the indirect impacts in presence of global spillovers / autocorrelation in the dependent variable.

For a better comparison, Figure 2 shows the bias of the indirect impacts in case of a GNS-like processes for different strengths of ρ and λ (for simplicity, we restrict the autocorrelation in covariates to zero and keep the local spillover effects fixed at $\boldsymbol{\theta} = (0.1, 0.8)^\intercal$). Respective numbers are shown in Appendix B. Three findings are particularly interesting. First, in a GNS-like situation, the bias in SDM grows with increasing autocorrelation in $\boldsymbol{y}$ (ρ) and increasing autocorrelation in the disturbances (λ). Second, the bias in SLX and SDEM increases with higher values of ρ, but is

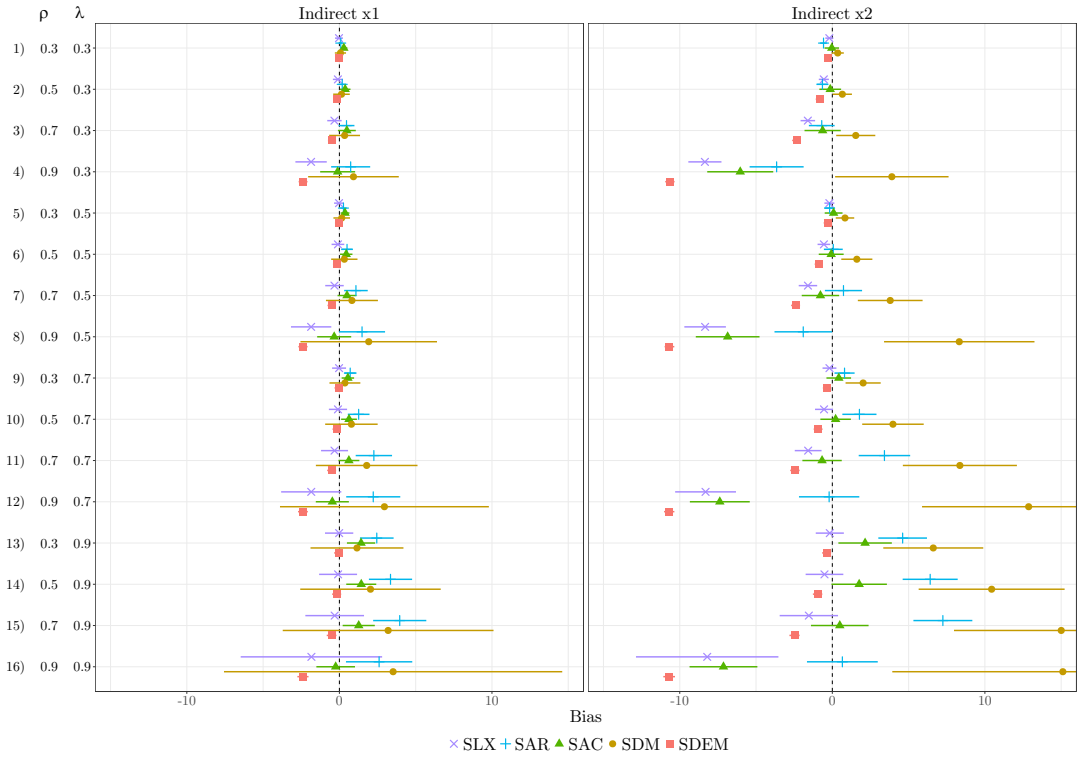

Figure 2. Bias of indirect impacts and 95% confidence interval of empirical standard deviation for different strengths of autocorrelation: $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$, $\boldsymbol{\delta} = (0, 0)^\intercal$, $\boldsymbol{\theta} = (0.1, 0.8)^\intercal$. ρ = autocorrelation in the dependent variable ($\boldsymbol{Wy}$); $\boldsymbol{\delta}$ = autocorrelation in the covariates ($\boldsymbol{x}_k = f(\boldsymbol{Wx}_k)$); λ = autocorrelation in the disturbances ($\boldsymbol{Wu}$); $\boldsymbol{\theta}$ = spatial spillover effects of covariates ($\boldsymbol{WX}$); $\boldsymbol{\gamma}$ = strength of omv.

unaffected from the strength of λ. Third, though SLX and SDEM suffer from the same problem, the bias from omitting global autocorrelation is less severe in SLX than in SDEM. Thus, the SLX outperforms SDEM. Furthermore, SLX outperforms SDM in most situations; only if the autocorrelation in the dependent variable is much stronger than the autocorrelation in the disturbances ($\rho = 0.9$, $\lambda = 0.3$), SDM yields lower bias than SLX.[6]

In sum, the results of the Monte Carlo experiments show that the three flexible model specifications of SLX, SDM, and SDEM offer an accurate estimate of the direct impacts, and all three specifications are relatively robust against misspecification regarding the direct impacts. However, results regarding the indirect impacts cast some doubt on the advice to consider only SDM or SDEM if no prior knowledge about the cause of spatial correlation is available. Especially in a 'mixed world' (where the true DGP is GNS), the results reveal that SLX offers a good alternative, which is more robust against misspecification in many situations.

4.2. *Results with omv*

So far, we have only considered the situation where $\boldsymbol{X}$ is perfectly exogenous in a non-spatial sense. However, in applied research one might often face situations in

which the covariates might be correlated with the disturbances. Thus, simulations in Figure 3 replicate previous simulations with an omitted variable correlated with $\boldsymbol{x}_1$ ($\boldsymbol{\gamma} = \begin{pmatrix} 0.3 & 0 \end{pmatrix}^\intercal$) in a non-spatial way.

Comparing column 1 of Figures 1 and 3 reveals that OLS suffers from a larger bias due to spatial autocorrelation if $\boldsymbol{x}$ is correlated with the disturbance, confirming results by Pace and LeSage (2010). Regarding the direct impacts of variable $\boldsymbol{x}_1$, which is affected by the omv bias, SEM exhibits the best estimation results. Yet, the good performance of SEM regarding one variable ($\boldsymbol{x}_1$) is somewhat offset by a relatively large bias in the second variable ($\boldsymbol{x}_2$), as the SEM tends to underestimate the impact of the second variable not affected by the omitted variable in case of positive spatial autocorrelation in $\boldsymbol{y}$. As in previous results, the bias in SLX, SDM, and SDEM is comparable and lower than in the remaining specifications.

Turning to the indirect impacts, SDEM reduces the bias in the variable affected by the omv compared to other model specifications. Still, it suffers (as is the case without omv) from a bad performance in the second variable if the DGP contains an autoregressive parameter of the dependent variable ($\rho > 0$). Furthermore, it becomes apparent that the underestimation of the indirect impacts in case of an GNS-like process without an omv bias (Figure 1) is somewhat counterbalanced by the positive omv bias leading to an overestimation of the impacts. For instance, in additional Monte Carlo experiments defining a downward bias ($\gamma_{x1} = -0.1$), SDEM amplifies the downward bias. Thus, it seems to depend on the constellation between the impacts and the omv bias whether SDEM reduces or increases the bias. However, a bias towards zero seems to be less severe than an upwardly biased parameter estimate. Another interesting finding is that the estimates of indirect impacts in SDM are affected most strongly by the omitted variable bias. Comparing column 3 in Figures 1 and 3 reveals that estimates in SDM show stronger changes than estimates of SLX and SDEM due to the non-spatial omitted variable (e.g. line 8 and 14). In sum, the second set of Monte Carlo simulations demonstrate that indirect impact estimates of SDEM and SLX are less affected by a non-spatial omv bias than SDM. Furthermore, in specific situations, SDEM may even help to reduce the non-spatial bias.

Additional simulations (available on request) reveal that the conclusions made above are robust to different variations of the parameters chosen in the DGP.[7] Obviously, as can be seen in Figure 2, the performance of SDM increases with increasing values of ρ relative to SDEM and SLX, and decreases with increasing values of λ (and vice versa). Still, SDM, SDEM, and SLX yield nearly equal biases regarding the direct impacts, while SLX and SDEM outperform the SDM in terms of the indirect impacts.

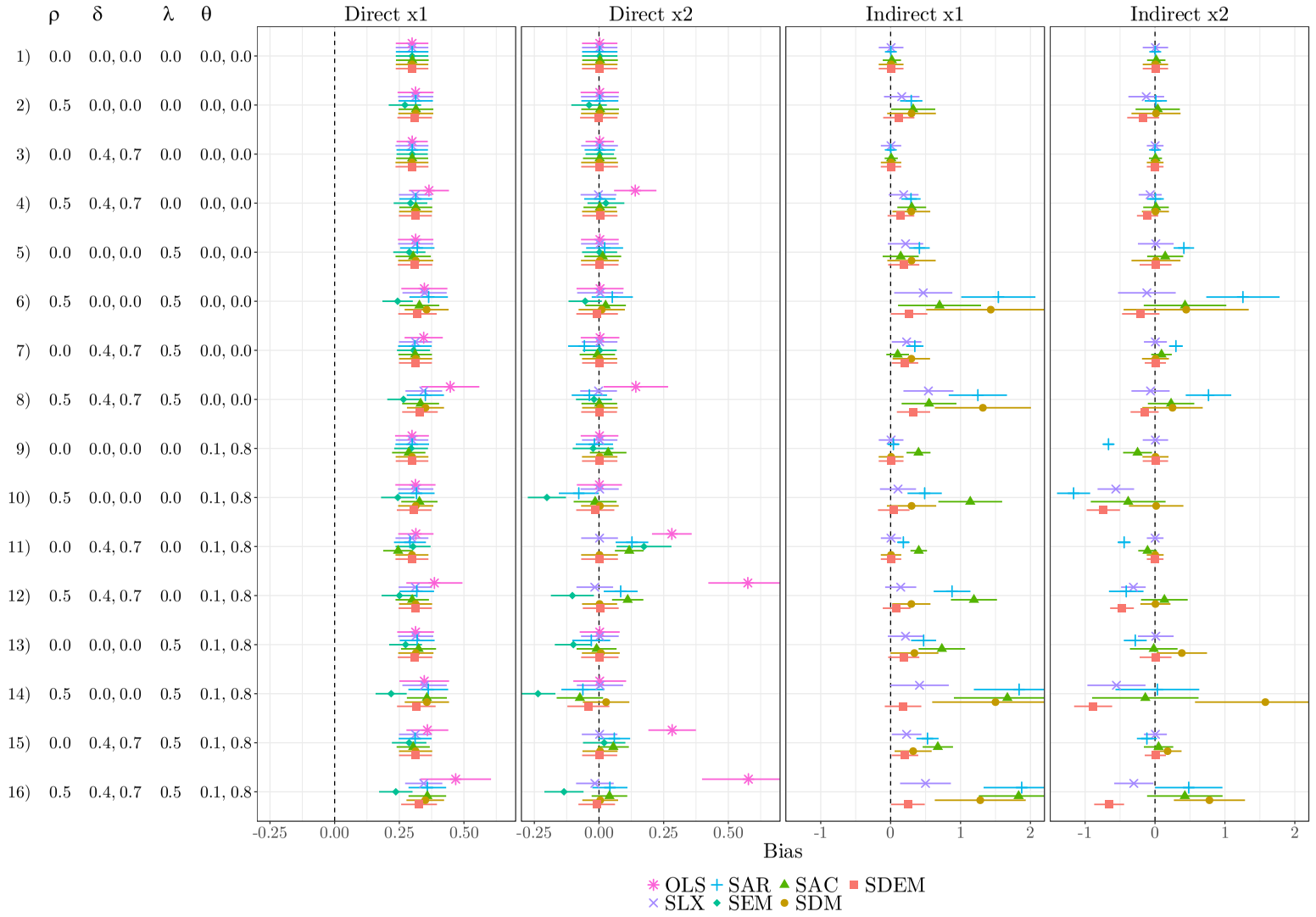

Figure 3. Bias of impacts and 95% confidence interval of empirical standard deviation with omv: $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0.3, 0)^\intercal$. ρ = autocorrelation in the dependent variable ($\boldsymbol{Wy}$); $\boldsymbol{\delta}$ = autocorrelation in the covariates ($\boldsymbol{x}_k = f(\boldsymbol{Wx}_k)$); λ = autocorrelation in the disturbances ($\boldsymbol{Wu}$); $\boldsymbol{\theta}$ = spatial spillover effects of covariates ($\boldsymbol{WX}$); $\boldsymbol{\gamma}$ = strength of omv.

5. Conclusion

The increasing availability of spatial or georeferenced data provides the possibility to investigate spatial research questions. However, analysing spatial data also requires careful consideration regarding the model specification. As has been shown theoretically and empirically, different constellations of the data exist which lead to biased estimates in non-spatial OLS models. Furthermore, non-spatial models disregard the spatial processes inherent in the data, thereby loosing interesting information. To overcome these problems, several spatial model specifications accounting for the spatial dependence can be employed.

Nevertheless, the variety of specifications also comes with the problem of selecting the correct specification in applied research, and specification tests are of little help in many situations. Therefore, this study employs a Monte Carlo experiment to systematically compare the bias of the most common spatial regression models in different situations of misspecification. In addition, this study extends previous simulations by relying on the impacts rather than the regression coefficients, as the impacts are the parameters of interest in applied research.

In line with previous studies (Elhorst, 2014; LeSage, 2014; LeSage & Pace, 2009), the Monte Carlo experiment reveals that the most commonly used SAR, SEM, and SAC specifications are outperformed by the more flexible specifications of SDM, SDEM, and SLX. Still, the results contradict the recommendation to consider only SDM and SDEM in applied research. While all three SDM, SDEM, and SLX show only marginal differences in the direct impacts, there are notable differences in the indirect spillover effects. Especially in a 'mixed world', in which the DGP follows a GNS-like structure, SLX produces less biased estimates of indirect impacts than SDM and SDEM. Note that this 'mixed world' resembles a very realistic scenario. In our example, one could argue in favour of all spatial processes: house prices in one district may directly influence house prices in neighbouring districts, house prices may depend on environmental quality in the focal and in contiguous districts, and clustered unobservables are likely to influence the house prices. Although SDEM can help to reduce an upward bias due to omitted variables in the presence of non-spatial endogeneity, this finding needs further investigation, as conclusions seem to depend on the constellation of impacts and omv bias.

In sum, the results of this paper support the claim recently made by Elhorst and Halleck Vega (2017); Halleck Vega and Elhorst (2015): if no theoretical reasons justify a specific model, it might be a better option to rely on the simple SLX specification rather than adopting the more complex SDM in applied research. Especially if it is not possible to eliminate one of the three sources of spatial dependence, SLX seems to be a good choice. Furthermore, the SLX is computational simple and intuitively interpretable. In contrast to global spillover effects in SDM, the local spillover effects can be interpreted as the effects of the spatially weighted neighbours (as defined by

$\boldsymbol{W}$). Beyond that, the spatial spillover effects in SLX can also be 'globalised' to some extent by including separate terms for second or even higher order neighbours, which, in addition, would be a more flexible function than higher order terms of one autoregressive parameter ρ. Moreover, Elhorst and Halleck Vega (2017) show that the SLX – in contrast to other specifications – allows to empirically parameterise the weights matrix $\boldsymbol{W}$, which can lead to improved estimates of the spillover effects. Taken together, SLX certainly provides a worthwhile alternative to SDM and SDEM.

Still, it is important to keep in mind that spatial models only give parameter estimates for (conditional) correlations. Simply estimating a model using cross-sectional observational data hardly tells anything about the causal mechanisms underlying these correlations. The causal process underlying the spatial correlation can be the result of either 1) spatial interdependence in the dependent variable, 2) spillover effects in the covariates, or 3) common unobserved shocks. To identify the (spatial) causal effects between two variables of interest, it is necessary to use designs or methods following the counterfactual approach like natural experiments (Angrist & Pischke, 2009, 2015; Morgan & Winship, 2015). However, this applies to all empirical research, including non-spatial observational studies. Thus, as is the case with all observational studies, spatial models can provide interesting insights into the correlational structure of the data, but can only be a first step in evaluating the causal mechanisms.

Notes

[1]The model intercept is omitted in all models for simplicity.

[2]Gibbons and Overman (2012) even argue that the SDM is only weakly identified in practice.

[3]Though previous literature (e.g. Halleck Vega & Elhorst, 2015; LeSage & Pace, 2009) has established the notation of direct and indirect impacts, it is important to note that also the direct impacts comprise a spatial 'multiplier' component if we specify an endogenous lagged depended variable, as a change in $\boldsymbol{x}_i$ influences $\boldsymbol{y}_i$, which influences $\boldsymbol{y}_j$, which in turn influences $\boldsymbol{y}_i$ (see also equation 10).

[4]A power series of $\sum_{k=0}^{\infty} \boldsymbol{W}^k$ converges to $(\boldsymbol{I}-\boldsymbol{W})^{-1}$ if the maximum absolute eigenvalue of $\boldsymbol{W} < 1$, which is ensured by standardising $\boldsymbol{W}$.

[5]See Figure D1 for results using inverse distance weighted 10-nearest neighbours and Figure D2 for results with maximum eigenvalue-normalised inverse distances weights (cut-off at 100 neighbours for computational efficiency) and a distance decay factor set to 1. Though, in line with previous research (Stakhovych & Bijmolt, 2009), the dispersion of the indirect impacts strongly increases with higher levels of connectivity (see Figure D2), the conclusions regarding the relative performance of the different specifications remain unchanged. Additional simulations using the Oklahoma census tract shape file (N=1,046) instead of a symmetric grid produced identical results (not shown).

[6]Note that the SAC yields relatively low biases for the indirect impacts in GNS-like processes, but at the same time produces relative large biases in the direct impacts (see Table B1).

[7]Variations were specified as follows: $\rho \in \{0.3, 0.7\}$, $\lambda \in \{0.3, 0.7\}$, $\boldsymbol{\theta} = \begin{pmatrix} 0.1 & 0.4 \end{pmatrix}^{\intercal}$, and $\boldsymbol{\gamma} \in \{\begin{pmatrix} -0.1 & 0 \end{pmatrix}^{\intercal}, \begin{pmatrix} -0.3 & 0 \end{pmatrix}^{\intercal}, \begin{pmatrix} 0.3 & 0.2 \end{pmatrix}^{\intercal}\}$.

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton Univ. Press.

Angrist, J. D., & Pischke, J.-S. (2015). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton: Princeton Univ. Press.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

Anselin, L. (2003). Spatial Externalities, Spatial Multipliers, and Spatial Econometrics. *International Regional Science Review*, *26*(2), 153–166.

Anselin, L., & Bera, A. K. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In A. Ullah & D. E. A. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237–289). New York: Dekker.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple Diagnostic Tests for Spatial Dependence. *Regional Science and Urban Economics*, *26*(1), 77–104.

Anselin, L., & Lozano-Gracia, N. (2008). Errors in Variables and Spatial Effects in Hedonic House Price Models of Ambient Air Quality. *Empirical Economics*, *34*(1), 5–34.

Betz, T., Cook, S. J., & Hollenbach, F. M. (2019). Spatial Interdependence and Instrumental Variable Models. *Political Science Research and Methods*, *60*, 1–16.

Bivand, R., & Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. *Journal of Statistical Software*, *63*(18), 1–36.

Burridge, P., Elhorst, J. P., & Zigova, K. (2016). Group Interaction in Research and the Use of General Nesting Spatial Models. In B. H. Baltagi, J. P. LeSage, & R. K. Pace (Eds.), *Spatial Econometrics: Qualitative and Limited Dependent Variables* (Vol. 37, pp. 223–258). Emerald Group Publishing Limited.

Cook, S. J., Hays, J. C., & Franzese, R. J. (2015). *Model Specification and Spatial Interdependence: Paper prepared for the 2015 Summer Methods Meeting.* Retrieved 09.07.2017, from `http://www.sas.rochester.edu/psc/polmeth/papers/Cook_Hays_Franzese.pdf`

Corrado, L., & Fingleton, B. (2012). Where is the Economics in Spatial Econometrics? *Journal of Regional Science*, *52*(2), 210–239.

Crowder, K., Hall, M., & Tolnay, S. E. (2011). Neighborhood Immigration and Native Out-Migration. *American Sociological Review*, *76*(1), 25–47.

Dokshin, F. A. (2016). Whose Backyard and What's at Issue? Spatial and Ideological Dynamics of Local Opposition to Fracking in New York State, 2010 to 2013. *American Sociological Review*, *81*(5), 921–948.

Downey, L. (2006). Using Geographic Information Systems to Reconceptualize Spatial Relationships and Ecological Context. *American Journal of Sociology*, *112*(2), 567–612.

Drukker, D. M., Egger, P., & Prucha, I. R. (2013). On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors. *Econometric Reviews*, *32*(5-6), 686–733.

Dubin, R. (2009). Spatial Weights. In A. S. Fotheringham & P. Rogerson (Eds.), *The Sage Handbook of Spatial Analysis* (pp. 125–158). Los Angeles and London: Sage.

Elhorst, J. P. (2014). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. Berlin and Heidelberg: Springer.

Elhorst, J. P., & Halleck Vega, S. (2017). The SLX Model: Extensions and the Sensitivity of Spatial Spillovers to W. *Papeles de Economía Española*, *152*, 34–50.

Florax, R., Folmer, H., & Rey, S. J. (2003). Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology. *Regional Science and Urban Economics*, *33*(5), 557–579.

Franzese, R. J., & Hays, J. C. (2007). Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data. *Political Analysis*, *15*(2), 140–164.

Friedrichs, J., Galster, G., & Musterd, S. (2003). Neighbourhood Effects on Social Opportunities: The European and American Research and Policy Context. *Housing Studies*, *18*(6), 797–806.

Gibbons, S., & Overman, H. G. (2012). Mostly Pointless Spatial Econometrics? *Journal of Regional Science*, *52*(2), 172–191.

Gibbons, S., Overman, H. G., & Patacchini, E. (2015). Spatial Methods. In G. Duranton, J. V. Henderson, & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics* (Vol. 5, pp. 115–168). Amsterdam: Elsevier.

Girard, D. A. (1989). A Fast 'Monte-Carlo Cross-Validation' Procedure for Large Least Squares Problems with Noisy Data. *Numerische Mathematik*, *56*(1), 1–23.

Halleck Vega, S., & Elhorst, J. P. (2015). The SLX Model. *Journal of Regional Science*, *55*(3), 339–363.

Kelejian, H. H., & Prucha, I. R. (1998). A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *The Journal of Real Estate Finance and Economics*, *17*(1), 99–121.

Kelejian, H. H., & Prucha, I. R. (2010). Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. *Journal of Econometrics*, *157*(1), 53–67.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, *75*(1), 83–119.

Lee, L.-f. (2004). Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica*, *72*(6), 1899–1925.

Legewie, J., & Schaeffer, M. (2016). Contested Boundaries: Explaining where Ethnoracial Diversity Provokes Neighborhood Conflict. *American Journal of Sociology*, *122*(1), 125–161.

LeSage, J. P. (2014). What Regional Scientists Need to Know about Spatial Econometrics. *The Review of Regional Studies*, *44*(1), 13–32.

LeSage, J. P., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Boca Raton: CRC Press.

LeSage, J. P., & Pace, R. K. (2014). The Biggest Myth in Spatial Econometrics. *Econometrics*, *2*(4), 217–249.

LeSage, J. P., & Pace, R. K. (2017). Spatial Econometric Monte Carlo Studies: Raising the Bar. *Empirical Economics*, in press.

Lichter, D. T., Parisi, D., & Taquino, M. C. (2015). Toward a New Macro-Segregation? Decomposing Segregation within and between Metropolitan Cities and Suburbs. *American Sociological Review*, *80*(4), 843–873.

Logan, J. R. (2012). Making a Place for Space: Spatial Thinking in Social Science. *Annual Review of Sociology*, *38*, 507–524.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531–542.

Morgan, S. L., & Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge Univ. Press.

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

Neumayer, E., & Plümper, T. (2016). W. *Political Science Research and Methods*, *4*(01), 175–193.

Ord, J. K. (1975). Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association*, *70*(349), 120–126.

Pace, R. K., & LeSage, J. P. (2010). Omitted Variable Biases of OLS and Spatial Lag Models. In A. Páez, J. Gallo, R. N. Buliung, & S. Dall'erba (Eds.), *Progress in Spatial Analysis* (pp. 17–28). Berlin and Heidelberg: Springer.

Pinkse, J., & Slade, M. E. (2010). The Future of Spatial Econometrics. *Journal of Regional Science*, *50*(1), 103–117.

Reardon, S. F., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The Geographic Scale of Metropolitan Racial Segregation. *Demography*, *45*(3), 489–514.

Rüttenauer, T. (2018). Neighbours Matter: A Nation-Wide Small-Area Assessment of Environmental Inequality in Germany. *Social Science Research*, *70*, 198–211.

Sampson, R. J., Morenoff, J. D., & Earls, F. (1999). Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children. *American Sociological Review*, *64*(5), 633–660.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual Review of Sociology*, *28*(1), 443–478.

Stakhovych, S., & Bijmolt, T. H. (2009). Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, *88*(2), 389–408.

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240.

Appendix A. Results of Monte Carlo experiment

Table A1. RMSE of direct impacts without omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$.

				OLS		SLX		SAR		SEM		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.0321	0.0350*	0.0322	0.0350	0.0321	0.0351	0.0321	0.0350	0.0323	0.0353	0.0321*	0.0351	0.0322	0.0350
0.5	0, 0	0.0	0, 0	0.0348	0.0376	0.0346	0.0373	0.0337*	0.0365*	0.0361	0.0514	0.0339	0.0375	0.0346	0.0374	0.0341	0.0369
0.0	0.4, 0.7	0.0	0, 0	0.0308*	0.0281*	0.0325	0.0363	0.0308	0.0304	0.0308	0.0282	0.0310	0.0354	0.0326	0.0363	0.0325	0.0363
0.5	0.4, 0.7	0.0	0, 0	0.0433	0.1459	0.0325	0.0353	0.0322*	0.0313*	0.0335	0.0471	0.0323	0.0338	0.0326	0.0347	0.0325	0.0348
0.0	0, 0	0.5	0, 0	0.0345	0.0372	0.0346	0.0373	0.0346	0.0427	0.0317*	0.0346*	0.0325	0.0373	0.0346	0.0372	0.0341	0.0370
0.5	0, 0	0.5	0, 0	0.0432	0.0456	0.0433	0.0455	0.0421	0.0663	0.0377	0.0623	0.0348*	0.0436	0.0433	0.0470	0.0380	0.0416*
0.0	0.4, 0.7	0.5	0, 0	0.0368	0.0379	0.0324	0.0352	0.0336	0.0655	0.0323*	0.0334*	0.0325	0.0344	0.0326	0.0348	0.0325	0.0347
0.5	0.4, 0.7	0.5	0, 0	0.0570	0.1547	0.0360	0.0363	0.0367	0.0513	0.0360	0.0392	0.0341*	0.0351*	0.0366	0.0352	0.0350	0.0353
0.0	0, 0	0.0	0.1, 0.8	0.0334	0.0372	0.0322	0.0350*	0.0333	0.0411	0.0336	0.0463	0.0317*	0.0504	0.0323	0.0350	0.0322	0.0350
0.5	0, 0	0.0	0.1, 0.8	0.0388	0.0441	0.0348	0.0377	0.0363	0.0875	0.0517	0.2033	0.0361	0.0422	0.0347	0.0373*	0.0344*	0.0402
0.0	0.4, 0.7	0.0	0.1, 0.8	0.0375	0.2843	0.0325	0.0363	0.0313	0.1256	0.0347	0.1820	0.0303*	0.1016	0.0325	0.0363*	0.0325	0.0363
0.5	0.4, 0.7	0.0	0.1, 0.8	0.0684	0.5806	0.0329	0.0397	0.0341	0.0861	0.0507	0.1096	0.0320*	0.1068	0.0326	0.0347*	0.0328	0.0360
0.0	0, 0	0.5	0.1, 0.8	0.0357	0.0393	0.0346	0.0373	0.0345	0.0478	0.0351	0.1047	0.0352	0.0390	0.0349	0.0379	0.0341*	0.0370*
0.5	0, 0	0.5	0.1, 0.8	0.0466	0.0510	0.0434	0.0458*	0.0396	0.0749	0.0555	0.2369	0.0390	0.0853	0.0440	0.0537	0.0388*	0.0583
0.0	0.4, 0.7	0.5	0.1, 0.8	0.0428	0.2866	0.0324*	0.0352	0.0325	0.0644	0.0347	0.0482	0.0327	0.0568	0.0327	0.0349	0.0325	0.0347*
0.5	0.4, 0.7	0.5	0.1, 0.8	0.0781	0.5844	0.0364	0.0404	0.0375	0.0533	0.0512	0.1381	0.0375	0.0532	0.0372	0.0356*	0.0355*	0.0370
Average				0.0434	0.1518	0.0348	0.0379	0.0347	0.0587	0.0386	0.0875	0.0336*	0.0517	0.0350	0.0377*	0.0340	0.0379
				0.0976		0.0363		0.0467		0.0631		0.0427		0.0363		0.0359*	

* Lowest RMSE for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A2. RMSE of indirect impacts without omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$.

				SLX		SAR		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.0903	0.0929	0.0121*	0.0302*	0.0366	0.0914	0.0909	0.0927	0.0907	0.0928
0.5	0, 0	0.0	0, 0	0.1371	0.1807	0.0449*	0.0843*	0.0876	0.2049	0.1766	0.1796	0.1302	0.2053
0.0	0.4, 0.7	0.0	0, 0	0.0745	0.0610	0.0113*	0.0280*	0.0240	0.0587	0.0748	0.0610	0.0748	0.0609
0.5	0.4, 0.7	0.0	0, 0	0.1143	0.1090	0.0406*	0.0644*	0.0572	0.1091	0.1368	0.0988	0.1098	0.1315
0.0	0, 0	0.5	0, 0	0.1292	0.1307	0.1728	0.4297	0.0570*	0.1415	0.1754	0.1789	0.1136	0.1165*
0.5	0, 0	0.5	0, 0	0.2150	0.2402*	0.5339	1.3185	0.1666	0.4097	0.4925	0.6443	0.1535*	0.2453
0.0	0.4, 0.7	0.5	0, 0	0.1082	0.0838	0.1377	0.3076	0.0317*	0.0760*	0.1363	0.0980	0.0968	0.0773
0.5	0.4, 0.7	0.5	0, 0	0.1837	0.1526*	0.3835	0.8000	0.0882*	0.1904	0.3674	0.3380	0.1374	0.1802
0.0	0, 0	0.0	0.1, 0.8	0.0903	0.0929	0.0467*	0.6617	0.1640	0.1566	0.0909	0.0953	0.0907	0.0928*
0.5	0, 0	0.0	0.1, 0.8	0.1648	0.5762	0.0820*	1.1564	0.3560	0.3315	0.1781	0.1999*	0.1771	0.7493
0.0	0.4, 0.7	0.0	0.1, 0.8	0.0745	0.0610	0.0341*	0.4238	0.1511	0.0726	0.0749	0.0618	0.0748	0.0609*
0.5	0.4, 0.7	0.0	0.1, 0.8	0.1373*	0.3220	0.1817	0.3776	0.3614	0.3485	0.1381	0.1104*	0.1500	0.4798
0.0	0, 0	0.5	0.1, 0.8	0.1292	0.1307	0.1388	0.2855	0.2987	0.2354	0.1831	0.4253	0.1136*	0.1165*
0.5	0, 0	0.5	0.1, 0.8	0.2328	0.5913	0.5372	0.3217*	0.4937	0.4227	0.5462	1.6884	0.2103*	0.8949
0.0	0.4, 0.7	0.5	0.1, 0.8	0.1082	0.0838	0.1610	0.1220	0.2438	0.1807	0.1404	0.2072	0.0968*	0.0773*
0.5	0.4, 0.7	0.5	0.1, 0.8	0.1983	0.3353*	0.5526	0.6058	0.5483	0.6037	0.3837	0.8393	0.1870*	0.6627
Average				0.1367	0.2028*	0.1919	0.4386	0.1979	0.2271	0.2116	0.3324	0.1254*	0.2652
				0.1697*		0.3153		0.2125		0.2720		0.1953	

* Lowest RMSE for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A3. RMSE of direct impacts with omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0.3, 0)^\intercal$.

				OLS		SLX		SAR		SEM		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.3011	0.0350*	0.3010*	0.0350	0.3014	0.0350	0.3011	0.0350	0.3021	0.0354	0.3010	0.0352	0.3010	0.0350
0.5	0, 0	0.0	0, 0	0.3148	0.0379	0.3152	0.0373	0.3152	0.0366*	0.2736*	0.0520	0.3162	0.0374	0.3154	0.0374	0.3112	0.0369
0.0	0.4, 0.7	0.0	0, 0	0.3014	0.0281*	0.3007	0.0363	0.3015	0.0301	0.3014	0.0282	0.3014	0.0330	0.3007*	0.0364	0.3007	0.0363
0.5	0.4, 0.7	0.0	0, 0	0.3668	0.1462	0.3137	0.0353	0.3153	0.0311*	0.2951*	0.0451	0.3156	0.0326	0.3147	0.0348	0.3137	0.0348
0.0	0, 0	0.5	0, 0	0.3151	0.0373	0.3153	0.0373	0.3213	0.0424	0.2906*	0.0347*	0.3056	0.0391	0.3154	0.0374	0.3128	0.0370
0.5	0, 0	0.5	0, 0	0.3499	0.0466	0.3504	0.0456	0.3655	0.0655	0.2453*	0.0627	0.3301	0.0474	0.3583	0.0469	0.3226	0.0416*
0.0	0.4, 0.7	0.5	0, 0	0.3464	0.0384	0.3141	0.0352	0.3121	0.0651	0.3067*	0.0336*	0.3130	0.0355	0.3147	0.0347	0.3141	0.0347
0.5	0.4, 0.7	0.5	0, 0	0.4507	0.1559	0.3466	0.0363	0.3527	0.0511	0.2680*	0.0405	0.3345	0.0352*	0.3531	0.0353	0.3318	0.0353
0.0	0, 0	0.0	0.1, 0.8	0.3009	0.0372	0.3010	0.0350*	0.3018	0.0408	0.2975	0.0463	0.2884*	0.0505	0.3009	0.0351	0.3010	0.0350
0.5	0, 0	0.0	0.1, 0.8	0.3149	0.0445	0.3152	0.0377	0.3184	0.0876	0.2464*	0.2050	0.3297	0.0449	0.3155	0.0373*	0.3094	0.0403
0.0	0.4, 0.7	0.0	0.1, 0.8	0.3159	0.2843	0.3007	0.0363	0.2929	0.1314	0.3051	0.1820	0.2472*	0.1203	0.3007	0.0363*	0.3007	0.0363
0.5	0.4, 0.7	0.0	0.1, 0.8	0.3896	0.5807	0.3135	0.0397	0.3195	0.0902	0.2530*	0.1114	0.3018	0.1151	0.3147	0.0347*	0.3131	0.0360
0.0	0, 0	0.5	0.1, 0.8	0.3149	0.0395	0.3153	0.0373	0.3210	0.0478	0.2762*	0.1057	0.3262	0.0405	0.3158	0.0380	0.3128	0.0370*
0.5	0, 0	0.5	0.1, 0.8	0.3499	0.0522	0.3505	0.0459*	0.3641	0.0754	0.2205*	0.2382	0.3588	0.0870	0.3593	0.0534	0.3181	0.0584
0.0	0.4, 0.7	0.5	0.1, 0.8	0.3609	0.2867	0.3141	0.0352	0.3128	0.0669	0.2900*	0.0463	0.3055	0.0633	0.3147	0.0349	0.3141	0.0347*
0.5	0.4, 0.7	0.5	0.1, 0.8	0.4732	0.5848	0.3464	0.0404	0.3603	0.0545	0.2388*	0.1409	0.3603	0.0535	0.3523	0.0356*	0.3283	0.0372
Average				0.3479	0.1522	0.3196	0.0379	0.3235	0.0595	0.2756*	0.0880	0.3148	0.0544	0.3217	0.0377*	0.3128	0.0379
				0.2501		0.1787		0.1915		0.1818		0.1846		0.1797		0.1754*	

* Lowest RMSE for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A4. RMSE of indirect impacts with omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0.3, 0)^\intercal$.

				SLX		SAR		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.0903	0.0929	0.0288*	0.0290*	0.0690	0.0693	0.0907	0.0927	0.0907	0.0928
0.5	0, 0	0.0	0, 0	0.2049	0.1807	0.3058	0.0818*	0.3601	0.1662	0.3477	0.1797	0.1629*	0.2053
0.0	0.4, 0.7	0.0	0, 0	0.0745	0.0610	0.0265*	0.0266*	0.0486	0.0482	0.0747	0.0611	0.0748	0.0609
0.5	0.4, 0.7	0.0	0, 0	0.2153	0.1091	0.2993	0.0615*	0.3197	0.0945	0.3266	0.0988	0.1764*	0.1315
0.0	0, 0	0.5	0, 0	0.2492	0.1307	0.4181	0.4194	0.1948*	0.1961	0.3467	0.1791	0.2189	0.1164*
0.5	0, 0	0.5	0, 0	0.5124	0.2405*	1.5642	1.2857	0.7623	0.5244	1.5075	0.6378	0.2954*	0.2454
0.0	0.4, 0.7	0.5	0, 0	0.2529	0.0839	0.3516	0.3032	0.1300*	0.1194	0.3261	0.0983	0.2270	0.0773*
0.5	0.4, 0.7	0.5	0, 0	0.5703	0.1533*	1.2656	0.7828	0.5845	0.2854	1.3652	0.3336	0.3490*	0.1807
0.0	0, 0	0.0	0.1, 0.8	0.0903	0.0929	0.0562*	0.6666	0.4080	0.2711	0.0910	0.0954	0.0907	0.0928*
0.5	0, 0	0.0	0.1, 0.8	0.1685	0.5761	0.5023	1.1710	1.1624	0.4710	0.3491	0.1994*	0.1234*	0.7496
0.0	0.4, 0.7	0.0	0.1, 0.8	0.0745*	0.0610	0.1867	0.4442	0.4065	0.1267	0.0750	0.0618	0.0748	0.0609*
0.5	0.4, 0.7	0.0	0.1, 0.8	0.1833	0.3220	0.8884	0.4307	1.2033	0.2167	0.3272	0.1098*	0.1325*	0.4804
0.0	0, 0	0.5	0.1, 0.8	0.2492	0.1307	0.4806	0.2946	0.7523	0.1748	0.3824	0.4252	0.2189*	0.1164*
0.5	0, 0	0.5	0.1, 0.8	0.4647	0.5912	1.8671	0.3076*	1.7147	0.4123	1.5691	1.6597	0.2219*	0.8959
0.0	0.4, 0.7	0.5	0.1, 0.8	0.2529	0.0839	0.5336	0.1372	0.6841	0.1186	0.3490	0.2067	0.2270*	0.0773*
0.5	0.4, 0.7	0.5	0.1, 0.8	0.5318	0.3355*	1.8952	0.5415	1.8538	0.5077	1.3240	0.8206	0.2762*	0.6655
Average				0.2616	0.2028*	0.6669	0.4365	0.6659	0.2376	0.5533	0.3287	0.1850*	0.2656
				0.2322		0.5517		0.4518		0.4410		0.2253*	

* Lowest RMSE for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A5. Bias of direct impacts without omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$.

				OLS		SLX		SAR		SEM		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	−0.0006	0.0027	−0.0007	0.0027	−0.0004	0.0030	−0.0007	0.0028	−0.0000*	0.0045	−0.0007	0.0026*	−0.0007	0.0027
0.5	0, 0	0.0	0, 0	−0.0003	0.0029	−0.0002	0.0031	−0.0005	0.0031	−0.0172	−0.0378	−0.0001*	0.0045	−0.0001	0.0033	−0.0019	−0.0009*
0.0	0.4, 0.7	0.0	0, 0	−0.0002	0.0029	−0.0011	0.0021	−0.0001*	0.0031	−0.0002	0.0029	−0.0002	0.0028	−0.0011	0.0021*	−0.0011	0.0022
0.5	0.4, 0.7	0.0	0, 0	0.0209	0.1401	−0.0010	−0.0021*	0.0000*	0.0035	−0.0080	0.0301	0.0001	0.0042	−0.0006	0.0023	−0.0010	0.0035
0.0	0, 0	0.5	0, 0	0.0001*	0.0034	−0.0001	0.0034	0.0073	0.0223	−0.0011	0.0023*	0.0009	0.0075	−0.0002	0.0033	−0.0003	0.0031
0.5	0, 0	0.5	0, 0	0.0007	0.0041*	0.0007*	0.0044	0.0192	0.0524	−0.0232	−0.0530	0.0024	0.0112	0.0033	0.0106	−0.0044	−0.0072
0.0	0.4, 0.7	0.5	0, 0	0.0006	0.0040	−0.0006	0.0023*	−0.0060	−0.0570	−0.0008	0.0029	−0.0002*	0.0026	−0.0006	0.0025	−0.0006	0.0025
0.5	0.4, 0.7	0.5	0, 0	0.0222	0.1422	−0.0003*	−0.0018	0.0035	−0.0376	−0.0176	−0.0168	0.0003	0.0046	0.0011	0.0037	−0.0028	0.0004*
0.0	0, 0	0.0	0.1, 0.8	−0.0010	0.0020*	−0.0007	0.0027	−0.0025	−0.0182	−0.0044	−0.0229	−0.0032	0.0363	−0.0006*	0.0026	−0.0007	0.0027
0.5	0, 0	0.0	0.1, 0.8	−0.0007	0.0015*	−0.0002	0.0024	−0.0068	−0.0781	−0.0400	−0.1998	0.0039	−0.0071	−0.0001*	0.0034	−0.0036	−0.0144
0.0	0.4, 0.7	0.0	0.1, 0.8	0.0140	0.2816	−0.0011	0.0021	−0.0003*	0.1214	0.0032	0.1737	−0.0137	0.0976	−0.0011	0.0019*	−0.0011	0.0022
0.5	0.4, 0.7	0.0	0.1, 0.8	0.0420	0.5754	−0.0012	−0.0164	0.0004*	0.0794	−0.0366	−0.1011	−0.0005	0.1023	−0.0005	0.0023*	−0.0014	0.0067
0.0	0, 0	0.5	0.1, 0.8	−0.0003	0.0026*	−0.0001*	0.0034	0.0033	−0.0298	−0.0138	−0.0981	0.0058	−0.0032	0.0003	0.0070	−0.0003	0.0031
0.5	0, 0	0.5	0.1, 0.8	0.0003*	0.0027*	0.0007	0.0036	0.0092	−0.0617	−0.0464	−0.2344	0.0075	−0.0718	0.0049	0.0280	−0.0086	−0.0410
0.0	0.4, 0.7	0.5	0.1, 0.8	0.0148	0.2828	−0.0006	0.0023*	−0.0016	0.0562	−0.0087	0.0237	−0.0061	0.0478	−0.0006*	0.0041	−0.0006	0.0025
0.5	0.4, 0.7	0.5	0.1, 0.8	0.0432	0.5776	−0.0006*	−0.0160	0.0061	0.0405	−0.0394	−0.1326	0.0065	0.0402	0.0014	0.0047*	−0.0056	−0.0085
Average (absolute bias)				0.0101	0.1268	0.0006*	0.0044*	0.0042	0.0417	0.0163	0.0709	0.0032	0.0280	0.0011	0.0053	0.0022	0.0065
				0.0684		0.0025*		0.0230		0.0436		0.0156		0.0032		0.0043	

* Lowest bias for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A6. Bias of indirect impacts without omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$.

				SLX		SAR		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.0064	0.0060	0.0013*	0.0030*	0.0105	0.0264	0.0067	0.0058	0.0065	0.0058
0.5	0, 0	0.0	0, 0	−0.0457	−0.1252	0.0030*	0.0110*	0.0175	0.0474	0.0120	0.0138	−0.0637	−0.1689
0.0	0.4, 0.7	0.0	0, 0	0.0055	0.0022	0.0006*	0.0008*	0.0043	0.0078	0.0057	0.0023	0.0055	0.0021
0.5	0.4, 0.7	0.0	0, 0	−0.0358	−0.0696	0.0016*	0.0051*	0.0073	0.0159	0.0098	0.0059	−0.0514	−0.1060
0.0	0, 0	0.5	0, 0	0.0086	0.0108	0.1679	0.4225	0.0243	0.0613	0.0127	0.0130	0.0068*	0.0076*
0.5	0, 0	0.5	0, 0	−0.0421*	−0.1169*	0.5107	1.2887	0.0857	0.2210	0.1971	0.4513	−0.0786	−0.2038
0.0	0.4, 0.7	0.5	0, 0	0.0070	0.0053	0.1335	0.3030	0.0104	0.0233	0.0102	0.0059	0.0067*	0.0044*
0.5	0.4, 0.7	0.5	0, 0	−0.0331	−0.0641	0.3629	0.7805	0.0256*	0.0600*	0.1546	0.2556	−0.0672	−0.1484
0.0	0, 0	0.0	0.1, 0.8	0.0064*	0.0060	−0.0427	−0.6605	0.1543	−0.1062	0.0065	0.0068	0.0065	0.0058*
0.5	0, 0	0.0	0.1, 0.8	−0.0995	−0.5606	−0.0322	−1.1496	0.3229	−0.1349	0.0119*	0.0165*	−0.1349	−0.7393
0.0	0.4, 0.7	0.0	0.1, 0.8	0.0055*	0.0022	0.0223	−0.4209	0.1448	−0.0178	0.0056	0.0025	0.0055	0.0021*
0.5	0.4, 0.7	0.0	0.1, 0.8	−0.0777	−0.3093	0.1544	−0.3544	0.3395	0.2994	0.0096*	0.0057*	−0.1110	−0.4720
0.0	0, 0	0.5	0.1, 0.8	0.0086	0.0108	0.1286	−0.2720	0.2828	0.1264	0.0570	0.3836	0.0068*	0.0076*
0.5	0, 0	0.5	0.1, 0.8	−0.0958*	−0.5523	0.5008	0.0649*	0.4480	−0.0739	0.3238	1.6066	−0.1637	−0.8841
0.0	0.4, 0.7	0.5	0.1, 0.8	0.0070	0.0053	0.1526	−0.0948	0.2333	0.1365	0.0413	0.1804	0.0067*	0.0044*
0.5	0.4, 0.7	0.5	0.1, 0.8	−0.0750*	−0.3038*	0.5222	0.5486	0.5172	0.5282	0.2208	0.7967	−0.1425	−0.6537
Average (absolute bias)				0.0350*	0.1344	0.1711	0.3988	0.1643	0.1179*	0.0678	0.2345	0.0540	0.2135
				0.0847*		0.2849		0.1411		0.1512		0.1338	

* Lowest bias for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A7. Bias of direct impacts with omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0.3, 0)^\intercal$.

				OLS		SLX		SAR		SEM		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.2994	0.0027	0.2993*	0.0027	0.2997	0.0030	0.2993	0.0028	0.3004	0.0040	0.2993	0.0027*	0.2993	0.0027
0.5	0, 0	0.0	0, 0	0.3128	0.0029	0.3133	0.0032	0.3133	0.0032	0.2718*	−0.0384	0.3143	0.0043	0.3135	0.0034	0.3093	−0.0009*
0.0	0.4, 0.7	0.0	0, 0	0.2998	0.0029	0.2989	0.0021*	0.2999	0.0030	0.2998	0.0029	0.2998	0.0026	0.2989*	0.0022	0.2989	0.0022
0.5	0.4, 0.7	0.0	0, 0	0.3647	0.1401	0.3120	−0.0021*	0.3137	0.0034	0.2933*	0.0267	0.3139	0.0036	0.3130	0.0024	0.3120	0.0035
0.0	0, 0	0.5	0, 0	0.3132	0.0033	0.3134	0.0034	0.3195	0.0217	0.2889*	0.0022*	0.3036	0.0145	0.3135	0.0033	0.3109	0.0031
0.5	0, 0	0.5	0, 0	0.3469	0.0041*	0.3477	0.0045	0.3635	0.0514	0.2435*	−0.0534	0.3277	0.0255	0.3557	0.0105	0.3204	−0.0072
0.0	0.4, 0.7	0.5	0, 0	0.3443	0.0041	0.3124	0.0023*	0.3104	−0.0566	0.3050*	0.0028	0.3112	−0.0064	0.3130	0.0025	0.3124	0.0025
0.5	0.4, 0.7	0.5	0, 0	0.4470	0.1423	0.3447	−0.0017	0.3508	−0.0373	0.2661*	−0.0194	0.3325	0.0012	0.3512	0.0036	0.3299	0.0004*
0.0	0, 0	0.0	0.1, 0.8	0.2990	0.0020*	0.2993	0.0027	0.2999	−0.0177	0.2956	−0.0229	0.2865*	0.0354	0.2992	0.0027	0.2993	0.0027
0.5	0, 0	0.0	0.1, 0.8	0.3124	0.0015*	0.3133	0.0024	0.3164	−0.0782	0.2442*	−0.2015	0.3277	−0.0150	0.3135	0.0034	0.3075	−0.0145
0.0	0.4, 0.7	0.0	0.1, 0.8	0.3140	0.2816	0.2989	0.0021	0.2912	0.1274	0.3032	0.1737	0.2454*	0.1169	0.2990	0.0019*	0.2989	0.0022
0.5	0.4, 0.7	0.0	0.1, 0.8	0.3857	0.5754	0.3117	−0.0163	0.3177	0.0838	0.2505*	−0.1030	0.3000	0.1109	0.3131	0.0022*	0.3114	0.0067
0.0	0, 0	0.5	0.1, 0.8	0.3128	0.0026*	0.3134	0.0034	0.3191	−0.0299	0.2743*	−0.0991	0.3244	−0.0094	0.3139	0.0070	0.3109	0.0031
0.5	0, 0	0.5	0.1, 0.8	0.3465	0.0027*	0.3477	0.0037	0.3620	−0.0624	0.2184*	−0.2358	0.3566	−0.0739	0.3566	0.0274	0.3158	−0.0412
0.0	0.4, 0.7	0.5	0.1, 0.8	0.3585	0.2828	0.3124	0.0023*	0.3112	0.0591	0.2881*	0.0200	0.3038	0.0554	0.3130	0.0040	0.3124	0.0025
0.5	0.4, 0.7	0.5	0.1, 0.8	0.4680	0.5776	0.3445	−0.0159	0.3584	0.0421	0.2365*	−0.1355	0.3584	0.0404	0.3504	0.0047*	0.3264	−0.0089
Average (absolute bias)				0.3453	0.1268	0.3177	0.0044*	0.3217	0.0425	0.2736*	0.0713	0.3129	0.0325	0.3198	0.0052	0.3110	0.0065
				0.2361		0.1611		0.1821		0.1724		0.1727		0.1625		0.1588*	

* Lowest bias for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table A8. Bias of indirect impacts with omv. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0.3, 0)^\intercal$.

				SLX		SAR		SAC		SDM		SDEM	
ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.0	0, 0	0.0	0, 0	0.0064	0.0060	0.0032*	0.0031*	0.0176	0.0177	0.0067	0.0059	0.0065	0.0058
0.5	0, 0	0.0	0, 0	0.1588	−0.1251	0.2947	0.0113*	0.3220	0.0388	0.2989	0.0143	0.1165*	−0.1689
0.0	0.4, 0.7	0.0	0, 0	0.0055	0.0022	0.0018*	0.0011*	0.0083	0.0064	0.0057	0.0021	0.0055	0.0021
0.5	0.4, 0.7	0.0	0, 0	0.1858	−0.0696	0.2910	0.0058	0.3019	0.0143	0.2960	0.0058*	0.1471*	−0.1060
0.0	0, 0	0.5	0, 0	0.2131	0.0109	0.4112	0.4127	0.1437*	0.1455	0.2985	0.0137	0.1871	0.0077*
0.5	0, 0	0.5	0, 0	0.4664	−0.1164*	1.5406	1.2576	0.6997	0.4293	1.4320	0.4455	0.2638*	−0.2040
0.0	0.4, 0.7	0.5	0, 0	0.2286	0.0053	0.3459	0.2990	0.1003*	0.0930	0.2961	0.0061	0.2054	0.0045*
0.5	0.4, 0.7	0.5	0, 0	0.5404	−0.0639*	1.2476	0.7650	0.5493	0.2298	1.3194	0.2504	0.3272*	−0.1488
0.0	0, 0	0.0	0.1, 0.8	0.0064*	0.0060	0.0396	−0.6655	0.3985	−0.2498	0.0066	0.0068	0.0065	0.0058*
0.5	0, 0	0.0	0.1, 0.8	0.1050	−0.5605	0.4865	−1.1647	1.1389	−0.3839	0.2992	0.0162*	0.0445*	−0.7396
0.0	0.4, 0.7	0.0	0.1, 0.8	0.0055*	0.0022	0.1817	−0.4417	0.4020	−0.1071	0.0056	0.0024	0.0055	0.0021*
0.5	0.4, 0.7	0.0	0.1, 0.8	0.1439	−0.3093	0.8781	−0.4118	1.1915	0.1341	0.2960	0.0057*	0.0855*	−0.4725
0.0	0, 0	0.5	0.1, 0.8	0.2131	0.0109	0.4718	−0.2825	0.7329	−0.0184	0.3400	0.3834	0.1871*	0.0077*
0.5	0, 0	0.5	0.1, 0.8	0.4127	−0.5518	1.8377	0.0338*	1.6697	−0.1388	1.4996	1.5784	0.1775*	−0.8851
0.0	0.4, 0.7	0.5	0.1, 0.8	0.2286	0.0053	0.5274	−0.1162	0.6751	0.0508	0.3217	0.1799	0.2054*	0.0045*
0.5	0.4, 0.7	0.5	0.1, 0.8	0.4985	−0.3036*	1.8747	0.4823	1.8309	0.4265	1.2818	0.7785	0.2474*	−0.6565
Average (absolute bias)				0.2137	0.1343*	0.6521	0.3971	0.6364	0.1553	0.5002	0.2309	0.1387*	0.2138
				0.1740*		0.5246		0.3958		0.3656		0.1763	

* Lowest bias for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Appendix B. Different combinations of autocorrelation

Table B1. Bias of direct impacts for different strengths of autocorrelation. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$, $\boldsymbol{\delta} = (0, 0)^\intercal$, $\boldsymbol{\theta} = (0.1, 0.8)^\intercal$.

		OLS		SLX		SAR		SEM		SAC		SDM		SDEM	
ρ	λ	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$				
0.3	0.3	-0.0004	0.0022*	−0.0001*	0.0031	−0.0004	−0.0518	−0.0271	−0.1503	0.0050	−0.0142	0.0004	0.0068	−0.0017	−0.0055
0.5	0.3	-0.0002	0.0021*	0.0003	0.0030	−0.0000*	−0.0741	−0.0441	−0.2219	0.0055	−0.0464	0.0017	0.0119	−0.0065	−0.0300
0.7	0.3	-0.0000*	0.0014*	0.0007	0.0024	0.0009	−0.1104	−0.0711	−0.3317	0.0012	−0.1109	0.0048	0.0245	−0.0217	−0.0996
0.9	0.3	-0.0006*	-0.0037	0.0009	−0.0027*	−0.0080	−0.2325	−0.1360	−0.5993	−0.0270	−0.3120	0.0116	0.0494	−0.0761	−0.3347
0.3	0.5	0.0000*	0.0027*	0.0003	0.0036	0.0061	−0.0465	−0.0301	−0.1686	0.0088	−0.0334	0.0019	0.0152	−0.0026	−0.0111
0.5	0.5	0.0003*	0.0027*	0.0007	0.0036	0.0092	−0.0617	−0.0464	−0.2344	0.0075	−0.0718	0.0049	0.0280	−0.0086	−0.0410
0.7	0.5	0.0007	0.0024*	0.0013	0.0034	0.0142	−0.0871	−0.0725	−0.3389	−0.0005*	−0.1470	0.0119	0.0578	−0.0245	−0.1127
0.9	0.5	0.0005*	-0.0016	0.0018	−0.0006*	0.0039	−0.2090	−0.1364	−0.6015	−0.0356	−0.3655	0.0235	0.0998	−0.0773	−0.3407
0.3	0.7	0.0007*	0.0037*	0.0009	0.0045	0.0203	−0.0256	−0.0327	−0.1844	0.0166	−0.0452	0.0061	0.0381	−0.0037	−0.0174
0.5	0.7	0.0011*	0.0040*	0.0015	0.0049	0.0279	−0.0290	−0.0484	−0.2453	0.0129	−0.0899	0.0130	0.0680	−0.0106	−0.0510
0.7	0.7	0.0018	0.0043*	0.0023	0.0054	0.0363	−0.0438	−0.0737	−0.3452	0.0005*	−0.1747	0.0263	0.1230	−0.0265	−0.1226
0.9	0.7	0.0022*	0.0027*	0.0033	0.0037	0.0157	−0.1861	−0.1367	−0.6032	−0.0422	−0.4083	0.0360	0.1528	−0.0778	−0.3445
0.3	0.9	0.0021*	0.0069*	0.0021	0.0077	0.0599	0.0492	−0.0348	−0.1975	0.0430	−0.0172	0.0201	0.1160	−0.0047	−0.0232
0.5	0.9	0.0029*	0.0083*	0.0030	0.0092	0.0681	0.0507	−0.0499	−0.2540	0.0346	−0.0723	0.0328	0.1660	−0.0119	−0.0583
0.7	0.9	0.0042*	0.0110*	0.0044	0.0120	0.0654	0.0152	−0.0744	−0.3495	0.0148	−0.1717	0.0465	0.2142	−0.0274	−0.1284
0.9	0.9	0.0059*	0.0183*	0.0065	0.0191	0.0224	−0.1736	−0.1366	−0.6039	−0.0409	−0.4287	0.0449	0.1814	−0.0775	−0.3454
Average		0.0015*	0.0049*	0.0019	0.0056	0.0224	0.0904	0.0719	0.3393	0.0185	0.1568	0.0179	0.0845	0.0287	0.1291
(absolute bias)		0.0032*		0.0037		0.0564		0.2056		0.0877		0.0512		0.0789	

* Lowest bias for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Table B2. Bias of indirect impacts for different strengths of autocorrelation. $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$, $\boldsymbol{\delta} = (0, 0)^\intercal$, $\boldsymbol{\theta} = (0.1, 0.8)^\intercal$.

		SLX		SAR		SAC		SDM		SDEM	
ρ	λ	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$
0.3	0.3	-0.0260*	-0.1971	0.0890	−0.5762	0.2900	−0.0351*	0.0686	0.3528	−0.0420	−0.2831
0.5	0.3	-0.0978*	-0.5567	0.1863	−0.6482	0.3776	−0.1366*	0.1374	0.6562	−0.1525	−0.8284
0.7	0.3	-0.3245*	-1.6044	0.4664	−0.6929	0.4893	−0.6352*	0.3418	1.5363	−0.4851	−2.3461
0.9	0.3	-1.8563	-8.3520	0.7387	−3.6504*	−0.1147*	−6.0352	0.9216	3.9014	−2.3812	−10.6470
0.3	0.5	-0.0245*	-0.1936	0.2670	−0.1768	0.3690	0.0839*	0.1520	0.8308	−0.0477	−0.3147
0.5	0.5	-0.0958*	-0.5523	0.5008	0.0649*	0.4480	−0.0739	0.3238	1.6066	−0.1637	−0.8841
0.7	0.5	-0.3219*	-1.5977	1.0849	0.7313*	0.4992	−0.7792	0.8223	3.7944	−0.4983	−2.4074
0.9	0.5	-1.8524	-8.3378	1.4847	−1.9049*	−0.3411*	−6.8599	1.9225	8.3204	−2.3870	−10.6740
0.3	0.7	-0.0219*	-0.1870*	0.7086	0.8009	0.5559	0.4223	0.3612	2.0194	−0.0536	−0.3465
0.5	0.7	-0.0925*	-0.5435	1.2625	1.7761	0.6306	0.2161*	0.7878	3.9738	−0.1731	−0.9316
0.7	0.7	-0.3174*	-1.5845	2.2593	3.4158	0.6228	−0.6721*	1.7894	8.3603	−0.5078	−2.4529
0.9	0.7	-1.8458	-8.3083	2.2187	−0.2088*	−0.4630*	−7.3746	2.9508	12.8685	−2.3900	−10.6909
0.3	0.9	-0.0166*	-0.1651*	2.4503	4.6096	1.4205	2.1489	1.1525	6.6166	−0.0586	−0.3737
0.5	0.9	-0.0860*	-0.5139*	3.3502	6.4141	1.4331	1.7591	2.0378	10.4425	−0.1795	−0.9651
0.7	0.9	-0.3087*	-1.5387	3.9543	7.2462	1.2674	0.4890*	3.1944	15.0044	−0.5123	−2.4786
0.9	0.9	-1.8342	-8.2010	2.6037	0.6582*	−0.2434*	−7.1352	3.5207	15.1102	−2.3891	−10.6950
Average		0.5701*	2.6521	1.4766	2.0984*	0.5979	2.1785	1.2803	5.9622	0.7763	3.5824
(absolute bias)		1.6111		1.7875		1.3882*		3.6212		2.1794	

* Lowest bias for $\boldsymbol{x}_k$ within the parameter combination. Number of observations=900, repetitions=1000.

Appendix C. Lagrange multiplier tests

Table C1. Rejection rates of H_0 (Lagrange multiplier test). $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$.

ρ	$\boldsymbol{\delta}$	λ	$\boldsymbol{\theta}$	LM_λ	LM_ρ	LM_λ^*	LM_ρ^*	$LM_{\lambda\rho}$
0.0	0, 0	0.0	0, 0	0.0410	0.0440	0.0520	0.0500	0.0580
0.4	0, 0	0.0	0, 0	1.0000	1.0000	0.0920	0.7750	1.0000
0.8	0, 0	0.0	0, 0	1.0000	1.0000	0.3880	0.9990	1.0000
0.0	0.4, 0.7	0.0	0, 0	0.0440	0.0520	0.0550	0.0590	0.0550
0.4	0.4, 0.7	0.0	0, 0	1.0000	1.0000	0.1340	0.9900	1.0000
0.8	0.4, 0.7	0.0	0, 0	1.0000	1.0000	0.9950	1.0000	1.0000
0.0	0, 0	0.4	0, 0	1.0000	1.0000	0.7190	0.1140	1.0000
0.4	0, 0	0.4	0, 0	1.0000	1.0000	0.8560	0.6830	1.0000
0.8	0, 0	0.4	0, 0	1.0000	1.0000	0.8380	0.9780	1.0000
0.0	0.4, 0.7	0.4	0, 0	1.0000	1.0000	0.9430	0.1100	1.0000
0.4	0.4, 0.7	0.4	0, 0	1.0000	1.0000	0.9970	0.9560	1.0000
0.8	0.4, 0.7	0.4	0, 0	1.0000	1.0000	1.0000	1.0000	1.0000
0.0	0, 0	0.8	0, 0	1.0000	1.0000	0.9990	0.2860	1.0000
0.4	0, 0	0.8	0, 0	1.0000	1.0000	0.9960	0.5190	1.0000
0.8	0, 0	0.8	0, 0	1.0000	1.0000	0.7450	0.7410	1.0000
0.0	0.4, 0.7	0.8	0, 0	1.0000	1.0000	1.0000	0.3000	1.0000
0.4	0.4, 0.7	0.8	0, 0	1.0000	1.0000	1.0000	0.7280	1.0000
0.8	0.4, 0.7	0.8	0, 0	1.0000	1.0000	1.0000	0.9410	1.0000
0.0	0, 0	0.0	0.1, 0.8	0.4240	0.9910	1.0000	1.0000	1.0000
0.4	0, 0	0.0	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0, 0	0.0	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000
0.0	0.4, 0.7	0.0	0.1, 0.8	0.8320	1.0000	1.0000	1.0000	1.0000
0.4	0.4, 0.7	0.0	0.1, 0.8	1.0000	1.0000	0.9960	1.0000	1.0000
0.8	0.4, 0.7	0.0	0.1, 0.8	1.0000	1.0000	0.4970	1.0000	1.0000
0.0	0, 0	0.4	0.1, 0.8	1.0000	1.0000	0.9900	1.0000	1.0000
0.4	0, 0	0.4	0.1, 0.8	1.0000	1.0000	0.9530	1.0000	1.0000
0.8	0, 0	0.4	0.1, 0.8	1.0000	1.0000	0.8520	1.0000	1.0000
0.0	0.4, 0.7	0.4	0.1, 0.8	1.0000	1.0000	0.9560	1.0000	1.0000
0.4	0.4, 0.7	0.4	0.1, 0.8	1.0000	1.0000	0.0930	1.0000	1.0000
0.8	0.4, 0.7	0.4	0.1, 0.8	1.0000	1.0000	0.9600	1.0000	1.0000
0.0	0, 0	0.8	0.1, 0.8	1.0000	1.0000	0.2590	0.9930	1.0000
0.4	0, 0	0.8	0.1, 0.8	1.0000	1.0000	0.2370	0.9940	1.0000
0.8	0, 0	0.8	0.1, 0.8	1.0000	1.0000	0.1530	0.9940	1.0000
0.0	0.4, 0.7	0.8	0.1, 0.8	1.0000	1.0000	0.9880	1.0000	1.0000
0.4	0.4, 0.7	0.8	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0.4, 0.7	0.8	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000

Number of observations=900, repetitions=1000. LM= Lagrange multiplier test, LM^*= Robust Lagrange multiplier test, each for H_0: $\lambda = 0$, H_0: $\rho = 0$, H_0: $\lambda, \rho = 0$.

Appendix D. Different specifications of W

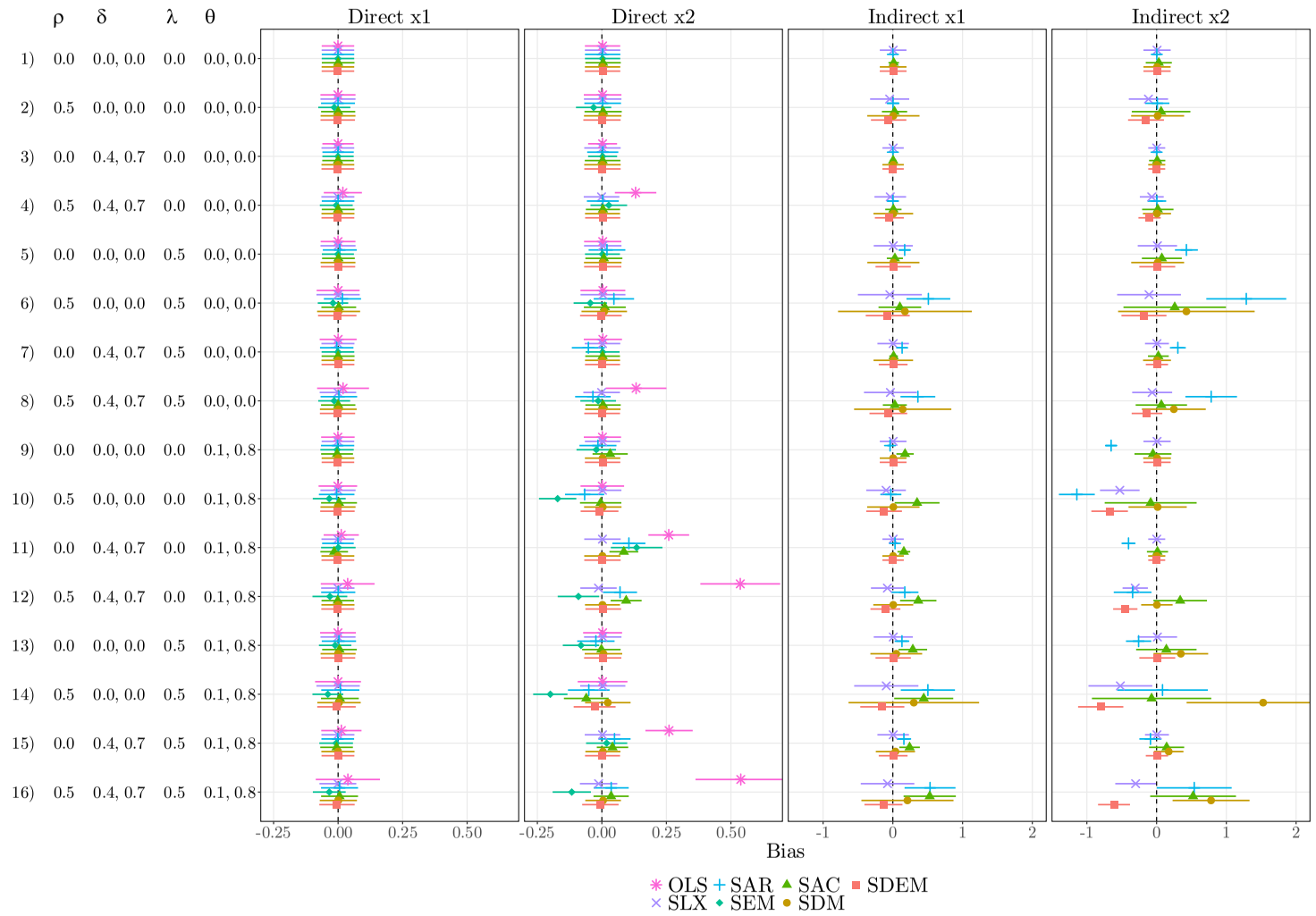

Figure D1. $\boldsymbol{W}$ = 10-nearest neighbours, inverse distance weighted, row-normalized. Bias of impacts and 95% confidence interval of empirical standard deviation without omv: $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$. ρ = autocorrelation in the dependent variable ($\boldsymbol{Wy}$); $\boldsymbol{\delta}$ = autocorrelation in the covariates ($\boldsymbol{x}_k = f(\boldsymbol{Wx}_k)$); λ = autocorrelation in the disturbances ($\boldsymbol{Wu}$); $\boldsymbol{\theta}$ = spatial spillover effects of covariates ($\boldsymbol{WX}$); $\boldsymbol{\gamma}$ = strength of omv.

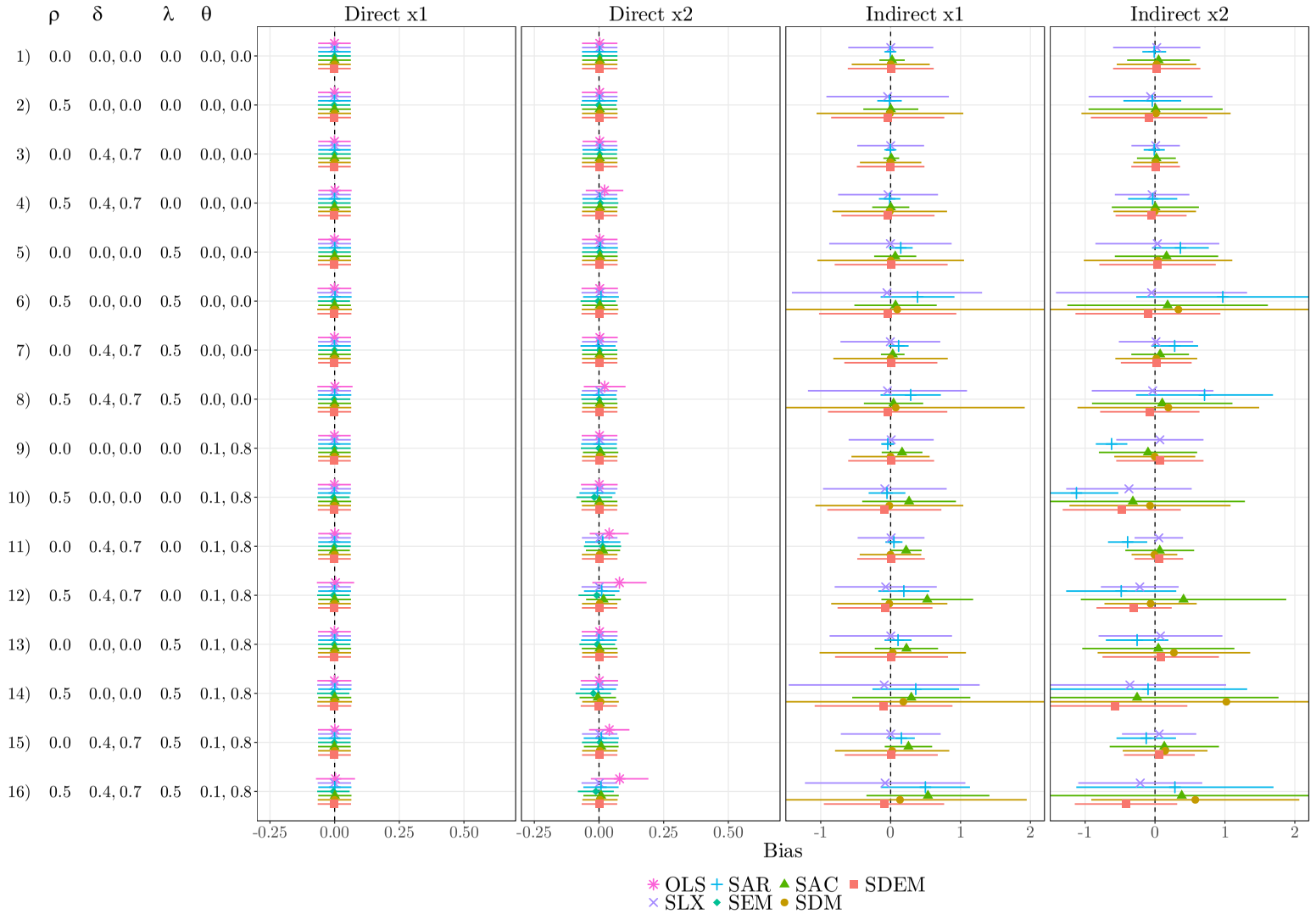

Figure D2. $\boldsymbol{W}$ = inverse distance weighted neighbours (cut-off: 100 neighbours), maximum eigenvalue-normalized. Bias of impacts and 95% confidence interval of empirical standard deviation without omv: $\boldsymbol{\beta} = (0.2, 0.5)^\intercal$, $\boldsymbol{\gamma} = (0, 0)^\intercal$. ρ = autocorrelation in the dependent variable ($\boldsymbol{Wy}$); $\boldsymbol{\delta}$ = autocorrelation in the covariates ($\boldsymbol{x}_k = f(\boldsymbol{Wx}_k)$); λ = autocorrelation in the disturbances ($\boldsymbol{Wu}$); $\boldsymbol{\theta}$ = spatial spillover effects of covariates ($\boldsymbol{WX}$); $\boldsymbol{\gamma}$ = strength of omv.