ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Fischer, Lorenz Benedikt

Article

The bias of avoiding spatial dynamic panel models: A tale of two research teams

REGION

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Fischer, Lorenz Benedikt (2021) : The bias of avoiding spatial dynamic panel models: A tale of two research teams, REGION, ISSN 2409-5370, European Regional Science Association (ERSA), Louvain-la-Neuve, Vol. 8, Iss. 1, pp. 153-180, https://doi.org/10.18335/region.v8i1.316

This Version is available at:
https://hdl.handle.net/10419/235830

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by-nc/4.0

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

REGION

The Journal of ERSA
Powered by WU

Volume 8, Number 1, 2021, 153–180
DOI: 10.18335/region.v8i1.316

journal homepage: region.ersa.org

The bias of avoiding spatial dynamic panel models A tale of two research teams*

Lorenz Benedikt Fischer[1]

[1] Johannes Kepler University Linz, Austria; University of Innsbruck, Innsbruck, Austria

Received: 8 April 2020/Accepted: 19 January 2021

Abstract. Many questions in urban and regional economics can be characterized as including both a spatial and a time dimension. However, often one of these dimensions is neglected in empirical work. This paper highlights the danger of methodological inertia, investigating the effect of neglecting the spatial or the time dimension when in fact both are important. A tale of two research teams, one living in a purely dynamic and the other in a purely spatial world of thinking, sets the scene. Because the research teams' choices to omit a dimension change the assumed optimal estimation strategies, the issue is more difficult to analyze than a typical omitted variables problem. First, the bias of omitting a relevant dimension is approximated analytically. Second, Monte Carlo simulations show that the neglected dimension projects onto the other, with potentially disastrous results. Interestingly, dynamic models are bound to overestimate autoregressive behavior whenever the spatial dimension is important. The same holds true for the opposite case. An application using the well-known, openly available cigarette demand data supports these findings.

Keywords: Spatial dynamic panel data Monte Carlo simulation Spatial interaction Dynamic model Omitted variable bias

JEL: C13 C23 R10

1 Introduction

In regional economics, it has long been understood that units of observation may not be considered independent from one another across space, often associated with Tobler's first law of geography. The beauty of one house may increase the perceived beauty of neighboring houses. A region may be forced to decrease local tax rates because its close-by regions chose to decrease theirs in order to retain its tax base. One region's negative employment shock may not only increase local unemployment but also unemployment in neighboring regions due to the mobile workers. From early influential contributions like Cliff, Ord (1972), a formidable spatial econometric literature has grown. Kelejian, Piras (2017) offer an extensive overview.

Economists are also aware about the dynamics of economic processes. The notion that the state of a variable in some period depends on the state in the previous period plays a major role in many fields. In general, any partial adjustment process may be seen as a

*Funding by the Austrian Science Fund (FWF) under project number P30729 is greatly appreciated. The author further wants to thank participants of the 2019 Winter Seminar of the Gesellschaft für Regionalforschung (GfR), as well as participants and reviewers of the 59th ERSA conference.

motivation. A classic example is the convergence literature, where the contribution of Islam (1995) is among the early ones to exploit panel data. Most readers have likely dealt with issues in estimating dynamic models themselves. This notion has too led to a massive body of econometric tools that enable researchers to use panel data to investigate dynamic processes. For example, sophisticated methods that deal with incidental parameter problems that come with short-in-time panel data have evolved and found widespread use. Hsiao (2014) offers a comprehensive overview of econometric methods for (dynamic) panel data. In the current discourse, generalized method of moments (GMM) approaches in the line of Arellano, Bond (1991) and Blundell, Bond (1998) are widely used, but have received considerable criticism about their suitability in empirical applications. While these GMM methods tend to suffer from instrument proliferation and are likely to break down when facing non-random initial observations or in proximity of unit roots, they remain popular due to their availability and the possibility to exploit internal and external instruments (Roodman 2009).

With the increasing availability of geo-referenced panel data, methods that take into account dependence over time and space have emerged. These methods are relatively new to the market, and have hardly been used in regional economics, at least in relative terms. Even if applied, dynamic models tend to be the main estimation methodology, while spatial dynamic models appear to be used mainly to assess robustness. It seems that spatial dynamic applications are more data-driven than formally theoretically justified. However, there are some exceptions worth mentioning that expanded more 'traditional' approaches to include spatial and dynamic properties. For example, Elhorst et al. (2013) investigate spatial diffusion of financial liberalization among countries using a spatial dynamic approach similar to the one discussed below. Expanding the famous Blanchard-Katz labor market model, Vega, Elhorst (2014) find highly significant spillover effects applying a spatial dynamic Durbin model. Similarly, Rios (2017) report significant indirect effects of employment growth with respect to unemployment rates among European NUTS 2 regions, which appear sizable especially in the long-run. One of the traditional motivations of (cross-sectional) spatial econometric approaches are house prices. Including a spatial time lag in a panel smooth transition regression model, Pijnenburg (2017) reports significant heterogeneity in spatial dependence. As a last example, Wanzenböck, Piribauer (2018) investigate R&D networks across NUTS 2 regions, highlighting strong spatial and dynamic effects in the course of a spatial Durbin model.

The aim of this study is to raise awareness about the problems research can encounter by neglecting dependence in either space or time from an econometric point of view. In order to meet this goal, the remainder of this paper builds on a tale of two research teams, where one team is located in a purely dynamic and the other in a purely spatial world of thinking. Both teams are handed the same data and each tries to explain the process as they see fit.

A few words of caution are in order. Even though the points made here are closely related to the typical omitted variables argument, the specific issues to be dealt with are a little different. An omitted variable usually does not change the optimal estimation strategy. Here, the preferred estimation methods of both teams will abstract from the optimal one. In fact, both teams aim for a suitable estimator *given* their beliefs about the underlying data generating process (DGP). Hence, 'Team Dynamic' fits a dynamic model, while 'Team Spatial' fits a (static) spatial panel model. Still, the mistakes made by both teams may be categorized as an omitted variable problem. However, they are more severe because they change the estimation method and, hence, the conceptualization of the DGP. In fact, both teams will make systematic errors in explaining the dynamics of the process. The question this paper tries to answer is whether the remaining dimension may still be estimated in an unbiased manner. In particular, is Team Dynamic able to estimate time dependence accurately? Is Team Spatial able to infer on spatial dependence without bias? The answer to both questions is *no*. In fact, it is possible to determine the direction of each bias.

As a first step, approximations of the mistakes made by both teams are derived analytically. The results show that the autoregressive parameter is overestimated in absolute terms whenever there is spatial dependence. Consequently, marginal effects of

covariates may also be severely biased, as discussed in Section 5. The other way around, results are less clear, but point in the same direction. Note that in the course of this paper, the term autoregressive always refers to the time dimension.

These approximations are put to the test using a Monte Carlo simulation. The results confirm the expectations and show potentially drastic outcomes. Lastly, a brief application using the openly available 'Baltagi cigarette demand' data (Baltagi, Li 2004) supports the findings.

2 A spatial dynamic process

As laid out in the introduction, the task is to infer the bias of omitting the time or spatial dimension. Naturally, there are multiple ways to include both types of dependence in a DGP that serves as a starting point. Leaning on Yu et al. (2008) and the examples stated before, assume a DGP that incorporates dependence over time as well as space in a seemingly separated manner. Specifically, define

$$Y_{nt} = \gamma_0 Y_{n,t-1} + \lambda_0 W_n Y_{nt} + X_{nt}\beta_0 + c_{n0} + d_{t0} l_n + V_{nt}, \tag{1}$$

where Y_{nt} is the $n \times 1$ vector of the outcome variable at time t that depends on its time lag $Y_{n,t-1}$, its contemporaneous spatial lag $W_n Y_{nt}$, an exogenous time-varying variable X_{nt}, and region specific fixed effects $\mathbf{c}_{n0}$ (all $n \times 1$). Further, the model includes a time-fixed effect given by a scalar term d_{t0} multiplied with the size-n vector of ones, l_n, and a $n \times 1$ vector of *i.i.d.* error terms V_{nt} with mean zero and variance σ_0^2. Scalars γ_0, λ_0, and β_0 are the coefficients of interest. The restriction to one exogenous variable X_{nt} is for notational convenience only. The $n \times n$ spatial weights matrix W_n is assumed non-stochastic and with diagonal values of zero. The common notion is that the more influence unit i has on another unit j, the larger the weight w_{ij}. As usual, let W_n be constant over time. For the following approximations it is useful to assume maximum-row-normalization, which preserves the symmetry of W_n. Note that whether one assumes local or global spillovers depends on the assumed DGP and how it translates into marginal effects, not the spatial weights matrix. Equation (1) implies the presence of global spillovers.

It has to be noted that the bulk – all except one – of the studies mentioned in the introduction as already applying spatial dynamic approaches use a spatial dynamic Durbin model (SDD), which would include a spatial time lag as well as a spatial lag of the covariate. This is motivated by LeSage, Pace (2009) for empirical applications due to preferable behavior in the presence of omitted variables or spatially correlated residuals, and Elhorst (2012) suggests a general-to-specific model selection approach where the Durbin model represents the most general approach.[1] For the ease of the argument, however, it is preferable to stick to seemingly 'separated' dimensions while noting that the DGP described in (1) is nested in a SDD model.

In the following, cases with negative time dependence are disregarded due to their near nonexistence in (regional) economics. Small samples in terms of the time dimension may require bias reduction procedures Team Dynamic and Team Spatial. In order to focus on the research question at hand, let us assume that the data provided to the teams has time dimension T large enough, such that this issue can too be disregarded.

Because explicit solutions of the biases are not available, either because of the presence of global spillovers or the estimation procedure, approximations are presented. The conclusion is that the direction of the bias can be determined, and that it can be approximated fairly simply.

Usually, exogenous variables are not of much interest in specifying DGPs, since the property of being exogenous suffices for proper estimation. Key to the derivations that follow in the next section is strict exogeneity of the covariate, implying that innovations in the DGP of the covariate are uncorrelated with contemporaneous and past innovations in the DGP of the dependent variable. Albeit not discussed in the literature, there is no reason to assume spatial behavior in the dependent variable but to deny it to the

[1] It is worth mentioning that the evolution of spatial dynamic panel models is still progressing. For example, Shi, Lee (2017) extend spatial dynamic panel models with interactive fixed effects, that is, different factor loadings of unobserved time-effects that may affect groups of units heterogeneously.

covariate. In fact, strict exogeneity is bound to break down if the covariate itself follows a dynamic DGP. The interested reader is relegated to Appendix D, where the case of a full spatial dynamic DGP of the covariate is investigated, after reading the following sections.

Lee, Yu (2010b) provide limited evidence on the effect of omitted time- and spatial-terms based on Monte Carlo simulations for spatial dynamic panel data models. Their focus, however, is on comparing the performance of ML estimation with respect to different types of panel transformations. Omitting the spatial lag leads to a potentially large bias in the estimated coefficient of the time lag and vice versa, while estimated coefficients of exogenous variables appear unaffected. In the following, the argument is formalized and the impact on marginal effects of the covariate is analyzed.

3 The tale

Before telling the tale of the two research teams – Team Dynamic and Team Spatial – it has to be stressed again that the analytical derivations of the estimates presented in this section are approximations, based on the first two elements of the Neumann series representation of $A_n = (I_n - \lambda_0 W_n)^{-1} = \sum_{k=0}^{\infty} (\lambda_0 W_n)^k$ with the usual constraint $|\lambda_0| < 1$, where I_n denotes an identity matrix of size n. This approximation is especially precise for low to moderate values of λ_0, as the share of A_n explained by the first two terms $I + \lambda_0 W_n$ equals $1 - \lambda_0^2$. As mentioned in the beginning, both research teams are handed the same data with the goal to explain the DGP as best as possible, thereby abstracting from usual omitted variable problems. While equally investigating the difference between estimated and true parameters and marginal effects by inserting the DGP in the estimate, the *estimator* is the not the one that would fit the DGP optimally.

Team Dynamic misreads the data, or is misguided by theory, such that the assumed DGP is purely dynamic. This implies that λ_0 is wrongly set to zero. A typical problem with dynamic panels is that the within-estimator is biased due to the correlation of the error terms and lagged dependent variable. This bias is of the order $O(1/T)$ (Nickell 1981). Therefore, the strategy is to take first differences, indicated by a bar. In the sense of Anderson, Hsiao (1982) and Arellano, Bover (1995), Team Dynamic may apply a two-stage least squares approach in which the endogenous lag of the first difference of the dependent variable ($\bar{Y}_-$) is instrumented by earlier lags to avoid the bias caused by taking first differences. Regardless of which of these moment estimators is applied, the estimation will be of the form

$$\boldsymbol{\phi}_{TD} = \left(\bar{\Gamma}' P_Z \bar{\Gamma}\right)^{-1} \bar{\Gamma}' P_Z \bar{Y}, \tag{2}$$

where the estimated parameters of Team Dynamic (TD) are collected in $\boldsymbol{\phi}_{TD} = (\gamma_{TD}, \beta_{TD})'$, $\bar{\Gamma} = \left[\bar{Y}_-, \bar{X}\right]$, and P_Z is the projection matrix of the instruments chosen by Team Dynamic. As already mentioned, let us assume that the time dimension is large. Hence, the Nickell problem may be disregarded such that using $P_Z = I_{nT}$ yields a sufficiently consistent estimator given the belief of Team Dynamic (Hsiao 2014). Using $W = I_T \otimes W_n$ and inserting the first-differenced reduced form of (1), the parameter bias of neglecting the spatial dimension is approximated by (see Appendix A)

$$\begin{aligned} \operatorname{plim}\left[\boldsymbol{\phi}_{TD} - \boldsymbol{\phi}_0\right] &\approx \lambda_0 \left(\tilde{\Gamma}'\tilde{\Gamma}\right)^{-1} \tilde{\Gamma}' W \tilde{\Gamma} \boldsymbol{\phi}_0 \\ &\approx \lambda_0 \left(\tilde{\Gamma}'\tilde{\Gamma}\right)^{-1} \left(\tilde{\Gamma}'\tilde{\Gamma} \odot \mathbf{M}\right) \boldsymbol{\phi}_0, \end{aligned} \tag{3}$$

where $\tilde{\Gamma} = P_Z \bar{\Gamma}$ and $\phi_0 = (\gamma_0, \beta_0)'$ captures the true parameter values.[2] The symmetric matrix $\mathbf{M}$ captures spatial correlation in the fashion of Moran's I. The off-diagonal values of $\mathbf{M}$ tend to zero as long as the covariate X remains exogenous, justifying the search for (quasi-) random assignments of covariates in the literature. In principle, one can also assume a non-spatial DGP for $\tilde{X} = P_Z \bar{X}$. The important characteristic, as in any dynamic model, is the strict exogeneity. Appendix A offers detailed derivations.

[2] Note that, in general, $\tilde{\Gamma}' W \bar{\Gamma} \neq \tilde{\Gamma}' W \tilde{\Gamma}$ even when P_Z and W are symmetric.

Because $\left(\tilde{\Gamma}'\tilde{\Gamma}\right)^{-1}$ is a concentration matrix, the elements η_{ij} are negative partial covariances, and diagonal elements η_i correspond to inverse residual variances of regressing all but one components of $\tilde{\Gamma}$ on the component of the corresponding row/column (Cox, Wermuth 1996). Under strict exogeneity, as spelled out, for example, in equation (D.4), these off-diagonal values will be zero unless the innovations in the DGP of $\tilde{X}$ are autocorrelated. The more complex case in which the covariate is not strictly exogenous is derived in Appendix A, equation (A.4), showing that potential 'cross-effects' might blur the main bias term as presented here. This case is further discussed in Section 5 and Appendix D.

Given these arguments, the approximated bias in both coefficients is easily derived as

$$\text{plim}\left(\boldsymbol{\phi}_{TD} - \boldsymbol{\phi}_0\right) \approx \lambda_0 \begin{bmatrix} \gamma_0 \eta_{\tilde{Y}_-} \sigma^2_{\tilde{Y}_-} m_{\tilde{Y}_-} \\ \beta_0 \eta_{\tilde{X}} \sigma^2_{\tilde{X}} m_{\tilde{X}} \end{bmatrix}, \tag{4}$$

where σ^2 indicates a variance, $m_{\tilde{Y}_-} = \tilde{Y}_-' W \tilde{Y}_- / \tilde{Y}_-' \tilde{Y}_-$ is Moran's I of $\tilde{Y}_-$, $m_{\tilde{X}}$ similarly represents Moran's I of the covariate, and in accordance to above notation, $\tilde{Y}_- = P_Z \bar{Y}_-$. Hence, the sign of the bias of each coefficient, additionally to its own sign, depends on λ_0 and the spatial dependence in its corresponding DGP. As shown in Appendix B, the sign of $m_{\tilde{Y}_-}$ in (3) is determined solely by λ_0 under the common assumption $|\lambda_0|, |\gamma_0| \leq 1$. This consequence is that Team Dynamic will always overestimate the true value of γ_0 in absolute terms and give too much weight to the time dimension whenever the spatial dimension is important. In cases where time dependence is already large, Team Dynamic will become more likely to report non-stationarity when in fact the DGP may be stationary.

Note that the measure of spatial dependence in the covariate, $m_{\tilde{X}}$, is based on the spatial structure defined in the true DGP, W_n, by construction. As long as $m_{\tilde{X}} = \tilde{X}' W \tilde{X} / \tilde{X}' \tilde{X}$ is nil, the coefficient estimate β_{TD} will not be affected. In any other case, the sign of the bias is given by the sign of the product of λ_0 and β_0 or in short: $sgn(\beta_{TD} - \beta_0) = sgn(\lambda_0 \beta_0)$. Therefore, this distortion changes on a case-by-case basis. As usual, whenever the spatial lag is unimportant ($\lambda_0 = 0$), spatial heterogeneity in explanatory variables will not lead to biased estimates, as (3) clarifies.

Interestingly, these results are in line with evidence of omitting spatial lag terms in cross-sectional and non-dynamic panel models. Regarding the cross section, Pace, LeSage (2008) report that the coefficient estimate of the covariate will show an asymptotic bias when the spatial lag term is neglected if there is spatial dependence in both, the covariate as well as the regressand. The authors further report that this bias is increasing in either spatial dependence. Noticably, the results regarding panel models reported in Franzese Jr, Hays (2007) follow suit, and show the same pattern regardless of cross-sectional size or the number of observed periods. Appendix D, where the covariate is allowed to exhibit spatial and time dependence, shows that this is also the case here, as illustrated in Table D.3. Effects on marginal effects are further investigated in Section 5.

Team Spatial is confident that the data represents a spatial-autoregressive process and believes $\gamma_0 = 0$. Here, fixed-effects are concentrated out by a usual within-transformation and a tilde indicates this transformation. For practical purposes, one may consider the transformation proposed in Lee, Yu (2010a), which does not create time dependence in the disturbances. The bias of neglecting the time dimension is approximated analytically by comparing the likelihood estimators of a full model, indicated by superscripted '0', and the model chosen by Team Spatial, indicated by 'TS', as derived in Appendix A. Given the strict exogeneity assumption on X, the quasi maximum likelihood (QML) estimate is

Table 1: Bias directions

dependence		*team*	
space (λ_0)	*time* (γ_0)	*dynamic*	*spatial*
+	+	overestimates time	unclear
-	+	overestimates time	underestimates space

Note: Spatial and time dependence in the data generating process and bias of each research team – Team Dynamic estimating a dynamic panel model, and Team Spatial estimating a static spatial panel model, while the true model is a spatial dynamic panel model.

given by (Yu et al. 2008)

$$\lambda_{TS} \approx \frac{\sigma^2_{\tilde{Y}} - m_{\tilde{Y},\tilde{X}} \frac{\sigma^2_{\tilde{X},\tilde{Y}}}{\sigma^2_{\tilde{X}}}}{\sigma^2_{\tilde{Y}} m^{W^2}_{\tilde{Y}} + \frac{1}{nT}\mathrm{tr}(W^2) - \frac{\sigma^2_{\tilde{X},\tilde{Y}}}{\sigma^2_{\tilde{X}}}(m_{\tilde{Y},\tilde{X}})^2} \tag{5}$$

$$\beta_{TS} \approx \left(\sigma_{\tilde{X},\tilde{Y}}\left(1 - m_{\tilde{X},\tilde{Y}}\lambda_{TS}\right)\right)/\sigma^2_{\tilde{X}}, \tag{6}$$

where $m_{\tilde{Y},\tilde{X}} = m_{\tilde{X},\tilde{Y}} = \tilde{Y}'W\tilde{X}/\tilde{Y}'\tilde{X}$ and $m^{W^2}_{\tilde{Y}} = \tilde{Y}'W^2\tilde{Y}/\tilde{Y}'\tilde{Y}$. Denoting β_{TS} as a function of the estimate of the spatial parameter clarifies that whenever the spatial parameter is biased, then so is β_{TS}. Because inserting the reduced form DGP into equations (5) and (6) as in an omitted variable analysis does not seem constructive, another angle is pursued. To see the fault in neglecting the time dimension, consider the QML estimator $\hat{\lambda}^0_{QML}$ of the full model:

$$\lambda^0_{QML} \approx \frac{m_{\tilde{Y}}\sigma^2_{\tilde{Y}} - m_{\tilde{X},\tilde{Y}} \frac{\sigma^2_{\tilde{Y},\tilde{X}}}{\sigma^2_{\tilde{X}}} - m_{\tilde{Y},\tilde{Y}_-} \frac{\sigma^2_{\tilde{Y},\tilde{Y}_{-1}}}{\sigma^2_{\tilde{Y}}}}{\sigma^2_{\tilde{Y}} m^{W^2}_{\tilde{Y}} + \frac{1}{nT}\mathrm{tr}(W^2) - \frac{\sigma^2_{\tilde{Y},\tilde{X}}}{\sigma^2_{\tilde{X}}}(m_{\tilde{Y},\tilde{X}})^2 - (m_{\tilde{Y},\tilde{Y}_-})^2}, \tag{7}$$

where $m_{\tilde{Y},\tilde{Y}_-} = \tilde{Y}'W\tilde{Y}_-/\tilde{Y}'\tilde{Y}_-$ is defined accordingly. Note that the estimate β^0_{QML} is determined equally in the full model by inserting equation (7) into (6). If there is no co-variation between contemporary and time-lagged values, the two estimates coincide. If there is co-variation, estimates differ. While the difference in the nominator of eqs. (5) and (7) depends on the sign of the dependence, it is negative in the denominator. Hence, the sign of $\lambda_{TS} - \lambda^0_{QML}$ is not entirely clear. Whether the estimate suffers from upward or downward bias hinges on $m_{\tilde{Y},\tilde{Y}_-}$, which reflects spatial correlation between the contemporary and time-lagged values of the dependent variable which again only depends on λ_0. The direction of the bias is only clear in cases where the nominator increases.

Table 1 summarizes the findings so far. Because all analytical derivations are based on the approximation $A_n \approx I_n + \lambda_0 W_n$, the next section attempts to confirm the expectations built here by means of a Monte Carlo simulation. In other words, the tale of the two research teams will be told many times. Note that the approximation of the spatial multiplier is not used at any other stage of the manuscript unless explicitly stated otherwise.

4 Monte Carlo simulation

There is little reason to expect finding explosive spatial dynamic processes in economic reality, at least on the regional level. Therefore, Monte Carlo simulations apply to stable DGPs. The condition is that the absolute eigenvalues of $A_n\gamma_0$ need to be smaller than unity (Lee, Yu 2010b). Applying the first-order approximation of the matrix inverse to the reduced form of (1), the condition simplifies to $1 > |\gamma_0|\cdot(1 + |\lambda_0|)$. Because it is not feasible to cover the whole set of parameters that obey this restriction, a subset of combinations as described in Table 2 is used. The underlying spatial structure is

Table 2: Parameter combinations.

team	λ_0, step size	γ_0, step size	combinations
dynamic	[-0.60; 0.60], 0.15	[0; 0.80], 0.05	144
spatial	[-0.60; 0.60], 0.05	[0; 0.80], 0.20	125

Note: Parameter combinations of the simulated data generating process as given in equation (1) used for the corresponding calculation of the bias in parameters and marginal effects of the corresponding research team. λ_0 denotes spatial dependence, γ_0 represents time dependence.

given by 35 Austrian NUTS 3 regions. A smaller number of regions is preferable from a computational point of view, as the spatial weights matrix is increasing quadratically in the number of regions. The spatial weights matrix is based on the road distance between the central cities of each NUTS 3 region. Most frequently researchers use either inverse distance measures or contiguity schemes. Given that user-friendly packages to calculate distances accurately for a larger number of connections without cost exist for most statistics programs, distance appears a more natural approach. Using a maximum-row normalized spatial weights matrix based on these inverse distance measures, the DGP is simulated for 30 periods according to equation (1). Pre-estimation periods ensure independence of spurious existence or absence of spatial correlation through randomly assigned start values (Lee, Yu 2010a). Periods 31 to 60 are handed to the research teams.

Team Dynamic uses a GMM procedure in the spirit of Arellano, Bover (1995) provided in the *xtabond2* package for Stata (Roodman 2006). As discussed above, Team Spatial uses a QML approach based on Lee, Yu (2010a), provided in the *xsmle* package for Stata (Belotti et al. 2017).

X_{nt} is strictly exogenous and the coefficient β_0 is set to 0.75. As expected, β_{TD} is largely unaffected by omitting uncorrelated terms by virtue of its strict exogeneity. As indicated by equation (6) the bias in the coefficient estimate of Team Spatial (β_{TS}) increases with the bias in λ_{TS}.[3]

Since we are interested in the interplay of λ_0 and γ_0, simulations are run for different parameter combinations, summarized in table 2. Each combination is simulated 500 times. Hence, results are based on 134,500 runs. Due to the large amount of information, results are summarized graphically in figures 1 and 2.

In general, the expectations formed in the former section find strong support. Simulation results for Team Dynamic are in line with table 1. However, with negative spatial dependence, the bias in estimating the autoregressive parameter is smaller compared to cases with positive spatial dependence. Still, both situations show that the autoregressive parameter is overestimated for every combination of λ_0 and γ_0, and high values of autocorrelation are disproportionately vulnerable to neglecting the spatial dimension, as the bias appears to explode at a range of $\lambda_0 \approx 0.65$. This is an alarming result given that many dynamic processes in regional economics show high levels of autocorrelation.

Figure 2 describes the bias of Team Spatial in estimating λ_{TS}. As discussed in the previous section, an underestimation in the case of negative spatial dependence (and positive time dependence) could be expected, and finds support in the simulation results. Notably, this overestimation in absolute terms pertains to the case of positive spatial dependence, implying that the importance of space will always be overstated. As before, this bias is larger the more important the omitted dimension is. In many (regional) economic investigations the time dependence is large, and the results here make clear that the importance of spillovers is likely to be overstated in cases where the 'within-dynamics' are not accounted for. It has to be stressed that accounting for potentially spatial heterogeneous (time-invariant) fixed effects does not help to counter this problem. Anselin, Arribas-Bel (2013), looking into the ability of 'spatial fixed effects' to control for spatial dependence in cross-sectional models, conclude that such spatial fixed effects

[3]Graphs included in the supplementary material. The do.-files for running the simulations in Stata are provided as supplementary material. A shape-file containing Austrian NUTS3 regions can be found, for example, at https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data

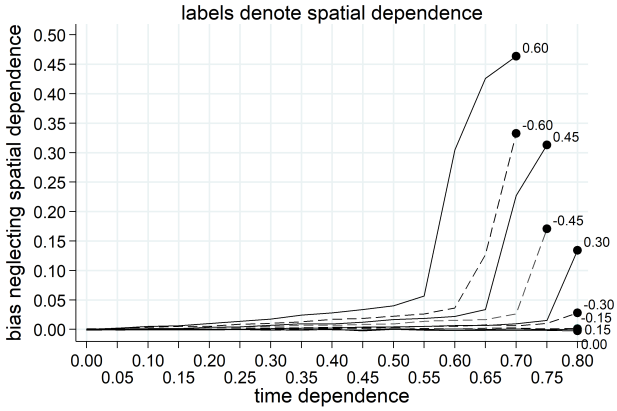

Figure 1: Bias of Team Dynamic neglecting the spatial dimension. The vertical axis describes $\gamma_{TD} - \gamma_0$, the horizontal axis depicts γ_0. Labels denote spatial dependence in the DGP (λ_0).

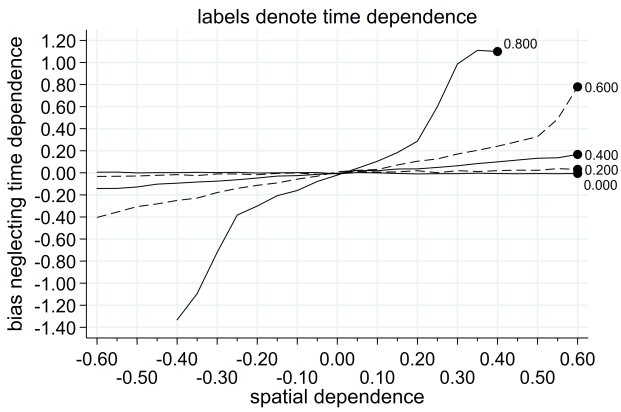

Figure 2: Bias of Team Spatial neglecting the time dimension. The vertical axis describes $\lambda_{TS} - \lambda_0$, the horizontal axis depicts λ_0. Labels denote spatial dependence in the DGP (γ_0).

might only filter out spatial dependence if there is no distance decay, implying group-wise dependencies. This argument may be extended to a panel set up, where such spatial-fixed effects would fall in line with panel-fixed effects. Hence, unless the true DGP features such group-wise structures, fixed effects do not suffice to account for spatial dependence.

Considerable effort has been given to (mis-)specification issues in spatial econometrics, and also in order to find empirical strategies to identify the underlying date generating processes (Florax et al. 2003, Anselin 2002, for example). As noted by Anselin (2010), typically one cannot differentiate between spatial heterogeneity and spatial dependence, an issue summarized as the inverse problem, closely related to Manski's reflection problem. Stressing these difficulties more rigorously, Gibbons, Overman (2012) suggest different strategies. First, one should aim at exploiting quasi-experimental set ups to pinpoint sources of exogenous variation. More importantly, the authors suggest using a reduced form spatial lag of X (SLX) approach to capture spatial dependence. This is a strategy that Team Dynamic might use, seeing that testing for spatial correlation in their residuals will most likely reject the null of no correlation. Graph E.2 demonstrates that this strategy does not relieve Team Dynamic of their bias completely, but attenuates it to a certain degree. Certainly, the extent to which WX is a 'good' replacement of WY in a reduced form depends on the explanatory power of the exogenous variable in determining Y.

In order to rule out that the results may be driven by the particular shape of Austria,

with a large longitudinal stretch (from 9.6 to 16.95 degrees) compared to latitudinal (from 46.53 to 48.82 degrees), the exercise is repeated using German NUTS 2 regions (see Figure E.1). Results of this robustness check are presented in Appendix E.

It can be expected that one will not find significant differences in marginal effects when comparing two estimates where one is based on maximum-row normalization and one is based on row-normalization, for example. While maximum-row normalization is used for representative convenience, row-normalization is the much-preferred specification in the literature. As LeSage, Pace (2014) argue, the correlation between spatial lag terms, for example W_nY_{nt}, based on two such spatial weights matrices that are themselves based on the same geographical information is shown to be very high. Therefore, it appears unlikely that any of the results, especially in terms of marginal effects of the covariate, differ due to the exact specification of W_n. Figure E.5 as well as Figure E.6 in Appendix E present simulation results using a row-normalized version of W_n, showing that this argument is indeed valid.

5 The marginal effect of the covariate

What if, instead of specifying the importance of time or space, the real interest lies in eliciting the effect of the covariate X? It does not suffice to compare parameter estimates, as both research teams assume different impact channels compared to the true DGP. In a similar vain, LeSage, Pace (2018) or Debarsy et al. (2012) argue that marginal effects should be the main target when comparing spatial econometric approaches in Monte Carlo studies, mainly because they are non-linear combinations of estimated parameters and often the main focus in applied research.[4] In order to elicit the respective marginal effects, differences in the assumed DGP need to be accounted for. In short, Team Spatial will not be able to differentiate between short- and long-run effects, and Team Dynamic will be unable to estimate indirect effects, the importance of feedback, and total effects. Therefore, it appears most appropriate to compare marginal effects of both teams with the corresponding correct ones.

The true marginal effects can be separated into direct, indirect, and total, and short-run or contemporaneous versus long-run or equilibrium effects. Regarding the short-run, they are based on the product of the coefficient and the spatial multiplier matrix:

$$\mathbf{y}_x^0 = \frac{\partial Y_{nt}}{\partial X_{nt}} = \beta_0 \left(I_n - \lambda_0 W_n\right)^{-1} \tag{8}$$

The direct marginal effect is calculated as the average diagonal element of $\mathbf{y}_x^0$ and the indirect effect is the average row sum of $\mathbf{y}_x^0$ excluding the diagonal element. The total effect follows as the sum of direct and indirect effects (LeSage, Pace 2009). Long-run or equilibrium effects are calculated accordingly using

$$\mathbf{y}_x^{0*} = \frac{\beta_0}{1-\gamma_0} \left(I_n - \varphi_0 W_n\right)^{-1}, \tag{9}$$

where $\varphi_0 = \lambda_0/(1-\gamma_0)$.

As discussed above, Team Dynamic is able to estimate β_0 without bias as long as $m_{\tilde{X}}$ is nil. Even though this assumption is unlikely to hold using observational data, it represents a case in which the covariate is assigned (quasi-)randomly. Because the team blocks out spatial interaction, it is only able to estimate a direct effect without any potential feedback captured in the spatial multiplier. Hence, the marginal short-run effect y_x^{TD} is simply β_{TS}. As such, one might expect an overestimation in the case of negative spatial dependence, and an underestimation otherwise, and that Team Dynamic is close to estimating the true direct effects as long as this feedback is small, implying $(1/n) \cdot \text{diag}(\mathbf{y}_x^0) l_n \approx \beta_0$ using the approximation of the spatial multiplier. As noted above, this approximation is more accurate for low to moderate values of λ_0. Formally, the bias

[4] It has to be noted that the focus of the authors' contribution is the precision of different estimation techniques in estimating marginal effects of covariates, instead of parameter estimates that have been the focus of earlier studies.

Table 3: Bias of marginal direct short- and long-run effects of Team Dynamic.

λ_0 \ γ_0	0.2	0.4	0.6	0.8
	PANEL A: DIRECT SHORT-RUN EFFECT			
-0.3	-0.005	-0.006	-0.004	0.000
0	-0.008	-0.008	-0.007	-0.006
0.3	-0.005	-0.005	-0.006	0.018
	PANEL B: DIRECT LONG-RUN EFFECT			
-0.3	0.001	0.001	0.005	0.108
0	0.000	-0.001	0.000	0.021
0.3	0.000	0.004	0.001	0.726

Note: Bias of Team Dynamic relative to corresponding marginal effects in the true model. λ_0 denotes spatial dependence, γ_0 represents time dependence.

of the marginal direct short-run effect is given by

$$\begin{aligned} y_x^{TD} - \frac{1}{n} l_n' \left(\text{diag}(\mathbf{y}_x^0)\right) l_n &= \beta_{TD} - \frac{1}{n} l_n \left(\text{diag}\left(I_n' - \lambda_0 W_n\right)^{-1}\right) l_n \beta_0 \\ &= (\beta_{TD} - \beta_0) - \beta_0 \frac{1}{n} \sum_{k=1}^{\infty} l_n' \text{diag}(\lambda_0 W_n)^k l_n, \end{aligned} \tag{10}$$

where the first term $(\beta_{TD} - \beta_0)$ may be labeled 'coefficient bias' and the latter term may accordingly be called the 'feedback bias'. Hence, even the coefficient estimate is unbiased, the feedback bias will lead Team Dynamic to a false marginal effect. With respect to long-run effects, the upward bias of Team Dynamic comes into play. Similarly to the short-run, the bias in the long-run direct effect can be separated into a coefficient bias and a feedback bias, where the autoregressive parameter enters directly. Formally, we get

$$\begin{aligned} y_x^{TD*} - \frac{1}{n} l_n' \left(\text{diag}(\mathbf{y}_x^{0*})\right) l_n = \\ \left(\frac{\beta_{TD}}{1-\gamma_{TD}} - \frac{\beta_0}{1-\gamma_0}\right) - \frac{\beta_0}{1-\gamma_0} \frac{1}{n} \sum_{k=1}^{\infty} l_n' \text{diag}(\frac{\lambda_0}{1-\gamma_0} W_n)^k l_n, \end{aligned} \tag{11}$$

where $y_x^{TD*} = \beta_{TD}/(1-\gamma_{TD})$. In principle, the overestimation of the time-autoregressive parameter might counteract the absence of feedback channels. Hence, it might be possible that Team Dynamic gets close to the direct long-run effects, even though working with a misspecified model. On the other hand, given that spatial dependence is large enough such that the estimated autoregressive parameter is close to unity, unreasonably large or even perverted effects might result.

Table 3 clarifies that marginal direct short-run effects can indeed be estimated fairly well for moderate values of spatial dependence. Only for cases with positive spatial dependence and high autocorrelation the relative bias surpasses the 1% mark. Panel B displays the relative bias of marginal direct long-run effects, and shows a similar picture as figure 1. Because the bias in the short-run coefficient is quite small, most of the bias in the long-run coefficient is carried by the bias in the autoregressive parameter. As mentioned before, this bias is larger the larger the autoregressive component and the larger the spatial component. Indeed, the last column in panel B shows that the bias is substantial for negative spatial dependence, and becomes even larger with $\lambda_0 > 0$. In the extreme, the direct long-run effect has an average upward bias of approximately 73%.

Team Spatial can estimate direct, indirect and total marginal effects, but cannot differentiate between the short- and the long-run and will interpret marginal effects as equilibrium, hence long-run, effects. The bias in either marginal effect is based on the

Table 4: Bias of marginal direct and indirect long-run effects of Team Spatial

λ_0 / γ_0	-0.5	-0.25	0	0.25	0.5
			DIRECT LONG-RUN EFFECT		
0	-0.038	-0.038	-0.039	-0.039	-0.039
0.6	-0.652	-0.642	-0.639	-0.641	-0.673
			INDIRECT LONG-RUN EFFECT		
0	0.026	0.065	-0.004	-0.089	-0.053
0.6	-0.749	-0.824	0.038	-0.644	-0.833

Note: Bias of Team Spatial relative to the corresponding marginal effects in the true model. Column three ($\lambda_0 = 0$) of indirect long-run effects measures absolute bias because the true effect is nil. λ_0 denotes spatial dependence, γ_0 represents time dependence.

difference

$$\begin{aligned}\mathbf{y}_x^{TS} - \mathbf{y}_x^{0*} &= (I_n - \lambda_{TS} W_n)^{-1} \beta_{TS} - \frac{\beta_0}{1-\gamma_0} (I_n - \varphi_0 W_n)^{-1} \\ &= (\beta_{TS} - \frac{1}{1-\gamma_0}\beta_0) + \sum_{k=1}^{\infty} W_n^k \left(\lambda_{TS}^k \beta_{TS} - \lambda_0^k/(1-\gamma_0)^{k+1}\beta_0\right), \qquad (12)\end{aligned}$$

showing three potential sources. These are differences between the estimated spatial parameter λ_{TS} and the true one, differences between the estimated coefficient of the covariate β_{TS} and the true one, and the omission of the autoregressive component γ. Equation (6) clarifies that a random assignment of the covariate such that $m_{\tilde{X},\tilde{Y}}$ is nil will lead to an unbiased estimate β_{TS}. Similar to the argument for Team Dynamic, (quasi-) random assignment across space appears crucial. However, the omission of the autoregressive component will not relieve Team Spatial of its inability to measure correct direct, indirect, or total effects. The overestimation of the spatial parameter in absolute terms increases in λ_0 as shown in Figure 2, however, equation (12) shows that this might off-set the bias to a certain degree. The bias can be expected to be small for low levels of time dependence. Interestingly, LeSage, Pace (2018) mention that cases where the estimates of β and λ are negatively correlated may mitigate the bias in the marginal effect. On the other hand, they may also be aggravated (p.22). However, even with unbiased estimation, both direct and indirect long-run effects are likely downwardly biased because Team Spatial neglects the autoregressive multiplier $(1/1(-\gamma_0))$ entirely, which the overestimation of λ_{TS} is unlikely to make up for assuming typical parameter values. Table 4 presents simulation results for a situation with no autoregressive component and with a parameter γ_0 of 0.6. In line with the results of figure 2, the bias is larger when λ_0 is greater in absolute value.

To sum this section up, it has to be stressed that (quasi-) random assignment of the covariate turns out to be a central requirement for both teams. Only then, Team Dynamic will have a proper estimate of the direct short-run effect, but will likely overestimate the long-run effect given that the autoregressive parameter is bound to be overestimated. Team Spatial, even though possibly able to get a proper estimate of the coefficient of the covariate, will suffer from neglecting the autoregressive multiplier $(1/(1-\gamma_0))$.

6 Application: Cigarette demand

The openly available cigarette demand data used in Baltagi, Li (2004) for all landlocked US states is a standard data set that is frequently used to illustrate issues in spatial, dynamic, or spatial dynamic estimation methods, as in Kelejian, Piras (2017) or Debarsy et al. (2012), who demonstrate the interpretation and calculation of marginal effects in space-time models in the short- and long-run. It covers 30 years of cigarette sales per capita, the average price per pack, and the average income per capita. The spatial weights matrix used for this illustration is given by a queen-contiguity scheme. Thereby, cigarette sales are hypothesized to depend on the average price P per pack, the average price in

neighboring states, and the average income as exogenous determinants (Debarsy et al. 2012, Elhorst 2014, Kelejian, Piras 2017). Referring to the tale in Section 3, table (5) presents the estimation results of Team Dynamic, Team Spatial, and of a 'full' specification according to the DGP assumed in equation (1). The estimation of the full model is carried out using the estimator of Yu et al. (2008), which has already been used in previous Sections. The fixed effects specification can be expected to work well given that the data is available for the years 1963 to 1992. Recently, Jin et al. (2020) propose a quasi-maximum likelihood estimator in first differences for panels with a short time horizon that would potentially be vulnerable to the incidental parameter problem, which would otherwise represent a viable alternative estimator of the full model.

The observations that can be drawn are in line with the expectations formed above. With a large autoregressive component, the spatial lag is highly overestimated when the former component is neglected, as a comparison of columns three ('Team Spatial') and four ('full model') shows. Indeed, while remaining highly significant in the full specification, Team Spatial is bound to overestimate the spatial lag parameter fourfold. Further, with a small spatial component, the potential bias in estimating the time component is small. Both estimators of Team Dynamic are very close to the full model in terms of the autoregressive parameter.

With respect to the covariates, parameter estimates themselves are not meaningful. Rather, marginal effects as discussed in section 5 need to be compared. In order to illustrate this comparison, estimated marginal effects of the average price are presented in panel 2 of Table 5. For Team Dynamic, it can be seen that both short-run and long-run (direct) effects are similar to those of the fully specified model. This is unsurprising given that the spatial component is small such that the autoregressive parameter can be estimated fairly well and the spatial multiplier is small. Because Team Spatial misses the autoregressive component in calculating long-run effects, neither the direct nor the indirect effect are similar to those in the full specification and are underestimated quite substantially. The indirect effect even shows the opposite sign, pointing towards the structural misspecification.

Regarding the observations made above, it seems quite unlikely that $m_{\tilde{X},\tilde{Y}}$ is zero in this setting such that the bias of the spatial-autoregressive parameter is transferred to the coefficient of the average price per pack. Team Spatial interprets marginal effects as long-run elasticities by construction and consequently underestimates them.

Table C.1 in Appendix (C) discusses some further possible modeling choices of both research teams. Team Spatial may argue that time-fixed effects already capture the time dimension. Given that state-fixed effects account for differences in levels, time-fixed effects would 'catch' average dynamics. Compared to the column (3) in table 5, not including time-fixed effects indeed induces a larger estimate of the spatial lag parameter, as column (1) illustrates. On the other hand, Team Dynamic might follow up on the argument of Gibbons, Overman (2012) and use a reduced-form spatial lag of X model. As shown in column (2), this approach appears to work reasonably well in this case, as both the autoregressive and the coefficient of the price are almost identical to those in the full specification. Note that the SLX specification allows calculating direct and indirect effects, albeit assuming local spillovers that do not account for potential feedback loops. Alternatively, Team Dynamic might simply run a dynamic spatial error model (SEM) as proposed by Su, Yang (2015).[5] As discussed in LeSage, Pace (2009), if covariates are the source of spatial heterogeneity, the corresponding dynamic SEM can be rewritten as a dynamic spatial Durbin model, in which case the results are naturally very similar to those reported in column (4) in Table 5.[6]

7 Conclusions

In a time when geo-referenced data becomes more available and stretches across longer time periods, researchers are able to account for dependencies in both dimensions, space and time. The aim of this study is to give insights to the problems one can expect if

[5] Unfortunately, this estimator is not implemented in statistical software.
[6] The author thanks an anonymous reviewer for these remarks.

Table 5: Empirical illustration; dependent variable: log consumption per capita

	Team Dynamic		Team Spatial	*full model*
	(1) fixed effects	(2) bias corrected	(3) QML	(4) QML
Panel 1: estimated parameters				
time lag	0.832***	0.862***		0.867***
	(32.60)	(58.55)		(24.63)
spatial lag			0.210***	0.049***
			(9.24)	(2.88)
price	-0.296***	-0.277***	-1.001***	-0.264***
	(-8.50)	(-11.96)	(-9.05)	(-4.95)
income	0.114***	0.101***	0.467***	0.090***
	(3.06)	(3.81)	(4.17)	(3.73)
W · price	0.084	0.086**	0.096	0.159**
	(1.64)	(2.48)	(0.59)	(2.17)
R^2	0.86		0.68	0.90
Panel 2: marginal effect of price				
short run:				
direct	-0.296***	-0.277***		-0.262***
	(72.25)	(143.04)		(146.23)
indirect				0.152***
				(12.81)
total				-0.110**
				(5.80)
long run:				
direct	-1.765***	-2.004***	-1.006***	-1.942***
	(39.80)	(88.82)	(848.22)	(129.16)
indirect			-0.144**	0.685
			(4.24)	(2.61)
total			-1.150***	-1.257**
			(197.25)	(6.39)

Observations: 1334 (N=46, T=29). Column 'fixed effects' follows a least squares dummy variable estimation. Column 'bias corrected' indicates the dynamic panel bias correction advocated by Kiviet (1995). Columns 'QML' and 'QML full' present results of a (dynamic) fixed effects spatial autoregressive model as outlined in Yu et al. (2008), where the latter applies a bias correction. Robust t-statistics in parentheses in panel 'estimated parameters'. Test statistics in panel 'marginal effects' are all of Wald-type for matters of comparability. All specifications include year-fixed effects. Data can be downloaded, for example, at spatial-panels.com/software. All variables in logs. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

either one dimension is neglected. In order to illustrate the issue at hand, a tale of two research teams is told, each one being agnostic of one dimension. The results show that the neglected dimension projects onto the other, thereby biasing coefficient estimates, and subsequently tests, marginal effects, and predictions. Neither purely dynamic nor purely (static) spatial estimation approaches are able to estimate 'their side' of the data appropriately. Analytical approximations combined with simulation results show that the direction of these biases in spatial- and time-autoregressive parameters can be determined. Even though signs are correctly estimated, the importance of the remaining dimension is overstated. Investigations of spatial dynamic processes that neglect space are likely to report slower convergence speeds, are more likely to 'find' unit roots when in fact processes are stable, and will show severely biased long-run effects. These results are amplified for positive spatial dependence. These biases further work their way through to marginal effects of covariates, which are shown to be equally affected.

Likewise, estimates of spatial dependence without considering time can be severely biased upwards in absolute terms. In these cases, the role of spillovers is likely to be overestimated. Because partial adjustment is ruled out when the time dimension is denied, marginal effects are bound to be underestimated given that autocorrelation is large. Both cases offer a wide range of research questions that may be reassessed considering these pitfalls. A short application to real data supports these findings. While the sizes of simulated biases are conditional on the spatial structure and weights matrix, the signs are robust. Figures E.3 and E.4 in Appendix E display qualitatively identical simulation

results of Team Dynamic and Team Spatial using German NUTS 2 regions as underlying spatial structure. Further, Figures E.5 and E.6 clarify that the result is not dependent on the normalization of the spatial weights matrix.

In terms of methodological recommendations, several aspects are worth noting. First of all, practitioners should have sound theoretical foundations that would exclude either partial adjustment or spatial interaction. More often, practitioners have used dynamic rather than spatial models. In that sense, tests for spatial residual correlation should be done routinely in regional and urban economics applications to check for the presence of any kind of spatial interaction. Of course, this also holds for non-dynamic applications. Likewise, spatial panel applications should routinely check for autoregressive residuals by the same argument. Coming back to the argument of Gibbons, Overman (2012) and the reflection issue, a reduced form spatial dynamic approach might be preferred. As in most issues regarding identification, sources of exogenous variation are most crucial. In the words of the tale of Team Dynamic and Team Spatial, as so often, it would be beneficial for all to put their heads together.

References

Anderson TW, Hsiao C (1982) Formulation and estimation of dynamic models using panel data. *Journal of econometrics* 18: 47–82. CrossRef.

Anselin L (2002) Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics* 27: 247–267. CrossRef.

Anselin L (2010) Thirty years of spatial econometrics. *Papers in regional science* 89: 3–25. CrossRef.

Anselin L, Arribas-Bel D (2013) Spatial fixed effects and spatial dependence in a single cross-section. *Papers in Regional Science* 92: 3–17

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies* 58: 277–297. CrossRef.

Arellano M, Bover O (1995) Another look at the instrumental variable estimation of error-components models. *Journal of econometrics* 68: 29–51. CrossRef.

Baltagi BH, Li D (2004) Prediction in the panel data model with spatial correlation. In: *Advances in spatial econometrics*. Springer, 283–295. CrossRef.

Belotti F, Hughes G, Mortari AP (2017) Spatial panel-data models using stata. *The Stata Journal* 17: 139–180. CrossRef.

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87: 115–143. CrossRef.

Cliff A, Ord K (1972) Testing for spatial autocorrelation among regression residuals. *Geographical analysis* 4: 267–284. CrossRef.

Cox D, Wermuth N (1996) *Multivariate Dependencies: Models, Analysis and Interpretation*, Volume 67. CRC Press. ISBN 9780367401375

Debarsy N, Ertur C, LeSage JP (2012) Interpreting dynamic space-time panel data models. *Statistical Methodology* 9: 158–171. CrossRef.

Elhorst P (2012) Dynamic spatial panels: Models, methods, and inferences. *Journal of geographical systems* 14: 5–28. CrossRef.

Elhorst P (2014) *Spatial econometrics: From cross-sectional data to spatial panels*, Volume 479. Springer

Elhorst P, Zandberg E, De Haan J (2013) The impact of interaction effects among neighbouring countries on financial liberalization and reform: A dynamic spatial panel data approach. *Spatial Economic Analysis* 8: 293–313. CrossRef.

Florax RJ, Folmer H, Rey SJ (2003) Specification searches in spatial econometrics: the relevance of hendry's methodology. *Regional Science and Urban Economics* 33: 557–579. CrossRef.

Franzese Jr RJ, Hays JC (2007) Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political analysis*: 140–164

Gibbons S, Overman HG (2012) Mostly pointless spatial econometrics? *Journal of Regional Science* 52: 172–191. CrossRef.

Hsiao C (2014) *Analysis of panel data*. Number 54. Cambridge university press. CrossRef.

Islam N (1995) Growth empirics: a panel data approach. *The Quarterly Journal of Economics* 110: 1127–1170. CrossRef.

Jin F, Lee Lf, Yu J (2020) First difference estimation of spatial dynamic panel data models with fixed effects. *Economics Letters* 189: 109010

Kelejian H, Piras G (2017) *Spatial econometrics*. Academic Press. ISBN 9780128133873

Kiviet JF (1995) On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics* 68: 53–78. CrossRef.

Lee Lf, Yu J (2010a) Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics* 154: 165–185. CrossRef.

Lee Lf, Yu J (2010b) Some recent developments in spatial panel data models. *Regional Science and Urban Economics* 40: 255–271. CrossRef.

LeSage JP, Pace RK (2009) *Introduction to spatial econometrics*. CRC press, Boca Raton, FL. CrossRef.

LeSage JP, Pace RK (2014) The biggest myth in spatial econometrics. *Econometrics* 2: 217–249

LeSage JP, Pace RK (2018) Spatial econometric Monte Carlo studies: Raising the bar. *Empirical Economics* 55: 17–34. CrossRef.

Nickell S (1981) Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*: 1417–1426. CrossRef.

Pace RK, LeSage JP (2008) Biases of ols and spatial lag models in the presence of an omitted variable and spatially dependent variables. *SSRN Electronic Journal* (1133438). CrossRef.

Pijnenburg K (2017) The spatial dimension of US house prices. *Urban Studies* 54: 466–481. CrossRef.

Rios V (2017) What drives unemployment disparities in European regions? A dynamic spatial panel approach. *Regional Studies* 51: 1599–1611. CrossRef.

Roodman D (2006) How to do xtabond2. In: *North American Stata Users' Group Meetings 2006*, Number 8. Stata Users Group

Roodman D (2009) A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics* 71: 135–158. CrossRef.

Shi W, Lee Lf (2017) Spatial dynamic panel data models with interactive fixed effects. *Journal of Econometrics* 197: 323–347. CrossRef.

Su L, Yang Z (2015) QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics* 185: 230–258

Vega SH, Elhorst JP (2014) Modelling regional labour market dynamics in space and time. *Papers in Regional Science* 93: 819–841. CrossRef.

Wanzenböck I, Piribauer P (2018) R&D networks and regional knowledge production in Europe: Evidence from a space-time model. *Papers in Regional Science* 97: S1–S24. CrossRef.

Yu J, De Jong R, Lee Lf (2008) Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics* 146: 118–134. CrossRef.

© 2021 by the authors. Licensee: REGION – The Journal of ERSA, European Regional Science Association, Louvain-la-Neuve, Belgium. This article is distributed under the terms and conditions of the Creative Commons Attribution, Non-Commercial (CC BY NC) license (http://creativecommons.org/licenses/by-nc/4.0/).

Appendix

A Approximation of biases

Team Dynamic runs two-stage least squares on first differences. The bias is determined by inserting the reduced form DGP for $\bar{y}$:

$$\boldsymbol{\phi}_{TD} - \phi_0 = \left(\tilde{\Gamma}'\tilde{\Gamma}\right)^{-1} \tilde{\Gamma}' \left[(I_{nT} - \lambda_0 W)^{-1} (\tilde{\Gamma}\phi_0 + \tilde{V})\right] - \phi_0 \tag{A.1}$$

This is the point where the approximation of the spatial multiplier enters. Multiplying out and exogenous residual terms yields the desired result of equation (3). Under strict exogeneity and assuming well-behaved innovations $A_{nT} = (I_{nT} - \lambda_0(I_T \otimes W_n))^{-1} \approx I_{nT} + \lambda_0 W$ as well as $P_Z = I_{nT}$,

$$\boldsymbol{\phi}_{TD} - \boldsymbol{\phi}_0 = \lambda_0 \left(\tilde{\Gamma}'\tilde{\Gamma}\right)^{-1} \tilde{\Gamma}' W \tilde{\Gamma} \boldsymbol{\phi}_0 \tag{A.2}$$

With K right-hand side variables ($k = 1$ indicating the time lag, $k = 2$ the first exogenous variable, and so on), one gets

$$\mathbf{Q} \equiv \left(\tilde{\Gamma}'\tilde{\Gamma}\right)^{-1} \left[\tilde{\Gamma}'\tilde{\Gamma} \odot \mathbf{M}\right] = \begin{bmatrix} \sum_{j=1}^{K} \eta_{1j}\sigma_{j1}m_{j1} & \dots & \sum_{j=1}^{K} \eta_{1j}\sigma_{jk}m_{jk} \\ \sum_{j=1}^{K} \eta_{2j}\sigma_{j2}m_{j2} & \dots & \sum_{j=1}^{K} \eta_{2j}\sigma_{jk}m_{jk} \\ \vdots & \vdots & \vdots \\ \sum_{j=1}^{K} \eta_{Kj}\sigma_{j1}m_{j1} & \dots & \sum_{j=1}^{K} \eta_{Kj}\sigma_{jK}m_{jK} \end{bmatrix} \tag{A.3}$$

where σ_{ij} (η_{ij}) is the row i, column j element in $\tilde{\Gamma}'\tilde{\Gamma}$ ($(\tilde{\Gamma}'\tilde{\Gamma})^{-1}$). Hence, the approximated bias is given as

$$\text{plim}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0) \approx \lambda_0 \mathbf{Q}\boldsymbol{\phi}_0 = \lambda_0 \begin{bmatrix} \sum_{k=1}^{K} \phi_{0,k} \sum_{j=1}^{K} \eta_{1j}\sigma_{jk}m_{jk} \\ \vdots \\ \sum_{k=1}^{K} \phi_{0,k} \sum_{j=1}^{K} \eta_{Kj}\sigma_{jk}m_{jk} \end{bmatrix} \tag{A.4}$$

Team Spatial estimates a spatial autoregressive model by QML. First consider a correctly specified model with first order conditions (Yu et al. 2008)

$$\tilde{\Gamma}'\tilde{V} = 0 \tag{A.5}$$

$$\tilde{Y}'W\tilde{V} - \text{tr}(G) = 0 \tag{A.6}$$

$$\tilde{V}'\tilde{V} = 0, \tag{A.7}$$

where $\tilde{\Gamma} = \left(\tilde{Y}_-, \tilde{X}\right)$, $\tilde{V} = (I - \lambda_0 W)\tilde{Y} - \tilde{\Gamma}\delta_0 = \tilde{Y} - \lambda_0 W\tilde{Y} - \tilde{\Gamma}\delta_0$, $G = W(I_{nT} - \lambda_0 W)^{-1}$, and $\delta_0 = \begin{bmatrix}\gamma_0 & \beta_0\end{bmatrix}'$. These conditions follow from the concentrated (within) likelihood function indicated by a tilde. The system of equations can easily be solved analytically by using the approximation of the spatial multiplier matrix in calculating the trace of G.

$$\text{tr}(G) = \text{tr}(W\,(I - \lambda_0 W)^{-1}) \approx \text{tr}(W\,(I + \lambda_0 W)) \approx \lambda_0 \text{tr}(W^2). \tag{A.8}$$

Condition (A.7) yields and estimate of the error variance and is not important for the approximation of the bias. The problem thus boils down to solving the linear equation system given by conditions (A.5) and (A.6):

$$\begin{bmatrix} \tilde{Y}_-'W\tilde{Y} & \tilde{Y}_-'\tilde{Y}_- & \tilde{Y}_-'\tilde{X} \\ \tilde{X}'W\tilde{Y} & \tilde{X}'\tilde{Y}_- & \tilde{X}'\tilde{X} \\ \tilde{Y}'W^2\tilde{Y} + \text{tr}(W^2) & \tilde{Y}'W\tilde{Y}_- & \tilde{Y}'W\tilde{X} \end{bmatrix} \begin{bmatrix} \hat{\lambda} \\ \hat{\gamma} \\ \hat{\beta} \end{bmatrix} \approx \begin{bmatrix} \tilde{Y}_-'\tilde{Y} \\ \tilde{X}'\tilde{Y} \\ \tilde{Y}'W\tilde{Y} \end{bmatrix} \tag{A.9}$$

Since all variables are demeaned, one can rewrite the problem in terms of (co-)variances, reflecting the notion of variance decomposition. According to the DGP defined above,

$\text{plim}(X'Y_-) = 0$. The resulting estimates are thus

$$\hat{\lambda}^0_{QML} \approx \frac{m_{\tilde{Y}}\sigma^2_{\tilde{Y}} - m_{\tilde{X},\tilde{Y}}\frac{\sigma^2_{\tilde{Y},\tilde{X}}}{\sigma^2_{\tilde{X}}} - m_{\tilde{Y},\tilde{Y}_-}\frac{\sigma^2_{\tilde{Y},\tilde{Y}_{-1}}}{\sigma^2_{\tilde{Y}}}}{\sigma^2_{\tilde{Y}} m^{W^2}_{\tilde{Y}} + \frac{1}{nT}\text{tr}(W^2) - \frac{\sigma^2_{\tilde{Y},\tilde{X}}}{\sigma^2_{\tilde{X}}}(m_{\tilde{Y},\tilde{X}})^2 - (m_{\tilde{Y},\tilde{Y}_-})^2} \tag{A.10}$$

$$\beta^0_{QML} \approx \frac{\sigma_{\tilde{Y},\tilde{X}}\left(1 - \hat{\lambda}^0_{QML} m_{\tilde{X},\tilde{Y}}\right)}{\sigma^2_{\tilde{X}}} \tag{A.11}$$

$$\gamma^0_{QML} = \frac{\sigma_{\tilde{Y},\tilde{Y}_{-1}}\left(1 - \hat{\lambda}^0_{QML} m_{\tilde{Y},\tilde{Y}_-}\right)}{\sigma^2_{\tilde{Y}}} \tag{A.12}$$

Team Spatial uses the same QML estimator under the assumption that $\gamma_0 = 0$. Hence the system of equations is given by

$$\begin{bmatrix} \tilde{X}'W\tilde{Y} & \tilde{X}'\tilde{X} \\ \tilde{Y}'W^2\tilde{Y} + \text{tr}(W^2) & \tilde{Y}'W\tilde{X} \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \beta_0 \end{bmatrix} \approx \begin{bmatrix} \tilde{X}'\tilde{Y} \\ \tilde{Y}'W\tilde{Y} \end{bmatrix} \tag{A.13}$$

The solution follows immediately as

$$\lambda_{TS} = \frac{\sigma^2_{\tilde{Y}} m_{\tilde{Y}} - m_{\tilde{Y},\tilde{X}}\frac{\sigma^2_{\tilde{X},\tilde{Y}}}{\sigma^2_{\tilde{X}}}}{\sigma^2_{\tilde{Y}} m^{W^2}_{\tilde{Y}} + \frac{1}{nT}\text{tr}(W^2) - \frac{\sigma^2_{\tilde{X},\tilde{Y}}}{\sigma^2_{\tilde{X}}}(m_{\tilde{Y},\tilde{X}})^2} \tag{A.14}$$

$$\beta_{TS} = \frac{\sigma_{\tilde{X},\tilde{Y}}\left(1 - m_{\tilde{X},\tilde{Y}}\lambda_{TS}\right)}{\sigma^2_{\tilde{X}}} \tag{A.15}$$

The result in equation (5) follows for $\sigma^2_{\tilde{Y}} = 1$.

B The sign of Moran's I

The MA(∞) representation of the DPG is given by

$$Y_{nt} = \gamma_0^t A_n^t Y_{n0} + \sum_{\tau=0}^{t} \gamma_0^\tau A_n^{\tau+1}\left(X_{n,t-\tau}\beta_0 + V_{n,t-\tau}\right) \tag{B.1}$$

For $\gamma_0 < 1$, the first term will vanish. Moran's I of the sample is the sum of the yearly values (divided by the square sum of the sample). For period t:

$$\begin{aligned} Y'_{nt}W_nY_{nt} = &\left(\sum_{\tau=0}^{t} \gamma_0^\tau A_n^{\tau+1}\left(X_{n,t-\tau}\beta_0 + V_{n,t-\tau}\right)\right)' W_n \\ &\left(\sum_{\tau=0}^{t} \gamma_0^\tau A_n^{\tau+1}\left(X_{n,t-\tau}\beta_0 + V_{n,t-\tau}\right)\right) \end{aligned} \tag{B.2}$$

assuming that cross-time values are nil we can rewrite

$$\begin{aligned} Y'_{nt}W_nY_{nt} = &\sum_{\tau=0}^{t} \gamma_0^{2\tau}\left(\beta_0^2 X'_{n,t-\tau}\left(A_n^{\tau+1}\right)' W_n A_n^{\tau+1} X_{n,t-\tau}\right. \\ &\left. + V'_{n,t-\tau}\left(A_n^{\tau+1}\right)' W_n A_n^{\tau+1} V_{n,t-\tau}\right) = \sum_{\tau=0}^{t} \gamma_0^{2\tau}\mathcal{M}_\tau \end{aligned} \tag{B.3}$$

Hence, $\text{plim}(Y'_{nt}W_nY_{nt}) = \text{plim}\sum_{\tau=0}^{t}\gamma_0^{2\tau}\mathcal{M}_\tau$ and γ_0 and β_0 cannot influence the sign. Using the approximation of the spatial multiplier matrices for the DGPs of Y and X, for

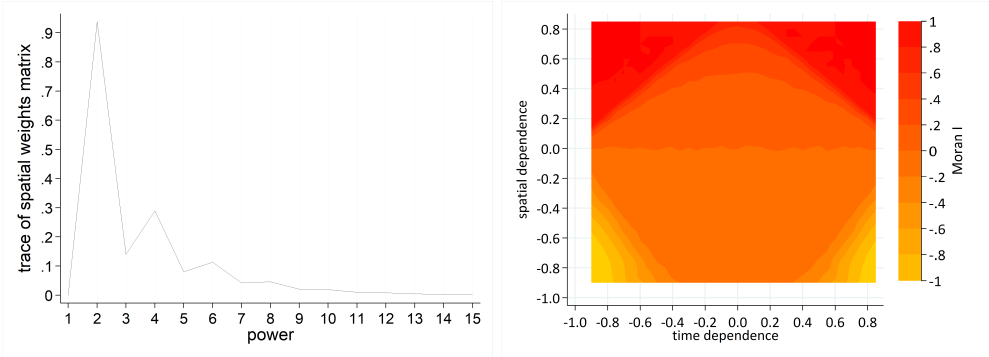

Figure B.1: Trace of powers of the spatial weights matrix (left) and simulated Moran's I values (right).

$\tau = 0$ it follows:

$$\begin{aligned}\mathcal{M}_0 \approx \beta_0^2 \big[&\sigma_\varepsilon \alpha_0 \left(2\mathrm{tr}(W_n^2) + \alpha_0 \mathrm{tr}(W_n^3)\right) + \\ &+ 2\lambda_0 \left(\sigma_\varepsilon^2 \alpha_0 \left(2\mathrm{tr}(W_n^3) + \alpha_0 \mathrm{tr}(W_n^4)\right) + \sigma_\varepsilon^2 \mathrm{tr}(W_n^2)\right) \\ &+ \lambda_0^2 \left(\sigma_\varepsilon^2 \alpha_0 \left(2\mathrm{tr}(W_n^4) + \alpha_0 \mathrm{tr}(W_n^5)\right) + \sigma_\varepsilon^2 \mathrm{tr}(W_n^3)\right)\big] \\ &+ \lambda_0 \sigma^2 \left(2\mathrm{tr}(W_2) + \lambda_0 \mathrm{tr}(W_n^3)\right)\end{aligned} \tag{B.4}$$

Evidently, it is difficult to get more insight by investigating the terms $\mathcal{M}_\tau$ in approximation. For $\alpha_0 = 0$, the term is greatly simplified to

$$\mathcal{M}_0 \approx \lambda_0 \sigma_\varepsilon^2 \beta_0^2 \left(2\mathrm{tr}(W_n^2) + \lambda_0 \mathrm{tr}(W_n^3)\right) + \lambda_0 \sigma^2 \left(2\mathrm{tr}(W_n^2) + \lambda_0 \mathrm{tr}(W_n^3)\right) \tag{B.5}$$

The symmetry of W_n implies $\mathrm{tr}(W_n^k) = \sum_{i=1}^n e_i^k$, where e_i is an eigenvalue of W_n. Since all eigenvalues are within the unit circle, $\max_{k \,\epsilon\, N} \mathrm{tr}(W_n^k) = \mathrm{tr}(W_n^2)$ (figure B.1, left). Even though one might assume that $\mathcal{M}_0$ dominates and that by $2\mathrm{tr}(W_n^2) + \lambda_0 \mathrm{tr}(W^3) > 0$ the sign is determined by λ_0, figure B.1 (right) confirms that indeed the sign of λ_0 determines the sign of Moran's I.

C Cigarette Demand: Further Specifications

Table C.1: Empirical illustration – Alternative specifications; dependent variable: log consumption

	Team Spatial	Team Dynamic			*full model*
	(1) LSDV	(2) No Space	(3) DSLX	(4) DSEM	(5)
Panel 1: estimated parameters					
time lag		0.858**	0.861***	0.865***	0.867***
		(57.33)	(58.42)	(65.04)	(24.63)
spatial lag	0.321***			0.049***	
	(11.06)			(2.88)	
price	-0.938***	-0.270***	-0.278***	-0.266***	-0.264***
	(-24.02)	(-11.72)	(-12.20)	(-13.19)	(-4.95)
income	0.348***	0.092***	0.109**	0.100***	0.090***
	(17.85)	(3.55)	(3.15)	(4.16)	(3.73)
W · price	0.476***		0.083**	0.170**	0.159**
	(10.33)		(2.34)	(3.66)	(2.17)
W · income			-0.020	-0.022	
			(-0.49)	(-0.87)	
spatial time lag				-0.015	
				(-0.29)	
Panel 2: marginal effect of log price					
short-run direct		-0.270***	-0.278***	-0.262***	-0.262***
		(-11.72)	(-12.20)	(-11.48)	(-12.09)
short-run indirect			0.083**	0.160***	0.152***
			(2.34)	(3.49)	(12.81)
long-run direct	-0.965***	-1.900***	-1.992***	-1.931***	-1.942***
	(-22.91)	(9.35)	(9.64)	(-9.59)	(-11.36)
long-run indirect	-0.420***		0.594**	0.610	0.685
	(-7.22)		(2.32)	(0.98)	(1.62)
time-fixed effects	No	Yes	Yes	Yes	Yes

Observations: 1334 (N=46, T=29). Robust t-statistics in parentheses in both panels. Test statistics in panel 'marginal effects' are all of Wald-type for matters of comparability. All specifications except column (1) include year-fixed effects. Column (1) estimates a least squares dummy variable model, outputs in columns (2) and (3) based on dynamic panel bias correction advocated by Kiviet (1995). Column (4) would ideally be estimated in the line of Su, Yang (2015), which is unfortunately not implemented in any (known to the author) statistics program. Results in column(4) replicated from Elhorst (2014) (p.114). Columns (4) and (5) use the estimator proposed by Yu et al. (2008). All variables in logs. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

D Spatial dynamic covariate

Let X_{nt} be generated as

$$X_{nt} = \psi_0 X_{n,t-1} + \alpha_0 W_n X_{nt} + \varepsilon_{nt} = (I_n - \alpha_0 W_n)^{-1} (\psi_0 X_{n,t-1} + \varepsilon_{nt}), \tag{D.1}$$

where ε_{nt} are independently and identically distributed random draws, α_0 represents spatial dependence, and parameter ψ_0 represents autocorrelation. Because all derivations above are conditional on within-transformed data, it is reasonable to assume a DGP like this for the transformed covariate. The assumption that there is a common spatial weights matrix W_n that reflects spatial linkages is most frequently used, in applied as well as theoretical work. In this section, we stick to this tradition. In principle one would be able to define different weight matrices for the spatial lag of the dependent variable and the DGP of the covariate. Besides having to bear additional notational burden of different weight matrices, there is very good reason to rely on one representation of space. As LeSage, Pace (2014) argue, spatial weights matrices are likely highly positively correlated.

Similar the equation (B.1), one can characterize the DGP of the covariate equally as

$$\begin{aligned} X_{nt} &= \psi_0^t A_{x,n}^t X_{n,0} + \sum_{\tau=0}^{t} \psi_0^\tau A_{x,n}^{\tau+1} \varepsilon_{n,t-\tau} \\ A_{x,n} &= (I_n - \alpha_0 W_n)^{-1} \end{aligned} \tag{D.2}$$

implying that the state of the covariate depends on its initial or start values, the start values of all other units, and all past residual terms, own and of all other units. For $\psi_0 < 1$, the first term will disappear. This gives rise to the same result as in Appendix B, stating that, indeed, α_0 governs the sign of spatial correlation of X.

Two additional parameters and the reduced form DGP of X greatly increase the difficulty in deriving approximations of the bias of the teams. For example, the approximated bias of Team Dynamic, as given in equation (A.4), in a set up with one covariate is given by

$$\text{plim}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0) \approx \lambda_0 \mathbf{Q} \boldsymbol{\phi}_0 = \lambda_0 \begin{bmatrix} \gamma_0 \left(\eta_{\tilde{Y}_-} \sigma^2_{\tilde{Y}_-} m_{\tilde{Y}_-} \right) + \beta_0 \left(\eta_{\tilde{X},\tilde{Y}_-} \sigma_{\tilde{X},\tilde{Y}_-} m_{\tilde{X},\tilde{Y}_-} \right) \\ \gamma_0 \left(\eta_{\tilde{X},\tilde{Y}_-} \sigma_{\tilde{X},\tilde{Y}_-} m_{\tilde{X},\tilde{Y}_-} \right) + \beta_0 \left(\eta_{\tilde{X}} \sigma^2_{\tilde{X}} m_{\tilde{X}} \right) \end{bmatrix} \tag{D.3}$$

Whereas term $m_{\tilde{X},\tilde{Y}_-}$ can be considered nil in the case of random assignment, this is not the case in this setting, as all random draws $\varepsilon_{n,0}$ to $\varepsilon_{n,t-1}$ affect both $\tilde{X}$ as well as $\tilde{Y}_-$. This implies further that $\tilde{X}_{nt}$ and $\tilde{Y}_{n,t-1}$ are in fact collinear, rendering $\eta_{\tilde{X},\tilde{Y}_-}$ and $\sigma_{\tilde{X},\tilde{Y}_-}$ nonzero. Hence, it appears quite unpromising to pursue further approximations, as potentially all parameters $(\gamma_0, \lambda_0, \alpha_0, \psi_0, \beta_0)$ may play a direct or indirect role in determining the bias in either estimate of Team Dynamic. The same argument can be applied to the approximated bias of Team Spatial.

In cases where $\psi_0 = 0$, equations (1) and (D.1) imply that the off-diagonal values of $\mathbf{M}$ $(m_{i,\neq i})$ tend to zero as long as innovations ε are not autocorrelated. Formally, the condition is given by

$$\text{plim} \left(\left[(I - \lambda_0 W)^{-1} (I - \alpha_0 W)^{-1} \varepsilon_{n,t-1} \right]' (I - \alpha_0 W)^{-1} \varepsilon_{nt} \right) = 0, \tag{D.4}$$

which holds by the *i.i.d.* assumption.

Simulation results of several possible cases are presented in the tables below, with results as discussed in the following two paragraphs.

Autoregressive Parameters The main conclusions remain for both teams. Team Dynamic will overestimate γ_0, and more so the larger spatial dependence (λ_0) in absolute value. Comparing positive and negative values of spatial dependence, the bias is smaller at the latter. For lower values of γ_0, autocorrelation in the covariate $(\psi_0 > 0)$ appears to

aggravate this bias, and also spatial depedence in the covariate (α_0) has a positive impact on the bias. This is shown in in Table D.2.

Regarding Team Spatial, Table D.5 reveals that the bias in estimating λ_0 is, as above, increasing in γ_0 in absolute terms conditional on α_0 and ψ_0. For a given parameterization of the DGP of Y_{nt}, interesting insights appear, however. First, the bias in increasing in ψ_0. Thus, it is possible that the sign of the bias may even change sign, even though this possibility seems confined to cases where the DGP of the dependent variable itself is highly autocorrelated. Similar to Team Dynamic, spatial dependence in the covariate seems to positively affect this bias. It has to be stressed, however, that the main effect is determined by λ_0 and γ_0 for both teams.

Marginal effect of the covariate As in the main text, Team Dynamic is able to measure direct long- and short-run effects *without* feedback, while Team Spatial interprets marginal effects as equilibrium effects without considering the time dimension. Because team *dynamic* misses feedback effects, it could be expected that the short-run marginal effect is more prone to underestimation the larger the potential for such feedbacks, indicated by large values of λ_0. This is indeed the case, and seems to be more pronounced the larger the autocorrelation (ψ_0) in the covariate. However, results show a slight overestimation for $\psi_0 = 0$ at high values of autocorrelation and positive spatial dependence in the dependent variable, as shown in Table D.3. The effect of α_0 seems limited, but suggest a shift of the bias, where the direction is given by the sign of λ_0

As in the main text, the balance between missing feedbacks and an overestimation of the auotregressive parameter plays a crucial role in the long-run. Table D.4 presents the result for varying parameter constellations in the DGP of the covariate, and shows that this interplay is influenced significantly. For example, while the long-run direct effect is slightly underestimated for $\gamma_0 = 0.4$, $\lambda_0 = 0.6$, $\alpha_0 = -0.3$, it turns out to be *over*estimated when α_0 is nil or positive. Autocorrelation in the covariate aggrevates the bias in either case.

Team spatial is most likely to underestimate equilibrium effects of the covariate due to the omission of the time dimension, and this finds strong support in Table D.6 and Table D.7. In the case of direct marginal effects in Table D.6, one can conclude that spatial correlation in the covariate has no visible effect on the bias, while autocorrelation $\psi_0 > 0$ of the covariate tends to dampen this bias mildly. Regarding indirect effects, as presented in Table D.7, this result remains unchanged with the exception that the sign of α_0 determines the sign of the bias when the DGP of the dependent variable in fact features no spatial dependence.

Table D.2: Bias of Team Dynamic – autoregressive parameter.

ψ_0 →		0			0.3			0.6		
α_0 →		-0.3	0	0.3	-0.3	0	0.3	-0.3	0	0.3
γ_0	λ_0									
0.00	-0.60	0.000	0.001	0.000	0.006	0.002	0.000	0.013	0.007	0.004
0.00	-0.30	0.001	-0.002	0.001	0.003	0.001	0.001	0.003	0.003	-0.001
0.00	0.00	0.000	0.000	-0.001	0.000	0.001	0.000	0.000	0.000	0.000
0.00	0.30	0.000	0.000	-0.001	-0.001	0.002	0.001	0.002	0.002	0.006
0.00	0.60	0.002	0.002	0.000	0.001	0.004	0.009	0.001	0.008	0.025
0.40	-0.60	0.021	0.016	0.011	0.027	0.020	0.016	0.049	0.028	0.018
0.40	-0.30	0.006	0.001	0.000	0.008	0.004	0.002	0.013	0.005	0.000
0.40	0.00	0.000	0.001	-0.001	0.000	0.001	0.000	0.001	-0.002	0.000
0.40	0.30	-0.001	0.003	0.004	0.001	0.006	0.007	0.002	0.010	0.017
0.40	0.60	0.020	0.027	0.035	0.021	0.032	0.051	0.029	0.049	0.098
0.80	-0.30	0.031	0.032	0.031	0.033	0.029	0.027	0.045	0.027	0.019
0.80	0.00	0.002	0.001	0.002	0.001	0.002	0.001	0.001	0.001	0.001
0.80	0.30	0.140	0.144	0.141	0.132	0.121	0.131	0.119	0.113	0.117

Note: Bias of the estimated autoregressive parameter γ_{TD} of Team Dynamic relative to the corresponding parameter in the true model. λ_0 and γ_0 represent spatial and time dependence in the DPG of the dependent variable Y, α_0 and ψ_0 represent spatial and time dependence in the DPG of the covariate X. Bias in absolute terms whenever $\gamma_0 = 0$.

Table D.3: Bias of Team Dynamic – direct short-run marginal effect of covariate.

ψ_0 →		0			0.3			0.6		
α_0 →		-0.3	0	0.3	-0.3	0	0.3	-0.3	0	0.3
γ_0	λ_0									
0.00	-0.60	0.010	0.000	-0.010	0.011	0.000	-0.012	0.009	-0.002	-0.020
0.00	-0.30	0.005	0.000	-0.003	0.007	0.001	-0.008	0.003	0.000	-0.007
0.00	0.00	0.000	0.001	0.003	-0.001	0.000	0.000	0.000	0.000	0.001
0.00	0.30	-0.003	0.000	0.005	-0.004	-0.001	0.008	-0.008	-0.002	0.005
0.00	0.60	-0.013	-0.001	0.018	-0.014	-0.002	0.016	-0.015	-0.005	0.009
0.40	-0.60	0.012	-0.001	-0.010	0.007	-0.003	-0.018	-0.005	-0.013	-0.028
0.40	-0.30	0.005	0.000	-0.004	0.003	0.000	-0.006	0.004	-0.001	-0.009
0.40	0.00	-0.001	0.001	-0.001	-0.002	-0.002	0.001	0.000	0.000	0.000
0.40	0.30	-0.004	-0.001	0.006	-0.006	-0.001	0.005	-0.011	-0.005	0.006
0.40	0.60	-0.011	0.001	0.016	-0.018	-0.005	0.008	-0.032	-0.024	-0.022
0.80	-0.30	0.007	0.006	0.001	0.005	-0.001	-0.009	-0.014	-0.014	-0.027
0.80	0.00	0.002	0.000	0.002	0.001	-0.001	0.002	-0.002	0.001	0.001
0.80	0.30	0.020	0.022	0.028	-0.019	-0.014	-0.007	-0.082	-0.068	-0.057

Note: Bias of the estimated direct short-run effect of Team Dynamic relative to the one in the true model. λ_0 and γ_0 represent spatial and time dependence in the DPG of the dependent variable Y, α_0 and ψ_0 represent spatial and time dependence in the DPG of the covariate X.

Table D.4: Bias of Team Dynamic – Long-run marginal effect of covariate.

ψ_0 →		0			0.3			0.6		
α_0 →		-0.3	0	0.3	-0.3	0	0.3	-0.3	0	0.3
γ_0	λ_0									
0.00	-0.60	0.011	0.002	-0.009	0.017	0.002	-0.012	0.022	0.005	-0.016
0.00	-0.30	0.006	-0.001	-0.002	0.010	0.002	-0.007	0.006	0.003	-0.007
0.00	0.00	0.000	0.001	0.002	0.000	0.001	0.000	-0.001	0.001	0.001
0.00	0.30	-0.003	0.000	0.003	-0.005	0.001	0.010	-0.007	0.000	0.011
0.00	0.60	-0.011	0.001	0.018	-0.013	0.003	0.025	-0.014	0.003	0.036
0.40	-0.60	0.028	0.006	-0.012	0.033	0.010	-0.012	0.062	0.014	-0.019
0.40	-0.30	0.011	-0.001	-0.007	0.013	0.003	-0.006	0.022	0.004	-0.013
0.40	0.00	0.000	0.002	-0.002	-0.001	0.000	0.001	0.002	-0.002	0.001
0.40	0.30	-0.009	0.000	0.009	-0.010	0.004	0.012	-0.012	0.008	0.031
0.40	0.60	-0.011	0.013	0.044	-0.017	0.014	0.065	-0.017	0.029	0.138
0.80	-0.30	0.099	0.106	0.101	0.100	0.069	0.043	0.170	0.024	-0.035
0.80	0.00	0.017	0.008	0.017	0.006	0.010	0.011	0.005	0.009	0.005
0.80	0.30	0.691	0.655	1.506	0.008	-0.107	0.054	-0.507	-0.511	-0.460

Note: Bias of the estimated direct long-run effect of Team Dynamic relative to the one in the true model. λ_0 and γ_0 represent spatial and time dependence in the DPG of the dependent variable Y, α_0 and ψ_0 represent spatial and time dependence in the DPG of the covariate X.

Table D.5: Bias of Team Spatial – spatial-autoregressive parameter.

ψ_0 →		0			0.3			0.6		
α_0 →		-0.3	0	0.3	-0.3	0	0.3	-0.3	0	0.3
γ_0	λ_0									
0.00	-0.60	0.018	-0.003	0.006	-0.006	0.007	0.013	0.005	0.003	0.003
0.00	-0.30	0.003	-0.004	-0.001	-0.007	0.000	0.007	-0.005	-0.001	-0.007
0.00	0.00	-0.006	0.008	-0.008	-0.006	-0.007	0.004	0.002	-0.003	-0.005
0.00	0.30	-0.003	-0.008	-0.006	-0.011	0.000	-0.006	-0.003	-0.001	-0.003
0.00	0.60	-0.010	-0.005	-0.008	-0.006	-0.003	-0.008	-0.015	-0.006	-0.011
0.40	-0.60	-0.194	-0.153	-0.094	-0.267	-0.202	-0.133	-0.363	-0.260	-0.149
0.40	-0.30	-0.114	-0.062	-0.026	-0.173	-0.102	-0.007	-0.251	-0.122	-0.012
0.40	0.00	-0.047	-0.006	0.051	-0.088	-0.003	0.079	-0.129	-0.010	0.113
0.40	0.30	0.015	0.069	0.116	0.021	0.079	0.167	0.002	0.121	0.247
0.40	0.60	0.119	0.167	0.203	0.147	0.209	0.271	0.182	0.261	0.367
0.80	-0.30	-0.737	-0.708	-0.648	-0.786	-0.653	-0.583	-0.928	-0.670	-0.384
0.80	0.00	-0.199	-0.063	0.140	-0.259	-0.013	0.241	-0.399	-0.015	0.405
0.80	0.30	0.981	1.004	1.006	0.956	0.922	0.985	0.914	0.931	0.964

Note: Bias of the estimated spatial-autoregressive parameter λ_{TS} of Team Dynamic relative to the corresponding parameter in the true model. λ_0 and γ_0 represent spatial and time dependence in the DPG of the dependent variable Y, α_0 and ψ_0 represent spatial and time dependence in the DPG of the covariate X. Bias in absolute terms whenever $\lambda_0 = 0$.

Table D.6: Bias of Team Spatial – marginal direct long-run effect.

ψ_0 →		0			0.3			0.6		
α_0 →		-0.3	0	0.3	-0.3	0	0.3	-0.3	0	0.3
γ_0	λ_0									
0.00	-0.60	0.000	0.000	0.000	0.000	0.000	-0.001	0.001	0.002	-0.001
0.00	-0.30	0.000	0.001	0.002	0.002	0.000	-0.002	-0.002	0.001	0.001
0.00	0.00	0.000	0.000	0.002	-0.001	0.000	0.000	-0.001	0.000	0.000
0.00	0.30	0.002	0.000	-0.001	0.001	-0.001	0.002	-0.001	-0.001	0.000
0.00	0.60	-0.003	-0.003	0.003	-0.003	-0.001	0.000	0.000	-0.001	-0.002
0.40	-0.60	-0.426	-0.425	-0.424	-0.351	-0.350	-0.349	-0.249	-0.250	-0.251
0.40	-0.30	-0.416	-0.416	-0.415	-0.343	-0.341	-0.339	-0.240	-0.240	-0.244
0.40	0.00	-0.415	-0.414	-0.414	-0.338	-0.340	-0.340	-0.240	-0.242	-0.241
0.40	0.30	-0.416	-0.415	-0.417	-0.341	-0.340	-0.341	-0.244	-0.242	-0.243
0.40	0.60	-0.431	-0.430	-0.431	-0.356	-0.355	-0.355	-0.258	-0.258	-0.254
0.80	-0.30	-0.842	-0.841	-0.839	-0.801	-0.800	-0.796	-0.720	-0.721	-0.720
0.80	0.00	-0.824	-0.823	-0.823	-0.778	-0.777	-0.778	-0.692	-0.689	-0.689
0.80	0.30	-0.975	-0.975	-0.975	-0.969	-0.970	-0.970	-0.957	-0.958	-0.959

Note: Bias of the estimated direct long-run effect of Team Spatial relative to the one in the true model. λ_0 and γ_0 represent spatial and time dependence in the DPG of the dependent variable Y, α_0 and ψ_0 represent spatial and time dependence in the DPG of the covariate X.

Table D.7: Bias of Team Spatial – marginal indirect long-run effect.

ψ_0 →		0			0.3			0.6		
α_0 →		-0.3	0	0.3	-0.3	0	0.3	-0.3	0	0.3
γ_0	λ_0									
0.00	-0.60	-0.027	0.005	-0.009	0.009	-0.011	-0.021	-0.006	-0.002	-0.005
0.00	-0.30	-0.011	0.012	0.004	0.024	0.000	-0.022	0.012	0.006	0.022
0.00	0.00	-0.001	0.002	-0.002	-0.001	-0.001	0.001	0.000	-0.001	-0.001
0.00	0.30	-0.008	-0.029	-0.024	-0.040	0.000	-0.021	-0.012	-0.004	-0.012
0.00	0.60	-0.025	-0.014	-0.014	-0.017	-0.008	-0.017	-0.033	-0.013	-0.026
0.40	-0.60	-0.534	-0.556	-0.585	-0.430	-0.468	-0.509	-0.274	-0.346	-0.423
0.40	-0.30	-0.507	-0.565	-0.604	-0.375	-0.459	-0.577	-0.172	-0.348	-0.508
0.40	0.00	-0.009	-0.001	0.010	-0.019	-0.001	0.018	-0.032	-0.003	0.030
0.40	0.30	-0.661	-0.592	-0.532	-0.609	-0.525	-0.394	-0.581	-0.384	-0.153
0.40	0.60	-0.672	-0.639	-0.614	-0.608	-0.557	-0.500	-0.517	-0.436	-0.302
0.80	-0.30	-0.896	-0.898	-0.903	-0.864	-0.879	-0.885	-0.783	-0.828	-0.874
0.80	0.00	-0.034	-0.011	0.026	-0.054	-0.003	0.059	-0.113	-0.005	0.147
0.80	0.30	-0.995	-0.994	-0.994	-0.994	-0.995	-0.993	-0.992	-0.992	-0.992

Note: Bias of the estimated indirect long-run effect of Team Spatial relative to the one in the true model. λ_0 and γ_0 represent spatial and time dependence in the DPG of the dependent variable Y, α_0 and ψ_0 represent spatial and time dependence in the DPG of the covariate X.

E Additional graphs and simulation results

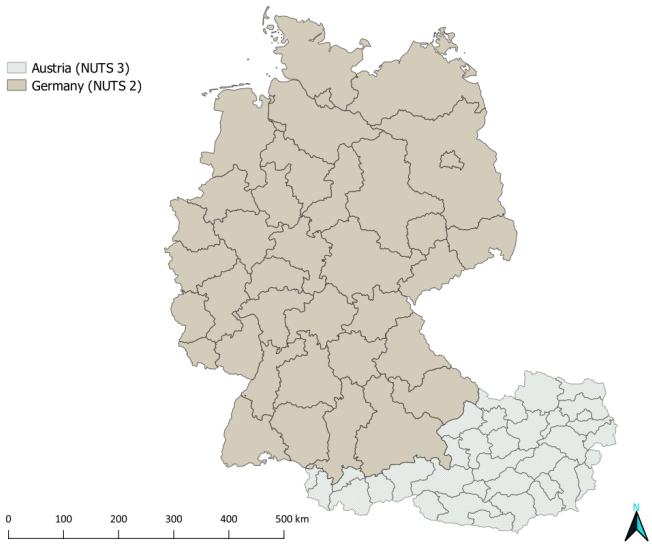

Figure E.1: NUTS regions used in the simulation.

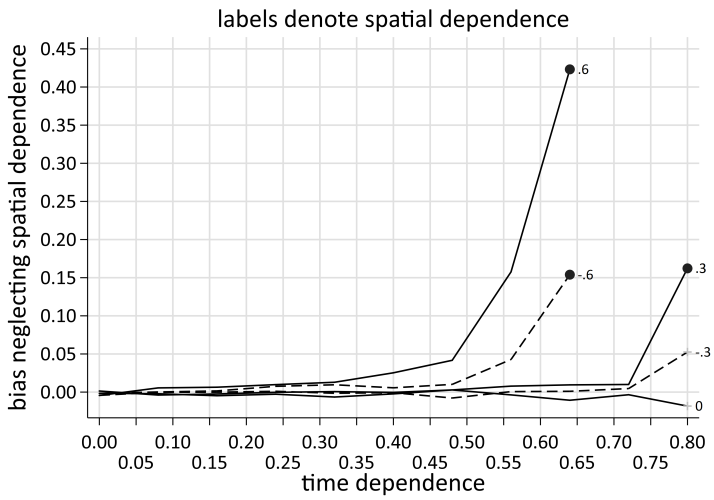

Figure E.2: Bias of Team Dynamic neglecting the spatial dimension, but applying a spatial lag of X specification. The vertical axis describes $\gamma_{TD} - \gamma_0$, the horizontal axis depicts γ_0. Labels denote spatial dependence in the DGP (λ_0).

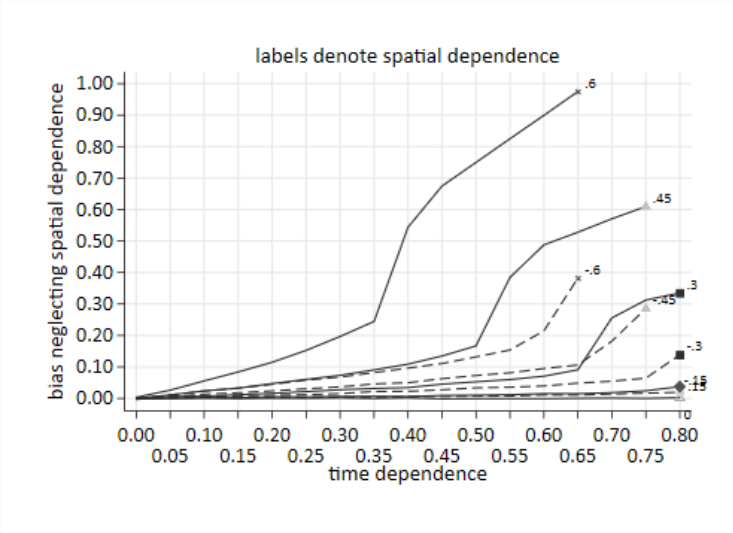

Figure E.3: Bias of Team Dynamic neglecting the spatial dimension. The vertical axis describes $\gamma_{TD} - \gamma_0$, the horizontal axis depicts γ_0. Labels denote spatial dependence in the DGP (λ_0). Based on German NUTS 2 regions.

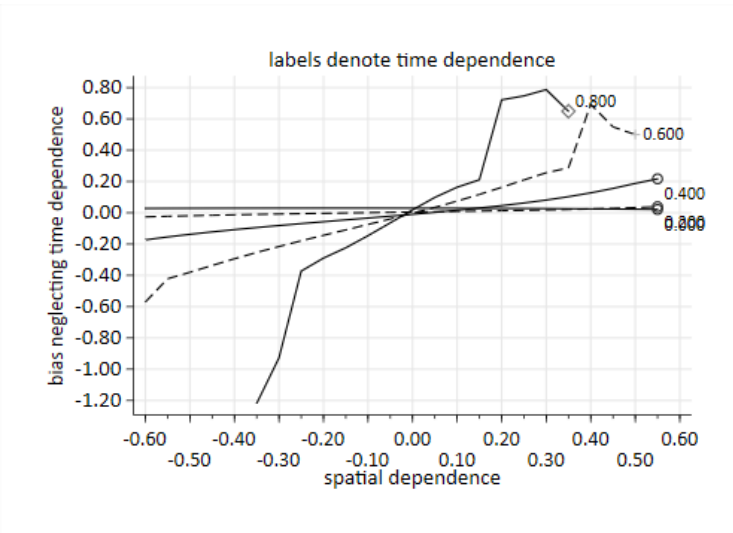

Figure E.4: Bias of Team Spatial neglecting the time dimension. The vertical axis describes $\lambda_{TS} - \lambda_0$, the horizontal axis depicts λ_0. Labels denote spatial dependence in the DGP (γ_0). Based on German NUTS 2 regions.

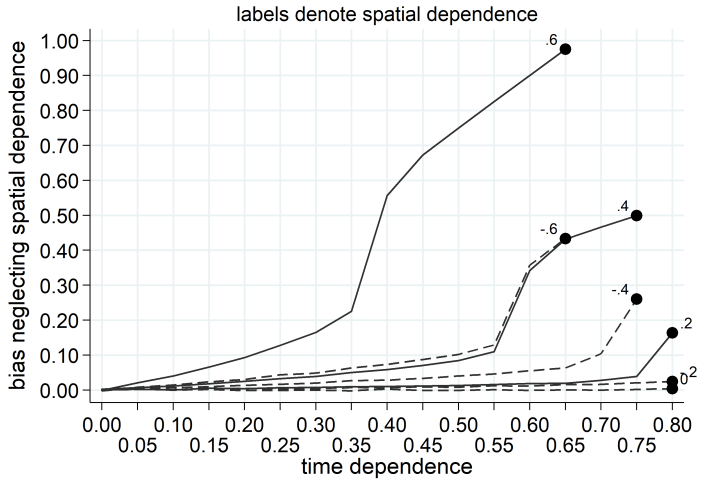

Figure E.5: Bias of Team Dynamic neglecting the spatial dimension. The vertical axis describes $\gamma_{TD} - \gamma_0$, the horizontal axis depicts γ_0. Labels denote spatial dependence in the DGP (λ_0). Based on row-normalized spatial weights matrix.

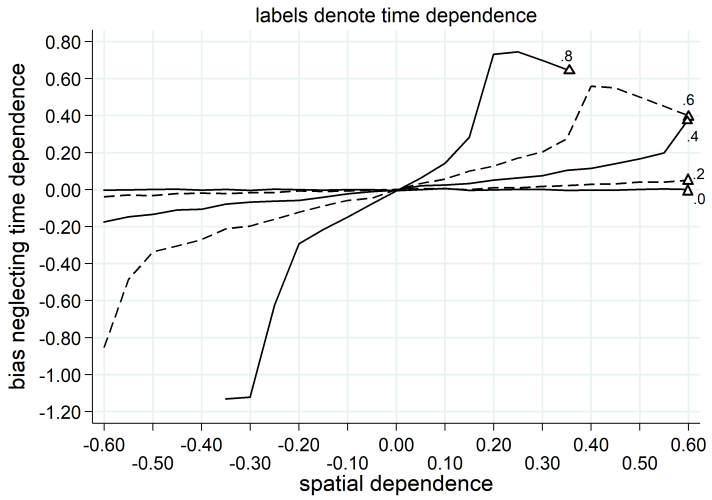

Figure E.6: Bias of Team Spatial neglecting the time dimension. The vertical axis describes $\lambda_{TS} - \lambda_0$, the horizontal axis depicts λ_0. Labels denote spatial dependence in the DGP (γ_0). Based on row-normalized spatial weights matrix.