PS
RM

ORIGINAL ARTICLE

Spatial interdependence and instrumental variable models

Timm Betz, Scott J. Cook and Florian M. Hollenbach[*]

Department of Political Science, Texas A&M University, College Station, TX 77843, USA
*Corresponding author. Email: fhollenbach@tamu.edu

Abstract
Instrumental variable (IV) methods are widely used to address endogeneity concerns. Yet, a specific kind of endogeneity – spatial interdependence – is regularly ignored. We show that ignoring spatial interdependence in the outcome results in asymptotically biased estimates even when instruments are randomly assigned. The extent of this bias increases when the instrument is also spatially clustered, as is the case for many widely used instruments: rainfall, natural disasters, economic shocks, and regionally- or globally-weighted averages. Because the biases due to spatial interdependence and predictor endogeneity can offset, addressing only one can increase the bias relative to ordinary least squares. We demonstrate the extent of these biases both analytically and via Monte Carlo simulation. Finally, we discuss a general estimation strategy – S-2SLS – that accounts for both outcome interdependence and predictor endogeneity, thereby recovering consistent estimates of predictor effects.

As political scientists increasingly focus on the identification of causal effects, instrumental variable (IV) models are becoming commonplace (e.g., Sovey and Green, 2011). IV models hinge on the validity of the instrument. While researchers are usually aware of conditional independence and relevance as general requirements for valid instruments, we identify a specific threat that is frequently ignored: spatial interdependence in the outcome variable. Our review of IV models in leading political science journals reveals that authors rarely discuss and never empirically address spatial interdependence as a threat to inference (see Figure 1), even as theories of spatial interdependence and diffusion proliferate across political science (see, e.g., Siverson and Starr 1990; Starr 1991; Ward and Gleditsch 2002; Ward and O'Loughlin 2002; Simmons, Dobbin and Garrett 2006; Franzese and Hays 2007; Plümper and Neumayer 2010).[1]

This is not a trivial oversight. We show that failing to model outcome interdependence produces estimates that are asymptotically biased, even when the instrument is randomly assigned. When, in addition, the instrument exhibits spatial dependence similar to that of the outcome, the bias in IV estimates increases and can even surpass that of ordinary least squares. This concern applies to many popular instruments, including geographic, meteorologic, and economic variables (see, e.g., Hansford and Gomez 2010; Ramsay 2011; Ahmed 2012), as well as any instrument measured at a higher level of aggregation than the outcome, such as regional or global economic, political, and institutional shocks (see, e.g., Stasavage 2005; Büthe and Milner

---

[1]We analyzed each article on the basis of whether prior theories of spatial interdependence or diffusion had been established for and could reasonably apply to the outcome of interest. The articles using IV models that are not at risk of the issues we discuss here include pure time-series analyses, survey experiments, most field experiments, etc.
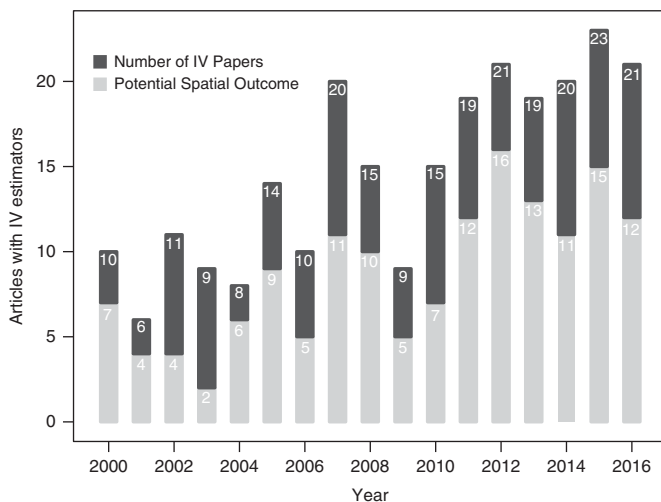
Figure 1. The plot shows the number of articles published in the APSR, AJPS, JOP, IO, BJPS, and World Politics between 2000 and 2016 that use IV models (light gray bars), and the number of those articles at risk of spatial interdependence in the outcome (dark gray bars).

2008; Boix 2011; Ramsay 2011). Because these instruments are not randomly distributed across space, they risk increased bias even when they are otherwise plausibly exogenous.

Our results connect more general findings in the otherwise distinct literatures on spatial interdependence and IVs. Ignored spatial interdependence constitutes an omitted variables problem (e.g., Franzese and Hays 2007). While IV models are commonly thought to be immune to omitted variable bias, and indeed frequently used to overcome it (Wooldridge, 2002), this intuition does not always hold. Instead, as we show below, IV models can augment omitted variables bias from unmodelled spatial interdependence.

Because of its reciprocal relationship with the outcome, ignored spatial interdependence also ensures any instrument violates the exclusion restriction. As is well known, even mild violations of the exclusion restriction can produce substantial bias (Bartels, 1991; Bound, Jaeger and Baker, 1995). When these violations are caused by spatial interdependence, however, solutions are available to recover asymptotically unbiased estimates if one is willing to make assumptions about the nature of spatial relationships in the outcome variable. Recent work in the spatial econometrics literature has generalized spatial models to allow for endogenous predictors (e.g., Kelejian and Prucha 2004; Anselin and Lozano-Gracia 2008; Fingleton and Le Gallo 2008; Drukker, Egger and Prucha 2013; Liu and Lee 2013). These same methods – hereafter spatial-two stage least squares (S-2SLS) – are useful when addressing endogenous predictors even when researchers are otherwise uninterested in spatial dependence theoretically.[2] In short, with S-2SLS researchers instrument for both the endogenous predictor and the spatial-lag of the outcome, thereby obtaining consistent estimates of the desired causal effect.

In addition to accounting for possible outcome interdependence, this approach has two attractive features. First, it nests the standard spatial-autoregressive (SAR) model and the standard IV model, allowing researchers to explicitly test restrictions rather than proceed by assumption.[3] Second, because it is an IVs approach, it should be straightforward to understand

---

[2]To clarify, Franzese and Hays (2007) and others have previously used S-2SLS to indicate a spatial autoregressive (SAR) model estimated via 2SLS. Here, we use this term more broadly to include instances where at least one of the non-spatial predictors is also endogenous.

[3]We focus on the spatial-autoregressive model (SAR) for two reasons. First, it is the most widely used spatial model in political science. Second, it is the interdependence in the outcome – as in the SAR – that induces the simultaneity that is at the

and implement for those already pursuing IV strategies. Our simulations demonstrate that this approach consistently outperforms estimation strategies that neglect interdependence – even under conditions unfavorable to spatial models.

We therefore advocate that researchers consider S-2SLS as a general, conservative strategy when confronting endogenous predictors and existing theories suggest the possibility of interdependence in the outcome variable. In the conclusion, we discuss some of the implications for the use of IV models in applied research.

OLS and multifarious endogeneity

In order to better understand the problems that arise from neglecting spatial interdependence in IV estimation, it is useful to first clarify that unmodelled interdependence is itself an omitted variables problem. Consider a simple linear-additive model

$$\mathbf{y} = \beta\mathbf{x} + \mathbf{e}, \tag{1}$$

where $\mathbf{y}$ is an n-length vector of outcomes, $\mathbf{x}$ the predictor, and $\mathbf{e}$ the disturbance. The OLS estimator of β is the sample covariance of $\mathbf{x}$ and $\mathbf{y}$ over the sample variance of $\mathbf{x}$,

$$\hat{\beta}_{\mathrm{OLS}} = \frac{\widehat{\mathrm{cov}}(\mathbf{x}, \mathbf{y})}{\widehat{\mathrm{var}}(\mathbf{x})}. \tag{2}$$

Substituting the right-hand side of equation (1) in for $\mathbf{y}$ yields the probability limit

$$\mathrm{plim}_{n\to\infty}\, \hat{\beta}_{\mathrm{OLS}} = \beta + \underbrace{\frac{\mathrm{cov}(\mathbf{x}, \mathbf{e})}{\mathrm{var}(\mathbf{x})}}_{\text{endogeneity bias}}, \tag{3}$$

showing that $\hat{\beta}_{\mathrm{OLS}}$ is asymptotically unbiased if $\mathrm{cov}(\mathbf{x}, \mathbf{e}) = 0$, that is, if $\mathbf{x}$ is exogenous.[4] This result should be familiar to readers. It is presented in any introductory econometrics textbook along with common sources of bias: confounding due to omitted variables, simultaneity, and measurement error in the variable of interest.

We are concerned with a special case of confounding: unmodelled interdependence between outcomes. Spatial, or cross-sectional, interdependence occurs when a unit's outcome affects the choices, actions, or decisions of other units (Kirby and Ward 1987; Ward and O'Loughlin 2002; Beck, Gleditsch and Beardsley 2006; Franzese and Hays 2007; Plümper and Neumayer 2010). Theories of interdependence are "ubiquitous, and often quite central, throughout the substance of political science" (Franzese and Hays 2007, p. 141): the contagion of conflict and crises, the spread of domestic institutions and ideologies, economic integration and resulting policy coordination, and participation in international agreements all provide examples. Ignoring this spatial interdependence induces cross-sectional correlation in the residuals and, more problematically, covariance between the predictors and the disturbances. As a consequence, coefficient estimates are both inefficient and biased; in the following, we focus on the latter concern.

To distinguish confounding due to spatial interdependence from other sources of endogeneity of $\mathbf{x}$, we decompose the error term in equation (1) as

$$\mathbf{e} = \rho\mathbf{W}\mathbf{y} + \mathbf{u}, \tag{4}$$

where ρ is the effect of outcomes $\mathbf{y}$ in surrounding units j on unit i, weighted by $\mathbf{W}$, an n-by-n connectivity matrix which identifies the relationship between units i and j. As usual in spatial

---

heart of the problem we discuss. Other models that contain a spatial lag of the outcome and additional features, such as autoregressive disturbances (the SAR-AR model), are extensions of the SAR model and could also be estimated. Drukker, Egger and Prucha (2013) discuss the estimation of a SAR-AR model with an endogenous predictor, which can be estimated using the same software routines we discuss below.

[4]When we discuss bias, we refer to asymptotic bias. All IV estimators have small-sample bias.

econometrics, we refer to $\mathbf{Wy}$ as the spatial lag, with $\mathbf{W}$ determining which other-unit outcomes y_j are likely to influence the choices, actions, behaviors of unit i.

Then, we can rewrite equation (3) as

$$\text{plim}_{n\to\infty}\,\hat{\beta}_{\text{OLS}}-\beta \;=\; \underbrace{\left[\frac{\text{cov}(\mathbf{x},\mathbf{u})}{\text{var}(\mathbf{x})}\right]}_{\text{Non-spatial endogeneity bias}} \;+\; \rho\underbrace{\left[\frac{\text{cov}(\mathbf{x},\mathbf{Wy})}{\text{var}(\mathbf{x})}\right]}_{\text{Spatial endogeneity bias}}. \tag{5}$$

Equation (5) separately identifies spatial and non-spatial endogeneity as two potential sources of bias in the OLS estimator.[5] First, bias can result from more familiar, non-spatial sources of endogeneity of $\mathbf{x}$, that is, correlation between $\mathbf{x}$ and $\mathbf{u}$. This is represented by the first term in equation (5), which drops out if $\text{cov}(\mathbf{x},\mathbf{u})$ is zero. Second, bias can arise from spatial interdependence in $\mathbf{y}$. As indicated by the second term on the right-hand side of equation (5), this bias drops out if $\rho = 0$; that is, when there is no spatial interdependence.[6] In what follows, we show that addressing the former while neglecting the latter fails to recover unbiased estimates of the effect. In many cases, it magnifies the bias relative to ordinary least squares.

Spatial bias in IV models

Following Sovey and Green (2011), we introduce IV estimation using familiar notation from structural equation models, assuming linear–additive relationships between the variables. Suppose a suitable instrument $\mathbf{z}$ is available, resulting in the following system of equations:

$$\mathbf{y} = \beta\mathbf{x}+\mathbf{e}, \tag{6}$$

$$\mathbf{x} = \gamma\mathbf{z}+\mathbf{v}. \tag{7}$$

As before, suppose that the disturbance can be decomposed as $\mathbf{e} = \rho\mathbf{Wy} + \mathbf{u}$ and interdependence is ignored in the estimation. Then, non-spatial endogeneity arises if $\text{cov}(\mathbf{u},\mathbf{v}) \neq 0$ and therefore $\text{cov}(\mathbf{x},\mathbf{u}) \neq 0$. We assume in the following that the variable $\mathbf{z}$ satisfies the usual assumptions for a valid instrument – $\text{cov}(\mathbf{z},\mathbf{x}) \neq 0$ and $\text{cov}(\mathbf{z},\mathbf{u}) = 0$ – such that $\mathbf{z}$ is correlated with the endogenous predictor $\mathbf{x}$ but uncorrelated with the disturbance $\mathbf{u}$.

The IV estimator is obtained as two-stage least squares (2SLS), such that

$$\hat{\beta}_{\text{2SLS}} = \frac{\text{cov}(\mathbf{y},\mathbf{z})}{\text{cov}(\mathbf{x},\mathbf{z})}. \tag{8}$$

Inserting the expression for $\mathbf{y}$ yields

$$\text{plim}_{n\to\infty}\,\hat{\beta}_{\text{2SLS}}-\beta = \frac{\rho\times\text{cov}(\mathbf{Wy},\mathbf{z})}{\text{cov}(\mathbf{x},\mathbf{z})} + \frac{\text{cov}(\mathbf{u},\mathbf{z})}{\text{cov}(\mathbf{x},\mathbf{z})}, \tag{9a}$$

$$= \rho\underbrace{\left[\frac{\text{cov}(\mathbf{Wy},\mathbf{z})}{\text{cov}(\mathbf{x},\mathbf{z})}\right]}_{\text{Spatial endogeneity bias}}, \tag{9b}$$

which shows that, by assumption, 2SLS does not suffer from the non-spatial endogeneity bias of OLS: because $\text{cov}(\mathbf{u},\mathbf{z}) = 0$ and $\text{cov}(\mathbf{x},\mathbf{z}) \neq 0$, the second term on the right-hand side of equation (9a)

---

[5]This derivation of the bias is only approximate, as $\mathbf{W}$ also increases in n.

[6]It is only when ρ is zero that this term drops out. $\text{cov}(\mathbf{x},\mathbf{Wy})$ is non-zero, because $\mathbf{Wy}$ is a function of $\mathbf{x}$. While the most obvious solution to address the bias from interdependence may be including $\mathbf{Wy}$ as a variable, this would not be sufficient, because $\mathbf{Wy}$ itself is endogenous in the outcome equation; see, e.g., Franzese and Hays (2007).

disappears. This result, of course, is well appreciated and motivates the use of 2SLS where $\mathbf{x}$ is suspected to be endogenous.

Less appreciated is that 2SLS is biased in the presence of (ignored and hence unmodelled) interdependence. In short, the instrument violates the exclusion restriction, because it is related to the outcome disturbances via the omitted interdependence term $\mathbf{Wy}$. To see why, note that after substituting and rearranging terms, equation (6) can be multiplied by $\mathbf{W}$ and written as

$$\mathbf{Wy} = \mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}[\beta\gamma\mathbf{z} + \beta\mathbf{v} + \mathbf{u}]. \tag{10}$$

That is, we can re-express the spatial lag, $\mathbf{Wy}$, in terms of the spatially weighted instrument $\mathbf{z}$ and stochastic terms $\mathbf{u}$ and $\mathbf{v}$. Substituting this expression into the definition of the spatial bias in 2SLS and rearranging, we obtain

$$\text{plim}_{n\to\infty}\, \hat{\beta}_{2\text{SLS}} - \beta = \beta\left[\rho\, \frac{\text{cov}(\mathbf{Wz}, \mathbf{z})}{\text{var}(\mathbf{z})}\right] + \beta\sum_{k=2}^{\infty}\left[\rho^k\, \frac{\text{cov}(\mathbf{W}^k\mathbf{z}, \mathbf{z})}{\text{var}(\mathbf{z})}\right]. \tag{11}$$

2SLS is biased unless the terms on the right-hand side are zero. For clarity in the following exposition, we have split the bias into two terms – the first representing the first-order bias and the second representing higher order terms. Both terms disappear if $\rho = 0$, such that no interdependence exists. If interdependence in the outcome does exist, such that $\rho \neq 0$, however, 2SLS is biased.

Notably, this bias persists even when $\mathbf{z}$ is randomly assigned and, therefore, independently distributed and otherwise exogenous. It is in this case that the two-term expression of the bias in equation (11) becomes useful. When $\mathbf{z}$ is independently distributed, the first term drops out, because independence in $\mathbf{z}$ implies that any specification of $\mathbf{W}$ yields $\text{cov}(\mathbf{Wz}, \mathbf{z}) = 0$.[7] That is, the value of $\mathbf{z}$ on unit i is uncorrelated with the value of $\mathbf{z}$ on any other unit (and their weighted-sum $\sum_j w_{ij}z_j$). However, this is not true of the second term in equation (11). While $\mathbf{W}$ is a hollow matrix – all elements along the diagonal equal zero – higher order multiples of $\mathbf{W}$ are not hollow matrices as ties between units are not unidirectional.[8] Because $\mathbf{W^k}$ has non-zero diagonal elements, it follows that $\mathbf{W}^k\mathbf{z}$ is, for unit i, a function of z_i, and therefore correlated with $\mathbf{z}$, regardless of the distribution of $\mathbf{z}$.

To gain more intuition for why this is the case, recall that $\mathbf{W}$ can be thought of as defining 'neighbors': non-zero entries indicate which units on the outcome variable are related to one another. Then, for each unit, the respective row of $\mathbf{W}$ defines a set of neighbors. Heuristically, powers of $\mathbf{W}$ then represent neighbors-of-neighbors. For example, the ith row of $\mathbf{W}^2$ indicates i's neighbors' neighbors. This is important because, intuitively, a unit always is a neighbor of its own neighbors. Consequently, if $\mathbf{W}$ links unit i to j and unit j to unit i, then $\mathbf{W}^2$ (and higher powers of $\mathbf{W}$) links unit i back to itself. Therefore, even under independence of $\mathbf{z}$, some $\mathbf{W}^k\mathbf{z} \not\perp \mathbf{z}$ as long as $\mathbf{W}$ is non-triangular. Put simply, even if unit i is not related to any of the neighbors defined by $\mathbf{W}$, unit i is always related to itself through these higher powers of $\mathbf{W}$.

That is, for $\rho \neq 0$, any instrument that is randomly assigned is (only) first-order unbiased, providing a lower bound on the spatial bias. While the bias is relatively mild, spatial

---

[7]Recall that $\mathbf{W}$ is the connectivity matrix of the outcome – based on, e.g., contiguity, neighbors, or inverse distance – defining how y_i is related to all $y_{j\neq i}$. In connectivity matrices like $\mathbf{W}$ the diagonal elements are always zero, that is, you cannot be a direct neighbor of yourself.

[8]If spatial ties were unidirectional – $\mathbf{W}$ is upper- or lower-triangular – the higher order multiples would remain independent of z_i. However, *inter*dependence generally rules out unidirectional ties. The importance of reciprocal relationships between units – i.e., interdependence – for our results can also be seen in the contrast to temporal dependence. With temporal dependence, the current value of the outcome is a function of past values of the outcome, but past outcomes are not a function of the current value. Hence, a randomly assigned instrument poses no problems under temporal dependence.

interdependence in the outcome variable renders IV models biased, even under conditions most favorable to IV models, such as experimental or quasi-experimental designs.

RESULT 1 *With unmodelled spatial interdependence in the outcome, 2SLS is asymptotically biased.*

However, the instruments often used in practice are not independently distributed, risking greater bias still. Specifically, the more the values z_i are similar to neighboring values $z_{j \neq i}$ (where neighboring values are defined by **W**, the matrix defining relationships among units for the outcome), the greater the bias will be: the first term in equation (11) no longer drops out, and all of the terms in the expression increase in magnitude.

To understand this result, it helps to think of 2SLS broken down into two stages. The first stage is a regression of the endogenous predictor **x** on the instrument **z**, which yields fitted values $\hat{\mathbf{x}}$. The second stage is a regression of the outcome variable **y** on the fitted values $\hat{\mathbf{x}}$. We make two observations. First, if **z** follows a spatial distribution, the projected values $\hat{\mathbf{x}}$ inherit some of that spatial pattern. Second, in a regression with an (erroneously) omitted spatial lag, the bias in coefficient estimates is reinforced for variables that have a spatial distribution similar to that of the outcome (see, e.g., Franzese and Hays 2007). It follows that the bias in 2SLS becomes most severe if the fitted values $\hat{\mathbf{x}}$ have a spatial distribution similar to that of the outcome – which, in turn, is the case if the instrument has a spatial distribution similar to the outcome.

It is not crucial that the instrument and the outcome follow identical spatial patterns, merely that the instrument and the outcome have some similarity in their spatial patterns. That is, the bias in 2SLS increases if the **W** that characterizes the relationships in **y** also partially characterizes the relationships in **z**. In practice, when considering the extent of the spatial bias in 2SLS, one can therefore remain largely agnostic about the nature of the spatial relationships on the instrument – in particular, it is not necessary to determine whether **z** and **y** are truly governed by identical **W**s or even which specific **W** applies to the instrument (and our empirical approach, detailed in the next section, is consistent with this view). Our point is much simpler: if the outcome is spatially interdependent, then the bias in 2SLS will be more severe for instruments with spatial patterns similar to that of the outcome.

These concerns apply to a large set of common instruments. Researchers often draw on geographic, meteorologic, or economic variables, such as natural disasters (Ramsay 2011), rainfall data (Hansford and Gomez 2010), or commodity price shocks (Ahmed 2012), where spatial dependence among units is likely – natural disasters, rainfall, and price shocks do not stop at territorial borders. The same problem arises for instruments that are measured at a higher level of aggregation than the endogenous predictor. If, for instance, the instrument is based on regional political or institutional shocks, such as waves of democratization (Stasavage 2005) or membership in international institutions in neighboring countries (Büthe and Milner 2008), the instrument induces spatial correlation in the projection $\hat{\mathbf{x}}$ by construction: the value of the instrument is identical or nearly identical for each of the lower-order observations within the cluster. Since many outcome variables of interest in political science also cluster regionally – e.g., democratization, economic growth, or policy – regional-level instruments are likely to reinforce the bias in 2SLS.

To illustrate, consider the use of meteorological variables as instruments for democratization (**z**) in models of economic development (**y**). Contiguous states (a widely used **W**) likely have both similar levels of development (**y**) and common weather patterns (**z**), where the former implies $\rho > 0$ and the latter implies $\text{cov}(\mathbf{Wz}, \mathbf{z}) > 0$. It is under these conditions that the bias will be most severe; as can be seen in equation (12), the bias increases in the strength of the interdependence in the outcome (ρ) and the strength of the spatial dependence in the instrument ($\text{cov}(\mathbf{Wz}, \mathbf{z})$).

RESULT 2 *With unmodelled spatial interdependence in the outcome, the more similar are the spatial distributions of the instrument and the outcome, the greater is the bias in 2SLS.*

We add three additional observations. First, these biases are usually inflationary, which can be seen from equation (11). The bias terms are multiplied by powers of ρ, which is positive in most applications (Franzese and Hays 2007). And, if $\mathbf{z}$ is governed by a similar pattern of spatial dependence as the outcome, the covariances between $\mathbf{W}^k\mathbf{z}$ and $\mathbf{z}$ are non-negative. Consequently, the right-hand side of equation (11) should have the same sign as β and be proportional to β. Thus, in most applications the bias in 2SLS that arises from spatial interdependence exaggerates the true parameter value – where β is negative, 2SLS produces smaller coefficient estimates, and where β is positive, 2SLS produces larger coefficient estimates.

Second, the spatial bias induced from the instrument can exceed the spatial bias in ordinary least squares. Consider the relative spatial bias of OLS (the left-hand side) and 2SLS (the right-hand side):

$$\frac{\text{cov}(\mathbf{Wy}, \mathbf{x})}{\text{var}(\mathbf{x})} \lesseqgtr \frac{\text{cov}(\mathbf{Wy}, \mathbf{z})}{\text{cov}(\mathbf{x}, \mathbf{z})}. \tag{12}$$

To focus on the comparison of the spatial bias between 2SLS and OLS, suppose that no non-spatial endogeneity exists. Re-expressing both terms, condition (12) becomes

$$\sum_{k=1}^{\infty}\left[\rho^k \frac{\text{cov}(\mathbf{W}^k\mathbf{x}, \mathbf{x})}{\text{var}(\mathbf{x})}\right] \lesseqgtr \sum_{k=1}^{\infty}\left[\rho^k \frac{\text{cov}(\mathbf{W}^k\mathbf{z}, \mathbf{z})}{\text{var}(\mathbf{z})}\right]. \tag{13}$$

Simply put, differences in the spatial distribution of the instrument and the endogenous variable inform the relative degree of spatial bias. This is similar to Bartels's (1991) recognition that, because $\mathbf{x}$ can be considered its own instrument, when using an invalid instrument $\mathbf{z}$ the gains relative to OLS are a function of the relative difference in how $\mathbf{z}$ and $\mathbf{x}$ covary with the disturbance of $\mathbf{y}$. Again thinking of the second stage in 2SLS as a regression of $\mathbf{y}$ on the projection $\hat{\mathbf{x}}$ further clarifies the role of spatial dependence in the instrument: the bias of 2SLS relative to OLS increases as the spatial distribution of the instrumented predictor, $\hat{\mathbf{x}}$, becomes more similar to the spatial distribution of the outcome than the original predictor, $\mathbf{x}$. Then, IV models augment the spatial bias, because $\hat{\mathbf{x}}$ is more similar to the omitted spatial lag than $\mathbf{x}$ is. The reverse, of course, also holds: if the instrument is randomly assigned, then the similarity between the spatial pattern of the instrumented predictor, $\hat{\mathbf{x}}$, and the outcome decreases, and the bias of 2SLS relative to OLS declines. Nonetheless, even in that case, as we emphasize in Result 1, 2SLS remains biased.

Finally, because spatial and non-spatial endogeneity biases may attenuate or reinforce each other, ignoring spatial interdependence in the outcome risks unpredictable and possibly greater overall bias than OLS. When the endogenous variable, $\mathbf{x}$, is spatially less clustered than the instrument, $\mathbf{z}$, the severity of the difference in the spatial biases may be sufficiently large to surmount the gains from addressing non-spatial endogeneity. And because the spatial and non-spatial bias may have different directions, resolving one of the biases may easily produce results further from the truth than resolving none. Perhaps most problematically, these offsetting effects mean that the OLS and 2SLS estimates will not even be sufficient to obtain bounds on the true parameter value.

Spatial models with additional endogenous predictors

What, then, can researchers concerned with endogeneity in a key predictor and spatial interdependence in the outcome do? The solution is actually quite simple: estimate a modified IV model. While early work in spatial econometrics assumed exogenous predictors, methods for estimating models with additional endogenous predictors have become increasingly common (Kelejian and Prucha 2004; Anselin and Lozano-Gracia 2008; Fingleton and Le Gallo 2008; Drukker, Egger and Prucha 2013; Liu and Lee 2013). To date, however, these models

have not received much attention in applied spatial work in political science, and even less so in contexts where researchers are not theoretically interested in spatial relationships. In short, to redress the concerns above, researchers need to account for the spatial interdependence of the outcome. Yet, including a spatial lagged-outcome produces a system of simultaneously-determined, non-separable equations. That is, $\mathbf{Wy}$ is itself is an endogenous predictor, no different than a simultaneously-determined $\mathbf{x}$. Consequently, in spatial modeling, researchers exploit the same strategies generally used when confronting endogenous predictors (such as maximum likelihood, 2SLS, or GMM). As such, one can simply extend the familiar IV framework, applying it to account for outcome interdependence and predictor endogeneity. As this is simply a special case of multiple endogenous variables, a spatial-two stage least squares (S-2SLS) model can be estimated as in standard IV analysis: instrumenting for $\mathbf{Wy}$ and $\mathbf{x}$ simultaneously.

While other solutions are also available – e.g., purging the spatial dependence of the outcome equation via eigenvector filtering – we prefer S-2SLS for several reasons.[9] First, the mechanics of estimating this model are already familiar to researchers using IV estimation for an endogenous variable $\mathbf{x}$, because the estimator, 2SLS, is the same. Second, S-2SLS nests the non-spatial IV model a researcher would have otherwise estimated. Rather than restrict ρ – the spatial effect – to be zero by assumption, as in 2SLS, S-2SLS allows researchers to explicitly test this. As we demonstrate in simulations, this nesting helps ensure that – even if no spatial interdependence is present and $\rho = 0$ – the model recovers the same estimates as the original 2SLS, with only minimal efficiency loss due to the additional parameter. Finally, the S-2SLS model, as well as several extensions, can be estimated in both Stata (spivreg) and R (sphet).

The only practical hurdles to estimating a S-2SLS are in the specification stage: (i) what are appropriate instruments for the spatial lag and (ii) what is the appropriate connectivity matrix $\mathbf{W}$ for the outcome variable. The first, instrument selection, is comparatively simple. While instruments for the endogenous predictor usually require finding additional exogenous variables, instruments for the spatial lag can typically be found from transformations to the existing data. Specifically, spatial lags of the exogenous predictors serve as instruments for the spatial lag of the outcome. To see the basic intuition for this, just multiply $\mathbf{W}$ by both sides of the simple linear-additive model – i.e., $\mathbf{y} = \beta \mathbf{x} + \mathbf{e} \Rightarrow \mathbf{Wy} = \beta \mathbf{Wx} + \mathbf{We}$. Just as $\mathbf{x}$ is related to $\mathbf{y}$, $\mathbf{Wx}$ is related to $\mathbf{Wy}$, the spatial lag.[10]

The second practical hurdle, the selection of $\mathbf{W}$, is already familiar to researchers with exposure to spatial models. For those less familiar, we briefly sketch out the basics. To undertake spatial econometric modeling, researchers must pre-specify how units are related to one another (i.e., the network). Geographic proximity (e.g., contiguity) is commonly used, though researchers should specify connections that are most theoretically appropriate for their data. These relational measures for 'space' are then supplied to the model as the elements in $\mathbf{W}$ – an n-by-n connectivity matrix which identifies the relationship between units i and j. S-2SLS clearly performs best when $\mathbf{W}$ reflects the true network, yet gains are still likely even when researchers do not have full information on the ties between units. First, in the worst-case (and unlikely) scenario that a researcher completely mischaracterizes $\mathbf{W}$, this would still do no worse in expectation than 2SLS – S-2SLS recovers a zero estimate of ρ due to misspecified $\mathbf{W}$, while 2SLS does so by

---

[9]Limited and full information estimators allowing for both spatial and non-spatial endogeneity have been established, with Kelejian and Prucha (2004) the first to derive formal large sample results; see also Drukker, Egger and Prucha (2013) for a GMM estimator. Franzese, Hays and Cook (2016) discuss the complications of modeling spatial interdependence in discrete-choice models.

[10]A more complete derivation can be seen by noting that the reduced form of the spatial-lag model discussed in section 3,

$$\mathbf{y} = (\mathbf{I} - \rho \mathbf{W})^{-1}[\mathbf{x}\beta_x + \mathbf{u}],$$

can be re-expressed using an infinite series and multiplied through by $\mathbf{W}$ to produce

$$\mathbf{Wy} = \mathbf{Wx}\beta_x + \rho \mathbf{W}^2 \mathbf{x}\beta_x + \rho^2 \mathbf{W}^3 \mathbf{x}\beta_x + \ldots + (\mathbf{I} - \rho \mathbf{W})^{-1}\mathbf{u},$$

indicating how spatial lags of $\mathbf{x}$ (and their higher order powers) effectively instrument for $\mathbf{Wy}$.

assumption. Second, due the high correlation across different possible network structures, even a misspecified **W** has power against the truth (LeSage and Pace 2014). We revisit this concern in the simulated experiments in the next section.[11]

Once specified, estimation of the S-2SLS model proceeds without additional complications. Because S-2SLS is estimated via the 2SLS estimator, it inherits the asymptotic and small-sample properties of 2SLS (including consistency, finite sample bias, and the sensitivity to weak instruments).[12] Similarly, standard variance estimators – robust to heteroskedasticity or non-independence, for instance – are easily applicable. We demonstrate the gains that can be realized from S-2SLS in the following sections.

Simulation experiments

To assess the performance of OLS, 2SLS, and S-2SLS, we undertake a series of Monte Carlo experiments with varying levels of spatial and non-spatial endogeneity. In particular, the data for our simulations are generated as follows:

$$\mathbf{y} = (\mathbf{I} - \rho_y \mathbf{W})^{-1}[\mathbf{x}\beta + \lambda_1 \mathbf{Q} + \mathbf{u}_1], \tag{14a}$$

$$\mathbf{x} = \gamma \mathbf{z} + \lambda_2 \mathbf{Q} + \mathbf{u}_2, \tag{14b}$$

$$\mathbf{z} = (\mathbf{I} - \rho_z \mathbf{W})^{-1}\mathbf{v}, \text{ where } \mathbf{v} \sim N(0, 1), \tag{14c}$$

where **y** is the outcome, **x** is the endogenous predictor, **Q** is a matrix of exogenous predictors, **W** is a row-standardized connectivity matrix, and **z** is the instrument.[13] Consistent with our discussion above, we only consider the consequences of spatial interdependence in **y** and **z**, which are the key attributes for bias in 2SLS.[14]

The extent of spatial interdependence in the outcome and the instrument is given by parameters ρ_y and ρ_z, respectively, with larger values of ρ_y and ρ_z resulting in greater spatial interdependence in **y** and **z**. We do not vary the specification of the **W** that governs the spatial pattern of **y** and **z**, respectively. Non-spatial endogeneity is induced through draws of $(u_1, u_2)^T = N(0, \Sigma)$, where Σ is the covariance matrix of a bivariate normal random variable. We decompose Σ such that we can specify the correlation (δ) between u_1 and u_2 directly. We vary δ to induce different degrees of non-spatial endogeneity. If $\delta = 0$, **x** is exogenous and OLS (or standard spatial) models should be preferred. With non-zero δ and non-zero ρ_y, the assumptions of neither OLS nor 2SLS hold.

This setup allows us to consider various scenarios that correspond to our results above. $\rho_y = \rho_z = 0$ produces the standard IV model with an i.i.d. instrument, such that 2SLS should

---

[11]Note that the researcher need not specify the spatial distribution of the instrument. **W** specifies connections among units with respect to the outcome variable. As we highlighted in the previous section, the extent of the bias in 2SLS depends on the similarity in the spatial pattern between the instrument and the outcome. But it is not necessary to determine the specific spatial pattern of the instrument.

[12]For an approach addressing weak instruments in IV models, see Betz (2013). Drukker, Egger and Prucha (2013) and Liu and Lee (2013) allow for both additional residual spatial error autocorrelation and/or heteroskedasticity. These extensions are GMM-plus-IV, implemented in Stata's spreg. While we do not discuss this at length here, the first step is the S-2SLS we present, which provides the initial, consistent estimates of the spatial interdependence in the outcome that can then be used in the second step estimation of the error autocorrelation, with successive iteration over both steps until convergence is obtained.

[13]Locations for the units are generated by twice taking *n* draws from a standard uniform, with the combined results producing xy-coordinate points. Connections between the units are then generated using a *k*-Nearest Neighbor algorithm with *k* = 5, returning a binary *n*-by-*n* matrix with each element in a row coded as 1 for the five closest units or 0 for all others (including zeros along the diagonal). The matrix is then row-standardized.

[14]As discussed above, the relative spatial pattern of **x** and **z** only matters for the performance of 2SLS relative to OLS. For the simulations, to illustrate Result 2, we only consider scenarios where 2SLS performs relatively poorly due to the spatial pattern in **z**.

Table 1. Parameter Values for Simulations

n	50	200	
ρ_y	0	0.3	0.6
ρ_z	0	0.3	0.6
γ		0.75	1.5
δ	−0.5	0	0.5

Note: Values in italics used in Figures 2 and 3.
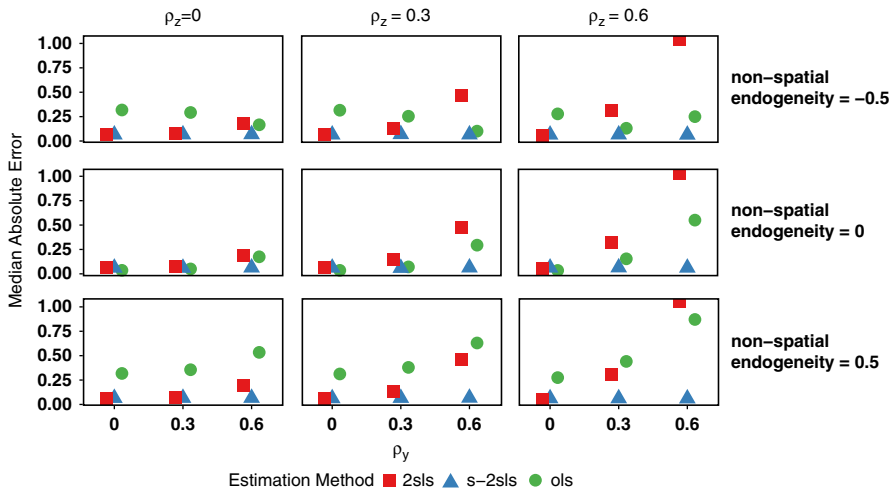

Figure 2. Median absolute error (MAE). Rows: δ, non-spatial endogeneity. Columns: ρ_z, spatial interdependence in the instrument. Horizontal axis within each plot: ρ_y, spatial interdependence in the outcome. Vertical axis within each plot: MAE.

perform well. $\rho_y \neq 0$ but $\rho_z = 0$ implies interdependence in the outcome but an i.i.d. instrument. Following Result 1, we should still observe some bias in 2SLS in this scenario, whereas S-2SLS should perform better. As ρ_z increases, the bias in 2SLS should increase, both in absolute terms (Result 2) and relative to OLS, because the instrument becomes more similarly distributed to the outcome relative to the predictor. Finally, varying δ, the extent of non-spatial endogeneity, allows us to evaluate scenarios under which OLS – which produces spatial and non-spatial endogeneity – should perform worse than 2SLS – which produces only spatial endogeneity.

The remaining parameters $\{\beta, \gamma, \lambda_1, \lambda_2\}$ are the coefficients of the predictors of **x** and **y**, respectively.[15] Our main focus is on the estimate of β, which we hold constant across experiments at 2. Table 1 shows the different parameter values which we use to create simulated data sets. There are 108 different combinations of the parameters shown in Table 1 (with the italic values indicating those used in the subsequent figures). For each combination we generate 1000 data sets, which results in a total of 108,000 simulation runs. On each dataset, we estimate β using OLS, 2SLS, and our preferred method, S-2SLS.

The results are presented in Figures 2 and 3, which report the median absolute error and coverage probabilities for the estimators, respectively.[16] The figures vary along three dimensions.

---

[15]In the first stage, we specify the intercept as 2. The two exogenous predictors have coefficients 3 and −2.5. For the second stage, the intercept is −2 and the exogenous predictors are −3 and 2.5. For the plots presented in the manuscript, we set the coefficient on the instrument to $\gamma = 1.5$ and the number of observations to $n = 200$. Observations for the predictors are drawn from standard normal distributions.

[16]We prefer MAE as it limits the influence of potential outliers. In the Online Appendix, we also present model performance in RMSE terms.
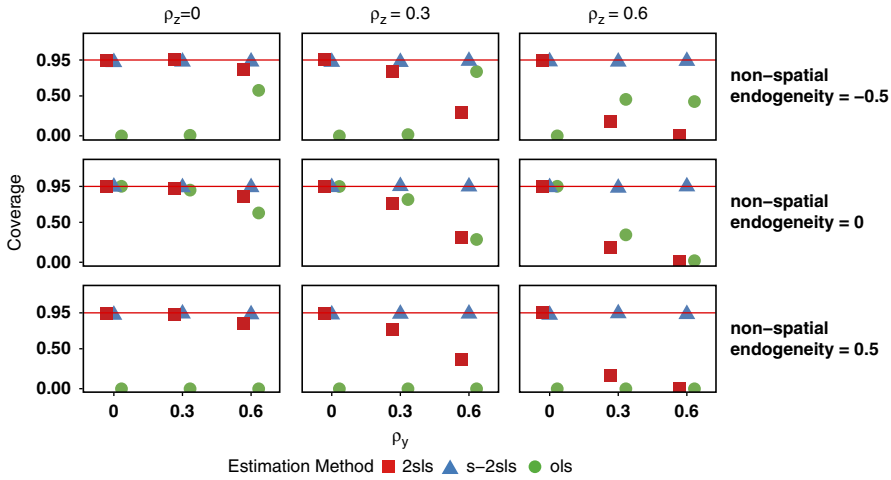
Figure 3. Coverage probabilities. Rows: δ, non-spatial endogeneity. Columns: ρ_z, spatial interdependence in the instrument. Horizontal axis within each plot: ρ_y, spatial interdependence in the outcome. Vertical axis within each plot: Coverage rate.

First, δ – the non-spatial endogeneity – increases across the three rows from −0.5 in the top row, to 0 in the middle row, to 0.5 in the bottom row. Second, each column shows results for a different value of ρ_z – the spatial pattern of the instrument – ranging from 0 in the column on the left over 0.3 in the middle to 0.6 on the right. Finally, within each individual plot, ρ_y – the spatial interdependence in the outcome – increases from left to right across the x-axis.

Several observations stand out from the plots. Turning to the median absolute error (MAE) in Figure 2 first, across all levels of non-spatial endogeneity (δ), the error of 2SLS grows as ρ_y increases, dramatically so as ρ_y and ρ_z increase together. This is consistent with our theoretical results: under interdependence in the outcome, the 2SLS model always returns biased estimates (Result 1), with the severity of these biases increasing in the similarity of the spatial pattern in the instrument and the outcome (Result 2). Importantly, when both the instrument and outcome are characterized by spatial dependence, a situation that in our view is not uncommon in the literature, the bias in 2SLS increases quickly. Conversely, the MAE of S-2SLS is stable, as its performance does not suffer under high interdependence in **y**, **z**, or both. In fact, S-2SLS weakly dominates 2SLS, besting it when spatial interdependence is present and matching it otherwise. Thus, when non-spatial endogeneity is present and IV models may be warranted, S-2SLS performs better than or as good as 2SLS. Across all scenarios considered in the simulations, 2SLS performs better in terms of MAE only when $\rho_y = 0$, and even then the maximum difference in median absolute error between 2SLS and S-2SLS is 0.03. While not surprising, this bolsters our claim that S-2SLS is a useful conservative specification, robust under non-spatial and spatial endogeneity, because it nests both cases.

The OLS estimator performs poorly when either non-spatial or spatial endogeneity is present. However, and as discussed above, the bias can be larger for 2SLS than for OLS, even in the case of strong non-spatial endogeneity where OLS should perform poorly. This occurs under higher levels of ρ_z and ρ_y – as we move from the left to the right in each box, and as we move from the left column to the right column – where the spatial and, in turn, total bias of 2SLS is greater due to the spatial interdependence of the instrument.

The top and middle rows of Figure 2 present two particularly interesting scenarios. In the top row, with negative non-spatial endogeneity and positive spatial interdependence, the relative performance of OLS improves, both in absolute terms and relative to 2SLS, as the spatial interdependence increases. The two biases are countervailing, combining to produce a result closer to the truth. Under these conditions, 2SLS produces relatively worse results, as it addresses

one type (and therefore direction) of bias, while neglecting the other. As a result, 2SLS produces *more* biased estimates even while – in fact, due to – addressing one of the sources of that bias.

In the middle row, we have no non-spatial endogeneity bias, and relying on 2SLS is unnecessary. Usually, using 2SLS instead of OLS is not much of a concern, aside from a slight efficiency loss. This changes with interdependence. If the instrument is spatially more similar to the outcome than the predictor (as in the second and third column), 2SLS produces more total bias than OLS. In this case, 2SLS not only was unnecessary, but results in worse estimates than OLS. (Of course, this result hinges on the simulation setup, which consistent with our discussion allowed for a spatial pattern in z but not in x – if the reverse was the case, 2SLS would perform relatively better.)

These results are particularly problematic, as researchers relying on 2SLS over OLS estimates will be more confident about results that are further from the truth and dismissive of results that were closer to it. Frequently, a difference between 2SLS and OLS estimates is accepted as evidence of suspected non-spatial endogeneity (such as measurement error or reverse causality) that was successfully removed by 2SLS. While 2SLS removes non-spatial endogeneity, such arguments ignore that 2SLS may come with biases of its own, and that these biases need not be less pronounced than the biases in OLS. Where outcomes are interdependent, there is no guarantee that 2SLS produces better estimates than OLS. S-2SLS, by contrast, does not suffer from this issue and consistently outperforms both OLS and 2SLS.

Figure 3 shows the coverage probabilities for each estimator. The coverage statistic measures the share of estimates for which the true parameter falls within the 95% confidence interval of the estimate. If perfectly calibrated, we would expect this to be true for 95% of cases. The results are generally consistent with our expectations. First, the coverage of OLS is generally poor under either spatial or non-spatial endogeneity. However, for the reason just discussed, when the spatial and non-spatial bias are oppositely signed (top row), the coverage of OLS improves with higher spatial interdependence. Second, with interdependence in the outcome, the 2SLS estimator undercovers, with the severity of this increasing ρ_z. Finally, S-2SLS has very good coverage throughout and is not affected by interdependence in z or y. In fact, the coverage of S-2SLS is consistently around 95%, ranging between 92% and 96%.

Robustness checks

In the Online Appendix, we discuss a series of additional experimental conditions. First, what if the researcher has incomplete information on the spatial network (i.e., the $\mathbf{W}$ matrix)? To evaluate this, we undertake additional experiments where we vary the level of misspecification – from no error to total error – of the spatial network used in estimation. As Figure A.1 and Figure A.2 in the Online Appendix show, S-2SLS weakly dominates 2SLS. This demonstrates what we articulated earlier: because S-2SLS nests 2SLS, it only suffers minor efficiency losses when it is the incorrect model. Next, we explore the consequences of a weak instrument (i.e., $\gamma = 0.75$). As expected, IV methods perform worse, yet the overall order in performance between the different methods does not change. Finally, we evaluate how the performance of the estimators varies with changes to sample size. All the results presented above hold.

Application

To illustrate how failing to account for spatial interdependence when using IV models can induce bias in published research, we replicate Ramsay's (2011) "Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy."[17] A long-standing literature in political science has considered the effects of natural resource revenues on political order. Ramsay (2011) identifies

---

[17]In the online Appendix, we provide an additional replication of Ashraf and Galor's article "Dynamics and Stagnation in the Malthusian Epoch" (2011).

reverse causality as one of the main threats to inference: changes in resource revenues may cause political change, but politics may also affect resource revenues.

The main independent variable of interest is a country's annual oil income per capita (specifically the price of crude oil times annual production divided by the population). The dependent variable is a country's level of democracy, measured as a normalized score of Polity IV. A valid instrument would have sufficient power to explain oil revenues; and fulfill the exclusion restriction such that it only affects changes in democracy via the path through oil revenues. In light of these requirements, Ramsay (2011) introduces out-of-region natural disasters as the IV (where regions are defined as Europe, Middle East, North Africa, sub-Saharan Africa, Asia, or the Americas). The rationale is that natural disasters, by reducing oil production in the affected countries, change world oil prices, and therefore oil revenues of individual countries; at the same time, natural disasters should have no direct effect on oil production in remote countries.

We highlight three concerns with these IV models. First, levels and changes of democracy cluster in space (Gleditsch and Ward 2006). Second, natural disasters, the instrument of choice, likely correlate in space. As Ramsay (2011) notes, the effects of disasters are likely to spill over, affecting neighboring states. Finally, Ramsay (2011) aggregates the variable to the regional level, inducing a spatial pattern by construction. By designing the instrument as "out of region disaster damage estimates" (Ramsay 2011, 514), all countries within each of the five regions have the same value on the instrument, thus inducing spatial correlation in the instrument by design. As we discussed above, with a spatially interdependent outcome and instrument, we generally expect inflationary bias in the non-spatial IV estimator.

The results from our analysis are presented in Table 2.[18] First, we estimate a linear model (via OLS), with our results (see Model 1) reproducing those found in Ramsay (2011).[19] Next, we determine whether these findings hold once we account for spatial interdependence. Before undertaking spatial analysis, we need to select the connectivity matrix, **W**.[20] Here, we use a row-standardized geographic binary contiguity matrix, as these are widely used in the literature.[21] Given **W**, we estimate a SAR model (Model 2), which returns a significant value for the spatial effect parameter ($\rho = 0.173$; p-value <0.05). The coefficient for oil income per capita is still negative and statistically significant.[22]

This indicates that there does appear to be spatial interdependence in the model, so now we consider the consequence for the IV estimator. In Model 3 of Table 2, we replicate the 2SLS model in Ramsay (2011).[23] Here the estimated coefficient of log oil income is −0.36 and statistically significant. That is, the 2SLS model presents an effect estimate that is almost *eight times larger* than the original OLS estimate. Finally, we estimate our preferred S-2SLS model, with the results given in Model 4 of Table 2.[24] The instrumented coefficient of logged oil income is now estimated to be −0.088. That is, while the effect is still significant and in the expected direction, its magnitude is much smaller than in the 2SLS model that ignores spatial interdependence.[25]

---

[18]As in Ramsay (2011), the predictors include log oil income per capita, GDP per capita, GDP growth, a lagged polity variable, and year fixed effects

[19]This is Model 4 in Ramsay (2011).

[20]A detailed discussion on **W** selection is beyond the scope of this paper, interested readers should consult Neumayer and Plümper (2016).

[21]We are unable to merge 4 observation to the shapefiles and therefore drop these from all models.

[22]Due to the non-linear nature of the SAR model, the average direct effect – that is, the average effect of a one-unit change in x_i on y_i – is not the coefficient estimate, but instead: $\mathrm{Tr}\{(\mathbf{I} - \rho\mathbf{W})^{-1}\beta_x\}/n$. For the variable of logged oil income per capita, this results in a value of −0.042, slightly smaller than the OLS effect estimate of −0.046.

[23]In Ramsay (2011), this is presented in column 4 of Table 3.

[24]**W** is the same as in Model 2.

[25]Table D.1 in the Appendix displays the results comparing the 2SLS model and the S-2SLS model for the robustness checks presented in Table 5 in Ramsay (2011). Again, the differences between 2SLS and S-2SLS are stark and reflect the same pattern as shown in Table 2.

Table 2. Replication of OLS and IV Results Table 1 and 3 in Ramsay (2011)

	(1) OLS Replication	(2) SAR Spatial	(3) 2SLS Replication	(4) S-2SLS Spatial
Log oil income per capita	−0.0457***	−0.0420***	−0.358**	−0.0878***
	(0.00575)	(0.00354)	(0.167)	(0.0114)
Log GDP per capita	0.0653***	0.0662***	0.356**	0.108***
	(0.00877)	(0.00533)	(0.155)	(0.0115)
GDP growth	−0.00363***	−0.00391***	−0.0118**	−0.00499***
	(0.00106)	(0.000921)	(0.00518)	(0.00102)
Polity at entry	0.666***	0.658***	−0.00517	0.564***
	(0.0280)	(0.0165)	(0.373)	(0.0288)
Constant	−0.171***	−0.237***	−0.962**	−0.330***
	(0.0518)	(0.0499)	(0.439)	(0.0591)
Spatial ρ_y		0.173***		0.119***
		(0.015)		(0.020)
Observations	1263	1263	1263	1263
Year dummies	Yes	Yes	Yes	Yes

Note: W matrix for spatial models based on contiguous neighbors. Instrumental variable: out-of-region natural disasters. Table shows coefficient estimates, standard errors in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Furthermore, we see substantial efficiency gains in the estimate – as indicated by the standard errors – once we account for spatial interdependence.

In sum, failing to account for spatial interdependence resulted in substantial inflationary bias in the estimates of interest. We do not overturn the central finding presented in Ramsay (2011) that oil revenue is negatively associated with the polity score, but the magnitude of the effect is reduced considerably and the purported gains from IV estimation are significantly reduced.

Conclusion

IV models are now a frequently used tool in political science research. IV methods are especially common in observational research, where endogeneity often threatens causal inference. However, observational data are also where concerns of spatial interdependence are the most salient and where instruments are unlikely to be randomly assigned. Consequently, IV methods are most widely used where the biases due to unmodelled outcome interdependence discussed above are the most likely to occur. This problem may be especially pronounced in published research as researchers are disproportionately likely to publish IV results where gains over OLS are the most pronounced – as can and will occur if outcome interdependence is present and unaddressed.[26] We discuss a simple strategy researchers should employ to avoid these biases: S-2SLS. This estimation strategy offers few complications for researchers already pursuing IV methods, inherits the properties of 2SLS familiar to those using IV models, and ensures results are robust to spatial interdependence. Our simulations evidence that S-2SLS performs well across a variety of situations and presents a conservative and robust alternative.

Our discussion adds to growing concerns over spatially dependent instruments (Cooperman 2017; Betz, Cook and Hollenbach 2018). While we have identified challenges to credible inference using observational data, we emphasize that we do not discourage analyses using these data. Instead, our purposes in this paper are twofold. First, we highlighted the unique problems posed by spatial interdependence for IV models. In our reading of the literature, these problems have largely been ignored by applied researchers. Second, we want to encourage researchers to

---

[26]This is a variant of the file drawer problem, with published findings suffering from substantial selection effects: biased IV estimates, driven by unmodeled spatial interdependence, are those which are most likely to be reported.

consider more carefully the potential drawbacks of IVs. Frequently, IV estimates are conjectured to be superior to results from ordinary least squares. This assumption is often wrong. The estimates obtained from IV models can quickly, and under fairly general circumstances, be worse than ordinary least squares, even with instruments that are plausibly exogenous. IVs can, under specific circumstances, identify causal effects. But these circumstances are more limited than is often realized, which should give researchers some pause in advocating the use of IV models. IVs may cause more problems than they solve.

Supplementary materials. To view supplementary material for this article, please visit https://doi.org/10.1017/psrm.2018.61

References

Ahmed FZ (2012) The Perils of Unearned Foreign Income: Aid, Remittances, and Government Corruption. *American Political Science Review* **106**(1), 146–65.

Anselin L and 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.

Ashraf Q and Galor O (2011) Dynamics and Stagnation in the Malthusian Epoch. *American Economic Review* **101**(5), 2003–41.

Bartels LM (1991) Instrumental and "Quasi-Instrumental" Variables. *American Journal of Political Science* **35**(3), 777–800.

Beck N, Gleditsch KS and Beardsley K (2006) Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy. *International Studies Quarterly* **50**(1), 27–44.

Betz T (2013) Robust Estimation with Nonrandom Measurement Error and Weak Instruments. *Political Analysis* **21**(1), 86–96.

Betz T, Cook SJ and Hollenbach F (2018) On the Use and Abuse of Spatial Instruments. *Political Analysis* **26**(4), 474–79.

Boix C (2011) Democracy, Development, and the International System. *American Political Science Review* **105**(04), 809–28.

Bound J, Jaeger DA and Baker RM (1995) Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak. *Journal of the American Statistical Association* **90** (430), 443–450.

Büthe T and Milner HV (2008) The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements? *American Journal of Political Science* **52**(4), 741–762.

Cooperman AD (2017) Randomization Inference with Rainfall Data: Using Historical Weather Patterns for Variance Estimation. *Political Analysis* **25**(3), 277–88.

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

Fingleton B and Le Gallo J (2008) Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties. *Papers in Regional Science* **87**(3), 319–39.

Franzese RJ Jr. and Hays JC (2007) Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data. *Political Analysis* **15**(2), 140–164.

Franzese RJ, Hays JC and Cook SJ (2016) Spatial- and Spatiotemporal-Autoregressive Probit Models of Interdependent Binary Outcomes. *Political Science Research and Methods* **4**(1), 151–73.

Gleditsch KS and Ward MD (2006) Diffusion and the International Context of Democratization. *International Organization* **60**(4), 911–33.

Hansford TG and Gomez BT (2010) Estimating the Electoral Effects of Voter Turnout. *American Political Science Review* **104**(02), 268–88.

Kelejian HH and Prucha IR (2004) Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations. *Journal of Econometrics* **118**(1), 27–50.

Kirby AM and Ward MD (1987) The Spatial Analysis of Peace and War. *Comparative Political Studies* **20**(3), 293–313.

LeSage JP and Pace RK (2014) The Biggest Myth in Spatial Econometrics. *Econometrics* **2**(4), 217–49.

Liu X and Lee L-F (2013) Two-Stage Least Squares Estimation of Spatial Autoregressive Models with Endogenous Regressors and Many Instruments. *Econometric Reviews* **32**(5-6), 734–53.

Neumayer E and Plümper T (2016) W. *Political Science Research and Methods* **4**(1), 175–93.

Plümper T and Neumayer E (2010) Model Specification in the Analysis of Spatial Dependence. *European Journal of Political Research* **49**(3), 418–42.

Ramsay KW (2011) Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy. *International Organization* **65**(3), 507–29.

Simmons BA, Dobbin F and Garrett G (2006) Introduction: The International Diffusion of Liberalism. *International Organization* **60**(4), 781–810.

Siverson RM and Starr H (1990) Opportunity, Willingness, and the Diffusion of War. *American Political Science Review* **84**(1), 47–67.

Sovey AJ and Green DP (2011) Instrumental Variables Estimation in Political Science: A Readers' Guide. *American Journal of Political Science* **55**(1), 188–200.

Starr H (1991) Democratic Dominoes: Diffusion Approaches to the Spread of Democracy in the International System. *Journal of Conflict Resolution* **35**(2), 356–81.

Stasavage D (2005) Democracy and Education Spending in Africa. *American Journal of Political Science* **49**(2), 343–58.

Ward MD and O'Loughlin J (2002) Spatial Processes and Political Methodology: Introduction to the Special Issue. *Political Analysis* **10**(3), 211–16.

Ward MD and Gleditsch KS (2002) Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace. *Political Analysis* **10**(3), 244–60.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.