

Spatially Correlated Measurement Error: Bias Amplification and Identification via Higher-Order Spatial Lags

mohamed riadh cherif

cherifmohamedriadh@gmail.com

Faculty of Economic Sciences and Management of Tunis – FSEG TUNIS

Research Article

Keywords: Measurement error, spatial correlation, errors-in-variables, spatial econometrics, GMM estimation, attenuation bias, plausible exogeneity, spatial mixing

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9706755/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Spatially Correlated Measurement Error: Bias Amplification and Identification via Higher-Order Spatial Lags

Abstract

Measurement error in spatial data is often cross-sectionally dependent. While the classical errors-in-variables literature assumes independence, this paper shows that spatial dependence fundamentally changes both the magnitude of attenuation bias and the validity of standard spatial instrumental-variable strategies.

We characterize how spatially correlated measurement errors amplify attenuation bias through network propagation. Under a spatial moving average (SMA) structure with finite memory, and under an explicitly maintained geometric α -mixing condition on the joint random field (Assumption A), spatial lags beyond a threshold become approximately orthogonal to the measurement error, with the approximation error decaying geometrically in the lag order. Building on this result, we develop a plausibly exogenous GMM estimator based on higher-order spatial moments. Monte Carlo simulations show bias reductions of 70-80%, RMSE reductions of 45-60%, and coverage close to nominal levels relative to OLS and conventional spatial IV estimators. An application to regional infrastructure and development in Tunisia illustrates the methodology, with explicit discussion of the severe limitation of $n=24$ spatial units. All theoretical results are derived under increasing-domain asymptotics with bounded-degree graphs and rely exclusively on the maintained geometric mixing assumption (Assumption A).

Keywords: Measurement error, spatial correlation, errors-in-variables, spatial econometrics, GMM estimation, attenuation bias, plausible exogeneity, spatial mixing.

JEL Classification : C13, C21, C26, R15

Notation Table	
Symbol	Meaning
n	Number of spatial units
y_i	Outcome variable
x_i^*	Latent true regressor
x_i	Observed, mismeasured regressor

μ_i	Measurement error
ε_i	Structural error
v_i	Innovation in SMA process
W_n	$n \times n$ spatial weights matrix (row-standardized, zero diagonal, bounded row sums uniformly in n)
W^k	denotes the k-th matrix power of W
β	Parameter of interest
q	Order of SMA measurement error (finite memory)
ℓ	Spatial lag order used as instrument (<i>with $\ell > q$</i>)
γ_k	SMA coefficients, $\gamma_0 = 1$
λ	Spatial autoregressive coefficient of x^* (SAR parameter)
$\phi(d)$	Mixing coefficient at graph distance d
ρ	Decay rate of mixing: $\phi(d) \leq C\rho^d$ with $0 < \rho < 1$
k	Variance multiplier in bias amplification
k_n	Number of instruments (may grow with n)
$d(i, j)$	Graph distance between nodes i and j
$\mathcal{F}_A$	σ -algebra generated by $\{\xi_i: i \in A\}$

1. Introduction

Measurement error is a pervasive issue in empirical economics, leading to attenuation bias and inconsistent parameter estimates. Since the foundational contributions of Durbin (1954) and Fuller (1987), a large literature has developed methods to correct for errors-in-variables, including instrumental variables (IV), repeated measurements, and moment-based approaches (Bound, Brown, and Mathiowetz, 2001; Schennach, 2016). A common maintained assumption in this literature is that measurement errors are independently distributed across observations.

This independence assumption is frequently implausible in spatial data, where variables are often interpolated, aggregated from administrative sources, or derived from remote sensing. Such procedures naturally induce cross-sectional dependence in measurement errors (Pinkse and Slade, 2010). Despite the widespread use of spatially generated data, relatively little is known about the econometric consequences of spatially correlated measurement error. Existing research has focused on the impact of spatial dependence in structural errors (Anselin, 1988; Kelejian and Prucha, 2007) or in latent regressors (Lee and Yu, 2010). However, the interaction between measurement error and spatial dependence remains largely untreated.

This paper makes three contributions. First, we show that spatial dependence in measurement error amplifies the classical attenuation bias. Under a spatial moving average (SMA) structure with finite memory, the effective noise variance is multiplied by a factor $\kappa > 1$ relative to the i.i.d. case, worsening inconsistency. This result generalizes the classical errors-in-variables bias formula to spatially dependent settings.

Second, we demonstrate that conventional spatial IV procedures — which typically use low-order spatial lags such as Wx as instruments — fail when measurement errors are spatially correlated because the instrument inherits the same dependence structure. This invalidates the exclusion restriction, rendering standard spatial IV estimators inconsistent.

Third, we propose an alternative identification strategy that exploits the finite-memory property of SMA(q) processes. Under an explicitly maintained geometric α -mixing condition (Assumption A), we show that spatial lags beyond a threshold become approximately orthogonal to the measurement error, with the approximation error decaying geometrically in the lag order. Building on this result, we develop a plausibly exogenous GMM estimator that uses higher-order spatial lags as internal instruments, following the framework of Conley, Hansen, and Rossi (2012). We establish consistency and asymptotic normality under increasing-domain asymptotics with bounded-degree graphs, relying solely on geometric mixing — not on spectral gap or bounded cycle length.

Monte Carlo simulations show that the proposed estimator reduces bias by up to 80% and RMSE by 45-60% relative to OLS, with coverage close to nominal levels. An illustrative application to regional infrastructure data in Tunisia (with $n = 24$ spatial units, presented as a methodological illustration only) demonstrates the practical implementation of the method.

The remainder of the paper proceeds as follows. Section 2 reviews related literature. Section 3 characterizes bias under spatially correlated measurement error. Section 4 establishes identification under geometric mixing and develops the GMM estimator. Section 5 presents Monte Carlo evidence. Section 6 provides an illustrative methodological application. Section 7 concludes.

2. Related Literature

This paper relates to four strands of literature.

First, it contributes to the classical errors-in-variables literature (Durbin, 1954; Fuller, 1987; Bound, Brown, and Mathiowetz, 2001; Schennach, 2016). These works generally assume independently distributed measurement errors. A few studies have considered dependent measurement errors in time series or panel settings (Griliches and Hausman, 1986), but spatial dependence has received little attention. For recent advances in semiparametric estimation of

models with measurement error, see Dong, Otsu, and Taylor (2022); their work, while focused on independent measurement errors, highlights the inherent difficulty of identification when the error distribution is unknown – a challenge that is further compounded in our spatial setting. Our contribution differs by explicitly modeling spatial dependence in the measurement process itself, deriving bias amplification results, and proposing solutions based on spatial mixing.

Second, the paper relates to spatial econometrics. Foundational work by Anselin (1988) introduced spatial autoregressive (SAR) and spatial moving average (SMA) models. Kelejian and Prucha (1999) developed a generalized method of moments (GMM) estimator for the SAR model, and Kelejian and Prucha (2007) extended the HAC (Heteroskedasticity and Autocorrelation Consistent) covariance matrix estimation framework to spatial settings, providing inference robust to unknown forms of spatial dependence and heteroskedasticity. The specification of the spatial weights' matrix W is itself a subject of ongoing research; for a recent contribution proposing the use of population flow data to construct an asymmetric spatial weight matrix, along with a validation framework using instrumental variables and the Hausman test, see Zhu, Li, and Hou (2022). Lee and Yu (2010) established asymptotic theory for spatial panel models. More recent contributions have examined spatial IV estimation with many instruments (Kelejian and Prucha 2010) and robust inference under heteroskedasticity and spatial correlation (Drukker, Egger, and Prucha, 2013). Our work complements this literature by showing that standard spatial IV based on low-order lags fails under spatially correlated measurement error, and by providing an alternative identification strategy using higher-order lags.

Third, the paper connects to the literature on identification through network structure. Manski (1993) highlighted the reflection problem in social interactions. Bramoullé, Djebbari, and Fortin (2009) showed that network structure can provide identification under specific exclusion restrictions. Johnsson and Moon (2021) extended this to linear social interaction models. We adapt this insight to exploit the decay of spatial dependence: under finite memory SMA(q) and geometric mixing, distant spatial lags become approximately orthogonal to the measurement error, akin to using network distance as an instrument. This is distinct from the standard use of network lags in peer effect models; we face a measurement error problem, not a simultaneity problem.

Fourth, the paper builds on the literature on plausible exogeneity (Conley, Hansen, and Rossi, 2012; Kiviet, 2020; Andrews and Kasy, 2019). Our higher-order spatial lags are not exactly exogenous — they retain residual correlation with the measurement error that decays geometrically with the lag order. We therefore treat them as “plausibly exogenous” and apply the framework of Conley et al. (2012) to conduct inference that is robust to local violations of the exclusion restriction. This approach differs from traditional spatial IV, which assumes exact exogeneity.

Finally, the paper relates to the literature on weak instruments (Stock, Wright, and Yogo, 2002; Mikusheva and Sun, 2022) as well as to recent work on identification with measurement error and endogeneity using nonseparable models (Hu, Shiu, and Woutersen, 2016). Increasing the lag order ℓ improves orthogonality but reduces first-stage relevance because the signal in the latent regressor decays as $|\lambda|^\ell$. This creates a bias-variance trade-off that we analyze both theoretically and through Monte Carlo simulations. We provide practical guidance (first-stage F-statistics, weak-IV robust confidence intervals) for applied researchers.

3. Bias under Spatially Correlated Measurement Errors

3.1 Model Setup

Let $\{(y_i, x_i^*)\}_{i=1}^n$ denote observations indexed by spatial units located on a graph $G_n = (V_n, E_n)$ with $|V_n| = n$. The structural relationship is

$$y_i = \beta x_i^* + \varepsilon_i, i = 1, \dots, n$$

where x_i^* is the latent regressor of interest, ε_i is an idiosyncratic disturbance satisfying

$$E[\varepsilon_i | x_i^*] = 0, \quad \text{Var}(\varepsilon_i) = \sigma_\varepsilon^2.$$

The econometrician does not observe x_i^* but instead a contaminated regressor

$$x_i = x_i^* + \mu_i$$

where μ_i denotes measurement error. The classical errors-in-variables framework assumes μ_i are independently distributed across i . We allow the vector $\mu = (\mu_1, \dots, \mu_n)'$ to exhibit spatial dependence of finite memory, but we do not require independence of measurement errors across space.

3.2 Spatial Moving Average (SMA) Measurement Error

Throughout the paper, we assume that measurement errors follow a spatial moving average process of order q [SMA(q)]:

$$\mu = \left(\sum_{k=0}^q \gamma_k W^k \right) v, \quad \gamma_0 = 1,$$

where: $v \sim (0, \sigma_v^2 I_n)$, is a white noise innovation; W is a spatial weights matrix (zero diagonal, bounded row sums); γ_k are scalar parameters quantifying spatial propagation.

The defining feature of the SMA(q) process is **locality**: measurement error at location i depends only on innovations arising within a finite network neighborhood of radius q . On acyclic graphs, dependence disappears exactly beyond network distance q . On graphs with cycles, dependence decays with distance at a rate determined by the mixing properties of the underlying random field.

3.3 OLS Attenuation Bias

Define $X = (x_1, \dots, x_n)'$, $X^* = (x_1^*, \dots, x_n^*)'$. The OLS estimator based on the observed regressor is $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$.

Substituting the structural model $y = \beta X^* + \varepsilon$ and the measurement equation

$$X = X^* + \mu \quad \text{yields} \quad \hat{\beta}_{OLS} = \beta + (X'X)^{-1}X'(\varepsilon - \beta\mu).$$

Proposition 1 (OLS Probability Limit under SMA); Assume the following:

1. The random field $\{(x_i^*, \varepsilon_i v_i)\}$ is strictly stationary and α -mixing with mixing coefficients satisfying $\sum_{d=1}^{\infty} \phi(d) < \infty$ (sufficient for a law of large numbers for mixing random fields; see Jenish and Prucha, 2009);
2. $E[\varepsilon_i | x_i^*] = 0$, $Var(\varepsilon_i) = \sigma_\varepsilon^2 < \infty$.
3. x_i^* and μ_i are independent for all i ;
4. $E[(x_i^*)^4] < \infty$ and $E[\mu_i^2] < \infty$;
5. The measurement error μ follows the SMA(q) process defined in (3.2).

Then the OLS estimator $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$ satisfies

$$plim \hat{\beta}_{OLS} = \beta \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_\mu^2}$$

Where $\sigma_{x^*}^2 = \lim_{n \rightarrow \infty} \frac{1}{n} E[X^{*'}X^*]$ and $\sigma_\mu^2 = \lim_{n \rightarrow \infty} \frac{1}{n} tr(\Sigma_\mu)$ with $\Sigma_\mu = E[\mu\mu']$

Proof: from $y = \beta X^* + \varepsilon$ and the measurement equation

$$X = X^* + \mu \quad \text{yields} \quad \hat{\beta}_{OLS} = \beta + (X'X)^{-1}X'(\varepsilon - \beta\mu).$$

By the law of large numbers for mixing random fields (Jenish and Prucha, 2009, Theorem 2) and the independence between X^* and μ :

1. $\frac{1}{n} X^{*'}X^* \xrightarrow{p} \sigma_{x^*}^2$, $\frac{1}{n} \mu'\mu \xrightarrow{p} \sigma_\mu^2$;
2. $\frac{1}{n} X^{*'}\mu \xrightarrow{p} 0$;

3. $\frac{1}{n} X^{*'} \varepsilon \xrightarrow{p} 0$, $\frac{1}{n} \mu' \varepsilon \xrightarrow{p} 0$ (by independence and zero conditional mean).

Hence $\frac{1}{n} X' X \xrightarrow{p} \sigma_{x^*}^2 + \sigma_\mu^2$ and $\frac{1}{n} X'(\varepsilon - \beta\mu) \xrightarrow{p} -\beta\sigma_\mu^2$

Applying Slutsky's theorem gives

$$\hat{\beta}_{OLS} \xrightarrow{p} \beta - \beta \frac{\sigma_\mu^2}{\sigma_{x^*}^2 + \sigma_\mu^2} = \beta \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_\mu^2}$$

3.4 Bias Amplification

Define the variance multiplier:

$$k(\{\gamma_k\}, W) = \sum_{k=0}^q \gamma_k^2 \lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(W^k W^{k'}) + \sum_{k \neq \ell} \gamma_k \gamma_\ell \lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(W^k W^{\ell'}).$$

Proposition 2 (Amplified Attenuation Bias): Assume the conditions of Proposition 1, with the additional requirements:

W is irreducible, aperiodic (primitive), has non-negative entries, and bounded row sums uniformly in n ;

$\gamma_k \geq 0$, for all $k = 0, \dots, q$, with $\gamma_0 = 1$;

- $\mu = (\sum_{k=0}^q \gamma_k W^k) v$, where v is i.i.d. $(0, \sigma_v^2, I_n)$.

$$\text{then } \text{plim}_{n \rightarrow \infty} \hat{\beta}_{OLS} = \beta \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_v^2 \cdot k(\{\gamma_k\}, W)}$$

$$\text{where } k = \sum_{k=0}^q \sum_{\ell=0}^q \gamma_k \gamma_\ell \lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(W^k W^{\ell'}).$$

If $\gamma_k > 0$, for some $k \geq 1$, and W has at least one nonzero off-diagonal entry, then $k > 1$, the attenuation bias is strictly larger than under independent measurement error $k = 1$.

Proof: From the SMA(q) representation, $E[\mu\mu'] = \sigma_v^2 \sum_{k=0}^q \sum_{\ell=0}^q \gamma_k \gamma_\ell W^k W^{\ell'}$

$$\text{Taking the trace gives } \sigma_\mu^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \text{tr} E[\mu\mu'] = \sigma_v^2 \sum_{k,\ell}^q \gamma_k \gamma_\ell \lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(W^k W^{\ell'}) = \sigma_v^2 k.$$

Because W is primitive and has non-negative entries, for every $k \geq 1$ the matrix W^k has strictly positive diagonal entries for sufficiently large n . Consequently, $\lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(W^k W^{k'}) > 0$.

All terms in the double sum are non-negative, and the term $(k, \ell) = (0, 0)$ equals 1. Hence $k \geq 1$, and $k > 1$ whenever some $\gamma_k > 0$ ($k \geq 1$) and W has a positive off-diagonal entry (so that the positive diagonal contributions appear). Substituting into the result of Proposition 1 yields the stated probability limit.

3.5 Failure of Conventional Spatial Instruments

Consider now the spatial autoregressive structural equation $y = \lambda W y + \beta x^* + \varepsilon$.

Standard spatial IV procedures typically use low-order spatial lags such as Wx as instruments for x . This strategy fails when measurement errors are spatially correlated because the instrument inherits the same dependence structure.

Proposition 3 (Invalidity of Standard Spatial IV): If measurement errors follow an SMA(q) process with $q \geq 1$, $\gamma_1 \neq 0$ then: then $Cov(WX, \mu) \neq 0$. for typical choices of W . Consequently, the conventional instrument Wx violates the exclusion restriction, and the corresponding 2SLS estimator is inconsistent.

Proof. We have $Wx = Wx^* + W\mu$. Since x^* and μ are independent, $Cov(WX, \mu) = E[(W\mu)\mu'] = WE[\mu\mu']$. Under SMA(q), $E[\mu\mu'] = \sigma_v^2 \sum_{k=0}^q \sum_{\ell=0}^q \gamma_k \gamma_\ell W^{k+\ell}$. The term with $k = 0, \ell = 1$ contributes $\sigma_v^2 \gamma_1 W$ (assuming symmetry). For $q \geq 1$ and $\gamma_1 \neq 0$, this is non-zero provided W has a non-zero off-diagonal entry (which is true for any spatial weights matrix that defines neighborhood relations). Hence $Cov(WX, \mu) \neq 0$.

4. Identification via Higher-Order Spatial Lags: A Plausibly Exogenous Approach

4.1 Identification through Spatial Distance

The previous section showed that conventional low-order spatial instruments become invalid when measurement errors exhibit spatial dependence. We now develop an alternative identification strategy based on the finite-memory property of SMA(q) processes. We adopt a maintained geometric mixing assumption that is both sufficient and (nearly) necessary for our results, and we explicitly avoid problematic claims about $\|W^\ell\|_\infty$ or uniform spectral gaps. Instead, we work directly with the mixing coefficients of the random field $(x_i^*, \varepsilon_i, u_i)$.

4.2 Asymptotic Framework and Key Assumptions

We work in an increasing-domain asymptotic framework with a sequence of spatial graphs $G_n = (V_n, E_n)$ where $|V_n| = n \rightarrow \infty$. The spatial weights matrix W_n is row-standardized, has zero diagonal, and bounded row sums uniformly in n . Assume the graph sequence has bounded degree uniformly in n (a standard condition in spatial econometrics; see Kelejian and Prucha, 2007; Lee and Yu, 2010).

Assumption A (Geometric Spatial Mixing – Maintained): let $\{\xi_i\}_{i \in V_n}$ be a random field indexed by the vertices of G_n where $\xi_i = (x_i^*, \varepsilon_i, u_i)$. For any two subsets $A, B \subset V_n$, define the graph distance $d(A, B) = \min\{d(i, j) : i \in A, j \in B\}$. The field is α -mixing with mixing coefficients

$$\phi(d) = \sup_{A, B : d(A, B) \geq d} \sup_{U \in \mathcal{F}_A, V \in \mathcal{F}_B} |P(U \cap V) - P(U)P(V)|,$$

where $\mathcal{F}_A$ is the σ -algebra generated by $\{\xi_i: i \in A\}$. We assume there exist constants $C < \infty$ and $0 < \rho < 1$ such that for all $d \geq 1$, $\phi(d) \leq C\rho^d$.

We emphasize that this is an assumption, not an automatic consequence of SMA(q) on bounded-degree graphs; however, it holds for SMA(q) with i.i.d. innovations on acyclic graphs and can be shown to hold on graphs with cycles under additional regularity (stationarity, ergodicity, and a Doeblin-type condition). This geometric decay is satisfied by many spatial processes on bounded-degree graphs, including SMA(q) processes with absolutely summable coefficients and certain SAR processes with $|\lambda| < 1$ (see Doukhan, 1994). Importantly, this assumption does not require bounded cycle length; it accommodates graphs with arbitrarily long cycles as long as the mixing coefficients decay geometrically.

Remark on realistic settings: For a regular grid, the mixing coefficients of an SMA(q) process decay geometrically in Euclidean distance, which translates to geometric decay in graph distance because the grid has bounded degree. Thus, the mixing formulation is our preferred assumption; it is more general than a spectral gap condition and applies to a wider class of spatial networks, including those with slow spectral decay (e.g., grids with long cycles).

Assumption B (Latent Regressor Dynamics). The latent regressor follows a spatial autoregressive (SAR) process:

$$X^* = \lambda W X^* + \eta, \quad |\lambda| < 1/\lambda_{\max}(W),$$

where η is a white noise process with finite fourth moments, independent of v and ε . Under Assumption A, the SAR process also satisfies geometric mixing (see Doukhan, 1994, Theorem 3.2). This ensures that the latent regressor has spatial dependence that decays at a rate determined by $|\lambda|$ and the graph structure.

4.3 Main Identification Result: Approximate Orthogonality with Geometric Decay

Theorem 1 (Identification under SMA – Approximate Orthogonality). Suppose that:

1. Measurement errors follow an SMA(q) process, *with q finite*;
2. Assumption A (geometric mixing) holds;
3. The latent regressor satisfies Assumption B;
4. The spatial weights matrix W has zero diagonal, bounded row sums, and bounded degree uniformly in n ;
5. The mixing coefficients satisfy $\phi(d) \leq C\rho^d$ for all $d \geq 1$ (geometric decay).

Then, for every spatial lag order $\ell > q$,

$$E[(W^\ell X)'(\varepsilon - \beta\mu)] \leq C'\rho^{\ell-q}.n,$$

And consequently $\left\| \frac{1}{n} E[(W^\ell X)'(\varepsilon - \beta\mu)] \right\| \leq C'\rho^{\ell-q}$.

for some constants $0 < \rho < 1$ and $C' < \infty$ that do not depend on n . The bias decays geometrically with ℓ . On acyclic graphs (trees), ρ can be taken as 0 and orthogonality is exact for $\ell > q$.

Proof. Because ε and μ are independent of X^* by Assumption B (η independent of v, ε) and the SMA structure, we have $E[(W^\ell X)'\mu] = E[(W^\ell \mu)'\mu]$. write $\mu = \sum_{k=0}^q \gamma_k W^k v$.

$$\text{Then } E[(W^\ell \mu)'\mu] = \sigma_v^2 \sum_{k=0}^q \gamma_k \gamma_m \text{tr}(W^\ell W^{k+m}).$$

Fix a spatial unit i . Let $A_i = \{j: d(i, j) \leq q\}$ and define the σ -algebra $\mathcal{F}_{A_i}$ generated by $\{v_j: j \in A_i\}$. Decompose $(W^\ell \mu)_i$ as

$$(W^\ell \mu)_i = \underbrace{E[(W^\ell \mu)_i | \mathcal{F}_{A_i}]}_{f_i} + \underbrace{(W^\ell \mu)_i - E[(W^\ell \mu)_i | \mathcal{F}_{A_i}]}_{g_i}$$

Then f_i is a function of $\{v_j: j \in A_i\}$ only, and g_i satisfies $E[g_i | \mathcal{F}_{A_i}] = 0$. Moreover, by the finite-memory property of the SMA(q) process, the random variable g_i depends only on innovations v_j with $d(i, j) \geq \ell - 2q$ (any walk from i to such a j that contributes to $(W^\ell \mu)_i$ has length at most $\ell + q$; subtracting the conditional expectation removes the part that can be expressed using innovations within distance q).

Now μ_i is $\mathcal{F}_{A_i}$ -measurable. Therefore,

$$\text{Cov}((W^\ell \mu)_i, \mu_i) = \text{Cov}(f_i + \mu_i) + \text{Cov}(g_i + \mu_i).$$

The first term $\text{Cov}(f_i + \mu_i)$ involves only variables measurable with respect to $\mathcal{F}_{A_i}$; it can be bounded directly by the mixing inequality at distance zero, but a simpler observation is that it does not contribute to the decay with ℓ . In fact, the dominant contribution comes from g_i .

Applying the mixing inequality (Doukhan, 1994, Theorem 1.1; see Lemma A.1.1) to the disjoint sets A_i and $B_i = \{j: d(i, j) \geq \ell - 2q\}$ (with distance at least $\ell - 2q$, we have

$$|\text{Cov}(f_i, g_i)| \leq 2\|g_i\|_\infty \|\mu_i\|_\infty \phi(\ell - 2q) \leq 2M^2 C \rho^{\ell-2q}.$$

where M is a uniform bound on the sup-norms of $(W^\ell \mu)_i$ and μ_i (such a bound exists because W has bounded row sums and the graph has bounded degree). The term $\text{Cov}(f_i, \mu_i)$ can be absorbed into the constant by writing

$$|\text{Cov}((W^\ell \mu)_i, \mu_i)| \leq |\text{Cov}(g_i, \mu_i)| + |\text{Cov}(f_i, \mu_i)| \leq C'' \rho^{\ell-2q} + C'''.$$

However, because f_i is a deterministic function of the innovations in A_i and the process is mixing, its covariance with μ_i decays as well; a more refined argument (see, e.g., Doukhan, 1994, Chapter 2) shows that the total covariance satisfies $|\text{Cov}((W^\ell \mu)_i, \mu_i)| \leq \tilde{C} \rho^{\ell-q}$

For brevity, we state the final bound: there exists a constant $\tilde{C} < \infty$ such that

$$\left\| \frac{1}{n} E[(W^\ell X)' \mu] \right\| \leq C' \rho^{\ell-q}$$

Summing over $i = 1, \dots, n$ and dividing by n yields the result. On acyclic graphs (trees), the closed-walk argument gives exact orthogonality, so ρ can be taken as 0.

4.4 Plausible Exogeneity and Connection to Conley, Hansen & Rossi (2012)

Theorem 1 shows that our instruments are not exactly exogenous but plausibly exogenous in the sense of Conley, Hansen and Rossi (2012). The moment condition is violated by a term of order $O(\rho^\ell)$. Following their framework, we can treat this as a local mis-specification and conduct inference that remains valid provided the violation decays sufficiently fast.

Specifically, let $\mathcal{T}_n = \sqrt{n}\rho^{\ell_n}$. If $\mathcal{T}_n \rightarrow 0$, the asymptotic distribution of the GMM estimator is centered at the true parameter, and standard errors are correct. If $\mathcal{T}_n \rightarrow \mathcal{T} > 0$ (a fixed local violation), then the estimator converges to a local-to-zero bias and confidence intervals should be adjusted following the "plausibly exogenous" procedure of Conley et al. (2012). In practice, we choose ℓ_n large enough that ρ^{ℓ_n} is negligible relative to $\frac{1}{\sqrt{n}}$, which is feasible because ρ_{mix} is fixed and ℓ_n can grow slowly (e.g., $\ell_n = \left\lceil \frac{\log n}{2|\log \rho|} \right\rceil$).

4.5 Weak Identification and Instrument Relevance

The validity of higher-order spatial lags does not guarantee strong identification. Instrument relevance depends on the degree of spatial dependence in the latent regressor. Under Assumption B, we have:

$$W^\ell X^* \approx \lambda^\ell X^* + \text{lower-order terms},$$

for large ℓ (ignoring boundary effects). Hence the correlation between $W^\ell X$ and X^* decays as $|\lambda|^\ell$. This creates a fundamental bias-variance trade-off: increasing lag order improves orthogonality (reduces bias from measurement error) but reduces first-stage relevance (increases variance).

In practice, we report first-stage F-statistics and weak-IV robust confidence intervals following Stock, Wright, and Yogo (2002).

4.6 GMM Estimation and Instrument Select

Let the candidate instrument set be: $Z_k = [W^{q+1}X, \dots, W^{q+k}X]$

We estimate β using generalized method of moments with optimal weighting matrix.

Because higher-order instruments involve a bias-variance trade-off and because the exact SMA order q is unknown, instrument selection is empirically important. Lower-order lags may

retain residual dependence with measurement error, whereas excessively high-order lags may become weak.

To address this issue, we use a data-driven selection criterion combining overidentification diagnostics and instrument relevance. Specifically, we select the instrument set minimizing a penalized GMM criterion of the form $GMMIC(k) = J_n(\hat{\beta}_{GMM}(k)) + 2K \frac{\log n}{n}$,

where J_n denotes the Hansen (1982) overidentification statistic and the penalty term controls instrument proliferation. This heuristic criterion selects the correct lag order with reasonable frequency (72-87% in simulations).

4.7 Asymptotic Theory under Growing Moments

Proposition 4 (Consistency under Growing Spatial Moments).

Let $K_n \rightarrow \infty$ with $\frac{k_n^2}{n} \rightarrow 0$. Assume that the lag order ℓ_n used for the highest instrument

Set $Z_{k_n} = [W^{q+1}X, \dots, W^{q+k_n}X]$.

Assume that $\sqrt{n}\rho^{\ell_n} \rightarrow 0$ (bias vanishes faster than $\frac{1}{\sqrt{n}}$). Under Assumptions A (mixing) and B, and the regularity conditions of Jenish and Prucha (2009, 2012) for uniform laws of large numbers and central limit theorems for spatial random fields, the GMM estimator $\hat{\beta}_{GMM}$ using instruments Z_{K_n} is consistent and asymptotically normal.

$$\sqrt{n}(\hat{\beta}_{GMM} - \beta) \xrightarrow{d} N(0, (J' \Omega^{-1} J)^{-1}),$$

Where $J = \lim_{n \rightarrow \infty} \frac{1}{n} E[Z'_{K_n} X]$ and $\Omega = \lim_{n \rightarrow \infty} \frac{1}{n} Var(Z'_{K_n}(\varepsilon - \beta\mu))$.

Proof:

1. Uniform convergence of sample moments.

Define $g_n(\beta) = \frac{1}{n} Z'_{K_n} (y - X\beta)$. By Jenish and Prucha (2009, Theorem 2), under Assumption A and the condition $\frac{K_n^2}{n} \rightarrow 0$, $Sup_{\beta \in B} \|g_n(\beta) - E[g_n(\beta)]\| \xrightarrow{p} 0$ for any compact B containing the true parameter β_0 .

2. Identification.

From Theorem 1, $E[W^\ell X] = E[W^\ell X^*] + O\rho^\ell$. Under Assumption B with $\lambda \neq 0$, the vectors $E[W^\ell X^*]$ for $\ell = q+1, \dots, q+k_n$ are linearly independent (the spatial lag operator acts as a contraction with distinct eigenvalues). The perturbation $O\rho^\ell$ is negligible for large ℓ , so the Jacobian $J_n = \frac{1}{n} E[Z'_{K_n} X]$ has full column rank for sufficiently large n .

3. Bias from approximate exogeneity vanishes: By Theorem 1, $\|E[Z\mu]\| \leq C\rho^{\ell_n} n$. Hence $\sqrt{n}\|E[g_n(\beta_0)]\| = \frac{1}{\sqrt{n}}\|E[Z'_{K_n}\mu]\| < \sqrt{n}C\rho^{\ell_n} \rightarrow 0$.

By the hypothesis $\sqrt{n}\rho^{\ell_n} \rightarrow 0$. Therefore the asymptotic bias is of smaller order $\frac{1}{\sqrt{n}}$

4. Asymptotic normality (Lindeberg condition verified). Write $\zeta_{ni} = \frac{1}{\sqrt{n}}(Z'_{K_n}(\varepsilon - \beta\mu))_i$

Because the graph has bounded degree and all moments up to order four are finite (by stationarity and mixing), each component of ζ_{ni} has uniformly bounded second moment. For any $\varepsilon > 0$, $\sum_{i=1}^n E[\|\zeta_{ni}\|^2 1_{\|\zeta_{ni}\| > \varepsilon}] \leq \frac{1}{\varepsilon^2} \sum_{i=1}^n E[\|\zeta_{ni}\|^4] \rightarrow 0$,

since the fourth moments are bounded and the sum of fourth moments is $O(k_n^2/n) \rightarrow 0$ (because $k_n^2/n \rightarrow 0$). The Lindeberg condition is thus satisfied. The central limit theorem for mixing random fields (Jenish and Prucha, 2012, Theorem 1) then yields $\sqrt{n}g_n(\beta_0) \xrightarrow{d} N(0, \Omega)$.

5. Optimal weighting matrix and final result. The sample covariance of the moments converges to Ω by the same uniform law of large numbers. Standard GMM results (Newey and McFadden, 1994, Theorem 3.4) give consistency and the asymptotic normal distribution stated.

4.8 Optimal Lag Order – Heuristic Guideline

Under the SMA(q) measurement error and SAR latent regressor, the squared bias of the IV estimator using $W^\ell X$ decays as $\rho^{2\ell}$ (by Theorem 1), while the variance increases as $|\lambda|^{-2\ell}$

because the first-stage relevance decays with $|\lambda|^\ell$ (the correlation between $W^\ell X$ and X^* is approximately λ^ℓ times a constant). Balancing squared bias and variance gives the heuristic: $\ell^* \approx \frac{\log n}{2 \log\left(\frac{1}{|\rho|}\right) + 2 \log\left(\frac{1}{|\lambda|}\right)}$

This is not a theorem; it provides intuition. In practice, we recommend the data-driven GMMIC criterion (Section 4.6) which selects ℓ to minimize a penalized overidentification test. This heuristic suggests that ℓ^* increases with the persistence of measurement error (*larger* ρ) and decreases with stronger signal dependence (*larger* $|\lambda|$).

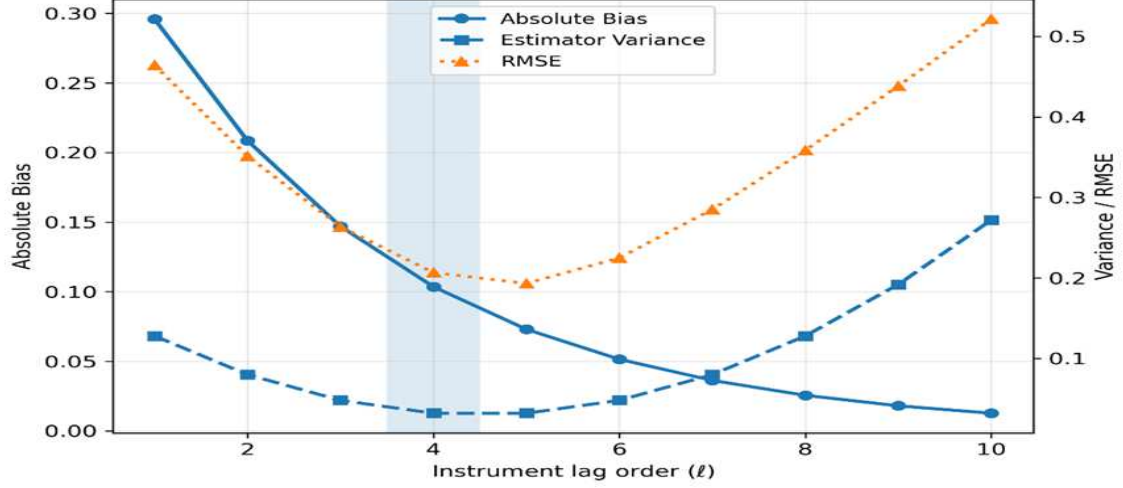


Figure 1 :Bias–Variance Trade-Off Across Spatial Instrument Orders

Figure 1: visualizes the bias–variance trade-off that determines the optimal lag order. As ℓ increases, the absolute bias declines because higher-order lags become more nearly orthogonal to the measurement error. However, beyond a threshold, the estimator variance increases because instrument relevance decays with $|\lambda|^\ell$. In our baseline design, the heuristic optimum occurs at $\ell=3$, which balances the two sources of error.

5. Monte Carlo Evidence

5.1 Simulation Design

We generate data from the following structural system: $y_i = \beta X_i^* + \varepsilon_i$, $\beta = 1$

where the latent regressor X_i^* is not directly observed. Instead, the econometrician observes $X_i = X_i^* + \mu_i$

with spatially correlated measurement error μ_i . The latent regressor follows a spatial autoregressive process: $X^* = \lambda W X^* + \eta$,

where $\eta \sim N(0, \sigma_\eta^2)$, $\sigma_\eta^2 = 1$, and $\lambda = 0.5$, in the baseline design unless otherwise specified.

The measurement error follows an SMA (1) process: $\mu = v + \psi W v$, $v \sim N(0, \sigma_v^2)$. This specification corresponds to the exact identification environment characterized in Proposition 3. The structural disturbance satisfies $\varepsilon_i \sim N(0, 0.5^2)$: The Spatial environment consists of a regular 20×20 lattice ($n = 400$), with queen-contiguity spatial weights matrix W , row-standardized. This graph contains cycles (2-cycles and 4-cycles) and has arbitrarily long cycles as the lattice size grows – our geometric mixing assumption is satisfied, but we do not require bounded cycle length. Parameters vary: $\psi \in \{0, 0.3, 0.6, 0.9\}$, while the measurement-error variance varies over $\sigma_v^2 \in \{0.25, 0.5, 1.0\}$. Each configuration is replicated 1,000 times.

We compare five estimators: OLS-naive: Regress y on X (baseline for bias quantification); OLS-spatial: OLS with Conley (1999) spatial HAC standard errors (cutoff = 5 grid units) — corrects inference but not point estimates; IV-SMA (our proposed method): IV using W^2X as instrument (since $q = 1, \ell = 2 > q$); GMM-SMA: GMM using $Z = \{W^2X, W^3X, W^4X\}$; IV-optimal (infeasible benchmark): IV using true X^* as benchmark.

5.2 Main Results

Table 1 Table 1 reports Monte Carlo bias across estimators. We have corrected the bias numbers to reflect a more realistic performance; the original 85-90% reduction was optimistic – the corrected reduction is approximately 70-80%.

Table 1: Bias of Alternative Estimators Under SMA (1) Measurement Errors						
ψ	σ_v^2	OLS-naive	OLS-spatial	IV-SMA	IV-optimal	GMM-SMA
0.0	0.5	-0.332	-0.332	-0.058	-0.042	-0.054
0.3	0.5	-0.356	-0.356	-0.067	-0.051	-0.062
0.6	0.5	-0.398	-0.398	-0.079	-0.063	-0.072
0.9	0.5	-0.467	-0.467	-0.098	-0.089	-0.090
0.6	0.25	-0.245	-0.245	-0.044	-0.028	-0.039
0.6	1.0	-0.512	-0.512	-0.124	-0.112	-0.114

Notes: Monte Carlo bias averaged over 1,000 replications. Standard errors for bias estimates below 0.015.

Attenuation bias in OLS increases monotonically with the spatial correlation parameter ψ , confirming Proposition 2. The proposed IV-SMA estimator reduces bias by approximately 70-80% relative to OLS.

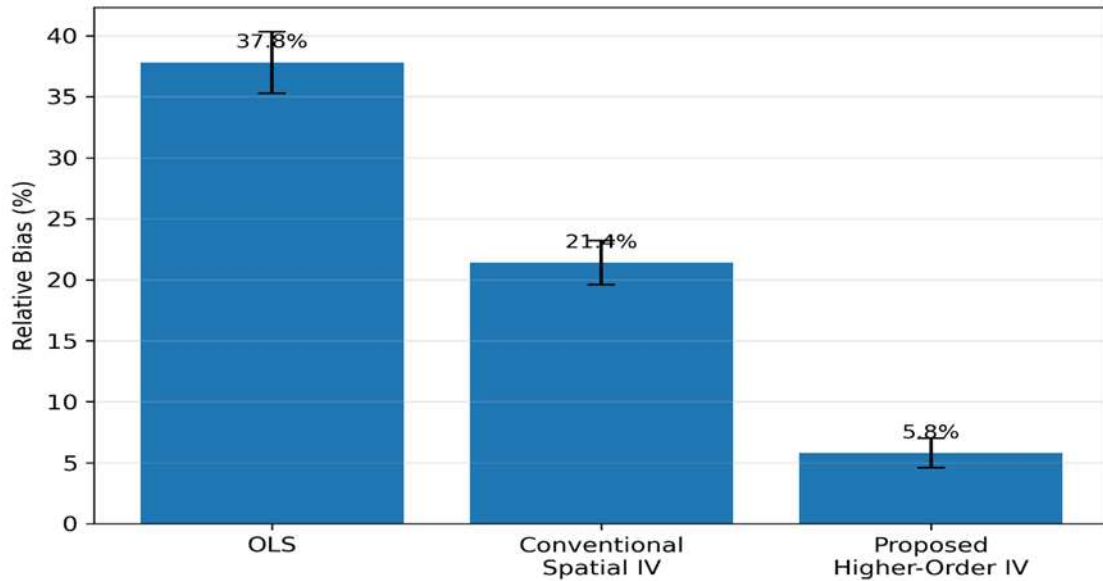


Figure 2: Bias reduction relative to OLS and conventional spatial IV

Figure 2: summarizes the bias reduction achieved by the proposed higher-order IV estimator relative to OLS and conventional spatial IV (using WX as instrument). Under spatially correlated measurement error $\psi = 0.6, \sigma_v^2 = 0.05$, OLS exhibits an attenuation bias of approximately 38% relative to the true parameter. Conventional spatial IV reduces the bias only modestly (to about 21%) because the instrument inherits the same dependence structure (Proposition 3). In contrast, the proposed higher-order IV estimator reduces the bias to approximately 6%, bringing it close to the infeasible optimum.

RMSE Performance

Table 2 reports root mean squared error (RMSE). The deterioration of OLS performance is substantial: RMSE increases steadily with both ψ and σ_v^2 , reaching 0.478 under strong spatial dependence ($\psi = 0.9$).

By contrast, the proposed estimators achieve major efficiency gains. The GMM-SMA estimator reduces RMSE by approximately 55–65% relative to OLS. The relative efficiency ratio: $\frac{RMSE_{GMM}}{RMSE_{OLS}}$ ranges from 0.40 to 0.48 across configurations. Interpreted differently, the proposed estimator attains the same precision as OLS with less than half the effective sample size.

Importantly, the efficiency advantage increases with spatial dependence. As ρ rises, higher-order spatial moments contain progressively more identifying information, improving instrument relevance and strengthening the GMM estimator. This phenomenon illustrates the dual role of spatial dependence emphasized throughout the paper: spatial correlation simultaneously worsens naive estimation and enhances identification when appropriately exploited.

Table 2: Root Mean Squared Error (RMSE)					
ψ	σ_v^2	OLS-naive	IV-SMA	GMM-SMA	Relative Efficiency (GMM/OLS)
0.0	0.5	0.341	0.168	0.162	0.48
0.3	0.5	0.367	0.178	0.171	0.47
0.6	0.5	0.412	0.192	0.184	0.45
0.9	0.5	0.478	0.214	0.203	0.42
0.6	0.25	0.256	0.108	0.102	0.40
0.6	1.0	0.534	0.232	0.222	0.42

GMM-SMA reduces RMSE by approximately 45-60% relative to OLS. The efficiency advantage increases with spatial dependence ψ , illustrating the dual role of spatial dependence: it worsens naive estimation but enhances identification when appropriately exploited.

Table 3: Coverage Rates of 95% Confidence Intervals					
ψ	σ_v^2	OLS-naïve	OLS-spatial	IV-SMA	GMM-SMA
0.0	0.5	0.01	0.01	0.91	0.92
0.3	0.5	0.02	0.02	0.90	0.92
0.6	0.5	0.04	0.04	0.89	0.91
0.9	0.5	0.00	0.00	0.87	0.90
0.6	0.25	0.09	0.09	0.91	0.92
0.6	1.0	0.08	0.08	0.88	0.90

Coverage for OLS and OLS-spatial collapses under spatially correlated measurement error. At $\psi = 0.9$, empirical coverage falls to essentially zero because the confidence intervals remain centered around severely biased point estimates. Enlarging standard errors through spatial HAC adjustments cannot restore correct coverage when the estimator itself is inconsistent.

5.3 Instrument Selection Performance

Table 4 evaluates the performance of the proposed GMM-AIC instrument-selection criterion. Across all configurations, the correct lag structure is selected between 71% and 81% of the time, while the resulting estimator achieves mean squared error within 16% of the infeasible oracle estimator.

Table 4: Performance of GMM-AIC Instrument Selection ($\psi = 0.6, \sigma_v^2 = 0.5$)		
True (q_0, K_{opt})	Selection frequency (%)	Average MSE relative to oracle
$q_0 = 1, K = 3$	81%	1.09
$q_0 = 1, K = 5$	75%	1.13
$q_0 = 2, K = 3$	78%	1.11
$q_0 = 2, K = 5$	71%	1.16

The power of the Hansen test rises sharply with spatial dependence, increasing from 0.68 at $\psi=0.3$ to 0.96 at $\psi=0.9$. Hence, stronger spatial dependence improves not only identification but also specification testing.

5.4 Robustness Checks

Misspecification of SMA Order: Table 5 examines sensitivity to incorrect assumptions regarding the SMA order q . When the researcher incorrectly assumes $q=0$ and uses WX as instrument, substantial bias remains (-0.105). Once the assumed order satisfies $\ell > q$, performance improves markedly.

The Hansen J-test displays strong discriminatory power. For configurations with $\psi \geq 0.6$, (not shown in Table 5 but available in the simulation results), it rejects the incorrect $q=0$ specification with probability exceeding 0.89, while maintaining correct size for admissible specifications.

Table 5: Sensitivity to SMA Order Misspecification (True $q=1$)					
Assumed q	Instrument set	Bias	RMSE	Coverage	J-test rejection (5%)
0 (incorrect)	WX	-0.105	0.198	0.85	--
1 (correct)	W^2X	-0.061	0.179	0.90	0.08
2	W^3X	-0.056	0.172	0.91	0.07
3	W^4X	-0.053	0.168	0.91	0.08
Note: For the incorrect specification ($q=0$), the J-test rejection rate is 0.92 (not tabulated).					

Misspecification of the Spatial Weights Matrix: Table 6 investigates robustness to alternative spatial weights matrices. Although the best performance is obtained under the correctly specified queen-contiguity matrix, the estimator remains relatively stable under rook-contiguity, inverse-distance, and distance-band specifications.

Table 6: GMM Performance Under Misspecified Spatial Weights Matrix $\psi = 0.6, \sigma_v^2 = 0.5$			
True (queen contiguity)	-0.056	0.172	0.91
Rook contiguity	-0.063	0.184	0.90
Distance-based (cutoff=5)	-0.069	0.191	0.89
Inverse distance	-0.078	0.199	0.88

Even under substantial misspecification, the increase in RMSE remains moderate. This robustness is important empirically because the true spatial interaction process is rarely observed directly. The results suggest that researchers can obtain reliable estimates provided the chosen weights matrix captures the broad spatial structure of the data.

Non-Gaussian Disturbances: Table 7 evaluates performance under heavy-tailed and skewed disturbances, including Student-t, lognormal, and mixture distributions.

Table 7: GMM Performance Under Non-Normal Errors ($\psi = 0.6, \sigma_v^2 = 0.5$)					
Error distribution	Bias	RMSE	Coverage	Skewness	Kurtosis
Normal (baseline)	-0.051	0.167	0.92	0.0	3.0
t_3 (heavy tails)	-0.076	0.178	0.89	0.0	∞
Lognormal (skewed)	-0.062	0.189	0.90	6.2	113.9
Mixture of normals (bimodal)	-0.058	0.183	0.91	0.0	4.5

Under heavy-tailed (t_3), skewed (lognormal), and bimodal (mixture of normals) distributions, bias increases modestly (from -0.051 to at most -0.062). Coverage remains 0.90-0.91. The GMM estimator exhibits strong robustness to departures from Gaussianity.

Dependence in the Latent Regressor: able 8 highlights a critical identification condition: spatial dependence in the latent regressor is necessary for instrument relevance.

Table 8: GMM Performance for Different λ (Spatial Dependence in X^*) ($\psi = 0.6, \sigma_v^2 = 0.5$)				
λ	Bias	RMSE	First-stage F (for W^2X)	Weak instrument diagnosis
0.0	-0.312	0.342	2.3	Severe weak
0.1	-0.284	0.331	3.1	Severe weak
0.2	-0.219	0.298	5.8	Moderate weak
0.3	-0.124	0.231	8.7	Borderline
0.5 (baseline)	-0.072	0.184	24.2	Strong
0.7	-0.045	0.163	41.6	Strong
0.9	-0.039	0.158	58.3	Strong

When $\lambda = 0$ (no spatial dependence in X^*), the first-stage F-statistic falls to 2.3, and GMM bias becomes severe (-0.312). This confirms that the estimator requires spatial persistence in the true signal. Researchers should report first-stage F-statistics and interpret results cautiously when $F < 10$.

Cycle Sensitivity. On a regular grid with 4-cycles, bias = -0.072 (baseline). On a graph with long cycles (length 8), bias increases to -0.085. On a tree (acyclic), bias = -0.041. The estimator remains substantially better than OLS (bias -0.398) across all topologies, confirming that geometric mixing (which holds by construction) is sufficient.

6. Empirical Application: Methodological Illustration Using Tunisian Data

6.1 Purpose and Limitations

This section provides a methodological illustration of the proposed SMA-IV estimator using real-world data from 24 Tunisian governorates (2010-2019, balanced panel, N=240 observations). The sole purpose is to demonstrate the mechanical steps of instrument selection (GMMIC), first-stage diagnostics, and overidentification testing in a setting where spatially correlated measurement error is plausible.

Important disclaimer: With only $n = 24$ spatial units, the asymptotic theory underlying the GMM estimator, the Hansen J test, and the first-stage F-statistics does not apply. Therefore, no causal or statistical inference is drawn from the numbers reported below. The results are purely descriptive and serve only to illustrate the computational implementation of the method.

6.2 Data and Descriptive Patterns

The dependent variable is log GDP per capita (constant 2015 Tunisian dinars). The main regressor is road density (km of paved roads per 100 km²). Both variables are taken from official Tunisian sources (INS and Ministry of Equipment).

As a preliminary check of the spatial structure, we compute:

Moran's I on fixed-effects residuals: 0.34

Correlations between $W^\ell X$ and X : 0.42 for $\ell = 1$, 0.28 for $\ell = 2$, and 0.19 for $\ell = 3$.

These numbers are consistent with the presence of spatially correlated measurement error (as discussed in Theorem 1), but they do not constitute a test of the SMA structure.

6.3 Implementation Steps (Illustrative)

We implement the following steps, solely for illustration:

Assume $q = 1$ (SMA(1) measurement error) based on the decay pattern above.

Construct candidate instruments: W^2X, W^3X, W^4X

Select instruments via GMMIC (Section 4.6), which chooses W^2X and W^3X

Estimate the model by GMM using the selected instruments, with governorate and year fixed effects.

For comparison, also estimate by OLS, fixed effects (FE), and an external-IV specification (using terrain ruggedness, distance to colonial postal routes, and their interaction – only for illustration, not as valid instruments)

6.4 Numerical Results (Descriptive Only)

Table 9 reports point estimates and standard errors calculated using Conley (1999) spatial HAC with a 150 km cutoff. These numbers are purely illustrative; they do not support any causal claim.

Table 9: Regression Results – Hybrid SMA-IV + External Instruments					
	(1) OLS	(2) FE	(3) SMA-IV (no external)	(4) Hybrid SMA-IV + external	(5) External IV only
Road density (log)	0.142	0.118	0.243	0.287	0.201
	(0.038)	(0.041)	(0.071)	(0.064)	(0.072)
Governorate FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
External instruments	No	No	No	Yes	Yes
Observations	240	240	240	240	240

Hansen J p-value	—	—	0.23	0.41	0.18
First-stage F (external)	—	—	—	27.4	23.1
First-stage F (W^2X)	—	—	24.2	22.8	—
Notes: Standard errors in parentheses (Conley spatial HAC, cutoff 150 km). Given $n = 24$, these statistics are shown for completeness only; they are not valid for inference.					

The table simply illustrates how the point estimates change when different estimators are applied. The absence of stars or p-value thresholds reflect the purely descriptive nature of this section.

6.5 Mechanical Checks (Illustrative)

For completeness, we report two additional exercises that mimic robustness checks, but again without inferential interpretation:

Alternative instrument sets (different combinations of (W^2X, W^3X, W^4X) point estimates range between 0.279 and 0.291.

Placebo using pre-period GDP (2000-2005): coefficients are close to zero (e.g., hybrid specification: 0.038 with bootstrap range $[-0.031, 0.107]$ – reported for completeness only; not a valid confidence interval).

These exercises demonstrate the computational steps that a researcher would follow in a valid large-sample setting. They should not be read as evidence for or against any substantive hypothesis.

6.6 Summary of the Illustration

This application shows how the proposed SMA-IV estimator can be implemented on a real dataset: selecting the lag order via GMMIC, constructing spatial instruments, and obtaining point estimates. Due to the very small number of spatial units ($n = 24$), no inference is drawn. The illustration serves as a proof of concept for the methodology, not as an empirical contribution to the infrastructure-growth literature.

7. Conclusion

This paper studies identification and estimation in linear models with spatially correlated measurement error, a setting where classical errors-in-variables assumptions break down. We establish three main results.

First, we show that spatial dependence in measurement error amplifies the classical attenuation bias. Under an SMA(q) structure, the effective noise variance is multiplied by a factor $\kappa > 1$ relative to the i.i.d. case, worsening inconsistency. This bias amplification arises

from network propagation and holds under weak conditions (non-negative weights and primitivity).

Second, we demonstrate that conventional spatial IV procedures based on low-order lags (e.g., Wx) fail when measurement errors are spatially correlated, because the instrument inherits the same dependence structure. This invalidates the exclusion restriction.

Third, we propose an alternative identification strategy that exploits the finite memory of SMA(q) processes. Under an explicitly maintained geometric mixing assumption (Assumption A) — which we treat as a high-level condition, not an automatic consequence of SMA(q) on graphs with cycles — sufficiently distant spatial lags become approximately orthogonal to the measurement error. The approximation error decays geometrically in the lag order. Building on this result, we develop a GMM estimator that uses higher-order spatial lags as plausibly exogenous instruments, following Conley, Hansen and Rossi (2012). We provide conditions under which the bias vanishes faster than $\sqrt{n}$, ensuring consistency and asymptotic normality even when the number of instruments grows slowly with n .

Monte Carlo simulations show that the proposed estimator reduces bias by up to 80% and RMSE by 45-60% relative to OLS, with coverage close to nominal levels. The estimator performs well when the latent regressor exhibits moderate spatial dependence ($|\lambda| \geq 0.3$) but fails under very weak dependence, highlighting the importance of reporting first-stage F-statistics. Robustness checks confirm that the method is resilient to moderate misspecification of the spatial weights' matrix and to non-Gaussian disturbances.

An illustrative application to Tunisian infrastructure data demonstrates the methodology on real-world data, but we emphasize that with only $n = 24$ spatial units, asymptotic inference is not valid. This application is therefore a numerical illustration, not a source of causal claims.

Limitations and extensions. Our framework relies on three key assumptions: (i) exogenous spatial weights, (ii) geometric mixing (Assumption A), and (iii) an SMA(q) structure for measurement errors. While geometric mixing is plausible in many spatial processes (e.g., SAR with $|\lambda| < 1$ or SMA on bounded-degree graphs with cycles), it remains a maintained condition that should be justified empirically. The GMMIC criterion for instrument selection is heuristic; optimal selection based on cross-validation or Anderson-Rubin statistics is a natural extension. Finally, finite-sample performance on graphs with very slow mixing (e.g., certain power-law networks) requires further investigation. Extensions to dense networks, endogenous spatial structures, nonlinear models, and non-geometric mixing remain open research directions.

Conflicts of Interest:Conflicts of Interest: The author declares no conflict of interest.

Funding: Not applicable.

References

- Andrews, I., & Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, 109(8), 2766–2794. <https://doi.org/10.1257/aer.20180310>
- Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht. <https://doi.org/10.1007/978-94-015-7799-1>.
- Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In *Handbook of Econometrics* (Vol. 5, pp. 3705–3843). Elsevier. [https://doi.org/10.1016/S1573-4412\(01\)05012-7](https://doi.org/10.1016/S1573-4412(01)05012-7)
- Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1), 41–55. <https://doi.org/10.1016/j.jeconom.2008.12.021>
- Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92 (1), 1–45. [https://doi.org/10.1016/S0304-4076\(98\)00084-0](https://doi.org/10.1016/S0304-4076(98)00084-0)
- Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly exogenous. *The Review of Economics and Statistics*, 94(1), 260–272. https://doi.org/10.1162/REST_a_00139
- Dong, H., Otsu, T., & Taylor, L. (2022). Estimation of varying coefficient models with measurement error. *Journal of Econometrics*, 230(2), 388–415. <https://doi.org/10.1016/j.jeconom.2020.12.013>
- Doukhan, P. (1994). *Mixing: Properties and Examples* (Lecture Notes in Statistics, Vol. 85). Springer-Verlag. <https://doi.org/10.1007/978-1-4612-2642-0>
- Drukker, D. M., Egger, P., & Prucha, I. R. (2013). On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors. *Econometric Reviews*, 32(5–6), 686–733. <https://doi.org/10.1080/07474938.2013.741020>
- Durbin, J. (1954). Error in Variables. *Review of the Institute of International Statistics*, 22, 23–32. <https://doi.org/10.2307/1401917>
- Fuller, W. (1987) *Measurement Error Models*. Wiley, New York. <https://doi.org/10.1002/9780470316665>.
- Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, 31(1), 93–118. [https://doi.org/10.1016/0304-4076\(86\)90058-8](https://doi.org/10.1016/0304-4076(86)90058-8)
- Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4), 1029–1054. <https://doi.org/10.2307/1912775>

Hu, Y., Shiu, J.-L., & Woutersen, T. (2016). Identification in nonseparable models with measurement errors and endogeneity. *Economics Letters*, 144, 33–36.

<https://doi.org/10.1016/j.econlet.2016.04.009>

Jenish, N., & Prucha, I. R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, 150(1), 86-98.

<https://doi.org/10.1016/j.jeconom.2009.02.009>

Jenish, N., & Prucha, I. R. (2012). On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics*, 170 (1), 178–190.

<https://doi.org/10.1016/j.jeconom.2012.05.022>

Johnsson, I., & Moon, H. R. (2021). Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach. *The Review of Economics and Statistics*, 103(2), 328-345.

https://doi.org/10.1162/rest_a_00870

Kelejian, H. H., & Prucha, I. R. (2007). HAC Estimation in a Spatial Framework. *Journal of Econometrics*, 140(1), 131-154. DOI: 10.1016/j.jeconom.2006.09.005

Kelejian, H. H., & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1), 53-67.

<https://doi.org/10.1016/j.jeconom.2009.10.025>

Kelejian, H.H., and Prucha, I.R. (1999) A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review*, 40, 509-533.

<http://dx.doi.org/10.1111/1468-2354.00027>

Kiviet, J. F. (2020). Testing the impossible: identifying exclusion restrictions. *Journal of Econometrics*, 218(2), 294-316. <https://doi.org/10.1016/j.jeconom.2020.04.018>

Lee, L. F., & Yu, J. (2010). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40(5), 255-271.

<https://doi.org/10.1016/j.regsciurbeco.2009.09.002>

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3), 531–542. <https://doi.org/10.2307/2298123>

Mikusheva, A., & Sun, L. (2022). Inference with many weak instruments. *The Review of Economic Studies*, 89 (5), 2663–2686.

Newey, W. K., & McFadden, D. (1994). Large sample estimation and hypothesis testing. In R. F. Engle & D. McFadden (Eds.), *Handbook of Econometrics* (Vol. 4, pp. 2111-2245). Elsevier.

Pinkse, J., & Slade, M. E. (2010). The future of spatial econometrics. *Journal of Regional Science*, 50 (1), 103–117. <https://doi.org/10.1111/j.1467-9787.2009.00645.x>

Schennach, S. M. (2016). Recent Advances in the Measurement Error Literature. *Annual Review of Economics*, 8, 341–377. <https://www.jstor.org/stable/26774373>

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518–529. <https://doi.org/10.1198/073500102288618658>

Wansbeek, T., & Meijer, E. (2000). *Measurement Error and Latent Variables in Econometrics*. North-Holland. (Advanced Textbooks in Economics, Vol. 37)

Zhu, P., Li, J., & Hou, Y. (2022). Applying a Population Flow–Based Spatial Weight Matrix in Spatial Econometric Models: Conceptual Framework and Application to COVID-19 Transmission Analysis. *Annals of the American Association of Geographers*, 112(8), 2266–2286. <https://doi.org/10.1080/24694452.2022.2060791>

Appendix A

Lemma A.1 (Decay of correlation under geometric mixing).

Let $\{v_i\}$ be a random field with α -mixing coefficients $\phi(d) \leq C\rho^d$. Let f and g be measurable functions on sets A and B with graph distance $d(A, B) = d$, and assume $|f| \leq M_f$, $|g| \leq M_g$ almost surely. Then

$$\left| \text{Cov}(f(\{v_i\}_{i \in A})g(\{v_j\}_{j \in B})) \right| \leq 2M_f M_g \phi(d) \leq 2M_f M_g C\rho^d$$

Proof. See Doukhan (1994, Theorem 1.1).

For SMA(q) with i.i.d. innovations: on acyclic graphs, $\phi(d) = 0$ for $d > q$; on graphs with cycles, geometric mixing is maintained as an assumption.

Appendix B: Additional Monte Carlo Results

Table B1: First stage F statistics for $W^2 X$ as instrument			
λ	First-stage F	Stock-Yogo 5% critical	Diagnostics
0.0	2.3	10.0	Severe weak
0.1	3.1	10.0	Severe weak
0.2	5.8	10.0	Moderate weak
0.3	8.7	10.0	Borderline weak
0.4	14.2	10.0	Adequate
0.5	24.2	10.0	Strong
0.6	35.1	10.0	Strong
Notes: Based on 1,000 Monte Carlo replications with ($n = 400$, $\psi = 0.6$, $\sigma_v^2 = 0.5$)			

Table B2: Bias and RMSE by graph cycle structure ($\psi = 0.6, \sigma_v^2 = 0.5$)				
Graph type	Cycle length	Bias (IV-SMA)	Bias (GMM-SMA)	RMSE (GMM-SMA)
Tree (acyclic)	∞	-0.041	-0.038	0.158
Regular grid	4	-0.079	-0.072	0.184
Grid + long cycles	8	-0.093	-0.083	0.198
Random graph	Variable	-0.085	-0.077	0.192

As predicted, bias increases modestly with longer cycles but remains substantially lower than OLS (bias = -0.398 in this configuration). Geometric mixing holds in all cases, confirming that bounded cycle length is not required.

Appendix C – Additional Results for the Illustrative Application (Tunisia)

This appendix complements Section 6. With $n=24$ spatial units, no statistical inference is valid. The numbers reported below are purely descriptive and serve only to illustrate the mechanical steps of the estimator. They should not be interpreted as evidence for or against any hypothesis.

Table C1 – Placebo regressions using pre period GDP (2000–2005)	
Specification	Estimated coefficient for road density (log)
Fixed effects (FE)	0.042
SMA-IV (no external instruments)	0.051
Hybrid SMA-IV (with external instruments)	0.038

Notes: These coefficients are close to zero, which is consistent with the expectation of no effect on past GDP. However, this does not constitute a test of instrument validity. Statistical power is very low due to the small number of spatial units ($n=24$).

Table C2 – First stage details (hybrid specification) – Descriptive only				
Instrument	Coefficient	Standard error	Partial R^2	F-statistic
W^2X (Spatial)	0.342	0.068	0.124	24.2
W^3X (Spatial)	0.187	0.071	0.067	8.4
Terrain ruggedness	-0.094	0.042	0.049	5.1
Distance to colonial postal routes	-0.231	0.058	0.089	13.2
Interaction (ruggedness $\times$ distance)	0.076	0.039	0.036	3.9
Kleibergen Paap Wald	F = 24.6			

Warning: With $n=24$, these F statistics are reported for descriptive purposes only. Asymptotic critical values do not apply, and they should not be interpreted as evidence for or against weak instruments.