Finding a Needle in a Haystack: the Problem of Dense Weight Matrices in Spatial Econometric Models.

Steven Farber [(1)]
Jesús Mur [(2)]
Manuel Ruiz [(3)]

Abstract

The weighting matrix occupies a central role in the specification of a spatial model. Its purpose is to reflect the spatial layout where the model is to be estimated. It is well-known that errors in the specification of this matrix, result in severe estimation biases and wrong inference. Other problem related with the matrix is that of the density, in the sense of number of connections reflected in the weighting matrix. As proved by Smith (2009), there is a direct link between higher density and worse estimation bias.
Our paper focuses in the extreme case of a maximal connectivity, where each spatial unit is connected, with the same weight, with all others elements in the system. We show that this is a very insidious problem whose consequences depend, to a great extent, on the presence of a constant term in the equation. In that case, the model is not identified and the maximum-likelihood estimators tend to an extreme point in the parametric space. If the model does not include a constant term, the bias is smaller although the maximum-likelihood estimators are inconsistent. We present analytical evidence of these consequences combined with a simulation experiment.

Keywords: Spatial models, weight matrix, density, identification.

JEL codes: R15, C15, C52

[(1)] Department of Geography, University of Utah, Salt Lake City, USA.
Email: Steven.Farber@geog.utah.edu
[(2)] Department of Economic Analysis, University of Zaragoza, Zaragoza, Spain.
Email: jmur@unizar.es
[(3)] Department of Quantitative Methods and Computing, Technical University of Cartagena, Cartagena, Spain. Email:manuel.ruiz@upct.es

A case of maximum,uniform connectivity

Let us assume the following autoregressive process:

$$y_r = \rho \sum_{\substack{s=1 \\ s \neq r}}^{n} \omega_{rs} y_s + \varepsilon_r \tag{1}$$

where each element interacts,with the same intensity, with all others elements in the system. Assume that ε_r is a random term normally distributed, $\varepsilon_r \sim iidN(0;\sigma^2)$;$\{\omega_{rs}; r,s - 1,2,\dots,n\}$ is a sequence of weights and W the so-called weighting matrix.As usual,$\{\omega_{rr} = 0; r = 1,2,\dots,n\}$ and, in this case, $\left\{\omega_{rs} = {}^{1}/_{(n-1)}; r,s = 1,2,\dots,n\right\}$. This is a spatially lagged dependent variable model without any exogenous factors or a constant; heretofore referred to as Model (i).

The summation in the right hand side, rhs in what follows, of (1) is almost the arithmetic mean of y, so we can rewrite y_r as a function of $\bar{y}$ and ε_r:

$$y_r = \rho \frac{n}{n-1}\bar{y} - \frac{\rho}{n-1} y_r + \varepsilon_r$$

$$y_r + \frac{\rho}{n-1} y_r = \rho \frac{n}{n-1}\bar{y} + \varepsilon_r$$

$$y_r \left(1 + \frac{\rho}{n-1}\right) = \rho \frac{n}{n-1}\bar{y} + \varepsilon_r$$

$$y_r = \rho \frac{n}{n-1+\rho}\bar{y} - \frac{1}{n-1+\rho}\varepsilon_r$$

(2)

Summing on both sides of (2) and rearranging terms, we get:

$$y = \frac{1}{1-\rho}\varepsilon \tag{3}$$

Note thatequation (2) offers a clue for an instrumental variable estimate for ρ. The regressor, $\bar{y}$, is correlated with the error term, ε_r, so the plain LS estimator of (2) is inconsistent. However a vector of ones, $\mathbf{1} = [1 \quad 1 \quad \dots \quad 1]'$, can be used as an instrument. The IV estimate of the composite parameter $\rho^* = \frac{\rho n}{n-1+\rho}$, is $\hat{\rho}^*_{IV} = \bar{y}$ which allows to obtain an estimation ofρas:

$$\hat{\rho}_{IV} = \frac{\bar{y}(n-1)}{n-\bar{y}}, \tag{4}$$

This is an unexpected result: there is an IV estimator in a situation where we could hardly have thought in this type of estimation algorithms. There is also a GMM (Generalized Method of Moments) estimate for the equation of (1), whose origin is the sample analogue of the identification condition:

$$E[\varepsilon' W \varepsilon] = 0 \Rightarrow \frac{y'(I-\rho W)' W (I-\rho W) y}{n} = 0 \tag{5}$$

Expanding the above identity condition reveals the nonlinear solution:

$$\frac{\varepsilon' W \varepsilon}{n} = 0 \Rightarrow \bar{y}^2 (1-\rho)^2 - \frac{S_y}{n-1}\left(1 + \frac{\rho}{n-1}\right)^2 = 0 \tag{6}$$

where $S_y = \frac{\sum_{r=1}^{n}(y_r - \bar{y})^2}{n}$ is the sampling variance of y.

Note that the special casewhen the sampling average of y is zero ($\bar{y} = 0$). In this case the IV estimator in (4) is zero, and the GMM estimator in (6) tends to the lower limit of the range of admissible values for ρ, $\hat{\rho}_{GMM} = -(n-1)$ as identified by {Smith 2009}. Thus, while in general IV and GMM estimates for Model (i) exist, they do not in the special case when $\bar{y} = 0$.

Next, we turn to the maximum likelihood estimator of ρ . Rearranging terms in (1-3), we can express $y_r$as a function of only random terms:

$$y_r = \theta_1 \varepsilon_r + \theta_2 \sum_{\substack{s=1 \\ s \neq r}}^{n} \varepsilon_s \rightarrow \begin{cases} \theta_1 = \dfrac{n-1}{n-1+\rho}\left[1 + \dfrac{1}{1-\rho}\dfrac{1}{n-1}\right] \\ \theta_2 = \dfrac{\rho}{(1-\rho)(n-1+\rho)} \end{cases} \tag{7}$$

Note thatif $n \rightarrow \infty$, for any finite value of ρ ,$\theta_1 \rightarrow 1$ and $\theta_2 \rightarrow 0$ .Thus, asymptotically, $V(y_r) = \sigma^2$ and $Cov(y_r; y_s) = 0$ since $\varepsilon_r \sim iidN(0; \sigma^2)$. In the finite sample case, we obtain from (7):

$$\left.\begin{aligned} V(y_r) &= \sigma^2\left(\frac{n-1}{n-1+\rho}\right)^2\left[1+\frac{\rho}{(n-1)(1-\rho)}\left(2+\frac{n\rho}{(n-1)(1-\rho)}\right)\right] \\ Cov(y_r;y_s) &= \sigma^2\frac{\rho}{(n-1+\rho)^2}\left[2+\frac{\rho}{(n-1)}\left(\frac{2}{(1-\rho)}+\frac{n-2}{(n-1)}\right)\right] \end{aligned}\right\} \quad (8)$$

The likelihood functionof the structural model of (1):

$$L(y) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{y'B'By}{2\sigma^2} + \ln|B| \quad (11)$$

where B=I-ρW. This matrix can be written: $B = \frac{n-1+\rho}{n-1}I - \frac{\rho}{n-1}11'$, whose inverse is $B^{-1} = \frac{n-1}{n-1+\rho}\left(\mathrm{I} + \frac{\rho}{n-1+\rho}\mathrm{ll'}\right)$. The score vector of partial derivatives is, again, nonlinear:

$$\begin{aligned} l_\rho &= \frac{\partial L}{\partial \rho} = \frac{y'W'By}{\sigma^2} - tr(B^{-1}W) \\ l_{\sigma^2} &= \frac{\partial L}{\partial \sigma^2} = -\frac{n}{\sigma^2} + \frac{y'B'By}{2\sigma^4} \end{aligned} \quad (12)$$

Note that $tr(B^{-1}W) = \frac{n}{n-1+\rho}\frac{\rho}{1-\rho}$. Using the first and second derivatives, the ML estimate of σ^2 is:

$$\hat{\sigma}^2 = \frac{y'B'By}{n} = \left(\frac{n-1+\rho}{n-1}\right)^2 S_y + (1-\rho)^2\bar{y}^2 \quad (13)$$

whichdepends on the ML estimate of ρ. From the first term of the score vector, $l_\rho=0$, we obtainan equation of order two:

$$\left(\frac{n-1+\rho}{n-1}\right)^2 S_y + (1-\rho)(n-1+\rho)\bar{y}^2 = \frac{\rho}{1-\rho} \quad (14)$$

which can be solved using numerical methods.

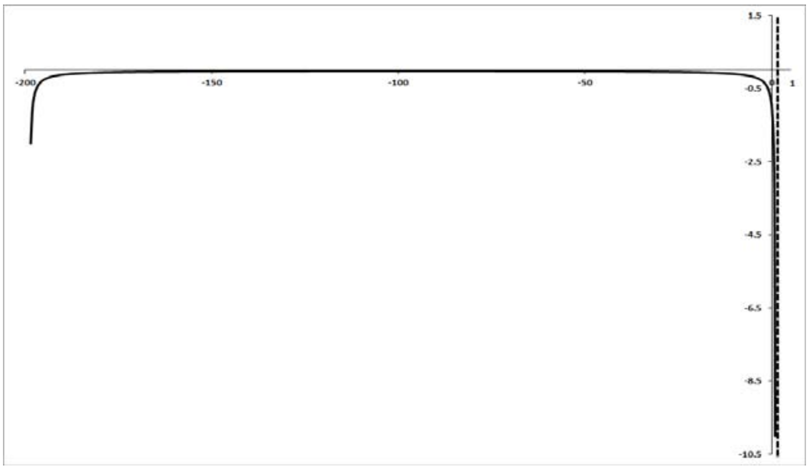
Note that, as with GMM and IV, there is no a ML estimator for ρ when the sampling average of y is zero, $\bar{y}=0$. Also, in this case the ML estimate of σ^2 is:

$\hat{\sigma}^2 = \left(\frac{n-1+\rho}{n-1}\right)^2 \left(\frac{y'y}{n}\right)$ and pluggin this estimate of the variance back into the first derivative of the likelihood function, we find the derivativeto be negative in the range of admissible values of ρ:

$$l_\rho = -\frac{\frac{n-1+\rho}{(n-1)^2} y'y}{\left(\frac{n-1+\rho}{n-1}\right)^2 \left(\frac{y'y}{n}\right)} - \frac{n}{n-1+\rho}\frac{\rho}{1-\rho} = -\frac{n}{n-1+\rho}\frac{1}{1-\rho} < 0;\; -(n-1) < \rho < 1 \quad (15)$$

The situation is clear from Figure 1 which shows ρ on the horizontal axis and the likelihood function on the vertical.

Figure 1: l_ρ for Model (i) in terms of ρ when $\bar{y} = 0$ (n=200)



In sum, the likelihood function is monotonically decreasing with ρ in the range of admissible values: there are no ML estimators.

The information matrix of Model (i) has a compact expression:

$$IM = \begin{bmatrix} 2tr(B^{-1}WB^{-1}W) & \frac{tr(B^{-1}W)}{\sigma^2} \\ \frac{tr(B^{-1}W)}{\sigma^2} & \frac{n}{2\sigma^4} \end{bmatrix} \quad (16)$$

where $tr(B^{-1}WB^{-1}W) = {}^{1}/_{(1-\rho)^2}$. The information matrix must be positive definite, $|IM| > 0$, which in our case implies that the range of admissible values of ρ must be further restricted, from $-(n-1) < \rho < 1$, as indicated by Smith (REFER),to $1 - \sqrt{n} < \rho < 1$. The determinant of (16) is negative outside this range.

From our point of view, the most important aspect of all these results is that, under the two restrictions derived above: (i)-$\bar{y} \neq 0$ and (ii)- $1-\sqrt{n} < \rho < 1$, there are well-defined ML estimates for Model (i), even for the extreme case of a maximally connected weighted matrix.

Table 1 below shows the bias obtained from a simple Monte Carlo of Model (i), with n=200, σ^2=1 and different values of parameter ρ (99.5% of connectivity means that each cell is contiguous to the other 199 other cells). Table 2 contains the results for the same experiment but with a connectivity of only 5% (each cell is contiguous to the nearest 10 cells). The 5%configuration will be treated as the control case.

Table 1: Connectivity vs bias in the ML estimates. MODEL (i). Connectivity=99.5%.

ρ	$\hat{\rho}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.7144	0.0071	0.0138	1.0071	1.0023	1.0000
-0.85	-0.6929	0.0075	0.0138	1.0066	1.0023	1.0000
-0.75	-0.6559	0.0079	0.0138	1.0061	1.0023	1.0000
-0.65	-0.6099	0.0084	0.0138	1.0056	1.0023	1.0000
-0.55	-0.5750	0.0089	0.0138	1.0051	1.0023	1.0000
-0.45	-0.5084	0.0095	0.0138	1.0046	1.0023	1.0000
-0.35	-0.4684	0.0102	0.0138	1.0041	1.0023	1.0000
-0.25	-0.3984	0.0110	0.0138	1.0036	1.0023	1.0000
-0.15	-0.3409	0.0120	0.0138	1.0031	1.0023	1.0000
-0.05	-0.2689	0.0131	0.0138	1.0026	1.0023	1.0000
0.05	-0.1953	0.0145	0.0138	1.0021	1.0023	1.0000
0.15	-0.1050	0.0162	0.0138	1.0016	1.0023	1.0000
0.25	-0.0194	0.0184	0.0138	1.0011	1.0023	1.0000
0.35	0.0682	0.0212	0.0138	1.0006	1.0023	1.0000
0.45	0.1757	0.0251	0.0138	1.0001	1.0023	1.0000
0.55	0.2832	0.0306	0.0138	0.9996	1.0023	1.0000
0.65	0.4158	0.0394	0.0138	0.9991	1.0023	1.0000
0.75	0.5598	0.0552	0.0138	0.9986	1.0023	1.0000
0.85	0.7133	0.0919	0.0138	0.9981	1.0023	1.0000
0.95	0.8823	0.2758	0.0138	0.9976	1.0023	1.0000

$\hat{\rho}$ stands for the ML estimate of ρ; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y;$\bar{\varepsilon}$and $r_{y,\varepsilon}$ stand for the sampling average and standard deviation of vector ε; $r_{y,\varepsilon}$ stands for the correlation coefficient between vectors y and ε.

Table 2: Connectivity vs bias in the ML estimates. MODEL (i). Connectivity=5%.

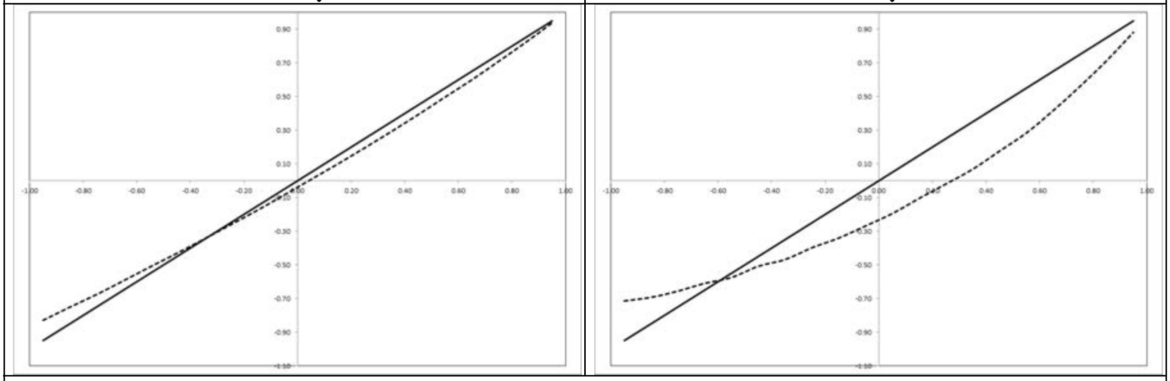
ρ	$\hat{\rho}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.8294	0.0082	0.0138	1.0862	1.0023	0.9768
-0.85	-0.7512	0.0086	0.0138	1.0715	1.0023	0.9808
-0.75	-0.6760	0.0090	0.0138	1.0581	1.0023	0.9845
-0.65	-0.5947	0.0094	0.0138	1.0461	1.0023	0.9878
-0.55	-0.5116	0.0099	0.0138	1.0354	1.0023	0.9908
-0.45	-0.4311	0.0105	0.0138	1.0259	1.0023	0.9935
-0.35	-0.3474	0.0110	0.0138	1.0179	1.0023	0.9958
-0.25	-0.2615	0.0117	0.0138	1.0112	1.0023	0.9977
-0.15	-0.1753	0.0125	0.0138	1.0062	1.0023	0.9991
-0.05	-0.0846	0.0133	0.0138	1.0031	1.0023	0.9999
0.05	0.0078	0.0143	0.0138	1.0022	1.0023	0.9999
0.15	0.0997	0.0155	0.0138	1.0040	1.0023	0.9988
0.25	0.1949	0.0169	0.0138	1.0095	1.0023	0.9961
0.35	0.2922	0.0188	0.0138	1.0200	1.0023	0.9911
0.45	0.3928	0.0211	0.0138	1.0382	1.0023	0.9825
0.55	0.4970	0.0244	0.0138	1.0685	1.0023	0.9679
0.65	0.5995	0.0292	0.0138	1.1211	1.0023	0.9427
0.75	0.7092	0.0374	0.0138	1.2208	1.0023	0.8970
0.85	0.8211	0.0549	0.0138	1.4528	1.0023	0.8056
0.95	0.9379	0.1303	0.0138	2.4153	1.0023	0.5823

$\hat{\rho}$ stands for the ML estimate of ρ; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y;$\bar{\varepsilon}$and $r_{y,\varepsilon}$ stand for the sampling average and standard deviation of vector ε; $r_{y,\varepsilon}$ stands for the correlation coefficient between vectors y and ε.

Figure 2illustrates the sequence of ML estimates for ρ on the vertical axis, versus the simulated coefficients of the DGP on the horizontal axis. The tendency to underestimate the parameter of spatial dependence is evident in the two cases, as is typical for the ML algorithm. The tendency is accentuated when a highly dense matrix is used in the DGP.

Figure 2: Connectivity vs ML estimation. MODEL (i)

Connectivity = 5.0%	Connectivity = 99.5%



rho estimated — perfect fit

Let us add a constant term to the model of (1):

$$y_r = \delta + \rho \sum_{\substack{s=1 \\ s\neq r}}^{n} \omega_{rs} y_s + \varepsilon_r \tag{17}$$

This is what we call Model (ii). Repeating all the calculus, we obtain that:

$$\rightarrow y_r = \frac{n-1}{n-1+\rho}\left[\frac{\delta}{\bar{y}} + \rho\frac{n}{n-1}\right]\bar{y} + \frac{n-1}{n-1+\rho}\varepsilon_r = \delta^{**} + \frac{n-1}{n-1+\rho}\varepsilon_r \quad \rightarrow \bar{y} = \frac{1}{1-\rho}[\delta + \bar{\varepsilon}]$$

(18)

The equivalent expression to the reduced form of (7) is:

$$y_r = \delta + \rho \sum_{\substack{s=1 \\ s\neq r}}^{n} \omega_{rs} y_s + \varepsilon_r \Rightarrow y_r = \delta^* + \theta_1 \varepsilon_r + \theta_2 \sum_{\substack{s=1 \\ s\neq r}}^{n} \varepsilon_s$$

$$\rightarrow \begin{cases} \delta^* = \delta \dfrac{n-1}{n-1+\rho}\left[1 + \dfrac{\rho}{1-\rho}\dfrac{n}{n-1}\right] \\ \theta_1 = \dfrac{n-1}{n-1+\rho}\left[1 + \dfrac{\rho}{1-\rho}\dfrac{1}{n-1}\right] \\ \theta_2 = \dfrac{\rho}{(1-\rho)(n-1+\rho)} \end{cases} \tag{19}$$

The constant term of the reduced form, assuming that $n \rightarrow \infty$, tends to $\delta^* \rightarrow \dfrac{\delta}{1-\rho}$ whereas $\theta_1 \rightarrow 1$ and $\theta_2 \rightarrow 0$, as before. Furthermore, the expressions for the variance and covariances remain the same, as in (8).

This reduced form does not allow estimating separately the two structural parameters. In factthe LS estimates of (19) would produce an estimation of the composite parameter δ^*.

Using (18) we can derive an IV estimator for the composite parameter, $\delta^{**} = \frac{n-1}{n-1+\rho}\left[\frac{\delta}{\bar{y}} + \rho\frac{n}{n-1}\right]$ but, as before, it is not possible to estimate separately δ from ρ.Even in the case that $\bar{y} = 0$, the corresponding IV estimator will be a mixture of both parameters: $\hat{\delta}^{**}_{IV} = \delta\frac{\widetilde{n-1}}{n-1+\rho}$. So in this case, there is no good IV estimator for the spatial parameter.

Nor is there a GMM solution to the problem of estimating the model of (17). In this case we need two conditions, namely, that the mean of the residual terms is zero (E[ε]=0) and that the residual terms are uncorrelated with their spatial lag (E[ε'Wε]=0). Using the sampling analogues of these two conditions, we get:

$$E\left[\frac{\sum_{r=1}^{n}\varepsilon_r}{n}\right] = 0 \rightarrow \varepsilon'l = (By - \delta l)'l = 0 \Rightarrow \delta = (1-\rho)\bar{y} \tag{20a}$$

$$E\left[\frac{\sum_{r=1}^{n}\varepsilon_r\omega_{rs}\varepsilon_s}{n}\right] = 0 \rightarrow \varepsilon'W\varepsilon = (By - \delta l)'W(By - \delta l) \Rightarrow \left(\frac{n-1+\rho}{n-1}\right)^2\frac{n}{n-1}S_y = 0 \tag{20b}$$

Condition (20a) implies that δ must be a function of ρ and $\bar{y}$ if the residuals are to have a mean of 0. This does not necessarily preclude the existence of GMM estimators. Condition (20b), however, can only be fulfilledwhen$\rho = -(n-1)$ which isthe extreme point of the range of admissible values for ρ. We therefore have no GMM estimate for ρorδ.

Model (ii) is not identified.It is important to stress that, in this case, the lack of identification is not a particular result, corresponding to a specific point in the sample space, $(\bar{y} = 0)$, as in the case of Model (i). The lack of identification affects Model (ii) whenever the contiguity matrix is fully connected.

The lack of identification also affects the ML estimators. The likelihood function of the structural equation of (17) using obvious matrix notation is:

$$l(y) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{(By-\delta l)'(By-\delta l)}{2\sigma^2} + ln|B| \tag{21}$$

The score vector composed of the three partial derivatives is nonlinear:

$$\left.\begin{aligned} l_{\rho} &= \frac{y'W(By-\delta l)}{\sigma^2} - trB^{-1}W \\ l_{\delta} &= \frac{-\delta n + y'Bl}{\sigma^2} \\ l_{\sigma^2} &= -\frac{n}{\sigma^2} + \frac{(By-\delta l)'(By-\delta l)}{2\sigma^4} \end{aligned}\right\} \qquad (22)$$

The ML estimator of δ is shown here to depend on ρ :

$$l_{\delta} = 0 \rightarrow \hat{\delta} = (1-\rho)\bar{y} \qquad (23)$$

The ML estimator of σ^2also depends on ρ :

$$l_{\sigma^2} - 0 \rightarrow \hat{\sigma}^2 = \frac{\left(By-\hat{\delta}l\right)'\left(By-\hat{\delta}l\right)}{n} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n} = \left(\frac{n-1+\rho}{n-1}\right)^2 S_y \qquad (24)$$

Using (23) and (24), the ML estimator of ρ comes from:

$$l_{\rho} = 0 \rightarrow \frac{\left[-\frac{1}{(n-1)(y-n\bar{y})}\right]'\left[\frac{n-1+\rho}{(n-1)(y-\bar{y})}\right]}{\hat{\sigma}^2} - \frac{n}{n-1+\rho}\frac{\rho}{1-\rho} = -\frac{n}{n-1+\rho} - \frac{n}{n-1+\rho}\frac{\rho}{1-\rho} \neq 0$$

(25a)

$$l_{\rho} = 0 \rightarrow -\frac{n}{(n-1+\rho)\left(1+\frac{\rho}{1-\rho}\right)} \neq 0 \qquad (25b)$$

This condition cannot be fulfilled in the usual range of admissible values for ρ, $-(n-1) < \rho < 1$. Moreover, it is evident that $l_{\rho} < 0$ in the admissible range of ρ , meaning that the function is monotonically decreasing in this range of values. The situation is similar to that depicted in Figure 1 and the ML estimate for ρ tends toward its minimum extreme point.

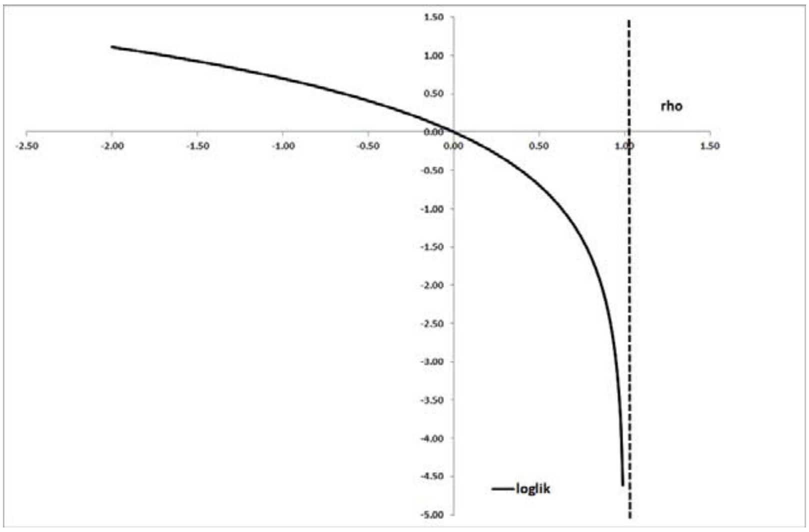
From another perspective, conditional on the ML estimates of δ and σ^2, we can obtain the ML estimation of ρ:

$$\left.\begin{aligned} l\left(y\,\middle|\,\hat{\delta},\hat{\sigma}^2\right) &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\left(\frac{n-1+\rho}{n-1}\right)^2 S_y\right) - \frac{n}{2} + \ln(1-\rho)\left(\frac{n-1+\rho}{n-1}\right)^{n-1} \\ &\rightarrow l\left(y\,\middle|\,\hat{\delta},\hat{\sigma}^2\right) \simeq -\ln\left(\frac{n-1+\rho}{n-1}\right) + \ln(1-\rho) \end{aligned}\right\}$$

(26)

This function is not bounded as is clear in Figure 3.

Figure 3: Log-likelihood of Model (ii) in terms of ρ (n=200)



In sum, there is no a ML estimator for ρ in Model (ii), expression(17). Obviously, the same applies for the estimates of δ and σ^2, which depend on ρ.

For the sake of completeness, the information matrix of Model (ii) would be:

$$IM = \begin{bmatrix} \frac{\delta^2}{\sigma^2(1-\rho)^2}\left(n + 2\frac{\sigma^2}{\delta^2}\right) & \frac{n\delta}{\sigma^2(1-\rho)} & \frac{n\rho}{\sigma^2(n-1+\rho)(1-\rho)} \\ \frac{n\delta}{\sigma^2(1-\rho)} & \frac{n}{\sigma^2} & 0 \\ \frac{n\rho}{\sigma^2(n-1+\rho)(1-\rho)} & 0 & \frac{n}{2\sigma^4} \end{bmatrix} \quad (27)$$

whose determinant is:

$$|IM| = \frac{n^2}{\sigma^6(1-\rho)^2}\left[1 - \frac{n\rho^2}{(n-1+\rho)^2}\right] > 0 \Rightarrow 1-\sqrt{n} < \rho < 1 \quad (28)$$

The new restriction on parameter ρ, obtained in relation toModel (i),also is in order here.

We may repeat the experiment of Tables 1 and 2, now including a constant in the DGP. Let us assume that $\delta = 1$. Main results are summarized in Tables 3 and 4 and in Figure 4.

Table 3: Connectivity vs bias in the ML estimates. MODEL (ii). Connectivity=99.5%.

ρ	$\hat{\rho}$	$\hat{\delta}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.9952	1.0239	0.5132	0.0007	1.0073	1.0025	1.0000
-0.85	-0.9963	1.0799	0.5409	0.0007	1.0068	1.0025	1.0000
-0.75	-0.9957	1.1412	0.5718	0.0007	1.0063	1.0025	1.0000
-0.65	-0.9958	1.2104	0.6065	0.0007	1.0057	1.0025	1.0000
-0.55	-0.9955	1.2884	0.6456	0.0007	1.0052	1.0025	1.0000
-0.45	-0.9956	1.3773	0.6902	0.0007	1.0047	1.0025	1.0000
-0.35	-0.9957	1.4794	0.7413	0.0007	1.0042	1.0025	1.0000
-0.25	-0.9960	1.5980	0.8006	0.0007	1.0037	1.0025	1.0000
-0.15	-0.9957	1.7366	0.8702	0.0007	1.0032	1.0025	1.0000
-0.05	-0.9955	1.9018	0.9531	0.0007	1.0027	1.0025	1.0000
0.05	-0.9957	2.1023	1.0534	0.0007	1.0022	1.0025	1.0000
0.15	-0.9960	2.3499	1.1773	0.0007	1.0017	1.0025	1.0000
0.25	-0.9955	2.6626	1.3343	0.0007	1.0012	1.0025	1.0000
0.35	-0.9963	3.0734	1.5396	0.0007	1.0007	1.0025	1.0000
0.45	-0.9956	3.6309	1.8195	0.0007	1.0002	1.0025	1.0000
0.55	-0.9956	4.4378	2.2238	0.0007	0.9997	1.0025	1.0000
0.65	-0.9957	5.7060	2.8592	0.0007	0.9992	1.0025	1.0000
0.75	-0.9956	7.9880	4.0029	0.0007	0.9987	1.0025	1.0000
0.85	-0.9961	13.3170	6.6715	0.0007	0.9982	1.0025	1.0000
0.95	-0.9957	39.9435	20.0144	0.0007	0.9977	1.0025	1.0000

$\hat{\rho}$ stands for the ML estimate of ρ; $\hat{\delta}$ stands for the ML estimate of δ; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y; $\bar{\varepsilon}$ and $\hat{\sigma}_\varepsilon$ stand for the sampling average and standard deviation of vector u; $r_{y,\varepsilon}$ stands for the correlation coefficient between vectors y and ε

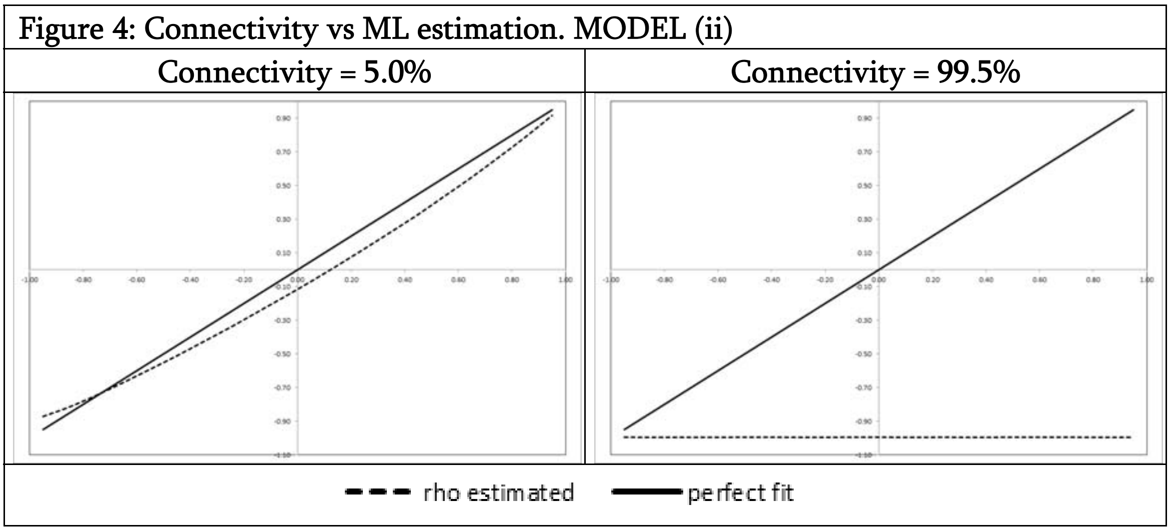
Figure 4 resumes the sequence of ML estimated parameters, depending on the connectivity. The situation depicted by this figure is similar to that described by Smith (2009): the ML estimate of the autoregressive coefficient is always in the proximities of -1, independently of the value used in the DGP, anda strong bias is affecting the ML estimate of the constant.

Table 4: Connectivity vs bias in the ML estimates. MODEL (ii). Connectivity=5.0%.

ρ	$\hat{\rho}$	$\hat{\delta}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.8147	0.9597	0.5122	0.0007	1.0921	1.0025	0.9775
-0.85	-0.7385	0.9804	0.5400	0.0007	1.0768	1.0025	0.9815
-0.75	-0.6580	0.9965	0.5710	0.0007	1.0629	1.0025	0.9850
-0.65	-0.5775	1.0110	0.6058	0.0007	1.0503	1.0025	0.9883
-0.55	-0.4944	1.0256	0.6450	0.0007	1.0390	1.0025	0.9912
-0.45	-0.4129	1.0409	0.6897	0.0007	1.0290	1.0025	0.9938
-0.35	-0.3293	1.0552	0.7409	0.0007	1.0203	1.0025	0.9960
-0.25	-0.2447	1.0716	0.8003	0.0007	1.0131	1.0025	0.9978
-0.15	-0.1591	1.0868	0.8700	0.0007	1.0075	1.0025	0.9991
-0.05	-0.0708	1.1046	0.9530	0.0007	1.0036	1.0025	0.9999
0.05	0.0181	1.1237	1.0534	0.0007	1.0019	1.0025	0.9999
0.15	0.1104	1.1430	1.1774	0.0007	1.0028	1.0025	0.9988
0.25	0.2021	1.1639	1.3345	0.0007	1.0072	1.0025	0.9963
0.35	0.2960	1.1874	1.5397	0.0007	1.0164	1.0025	0.9915
0.45	0.3948	1.2168	1.8195	0.0007	1.0328	1.0025	0.9833
0.55	0.4937	1.2420	2.2234	0.0007	1.0608	1.0025	0.9691
0.65	0.5951	1.2815	2.8577	0.0007	1.1102	1.0025	0.9445
0.75	0.7016	1.3225	3.9986	0.0007	1.2051	1.0025	0.8992
0.85	0.8093	1.4034	6.6576	0.0007	1.4277	1.0025	0.8077
0.95	0.9254	1.6371	19.9262	0.0007	2.3383	1.0025	0.5830

$\hat{\rho}$ stands for the ML estimate of ρ; $\hat{\delta}$ stands for the ML estimate of δ; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y; $\bar{\varepsilon}$ and $\hat{\sigma}_\varepsilon$ stand for the sampling average and standard deviation of vector ε; $r_{y,\varepsilon}$ stands for the correlation coefficient between vectors y and ε

Figure 4: Connectivity vs ML estimation. MODEL (ii)



These results are coherent with our findings. The ML estimate of ρ tends to –(n-1); given that the ML algorithms usually restrict the optimization search to the interval

(-1;1), the value of -1 is the logical consequence. The ML estimate of δ, according to (24), is twice the sampling mean of y ($\hat{\delta} = (1-\rho)\bar{y} \approx 2\bar{y}$) and the standard deviation of the residual terms should be a bit smaller than the standard deviation of y ($\hat{\sigma} = \left(\frac{n-2}{n-1}\right)\sqrt{S_y}$). This is what is called *density bias* in the literature. According to our results, this bias should be attributed to an identification problem in Model (ii) which results in a bad-behaved likelihood function.

If, in Model (ii) we substitute the constant term by another variable which has some variability, the identification problems disappear:

$$y_r = \beta x_r + \rho \sum_{\substack{s=1 \\ s \neq r}}^{n} \omega_{rs} y_s + \varepsilon_r ; \quad V(x_r) > 0 \tag{29}$$

Table 5: Connectivity vs bias in the ML estimates. MODEL (iii). Connectivity=99.5%.

ρ	$\hat{\rho}$	$\hat{\beta}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.8525	1.0031	0.0045	-0.0065	1.4151	0.9950	0.7061
-0.85	-0.8186	1.0029	0.0047	-0.0065	1.4144	0.9950	0.7061
-0.75	-0.7954	1.0026	0.0050	-0.0065	1.4137	0.9950	0.7061
-0.65	-0.7617	1.0024	0.0053	-0.0065	1.4130	0.9950	0.7061
-0.55	-0.7128	1.0022	0.0056	-0.0065	1.4123	0.9950	0.7061
-0.45	-0.6767	1.0020	0.0060	-0.0065	1.4116	0.9950	0.7061
-0.35	-0.6361	1.0018	0.0065	-0.0065	1.4109	0.9950	0.7061
-0.25	-0.5877	1.0017	0.0070	-0.0065	1.4102	0.9950	0.7061
-0.15	-0.5316	1.0015	0.0076	-0.0065	1.4094	0.9950	0.7061
-0.05	-0.4856	1.0014	0.0083	-0.0065	1.4087	0.9950	0.7061
0.05	-0.4128	1.0014	0.0092	-0.0065	1.4080	0.9950	0.7061
0.15	-0.3482	1.0013	0.0103	-0.0065	1.4073	0.9950	0.7061
0.25	-0.2657	1.0014	0.0117	-0.0065	1.4066	0.9950	0.7061
0.35	-0.1782	1.0014	0.0134	-0.0065	1.4059	0.9950	0.7061
0.45	-0.0684	1.0016	0.0159	-0.0065	1.4052	0.9950	0.7061
0.55	0.0607	1.0019	0.0194	-0.0065	1.4045	0.9950	0.7061
0.65	0.2107	1.0023	0.0250	-0.0065	1.4038	0.9950	0.7061
0.75	0.3769	1.0026	0.0350	-0.0065	1.4031	0.9950	0.7061
0.85	0.5722	1.0031	0.0583	-0.0065	1.4024	0.9950	0.7061
0.95	0.8063	1.0037	0.1748	-0.0065	1.4017	0.9950	0.7061

$\hat{\rho}$ stands for the ML estimate of ρ; $\hat{\beta}$ stands for the ML estimate of β; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y; $\bar{\varepsilon}$ and $\hat{\sigma}_\varepsilon$ stand for the sampling average and standard deviation of vector ε; $r_{y,\varepsilon}$ stands for the correlation coefficient between vectors y and ε

The equivalence between the equations of the reduced form of Model (i) and (ii) no longer holds in relation to Model (iii) of expression (29). We can repeat the experiment of ML estimation of Model (iii), with the results shown in Tables 5, 6 and in Figure 5. Note that parameter β is equal to 1.

Table 6: Connectivity vs bias in the ML estimates. MODEL (iii). Connectivity=5.0%.

ρ	$\hat{\rho}$	$\hat{\beta}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.8420	1.0074	0.0033	-0.0065	1.5177	0.9950	0.6891
-0.85	-0.7638	1.0057	0.0036	-0.0065	1.4977	0.9950	0.6920
-0.75	-0.6768	1.0045	0.0039	-0.0065	1.4795	0.9950	0.6947
-0.65	-0.5885	1.0037	0.0043	-0.0065	1.4633	0.9950	0.6972
-0.55	-0.4962	1.0032	0.0047	-0.0065	1.4489	0.9950	0.6994
-0.45	-0.4088	1.0026	0.0052	-0.0065	1.4364	0.9950	0.7014
-0.35	-0.3186	1.0022	0.0057	-0.0065	1.4259	0.9950	0.7031
-0.25	-0.2277	1.0019	0.0064	-0.0065	1.4176	0.9950	0.7045
-0.15	-0.1366	1.0017	0.0072	-0.0065	1.4118	0.9950	0.7055
-0.05	-0.0440	1.0016	0.0082	-0.0065	1.4087	0.9950	0.7061
0.05	0.0491	1.0016	0.0094	-0.0065	1.4090	0.9950	0.7060
0.15	0.1435	1.0017	0.0109	-0.0065	1.4136	0.9950	0.7051
0.25	0.2393	1.0019	0.0129	-0.0065	1.4239	0.9950	0.7029
0.35	0.3349	1.0023	0.0156	-0.0065	1.4422	0.9950	0.6989
0.45	0.4335	1.0026	0.0193	-0.0065	1.4726	0.9950	0.6920
0.55	0.5327	1.0031	0.0249	-0.0065	1.5228	0.9950	0.6802
0.65	0.6340	1.0035	0.0340	-0.0065	1.6095	0.9950	0.6599
0.75	0.7341	1.0042	0.0506	-0.0065	1.7755	0.9950	0.6229
0.85	0.8381	1.0046	0.0890	-0.0065	2.1702	0.9950	0.5498
0.95	0.9446	1.0044	0.2577	-0.0065	3.8604	0.9950	0.3810

$\hat{\rho}$ stands for the ML estimate of ρ; $\hat{\beta}$ stands for the ML estimate of β; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y; $\bar{\varepsilon}$ and $\hat{\sigma}_\varepsilon$ stand for the sampling average and standard deviation of vector ε; $r_{y,\varepsilon}$ stands for the correlation coefficient between vectors y and ε

Figure 5: Connectivity vs ML estimation. MODEL (iii)

Connectivity = 5.0%	Connectivity = 99.5%

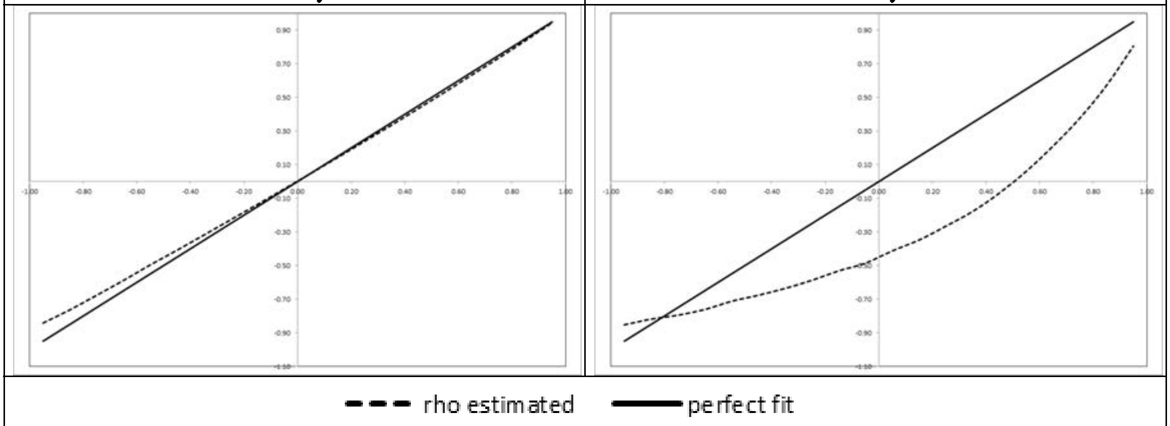


Figure 5 confirms that there is a tendency to underestimate the parameter of spatial dependence, rather more severe in the case of extremely dense weighting matrices. The underestimation is also slightly present in the case of a 5% standard connectivity.

Finally, Model (iv) combines model (ii) and (iii) by includinga constant term as well as the exogenous variable X in the equation of the DGP:

$$y_r = \delta + \beta x_r + \rho \sum_{\substack{s=1 \\ s \neq r}}^{n} \omega_{rs} y_s + \varepsilon_r ; \quad V(x_r) > 0 \tag{30}$$

The consequences are well known: the model is not identified. The log-likelihood function for this general model is:

$$l(y) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{(By - x\beta)'(By - x\beta)}{2\sigma^2} + \ln|B| \tag{31}$$

where x is a (nxk) matrix of observations of the k regressors (assume that the first is a column of ones, associated to the constant term) and β is a (kx1) vector of parameters. The score vector is:

$$\left.\begin{array}{c} \frac{\partial l}{\partial \rho} = -\frac{y'W(By - x\beta)}{\sigma^2} - tr B^{-1}W \\ \frac{\partial l}{\partial \beta} = -\frac{x'(By - x\beta)}{\sigma^2} \\ \frac{\partial l}{\partial \beta} = -\frac{n}{2\sigma^2} + \frac{(By - x\beta)'(By - x\beta)}{2\sigma^4} \end{array}\right\} \tag{32}$$

The ML estimate of β is:

$$\hat{\beta}_{ML} = (x'x)^{-1}x'By = \left(\frac{n-1+\rho}{n-1}\right)(x'x)^{-1}x'y - \rho\frac{n}{n-1}\bar{y}u_k = \left(\frac{n-1+\rho}{n-1}\right)\hat{\beta}_{LS} - \rho\frac{n}{n-1}\bar{y}u_k$$
(33)

where $\hat{\beta}_{LS}$ is the vector of LS estimates and u_k is a (kx1) vector of zeros except the first element which is a 1. The ML residuals are proportional to the LS residuals:

$$\hat{u}_{ML} = By - x\hat{\beta}_{ML} = \left(\frac{n-1+\rho}{n-1}\right)\hat{u}_{LS} \tag{34}$$

This relation is maintained in terms of the estimation of σ^2:

$$\hat{\sigma}^2_{ML} = \left(\frac{n-1+\rho}{n-1}\right)^2 \hat{\sigma}^2_{LS} \tag{35}$$

Finally, substituting (35) and (33) in the log-likelihood of (31):

$$\left.\begin{array}{c} \left(y\,\middle|\,\hat{\beta}_{ML},\hat{\sigma}^2_{ML}\right) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\hat{\sigma}^2_{ML}\right) - \frac{n}{2\sigma^2} + \ln(1-\rho)\left(\frac{n-1+\rho}{n-1}\right)^{n-1} \\ \left(y\,\middle|\,\hat{\beta}_{ML},\hat{\sigma}^2_{ML}\right) \simeq -\ln\left(\frac{n-1+\rho}{n-1}\right) + \ln(1-\rho) - ln\hat{\sigma}^2_{LS} \end{array}\right\}$$
(36)

Note that the LS estimation of σ^2 does not depend on ρ. The situation is similar of that of Model (ii), expression (26) and Figure 3: the likelihood function is unbounded with respect to ρ.

Tables 7 and 8 and Figure 6illustrate the main results of the corresponding simulation experiment of this case, Model (iv), using only one regressor combined with a constant term (the parameters associated to both terms are 1).

Table 7: Connectivity vs bias in the ML estimates. MODEL (iv). Connectivity=99.5%. $\hat{\beta}$

ρ	$\hat{\rho}$	$\hat{\delta}$	$\hat{\beta}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.9960	1.0171	1.0025	0.5113	-0.0057	1.4344	1.0067	0.7054
-0.85	-0.9959	1.0722	1.0020	0.5389	-0.0057	1.4336	1.0067	0.7054
-0.75	-0.9960	1.1337	1.0015	0.5697	-0.0057	1.4329	1.0067	0.7054
-0.65	-0.9956	1.2024	1.0010	0.6042	-0.0057	1.4322	1.0067	0.7054
-0.55	-0.9957	1.2802	1.0005	0.6432	-0.0057	1.4315	1.0067	0.7054
-0.45	-0.9960	1.3690	1.0000	0.6876	-0.0057	1.4307	1.0067	0.7054
-0.35	-0.9959	1.4706	0.9995	0.7385	-0.0057	1.4300	1.0067	0.7054
-0.25	-0.9960	1.5885	0.9990	0.7976	-0.0057	1.4293	1.0067	0.7054
-0.15	-0.9964	1.7274	0.9985	0.8669	-0.0057	1.4286	1.0067	0.7054
-0.05	-0.9959	1.8917	0.9980	0.9495	-0.0057	1.4279	1.0067	0.7054
0.05	-0.9956	2.0908	0.9975	1.0494	-0.0057	1.4271	1.0067	0.7054
0.15	-0.9956	2.3373	0.9970	1.1729	-0.0057	1.4264	1.0067	0.7054
0.25	-0.9958	2.6496	0.9965	1.3293	-0.0057	1.4257	1.0067	0.7054
0.35	-0.9961	3.0583	0.9960	1.5338	-0.0057	1.4250	1.0067	0.7054
0.45	-0.9956	3.6140	0.9955	1.8127	-0.0057	1.4243	1.0067	0.7054
0.55	-0.9956	4.4180	0.9950	2.2155	-0.0057	1.4236	1.0067	0.7054
0.65	-0.9959	5.6819	0.9945	2.8485	-0.0057	1.4229	1.0067	0.7054
0.75	-0.9960	7.9563	0.9940	3.9879	-0.0057	1.4221	1.0067	0.7054
0.85	-0.9961	13.2635	0.9935	6.6465	-0.0057	1.4214	1.0067	0.7054
0.95	-0.9960	39.7949	0.9930	19.9394	-0.0057	1.4207	1.0067	0.7054

$\hat{\rho}$ stands for the ML estimate of ρ; $\hat{\delta}$ and $\hat{\beta}$ stand for the ML estimate of δ and β; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y; $\bar{u}$ and $\hat{\sigma}_u$ stand for the sampling average and standard deviation of vector ε; $r_{y,u}$ stands for the correlation coefficient between vectors y and ε

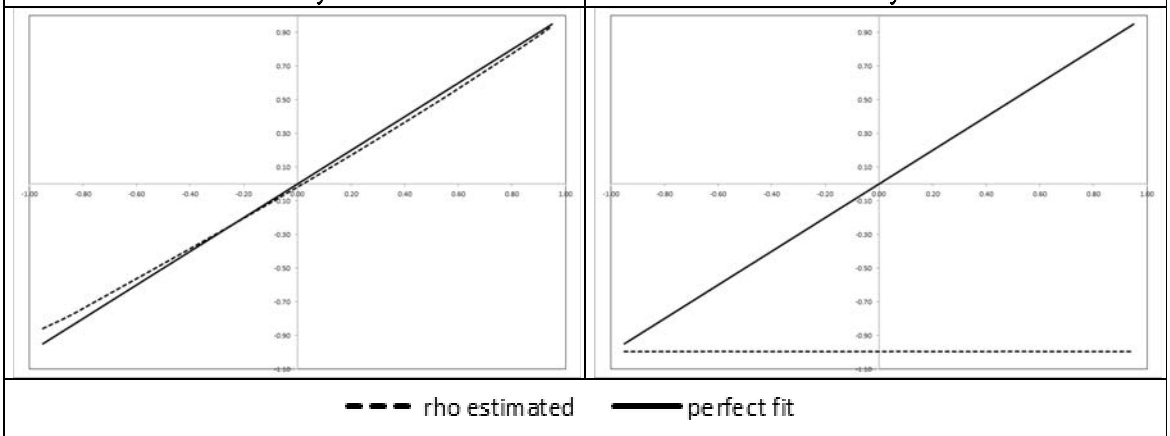
Table 8: Connectivity vs bias in the ML estimates. MODEL (iv). Connectivity=5.0%.

ρ	$\hat{\rho}$	$\hat{\delta}$	$\hat{\beta}$	$\bar{y}$	$\bar{\varepsilon}$	$\hat{\sigma}_y$	$\hat{\sigma}_\varepsilon$	$r_{y,\varepsilon}$
-0.95	-0.8592	0.9463	1.0066	0.5133	-0.0057	1.5395	1.0067	0.6888
-0.85	-0.7831	0.9567	1.0048	0.5407	-0.0057	1.5189	1.0067	0.6918
-0.75	-0.6954	0.9617	1.0039	0.5714	-0.0057	1.5004	1.0067	0.6945
-0.65	-0.6068	0.9666	1.0032	0.6057	-0.0057	1.4838	1.0067	0.6969
-0.55	-0.5179	0.9719	1.0026	0.6445	-0.0057	1.4691	1.0067	0.6991
-0.45	-0.4286	0.9778	1.0021	0.6887	-0.0057	1.4563	1.0067	0.7010
-0.35	-0.3392	0.9844	1.0018	0.7394	-0.0057	1.4457	1.0067	0.7027
-0.25	-0.2492	0.9917	1.0015	0.7983	-0.0057	1.4372	1.0067	0.7040
-0.15	-0.1570	0.9984	1.0014	0.8674	-0.0057	1.4311	1.0067	0.7049
-0.05	-0.0647	1.0061	1.0015	0.9497	-0.0057	1.4279	1.0067	0.7054
0.05	0.0285	1.0147	1.0017	1.0493	-0.0057	1.4281	1.0067	0.7052
0.15	0.1234	1.0232	1.0020	1.1723	-0.0057	1.4325	1.0067	0.7041
0.25	0.2187	1.0336	1.0025	1.3283	-0.0057	1.4426	1.0067	0.7018
0.35	0.3170	1.0426	1.0029	1.5322	-0.0057	1.4606	1.0067	0.6977
0.45	0.4160	1.0534	1.0036	1.8102	-0.0057	1.4903	1.0067	0.6907
0.55	0.5154	1.0685	1.0043	2.2120	-0.0057	1.5391	1.0067	0.6791
0.65	0.6187	1.0805	1.0052	2.8434	-0.0057	1.6226	1.0067	0.6593
0.75	0.7221	1.1028	1.0061	3.9800	-0.0057	1.7797	1.0067	0.6236
0.85	0.8269	1.1450	1.0072	6.6325	-0.0057	2.1438	1.0067	0.5534
0.95	0.9379	1.2312	1.0071	19.8951	-0.0057	3.6508	1.0067	0.3874

$\hat{\rho}$ stands for the ML estimate of ρ; $\hat{\delta}$ and $\hat{\beta}$ stand for the ML estimate of δ and β; $\bar{y}$ and $\hat{\sigma}_y$ stand for the sampling average and standard deviation of vector y; $\bar{u}$ and $\hat{\sigma}_u$ stand for the sampling average and standard deviation of vector ε; $r_{y,u}$ stands for the correlation coefficient between vectors y and ε

Figure 6: Connectivity vs ML estimation. MODEL (iv)

Connectivity = 5.0% | Connectivity = 99.5%



For the sake of completeness, in Figure 7 we include a brief summary of the results of the most popular tests of spatial correlation, namely, the Moran's I and the

Lagrange Multiplier for a lag structure in the variable. The expressions of the two statistics are well-known:

$$I = \frac{\sum_{r,s=1}^{n}(y_r - \bar{y})\omega_{rs}(y_s - \bar{y})}{\sum_{r=1}^{n}(y_r - \bar{y})^2} \tag{37}$$

$$LM_{LAG} = \frac{\left(\hat{\varepsilon}'Wy/_{\hat{\sigma}^2}\right)^2}{\frac{\left(Wx\hat{\beta}\right)' M \left(Wx\hat{\beta}\right)}{\hat{\sigma}^2} + tr(W'W + WW)} \tag{38}$$

with $\hat{\beta}$ being the LS estimate of the parameters of the rhs of the equation and M the matrix $M - I - x(x'x)^{-1}x'$. Moran's I, once normalized, is distributed asymptotically as a standard normal distribution whereas the Lagrange Multiplier is a chi-square with one degree of freedom.

Smith (2009) shows that Moran's I, in the case of totally connected matrices,is equal to its expected value, $-\frac{1}{n-1}$, whatever the characteristics of DGP under consideration (it should stressed that the value of ρ does not matter). The result is clear just looking at the architecture of the statistic:

$$I = \frac{(y-\bar{y})W(y-\bar{y})}{(y-\bar{y})(y-\bar{y})} = \frac{(y-\bar{y})\left[-\frac{1}{n-1}(I-11')\right](y-\bar{y})}{(y-\bar{y})(y-\bar{y})} = -\frac{1}{n-1}\frac{(y-\bar{y})[(I-11')](y-\bar{y})}{(y-\bar{y})(y-\bar{y})} = -\frac{1}{n-1} \tag{39}$$

The Moran's I test is always zero for this specific type of matrix.

The discussion about the Lagrange Multiplier depends on the knowledge we assume about the DGP. We may consider the following cases:

CASE 1: Model (i), as it appears in (1),there are neither a constant norregressors in the rhs of the equation.

The LM for this case is the following:

$$LM_{LAG} = \frac{1}{2}\frac{n}{n-1}\left[n\frac{\bar{y}^2}{S_y + \bar{y}^2} - 1\right]^2 \sim \chi^2(1) \tag{40}$$

It is well–behaved for the range of positive values of ρ (first panel of Figure 7) but it lacks power for negative values of ρ.

CASE 2: Model (ii), as it appears in (17),there is only a constant in the rhs of the equation.

The LM for this case is the following:

$$LM_{LAG} = \frac{1}{2}\left[\frac{n}{n-1}\right]^2 \sim \chi^2(1) \tag{41}$$

The LM does not depend on ρ, its value tends asymptotically to 0.5 and its power is zero (second panel of Figure 7).

CASE 3: Model (iii), as it appears in (29), there are regressors in the rhs of the equation which have a non-zero variance.

The expression of the LM corresponds to the general case of (38). No significant simplifications can be obtained.

CASE 4: Model (iv), as it appears in (30), there is a constant term plus some other regressors with a non-zero variance in the rhs of the equation.

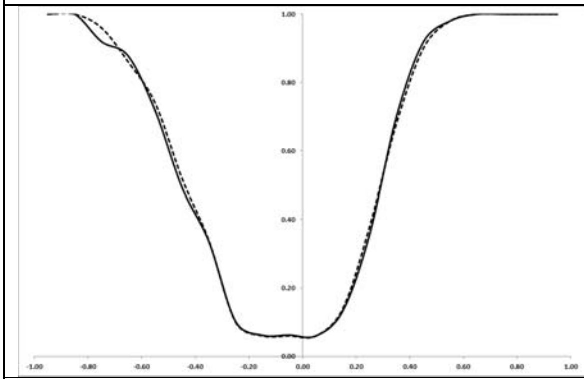
The expression of the LM appears in (41).

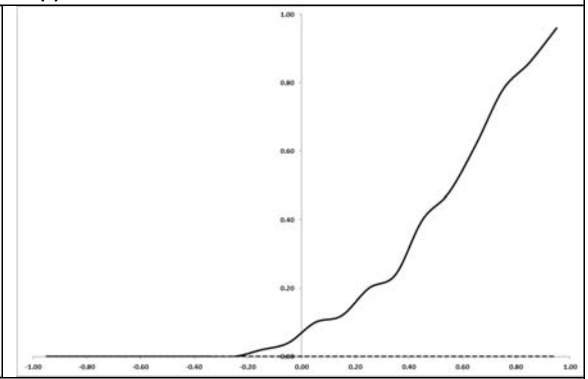
In sum, for a totally connected matrix the Moran's I statistic is always equal to zero irrespective of the characteristics of the DGP. This test must not be used when a highly dense matrix is in the testing framework. The behaviour of the LM_{LAG} depends on the use of a constant in the testing equation (or, what is the same, in the assumed DGP). If a constant term is introduced in the DGP, the value of the LM_{LAG} is always $\frac{1}{2}\left[\frac{n}{n-1}\right]^2$. The LM should not be use either in this case. The behaviour of the LM_{LAG}improves if the constant term is dropped from the testing equation; however, the power of the test, although satisfactory for the range of positive values of ρ, is almost zero for negative autocorrelation.The LM test offers only a limited coverage to testing for the presence of spatial autocorrelation in the case of fully connected matrices.

Figure 7: Spatial Autocorrelation tests: Moran's I and the LM.
Estimated power

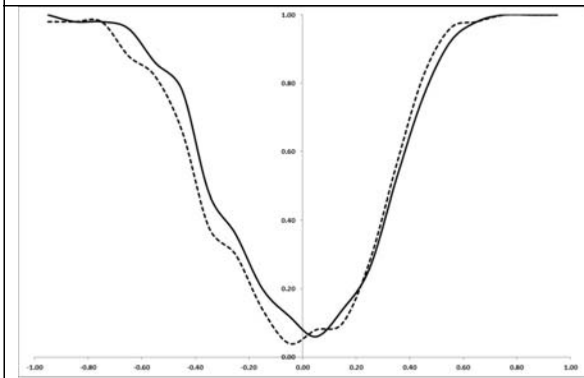
Connectivity = 5.0%	Connectivity = 99.5%

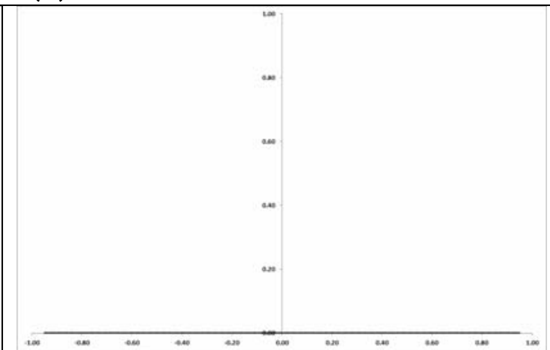
Model (i)



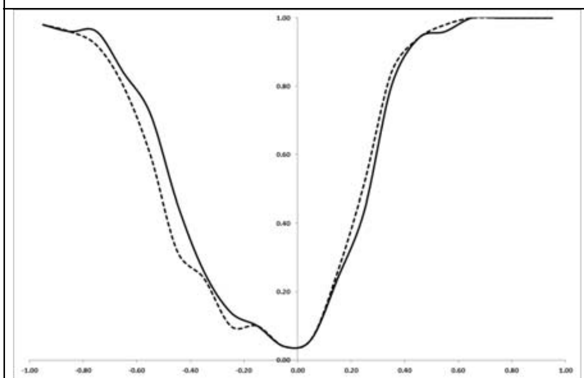


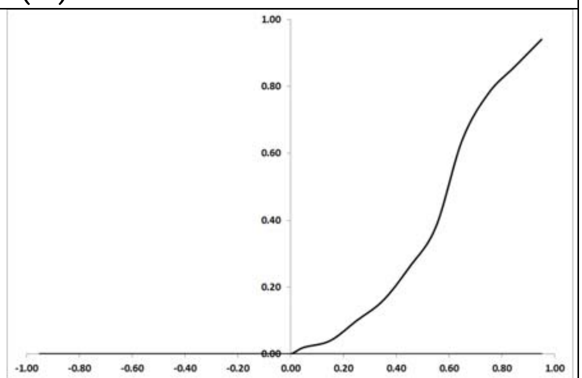
Model (ii)



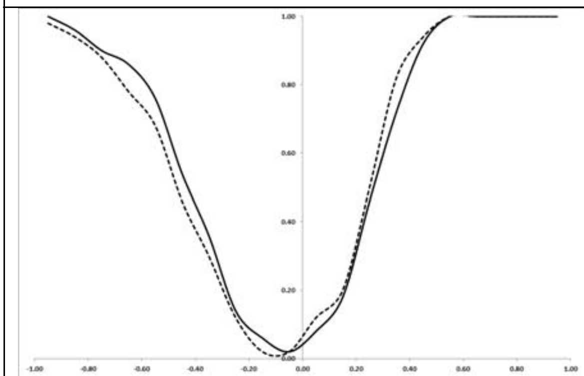


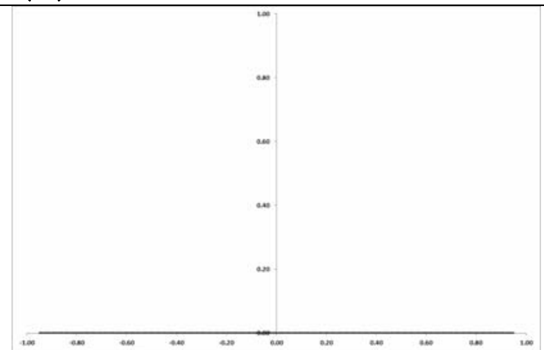
Model (iii)





Model (iv)





- - - Moran's —— LM