Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Article

A Comparison Study on Criteria to Select the Most Adequate Weighting Matrix

Marcos Herrera [1,‡], Jesus Mur [2,‡] and Manuel Ruiz [3,*]

1 CONICET-IELDE, Universidad Nacional de Salta (Argentina); mherreragomez@gmail.com
2 Universidad de Zaragoza (Spain); jmur@unizar.es
3 Universidad Politécnica de Cartagena (Spain); manuel.ruiz@upct.es
* Correspondence:manuel.ruiz@upct.es; Tel.: +34-968325901
‡ These authors contributed equally to this work.

Version November 8, 2018 submitted to Preprints

Abstract: The practice of spatial econometrics revolves around a weighting matrix, which is often supplied by the user on previous knowledge. This is the so called **W** issue. Probably, the aprioristic approach is not the best solution although, nowadays, there few alternatives for the user. Our contribution focuses on the problem of selecting a **W** matrix from among a finite set of matrices, all of them considerer appropriate for the case. We develop a new and simple method based on the Entropy corresponding to the distribution of probability estimated for the data. Other alternatives, which are common in current applied work, are also reviewed. The paper includes a large Monte Carlo to calibrate the effectiveness of our approach compared to the others. A well-known case study is also included.

Keywords: Weights matrix, Model Selection, Entropy, Monte Carlo

---

1. Introduction

Let us begin with a mantra: the weighting matrix is the most characteristic element in a spatial model. Most scholars agree with this popular commonplace. In fact, spatial models deal primarily with phenomena such as spillovers, trans-boundary competition or cooperation, flows of trade, migration, knowledge, etc. in complex networks. Rarely does the user know about how these events operate in practice. Indeed, they are mostly unobservable phenomena which are, however, required to build the model. Traditionally the gap has been solved by providing externally this information, in the form of a weighting matrix. As an additional remark, we should note that **W** is not the unique solution to deal with such kind of unobservables (1, for example, develop a latent variables approach that does not need of **W**), but is the most simple.

Hays *et al.* [2] give a sensible explanation about the preference for a predefined **W**. Network analysts are more interested in the formation of networks, taking units attributes and behaviors as given. This is spatial dependence due to selection, where relations of homophily and heterophily are crucial. The spatial econometricians are more interested in what they call *'computing the effects of alters actions on ego's actions through the network'*; in this case, the patterns of connectivity are taken as given. This form of spatial dependence is due to the influence between the individuals, and the notions of contagion and interdependence are capital. So, if the network is predefined, why not supplying it externally?

However, beyond this point, the **W** issue han been frequent cause of dispute. In the early stages, terms like 'join' or 'link' were very common (for instance, in 3, or 4). The focus at that time was mainly on testing for the presence of spatial effects, for which is not so important the specification of a highly detailed weighting matrix; contiguity, nearness, rough measures of separation may be appropriate notions for that purpose. The work of Ord [5] is a milestone in the evolution of this issue because of its strong emphasis on the task of modelling spatial relationships. It is evident that the weights matrix needs more attention if we want to avoid estimation biases and wrong inference. Anselin [67] puts

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

the **W** matrix in the center of the debate about specification of spatial models, but, after decades of practicing, the question still remains unclear.

The purpose of the so-called **W** is to '*determine which ... units in the spatial system have an influence on the particular unit under consideration ... expressed in notions of neighborhood and nearest neighbor*' relations, in words of Anselin [6, p.16] or '*to define for any set of points or area objects the spatial relationships that exist between them*' as stated by Haining [8, p. 74]. The problem is how should it be done.

Roughly speaking, we may distinguish two approaches: (i) specifying **W** exogenously; (ii) estimating **W** from data. The exogenous approach is by far the most popular and includes, for example, use of a binary contiguity criterion, k-nearest neighbours, kernel functions based on distance, etc. The second approach uses the topology of the space and the nature of the data, and takes many forms.We find ad-hoc procedures in which a predefined objective guides the search such as the maximization of Moran's I in Kooijman [9] or the local statistical model of Getis and Aldstadt [10]. Benjanuvatra and Burridge [11] develop a quasi maximum-likelihood, QML, algorithm to estimate the weights in **W** assuming partial knowledge about the form of the weights. More flexible approaches are possible if we have panel information such as in Bhattacharjee and Jensen-Butler [12] or Beenstock and Felsenstein [13]. Endogeneity of the weight matrix is another topic introduced recently in the field (i.e., 14), which connects with the concept of *coevolution* put forward by Snijders *et al.* [15] and based on the assumption that, in the long run, network connectivity must evolve endogenously with the model. Much of the recent literature on spatial econometrics revolves around endogeneity, but our contribution pertains to the exogenous approach where remains most part of the applied research.

Before continue, we may wonder if the **W** issue, even in our context of pure exogeneity, is really a problem to worry for or it is the *biggest myth* of the discipline as claimed by LeSage and Pace [16]. Their argument is that only dramatic different choices for **W** would lead to significant differences in the estimates or in the inference. We partly agree with them in the sense that is a bit silly to argue whether it is better the 5 or the 6 nearest-neighbor matrix; surely there will be almost no difference between the two. However, there is consistent evidence, obtained mainly by Monte Carlo [17–20] showing that the misspecification of **W** has a non-negligeable impact on the inference of the coefficients of spatial dependence and other terms in the model. Moreover, the magnitude of the bias increases for the estimates of the marginal direct/indirect effects. So, we are not pretty sure that '*far too much effort has gone into fine-tunning spatial weight matrices*' as stated by LeSage and Pace [16]. Our impression is that any useful result should be welcomed in this field and, especially, we need practical, clear guides to approach the problem.

Another question of concern are the criticisms of Gibbons and Overman [21]. As said, it is common in spatial econometrics to assume that the weighting matrix is known, which is the cause of identification problems; this flaw extends to the instruments, moment conditions, etc. There is little to say in relation to this point. In fact, spatial parameters (i.e., ρ) and weighting matrix, **W**, are only jointly identified (we do estimate $\rho\mathbf{W}$). Hays *et al.* [2] and Bhattacharjee and Jensen-Butler [12] agree in this point.

Bavaud [22, p. 153], given this controversial debate, was very skeptic, '*there is no such thing as "true", "universal" spatial weights, optimal in all situations*' and continues by stating that the weighting matrix '*must reflect the properties of the particular phenomena, properties which are bound to differ from field to field*'. We share his skepticism; perhaps it would suffice with a 'reasonable' weighting matrix, the best among those considered. In practical terms, this means that the problem of selecting a weighting matrix can be interpreted as a problem of model selection. In fact, different weighting matrices result in different spatial lags of the variables included in the model and different equations with different regressors amounts to a model selection problem.

As said, our intention is to offer new evidence to help the user to select the most appropriate **W** matrix for the specification. Section 2 revises four selection criteria that fit well into the problem of selecting a weighting matrix from among a finite set of them. Section 3 presents the main features of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

the Monte Carlo solved in the fourth Section. Section 5 includes a well known case study which is revised in the light of our findings. Sixth Section concludes.

2. Criteria to select a W matrix from among a finite set

The **W** issue has been present in the literature on spatial econometrics since very early. However the case of choosing one matrix from among a finite set of them is relatively recent. First, we review the literature devoted to the *J* test and then we moved to the selection criteria, Bayesian methods and a new procedure based on Entropy.

Anselin [23] poses formally the problem suggesting a *Cox* statistic derived in a framework of non-nested models. Leenders [24], on this basis, elaborates a *J*-test using classical augmented regressions. Later on, Kelejian [25] extends the approach of Leenders to a *SAC* model, with spatial lags of the endogenous variable and in the error terms, using *GMM* estimates. Piras and Lozano [26] confirm the adequacy of the *J*-test to compare different weighting matrices stressing that we should make use of a full set of instrument to increase *GMM* accuracy. Burridge and Fingleton [27] show that the Chi-square asymptotic approximations for the *J*-tests produces irregular results, excessively liberal or conservative in a series of leading cases; they suggest a bootstrap resampling approach. Burridge [28] focuses on the propensity of the spatial *GMM* algorithm to deliver spatial parameter estimates lying outside the invertibility region which, in turn, affects the bootstrap; he suggest the use of a *QML* algorithm to remove the problem. Kelejian and Piras [29] generalized and modify the original version of Kelejian to account for all the available information, according to the findings of Piras and Lozano. Finally, Kelejian and Piras [30] adapt the *J* test to a panel data setting with unobserved fixed effects and additional endogenous variables which reinforces the adequacy of the *GMM* framework. Another milestone in the *J* test literature is Hagemann [31], who copes with the reversion problem originated by the lack of a well defined null hypothesis in the test. He introduces the minimum *J* test, *MJ*. His approach is based on the idea that if there is a finite set of competing models, only the model with the smallest *J* statistic can be the correct one. In this case, the *J* statistic will converge to the Chi-square distribution but will diverge if none of the models is correct. The author proposes a wild bootstrap to test if the model with the minimum *J* is correct. This approach has been applied by Debarsy and Ertur [20] to a spatial setting with good results.

In the Monte Carlo that follows, we know that there is a correct model so are going to use only the first part of the procedure of Hagemann. Assuming a collection of m different weighting matrices, such as: $\mathcal{W} = \{\mathbf{W}_1; \mathbf{W}_2; ...; \mathbf{W}_m\}$, the MJ approach works as follows:

1. We need the estimates of the m models; in each case, the same equation is employed but with a different weighting matrix belonging to $\mathcal{W}$. Following Burridge [28] and given that our interest lies on the small sample case, the models are estimated by *ML*.
2. For each model, we obtain the battery of *J* statistics as usual, after estimating, also by *ML*, the corresponding extended equations.
3. The chosen matrix is the one associated with the minimum *J* statistic. We do not test if this matrix is really the correct matrix.

Another popular method for choosing between models deals with the so-called *Information Criteria*. Most are developed around a loss function, such as the *Kullback-Leibler*, *KL*, quantity of information which measures the closeness of two density functions. One of them corresponds to the true probability distribution that generated the data, obviously not known, the other is the distribution estimated from the data. The criteria differ in the role assigned to the aprioris and in the way of solving the approximation to the unknown true density function [32]. The two most common procedures are the *AIC* [33] and the Bayesian *BIC* criteria [34]. The first compares the models on equal basis whereas the second incorporates the notion of model of the null. Both criteria are characterized by their lack of specificity in the sense that the selected model is the closest to the true model, as measured by *KL*. We should note that, as indicated by Potscher [35], a good global fit does not mean that the model is the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

best alternative to estimate the parameters of interest. *AIC* and *BIC* lead to single expressions that depend on the accuracy of the *ML* estimation plus a penalty term related to the number of parameters entering the model; that is:

$$\left.\begin{array}{ll} AIC(k): & -2l(\widetilde{\gamma}) + 2k, \\ BIC(k): & -2l(\widetilde{\gamma}) + k\log(n), \end{array}\right\} \tag{1}$$

where $l(\widetilde{\gamma})$ means the estimated log-likelihood at the *ML* estimates, $\widetilde{\gamma}$, k is the number of non-zero parameters in the model and n the number of observations. For the case that we are considering the models only differ in the weighting matrix, so k and n are the same in every case. This means that the decision depends on the estimated log-likelihood, or on the balance between the estimated variance and the Jacobian term. Note that, for a standard spatial model of, i.e., *SLM* type we can write: $l(\widetilde{\gamma}) \propto \log\left[\frac{1}{\widetilde{\sigma}^n}\left|I - \widetilde{\rho}\mathbf{W}\right|\right]$, being σ the standard deviation and ρ the corresponding spatial dependence coefficient. To minimize any of the two statistics in (1) the objective is to maximize the concentrated estimated log-likelihood, $l(\widetilde{\gamma})$. The same as before, the *Information Criteria* approach implies:

1. Estimate by *ML* of the m models corresponding to each weighting matrix in $\mathcal{W}$.
2. For each model, we obtain the corresponding *AIC* statistic (*BIC* produces the same results).
3. The matrix in the model with minimum *AIC* statistic should be chosen.

An important part of the recent literature on spatial econometrics has Bayesian basis; this extends also to the topic of choosing a weighting matrix. The Bayesians are well equipped to cope with these type of problems using the concept of *posterior probability* as the basis for taking a decision. There are excellent reviews in Hepple [363738], Besag and Higdon [39] and especially, LeSage and Pace [40]. For the sake of completeness, let us highlight the main points in this approach.

The analysis is made conditional to a model, which is not under discussion. Moreover, we have a collection of m weighting matrices in $\mathcal{W}$, a set of k parameter in γ, some of which are of dispersion, σ, others of position, β, and others of spatial dependence, ρ and θ, and a panel data set with nT observations in y. The point of departure is the joint probability of data, parameters and matrices:

$$p\left(\mathbf{W}_i;\gamma;y\right) = \pi\left(\mathbf{W}_i\right)\pi\left(\gamma \mid \mathbf{W}_i\right)L\left(y \mid \gamma;\mathbf{W}_i\right), \tag{2}$$

where $\pi(\cdot)$ are the prior distributions and $L\left(y \mid \gamma;\mathbf{W}_i\right)$ the likelihood for y conditional on the parameters and the matrix. Bayes' rule leads to the posterior joint probability for matrices and parameters:

$$p\left(\mathbf{W}_i;\gamma \mid y\right) = \frac{\pi\left(\mathbf{W}_i\right)\pi\left(\gamma \mid \mathbf{W}_i\right)L\left(y \mid \gamma;\mathbf{W}_i\right)}{L\left(y\right)}, \tag{3}$$

whose integration over the space of parameters, $\gamma \in \Upsilon$, produces the posterior probability for matrix $\mathbf{W}_i$:

$$p\left(\mathbf{W}_i \mid y\right) = \int_{\Upsilon} p\left(\mathbf{W}_i;\gamma \mid y\right)d\gamma. \tag{4}$$

The presence of spatial structures in the model complicates the resolution of (4) which usually requires of numerical integration. The priors are always a point of concern and, usually, practitioners prefer diffuse priors. LeSage and Pace [40, Section 6.3] suggest $\pi\left(\mathbf{W}_i\right) = \frac{1}{m}\ \forall i$, a *NIG* conjugate prior for β and σ where $\pi_\beta\left(\beta \mid \sigma\right) \sim N\left(\beta_0;\sigma^2\left(\kappa X'X\right)^{-1}\right)$, being X the matrix of the exogenous variables in the model, and $\pi\left(\sigma\right)$ a inverse gamma, $IG(a,b)$. For the parameter of spatial dependence they suggest a $Beta(d,d)$ distribution, being d the amplitude of the sampling space of ρ. The defaults in the MATLAB codes of LeSage [41] are $\beta_0 = 0$, $\kappa = 10^{-12}$ and $a = b = 0$. In sum, the *Bayesian* approach implies the following:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

1. Fix the priors for all the terms appearing in the equation. In this point, we have followed the suggestions of LeSage and Pace.
2. For each matrix, obtain the corresponding posterior probability of (4) for which we need (i) solve the *ML* estimation of the corresponding model and (ii) solve the numerical integration of (4).
3. The matrix chosen will be that associated with the highest posterior probability.

This paper advocates for a selection procedure based on the notion of *Entropy*, which is used as a measure of the information contained in a distribution of probability. Let us assume an univariate continuous variable, y, whose probability density function is $p(y)$; then, *Entropy* is defined as:

$$h(p) = -\int_I p(y) \log p(y) dy, \tag{5}$$

being I the domain of the random variable y. As known, higher *Entropy* means less information or, what is the same, more uncertainty about y. Our case fits with Shannon's framework (42): we observe a random variable, y, and there is a finite set of rival distribution functions capable of having generated the data. Our decision problem is well defined because each distribution function differs from the others only in the weighting matrix; the other elements are the same. However, we are not interested in the Laplacian principle of indifference (select the density with maximum *Entropy*, less informative, to avoid uncertain information). Quite the opposite: in our case there is no uncertain information and we are looking for the more informative probability distribution so our objective is to minimize *Entropy*.

As with the other three cases, the application of this principle requires the complete specification of the distribution function, which means that the user knows the form of the model (equations 7 to 9 below, except the **W** matrix); additionally we add a Gaussian distribution. Next, we should remind that for the case of a $(n \times 1)$ multivariate normal random variable, $y \sim N(\mu; \Sigma)$, the entropy is: $h(y) = \frac{1}{2}\left[n + \log\left((2\pi)^n |\Sigma|\right)\right]$. This measure does not depend, directly, on first order moments (parameters of position of the model) but on second order moments (dependence and dispersion parameters). For example, in the case of the *SLM* of (9) below, the entropy is:

$$h(y)_{SDM} = \frac{1}{2}\left(nT + \log((2\pi\sigma^2)^{nT}\left|(B'B)^{-1}\right|)\right) \tag{6}$$

where $B = (I - \rho\mathbf{W})$. Note that the covariance matrix for y in the *SDM* is $V(y) = B^{-1}V(u)B'^{-1}$. If u is indeed a white noise random term with variance σ^2, the covariance matrix of y is simply $V(y) = \sigma^2 (B'B)^{-1}$. Let us note that the covariance matrix of y in the *SDM* of (7) coincides with that of the *SLM* case. The covariance matrix for the *SDEM* equation is $V(y) = \sigma^2 (B'B)$, everything else remains the same.

In order to apply the *Entropy* criterion we must must go through the following steps:

1. Estimate each one of the m versions of the model that we are considering. As said, each models differs only in the weighting matrix. We obtain the *ML* estimates for reasons given above.
2. For each model, we obtain the corresponding value of the *Entropy*, in the h_i; $i = 1, 2, ..., m$ statistic.
3. The decision criterion consists in choosing the weighting matrix corresponding to the model with minimum value of the *Entropy*.

3. Description of the Monte Carlo

This part of the paper is devoted to the design of the Monte Carlo conducted in the next Section in order to to calibrate the performance of the four criteria presented so far for selecting **W**: the *MJ* procedure, the *Bayesian* approach, the *AIC* criterion and the *Entropy* measure. The objective of the analysis is to identify and select the matrix that intervened in the generation of the data. Moreover, our focus is on small sample sizes. As will be clear below, the four criteria have good behaviour even in small samples, so it is not necessary to employ very large sample sizes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

We are going to simulate a panel setting, with three of the most common *DGPs* in the applied literature on spatial econometrics; namely, the spatial Durbin Model, *SDM* of (7), the spatial Durbin error model, *SDEM* in expression (8) and the spatial lag model of (9), *SLM*.[1]

$$y_{it} = \beta_0 + \rho \sum_{j=1}^{n} \omega_{ij} y_{jt} + x_{it}\beta_1 + \theta \sum_{j=1}^{n} \omega_{ij} x_{jt} + \varepsilon_{it}, \tag{7}$$

$$y_{it} = \beta_0 + x_{it}\beta_1 + \theta \sum_{j=1}^{n} \omega_{ij} x_{jt} + u_{it},\ u_{it} = \rho \sum_{j=1}^{n} \omega_{ij} u_{jt} + \varepsilon_{it}. \tag{8}$$

$$y_{it} = \beta_0 + \rho \sum_{j=1}^{n} \omega_{ij} y_{jt} + x_{it}\beta_1 + \varepsilon_{it}, \tag{9}$$

Only one exogenous regressor, x variable, appears in the right hand side of the equations whose observations are obtained from a normal distribution, $x_{it} \sim i.i.d.N\left(0;\sigma_x^2\right)$, where $\sigma_x^2 = 1$; the same applies with respect to the error terms: $\varepsilon_{it} \sim i.i.d.N\left(0;\sigma_\varepsilon^2\right)$, where $\sigma_\varepsilon^2 = 1$. The two variables are not related, $E\left(x_{it}\varepsilon_{it}\right) = 0$. Our space is made of hexagonal pieces which are arranged regularly, one next to the others without discontinuities nor empty spaces.

One weighting matrix appears in the three equations, which plays a central role in the functioning of the model. As said before, the weighting matrix is not observable and the user must take decisions to resolve the uncertainty. The problem consists in choosing one matrix from among a finite set of alternatives which in our simulation are composed by three candidates: $\mathbf{W}_1$ is built using the traditional contiguity criterion between spatial units; the weights in $\mathbf{W}_2$ are the inverse of the distance between the centroids of the spatial units, $\mathbf{W}_2 = \left\{\omega_{ij} = \frac{1}{d_{ij}}; i \neq j\right\}$; whereas $\mathbf{W}_3$ incorporates a cut-off point to the connections in $\mathbf{W}_2$, so that $\mathbf{W}_3 = \left\{\omega_{ij} = \frac{1}{d_{ij}}; i \neq j\ if\ j \in N_4(i);\ 0\ otherwise\right\}$ being $N_4(i)$ the set of 4 nearest neighbors to i. To keep things simple, the same weighting matrix plays with the endogenous and exogenous variables in (7) and with the exogenous and error terms in (8). Following usual practice, every matrix has been row-standardized. Due to the row-standardization, the three matrices are non nested in the sense that the sequence of weights are different among them.

Three different small cross-sectional sample sizes, n, have been used $n \in \{25, 49, 100\}$; that is enough because higher values of this parameter only improves marginally the results. For the same reason, the number of cross-sections in the panel, T, are limited to only three, $T \in \{1, 5, 10\}$. The values for the coefficient of spatial dependence, ρ, ranges from negatives to positives, $\rho = \{-0.8, -0.5, -0.2, 0.2, 0.5, 0.8\}$. Other global parameters are those associated with the constant term, $\beta_0 = 1$, the x variable, $\beta_1 \in \{1, 5\}$, and its spatial lag, $\theta \in \{1, 5\}$.

In sum, each case consists in:

- Generate the data using a given weighting matrix, $\mathbf{W}_k$, $k = 1, 2, 3$ and a spatial equation, *SLM*, *SDM* or *SDEM*. There are 216 cases of interest for each equation (6 values in ρ, 3 values in n, 3 values in T, 2 values in β_1 and 2 values in θ).
- The spatial equation is assumed to be known so the model can be estimated by maximum likelihood, *ML*, once the user supplies a **W** matrix.
- Compute the four selection criteria, *MJ*, *Posterior probability*, *Entropy* and *AIC* for the three alternative weighting matrices for each draw.
- Select the corresponding matrix according to each criterion and compare the result with the *true* matrix in the *DGP*.
- The process has been replicated 1,000 times.

[1] Main conclusions can be extended to other processes like the spatial error model, which are not replicated here (details on request from the authors).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Note that the selection of the matrix is made conditional on a correct specification of the equation. We are perfectly aware that this dichotomy is artificial; in fact, both decisions are intimately related because the same matrix, but in different equations, plays different roles and bears different information. However, this point is not further developed in the present paper. In order to give some intuition, we include the results corresponding to the case of a wrong specification (i.e, estimate a SDM model whereas the true model in the DGP is a $SDEM$).

4. Results of the Monte Carlo

This Section summarizes the results obtained in the Monte Carlo. Let us advance an little spicy: in strictly quantitative terms, the *Entropy* measure is the best criterion. What is more surprising, the *Bayesian* approach is marginally better than the AIC, but only when the amount of information is large and there is positive spatial correlation. Finally, the MJ approach is the worse alternative among the four criteria. The last two observations are a bit surprising given the strong support that the two procedures have received in the literature. Table 1 presents the percentage of correct selections attained by each criterion after aggregating all the experiments in the Monte Carlo. A cell in bold indicates that the respective criterion reaches the maximum rate of correct selections.

Table 1. Percentage of correct selections. Aggregated results

ρ	$h(y)$	*Bayes*	MJ	AIC
−0.8	83.8	83.2	50.7	**84.4**
−0.5	**71.4**	69.7	52.8	**71.4**
−0.2	**55.9**	49.4	54.2	54.6
0.2	**60.8**	54.6	58.3	60.5
0.5	**75.7**	73.6	58.2	73.5
0.8	**85.9**	85.4	53.6	78.7
$AVERAGE$	**72.3**	69.3	54.6	70.5

Entropy dominates in 5 out of the 6 cases presented in the Table, and is the second in the sixth case; AIC leads in two cases, is second in two and third in another two cases. *Bayes* does not do very well for small values of the spatial coefficient (is fourth in ± 0.2) and the curve of correct selections of the MJ is very flat.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Figure 1. Percentages of correct selections, disaggregated by n and T

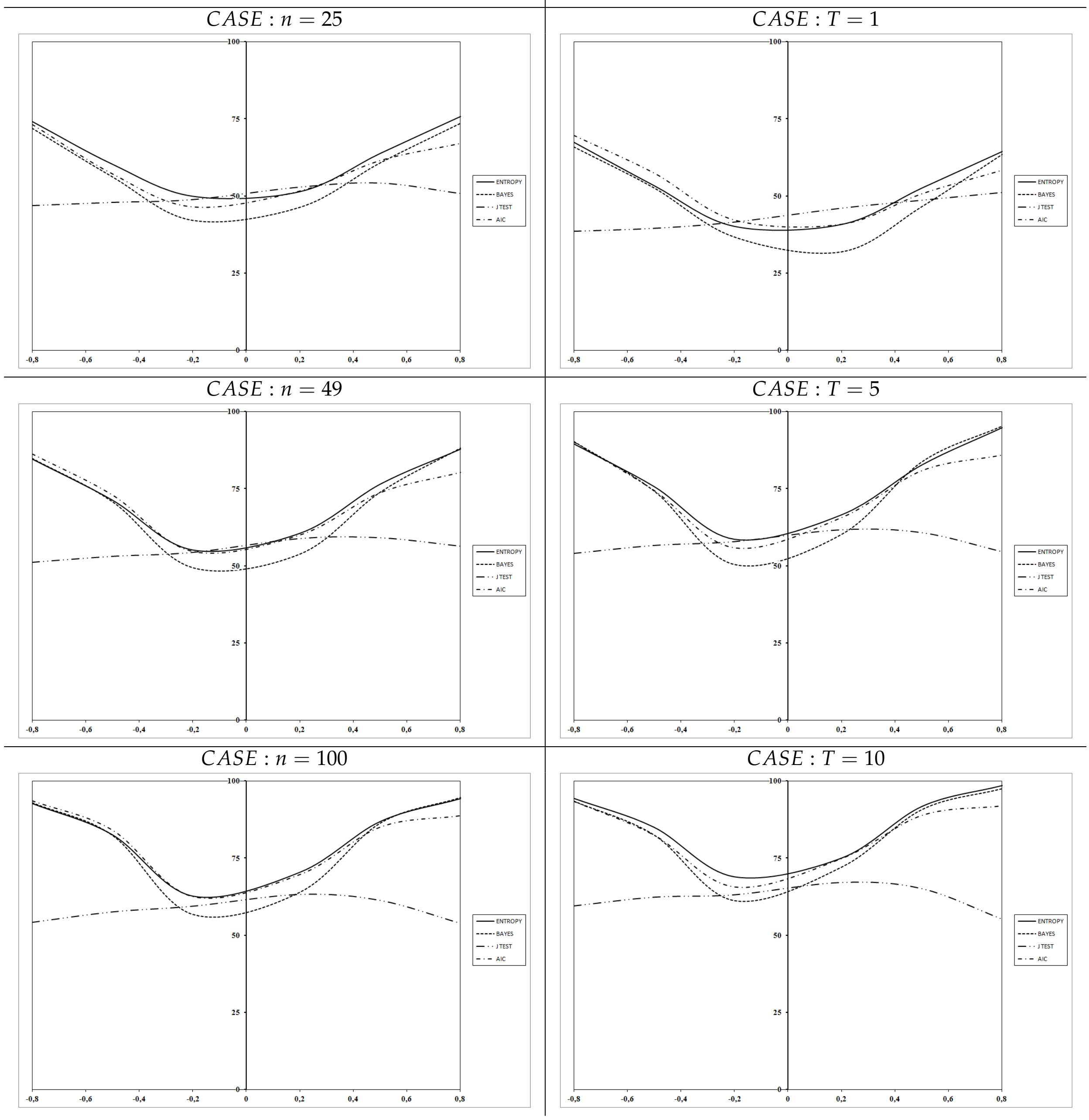

Figure 1 disaggregates the accumulated percentages by number of spatial units, left, or number of cross-sections, right. Note that in each case, the data represent aggregated percentages (i.e, in the case $n = 25$ we aggregate the three cross-sections corresponding to $T = 1$, $T = 5$ and $T = 10$). These courves ratifies the ordering set out above. Note the asymmetry in all the curves and the strange behaviour of the MJ criterion that produces worst results at the extremes of the interval for ρ. The other three criteria react positively to increases in the sample size (both in n or in T). Overall, the improvement is more relevant according to T than to n, specially for high values of the spatial coefficient.

Tables 2 to Table 5 present the details by type of DGP. A quick look at the Tables reveals that bold percentages are concentrated, mainly, in the *Entropy* and AIC columns.

The prevalence of the *Entropy* criterion is quite regular (the exception is the $SDEM$ process where AIC has better results). The preference extends to the case of correctly specified models, as in Tables

2, 3 and 4, and also for misspecified equations, as in Table 5, for negative and especially for positive values of the spatial coefficient, for small and large number of individuals in the sample (n) and for simple to large panels (T). Overall, *Entropy* attains the highest rate in 48% of the 180 cases in Tables 2 to 5.

The complete relation of results for the 864 different experiments in the MC (3 ns, 3 Ts, 6 ρs, 2 βs, 2 θs and four configurations for the DGP/estimated equation pair) appear in Tables 10 to 21 in the Appendix. Let us note the good results attained in the case of small samples ($n = 25$ and $T = 1$) where the average rate of correct selections for *Entropy* and AIC is above 40% criteria (a little worse for the other two). The percentage exceeds 60% at the extremes of the spatial parameter interval, ± 0.8. The average rate improves upto 65% - 75%, for the case of $n = 25$ and $T = 5$ and continues improving when $T = 10$, where most cases have a rate of correct selections above 90%. In general, the rate of correct selections is nearly 100%, using 5 to 10 cross-sections.

Table 2. Average percentage of correct selections. DGP: SDM. Equation estimated: SDM.

Aggregated by cross-section, sample size (n)						Aggregated by time series, sample size (T)					
	ρ	$h(y)$	*Bayes*	MJ	AIC		ρ	$h(y)$	*Bayes*	MJ	AIC
$n = 25$	−0.8	78.1	77.8	52.4	**79.6**	$T = 1$	−0.8	67.4	66.2	39.4	**68.8**
	−0.5	**62.9**	62.5	52.0	61.8		−0.5	54.4	54.3	38.5	**57.5**
	−0.2	**53.5**	48.7	53.1	50.2		−0.2	41.1	38.4	40.0	**41.7**
	0.2	61.5	59.8	**65.0**	61.2		0.2	43.2	35.8	**48.4**	40.8
	0.5	**74.7**	56.5	50.8	72.1		0.5	**56.5**	50.8	55.2	54.3
	0.8	**84.3**	81.7	74.5	75.5		0.8	**69.7**	68.1	63.4	63.8
$n = 49$	−0.8	88.9	88.7	57.6	**90.1**	$T = 5$	−0.8	91.9	93.0	62.3	**93.5**
	−0.5	76.4	77.5	58.6	**78.7**		−0.5	79.4	**80.2**	63.7	79.5
	−0.2	**59.6**	55.5	58.6	58.8		−0.2	**63.3**	57.3	62.4	60.1
	0.2	71.0	67.9	**73.1**	70.0		0.2	**79.1**	76.6	78.1	76.1
	0.5	**84.1**	81.7	81.6	82.0		0.5	92.3	**92.5**	87.1	89.8
	0.8	93.3	**93.8**	88.1	87.4		0.8	**98.4**	98.3	88.2	89.9
$n = 100$	−0.8	94.4	94.3	63.9	**95.2**	$T = 10$	−0.8	97.3	97.4	69.2	**97.9**
	−0.5	87.3	87.2	66.6	**88.7**		−0.5	**88.8**	88.7	72.1	87.9
	−0.2	**67.6**	61.9	62.8	66.6		−0.2	**72.3**	66.5	68.2	69.8
	0.2	**80.5**	76.4	79.4	77.1		0.2	86.6	**87.7**	86.4	87.4
	0.5	**91.9**	90.5	85.6	89.7		0.5	97.0	**97.5**	92.9	95.4
	0.8	**97.3**	96.3	89.5	92.4		0.8	**99.8**	**99.8**	94.6	95.8

Table 3. Average percentage of correct selections. DGP: SDEM. Equation estimated: SDEM.

Aggregated by cross-section, sample size (n)						Aggregated by time series, sample size (T)					
	ρ	$h(y)$	*Bayes*	MJ	AIC		ρ	$h(y)$	*Bayes*	MJ	AIC
$n = 25$	−0.8	80.5	77.3	56.7	**82.5**	$T = 1$	−0.8	66.7	65.3	42.5	**70.4**
	−0.5	**69.6**	65.2	57.5	**69.6**		−0.5	55.4	55.6	44.0	**62.1**
	−0.2	**59.6**	52.5	56.5	58.2		−0.2	42.5	42.8	43.8	**49.3**
	0.2	55.6	52.4	56.9	**57.7**		0.2	39.5	36.1	**45.7**	43.4
	0.5	63.5	62.6	55.7	**63.7**		0.5	**49.5**	45.4	46.5	48.7
	0.8	**74.4**	73.8	54.0	67.0		0.8	**59.3**	58.1	48.9	53.2
$n = 49$	−0.8	88.1	88.5	64.5	**91.0**	$T = 5$	−0.8	94.0	94.7	71.1	**95.4**
	−0.5	78.2	78.8	65.6	**81.9**		−0.5	84.1	84.8	72.5	**84.9**
	−0.2	64.6	62.6	65.2	**66.3**		−0.2	70.2	67.1	**71.6**	69.6
	0.2	64.8	61.4	65.2	**65.9**		0.2	71.2	69.3	70.3	**73.1**
	0.5	**78.0**	75.7	64.5	75.1		0.5	83.1	**85.5**	67.8	83.7
	0.8	**88.0**	87.1	64.0	79.6		0.8	94.7	**95.4**	64.6	86.8
$n = 100$	−0.8	95.1	95.8	75.1	**96.4**	$T = 10$	−0.8	97.7	98.2	78.9	**98.6**
	−0.5	88.9	91.3	76.4	**92.1**		−0.5	**92.4**	91.6	79.3	91.7
	−0.2	74.2	75.1	76.4	**77.2**		−0.2	**81.5**	77.4	78.8	78.3
	0.2	75.6	74.9	75.1	**78.3**		0.2	81.4	80.3	77.3	**81.7**
	0.5	87.9	**89.1**	72.9	88.1		0.5	93.3	**93.5**	75.2	91.3
	0.8	94.3	**95.6**	69.4	90.7		0.8	99.1	**99.4**	70.0	93.9

Table 4. Average percentage of correct selections. DGP: SLM. Equation estimated: SLM.

Aggregated by cross-section, sample size (n)						Aggregated by time series, sample size (T)					
	ρ	$h(y)$	*Bayes*	*MJ*	*AIC*		ρ	$h(y)$	*Bayes*	*MJ*	*AIC*
$n = 25$	-0.8	58.7	**59.7**	27.3	58.3	$T = 1$	-0.8	54.0	53.1	23.7	**54.5**
	-0.5	**41.9**	36.2	27.0	38.4		-0.5	36.3	33.3	24.3	**37.7**
	-0.2	**28.3**	15.8	**28.3**	26.5		-0.2	22.6	14.2	**28.3**	22.5
	0.2	33.4	21.0	30.2	**33.5**		0.2	30.8	12.6	**32.5**	30.0
	0.5	54.0	49.6	31.8	**54.2**		0.5	**46.3**	37.4	34.6	45.4
	0.8	**72.4**	70.8	31.9	70.0		0.8	**61.0**	**61.0**	36.2	56.6
$n = 49$	-0.8	73.6	73.2	22.1	**74.4**	$T = 5$	-0.8	79.7	**81.4**	19.9	80.7
	-0.5	**53.2**	47.7	25.4	51.5		-0.5	**57.9**	53.0	24.1	55.9
	-0.2	**32.3**	17.9	28.9	30.5		-0.2	**32.8**	15.1	28.1	30.2
	0.2	**41.7**	24.9	31.5	39.9		0.2	**44.4**	27.4	29.2	43.0
	0.5	**68.8**	64.3	26.9	67.7		0.5	**73.0**	72.8	24.4	71.5
	0.8	86.8	**87.1**	26.2	82.1		0.8	**93.5**	93.3	24.8	88.3
$n = 100$	-0.8	86.7	87.0	12.0	**87.6**	$T = 10$	-0.8	**85.4**	85.3	17.7	85.1
	-0.5	68.0	65.2	18.0	**68.3**		-0.5	**68.8**	62.8	22.0	64.7
	-0.2	**37.8**	22.2	27.1	36.3		-0.2	**43.0**	26.6	28.0	40.6
	0.2	**51.3**	35.4	27.6	50.6		0.2	**51.1**	41.3	27.6	**51.1**
	0.5	**81.4**	79.8	20.6	79.0		0.5	**84.9**	83.4	20.2	84.0
	0.8	92.3	**92.9**	20.2	86.9		0.8	**97.0**	96.5	17.3	94.0

Table 5. Average percentage of correct selections. DGP: SDEM. Equation estimated: SDM.

Aggregated by cross-section, sample size (n)						Aggregated by time series, sample size (T)					
	ρ	$h(y)$	*Bayes*	*MJ*	*AIC*		ρ	$h(y)$	*Bayes*	*MJ*	*AIC*
$n = 25$	-0.8	**79.2**	77.2	51.4	77.6	$T = 1$	-0.8	66.2	66.0	38.1	**68.9**
	-0.5	**66.5**	64.1	55.1	62.3		-0.5	54.3	55.5	40.8	**58.1**
	-0.2	**57.9**	52.4	57.4	52.7		-0.2	42.2	40.2	42.3	**42.7**
	0.2	54.8	54.0	**59.1**	55.4		0.2	38.2	32.3	**44.8**	37.7
	0.5	**62.8**	62.4	55.9	59.4		0.5	**46.4**	42.3	45.4	43.6
	0.8	71.9	**72.2**	42.5	59.4		0.8	**55.1**	54.1	42.5	47.3
$n = 49$	-0.8	87.9	88.6	60.4	**89.9**	$T = 5$	-0.8	**92.3**	**92.3**	64.3	91.7
	-0.5	77.5	78.9	62.7	**79.8**		-0.5	**83.3**	81.6	67.4	80.0
	-0.2	64.2	61.3	**64.7**	63.1		-0.2	69.4	63.4	**69.5**	64.5
	0.2	64.8	60.7	**65.5**	64.1		0.2	**71.9**	68.5	69.4	70.9
	0.5	**75.1**	73.5	63.2	70.2		0.5	82.2	**83.9**	65.7	78.2
	0.8	**84.3**	**84.3**	47.1	72.1		0.8	92.2	**93.7**	44.3	77.9
$n = 100$	-0.8	94.9	95.0	67.0	**95.8**	$T = 10$	-0.8	**97.2**	**97.2**	72.6	96.9
	-0.5	87.9	88.6	70.4	**89.8**		-0.5	**89.9**	**89.9**	76.1	88.8
	-0.2	**72.4**	68.9	71.5	71.3		-0.2	**78.6**	74.9	77.7	75.6
	0.2	**74.9**	70.0	71.1	73.4		0.2	**80.8**	80.2	77.0	80.5
	0.5	**85.9**	**85.9**	68.2	83.3		0.5	91.8	**92.5**	72.0	87.9
	0.8	92.7	**93.3**	39.6	84.4		0.8	98.3	**98.8**	38.3	87.6

In a similar vein, the increase in the cross-sectional size, n, maintaining constant the number of cross-sections, T, also has positive effects in the four criteria. The rate of correct selections for the case of a hundred of spatial units is above 70%, on average, for the case of a single cross-section ($T = 1$), but these percentages improve quickly if the time dimension of the panel increases.

The value of parameter β_1, as expected, has a weak impact in the four criteria; on the contrary, the signal of θ_1 plays a crucial role in the $SDEM$ case. Another interesting feature is the asymmetry of the selection curves, that tends to be diluted with T. Negative spatial dependence helps to detect the correctly weighting matrix, especially when the number of time cross-sections is small. The asymmetry exists in Entropy, Bayes and AIC. However, the behavior of the MJ worsens in case of negative values in parameter ρ.

To complete the picture, we estimate a *response-surface* for each DGP/Estimated-equation combination, with the aim of modelling the empirical probability of choosing the correct weighting

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

matrix for each criterion, p_i. As usual, a logit transformation of the empirical probabilities is carried out, so the estimated equation is:

$$\log\left(\frac{p_i + (2r)^{-1}}{1 - p_i + (2r)^{-1}}\right) = p_i^* = \eta + z_i\varphi + \epsilon_i, \tag{10}$$

where p_i^* is the logit transformation, often known as the *logit*, r the number of replications of each experiment (1000 in all the cases); $(2r)^{-1}$ assures that the *logit* is defined even when the probability of correct selection is 0 or 1 (43); η is an intercept term, z_i the design matrix whose columns reflect the conditions of each experiment, φ is a vector of parameters and ϵ_i the error term assumed to be independent and identically distributed (this assumption is reasonable if all experiments come from the same study, as ours, and are obtained under identical circumstances; 44). Let us remind that the number of observations for each *response-surface* equation is 216 (so $i = 1, 2, ..., 216$). Table 6 shows the results for the four DGP/Estimated-equation combinations.

In general, the estimates confirm previous facts. The main factor influencing the empirical probability of choosing the correct weights matrix is the spatial parameter, absolute value of ρ in Table 6. Its contribution is crucial in the case of the *Bayesian* criteria and, to a lesser extend, also in the cases of *Entroy* and AIC. This parameter is not significant, for the case of the MJ approach and $SDEM$ processes whereas its contribution is negative in the SLM and in misspecified equations. The second more influential factor is the parameter θ, associated to spatial spillovers. Its impact is beneficial for all the cases though it appears to be more important for the MJ; the other three criteria are a bit less sensitive. Sample size is also relevant in all the cases and T has a relatively higher impact than n. Finally, as said before, parameter β_1 is not significant in any circumstance, with the exception of the SLM case; this means that the *signal-to-noise* ratio should not be a major factor to consider when the problem is select the best weighting matrix.

Table 7 completes the *response-surface* analysis with the F tests of equality in the coefficients of the estimates of Table 6. According to the sequence of F tests, the most dissimilar method is the MJ approach, and then *Bayes*. On the other hand, *Entropy* and AIC emerge as similar strategies to compare weighting matrices; in fact, in what respect this simple *response-surface* analysis, they are almost indistinguishable in the four types of $DGPs$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Table 6. Estimated response surfaces.

SDEM case	constant	n	T	β_1	θ	$\lvert\rho\rvert$	R^2	F_{AV}
Entropy	-5.9410	0.0037	0.0566	0.0005	0.0748	0.5568	0.74	117.90
	(0.0000)	(0.0000)	(0.0000)	(0.9402)	(0.0000)	(0.0000)		(0.0000
Bayes	-6.2233	0.0051	0.0660	-0.0017	0.0904	0.6813	0.66	81.57
	(0.0000)	(0.0000)	(0.0000)	(0.8553)	(0.0000)	(0.0000)		(0.0000
MJ test	-6.1295	0.0044	0.0520	0.0106	0.1569	-0.0377	0.82	196.74
	(0.0000)	(0.0000)	(0.0000)	(0.0910)	(0.0000)	(0.4612)		(0.0000
AIC	-5.9177	0.0043	0.0506	0.0044	0.0795	0.4590	0.67	87.21
	(0.0000)	(0.0000)	(0.0000)	(0.5407)	(0.0000)	(0.0000)		(0.0000
SDM case	constant	n	T	β_1	θ	$\lvert\rho\rvert$	R^2	F_{AV}
Entropy	-5.8902	0.0033	0.0481	0.0053	0.0702	0.06348	0.66	83.35
	(0.0000)	(0.0000)	(0.0000)	(0.4614)	(0.0000)	(0.0000)		(0.0000
Bayes	-6.1117	0.0033	0.0548	0.0052	0.0861	0.8116	0.60	63.33
	(0.0000)	(0.0000)	(0.0000)	(0.5974)	(0.0000)	(0.0000)		(0.0000
MJ test	-5.8998	0.0024	0.0476	0.0186	0.1036	0.1668	0.47	36.74
	(0.0000)	(0.0004)	(0.0000)	(0.0813)	(0.0000)	(0.0552)		(0.0000
AIC	-5.9339	0.0034	0.0479	0.0092	0.0722	0.6301	0.67	83.61
	(0.0000)	(0.0000)	(0.0000)	(0.2051)	(0.0000)	(0.0000)		(0.0000
SLM case	constant	n	T	β_1		$\lvert\rho\rvert$	R^2	F_{AV}
Entropy	-6.3435	0.0049	0.0613	-0.0390		1.2505	0.81	113.60
	(0.0000)	(0.0000)	(0.0000)	(0.0001)		(0.0000)		(0.0000
Bayes	-7.0854	0.0054	0.0786	-0.0709		2.2207	0.83	122.53
	(0.0000)	(0.0000)	(0.0000)	(0.0000)		(0.0000)		(0.0000
MJtest	1.3129	-0.0131	-0.0896	-0.2215		-1.4089	0.40	16.92
	(0.0000)	(0.0000)	(0.0006)	(0.0000)		(0.0004)		(0.0000
AIC	-6.3808	0.0050	0.0599	-0.0396		1.2678	0.79	96.74
	(0.0000)	(0.0000)	(0.0000)	(0.0003)		(0.0000)		(0.0000
MISS case	constant	n	T	β_1	θ	$\lvert\rho\rvert$	R^2	F_{AV}
Entropy	-5.9736	0.0039	0.0583	-0.0004	0.0745	0.5505	0.72	109.13
	(0.0000)	(0.0000)	(0.0000)	(0.9511)	(0.0000)	(0.0000)		(0.0000
Bayes	-6.1882	0.0040	0.0648	-0.0001	0.0916	0.7103	0.67	85.13
	(0.0000)	(0.0000)	(0.0000)	(0.9887)	(0.0000)	(0.0000)		(0.0000
MJ test	-5.6677	0.0020	0.0379	0.0007	0.1162	-0.3854	0.55	50.92
	(0.0000)	(0.0000)	(0.0000)	(0.9431)	(0.0000)	(0.0000)		(0.0000
AIC	-5.9741	0.0043	0.0558	-1.9169	0.0696	0.4728	0.68	88.38
	(0.0000)	(0.0000)	(0.0000)	(0.9979)	(0.0000)	(0.0000)		(0.0000

Note: pvalue appear between brackets. F_{AV} means F test of the null that all coefficients are zero except the constant. MISS means that the model in the DGP is a SDEM but we estimate a SDM equation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Table 7. F test for the equality of coefficients in the response-surface estimates

SDEM case	*Bayes*	*MJ test*	*AIC*
Entropy	55.000 (0.00)	87.331 (0.00)	**1.535 (0.17)**
Bayes	–	34.720 (0.00)	4.558 (0.00)
MJ test	–	–	61.774 (0.00)
SDM case	*Bayes*	*MJ test*	*AIC*
Entropy	4.699 (0.00)	34.886 (0.00)	**0.471 (0.83)**
Bayes	–	14.791 (0.00)	3.300 (0.00)
MJ test	–	–	28.553 (0.00)
SLM case	*Bayes*	*MJ test*	*AIC*
Entropy	61.544 (0.00)	8685.34 (0.00)	**0.500 (0.78)**
Bayes	–	432.170 (0.00)	45.475 (0.00)
MJ test	–	–	7423.01 (0.00)
MISS case	*Bayes*	*MJ test*	*AIC*
Entropy	4.454 (0.00)	118.882 (0.00)	**2.056 (0.06)**
Bayes	–	65.420 (0.00)	5.171 (0.00)
MJ test	–	–	85.234 (0.00)

Note: p-value appear between brackets.

5. Empirical application

The case studied in this section is based on a well-known economic model. It is a model of economic growth estimated by Ertur and Koch (2007) using a cross-section of 91 countries for the period 1960–1995. The purpose of this section is to illustrate the use of the selection procedures discussed before.

5.1. Study case: Ertur and Koch (2007)

Ertur and Koch [45] build a growth equation to model technological interdependence between countries using spatial externalities. The main hypotheses of interaction is that the stock of knowledge in one country produces externalities that cross national borders and spill over into neighboring countries, with an intensity which decreases with distance. The authors use a geographical distance measure.

The benchmark model assumes an aggregated Cobb-Douglas production function with constant returns to scale in labour and physical capital:

$$Y_i(t) = A_i(t)K_i^{\alpha}(t)L_i^{1-\alpha}(t), \tag{11}$$

where $Y_i(t)$ is output, $K_i(t)$ is the level of reproducible physical capital, $L_i(t)$ is the level of labour, and $A_i(t)$ is the aggregate level of technology specified as:

$$A_i(t) = \Omega(t)k_i^{\phi}(t)\prod_{j\neq i}^{n} A_i^{\delta\omega_{ij}}(t). \tag{12}$$

The aggregate level of technology $A_i(t)$ in a country i depends on three elements. First, a certain proportion of technological progress is exogenous and identical in all countries: $\Omega(t) = \Omega(0)e^{\mu t}$, where μ is a constant rate of technological growth. Second, each country's aggregate level of technology increases with the aggregate level of physical capital per worker $k_i^{\phi}(t) = (K_i(t)/L_i(t))^{\phi}$ with parameter $\phi \in [0;1]$ capturing the strength of home externalities by physical capital accumulation. Finally, the third term captures the external effects of knowledge embodied in capital located in a different country, whose impact crosses national borders at a diminishing intensity, $\delta \in [0;1]$. The terms ω_{ij} represent the connectivity between country i and its neighbours; these weights are assumed to be exogenous, non-negative and finite.

Following Solow, the authors assume that a constant fraction of output s_i, in every country i, is saved and that labour grows exogenously at the rate n_i. Also, they assume a constant and identical

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

annual rate of depreciation of physical capital for all countries, denoted τ. The evolution of output per worker in country i is governed by the usual fundamental dynamics of the Solow equation which, after some manipulations, lead to a steady-state real income per worker [45, p. 1038, eq. 9]:

$$y = \Omega + (\alpha + \phi)\, k - \alpha\delta \mathbf{W}k + \delta \mathbf{W}y. \tag{13}$$

This is a spatially augmented Solow model and coincides with the predictor obtained by Solow adding spillover effects. In terms of spatial econometrics, we have a *Spatial Durbin Model*, SDM, which can be expressed as:

$$y = x\beta + \rho \mathbf{W}y + \mathbf{W}x\theta + \varepsilon. \tag{14}$$

Equation (14) is estimated using information on real income, investment and population growth for a sample of 91 countries for the period $1960 - 1995$. Regarding the spatial weighting matrix, Ertur and Koch consider two geographical distance functions: the inverse of squared distance (which is the main hypothesis) and the negative exponential of squared distance (to check robustness in the specification). We also consider a third matrix based on the inverse of the distance.

Let us call the three weighting matrices as $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{W}_3$ which are row-standardized: $\omega_{hij} = \omega^*_{hij} / \sum_{j=1}^{n} \omega^*_{hij}$; $h = 1, 2, 3$ where:

$$\omega^*_{1ij} = \begin{cases} 0 & if\ i = j \\ d_{ij}^{-2} & otherwise \end{cases} ; \quad \omega^*_{2ij} = \begin{cases} 0 & if\ i = j \\ e^{-2d_{ij}} & otherwise \end{cases} ; \quad \omega^*_{3ij} = \begin{cases} 0 & if\ i = j \\ d_{ij}^{-1} & otherwise \end{cases} , \tag{15}$$

with d_{ij} as the great-distance between the capitals of countries i and j.

The authors analyze several specifications checking for different theoretical restrictions and alternative spatial equations. We concentrate our revision in the so-called non-restricted equation, in the sense that it includes more coefficients than advised by theory. Table 8 presents the SDM version of this equation using the three alternative weighting matrices specified before (the first two columns coincide with those in Table I, columns 3-4, pp. 1047, of 45). The last four rows in the Table show the value of the selection criteria corresponding to each case.

The preferred model by Ertur and Koch is the $SDM/\mathbf{W}_1$ which coincides with the selection attained by minimum *Entropy*, the *Bayesian* posterior probability and AIC. The selection of the MJ approach is $\mathbf{W}_2$.

Other results in Ertur and Koch refer to the Spatial Error Model version of the steady-state equation of (13), or SEM model. The intention of the authors is to visualize the presence of spatial correlation in the traditional non spatial Solow equations; we use this case as an example of selection of weighting matrices in misspecified models. The main results appear in Table 9 (which can be compared with columns 2-3 of Table II, in 45, p. 1048).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Table 8. Ertur & Koch case. Unrestricted SDM estimates

Model/Weight matrix	*SDM / W1*	*SDM / W2*	*SDM / W3*
constant	0.967 (0.51)	0.499 (0.27)	5.197 (0.99)
$\log(s)$	0.825 (8.26)	0.792 (7.62)	0.910 (8.49)
$\log(n+0.05)$	−1.498 (−2.64)	−1.451 (−2.62)	−1.710 (−2.67)
$\mathbf{W} \times \log(s)$	−0.326 (−1.78)	−0.378 (−2.29)	0.500 (1.25)
$\mathbf{W} \times \log(n+0.05)$	0.574 (0.68)	0.141 (0.18)	2.150 (1.01)
$\mathbf{W} \times \log(y)$	0.742 (10.70)	0.661 (9.01)	0.883 (11.60)
Selection Criteria			
Entropy	28.001	29.615	34.615
Bayesian	0.864	0.133	0.003
MJ	11.158	9.388	10.208
AIC	95.885	99.100	109.132

Note: t-ratios appear between brackets.

Table 9. Ertur & Koch case. Unrestricted SEM estimates

Model/Weight matrix	*SEM / W1*	*SEM / W2*	*SEM / W3*
constant	6.458 (4.23)	6.706 (4.62)	5.892 (3.02)
$\log(s_i)$	0.828 (8.37)	0.804 (7.88)	0.992 (8.95)
$\log(n_i+0.05)$	−1.702 (−3.03)	−1.553 (−2.85)	−2.269 (−3.65)
$\mathbf{W} \times \varepsilon_i$	0.823 (15.69)	0.737 (12.19)	0.937 (22.08)
Selection Criteria			
Entropy	30.973	31.734	42.049
Bayesian	0.690	0.310	0.000
MJ	$0.171e^{-12}$	$0.043e^{-12}$	$0.085e^{-12}$
AIC	97.870	99.391	120.021

Note: t-ratios appear between brackets.

The selection of the most adequate **W** matrix does not change. Using the values of *Entropy* criterion we select the model in which intervenes the matrix $\mathbf{W}_1$, the same as with the Bayesian approach and the AIC criterion; MJ continues selecting $\mathbf{W}_2$.

6. Conclusion

Much of the applied spatial econometrics literature seems to prefer an exogenous approximation to the **W** matrix. Implicitly, it is assumed that the user has relevant knowledge with respect to the way individuals in the sample interact. In recent years, new literature advocates for a more data driven approach to the **W** issue. We strongly support this tendency, which should be dominant in the future; however, our focus in this paper is on the exogenous approach.

The problem posed in the paper is very frequent in applied work: the user has a finite collection of weighting matrices, they all are coherent with the case of study, and one needs to select one of them. Which is the best **W**? We can address this question using different proposals: the *Bayesian* posterior probability, the J approach with all its variants, by means of simple model selection criteria, such as AIC or BIC and several other alternatives not used in this study. We add a fourth one, based on the *Entropy* of the estimated distribution function. This new criterion *h(y)* is a measure of uncertainty, and fits well with the **W** decision problem. The *h(y)* statistics depends on the estimated covariance matrix of the corresponding model offering a more complete picture of the suitability of the distribution function (related to a particular choice of **W**), to deal with the data at hand.

The conclusions of our Monte Carlo are very illuminating. First, we can confirm that it is possible to identify, with confidence, the true weighting matrix (if it exists); in this sense, the selection criteria do a good job. However, the four criteria should not be taken as indifferent, especially in samples of small size (n or T). The ordering is clear: *Entropy* in first place, AIC and *Bayesian* posterior probability

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

slightly worse, and then MJ in the fourth position. As shown in the paper, the value of the spatial parameter has a great impact to guarantee a correct selection, but this aspect is unobservable to the researcher. However, the user effectively controls the amount of information involved in the exercise, and this is also a key factor. The advice is clear: use as much information as you have because the quality of the decision improves with the amount of information. Once again, the way the information accrues is not neutral: the length of the time series in the panel is more relevant than the number of cross-sectional units in the sample.

Our final recommendation for applied researchers is to care for the adequacy of the weighting matrix and, in case of having various candidates, take a decision using well-defined criteria such as the *Entropy*. The empirical application presented in Section 5 illustrates the procedure.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Author Contributions: conceptualization; methodology; formal analysis; writing original draft, review and editing: Marcos Herrra, Jesus Mur and Manuel Ruiz

Funding: This research was funded by the Spanish Ministry of Economics grants numbers ECO2015-65758-P and ECO2015-65637-P and Government of Aragon grant number S37_17R. This study is part of the collaborative activities carried out under the program Groups of Excellence of the Region of Murcia, the Fundación Séneca, Science and Technology Agency of the Region of Murcia Project 19884/GERM/15.

Conflicts of Interest: The authors declare no conflict of interest.

7. References

1. Oud, J.H.; Folmer, H. A structural equation approach to models with spatial dependence. *Geographical Analysis* **2008**, *40*, 152–166.
2. Hays, J.C.; Kachi, A.; Franzese, R.J. A spatial model incorporating dynamic, endogenous network interdependence: A political science application. *Statistical Methodology* **2010**, *7*, 406–428.
3. Moran, P.A. The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)* **1948**, *10*, 243–251.
4. Whittle, P. On Stationary Processes in the Plane. *Biometrika* **1954**, *41*, 434–449.
5. Ord, K. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* **1975**, *70*, 120–126.
6. Anselin, L. *Spatial econometrics: Methods and models*; Vol. 4, Kluwer Academic Publishers, Dordrecht, 1988.
7. Anselin, L. Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics* **2002**, *27*, 247–267.
8. Haining, R.P. *Spatial data analysis: theory and practice*; Cambridge University Press, 2003.
9. Kooijman, S. Some remarks on the statistical analysis of grids especially with respect to ecology. *Annals of Systems Research* **1976**, *5*, 113–132.
10. Getis, A.; Aldstadt, J. Constructing the Spatial Weights Matrix Using a Local Statistic. *Geographical Analysis* **2004**, *36*, 90–104.
11. Benjanuvatra, S.; Burridge, P. Qml estimation of the spatial weight matrix in the mr-sar model. *Working Paper 15/24. Department of Economics and Related Studies, University of York* **2015**.
12. Bhattacharjee, A.; Jensen-Butler, C. Estimation of the spatial weights matrix under structural constraints. *Regional Science and Urban Economics* **2013**, *43*, 617–634.
13. Beenstock, M.; Felsenstein, D. Nonparametric estimation of the spatial connectivity matrix using spatial panel data. *Geographical Analysis* **2012**, *44*, 386–397.
14. Qu, X.; f. Lee, L. Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics* **2015**, *184*, 209–232.
15. Snijders, T.A.; Steglich, C.E.; Schweinberger, M., Longitudinal models in the Behavioural ond Related Sciences; Lawrence Erlbaum, 2007; chapter Modeling the co-evolution of networks and behavior, pp. 41–71.
16. LeSage, J.P.; Pace, R.K. The biggest myth in spatial econometrics. *Econometrics* **2014**, *2*, 217–249.
17. Florax, R.J.; Rey, S. The impacts of misspecified spatial interaction in linear regression models. In *New directions in spatial econometrics*; Anselin, L.; Florax, R., Eds.; Springer, 1995; pp. 111–135.
18. Franzese Jr, R.J.; Hays, J.C. Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis* **2007**, *15*, 140–164.
19. Lee, L.f.; Yu, J. QML estimation of spatial dynamic panel data models with time varying spatial weights matrices. *Spatial Economic Analysis* **2012**, *7*, 31–74.
20. Debarsy, N.; Ertur, C. Interaction matrix selection in spatial econometrics with an application to growth theory. *Document de Recherche du Laboratoire d'Economie d'Orleans* **2016**.
21. Gibbons, S.; Overman, H.G. Mostly pointless spatial econometrics? *Journal of Regional Science* **2012**, *52*, 172–191.
22. Bavaud, F. Models for spatial weights: a systematic look. *Geographical analysis* **1998**, *30*, 153–171.
23. Anselin, L. Specification Tests on the Structure of Interaction in Spatial Econometric Models. *Papers of the Regional Science Association* **1984**, *54*, 165–182.

24. Leenders, R. Modeling social influence through network autocorrelation: constructing the weight matrix. *Social Networks* **2002**, *24*, 21–47.
25. Kelejian, H.H. A spatial J-test for model specification against a single or a set of non-nested alternatives. *Letters in Spatial and Resource Sciences* **2008**, *1*, 3–11.
26. Piras, G.; Lozano, N. Spatial J-test: some Monte Carlo evidence. *Statistics and Computing* **2012**, *22*, 169–183.
27. Burridge, P.; Fingleton, B. Bootstrap inference in spatial econometrics: the J-test. *Spatial Economic Analysis* **2010**, *5*, 93–119.
28. Burridge, P. Improving the J test in the SARAR model by likelihood-based estimation. *Spatial Economic Analysis* **2012**, *7*, 75–107.
29. Kelejian, H.H.; Piras, G. An extension of Kelejian's J-test for non-nested spatial models. *Regional Science and Urban Economics* **2011**, *41*, 281–292.
30. Kelejian, H.H.; Piras, G. An Extension of the J-Test to a Spatial Panel Data Framework. *Journal of Applied Econometrics* **2016**, *31*, 387–402.
31. Hagemann, A. A simple test for regression specification with non-nested alternatives. *Journal of Econometrics* **2012**, *166*, 247–254.
32. Hansen, B.E. Challenges for econometric model selection. *Econometric Theory* **2005**, *21*, 60–68.
33. Akaike, H., 2nd International Symposium on Information Theory; Akademiai Kiodo, 1973; chapter Information Theory and an Extension of the Maximum Likelihood Principle, pp. 267–281.
34. Schwarz, G. Estimating the dimension of a model. *The annals of statistics* **1978**, *6*, 461–464.
35. Potscher, B. Effects of model selection on inference. *Econometric Theory* **1991**, *7*, 163–185.
36. Hepple, L.W. Bayesian techniques in spatial and network econometrics: 1. Model comparison and posterior odds. *Environment and Planning A* **1995**, *27*, 447–469.
37. Hepple, L.W. Bayesian techniques in spatial and network econometrics: 2. Computational methods and algorithms. *Environment and Planning A* **1995**, *27*, 615–644.
38. Hepple, L.W. Bayesian model choice in spatial econometrics. In *Advances in Econometrics: Spatial and Spatiotemporal Econometrics*; LeSage, J.; Pace, K., Eds.; Emerald Group Publishing Limited, 2004; Vol. 18, pp. 101–126.
39. Besag, J.; Higdon, D. Bayesian analysis of agricultural field experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **1999**, *61*, 691–746.
40. LeSage, J.; Pace, R. *Introduction to spatial econometrics*; Statistics: A Series of Textbooks and Monographs, Chapman and Hall, CRC press, 2009.
41. LeSage, J. Spatial econometrics toolbox. *available at: www. spatial-econometrics. com* **2007**.
42. Shannon, C.; Weaver, C. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, 1949.
43. Maddala, G. *Qualitative and limited dependent variable models in econometrics*; Cambridge: Cambridge University Press, 1983.
44. Florax, R.J.; De Graaff, T. The performance of diagnostic tests for spatial dependence in linear regression models: a meta-analysis of simulation studies. In *Advances in spatial econometrics*; Anselin, L.; Florax, R.; Rey, S., Eds.; Springer, 2004; pp. 29–65.
45. Ertur, C.; Koch, W. Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of applied econometrics* **2007**, *22*, 1033–1062.
46. Munnell, A.H.; Cook, L.M.; others. How does public infrastructure affect regional economic performance? *New England economic review* **1990**, pp. 11–33.
47. Millo, G.; Piras, G.; others. splm: Spatial panel data models in R. *Journal of Statistical Software* **2012**, *47*, 1–38.
48. Alvarez, I.C.; Barbero, J.; Zofio, J.L. A Panel Data Toolbox for MATLAB. *Journal of Statistical Software* **2017**, *76*, 1–27.

8. Appendix

Table 10. Percentage of correct selections. DGP: SDM; Estimated equation SDM. T=1

				CASE n=25				CASE n=49				CASE n=100			
T	ρ	β_1	θ	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
1	-0.8	1	1	53.3	54.0	35.7	**57.7**	67.0	68.3	28.3	**70.6**	80.7	79.8	34.2	**81.9**
1	-0.5	1	1	45.8	42.7	32.1	**48.5**	52.3	51.1	29.8	**54.9**	59.1	57.4	32.3	**63.2**
1	-0.2	1	1	33.7	33.2	34.9	**38.2**	**36.0**	30.9	32.9	34.7	30.8	25.0	30.2	**33.0**
1	0.2	1	1	24.3	23.1	**35.7**	25.5	29.5	18.3	37.1	28.1	**37.9**	18.6	**37.6**	34.9
1	0.5	1	1	**27.8**	23.3	41.6	27.6	45.2	33.5	**46.1**	40.6	**63.4**	57.1	42.1	60.7
1	0.8	1	1	36.8	31.1	**40.7**	33.4	**62.8**	61.7	53.4	55.1	80.5	**82.3**	55.0	67.3
1	-0.8	1	5	**65.9**	57.7	43.9	61.9	**81.4**	75.3	53.1	77.7	90.5	91.0	69.4	**92.9**
1	-0.5	1	5	55.0	53.5	45.5	**57.2**	67.4	71.5	55.3	**72.6**	81.4	83.9	71.1	**84.3**
1	-0.2	1	5	**51.2**	48.2	47.1	49.7	**57.8**	56.5	56.6	54.4	**72.7**	70.9	72.1	71.0
1	0.2	1	5	**45.9**	45.1	53.2	43.3	57.6	51.6	**63.6**	53.8	**73.4**	65.7	74.2	69.8
1	0.5	1	5	52.7	44.8	**57.4**	46.4	66.2	61.5	**68.6**	67.6	**83.3**	79.0	74.7	79.8
1	0.8	1	5	**62.1**	51.5	60.1	56.6	81.1	**87.6**	82.6	78.8	**94.9**	93.8	77.9	84.4
1	-0.8	5	1	59.8	57.7	42.6	**62.2**	75.8	74.8	46.4	**76.6**	84.0	85.7	55.3	**88.2**
1	-0.5	5	1	46.4	46.1	34.2	**50.4**	52.7	55.6	33.7	**58.7**	65.6	63.5	38.5	**68.1**
1	-0.2	5	1	34.7	31.0	31.9	**36.2**	28.5	28.6	29.1	**33.0**	28.6	22.3	26.9	**32.5**
1	0.2	5	1	29.0	27.4	**41.3**	30.1	31.9	23.4	**44.8**	30.9	**46.7**	26.9	45.8	39.4
1	0.5	5	1	39.6	35.7	**49.3**	37.5	**55.7**	48.7	58.1	55.1	**74.6**	70.3	59.2	69.8
1	0.8	5	1	**58.0**	49.1	55.2	50.6	76.7	**85.7**	78.2	76.0	**93.3**	92.0	73.9	81.3
1	-0.8	5	5	**56.8**	51.1	36.1	55.1	67.7	69.8	30.1	**73.6**	**82.6**	80.2	33.7	82.0
1	-0.5	5	5	48.1	47.1	34.9	**51.1**	56.6	56.4	39.6	**59.0**	70.6	70.2	49.5	**73.6**
1	-0.2	5	5	47.9	46.8	46.4	**48.6**	53.7	53.0	**54.0**	53.7	**66.0**	60.8	64.0	63.6
1	0.2	5	5	49.6	47.1	**56.5**	48.4	65.2	55.5	**67.5**	58.9	**76.7**	74.5	79.5	74.6
1	0.5	5	5	58.1	48.5	**62.5**	52.1	77.0	67.6	77.4	**78.2**	**92.5**	88.4	87.9	88.7
1	0.8	5	5	**72.4**	67.1	70.0	69.9	90.7	**94.6**	91.8	90.4	**98.9**	87.6	92.0	91.7

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018
doi:10.20944/preprints201811.0188.v1

Table 11. Percentage of correct selections. DGP: SDM; Estimated equation SDM. T=5

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	r	b1	q1	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC
5	-0.8	1	1	75.4	77.0	30.7	**78.3**	92.2	92.6	35.2	**93.9**	98.2	98.2	33	**99**
5	-0.5	1	1	**51.0**	46.5	33.6	46.7	65.7	69.7	35.7	**72.4**	88.5	88.0	34.6	**89.7**
5	-0.2	1	1	**34.4**	22.6	34.1	26.9	38.3	30.4	**40.1**	37.7	**51.3**	39.7	45.8	48.5
5	0.2	1	1	**41.5**	33.8	39.6	41.1	53.9	49.5	51.7	**55.8**	**71.3**	67.1	63.0	48.5
5	0.5	1	1	**68.5**	66.9	48.2	64.6	80.7	**82.0**	64.1	74.9	92.6	**93.7**	76.7	88.6
5	0.8	1	1	**88.3**	89.1	52.4	72.9	**97.1**	96.5	72.2	79.6	**99.7**	99.5	81.8	91.1
5	-0.8	1	5	90.3	**92.6**	79.8	91.8	97.9	98.9	93.9	99.3	99.8	**99.9**	98.3	**99.9**
5	-0.5	1	5	84.6	**85.8**	82.2	79.5	96.6	96.7	94.7	**97.0**	99.7	99.6	99.3	**99.8**
5	-0.2	1	5	82.3	80.8	**85.0**	71.8	92.3	**95.1**	**95.1**	93.3	**99.6**	**99.6**	99.4	99.4
5	0.2	1	5	88.5	85.9	**89.7**	85.7	96.6	97.3	**97.5**	95.4	99.8	**100**	**100**	99.5
5	0.5	1	5	**94.6**	92.7	90.5	90.4	98.5	**99.6**	99.3	96.4	**100**	**100**	**100**	99.6
5	0.8	1	5	**98.3**	97.4	84.1	83.0	**99.9**	**99.9**	99.4	93.5	**100**	**100**	**100**	98.5
5	-0.8	5	1	85.3	89.1	70.3	**89.5**	96.0	96.4	79.6	**98.0**	99.4	99.4	91.3	**99.8**
5	-0.5	5	1	54.3	**55.1**	44.4	51.7	73.2	76.3	52.0	**77.0**	92.2	93.8	68.6	**94.2**
5	-0.2	5	1	**27.9**	17.6	27.0	24.5	30.0	18.3	28.9	**31.8**	39.1	25.7	29.3	**39.3**
5	0.2	5	1	**56.4**	45.6	53.1	51.3	66.1	66.6	68.8	**69.1**	83.7	**85.5**	84.1	82.6
5	0.5	5	1	**83.5**	83.3	79.2	**83.5**	93.0	**95.8**	92.9	89.0	98.3	**98.8**	97.8	95.8
5	0.8	5	1	97.5	**97.8**	78.5	81.7	**99.7**	**99.7**	98.5	89.5	**100**	**100**	**100**	97.9
5	-0.8	5	5	78.0	79.7	38.7	**80.4**	93.0	93.8	44.0	**94.1**	97.7	98.4	52.6	**98.5**
5	-0.5	5	5	**67.2**	66.6	61.8	60.8	83.3	87.2	72.4	**87.6**	96.9	96.7	85.3	**97.1**
5	-0.2	5	5	**78.6**	70.3	76.4	65.6	88.0	89.0	**90.6**	85.3	97.2	**98**	97.3	97.4
5	0.2	5	5	**93.5**	88.8	90.9	86.9	98.2	**99.3**	99.1	97.6	**100**	**100**	**100**	**100**
5	0.5	5	5	**98.5**	96.8	96.3	96.1	99.6	**100**	99.9	98.5	**100**	**100**	**100**	99.7
5	0.8	5	5	**99.8**	**99.8**	91.3	92.9	**100**	99.9	**100**	98.6	**100**	**100**	**100**	**100**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018
doi:10.20944/preprints201811.0188.v1

Table 12. Percentage of correct selections. DGP: SDM; Estimated equation SDM. T=10

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	**r**	**b1**	**q1**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
10	-0.8	1	1	89.7	89.9	32.8	**92.2**	98.2	97.2	34.0	**98.6**	99.7	99.6	35.3	**100**
10	-0.5	1	1	**62.6**	61.8	32.9	61.3	**84.5**	81.5	37.7	82.9	96.1	95.6	41.2	**96.2**
10	-0.2	1	1	**40.0**	29.1	38.5	36.7	**53.4**	41.0	48.5	47.4	**68.0**	59.2	60.3	61.2
10	0.2	1	1	49.7	52.9	53.2	**54.6**	**69.6**	69.1	64.5	68.9	81.8	**84.7**	77.1	84.3
10	0.5	1	1	80.7	**84.1**	61.6	80.0	**93.3**	92.7	75.1	88.4	98.6	**99.0**	89.5	94.4
10	0.8	1	1	**98.3**	98.0	67.0	84.1	99.3	**99.7**	81.3	91.3	**100**	**100**	92.9	97.2
10	-0.8	1	5	98.3	**98.8**	93.5	98.4	**100**	99.9	98.1	99.9	**100**	**100**	99.9	**100**
10	-0.5	1	5	93.1	**96.2**	94.5	93.7	99.6	99.8	99.2	**99.9**	**100**	99.9	99.9	**100**
10	-0.2	1	5	91.5	94.2	**95**	90.8	99.3	**99.4**	99.0	99.1	**100**	**100**	**100**	**100**
10	0.2	1	5	95.6	98.1	**98.3**	97.7	**99.9**	99.3	99.4	98.8	**100**	**100**	**100**	**100**
10	0.5	1	5	99	**99.7**	99.4	97.8	**100**	**100**	**100**	**100**	**100**	**100**	**100**	99.9
10	0.8	1	5	**100**	**100**	98.5	92.8	**100**	**100**	99.9	98.6	**100**	**100**	**100**	99.9
10	-0.8	5	1	95.5	**97.1**	82.9	**97.1**	99.2	99.5	92.6	**99.8**	**100**	**100**	98.3	**100**
10	-0.5	5	1	67.5	**68.9**	53.5	66.5	**89.9**	87.5	66.0	86.4	97.9	97.8	81.7	**98**
10	-0.2	5	1	**33.4**	21.0	28.0	29.7	**40.8**	25.4	31.0	38.0	**57.9**	41.4	28.3	53.7
10	0.2	5	1	66.5	70.7	69.8	**71.3**	83.5	**85.2**	83.4	83.3	**94.4**	93.9	92.0	91.0
10	0.5	5	1	93.8	**95.7**	91.9	90.8	**99.6**	98.8	97.8	95.8	**100**	**100**	99.7	99.0
10	0.8	5	1	99.7	**100**	96.7	90.6	**100**	**100**	99.8	98.0	**100**	**100**	**100**	99.5
10	-0.8	5	5	89.0	88.9	42.3	**90.4**	97.8	98.3	55.5	**98.5**	99.7	99.7	65.6	**99.9**
10	-0.5	5	5	79.3	**80.0**	74.4	74.2	95.3	**96.1**	87.2	95.4	**99.8**	**99.8**	97.0	**99.8**
10	-0.2	5	5	86.1	89.5	**92.3**	84	97.6	**98.1**	97.7	97.1	99.9	99.9	99.8	**100**
10	0.2	5	5	97.9	**99.2**	99.1	98.5	**100**	99.8	99.8	99.8	**100**	**100**	**100**	**100**
10	0.5	5	5	99.0	**100**	**100**	98.9	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
10	0.8	5	5	**100**	**100**	99.5	97.6	**100**	**100**	**100**	99.9	**100**	**100**	**100**	**100**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018 doi:10.20944/preprints201811.0188.v1

Table 13. Percentage of correct selections. DGP: SDEM; Estimated equation SDEM. T=1

				CASE n=25				CASE n=49				CASE n=100			
T	ρ	β_1	θ	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
1	-0.8	1	1	56.6	52.3	23.1	**58.1**	64.6	67.0	24.3	**70.5**	81.2	80.5	28.2	**82.8**
1	-0.5	1	1	45.8	42.3	22.4	**48.9**	46.0	51.1	24.0	**56.7**	60.7	60.0	31.2	**65.3**
1	-0.2	1	1	36.9	31.8	23.8	**38.4**	30.4	31.9	26.9	**37.0**	33.6	26.6	31.0	**34.9**
1	0.2	1	1	**28.3**	20.7	24.4	26.5	28.3	17.0	**33.1**	26.4	**36.0**	17.9	35.9	34.1
1	0.5	1	1	29.1	21.0	**30.2**	26.1	**43.9**	30.6	42.2	37.8	**61.5**	55.8	37.1	58.8
1	0.8	1	1	31.9	26.3	**35.9**	28.0	**61.5**	56.2	46.9	48.9	77.1	**79.4**	44.0	67.7
1	-0.8	1	5	63.6	54.7	40.0	**64.7**	75.6	77.5	63.4	**81.8**	92.2	91.8	75.1	**92.7**
1	-0.5	1	5	**58.7**	48.8	43.0	58.6	66.5	73.5	64.7	**77.6**	84.2	85.2	77.6	**86.2**
1	-0.2	1	5	51.5	42.7	42.2	**51.8**	49.8	59.9	**62.3**	61.8	**76.3**	74.5	76.1	76.2
1	0.2	1	5	44.3	39.1	**46.1**	45.2	49.4	55.4	**64.3**	58.5	72.8	67.2	74.4	**72.9**
1	0.5	1	5	44.6	38.9	**44.7**	40.3	58.0	59.7	**61.2**	**61.2**	**80.8**	78.6	71.6	79.5
1	0.8	1	5	**44.6**	37.6	43.8	36.8	**72.3**	71.4	63.8	68.4	89.8	**91.6**	66.1	83.7
1	-0.8	5	1	**57.9**	40.7	22.6	56.2	63.7	62.8	23.6	**70.2**	80.9	91.8	75.1	**92.7**
1	-0.5	5	1	47.6	31.1	23.4	**49.5**	46.6	41.8	26.5	**55.8**	63.5	85.2	77.6	**86.2**
1	-0.2	5	1	**39.4**	19.4	21.5	38.1	29.4	22.9	26.7	**35.5**	32.4	74.5	76.1	**76.2**
1	0.2	5	1	**30.0**	13.5	27.7	27.5	28.7	15.3	**32.3**	27.3	36.2	67.2	**74.4**	72.9
1	0.5	5	1	**30.2**	14.9	28.5	24.2	**46.8**	30.9	41.8	36.3	60.3	78.6	71.6	**79.5**
1	0.8	5	1	30.6	25	**36.3**	26.6	**67.2**	59.5	51.9	48.8	76.6	**91.6**	66.1	83.7
1	-0.8	5	5	64.2	42.0	46.0	**66.2**	72.8	75.0	62.1	**83.9**	**91.0**	88.9	72.4	90.6
1	-0.5	5	5	57.8	36.0	44.9	**60.0**	61.2	64.9	62.7	**73.3**	84.1	83.2	74.7	**86.9**
1	-0.2	5	5	50.0	34.5	46.7	**53.1**	53.4	56.5	64.3	**66.1**	**77.1**	72.8	74.6	75.4
1	0.2	5	5	46.0	34.8	**46.3**	44.0	48.1	52.8	**62.9**	58.7	71.9	67.2	**73.4**	71.1
1	0.5	5	5	41.9	35.5	**42.8**	39.3	57.3	57.8	**59.3**	59.1	80.9	78.5	70.2	**81.1**
1	0.8	5	5	41.7	44.2	**47.5**	40.3	70.2	**70.8**	64.2	64.7	**90.3**	87.6	67.7	81.1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 14. Percentage of correct selections. DGP: SDEM; Estimated equation SDEM. T=5

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	**r**	**b1**	**q1**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
5	-0.8	1	1	78.3	78.5	36.6	**81.0**	92.9	93.3	44.3	**94.9**	98.5	98.5	53.9	**99.2**
5	-0.5	1	1	**52.2**	50.2	39.7	50.6	72.4	73.5	47.5	**74.5**	90.5	90.9	56.6	**92.8**
5	-0.2	1	1	34.3	26.7	**35.8**	30.6	44.9	38.8	**46.0**	43.9	**58.1**	52.2	56.8	57.8
5	0.2	1	1	34.6	31.7	37.1	**37.9**	49.5	42.3	45.3	**52.6**	**67.9**	58.8	52.4	64.0
5	0.5	1	1	55.0	60.2	37.3	**62.7**	77.1	**77.3**	41.5	72.2	88.7	**92.7**	50.9	86.5
5	0.8	1	1	84.0	**86.5**	35.8	71.6	93.2	**94.3**	45.6	78.8	**99.1**	98.9	48.5	90.6
5	-0.8	1	5	96.3	**98**	94.4	97.9	99.2	**99.8**	98.8	99.7	99.9	**100**	99.8	**100**
5	-0.5	1	5	90.6	**94.6**	93.9	93.0	98.6	**99.0**	98.6	98.9	99.8	99.9	**100**	99.9
5	-0.2	1	5	85.2	89.9	**92.9**	88.7	96.7	**98.5**	98.4	98.2	99.8	**100**	**100**	**100**
5	0.2	1	5	78.9	87.1	**90**	89.0	95.9	**96.8**	**96.8**	96.4	99.4	**99.7**	**99.7**	99.5
5	0.5	1	5	81.8	**87.8**	84.2	87.5	**97.6**	97.1	93.2	96.9	**99.9**	**99.9**	99.2	99.7
5	0.8	1	5	90.5	**93.3**	75.9	86.5	99.1	**99.4**	88.0	95.0	**100**	**100**	96.1	98.6
5	-0.8	5	1	77.9	81.0	32.1	**81.9**	93.0	91.5	45.7	**93.8**	97.7	98.1	54.7	**98.6**
5	-0.5	5	1	**52.2**	51.9	36.5	49.6	71.6	70.9	47.2	**73.6**	89.8	92.9	57.7	**93.5**
5	-0.2	5	1	35.5	26.9	**36**	32.2	44.0	35.0	**45.3**	42.0	**59.3**	51.8	59.4	57.7
5	0.2	5	1	35.2	30.2	37.0	**37.7**	**50.7**	41.3	44.6	49.2	65.2	62.1	56.1	**68.1**
5	0.5	5	1	55.7	**57.2**	37.2	56.6	76.8	**77.8**	45.4	73.4	88.1	**91.1**	49.3	85.9
5	0.8	5	1	**86.5**	86.3	37.6	71.1	94.2	**95.4**	43.3	77.8	**98.8**	98.6	46.8	90.7
5	-0.8	5	5	95.2	98.0	94.3	98.1	**99.3**	99.5	98.6	**99.6**	**100**	**100**	**100**	**100**
5	-0.5	5	5	92.1	**94.8**	94.0	93.0	**98.9**	99.3	98.3	98.9	**100**	99.9	99.9	**100**
5	-0.2	5	5	87.8	87.7	**90.7**	86.5	97.4	**98.4**	98.1	98.2	99.7	**99.8**	99.5	**99.8**
5	0.2	5	5	80.3	85.5	**88.7**	88.0	**97.3**	96.3	96.4	95.5	99.5	**99.9**	99.7	99.8
5	0.5	5	5	80.8	**87.9**	84	87.2	95.7	**97.2**	93	96.4	**99.8**	99.7	98.9	99.4
5	0.8	5	5	91.2	**93.1**	75.5	87.0	**99.7**	99.4	86.5	95.4	**100**	**100**	95.1	99

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 15. Percentage of correct selections. DGP: SDEM; Estimated equation SDEM. T=10

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	**r**	**b1**	**q1**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
10	-0.8	1	1	89.6	92.7	47.4	**94.2**	98.4	97.8	56.7	**98.9**	**100**	99.9	70.8	**100**
10	-0.5	1	1	68.0	**68.8**	48.2	68.6	**88.2**	85.7	58.3	86.7	97.0	**97.7**	71.1	97.6
10	-0.2	1	1	**49.4**	36.9	43.8	41.5	**65.3**	54.2	56.6	56.3	**75.9**	74.1	71.7	73.3
10	0.2	1	1	46.9	45.3	45.6	**49.8**	**67.1**	60.8	54.2	63.7	78.0	**80.9**	69.3	80.8
10	0.5	1	1	**76.8**	76.3	45.6	73.3	**91.4**	89.8	49.6	85.3	**97.3**	96.8	62.1	93.0
10	0.8	1	1	95.5	**97.2**	43.1	83.0	99.4	**99.7**	43.6	90.5	99.9	**100**	51.9	97.1
10	-0.8	1	5	99.5	99.7	99.0	**99.9**	**100**	**100**	99.8	**100**	**100**	**100**	**100**	**100**
10	-0.5	1	5	99.4	**99.6**	99.1	99.3	**100**	99.9	99.9	99.9	**100**	**100**	**100**	**100**
10	-0.2	1	5	98.1	**98.3**	98.2	98.0	**99.8**	99.7	99.6	99.7	**100**	**100**	**100**	**100**
10	0.2	1	5	95.9	98.0	97.9	**98.1**	**99.8**	99.1	99.0	98.8	**100**	**100**	**100**	**100**
10	0.5	1	5	95.5	**98.1**	96.1	97.2	99.5	**99.7**	98.8	99.6	**100**	**100**	**100**	**100**
10	0.8	1	5	**99.4**	**99.4**	88.9	97.0	**100**	**100**	95.1	99.2	**100**	100	99.8	**100**
10	-0.8	5	1	87.4	90.5	46.0	**92.4**	97.8	97.9	57	**98.3**	99.7	**99.8**	70.6	**99.8**
10	-0.5	5	1	**71.9**	63.9	45.4	64.8	**88.5**	86.5	59.3	87.2	**96.8**	**96.8**	70.7	**96.8**
10	-0.2	5	1	**49.3**	36.9	47.9	40.8	**64.6**	55.6	57.7	56.9	**78**	74.6	71.3	74.7
10	0.2	5	1	50.8	46.1	46.0	**52**	62.7	59.4	54.0	**64.6**	**79.8**	78.4	66.1	76.8
10	0.5	5	1	74.6	**76.4**	43.3	73.2	**91.7**	90.7	49.3	83.9	97.1	**97.2**	64.2	93.7
10	0.8	5	1	96.8	**97.7**	40.5	80.8	98.6	**99.5**	43.5	89.0	**100**	99.8	51.5	95.9
10	-0.8	5	5	**99.9**	99.8	99.0	99.8	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
10	-0.5	5	5	99.3	**100**	99.7	99.5	**100**	**100**	99.9	**100**	**100**	**100**	**100**	**100**
10	-0.2	5	5	98.1	98.5	**98.6**	98.5	**99.9**	99.8	**99.9**	99.7	**100**	**100**	**100**	**100**
10	0.2	5	5	95.7	**96.2**	**96.2**	**96.2**	99.8	**99.9**	99.7	99.5	99.9	**100**	**100**	**100**
10	0.5	5	5	95.9	**97.4**	94.1	96.9	**99.8**	99.5	99.1	99.2	**100**	**100**	99.9	**100**
10	0.8	5	5	**99.7**	99.0	87.1	95.2	**100**	**100**	95.7	98.9	**100**	**100**	99.6	**100**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018
doi:10.20944/preprints201811.0188.v1

Table 16. Percentage of correct selections. DGP: SLM; Estimated equation SLM. T=1

				CASE n=25				CASE n=49				CASE n=100			
T	ρ	β_1	θ	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
1	-0.8	1	1	49.6	47.5	29.6	**50.2**	64.5	62.4	30.1	**66.4**	78.8	79.5	26.5	**80.3**
1	-0.5	1	1	33.3	32.8	25.9	**36.3**	42.3	40.6	25.2	**45.2**	54.4	54.2	25.4	**59.6**
1	-0.2	1	1	24.5	20.6	**28.3**	26.7	24.2	19.8	**29.4**	25.6	26.8	17.5	**29.0**	28.0
1	0.2	1	1	23.7	16.2	**32.5**	25.2	29.1	11.8	**34.9**	28.1	**36.6**	15.5	31.4	35.4
1	0.5	1	1	26.1	19.8	**35.9**	26.9	**47.0**	37.2	41.5	45.2	**62.6**	60.3	39.5	59.8
1	0.8	1	1	37.6	35.1	**47.0**	36.6	62.2	**63.2**	45.6	57.1	76.7	**78.5**	41.0	64.0
1	-0.8	1	5	33.3	**34.7**	22.3	31.0	42.7	40.4	24.0	**42.9**	55.0	54.0	9.6	**55.9**
1	-0.5	1	5	22.7	18.4	**23.3**	19.9	31.0	26.3	25.7	**31.1**	**34.3**	27.7	20.1	33.9
1	-0.2	1	5	20.2	10.1	**27.2**	17.0	21.3	9.4	**28.5**	19.7	18.7	7.5	**27.4**	17.7
1	0.2	1	5	27.6	9.5	**31.3**	27.0	32.7	10.9	**34.5**	31.2	**35.2**	11.4	30.1	32.8
1	0.5	1	5	34.9	22.8	**37.7**	34.9	**50.0**	36.0	31.6	49.9	**57.2**	48.4	21.4	55.6
1	0.8	1	5	**42.3**	39.3	38.3	41.6	69.3	**69.8**	30.3	69.5	77.9	**80.0**	14.8	70.7
1	-0.8	5	1	49.6	47.5	29.6	**50.2**	64.5	62.4	30.1	**66.4**	78.8	79.5	26.5	**80.3**
1	-0.5	5	1	33.3	32.8	25.9	**36.3**	42.3	40.6	25.2	**45.2**	54.4	54.2	25.4	**59.6**
1	-0.2	5	1	24.5	20.6	**28.3**	26.7	24.2	19.8	29.4	**25.6**	26.8	17.5	29.0	**28.0**
1	0.2	5	1	23.7	16.2	**32.5**	25.2	29.1	11.8	**34.9**	28.1	**36.6**	15.5	31.4	35.4
1	0.5	5	1	26.1	19.8	**35.9**	26.9	**47.0**	37.2	41.5	45.2	**62.6**	60.3	39.5	59.8
1	0.8	5	1	37.6	35.1	**47.0**	36.6	62.2	**63.2**	45.6	57.1	76.7	**78.5**	41.0	64
1	-0.8	5	5	33.3	**34.7**	22.3	31.0	42.7	40.4	24.0	**42.9**	55.0	54.0	9.6	**55.9**
1	-0.5	5	5	22.7	18.4	**23.3**	19.9	31.0	26.3	25.7	**31.1**	**34.3**	27.7	20.1	33.9
1	-0.2	5	5	20.2	10.1	**27.2**	17.0	21.3	9.4	**28.5**	19.7	18.7	7.5	**27.4**	17.7
1	0.2	5	5	27.6	9.5	**31.3**	27.0	32.7	10.9	**34.5**	31.2	**35.2**	11.4	30.1	32.8
1	0.5	5	5	34.9	22.8	**37.7**	34.9	**50.0**	36.0	31.6	49.9	**57.2**	48.4	21.4	55.6
1	0.8	5	5	**42.3**	39.3	38.3	41.6	69.3	69.8	30.3	**69.5**	77.9	**80.0**	14.8	70.7

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 17. Percentage of correct selections. DGP: SLM; Estimated equation SLM. T=5

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	r	b1	q1	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC
5	-0.8	1	1	78.5	**82.3**	29.8	81.1	90.8	91.9	25.7	**92.7**	98.2	**98.3**	17.7	98.2
5	-0.5	1	1	**52.3**	48.4	28.6	49.8	64.9	64.3	28.0	**65.9**	86.0	86.6	24.1	**87.9**
5	-0.2	1	1	**28.3**	13.8	27.3	25.5	**34.6**	17.5	28.5	31.3	**44.2**	24.8	27.5	42.1
5	0.2	1	1	**36.0**	22.0	28.4	35.4	**49.1**	26.6	30.2	42.1	56.8	44.3	29.4	**57.5**
5	0.5	1	1	60.0	**60.7**	32.2	58.8	76.1	**77.9**	34.4	73.1	87.3	**89.0**	29.8	82.7
5	0.8	1	1	86.6	**87.8**	38.9	72.7	95.7	**96.5**	42.2	81.8	**99.3**	99.0	39.8	90.5
5	-0.8	1	5	52.7	**55.2**	24.7	51.8	66.9	**68.2**	18.8	67.1	91.2	92.7	2.5	**93.2**
5	-0.5	1	5	**37.1**	29.4	27.2	32.4	**44.6**	33.6	25.8	39.3	**62.6**	55.4	10.9	59.8
5	-0.2	1	5	**27.9**	11.6	26.5	24.6	31.2	12.4	**31.7**	29.4	**30.3**	10.6	27.1	28.2
5	0.2	1	5	**33.9**	18.9	29.9	**33.9**	**40.0**	22.3	31.1	39.2	**50.5**	30.0	26.4	49.9
5	0.5	1	5	60.1	57.8	27.7	**61.5**	**66.9**	64.2	17.4	66.7	**87.5**	87.1	5.1	86.2
5	0.8	1	5	84.8	82.6	21.1	**89.8**	94.6	93.9	6.3	**95.5**	**100**	**100**	0.2	99.7
5	-0.8	5	1	78.5	**82.3**	29.8	81.1	90.8	91.9	25.7	**92.7**	98.2	**98.3**	17.7	98.2
5	-0.5	5	1	**52.3**	48.4	28.6	49.8	64.9	64.3	28.0	**65.9**	86.0	86.6	24.1	**87.9**
5	-0.2	5	1	**28.3**	13.8	27.3	25.5	**34.6**	17.5	28.5	31.3	**44.2**	24.8	27.5	42.1
5	0.2	5	1	**36.0**	22.0	28.4	35.4	**49.1**	26.6	30.2	42.1	56.8	44.3	29.4	**57.5**
5	0.5	5	1	60.0	**60.7**	32.2	58.8	76.1	**77.9**	34.4	73.1	87.3	**89.0**	29.8	82.7
5	0.8	5	1	86.6	**87.8**	38.9	72.7	95.7	**96.5**	42.2	81.8	**99.3**	99.0	39.8	90.5
5	-0.8	5	5	52.7	**55.2**	24.7	51.8	66.9	**68.2**	18.8	67.1	91.2	92.7	2.5	**93.2**
5	-0.5	5	5	**37.1**	29.4	27.2	32.4	**44.6**	33.6	25.8	39.3	**62.6**	55.4	10.9	59.8
5	-0.2	5	5	**27.9**	11.6	26.5	24.6	31.2	12.4	**31.7**	29.4	**30.3**	10.6	27.1	28.2
5	0.2	5	5	**33.9**	18.9	29.9	**33.9**	**40.0**	22.3	31.1	39.2	**50.5**	30.0	26.4	49.9
5	0.5	5	5	60.1	57.8	27.7	**61.5**	**66.9**	64.2	17.4	66.7	**87.5**	87.1	5.1	86.2
5	0.8	5	5	84.8	82.6	21.1	**89.8**	94.6	93.9	6.3	**95.5**	**100**	**100**	0.2	99.7

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 18. Percentage of correct selections. DGP: SLM; Estimated equation SLM. T=10

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	r	b1	q1	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC
10	-0.8	1	1	89.3	89.1	29.7	**90.1**	97.9	97.7	24.0	**98.3**	**99.7**	**99.7**	15.0	**99.7**
10	-0.5	1	1	**64.7**	58.3	27.0	58.2	**80.9**	78.1	23.6	79.8	95.5	95.6	20.3	**95.8**
10	-0.2	1	1	**37.1**	21.8	30.0	34.7	**45.9**	27.3	25.2	42.1	**60.2**	45.7	26.8	57.7
10	0.2	1	1	40.3	30.8	28.7	**40.7**	**53.6**	43.1	30.3	52.8	**69.7**	64.5	29.9	68.6
10	0.5	1	1	**76.7**	75.9	34.8	73.5	90.0	**90.4**	29.3	84.8	97.0	**97.2**	26.1	92.2
10	0.8	1	1	**96.1**	96.0	36.5	84.6	99.7	**99.8**	31.5	88.6	**100**	**100**	25.1	96.2
10	-0.8	1	5	49.0	**49.2**	27.6	45.5	**79.0**	78.5	9.7	78.8	97.4	97.8	0.4	**98.2**
10	-0.5	1	5	**41.5**	29.9	30.1	33.8	**55.2**	43.4	24.1	47.9	**75.0**	71.5	7.0	72.7
10	-0.2	1	5	**31.8**	16.9	30.7	30.2	**36.7**	20.7	30.3	34.8	**46.5**	27.3	24.7	44.1
10	0.2	1	5	38.6	28.3	30.4	**38.8**	45.9	34.7	27.7	**46.2**	58.7	46.6	18.4	**59.6**
10	0.5	1	5	66.1	60.5	22.2	**69.6**	82.7	79.9	7.2	**86.3**	96.8	96.7	1.4	**97.3**
10	0.8	1	5	86.7	83.9	9.6	**94.8**	99.2	99.1	1.2	**100**	**100**	**100**	0.0	**100**
10	-0.8	5	1	89.3	89.1	29.7	**90.1**	97.9	97.7	24.0	**98.3**	**99.7**	**99.7**	15.0	**99.7**
10	-0.5	5	1	**64.7**	58.3	27.0	58.2	**80.9**	78.1	23.6	79.8	95.5	95.6	20.3	**95.8**
10	-0.2	5	1	**37.1**	21.8	30.0	34.7	**45.9**	27.3	25.2	42.1	**60.2**	45.7	26.8	57.7
10	0.2	5	1	40.3	30.8	28.7	**40.7**	**53.6**	43.1	30.3	52.8	**69.7**	64.5	29.9	68.6
10	0.5	5	1	**76.7**	75.9	34.8	73.5	90.0	**90.4**	29.3	84.8	97.0	**97.2**	26.1	92.2
10	0.8	5	1	**96.1**	96.0	36.5	84.6	99.7	**99.8**	31.5	88.6	**100**	**100**	25.1	96.2
10	-0.8	5	5	49.0	**49.2**	27.6	45.5	**79.0**	78.5	9.7	78.8	97.4	97.8	0.4	**98.2**
10	-0.5	5	5	**41.5**	29.9	30.1	33.8	**55.2**	43.4	24.1	47.9	**75.0**	71.5	7.0	72.7
10	-0.2	5	5	**31.8**	16.9	30.7	30.2	**36.7**	20.7	30.3	34.8	**46.5**	27.3	24.7	44.1
10	0.2	5	5	38.6	28.3	30.4	**38.8**	45.9	34.7	27.7	**46.2**	58.7	46.6	18.4	**59.6**
10	0.5	5	5	66.1	60.5	22.2	**69.6**	82.7	79.9	7.2	**86.3**	96.8	96.7	1.4	**97.3**
10	0.8	5	5	86.7	83.9	9.6	**94.8**	99.2	99.1	1.2	**100**	**100**	**100**	0.0	**100**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 19. Percentage of correct selections. DGP: SDEM; Estimated equation SDM. T=1

				CASE n=25				CASE n=49				CASE n=100			
T	ρ	β_1	θ	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**	**Entropy**	**Bayes**	**MJ**	**AIC**
1	-0.8	1	1	53.5	53.5	36.1	**59.8**	65.3	67.0	28.2	**70.4**	81.7	80.6	34.2	**82.5**
1	-0.5	1	1	46.4	44.5	32.8	**49.5**	49.6	51.7	29.1	**56.3**	61.0	59.8	34.3	**63.6**
1	-0.2	1	1	34.9	33.7	35.1	**38.2**	30.8	32.4	34.0	**35.8**	33.1	26.1	32.7	**34.3**
1	0.2	1	1	23.9	21.9	**34.0**	25.0	25.1	16.7	**37.1**	26.3	**35.7**	16.7	34.8	33.3
1	0.5	1	1	25.7	22.3	**39.8**	25.3	**42.0**	29.1	41.1	36.2	**59.6**	55.3	42.0	56.9
1	0.8	1	1	29.3	26.8	**37.5**	28.3	**56.2**	55.3	48.1	48.7	77.7	**79.4**	52.3	65.7
1	-0.8	1	5	**66.0**	60.2	44.0	64.8	73.8	72.9	47.9	**74.1**	90.0	91.2	59.9	**93.1**
1	-0.5	1	5	53.8	53.8	49.2	**57.0**	64.0	**68.2**	54.0	68.1	78.9	80.7	70.3	**82.0**
1	-0.2	1	5	**51.2**	48.4	48.1	48.1	57.4	55.7	**58.1**	53.5	71.6	69.1	**72.9**	69.4
1	0.2	1	5	45.0	43.8	**50.2**	43.1	50.7	51.3	**62.6**	53.3	70.6	64.6	**72.9**	66.8
1	0.5	1	5	44.2	39.3	**51.0**	37.8	54.2	54.4	**57.8**	52.3	**74.4**	73.7	64.2	71.0
1	0.8	1	5	41.0	35.9	**44.5**	35.2	**63.6**	62.5	49.5	53.6	81.0	**83.5**	43.8	70.3
1	-0.8	5	1	54.0	52.2	35.8	**57.7**	66.4	67.7	31.4	**71.5**	81.6	83.4	33.1	**85.4**
1	-0.5	5	1	45.1	44.9	32.7	**50.5**	46.4	50.4	30.8	**55.6**	61.8	63.4	33.3	**66.4**
1	-0.2	5	1	35.6	33.3	35.3	**37.2**	31.8	30.1	31.0	**34.2**	31.6	25.8	31.6	**34.5**
1	0.2	5	1	23.7	22.4	**38.5**	26.0	25.8	16.6	**38.2**	26.3	**35.8**	15.1	35.7	30.9
1	0.5	5	1	24.7	21.1	**36.7**	22.7	43.9	29.1	**44.6**	35.5	59.3	56.4	41.5	**60.4**
1	0.8	5	1	29.9	26.6	**40.6**	27.4	**59.4**	58.1	48.9	49.2	77.4	**81.1**	49.8	69.4
1	-0.8	5	5	67.5	64.1	45.5	**67.8**	73.3	74.9	47.7	**76.7**	89.2	87.9	58.3	**90.4**
1	-0.5	5	5	53.7	55.9	46.8	**59.0**	64.6	**66.9**	54.0	64.5	79.4	81.5	69.0	**83.3**
1	-0.2	5	5	51.4	51.2	49.1	**51.6**	54.8	57.5	**58.5**	57.1	**73.2**	69.8	70.2	69.6
1	0.2	5	5	44.5	44.1	**52.2**	45.4	52.0	52.2	**61.8**	53.7	69.7	65.8	**72.3**	68.0
1	0.5	5	5	42.3	36.2	**50.4**	37.2	54.6	52.6	**60.2**	52.8	**73.9**	**73.9**	65.4	72.7
1	0.8	5	5	40.1	40.0	**48.7**	37.3	**64.4**	59.5	51.4	50.7	**80.7**	80.2	43.2	69.1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 20. Percentage of correct selections. DGP: SDEM; Estimated equation SDM. T=5

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	r	b1	q1	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC
5	-0.8	1	1	76.1	76.6	36.1	**77.8**	90.8	94.2	41.7	**94.9**	98.6	98.7	48.3	**99.3**
5	-0.5	1	1	**54.4**	49.0	38.9	48.1	71.4	71.9	42.9	**74.3**	90.4	90.3	50.0	**92.3**
5	-0.2	1	1	**38.1**	25.0	36.7	28.9	41.4	35.3	**46.0**	42.3	**55.5**	48.6	53.8	53.4
5	0.2	1	1	**36.9**	30.0	36.5	36.7	50.9	41.3	43.8	**51.2**	**67.7**	58.7	51.2	63.7
5	0.5	1	1	**61.9**	59.7	36.6	60.9	**79.0**	76.8	43.3	69.7	88.2	**91.8**	48.5	85.9
5	0.8	1	1	85.7	**86.7**	43.8	69.7	94.3	**94.4**	47.2	77.7	**98.9**	**98.9**	48.9	88.7
5	-0.8	1	5	**93.3**	84.4	66.0	77.1	99.0	**99.6**	93.6	99.5	99.8	**100**	99.3	**100**
5	-0.5	1	5	**87.8**	78.3	75.7	67.2	96.4	**98.1**	96.8	97.4	99.9	**100**	99.7	99.9
5	-0.2	1	5	**85.2**	78.4	83.2	70.2	96.6	95.8	**96.8**	93.7	**100**	99.8	99.8	99.5
5	0.2	1	5	82.6	86.3	**88.7**	83.0	94.8	95.0	**95.8**	92.6	98.8	99.2	**99.3**	98.3
5	0.5	1	5	80.6	**84.4**	78.1	77.4	88.7	**93.6**	88.6	84.4	97.5	**97.9**	95.2	93.4
5	0.8	1	5	83.6	**87.5**	43.4	66.4	91.6	**94.2**	49.6	75.0	99.2	**99.6**	32.2	88.5
5	-0.8	5	1	75.1	78.3	36.4	**79.0**	91.7	91.7	42.5	**94.2**	98.1	98.1	47.9	**98.9**
5	-0.5	5	1	**55.1**	49.7	37.4	48.9	68.4	71.8	43.1	**73.7**	89.5	92.6	53.2	**93.6**
5	-0.2	5	1	**38.1**	25.4	36.6	29.4	39.7	32.8	45.3	**39.8**	**55.1**	47.8	57.7	53.9
5	0.2	5	1	**38.6**	29.7	37.6	36.7	**49.5**	40.7	44.0	**49.5**	65.5	62.5	54.6	**66.3**
5	0.5	5	1	**58.8**	57.9	39.9	56.7	75.4	**77.7**	43.6	71.0	87.4	**90.2**	51.9	85.0
5	0.8	5	1	83.5	**87.1**	42.5	69.2	94.1	**95.0**	48.9	76.6	**98.8**	98.5	49.4	90.5
5	-0.8	5	5	**93.0**	86.5	65.6	80.4	99.0	99.5	95.0	**99.8**	**100**	**100**	99.3	**100**
5	-0.5	5	5	**88.6**	79.3	75.2	68.1	97.4	**98.0**	95.8	96.9	99.8	**100**	99.7	**100**
5	-0.2	5	5	**87.9**	76.5	81.5	69.0	95.6	95.6	**96.6**	93.8	99.6	99.8	99.7	**99.9**
5	0.2	5	5	82.9	84.9	**86.7**	81.6	**95.6**	94.6	94.7	92.4	98.8	**99.6**	99.4	98.8
5	0.5	5	5	80.9	**86.2**	79.6	77.0	89.4	**92.7**	88.5	83.3	**98.4**	98.3	94.9	93.7
5	0.8	5	5	84.9	**89.0**	43.7	67.9	92.5	**94.7**	47.4	75.2	**99.2**	98.7	35.1	88.9

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 November 2018

doi:10.20944/preprints201811.0188.v1

Table 21. (continues) Percentage of correct selections. DGP: SDEM; Estimated equation SDM. T=10

Other parameters				CASE n=25				CASE n=49				CASE n=100			
T	r	b1	q1	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC	Entropy	Bayes	MJ	AIC
10	-0.8	1	1	89.1	91.2	42.1	**92.6**	97.7	97.9	50.5	**99.1**	**100**	99.9	61.1	**100**
10	-0.5	1	1	**64.3**	64.0	44.7	63.7	**86.2**	85.1	52.8	85.9	96.9	97.9	67.8	**98.3**
10	-0.2	1	1	43.6	33.6	**45.5**	39.1	**62.2**	49.9	54.8	53.7	**73.3**	69.2	69.8	69.6
10	0.2	1	1	45.6	45.9	45.8	**48.2**	**68.6**	62.2	55.3	63.9	76.7	**79.6**	67.4	79.4
10	0.5	1	1	75.9	**76.8**	44.2	73.0	**90.4**	89.6	48.4	84.1	**96.7**	96.5	60.8	91.6
10	0.8	1	1	**97.7**	97.3	47.0	81.9	99.1	**99.6**	48.3	89.8	99.9	**100**	47.6	96.1
10	-0.8	1	5	**96.5**	95.4	82.8	92.7	**100**	**100**	98.6	**100**	**100**	**100**	**100**	**100**
10	-0.5	1	5	91.5	**94.4**	93.0	88.4	**100**	**100**	99.9	99.9	**100**	**100**	**100**	**100**
10	-0.2	1	5	93.4	94.2	**95.1**	90.1	**99.9**	99.7	99.7	99.6	**100**	**100**	**100**	**100**
10	0.2	1	5	93.0	**97.3**	97.1	95.2	**99.8**	98.2	98.5	97.9	**100**	**100**	99.9	99.8
10	0.5	1	5	90.4	**94.7**	86.5	85.6	**98.5**	98.4	95.0	95.2	99.8	**99.9**	98.9	99.2
10	0.8	1	5	94.9	**97.0**	38.7	73.8	98.6	**99.6**	39.0	90.0	99.8	**100**	13.2	95.8
10	-0.8	5	1	89.9	89.4	42.2	**90.5**	98.0	97.7	48.4	**98.2**	99.7	99.6	62.7	**99.8**
10	-0.5	5	1	**63.5**	61.4	43.6	61.0	**85.7**	84.2	53.0	85.5	97.0	97.1	67.4	**97.8**
10	-0.2	5	1	42.6	35.2	**45.8**	39.7	**59.9**	51.4	55.7	54.3	**76.2**	71.2	69.7	71.6
10	0.2	5	1	48.1	46.2	45.1	**51.2**	**65.9**	59.9	54.5	63.7	**79.7**	78.2	65.3	75.3
10	0.5	5	1	75.4	**75.7**	43.1	71.4	88.1	**89.8**	51.2	82.7	96.3	**96.8**	56.8	91.7
10	0.8	5	1	96.9	**97.5**	42.2	81.1	99.0	**99.7**	47.7	89.0	**100**	99.8	47.9	95.1
10	-0.8	5	5	**95.8**	95.1	83.7	90.4	**100**	**100**	98.8	**100**	**100**	**100**	**100**	**100**
10	-0.5	5	5	93.7	**94.5**	91.6	85.7	**100**	**100**	99.7	99.9	**100**	**100**	**100**	**100**
10	-0.2	5	5	92.6	94.4	**96.8**	90.6	**99.9**	99.7	99.5	99.1	**100**	**100**	**100**	**100**
10	0.2	5	5	92.6	95.7	**96.4**	93.2	99.4	**99.5**	99.1	98.3	99.7	**100**	**100**	99.9
10	0.5	5	5	93.0	**94.1**	85.2	87.5	97.2	**98.0**	95.6	94.6	99.7	**99.9**	97.9	97.8
10	0.8	5	5	**95.5**	**95.5**	37.2	74.3	98.3	**99.0**	38.7	89.5	**100**	**100**	12.1	94.8