262
Reihe Ökonomie
Economics Series

Sensitivity Analysis of SAR Estimators

A numerical approximation

Shuangzhe Liu, Wolfgang Polasek, Richard Sellner

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

262
Reihe Ökonomie
Economics Series

Sensitivity Analysis of SAR Estimators

A numerical approximation

Shuangzhe Liu, Wolfgang Polasek, Richard Sellner

January 2011

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Shuangzhe Liu
University of Canberra
Kirinari Street
Canberra ACT 2601, Australia
Shuangzhe.Liu@canberra.edu.au

Wolfgang Polasek – *Corresponding author*
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
email: wolfgang.polasek@ihs.ac.at

Richard Sellner
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
email: richard.sellner@ihs.ac.at

Abstract

Estimators of spatial autoregressive (SAR) models depend in a highly non-linear way on the spatial correlation parameter and least squares (LS) estimators cannot be computed in closed form. We first compare two simple LS estimators by distance and covariance properties and then we study the local sensitivity behavior of these estimators using matrix derivatives. These results allow us to calculate the Taylor approximation of the least squares estimator in the spatial autoregression (SAR) model up to the second order. Using Kantorovich inequalities, we compare the covariance structure of the two estimators and we derive efficiency comparisons by upper bounds. Finally, we demonstrate our approach by an example for GDP and employment in 239 European NUTS2 regions. We find a good approximation behavior of the SAR estimator, evaluated around the non-spatial LS estimators. These results can be used as a basis for diagnostic tools to explore the sensitivity of spatial estimators.

Keywords

Spatial autoregressive models, least squares estimators, sensitivity analysis, Taylor Approximations, Kantorovich inequality

JEL Classification

C11, C15, C52, E17, R12

Contents

1 Introduction

Spatial autoregression (SAR) models for the error covariance structure have been studied and applied to a wide range of areas in e.g. economics, demography, geography, biology, epidemiology, statistics and scientific modeling. See e.g. Anselin (1988, 1999), LeSage (1997, 1998), LeSage and Pace (2004, 2009), LeSage and Polasek (2008) and Polasek, Sellner and Llano (2009). On the other hand, sensitivity analysis for least-squares (LS) estimators in several models has been established in recent years. One of the approaches for such results is to use Taylor series to approximate the estimator under study, as used e.g. in a repeated multivariate sampling model by Wang et al. (1994). To our knowledge, however, a sensitivity analysis for LS estimators in spatial models has not been reported.

Spatial models have become popular in recent decades, but classical estimators even for a simple SAR model can suffer from numerical computation problems especially for large dimensional problems. For large cross sections the introduction of a spatial lag requires the inversion of the large spread matrix $R = I_n - \rho W$ (see below), which is of the dimension n, the number of observations. Thus it would be desirable to examine if simple approximations of spatial estimators can be found without inverting the spread matrix. Also, for Bayesian MCMC estimation such approximations can be used in case we have to use a proposal density in a Metropolis step.

In the present paper, we consider a spatial model of order 1, the SAR(1) model, for the error covariance structure, and we are interested in the sensitivity analysis of the LS estimators in this model. We will use a Taylor approximation with respect to the spatial correlation parameter ρ, similar to the approach of the repeated sampling model with unequal sample size that was studied by Wang et al. (1994). The variance matrix of the LS estimator of the SAR model has a variance matrix that is a non-linear function of the spatial autocorrelation parameter ρ. In this respect the SAR models are different from the repeated sampling models, where the LS estimator is a linear function of the additional correlation parameter. The numerical disadvantage in the estimation of SAR models is that the spread matrix R (which is a function of the large spatial neighborhood matrix W) depends on ρ and needs to be inverted, which can be time consuming. The question is if the inversion can be avoided and do good approximations exist, and if so, what estimators and what approaches should be used?

First we propose the 'pseudo' LS estimator and we show that be expanded in a Taylor series around the non-spatial LS estimator of a linear regression model. Then we discuss how to measure the distance between the LS estimator and the 1st or 2nd order Taylor approximation of the pseudo LS estimator. Also, we discuss how these covariance matrix of these estimators can be evaluated by the Kantorovich inequality.

The structure of the next sections of the paper is as follows. In section 2 we introduce the SAR model and the possible estimators. We continue with making sensitivity analysis and efficiency comparisons in section 3. The Taylor approximations of the estimators are established in section 4. The results are illustrated by a spatial estimation example involving European regional economic data in section 5. Finally some concluding remarks are made in section 6.

2 LS estimators in the SAR model

Let us consider here the following notation for SAR models, i.e. for the $n \times 1$ cross-sectional observations y of the form

$$y = \rho W y + X\beta + u, \quad u \sim N[0, \sigma^2 I_n], \tag{1}$$

where ρ is the spatial autocorrelation parameter (a scalar), W is a $n \times n$ spatial weight matrix normalized with row sums 1, β is a $n \times 1$ parameter vector, I_n is a $n \times n$ identity matrix, u is a $n \times 1$ error vector and follows a normal distribution with a $n \times 1$ mean vector centered at 0 and a $n \times n$ variance matrix $\sigma^2 I_n$.

The SAR model (1) can be written for known spatial autocorrelation ρ in the spatial filter (SF) form

$$Ry = X\beta + u, \quad u \sim N(0, \sigma^2 I_n). \tag{2}$$

By inversion of the spread matrix $R = I_n - \rho W$ we get the reduced form (RF) of the SAR model

$$y = R^{-1} X\beta + R^{-1} u = Z\beta + v, \quad v \sim N[0, \sigma^2 \Sigma(\rho) = \sigma^2 (R'R)^{-1}], \tag{3}$$

where the reduced form (RF) can be also written by the following transformed variables of the SAR model:

$$Z = R^{-1} X, \quad v = R^{-1} u. \tag{4}$$

The RF implies a heteroskedastic model with covariance matrix $Cov(y) = Cov(v) = \sigma^2 \Sigma(\rho) = \sigma^2 (R'R)^{-1}$. Obviously, the variance matrix of the reduced form $\Sigma = \Sigma(\rho)$ is a non-linear function of the spatial correlation parameter ρ.

In the following we list LS estimators for the β coefficients in the SAR models, that follow from these different ways of looking at the SAR model.

1. First, we get the ordinary LS (OLS) estimator of β if we set $\rho = 0$ in the SAR model (1), i.e. the $SAR(\rho = 0)$ model is just the linear regression model $y = X\beta + u$ and is given by

$$b_0 = (X'X)^{-1}X'y. \tag{5}$$

The covariance matrix of this LS estimate is the same as in the ordinary regression model.

$$\begin{aligned} Cov(b_0) &= Cov[(X'X)^{-1}X'y] \\ &= \sigma^2(X'X)^{-1}X'Cov(y)X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}. \end{aligned} \tag{6}$$

2. Second, conditionally on a known $\rho \in (-1, 1)$ for row-normalized W, we find the spatial filter (SF) form (2) also known as the SF model $Ry \sim N[X\beta, \sigma^2 I_n]$ and we obtain the LS estimator b_r for β or in brief the SF-LS estimator

$$b_r = (X'X)^{-1}X'Ry. \tag{7}$$

This estimator b_r differs from the OLS estimator b_0 only by the spatial filter transformation Ry, which replaces the dependent variable y. The covariance matrix of this estimator is

$$\begin{aligned} Cov(b_r) &= Cov[(X'X)^{-1}X'Ry] \\ &= (X'X)^{-1}X'Cov(Ry)X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}X'RR'X(X'X)^{-1} \end{aligned} \tag{8}$$

Note that $b_r = b_r(\rho)$ also reduces to the OLS estimator $b_0 = b_r(0)$ for $\rho = 0$, because $R = I$.

3. Third, we consider a 'pseudo' LS estimator b_z of β for the reduced form model (3), ignoring the covariance structure

$$b_z = (Z'Z)^{-1}Z'y = H^{-1}h \tag{9}$$

with the transformed (spatially filtered) regressors $Z = R^{-1}X$ (but untransformed y) and we define the 2 components

$$H = Z'Z = X'(RR')^{-1}X, \quad h = Z'y = X'R'^{-1}y. \tag{10}$$

The covariance matrix of this 'pseudo' LS estimator is

$$\begin{aligned} Cov(b_z) &= Cov[(Z'Z)^{-1}Z'y] \\ &= (Z'Z)^{-1}Z'Cov(y)Z(Z'Z)^{-1} \\ &= \sigma^2(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1}, \end{aligned} \tag{11}$$

where we have used the correct covariance $Cov(y) = \sigma^2\Sigma$. In case of $\Sigma = I_n$ we have the result

$$Cov(b_z) = \sigma^2(Z'Z)^{-1}.$$

Note that $b_z = b_z(\rho)$ also reduces to b_0 for $\rho = 0$.

In addition, for the reduced form (RF) model (3) the correct GLS estimator b_{GLS} of β is given by

$$\begin{aligned} b_{GLS} &= (Z'\Sigma^{-1}Z)^{-1}Z'\Sigma^{-1}y = [Z'(R'R)Z]^{-1}Z'(R'R)y \\ &= (X'X)^{-1}X'Ry \\ &= b_r, \end{aligned} \tag{12}$$

which is the same as the LS of the SF model in (7). The covariance matrix of the GLS estimator is therefore

$$\begin{aligned} Cov(b_{GLS}) &= Cov(b_r) = \sigma^2(X'X)^{-1} \\ &= \sigma^2(Z'\Sigma^{-1}Z)^{-1} \end{aligned} \tag{13}$$

It is clear that ρ plays an important role for spatial modeling and estimation. The behavior of the estimators when the value of ρ changes around zero or the relationship between the estimators should be important information for spatial models. Therefore the sensitivity of the estimators with respect to ρ is studied in the next section.

3 Local SAR sensitivity analysis

For the local sensitivity analysis for the SAR models we will use the following estimators, which we summarize with their corresponding regression models:

- $b_0 = b_{OLS}$ is the LS estimator in the model $y = X\beta + u,\ u \sim N[0, \sigma^2 I_n]$. The covariance matrix is $Cov(b_0) = \sigma^2(X'X)^{-1}$.
- b_1 is the LS estimator in the spatial lag-1 model $Wy = X\beta + u,\ u \sim N[0, \sigma^2 I_n]$, the basic linear regression model explaining the spatial lag Wy. The covariance matrix is the same as before: $Cov(b_1) = \sigma^2(X'X)^{-1}$.
- b_r is the LS (or SF-LS) estimator in the spatial filter (SF) model $Ry = X\beta + u,\ u \sim N[0, \sigma^2 I_n]$, the linear regression model explaining the spatial filter Ry, where y is 'filtered' by the spread matrix $R = I_n - \rho W$. The covariance matrix is $Cov(b_r) = \sigma^2(X'X)^{-1}$.
- b_z is the 'pseudo' LS estimator in the reduced form (RF) model $y = Z\beta + v$ with $Z = R^{-1}X$ and instead of $v \sim N[0, \sigma^2\Sigma]$ we impose the uncorrelated error matrix $\Sigma = I_n$. The covariance matrix is $Cov(b_z) = \sigma^2(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1} = \sigma^2(Z'Z)^{-1}$.

3.1 Sensitivity analysis for the spatial filter estimator b_r

For the spatial filter (SF-LS) estimator b_r we find a simple linear relationship, which shows the difference to the LS estimator b_0.

Theorem 1 (The SF-LS estimator b_r) *The spatial filter (SF) estimator in the SAR model (conditional on ρ) can be expressed as linear combination of two simpler LS estimators*

$$b_r = b_0 - \rho(X'X)^{-1}X'Wy = b_0 - \rho b_1, \tag{14}$$

and therefore the squared distance of b_r to the LS estimator b_0 is given by

$$||b_r - b_0||^2 = \rho^2 b_1' b_1 = \rho^2 y'W'X(X'X)^{-2}X'Wy.$$

Proof 1 *For b_r we can make the following substitutions*

$$\begin{aligned} b_r &= (X'X)^{-1}X'Ry \\ &= (X'X)^{-1}X'(I - \rho W)y \\ &= (X'X)^{-1}X'y - \rho(X'X)^{-1}X'Wy \\ &= b_0 - \rho(X'X)^{-1}X'Wy = b_0 - \rho b_1. \end{aligned}$$

We see that the difference between the OLS and the SF-LS estimator $b_0 - b_r = \rho b_1$ is proportional to the spatial parameter ρ and the 1st order spatial lag-1 LS estimator b_1. Next, we want to find the derivative of b_r with respect to ρ, which measures actually the sensitivity of b_r upon a small change of ρ. For analytical and mathematical convenience, we use the differential notation which is mathematically equivalent to the derivative. The notation of the matrix calculus follows Magnus and Neudecker (1988/1999).

Theorem 2 (The derivative of the SF estimator b_r) *The derivative of the spatial filter b_r in (7) with respect to ρ is the negative LS estimator in the linear model for explaining the 1st order spatial lag:*

$$\partial b_r / \partial \rho = -(X'X)^{-1} X'Wy = -b_1. \tag{15}$$

Proof 2 *Using the result $b_r = b_0 - \rho b_1$ in (14), we get the differential of the b_r estimator with respect to ρ:*

$$db_r = -b_1 d\rho = -(X'X)^{-1} X'Wy d\rho.$$

By rearranging terms, we establish the derivative.

We can interpret this remarkable result that the direction of the first order correction is the OLS estimator with respect to the spatial neighbors. This results follows from the presence of the spread matrix R in the b_r estimator. The spread matrix can be interpreted as a correction of the identity matrix with respect to the neighborhood structure W of the cross section model. It is the direction of this ' covariance correction' that we get as the result of the differencing operation. Thus, a spatial lag-1 model explains the direction of the correction in a SAR model and is estimated by b_1. The spatial ρ is just the length of this direction.

For the pseudo LS estimator b_z we cannot get results that can be presented in a similar simple way. However, we will use the matrix differential technique and the Taylor approximation to get a similar result, as it is shown in the next subsection.

3.2 First order sensitivity analysis for the estimator b_z

This section gives the sensitivity of the pseudo LS estimator of the reduced form of the SAR model (3).

Theorem 3 (Sensitivity analysis of the pseudo LS estimator b_z) *The derivative of b_z with respect to ρ takes into account the transformed variables of the estimator*

$$\begin{aligned}\partial b_z/\partial\rho &= H^{-1}[\partial h/\partial\rho - (\partial H/\partial\rho)b_z)] \\ &= H^{-1}[X'R'^{-1}W'R'^{-1}y - (X'R'^{-1}(W'R'^{-1} + R^{-1}W)R^{-1}X)b_z] \\ &= H^{-1}[h_r - H_r b_z] \\ &= P \qquad (16)\end{aligned}$$

with H given in (10) and we define the two auxiliary quantities

$$\begin{aligned}h_r &= X'R'^{-1}W'R'^{-1}y \quad \text{and} \\ H_r &= X'R'^{-1}(W'R'^{-1} + R^{-1}W)R^{-1}X.\end{aligned}$$

Proof 3 *For the sensitivity analysis of $b_z = H^{-1}h$, we need the differential of b_z with respect to ρ:*

$$\begin{aligned}d\ b_z &= (d\ H^{-1})h + H^{-1}(d\ h) \\ &= -H^{-1}(d\ H)H^{-1}h + H^{-1}d\ h \\ &= H^{-1}[d\ h - H^{-1}(dH)H^{-1}h] \qquad (17)\end{aligned}$$

where we used differentials and partial derivatives that are given by

$$\begin{aligned}dR &= -W d\rho \\ dR^{-1} &= R^{-1}WR^{-1}d\rho \\ dR'^{-1} &= R'^{-1}W'R'^{-1}d\rho \\ dH &= d(X'R'^{-1}R^{-1}X) = X'(dR'^{-1})R^{-1}X + X'R'^{-1}(dR^{-1})X \\ &= X'R'^{-1}(W'R'^{-1} + R^{-1}W)R^{-1}X d\rho \\ \partial H/\partial\rho &= X'R'^{-1}(W'R'^{-1} + R^{-1}W)R^{-1}X \\ dh &= X'R'^{-1}W'R'^{-1}y d\rho \\ \partial h/\partial\rho &= X'R'^{-1}W'R'^{-1}y \qquad (18)\end{aligned}$$

with the spread matrix $R = I_n - \rho W$.

Next we evaluate the derivatives at $\rho = 0$. Because $R(\rho = 0) = I_n$, we get $H(\rho = 0) = X'X$ and

$$\begin{aligned}h_r(\rho = 0) &= X'R'^{-1}W'R'^{-1}y = X'W'y = h_0^r \\ H_r(\rho = 0) &= X'R'^{-1}(W'R'^{-1} + R^{-1}W)R^{-1}X = X'(W' + W)X = H_0^r.\end{aligned}$$

Note that $W' + W$ is symmetric and takes the role of a precision matrix.

Corollary 1 *As a special case we get the derivative in the uncorrelated case, denoted by $P(\rho = 0) = P_0$, which leads to the following expression:*

$$
\begin{aligned}
P_0 &= H^{-1}[h_r - H_r b_z] \\
&= (X'X)^{-1}[X'W'y - X'(W'+W)Xb_0].
\end{aligned} \tag{19}
$$

Theorem 4 (Distance between the estimators b_z and b_0) *The distance between the pseudo LS estimator b_z and the OLS estimator b_0 is given by*

$$||b_z - b_0||^2 = y'Vy, \tag{20}$$

where we have defined $b_z = Z^+y$, $b_0 = X^+y$, $V = Z'^+Z^+ + X'^+X^+ - Z'^+X^+ - X'^+Z^+$, $Z^+ = (Z'Z)^{-1}Z'$, and $X^+ = (X'X)^{-1}X'$.

Proof 4 *This follows by simplifications using the definitions of the estimators b_z in (9) and b_0.*

3.3 Second order sensitivity analysis for b_z

The second order local sensitivity derivative of the pseudo LS estimator (9) of the reduced form of the SAR model (3) is given in the next theorem.

Theorem 5 (2nd order sensitivity of the pseudo RF-LS estimator b_z) *For the peudo LS estimator $b_z = H^{-1}h$ in the RF model (3) we find*

$$
\begin{aligned}
Q &= \partial^2 b_z/\partial\rho^2 \\
&= -H^{-1}(\partial H/\partial\rho)H^{-1}[\partial h/\partial\rho - (\partial H/\partial\rho)b_z] \\
&\quad + H^{-1}[(\partial^2 h/\partial\rho^2) - (\partial^2 H/\partial\rho^2)b_z - (\partial H/\partial\rho)(\partial b_z/\partial\rho)] \\
&= -H^{-1}(\partial H/\partial\rho)H^{-1}[\partial h/\partial\rho - (\partial H/\partial\rho)b_z] \\
&\quad + H^{-1}[(\partial^2 h/\partial\rho^2) - (\partial^2 H/\partial\rho^2)b_z] \\
&\quad + H^{-1}(\partial H/\partial\rho)H^{-1}[(\partial H/\partial\rho)b_z - \partial h/\partial\rho] \\
&= -H^{-1}H_rH^{-1}[h_r - H_r b_z] \\
&\quad + H^{-1}[h_{rr} - H_{rr}b_z] \\
&\quad + H^{-1}H_rH^{-1}[H_r b_z - h_r]
\end{aligned} \tag{21}
$$

with the 2nd derivatives $h_{rr} = \partial^2 h/\partial\rho^2$ and $H_{rr} = \partial^2 H/\partial\rho^2$.

Proof 5 *We compute the first differential db_z (of the pseudo LS estimator $b_z = H^{-1}h$) in (9) of Theorem 3 and get*

$$
\begin{aligned}
d^2 b_z = & -H^{-1}(dH)H^{-1}[(dh) - (dH)H^{-1}h] \\
& +H^{-1}[(d^2 h) - (d^2 H)H^{-1}h - (dH)db_z] \\
= & -H^{-1}(dH)H^{-1}[(dh) - (dH)b_z] \\
& +H^{-1}[(d^2 h) - (d^2 H)b_z] \\
& +H^{-1}(dH)H^{-1}[(dH)b_z - (dh)]. \qquad (22)
\end{aligned}
$$

From the differentials in (22) and dh, dH and db, we establish the derivative results with $h_r = \partial h/\partial \rho$ and $H_r = \partial H/\partial \rho$ as given above. Now we find

$$
\begin{aligned}
\partial^2 h/\partial \rho^2 = & \, 2X'W'^2 R'^{-2} y \\
= & \, h_{rr}, \\
\partial^2 H/\partial \rho^2 = & \, 2X'[W'^2 R'^{-3} R^{-1} + W' R'^{-2} R^{-2} W + R'^{-1} R^{-3} W^2] X \\
= & \, H_{rr}, \qquad (23)
\end{aligned}
$$

where the last two equalities are obtained by taking the differentials of dh and dH, which are given in (18) above.

A simplified version of the second derivative for the $\rho = 0$ case is found in the following way: We compute the simplified second derivatives in (23) by

$$
\begin{aligned}
h_{rr}(\rho = 0) = & \, 2X'W'^2 y \\
H_{rr}(\rho = 0) = & \, 2X'[W'^2 + W'W + W^2]X \\
= & \, 2X'W^{\oplus} X \qquad (24)
\end{aligned}
$$

with the extended 'second order' weight matrix $W^{\oplus} = W'^2 + W'W + W^2$, which is symmetric.

Corollary 2 *With the simplified first order derivatives in (19) we get*

$$
\begin{aligned}
Q_0 &= Q(\rho = 0) \\
&= -H^{-1}H_rH^{-1}[h_r - H_rb_z] \\
&\quad +H^{-1}[h_{rr} - H_{rr}b_z] \\
&\quad +H^{-1}H_rH^{-1}[H_rb_z - h_r] \\
&= H^{-1}[h_{rr} - H_{rr}b_z] \\
&\quad +H^{-1}H_rH^{-1}[-h_r + H_rb_z + H_rb_z - h_r] \\
&= H^{-1}[h_{rr} - H_{rr}b_z] \\
&\quad +2H^{-1}H_rH^{-1}[H_rb_z - h_r] \\
&= (X'X)^{-1}[h_{rr} - H_{rr}b_0] \\
&\quad +2(X'X)^{-1}H_r(X'X)^{-1}[H_rb_0 - h_r] \\
&= 2(X'X)^{-1}[X'W'^2y - X'W^{\oplus}Xb_0] \\
&\quad +2(X'X)^{-1}X'(W'+W)X(X'X)^{-1}[X'(W'+W)Xb_0 - X'W'y]. \qquad (25)
\end{aligned}
$$

3.4 Efficiency comparisons

The results of the next theorem allow to make the main comparison between the estimators b_z and b_r.

Theorem 6 (Kantorovich inequality for b_z and b_r) *The covariance matrices of the two estimators b_z in (9) and b_r in (8) can be compared and establish the efficiency of b_z in terms of the covariance matrix of the b_r estimator:*

$$
\begin{aligned}
Cov(b_r) \le Cov(b_z) &\le k_1 Cov(b_r), \\
k_1 &= \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}; \qquad (26) \\
\Sigma_D = Cov(b_z) - Cov(b_r) &\le k_2\sigma^2(Z'Z)^{-1}, \\
k_2 &= (\sqrt{\lambda_1} - \sqrt{\lambda_n})^2; \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
|Cov(b_z)[Cov(b_r)]^{-1}| &\le k_3, \\
k_3 &= \prod_{j=1}^{n} \frac{(\lambda_j + \lambda_{n-j+1})^2}{4\lambda_j\lambda_{n-j+1}}; \qquad (28) \\
tr\ Cov(b_z)[Cov(b_r)]^{-1} &\le k_4, \\
k_4 &= \sum_{j=1}^{n} \frac{(\lambda_j + \lambda_{n-j+1})^2}{4\lambda_j\lambda_{n-j+1}}, \qquad (29)
\end{aligned}
$$

where $\lambda_1 \ge ... \ge \lambda_n > 0$ are the eigenvalues of $R'R$.

Note that the constant k_1 can be interpreted as least squares inefficiency, since it compares the RF estimator with the SAR spatial filter model, and k_2 is the upper bound or a size constant for the difference between the two covariance matrices of b_z and b_r. The constant k_3 is the upper bound for the ratio of determinants while k_4 is a bound for the trace of the 'ratio' of covariance matrices.

Proof 6 *The covariance matrices of the two estimators are given by* $Cov(b_z) = \sigma^2(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1}$ *as in (11) and* $Cov(b_r) = \sigma^2(Z'\Sigma^{-1}Z)^{-1}$ *as in (8), where the transformed variables of the reduced form are given in (3). Comparing them we find* $Cov(b_r) \leq Cov(b_z)$ *due to the Cauchy-Schwarz inequality*

$$(Z'\Sigma^{-1}Z)^{-1} \leq (Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1},$$

and $Cov(b_z) \leq k_1 Cov(b_r)$ *due to the Kantorovich inequality*

$$(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1} \leq k_1(Z'\Sigma^{-1}Z)^{-1},$$

where the constant k_1 *is given in (26). This constant* k_1 *was derived by using* $V = Z(Z'Z)^{-1/2}$ *and* $V'\Sigma V \leq k_1(V'\Sigma^{-1}V)^{-1}$*, for* $V'V = I$*; see e.g. Proposition 1 of Liu (1995, page 48).*

The difference between the covariance matrices of the pseudo RF-LS and the SF-LS estimators can be derived in closed form by

$$\begin{aligned} Cov(b_z) - Cov(b_r) &= \sigma^2(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1} - \sigma^2(Z'\Sigma^{-1}Z)^{-1} \\ &= \sigma^2(Z'Z)^{-1/2}[(Z'Z)^{-1/2}Z'\Sigma Z(Z'Z)^{-1/2} \\ &\quad -((Z'Z)^{-1/2}Z'\Sigma^{-1}Z(Z'Z)^{-1/2})^{-1}](Z'Z)^{-1/2} \\ &\leq M = k_2\sigma^2(Z'Z)^{-1} \end{aligned}$$

and we get the constant k_2 *from* $V'\Sigma V - (V'\Sigma^{-1}V)^{-1} \leq k_2 I$*, for* $V'V = I$ *with* $V = Z(Z'Z)^{-1/2}$*, see Liu and Neudecker (1994). Note that the difference is non-negative definite and the upper bound comes with* k_2*.*

The other 2 constants k_3 *and* k_4 *can be obtained in the following way:*

$$
\begin{aligned}
Cov(b_z)[Cov(b_r)]^{-1} &= (Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1}Z'\Sigma^{-1}Z, \\
|Cov(b_z)[Cov(b_r)]^{-1}| &= |(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1}Z'\Sigma^{-1}Z| \\
&= |(Z'Z)^{-1/2}Z'\Sigma Z(Z'Z)^{-1}Z'\Sigma^{-1}Z(Z'Z)^{-1/2}| \\
&\leq k_3, \qquad (30) \\
trCov(b_z)[Cov(b_r)]^{-1} &= tr(Z'Z)^{-1}Z'\Sigma Z(Z'Z)^{-1}Z'\Sigma^{-1}Z \\
&= tr(Z'Z)^{-1/2}Z'\Sigma Z(Z'Z)^{-1}Z'\Sigma^{-1}Z(Z'Z)^{-1/2} \\
&\leq k_4. \qquad (31)
\end{aligned}
$$

The results in (30) and (31) rely on the following inequalities $|V'\Sigma VV'\Sigma^{-1}V| \leq k_3$ *and* $tr(V'\Sigma VV'\Sigma^{-1}V) \leq k_4$*, for* $V'V = I$*, and these inequalities were shown e.g. in Theorem 1 in Liu (2000).*

Furthermore, we can derive an inequality for the difference of the covariance matrices of the pseudo and ordinary LS point predictions, i.e. $y_z = Zb_z$ and $y_r = Zb_r$:

$$
\begin{aligned}
Cov(Zb_z) - Cov(Zb_r) &\leq k_2\sigma^2 Z(Z'Z)^{-1}Z' \\
&\leq k_2\sigma^2 I_n, \qquad (32)
\end{aligned}
$$

where k_2 is the same constant as given above and $Z(Z'Z)^{-1}Z' \leq I$ is known from linear regression theory. Thus, the increase in uncertainty or efficiency loss also turns over to the same type of efficiency loss if it comes to predictions with the pseudo RF or SF estimator that are based on the SAR model.

Note that the constants k_1 and k_2 depend on the minimum and maximum eigenvalues of $R'R$, and therefore on the spread matrix R, which implies on ρ and the neighborhood matrix W. The dependence on W increases if the eigenvalues of W start dominating the size of ρ. In fact, the constant k_1 depends on $(c+1)^2/4c$, where $c = \lambda_1/\lambda_n$ is the condition number of $R'R$.

Therefore any $R'R$ matrix that increases the condition number will increase the constant k_1. We see that the covariance matrix of b_z can be almost as good as the one of b_r (and therefore b_z can be 'almost as good' as b_r) if k_1 is close enough to one. In particular, the variances on the main diagonal of Cov(b_z) have upper bounds by k_1 times the variances on the main diagonal of Cov(b_r):

$$Var(b_z(i)) \leq k_1 Var(b_r(i)) \quad \text{for} \quad i = 1, ..., k, \qquad (33)$$

where $b_z(i)$ and $b_r(i)$ are the elements of the vectors b_z and b_r.

Furthermore, we conclude that the covariance matrix of the pseudo LS estimator b_z can be almost as good as that one of b_r if the constant k_2 is close enough to zero. In particular, the differences of the variances on the main diagonal of Cov(b_z) and Cov(b_r) are upper bounded by k_2 times the main diagonal of $(Z'Z)^{-1}$ (besides σ^2). Equivalently, the main diagonal of the difference of Cov(Zb_z) and Cov(Zb_r) is upper bounded by k_2, apart from σ^2. In other words, the efficiency loss of the point predictors is measured by the covariance matrices of the predictors and is at most k_2.

Since the difference is positive for all type of comparisons we conclude that the pseudo RF estimator comes (necessarily) with more uncertainty than the SF-LS estimator in the spatial filter model. Knowing that ρ can reduce the uncertainty of all the estimated regression coefficients, the result is independent of the way efficiency comparison is made. The unknown ρ blows up the correlation structure of the residuals, and this property creates additional heteroskedasticity and does not reduce the uncertainty in the covariance matrix.

We are interested how the above findings of the approximate SAR estimators can translate into the questions as how good are predictions that are made by approximate SAR estimators. Let z be a vector of known regressor values where we make the prediction with the estimator b of β by $\hat{y} = z'b$ then we have $Var(\hat{y}) = Var(z'b) = z'Cov(b)z$, and we look at the difference of the covariance matrices of the predictions made by b_z and b_r

$$\begin{aligned} Cov(z'b_z) - Cov(z'b_r) &= z'Cov(b_z)z - z'Cov(b_r)z \\ &= z'[Cov(b_z) - Cov(b_r)]z \\ &\leq z'Mz, \end{aligned} \tag{34}$$

where $M = k_2\sigma^2(Z'Z)^{-1}$ is the upper bound matrix of the difference $\Sigma_D = Cov(b_z) - Cov(b_r)$.

4 Taylor approximation for the SAR estimator

Based on the first and second order derivative results for the pseudo LS estimator b_z of the SAR model from the previous section we find for the Taylor expansion of the LS estimator around the OLS location $b_0 = b_z(\rho = 0)$ by the mean value theorem of calculus (see Magnus and Neudecker 1988/1999, p. 113). The twice differentiable functions are:

$$\phi(c + u) = \phi(c) + d\phi(c; u) + d^2\phi(c + \theta u; u)/2 \quad \text{for} \ \ 0 < \theta < 1$$

, where the first derivative is $d\phi(c; u) = (db_z)\rho$ and the second derivative is $d^2\phi(c + \theta u; u) = (d^2 b_z)\rho^2$. In our case c denotes the point of the OLS location $(\rho = 0)$ and u is the value of the ρ parameter around 0.

$$b_z(\rho) = b_0 + (db_z/d\rho)\rho + (d^2b_z/d\rho)\rho^2\theta/2, \quad \text{for} \;\; 0 < \theta < 1, \tag{35}$$

and the first and second order differentials, db_z and d^2b_z, are given in Theorems 3 and 5, respectively.

4.1 First and second order Taylor approximation for SAR models

This section develops the Taylor approximation for the SAR model.

Theorem 7 (First and second order Taylor approximation) *The first order Taylor approximation of the $b_z(\rho)$ estimator around $b_0 = \hat{\beta}$ is:*

$$b_z(\rho) = b_0 + P_0\rho + O(\rho^2) \tag{36}$$

with P_0 given in (19). The 2nd order Taylor approximation for the SF-OLS estimator b_z around the OLS location $b_0 = \hat{\beta}$ is given by

$$b_z(\rho) = b_0 + P_0\rho + Q_0\rho^2/2 + O(\rho^3), \tag{37}$$

where the vectors P_0 and Q_0 are as given in (19) and (25), respectively. They are the first and second order derivatives (obtained in Theorems 3 and 5), evaluated at the uncorrelated case $\rho = 0$.

Proof 7 *: The result is obtained by plugging (16) and (21) into (35).*

5 Example: European NUTS2 regional data

In this section we demonstrate our approach by a numerical example. We make simulations by setting ρ to range from -0.3 to 0.3 in steps of 0.05 from the SAR model

$$y = \rho Wy + X\beta + u, \qquad u \sim N(0, \sigma^2).$$

and we apply their SAR sensitivity analysis to an empirical regional example with data from Europe (Eurostat). We estimate an simple GDP - employment relationship for 239 European NUTS2 [1] regions. Our variable of interest is

[1] The Nomenclature of Territorial Units for Statistics (NTUS), for the French nomenclature d'unités territoriales statistiques) is a geocode standard for referencing the subdivisions of countries for statistical purposes. The standard is developed and regulated by statistical office of the European Union, Eurostat.

the log of GDP in the year 2005 and as regressors we use the log of employment in 2005 and the population density. All data are taken from the Eurostat regional database. The spatial weight matrix W is row-normalized and constructed by the inverse distances between the regions. As distances we have used car travel times in minutes between the main locations of the NUTS2 regions, whereas we define the spatial cut-off point by 300 minutes. So if the car travel time between two regions exceeds 300 minutes, those regions are not defined as neighbors. The results of the SAR estimation program using the spatial-econometrics library in Matlab (see: www.spatial-econometrics.com) is shown in Table 1.

Table 1
SAR estimation results

Dependent Variable: $\log GDP_{2005}$			
coef	estimate (SD)	t-stat	p-prob
constant	4.241 (0.370)	11.468	0.0000
$\log employment_{2005}$	0.869 (0.050)	17.470	0.0000
$\log populationdensity_{2005}$	0.200 (0.036)	5.562	0.0000
ρ	0.062 (0.020)	3.172	0.0015
R^2			0.6300
$\bar{R}^2$			0.6269
observations			239
log-likelihood			-139.677

We have used the ρ of the estimation above as a comparison value and constructed the measures. The results given in Figures 1, 5 and 6 indicate the second-order approximation value is better than the first-order approximation and the approximation error is close to zero.

Figure 1 shows the Euclidean [2] distance between b_z and the OLS estimator and the Taylor series approximations as function of ρ's in the interval -0.3 to 0.3. Interestingly, the graphs are quite symmetric and diverge quadratically at both sides. While the first order approximation is only a good for very small ρ's (less than .1), the second-order approximation is good up to a rho value of 0.25, then it quickly deteriorates. The difference of the constants k_1 and k_2

[2] The values have been transformed for clarity of the depiction. First the distances have been multiplied by $1E^9$ and then logs have been taken, whereas the zero values have been set to the non-transformed value, which is zero.

lies in the fact that k_2 is the constant for the difference between 2 estimators where the upper bound matrix are quite different for these 2 cases.

Fig. 1. Distances (log transformation) between b_z and the OLS estimator and the Taylor series approximations for different ρ's

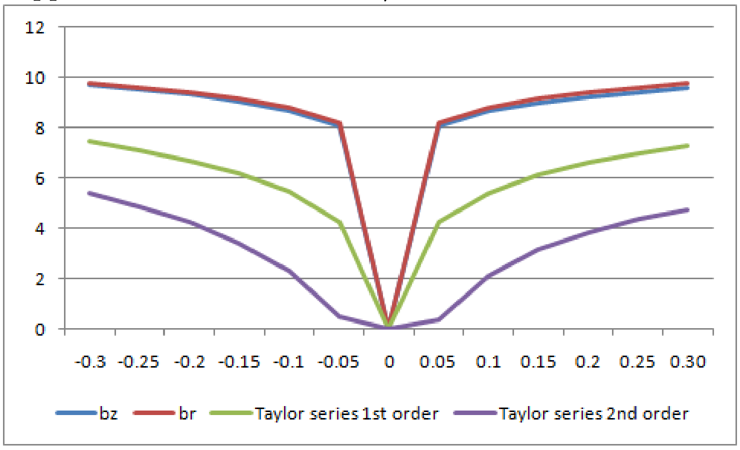


Figure 2 shows the behavior of the coefficient estimates in dependence of ρ. The SAR estimates of the model is about histogram at the $\rho = 0.05$ value of the chart. Moreover, all curves show a monotonic decreasing trend and the coefficient population density, which is the smallest, has also the flattest slope (least sensitive to ρ). The standard deviation of the employment coefficient is 0.05 and this is the change in the SAR estimate if we move say from $\rho = 0.2$ to $\rho = 0.1$.

Fig. 2. b_z for different ρ's

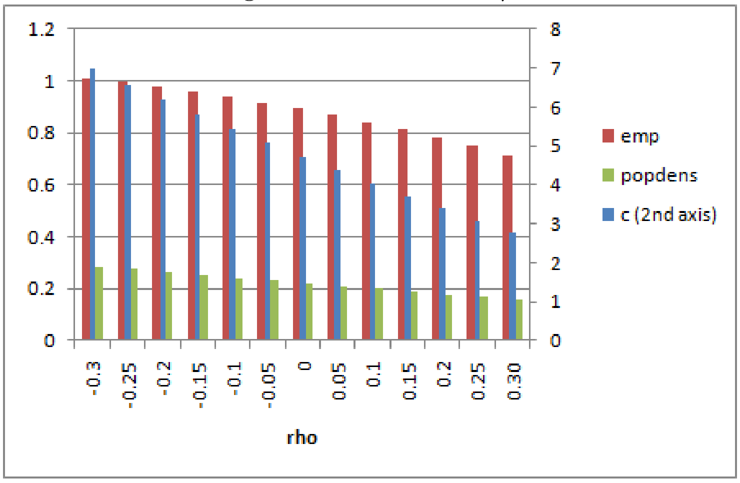


Figure 3 plots the Kantorovich constants k_1 and k_2 for different ρ's. Both curves are symmetric around 0 and show a quadratic behavior. Both curves show an increase of (efficiency loss) about 50% when rho approaches $+/-0.3$.

Figure 4 displays the Kantorovich constants k_3 and k_4 for different ρ's.

Fig. 3. The Kantorovich constants $k1$ and $k2$ for different ρ's

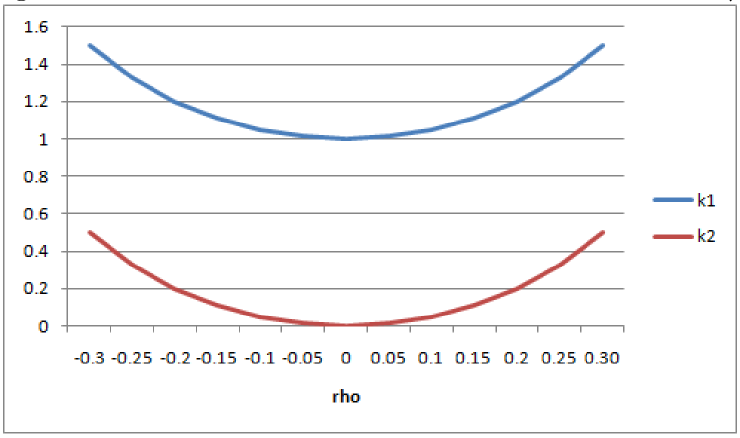


Fig. 4. The constants $k3$ and $k4$ for different ρ's

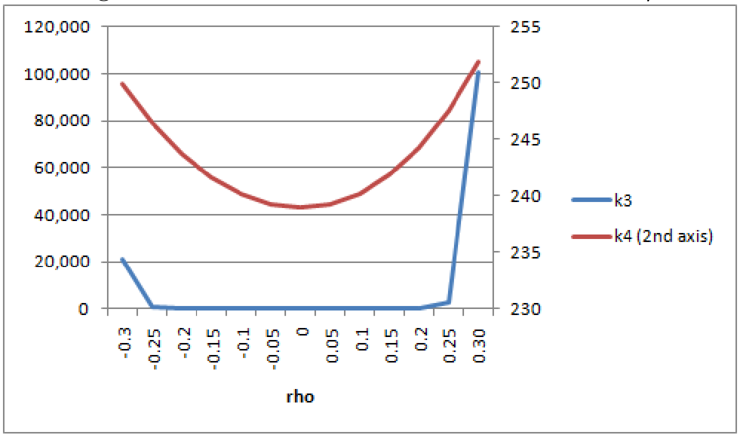
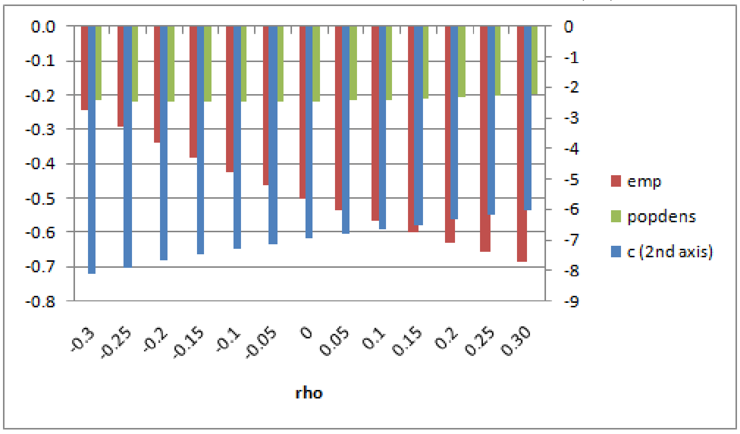
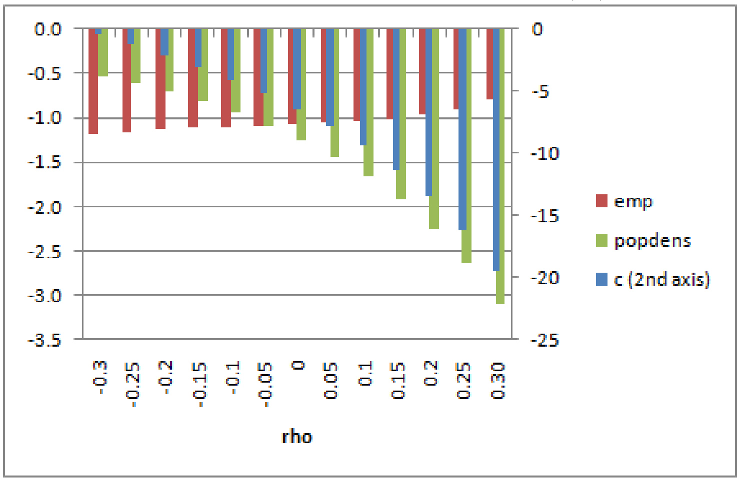


Fig. 5. NUTS2 example: The first order derivative (P) for different ρ's



Figures 5 and 6 show the first and second order derivatives of the b_z estimator of the regional NUTS2 example, respectively.

Fig. 6. NUTS2 example: Second order derivative (Q) for different ρ's



6 Conclusions

In this paper we have derived several results for the sensitivity analysis of the LS estimators in the spatial autoregressive (SAR) model. We used new results on the Kantorovich inequality to examine the quality of the approximation with respect to the difference of the covariance matrices. The main goal was to find the Taylor approximation with respect to ρ in the SAR model and to measure the difference of the LS estimator, which is non-linear function in ρ, from the ordinary LS estimator. In an empirical example involving regional economic data in Europe we have shown that the Taylor approximation of the LS estimator of the SAR model gives good approximation results for small ρ's, say up to $+/-0.4$. The efficiency loss according to the Kantorovich inequality is about 50% when ρ approaches $+/-0.3$. These values were found in an example involving regional economic growth in Europe.

There is an interesting result that needs further research: The Taylor approximation based on the pseudo OLS estimator b_z gives better results than for the spatial filter estimator b_r, as was shown in Figure 1. While the reason for this result is given in Theorems 1 and 4, this is rather surprising since b_r is more efficient than b_z. Thus, for approximating non-linear estimators it can be sometimes useful to look for simpler estimators that can be better approximated by a Taylor series. The approximation results are encouraging since they allow good first step approximations for non-linear LS methods or can be used as a proposal densities in the Metropolis step of a MCMC algorithm. An open question is if other or better approximations can be found for medium or large ρ values. In a further research paper, we intend to show that these approximations results can be applied to a larger class of spatial models.

Acknowledgements

The research had been initiated when Wolfgang Polasek visited Shuangzhe Liu at University of Canberra in 2008, and continued when Shuangzhe Liu took his OSP leave in July-Dec. 2009.

References

[1] Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht: Kluwer Academic Publishers.

[2] Anselin, L. (1999). Spatial Econometrics. Working Paper, Bruton Center, School of Social Science, University of Texas, Dallas.

[3] LeSage, J. P. (1997). Bayesian estimation of spatial autoregressive models. International Regional Science Review, 20: 113-129.

[4] LeSage, J. P. (1998). Spatial Econometrics. Technical report, University of Toledo. www.spatial-econometrics.com

[5] LeSage, J. P. and Pace, R. K. (2004). Models for spatially dependent missing data. Journal of Real Estate Finance and Economics, 29: 233-254.

[6] LeSage, J. P. and Pace, R. K. (2009). Introduction to Spatial Econometrics. New York: CRC Press.

[7] LeSage, J. P. and Polasek, W. (2008). Incorporating transportation network structure in spatial econometric models of commodity flows. Spatial Economic Analysis, 3(2): 225-245.

[8] Liu, S. (1995). Contributions to Matrix Calculus and Applications in Econometrics. Amsterdam: Thesis Publishers.

[9] Liu, S. (2000). On matrix trace Kantorovich-type inequalities. In Innovations in Multivariate Statistical Analysis-A Festschrift for Heinz Neudecker, edited by R.D.H. Heijmans, D.S.G. Pollock and A. Satorra. Dordrecht: Kluwer Academic Publishers. pp 39-50.

[10] Liu, S. and Neudecker, H. (1994). Several matrix Kantorovich-type inequalities. J. Math. Anal. Appl.

[11] Magnus, J. R. and Neudecker, H. (1988/1999). Matrix Differential Calculus with Applications in Statistics and Econometrics, revised edition, Chichester: John Wiley and Sons.

[12] Polasek, W., Sellner, R. and Llano, C. (2009). Interpolating Spanish Nuts 3 data by spatial Chow-Lin. forthcoming.

[13] Wang, S. G., Chow, S-C. and Tse, S-K. (1994). On ordinary least-squares methods for sample surveys. Statistics & Probablity Letters 20: 173-182.

Authors: Shuangzhe Liu, Wolfgang Polasek, Richard Sellner

Title: Sensitivity Analysis of SAR Estimators: A numerical approximation

Reihe Ökonomie / Economics Series 262

Editor: Robert M. Kunst (Econometrics)
Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996

Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 59991-555 • http://www.ihs.ac.at

ISSN: 1605-7996