Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ECONOMIA E STATISTICA AGROALIMENTARE

Ciclo XXV

MODELING THE SPATIAL DYNAMICS OF ECONOMIC MODELS

Settore concorsuale di afferenza: 13/A2 – Politica Economica
Settore scientifico-disciplinare: SECS-P/02 – Politica Economica

Presentata da: ***Annachiara Saguatti***

Coordinatore Dottorato	**Relatore**
Prof. Roberto Fanfani	**Prof. Luciano Gutierrez**

Esame finale anno 2013

Contents

List of tables

List of Figures

Acknowledgments

This work represents the conclusion of a long journey that led me to experience the most various feelings and strongly challenged me both from a professional and a personal point of view. Looking back at the experience of the last three years I feel the satisfaction for the long way covered and the lightness that comes from having successfully and finally completed a heavy task.

I have now the opportunity to thank the many people who supported, helped and taught me during this experience as a PhD student.

First and foremost I wish to thank Cosimo, who was by my side in the numerous moments of discouragement, offering warmth and support, patiently explaining me all the algebra I needed and revealing me all the secrets of Matlab. There is no other way in which I can say what he means to me but with a kiss. We will surely face new and bigger challenges in the future, but at least this level appears to be finally completed! “Estote parati” for what comes next!

A special and lovely thought goes to our daughter Alina who generously shared my energies with this work and unconsciously showed me the direction. I am sure we will have plenty of time to share exclusively with each other from now on.

My parents and sister followed this long journey discreetly, supporting me when I needed their help and being able to step back when Cosimo and I needed to face difficulties on our own. I know they are always there for us.

Thanks to my friends, who were ready to make my efforts lighter with their presence and supported me in the final rush, tolerating my being constantly unavailable.

I cannot forget the friends and colleagues I came across in these years. With some of them I shared the office and chatting (almost) every day in the last three years during the coffee breaks that I did not take because of the lack of milk, with some others I shared the frustrations that any PhD student feels, particularly during the typical “second-year crisis”. I am particularly grateful to Maria Grazia and Nica, who were there for supporting me when the discouragement shaded the whole picture, despite the distance and their own personal situations. I really felt welcome every time I needed to talk. To all of them I am grateful because their presence made me feel less lonely during the lonely work of a PhD student.

I am most grateful to Professor Luciano Gutierrez, my PhD supervisor, for the confidence he placed in me, for having accepted to be involved in my research interests and sharing his with me. He agreed to follow me in my final race against time despite his numerous tasks and guided me to cross the finish line.

I wish to express my warm and sincere thanks to Dr. Kenneth Erickson for providing me with the dataset on which I worked in the third chapter of this thesis. However, his contribution to this work went much further, from sending me references to encouraging me to ask more and more questions. The image of the gold miner attached to one of his emails inspired me a lot.

Many thanks to Professor Jihai Yu for providing me with his Matlab code file for his QML estimator.

My tutor, Professor Cristina Brasili gave me the opportunity to work autonomously and challenged me by involving me in different research fields. I thank her for letting me attend the courses that mostly influenced my research: the Spatial Econometrics Advanced Institute in Rome and the course of panel data econometrics in Bertinoro.

The coordinator of the PhD program, Professor Roberto Fanfani, shared with his PhD students his immense experience and tried dearly to teach us important life lessons. Hopefully we learned at least how to properly format a report.

A thankful thought goes to Professor Francesca Bruno and Professor Mario Mazzocchi, for being encouraging every time I needed to be pushed and for showing me what it means being passionate for what one does.

Last but not least, I wish to thank the Department of Statistics of the University of Bologna and the Professors who work there, for offering me an office, some economic support and a stimulating environment, and the Italian Ministry of Education, Universities and Research for covering my PhD grant.

The responsibility for any error or imprecision that may occur in this work is entirely my own.

Introduction

This thesis is composed of three rather autonomous chapters, linked by the common interest in spatial econometric techniques, that address the topic from different points of view. The recent advances that have been characterizing the subject in recent years are mostly theoretical and have not found an extensive empirical application yet. In this work we aim at supplying, as exhaustively as possible, a review of the main tools of spatial econometrics and to show empirical applications of the most recently introduced estimators.

Since the late 1970s, spatial econometrics has been growing as a distinct branch of econometrics, originally confined in the domain of regional sciences, in which the spatial dimension of the data, in both its forms of spatial dependence and heterogeneity, is more evident. The rather young age of the discipline makes spatial econometrics a field in which theoretical advances are numerous and ongoing. They are currently concentrated on the treatment of spatiotemporal data. This is due, among other reasons, to the growing availability of datasets in which the informative contents of both the cross-sectional and the temporal dimensions of the data can be explored.

Although a quite wide literature has been devoted to reviewing the techniques of spatial econometrics in the last 25 years (Anselin 1988a; Getis et al. 2004; Arbia 2006; LeSage and Pace 2009 among others), an updated dissertation, capable of collecting all the most recent theoretical advances in the discipline together with its bases is still lacking. With this purpose, the first chapter contains a review of the widely known spatial cross-sectional models and a taxonomy for the less common models for panel data. Together with the description of the models, also the main estimation techniques are discussed, highlighting the advantages and disadvantages that characterize each one of them. We argue that, despite the numerous alternatives that the econometric theory provides for the treatment of spatial (and spatiotemporal) data, empirical analyses are still limited by the lack of availability of the correspondent routines in statistical and econometric software.

In chapter 2 we focus on the estimation of spatiotemporal models. We overcome the lack of readily-available software by an autonomous programming of the routines that, albeit necessary to our analysis, were not available in any ready-to-use software packages.

Spatiotemporal modeling represents one of the most recent developments in spatial econometric theory and the finite sample properties of the estimators that have been proposed are currently being tested in the literature (Yu et al. 2008; Kukenova and Monteiro 2009; Jacobs et al. 2009). Our purpose is to provide a comparison between some estimators for a dynamic panel data model under certain conditions, by means of a Monte Carlo simulation analysis. Holding the assumption of homoskedasticity of the errors, we focus on different settings, which are characterized either by fully stable or quasi-unit root series. We also investigate the extent of the bias that is caused by a non-spatial estimation of a model when the data are characterized by different degrees of spatial dependence. To our knowledge, although the theoretical consequences of ignoring spatial dependence have been extensively studied, no empirical study is available for the assessment of the effects of such a misspecification in terms of bias of the estimates of the model coefficients.

Finally, chapter 3 provides an empirical application of what the previous chapters only theoretically of fictionally study. This is done by choosing a relevant and prolific field of analysis, in which spatial econometrics has only found limited space so far, in order to fully explore the value-added of considering the spatial dimension of the data. In particular, we estimate a spatial dynamic panel data model that studies the determinants of cropland value in Midwestern U.S.A. in the years 1971-2009. We adopt the present value model as the theoretical framework, and therefore focus on the relationship between land value and cash rents, expecting to find a positive one. This would be consistent with the present value model, which considers that the value of an income-producing asset such as land is the capitalized value of the current and future stream of earnings from owing that asset. We believe the employed dataset represents an improvement with respect to earlier studies because it refers to a rather homogeneous sample of States and only to cropland rather than farmland in general, thus excluding the value of buildings from the analysis. This appears to be a favorable situation, because buildings are usually excluded from the statistics on cash rents. Although the conclusions that we present should only be considered as preliminary results, we argue that they are already apt to convey the importance of taking spatial effects into consideration when addressing this field of analysis.

1 Spatial econometric models and estimation strategies

1.1 Introduction

The development of spatial econometrics as a distinct branch of econometrics dates back to 1970s. If a year is to be given as a conventional birth date for this rather young discipline, this is 1979, when Paelinck and Klaassen (1979) published the volume “Spatial Econometrics”. As the authors mention in the introduction of that volume, spatial econometrics was born in the context of regional and urban econometric modeling. This is indeed a context where the spatial (geographical) dimension of the data is easily conceivable and this is also the realm in which spatial econometrics remained confined until more recent years. Still at the end of the 1980s, Anselin defined spatial econometrics as closely related to the requirements of modeling in regional sciences:

> “I will consider the field of spatial econometrics to consist of those methods and techniques that, based on formal representation of the structure of spatial dependence and spatial heterogeneity, provide the means to carry out the proper specification, estimation, hypothesis testing, and prediction for models in regional science”. (Anselin 1988a, page 10)

Only in more recent years and in parallel with the growth of software availability, spatial econometrics has entered the general toolbox of applied econometrics. Theoretical econometrics has also started to deal with spatial issues and this has resulted in great advances that, in the last ten years, have focused on the field of space-time analysis, which appears to be the current frontier in spatial econometrics for what it concerns both modeling and testing. Anselin (2010) provides a thorough, albeit personal, analysis of the development of spatial econometrics in the past 30 years.

This chapter focuses on spatial econometrics with particular interest in the topics of modeling and estimation of spatial models whereas only minor reference is made to the literature about specification testing methods. Section 1.2 focuses on the definition of the spatial effects as the elements that justify the existence of spatial econometrics as a distinct discipline according to Anselin (1988a). The specification of the spatial weight matrix as the main econometric tool that allows to model spatial interactions is treated in section 1.3. A taxonomy is then provided for cross-sectional spatial models (section 1.4) and the main estimation strategies suggested by the literature are discussed (section 1.5). Section 1.6 focuses on the testing methods that have been developed in order to discriminate between the different model specifications in a cross-sectional context, before turning to space-time analysis with a review of the literature on the specification of spatial panel data models (section 1.7) and the most common estimation techniques (section 1.8). The last section (1.9) addresses the topic of software availability, which appears to be a major issue in determining the extent to which spatial econometrics techniques are applied by empirical researches.

1.2 Spatial effects

Spatial econometrics techniques are specifically designed in order to deal with the spatial dimensions of data, which can take the form of spatial interaction (*spatial autocorrelation*) and spatial structure (*spatial heterogeneity*), which have been described in detail by Anselin (1988a; 2001).

1.2.1 Spatial autocorrelation

Spatial dependence is defined as "the existence of a functional relationship between what happens at one point in space and what happens elsewhere" and has to do with the concept of relative location of a spatial unit *i* with respect to other spatial units (Anselin 1988a, page 11; Abreu et al. 2005).

Spatial dependence is often called spatial autocorrelation and, although the two concepts do not totally overlap, spatial autocorrelation being a weaker form of spatial dependence. However, following Anselin (2001), we will use the two terms interchangeably.

Two main sources of spatial dependence can be listed (Le Gallo 2002): the first one is related to the spatial dimension of the data originated by interactions between units. In statistical terms, we define a spatial stochastic process, or spatial random field, as a collection of random variables Y indexed by location: $\{Y_i, i \in L\}$, where the index L is either a continuous surface or a finite set of discrete locations. In this work, we only take into account the case in which L is a finite set of discrete locations, $L = \{1,2, \dots, N\}$, where $N \in \mathbb{N}$. Notice that we identify each spatial unit with an index in L: for example, if $L = \{1,2\}$, we identify the two spatial units as unit 1 and unit 2. Spatial autocorrelation is defined by the moment condition: $Cov[y_i y_j] \neq 0$ for $i \neq j$. The second source of spatial dependence is model misspecification, which can be caused by omitted spatially autocorrelated variables (Fingleton 1999), a wrong functional form or the presence of measurement errors (Luc Anselin 1988a). In particular, a difference in the spatial scope of the phenomenon under study and the spatial level of observation can easily result in spill-overs across different spatial units and subsequent autocorrelation of the errors.

The magnitude of spatial dependence is not constant throughout all spatial units. Spatial autocorrelation between unit i and j depends on their relative location, according to Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970, page 236). Therefore positive spatial autocorrelation occurs when high/low values of a random variable are concentrated in neighboring spatial units. Differently, negative spatial autocorrelation occurs when high/low realizations of a random variable are surrounded by low/high values of that random variable in neighboring spatial units.

Differently from temporal autocorrelation, whose causal direction can be easily defined as going from past to present (to future), spatial autocorrelation has a multidirectional nature which makes econometric modeling more complicated.

1.2.2 Spatial heterogeneity

Spatial heterogeneity is related to the "lack of stability over space of the behavioral or other relationships under study" (Anselin 1988a) and is also called effect of "absolute location" which pertains to being located at a particular point in space (Abreu et al. 2005).

This type of structural instability can take the form of heteroskedasticity or parameter instability over space (Anselin 2001). Given a set of spatial units L, partitioned into R non overlapping subsets L_r , with $r = 1, 2, \dots, R$, heteroskedasticity consists in non-constant error variances that can be formally expressed as $Var[\varepsilon_i] = \sigma_r^2$ when $i \in L_r$; this problem can be due to different causes, like omitted variables or other misspecifications and can be addressed by standard econometric tools. The most popular form of parameter instability is specified as varying regression coefficients across spatial regimes, $y_i = f(x_i, \beta_r)$.

These two forms of spatial heterogeneity can be jointly present and can also be associated to the presence of spatial autocorrelation. A further difficulty is represented by the fact that spatial autocorrelation and heterogeneity might be "observationally equivalent" (Anselin 2001) and spatial autocorrelation of the residuals may be provoked by unmodeled spatial heterogeneity (Ertur et al. 2006).

1.3 The spatial weight matrix and spatial lag variables

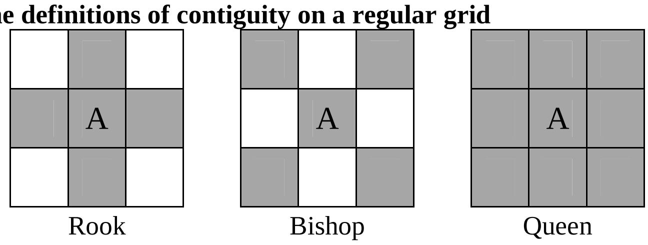
The way in which connectedness in space is to be incorporated in an econometric model is one of the main issues in spatial econometrics

(Anselin 1988a; Arbia 2006). This is usually done by means of the so-called spatial weight matrix. A weight matrix $\boldsymbol{W}$ is a square ($N \times N$), non-stochastic and symmetric matrix, whose elements w_{ij} measure the intensity of the spatial connection between units i and j and take on a finite and non-negative value. By convention, $w_{ii} = 0$. This is the main econometric tool for modeling spatial interactions among neighboring units and taking spatial dependence into account in econometric modeling. The concept of neighborhood and its several definitions are therefore at the basis of the construction of a spatial weight matrix.

1.3.1 Definitions of neighborhood

The most common definition of neighborhood is that of contiguity-based neighborhood. When spatial units are territories on a map, as it is often the case, contiguity is straightforwardly detected as the sharing of common boundaries. By approximating irregular polygons on a map by a regular grid, several kinds of contiguity can be defined after the game of chess (Figure 1.1): in case of "rook contiguity", the set of neighbors of unit A are those that share a common edge; "bishop contiguity" requires the sharing of a vertex; finally, according to the "queen contiguity" criterion, the neighbors of unit A are defined as those that share a vertex or an edge with it. Similarly, several orders of contiguity can be defined in a recursive way.

Figure 1.1. The definitions of contiguity on a regular grid



When spatial units are points instead of areas (cities, centroids, firms, etc.) different contiguity criteria can be employed: points can be considered to be neighbors if they are within a maximum distance from each other or boundaries can be generated by various spatial tessellations (Anselin 1988a). When neighboring units are to be identified on the basis of the distance that exists between them, a cut-off distance can be defined. Two spatial units i and j are considered to be neighbors if $0 \leq d_{ij} \leq D$, with d_{ij} an appropriate distance measure and D the cut-off distance above which any interaction is considered to be negligible.

Finally, neighborhood can be defined in terms of nearest neighbors. In this case, two spatial units i and j are said to be neighbors if $d_{ij} = \min_{1 \leq k \leq N}(d_{ik})$.

1.3.2 Spatial weight matrices

The most commonly used kind of spatial weight matrix is the contiguity matrix, which is based on the notion of binary contiguity and expresses the structure of neighbors as:

$$w_{ij} = \begin{cases} 1 & \text{if region } i \text{ is contiguous to region } j \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

Binary spatial weight matrices are therefore commonly constructed following the contiguity-based definition of neighborhood. This is also the simplest structure for a spatial weight matrix; yet, it appears to be able to provide only a restrictive representation of the spatial interactions.

Greater flexibility is possible when considering generalized sets of spatial weights. Cliff and Ord (1981) originally suggested the definition of the elements of $\boldsymbol{W}$ as a combination of distance measures and the relative length of common borders. Generally, especially in the regional sciences literature (Fingleton 1999; Ertur et al. 2006; Le Gallo and Dall'Erba 2006; Dall'Erba and Le Gallo 2008; Ramajo et al. 2008, among the others), the spatial weights are defined as an inverse function of distance:

$$\begin{cases} w_{ij} = 0 & \text{if} \quad i = j \\ w_{ij} = 1/d_{ij}^{\alpha} & \text{if} \quad d_{ij} \leq D \\ w_{ij} = 0 & \text{if} \quad \text{otherwise} \end{cases} \tag{1.2}$$

with α a parameter determined a-priori by the researcher, d_{ij} and D defined as in the previous section. When considering areal spatial units, distance can measured typically between centroids or capital cities, but other possibilities are also present in the literature (Arbia 2006).

The definition of distance is also crucial. Different distance metrics can be used in order to model geographical links between units, such as Euclidean distance, Manhattan distance, Minkowski distance. Alternative possible but less used measures that approximate geographical distance are travel time or transport costs. Reference to geographical localization of units can also be ignored in favor of other measures of distance: social distance (Doreian 1980), cultural distance (Eff 2008), socio-economic distance (Case et al. 1993), institutional distance also combined with geographical distance as in Arbia et al. (2007). When considering such alternative specifications of the weight matrix, it is important to preserve the exogeneity of the weights in order to avoid identification problems (Manski 1993): as Anselin (2002) warns, "if the same variables are used to compute a general distance metric as are included in the model, the weights are unlikely to remain exogenous" (page 18) and this should be taken into account.

One of the main criticisms that have been moved against spatial econometrics is that the choice of the spatial weight matrix to use is to some extent arbitrary. It is therefore always recommended to be driven by theoretical reasoning and to test the robustness of the results to the choice of $\boldsymbol{W}$.

The spatial weight matrix is often row-standardized: each element w_{ij} is divided by the row-sum $\sum_i w_{ij}$, so as to take values between 0 and 1. This makes the spatial parameter comparable between different models, but also implies a different interpretation of the spatial weights. In the most common case of a binary contiguity matrix, for example, the strength of the spatial connection between two units depends on the number of neighbors of each unit, the effect of any individual neighbor decreases as the number of neighbors increases and $\boldsymbol{W}$ is not symmetric anymore. The main reason for row-standardizing a spatial weight matrix is to ensure that some assumptions on the parameter space of the most common model specifications are verified, as we shall see in the following sections.

1.3.3 Spatial lag variables

Spatial weight matrices are to be used in the construction of spatially lagged variables. In the context of time series, it is straightforward to define a temporal lag of order k of a random variable y as $Lag_k(y) = y_{t-k}$. It is more difficult in spatial econometrics to define the lagged value of a variable y_i in space, because of the multidirectional nature of spatial proximity.

The spatial lag of the $N \times 1$ vector of observations of the random variable Y is defined in the spatial econometric literature as $\boldsymbol{W}y$. It is important to notice that, when $\boldsymbol{W}$ is row-standardized, the spatial lag of y_i, is the average of the values of that variable observed in the neighborhood of i: $\sum_j \mathrm{w}_{ij} \cdot y_j = \boldsymbol{W}y_i$.

1.4 Cross-sectional models with spatial autocorrelation

When spatial autocorrelation is present in the data, the hypothesis of independence between the observations is violated and inference based on Ordinary Least Squares (OLS) estimation is therefore not reliable. This is the reason for the need to pay great attention to the presence of spatial autocorrelation when estimating econometric models. In order to address this issue, taxonomy for the most popular cross-sectional spatial models is given, followed by a discussion of the estimation strategies that are usually employed.

1.4.1 A taxonomy for cross-sectional spatial models

Given a classical linear regression model, such as

$$y = \boldsymbol{X}\beta + \varepsilon \tag{1.3}$$

where y is the $N \times 1$ vector of the dependent variable, $\boldsymbol{X}$ is the $N \times K$ matrix of observations for the K independent variables, β is the $K \times 1$ vector of unknown coefficients and ε is the $N \times 1$ vector of errors, OLS estimation is based on the following assumptions, that make it the Best Linear Unbiased Estimator (BLUE):

OLS_1. Exogeneity of the regressors $\boldsymbol{X}$.

OLS_2. $E(\varepsilon) = 0$

OLS_3. $E(\varepsilon'\varepsilon) = \sigma^2\boldsymbol{I}$.

The presence of spatial dependence causes the violation of some of these hypotheses as the following sections will make clear, thus making OLS estimates inefficient or even biased. Spatial dependence can be incorporated in the specification of a linear regression model in different ways, particularly either in the form of a spatially lagged variable (the spatial lag of the dependent variable, $\boldsymbol{W}y$, or the spatial lag of an exogenous variable, $\boldsymbol{W}x$), or in the error structure, so that $E(\varepsilon_i\varepsilon_j) \neq 0$. These two forms can also be combined in the more complex Cliff-Ord model.

1.4.2 SAR models

The Spatial Autoregressive (SAR) model incorporates spatial dependence through a spatial lag of the dependent variable:

$$y = \lambda_1\boldsymbol{W}y + \boldsymbol{X}\beta + \varepsilon \tag{1.4}$$

where λ_1 is the so called spatial autoregressive coefficient and the other notation is unchanged. For the sake of simplicity the error terms are assumed to be $i.i.d.$ although heteroskedasticity can be variously incorporated (Anselin 1988a).

The introduction of the spatial lag of the dependent variable allows one to evaluate the effects of spatial dependence once the effects of the other regressors are controlled for; on the other hand, it also allows evaluating the impact of the other regressors once the effects of spatial dependence are wiped out.

It is important to notice that the term $\boldsymbol{W}y$ is correlated with the error terms in model (1.4), thus resulting in an endogenous regressor that causes bias and inconsistency in a-spatial OLS estimates. This becomes clear when one considers the following rearrangement of equation (1.4):

$$y = (\boldsymbol{I} - \lambda_1 \boldsymbol{W})^{-1}(\boldsymbol{X}\beta + \varepsilon). \tag{1.5}$$

Expression (1.5) shows how a shock occurring in unit i affects not only the value of y in that unit, but also that of the other units through the inverse spatial transformation (Anselin 2001). The matrix $(\boldsymbol{I} - \lambda_1 \boldsymbol{W})^{-1}$ also determines the parameter space for this model, because it is required to be a non-singular matrix in order to be inverted. When the spatial weight matrix is row-standardized, this is always true for $|\lambda_1| < 1$[1].

1.4.3 SARE models

When spatial dependence is incorporated in the error term, ε becomes non-spherical and the structure of the spatial dependence is expressed by the off-diagonal elements of the covariance matrix. The OLS estimates are therefore unbiased but inefficient. This type of model, the spatial error

[1] Since the diagonal elements of $\boldsymbol{W}$ are equal to 0, the diagonal elements of $(\boldsymbol{I} - \lambda_1 \boldsymbol{W})$ are 1 and, under the condition $|\lambda_1| < 1$, strictly exceed the sum of the other elements in the row, which equals λ_1. This makes the matrix $(\boldsymbol{I} - \lambda_1 \boldsymbol{W})$ strictly diagonally dominant and therefore always invertible.

model, can be specified in different ways. The most common ones incorporate spatial dependence in the error terms by defining them as spatial moving average or spatial autoregressive error (SARE) processes. The latter is probably the most widely used and is specified as

$$\begin{aligned} y &= \boldsymbol{X}\beta + \varepsilon \\ \varepsilon &= \lambda_2 \boldsymbol{W}\varepsilon + \upsilon \end{aligned}, \tag{1.6}$$

where υ is an *i.i.d.* error term and λ_2 is a spatial coefficient that measures spatial dependence between the errors ε.

The reduced form for model (1.6) is expressed as:

$$y = \boldsymbol{X}\beta + (\boldsymbol{I} - \lambda_2 \boldsymbol{W})^{-1}\upsilon \tag{1.7}$$

and requires the matrix $(\boldsymbol{I} - \lambda_2 \boldsymbol{W})$ to be a non-singular matrix. This condition is always verified under the assumption $|\lambda_2| < 1$ when the spatial weight matrix is row-standardized. It follows that $\varepsilon = (\boldsymbol{I} - \lambda_2 \boldsymbol{W})^{-1}\upsilon$ and therefore $E(\varepsilon) = 0$ and $E(\varepsilon'\varepsilon) = \boldsymbol{\Omega}_{(\lambda_2)}$, where $\boldsymbol{\Omega}_{(\lambda_2)}$ depends on the value of λ_2:

$$\boldsymbol{\Omega}_{(\lambda_2)} = \sigma^2[(\boldsymbol{I} - \lambda_2 \boldsymbol{W})'(\boldsymbol{I} - \lambda_2 \boldsymbol{W})]^{-1} \tag{1.8}$$

An important feature of this kind of models regards a possible interpretation of the presence of spatial autocorrelation in the error terms as the effect of relevant spatially autocorrelated omitted variables (Fingleton 1999), which will likely result in biased estimates if not properly modeled. In this perspective the SARE model is capable of capturing the effect of omitted variables which is a common problem for economic modeling.

Model (1.6) can also be rewritten in a way such that a spatial lag of the dependent variable appears, as:

$$y = \lambda_2 \boldsymbol{W}y + \boldsymbol{X}\beta - \lambda_2 \boldsymbol{W}\boldsymbol{X}\beta + \upsilon. \tag{1.9}$$

This is the so-called “Spatial Durbin model” (Anselin 1988a), which imposes some non-linear constraints on the coefficients. The presence of a spatial lag of the dependent variable in this specification complicates the testing procedure for spatial autocorrelation, making it difficult to distinguish between the spatial lag and the spatial error alternatives.

1.4.4 Cross-regressive models

When a spatial lag of the exogenous variable(s) is included into a classical linear regression, a cross-regressive model is specified as

$$y = \boldsymbol{X}\beta + \boldsymbol{WZ}\delta + \varepsilon, \tag{1.10}$$

where $\boldsymbol{Z}$ is an $N \times L$ matrix of exogenous variables which may correspond, totally or partially, to the variables included in $\boldsymbol{X}$ and δ is a row-vector of L spatial parameters. This kind of model is particularly useful for measuring the effects on y of spatial spill-overs of exogenous variables.

For what it concerns the estimation of a cross-regressive model, it must be noticed that, as $\boldsymbol{Z}$ only contains exogenous variables, model (1.10) can be estimated via OLS, as long as assumption OLS_1 holds for the matrix $\boldsymbol{X}^* = [\boldsymbol{X}\ \ \boldsymbol{WZ}]$ and assumptions OLS_2 and OLS_3 hold for the error terms ε. Cross-regressive terms can also be added to previous specifications.

1.4.5 Spatial Cliff-Ord model

The Cliff and Ord type models, also known as SARAR(1,1) in analogy with time series literature, contains both a spatial lag of the dependent variable and of the error term (Kelejian and Prucha 1998):

$$\begin{aligned} y &= \lambda_1 \boldsymbol{W}_1 y + \boldsymbol{X}\beta + \varepsilon, \quad && |\lambda_1| < 1 \\ \varepsilon &= \lambda_2 \boldsymbol{W}_2 \varepsilon + \upsilon, && |\lambda_2| < 1 \end{aligned} \tag{1.11}$$

where $\mathbf{W}_1$ and $\mathbf{W}_2$ may be the same spatial weight matrix or not. In particular, the two must be different from each other as a requirement for identification when applying Maximum Likelihood (ML) estimators[2],

[2] These identification problems that may arise in the ML estimation of this kind of model are such that almost no empirical application exists.

whereas an advantage of Instrumental Variables (IV) / Generalized Method of Moments (GMM) estimators is that the same spatial weight matrix can be used (Elhorst 2010). The model may also contain cross-regressive terms.

1.5 Estimation of spatial cross-sectional models

As it was made clear in sections 1.4.2 and 1.4.3, when spatial autocorrelation is present the OLS assumptions are violated and OLS estimators are biased and inconsistent (when a spatial lag of the dependent variable is included) or at least inefficient (in presence of error spatial autocorrelation).

The most commonly used efficient, unbiased and consistent estimator for cross-sectional spatial models is the maximum likelihood estimator, while other possible choices may be instrumental variables or the generalized method of moments.

1.5.1 Maximum likelihood estimation

The main assumption on which ML estimation of SAR and SARE models relies is that of normality for the error terms. In many circumstances this is a quite a strong one.

The loglikelihood for a SARE model as in equation (1.6) follows from $\varepsilon \sim MVN(0, \boldsymbol{\Omega}_{(\lambda_2)})$ (Anselin 2001):

$$\ln L = -\frac{\mathrm{N}}{2}\ln(2\pi\sigma^2) + ln|\boldsymbol{I} - \lambda_2\boldsymbol{W}| + (\frac{1}{2\sigma^2})\upsilon\,'(\boldsymbol{I} - \lambda_2\boldsymbol{W})'(\boldsymbol{I} - \lambda_2\boldsymbol{W})\varepsilon. \tag{1.12}$$

Conditional upon λ_2,

$$\hat{\beta}_{ML} = [(\boldsymbol{X} - \lambda_2\boldsymbol{WX})'(\boldsymbol{X} - \lambda_2\boldsymbol{WX})]^{-1}(\boldsymbol{X} - \lambda_2\boldsymbol{WX})'(y - \lambda_2\boldsymbol{W}y) \tag{1.13}$$

and

$$\hat{\sigma}^2{}_{ML} = (e - \lambda_2\boldsymbol{W}e)'(e - \lambda_2\boldsymbol{W}e)/N \tag{1.14}$$

with $e = y - \boldsymbol{X}\hat{\beta}_{ML}$. The estimator for λ_2 must be obtained from an explicit maximization of a concentrated likelihood function (Anselin 1988a).

Following Anselin (2001), the loglikelihood for a SAR model is:

$$\ln L = -\frac{N}{2}\ln(2\pi\sigma^2) + ln|\boldsymbol{I} - \lambda_1\boldsymbol{W}| + (\frac{1}{2\sigma^2})\varepsilon'\varepsilon. \tag{1.15}$$

Estimators for the parameters are obtained from an explicit maximization of the likelihood. Conditional upon λ_1,

$$\hat{\beta}_{ML} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(y - \lambda_1\boldsymbol{W}y) \tag{1.16}$$

and

$$\hat{\sigma}^2{}_{ML} = (e_0 - \lambda_1 e_L)'(e_0 - \lambda_1 e_L)/N \tag{1.17}$$

with $\hat{\beta}_0 = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'y$, $e_0 = y - \boldsymbol{X}\hat{\beta}_0$, $\hat{\beta}_L = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}y$ and $e_L = y - \boldsymbol{X}\hat{\beta}_L$.

This results in a concentrated likelihood in a single parameter that is optimized by means of direct search techniques.

The classical properties of consistency, asymptotic normality and asymptotic efficiency of ML estimators do not straightforwardly hold when spatial dependence is present, even in the case of normally distributed error terms (Kelejian and Prucha 1999).

The forms taken by the loglikelihood in equations (1.12) and (1.15) also define the parameter space for λ_1 and λ_2. The main problem raised by the estimation of spatial models via ML concerns the presence of the Jacobian matrix in the loglikelihood function, that is equal to $|\boldsymbol{I} - \lambda_1\boldsymbol{W}|$ in SAR models and to $|\boldsymbol{I} - \lambda_2\boldsymbol{W}|$ in SARE models. The maximization of the function of loglikelihood requires the evaluation of the determinant of the Jacobian matrix for each value of λ_1 or λ_2 and, since in practice spatial weight matrices are not symmetric (because $\boldsymbol{W}$ is commonly row-standardized), the procedure may be computationally very complex with very large datasets. In order to avoid singularity, the parameter space is generally restricted to the interval $(-1, 1)$. However, it is important to notice that, given a generic spatial autoregressive coefficient λ (which may be either λ_1 for a SAR process or λ_2 for a SARE process), its parameter space is defined as

$$1/\omega_{min} < \lambda < 1/\omega_{max}, \tag{1.18}$$

where ω_{min} and ω_{max} are respectively the smallest and largest eigenvalues of the spatial weight matrix. When $\boldsymbol{W}$ is row-standardized, $\omega_{max} = 1$ and $\omega_{min} > -1$ (Anselin 1988a; 2001; Elhorst 2010).

1.5.2 Other estimators (IV, GMM)

A possible estimation strategy, alternative to ML, is that of IV which is suitable for addressing the endogeneity of $\boldsymbol{W}y$ in a SAR model of type (1.4) (Anselin 1988a). The general principle behind this approach is based on the existence of a set of M instruments $\boldsymbol{Q}$ (with $M \geq K + 1$), which are correlated to the regressors of the SAR model $\boldsymbol{X}^* = [\mathbf{W}\text{y } \mathbf{X}]$ but asymptotically uncorrelated with the error term. When $M > K + 1$ no exact solution exists. This problem is addressed by defining the estimator for the model coefficients $\theta = [\lambda_1 \;\; \beta]$ as:

$$\hat{\theta}_{IV} = \left(\boldsymbol{X}^{*\prime}\boldsymbol{P_Q}\boldsymbol{X}^*\right)^{-1}\boldsymbol{X}^{*\prime}\boldsymbol{P_Q}.y \tag{1.19}$$

with $\boldsymbol{P_Q} = \boldsymbol{Q}(\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'$ a symmetric idempotent matrix.

Under a set of assumptions discussed by Kelejian and Robinson (1993) and Kelejian and Prucha (1998), the spatial two-stage least squares (2SLS) estimator can be proved to achieve consistency and asymptotic normality as the standard 2SLS.

When it comes to selecting the set of instruments $\mathbf{Q}$, the exogenous regressors should always be included. Proper instruments for the spatial lag of the dependent variable are the spatial lags of the exogenous regressors, **WX** (Kelejian and Prucha 1998). This procedure can be easily extended to more complex models (Anselin 1988a), but it is not suitable for estimating

SARE models (Anselin 2001), unless properly generalized as in Kelejian and Prucha (1998).

Another possible approach that is suitable for estimating a SARE model (specified as model (1.6)) is the GMM estimator presented in Kelejian and Prucha (1999). Under the assumptions concerning the error term ε in model (1.6), the authors specify three moment conditions on which the GMM estimator is based:

$$E\left[\frac{1}{N}\upsilon'\upsilon\right] = \sigma^2;\; E\left[\frac{1}{N}\upsilon'\boldsymbol{W}'\boldsymbol{W}\upsilon\right] = \sigma^2 N^{-1} tr(\boldsymbol{W}'\boldsymbol{W});\; E\left[\frac{1}{N}\upsilon'\boldsymbol{W}\upsilon\right] = 0 \quad (1.20)$$

By replacing υ with $\varepsilon - \lambda_2\boldsymbol{W}\upsilon$ and considering the sample analogue of ε (the vector of residuals after a consistent estimation, usually obtained by OLS: e), a three-equation system is given for parameters λ_2, ${\lambda_2}^2$ and σ^2, implied by equations (1.6) and (1.20). Given $\bar{\varepsilon} = \boldsymbol{W}\varepsilon$ and $\bar{\bar{\varepsilon}} = \boldsymbol{W}\bar{\varepsilon}$ (and consequently $\bar{e} = \boldsymbol{W}e$ and $\bar{\bar{e}} = \boldsymbol{W}\bar{e}$ as sample analogues), consider

$$\Gamma\left[\lambda_2,\ {\lambda_2}^2, \sigma^2\right]' - \gamma = 0 \quad (1.21)$$

where $\Gamma = \begin{bmatrix} \frac{2}{N}E(\varepsilon'\bar{\varepsilon}) & \frac{-1}{N}E(\bar{\varepsilon}'\bar{\varepsilon}) & 1 \\ \frac{2}{N}E(\bar{\bar{\varepsilon}}'\bar{\varepsilon}) & \frac{-1}{N}E(\bar{\bar{\varepsilon}}'\bar{\bar{\varepsilon}}) & \frac{1}{N}tr(\boldsymbol{W}'\boldsymbol{W}) \\ \frac{1}{N}E(\varepsilon'\bar{\bar{\varepsilon}} + \bar{\varepsilon}'\bar{\varepsilon}) & \frac{-1}{N}E(\bar{\varepsilon}'\bar{\bar{\varepsilon}}) & 0 \end{bmatrix}$ and $\gamma = \begin{bmatrix} \frac{1}{N}E(\varepsilon'\varepsilon) \\ \frac{1}{N}E(\bar{\varepsilon}'\bar{\varepsilon}) \\ \frac{1}{N}E(\varepsilon'\bar{\varepsilon}) \end{bmatrix}$.

Its sample analogue is:

$$G\left[\lambda_2,\ {\lambda_2}^2, \sigma^2\right]' - g = \nu(\lambda_2, \sigma^2) \quad (1.22)$$

where $G = \begin{bmatrix} \frac{2}{N}E(e'\bar{e}) & \frac{-1}{N}E(\bar{e}'\bar{e}) & 1 \\ \frac{2}{N}E(\bar{\bar{e}}'\bar{e}) & \frac{-1}{N}E(\bar{\bar{e}}'\bar{\bar{e}}) & \frac{1}{N}tr(\boldsymbol{W}'\boldsymbol{W}) \\ \frac{1}{N}E(e'\bar{\bar{e}} + \bar{e}'\bar{e}) & \frac{-1}{N}E(\bar{e}'\bar{\bar{e}}) & 0 \end{bmatrix}$ and $g = \begin{bmatrix} \frac{1}{N}E(e'e) \\ \frac{1}{N}E(\bar{e}'\bar{e}) \\ \frac{1}{N}E(e'\bar{e}) \end{bmatrix}$.

The GMM estimator for λ_2 and σ^2 is then defined as the nonlinear least squares estimator corresponding to equation (1.22). According to this approach, λ_2 is considered a nuisance parameter whose significance does not need to be tested. Once $\hat{\lambda}_2$ and $\hat{\sigma}^2$ are obtained, the vector of parameters β can be estimated by feasible generalized least squares (FGLS), as:

$$\hat{\beta}_{FGLS} = \left[\boldsymbol{X}'\boldsymbol{\Omega}_{(\hat{\lambda}_2)}{}^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}'\boldsymbol{\Omega}_{(\hat{\lambda}_2)}{}^{-1}y, \tag{1.23}$$

where $\boldsymbol{\Omega}_{(\hat{\lambda}_2)}$ is the covariance matrix of the error ε corresponding to the GMM estimate of λ_2 and σ^2.

Kelejian and Prucha (1998) also propose a combination of 2SLS and GMM that allows obtaining unbiased estimates for the parameter of a Cliff and Ord type of model (as defined in equation 1.11).

An advantage of IV/GMM estimator is that, differently from the ML estimators, they do not rely on the assumption of normality of the errors, although they do assume, just like ML estimators, that the errors are *i.i.d.* with 0 mean. However, a disadvantage of IV/GMM estimator is that it is possible to obtain an estimate of the spatial parameter which is outside its parameter space $(1/\omega_{min}, 1/\omega_{max})$, since these estimators ignore the Jacobian term which restricts λ_1 (or λ_2) to its parameters space in the log-likelihood function of ML estimators (Elhorst 2010).

1.6 Testing for the presence of spatial effects

Testing for the presence of residual spatial effects is of utmost importance as statistical inference based on OLS estimation may not be reliable when spatial dependence or heterogeneity is present (Ertur et al. 2006).

1.6.1 Moran's I test of spatial autocorrelation

Among the tests for the detection of spatial autocorrelation, the one based on the computation of Moran's I statistics (Moran 1950) is the most common:

$$I = \left(\frac{n}{s}\right) y'\boldsymbol{W}y(y'y)^{-1} \tag{1.24}$$

where S is the sum of all elements of the spatial weight matrix $\boldsymbol{W}$ and y is the vector of the n observations for the considered variable. The null hypothesis is the absence of spatial dependence, but the nature of the underlying spatial stochastic process is not specified under the alternative hypothesis. Inference is based on a normal approximation, using the standardized value $z_I = \{I - E[I]\}/\{\sqrt[2]{V[I]}\}$, which is obtained from expressions for the mean and variance (Cliff and Ord, 1981).

When testing for residual spatial autocorrelation, y is substituted by the vector of residuals of the OLS-estimated model ($e = y - X\hat{\beta}$) (Cliff and Ord 1972), under the assumption of independent and identical normal distribution of the errors:

$$I = \left(\frac{n}{S}\right) e'\boldsymbol{W}e(e'e)^{-1} \tag{1.25}$$

In this case, the expressions for the moments become more complicated (see Anselin 1988a, page 102).

1.6.2 The Lagrange Multiplier tests

When spatial regression models are estimated via Maximum Likelihood, the Lagrange Multiplier tests are particularly useful when searching for the best specification of the model, because they give insight on the form of spatial autocorrelation that should be considered in the model and they do not require the estimation of a spatial model for testing purposes. Different simple null hypothesis are tested through different test specifications (Anselin 2001):

Lagrange Multiplier test for error spatial autocorrelation (LM_{err})

The LM_{err} test, first introduced by Burridge (1980) and then extensively treated in Anselin (2001) and Anselin and Florax (1995), tests the null hypothesis $H_0\!: \lambda_2 = 0$, therefore testing a non-spatial specification of a linear regression model against a SARE model specification. The LM_{err} test is written as

$$LM_{err} = (e'\boldsymbol{W}e/\hat{\sigma}^2)^2/T \tag{1.26}$$

where e is a vector of OLS residuals, $\hat{\sigma}^2$ is an estimate of σ^2 under the null hypothesis equal to $e'e/N$, T= $tr[\boldsymbol{W'W} + \boldsymbol{W^2}]$. Under the null hypothesis, $LM_{err} \rightarrow \chi_1^2$. When the null hypothesis is rejected, error spatial correlation should be included in the model. However, no clear indication is given about whether it should be specified as a spatial autoregressive process ($\varepsilon = \lambda_2 \boldsymbol{W\varepsilon} + \upsilon$) or as a moving average process ($\varepsilon = \lambda_2 \boldsymbol{W}\varepsilon + \upsilon$), the test being the same for both cases.

Lagrange Multiplier test for spatial autocorrelation of the dependent variable (LM_{lag})

The LM_{lag} test for $H_0\colon \lambda_1 = 0$ against a spatial lag alternative takes the form

$$LM_{lag} = (e'\boldsymbol{W}y/\hat{\sigma}^2)^2/[(\boldsymbol{R}/\hat{\sigma}^2) + T] \tag{1.27}$$

where **R**=$(\boldsymbol{WX}\beta)'(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X}')(\boldsymbol{WX}\beta)$. This test is also asymptotically distributed as a χ^2 with 1 degree of freedom (Anselin 1988b; Anselin and Florax 1995).

When performing the LM_{err} or the LM_{lag} test it is important to account for possible spatial dependence of the other form (i.e. spatial lag dependence when testing for spatial error and vice versa), by means of either a joint test, which takes a rather complicated specification (Anselin 1988b), or tests that are robust to the presence of local misspecification of the other form (Anselin et al. 1996).

LM joint test

The LM joint test for spatial lag and spatial moving average error (*SARMA*) allows testing the joint null hypothesis $H_0\colon \lambda_1 = \lambda_2 = 0$, by taking the following form in the simplified case in which $\boldsymbol{W}$ is the same for both the spatial (moving average) error and spatial lag autoregressive processes (Anselin 1988b):

$$SARMA = (e'\boldsymbol{W}y/\hat{\sigma}^2 - e'\boldsymbol{W}e/\hat{\sigma}^2)^2/\boldsymbol{R} + (e'\boldsymbol{W}e/\hat{\sigma}^2)^2/T \qquad (1.28)$$

Under the null hypothesis, SARMA $\rightarrow \chi_2^2$. The rejection of the null hypothesis, however, does not give clear evidence about the nature of the spatial dependence which is detected. The SARMA test is identical in its formula to a test for a joint spatial lag and spatial autoregressive error, except for the fact that such a SARAR process with identical spatial weight matrices for the spatial error and the spatial lag is not identified, whereas a SARMA process is (Anselin and Florax 1995).

Robust LM tests

The robust versions of the LM_{err} and the LM_{lag} tests are adjusted to be robust to local misspecifications. The RLM_{err} test is adjusted so as to maintain a χ^2 asymptotic distribution even when $\lambda_1 \neq 0$; similarly, the RLM_{lag} allows to test $H_0: \lambda_1 = 0$ even in presence of $\lambda_2 \neq 0$. The complete specifications of these two tests can be found in Anselin et al. (1996).

1.6.3 The choice of the correct model specification

Once the presence of spatial dependence is detected by the Moran's I test and/or the SARMA test, it is possible to try to reduce it by including additional exogenous variables and/or their spatial lags in the model specification. If spatial autocorrelation is still present, proceeding from the results of the LM tests, this can be accounted for by means of a spatial model. The correct specification can be chosen following the criteria indicated by Anselin and Florax (1995) and Anselin (2005):

- If the LM_{err} does not reject the null hypothesis and the LM_{lag} rejects the null hypothesis, then a spatial lag of the dependent variable should be included;
- If the LM_{lag} does not reject the null hypothesis and the LM_{err} rejects the null hypothesis or both tests reject the null hypothesis, but the LM_{err} is more significant than the LM_{lag} test, then a spatial error model should be preferred;

If both non-robust LM tests reject the null hypothesis, their robust versions should be considered:

- If RLM_{lag} rejects the null hypothesis whereas RLM_{err} does not or the RLM_{lag} is more significant than the RLM_{err} test, then the spatial lag of the dependent variable should be included;
- If RLM_{err} rejects the null hypothesis whereas RLM_{lag} does not or the RLM_{err} is more significant than the RLM_{lag} test, then the spatial lag of the error term should be included.

Once a spatial model is estimated, additional conditional LM tests can be performed in order to exclude the need to include additional spatial autoregressive terms, such as a spatial error term in a spatial lag model or vice versa. If uncertainty about the best spatial specification persists, the choice can be done according the classical information criteria (AIC, BIC).

1.7 Spatial panel data models

The interest of spatial econometrics literature in the estimation of panel data models has been growing in recent years.

The main reason for this, together with the growing availability of datasets at micro and macro level, is that using panel data yields a number of benefits such as greater variability and less collinearity among the variables, more degrees of freedom and efficiency; the possibility of studying the dynamic and the individual heterogeneity at the same time; the ability to identify and measure effects that cross-section and time-series data cannot detect; a better control over unobservable/unobserved heterogeneity (individual or time-invariant characteristics) (Baltagi 2005).

1.7.1 A taxonomy for spatial panel data models

Similarly to cross-section models, panel data models can be specified as spatial models by controlling for spatial effects. Spatial dependence is the most problematic spatial effect to model, since most econometric aspects of spatial heterogeneity can be handled by means of non-spatial panel data.

As a non-purely notational purpose, it must be noted that spatial panel data models are stacked by cross-sections, rather than individual time series. This means that observations are sorted first by time and then by cross-sectional units, so that, for example, the $NTx1$ vector y is organized as $y' = [y_{11} \quad y_{21} \quad \cdots \quad y_{N1} \quad y_{1t} \quad \cdots \quad y_{Nt} \quad \cdots \quad y_{1T} \quad \cdots \quad y_{NT}]$, with $t = 1, 2, \ldots, T$ indexing time periods. Cross-sectional units will be indexed by the index i, such that for fixed i and t, element y_{it} will be the observation of variable Y in the i-th unit at time t.

1.7.2 Static models

Following Anselin et al. (2008), spatial dependence is generally considered as a cross-sectional non-zero correlation among different units according to certain spatial ordering, so that error autocorrelation only pertains to the same time period t. Spatial dependence is modeled by means of a spatial weight matrix that is assumed to be constant over time, so that the full $NT \times NT$ spatial weight matrix (W_{NT}) is defined as

$$\boldsymbol{W}_{NT} = \boldsymbol{I}_T \otimes \boldsymbol{W}_N = \begin{bmatrix} \boldsymbol{W}_N & \cdots & \boldsymbol{0}_N \\ \vdots & \ddots & \vdots \\ \boldsymbol{0}_N & \cdots & \boldsymbol{W}_N \end{bmatrix}_{NT} \tag{1.29}$$

where $\boldsymbol{I}_T$ is a $T \times T$ identity matrix, $\boldsymbol{W}_N$ is a $N \times N$ spatial weight matrix defined as in section 1.3.2 and $\mathbf{0}_N$ is a $N \times N$ zero matrix. The particular structure of data, stacked by cross-sections, permits to build the spatial lag of a variable by multiplying its observations for each time period by the spatial weight matrix.

SAR model

Similarly to the cross-sectional case, the basic spatial lag model specification for panel data is

$$y = \lambda_1 \boldsymbol{W}_{NT} y + \boldsymbol{X}\beta + \varepsilon \tag{1.30}$$

where the observations are stacked in successive cross-sections for $t = 1, \ldots, T$. Therefore y is a $NT \times 1$ vector, λ_1 is a spatial autoregressive parameter, $\boldsymbol{X}$ is a $NT \times K$ matrix with K equal the number of regressors, β is a $K \times 1$ vector and ε is a $NT \times 1$ vector of $i.i.d.$ errors.

Again similarly to what happens in cross-sectional models, $\boldsymbol{W}_{NT} y$ appears to be endogenous as the result of the joint determination of the values of y in the spatial system as a function of the explanatory variables and the error terms at all locations in the system. The reduced form of equation (1.30) makes it clear for each $N \times 1$ cross-section at time t:

$$\begin{aligned} y_t = {} & \boldsymbol{X}_t\, \beta + \lambda_1 \boldsymbol{W}_N\, \boldsymbol{X}_t\, \beta + \lambda_1^2 \boldsymbol{W}_N^{\,2}\, \boldsymbol{X}_t\, \beta + \cdots + \varepsilon_t + \lambda_1 \boldsymbol{W}_N\, \varepsilon_t + \\ & + \lambda_1^2 \boldsymbol{W}_N^{\,2}\, \varepsilon_t + \ \ldots \end{aligned} \tag{1.31}$$

or equivalently,

$$y_t = (\boldsymbol{I}_N - \lambda_1 \boldsymbol{W}_N)^{-1} (\boldsymbol{X}\beta + \varepsilon). \tag{1.32}$$

In this simple pooled model, the spatial multiplier effect is only limited to each cross-section.

SARE model

The spatial error specification is characterized by a non-spherical error covariance matrix. In a panel data setting with $N < T$, an unconstrained non-spherical error covariance matrix contains $N \times (N-1)/2$ parameters. In order to estimate them when $N \gg T$, a structure must be imposed in order to turn the complex structure of the error covariance matrix into a function of a set of parameters. Four main approaches have been suggested in the literature and reviewed by Anselin et al. (2008):

- The *direct representation* approach is rooted in the geostatistical field (Cressie 1993) and is based on the specification of the covariance between two observations as a direct function of the distance that separates them. Given the $N \times N$ time-invariant error covariance structure $\boldsymbol{\Omega}_N$ for each cross-section and σ^2 a scalar variance term, the overall matrix can be defined as $\boldsymbol{\Sigma}_{NT} = \sigma^2[\boldsymbol{I}_T \otimes \boldsymbol{\Omega}_N]$.
- *Spatial error processes* are based on a formal relation between a location and its neighbors (not between all pairs of observations), through a spatial weight matrix. As it was already made clear, analogously to time-series analysis, the most common models for spatial processes are the autoregressive and the moving average specifications. In a panel data setting, a SARE process is specified as:

$$\begin{aligned} y &= \boldsymbol{X}\beta + \varepsilon \\ \varepsilon &= \lambda_2 \boldsymbol{W}_{NT}\varepsilon + \upsilon \end{aligned} \tag{1.33}$$

  where υ is a $NT \times 1$ vector of *i.i.d.* errors. The full error covariance matrix, again assumed not to vary over time, and then is equal to:

$$\boldsymbol{\Sigma}_{NT} = \sigma_\upsilon^2[\boldsymbol{I}_T \otimes [(\boldsymbol{I}_N - \lambda_2 \boldsymbol{W}_N)'(\boldsymbol{I}_N - \lambda_2 \boldsymbol{W}_N)]^{-1}] \,. \tag{1.34}$$

- The *spatial error components* specification, proposed by Kelejian and Robinson (1995), decomposes the error term into a local and a spillover effect, which are assumed to be uncorrelated. Each component is assumed to be *i.i.d.*, with a specific variance. The time-invariant overall error covariance matrix results from the sum of the covariance matrices of the two components.

- The standard two-way error component model (Baltagi 2005) specifies the error terms of the regression model as the sum of an unobserved individual component, a time-specific component and an idiosyncratic error term. The common time component results in a particular form of cross-sectional (hence spatial) correlation. This kind of model specification was recently extended into the *common factor* model approach by expressing the time component as an unobserved time-specific common factor to which all cross-sectional units are exposed (each of them having a distinct factor loading on this common factor).

Therefore, spatial dependence is typically taken into account by including the spatial lag of the dependent variable or a spatial autoregressive error term into the model (Anselin et al. 2008; Elhorst 2010), such as in the cross-sectional framework. As described in section 1.5.1, in both cases, stationarity requires that the value of the spatial parameter is included between the smallest and largest eigenvalues of the spatial weight matrix (Elhorst 2010).

1.7.3 Temporal and spatial heterogeneity

Still following Anselin et al. (2008), the homogeneous specifications of spatial panel data model outlined so far can be extended so as to introduce heterogeneity both over time and across space. Many different specifications can be theoretically introduced. However, most of them suffer from identification problems and do not find empirical application; therefore they are not described in what follows.

Temporal heterogeneity can be introduced straightforwardly in spatial panel data models by allowing for time-specific parameters. The cross-sectional error terms can be allowed to be correlated over time periods in what is called a *spatial Seemingly Unrelated Regression* (SUR) model. In a spatial lag model, for each cross-section at time $t = 1, 2, \dots, T$, the standard specification is enriched by a time-specific spatial autoregressive coefficient (λ_{1t}), thus becoming:

$$y_t = \lambda_{1t}\boldsymbol{W}_N y_t + \boldsymbol{X}_t \beta_t + \varepsilon_t. \tag{1.35}$$

The error covariance matrix for model (1.35) is $E[\varepsilon\varepsilon'] = \boldsymbol{\Sigma}_T \otimes \boldsymbol{I}_N$, with $\boldsymbol{\Sigma}_{\mathrm{T}}$ as a $T \times T$ temporal covariance matrix with elements σ_{ts} (equal to the temporal covariance between period s and t, for $s \neq t$). Spatial error autocorrelation can also be introduced in spatial SUR models. If this is done by means of a spatial autoregressive error process, for each cross-section $N \times 1$ at time $t = 1, 2, \dots, T$ the model is specified as

$$\begin{aligned} y_t &= \boldsymbol{X}_t \beta + \varepsilon_t \\ \varepsilon_t &= \lambda_{2t}\boldsymbol{W}_N \varepsilon_t + \upsilon_t \end{aligned}. \tag{1.36}$$

Spatial heterogeneity is usually unobserved heterogeneity that is included in the model specification either as fixed effects or random effects. In a spatial econometric framework, both these kinds of models can be extended to the SAR and SARE specification.

The classic specification of a *fixed effects* model captures unobserved heterogeneity through an individual specific, time-constant term (c_i) that is not assumed to be orthogonal to non-stochastic regressors $\boldsymbol{X}$: $y = \boldsymbol{X}\beta + \boldsymbol{D}_c c + \varepsilon$, where $\boldsymbol{D}_c$ is a $NT \times N$ matrix of individual dummies and c is a $N \times 1$ vector of fixed parameters (individual effects). The disturbances are assumed to be $i.i.d.(0, \sigma_\varepsilon^2)$. The spatial lag extension of this approach is not straightforward. A fixed effects spatial lag model in stacked form is specified as:

$$y = \lambda_1(\boldsymbol{I}_T \otimes \boldsymbol{W}_N)y + \boldsymbol{X}\beta + (\iota_T \otimes c) + \varepsilon. \tag{1.37}$$

The estimation of model (1.37) requires the use of a demeaned form, in order to overcome the incidental parameter problem. The demeaned form is obtained by subtracting the average over the time dimension for each cross-sectional unit, thus wiping out the fixed effects and the constant term.

However, the singularity of the demeaning operator is still a problem for ML estimation (Anselin et al. 2008).

The standard specification of the error term for each cross-section in a *one-way error component* model is

$$\varepsilon_t = c + \nu_t \ , \tag{1.38}$$

where c is a $N \times 1$ vector of individual random components with $c_i \sim i.i.d.(0, \sigma_c^2)$ and $\nu_{it} \sim i.i.d.(0, \sigma_\nu^2)$. Spatial error autocorrelation can be incorporated into the standard random effects model in some different ways, which are described in detail in Anselin et al. (2008). One possible approach is the specification of a SAR process for the idiosyncratic error component of equation (1.38), as Baltagi et al. (2003) do, so that for each cross-section,

$$\nu_t = \lambda_2 \boldsymbol{W}_N \nu_t + \upsilon_t \ . \tag{1.39}$$

A second possible specification is the one adopted in Kapoor et al. (2007), that first applies a SAR process to the error term ε and then specifies the vector of innovations υ as an error component model:

$$\begin{aligned} \varepsilon_t &= \lambda_2 \boldsymbol{W}_N \varepsilon_t + \upsilon_t \\ \upsilon_t &= c + \nu_t \end{aligned} \ . \tag{1.40}$$

1.7.4 Dynamic models

Dynamic panel data models incorporate dependence both in time and space. Dynamics in time is embodied in the model through the inclusion of a lagged dependent variable ($y_{i,t-1}$) among the regressors (Baltagi 2005), whereas spatial dynamics can be included in the usual ways in a SAR or SARE framework.

Anselin (2001) distinguishes spatial dynamic models into some broad categories. Space-time dependence in the error term is ignored at first and the focus is on models where the cross-sectional dimension is bigger than the time-dimension ($N \gg T$). The taxonomy provided for spatial dynamic models with lag dependence is the following (for ease of exposition the models are expressed as a $N \times 1$ cross-section at time $t = 1, \dots, T$).

Pure space-recursive models:

$$y_t = \varrho \boldsymbol{W}_N y_{t-1} + \boldsymbol{X}_t \beta + \varepsilon_t, \tag{1.41}$$

where ϱ is the space-time autoregressive parameter and $\boldsymbol{W}_N y_{t-1}$ is the $N \times 1$ vector of observations of the spatially lagged dependent variable at time $t - 1$. The model can be easily extended so as to include the time or spatial lag of the explanatory variables, although their space-time lag should not be included in order to avoid identification problems (Anselin et al. 2008). In this kind of model, the dependence only pertains to neighboring units in a previous period. This means that it takes one period for spatial dependence to manifest itself.

Time-space recursive models:

$$y_t = \gamma y_{t-1} + \varrho \boldsymbol{W}_N y_{t-1} + \boldsymbol{X}_t \beta + \varepsilon_t, \tag{1.42}$$

where γ is a serial (i.e. time) autoregressive parameter. Spatially lagged contemporaneous explanatory variables ($\boldsymbol{W}_N \boldsymbol{X}_t$) can also be included, but no time or space-time lags of the vector $\boldsymbol{X}_t$ should be added because of identification problems.

Time-space simultaneous models:

$$y_t = \gamma y_{t-1} + \lambda_1 \boldsymbol{W}_N y_t + \boldsymbol{X}_t \beta + \varepsilon_t, \tag{1.43}$$

where, in accordance with previous notation, λ_1 is the spatial autoregressive parameter. In this model, the inclusion of any kind of spatial lag of the explanatory variables would be problematic, because the effect of it is already included in the combined effect of the spatial multiplier and the space-time multiplier that follows from the presence of y_{t-1} among the regressors (Anselin et al. 2008).

Time-space dynamic models include all three possible lags of the dependent variable:

$$y_t = \gamma y_{t-1} + \lambda_1 \boldsymbol{W}_N y_t + \varrho \boldsymbol{W}_N y_{t-1} + \boldsymbol{X}_t \beta + \varepsilon_t. \tag{1.44}$$

This is the more general specification, but its estimation may be complicated because of identification problems. Nevertheless, this model has been extensively studied in recent literature (Yu et al. 2008; Lee and Yu 2010a).

Space-time dependence can also be included in the error terms following what is called the "error component approach" (Anselin et al. 2008). The starting point is the spatial random effects model as specified in equations (1.38) and (1.39), in which the idiosyncratic component u_t is substituted by a serially correlated term (ξ_t):

$$\begin{aligned} \varepsilon_t &= c + \nu_t \\ \nu_t &= \lambda_2 \boldsymbol{W}_N \nu_t + \xi_t \\ \xi_t &= \phi \xi_{t-1} + u_t \end{aligned} \quad . \tag{1.45}$$

1.8 Estimation of spatial panel data models

The estimation of spatial panel data models needs to deal with the problems caused by autocorrelation in space, already described in section 1.2.1: although the panel data framework appears to be more complex than the cross-sectional one, the basic reasoning is analogous. When considering a spatial lag model, the simultaneity between $\boldsymbol{W}_{NT} y$ and ε must be taken into account.

When panel data models incorporate dependence both in time and space, such as spatial dynamic panel data models do, it is also convenient to focus briefly on time dynamics as a second source of autocorrelation. If individual (spatial) heterogeneity is present in the model as a one-way error component (also called random effects) and since $y_{i,t}$ is a function of the individual-specific term c_i, it follows that that $y_{i,t-1}$ is also a function of c_i. Therefore one of the regressors is correlated with the error term and the OLS estimator is biased and inconsistent and so are the fixed effects estimator and the GLS random effects estimator (Baltagi 2005).

The main ways that have been proposed in order to deal with these two sources of autocorrelation are instrumentation-based IV/GMM estimation procedures and ML estimation, which specifies a complete distributional model. Nevertheless, it should be noted that different model specifications have given birth to different estimation strategies in the literature, each one dealing with the peculiar econometric problems that the estimation of that model presents.

Recent literature has developed theoretical properties for spatial panel data models estimators. Kapoor et al. (2007) contributed to the GMM approach deriving a GMM estimator for a spatially autocorrelated error static panel data model with individual effects.

Quasi-maximum likelihood (QML) estimators were also proposed: Yu et al. (2008) studied the asymptotic behavior of a QML estimator for a dynamic spatial autoregressive panel data model with only individual fixed effects when both N and T are large; this was later extended to two-way error component models, where both time and individual effects are present (Lee and Yu 2010a). A Least-Squares Dummy Variable (LSDV) estimator for a “time-space recursive” model with fixed individual effects which also allows for endogenous regressors was proposed by Korniotis (2010). Some of the main estimators recently proposed in the literature are reviewed in the following sections[3].

[3] The notation used in the following sections might slightly differ from the one in the original reviewed papers. This is done for the sake of consistency.

1.8.1 The KKP estimator for a SARE static panel data model

A fundamental contribution to the literature on the estimation of spatial panel data model is that by Kapoor et al. (2007), who introduced the so-called "*KKP*" (Kapoor, Kelejian and Prucha's) estimator. The authors consider a static panel data model that allows for the disturbances to be correlated both over time and across space. Spatial dependence is modeled as a first order spatial autoregressive process in the error term and correlation over time is obtained through the $N \times 1$ vector of individual effects c, as in equation (1.40). Stacking the observations, the model can written as

$$\begin{aligned} y &= \boldsymbol{X}\beta + \varepsilon \\ \varepsilon &= \lambda_2(\boldsymbol{I}_T \otimes \boldsymbol{W}_N)\varepsilon + \upsilon. \\ \upsilon &= (\iota_T \otimes \boldsymbol{I}_N)c + \nu \end{aligned} \tag{1.46}$$

In this specification, υ corresponds to a classical one-way error component (Baltagi 2005); ν is an $i.i.d.$ error term with zero mean, variance σ_ν^2 and finite fourth moments; the unit specific error components c are also $i.i.d.$ with zero mean, variance σ_c^2 and finite fourth moments. The processes ν and c are independently distributed. The spatial weight matrix is defined as usual with null diagonal elements. Moreover, in order for the matrix $(\boldsymbol{I}_N - \lambda_2 \boldsymbol{W}_N)$ to be non-singular, it is also assumed that $|\lambda_2| < 1$.

This model specification implies that the innovations υ are autocorrelated over time but not across spatial units. Their variance-covariance matrix is defined as:

$$\boldsymbol{\Omega}_\upsilon = E(\upsilon\upsilon') = \sigma_c^2(\boldsymbol{J}_T \otimes \boldsymbol{I}_N) + \sigma_\nu^2 \boldsymbol{I}_{NT}, \tag{1.47}$$

Differently, the model disturbances ε are correlated both over time and across space and are such that $E(\varepsilon) = 0$ and

$$\boldsymbol{\Omega}_\varepsilon = E(\varepsilon\varepsilon') = [\boldsymbol{I}_T \otimes (\boldsymbol{I}_N - \lambda_2 \boldsymbol{W}_N)^{-1}]\boldsymbol{\Omega}_\upsilon[\boldsymbol{I}_T \otimes (\boldsymbol{I}_N - \lambda_2 \boldsymbol{W'}_N)^{-1}]. \tag{1.48}$$

The *KKP* procedure is defined for the case in which T is fixed and $N \to \infty$. Three GMM estimators are proposed for λ_2, σ_c^2 and σ_ν^2, in terms of six moment conditions that generalize the moment conditions introduced in Kelejian and Prucha (1998, 1999). The first set of estimators provides initial estimates for λ_2, and σ_ν^2 as the unweighted nonlinear least squares estimators based on a subset of the moment conditions and the residuals calculated after the OLS estimation of model $y = \boldsymbol{X}\beta + \varepsilon$. The estimates obtained ($\tilde{\lambda}_2$, and $\tilde{\sigma}_\nu^2$) are then used in order to provide an estimate for $\sigma_1^2 = \sigma_\nu^2 + T\sigma_c^2$ based on the fourth moment condition.

Under the assumption of normality of innovations υ, the authors derive the variance-covariance matrix of the sample moments at the true parameter values ($\boldsymbol{\Xi}$, consistently estimated by $\tilde{\boldsymbol{\Xi}}$), whose inverse is to be used as the optimal weighting matrix in a GMM estimator. Therefore the use of the weighting matrix proposed for this procedure will not be strictly optimal when the normality assumption for υ does not hold.

The second GMM estimator is then defined as the nonlinear least squares estimator based on the moment conditions weighted by $\tilde{\boldsymbol{\Xi}}^{-1}$. The third GMM estimator is proposed mainly because of computational considerations and is based on a simpler weighting matrix, which places the same weight on each of the first three moment conditions and the same – but different from the previous – weight on each of the last three moment conditions. This partially weighted GMM estimator is also proved to be a consistent estimator for λ_2, σ_c^2 and σ_ν^2.

Finally, the authors provide a Feasible Generalized Least Squares (FGLS) estimator for β based on the estimates obtained for λ_2, σ_c^2 and σ_ν^2, which is proved to be consistent, asymptotically normal and to have the same asymptotic distribution as the real GLS estimator.

1.8.2 A Quasi-Maximum likelihood estimator for a time-space dynamic model

Yu et al. (2008) investigate the asymptotic properties of a QML estimator for a *time-space dynamic* model with fixed individual effects when both the cross sectional dimension (N) and the time dimension (T) go to infinity (either at a proportional rate or not) and also propose a bias correction. The model considered is the most general model proposed in Anselin's taxonomy (2001), therefore the asymptotic results presented in this paper are also applicable to the other categories of models as special cases in which some of the parameters are equal to zero.

For each time period $t = 1,2, \dots, T$, the model is

$$y_t = \lambda \boldsymbol{W}_N y_t + \gamma y_{t-1} + \varrho \boldsymbol{W}_N y_{t-1} + \boldsymbol{X}_t \beta + c + \varepsilon_t, \tag{1.49}$$

where $y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{nt})'$ are $N \times 1$ vectors and ε_{it} is $i.i.d.$ across i and t with zero mean and variance σ^2; $\boldsymbol{W}_N$ is the $N \times N$ spatial weight matrix, $\boldsymbol{X}_t$ is an $N \times K$ matrix of non-stochastic regressors and c is the $N \times 1$ vector of fixed individual effects. The total number of parameters in this model is therefore equal to $N + K + 4$.

Following Yu et al. (2008), define $\boldsymbol{S}_N = (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)$ and, assuming that $\boldsymbol{S}_N$ is invertible, $\boldsymbol{A}_N = {\boldsymbol{S}_N}^{-1}(\gamma \boldsymbol{I}_N - \varrho \boldsymbol{W}_N)$. Model (1.49) can be rewritten as

$$y_t = \boldsymbol{A}_N y_{t-1} + {\boldsymbol{S}_N}^{-1} \boldsymbol{X}_t \beta + {\boldsymbol{S}_N}^{-1} c + {\boldsymbol{S}_N}^{-1} \varepsilon_t \tag{1.50}$$

and, assuming that the infinite sums are well defined, by continuous substitution of (1.50) we obtain

$$y_t = \sum_{h=0}^{\infty} \boldsymbol{A}_N {\boldsymbol{S}_N}^{-1} (c + \boldsymbol{X}_{t-h} \beta_N + \varepsilon_{t-h}). \tag{1.51}$$

The likelihood function of model (1.49) is given by

$$lnL(\theta, c) = -\frac{NT}{2} ln 2\pi - \frac{NT}{2} ln\sigma^2 + T ln|\boldsymbol{S}_N \lambda| - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \varepsilon_t'(\zeta) \varepsilon_t(\zeta) \quad , \tag{1.52}$$

where $\theta = (\delta', \lambda, \sigma^2)'$ and $\zeta = (\delta', \lambda, c')'$, $\delta = (\gamma, \varrho, \beta')'$, $\varepsilon_t(\zeta) = \boldsymbol{S}_N(\lambda) y_t - \gamma y_{t-1} - \varrho \boldsymbol{W}_N y_{t-1} - \boldsymbol{X}_t \beta - c$. The QML estimators for θ and c are the extreme estimators ($\hat{\theta}$ and $\hat{c}$) derived from the maximization of the likelihood function (1.52). When the error terms (ε_t) are normally distributed, $\hat{\theta}$ and $\hat{c}$ are ML estimators; when the errors are not normally distributed, then we have QML estimators.

Because the number of the parameters to be estimated goes to infinity as the cross-sectional dimension goes to infinity, the authors also propose a likelihood function that concentrates c out. Define $\tilde{y}_t = y_t - \bar{y}_T$ and $\tilde{y}_{t-1} = y_{t-1} - \bar{y}_{T-1}$ for $t = 1,2, \dots, T$, where $\bar{y}_T = \sum_{t=1}^{T} y_t / T$ and $\bar{y}_{T-1} = \sum_{t=1}^{T} y_{t-1} / T$. Similarly, $\widetilde{\boldsymbol{X}}_t$ and $\tilde{\varepsilon}_t$ are defined. Finally, $\widetilde{\boldsymbol{Z}}_t = (y_{t-1} - \bar{y}_T, \boldsymbol{W}_N y_{t-1} - \boldsymbol{W}_N \bar{y}_{T-1}, \boldsymbol{X}_t - \overline{\boldsymbol{X}}_T)$. The resulting concentrated likelihood function is

$$lnL(\theta) = -\frac{NT}{2} ln2\pi - \frac{NT}{2} ln\sigma^2 + Tln|\boldsymbol{S}_N \lambda| - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \tilde{\varepsilon}_t{}'(\zeta) \tilde{\varepsilon}_t(\zeta) \, , \quad (1.53)$$

where $\tilde{\varepsilon}_t(\zeta) = \boldsymbol{S}_N(\lambda) \tilde{y}_t - \widetilde{\boldsymbol{Z}}_t \delta$. The QML estimator for θ maximizes the concentrated likelihood function (1.53). By this approach, it is also possible to recover the estimated individual effects, which is not the case when other ML estimators are considered such as, for example, those proposed by Elhorst (2005) for either a SARE or a SAR dynamic panel data model with fixed effects.

The asymptotic properties of the QML estimators are based on the following assumptions:

QML_1. $\boldsymbol{W}_N$ is a constant $N \times N$ spatial weight matrix whose diagonal elements are equal to 0;

QML_2. The error terms ε_{it} are $i.i.d.$ across i and t with zero mean, variance σ^2 and at least one moment of order > 4 which is finite;

QML_3. $\boldsymbol{S}_N(\lambda)$ is invertible for all $\lambda \in \Lambda$. Furthermore Λ is compact and λ is in the interior of Λ;

QML_4. The elements of $\boldsymbol{X}_t$ are non-stochastic and uniformly bounded. Also $\lim_{T\to\infty}\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t{}'\widetilde{\boldsymbol{X}}_t$ exists and is nonsingular;

QML_5. $\boldsymbol{W}_N$ is uniformly bounded in row and column sums in absolute value. Also $\boldsymbol{S}_N{}^{-1}(\lambda)$ is uniformly bounded, uniformly in $\lambda \in \Lambda$;

QML_6. $\sum_{h=1}^{\infty} abs(\boldsymbol{A}_N^h)$ is uniformly bounded;

QML_7. N is a non-decreasing function of T and T goes to infinity.

For assumption QML_3 to be verified, in empirical applications where $\boldsymbol{W}_N$ is row-normalized, the parameter space for λ is just $(-1, 1)$. Moreover, in order to justify the absolute summability of $\boldsymbol{A}_N$, a sufficient condition is $\|\boldsymbol{A}_N\| < 1$, where the matrix norm is the row sum norm or the column sum norm. If $\boldsymbol{W}_N$ is row-normalized, one usually considers the spatial and temporal parameters satisfying the constraint $|\lambda| + |\gamma| + |\varrho| < 1$.

The proofs provided by the authors show that the concentrated QML estimator is consistent and asymptotically normal, but the limit distribution is not centered around zero. In order to overcome this, a bias reduction procedure is proposed which has a better performance than the standard QML estimator especially when $N \gg T$.

1.8.3 Least Squares Dummy Variable Estimator

Korniotis (2010) introduces a new bias-corrected estimator which is suitable for estimating a dynamic panel data model with fixed effects by Least Squares Dummy Variable (LSDV) and allows for spatial effects and endogenous regressors. The model to be estimated includes a time-lagged and a spatially lagged dependent variable, in a *time-space recursive* framework, and fixed effects (c). For each time period $t = 1,2, \dots, T$:

$$y_t = \gamma y_{t-1} + \varrho \boldsymbol{W}_N y_{t-1} + \boldsymbol{X}_t \beta + c + \varepsilon_t, \tag{1.54}$$

where ε_t are $i.i.d.$ error terms with zero mean, variance σ^2 and at least one moment of order > 2 which is finite. The set of control variables is made of endogenous regressors and can include both contemporaneous and time-lagged regressors. Moreover, N and T are assumed to grow at a finite rate and the usual assumptions on the spatial weight matrix, that needs to be row-standardized, hold. For details on the model assumptions, please refer to the original paper.

The standard LSDV estimate for γ is

$$\hat{\gamma}_{LSDV} = \left[\frac{1}{NT}\sum_{1=1}^{N}\sum_{t=1}^{T}(\tilde{X}_{i,t-1}^{d})'\tilde{X}_{i,t-1}^{d}\right]^{-1}\left[\frac{1}{NT}\sum_{1=1}^{N}\sum_{t=1}^{T}(\tilde{X}_{i,t-1}^{d})'\tilde{y}_{i,t}^{d}\right], \quad (1.55)$$

where the superscript d denotes that the data are de-meaned to have zero mean; $\tilde{X}_{i,t-1} = [y_{i,t-1}\ \boldsymbol{W}_i y_{t-1}\ X_{it}]$ is a vector of dimensions $1 \times (K+2)$ where X_{it} is a $1 \times K$ vector, $\boldsymbol{W}_i$ is the i-th row of the spatial weight matrix and $y_{t-1} = [y_{1,t-1}, y_{2,t-1}, \dots, y_{N,t-1}]'$. However, the LSDV estimator of γ is biased by the presence of fixed effects and endogenous regressors. Therefore a hybrid estimator is proposed which instruments the endogenous regressors and transforms equation (1.55) into

$$\hat{\gamma}_{b} = \left[\frac{1}{NT}\sum_{1=1}^{N}\sum_{t=1}^{T}(\tilde{Z}_{i,t-1}^{d})'\tilde{X}_{i,t-1}^{d}\right]^{-1}\left[\frac{1}{NT}\sum_{1=1}^{N}\sum_{t=1}^{T}(\tilde{Z}_{i,t-1}^{d})'\tilde{y}_{i,t}^{d}\right], \quad (1.56)$$

where $\tilde{Z}_{i,t-1} = [y_{i,t-1}\ \boldsymbol{W}_i y_{t-1}\ Z_{it-1}]$ and Z_{it-1} is a $1 \times K$ vector of instruments for X_{it}. The instruments are assumed to be contemporaneously correlated with the error term, but their time-lagged values are taken to be independent from the errors. This hybrid estimator is proved to have a finite asymptotic bias and to converge to a normal distribution. A bias-corrected estimator is then defined as

$$\hat{\gamma}_{c} = \left[\frac{1}{NT}\sum_{1=1}^{N}\sum_{t=1}^{T}(\tilde{Z}_{i,t-1}^{d})'\tilde{X}_{i,t-1}^{d}\right]^{-1}\left[\frac{1}{NT}\sum_{1=1}^{N}\sum_{t=1}^{T}(\tilde{Z}_{i,t-1}^{d})'\tilde{y}_{i,t}^{d} - \frac{\hat{B}_{NT}}{\sqrt{NT}}\right], \quad (1.57)$$

where $\widehat{B}_{NT}$ is the value of the asymptotic bias under a consistent estimator of γ (see Korniotis (2010) for greater details and the implementation procedure). The bias-corrected estimator only instruments the endogenous regressors whereas a pure IV estimator needs to instruments also the time and space-lagged dependent variable, therefore being more exposed to the weak instruments biases. It is proved to be asymptotically unbiased and asymptotically normal and to have good small sample properties.

1.9 Software availability for the estimation of spatial models

Theoretical contributions in the field of spatial econometrics are now numerous and cover a wide range of models and testing procedures, both for cross-sectional models and panel data models, as previous sections outlined. Although the subject appears to have reached a phase of maturity (Anselin 2010), empirical applications are still limited by the lack of software availability, particularly in the field of spatial panel data models.

1.9.1 Software availability for the estimation of spatial cross-sectional models

The availability of codes for cross-sectional spatial analysis is such that applied researchers can enjoy enough flexibility in the choice of the model to estimate.

A Cliff and Ord model of type (1.11) which may also contain endogenous regressor and heteroskedastic error terms can be estimated either via ML or GMM/IV by the `spreg` and `spivreg` Stata functions that also flexibly allow the estimation of SAR models of type (1.4) and SARE models of type (1.6) (Drukker et al. 2011; 2011a; 2011b).

The spatial econometrics toolbox for MATLAB provided by LeSage (1999) on his website[4] contains functions that are suitable for ML

[4] http://www.spatial-econometrics.com

estimation of Cliff and Ord models (1.11), SAR models (1.4) and SARE models (1.6), together with some functions that allow Bayesian estimation of spatial models.

The most appealing software for cross-sectional analysis is R, which offers a wide range of estimation choices with the packages `spdep` and `sphet` (Bivand 2006, 2013; Piras 2010): among others, we recall the ML estimation of a SARE model of type (1.6) (function `errorsarlm`), the estimation of a Cliff and Ord model as described in equation (1.11) either via ML (function `sacsarlm`) or GMM (function `gstslshet`), the estimation of a SAR model (1.4) via Spatial Two-Stage Least Squares (function `stsls`) or ML (function `lagsarlm`). R also provides plenty of functions that are suitable for exploratory spatial data analysis and testing.

1.9.2 Software availability for the estimation of spatial panel models

Procedures to estimate spatial panel data models are less numerous and this is still hindering these models to be applied in empirical studies. Some MATLAB routines are available on Paul Elhorst's website[5] for the ML estimation of static fixed effects and random effects SAR models, as described respectively in equations (1.37) and (1.38), and SARE models (Elhorst 2010), extended in order to include the bias correction procedure proposed by Lee and Yu (2010b). Some testing procedures are also available which can be used to test for a spatially lagged dependent variable or spatial error autocorrelation in a spatial static panel data model using (robust) LM tests (Elhorst 2010).

Ingmar Prucha also provides some Stata codes on his website[6], one of which is suitable for the estimation of a spatial error static panel data model as specified in equation (1.40) via the KKP estimator.

[5] http://www.regroningen.nl/elhorst/software.shtml
[6] http://econweb.umd.edu/~prucha/Research_Prog.htm

As in the cross-sectional case, R is the software that offers the richest although not yet comprehensive choice of estimating routines. The R package for spatial panel data estimation is `splm` (Millo and Piras 2012) and it allows ML estimation of Cliff and Ord type models (that include both a spatially lagged dependent variable and a spatially autocorrelated error term) either with spatially uncorrelated individual effects, as in equations (1.38) and (1.39), or where spatial correlation is present both in the individual and the error component, as in equation (1.40); both the fixed and random effects models are implemented. The KKP estimation procedure can be applied to a Cliff and Ord type of model with the error terms specified according to equation (1.40). Again, both the random and fixed effects models are considered.

The options for estimating spatial static panel data models are now numerous and, although no comprehensive packages have been implemented yet, empirical researches are offered a good choice. The estimation of spatial dynamic panel data models is instead made particularly difficult and hence less frequent in empirical analyses by the lack of readily available routines implemented in statistical and econometric software. To our knowledge, the only available code is the one published by Elhorst for the ML estimation of a dynamic panel data model including a serially lagged dependent variable, regional fixed effects and spatial error autocorrelation (Elhorst 2005), although some more codes are in preparation his website for the estimation of a dynamic panel data model including a serially lagged dependent variable, a spatially lagged dependent variable and individual fixed effects.

2 A Monte Carlo Investigation

2.1 Introduction

The literature about GMM estimation of spatial panel data models is still limited. As it was reviewed in chapter 1, the KKP estimator was introduced in order to estimate SARE static panel data models with random effects (Kapoor et al. 2007). A more recent contribution by Kukenova and Monteiro (2009) presents a system-GMM estimator for a time-space simultaneous model with fixed effects and additional endogenous covariates and compares its performance in finite samples against other spatial estimators; the authors find that the system-GMM estimator, although generally less efficient, tends to exhibit the smallest bias for the spatial autoregressive parameter and that it decreases as N and/or T increase. To our knowledge, however, no published paper has yet studied extensively the finite sample estimation of a space-time dynamic panel data model, which also includes a time lag of the spatially lagged dependent variable among the regressors, by a GMM-type estimator.

We therefore perform a Monte Carlo (MC) simulation exercise that permits an assessment of the performance of the most common estimators for dynamic panel data models for different temporal and cross-sectional dimensions and different degrees of spatial, temporal and spatiotemporal dependence.

Moreover, we also aim at studying the bias resulting from a non-spatial estimation of a dynamic panel data model that ignores the spatial effects that characterize the data, for different degrees of spatial dependence. To our knowledge, although the theoretical consequences of ignoring spatial dependence have been extensively studied (as described in the previous chapter), no empirical study is available for the assessment of the effects of such a misspecification in terms of bias of the estimates of the model coefficients.

The structure of the chapter is as follows. Section 2.2 offers a description of the simulation model which includes an introduction to the Monte Carlo principle, a description of the objectives of the present analysis and how it was implemented. Section 2.3 introduces the GMM-type estimator whose small sample properties we investigated through the MC analysis. The results of our simulation analysis are described in section 2.4, for various scenarios that differ in terms of degree of spatial, temporal and spatiotemporal dependence included in the model. Lastly, we will come to our conclusions in section 2.5.

2.2 Simulation model

2.2.1 The Monte Carlo principle

Monte Carlo simulations are based on the empirical tracking of a statistic's behavior in random samples drawn from known populations of simulated data. The strategy is to create an artificial world that resembles the real world as much as possible, whose characteristics are known to the researcher. This artificial world is called *pseudo-population*.

Following Mooney (1997), a basic MC procedure can be described as follows. First the Data Generating Process (DGP) is specified such that it describes all the characteristics of the *pseudo-population*. A computer algorithm needs to be developed in order to be able to generate data according to the specified DGP. Once the DGP is implemented, a *pseudo-sample* is generated such that it reflects certain characteristics of the sample that we want to put under investigation (e.g. sample size). The statistic(s) whose properties are being studied in the pseudo-environment is calculated in the *pseudo-sample* and stored. The sampling and estimating steps are repeated for a K number of trials, so that K values for each statistic under consideration are calculated and stored in a vector $\hat{\theta}$. The distribution of the resulting K values of the statistics is the MC estimate of the sampling distribution of $\hat{\theta}$ under the conditions specified for the *pseudo-population* and the *pseudo-sample*.

The rationale behind MC simulation is therefore quite simple to grasp. What is more difficult in practice is the specification of a proper DGP, the implementation of a computer code to implement it and the interpretation of the estimated sampling distribution.

2.2.2 A MC study for investigating the small sample properties of some estimators for a time-space dynamic model

The MC methodology is applied in order to investigate the finite sample properties of some estimators for a time-space dynamic model (section 1.7.4) with comparison purposes. The first one is the QML estimator developed by Yu et al. (2008), whose asymptotic properties have been analytically analyzed in their paper. Secondly, we take into consideration some GMM estimators for a dynamic panel data model à la Arellano and Bond (1991), extended so as to include a spatial lag of the dependent variable among the regressors. In order to deal with some econometric issues that arise in relation to the GMM procedure, two different GMM estimators are proposed.

All simulations are performed using Matlab R2011b[7]. Since the QML estimator was already thoroughly described in the previous chapter (section 1.8.2), no additional details will be provided on that. Differently, although a GMM estimators for spatial dynamic panel data models have been proposed in a working papers (Yu and Lee 2010), the one we are going to analyze has been developed autonomously and autonomously implemented in Matlab. A detailed description of this estimator is therefore provided in the following sections.

[7] The code for the QML estimator was kindly provided by Jihai Yu upon request specifically for this project.

2.2.3 Data Generating Process

The first step of the Monte Carlo simulation consists in the definition of the DGP. Here we take into consideration a time-space dynamic panel model with fixed-effects; hence, following the notation already presented in the previous sections, *pseudo-samples* are generated from:

$$y_t = \lambda \boldsymbol{W}_N y_t + \gamma y_{t-1} + \varrho \boldsymbol{W}_N y_{t-1} + x_t \beta + c + \varepsilon_t \tag{2.1}$$

$$x_t = x_{t-1}\delta + \eta_t \tag{2.2}$$

where $y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{nt})'$ and $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ are $N \times 1$ vectors and ε_{it} is $i.i.d.$ across i and t with zero mean and variance σ_ε^2; additionally, $c \sim N(0, \iota_N \sigma_c^2)$; $\varepsilon_t \sim N(0, \iota_N \sigma_\varepsilon^2)$; $\eta_t \sim N(0, \iota_N)$, with ι_N an N-dimensional vector of ones. The spatial weight matrix $\boldsymbol{W}_N$ is a $N \times N$ row-standardized rook matrix. The subscript N is dropped from now on for a more compact notation.

In order to prevent results from being influenced by initial observations, initial values for y_0 are drawn from a random standard normal distribution and the vector x_0 is initialized as a $N \times 1$ vector of zeros. Moreover, each variable is generated 60 times and the first $(60 - T)$ observations are then discarded.

The dependent variable is then generated according to the reduced form of equation (2.1):

$$y_t = (\boldsymbol{I}_N - \lambda \boldsymbol{W})^{-1}(\gamma y_{t-1} + \varrho \boldsymbol{W} y_{t-1} + x_t \beta + c + \varepsilon_t). \tag{2.3}$$

2.2.4 MC design

The aim of the present MC simulation is to assess the finite sample properties of the considered estimators for different values of N and T and to compare their performances since the true model is known. Sample sizes and degrees of spatial dependence in the data are varied in order to determine the conditions under which one estimator should be preferred to the others. Therefore, some different scenarios are simulated, each one described by different values for N and T and different values for the model parameters.

The first scenario is such that stationarity is ensured by the restriction $|\lambda| + |\gamma| + |\varrho| < 1$ (Yu et al. 2012). The MC experiment for stationary data relies on the following designs:

$$\begin{gathered} T \in \{5, 10, 50\}; \quad N \in \{16, 49, 121\}; \\ \gamma \in \{0.2, 0.4\}; \quad \varrho \in \{0.2, 0.4\}; \\ \lambda = 0.2; \quad \beta = 1; \quad \delta = 0.5 \\ \sigma_c^2 = 1; \quad \sigma_\varepsilon^2 = 1 \end{gathered} . \tag{2.4}$$

All combination of values for N and T are considered. When considering spatial and temporal parameters, not all combinations of the values in (2.4) are admitted, but only those that ensure model stationarity. Therefore the settings used for our MC simulations in a stationary context, with regard to λ, γ and ϱ, are only those summarized in Table 2.1.

Table 2.1. Setting of spatial and temporal parameters in the simulation model in a stationary context

γ	0.2	0.4	0.2
λ	0.2	0.2	0.2
ϱ	0.2	0.2	0.4
$\lvert\lambda\rvert + \lvert\gamma\rvert + \lvert\varrho\rvert$	0.6	0.8	0.8

A second scenario that is taken into consideration in the MC experiment is that of a quasi-unit root context, in which $|\lambda| + |\gamma| + |\varrho|$ is close to 1. The following designs are adopted:

$$\begin{gathered} T \in \{5, 10, 50\}; \quad N \in \{16, 49, 121\}; \\ \gamma \in \{0.2, 0.58\}; \quad \varrho \in \{0.2, 0.58\}; \\ \lambda = 0.2; \quad \beta = 1; \quad \delta = 0.5 \\ \sigma_c^2 = 1; \quad \sigma_\varepsilon^2 = 1 \end{gathered} . \tag{2.5}$$

Again, all combinations of N and T are taken into consideration, but only certain combinations of values for the spatial and temporal parameters are admitted in the experimental setting, as described in Table 2.2.

Table 2.2. Setting of spatial and temporal parameters in the simulation model in a quasi-non-stationary context

γ	0.2	0.58
λ	0.2	0.2
ϱ	0.58	0.2
$\lvert\lambda\rvert + \lvert\gamma\rvert + \lvert\varrho\rvert$	0.98	0.98

Finally, a MC simulation was conducted so as to measure the bias that a non-spatial estimation of a spatial model would cause, which is in fact one of the most important reasons for sponsoring spatial econometrics techniques. This is done in a context where the stationarity condition is met and spatial dependence is not too strong, according to the following designs:

$$\begin{array}{c} T \in \{5, 10, 50\}; \quad N \in \{16, 49, 121\}; \\ \varrho \in \{0.1, 0.3\}; \quad \lambda \in \{0.1, 0.3\}; \\ \gamma = 0.3; \quad \beta = 1; \quad \delta = 0.5 \\ \sigma_c^2 = 1; \quad \sigma_\varepsilon^2 = 1 \end{array} . \tag{2.6}$$

In this third scenario, all parameter combinations are taken into consideration.

For each of the described designs, 999 trials were performed. For each set of generated pseudo-observations and each estimator we report some statistics which are suitable for assessing the properties of the estimators and comparing them:

- Mean: it is the average point estimates for the coefficients over the 999 iterations ($niter$). For each generic model coefficient θ, it is calculated as: $\hat{\hat{\theta}} = \frac{1}{niter} \cdot \sum_{k=1}^{niter} \hat{\theta}_k$;
- Standard Deviation (SD), calculated as $\left[\frac{1}{niter} \cdot \sum_{k=1}^{niter} (\hat{\theta}_k - \hat{\hat{\theta}})\right]^{1/2}$;
- Bias, calculated as $\hat{\hat{\theta}} - \theta$;
- Root Mean Square Error (RMSE): it is a measure of consistency and is suitable for assessing the quality of an estimator in terms of variability and bias, being defined as $\left[\frac{1}{niter} \cdot \sum_{k=1}^{niter} (\hat{\theta}_k - \theta)^2\right]^{1/2}$.

2.2.5 Relevant econometric issues

Having defined the design of the empirical MC simulations, it is important to draw the attention on some econometric issues that arise when a time-space dynamic model with fixed-effects as specified in equation (2.1) is to be estimated. The most relevant issues are relative to the presence of the spatial, temporal and spatiotemporal lags of the dependent variable. In order not to further complicate the model, the error terms are assumed to be homoskedastic.

Specific remedies are needed in order to account for the different sources of endogeneity in this kind of model. As we extensively discussed in the previous chapter (sections 1.4.2 and 1.7.2), being correlated with the error terms, the contemporaneous spatial lags of the dependent variable are endogenous regressors. The dynamic specification of the model introduces also a different kind of endogeneity: the time lagged dependent variable is correlated with the fixed effects and with past values of the error terms. Therefore it cannot be considered to be strictly exogenous but only sequential exogeneity holds, conditional on the unobserved effect c_i (Wooldridge 2010). For the same reason, the presence of a spatiotemporal lag of the dependent variable is also a source of endogeneity.

The most popular estimator for this kind of space-time dynamic model is the QML estimator proposed by Yu et al. (2008), which was previously reviewed (section 1.8.2). This is therefore one of the estimators that our MC analysis will consider.

Another possible approach to the estimation of this model consists in a GMM approach, which can address the different sources of endogeneity included in the model by instrumenting the regressors. Nevertheless, the definition of the instruments raises some econometric issues. The first of them is the instrument proliferation problem which could cause a bias in the GMM estimates (Roodman 2009). In the next section (2.3), the GMM estimators that are considered in the MC analysis are thoroughly described together with the strategies that were adopted in order to overcome the main econometric problems that we met.

Another possible source of bias when dealing with the estimation of this kind of model is caused by overlooking the spatial dependence that is present in the data. In this case, typically, the empirical researcher would estimate a non-spatial model such as

$$y_t = \gamma y_{t-1} + \boldsymbol{X}_t\beta + c + \varepsilon_t, \qquad |\gamma| < 1, \tag{2.7}$$

where $\boldsymbol{X}_t$ is a $N \times K$ matrix of independent regressors, not serially correlated with ε_t but correlated with the fixed effects c. The error terms are assumed to have finite moments and in particular $E(\varepsilon_{it}) = E(\varepsilon_{it}\varepsilon_{is}) = 0$ for $t \neq s$. Since this type of misspecification does not seem to be unlikely to be encountered, our MC simulations are also designed for evaluating its consequences in terms of bias. There is no need to point out that a first drawback of such a model misspecification consists in missing estimates for the spatial parameters and therefore missing evaluation of the spatial effects. This, however, may not necessarily result in a terrible bias associated to the estimates of the other model coefficients and may not prevent a meaningful assessment of the effects of the non-spatial regressors. If this is the case and the main interest of the researcher focuses on the β coefficients, the estimation of a simple non spatial model by a suitable estimator may not be a bad choice. Empirical evidence, however, is still lacking and our MC exercise precisely aims at providing some.

2.3 A GMM estimator for spatial dynamic panel data models

The spatial GMM estimator for model (2.1) that we propose stems from the Arellano and Bond difference-GMM estimator for a dynamic panel data model with fixed effects, both in the one-step and two-steps versions.

2.3.1 The Arellano and Bond (1991) difference-GMM estimator

When dealing with a model such as model (2.7), Arellano and Bond (1991) propose a GMM estimator that estimates the model in first differences, so that the time-invariant fixed effects are canceled out.

The set of moment conditions that can be used in order to define a set of instruments depends on whether the covariates are sequentially exogenous, so that $E(\varepsilon_{it}|y_i^{t-1},\boldsymbol{X}_i^t,c_i) = 0$ for $t = 1, \dots T$, $y_i^{t-1} = (y_{i1}, y_{i2}, \dots, y_{it-1})'$, and $\boldsymbol{X}_i^t = (\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \dots, \boldsymbol{X}_{it})'$, or strictly exogenous, so that $E(\varepsilon_{it}|y_i^{t-1},\boldsymbol{X}_i^T, c_i) = 0$ for $t = 1, \dots T$ and $\boldsymbol{X}_i^T = (\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \dots, \boldsymbol{X}_{iT})'$.

In case of sequential exogeneity of the covariates, $E(\boldsymbol{X}_{it}\varepsilon_{is}) \neq 0$ for $s < t$ and $E(\boldsymbol{X}_{it}\varepsilon_{is}) = 0$ for $s \geq t$, then only $\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \dots, \boldsymbol{X}_{is-1}$ are valid instruments in the differenced equation for period s. With regards to the time lag of the dependent variable, only the values of y lagged two periods or more are to be considered as valid instruments, being its time lag an endogenous regressor. Following Arellano and Bond (1991, page 280), the optimal matrix of instruments is therefore a $(T-2) \times (T-2)[(K-1)(T+1) + (T-1)]/2$ sparse matrix defined as

$$\boldsymbol{Z}_i = diag(y_{i1} \dots y_{is}\boldsymbol{X}_{i1}{}' \dots \boldsymbol{X}_{is+1}{}'), \qquad s = 1, \dots, T-2, \tag{2.8}$$

where each row corresponds to a time period for which instruments are available (see Annex 1 for the extended notation). The first row includes the valid instruments that are available for the first-differenced model for period $t = 3$, the last row is for period $t = T$.

Differently, if the covariates are all strictly exogenous, which means that $E(\boldsymbol{X}_{it}\varepsilon_{is}) = 0$ for all s and t, then their values in all time periods are valid instruments and the sparse matrix of instruments is defined as

$$\boldsymbol{Z}_i = diag(y_{i1} \dots y_{is}\boldsymbol{X}_{i1}{}' \dots \boldsymbol{X}_{iT}{}'), \qquad s = 1, \dots, T-2. \tag{2.9}$$

Let us also define $\tilde{y}_{it} = y_{it} - y_{it-1}$ and apply the same notation to the other variables in the model. Once a proper matrix of instruments has been defined, the choice of the weighting matrix to be employed in the GMM estimation procedure distinguishes between the Arellano and Bond one-step difference-GMM estimator (AB1) and two-step estimator (AB2).

For the one-step estimator, the chosen weighting matrix is given by

$$\boldsymbol{A}_N = (N^{-1} \sum_{i=1}^{N} \boldsymbol{Z}_i{}' \boldsymbol{H} \boldsymbol{Z}_i)^{-1} \tag{2.10}$$

where $\boldsymbol{H}$ is a $(T-2) \times (T-2)$ matrix defined as:

$$\boldsymbol{H} = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 \end{pmatrix}. \tag{2.11}$$

$\boldsymbol{A}_N$ is the optimal weighting matrix when homoscedasticity and absence of serial correlation is assumed. The AB1 estimator for the model parameters $\theta = (\gamma, \beta)$ is then defined as

$$\hat{\theta}_{AB1} = (\widetilde{\boldsymbol{X}}^{*\prime} \boldsymbol{Z} \boldsymbol{A}_N \boldsymbol{Z}' \widetilde{\boldsymbol{X}}^{*})^{-1} \widetilde{\boldsymbol{X}}^{*\prime} \boldsymbol{Z} \boldsymbol{A}_N \boldsymbol{Z}' \tilde{y} \ , \tag{2.12}$$

where $\boldsymbol{X}^* = (y_{it-1}, \boldsymbol{X}_{it}{}')'$, $\widetilde{\boldsymbol{X}}^*$ is a $(T-2)N \times K$ matrix and $\boldsymbol{Z}$ is the proper matrix of instruments.

The two-step estimator is given instead by the optimal choice of $\boldsymbol{A}_N$, which is $\boldsymbol{V}_N^{-1}$:

$$\hat{\theta}_{AB1} = (\widetilde{\boldsymbol{X}}^{*\prime} \boldsymbol{Z} \, \boldsymbol{V}_N^{-1} \boldsymbol{Z}' \widetilde{\boldsymbol{X}}^{*})^{-1} \widetilde{\boldsymbol{X}}^{*\prime} \boldsymbol{Z} \, \boldsymbol{V}_N^{-1} \boldsymbol{Z}' \tilde{y} \tag{2.13}$$

where

$$\boldsymbol{V}_N^{-1} = (N^{-1} \sum_{i=1}^{N} \boldsymbol{Z}_i{}' \, \tilde{e}_i \tilde{e}_i{}' \boldsymbol{Z}_i)^{-1} \tag{2.14}$$

with $\tilde{e}_i$ being the $(T-2) \times 1$ vector of residuals of the first step estimation.

The estimator we are considering is consistent for $N \to \infty$ and fixed T. According to the findings in Arellano and Bond (1991), it exhibits only a small finite-sample downwards bias, thus not surprisingly outperforming OLS and within-group estimators, and represents a gain in efficiency when compared to the Anderson and Hsiao (AH) IV estimators. A well-known drawback of the AB2 estimator is that it returns downwards biased estimated standard errors, particularly in finite samples and therefore requires some correction such as the Windmeijer (2005) correction. Nevertheless, the AB estimator has been found to suffer also from a severe finite-sample bias when the instruments for the differenced equation are weak (Blundell and Bond 1998).

2.3.2 The instrument proliferation

The number of moment conditions on which the AB difference-GMM estimator is based grows rapidly as T increases, so that the instrument count gets quickly very large. In particular, when estimating a model such as model (2.7), the number of available instruments is equal to $0.5(T-1)(T-2)+0.5(T+1)(T-2)K$. Despite the great popularity of the AB estimator, instrument proliferation is an often underestimated problem that is also shared by the system-GMM estimator introduced by Blundell and Bond (1998). The number of instruments increases as T increase and grows large relative to N, causing a number of undesirable outcomes. Roodman (2009) focuses on this issue and describes the main problems that arise from instrument proliferation in small samples. The first failure is represented by the overfitting of the endogenous variables, which biases the estimates towards the OLS estimates. Unfortunately, no testing procedure is available against the overfitting bias, although the problem has been studied in a number of contributions (Ziliak 1997, Windmeijer 2005). A second problem that is caused by instrument proliferation is the imprecise estimation of the optimal weighting matrix ($\boldsymbol{V}_N^{-1}$): the estimates of the parameters are still consistent, but efficiency is often affected and the already mentioned downward bias of the AB2 estimator is one of the consequences (Windmeijer 2005). Finally, but most importantly, the

Hansen J-test[8] for instrument validity is weakened and its p-value is not reliable: a high p-value is considered to be an indication in favor of the validity of the GMM results, but instrument proliferation may alter the result in the direction of increasing the p-value associated to the test as T and the number of instruments increase, because the two-step standard errors enter the Hansen test formula (Bowsher 2002).

Unfortunately, there is no clear guidance on what is a safe number of instruments. Although a general rule of thumb is considered to be keeping the instrument count lower than N, this cannot be viewed as a completely safe solution. A second suggestion comes from Roodman (2009) who considers a p-value greater than 0.25 for the Hansen test to be viewed with concern.

Some techniques have been proposed as a solution to the instrument proliferation problem, for which a good review can be found in Roodman (2009). The first strategy to limit the number of instruments is to use only some of them, up to a certain lag, instead of all available lags. The instrument count stops being more than proportional with respect to T and becomes linear in T. The strategy of truncating the instruments is quite common in empirical applications, although it should be stated that the definition of the cut-off lag is often arbitrary and does not follow an economic explanation.

The second approach combines the instruments through addition into smaller sets, without dropping any of the lags. It is also known as "collapse" after Roodman's terminology. The collapsed matrix of instruments for the equation in first-differences, for predetermined covariates, is

[8] The Hansen (1982) J-test is also called test for over-identifying restrictions and can be thought as a test of instrument validity, as it tests whether all the restrictions imposed by the model are jointly satisfied.

$$\boldsymbol{Z}_i = \begin{pmatrix} y_{i1} & \boldsymbol{X}_{i1}{}' & \boldsymbol{X}_{i2}{}' & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ y_{i1} & y_{i2} & \boldsymbol{X}_{i1}{}' & \boldsymbol{X}_{i2}{}' & \boldsymbol{X}_{i3}{}' & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots \\ y_{i1} & y_{i2} & \dots & \dots & \dots & \dots & y_{iT-2} & \boldsymbol{X}_{i1}{}' & \dots & \boldsymbol{X}_{iT-1}{}' \end{pmatrix}. \tag{2.15}$$

"Collapsing" is also a method that makes the instrument count linear in T and retains more information than the truncation method.

These two strategies can also be combined: the matrix of instruments can be collapsed and lag depth reduced. For example, in case of sequentially exogenous covariates and limiting the lag depth to 3, the collapsed and truncated matrix of instruments is equal to

$$\boldsymbol{Z}_i = \begin{pmatrix} y_{i1} & \boldsymbol{X}_{i1}{}' & \boldsymbol{X}_{i2}{}' & 0 & 0 & 0 \\ y_{i1} & y_{i2} & \boldsymbol{X}_{i1}{}' & \boldsymbol{X}_{i2}{}' & \boldsymbol{X}_{i3}{}' & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{T-4} & y_{iT-3} & y_{iT-2} & \boldsymbol{X}_{iT-3}{}' & \boldsymbol{X}_{iT-2}{}' & \boldsymbol{X}_{iT-1}{}' \end{pmatrix}. \tag{2.16}$$

Despite the amount of instrument reduction strategies that have been proposed in the literature, no clear indication has been given about what a safe way to proceed is and the robustness of the estimates to alternative specifications of the GMM estimator has not been tested extensively yet (see Bontempi and Mammi, 2012, for a first discussion of this topic).

2.3.3 A spatial difference-GMM estimator

Given what was discussed above, we autonomously developed a spatial difference-GMM (SAB) estimator which is suitable for estimating a space-time dynamic model, following the non-spatial estimator developed by Arellano and Bond (1991). The instrument proliferation problem has also been taken into account by applying both the instrument reduction strategies that were discussed. The estimation with the full set of instruments was not performed because it was computationally unfeasible given the high instrument count for this kind of model, especially as T increased.

According to the DGP that was adopted for the MC simulation exercise, described in section 2.2.3, the definition the proper instruments for the differenced equation followed from the literature and autonomous thinking, as reported in Table 2.3.

Table 2.3. Variables and available instruments for a spatial difference-GMM estimator

Variable	Moment conditions	Instruments
y_{t-1} Endogenous	$E(y_{it-s}\tilde{\varepsilon}_{it})$ $t = 3, \dots, T$ $s = 2, \dots, (t-1)$ (Arellano and Bond 1991)	$\{y_1, \dots, y_{t-2}\}$ $t = 3, \dots, T$
$\boldsymbol{W}y_t$ Endogenous	$E(\boldsymbol{W}y_{it-s}\tilde{\varepsilon}_{it})$ $t = 3, \dots, T$ $s = 2, \dots, (t-1)$ (Kukenova and Monteiro 2009)	$\{\boldsymbol{W}y_1, \dots, \boldsymbol{W}y_{t-2}\}$ $t = 3, \dots, T$
$\boldsymbol{W}y_{t-1}$ Endogenous	$E(\boldsymbol{W}y_{it-s}\tilde{\varepsilon}_{it})$ $t = 3, \dots, T$ $s = 2, \dots, (t-1)$	$\{\boldsymbol{W}y_1, \dots, \boldsymbol{W}y_{t-2}\}$ $t = 3, \dots, T$
x_t Sequentially exogenous	$E(x_{it-s}\tilde{\varepsilon}_{it})$ $t = 3, \dots, T$ $s = 1, \dots, (t-1)$ (Arellano and Bond 1991)	$\{x_1, \dots, x_{t-1}\}$ $t = 3, \dots, T$

Since in empirical applications it is usually unknown whether the covariates are strictly exogenous, predetermined or even endogenous, we treat the covariate as a sequentially exogenous variable, in order to be conservative with respect to the choice of strict exogeneity.

We consider both the one-step and the two-step SAB estimator, based on a collapsed matrix of instruments. We also limit the choice of instruments to the third lag, in order to avoid instrument proliferation. The matrix of instruments is therefore defined as:

$\boldsymbol{Z}_i =$

$$\begin{pmatrix} y_{i1} & \boldsymbol{W}y_{i1} & x_{i1} & x_{i2} & 0 & 0 & 0 & 0 & 0 \\ y_{i1} & y_{i2} & \boldsymbol{W}y_{i1} & \boldsymbol{W}y_{i2} & x_{i1} & x_{i2} & x_{i3} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{T-4} & y_{iT-3} & y_{iT-2} & \boldsymbol{W}y_{iT-4} & \boldsymbol{W}y_{iT-3} & \boldsymbol{W}y_{iT-2} & x_{iT-3} & x_{iT-2} & x_{iT-1} \end{pmatrix} \tag{2.17}$$

We also considered a second, extended, specification of a spatial GMM estimator (ESAB) which is based on the additional moment condition:

$$E(\boldsymbol{W}x_{it-s}\tilde{\varepsilon}_{it}) \text{ for } t = 3, \dots, T \text{ and } s = 1, \dots, (t-1), \tag{2.18}$$

that leads to the definition of an additional set of instruments: $\{\boldsymbol{W}x_{t-3}, \boldsymbol{W}x_{t-2}, \boldsymbol{W}x_{t-1}\}$ for $t = 3, \dots, T$ that is added to each row of the instrument matrix. In this case, the total instrument count is limited to $(T-2) + 4 \cdot nlags \cdot K$, where $nlags$ is the reduced lag depth (and it is equal to 3 in our case) and K is the number of covariates.

2.4 Results

The MC simulations were performed according to the designs previously described (see section 2.2.4). The finite sample performance of the QML estimator by Yu et al. (2008), the one-step and two-step SAB and the one-step and two-step ESAB were tested for various values of N and T and for different values of the parameters, according to the different scenarios of stationarity and quasi-unit root.

We also performed a non-spatial estimation of the same spatial data (generated according to the described DGP), through a difference-GMM à la Arellano and Bond (1991) and compare the results against the QML estimator and the ESAB estimator (both one and two-steps), in order to assess the bias that the estimates suffer if the spatial dimension of the data is not properly empirically modeled.

2.4.1 Stationary scenario

The first scenario that is taken into consideration is that of stationarity of the data, in which $|\gamma| + |\lambda| + |\varrho| < 1$ (Yu et al. 2012).

When spatiotemporal dependence is limited, for $|\gamma| + |\lambda| + |\varrho| = 0.6$, all the estimates obtained in our MC simulation are well centered around the true value of the parameters (Figure 2.37 to Figure 2.45 in Annex 2). The RMSE error of the estimates for all considered coefficients $(\gamma, \lambda, \varrho, \beta)$ decreases as T increases from 5 to 50 in our simulations and, for each value of T, it also decreases as N increases for all the estimators included in the simulation (Figure 2.1 to Figure 2.4). The estimator that appears to have the best performance in terms of RMSE for all the T and N under consideration is the QML estimator, while the SAB estimator is generally outperformed by its extended version which also includes the spatial lag of the covariate among the regressors. Only with respect to the estimation of parameter γ, which is the time-lag parameter for the autoregressive term of order 1, the QML estimator is outperformed in terms of RMSE by the GMM estimators as N increases. It is worth noting that, with respect to this parameter, even if the general performance of the QML estimator improves as the cross-section dimension grows, it is outperformed by the GMM estimators that, in their non-spatial version à la Arellano and Bond, are consistent for $N \to \infty$ and fixed T. The GMM estimation of the AR(1) parameter γ is therefore not surprisingly particularly good in terms of bias and RMSE with small T and growing N (Table 2.4 to Table 2.6 in Annex 2).

Differently, particularly when T is small, the spatial GMM estimators appear to produce less reliable estimates for the other considered parameters (the spatial parameters and the coefficient of the covariate), mainly because of a greater variability of the estimates and therefore higher standard deviations (see Table 2.4 to Table 2.6 and Figure 2.37 to Figure 2.45 in Annex 2).

With respect to the estimates of the spatial autoregressive parameter λ, our simulations find that the bias associated to the ESAB estimator (both one and two-step) tends to decrease, as T and N increase, up to being comparable to the bias associated to the QML estimates when $T = 50$. Nevertheless, the ESAB estimates of λ are always associated to a higher variability. This result is in line with what Kukenova and Monteiro (2009) find in their simulation exercise for a time-space simultaneous model and a spatial system-GMM estimator, relatively to the estimation of the spatial autoregressive parameter.

Growing spatiotemporal dependence, for $|\gamma| + |\lambda| + |\varrho| = 0.8$[9], does not change dramatically the performance of the considered estimators. Growing T and N improve the estimates in terms of RMSE for all the estimators and all parameters (see Figure 2.5 to Figure 2.8 and Figure 2.9 to Figure 2.12). As previously noticed, the QML estimator is still the best performing estimator for all T and N and all parameters, with the only exception of parameter γ. In these settings as in the previous one, as N grows, the spatial GMM estimators of γ perform better than the QML estimator (see Figure 2.5 and Figure 2.9). As for what it concerns the estimation of the spatial coefficients (ϱ and λ) and of the coefficient of the covariate (β), the QML appears to produce more reliable estimates than spatial GMM the ESAB and, particularly, the SAB estimators, especially when N is small. This is due to both a bigger bias and standard deviation of the estimates of these parameters when the GMM approach is adopted (see Table 2.7 to Table 2.9 and Table 2.10 to Table 2.12 in Annex 2).

[9] In our MC design, this is the case for the following two parameter settings: (a) $\gamma = 0.4$, $\lambda = \varrho = 0.2$, $\beta = 1$ and (b) $\varrho = 0.4$, $\lambda = \gamma = 0.2$, $\beta = 1$.

Figure 2.1. RMSE of γ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

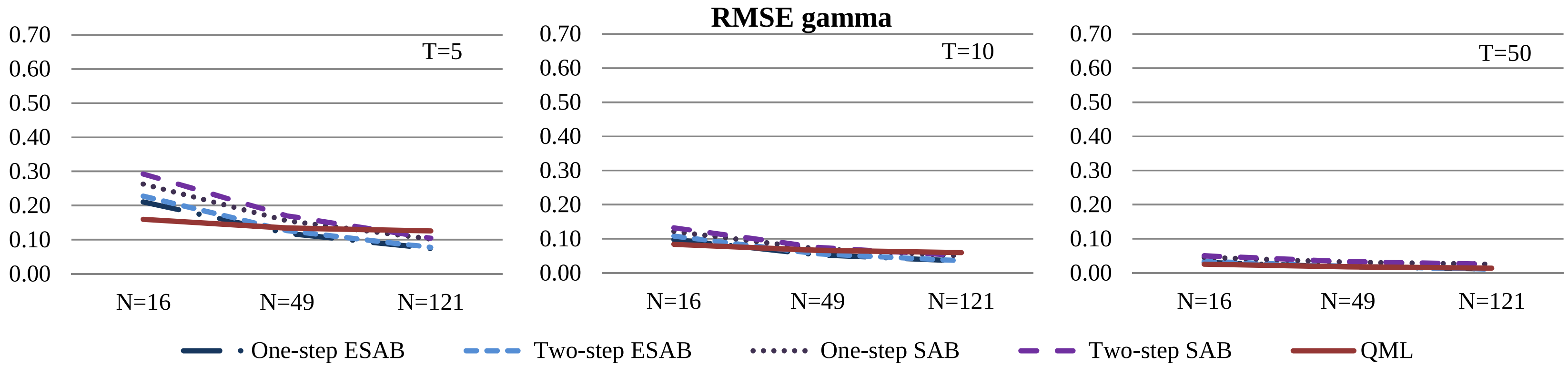


Figure 2.2. RMSE of λ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

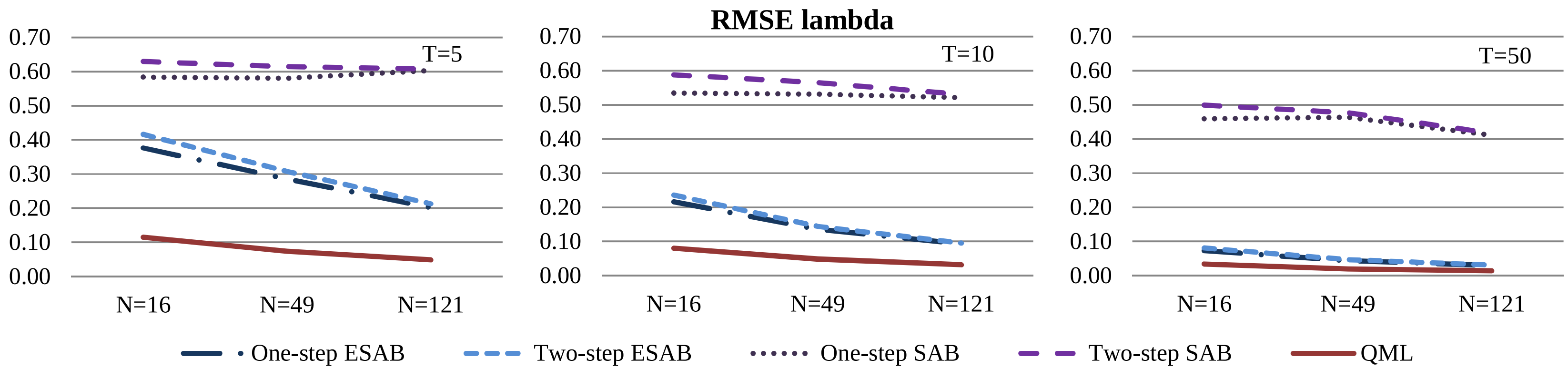


Figure 2.3. RMSE of ϱ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

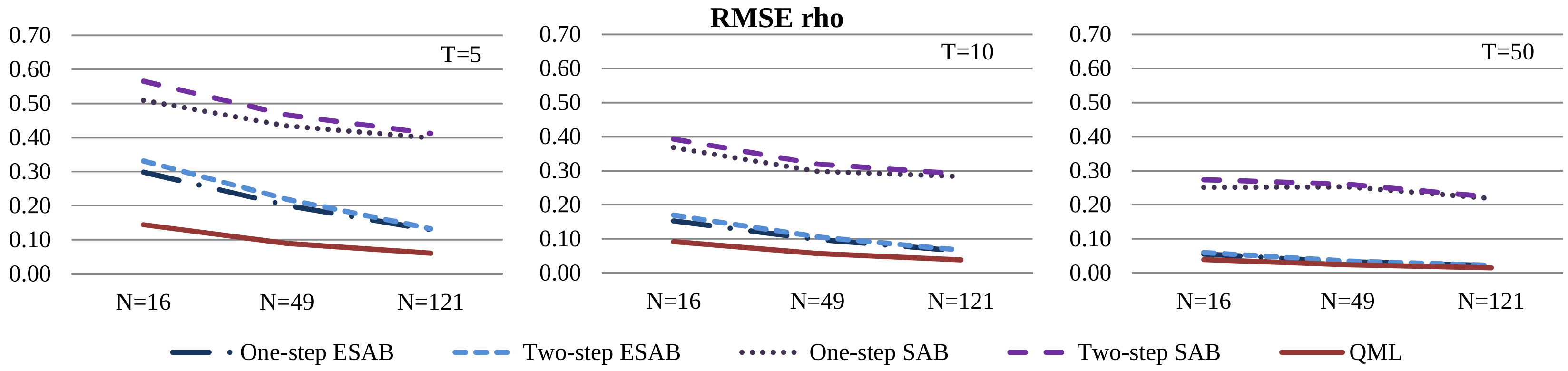


Figure 2.4. RMSE of β for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

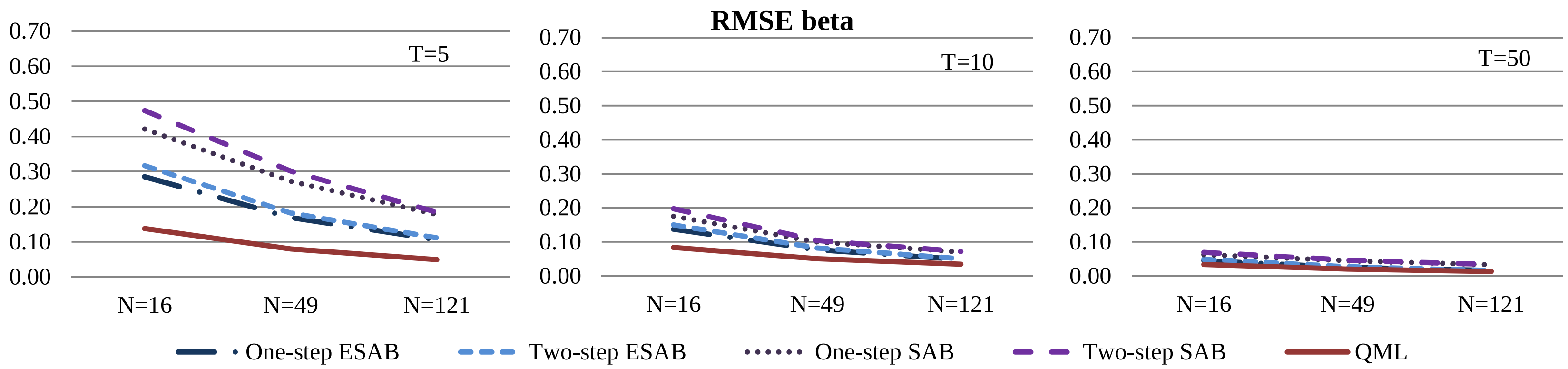


Figure 2.5. RMSE of γ for various spatial estimators, for $\gamma = 0.4$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

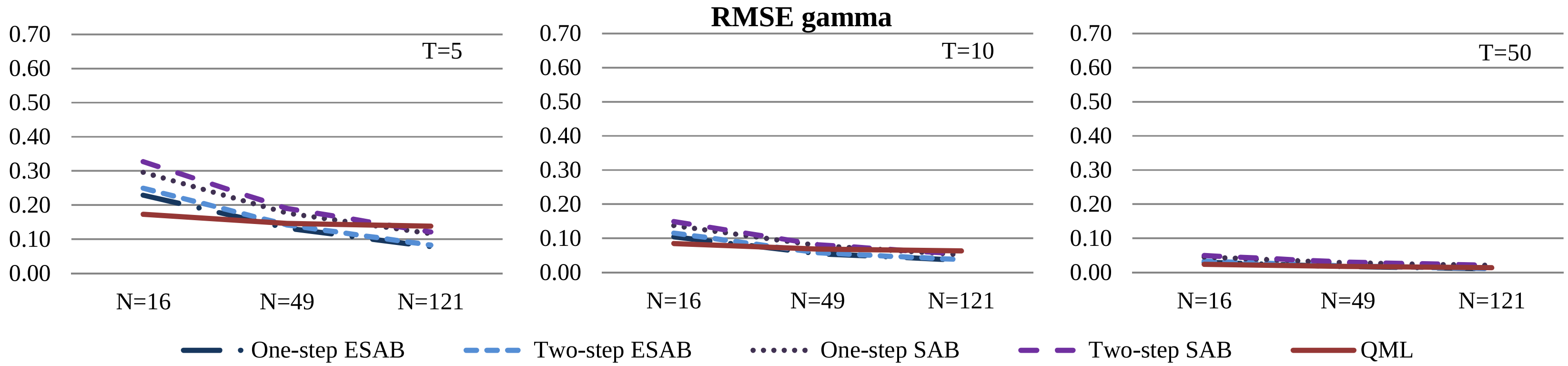


Figure 2.6. RMSE of λ for various spatial estimators, for $\gamma = 0.4$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

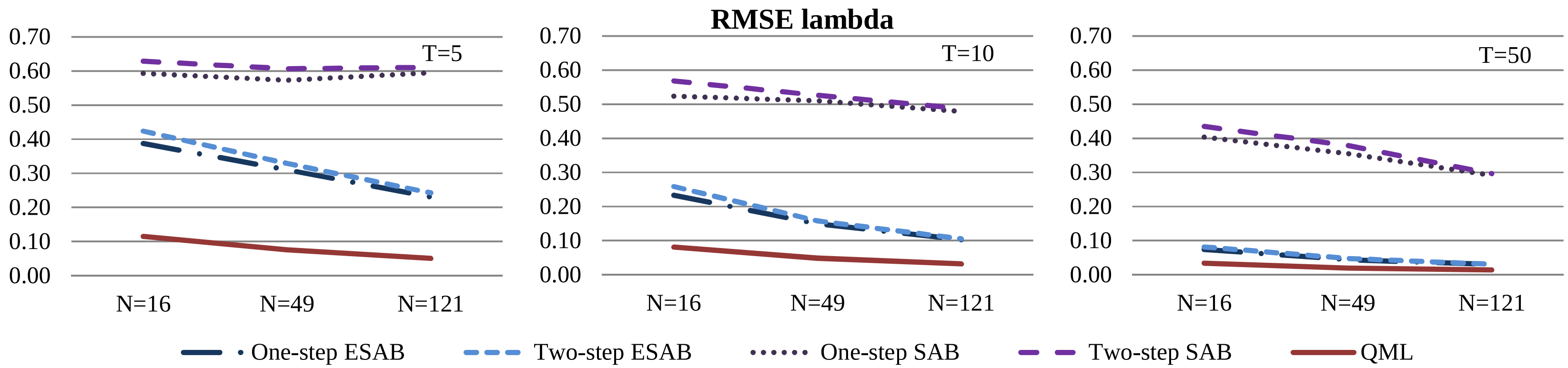


Figure 2.7. RMSE of ϱ for various spatial estimators, for $\gamma = 0.4$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

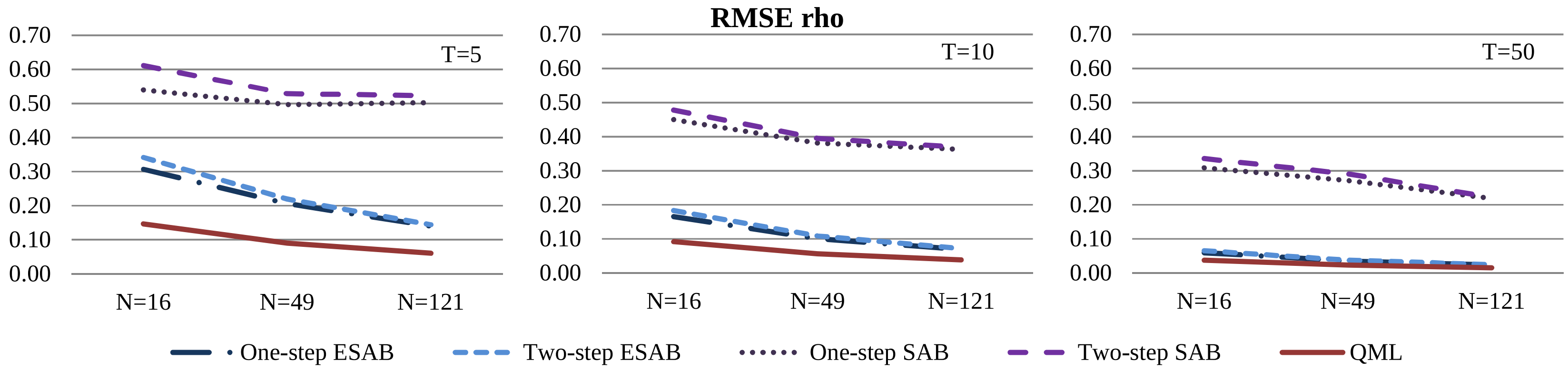


Figure 2.8. RMSE of β for various spatial estimators, for $\gamma = 0.4$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

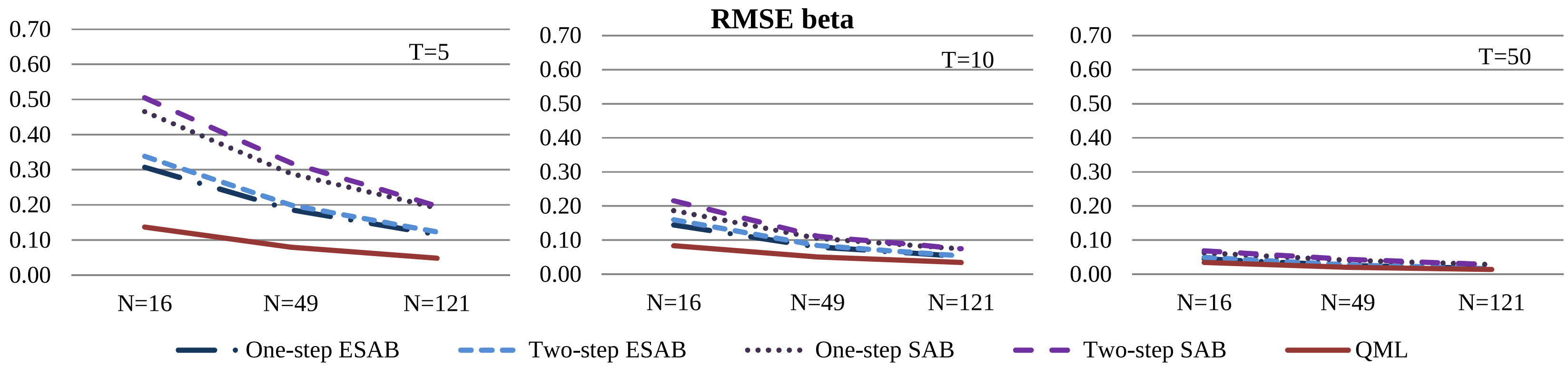


Figure 2.9. RMSE of γ for various spatial estimators, for $\varrho = 0.4$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

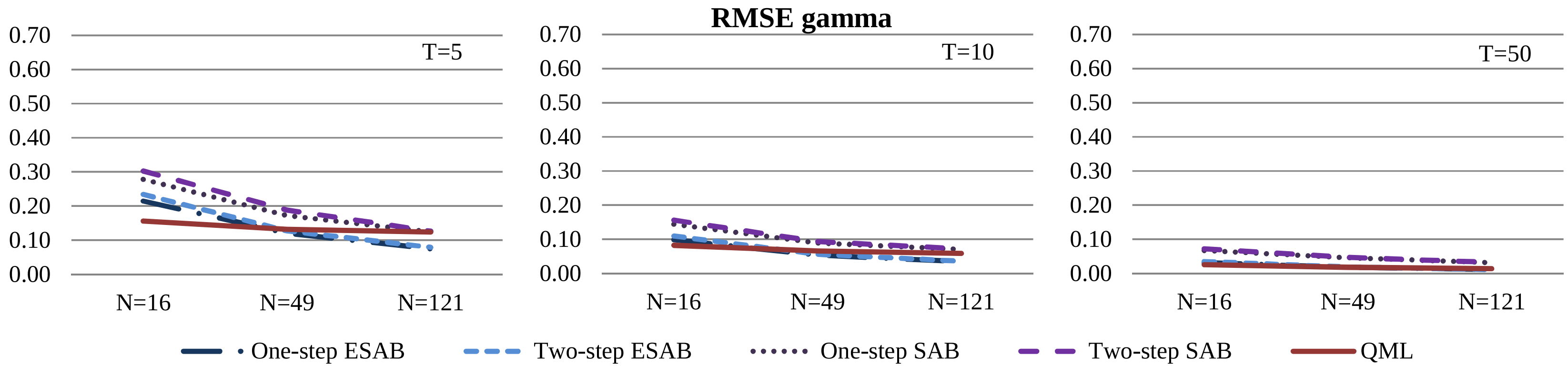


Figure 2.10. RMSE of λ for various spatial estimators, for $\varrho = 0.4$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

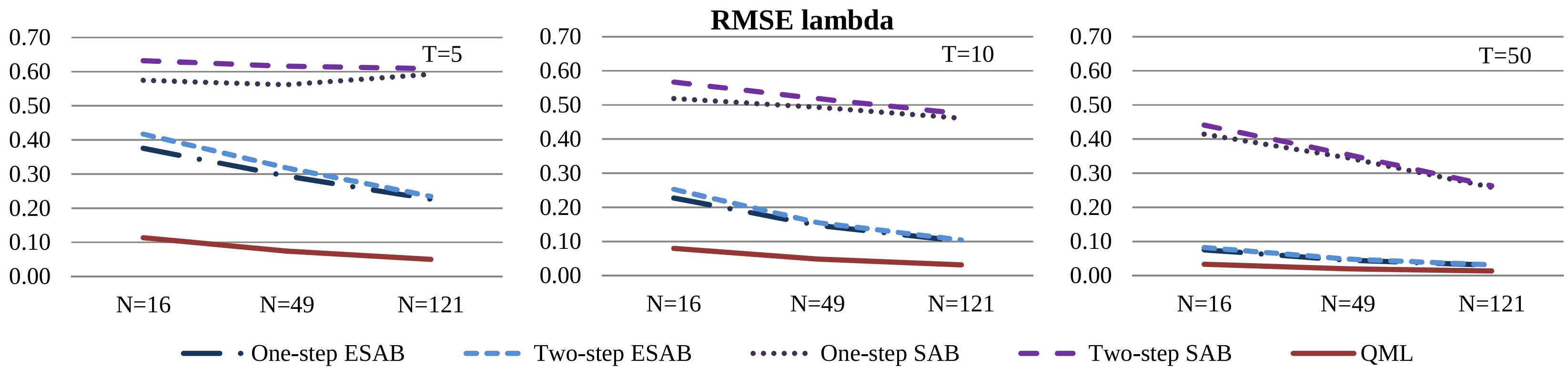


Figure 2.11. RMSE of ϱ for various spatial estimators, for $\varrho = 0.4$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

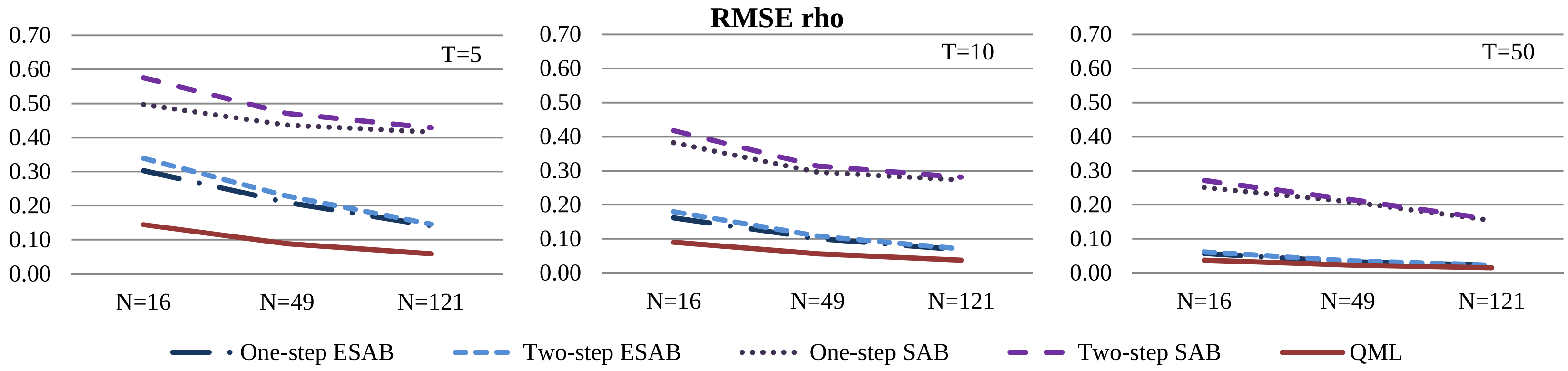
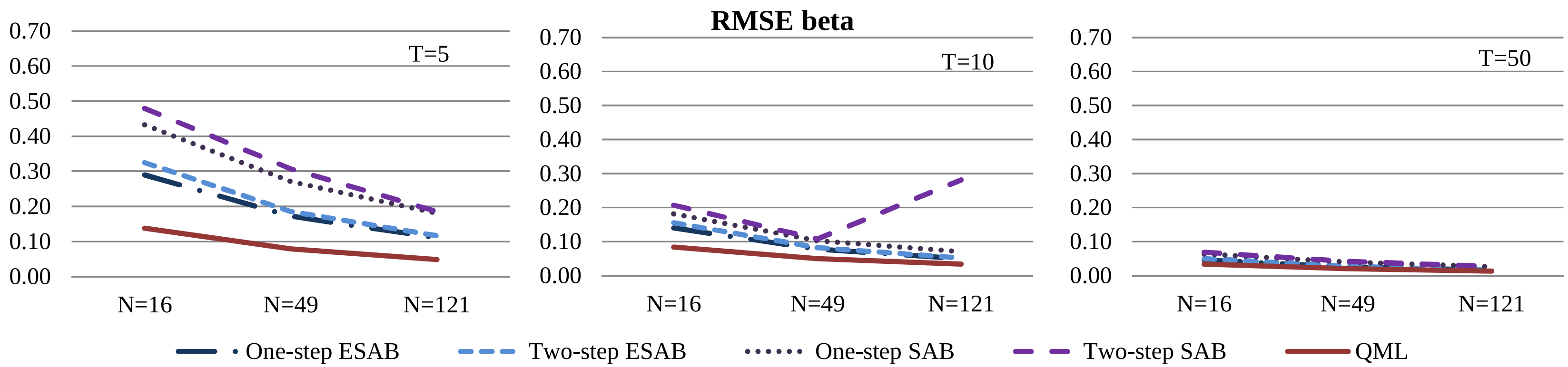


Figure 2.12. RMSE of β for various spatial estimators, for $\varrho = 0.4$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations



2.4.2 Quasi-unit root scenario

A quasi-non stationary scenario was also simulated, in which $|\gamma| + |\lambda| + |\varrho| = 0.98$[10], thus generating a quasi-unit root panel pseudo-dataset.

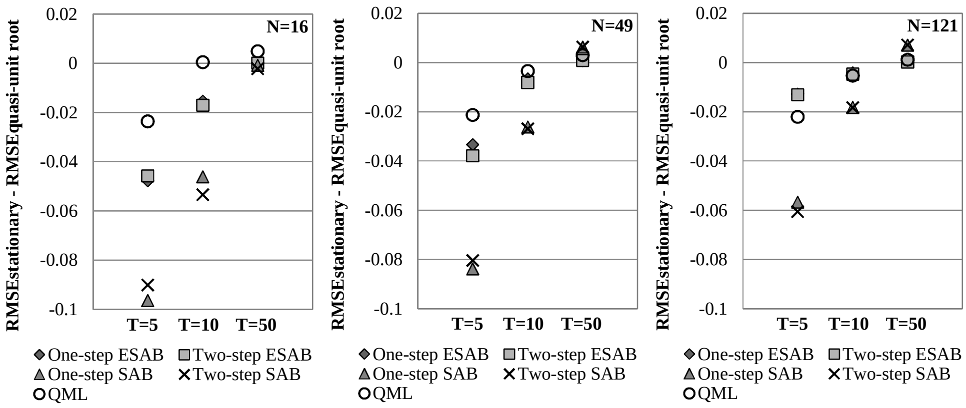
The QML estimator is confirmed to be the one that produces the smallest RMSE for all estimated parameters, with the only exception of the AR(1) coefficient γ, for which the GMM estimates are associated to a smaller RMSE for growing values of N and fixed T (see Figure 2.14 and Figure 2.18). On the contrary, the spatial GMM estimators return estimates for the other coefficients of the model which are generally more biased and affected by higher variability than the QML estimates (see Table 2.13 to Table 2.18 in Annex 2): it is so for all parameters and all values of T and N with the only exceptions of parameters ϱ and β when T and N get larger ($T = 50$ and $N = 121$), although the effect of a smaller bias is wiped out by a larger variability, that causes the RMSE associated to the GMM estimates to be larger than the one associated to the QML estimates (Table 2.15 and Table 2.18 in Annex 2).

If compared to the stationary scenario, the RMSE associated to the estimates of parameters γ and ϱ (the parameters that measure time-dependence) show higher values particularly for smaller values of T. Moreover, it should be noted that the RMSE associated to the parameter that is set to take on the value of 0.58 becomes smaller in the quasi-unit root setting with respect to the stationary setting as T and N increase. For example, let us consider the two alternative settings in which $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$ and $\gamma = 0.58$, $\lambda = \varrho = 0.2$, $\beta = 1$ and compare the RMSEs associated to parameter γ for different T and N by subtracting the RMSE calculated in the second setting to the one calculated in the first setting

[10] The following parameter settings were adopted: (a) $\gamma = 0.58$, $\lambda = \varrho = 0.2$, $\beta = 1$ and (b) $\varrho = 0.58$, $\lambda = \gamma = 0.2$, $\beta = 1$.

(thus a negative number indicates a higher RMSE associated to the quasi-unit root scenario).

Figure 2.13. Difference in RMSE for parameter γ in a stationary and a quasi-unit root scenario for different values of *N* and *T*



It is evident from Figure 2.13 not only that the difference between this statistic calculated for the two considered settings becomes smaller as T increases, but also that the RMSE for the quasi-unit root scenario becomes smaller than that of the stationary scenario for the majority of the estimators as T and N increase.

Figure 2.14. RMSE of γ for various spatial estimators, for $\gamma = 0.58$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

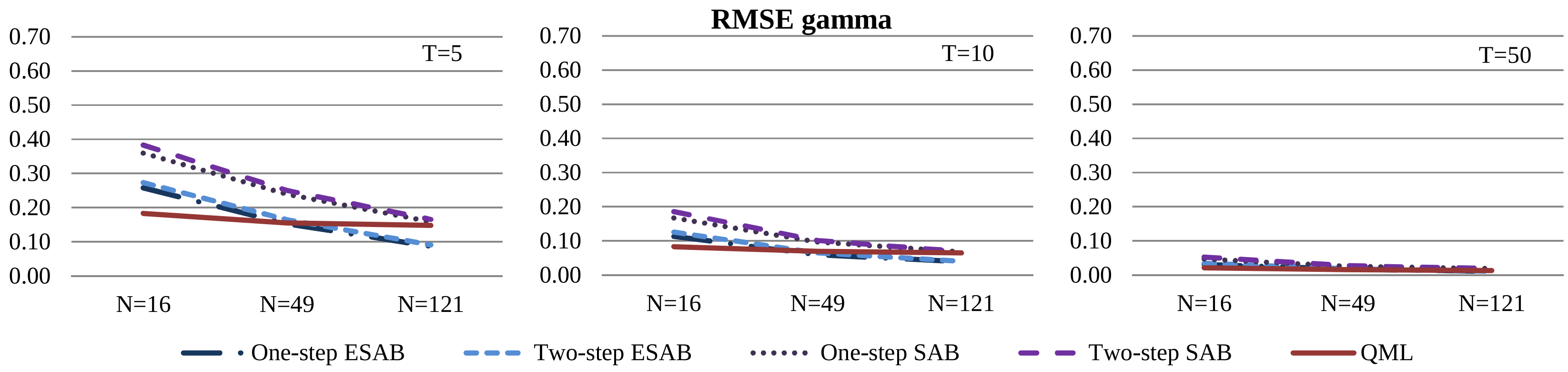


Figure 2.15. RMSE of λ for various spatial estimators, for $\gamma = 0.58$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

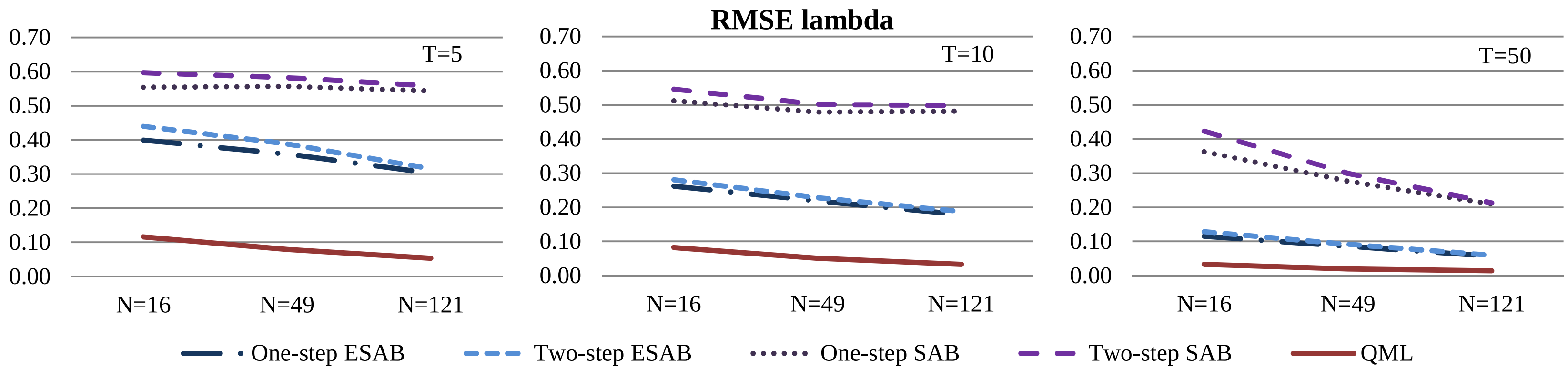


Figure 2.16. RMSE of ϱ for various spatial estimators, for $\gamma = 0.58$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

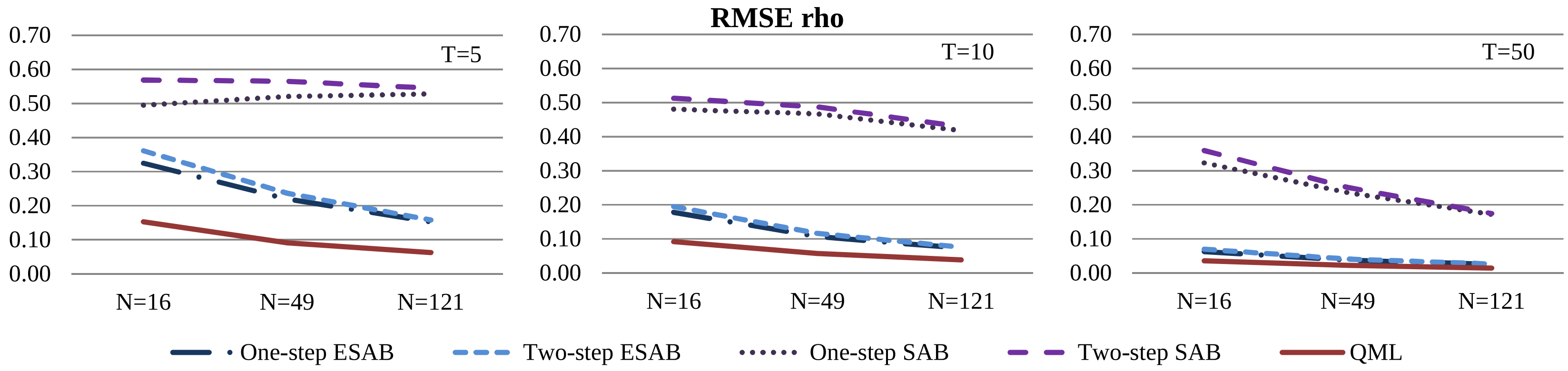


Figure 2.17. RMSE of β for various spatial estimators, for $\gamma = 0.58$, $\lambda = \varrho = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

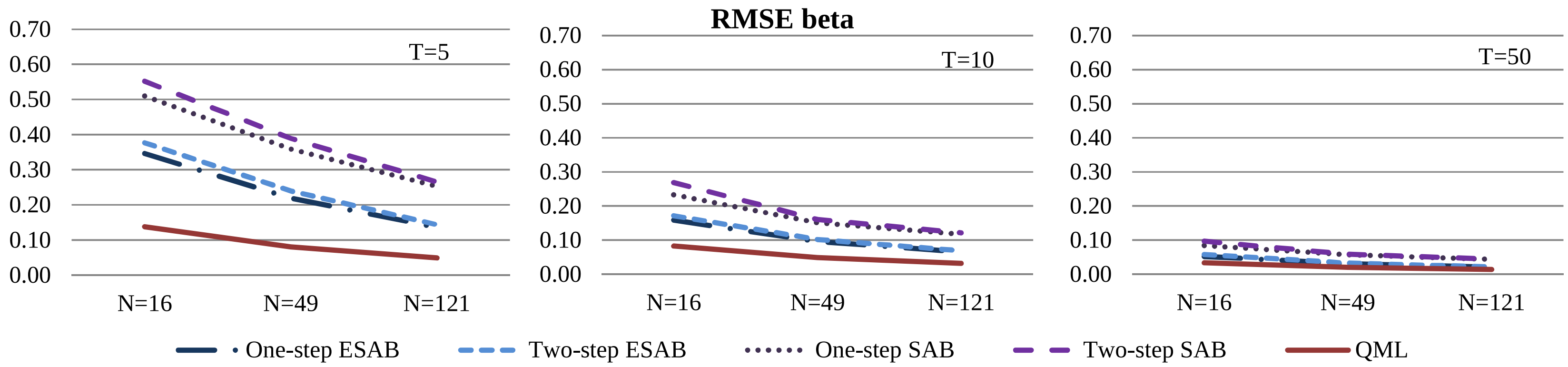


Figure 2.18. RMSE of γ for various spatial estimators, for $\varrho = 0.58$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

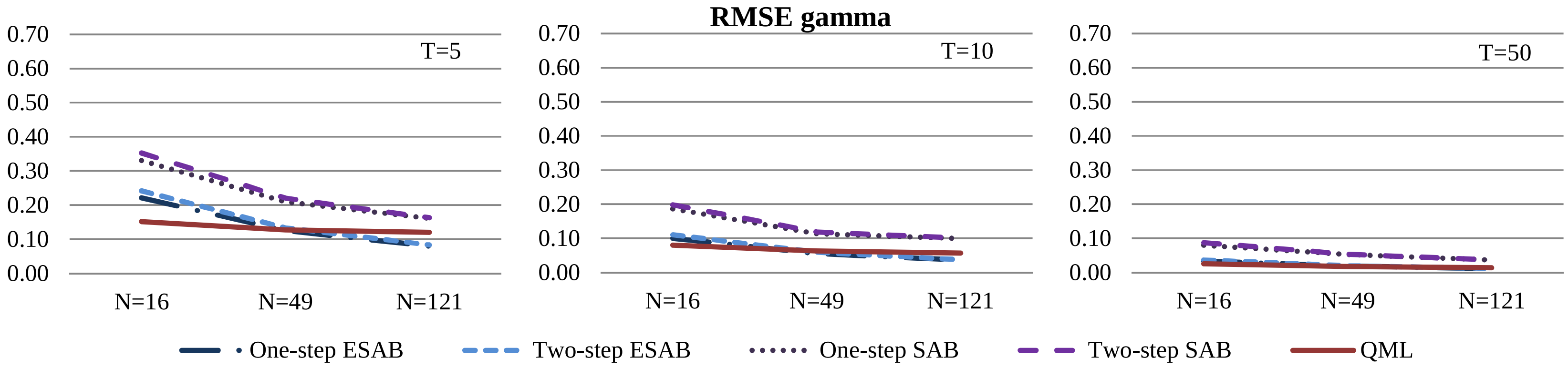


Figure 2.19. RMSE of λ for various spatial estimators, for $\varrho = 0.58$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

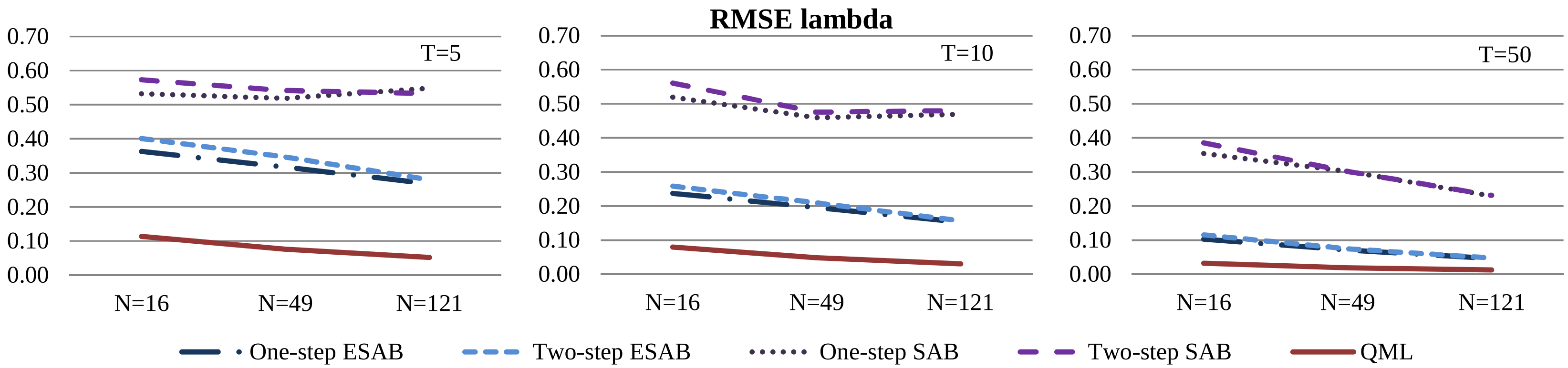


Figure 2.20. RMSE of ϱ for various spatial estimators, for $\varrho = 0.58$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations

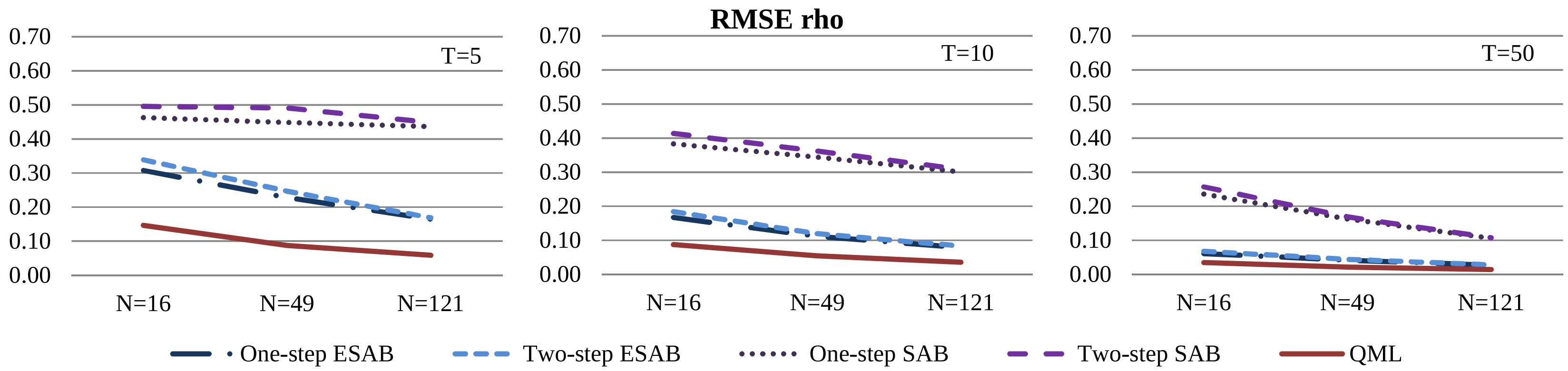
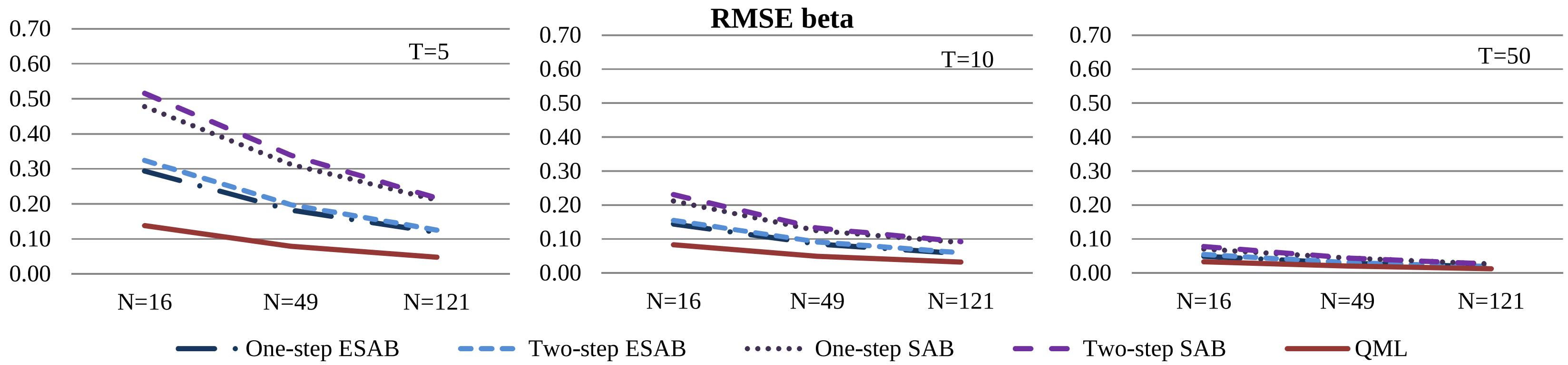


Figure 2.21. RMSE of β for various spatial estimators, for $\varrho = 0.58$, $\lambda = \gamma = 0.2$, $\beta = 1$ and various T and N, over 999 iterations



2.4.3 Spatial – Non spatial scenario

The performance of non-spatial GMM estimators is compared to that of the ESAB and the QML estimators in order to assess the risks that an empirical researcher faces when s/he fails to take the spatial dimension of the data into account.

In particular, we imagine that a researcher who is not aware (or not convinced) of the need to apply the spatial econometrics tools when spatial dependence is present in the data will be only interested in the estimation of the coefficient(s) β of the model, in order to evaluate the effects of the covariate(s) on the dependent variable. In particular, we expect that a model specification such as the space-time dynamic model described in equation (2.1) might be reduced to a non-spatial dynamic panel data model with fixed effects such as

$$y_t = \gamma y_{t-1} + x_t \beta + c + \varepsilon_t, \quad \varepsilon_t \sim N(0, \iota_N \sigma_\varepsilon^2). \tag{2.19}$$

Given the popularity of the GMM approach for the estimation of dynamic panel data models, we suppose that a "non-spatial" researcher will often choose a GMM estimation strategy when s/he needs to treat a model specified as in equation (2.19): this success is mainly due to the flexibility of the estimator, to the availability of internal instruments and to the easy implementation of GMM estimation in the most popular econometric and statistical software packages, such as R, Stata and Matlab.

Even if the effects that ignoring spatial dependence may have on the estimates in terms of bias and efficiency have been thoroughly identified and described in the spatial econometrics literature and reviewed in previous sections (chapter 1), to our knowledge no empirical study has been published that tries to quantify the bias that may affect the estimated β coefficient(s) when spatial dependence is ignored.

Our simulations consider different degrees of spatiotemporal dependence. The first setting is characterized by the following parameter values: $\gamma = 0.3$, $\lambda = \varrho = 0.1$, $\beta = 1$, that define a situation of little spatial and spatiotemporal dependence. Clearly, the non-spatial difference-GMM estimation à la Arellano and Bond (1991) does not return any estimated value for parameters λ and ϱ. Since an empirical researcher is probably primarily interested in the value of the coefficient β, we will focus on the performances of the considered estimators relatively to this parameter.

When T is small ($T = 5$ and $T = 10$) and for all values of N, the non-spatial GMM estimators (both one-step and two-step estimators) (Table 2.19 and Table 2.20) are those that show the smallest RMSEs associated to the estimates of β. Nevertheless, all estimates appear to be quite well centered around the true value (Figure 2.24 and Figure 2.25). Differently, when $T = 50$, the ESAB estimates of β are associated to the smallest bias for all N, although the results in terms of RMSE are comparable among all five estimators considered when T is large (see Figure 2.23, third panel). However, the advantage in terms of smaller bias associated to non-spatial estimates of β is reduced by the effects of higher variability, that result in smaller RMSEs for QML estimates for all values of T and N.

Lastly, it should be noticed that the differences in variability decrease as T increases (see Figure 2.24 to Figure 2.26), thus primarily contributing to the reduction of the gap between the RMSEs in correspondence of a larger time dimension.

Figure 2.22. RMSE of γ for various spatial estimators, for $\gamma = 0.3$, $\lambda = \varrho = 0.1$, $\beta = 1$ and various T and N, over 999 iterations

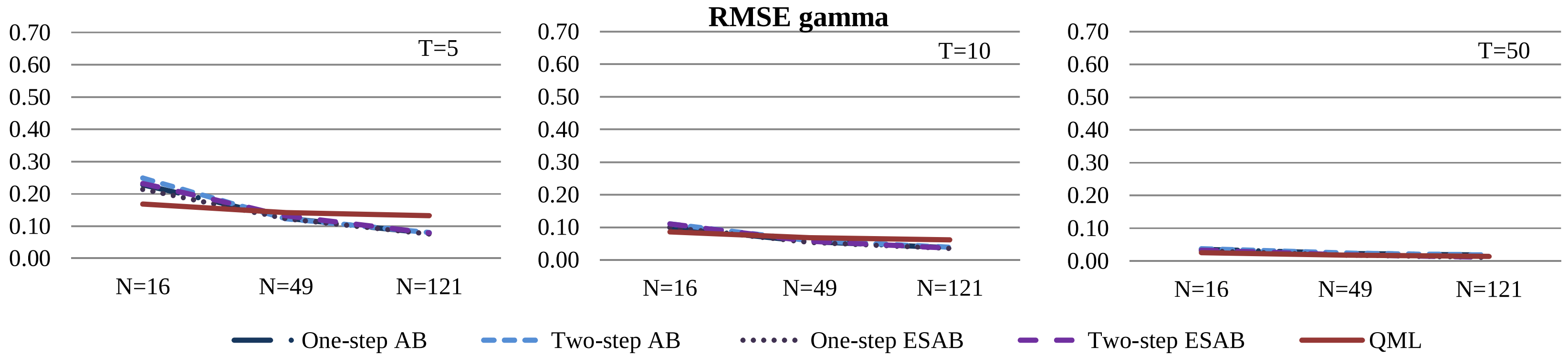


Figure 2.23. RMSE of β for various spatial estimators, for $\gamma = 0.3$, $\lambda = \varrho = 0.1$, $\beta = 1$ and various T and N, over 999 iterations

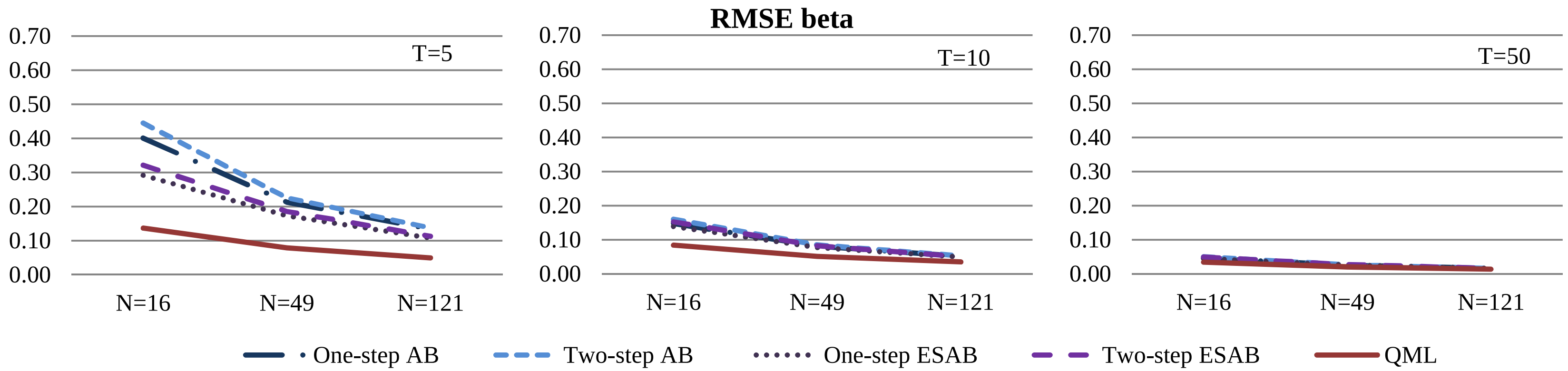


Figure 2.24. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = 0.3$, $\lambda = \varrho = 0.1$, $\beta = 1$, various T and $N = 16$ over 999 iterations

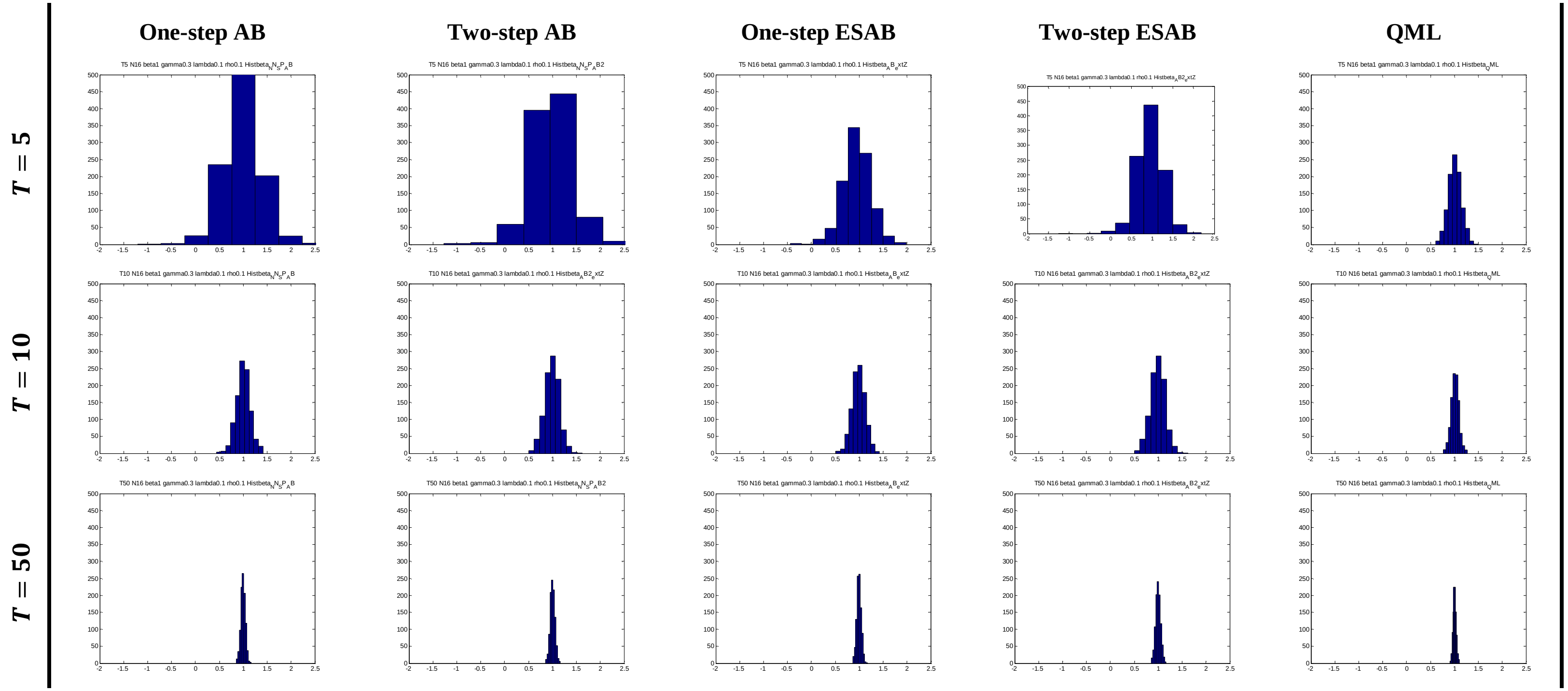


Figure 2.25. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = 0.3$, $\lambda = \varrho = 0.1$, $\beta = 1$, various T and $N = 49$ over 999 iterations

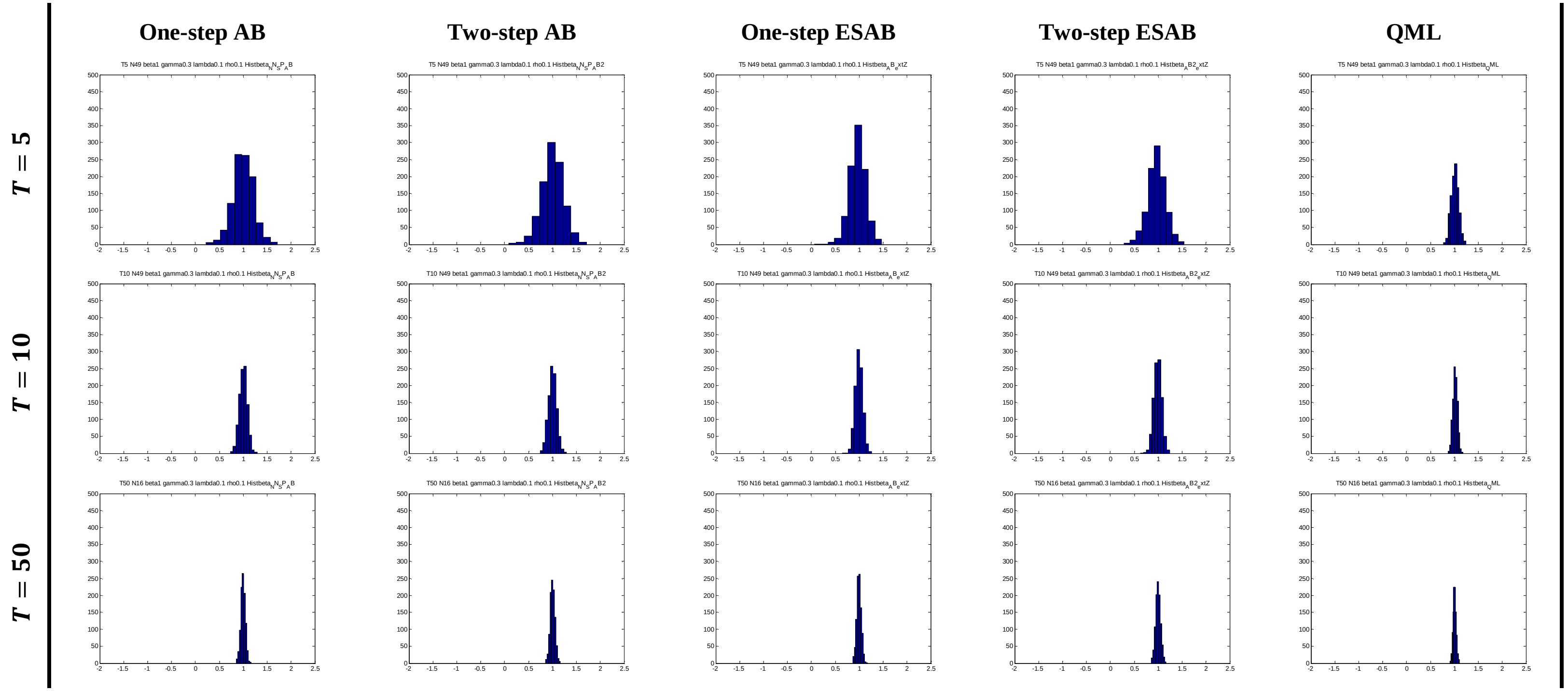
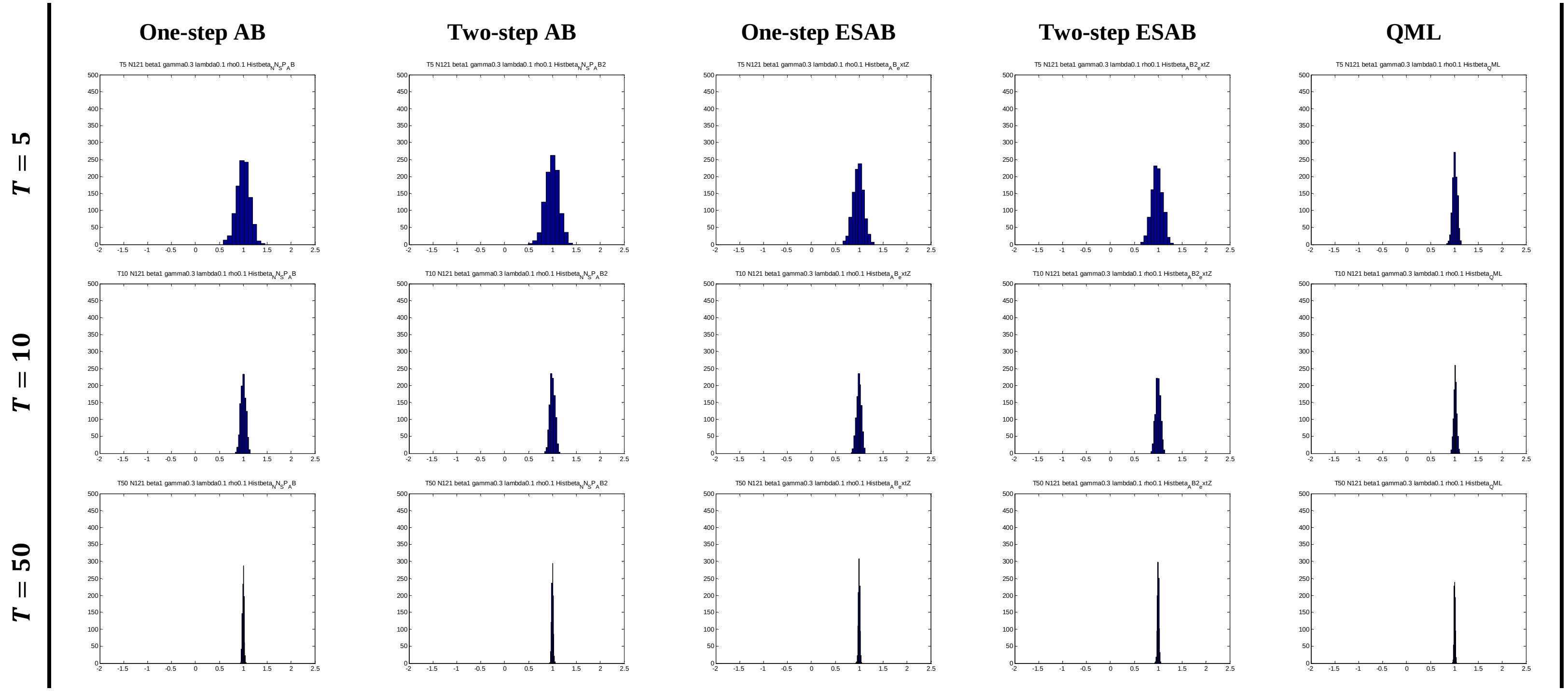


Figure 2.26. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = 0.3$, $\lambda = \varrho = 0.1$, $\beta = 1$, various T and $N = 121$ over 999 iterations



The degree of spatial dependence is then increased first by setting the value of the spatiotemporal parameter to 0.3 and then by doing the same for the spatial autoregressive parameter. Our MC simulations therefore consider the following two settings: $\gamma = \varrho = 0.3$, $\lambda = 0.1$, $\beta = 1$ or $\gamma = \lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$.

In both cases, there is evidence of higher reliability in the estimates of β obtained via the spatial estimation procedures, which are always associated to the smallest bias and variability (Figure 2.31 to Figure 2.36).

Among the spatial estimators, the ESAB outperforms the QML in terms of bias reduction as the time and cross-sectional dimensions increase, although its higher variability leads to better RMSE results for the QML estimator for all T and N (see Table 2.22 to Table 2.24 in Annex 2). Moreover, the QML estimator outperforms the ESAB in the estimation of all the other model coefficients in terms of RMSE, except for the estimation of γ as N increases for all values of T. Parameter γ is therefore confirmed to be the most troublesome for the QML estimator.

When spatial dependence increases, then, the performance of the non-spatial AB estimator is significantly worse than that considered of the spatial procedures, as expected. We also expect that any further increase in the degree of spatial dependence in the data would lead to a worsening of the relative performance of non-spatial estimators. On the other hand, if the reliability of non-spatial estimates was to be assessed according to the results just discussed, one should probably conclude that the bias associated to non-spatial estimates of β (which is in the range of 1.5% to 4% at worst) is not tremendous. What is certainly a drawback of estimating a non-spatial model when a spatial model should be specified instead is that it prevents from estimating the spatial and spatiotemporal effects that are present in the data, thus hiding important spatial spillover effects that may take place.

Figure 2.27. RMSE of γ for various spatial estimators, for $\varrho = \gamma = 0.3$, $\lambda = 0.1$, $\beta = 1$ and various T and N, over 999 iterations

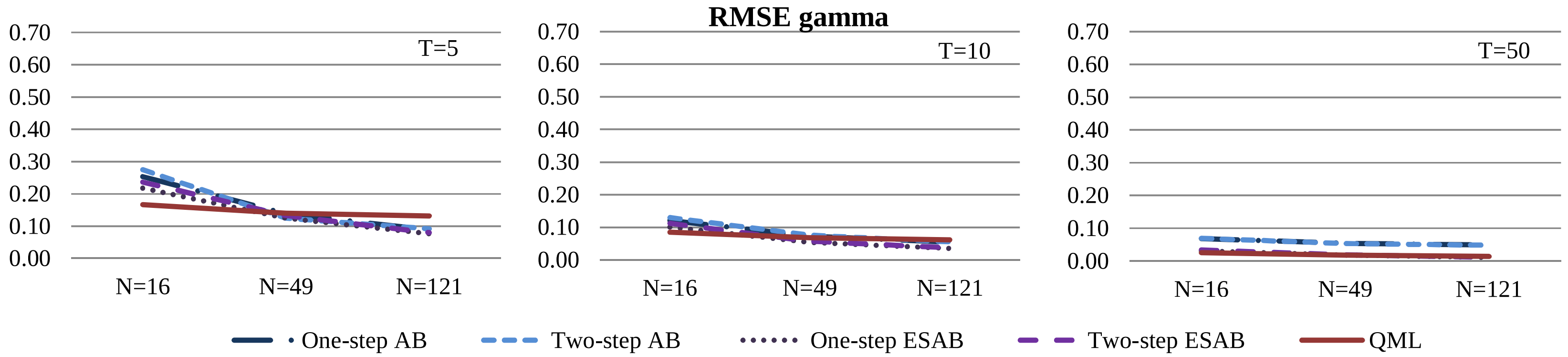


Figure 2.28. RMSE of β for various spatial estimators, for $\varrho = \gamma = 0.3$, $\lambda = 0.1$, $\beta = 1$ and various T and N, over 999 iterations

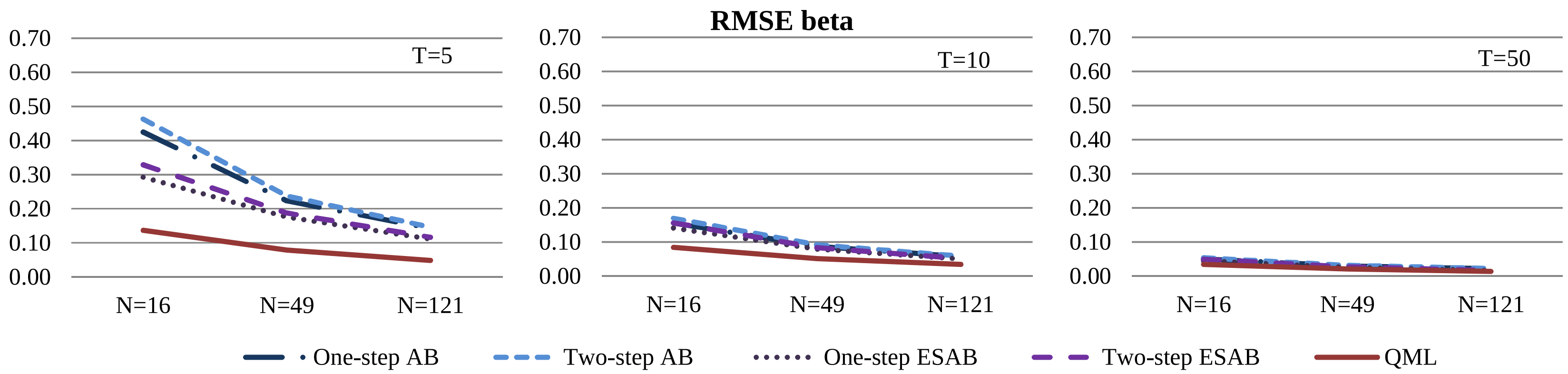


Figure 2.29. RMSE of γ for various spatial estimators, for $\gamma = \lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$ and various T and N, over 999 iterations

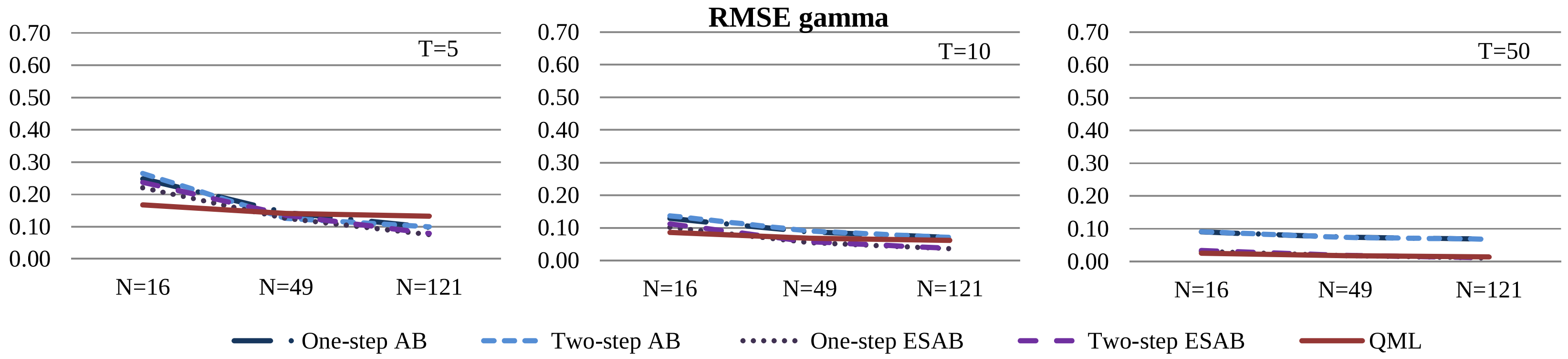


Figure 2.30. RMSE of β for various spatial estimators, for $\gamma = \lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$ and various T and N, over 999 iterations

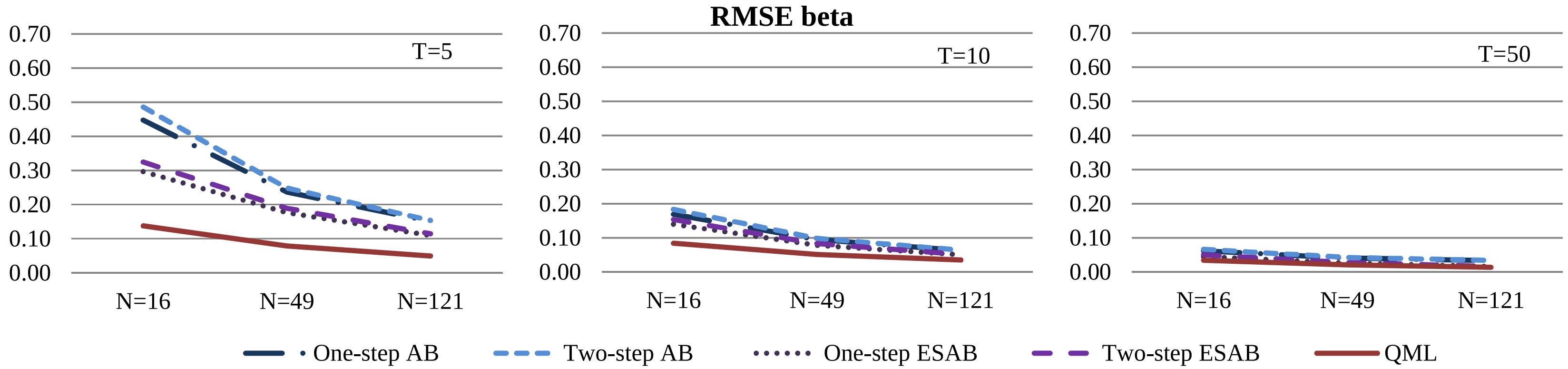


Figure 2.31. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = \varrho = 0.3$, $\lambda = 0.1$, $\beta = 1$, various T and $N = 16$ over 999 iterations

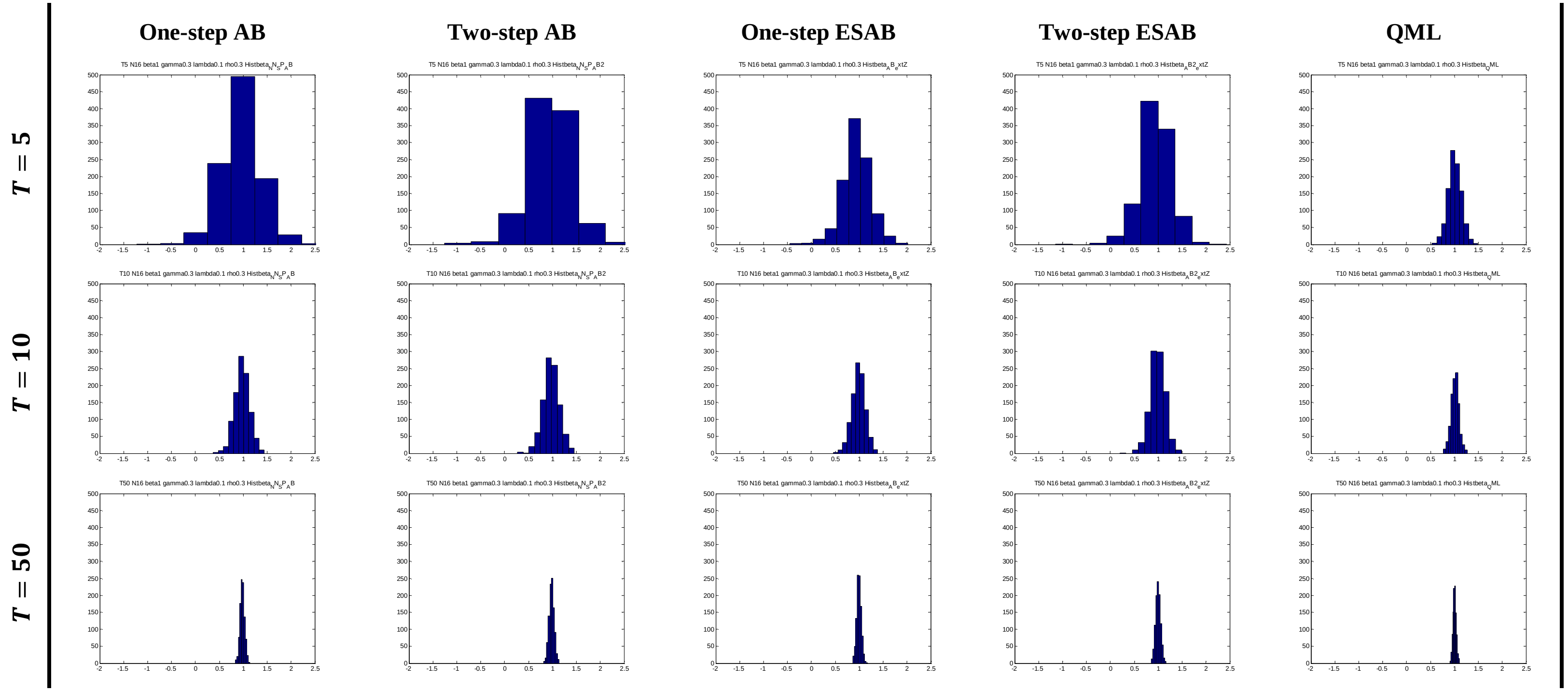


Figure 2.32. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = \varrho = 0.3$, $\lambda = 0.1$, $\beta = 1$, various T and $N = 49$ over 999 iterations

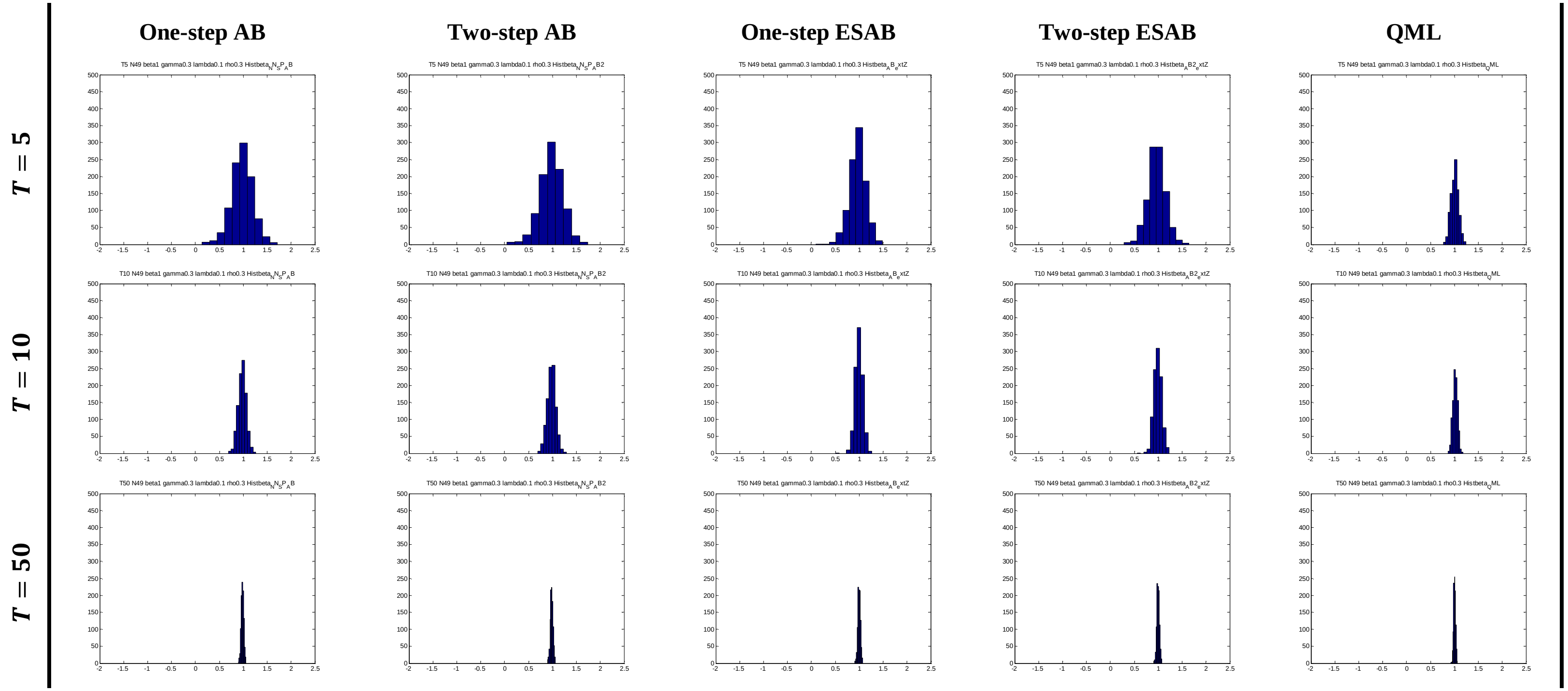


Figure 2.33. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = \varrho = 0.3$, $\lambda = 0.1$, $\beta = 1$, various T and $N = 121$ over 999 iterations

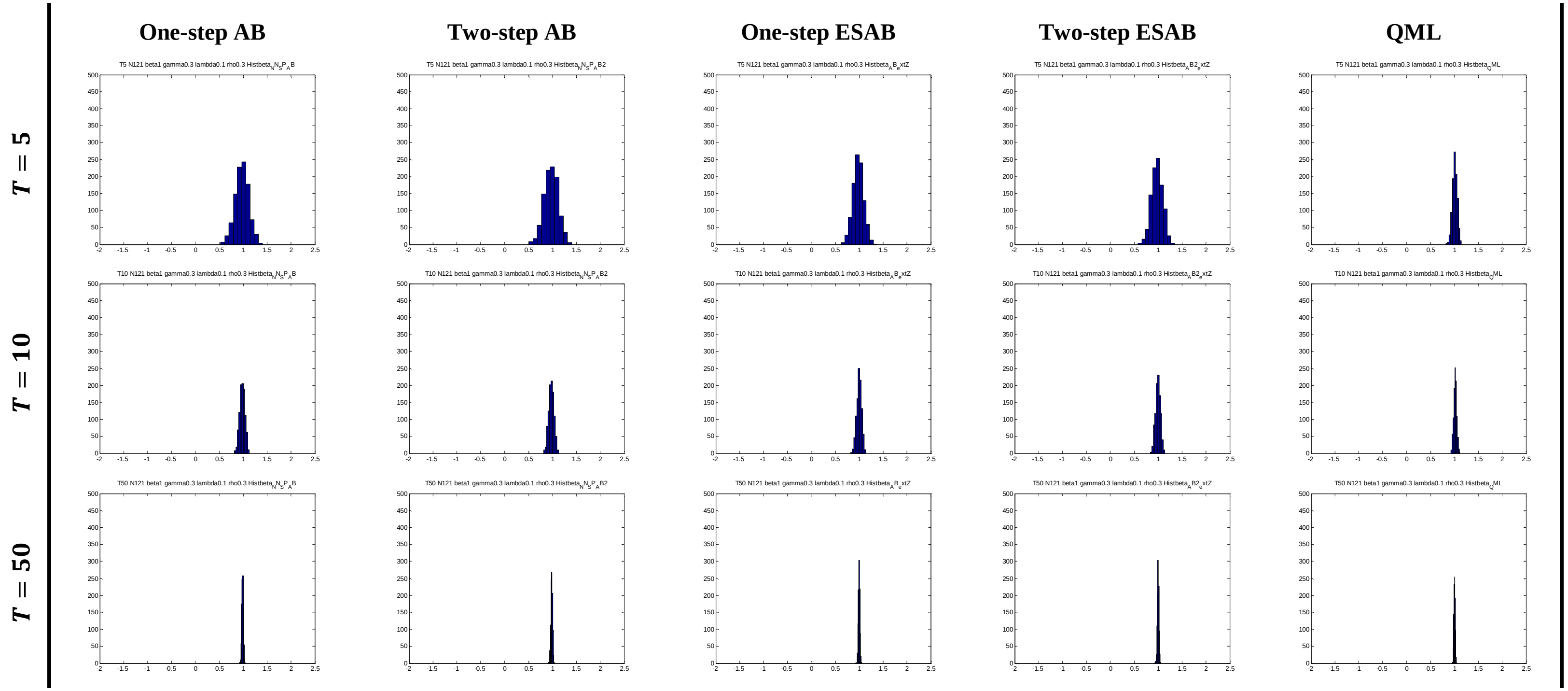


Figure 2.34. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = \lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$, various T and $N = 16$ over 999 iterations

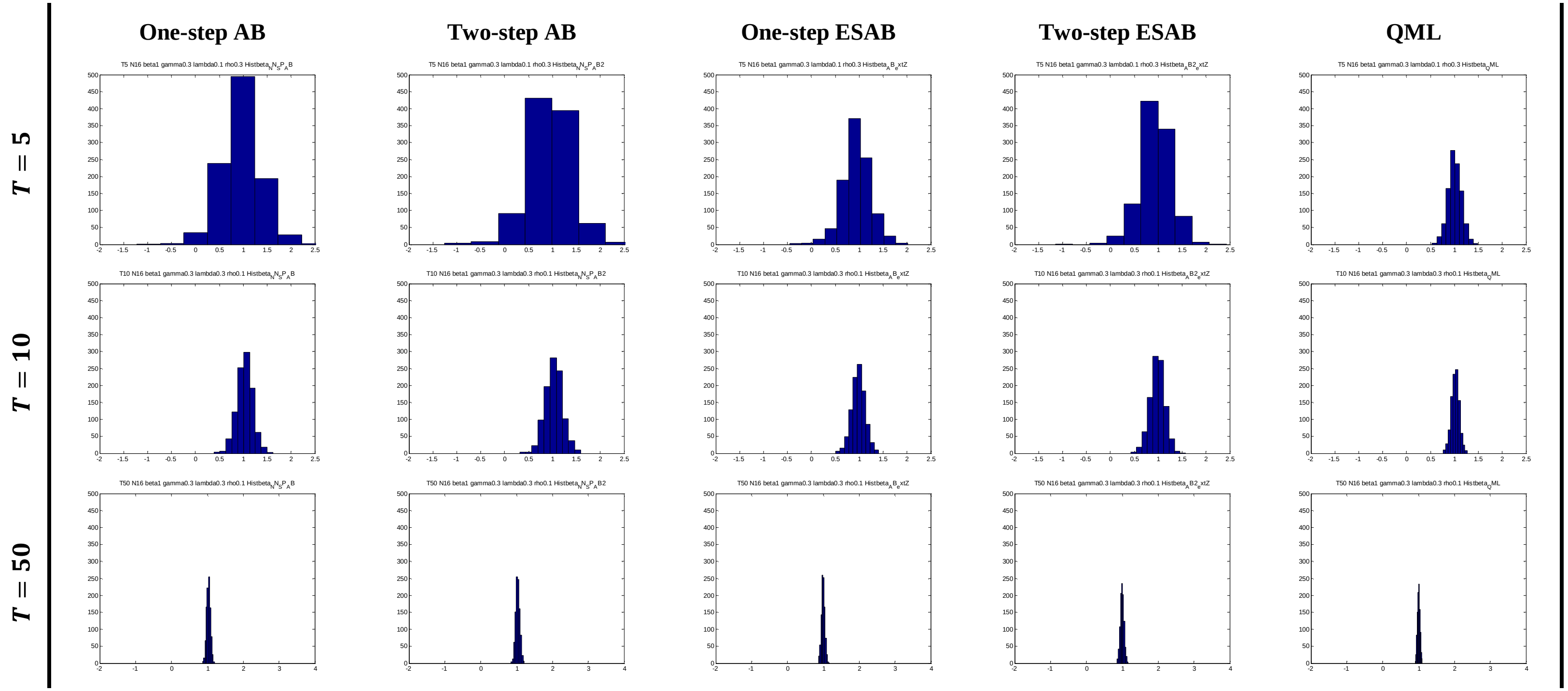


Figure 2.35. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = \lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$, various T and $N = 49$ over 999 iterations

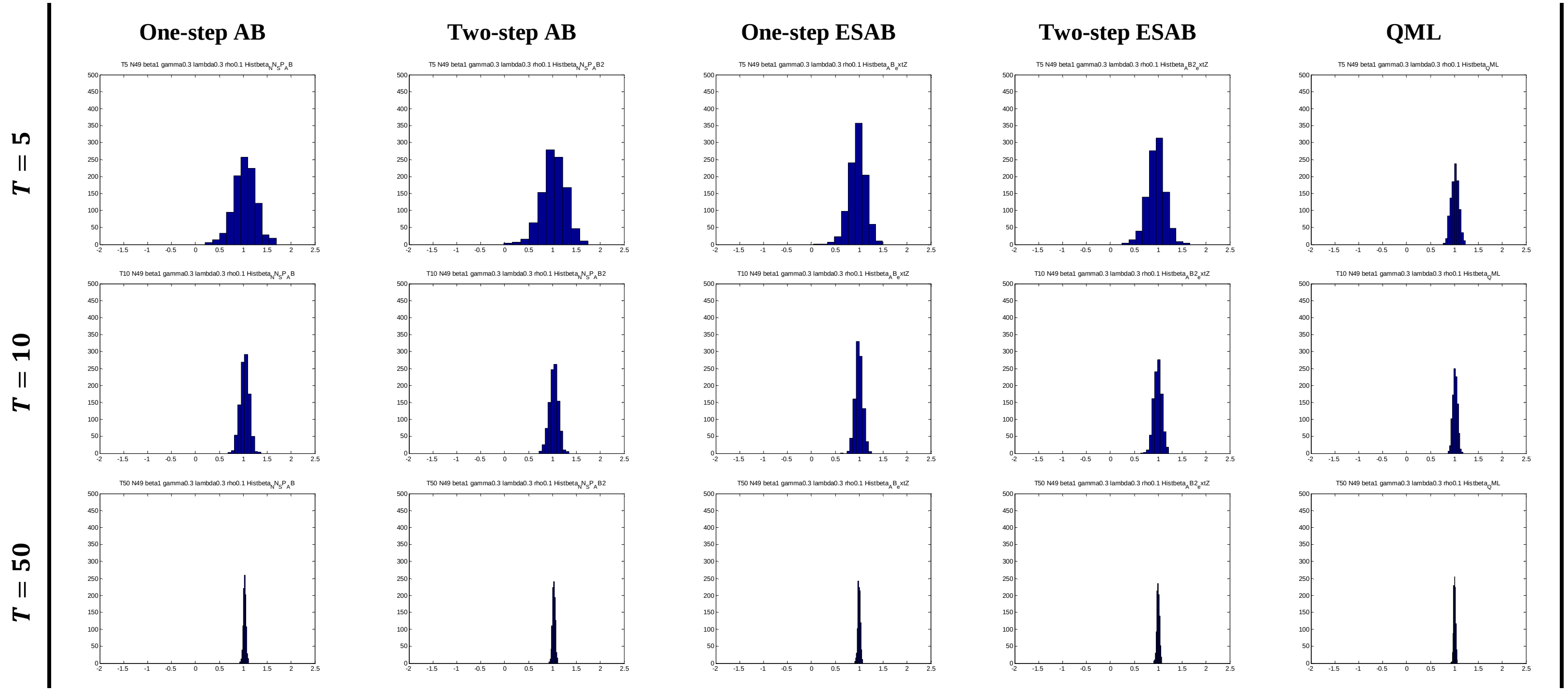
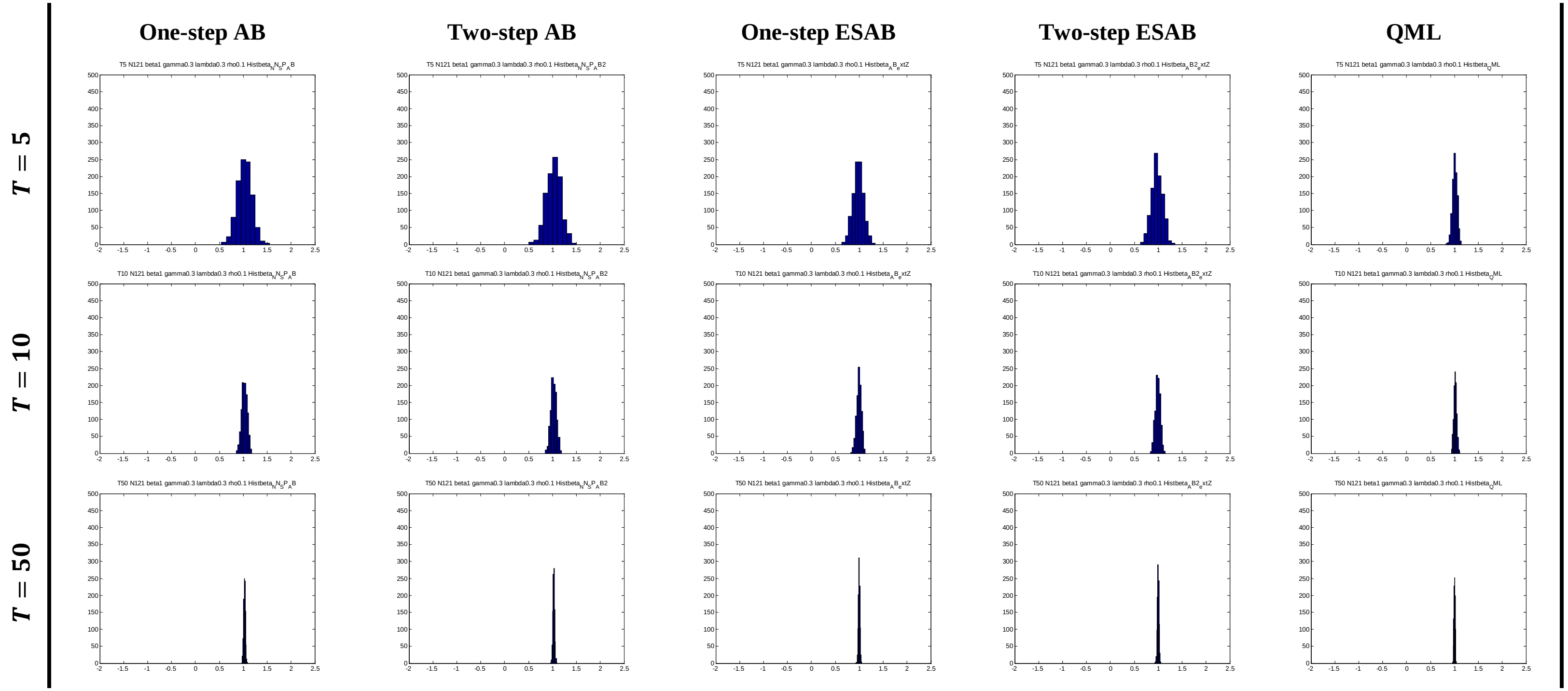


Figure 2.36. Distribution of the estimates of parameter β for various non-spatial and spatial estimators, for $\gamma = \lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$, various T and $N = 121$ over 999 iterations



2.5 Concluding remarks

The empirical researcher who needs to estimate a time-space dynamic panel data model is not facing an easy task, both because of a lack of ready-to-use software routines and because the literature on this kind of models is still quite limited. With the present analysis we aim at providing some evidence on the small sample properties of a number of estimators for time-space dynamic panel data models with fixed effects that the empirical researcher may decide to apply: the popular QML estimator by Yu et al. (2008) and a few spatial and non-spatial difference-GMM estimators. Among these, a GMM-type estimator (ESAB) was proposed and its small sample performance investigated. We do not aim at setting guidelines for the estimation of this kind of models, but nevertheless our analysis can suggest some general comments.

Differently from the artificial world of a MC simulation analysis, empirical researchers are usually not aware of the DGP that characterizes their data. Therefore an accurate exploratory analysis of the data should always be performed in order to identify the more suitable model specification. Given the uncertainty that the researcher faces on the nature of the DGP underlying his data, it is also useful to have the best knowledge possible on the hypothesis on which the consistency of the available estimating procedures relies and their virtues and drawbacks depending on the time and cross-sectional dimensions of the dataset.

As a first remark, it should be noticed that the consistency of both the estimation approaches that were considered in this chapter does not require any assumption on the normality of errors to be verified. Secondly, whether the data are characterized by a stationary or a quasi-unit nature, the RMSE and bias associated to the estimates of the coefficients not surprisingly decrease as the time and cross-sectional dimension of the dataset increase. This is particularly evident with respect to the difference-GMM estimators, that are found consistent for fixed T and growing N (Arellano and Bond 1991). Moreover, the spatial difference-GMM estimator that also includes the spatial lag of the covariates among the instruments (ESAB) performs significantly better than the SAB estimator. A further step of this analysis might therefore concern the testing of the validity of the instruments for the different GMM procedures that have been proposed through a J test.

Focusing on the study of a stationary scenario, presented in section 2.4.1, the QML estimator showed the best small-sample performance in terms of RMSE for all T and N and with respect to almost all the considered coefficients, mainly thanks to a considerably lower variability. Only the estimation of parameter γ, associated to the temporal lag of the dependent variable, appears to be more problematic with the QML estimator, which is generally outperformed by GMM-type estimators as the cross-sectional dimension grows for fixed T.

The relative performance of the considered estimators does not change if we consider a quasi-unit root scenario instead. Only in comparison to the stationary scenario some differences should be highlighted. The parameters that measure time-dependence (γ and ϱ) show higher RMSE in the quasi-unit root context for smaller values of T. As T and N increase, however, the RMSE associated to the parameter that is set equal to 0.58 in the quasi-unit root context becomes smaller than the one observed in the stationary scenario.

The evidence from the present analysis therefore suggests that, QML estimation is probably the safest choice in both situations. However, an element that should not be overlooked, is the fact that we treated only an exogenous covariate in our simulations. A further extension will need to consider the performance of these estimators when an endogenous covariate is included in the model, which is not such an unlikely situation in economics. Some evidence, referred to time-space simultaneous panel data model shows a better performance of GMM estimator with respect to the estimation of the parameters of endogenous covariates (Kukenova and Monteiro 2009), but to our knowledge no evidence is still available with respect to the estimation of time-space dynamic panel data models.

A final observation concerns the risks implied by ignoring the spatial dependence that characterizes the data. In fact, to our knowledge no empirical evidence is available that quantifies the bias that may affect the estimated β coefficient(s) when spatial dependence, although present, is ignored. Our analysis suggests that, when spatial dependence is limited, a non-spatial difference-GMM provides reliable estimates for β, particularly when T is small. This good performance in terms of limited bias, however, is limited by a high variability of the non-spatial estimates. When spatial dependence increases, however, the performance of non-spatial estimators becomes significantly worse than that of spatial estimators, although the bias is not tremendous in absolute terms, especially as T gets larger. In conclusion, the probably time-saving choice of the empirical researcher who ignores the presence of spatial dependence in the data may not necessarily bring to tremendous drawbacks in terms of biased estimates of the parameters of the covariates, although the bias tends to increase as the extent of spatial dependence increases. Nevertheless, the main failure of non-spatial estimation, which should not be neglected is the fact that it prevents the identification and estimation of spatial spillover effects when present, thus considerably limiting the information that can be drawn from the data.

Annex 1

The matrix of instruments described in equation (2.8) is:

$$\boldsymbol{Z}_i = \begin{pmatrix} y_{i1} & \boldsymbol{X}_{i1}' & \boldsymbol{X}_{i2}' & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & y_{i1} & y_{i2} & \boldsymbol{X}_{i1}' & \boldsymbol{X}_{i2}' & \boldsymbol{X}_{i3}' & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{iT-2} & \boldsymbol{X}_{i1}' & \dots & \boldsymbol{X}_{iT-1}' \end{pmatrix} \tag{2.20}$$

The matrix of instruments described in equation (2.9) is:

$$\boldsymbol{Z}_i = \begin{pmatrix} y_{i1} & \boldsymbol{X}_{i1}' & \boldsymbol{X}_{i2}' & \boldsymbol{X}_{i3}' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & y_{i1} & y_{i2} & \boldsymbol{X}_{i1}' & \boldsymbol{X}_{i2}' & \boldsymbol{X}_{i3}' & \boldsymbol{X}_{i4}' & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{iT-2} & \boldsymbol{X}_{i1}' & \dots & \boldsymbol{X}_{iT}' \end{pmatrix} \tag{2.21}$$

The matrix of instruments when only certain lags (e.g. only lags up to 3 time periods) are used and the covariates are sequentially exogenous is equal to:

$$\boldsymbol{Z}_i = \begin{pmatrix} y_{i1} & \boldsymbol{X}_{i1}' & \boldsymbol{X}_{i2}' & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & y_{i1} & y_{i2} & \boldsymbol{X}_{i1}' & \boldsymbol{X}_{i2}' & \boldsymbol{X}_{i3}' & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & y_{iT-4} & y_{iT-3} & y_{iT-2} & \boldsymbol{X}_{iT-3}' & \boldsymbol{X}_{iT-2}' & \boldsymbol{X}_{iT-1}' \end{pmatrix} \tag{2.22}$$

Annex 2

Table 2.4. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 5$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML
Gamma=0.2	**Mean**	0.1119	0.1111	0.0821	0.0729	0.0688	0.1626	0.1623	0.1585	0.1588	0.0762	0.1844	0.1830	0.1745	0.1742	0.0784
	Bias	-0.0882	-0.0889	-0.1179	-0.1271	-0.1312	-0.0374	-0.0377	-0.0415	-0.0412	-0.1238	-0.0156	-0.0170	-0.0255	-0.0258	-0.1216
	SD	0.1906	0.2093	0.2354	0.2639	0.0910	0.1130	0.1212	0.1498	0.1644	0.0509	0.0723	0.0757	0.0986	0.1014	0.0317
	RMSE	0.2100	0.2274	0.2633	0.2929	**0.1597**	**0.1190**	0.1269	0.1555	0.1695	0.1339	**0.0740**	0.0776	0.1018	0.1046	0.1256
Lambda=0.2	**Mean**	0.2978	0.2923	0.3134	0.3158	0.1908	0.2563	0.2589	0.3086	0.3061	0.1911	0.2323	0.2339	0.3113	0.3146	0.1961
	Bias	0.0978	0.0923	0.1134	0.1158	-0.0092	0.0563	0.0589	0.1086	0.1061	-0.0089	0.0323	0.0339	0.1113	0.1146	-0.0039
	SD	0.3635	0.4058	0.5724	0.6189	0.1148	0.2798	0.3024	0.5698	0.6050	0.0738	0.1993	0.2103	0.5932	0.5966	0.0488
	RMSE	0.3764	0.4162	0.5835	0.6297	**0.1151**	0.2855	0.3081	0.5801	0.6142	**0.0743**	0.2019	0.2130	0.6036	0.6075	**0.0490**
Rho=0.2	**Mean**	0.1837	0.1857	0.2168	0.2225	0.2052	0.1849	0.1887	0.1611	0.1577	0.2053	0.1914	0.1921	0.1643	0.1627	0.2077
	Bias	-0.0163	-0.0143	0.0168	0.0225	0.0052	-0.0151	-0.0113	-0.0389	-0.0423	0.0053	-0.0086	-0.0079	-0.0358	-0.0373	0.0077
	SD	0.2982	0.3314	0.5088	0.5649	0.1446	0.2007	0.2184	0.4319	0.4645	0.0889	0.1285	0.1320	0.3978	0.4102	0.0597
	RMSE	0.2987	0.3317	0.5090	0.5654	**0.1447**	0.2013	0.2187	0.4337	0.4665	**0.0891**	0.1288	0.1322	0.3994	0.4119	**0.0602**
Beta=1	**Mean**	0.9563	0.9556	0.9173	0.9080	1.0131	0.9719	0.9716	0.9676	0.9682	1.0124	0.9853	0.9806	0.9720	0.9698	1.0132
	Bias	-0.0437	-0.0444	-0.0828	-0.0920	0.0131	-0.0281	-0.0284	-0.0324	-0.0318	0.0124	-0.0147	-0.0194	-0.0280	-0.0302	0.0132
	SD	0.2823	0.3136	0.4130	0.4644	0.1373	0.1667	0.1796	0.2703	0.2992	0.0787	0.1057	0.1099	0.1757	0.1822	0.0480
	RMSE	0.2857	0.3168	0.4212	0.4735	**0.1380**	0.1690	0.1818	0.2723	0.3009	**0.0797**	0.1067	0.1116	0.1779	0.1847	**0.0497**

Table 2.5. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 10$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					**N=49**					**N=121**				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.2	**Mean**	0.1778	0.1797	0.1698	0.1727	0.1382	0.1928	0.1953	0.1879	0.1908	0.1425	0.1956	0.1968	0.1909	0.1918	0.1438
	Bias	-0.0222	-0.0203	-0.0302	-0.0273	-0.0618	-0.0072	-0.0047	-0.0121	-0.0092	-0.0575	-0.0044	-0.0032	-0.0091	-0.0082	-0.0562
	SD	0.0950	0.1063	0.1176	0.1295	0.0571	0.0525	0.0562	0.0696	0.0740	0.0331	0.0347	0.0356	0.0492	0.0505	0.0204
	RMSE	0.0975	0.1082	0.1214	0.1323	**0.0841**	**0.0530**	0.0564	0.0707	0.0746	0.0663	**0.0349**	0.0357	0.0500	0.0511	0.0598
Lambda=0.2	**Mean**	0.2329	0.2338	0.3076	0.2990	0.1951	0.2173	0.2178	0.3161	0.3159	0.1980	0.2093	0.2099	0.2814	0.2863	0.2005
	Bias	0.0329	0.0338	0.1076	0.0990	-0.0049	0.0173	0.0178	0.1161	0.1159	-0.0020	0.0093	0.0099	0.0814	0.0863	0.0005
	SD	0.2139	0.2336	0.5237	0.5796	0.0797	0.1347	0.1435	0.5189	0.5530	0.0485	0.0921	0.0950	0.5147	0.5236	0.0319
	RMSE	0.2164	0.2361	0.5346	0.5880	**0.0798**	0.1358	0.1446	0.5317	0.5651	**0.0486**	0.0925	0.0955	0.5211	0.5307	**0.0319**
Rho=0.2	**Mean**	0.1913	0.1926	0.1612	0.1580	0.2029	0.1951	0.1925	0.1457	0.1437	0.2030	0.1998	0.1983	0.1664	0.1618	0.2051
	Bias	-0.0087	-0.0074	-0.0388	-0.0420	0.0029	-0.0049	-0.0075	-0.0543	-0.0563	0.0030	-0.0002	-0.0017	-0.0336	-0.0382	0.0051
	SD	0.1529	0.1699	0.3655	0.3909	0.0918	0.0981	0.1059	0.2932	0.3149	0.0573	0.0651	0.0677	0.2812	0.2882	0.0386
	RMSE	0.1531	0.1700	0.3676	0.3931	**0.0918**	0.0982	0.1061	0.2982	0.3199	**0.0574**	0.0651	0.0677	0.2832	0.2908	**0.0389**
Beta=1	**Mean**	0.9881	0.9886	0.9784	0.9801	1.0170	0.9950	0.9923	0.9895	0.9913	1.0185	0.9971	0.9966	0.9922	0.9924	1.0185
	Bias	-0.0119	-0.0114	-0.0216	-0.0199	0.0170	-0.0050	-0.0077	-0.0105	-0.0087	0.0185	-0.0029	-0.0034	-0.0078	-0.0076	0.0185
	SD	0.1371	0.1499	0.1739	0.1969	0.0826	0.0771	0.0817	0.0996	0.1045	0.0478	0.0491	0.0509	0.0702	0.0720	0.0297
	RMSE	0.1376	0.1504	0.1753	0.1979	**0.0844**	0.0773	0.0821	0.1002	0.1049	**0.0512**	0.0492	0.0510	0.0707	0.0724	**0.0350**

Table 2.6. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 50$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML
Gamma=0.2	Mean	0.1975	0.1984	0.1931	0.1925	0.1881	0.1986	0.1994	0.1946	0.1953	0.1879	0.1996	0.2000	0.1950	0.1953	0.1889
	Bias	-0.0025	-0.0016	-0.0069	-0.0075	-0.0119	-0.0014	-0.0006	-0.0054	-0.0047	-0.0121	-0.0004	0.0000	-0.0050	-0.0047	-0.0111
	SD	0.0307	0.0342	0.0457	0.0499	0.0232	0.0180	0.0192	0.0317	0.0332	0.0138	0.0110	0.0114	0.0252	0.0256	0.0085
	RMSE	0.0308	0.0342	0.0462	0.0505	**0.0261**	**0.0180**	0.0193	0.0321	0.0335	0.0184	**0.0110**	0.0114	0.0257	0.0261	0.0139
Lambda=0.2	Mean	0.2062	0.2046	0.2804	0.2951	0.2013	0.2013	0.2020	0.2662	0.2658	0.2012	0.2012	0.2009	0.2829	0.2831	0.2012
	Bias	0.0062	0.0046	0.0804	0.0951	0.0013	0.0013	0.0020	0.0662	0.0658	0.0012	0.0012	0.0009	0.0829	0.0831	0.0012
	SD	0.0733	0.0811	0.4525	0.4896	0.0343	0.0438	0.0469	0.4588	0.4727	0.0201	0.0303	0.0310	0.4023	0.4088	0.0142
	RMSE	0.0736	0.0812	0.4596	0.4987	**0.0344**	0.0438	0.0469	0.4635	0.4773	**0.0201**	0.0303	0.0310	0.4107	0.4171	**0.0142**
Rho=0.2	Mean	0.1975	0.1970	0.1571	0.1493	0.2011	0.1985	0.1982	0.1640	0.1638	0.2003	0.2001	0.2001	0.1564	0.1561	0.2006
	Bias	-0.0025	-0.0030	-0.0429	-0.0507	0.0011	-0.0015	-0.0018	-0.0360	-0.0362	0.0003	0.0001	0.0001	-0.0436	-0.0439	0.0006
	SD	0.0551	0.0605	0.2472	0.2690	0.0394	0.0332	0.0352	0.2500	0.2579	0.0237	0.0224	0.0231	0.2146	0.2184	0.0159
	RMSE	0.0551	0.0606	0.2509	0.2737	**0.0394**	0.0332	0.0353	0.2526	0.2605	**0.0237**	0.0224	0.0231	0.2190	0.2228	**0.0159**
Beta=1	Mean	0.9976	0.9983	0.9922	0.9924	1.0052	1.0001	1.0005	0.9963	0.9965	1.0061	1.0001	1.0000	0.9946	0.9946	1.0053
	Bias	-0.0024	-0.0017	-0.0078	-0.0076	0.0052	0.0001	0.0005	-0.0037	-0.0035	0.0061	0.0001	0.0000	-0.0054	-0.0054	0.0053
	SD	0.0454	0.0498	0.0629	0.0697	0.0338	0.0261	0.0277	0.0450	0.0465	0.0200	0.0164	0.0168	0.0328	0.0335	0.0127
	RMSE	0.0455	0.0498	0.0634	0.0701	**0.0342**	0.0261	0.0277	0.0452	0.0466	**0.0209**	0.0164	0.0168	0.0333	0.0339	**0.0138**

Figure 2.37. Distribution of the estimates of parameter γ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N $= 16$ over 999 iterations

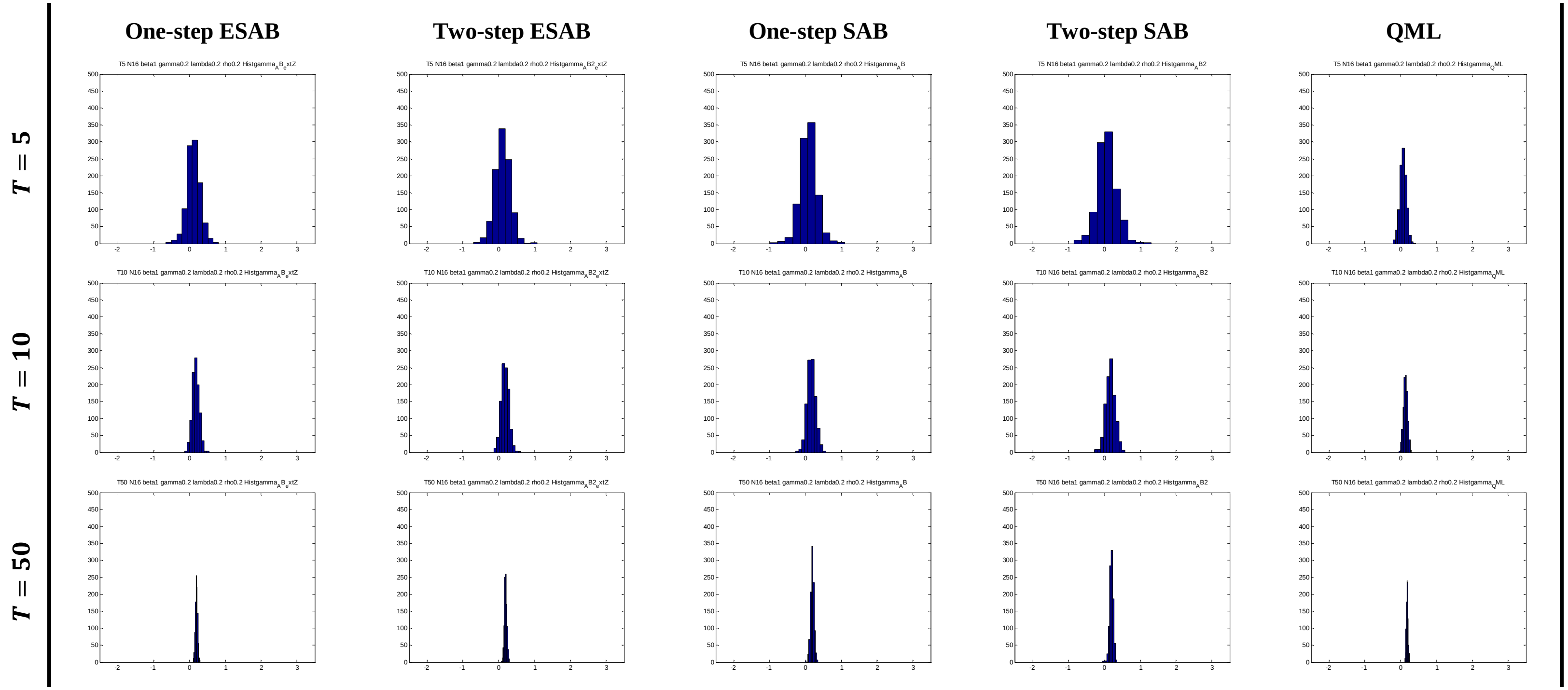


Figure 2.38. Distribution of the estimates of parameter γ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N = 49 over 999 iterations

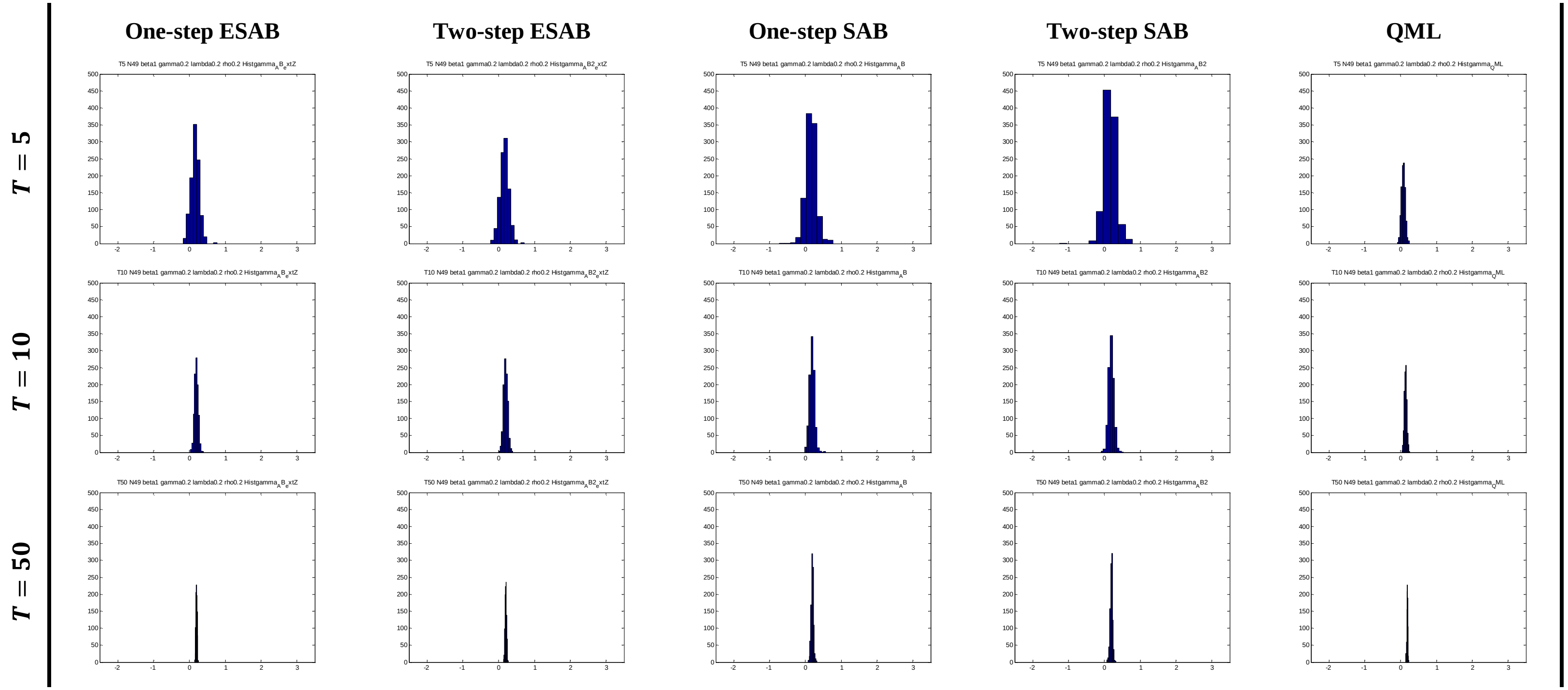


Figure 2.39. Distribution of the estimates of parameter γ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and $\mathrm{N} = 121$ over 999 iterations

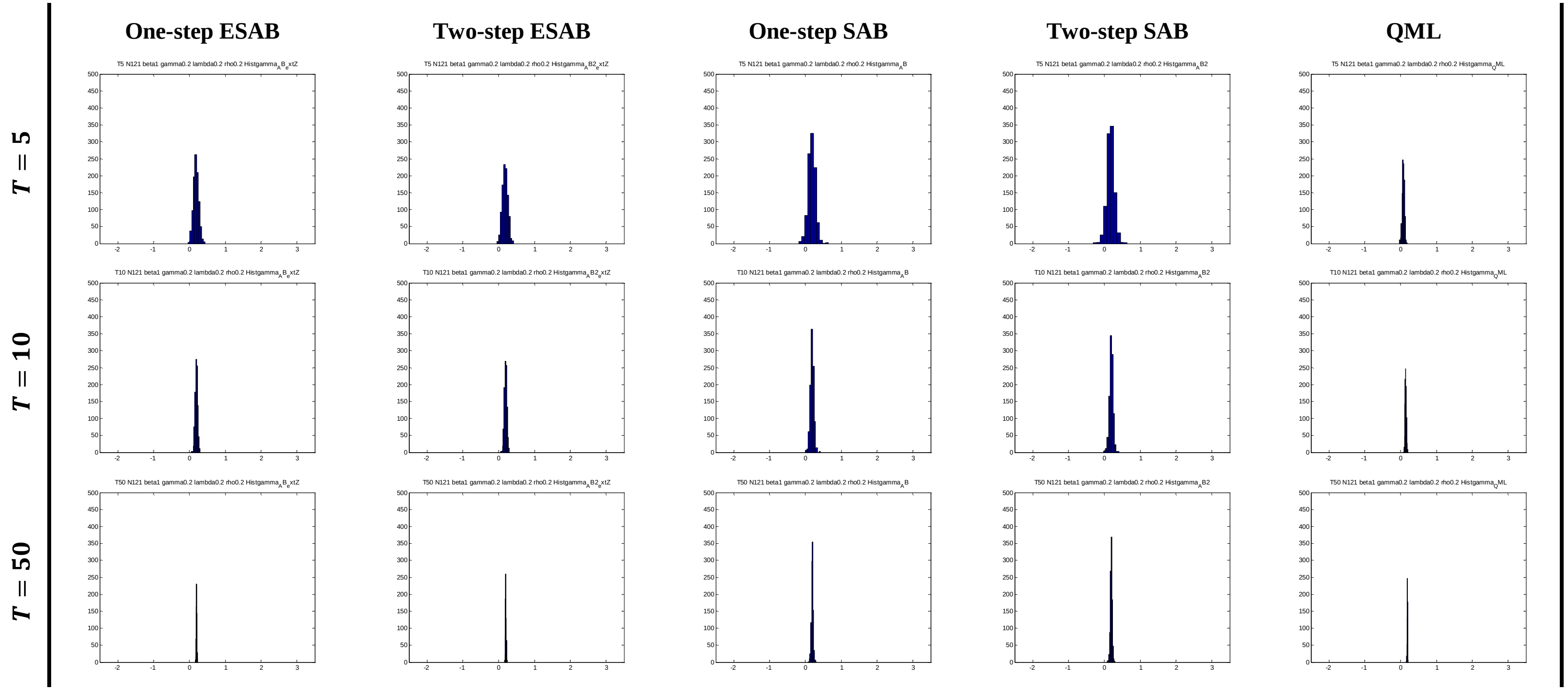


Figure 2.40. Distribution of the estimates of parameter λ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N = 16 over 999 iterations

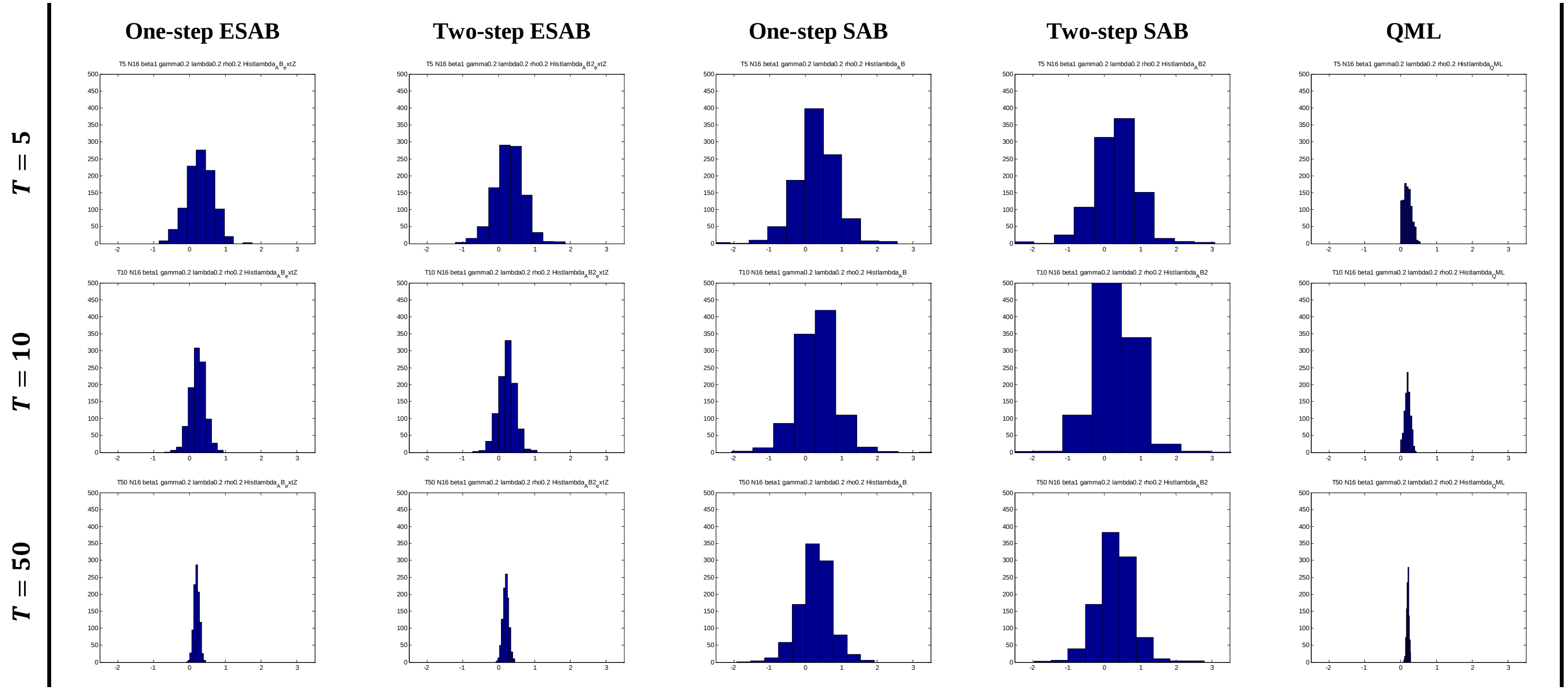


Figure 2.41. Distribution of the estimates of parameter λ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N = 49 over 999 iterations

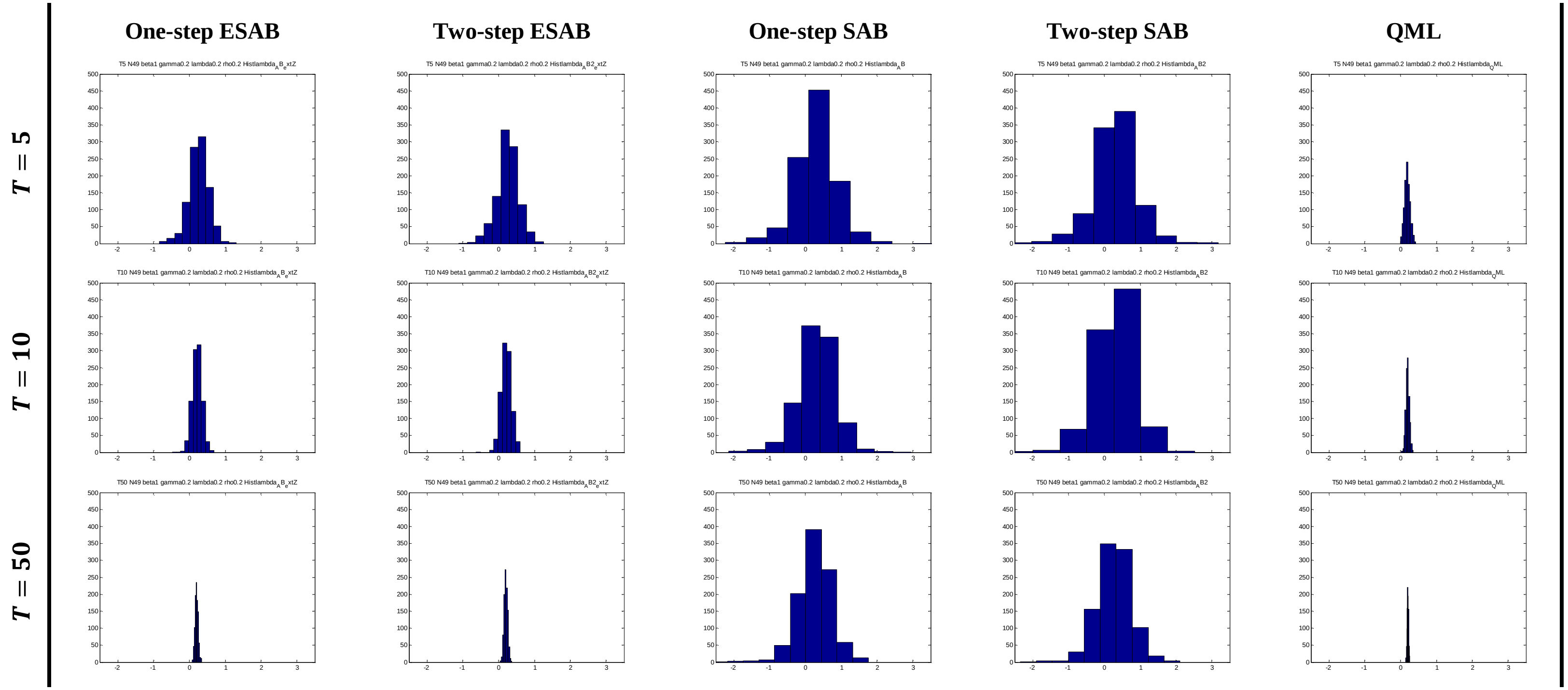


Figure 2.42. Distribution of the estimates of parameter λ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and $N = 121$ over 999 iterations

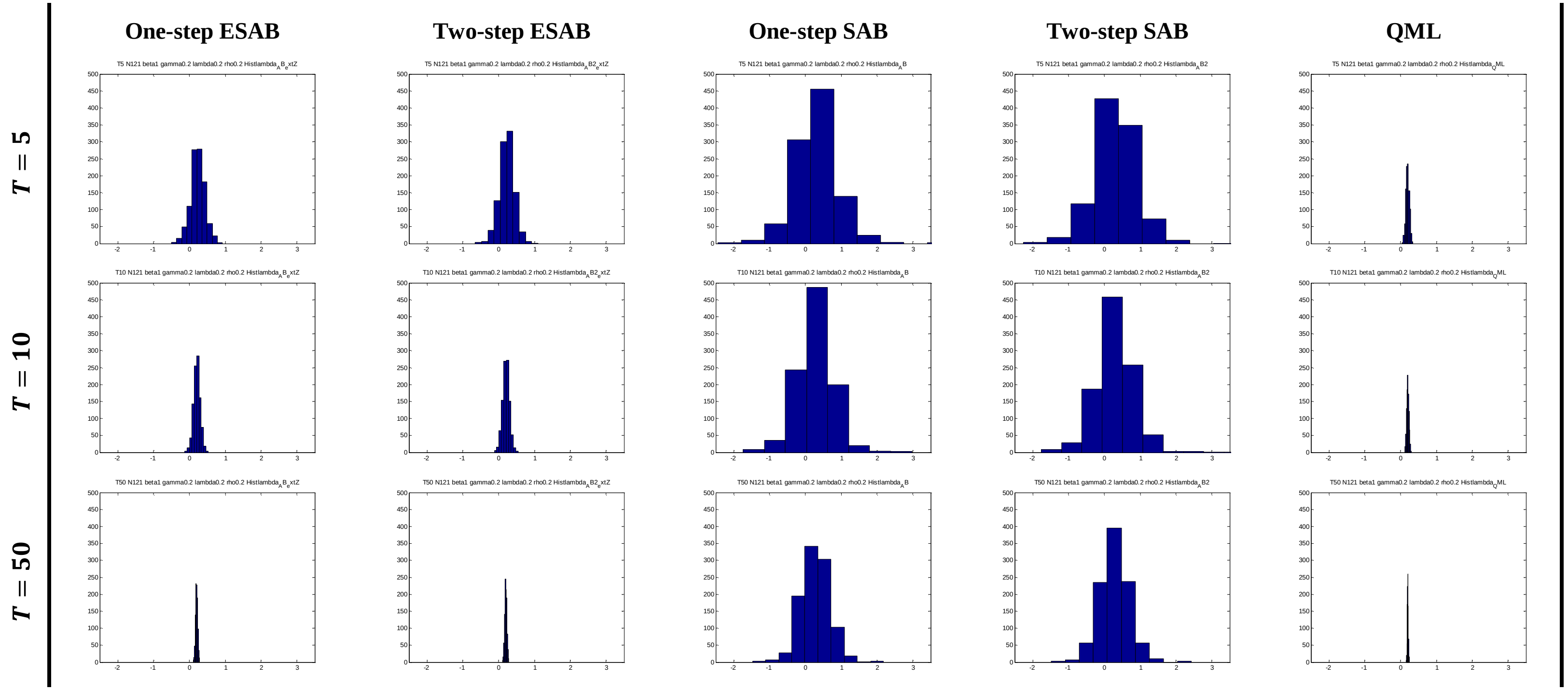


Figure 2.43. Distribution of the estimates of parameter ϱ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N $= 16$ over 999 iterations

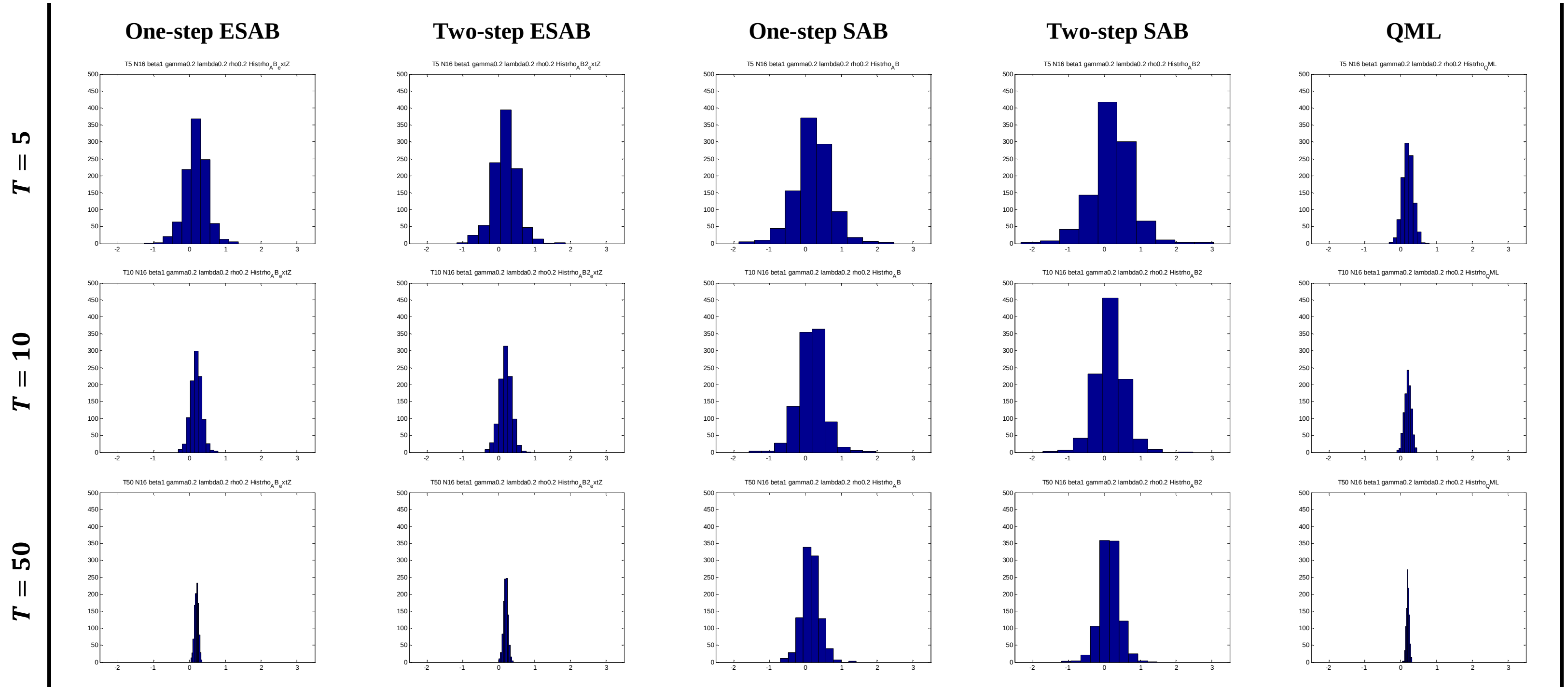


Figure 2.44. Distribution of the estimates of parameter ϱ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N = 49 over 999 iterations

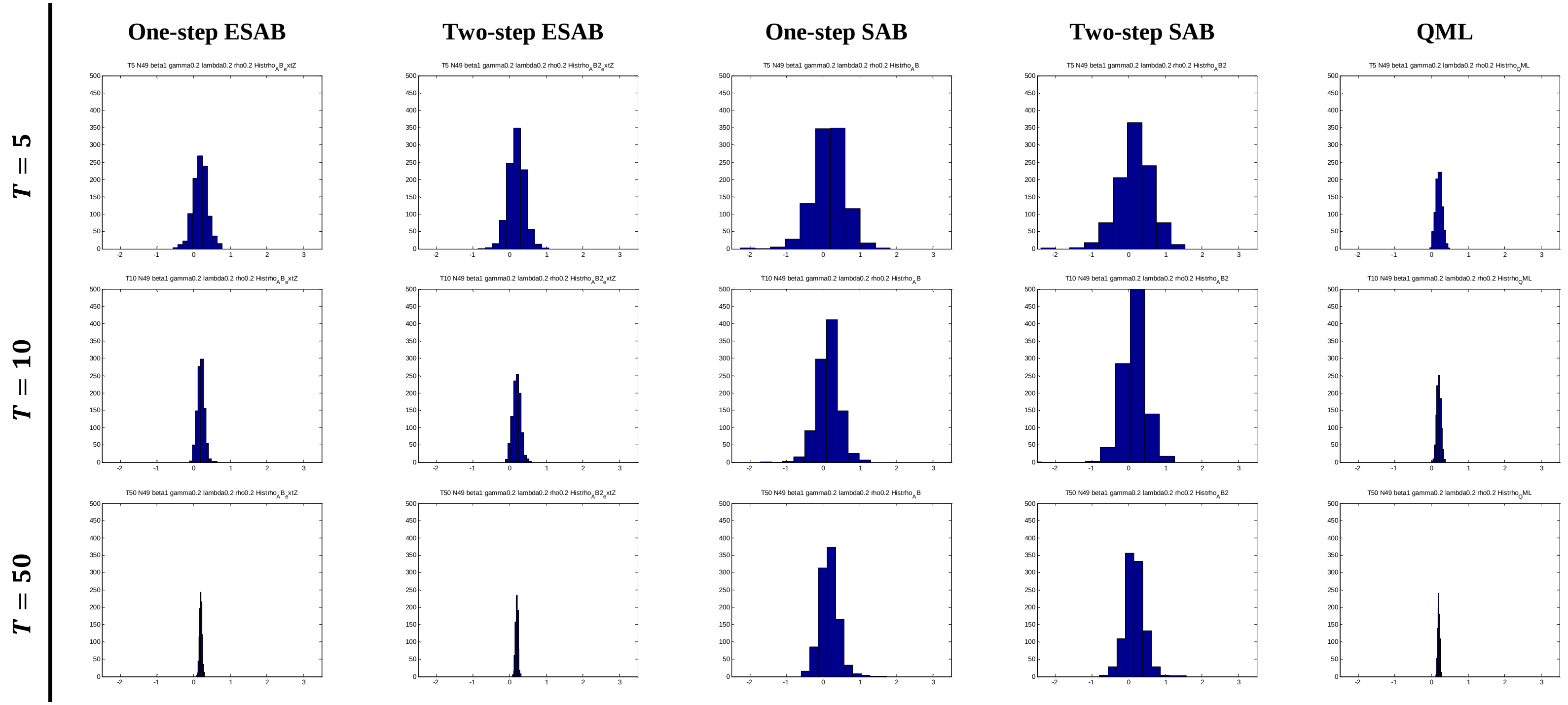


Figure 2.45. Distribution of the estimates of parameter ϱ for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and $N = 121$ over 999 iterations

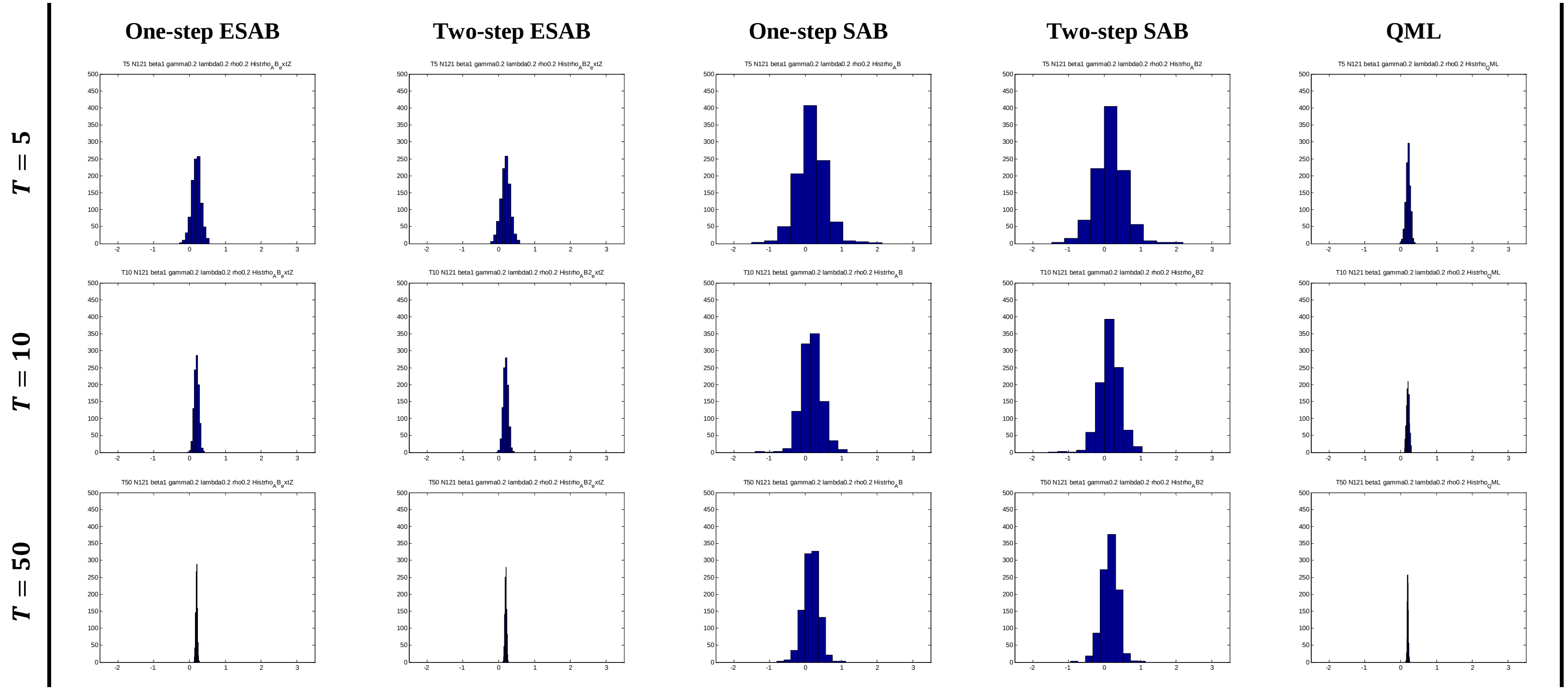


Figure 2.46. Distribution of the estimates of parameter β for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N = 16 over 999 iterations

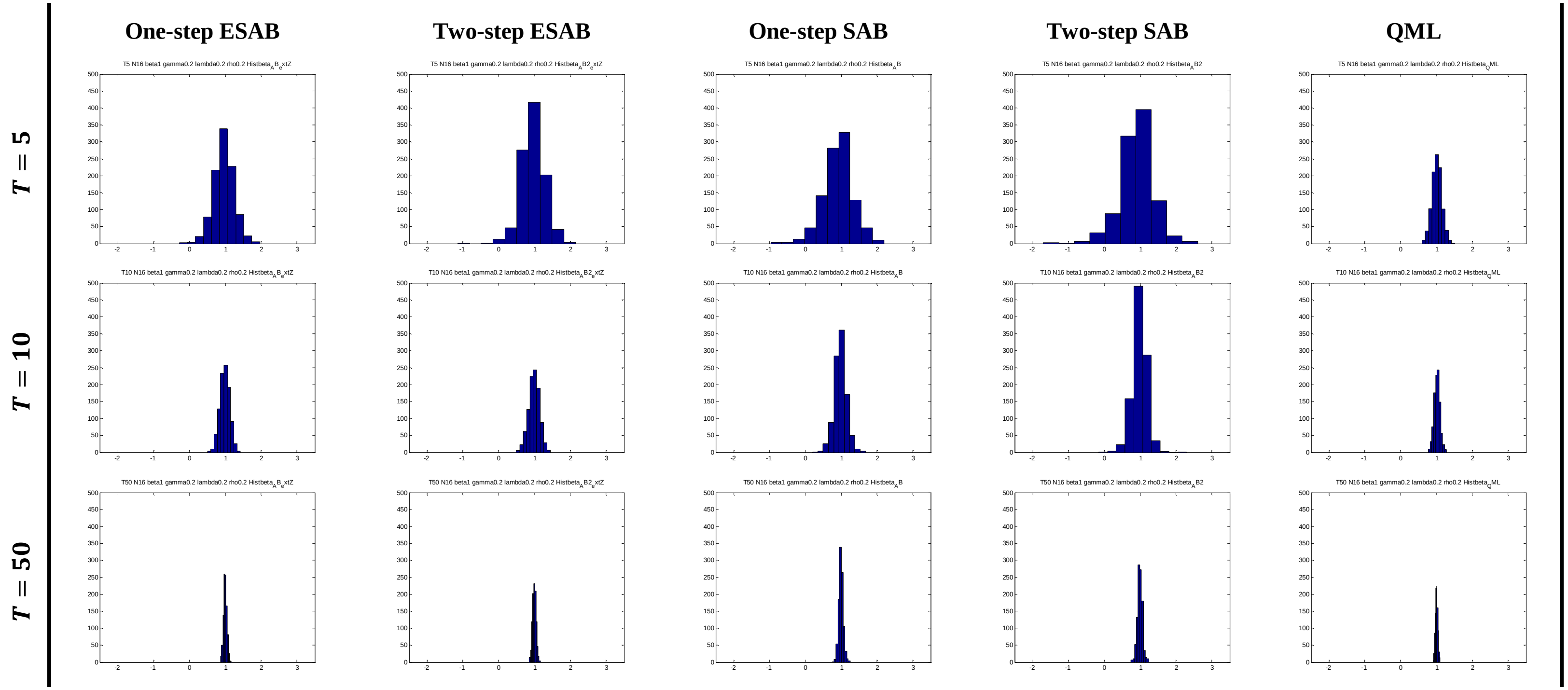


Figure 2.47. Distribution of the estimates of parameter β for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and N $= 49$ over 999 iterations

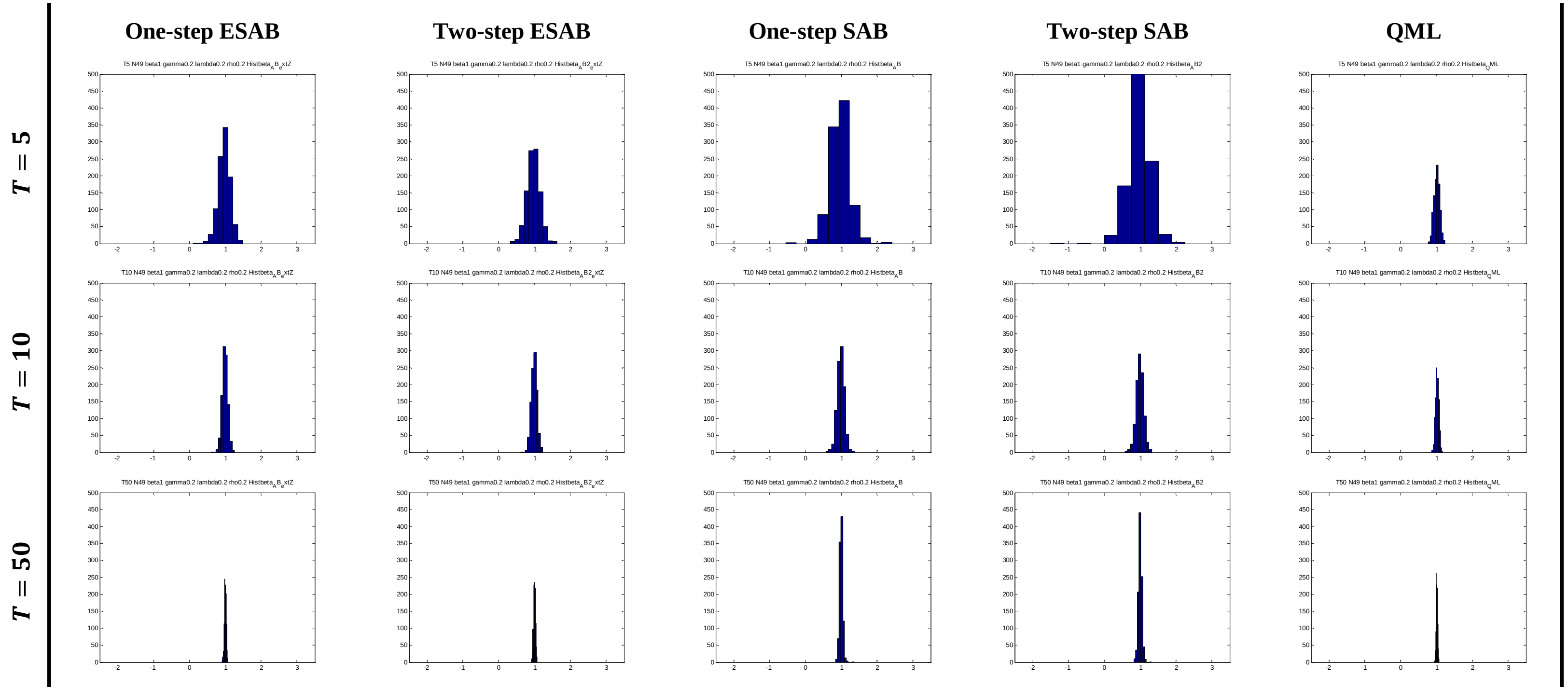


Figure 2.48. Distribution of the estimates of parameter β for various spatial estimators, for $\gamma = \lambda = \varrho = 0.2$, $\beta = 1$, various T and $N = 121$ over 999 iterations

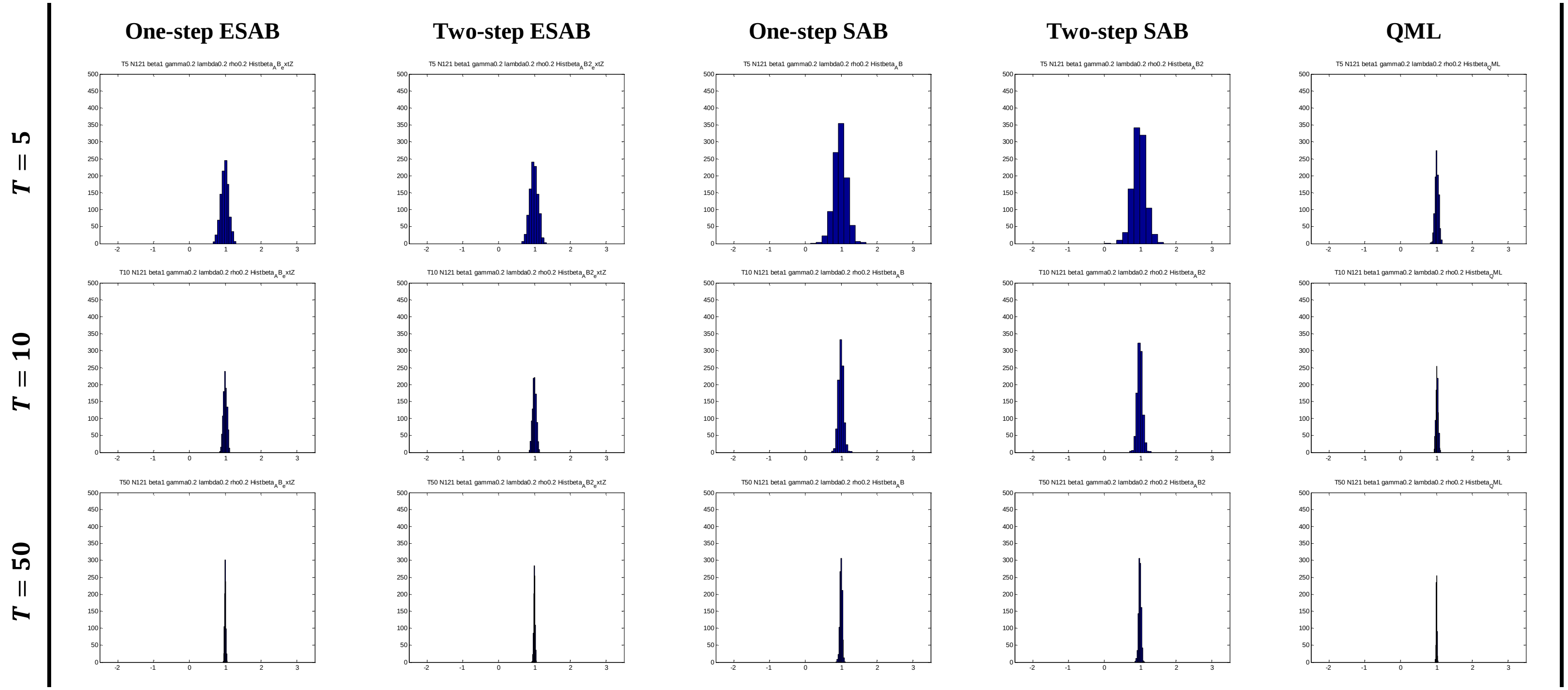


Table 2.7. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $\mathbf{T = 5}$, $\boldsymbol{\gamma = 0.4}$, $\boldsymbol{\lambda = 0.2}$, $\boldsymbol{\varrho = 0.2}$, $\boldsymbol{\beta = 1}$ and various N, over 999 iterations

		N=16					**N=49**					**N=121**				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.4	**Mean**	0.2908	0.2895	0.2512	0.2407	0.2530	0.3528	0.3517	0.3423	0.3414	0.2629	0.3801	0.3780	0.3656	0.3639	0.2654
	Bias	-0.1092	-0.1105	-0.1488	-0.1593	-0.1470	-0.0472	-0.0483	-0.0577	-0.0586	-0.1371	-0.0199	-0.0220	-0.0344	-0.0361	-0.1346
	SD	0.2017	0.2233	0.2559	0.2860	0.0913	0.1240	0.1333	0.1671	0.1806	0.0507	0.0760	0.0794	0.1115	0.1156	0.0311
	RMSE	0.2294	0.2491	0.2961	0.3274	**0.1730**	**0.1326**	0.1418	0.1768	0.1899	0.1462	**0.0785**	0.0824	0.1167	0.1211	0.1382
Lambda=0.2	**Mean**	0.3179	0.3146	0.3610	0.3644	0.1849	0.2726	0.2767	0.3694	0.3624	0.1854	0.2405	0.2425	0.3480	0.3522	0.1900
	Bias	0.1179	0.1146	0.1610	0.1644	-0.0151	0.0726	0.0767	0.1694	0.1624	-0.0146	0.0405	0.0425	0.1480	0.1522	-0.0100
	SD	0.3691	0.4082	0.5711	0.6067	0.1140	0.3016	0.3201	0.5476	0.5847	0.0741	0.2273	0.2389	0.5758	0.5920	0.0493
	RMSE	0.3874	0.4240	0.5933	0.6285	**0.1150**	0.3102	0.3292	0.5732	0.6068	**0.0755**	0.2309	0.2426	0.5945	0.6113	**0.0503**
Rho=0.2	**Mean**	0.1586	0.1577	0.1469	0.1539	0.2105	0.1706	0.1703	0.0969	0.0999	0.2117	0.1826	0.1825	0.0916	0.0908	0.2119
	Bias	-0.0414	-0.0423	-0.0531	-0.0461	0.0105	-0.0294	-0.0297	-0.1031	-0.1001	0.0117	-0.0174	-0.0175	-0.1084	-0.1092	0.0119
	SD	0.3040	0.3393	0.5371	0.6088	0.1462	0.2040	0.2177	0.4860	0.5190	0.0893	0.1389	0.1430	0.4906	0.5115	0.0596
	RMSE	0.3069	0.3419	0.5397	0.6106	**0.1466**	0.2061	0.2197	0.4968	0.5285	**0.0901**	0.1400	0.1441	0.5024	0.5230	**0.0608**
Beta=1	**Mean**	0.9344	0.9368	0.8781	0.8713	1.0038	0.9590	0.9577	0.9443	0.9438	1.0032	0.9804	0.9760	0.9586	0.9549	1.0045
	Bias	-0.0656	-0.0632	-0.1219	-0.1287	0.0038	-0.0410	-0.0423	-0.0557	-0.0562	0.0032	-0.0196	-0.0240	-0.0414	-0.0451	0.0045
	SD	0.3002	0.3319	0.4492	0.4880	0.1370	0.1813	0.1946	0.2835	0.3139	0.0788	0.1141	0.1201	0.1868	0.1915	0.0478
	RMSE	0.3072	0.3379	0.4655	0.5047	**0.1370**	0.1859	0.1992	0.2889	0.3189	**0.0789**	0.1158	0.1225	0.1914	0.1968	**0.0480**

Table 2.8. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 10$, $\gamma = 0.4$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.4	**Mean**	0.3730	0.3744	0.3597	0.3613	0.3337	0.3910	0.3926	0.3826	0.3846	0.3385	0.3945	0.3954	0.3883	0.3887	0.3397
	Bias	-0.0270	-0.0256	-0.0403	-0.0387	-0.0663	-0.0090	-0.0074	-0.0174	-0.0154	-0.0615	-0.0055	-0.0046	-0.0117	-0.0113	-0.0603
	SD	0.1013	0.1126	0.1317	0.1440	0.0530	0.0546	0.0579	0.0770	0.0799	0.0311	0.0357	0.0369	0.0511	0.0526	0.0191
	RMSE	0.1048	0.1154	0.1377	0.1491	**0.0849**	**0.0554**	0.0583	0.0790	0.0814	0.0689	**0.0361**	0.0372	0.0524	0.0538	0.0632
Lambda=0.2	**Mean**	0.2431	0.2478	0.3397	0.3403	0.1916	0.2208	0.2232	0.3509	0.3515	0.1951	0.2096	0.2097	0.2895	0.2910	0.1977
	Bias	0.0431	0.0478	0.1397	0.1403	-0.0084	0.0208	0.0232	0.1509	0.1515	-0.0049	0.0096	0.0097	0.0895	0.0910	-0.0023
	SD	0.2289	0.2545	0.5043	0.5506	0.0800	0.1482	0.1559	0.4877	0.5041	0.0487	0.1020	0.1054	0.4704	0.4789	0.0318
	RMSE	0.2329	0.2589	0.5233	0.5682	**0.0804**	0.1496	0.1576	0.5105	0.5264	**0.0490**	0.1025	0.1058	0.4788	0.4874	**0.0319**
Rho=0.2	**Mean**	0.1839	0.1833	0.1126	0.1066	0.2063	0.1914	0.1887	0.1017	0.1026	0.2057	0.1961	0.1941	0.1387	0.1346	0.2069
	Bias	-0.0161	-0.0167	-0.0874	-0.0934	0.0063	-0.0086	-0.0113	-0.0983	-0.0974	0.0057	-0.0039	-0.0059	-0.0613	-0.0654	0.0069
	SD	0.1647	0.1824	0.4417	0.4688	0.0918	0.1012	0.1085	0.3689	0.3828	0.0561	0.0687	0.0710	0.3572	0.3628	0.0375
	RMSE	0.1655	0.1832	0.4502	0.4780	**0.0920**	0.1015	0.1091	0.3818	0.3950	**0.0564**	0.0688	0.0713	0.3624	0.3686	**0.0381**
Beta=1	**Mean**	0.9842	0.9855	0.9653	0.9630	1.0160	0.9935	0.9915	0.9854	0.9871	1.0174	0.9957	0.9952	0.9888	0.9882	1.0176
	Bias	-0.0158	-0.0145	-0.0347	-0.0370	0.0160	-0.0065	-0.0085	-0.0146	-0.0129	0.0174	-0.0043	-0.0048	-0.0112	-0.0118	0.0176
	SD	0.1435	0.1592	0.1836	0.2126	0.0821	0.0799	0.0839	0.1044	0.1108	0.0475	0.0517	0.0533	0.0731	0.0743	0.0295
	RMSE	0.1444	0.1599	0.1868	0.2158	**0.0837**	0.0801	0.0844	0.1054	0.1116	**0.0506**	0.0519	0.0536	0.0740	0.0753	**0.0344**

Table 2.9. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 50$, $\gamma = 0.4$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML
Gamma=0.4	Mean	0.3973	0.3979	0.3957	0.3958	0.3878	0.3988	0.3996	0.3988	0.4000	0.3878	0.3995	0.3998	0.4006	0.4008	0.3886
	Bias	-0.0027	-0.0021	-0.0043	-0.0042	-0.0122	-0.0012	-0.0004	-0.0012	0.0000	-0.0122	-0.0005	-0.0002	0.0006	0.0008	-0.0114
	SD	0.0303	0.0334	0.0448	0.0489	0.0209	0.0176	0.0188	0.0286	0.0302	0.0122	0.0108	0.0112	0.0210	0.0213	0.0075
	RMSE	0.0304	0.0335	0.0450	0.0491	**0.0242**	0.0176	0.0188	0.0286	0.0302	**0.0172**	**0.0108**	0.0112	0.0210	0.0213	0.0137
Lambda=0.2	Mean	0.2056	0.2038	0.2391	0.2496	0.2009	0.2009	0.2019	0.1999	0.1915	0.2009	0.2011	0.2008	0.1792	0.1798	0.2013
	Bias	0.0056	0.0038	0.0391	0.0496	0.0009	0.0009	0.0019	-0.0001	-0.0085	0.0009	0.0011	0.0008	-0.0208	-0.0202	0.0013
	SD	0.0739	0.0814	0.4013	0.4318	0.0340	0.0442	0.0475	0.3549	0.3788	0.0199	0.0304	0.0312	0.2909	0.2961	0.0141
	RMSE	0.0741	0.0814	0.4032	0.4347	**0.0340**	0.0442	0.0476	0.3549	0.3789	**0.0199**	0.0304	0.0312	0.2917	0.2968	**0.0142**
Rho=0.2	Mean	0.1965	0.1959	0.1710	0.1621	0.2014	0.1985	0.1981	0.1996	0.2057	0.2011	0.1998	0.1997	0.2168	0.2163	0.2008
	Bias	-0.0035	-0.0041	-0.0290	-0.0379	0.0014	-0.0015	-0.0019	-0.0004	0.0057	0.0011	-0.0002	-0.0003	0.0168	0.0163	0.0008
	SD	0.0596	0.0654	0.3070	0.3335	0.0379	0.0352	0.0377	0.2716	0.2910	0.0230	0.0242	0.0249	0.2182	0.2220	0.0155
	RMSE	0.0597	0.0656	0.3083	0.3357	**0.0380**	0.0352	0.0377	0.2716	0.2911	**0.0230**	0.0242	0.0249	0.2189	0.2226	**0.0156**
Beta=1	Mean	0.9974	0.9980	0.9957	0.9955	1.0060	1.0000	1.0007	1.0008	1.0019	1.0067	1.0000	1.0000	1.0018	1.0017	1.0060
	Bias	-0.0026	-0.0020	-0.0043	-0.0045	0.0060	0.0000	0.0007	0.0008	0.0019	0.0067	0.0000	0.0000	0.0018	0.0017	0.0060
	SD	0.0453	0.0497	0.0628	0.0687	0.0336	0.0257	0.0271	0.0410	0.0433	0.0198	0.0162	0.0167	0.0280	0.0285	0.0127
	RMSE	0.0453	0.0498	0.0630	0.0689	**0.0342**	0.0257	0.0271	0.0410	0.0433	**0.0209**	0.0162	0.0167	0.0281	0.0285	**0.0140**

Table 2.10. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 5$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.4$, $\beta = 1$ and various N, over 999 iterations

		N=16					**N=49**					**N=121**				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.2	**Mean**	0.1053	0.1046	0.0648	0.0592	0.0720	0.1594	0.1590	0.1391	0.1389	0.0787	0.1824	0.1809	0.1603	0.1604	0.0802
	Bias	-0.0947	-0.0954	-0.1352	-0.1408	-0.1280	-0.0406	-0.0410	-0.0609	-0.0611	-0.1213	-0.0176	-0.0191	-0.0397	-0.0396	-0.1198
	SD	0.1929	0.2141	0.2433	0.2679	0.0898	0.1139	0.1207	0.1615	0.1785	0.0508	0.0732	0.0770	0.1171	0.1200	0.0316
	RMSE	0.2149	0.2344	0.2783	0.3026	**0.1564**	**0.1210**	0.1275	0.1726	0.1887	0.1315	**0.0753**	0.0793	0.1236	0.1263	0.1239
Lambda=0.2	**Mean**	0.3059	0.3065	0.3672	0.3663	0.1833	0.2634	0.2639	0.3814	0.3770	0.1844	0.2391	0.2404	0.3656	0.3657	0.1898
	Bias	0.1059	0.1065	0.1672	0.1663	-0.0167	0.0634	0.0639	0.1814	0.1770	-0.0156	0.0391	0.0404	0.1656	0.1657	-0.0102
	SD	0.3605	0.4032	0.5496	0.6101	0.1130	0.2871	0.3114	0.5313	0.5896	0.0736	0.2231	0.2315	0.5688	0.5865	0.0491
	RMSE	0.3758	0.4170	0.5745	0.6324	**0.1142**	0.2940	0.3179	0.5614	0.6156	**0.0752**	0.2265	0.2350	0.5925	0.6095	**0.0502**
Rho=0.4	**Mean**	0.3496	0.3514	0.3375	0.3413	0.3891	0.3675	0.3673	0.2974	0.2961	0.3933	0.3807	0.3804	0.3061	0.3031	0.3945
	Bias	-0.0504	-0.0486	-0.0625	-0.0587	-0.0109	-0.0325	-0.0327	-0.1026	-0.1039	-0.0067	-0.0193	-0.0196	-0.0939	-0.0969	-0.0055
	SD	0.2989	0.3357	0.4931	0.5725	0.1439	0.2073	0.2255	0.4246	0.4584	0.0880	0.1402	0.1450	0.4050	0.4176	0.0586
	RMSE	0.3031	0.3392	0.4970	0.5755	**0.1443**	0.2098	0.2279	0.4368	0.4700	**0.0883**	0.1415	0.1464	0.4158	0.4287	**0.0589**
Beta=1	**Mean**	0.9526	0.9542	0.9021	0.8946	1.0096	0.9700	0.9698	0.9547	0.9523	1.0091	0.9841	0.9799	0.9648	0.9625	1.0101
	Bias	-0.0474	-0.0458	-0.0979	-0.1054	0.0096	-0.0300	-0.0302	-0.0453	-0.0477	0.0091	-0.0159	-0.0201	-0.0352	-0.0375	0.0101
	SD	0.2856	0.3222	0.4219	0.4676	0.1375	0.1698	0.1829	0.2666	0.3034	0.0788	0.1091	0.1146	0.1773	0.1818	0.0479
	RMSE	0.2895	0.3254	0.4331	0.4793	**0.1379**	0.1724	0.1853	0.2704	0.3071	**0.0793**	0.1102	0.1164	0.1807	0.1857	**0.0490**

Table 2.11. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 10$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.4$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML
Gamma=0.2	Mean	0.1765	0.1784	0.1535	0.1542	0.1400	0.1921	0.1944	0.1756	0.1782	0.1436	0.1949	0.1959	0.1834	0.1836	0.1446
	Bias	-0.0235	-0.0216	-0.0465	-0.0458	-0.0600	-0.0079	-0.0056	-0.0244	-0.0218	-0.0564	-0.0051	-0.0041	-0.0166	-0.0164	-0.0554
	SD	0.0961	0.1076	0.1363	0.1494	0.0569	0.0530	0.0563	0.0862	0.0908	0.0328	0.0347	0.0358	0.0684	0.0703	0.0202
	RMSE	0.0989	0.1097	0.1440	0.1563	**0.0827**	**0.0536**	0.0566	0.0896	0.0933	0.0653	**0.0351**	0.0360	0.0704	0.0722	0.0590
Lambda=0.2	Mean	0.2366	0.2371	0.3547	0.3566	0.1918	0.2190	0.2204	0.3514	0.3505	0.1960	0.2088	0.2093	0.3003	0.3053	0.1988
	Bias	0.0366	0.0371	0.1547	0.1566	-0.0082	0.0190	0.0204	0.1514	0.1505	-0.0040	0.0088	0.0093	0.1003	0.1053	-0.0012
	SD	0.2242	0.2497	0.4951	0.5452	0.0794	0.1467	0.1543	0.4694	0.4966	0.0482	0.1009	0.1041	0.4504	0.4626	0.0316
	RMSE	0.2272	0.2525	0.5187	0.5672	**0.0798**	0.1480	0.1557	0.4932	0.5189	**0.0484**	0.1013	0.1045	0.4614	0.4744	**0.0316**
Rho=0.4	Mean	0.3833	0.3851	0.3196	0.3163	0.3977	0.3919	0.3890	0.3249	0.3243	0.3988	0.3971	0.3950	0.3488	0.3438	0.4006
	Bias	-0.0167	-0.0149	-0.0804	-0.0837	-0.0023	-0.0081	-0.0110	-0.0751	-0.0757	-0.0012	-0.0029	-0.0050	-0.0512	-0.0562	0.0006
	SD	0.1613	0.1791	0.3735	0.4091	0.0903	0.1008	0.1081	0.2859	0.3048	0.0560	0.0678	0.0702	0.2683	0.2756	0.0377
	RMSE	0.1621	0.1798	0.3821	0.4175	**0.0903**	0.1011	0.1087	0.2956	0.3141	**0.0560**	0.0678	0.0704	0.2732	0.2813	**0.0377**
Beta=1	Mean	0.9878	0.9883	0.9672	0.9667	1.0155	0.9948	0.9922	0.9866	0.9888	1.0171	0.9966	0.9959	0.9896	0.3438	1.0172
	Bias	-0.0122	-0.0117	-0.0328	-0.0333	0.0155	-0.0052	-0.0078	-0.0134	-0.0112	0.0171	-0.0034	-0.0041	-0.0104	-0.0562	0.0172
	SD	0.1392	0.1551	0.1785	0.2046	0.0824	0.0783	0.0820	0.1016	0.1076	0.0476	0.0502	0.0518	0.0702	0.2756	0.0296
	RMSE	0.1397	0.1555	0.1815	0.2073	**0.0838**	0.0785	0.0824	0.1025	0.1082	**0.0506**	0.0503	0.0520	0.0710	0.2813	**0.0343**

Table 2.12. Mean, bias, standard deviation and root mean square error of various spatial estimators in a stationary scenario, for $T = 50$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.4$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.2	**Mean**	0.1974	0.1979	0.1862	0.1856	0.1884	0.1987	0.1995	0.1914	0.1923	0.1882	0.1996	0.1999	0.1935	0.1937	0.1890
	Bias	-0.0026	-0.0021	-0.0138	-0.0144	-0.0116	-0.0013	-0.0005	-0.0086	-0.0077	-0.0118	-0.0004	-0.0001	-0.0065	-0.0063	-0.0110
	SD	0.0315	0.0347	0.0662	0.0709	0.0233	0.0181	0.0193	0.0450	0.0467	0.0136	0.0112	0.0116	0.0314	0.0319	0.0084
	RMSE	0.0316	0.0348	0.0676	0.0723	**0.0260**	0.0181	0.0193	0.0459	0.0473	**0.0180**	**0.0112**	0.0116	0.0321	0.0325	0.0138
Lambda=0.2	**Mean**	0.2049	0.2035	0.2863	0.2947	0.2011	0.2006	0.2015	0.2612	0.2586	0.2016	0.2008	0.2005	0.2533	0.2536	0.2015
	Bias	0.0049	0.0035	0.0863	0.0947	0.0011	0.0006	0.0015	0.0612	0.0586	0.0016	0.0008	0.0005	0.0533	0.0536	0.0015
	SD	0.0754	0.0827	0.4052	0.4308	0.0333	0.0452	0.0484	0.3399	0.3496	0.0197	0.0309	0.0318	0.2535	0.2577	0.0138
	RMSE	0.0756	0.0827	0.4142	0.4411	**0.0333**	0.0452	0.0484	0.3453	0.3544	**0.0197**	0.0309	0.0318	0.2590	0.2632	**0.0139**
Rho=0.4	**Mean**	0.3965	0.3959	0.3461	0.3405	0.4003	0.3984	0.3982	0.3618	0.3631	0.3999	0.3998	0.3997	0.3682	0.3678	0.4001
	Bias	-0.0035	-0.0041	-0.0539	-0.0595	0.0003	-0.0016	-0.0018	-0.0382	-0.0369	-0.0001	-0.0002	-0.0003	-0.0318	-0.0322	0.0001
	SD	0.0566	0.0623	0.2448	0.2649	0.0378	0.0338	0.0359	0.2060	0.2134	0.0231	0.0229	0.0236	0.1515	0.1545	0.0154
	RMSE	0.0567	0.0624	0.2506	0.2715	**0.0378**	0.0338	0.0359	0.2095	0.2166	**0.0231**	0.0229	0.0236	0.1548	0.1578	**0.0154**
Beta=1	**Mean**	0.9975	0.9982	0.9905	0.9911	1.0052	1.0000	1.0005	0.9955	0.9961	1.0060	1.0001	1.0000	0.9960	0.9960	1.0053
	Bias	-0.0025	-0.0018	-0.0095	-0.0089	0.0052	0.0000	0.0005	-0.0045	-0.0039	0.0060	0.0001	0.0000	-0.0040	-0.0040	0.0053
	SD	0.0457	0.0499	0.0628	0.0683	0.0337	0.0261	0.0277	0.0401	0.0415	0.0198	0.0164	0.0169	0.0266	0.0272	0.0127
	RMSE	0.0457	0.0499	0.0636	0.0689	**0.0341**	0.0261	0.0277	0.0404	0.0417	**0.0207**	0.0164	0.0169	0.0269	0.0275	**0.0137**

Table 2.13. Mean, bias, standard deviation and root mean square error of various spatial estimators in a quasi-unit root scenario, for $T = 5$, $\gamma = 0.58$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.58	**Mean**	0.4433	0.4457	0.3638	0.3623	0.4205	0.5209	0.5199	0.4737	0.4726	0.4332	0.5541	0.5509	0.5130	0.5108	0.4354
	Bias	-0.1367	-0.1343	-0.2162	-0.2177	-0.1595	-0.0591	-0.0601	-0.1063	-0.1074	-0.1468	-0.0259	-0.0291	-0.0670	-0.0692	-0.1446
	SD	0.2186	0.2380	0.2876	0.3153	0.0904	0.1405	0.1535	0.2144	0.2257	0.0502	0.0829	0.0859	0.1437	0.1500	0.0304
	RMSE	0.2578	0.2733	0.3598	0.3831	**0.1834**	**0.1524**	0.1648	0.2394	0.2500	0.1552	**0.0869**	0.0907	0.1586	0.1652	0.1477
Lambda=0.2	**Mean**	0.3565	0.3667	0.4579	0.4577	0.1764	0.3324	0.3314	0.4707	0.4748	0.1770	0.3033	0.3053	0.4726	0.4621	0.1808
	Bias	0.1565	0.1667	0.2579	0.2577	-0.0236	0.1324	0.1314	0.2707	0.2748	-0.0230	0.1033	0.1053	0.2726	0.2621	-0.0192
	SD	0.3669	0.4072	0.4909	0.5382	0.1142	0.3332	0.3657	0.4866	0.5128	0.0754	0.2845	0.2971	0.4698	0.4927	0.0500
	RMSE	0.3988	0.4400	0.5545	0.5967	**0.1166**	0.3585	0.3886	0.5569	0.5818	**0.0788**	0.3026	0.3152	0.5431	0.5580	**0.0535**
Rho=0.2	**Mean**	0.1172	0.1195	0.0692	0.0667	0.2198	0.1366	0.1354	-0.0332	-0.0409	0.2222	0.1608	0.1610	-0.0653	-0.0677	0.2197
	Bias	-0.0828	-0.0805	-0.1308	-0.1333	0.0198	-0.0634	-0.0646	-0.2332	-0.2409	0.0222	-0.0392	-0.0390	-0.2653	-0.2677	0.0197
	SD	0.3136	0.3520	0.4775	0.5526	0.1519	0.2099	0.2278	0.4656	0.5109	0.0879	0.1476	0.1530	0.4562	0.4745	0.0598
	RMSE	0.3243	0.3611	0.4951	0.5684	**0.1532**	0.2193	0.2368	0.5207	0.5649	**0.0907**	0.1527	0.1579	0.5278	0.5448	**0.0629**
Beta=1	**Mean**	0.8916	0.8895	0.8001	0.7981	0.9858	0.9452	0.9431	0.8714	0.8696	0.9868	0.9815	0.9767	0.9090	0.9041	0.9883
	Bias	-0.1084	-0.1105	-0.1999	-0.2019	-0.0142	-0.0548	-0.0569	-0.1286	-0.1304	-0.0132	-0.0185	-0.0233	-0.0910	-0.0959	-0.0117
	SD	0.3291	0.3601	0.4691	0.5135	0.1369	0.2120	0.2329	0.3352	0.3664	0.0788	0.1346	0.1411	0.2353	0.2468	0.0478
	RMSE	0.3465	0.3767	0.5099	0.5518	**0.1376**	0.2189	0.2397	0.3591	0.3889	**0.0799**	0.1358	0.1430	0.2523	0.2648	**0.0492**

Table 2.14. Mean, bias, standard deviation and root mean square error of various spatial estimators in a quasi-unit root scenario, for $T = 10$, $\gamma = 0.58$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.58	**Mean**	0.5456	0.5435	0.5131	0.5141	0.5117	0.5678	0.5680	0.5447	0.5457	0.5163	0.5734	0.5740	0.5540	0.5549	0.5173
	Bias	-0.0344	-0.0365	-0.0669	-0.0659	-0.0683	-0.0122	-0.0120	-0.0353	-0.0343	-0.0637	-0.0066	-0.0060	-0.0260	-0.0251	-0.0627
	SD	0.1077	0.1199	0.1536	0.1737	0.0484	0.0584	0.0634	0.0902	0.0955	0.0286	0.0386	0.0399	0.0631	0.0647	0.0175
	RMSE	0.1130	0.1253	0.1676	0.1858	**0.0837**	**0.0597**	0.0645	0.0969	0.1015	0.0698	**0.0391**	0.0404	0.0682	0.0694	0.0651
Lambda=0.2	**Mean**	0.2607	0.2633	0.4291	0.4165	0.1842	0.2425	0.2450	0.4264	0.4241	0.1877	0.2290	0.2322	0.4186	0.4113	0.1909
	Bias	0.0607	0.0633	0.2291	0.2165	-0.0158	0.0425	0.0450	0.2264	0.2241	-0.0123	0.0290	0.0322	0.2186	0.2113	-0.0091
	SD	0.2550	0.2732	0.4580	0.5008	0.0812	0.2146	0.2234	0.4221	0.4487	0.0494	0.1766	0.1843	0.4288	0.4508	0.0318
	RMSE	0.2621	0.2804	0.5121	0.5456	**0.0827**	0.2188	0.2278	0.4790	0.5016	**0.0509**	0.1790	0.1871	0.4813	0.4978	**0.0331**
Rho=0.2	**Mean**	0.1642	0.1631	0.0125	0.0156	0.2129	0.1819	0.1788	-0.0316	-0.0320	0.2106	0.1894	0.1881	0.0143	0.0158	0.2112
	Bias	-0.0358	-0.0369	-0.1875	-0.1844	0.0129	-0.0181	-0.0212	-0.2316	-0.2320	0.0106	-0.0106	-0.0119	-0.1857	-0.1842	0.0112
	SD	0.1746	0.1913	0.4423	0.4788	0.0907	0.1053	0.1148	0.4058	0.4280	0.0557	0.0727	0.0746	0.3739	0.3883	0.0366
	RMSE	0.1782	0.1948	0.4804	0.5131	**0.0916**	0.1068	0.1168	0.4672	0.4869	**0.0567**	0.0735	0.0755	0.4175	0.4298	**0.0382**
Beta=1	**Mean**	0.9748	0.9691	0.9313	0.9277	1.0091	0.9921	0.9909	0.9523	0.9531	1.0111	0.9963	0.9960	0.9711	0.9699	1.0116
	Bias	-0.0252	-0.0309	-0.0687	-0.0723	0.0091	-0.0079	-0.0091	-0.0477	-0.0469	0.0111	-0.0037	-0.0040	-0.0289	-0.0301	0.0116
	SD	0.1570	0.1687	0.2226	0.2589	0.0817	0.0954	0.1013	0.1437	0.1530	0.0473	0.0659	0.0691	0.1134	0.1176	0.0294
	RMSE	0.1590	0.1715	0.2329	0.2688	**0.0822**	0.0957	0.1017	0.1514	0.1601	**0.0485**	0.0660	0.0692	0.1170	0.1214	**0.0316**

Table 2.15. Mean, bias, standard deviation and root mean square error of various spatial estimators in a quasi-unit root scenario, for $T = 50$, $\gamma = 0.58$, $\lambda = 0.2$, $\varrho = 0.2$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML
Gamma=0.58	Mean	0.5775	0.5781	0.5742	0.5743	0.5684	0.5789	0.5794	0.5784	0.5793	0.5684	0.5793	0.5795	0.5794	0.5794	0.5690
	Bias	-0.0025	-0.0019	-0.0058	-0.0057	-0.0116	-0.0011	-0.0006	-0.0016	-0.0007	-0.0116	-0.0007	-0.0005	-0.0006	-0.0006	-0.0110
	SD	0.0306	0.0341	0.0468	0.0525	0.0178	0.0173	0.0184	0.0258	0.0272	0.0102	0.0108	0.0112	0.0186	0.0188	0.0064
	RMSE	0.0307	0.0342	0.0471	0.0528	**0.0213**	0.0173	0.0185	0.0259	0.0272	**0.0154**	**0.0108**	0.0112	0.0186	0.0189	0.0128
Lambda=0.2	Mean	0.2059	0.2079	0.2416	0.2578	0.1998	0.2020	0.2022	0.2076	0.2025	0.2003	0.2016	0.2014	0.2046	0.2027	0.2005
	Bias	0.0059	0.0079	0.0416	0.0578	-0.0002	0.0020	0.0022	0.0076	0.0025	0.0003	0.0016	0.0014	0.0046	0.0027	0.0005
	SD	0.1151	0.1278	0.3602	0.4186	0.0337	0.0857	0.0918	0.2763	0.2997	0.0196	0.0575	0.0592	0.2094	0.2127	0.0137
	RMSE	0.1152	0.1280	0.3626	0.4226	**0.0337**	0.0857	0.0918	0.2764	0.2997	**0.0196**	0.0575	0.0593	0.2094	0.2127	**0.0137**
Rho=0.2	Mean	0.1952	0.1947	0.1588	0.1506	0.2034	0.1982	0.1974	0.1936	0.1971	0.2035	0.1992	0.1992	0.2010	0.2007	0.2031
	Bias	-0.0048	-0.0053	-0.0412	-0.0494	0.0034	-0.0018	-0.0026	-0.0064	-0.0029	0.0035	-0.0008	-0.0008	0.0010	0.0007	0.0031
	SD	0.0626	0.0694	0.3200	0.3559	0.0361	0.0379	0.0412	0.2356	0.2514	0.0221	0.0261	0.0269	0.1723	0.1748	0.0146
	RMSE	0.0628	0.0696	0.3226	0.3594	**0.0362**	0.0379	0.0412	0.2357	0.2514	**0.0224**	0.0262	0.0269	0.1724	0.1748	**0.0150**
Beta=1	Mean	0.9969	0.9977	0.9923	0.9940	1.0056	0.9997	1.0004	0.9997	1.0006	1.0064	0.9998	0.9999	1.0008	1.0003	1.0058
	Bias	-0.0031	-0.0023	-0.0077	-0.0060	0.0056	-0.0003	0.0004	-0.0003	0.0006	0.0064	-0.0002	-0.0001	0.0008	0.0003	0.0058
	SD	0.0520	0.0575	0.0840	0.0968	0.0332	0.0310	0.0328	0.0571	0.0585	0.0195	0.0204	0.0212	0.0433	0.0445	0.0125
	RMSE	0.0521	0.0576	0.0843	0.0970	**0.0337**	0.0310	0.0328	0.0571	0.0585	**0.0206**	0.0204	0.0212	0.0433	0.0445	**0.0138**

Table 2.16. Mean, bias, standard deviation and root mean square error of various spatial estimators in a quasi-unit root scenario, for $T = 5$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.58$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML	One-step ESAB	Two-step ESAB	One-step SAB	Two-step SAB	QML
Gamma=0.2	**Mean**	0.0983	0.0960	0.0254	0.0232	0.0770	0.1551	0.1573	0.0883	0.0898	0.0829	0.1771	0.1756	0.1155	0.1174	0.0835
	Bias	-0.1017	-0.1040	-0.1746	-0.1768	-0.1230	-0.0449	-0.0427	-0.1117	-0.1102	-0.1171	-0.0229	-0.0244	-0.0845	-0.0826	-0.1165
	SD	0.1965	0.2185	0.2806	0.3056	0.0887	0.1171	0.1265	0.1771	0.1909	0.0511	0.0765	0.0800	0.1375	0.1401	0.0316
	RMSE	0.2212	0.2420	0.3305	0.3531	**0.1516**	**0.1254**	0.1335	0.2094	0.2204	0.1277	**0.0799**	0.0836	0.1614	0.1627	0.1207
Lambda=0.2	**Mean**	0.3149	0.3178	0.4328	0.4304	0.1752	0.2857	0.2827	0.4645	0.4590	0.1777	0.2676	0.2688	0.4732	0.4600	0.1823
	Bias	0.1149	0.1178	0.2328	0.2304	-0.0248	0.0857	0.0827	0.2645	0.2590	-0.0223	0.0676	0.0688	0.2732	0.2600	-0.0177
	SD	0.3442	0.3835	0.4786	0.5243	0.1114	0.3050	0.3364	0.4459	0.4757	0.0734	0.2593	0.2714	0.4758	0.4657	0.0488
	RMSE	0.3629	0.4011	0.5322	0.5727	**0.1141**	0.3168	0.3464	0.5184	0.5416	**0.0767**	0.2679	0.2800	0.5486	0.5334	**0.0520**
Rho=0.58	**Mean**	0.4938	0.5009	0.4345	0.4420	0.5570	0.5177	0.5168	0.3708	0.3644	0.5650	0.5439	0.5440	0.3600	0.3567	0.5646
	Bias	-0.0862	-0.0791	-0.1455	-0.1380	-0.0230	-0.0623	-0.0632	-0.2092	-0.2156	-0.0150	-0.0361	-0.0360	-0.2200	-0.2233	-0.0154
	SD	0.2958	0.3293	0.4394	0.4759	0.1447	0.2198	0.2388	0.3966	0.4412	0.0862	0.1608	0.1653	0.3772	0.3883	0.0571
	RMSE	0.3081	0.3387	0.4629	0.4955	**0.1465**	0.2284	0.2470	0.4484	0.4911	**0.0875**	0.1648	0.1692	0.4367	0.4479	**0.0591**
Beta=1	**Mean**	0.9473	0.9452	0.8645	0.8697	1.0027	0.9657	0.9665	0.9043	0.9005	1.0037	0.9853	0.9822	0.9308	0.9299	1.0047
	Bias	-0.0527	-0.0548	-0.1355	-0.1303	0.0027	-0.0343	-0.0335	-0.0957	-0.0995	0.0037	-0.0147	-0.0178	-0.0692	-0.0701	0.0047
	SD	0.2889	0.3199	0.4582	0.4993	0.1374	0.1785	0.1951	0.2980	0.3238	0.0789	0.1181	0.1238	0.1993	0.2048	0.0479
	RMSE	0.2936	0.3246	0.4778	0.5160	**0.1374**	0.1817	0.1979	0.3130	0.3388	**0.0790**	0.1191	0.1251	0.2110	0.2165	**0.0481**

Table 2.17. Mean, bias, standard deviation and root mean square error of various spatial estimators in a quasi-unit root scenario, for $\boldsymbol{T = 10}$, $\boldsymbol{\gamma = 0.2}$, $\boldsymbol{\lambda = 0.2}$, $\boldsymbol{\varrho = 0.58}$, $\boldsymbol{\beta = 1}$ and various $\boldsymbol{N}$, over 999 iterations

		N=16					**N=49**					**N=121**				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.2	**Mean**	0.1740	0.1751	0.1128	0.1159	0.1435	0.1898	0.1919	0.1497	0.1509	0.1457	0.1932	0.1939	0.1579	0.1581	0.1463
	Bias	-0.0260	-0.0249	-0.0872	-0.0841	-0.0565	-0.0102	-0.0081	-0.0503	-0.0491	-0.0543	-0.0068	-0.0061	-0.0421	-0.0419	-0.0537
	SD	0.0970	0.1079	0.1653	0.1797	0.0567	0.0545	0.0593	0.1026	0.1088	0.0325	0.0361	0.0376	0.0898	0.0921	0.0200
	RMSE	0.1004	0.1107	0.1869	0.1984	**0.0800**	**0.0555**	0.0598	0.1143	0.1193	0.0633	**0.0368**	0.0381	0.0991	0.1012	0.0573
Lambda=0.2	**Mean**	0.2403	0.2427	0.4378	0.4333	0.1864	0.2247	0.2239	0.4074	0.4065	0.1917	0.2188	0.2211	0.4024	0.4045	0.1951
	Bias	0.0403	0.0427	0.2378	0.2333	-0.0136	0.0247	0.0239	0.2074	0.2065	-0.0083	0.0188	0.0211	0.2024	0.2045	-0.0049
	SD	0.2343	0.2549	0.4622	0.5100	0.0789	0.1945	0.2082	0.4101	0.4281	0.0478	0.1518	0.1559	0.4239	0.4351	0.0307
	RMSE	0.2378	0.2584	0.5198	0.5608	**0.0800**	0.1960	0.2095	0.4595	0.4753	**0.0485**	0.1530	0.1573	0.4697	0.4808	**0.0311**
Rho=0.58	**Mean**	0.5465	0.5478	0.4215	0.4148	0.5744	0.5648	0.5608	0.4258	0.4255	0.5756	0.5735	0.5716	0.4702	0.4657	0.5778
	Bias	-0.0335	-0.0322	-0.1585	-0.1652	-0.0056	-0.0152	-0.0192	-0.1542	-0.1545	-0.0044	-0.0065	-0.0084	-0.1098	-0.1143	-0.0022
	SD	0.1642	0.1815	0.3487	0.3796	0.0877	0.1107	0.1181	0.3074	0.3271	0.0542	0.0793	0.0829	0.2806	0.2866	0.0360
	RMSE	0.1676	0.1844	0.3830	0.4140	**0.0879**	0.1117	0.1196	0.3439	0.3618	**0.0544**	0.0796	0.0833	0.3013	0.3086	**0.0361**
Beta=1	**Mean**	0.9838	0.9833	0.9402	0.9406	1.0114	0.9933	0.9912	0.9656	0.9650	1.0137	0.9965	0.9956	0.9812	0.9802	1.0142
	Bias	-0.0162	-0.0167	-0.0598	-0.0594	0.0114	-0.0067	-0.0088	-0.0344	-0.0350	0.0137	-0.0035	-0.0044	-0.0188	-0.0198	0.0142
	SD	0.1431	0.1537	0.2037	0.2228	0.0822	0.0855	0.0917	0.1197	0.1281	0.0474	0.0572	0.0593	0.0878	0.0904	0.0296
	RMSE	0.1441	0.1546	0.2123	0.2306	**0.0830**	0.0858	0.0921	0.1245	0.1328	**0.0493**	0.0573	0.0594	0.0898	0.0925	**0.0328**

Table 2.18. Mean, bias, standard deviation and root mean square error of various spatial estimators in a quasi-unit root scenario, for $T = 50$, $\gamma = 0.2$, $\lambda = 0.2$, $\varrho = 0.58$, $\beta = 1$ and various N, over 999 iterations

		N=16					**N=49**					**N=121**				
		One-step ESAB	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**	**One-step ESAB**	**Two-step ESAB**	**One-step SAB**	**Two-step SAB**	**QML**
Gamma=0.2	**Mean**	0.1975	0.1978	0.1781	0.1775	0.1895	0.1988	0.1994	0.1883	0.1889	0.1888	0.1994	0.1997	0.1935	0.1938	0.1894
	Bias	-0.0025	-0.0022	-0.0219	-0.0225	-0.0105	-0.0012	-0.0006	-0.0117	-0.0111	-0.0112	-0.0006	-0.0003	-0.0065	-0.0062	-0.0106
	SD	0.0334	0.0368	0.0774	0.0848	0.0232	0.0187	0.0200	0.0514	0.0529	0.0134	0.0117	0.0120	0.0359	0.0362	0.0084
	RMSE	0.0335	0.0369	0.0804	0.0877	**0.0254**	0.0187	0.0200	0.0527	0.0540	**0.0175**	**0.0117**	0.0120	0.0365	0.0367	0.0135
Lambda=0.2	**Mean**	0.2053	0.2055	0.2960	0.3026	0.2004	0.1999	0.2007	0.2601	0.2593	0.2017	0.2013	0.2005	0.2398	0.2380	0.2018
	Bias	0.0053	0.0055	0.0960	0.1026	0.0004	-0.0001	0.0007	0.0601	0.0593	0.0017	0.0013	0.0005	0.0398	0.0380	0.0018
	SD	0.1027	0.1152	0.3408	0.3722	0.0322	0.0717	0.0752	0.2960	0.2957	0.0185	0.0468	0.0482	0.2264	0.2288	0.0133
	RMSE	0.1028	0.1153	0.3541	0.3861	**0.0322**	0.0717	0.0752	0.3021	0.3016	**0.0186**	0.0469	0.0482	0.2299	0.2320	**0.0134**
Rho=0.58	**Mean**	0.5757	0.5752	0.5112	0.5066	0.5811	0.5777	0.5772	0.5426	0.5416	0.5811	0.5795	0.5794	0.5612	0.5609	0.5809
	Bias	-0.0043	-0.0048	-0.0688	-0.0734	0.0011	-0.0023	-0.0028	-0.0374	-0.0384	0.0011	-0.0005	-0.0006	-0.0188	-0.0191	0.0009
	SD	0.0610	0.0682	0.2254	0.2462	0.0348	0.0418	0.0443	0.1585	0.1643	0.0219	0.0275	0.0284	0.1045	0.1061	0.0144
	RMSE	0.0612	0.0683	0.2357	0.2570	**0.0348**	0.0419	0.0444	0.1628	0.1687	**0.0219**	0.0275	0.0284	0.1062	0.1078	**0.0144**
Beta=1	**Mean**	0.9971	0.9978	0.9868	0.9872	1.0044	0.9994	1.0000	0.9960	0.9961	1.0053	1.0001	1.0000	0.9987	0.9986	1.0047
	Bias	-0.0029	-0.0022	-0.0132	-0.0128	0.0044	-0.0006	0.0000	-0.0040	-0.0039	0.0053	0.0001	0.0000	-0.0013	-0.0014	0.0047
	SD	0.0496	0.0544	0.0695	0.0769	0.0335	0.0292	0.0308	0.0443	0.0439	0.0196	0.0181	0.0188	0.0265	0.0269	0.0126
	RMSE	0.0497	0.0544	0.0707	0.0780	**0.0337**	0.0292	0.0308	0.0445	0.0441	**0.0203**	0.0181	0.0188	0.0266	0.0270	**0.0135**

Table 2.19. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $\boldsymbol{T = 5}$, $\boldsymbol{\gamma = 0.3}$, $\boldsymbol{\lambda = 0.1}$, $\boldsymbol{\varrho = 0.1}$, $\boldsymbol{\beta = 1}$ and various $\boldsymbol{N}$, over 999 iterations

		N=16					**N=49**					**N=121**				
		One-step AB	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**	**One-step AB**	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**	**One-step AB**	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**
Gamma=0.3	**Mean**	0.2556	0.2598	0.2072	0.2047	0.1590	0.2917	0.2918	0.2601	0.2597	0.1680	0.3038	0.3033	0.2836	0.2820	0.1710
	Bias	-0.0444	-0.0402	-0.0928	-0.0953	-0.1410	-0.0083	-0.0082	-0.0399	-0.0403	-0.1320	0.0038	0.0033	-0.0164	-0.0180	-0.1290
	SD	0.2235	0.2456	0.1932	0.2109	0.0916	0.1241	0.1331	0.1157	0.1257	0.0507	0.0771	0.0795	0.0734	0.0766	0.0313
	RMSE	0.2279	0.2489	0.2143	0.2315	**0.1681**	0.1244	0.1333	**0.1224**	0.1320	0.1414	0.0772	0.0795	**0.0752**	0.0787	0.1328
Lambda=0.1	**Mean**			0.1577	0.1480	0.1106			0.1316	0.1346	0.0985			0.1171	0.1186	0.0987
	Bias			0.0577	0.0480	0.0106			0.0316	0.0346	-0.0015			0.0171	0.0186	-0.0013
	SD			0.3886	0.4335	0.1029			0.2945	0.3158	0.0694			0.2024	0.2150	0.0489
	RMSE			0.3928	0.4361	**0.1035**			0.2962	0.3177	**0.0694**			0.2031	0.2158	**0.0489**
Rho=0.1	**Mean**			0.0930	0.0946	0.1020			0.0914	0.0959	0.1019			0.0948	0.0956	0.1045
	Bias			-0.0070	-0.0054	0.0020			-0.0086	-0.0041	0.0019			-0.0052	-0.0044	0.0045
	SD			0.3049	0.3390	0.1465			0.2004	0.2164	0.0902			0.1294	0.1330	0.0606
	RMSE			0.3050	0.3390	**0.1465**			0.2006	0.2165	**0.0903**			0.1295	0.1331	**0.0608**
Beta=1	**Mean**	0.9904	0.9838	0.9501	0.9482	1.0111	1.0019	0.9999	0.9681	0.9667	1.0101	0.9996	0.9979	0.9844	0.9796	1.0112
	Bias	-0.0096	-0.0162	-0.0499	-0.0518	0.0111	0.0019	-0.0001	-0.0319	-0.0333	0.0101	-0.0004	-0.0021	-0.0156	-0.0204	0.0112
	SD	0.4012	0.4444	0.2874	0.3175	0.1366	0.2129	0.2258	0.1702	0.1826	0.0786	0.1342	0.1373	0.1071	0.1115	0.0478
	RMSE	0.4013	0.4447	0.2917	0.3217	**0.1370**	0.2130	0.2258	0.1731	0.1856	**0.0793**	0.1342	0.1373	0.1082	0.1133	**0.0491**

Table 2.20. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 10$, $\gamma = 0.3$, $\lambda = 0.1$, $\varrho = 0.1$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML
Gamma=0.3	Mean	0.3015	0.3016	0.2767	0.2787	0.2347	0.3105	0.3123	0.2925	0.2948	0.2398	0.3108	0.3112	0.2956	0.2967	0.2413
	Bias	0.0015	0.0016	-0.0233	-0.0213	-0.0653	0.0105	0.0123	-0.0075	-0.0052	-0.0602	0.0108	0.0112	-0.0044	-0.0033	-0.0587
	SD	0.1001	0.1101	0.0969	0.1083	0.0550	0.0547	0.0573	0.0530	0.0567	0.0322	0.0360	0.0365	0.0351	0.0360	0.0198
	RMSE	0.1001	0.1101	0.0997	0.1104	**0.0854**	0.0557	0.0586	**0.0535**	0.0569	0.0683	0.0376	0.0382	**0.0354**	0.0362	0.0619
Lambda=0.1	Mean			0.1177	0.1185	0.1025			0.1113	0.1114	0.0987			0.1066	0.1070	0.1003
	Bias			0.0177	0.0185	0.0025			0.0113	0.0114	-0.0013			0.0066	0.0070	0.0003
	SD			0.2226	0.2437	0.0740			0.1369	0.1459	0.0488			0.0937	0.0967	0.0322
	RMSE			0.2233	0.2444	**0.0741**			0.1373	0.1464	**0.0488**			0.0939	0.0970	**0.0322**
Rho=0.1	Mean			0.0974	0.0988	0.0997			0.0975	0.0952	0.1006			0.1002	0.0990	0.1027
	Bias			-0.0026	-0.0012	-0.0003			-0.0025	-0.0048	0.0006			0.0002	-0.0010	0.0027
	SD			0.1563	0.1736	0.0922			0.0994	0.1074	0.0575			0.0662	0.0687	0.0386
	RMSE			0.1564	0.1736	**0.0922**			0.0994	0.1075	**0.0575**			0.0662	0.0687	**0.0387**
Beta=1	Mean	1.0023	1.0018	0.9882	0.9895	1.0176	1.0017	1.0015	0.9952	0.9927	1.0189	1.0002	1.0001	0.9971	0.9967	1.0190
	Bias	0.0023	0.0018	-0.0118	-0.0105	0.0176	0.0017	0.0015	-0.0048	-0.0073	0.0189	0.0002	0.0001	-0.0029	-0.0033	0.0190
	SD	0.1477	0.1604	0.1384	0.1523	0.0825	0.0817	0.0850	0.0771	0.0821	0.0476	0.0521	0.0529	0.0493	0.0511	0.0296
	RMSE	0.1477	0.1604	0.1389	0.1527	**0.0844**	0.0818	0.0850	0.0772	0.0824	**0.0512**	0.0521	0.0529	0.0494	0.0512	**0.0352**

Table 2.21. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 50$, $\gamma = 0.3$, $\lambda = 0.1$, $\varrho = 0.1$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**	**One-step AB**	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**	**One-step AB**	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**
Gamma=0.3	**Mean**	0.3168	0.3170	0.2975	0.2985	0.2878	0.3145	0.3148	0.2987	0.2995	0.2876	0.3144	0.3146	0.2996	0.2999	0.2886
	Bias	0.0168	0.0170	-0.0025	-0.0015	-0.0122	0.0145	0.0148	-0.0013	-0.0005	-0.0124	0.0144	0.0146	-0.0004	-0.0001	-0.0114
	SD	0.0312	0.0331	0.0301	0.0335	0.0219	0.0183	0.0192	0.0177	0.0190	0.0131	0.0111	0.0114	0.0107	0.0112	0.0080
	RMSE	0.0354	0.0372	0.0302	0.0336	**0.0251**	0.0234	0.0243	**0.0178**	0.0190	0.0180	0.0182	0.0185	**0.0107**	0.0112	0.0139
Lambda=0.1	**Mean**			0.1044	0.1028	0.1010			0.1007	0.1014	0.1009			0.1010	0.1007	0.1012
	Bias			0.0044	0.0028	0.0010			0.0007	0.0014	0.0009			0.0010	0.0007	0.0012
	SD			0.0743	0.0826	0.0351			0.0443	0.0475	0.0204			0.0306	0.0314	0.0144
	RMSE			0.0744	0.0826	**0.0351**			0.0443	0.0475	**0.0204**			0.0306	0.0314	**0.0144**
Rho=0.1	**Mean**			0.0987	0.0987	0.1008			0.0988	0.0985	0.0996			0.1003	0.1003	0.0999
	Bias			-0.0013	-0.0013	0.0008			-0.0012	-0.0015	-0.0004			0.0003	0.0003	-0.0001
	SD			0.0565	0.0617	0.0394			0.0336	0.0359	0.0235			0.0228	0.0235	0.0160
	RMSE			0.0565	0.0617	**0.0394**			0.0336	0.0359	**0.0235**			0.0228	0.0235	**0.0160**
Beta=1	**Mean**	1.0016	1.0023	0.9980	0.9986	1.0056	1.0021	1.0024	1.0001	1.0007	1.0065	1.0017	1.0017	1.0001	1.0000	1.0057
	Bias	0.0016	0.0023	-0.0020	-0.0014	0.0056	0.0021	0.0024	0.0001	0.0007	0.0065	0.0017	0.0017	0.0001	0.0000	0.0057
	SD	0.0464	0.0499	0.0451	0.0495	0.0337	0.0257	0.0267	0.0257	0.0271	0.0200	0.0165	0.0167	0.0162	0.0166	0.0127
	RMSE	0.0464	0.0500	0.0451	0.0495	**0.0342**	0.0258	0.0268	0.0257	0.0271	**0.0210**	0.0166	0.0168	0.0162	0.0166	**0.0140**

Table 2.22. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 5$, $\gamma = 0.3$, $\lambda = 0.1$, $\varrho = 0.3$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML
Gamma=0.3	Mean	0.2787	0.2777	0.2023	0.2010	0.1604	0.3138	0.3133	0.2580	0.2571	0.1695	0.3261	0.3249	0.2824	0.2808	0.1721
	Bias	-0.0213	-0.0223	-0.0977	-0.0990	-0.1396	0.0138	0.0133	-0.0420	-0.0429	-0.1305	0.0261	0.0249	-0.0176	-0.0192	-0.1279
	SD	0.2522	0.2738	0.1941	0.2143	0.0906	0.1376	0.1480	0.1170	0.1253	0.0507	0.0848	0.0880	0.0741	0.0778	0.0311
	RMSE	0.2531	0.2747	0.2173	0.2361	**0.1664**	0.1383	0.1485	**0.1243**	0.1325	0.1400	0.0887	0.0915	**0.0762**	0.0801	0.1316
Lambda=0.1	Mean			0.1685	0.1649	0.1039			0.1344	0.1375	0.0923			0.1211	0.1226	0.0924
	Bias			0.0685	0.0649	0.0039			0.0344	0.0375	-0.0077			0.0211	0.0226	-0.0076
	SD			0.3836	0.4269	0.1006			0.2979	0.3199	0.0682			0.2191	0.2309	0.0489
	RMSE			0.3897	0.4318	**0.1007**			0.2999	0.3221	**0.0686**			0.2201	0.2320	**0.0495**
Rho=0.3	Mean			0.2624	0.2607	0.2854			0.2764	0.2782	0.2895			0.2862	0.2866	0.2916
	Bias			-0.0376	-0.0393	-0.0146			-0.0236	-0.0218	-0.0105			-0.0138	-0.0134	-0.0084
	SD			0.3065	0.3451	0.1465			0.2065	0.2225	0.0897			0.1359	0.1399	0.0599
	RMSE			0.3088	0.3473	**0.1472**			0.2078	0.2235	**0.0903**			0.1366	0.1406	**0.0605**
Beta=1	Mean	0.9695	0.9594	0.9490	0.9495	1.0087	0.9772	0.9744	0.9667	0.9654	1.0078	0.9760	0.9733	0.9832	0.9788	1.0089
	Bias	-0.0305	-0.0406	-0.0510	-0.0505	0.0087	-0.0228	-0.0256	-0.0333	-0.0346	0.0078	-0.0240	-0.0267	-0.0168	-0.0212	0.0089
	SD	0.4239	0.4609	0.2880	0.3253	0.1368	0.2216	0.2357	0.1721	0.1847	0.0787	0.1404	0.1436	0.1097	0.1147	0.0478
	RMSE	0.4250	0.4627	0.2925	0.3292	**0.1371**	0.2228	0.2371	0.1753	0.1879	**0.0791**	0.1424	0.1460	0.1109	0.1166	**0.0486**

Table 2.23. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 10$, $\gamma = 0.3$, $\lambda = 0.1$, $\varrho = 0.3$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML
Gamma=0.3	Mean	0.3305	0.3287	0.2760	0.2780	0.2355	0.3377	0.3380	0.2921	0.2942	0.2405	0.3365	0.3355	0.2951	0.2961	0.2418
	Bias	0.0305	0.0287	-0.0240	-0.0220	-0.0645	0.0377	0.0380	-0.0079	-0.0058	-0.0595	0.0365	0.0355	-0.0049	-0.0039	-0.0582
	SD	0.1166	0.1269	0.0979	0.1098	0.0549	0.0629	0.0661	0.0534	0.0567	0.0320	0.0410	0.0423	0.0351	0.0362	0.0197
	RMSE	0.1205	0.1301	0.1008	0.1120	**0.0847**	0.0733	0.0762	**0.0540**	0.0570	0.0676	0.0549	0.0552	**0.0355**	0.0364	0.0615
Lambda=0.1	Mean			0.1179	0.1190	0.0998			0.1100	0.1112	0.0969			0.1056	0.1060	0.0989
	Bias			0.0179	0.0190	-0.0002			0.0100	0.0112	-0.0031			0.0056	0.0060	-0.0011
	SD			0.2301	0.2550	0.0731			0.1470	0.1544	0.0487			0.0993	0.1024	0.0322
	RMSE			0.2308	0.2557	**0.0731**			0.1474	0.1548	**0.0488**			0.0994	0.1026	**0.0323**
Rho=0.3	Mean			0.2894	0.2905	0.2945			0.2944	0.2917	0.2966			0.2985	0.2966	0.2986
	Bias			-0.0106	-0.0095	-0.0055			-0.0056	-0.0083	-0.0034			-0.0015	-0.0034	-0.0014
	SD			0.1607	0.1790	0.0910			0.1013	0.1088	0.0568			0.0678	0.0703	0.0383
	RMSE			0.1611	0.1793	**0.0912**			0.1014	0.1092	**0.0569**			0.0678	0.0704	**0.0383**
Beta=1	Mean	0.9777	0.9758	0.9873	0.9883	1.0164	0.9801	0.9786	0.9946	0.9922	1.0180	0.9801	0.9792	0.9968	0.9962	1.0181
	Bias	-0.0223	-0.0242	-0.0127	-0.0117	0.0164	-0.0199	-0.0214	-0.0054	-0.0078	0.0180	-0.0199	-0.0208	-0.0032	-0.0038	0.0181
	SD	0.1539	0.1679	0.1402	0.1558	0.0823	0.0853	0.0889	0.0788	0.0831	0.0475	0.0540	0.0551	0.0502	0.0518	0.0296
	RMSE	0.1555	0.1697	0.1407	0.1563	**0.0839**	0.0876	0.0914	0.0790	0.0834	**0.0508**	0.0575	0.0589	0.0503	0.0520	**0.0347**

Table 2.24. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 50$, $\gamma = 0.3$, $\lambda = 0.1$, $\varrho = 0.3$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**	**One-step AB**	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**	**One-step AB**	**Two-step AB**	**One-step ESAB**	**Two-step ESAB**	**QML**
Gamma=0.3	**Mean**	0.3580	0.3570	0.2975	0.2982	0.2879	0.3498	0.3485	0.2988	0.2996	0.2878	0.3474	0.3463	0.2996	0.2999	0.2887
	Bias	0.0580	0.0570	-0.0025	-0.0018	-0.0121	0.0498	0.0485	-0.0012	-0.0004	-0.0122	0.0474	0.0463	-0.0004	-0.0001	-0.0113
	SD	0.0362	0.0387	0.0306	0.0338	0.0220	0.0207	0.0221	0.0177	0.0189	0.0130	0.0126	0.0133	0.0109	0.0113	0.0080
	RMSE	0.0683	0.0689	0.0307	0.0339	**0.0251**	0.0540	0.0533	**0.0177**	0.0189	0.0178	0.0490	0.0482	**0.0109**	0.0113	0.0138
Lambda=0.1	**Mean**			0.1032	0.1015	0.1009			0.1001	0.1009	0.1011			0.1007	0.1004	0.1013
	Bias			0.0032	0.0015	0.0009			0.0001	0.0009	0.0011			0.0007	0.0004	0.0013
	SD			0.0756	0.0835	0.0346			0.0453	0.0486	0.0202			0.0311	0.0319	0.0141
	RMSE			0.0757	0.0835	**0.0346**			0.0453	0.0486	**0.0203**			0.0311	0.0319	**0.0142**
Rho=0.3	**Mean**			0.2980	0.2976	0.3002			0.2987	0.2985	0.2994			0.3001	0.3001	0.2995
	Bias			-0.0020	-0.0024	0.0002			-0.0013	-0.0015	-0.0006			0.0001	0.0001	-0.0005
	SD			0.0574	0.0631	0.0386			0.0340	0.0362	0.0232			0.0232	0.0239	0.0156
	RMSE			0.0574	0.0631	**0.0386**			0.0340	0.0362	**0.0233**			0.0232	0.0239	**0.0156**
Beta=1	**Mean**	0.9831	0.9835	0.9979	0.9986	1.0056	0.9854	0.9854	1.0001	1.0007	1.0064	0.9860	0.9858	1.0001	1.0001	1.0057
	Bias	-0.0169	-0.0165	-0.0021	-0.0014	0.0056	-0.0146	-0.0146	0.0001	0.0007	0.0064	-0.0140	-0.0142	0.0001	0.0001	0.0057
	SD	0.0474	0.0516	0.0452	0.0495	0.0337	0.0260	0.0281	0.0257	0.0272	0.0199	0.0168	0.0176	0.0162	0.0167	0.0127
	RMSE	0.0503	0.0541	0.0453	0.0495	**0.0341**	0.0298	0.0317	0.0257	0.0272	**0.0209**	0.0218	0.0226	0.0162	0.0167	**0.0139**

Table 2.25. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 5$, $\gamma = 0.3$, $\lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML
Gamma=0.3	**Mean**	0.2973	0.2988	0.2027	0.2008	0.1604	0.3345	0.3330	0.2582	0.2574	0.1688	0.3449	0.3437	0.2826	0.2810	0.1712
	Bias	-0.0027	-0.0012	-0.0973	-0.0992	-0.1396	0.0345	0.0330	-0.0418	-0.0426	-0.1312	0.0449	0.0437	-0.0174	-0.0190	-0.1288
	SD	0.2478	0.2651	0.1973	0.2155	0.0921	0.1395	0.1499	0.1188	0.1285	0.0511	0.0872	0.0898	0.0738	0.0771	0.0317
	RMSE	0.2478	0.2651	0.2200	0.2372	**0.1672**	0.1437	0.1535	**0.1259**	0.1354	0.1408	0.0981	0.0999	**0.0758**	0.0794	0.1326
Lambda=0.3	**Mean**			0.4352	0.4316	0.2825			0.3834	0.3871	0.2888			0.3474	0.3493	0.2945
	Bias			0.1352	0.1316	-0.0175			0.0834	0.0871	-0.0112			0.0474	0.0493	-0.0055
	SD			0.3435	0.3778	0.1158			0.2723	0.2929	0.0704			0.1950	0.2058	0.0457
	RMSE			0.3691	0.4001	**0.1172**			0.2847	0.3056	**0.0713**			0.2007	0.2116	**0.0460**
Rho=0.1	**Mean**			0.0814	0.0833	0.1277			0.0796	0.0816	0.1254			0.0890	0.0897	0.1269
	Bias			-0.0186	-0.0167	0.0277			-0.0204	-0.0184	0.0254			-0.0110	-0.0103	0.0269
	SD			0.2968	0.3280	0.1425			0.1977	0.2147	0.0875			0.1304	0.1337	0.0588
	RMSE			0.2974	0.3285	**0.1452**			0.1987	0.2155	**0.0911**			0.1309	0.1341	**0.0647**
Beta=1	**Mean**	1.0283	1.0152	0.9427	0.9434	1.0108	1.0329	1.0288	0.9636	0.9626	1.0098	1.0283	1.0248	0.9821	0.9775	1.0105
	Bias	0.0283	0.0152	-0.0573	-0.0566	0.0108	0.0329	0.0288	-0.0364	-0.0374	0.0098	0.0283	0.0248	-0.0179	-0.0225	0.0105
	SD	0.4468	0.4857	0.2914	0.3197	0.1373	0.2344	0.2464	0.1725	0.1856	0.0789	0.1486	0.1513	0.1081	0.1127	0.0480
	RMSE	0.4477	0.4859	0.2970	0.3247	**0.1378**	0.2367	0.2481	0.1763	0.1894	**0.0795**	0.1513	0.1534	0.1096	0.1149	**0.0491**

Table 2.26. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 10$, $\gamma = 0.3$, $\lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML
Gamma=0.3	**Mean**	0.3553	0.3535	0.2755	0.2774	0.2355	0.3599	0.3607	0.2920	0.2941	0.2402	0.3573	0.3566	0.2952	0.2963	0.2414
	Bias	0.0553	0.0535	-0.0245	-0.0226	-0.0645	0.0599	0.0607	-0.0080	-0.0059	-0.0598	0.0573	0.0566	-0.0048	-0.0037	-0.0586
	SD	0.1163	0.1260	0.0983	0.1096	0.0555	0.0635	0.0659	0.0537	0.0573	0.0325	0.0411	0.0418	0.0353	0.0364	0.0200
	RMSE	0.1287	0.1369	0.1013	0.1119	**0.0851**	0.0873	0.0895	**0.0543**	0.0576	0.0680	0.0705	0.0704	**0.0356**	0.0365	0.0619
Lambda=0.3	**Mean**			0.3527	0.3562	0.2922			0.3252	0.3264	0.2964			0.3127	0.3132	0.2989
	Bias			0.0527	0.0562	-0.0078			0.0252	0.0264	-0.0036			0.0127	0.0132	-0.0011
	SD			0.2060	0.2255	0.0762			0.1319	0.1397	0.0455			0.0898	0.0928	0.0292
	RMSE			0.2126	0.2324	**0.0766**			0.1343	0.1421	**0.0456**			0.0907	0.0937	**0.0292**
Rho=0.1	**Mean**			0.0863	0.0867	0.1127			0.0923	0.0897	0.1120			0.0979	0.0961	0.1135
	Bias			-0.0137	-0.0133	0.0127			-0.0077	-0.0103	0.0120			-0.0021	-0.0039	0.0135
	SD			0.1540	0.1717	0.0908			0.0977	0.1054	0.0559			0.0656	0.0681	0.0374
	RMSE			0.1546	0.1722	**0.0917**			0.0980	0.1059	**0.0572**			0.0656	0.0682	**0.0397**
Beta=1	**Mean**	1.0343	1.0342	0.9841	0.9848	1.0178	1.0294	1.0283	0.9935	0.9910	1.0190	1.0267	1.0262	0.9961	0.9957	1.0189
	Bias	0.0343	0.0342	-0.0159	-0.0152	0.0178	0.0294	0.0283	-0.0065	-0.0090	0.0190	0.0267	0.0262	-0.0039	-0.0043	0.0189
	SD	0.1658	0.1801	0.1390	0.1528	0.0825	0.0902	0.0943	0.0782	0.0827	0.0478	0.0568	0.0581	0.0500	0.0519	0.0297
	RMSE	0.1693	0.1833	0.1399	0.1535	**0.0844**	0.0949	0.0984	0.0785	0.0832	**0.0514**	0.0628	0.0637	0.0502	0.0520	**0.0352**

Table 2.27. Mean, bias, standard deviation and root mean square error of various non-spatial and spatial estimators in a stationary scenario, for $T = 50$, $\gamma = 0.3$, $\lambda = 0.3$, $\varrho = 0.1$, $\beta = 1$ and various N, over 999 iterations

		N=16					N=49					N=121				
		One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML	One-step AB	Two-step AB	One-step ESAB	Two-step ESAB	QML
Gamma=0.3	Mean	0.3824	0.3825	0.2974	0.2983	0.2879	0.3713	0.3704	0.2987	0.2995	0.2878	0.3674	0.3667	0.2996	0.2999	0.2887
	Bias	0.0824	0.0825	-0.0026	-0.0017	-0.0121	0.0713	0.0704	-0.0013	-0.0005	-0.0122	0.0674	0.0667	-0.0004	-0.0001	-0.0113
	SD	0.0370	0.0392	0.0306	0.0339	0.0224	0.0211	0.0222	0.0180	0.0192	0.0132	0.0128	0.0133	0.0109	0.0114	0.0081
	RMSE	0.0903	0.0913	0.0307	0.0340	**0.0254**	0.0744	0.0738	0.0180	0.0192	**0.0180**	0.0686	0.0680	**0.0110**	0.0114	0.0139
Lambda=0.3	Mean			0.3085	0.3071	0.3001			0.3021	0.3028	0.3004			0.3016	0.3013	0.3006
	Bias			0.0085	0.0071	0.0001			0.0021	0.0028	0.0004			0.0016	0.0013	0.0006
	SD			0.0699	0.0772	0.0317			0.0419	0.0450	0.0185			0.0290	0.0297	0.0132
	RMSE			0.0704	0.0776	**0.0317**			0.0419	0.0451	**0.0185**			0.0290	0.0297	**0.0132**
Rho=0.1	Mean			0.0961	0.0955	0.1031			0.0982	0.0978	0.1023			0.0998	0.0997	0.1022
	Bias			-0.0039	-0.0045	0.0031			-0.0018	-0.0022	0.0023			-0.0002	-0.0003	0.0022
	SD			0.0562	0.0619	0.0379			0.0337	0.0359	0.0228			0.0229	0.0236	0.0154
	RMSE			0.0564	0.0620	**0.0381**			0.0338	0.0360	**0.0230**			0.0229	0.0236	**0.0155**
Beta=1	Mean	1.0343	1.0356	0.9971	0.9978	1.0058	1.0303	1.0312	0.9999	1.0005	1.0066	1.0282	1.0287	0.9999	0.9999	1.0057
	Bias	0.0343	0.0356	-0.0029	-0.0022	0.0058	0.0303	0.0312	-0.0001	0.0005	0.0066	0.0282	0.0287	-0.0001	-0.0001	0.0057
	SD	0.0515	0.0559	0.0454	0.0500	0.0338	0.0277	0.0292	0.0261	0.0276	0.0201	0.0178	0.0186	0.0164	0.0169	0.0128
	RMSE	0.0619	0.0663	0.0455	0.0501	**0.0343**	0.0410	0.0427	0.0261	0.0276	**0.0211**	0.0333	0.0342	0.0164	0.0169	**0.0140**

3 The determinants of cropland values in Midwestern U.S.A.

3.1 Introduction

Farm real estate represents a dominant asset on the farm sector balance sheet in the U.S.A. (it accounted for nearly 84% of total U.S. farm assets in 2009) and is usually the largest investment in the farmers' portfolio: it is therefore considered to be an important indicator of the performance of the sector and of the producers' welfare (Nickerson et al. 2012). The real values of agricultural land have been increasing dramatically in recent years, particularly starting from the second half of 2000s, raising many questions about their macroeconomic determinants and whether the boom will turn into a bust (Gloy 2013), especially after the financial crisis that invested the U.S.A. and the rest of the world in 2007. The analysis of land values also raises a number of policy issues, regarding government support, taxation and environmental protection.

For all these reasons, the empirical literature on the determinants of agricultural land values is wide and the economic theory has frequently addressed the topic. The relationship between the farmland prices and the expected future returns on this asset have been extensively investigated in the past (see, for example, Falk 1991; Engsted 1998; Lence and Miller 1999) and the topic is currently widely addressed. However, despite the great amount of economic research efforts, most economic theories have only met small empirical evidence (Gutierrez et al. 2007). Among the most popular theoretical economic models that have addressed the topic of land values behavior in the long-run, one is the Present Value Model (PVM), which is reviewed in section 3.2 both from a theoretical and an empirical point of view.

The purpose of this chapter is to investigate the spatial effects that may characterize the process of determination of agricultural land values in

Midwestern U.S.A., by adopting the PVM as the theoretical framework. In order to do so, we choose to conveniently specify and estimate a time-space dynamic model that relates land value to its determinants. The employed dataset is presented in section 3.3 and the spatial characteristics of the data are explored in section 3.4. The model is then estimated by the QML estimator that was extensively analyzed in chapters 1 and 2 and the results are given and discussed in section 3.5. Section 3.6 presents the necessary checks of the stability conditions for the estimated model and the computation of long-run elasticities of cropland value with respect to the included regressors. Section 3.7 contains the final concluding remarks and the discussion of possible future developments.

3.2 The present value model

3.2.1 The theoretical model

The PVM (Campbell and Shiller 1988; Campbell et al. 1997) is a financial model that relates the price of a stock to its expected future returns discounted to the present using a constant or time-varying discount rate. It is a model that deals with long-horizon asset returns: since dividends in all future periods enter the present-value model, the dividend in any single period is only a small component of the price and therefore persistent movements have much more influence on prices than short-term, temporary variations do. When applied to the analysis of land values, we consider the price of the stock to be the price of land (in our case, the value of cropland, CV); the dividends are measured as cash rents (CR) received by the land owners. The value of cropland is therefore related to the capitalized value of the current and future stream of cash rents.

Let the *net simple return* (R_{t+1}) on a stock be defined as

$$R_{t+1} \equiv (CV_{t+1} + CR_{t+1})/CV_t - 1, \tag{3.1}$$

where we assume that the dividend (CR_{t+1}) is paid just before the price (CV_{t+1}) is recorded, so that CV_{t+1} is taken to be an *ex-dividend* price at time $t+1$, for $t = 1, \dots, T$. The simple gross return is defined as $1 + R_{t+1}$ and this makes clear that an asset's gross return over the K more recent periods, $1 + R_{t+1}(K)$, is defined as the product of the K single period returns from $t - K + 1$ to t (*compound returns*):

$$1 + R_{t+1}(K) = (1 + R_t) \cdot (1 + R_{t-1}) \cdot \dots \cdot (1 + R_{t-K+1}). \tag{3.2}$$

Because of the presence of a ratio in equation (3.1), any averaging would require a geometric averaging. This motivates the alternative definition of *continuously compounded returns* or *log returns* of an asset, which is defined as the natural logarithm of its gross return:

$$r_{t+1} \equiv log(1 + R_{t+1}). \tag{3.3}$$

The lowercases letters will denote natural logarithms of the variables from now on.

If we assume constant expected returns, such that $\mathrm{E}_t(R_{t+1}) = R$, we obtain an equation that relates the current stock prices to the stock price and future payoffs in the next period:

$$CV_t = \mathrm{E}_t\left(\frac{CV_{t+1} + CR_{t+1}}{1+R}\right). \tag{3.4}$$

In order to eliminate future-dated expectations, equation (3.4) should be solved by repeatedly substituting out future prices and using the Law of Iterated Expectations ($E_t\left[E_{t+1}[X]\right] = E_t[X]$). After solving for K periods, we have:

$$CV_t = \mathrm{E}_t\left[\sum_{i=1}^{K}\left(\frac{1}{1+R}\right)^i CR_{t+i}\right] + \mathrm{E}_t\left[\left(\frac{1}{1+R}\right)^K CV_{t+K}\right], \tag{3.5}$$

where the second term on the RHS of equation (3.5) represents the discounted value of the stock price, K periods from the present. Assuming that, as the time horizon increases, this term shrinks to zero and that $K \rightarrow$

∞,[11] we can express the stock price as the expected present value of future dividends (CV_{CRt}) out to the infinite future, discounted at a constant rate:

$$CV_t = CV_{CRt} \equiv \mathrm{E}_t\left[\sum_{i=1}^{\infty}\left(\frac{1}{1+R}\right)^i CR_{t+i}\right]. \tag{3.6}$$

The stock price CV_t will follow a linear process with a unit root (also known as *integrated* process) if the dividend CR_t itself follows a linear process with a unit root, which means that shocks would have permanent effects on the level of the variable but not on the change in the variable. If this is the case, the formula in equation (3.6) can be transformed to a relation between two stationary variables by subtracting a multiple of the dividend from both sides of the equation:

$$CV_t - \frac{CR_t}{R} = \left(\frac{1}{R}\right)\mathrm{E}_t\left[\sum_{i=1}^{\infty}\left(\frac{1}{1+R}\right)^i \Delta CR_{t+1+i}\right]. \tag{3.7}$$

In this case, even if the dividend and the price processes are not stationary, there is a stationary linear combination of prices and dividends that makes the two series *cointegrated*.

Differently and more realistically, when we assume time-varying expected stock returns, the relationship between prices and returns is non-linear, therefore a log-linear approximation of the model should be considered to be more appropriate. According to the model proposed by Campbell and Shiller (1988) and equation (3.3), we define the log of the gross real rate of return as

$$r_{t+1} \equiv log(CV_{t+1} + CR_{t+1}) - log(CV_t) \tag{3.8}$$

or equivalently

$$r_{t+1} \equiv cv_{t+1} - cv_t + log(1 + exp(s_{t+1})), \tag{3.9}$$

[11] This assumption imposes a *transversality condition* that excludes the presence of a rational bubble. The second term on the RHS of equation (3.5) is indeed consistent with rational expectations and constant expected returns. Excluding a rational bubble means to exclude "financial exuberance episodes in which investors appeared to be betting that other investors drive prices even higher in the future, far higher than explained by fundamentals" (Gutierrez 2011).

where $s_{t+1} = cr_{t+1} - cv_{t+1}$ is the natural logarithm of the dividend-price ratio (CR_{t+1}/CV_{t+1}), which is also called *spread* in financial literature.

Equation (3.9) can be linearized using a first-order Taylor expansion into

$$r_{t+1} \approx k + s_t - \rho s_{t+1} + \Delta cr_{t+1}, \tag{3.10}$$

where $k = -log(\rho) - (1-\rho) \cdot log(1/\rho - 1)$ and $\rho = 1/(1 + CR/CV)$. One should notice that equation (3.10) is a linear difference equation for the log stock price analogous to the one that was obtained in (3.4) under the assumption of constant expected returns. It can be solved forwardly and, under the condition that $\lim_{j\to\infty} \rho^j\, s_{t+j} = 0$, we obtain

$$s_t \approx -k/(1-\rho) - \sum_{j=0}^{\infty} \rho^j\, (\Delta cr_{t+1+j} - r_{t+1+j}). \tag{3.11}$$

According to equation (3.11), if the stock price is high today, then there must be some combination of high dividends and low stock returns in the future (Campbell et al. 1997, page 263). This relation holds ex-ante as much as ex-post, therefore taking expectations we obtain

$$s_t + k/(1-\rho) \approx -\mathrm{E}_t\left[\sum_{j=0}^{\infty} \rho^j\, (\Delta cr_{t+1+j} - r_{t+1+j})\right]. \tag{3.12}$$

The rationale of the PVM is embodied in equation (3.12) as it expresses the current value of the dividend-price ratio in terms of the present discounted value of expected future values of Δcr_{t+1} and r_{t+1} (Gutierrez et al. 2007, page 164). The log dividend-price ratio is high only when dividends are expected to grow slowly or the expected stock returns are high and, when the dividend follows a log-linear unit-root process, the log dividend-price ratio is stationary provided that the expected stock return is stationary (Campbell et al. 1997). According to the PVM, if the agents are fully rational, then the asset prices (e.g. farmland values) and the dividends generated from that asset (e.g. cash rents) cannot drift persistently far apart from each other.

Let us also assume that the return to our asset $E_t[r_t]$ exceeds the expected return of another asset $E_t[g_t]$ by a constant r that represents the risk premium on investments on our asset; the PVM reduces to

$$s_t + (k - r)/(1 - \rho) \approx \mathrm{E}_t\left[\sum_{j=0}^{\infty} \rho^j \left(g_{t+1+j} - \Delta cr_{t+1+j}\right)\right]. \quad (3.13)$$

By supposing further that the expected rate of return on the alternative asset is stationary and that the logs of dividends and prices are non-stationary but their differences are, then it should be concluded that the RHS of equation (3.13) is stationary too and the constant excess returns PVM holds. According to this finding, the PVM has been tested in the literature by estimating and then testing for cointegration the following equation

$$cv_t = \alpha + \beta cr_t + \varepsilon_t, \quad (3.14)$$

where $\alpha = -(k - r)/(1 - \rho)$ and ε is a zero-mean disturbance, or equivalently

$$s_t - \alpha = (1 - \beta) cr_t - \varepsilon_t. \quad (3.15)$$

If $\beta = 1$, intuitively, the log prices move one-to-one with log dividends and their unit-root components cancel out thus leaving the spread unaffected. On the contrary, if $\beta \neq 1$, then $(1 - \beta)cr_t$ does not disappear and the spread is non-stationary (Gutierrez et al. 2007).

3.2.2 *Empirical literature on the PVM and farmland prices*

Many empirical studies on the determinants of farmland prices refer to the PVM as their theoretical framework. According to it, the value of an income-producing asset such as farmland is the capitalized value of the current and future stream of earnings from owing that asset (often measured, not exclusively, as cash rents). In other words, land values should equal the present value of all future expected cash flows stemming from a productive use of that land and therefore changes in expected returns to farming should explain changes in farmland prices (Du et al. 2007).

The empirical testing of the PVM has consisted in estimating equation (3.14) for each cross-sectional unit i and then testing the stationarity of the residuals by means of conventional cointegration tests. However, the empirical results do not fully support the PVM as the most appropriate for explaining farmland values. Among the empirical studies on this topic, we recall the analysis on farmland prices in Iowa conducted by Falk (1991), that ended up rejecting the PVM because, although highly correlated, farmland price and rent movements are not consistent with that. Clark et al. (1993) found similar results for Illinois, Tegene and Kuchler (1993) and Engsted (1998) for three U.S. regions (the Lake States, the Corn Belt and the Northern Plains). The failure to find cointegration is addressed by Gutierrez et al. (2007) by allowing structural breaks in the cointegrated relationship that represent a shifting risk premium on farmland investments, thus finding results in favor of the PVM.

Moving from the classical literature on PVM, some other trends have been gaining popularity in the analysis of farmland value. Some researchers concentrated on the influence of urbanization (Hardie et al. 2001; Plantinga et al. 2002; Livanis et al. 2006 among others); others focused on the testing of the PVM in presence of transaction costs (Lence and Miller 1999; de Fontnouvelle and Lence 2001). Important contributions tended to make distinctions among the streams of rents, particularly by arguing that farmland rents do not only consist in cash rents and that government payments should be considered as rent sources, but also distinguishing between different types of public subsidies (Clark et al. 1993; Weersink et al. 1999; Goodwin et al. 2003 among the others).

3.3 The data

All the employed data for the agricultural sector are made available by the United States Department of Agriculture (USDA), National Agricultural Statistics Service[12] (NASS) and Economic Research Service[13]. The estimates of land values are based on annual survey data and report the market value[14] per acre of cropland only (in current dollars), so that problems arising from heterogeneity in land quality and use are limited (pastureland, for example, is not included). Cropland only includes the land used to grow field crops, vegetables or land harvested for hay. This also permits to exclude the value of farm buildings and take the value of land only into consideration.

Net cash rents are also estimated only for cropland from data on gross cash rents (in current dollars). Net cash rents are used to measure returns from land, that is from agricultural production, and can be interpreted as a Ricardian land rent. Besides this type of rent, agricultural support programs also represent a land return which may capitalize into land value. Direct government payments per acre of cropland, as estimated by the USDA-Economic Research Service, are therefore used as explanatory variables.

All monetary variables were deflated using the GDP implicit price deflator (reference year 2005) from the U.S. Department of Commerce, Bureau of Economic Analysis.

Population density, calculated from the annual estimate of population from the U.S. Department of Commerce, Bureau of Census, is included among the covariates of the model as a proxy for urban pressure, that represents competing demand for land for non-agricultural use (Feichtinger and Salhofer 2011).

[12] http://www.nass.usda.gov/Quick_Stats/

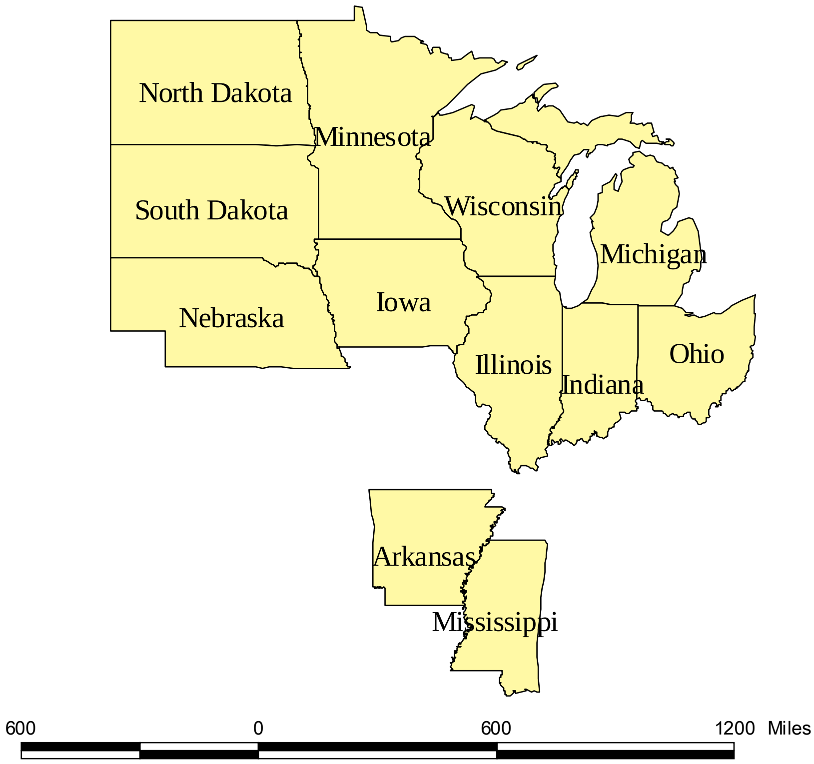
[13] We thank Doctor Kenneth Erickson for making the dataset available for this research through a patient and thorough collection and check of the data.

[14] The land value is the value at which the land used for agricultural production can be sold under current market conditions, if allowed to remain on the market for a reasonable amount of time (USDA-NASS 2012).

The employed dataset is a panel of annual observations for 12 U.S. States and 39 years, between 1971 and 2009. The considered States are part of the Midwestern United States (Lake States, Corn Belt States, Northern Plains and Delta States) (Figure 3.1), for which more homogeneous data are available, less affected by urban influence (like those for Northeaster States). Moreover, cropland is mostly found in the Midwest States, while the Western States, that have lower shares of cropland to total farmland, are less heavily surveyed by NASS for cash rents and the data on cropland per acre are either thinner or not available because sometimes limited only to either irrigated or non-irrigated cropland.

The availability of data on cropland value per acre for the selected variables turned out to be constraint that led to the exclusion of States such as Louisiana, Missouri and Kansas form the original dataset. The availability of data on cash rents, only limited to 2009 for South Dakota, determined the time-span.

Figure 3.1. Map of the States included in the regression analysis



3.4 Exploratory Spatial Data Analysis

Any spatial analysis requires the definition of a spatial weight matrix. As a robustness check for all the results, we employed three different definitions of neighborhood that led to the specification of some different spatial weight matrices, all of them based on geographical proximity between States. In particular, the rook, queen and distance criteria of proximity where alternatively adopted. All spatial weight matrices are row-standardized.

The elements of the distance-based spatial weight matrices before row-standardization are defined as the inverse of the squared arc distance between States i and j:

$$\begin{cases} w_{ij} = 0 & \text{if} \quad i = j \\ w_{ij} = 1/d_{ij}^2 & \text{if} \quad d_{ij} \leq D \\ w_{ij} = 0 & \text{otherwise.} \end{cases} \tag{3.16}$$

Different values were taken as cut-off distance (D): the minimum distance that allows each State to have at least one neighbor; 300 miles (about 400 km); the first quartile of the distance distribution (378.7 miles, about 609 km); the second quartile of the distance distribution (524.6 miles, about 844 km); the third quartile of the distance distribution (729.6 miles, about 1174 km).

Table 3.1. Connectivity schemes resulting from the specification of different spatial weight matrices

		Number of neighbors										
	Spatial Weight Matrix	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
Cut-off distance	*Minimum*	8	4									
	300 miles	6	4	2								
	I quartile (379 miles)	2	4	1	4	1						
	II quartile (525 miles)		2		4	1	2	2	1	1		
	III quartile (730 miles)					1	1	2	3	2	2	1
	Rook	2	3	3	3	1						
	Queen	1	5	4	3	1						

Although the resulting connectivity schemes were quite different (Table 3.1), the results of the ESDA proved to be quite robust to the choice of **W**. We therefore choose to present only the results for a distance-based spatial weight matrix with a cut-off distance fixed at the 1st quartile of the distance distribution, since we believe it represents a good average picture.

A first step in the ESDA is to determine whether there is overall spatial dependence among the observed cropland values. This is assessed through the well-known Moran's I index and scatterplot. The Moran's I index (Table 3.2) shows significant positive values for all considered years thus leading to reject the null hypothesis of no spatial dependence in favor of positive spatial dependence in the distribution of cropland values.

Table 3.2. Results for the Moran's I index for observed cropland value (1971 - 2009)

Year	*Moran's I*	*p-value*	*Year*	*Moran's I*	*p-value*
1971	0.439	0.026	*1991*	0.499	0.014
1972	0.465	0.019	*1992*	0.478	0.018
1973	0.454	0.022	*1993*	0.539	0.010
1974	0.441	0.024	*1994*	0.501	0.010
1975	0.425	0.030	*1995*	0.585	0.007
1976	0.437	0.028	*1996*	0.568	0.007
1977	0.519	0.013	*1997*	0.707	0.002
1978	0.543	0.010	*1998*	0.720	0.002
1979	0.554	0.009	*1999*	0.748	0.001
1980	0.539	0.011	*2000*	0.770	0.001
1981	0.506	0.015	*2001*	0.778	0.001
1982	0.465	0.020	*2002*	0.783	0.001
1983	0.440	0.024	*2003*	0.522	0.012
1984	0.536	0.010	*2004*	0.778	0.001
1985	0.646	0.003	*2005*	0.762	0.001
1986	0.669	0.002	*2006*	0.761	0.001
1987	0.708	0.001	*2007*	0.756	0.001
1988	0.598	0.005	*2008*	0.733	0.001
1989	0.477	0.017	2009	0.733	0.001
1990	0.511	0.013			

The same information is displayed by the Moran scatterplots (Figure 3.2).

Figure 3.2. Moran scatterplots for observed cropland value (years 1971 to 2009)

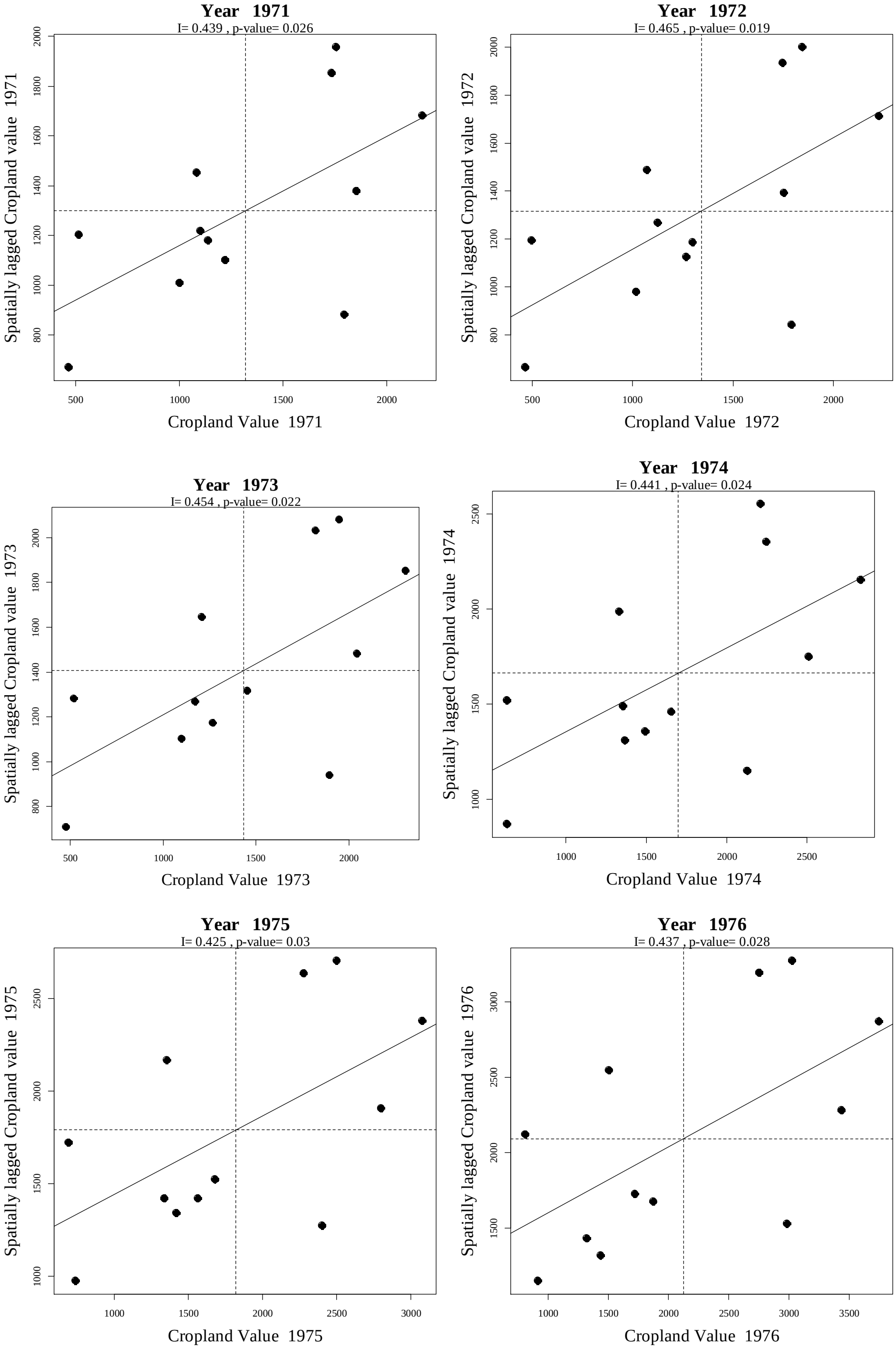


Figure 3.2. (continued)

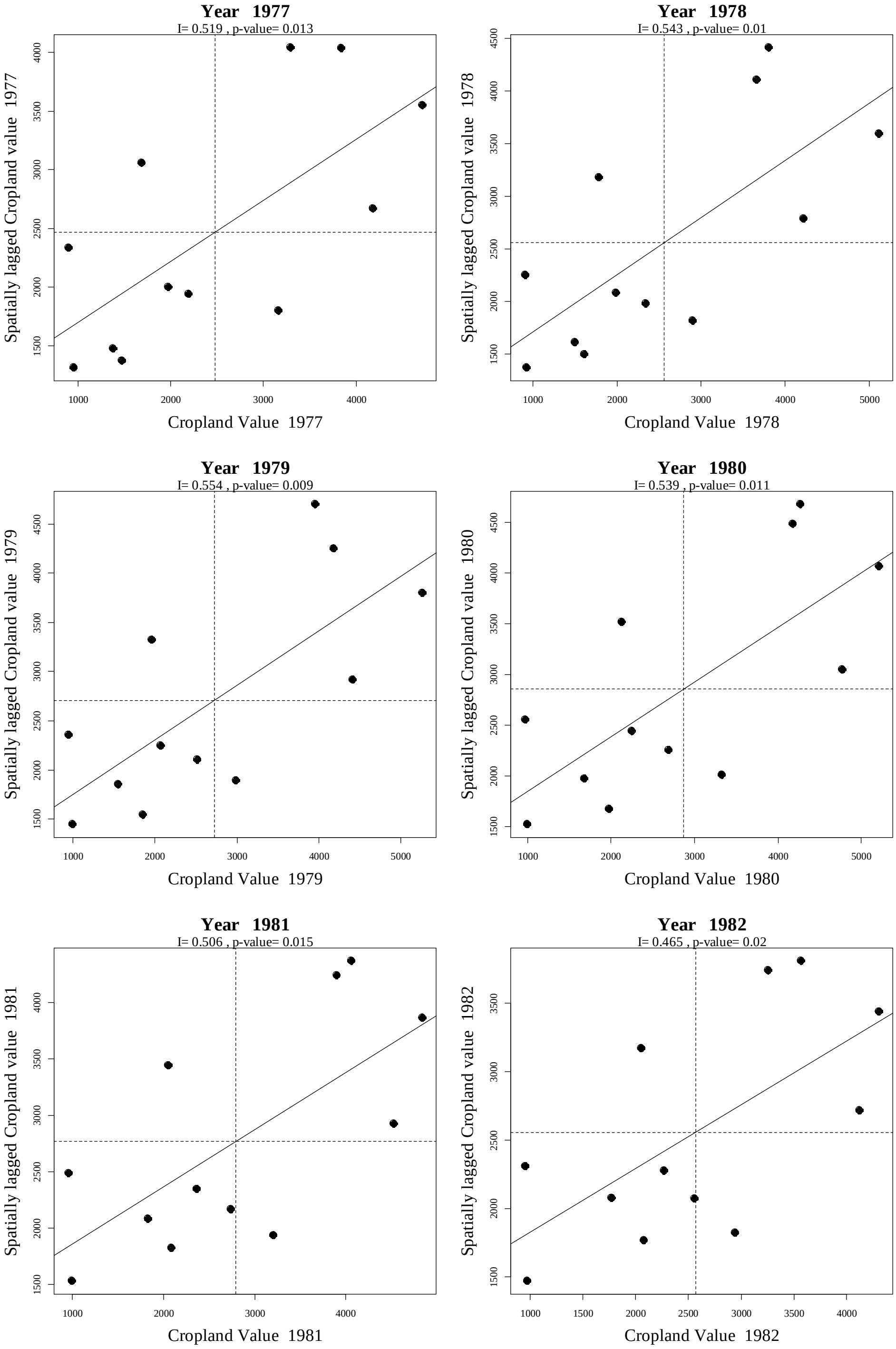


Figure 3.2. (continued)

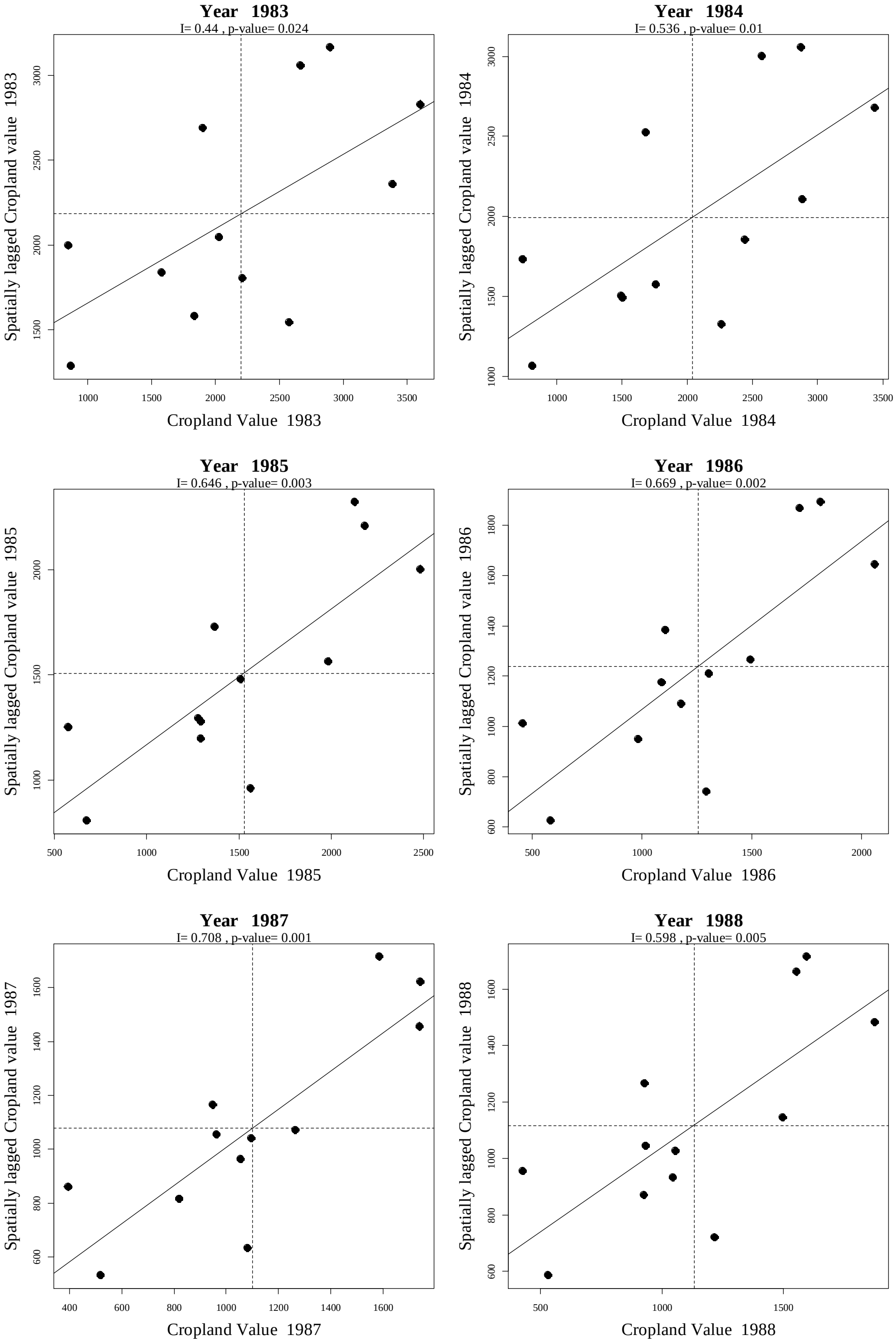


Figure 3.2. (continued)

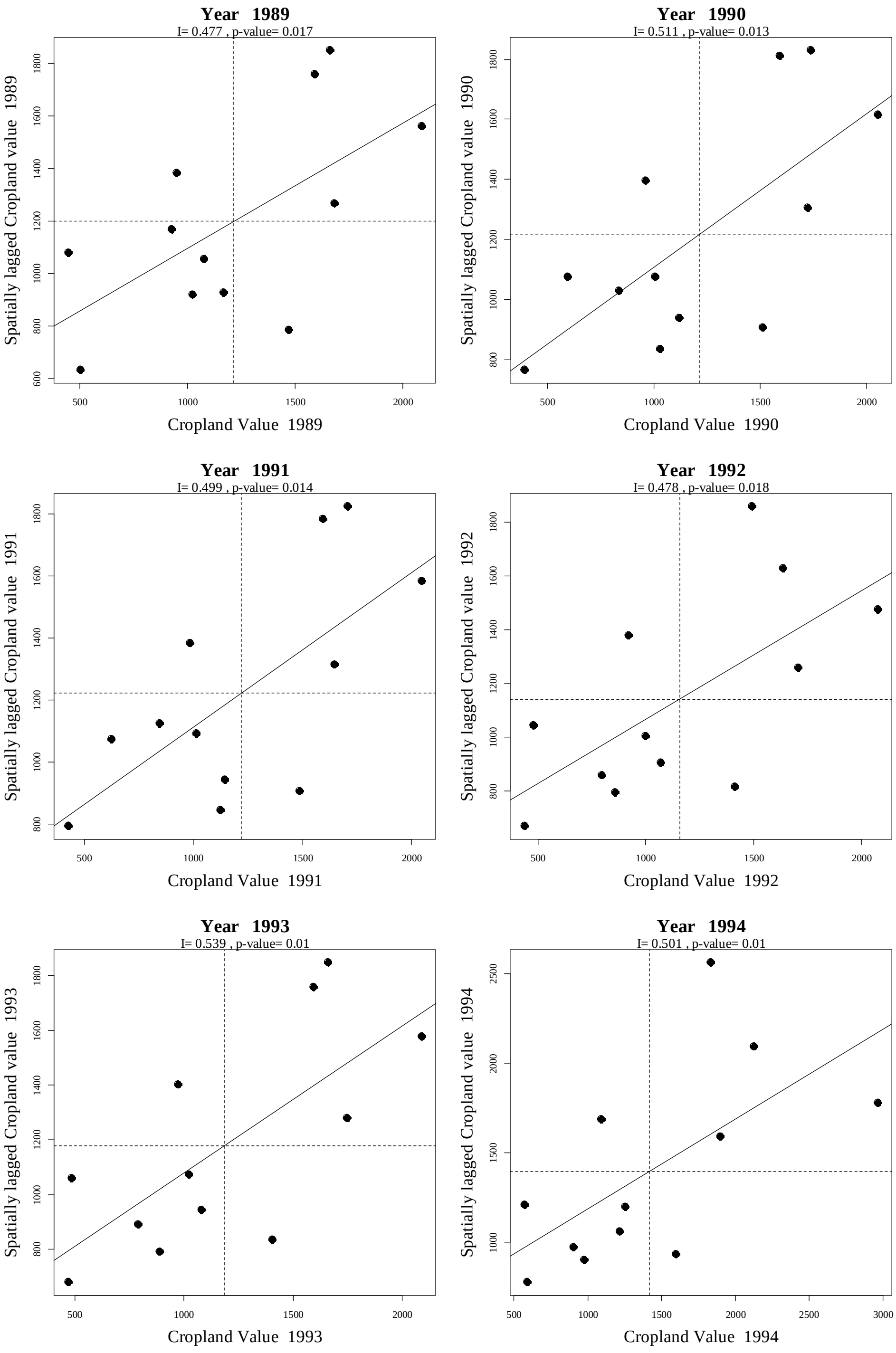


Figure 3.2. (continued)

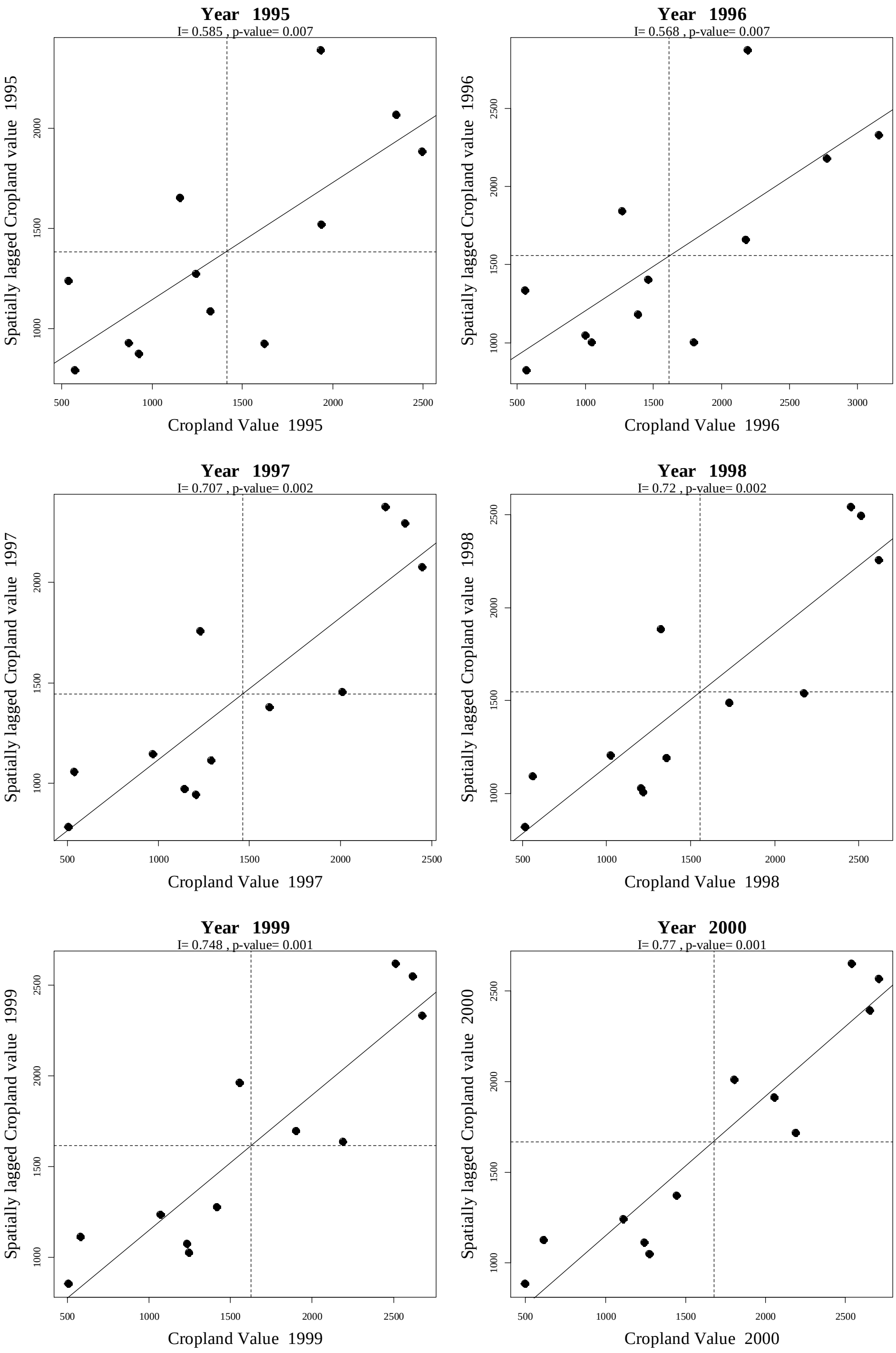


Figure 3.2. (continued)

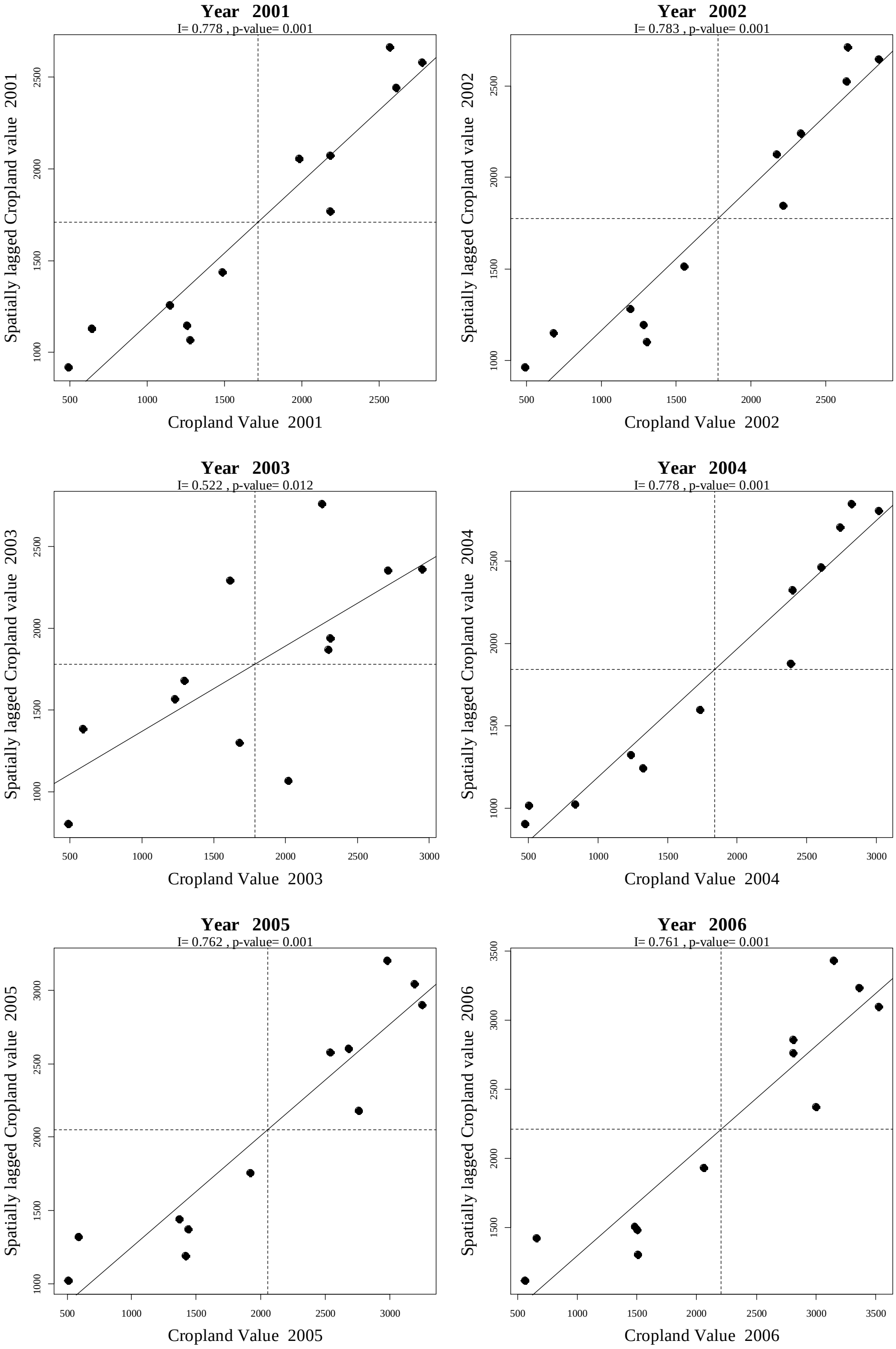
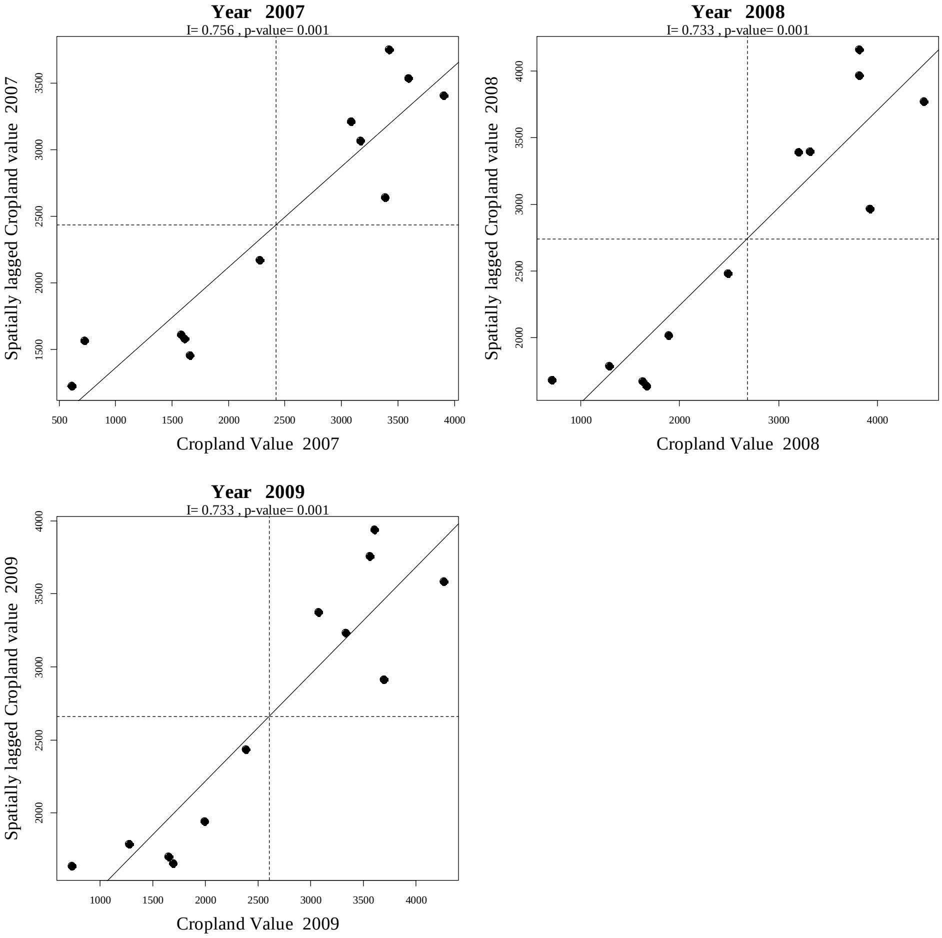


Figure 3.2. (continued)



Each quadrant in the scatterplot corresponds to a particular kind of spatial association between the observations of a given variable in a State and in its neighbors: the first and third quadrants are characterized by positive dependence between, respectively, the high/low values of the variable in one State and its neighbors. The second and fourth quadrants, on the other hand, show negative dependence. Albeit present in all considered years, spatial dependence appears to be stronger starting from the years 2000s.

The results of the ESDA therefore give clear indication in favor of the estimation of a spatial model, capable of taking the spatial dependence among the observations of the dependent variable into account.

3.5 Results and discussion

The analysis on the determinants of cropland values in 12 U.S. States over the period 1971-2009 is conducted by estimating a time-space dynamic model. Fixed individual effects are added to the specification in order to take into account unobserved time-invariant sources of heterogeneity such as climate and land quality (Kirwan 2009). Different sets of covariates were included, as described in equations (3.17) and (3.18):

$$cv_{it} = \lambda \boldsymbol{W} cv_{it} + \gamma cv_{it-1} + \varrho \boldsymbol{W} cv_{it-1} + \beta_1 cr_{it} + \beta_2 pd_{it} + c_i + \varepsilon_{it}; \quad (3.17)$$

$$cv_{it} = \lambda \boldsymbol{W} cv_{it} + \gamma cv_{it-1} + \varrho \boldsymbol{W} cv_{it-1} + \beta_1 cr_{it} + \beta_2 pd_{it} + \beta_3 gp_{it} + c_i + \varepsilon_{it}, \quad (3.18)$$

where cv is the real cropland value, cr is the real net cash rent for cropland, pd is the population density and gp are real direct government payments. All variables are included in the model after a natural logarithm transformation.

Given that in our dataset $\mathrm{T} > \mathrm{N}$ and that, according to the MC analysis performed in chapter 2, spatial GMM-type estimators perform better for $\mathrm{N} > \mathrm{T}$, models (3.17) and (3.18) are estimated by the QML estimator by Yu et al. (2008) for different spatial weight matrices $\mathbf{W}$, showing results that appear to be robust to the choice of the weighting scheme.

3.5.1 *The effects of net cash rents and population density on cropland values*

According to the PVM, we expect net cash rents to have a positive impact on cropland value. The estimation of model (3.17) (Table 3.3) indicates a significant, albeit limited, coefficient for the expected net cash rents (0.07-0.08 depending on the spatial weight matrix). Population density shows a higher positive coefficient (0.3). Indeed, increasing population density may increase the demand for agricultural goods and therefore agricultural land

and, at the same time, it may be sign of increasing urban pressure that enhances competing demand for land for non-agricultural use. A stronger effect of changes in population than of returns to farmland on farmland values has already been found for some U.S. regions by applying an entropy-based information approach: Salois et al. (2011) find that, although changes in farmland values are more strongly associated with changes in returns to farmland at the national level, the relationship appears to change over time and region and for some regions (Northeast, Corn Belt, Appalachia, Mountain and Pacific) population has become more informative.

Table 3.3. QML estimates for the coefficients of model (3.17)

Model (3.17)	*Rook matrix*		*Distance-based matrix (I quartile)*	
Coeff.	*Estimate*	*t-stat*	*Estimate*	*t-stat*
λ	0.382	8.899***	0.382	8.986***
γ	0.734	19.824***	0.766	21.616***
ϱ	-0.182	-3.254***	-0.214	-3.906***
$\beta_1\ (cr)$	0.079	2.720***	0.072	2.443**
$\beta_2\ (pd)$	0.328	3.426***	0.320	3.370***

Significance level: ***=1% ($|\text{t-stat}| > 2.58$); **=5% ($|\text{t-stat}| > 1.96$) ; *=10% ($|\text{t-stat}| > 1.64$).

The reasons for such limited effects of the covariates may be numerous. One possible explanation relies in the inclusion of State-specific fixed effects; some results in the literature already support the idea that they may absorb part of the cross-sectional effect of the expected land rent, thus suggesting that structural determinants of the expected rents are more effective in determining cropland value than short-run expected fluctuations (see Duvivier et al. 2005 for a study on a Belgian case). The high and highly significant coefficients obtained for the spatial and temporal autoregressive coefficients (λ and γ) suggest that these may also absorb part of the effects of the covariates. The time-space autoregressive coefficient is also significant (ϱ), albeit negative and smaller in absolute value.

3.5.2 The inclusion of government payments

The inclusion of government payments as a covariate into the model does not return straightforward results (Table 3.4). First, the coefficient associated to direct government payments is significant and negative, indicating a negative impact of public subsidies on cropland value. This result is unexpected and requires deeper analysis and interpretation. Then, when we consider the effects on the other coefficients, it should be noted that the spatial and temporal effects are not significantly affected, whereas the inclusion of government payments enhances the impact of population density (whose coefficient rises from 0.3 to 0.5). Yet the most remarkable consequence is that caused on the estimates of β_1, that turn to be negative and not significant.

Table 3.4. QML estimates for the coefficients of model (3.18)

Model (3.18)	*Rook matrix*		*Distance-based matrix (I quartile)*	
Coeff.	*Estimate*	*t-stat*	*Estimate*	*t-stat*
λ	0.382	9.074***	0.382	9.161***
γ	0.713	20.359***	0.747	22.278***
ϱ	-0.187	-3.529***	-0.221	-4.248***
β_1 (cr)	-0.012	-0.415	-0.018	-0.576
β_2 (pd)	0.548	5.659***	0.538	5.598***
β_3 (gp)	-0.048	-6.906***	-0.047	-6.864***

Significance level: ***=1% ($|$t-stat$|$ > 2.58); **=5% ($|$t-stat$|$ > 1.96) ; *=10% ($|$t-stat$|$ > 1.64).

The empirical literature has already addressed the issue in various contributions that led to very different conclusions. A central point that should be taken into consideration concerns the fact that agricultural support policy instruments are thought to be highly correlated with land rents, so that part of the literature concentrates on explaining the relationship between these two variables rather that their effect on land values, trying to assess whether agricultural policy benefits landowners of farmers the most (see, for example, Roberts et al. 2003; Lence and Mishra 2003; Goodwin et al. 2004; Latruffe and Le Mouël 2009; Kirwan 2009).

Moreover, different types of subsidies are expected to have different impacts on cash rents and land values, therefore a distinction between the programs of agricultural support appears to be necessary in order to provide more accurate information. Lence and Mishra (2003), for example, found that alternative farm programs have different effects on cash rents in Iowa, with positive effects of market loss assistance and production flexibility contracts, no effects of conservation reserve programs and a negative impact of deficiency payments. Similar results are found by Goodwin et al. (2003). Feichtinger and Salhofer (2011) also found different capitalization rates for particular types of payments, with lower elasticity for agro-environmental payments, that often cause land rents to decrease.

The sources of bias when including government payments in the model are therefore numerous and the results obtained through model (3.18) can only be considered as an indication of the need of further research that takes into account the evolutions of agricultural policy in time and the differences in types of agricultural subsidies.

3.6 Short run and long run land value elasticity

The coefficients β_1 and β_2 estimated in sub-section 3.5.1 cannot be interpreted exactly as the elasticity of land value to, respectively, cash rents and population density, because of the presence of the variable cv on the RHS of model (3.17). Another contribution we make is therefore to provide an estimation of the impact and long-run elasticity of cropland values in response to changes in net cash rents and population density.

Before applying long-run value effect analysis, we test the series stationarity, in order to be sure that the process we are analyzing is not an explosive one. In order to do so, from equation (3.17) we define the $N \times N$ matrix

$$\boldsymbol{A} = (\boldsymbol{I} - \lambda \boldsymbol{W})^{-1}(\gamma \boldsymbol{I} + \varrho \boldsymbol{W}) \tag{3.19}$$

where $\boldsymbol{I}$ is an $N \times N$ identity matrix and $\boldsymbol{W}$ is an exogenous spatial weight matrix of the same dimensions.

Using $\boldsymbol{A}$ we can re-write model (3.17) as

$$cv_{it} = \boldsymbol{A}cv_{t-1} + (\boldsymbol{I} - \lambda\boldsymbol{W})^{-1}(\beta_1 cr_{it} + \beta_2 pd_{it} + c_i + \varepsilon_{it}) \tag{3.20}$$

The stability conditions of the process described in equation (3.20) can be now analyzed by computing the eigenvalues of the $\boldsymbol{A}$ matrix.

Depending on the eigenvalues, i.e. the characteristic roots of $\boldsymbol{A}$, we have three possible cases. When all the roots are less than 1 in absolute value, we call it a stable case. When all the roots are equal to 1, we term it a pure unit root case, which generalizes the unit root dynamic panel data model in the time series literature to include spatial elements. When some of the roots (but not all) are equal to 1, we define it as a spatial cointegration case, where the unit roots in the process are generated with mixed time and spatial dimensions.

Using the estimates obtained in section 3.5.1 for the autoregressive parameters by using a rook spatial weight matrix[15] ($\hat{\gamma} = 0.734$; $\hat{\lambda} = 0.382$; $\hat{\varrho} = -0.182$), we find the following eigenvalues of matrix $\boldsymbol{A}$ [$0.893, 0.850, 0.773, 0.759, 0.735, 0.710, 0.681, 0.696, 0.693, 0.692, 0.893, 0.663$]. Since all the values are less than 1, we can conclude that the system is stable. Hence the computation of elasticities for cash rents and population density is possible and can be easily done by solving the dynamic equation (3.20), i.e.

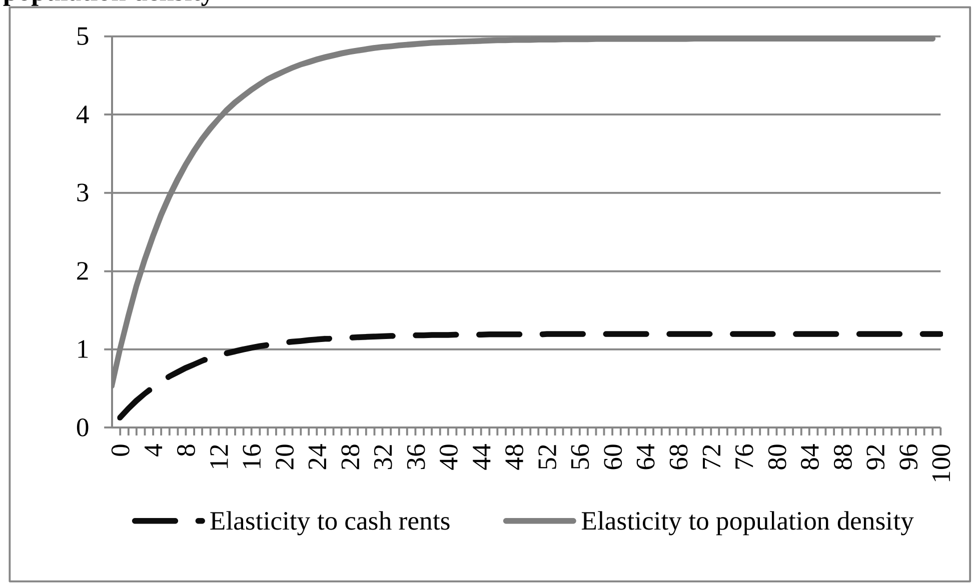
$$cv_{it} = (\boldsymbol{I} - \boldsymbol{A}\mathrm{L})^{-1}(\boldsymbol{I} - \lambda\boldsymbol{W})^{-1}(\beta_1 cr_{it} + \beta_2 pd_{it} + c_i + \varepsilon_{it}). \tag{3.21}$$

where L is the lag operator, that operates on an element of a time series to produce the previous element, such that, given $X = \{X_1, X_2, X_3, \dots\}$, $X_{it}L = X_{t-1}$, for all $t > 1$.

[15] The results lead to the same conclusions when the estimates obtained by using the other spatial weight matrices (Table 3.3) are used in the computations.

Using the estimates $\hat{\beta}_1$=0.079 and $\hat{\beta}_2$=0.328 and $t = 0, \dots, 100$, we find that the impact elasticity of cropland value (i.e. the elasticity calculated at $t = 0$) is equal to 0.13 with respect to cash rents and 0.53 with respect to population density. These values represent the expected immediate percentage changes that a 1% percent change in, respectively, cash rents and population density would cause on cropland values.

Figure 3.3. Long-run elasticity of cropland value with respect to net cash rents and population density



Considering long-run impacts instead, the calculated long-run elasticity of cropland value with respect to a 1% increase in cash rents is equal to 1.2, while the long-run elasticity of cropland value with respect to a 1% increase in population density is equal to 4.97 (Figure 3.3). About 50% of the long-run impact of both cash rents and population density on cropland value is already reached after 6 years and the percentage increases up to 90% after 21 years. Therefore in the long-run, the effect of population density (hence, according to our assumptions, of urban pressure and competing land uses) is significantly higher than that of cash rents in determining cropland values.

Such a close-to-unity estimated long-run elasticity of cropland values to cash rents is close to what one would expect according to the PVM and that is usually not verified in empirical analyses. Gutierrez et al. (2007) find similar results by allowing for structural breaks in the cointegration relationship between the two time series, for a large panel of 31 U.S. States for the period 1960-2000. Previous empirical contributions, mainly based on time-series analysis, lead to different conclusions and, as previously said, end up rejecting the PVM and generally finding evidence of divergence between the present value of future cash flows and the market price of farmland (Falk 1991; Clark et al. 1993a; Engsted 1998).

3.7 Concluding remarks

The analysis of the determinants of land value in the U.S.A. is a relevant field of study given the importance of farm real estate on the farm balance sheet and because of the great number of policy issues that it raises. We adopted the PVM framework, according to which the value of land is the capitalized value of the current and future stream of earnings from owing that asset. In order to consider a more homogeneous dataset, only 12 States of Midwestern U.S.A., for which more reliable agricultural data are available, were included in the analysis and only cropland was taken into consideration when collecting data on land value and cash rents. Our model also introduced population density among the regressors as a proxy for urban pressure, in order to take into account the effects that competing alternative land uses might exert.

Although a fairly large body of literature has been devoted to this topic, spatial econometrics has only found limited application in this empirical field so far. We believe, as the ESDA confirmed, that data on land values are characterized by effects of spatial dependence that should be taken into account in estimating an econometric model that aims at explaining the factors that contribute to land value formation. In order to do so, we chose to estimate a model in which a spatial lag of the dependent variable is included. The temporal dynamics is described as an autoregressive process of first order and a spatiotemporal lag was also introduced so as to make our model a truly time-space dynamic model.

The results that we obtained confirm the existence of significant spatial and temporal dependence and therefore the need to take them into consideration. Our estimate of the long-run elasticity of cropland value with respect to net cash rents, which is close to unity, is an element favorable to the validity of the PVM assumptions. This is a result that has found only limited support in the literature on land values, which generally ends up rejecting the PVM. Gutierrez et al. (2007) find similar evidence in favor of the theoretical model when allowing for structural breaks in the time series. However, further checks on the estimated elasticity of 1.2 are required before drawing a conclusion on this.

The effect of cash rents in determining land values is smaller than that of population density, which also has a positive significant effect on cropland values. Both variables appear to exert the biggest part of their influence on land values in about 20 years, as the computation of long-run elasticities revealed, even if about half of that impact is already reached after about 6 years.

The inclusion of government payments among the regressors is motivated by the fact that they can also be considered as an expected future stream of earnings from owing land, with relevant policy implications. However, the results that we have obtained so far do not allow to draw final conclusions on the impact of agricultural support programs on cropland values. As suggested by the vast literature on this topic, a deeper reasoning and more disaggregated data are needed in order to provide a better model specification, capable of taking into account the evolution of U.S. agricultural policy in time and the differences between different instruments of government intervention.

Future developments of this analysis should therefore follow two main paths. On the methodological point of view, the econometric model that was estimated is one that has not been widely employed in empirical analyses, because of the complexity of its estimation and the lack of already available routines in econometric software. No standard and widely known testing procedures are available yet. Nevertheless we consider running precise specification testing as a priority in order to complete the present analysis. Moreover, following Gutierrez et al. (2007), the model should also be tested for structural breaks that may occur in the time series. This is not only a methodological extension of the study because detecting and allowing for structural breaks may also serve as a means for adding to the analysis of government support intervention. This is indeed a second direction that a more in-depth analysis should follow in the future.

Conclusions

Since the late 1970s, when spatial econometrics started to grow as an autonomous branch of the econometric discipline, it has been characterized by important developments both from a theoretical and from an empirical point of view. On the one hand, an increasing number of testing procedures have been proposed able to detect the presence of spatial effects; the spatial econometric modeling, which used to focus mainly on cross-sectional data, has turned its interest onto the econometrics of panel data, static and dynamic; the estimation procedures have addressed an increasing number of issues and by now many different estimators have been developed and tested in their large asymptotic properties, each one suitable for a different model specification. On the other hand, the empirical field of application of these methodologies has extended from regional and urban studies to other fields, such as environmental studies or other branches of economics. Nevertheless, the gap between theoretical advances and empirical applications is still wide.

The aim of this work, after a comprehensive review of the main tools of spatial econometricians, is therefore to provide an empirical application of the most recently introduced techniques of spatial analysis to a field of study in which the potential of spatial econometrics has not been fully explored yet.

In chapter 1 we provided a review not only of the most widely known and applied techniques of analysis of spatial cross-sectional data, but also of the most recent improvements that mainly regard the spatial econometrics of panel data. We highlighted the difficulties that the temporal autocorrelation of data adds to the estimation procedures once the spatial autocorrelation has been treated. In order to address the issue of the gap that exists between theoretical and empirical advances, we focused on software availability, showing that the lack of ready-to-use routines contributes to hindering the application of new techniques. For example, we showed that some programming skills are needed in order to be able to estimate spatial dynamic panel data models.

Our contribution to the empirical literature that applies the most recent spatial estimators comprised two different approaches.

Chapter 2 provided an analysis of the small sample properties of some estimators (the QML estimator by Yu et al. (2008) and some difference-GMM estimators) for a time-space dynamic panel data model, for different temporal and cross-sectional dimensions and different degrees of spatial, temporal and spatiotemporal dependence. The reason for conducting such an analysis relies on the fact that the empirical researcher is usually unaware of what the correct model specification is for his/her data and, although these estimators have been proved to be asymptotically consistent, one cannot usually count on datasets of dimensions such that they make small sample biases only remote threat. Indeed, the RMSE and bias associated to the estimates of the coefficients not surprisingly decrease as the time and cross-sectional dimension of the dataset increase; this is particularly true for GMM estimators, as expected. The QML estimator was found to show the best small-sample performance in terms of RMSE for all values of T and N, mainly thanks to a considerably lower variability, both in a static and a quasi-unit root scenario.

We also focused on the assessment of the risks implied by ignoring the spatial dependence that characterizes the data. To our knowledge, empirical researchers, although fully warned on the theoretical consequences of such a misspecification, are not provided with a quantitative estimation of the bias that may characterize the estimates of the regression coefficients. Our conclusion is that the time-saving choice of ignoring the presence of spatial dependence in the data may not necessarily bring to tremendous drawbacks in terms of biased estimates of the parameters of the covariates, although the bias tends to increase as the extent of spatial dependence increases. Nevertheless, the main failure of non-spatial estimation is the fact that it prevents the identification and estimation of spatial spillover effects when present, thus limiting the information that can be drawn from the data.

The final chapter proposes an empirical application of one among the most recent estimating procedures proposed in the spatial econometrics literature to a field of analysis in which spatial econometrics has only found limited application so far. We approached the study of the determinants of agricultural land values in 12 Midwestern U.S.A. in the period 1971-2009 by choosing the PVM as the reference theoretical framework and by estimating a time-space dynamic panel data model with fixed effects. After having taken into account the spatial dependence evidenced by the ESDA, our purpose was therefore to test the assumption at the basis of the PVM, according to which land values should equal the present value of all future expected cash flows stemming from a productive use of that land. In order to do so, we regressed cropland values per acre on the net cash rents per acre and population density, all variables expressed in natural logarithms.

The results that we obtained confirmed the need to take spatial and temporal dependence into consideration and the PVM assumptions are confirmed by the estimates of the long-elasticity of cropland value with respect to this variable close to unity. This is a result that has found only limited support in the literature on land values, which generally ends up rejecting the PVM and may be at least partially due to the inclusion of spatial effects in the model specification.

Population density, that was included as a proxy for urban pressure, proved to be an important determinant of agricultural land values. The inclusion of government payments among the regressors, motivated by the fact that they can also be considered as an expected future stream of earnings from owing land, with relevant policy implications, does not lead to final conclusions on the impact of agricultural support programs on cropland values: a deeper reasoning and more disaggregated data are needed in order to provide a better model specification, capable of taking into account the evolution of U.S. agricultural policy in time and the differences between different instruments of government intervention. Together with some methodological improvements, this represents a path for future developments of our empirical analysis.

In conclusion, we found that the application of the most recently introduced tools of spatial econometrics to new empirical fields of analysis is capable of opening new research streams, still very poorly explored. The main factors that prevent this are due to the lack of already available routines for estimating spatial dynamic panel data models, thus requiring the empirical researchers to have some programming skills. Nevertheless, the empirical application of this rather recent econometric technique appears to be important in order to fully explore their potential contribution to a deeper understanding of many economic issues and, at the same time, highlight possible unexpected small sample biases that may arise.

References

Abreu, Maria, Henry L. F. De Groot, and Raymond J. G. M. Florax. 2005. "Space and Growth: a Survey of Empirical Evidence and Methods". *Région et Développement* 21: 13–44.

Anselin, Luc. 1988a. *Spatial Econometrics : Methods and Models*. Dordrecht; Boston: Kluwer Academic Publishers.

———. 1988b. "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity." *Geographical Analysis* 20 (1): 1–17.

———. 1999. "Spatial Econometrics." (Working Paper) https://csiss.ncgia.ucsb.edu/aboutus/presentations/files/baltchap.pdf.

———. 2001. "Spatial Econometrics." In *A Companion to Theoretical Econometrics*, edited by Badi H. Baltagi. Blackwell Companions to Contemporary Economics. Malden, Mass: Blackwell.

———. 2002. "Under the Hood Issues in the Specification and Interpretation of Spatial Regression Models". *Agricultural Economics* 27: 247–267.

———. 2005. "Exploring Spatial Data with GeoDa: A Workbook". Centre for Spatially Integrated Social Sciences. https://geodacenter.asu.edu/system/files/geodaworkbook.pdf.

———. 2010. "Thirty Years of Spatial Econometrics." *Papers in Regional Science* 89 (1): 3–25.

Anselin, Luc, Anil K. Bera, Raymond Florax, and Mann J. Yoon. 1996. "Simple Diagnostic Tests for Spatial Dependence". *Regional Science and Urban Economics* 26 (1): 77–104.

Anselin, Luc, and Raymond J. G. M. Florax. 1995. "Small Sample Properties of Tests for Spatial Dependence in Regression Models: Some Further Results." In *New Directions in Spatial Econometrics*, edited by Luc Anselin and Raymond J. G. M. Florax, 21–74. Berlin, Heidelberg: Springer Berlin Heidelberg.

Anselin, Luc, Julie Le Gallo, and Hubert Jayet. 2008. "Spatial Panel Econometrics." In *The Econometrics of Panel Data*, edited by László Mátyás and Patrick Sevestre, 46: 625–660. Berlin, Heidelberg: Springer Berlin Heidelberg.

Arbia, Giuseppe. 2006. *Spatial Econometrics Statistical Foundations and Applications to Regional Convergence*. New York: Springer Berlin Heidelberg.

Arbia, Giuseppe, Michele Battisti, and Gianfranco Di Vaio. 2007. “Do Institutions Matter More Than Geography? Empirical Tests of Spatial Growth Regressions for the Regions of Europe.” In . http://ricerca.economiaefinanza.luiss.it/dptea/files/llwp57.pdf.

Arellano, Manuel, and Stephen Bond. 1991. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *The Review of Economic Studies* 58 (2): 277.

Baltagi, Badi H. 2005. *Econometric Analysis of Panel Data*. 3rd ed. Chichester ; Hoboken, NJ: J. Wiley & Sons.

Baltagi, Badi H., Seuck Heun Song, and Won Koh. 2003. “Testing Panel Data Regression Models with Spatial Error Correlation.” *Journal of Econometrics* 117 (1): 123–150.

Bivand, Roger. 2006. “Implementing Spatial Data Analysis Software Tools in R.” *Geographical Analysis* 38 (1): 23–40.

———. 2013. “‘Spdep’. Spatial Dependence: Weighting Schemes, Statistics and Models.” http://cran.r-project.org/web/packages/spdep/spdep.pdf.

Blundell, Richard, and Stephen Bond. 1998. “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models.” *Journal of Econometrics* 87 (1): 115–143.

Bontempi, Maria Elena, and Irene Mammi. 2012. “A Strategy to Reduce the Count of Moment Conditions in Panel Data GMM.” *SSRN Electronic Journal*.

Bowsher, Clive G. 2002. “On Testing Overidentifying Restrictions in Dynamic Panel Data Models.” *Economics Letters* 77 (2): 211–220.

Burridge, Peter. 1980. “On the Cliff-Ord Test for Spatial Autocorrelation.” *Journal of the Royal Statistical Society B* 42: 107–108.

Campbell, J. Y., and R. J. Shiller. 1988. “The Dividend-price Ratio and Expectations of Future Dividends and Discount Factors.” *Review of Financial Studies* 1 (3): 195–228.

Campbell, John Y, Andrew W Lo, and Archie Craig MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, N.J.: Princeton University Press.

Case, Anne C., Harvey S. Rosen, and James R. Hines. 1993. “Budget Spillovers and Fiscal Policy Interdependence”. Journal of Public Economics 52 (3): 285–307.

Clark, J. Stephen, Murray Fulton, and John T. Scott. 1993a. “The Inconsistency of Land Values, Land Rents, and Capitalization

Formulas". *American Journal of Agricultural Economics* 75 (1): 147-155.

Clark, J. Stephen, K. K. Klein, and Shelley J. Thompson. 1993b. "Are Subsidies Capitalized into Land Values? Some Time Series Evidence from Saskatchewan." *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie* 41 (2): 155–168.

Cliff, A. D, and J. K Ord. 1981. *Spatial Processes : Models & Applications*. London: Pion.

Cliff, Andrew, and Keith Ord. 1972. "Testing for Spatial Autocorrelation Among Regression Residuals." *Geographical Analysis* 4 (3): 267–284.

Cressie, Noel A. C. 1993. *Statistics for Spatial Data*. Rev. ed. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

Dall'Erba, Sandy, and Julie Le Gallo. 2008. "Regional Convergence and the Impact of European Structural Funds over 1989-1999: A Spatial Econometric Analysis". *Papers in Regional Science* 87 (2): 219–244.

De Fontnouvelle, Patrick, and Sergio H. Lence. 2001. "Transaction Costs and the Present Value 'Puzzle' of Farmland Prices." *Southern Economic Journal* 68 (3): 549–565.

Doreian, P. 1980. "Linear Models with Spatially Distributed Data: Spatial Disturbances or Spatial Effects?". *Sociological Methods & Research* 9 (1): 29–60.

Drukker, David M., Hua Peng, Ingmar R. Prucha, and Rafal Raciborski. 2011a. "Creating and Managing Spatial-weighting Matrices Using the Spmat Command". Technical report, Stata.

Drukker, David M., Ingmar R. Prucha, and Rafal Raciborski. 2011b. "A Command for Estimating Spatial-autoregressive Models with Spatial-autoregressive Disturbances and Additional Endogenous Variables." Technical report, Stata.

———. 2011c. "Maximum-likelihood and Generalized Spatial Two-stage Least-squares Estimators for a Spatial-autoregressive Model with Spatial-autoregressive Disturbances". Technical report, Stata.

Du, Xiaodong, David A. Hennessy, and William M. Edwards. 2007. "Determinants of Iowa Cropland Cash Rental Rates: Testing Ricardian Rent Theory". Center for Agricultural and Rural Development Iowa State University, Working Paper 07-WP 454.

Duvivier, Romain, Frédéric Gaspart, and Bruno Henry de Frahan. 2005. "A Panel Data Analysis of the Determinants of Farmland Price: An Application to the Effects of the 1992 CAP Reform in Belgium." In

99th Seminar of the EAAE (European Association of Agricultural Economists). Copenhagen, Denmark.

Eff, Anthon E. 2008. “Weight Matrices for Cultural Proximity: Deriving Weights from a Language Phylogeny.” *Structure and Dynamics,* 3 (2).

Elhorst, J. Paul. 2005. “Unconditional Maximum Likelihood Estimation of Linear and Log-Linear Dynamic Models for Spatial Panels.” *Geographical Analysis* 37 (1): 85–106.

———. 2010. “Spatial Panel Data Models.” In *Handbook of Applied Spatial Analysis*, edited by Manfred M. Fischer and Arthur Getis, 377–407. Berlin, Heidelberg: Springer Berlin Heidelberg.

Engsted, Tom. 1998. “Do Farmland Prices Reflect Rationally Expected Future Rents?” *Applied Economics Letters* 5 (2): 75–79.

Ertur, Cem, Julie Le Gallo, and Catherine Baumont. 2006. “The European Regional Convergence Process, 1980-1995: Do Spatial Regimes and Spatial Dependence Matter?” . *International Regional Science Review* 29 (1): 3–34.

Falk, Barry. 1991. “Formally Testing the Present Value Model of Farmland Prices.” *American Journal of Agricultural Economics* 73 (1): 1.

Feichtinger, Paul, and Klaus Salhofer. 2011. “The Valuation of Agricultural Land and the Influence of Government Payments.” Edited by Belgium) Centre for European Policy Studies (Brussels. Centre for European Policy Studies (CEPS), Factor Markets project.

Fingleton, B. 1999. “Estimates of Time to Economic Convergence: An Analysis of Regions of the European Union”. *International Regional Science Review* 22 (1): 5–34.

Getis, Arthur, Jesús Mur, and Henry Zoller, ed. 2004. *Spatial Econometrics and Spatial Statistics*. Applied Econometrics Association Series. New York: Palgrave Macmillan.

Gloy, Brent. 2013. “Farmland Values - Will the Boom Turn Bust?” January 30, Center for Commercial Agriculture, Purdue University. https://www.agecon.purdue.edu/commercialag/resources/finance/farmlandvalues/Gloy%20Top%20Producer%20Farmland%20ValuesFINAL.pdf.

Goodwin, Barry K., Ashok K. Mishra, and François N. Ortalo-Magné. 2003. “What’s Wrong with Our Models of Agricultural Land Values?” *American Journal of Agricultural Economics* 85 (3): 744–752.

———. 2004. “Landowners’ Riches: The Distribution of Agricultural Subsidies”. Working Paper, School of Business, University of

Wisconsin, Madison. http://ww.bus.wisc.edu/realestate/documents/04LandownersRiches.pdf.

Gutierrez, L., J. Westerlund, and K. Erickson. 2007. “Farmland Prices, Structural Breaks and Panel Data.” *European Review of Agricultural Economics* 34 (2) : 161–17.

Gutierrez, Luciano. 2011. “Looking for Rational Bubbles in Agricultural Commodity Markets.” In *European Association of Agricultural Economists, 2011 International Congress (August 3- September 2)*. Zurich, Switzerland. http://purl.umn.edu/120377.

Hansen, Lars Peter. 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica* 50 (4): 1029–1054.

Hardie, Ian W., Tulika A. Narayan, and Bruce L. Gardner. 2001. “The Joint Influence of Agricultural and Nonfarm Factors on Real Estate Values: An Application to the Mid-Atlantic Region.” *American Journal of Agricultural Economics* 83 (1): 120–132.

Jacobs, Jan P. A. M., Jenny E. Ligthart, and Hendrik Vrijburg. 2009. “Dynamic Panel Data Models Featuring Endogenous Interaction and Spatially Correlated Errors.” *SSRN Electronic Journal*. http://www.ssrn.com/abstract=1517257.

Kapoor, Mudit, Harry H. Kelejian, and Ingmar R. Prucha. 2007. “Panel Data Models with Spatially Correlated Error Components.” *Journal of Econometrics* 140 (1): 97–130.

Kelejian, Harry H., and Ingmar R. Prucha. 1998. “A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.” *Journal of Real Estate Finance and Economics* 17 (1): 99–121.

———. 1999a. “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model.” *International Economic Review* 40 (2): 509–533.

———. 1999b. “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model.” *International Economic Review* 40 (2): 509–533.

Kelejian, Harry H., and Dennis P. Robinson. 1993. “A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a County Expenditure Model.” *Papers in Regional Science* 72 (3): 297–312.

———. 1995. “Spatial Correlation: A Suggested Alternative to the Autoregressive Model.” In *New Directions in Spatial Econometrics*, edited by Luc Anselin and Raymond J. G. M. Florax, 75–95. Berlin, Heidelberg: Springer Berlin Heidelberg.

Kirwan, Barrett E. 2009. “The Incidence of U.S. Agricultural Subsidies on Farmland Rental Rates.” *Journal of Political Economy* 117 (1): 138–164.

Korniotis, George M. 2010. “Estimating Panel Models With Internal and External Habit Formation.” *Journal of Business & Economic Statistics* 28 (1): 145–158.

Kukenova, Madina, and Jose-Antonio Monteiro. 2009. “Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation.” *MPRA Paper No. 14319*.

Le Gallo, Julie. 2002. “Econométrie Spatiale: L’autocorrélation Spatiale Dans Les Modèles de Régression Linéaire”. *Economie & prévision* 4 (155): 139–157.

Le Gallo, Julie, and Sandy Dall’Erba. 2006. “Evaluating the Temporal and Spatial Heterogeneity of the European Convergence Process, 1980-1999”. *Journal of Regional Science* 46 (2): 269–288.

Lee, Lung-fei, and Jihai Yu. 2010a. “A Spatial Dynamic Panel Data Model with Both Time and Individual Fixed Effects.” *Econometric Theory* 26 (2): 564–597.

———. 2010b. “Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects.” *Journal of Econometrics* 154 (2): 165–185.

Lence, Sergio H., and Douglas J. Miller. 1999. “Transaction Costs and the Present Value Model of Farmland: Iowa, 1900-1994.” *American Journal of Agricultural Economics* 81 (2): 257-272.

Lence, Sergio H., and Ashok K. Mishra. 2003. “The Impacts of Different Farm Programs on Cash Rents.” *American Journal of Agricultural Economics* 85 (3): 753–761.

LeSage, James P, and R. Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton: CRC Press.

LeSage, James P. 1999. “The Theory and Practice of Spatial Econometrics.” http://www.spatial-econometrics.com/.

Livanis, G., C. B. Moss, V. E. Breneman, and R. F. Nehring. 2006. “Urban Sprawl and Farmland Prices.” *American Journal of Agricultural Economics* 88 (4): 915–929.

Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem”. *The Review of Economic Studies* 60 (3): 531–542.

Millo, Giovanni, and Gianfranco Piras. 2012. "Splm: Spatial Panel Data Models in R." *Journal of Statistical Software* 47 (1): 1–38.

Mooney, Christopher Z. 1997. *Monte Carlo Simulation*. Sage University Papers Series no. 07-116. Thousand Oaks, Calif: Sage Publications.

Moran, P. A. P. 1950. "Notes on Continuous Stochastic Phenomena". *Biometrika* 37 (1/2): 17-23.

Nickerson, Cynthia, Morehart Mitchell, Todd Kuethe, Jayson Backman, Jennifer Ifft, and Ryan Williams. 2012. "Trends in U.S. Farmland Values and Ownership". *Economic Information Bulletin* N. 92, U.S. Dept. of Agriculture, Economic Research Service.

Paelinck, Jean H. P., and Leo H. Klaassen. 1979. *Spatial Econometrics*. Studies in Spatial Analysis. Farnborough, Eng: Saxon House.

Piras, Gianfranco. 2010. "Sphet: Spatial Models with Heteroskedastic Innovations in R." *Journal of Statistical Software* 35 (1): 1–21.

Plantinga, Andrew J., Ruben N. Lubowski, and Robert N. Stavins. 2002. "The Effects of Potential Land Development on Agricultural Land Prices." *SSRN Electronic Journal*. http://www.ssrn.com/abstract=305498.

Ramajo, Julián, Miguel A. Márquez, Geoffrey J.D. Hewings, and María M. Salinas. 2008. "Spatial Heterogeneity and Interregional Spillovers in the European Union: Do Cohesion Policies Encourage Convergence Across Regions?". *European Economic review* 52 (3): 551–567.

Roberts, Michael J., Barrett Kirwan, and Jeffrey Hopkins. 2003. "The Incidence of Government Program Payments on Agricultural Land Rents: The Challenges of Identification." *American Journal of Agricultural Economics* 85 (3): 762–769.

Roodman, David. 2009. "A Note on the Theme of Too Many Instruments." *Oxford Bulletin of Economics and Statistics* 71 (1): 135–158.

Salois, M., C. Moss, and K. Erickson. 2011. "Farm Income, Population and Farmland Prices: a Relative Information Approach." *European Review of Agricultural Economics* 39 (2): 289–307.

Tegene, A., and F. Kuchler. 1993. "A Regression Test of the Present Value Model of US Farmland Prices." *Journal of Agricultural Economics* 44 (1): 135–143.

Tobler, W. R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46: 234-240.

United States Department of Agriculture. National Agricultural Statistics Service. 2012. "Land Values, 2012 Summary."

Weersink, Alfons, Stephen J. Clark, Calum G. Turvey, and Rakhal Sarkar. 1999. "The Effect of Agricultural Policy on Farmland Values." *Land Economics* 75 (3): 425–439.

Windmeijer, Frank. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators." *Journal of Econometrics* 126 (1): 25–51.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, Mass: MIT Press.

Yu, Jihai, Robert de Jong, and Lung-fei Lee. 2008. "Quasi-maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T Are Large." *Journal of Econometrics* 146 (1): 118–134.

———. 2012. "Estimation for Spatial Dynamic Panel Data with Fixed Effects: The Case of Spatial Cointegration." *Journal of Econometrics* 167 (1): 16–37.

Yu, Jihai, and Lung-fei Lee. 2010. "Efficient GMM Estimation of Spatial Dynamic Panel Data Models with Fixed Effects." In *Econometric Society 10th World Congress*. Shanghai, China.

Ziliak, James P. 1997. "Efficient Estimation With Panel Data When Instruments Are Predetermined: An Empirical Comparison of Moment-Condition Estimators." *Journal of Business & Economic Statistics* 15 (4): 419–431.