SPATIAL CORRELATION ROBUST INFERENCE WITH ERRORS IN LOCATION OR DISTANCE

Timothy G. Conley
Francesca Molinari

Spatial Correlation Robust Inference
with Errors in Location or Distance*

Timothy G. Conley

Graduate School of Business

University of Chicago

Francesca Molinari

Department of Economics

Cornell University

October 15, 2005

Abstract

This paper presents results from a Monte Carlo study concerning inference with spatially dependent data. We investigate the impact of location/distance measurement errors upon the accuracy of parametric and nonparametric estimators of asymptotic variances. Nonparametric estimators are quite robust to such errors, method of moments estimators perform surprisingly well, and MLE estimators are very poor. We also present and evaluate a specification test based on a parametric bootstrap that has good power properties for the types of measurement error we consider.

1 Introduction

Spatial econometric models have proven useful in many areas of economics.[1] Economic models underpinning empirical work in urban, environmental, development, industrial organization, and growth frequently suggest that observed agents will have outcomes that are not independent. Often these models suggest a suitable metric or a set of locations in some space that characterizes the structure of dependence among agents. A spatial model is simply a data generating model that utilizes such a set of locations or distances

---

[1]Examples of work applying spatial models include Case (1991), Kelejian and Robinson (1992), Case, Hines, and Rosen (1993), Elliott (1993), Moreno and Trehan (1997), Bollinger and Ihlanfeldt (1997), Bell and Bockenstael (2000), Conley and Topa (2002), Topa (2001), Pinkse, Slade, and Brett (2002), Conley and Dupor (2003), and Kim, Phipps, and Anselin (2003).

to define the relationships between agents' variables. The notion of space can be general and is certainly not confined to physical or geographic space.

Typical spatial models are parametric models of the dependence between agents, examples include Whittle (1954), Ord (1975), Anselin and Griffith (1988), Case (1991), and Kelejian and Prucha (1999). The most prevalent models are for Gaussian data with a covariance structure that is a parametric function of known locations. A smaller literature has focused on nonparametric methods for estimating covariance structure both as a direct object of interest and to conduct inference about conditional mean estimates, see e.g. Grenander and Rosenblatt (1957), Hall, Fisher, and Hoffman (1994), Hall and Patil (1994), Conley (1995, 1999), Conley and Dupor (2003). These nonparametric methods estimate covariances or their sum with local averages. Some of these methods can be viewed as smoothed periodograms (e.g. Grenander and Rosenblatt (1957) and Conley (1995, 1999)) and hence are directly related to the extensive literature on spectral representations for time series and random fields.[2]

The key ingredient in any spatial model is the choice of metric space and locations for the observed agents. However, it is routinely the case that agents' locations are not known with certainty in data available to the econometrician. It is very common for information about agents' physical locations to be imprecise, e.g. locations to be known only within an area– census tract, zip code, county, or SMSA. At best this will result in imprecise distance information between agents and if inter-agent distances are approximated with measurements based on these areas, e.g. distance between centroids, errors will result. Moreover, in many applications the most appropriate metric is not physical distance and must be either estimated or approximated in some way. For example, the travel time between locations is often an appropriate metric; it must be estimated or approximated and cannot be known with certainty. Thus it is common for the econometrician's measurements of locations/distances to be imprecise or measured with error. In this paper we focus on the consequences of measurement errors in locations/distances for inference, leaving an investigation of the consequences of imprecise location/distance information for future research.

Measurement error in locations/distances creates problems for parametric models of spatial covariance. Unless they include an explicit treatment of the measurement error process, parametric models will generally be misspecified and inconsistent when locations/distances are measured with error.[3] There is a small but growing body of work in spatial statistics that does explicitly model measurement error in physical locations. Examples in geostatistics include Gabrosek and Cressie (2002), Cressie and Kornak (2003), and Zhao and Wall (2004) who are concerned with the consequences of physical location measurement error for prediction and estimation of covariances/variograms.[4] To our knowledge, such an explicit modeling of

---

[2]See, e.g., Priestley (1981) for an excellent discussion of much of this literature.

[3]See Griffith and Lagona (1998) for results on the inconsistency of MLE estimators of spatial correlations when locations are misspecified.

[4]Modeling errors in location has also been done in biostatistics in the study of the geometric form of organisms through modeling deformations of "landmark" locations, e.g. points on the skull, see for example Bookstein (1986). Deformations of mapped locations via GIS processing have been studied in the geography literature. See, e.g., Arbia, Griffith, and Haining

measurement errors has not been done in the econometrics literature.

Modeling errors in 'economic distance' is difficult and presents challenges distinct from those in geostatistics. In econometrics, measuring economic distance is very often not solely a matter of measuring the physical coordinates of a location. Many metrics that are very well-motivated by the underlying economic theory are inherently unobservable, making it impossible to obtain validation studies/experiments like those that are potentially feasible for devices that measure physical locations. For example, economic distances between observed firms might be based on constructed measures of the similarity of their local labor market conditions, the similarity of their technology, or their product markets.[5] Moreover, measurement errors in economic distances will routinely be endogenous, not independent of unobservables influencing outcome variables of interest. For example, unobserved aspects of urban density/congestion may influence both outcomes like wage and rent differentials and economic distances based on travel time estimates. Thus even for economic distances that are potentially observable like travel costs, endogenous measurement errors may require estimation of a joint model for outcomes, true locations, and measured locations. This is a substantially more complicated task than estimating the joint distribution of true and measured physical locations. Finally, it is often sufficient in economic applications to estimate a sum of an autocovariance function, an easier object to estimate in a manner that is robust to distance/location errors than the whole covariance function itself which is often required in geostatistics.

The econometric difficulties in fully modeling economic distance/location measurement errors motivate the development of inference methods that are robust to such measurement errors. In contrast to parametric methods, the nonparametric inference procedure in Conley (1995, 1999) is robust to measurement error in distances/locations. This class of asymptotic covariance matrix estimators remain consistent with bounded, potentially endogenous measurement errors and are robust in practice. These estimators, analogous to kernel Heteroskedasticity Autocorrelation Consistent (HAC) estimators in time series, can be viewed as weighted sums of cross products of observations (sample covariances) and remain consistent because bounded measurement errors change the form of the weights in the sum, but the altered form still satisfies the requisite conditions for consistency. Robustness in practice results from the qualitative features of the weights being largely unaffected by small to moderately-sized location/distance mistakes, even endogenous ones.

This paper presents a Monte Carlo study that investigates the impact of location/distance measurement errors upon the accuracy of estimators of the asymptotic variance of a sample average, V, and the performance of two new specification tests for parametric estimators of V. Such asymptotic covariance matrix estimators are a fundamental component of the large-sample approximations used in a large share of applications. We compare the performance of two parametric V estimators to the nonparametric V estimator of Conley (1995, 1999) when agents' locations are measured with differing amounts of error. The

---

(1998)) who study the propagation of location errors for sequences of maps resulting from overlay operations.

[5]See Conley and Dupor (2003) for examples of economic metrics based on output market and technological similarity.

parametric estimators we consider are the maximum likelihood estimator (MLE) and a method of moments (MM) estimator with spatial correlations and variances as moments. Despite the fact that the parametric models will be inconsistent, we expect them to outperform the nonparametric model in finite samples with amounts of measurement error that are small enough. We also anticipate that the MM estimator may be considerably more robust to distance errors than the MLE, albeit at the cost of efficiency when distances are close to perfectly measured.

We present and investigate the performance of two types of specification tests.[6] The first type is based on a comparison of the parametric and nonparametric estimates of V using a limiting distribution (in the spirit of a Durbin-Wu-Hausman test). It remains feasible with only a partial specification of the data generating process (DGP), i.e. the autocovariance function. The second type of test is applicable when the full data generating process is specified. This test uses parameter estimates (which could come from MLE or MM estimators) to conduct a parametric bootstrap, simulating the finite sample distribution of the nonparametric estimator of V, which provides an acceptance region for the nonparametric estimate under the null hypothesis of a correct parametric specification and no measurement errors in locations/distances. It is important to note that these specification tests are for the joint hypothesis of a correct parametric model and no measurement error in distances, so they cannot be used as a test just for measurement error in distances.

We use DGPs that are different from the simultaneous spatial autoregressive (SAR) model that is most typically used in the spatial econometrics literature. Instead of taking the basic SAR approach of specifying a known spatial weights matrix (and scalar parameter) in a simultaneous equations model, we first specify agents' locations on a lattice and then specify the DGP for agents' variables in terms of their lattice locations. We consider stationary, mixing data DGPs on both a one and a two-dimensional lattice and use an increasing domain asymptotic approach. Specifically, our DGPs are finite-order moving averages with geometrically decreasing weights. Asymptotic covariances for averages of spatial data are sums of spatial autocovariances, analogous to the asymptotic variance of averages of covariance stationary time series. The key conditions for consistency for our nonparametric estimator are that the data are mixing and measurement errors are limited, with bounded errors being a sufficient condition (see Conley (1999)). The assumption of stationarity is not necessary and analogous HAC methods can be applied to weakly dependent but nonstationary data (see Pinkse, Slade, and Brett (2002) and Kelejian and Prucha (2003)).

We don't use the typical approach of specifying a simultaneous equations SAR weight matrix for three main reasons. First and foremost, directly modeling agents' lattice locations facilitates our description of the DGP for location/distance measurement errors. A clear description of the candidate models for

---

[6]There is of course a large literature on specification testing in spatial models. However, most of the work known to us involves testing for the presence of spatial correlation in cross section or panel models (e.g., Moran (1950), Kelejian and Prucha (2001), Pinkse (1999), Baltagi et al. (2003)). We are unaware of prior work on specification tests focusing on the asymptotic variance of a sample mean.

measurement error is of course crucial for this paper. Second, this is the most straightforward approach when estimation is going to be nonparametric using a local average, or smoothed periodogram approach, or a parameterization of only a covariance function rather than a full DGP.[7] Finally, there are many applications where the "two-step" modeling process of first specifying agents' locations/distances and then modeling their random variables' dependence as a function of these locations is much easier than simultaneously specifying an implicit functional form for covariances and set of locations as in the typical SAR approach.[8] For example, locations/metrics are often naturally suggested by the economics of the application but plausible functional forms are not. In such circumstances, it is often much easier to experiment with covariance specifications given the metric rather than indirectly specifying both through simultaneous equations.

The effect of the types of location errors we consider can be thought of as changing the true DGP's autocorrelation function. The general effect of measurement error in distances will be to smooth the true covariance function and extend the range of measured distances with nonzero covariances for finite order MAs. Our results show that measurement errors impact the performance of our estimators of the asymptotic variance in different degrees, with the MLE estimator being the most sensitive to location errors, the nonparametric estimators being the least sensitive, and the MM estimator performing surprisingly well for some specifications. Our parametric bootstrap test displays good power properties, and performs well across a range of kernel bandwidth choices for our nonparametric estimator.

The remaining sections of the paper are organized as follows. Section two presents the data generating model and our estimators. Section three presents our design of data generating processes for data and location/distance errors, as well as the specific forms for estimators and specification tests. Our simulation results are presented in Section four. We conclude by discussing future research suggested by our experiments in Section five.

2   Econometric Model and Estimation Problem

The econometric model we use assumes there is a population of agents residing at d-dimensional integer lattice locations with one individual per location. We focus on an expectation zero process $X_{\mathbf{s}}$ indexed on this lattice that is assumed to be mixing ($X_{\mathbf{s}}$ and $X_{\mathbf{r}}$ approach independence as the distance between $\mathbf{s}$ and $\mathbf{r}$ grows). For simplicity, we also assume the process is stationary: the joint distribution of $X_{\mathbf{s}}$ for a collection of locations is invariant to translation and so, assuming second moments exist, $E\{X_{\mathbf{s}}X_{\mathbf{s}+\mathbf{h}}\} = C(\mathbf{h})$. The

---

[7]There are of course also applications where it is more natural to take the SAR simultaneous modeling approach, e.g. Pinkse, Slade, and Brett (2002). See Lee (2001a, 2001b, 2004) for an extensive characterization of SAR models under various forms of asymptotic approximations.

[8]Applications adopting this "two-step" approach include Pulvino (1998), Conley, Flyer, Tsiang (1999), Manuszak (2001), Conley and Dupor (2003), Bronnenberg, Dube, and Dhar (2003), Greenstone and Deschenes (2003), Vigfusson (2003), and Rappaport and Sachs (2003).

econometrician's sample consists of realizations of agents' random variables $X_{\mathbf{s}}$ at a collection of locations $\{\mathbf{s}_i\}$ inside a sample region Λ_τ. We use the notation $|\Lambda_\tau|$ to denote the number of agents in our sample region and, for simplicity, assume that all locations in Λ_τ are sampled. When taking limits, we view Λ_τ as one of a sequence of regions indexed by τ that grow to include the whole lattice, an increasing domain approach to asymptotic approximations.

We are interested in conducting inference about EX using the usual large-sample distribution approximations for the sample average of points in $\Lambda_\tau : \bar{X} = \frac{1}{|\Lambda_\tau|} \sum_{i=1}^{|\Lambda_\tau|} X_{\mathbf{s}_i}$. To do this, we need to estimate the asymptotic variance of a normalized sample mean. Using, for example, the central limit theorem due to Bolthausen (1982) for stationary, mixing random fields on regular lattices, we know that (under mixing and moment conditions[9]) the normalized sample mean has a limiting normal distribution:

$$\frac{1}{\sqrt{|\Lambda_\tau|}} \sum_{i=1}^{|\Lambda_\tau|} X_{\mathbf{s}_i} \Rightarrow N(0, V).$$

The general form for the asymptotic covariance V is as an infinite sum of an autocovariance function $C(\mathbf{h})$. Referring to the entries of the vector $\mathbf{h}$ individually as $h_1, h_2, ..., h_d$, V has the form:

$$V = \sum_{h_1=-\infty}^{\infty} ... \sum_{h_d=-\infty}^{\infty} C(h_1, h_2, ..., h_d).$$

Thus if $d = 1$, the expression for V coincides with the asymptotic variance of a sample mean for a covariance stationary time series: $V = \sum_{h_1=-\infty}^{\infty} C(h_1)$.

We are interested in comparing the performance of parametric and nonparametric estimators of V when locations are measured with error. We examine an MM estimator that corresponds to an assumption that the covariance function is known up to a finite-dimensional parameter vector, so it can be written as $C(\mathbf{h}; \theta)$. We compute a minimum distance estimator $\hat{\theta}$ and then compute a V estimator by plugging in the estimate $\hat{\theta}$, and calculating the sum of $C(\mathbf{h}; \hat{\theta})$. For notational simplicity, we suppress reference to τ and refer to this estimator as $\hat{V}_{MM}$. We also examine the performance of the estimator implied by the MLE for the process and (again suppressing τ) we let $\hat{V}_{MLE}$ be the corresponding estimate of V. In the presence of measurement error in distances, $C(\mathbf{h}; \theta)$ and the likelihood will generally be misspecified and the resulting estimators $\hat{V}_{MM}$ and $\hat{V}_{MLE}$ inconsistent. However, in finite samples with small enough amounts of location measurement error one or both may still be preferable to a consistent but less precise nonparametric estimator.

Our main nonparametric estimator of V is that proposed by Conley (1999). This method is a straight-

---

[9]Bolthausen (1982) provides sets of regularity conditions for α and ρ-mixing processes. These conditions are reported in Appendix A for the interested reader. Central limit results under different mixing conditions are also available, see e.g. Takahata (1983) or Goldie and Greenwood (1986). A survey of random field central limit results is given by Goldie and Morrow (1986).

forward generalization of well-known smoothed periodogram spectral density estimators.[10] We estimate V as:

$$\hat{V}_{NP} = \frac{1}{|\Lambda_\tau|} \sum_{i=1}^{|\Lambda_\tau|} \sum_{j=1}^{|\Lambda_\tau|} K_\tau \left(\mathbf{s}_i - \mathbf{s}_j\right) \cdot \left(X_{\mathbf{s}_i} - \bar{X}\right) \cdot \left(X_{\mathbf{s}_j} - \bar{X}\right)$$

where the dependence of $\hat{V}_{NP}$ on sample size is notationally suppressed. $K_\tau(\cdot)$ is a kernel which will be used to weight the observations, and is such that $K_\tau(\mathbf{h}) \to 1$ for all $\mathbf{h}$ as $\tau \to \infty$, slowly enough so that the variance of $\hat{V}_{NP}$ collapses to zero. If only distances and not locations were known, the kernel could of course be chosen so that $K_\tau(\mathbf{s}_i - \mathbf{s}_j)$ just depends on $\|\mathbf{s}_i - \mathbf{s}_j\|$.[11]

For the case of perfectly observed locations $\{\mathbf{s}_i\}$, there are a variety of sets of sufficient conditions on the process $X_{\mathbf{s}}$ and kernel $K_\tau(\cdot)$ that imply that $\hat{V}_{NP}$ will be consistent. For example, Grenander and Rosenblatt (1957) provide conditions for Gaussian processes (see also Priestley (1981)). Conley (1999) provides sufficient conditions for stationary, alpha-mixing processes for kernels $K_\tau(\mathbf{h})$ that are uniformly bounded and equal to zero after a cutoff distance in each dimension. Letting $L_{i,\tau}$ denote a cutoff distance and $N_{i,\tau}$ denote the width of the sample region Λ_τ in one coordinate dimension, Conley demonstrates consistency of $\hat{V}_{NP}$ when $L_{i,\tau} = o(N_{i,\tau}^{1/3})$ for $X_{\mathbf{s}}$ processes that have slightly more than fourth moments and satisfy a mixing rate condition.[12] Neither stationarity nor our specific sampling framework is required to show consistency for analogous HAC estimators. Recent papers providing conditions for consistency of closely related HAC estimators with other forms of weakly dependent data, including SAR models, are Pinkse, Slade, and Brett (2002) and Kelejian and Prucha (2003).

For the case of bounded location errors, Conley (1999) provides a consistency result for $\hat{V}_{NP}$ with 'scale parameter' kernels $K_\tau(\mathbf{h})$ for processes on the plane $K_\tau(h_1, h_2) = K\left[\frac{h_1}{L_{1,\tau}}, \frac{h_2}{L_{2,\tau}}\right]$. We restate and slightly extend this result to allow for a uniform kernel in Proposition 1, again confining attention to the plane for ease of exposition.

Proposition 1 *Suppose errors in locations are bounded and: (a) $X_{\mathbf{s}}$ is a stationary, mixing process with $(4 + \delta)$ th moments, $\delta > 0$, and with alpha mixing coefficient $\alpha_{\infty,\infty}(m)$, defined in Appendix A, s.t. $\alpha_{\infty,\infty}(m)^{\delta/(2+\delta)} = o(m^{-4})$; (b) Each $L_{i,\tau} = o(N_{i,\tau}^{1/3})$, and $K(\cdot)$ is a continuous bounded function on $[-1,1]^2$ with $K(0,0) = 1$, and such that either (i) $K(\cdot)$ has absolutely summable Fourier coefficients, or (ii) $K(\cdot)$ is a uniform function on $[-1,1]^2$. Then*

$$\hat{V}_{NP} \xrightarrow{p} V.$$

---

[10]See Priestley (1981) for an extensive discussion of the vast literature on spectral methods in time series, and some extensions to random fields. Spectral methods for random fields/spatial processes date back to at least the 1950s, e.g. Whittle (1954), Bartlett (1955), Grenander and Rosenblatt (1957), Priestley (1962). More recent contributions of kernel methods in covariance estimation include Hall, Fisher, and Hoffman (1994) and Oehlert (1993).

[11]If the process were isotropic, it would be natural to specify K_τ as depending only on distances. However, isotropy is not required for consistency with K_τ of this form.

[12]The specific regularlity conditions use Bolthausen's specification of alpha mixing coefficients presented in Appendix A. They are that the process have $(4 + \delta)$ th moments and $\alpha_{\infty,\infty}(m)^{\delta/(2+\delta)} = o(m^{-4})$.

Proof. See Appendix B. ∎

$\hat{V}_{NP}$ remains consistent in the presence of bounded measurement error because all locations' displacements h will eventually have a weight approaching one. Note that the location measurement errors need not be exogenous, merely bounded for this result to hold. This estimator will also be robust to moderate location/distance measurement errors in practice as mismeasured locations' weightings under the kernels will often be close to the weight they would get with perfectly measured locations. The exceptions will occur for pairs of observations near the cutoff parameters $L_{i,\tau}$. For example, if K_τ is a uniform kernel equal to one only for displacements with length less than L_τ in each dimension, only those pairs of observations whose true displacement lengths are in a neighborhood around L_τ will have different weights from those for true displacements. Typically these misweighted observations are small fraction of the total with moderate measurement errors (See Figure 5).

It is well known that the kernel K_τ can be chosen so that $\hat{V}_{NP}$ will be nonnegative in sample by choosing from a class of kernels with nonnegative Fourier transforms.[13] However, we have two sources of motivation to investigate the small sample properties of a uniform kernel that is outside this class. The first is that this kernel greatly facilitates both derivation and implementation for one of our specification tests. Second, it is a natural choice of kernel for our finite–order moving average DGPs. It facilitates choice of a cutoff that is a little too small and one that is a little too large relative to the order of the moving average. This allows us to concisely address the important issue of estimator performance for differing smoothing parameter choices. The potential drawback of this kernel returning negative estimates did not occur in any of our approximately 50,000 simulations. We note that if our interest was in the entire covariance structure, rather than just V, guaranteeing nonnegativity by kernel choice would be much more important.

We also consider an approximately unbiased estimator that is analogous to $\hat{V}_{NP}$. The bias in $\hat{V}_{NP}$ depends on the dimension d and is of order $|\Lambda_\tau|^{-1/d}$ which, while negligible on the line, can be of great importance for $d \geq 2$ (See e.g. Guyon (1982), Politis and Romano (1996)). In our simulations setup, it is straightforward to construct an approximately unbiased estimator that simply adds up approximately unbiased sample covariances, given by partial sums of demeaned observations at a given lag divided by the number of observations at that lag. In practice, unbiased estimators may be difficult to construct and/or perform poorly since the true form of V is defined by the lattice structure which will be unknown and distorted by measurement errors in locations/distances. In contrast $\hat{V}_{NP}$ is easily constructed in practice, therefore it is interesting to compare its performance with its unbiased version in our simulations.

---

[13]See e.g. Priestley (1981).

3   Data Generating Processes for Simulations

This section describes the data generating processes (DGP) for $X_{\mathbf{s}}$ and the measurement error process for locations that we use in our simulation experiments. We consider lattice processes indexed in both one and two dimensions. While we expect that our results on the plane are more relevant for the majority of applications, we found it very useful in understanding the nature of our measurement error process to start with the simpler case on the line. We simulate a region sufficiently larger than Λ_τ and cut Λ_τ from its interior to insure that boundary points have the appropriate marginal distribution. Rather than calibrate a particular DGP and measurement error process to a single specific application, we use a simple finite-order moving average and a simple, stylized model of local mistakes in measured lattice locations. This model is well motivated by an important class of applications (discussed below) and we hope that by keeping the processes simple our results will be a useful starting point for thinking about the effects of location measurement error in other applications.

3.1   DGPs for X

The DGPs we consider for X_s are expectation zero finite-order moving averages. Our DGP for X_s on the line is a finite-order two-sided moving average with geometrically declining weights:

$$X_s = \rho^m u_{s-m}... + \rho^2 u_{s-2} + \rho u_{s-1} + u_s + \rho u_{s+1} + \rho^2 u_{s+2} + ...\rho^m u_{s+m}$$

where u_s is IID $N(0,\sigma^2)$. Since this process is a finite-order moving average, $V = \sum\limits_{k=-2m}^{2m} C(k)$ with $C(k)$ being given by:

$$C\left(k;\rho,\sigma\right) = \begin{cases} \sigma^2\left(1+2\sum\limits_{j=1}^m \rho^{2j}\right) & \text{if } k=0, \\[2em] \sigma^2\left[(|k|+1)\rho^{|k|}+2\rho^{2+|k|}\sum\limits_{j=0}^{m-|k|-1}\rho^{2j}\right] & \text{if } |k|\le m-1, \\[2em] \sigma^2\left[(m+1)\rho^m\right] & \text{if } |k|=m, \\[1em] \sigma^2(2m-|k|+1)\rho^{|k|} & \text{if } m+1\le |k|\le 2m, \\[1em] 0 & \text{otherwise.} \end{cases} \tag{1}$$

The DGP we consider for $X_{\mathbf{s}}$ on the plane is also a finite-order moving average with weights that decline geometrically with distance from $\mathbf{s}$, until a distance of m, and then are set to zero:

$$X_{\mathbf{s}} = \sum_{\mathbf{r}:\|\mathbf{s}-\mathbf{r}\|\le m} \rho^{\|\mathbf{s}-\mathbf{r}\|} u_{\mathbf{s}-\mathbf{r}}$$

where $u_{\mathbf{s}}$ is IID $N(0, \sigma^2)$. Explicit expressions for V and $C(\mathbf{k})$ for the process that we use for our simulations are provided in Appendix C.

Sample Size and Parameter Choices

We investigate sample sizes of 500 on the line and a square grid of 1600 observations on the plane. These choices of sample sizes are meant to roughly correspond to those that would occur in applications using census tract socioeconomic data in a medium to large US city, or typical household-level marketing data for such a city, or firm level data within certain industrial sectors.

In both our simulations on the line and the plane we choose $m = 3$. On the line, the explicit expression for V for this process is:

$$V = \sigma^2 \left(4\rho^6 + 8\rho^5 + 12\rho^4 + 12\rho^3 + 8\rho^2 + 4\rho + 1\right). \tag{2}$$

The expression for V on the plane is sufficiently inelegant that it is relegated to Appendix C. We illustrate how V varies with the decay parameter ρ for both our process on the line and plane in Figures 1a and 1b. The dotted and dot-dashed lines plot V as a function of ρ when $\sigma^2 = 1$ for our process on the line and plane, respectively. For comparison, the solid line plots the asymptotic variance for a one-sided first-order autoregression with correlation parameter ρ and an innovation variance of one.

We investigate ρ values of .3 and .45. Thus our DGPs are 6th order moving averages with $X_{\mathbf{s}}$ and $X_{\mathbf{r}}$ independent when $\|\mathbf{s} - \mathbf{r}\| > 6$. Our choices of m and ρ are meant to reflect small to medium spatial correlation and are motivated by empirical work that suggests that this range is very relevant. In econometric applications, spatial correlation in regression residuals is often only small to moderate, even when there is a great deal of spatial correlation in the outcome, because typical regressors capture a good deal of the correlation across space. Examples of applications where this occurs include Conley, Flyer, and Tsiang (2003), Conley and Topa (2002), and Conley and Ligon (2002).

3.2 Measurement Error

We run simulation experiments with each DGP for $X_{\mathbf{s}}$ for different levels of measurement error in locations. We model errors in locations/distances as erroneously measured positions $\{\mathbf{s}_i\}$. We limit ourselves to studying exogenous location errors in this paper, though one of the main motivations for nonparametric estimators of V is that they are robust to endogenous distance/location errors. This is solely because we want to understand exogenous measurement errors first, before moving on to more involved modeling of endogenous errors. Also, we retain the feature that mismeasured lattice locations still have each agent occupying a distinct location as it greatly facilitates both exposition of the nature of measurement error and computations in our simulation experiments.[14] This comes at the cost of not allowing us to investigate

---

[14]V can be represented as an integral of a covariance function against a measure on various lags in locations. Retaining the same set of distinct locations under measurement error makes the 'lag measure' equivalent under the true locations and mismeasured locations. So the effect of measurement error can be described by changes in the covariance function rather than

the benchmark of independent location errors.[15] In an effort to remain close to this benchmark we use a process that independently perturbs agents locations by adding a bounded measurement error, but then remaps each back to a distinct integer lattice coordinate. Position shifts of agents will necessarily be dependent due to requiring one agent per location, but they will be bounded.

We model location measurement errors on the line by perturbing locations with the following algorithm. Agents are at consecutive integer coordinates from 1 to $|\Lambda_\tau|$. Each agent's integer location is independently perturbed by adding a random amount ξ from a uniform distribution on $[-v,v]$. In other words agent i is given a perturbed location $\tilde{s}_i = s_i + \xi_i$. Then, each agent's measured location is defined by assigning the perturbed locations $\{\tilde{s}_i\}$ to integers from 1 to $|\Lambda_\tau|$, according to the rank order of the $\{\tilde{s}_i\}$ from smallest to largest. The resulting maximum shift in each coordinate is bounded by 2v. We vary the amount of reshuffling of agents' locations by examining seven different values for v.

We use an analogous process on the plane. Agents true locations are points in a square integer lattice and are subjected to the same perturbation method as used on the line, independently for each coordinate. Specifically, agent i's horizontal coordinate, s_{1i}, is perturbed by adding a bounded error, so that $\tilde{s}_{1i} = s_{1i} + \xi_{1i}$, with $\xi_{1i} \overset{IID}{\sim} Unif[-v,v]$.[16] Then the $\{\tilde{s}_{i1}\}$ are mapped back to integer coordinates. For observations with common true vertical coordinate, s_2, (potentially) mismeasured horizontal coordinates are assigned according to the rank order of their $\tilde{s}_{i1}$. Agents' vertical coordinates are then perturbed using the same approach. The resulting maximum shift in each coordinate is bounded by 2v. Again, we vary the amount of measurement error by choosing seven different values of v (distinct from those on the line).

This DGP's properties of independence across coordinates, and the measured coordinates being perturbed ranks are directly relevant for many applications with agents indexed by characteristics other than their physical location. Such characteristics are routinely quite different things, measured in such different ways, or even come from different data sources so that independence across their measurement errors is plausible. For example, a study of firm-level productivity could use coordinates derived from characteristics like the input share of technology-intensive inputs for its SIC code, its investment in R&D, or computer technology utilization measures. These three characteristics might well be measured using data from different sources, e.g.: benchmark input-output data, census of manufacturing data, and employee surveys. Other examples include firm coordinates of book to market and sales, or country coordinates based on trade costs and ethnic/demographic composition. As for ranks being the object of interest, it is commonly the case that agent characteristics can be measured in more than one way, e.g. firm size by revenue, market capitalization, or labor force. In many such situations, it is plausible that the researcher has much more confidence in the approximate rank order of the agents than the measurement of the cardinal gaps between

---

changes in both this function and the 'lag measure.'

[15] In the statistics literature known to us, independent location errors are the most studied, see e.g. Gabrosek and Cressie (2002).

[16] Uniform location errors, independent across coordinates are also investigated by Gabrosek and Cressie (2002), though they work direcly with an analog of $\tilde{\mathbf{s}}_i$ instead of mapping to integer locations.

them, as the latter would be much more likely to vary with different ways of measuring the characteristic. Our DGP is a natural choice for localized errors in ranks.

We think our results will still be useful, even when our DGP is not of direct interest for the application. We anticipate that the key question for most if not all applications will be the relative magnitude of measurement error in locations/distances versus the strength of the spatial correlation in the data. By using a simple DGP and location error process we hope to get a handle on at least rough thresholds for this relative magnitude that determine relative ranking of our estimators. This information on rough thresholds in terms of relative magnitude of measurement error (versus dependence) may prove valuable in disparate applications, as researchers will often have a satisfactory idea of the precision of their constructed distance measure relative to the strength of dependence in the data. For example, a researcher might have an idea of the precision of travel-time-based economic distances and typical maximum commute times might serve as a plausible upper bound for lags with appreciable spatial correlation.

Illustration of Measurement Error Processes

Table 1 and Figure 2 are meant to provide some sense of how much change in locations on the line is induced by each level of measurement error. The percentages of agents' measured locations that are at different displacements from their true locations are given in Table 1. Our smallest level of measurement error leaves 75% of locations unchanged and the remainder moved only one unit. For a sixth-order moving average, we think this is reasonably thought of as a small level of error. On the other extreme, our level 7 measurement errors shift a majority of the locations by two or more units and about 30% are shifted 3 or more units. We consider this to be a substantial amount of error for an MA(6).

The effect of location measurement error can be partially characterized as changing the autocorrelation function of a process. Since our DGP for X_s has no trend (it is mean zero) the first moment of the process at mismeasured locations will be of course unchanged. The sample covariance at a given distance will converge to a weighted average of the true covariances of the process. The relative contributions of covariances at other distances will depend on the measurement error process. For example, under level 1 measurement errors, observations measured as being 2 units apart will consist of many pairs that are really 2 units apart and some pairs whose true distances are 1, 3, and 4. Thus, the general effect of measurement error in distances will be to smooth out the true covariance function and extend the range of measured distances with nonzero covariances for finite order MAs.

This general effect of the altered autocovariances flattening out as measurement error increases is illustrated in Figure 2. In this Figure we plot the true autocorrelations for X_s with an approximation for the autocorrelations at each measured distance under levels 1, 4, and 7 of measurement error (obtained by Monte Carlo integration) for $\rho = .3$.[17] The flattening out of the correlation function is consistent with the results obtained by Gabrosek and Cressie (2002) and Cressie and Kornak (2003) for different DGPs.

---

[17]For location errors of level 1, Appendix D reports the exact analytic expression for the covariance function of the mismeasured process.

Table 2 and Figures 3 and 4 present summaries of our measurement error process on the plane, analogous to those above. Table 2 presents the values of v we use on the plane and a description of the Euclidean distances from their true locations. The same value of v is used for both the vertical and horizontal dimension. Our level 1 measurement error shifts only 20% of locations, and the new location is at most at distance $\sqrt{2}$ from the original one. The discrepancy in the correlations at lags that differ in this range can be inferred by looking at Figures 3 and 4. We consider this to be a small amount of error. On the other extreme, our largest level of measurement errors shift 89% of the locations, and move about 40% of them to new location that are at distance greater than 2 from the original ones. Given the true locations correlation functions plotted in Figures 3 and 4, we consider this to be a substantial amount of error.

Our DGP on the plane is not isotropic, the covariance at distances that correspond to "diagonals" are not the same as straight vertical or horizontal displacements, e.g. $C([3\ 4]) \neq C([0\ 5])$ (see Appendix C). But the average covariances still offer a useful illustration of the effect of measurement errors. Thus, we plot the averages of correlations at each distance for the truth and measurement error of levels 1, 4 and 7 in Figure 3 for $\rho = .3$ and Figure 4 for $\rho = .45$ (error versions are obtained by Monte Carlo integration). Again, these Figures illustrate the effect of increases in this type of measurement error flattening out and extending the altered autocorrelation function relative to the true autocorrelation function.

It appears that few generalizations can be made about the position of the altered covariance function relative to the true covariance function at any given distance. If the true covariances are monotonically decreasing, then the altered covariance at the shortest distance will of course be below the true covariance at that distance for any measurement error process that is independent of the realizations of X_s. The altered covariance can only include pairs of observations truly at the minimum distance or greater. But beyond the fact that the altered covariance function must have the same integral as the true covariance function (V is invariant to measurement error in locations), there does not seem to be any necessary relationship between true and altered covariances for a general measurement error process and DGP.

3.3   V Estimators and Specification Tests

Parametric Estimators

We compute the MLE for our Gaussian DGPs and an MM estimator.

On both the line and on the plane, our estimate of V for the MLE, $\hat{V}_{MLE}$, is formed by plugging the MLE point estimates for σ and ρ into the analytic expression for V. Our process and sampling framework ensure consistency and asymptotic normality of the MLE estimator.[18] In particular, the following Proposition holds:

---

[18] Other conditions for consistency and asymptotic normality of likelihood based estimators are provided, among others, by Heijmans and Magnus (1986) and Cressie and Lahiri (1996).

Proposition 2 *For correctly measured distances, the Gaussian DGPs and sampling framework for the* $\{X_{\mathbf{s}}\}$ *process satisfy the conditions in Theorem 3 of Mardia and Marshall (1984). Hence the MLE estimator is consistent and asymptotically normal.*

Proof. See Appendix B. ∎

The moments in the MM estimator are simply the nonzero autocorrelations of X and its second moment. For the DGP on the line, these correlations are: $\Gamma(\rho) = \left[\frac{C(1;\rho,\sigma)}{C(0;\rho,\sigma)}, ..., \frac{C(6;\rho,\sigma)}{C(0;\rho,\sigma)}\right]$. We obtain sample correlations using unbiased covariance estimates $\hat{C}(\cdot)$ and form $\hat{\Gamma} = \left[\frac{\hat{C}(1)}{\hat{C}(0)}, \ldots, \frac{\hat{C}(6)}{\hat{C}(0)}\right]$, estimating $\hat{\rho}_{MM}$ as:

$$\hat{\rho}_{MM} = \arg\min_{\rho} \left[\Gamma(\rho) - \hat{\Gamma}\right]' \left[\Gamma(\rho) - \hat{\Gamma}\right].$$

Once we have the estimate $\hat{\rho}_{MM}$, we can estimate σ^2 by means of equation (1). In particular, we can use the sample variance of X_s as an estimate of $C(0)$, and then estimate σ^2 with $\hat{\sigma}^2_{MM} = \hat{C}(0)\left[1 + 2\sum_{j=1}^m \hat{\rho}_{MM}^{2j}\right]^{-1}$. Once we have these estimates, we can get $\hat{V}_{MM}$ by plugging $\hat{\rho}_{MM}$ and $\hat{\sigma}^2_{MM}$ in (2).

The estimator used for the process on the plane is constructed in the same manner. The moments used are again the nonzero correlations. Denoting these nonzero correlations by $\Upsilon(\rho)$ and their sample analogs by $\hat{\Upsilon}$, we get our estimate $\hat{\rho}_{MM}$ by solving:

$$\hat{\rho}_{MM} = \arg\min_{\rho} \left[\Upsilon(\rho) - \hat{\Upsilon}\right]' \left[\Upsilon(\rho) - \hat{\Upsilon}\right]$$

Once we have the estimate $\hat{\rho}_{MM}$, we again estimate σ^2 by using the second moment and then obtain the plug-in $\hat{V}_{MM}$ estimator using $\hat{\rho}_{MM}$ and $\hat{\sigma}^2_{MM}$. These MM estimators are special cases of GMM estimators whose consistency is demonstrated in Conley (1999).

Nonparametric Estimators

The nonparametric estimator $\hat{V}_{NP}$ takes the form

$$\hat{V}_{NP} = \frac{1}{|\Lambda_\tau|} \sum_{i=1}^{|\Lambda_\tau|} \sum_{j=1}^{|\Lambda_\tau|} K_\tau(\mathbf{s}_i - \mathbf{s}_j) \cdot (X_{\mathbf{s}_i} - \bar{X}) \cdot (X_{\mathbf{s}_j} - \bar{X})$$

with

$$K_\tau(s_i - s_j) = \begin{cases} 1 & \text{if } |s_i - s_j| \le L_\tau, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

for the process on the line, and

$$K_\tau(\mathbf{s}_i - \mathbf{s}_j) = \begin{cases} 1 & \text{if } |s_{i1} - s_{j1}| \le L_\tau, |s_{i2} - s_{j2}| \le L_\tau, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

for the process on the plane.

Figure 5 illustrates the effect measurement error in distances has upon the weighting function K_τ for locations on the line. Similar effects occur for locations on the plane. The weighting function with true distances will be a uniform kernel putting weight 1 on all distances up to the cutoff amount L_τ and then zero for larger distances. The introduction of measurement error can be interpreted as a change in the kernel K_τ from one that puts the same weight on any pair of observations that have the same true distance to a kernel that puts the weight 1 on a fraction of observations at each true distance given by the graph. With distance measurement error, some observations will get a weight of 1 early relative to when they would get this weight with true locations as τ grows, and some will get a weight of 1 late. Eventually, all observations will get weight 1 as $\tau \to \infty$, so $\hat{V}_{NP}$ remains consistent. The source of the robustness of $\hat{V}_{NP}$ to measurement error is also illustrated in this Figure, as even for higher levels of measurement error, the large majority of observations will get the same weight with measurement error as they would with the true distances.

Specification Tests

We present two types of specification tests that are joint tests for proper specification of the parametric estimators and correctly measured distances. The first uses the asymptotic distribution of the difference between parametric and nonparametric estimators and the second uses a parametric bootstrap procedure to sample from the process implied by our MLE point estimates to approximate the finite sample distribution of our nonparametric estimator under the null hypothesis.

Our t-tests are based on the asymptotic distribution of the difference between the parametric estimator and a nonparametric estimator of V ($\hat{V}_{NP}$ or its approximately unbiased analog, which with abuse of terminology we will refer to as $\hat{V}_{NP}$ unbiased). The parametric estimators converge at a faster rate than the nonparametric estimators, therefore only the sampling variation in the nonparametric estimator is relevant for the first-order asymptotic distribution of the difference between the parametric and nonparametric estimator. On the line we can apply the results of Anderson (1994), Theorem 9.4.1 and Corollary 9.4.1 to obtain the the limiting distribution of the difference between $\hat{V}_{NP}$ and $\hat{V}_i$. On the plane, the asymptotic distribution of the difference between the parametric estimator and the nonparametric estimator of V is derived building on Rosenblatt (1985, Chapter 5, Theorem 7) and Mardia and Marshall (1984, Theorem 3). We state these limiting distribution results as Proposition 3:

Proposition 3 *Given our DGP for X_s on the line and correctly measured distances, the uniform kernel function $K_\tau(x)$ in (3), and choosing L_τ so that $L_\tau \to \infty$ and $\frac{L_\tau}{|\Lambda_\tau|} \to 0$ as $|\Lambda_\tau| \to \infty$, it follows that:*

$$\sqrt{\frac{|\Lambda_\tau|}{L_\tau}} \left(\hat{V}_{NP} - \hat{V}_i \right) \to N\left(0, 4V^2\right), \quad i = MLE, MM.$$

Given our DGP for $X_\mathbf{s}$ on the plane and correctly measured distances, the uniform kernel function $K_\tau(\mathbf{x})$

in (4), and choosing L_τ so that $L_\tau \to \infty$ and $\frac{L_\tau}{|\Lambda_\tau|^{1/2}} \to 0$ as $|\Lambda_\tau| \to \infty$, it follows that:

$$\sqrt{\frac{|\Lambda_\tau|}{L_\tau^2}} \left(\hat{V}_{NP} - \hat{V}_i \right) \xrightarrow{d} N\left(0, 8V^2\right), \quad i = MLE, \ MM.$$

Proof. See Appendix B. ∎

Based on this result, we construct two types of $t-$tests. On the plane they take the form:

$$t_1 = \frac{\sqrt{\frac{|\Lambda_\tau|}{L_\tau^2}} \left(\hat{V}_{NP} - \hat{V}_i \right)}{\sqrt{8\hat{V}_i^2}}, \quad i = MLE, \ MM,$$

$$t_2 = \frac{\sqrt{\frac{|\Lambda_\tau|}{L_\tau^2}} \left(\hat{V}_{NP} - \hat{V}_i \right)}{\sqrt{8\hat{V}_{NP}^2}}, \quad i = MLE, \ MM;$$

their form on the line is analogous. These two $t-$statistics differ in the estimator used for the variance of the limiting distribution. Under the joint null of proper specification and correctly measured distances both $\hat{V}_{NP}$ are $\hat{V}_i$ are of course consistent, so either could be used to estimate the denominator. Typical practice in related tests is to use t_1, due to its denominator's faster convergence rate. We are motivated to investigate t_2 as well because, in the results below, with distance measurement error there is substantial downward bias of the parametric V estimators. This heavily influences the finite sample distribution of the test statistic, inflating the value of t_1. The contribution to the power of the test from this downward bias might not occur with other DGPs. $\hat{V}_{NP}$ has better and more stable (as level of errors changes) bias properties so the power properties of t_2 may be more generalizable to other DGPs.

The second test is based on a parametric bootstrap.[19] Under the joint null hypothesis of a correctly specified MLE and perfect location/distance information, our MLE estimates are consistent. Since we have the full likelihood specification, under the null hypothesis we can use our MLE estimates to obtain simulated samples from a consistent estimate of the true DGP formed from the parametric distribution specification evaluated at the MLE point estimates. We use this distribution to simulate a large number of independent bootstrap draws of the same size as the original data. We then obtain nonparametric V estimates for each bootstrap sample. The resulting set of bootstrap V estimates provides critical values for an acceptance region for the nonparametric estimate using the original data. In other words, to get a 90% acceptance region for $\hat{V}_{NP}$ we apply this estimator to each of our bootstrap simulations and our acceptance region is between the 5th and 95th percentiles of these bootstrap estimates. The specification test then consists of simply observing whether $\hat{V}_{NP}$ estimated with the original data is in this range. If $\hat{V}_{NP}$ is in this range, it is consistent with data generated from what is, under the null hypothesis, a consistent estimate of the true DGP. If $\hat{V}_{NP}$ lies outside the range, we reject the joint null hypothesis that the MLE is properly specified and locations/distances are correct.

---

[19]A similar parametric bootstrap is used by Conley, Hansen, and Liu (1997) in the context of diffusion models.

This parametric bootstrap has an advantage over tests based on the limiting distribution as the quality of acceptance regions are not dependent on a good choice for bandwidth/kernel. Any choice of bandwidth or kernel used in $\hat{V}_{NP}$ will automatically be accounted for in the acceptance region as the same choices will be applied to the bootstrap draws as used for the real data. Of course, the power of the test will be affected by bandwidth choice, so it is not irrelevant, but at least the reference distribution itself will be appropriate regardless of bandwidth choice. Of course, the downside of this approach is that it requires that the full DGP be specified, rather than just the autocovariance function.

4   Simulation Results

This Section reports the results of 1000 repetitions of a Monte Carlo experiment based on a sample of size $|\Lambda_\tau|$ equal to 500 on the line, and equal to a 40 by 40 grid on the plane. We present simulation results for $\rho = .3$ on the line, and for $\rho = .3, .45$ on the plane, with σ in each case adjusted so that the true value of $V = 1$ in all simulations.[20]

Choice of smoothing parameter is a key decision in the application of kernel HAC estimators like $\hat{V}_{NP}$, so we are motivated to examine estimators performance under different smoothing parameters, L_τ. On the plane we experiment with cutoff points from $L_\tau = 3$ to $L_\tau = 7$. This range includes cutoffs that are 'too small' to those that are 'too large' given our DGP. In particular, with $L_\tau = 3$ several cross products of observations with non-zero covariance are left out of the weighted average giving the estimate of V.[21] On the other hand, $L_\tau = 5$ is roughly the ideal cutoff value: the only cross products with non-zero covariance left out of the weighted average are those between $X_\mathbf{s}$ and $X_{\mathbf{s}\pm[0,6]}$, or $X_\mathbf{s}$ and $X_{\mathbf{s}\pm[6,0]}$; each observation has only four neighbors at these locations, and the corresponding covariance is very close to zero both for $\rho = .3$ and $\rho = .45$. Additionally, just a few cross-products with covariance equal to zero are added to the weighted average.[22] Hence, we expect $L_\tau = 5$ to be the cutoff point giving us the best performance for the nonparametric estimator. On the other hand, $L_\tau = 7$ is a cutoff that is a little too large relative to the order 6 of our moving average. In particular, a relatively large number of cross-products with covariance equal to zero are added to the weighted average.[23] To conserve space, we report only results on the line for a cutoff of $L_\tau = 8$ a little 'too large' relative to our DGP.

---

[20]Additionally, we experimented with values of $\rho = .6, .9$ and a sample size of 1000 and with different cutoff points on the line, and with $\rho = .6, .9$ on the plane. Qualitatively, as we change sample size and amount of spatial correlation, the results are the same as those obtained with smaller sample size on the line, and those obtained with medium amount of spatial correlation on the plane. These results are available from the authors upon request.

[21]To be precise, 53% of the cross products involving a sufficiently interior point (at least 6 units from the boundary) with non-zero expectation are left out of the weighted average with $L_\tau = 3$.

[22]16% of the cross products involving a sufficiently interior point and entering the weighted average with $L_\tau = 5$ have expectation equal to zero.

[23]49% of the cross products involving a sufficiently interior point and entering the weighted average with $L_\tau = 7$ have expectation equal to zero.

4.1 Locations On the Line

V Estimator Performance

Table 3 collects the results obtained when the locations are perfectly measured and when we have location errors of level one through seven, for $\rho = .3$. The Table reports Bias and Root Mean Squared Error (RMSE) for $\hat{V}_{MLE}$, $\hat{V}_{MM}$, $\hat{V}_{NP}$ and its unbiased analog, as well as coverage probabilities for 95% confidence intervals for EX constructed using the alternative variance estimators.

The first four columns of Table 3 show that when the true locations are used, the bias associated with $\hat{V}_{NP}$ is bigger (in absolute terms) than that associated both with $\hat{V}_{MLE}$ and $\hat{V}_{MM}$. As expected, $\hat{V}_{MLE}$ has the smallest bias (in absolute terms) among the estimators we are considering, and the unbiased analog of $\hat{V}_{NP}$ has a smaller bias than $\hat{V}_{NP}$, although this difference is negligible on the line (See Guyon (1982)). As the level of the location errors ranges from one to seven, the bias of $\hat{V}_{MLE}$ increases sharply (in absolute terms), and with errors of level 1 it is already substantially bigger than that of $\hat{V}_{NP}$ and $\hat{V}_{MM}$. The bias of $\hat{V}_{MM}$ also increases sharply with the level of location errors, although not as rapidly as that of $\hat{V}_{MLE}$, and it surpasses the bias of $\hat{V}_{NP}$ (and its unbiased analog) for errors of level 2 and higher. On the other hand, the bias of the nonparametric estimators is relatively constant with respect to the different levels of location errors.

Similar patterns can be observed when looking at the RMSE: the RMSE associated with $\hat{V}_{NP}$ (and its unbiased analog) is higher than that associated with $\hat{V}_{MLE}$ and $\hat{V}_{MM}$ with true locations. However, if locations are incorrectly measured, the nonparametric estimators' performance varies little as the level of location errors increases. In contrast, the RMSE of $\hat{V}_{MLE}$ deteriorates rapidly, and at errors of level 2 and higher it becomes worse than that of the nonparametric estimator. The RMSE of $\hat{V}_{MM}$ does not increase as rapidly as that of $\hat{V}_{MLE}$. However, as soon as location errors of level 4 and higher are introduced, both the bias and the RMSE of $\hat{V}_{MM}$ get worse than that of the nonparametric estimators.

The behavior of the RMSE of the four estimators can be better understood by looking at the deciles of their distributions. When the locations are accurately measured and $\rho = .3$, the 10th percentile-90th percentile range for $\hat{V}_{MLE}$ is $[0.84, 1.18]$, that for $\hat{V}_{MM}$ is $[0.81, 1.21]$, and that for $\hat{V}_{NP}$ and its unbiased analog are both $[0.67, 1.30]$. Once we introduce location errors of level 7, the 10th-90th percentile range becomes $[0.42, 0.61]$ for $\hat{V}_{MLE}$, $[0.45, 0.79]$ for $\hat{V}_{MM}$, while that for $\hat{V}_{NP}$ and its unbiased analog is virtually unchanged at $[0.67, 1.28]$. The distribution of $\hat{V}_{MLE}$ remains tight, but shifts to the left, with a downward bias for the estimate of V. The distributions of the $\hat{V}_{MM}$ also shifts to the left, but less than that of $\hat{V}_{MLE}$. The distributions of the nonparametric estimators are relatively unaffected by the introduction of the location errors.

Perhaps the best measure of these estimators' performance is their coverage probabilities. The 95% confidence intervals for EX constructed using $\hat{V}_{NP}$ cover its true value of zero in approximately 95% of the Monte Carlo draws for all levels of location errors. In contrast, the coverage probabilities of the 95% confidence intervals constructed using $\hat{V}_{MLE}$ deteriorate substantially with a rise in the level of the

location errors. While the coverage probability is 96.5% with correctly measured locations, it goes down to 84.8% with location errors of level seven. At correctly measured locations, the coverage probability of 95% confidence intervals constructed using $\hat{V}_{MM}$ is 96.4%. In the presence of location errors of level 7, the results mirror the 10-90th percentile ranges described above: the coverage probability goes down to 88%.

Specification Tests

Table 4 reports the results of our specification tests. The left section of this table reports the results of the $t-$tests, both when the variance of the limiting distribution is estimated using $\hat{V}_i$, $i = MLE, MM$, and when it is estimated using $\hat{V}_{NP}$. The t-test comparing the nonparametric and the MM estimates performs poorly, both in terms of size and power of the test, especially when the variance of the limiting distribution is estimated using $\hat{V}_{NP}$. This result is not surprising, considering that $\hat{V}_{MM}$ is not greatly affected by the presence of moderate levels of location errors. On the other hand, the t-test comparing $\hat{V}_{NP}$ and $\hat{V}_{MLE}$ seems to have desirable properties when the variance of the limiting distribution is estimated using $\hat{V}_{MLE}$. However, as can be seen by looking at the same test when the variance of the limiting distribution is estimated using $\hat{V}_{NP}$, this result is in part driven by the fact that in the presence of location errors, $\hat{V}_{MLE}$ tends to underestimate V.[24]

The right section of Table 4 presents parametric bootstrap test results. Our specific procedure for each Monte Carlo simulation, is to first use the corresponding estimates $(\hat{\rho}_{MLE}, \hat{\sigma}_{MLE})$ to construct the variance-covariance matrix of the vector $[X_s, s \in \Lambda_\tau]$, Σ. We then draw $|\Lambda_\tau|$ observations from a distribution $N(0, \Sigma)$, and estimate $\hat{V}_{NP}^B$ for the simulated bootstrap sample B. Following this procedure, we construct 200 bootstrap samples for each Monte Carlo repetition, and check whether $\hat{V}_{NP}$ estimated with that repetition's original data is in the acceptance region given by the 5th to 95th percentile range of the bootstrap estimates. In contrast to the tests using the asymptotic distribution, the parametric bootstrap test comparing $\hat{V}_{NP}$ and $\hat{V}_{MLE}$ appears to have good properties. Both using the nonparametric estimator and its unbiased analog, the test has a size closer to the desired one, and has good power properties.

4.2 Locations On Plane

V Estimator Performance

Tables 5 and 6 collect the results obtained when the locations are perfectly measured and when we have location errors of level one through seven, for $\rho = .3$ and $\rho = .45$ respectively. Each table reports Bias and Root Mean Squared Error (RMSE) for $\hat{V}_{MLE}$, $\hat{V}_{MM}$, $\hat{V}_{NP}$ and its unbiased analog estimated with three cutoff points, $L_\tau = 3, 5, 7$, as well as coverage probabilities for 95% confidence intervals for EX constructed using the alternative variance estimators. Several of the results reported below are similar to those obtained with locations on the line.

---

[24]An analogous downward bias is also evident with $\hat{V}_{MM}$. However the $t-$test involving $\hat{V}_{MM}$ does not have good power properties even when $\hat{V}_{MM}$ is used in the denominator of the $t-$statistic.

Tables 5a and 6a show that when the true locations are used, the bias associated with $\hat{V}_{NP}$ is bigger (in absolute terms) than that associated both with $\hat{V}_{MLE}$ and $\hat{V}_{MM}$, both for small and medium amounts of spatial correlation. As expected, $\hat{V}_{MLE}$ has the smallest bias (in absolute terms) among the estimators we are considering. The unbiased analog of $\hat{V}_{NP}$ has a smaller bias than $\hat{V}_{NP}$, and now that the locations are on the plane, this difference is more substantial (See Guyon (1982)). As the level of the location errors ranges from one to seven, the bias of $\hat{V}_{MLE}$ increases sharply (in absolute terms), and even with errors of level 1 it is substantially bigger than that of $\hat{V}_{NP}$ and $\hat{V}_{MM}$. This occurs both for $\rho = .3$ and $\rho = .45$. At small levels of spatial correlation ($\rho = .3$), the bias of $\hat{V}_{MM}$ also increases sharply with the level of location errors, although not as rapidly as that of $\hat{V}_{MLE}$, and it surpasses the bias of $\hat{V}_{NP}$ (and its unbiased analog) for errors of level 3 and higher. For medium levels of spatial correlation ($\rho = .45$) the bias of $\hat{V}_{MM}$ remains surprisingly low as the level of location errors increases. Both for $\rho = .3$ and $\rho = .45$, the bias of the nonparametric estimators with $L_\tau \geq 5$ is relatively constant with respect to the different levels of location errors. When the cutoff parameter is equal to 3, the bias of $\hat{V}_{NP}$ is more sensitive to errors in locations, and it increases with the level of locations error, especially for $\rho = 0.45$, although not as sharply as that of $\hat{V}_{MLE}$ (or $\hat{V}_{MM}$ with $\rho = 0.3$). This is because with $L_\tau = 3$ many cross-products with non-zero covariances are left out of the weighted average giving the estimate of V. With correctly measured locations, the covariances of the cross products left out are relatively small, and therefore their influence on the bias is not much higher than what we have with larger cutoff points. However, when locations are measured with error some of the cross products left out have a fairly high covariance, and therefore the bias is higher.

The RMSE associated with $\hat{V}_{NP}$ (and its unbiased analog) is higher than that associated with $\hat{V}_{MLE}$ and $\hat{V}_{MM}$ with true locations, both at small and medium levels of spatial correlation. However, if locations are incorrectly measured, the nonparametric estimators' performance varies little as the level of location errors increases (especially for $L_\tau \geq 5$). In contrast, the RMSE of $\hat{V}_{MLE}$ deteriorates rapidly, and at errors of level 1-2 and higher it becomes worse than that of the nonparametric estimators. The RMSE of $\hat{V}_{MM}$ does not increase as rapidly as that of $\hat{V}_{MLE}$. However, for $\rho = .3$, as soon as location errors of level 4 and higher are introduced, both the bias and the RMSE of $\hat{V}_{MM}$ become worse than that of the nonparametric estimators with cutoff point $L_\tau \leq 5$; the same happens for errors of level 6 and 7 when the cutoff point of the nonparametric estimators is $L_\tau = 7$. For $\rho = .45$, the RMSE of $\hat{V}_{MM}$ becomes worse than that of both nonparametric estimators with cutoff point $L_\tau \leq 5$ when location errors of level 4 and higher are introduced. However, it remains similar or better than that of the nonparametric estimators with cutoff point $L_\tau = 7$ despite the presence of location errors.

Again it is informative to look at the deciles of our estimators' distributions. When the locations are accurately measured and $\rho = .45$, the 10th-90th percentile range for $\hat{V}_{MLE}$ is $[0.93, 1.06]$ and that for $\hat{V}_{MM}$ is $[0.74, 1.25]$. The corresponding ranges for $\hat{V}_{NP}$ across the different cutoff parameters are $[0.64, 1.06]$, $[0.55, 1.25]$ and $[0.39, 1.32]$ for $L_\tau = 3, 5, 7$ respectively. With location errors of level 7, the

10th-90th percentile range becomes $[0.15, 0.23]$ for $\hat{V}_{MLE}$ and $[0.41, 1.41]$ for $\hat{V}_{MM}$. For $\hat{V}_{NP}$ they become $[0.49, 0.85]$, $[0.54, 1.18]$, and $[0.41, 1.30]$ for $L_\tau = 3, 5, 7$ respectively. Hence, as we observed for locations on the line, when location errors are introduced the distribution of $\hat{V}_{MLE}$ remains tight, but shifts dramatically to the left. In contrast, the distribution of $\hat{V}_{MM}$ spreads out, but it remains relatively centered around the true value of V. For $L_\tau = 5, 7$ the percentile ranges of $\hat{V}_{NP}$ are also centered around the true value of V. These ranges are relatively unaffected by the introduction of location errors. On the other hand, for $L_\tau = 3$ the 10th-90th percentile range is centered below the true value of V, and this feature becomes more pronounced as location errors are introduced, but of course not as severely as for $\hat{V}_{MLE}$. Unsurprisingly, the spread of the 10th-90th percentile range for $\hat{V}_{NP}$ increases with L_τ.[25]

When $\rho = .3$, the distribution of $\hat{V}_{MM}$ is more affected by measurement error. With correct locations, the 10th-90th percentile range for $\hat{V}_{MM}$ is $[0.78, 1.20]$, while that of $\hat{V}_{MLE}$ is $[0.86, 1.12]$. The 10th-90th percentile ranges of $\hat{V}_{NP}$ are $[0.67, 1.12]$, $[0.55, 1.27]$, and $[0.38, 1.34]$ for $L_\tau = 3, 5, 7$ respectively. When errors of level 7 are introduced, the 10th-90th percentile range for $\hat{V}_{MLE}$ becomes $[0.19, 0.28]$, while those for $\hat{V}_{NP}$ become $[0.55, 0.93]$, $[0.54, 1.18]$, and $[0.42, 1.33]$ for $L_\tau = 3, 5, 7$ respectively.[26] This behavior is similar to what we observed for $\rho = .45$. On the other hand, the 10th-90th percentile range for $\hat{V}_{MM}$ becomes $[0.37, 0.71]$. Hence the distribution of $\hat{V}_{MM}$ remains tight, but in contrast to what happened for $\rho = .45$, it shifts to the left, with a downward bias for the estimate of V.

Looking at the coverage probabilities of these estimators, we observe that the 95% confidence intervals for EX constructed using $\hat{V}_{NP}$ cover zero in approximately 89 to 94% of the Monte Carlo draws when $\rho = .3$, and in approximately 90 to 94% for $\rho = .45$. Most of this variation is across cutoff parameters, and does not depend on the level of location errors. In contrast, the coverage probabilities of the 95% confidence intervals constructed using $\hat{V}_{MLE}$ deteriorate with a rise in the level of the location errors. While the coverage probability is approximately 95% with correctly measured locations, it goes down to 68% with location errors of level seven when $\rho = .3$, and to 62% when $\rho = .45$. The performance of the MM estimator varies substantially with the value of ρ. When $\rho = .3$ the coverage probability deteriorates with the increase of the level of measurement error, though not as severely as $\hat{V}_{MLE}$ (going from 95.3% with true locations to 84.3% with level 7 location errors). When $\rho = .45$ the coverage probability falls only slightly, going from 95.3% to 92.1%.

We find it helpful in understanding the reason the MM estimator performs so surprisingly well for $\rho = .45$ to think of V as the product of two terms: $V = C(0)\left[\frac{V}{C(0)}\right]$. The first term is $C(0)$ which depends on both σ and ρ; the second term is the ratio $\frac{V}{C(0)}$, which is a function of ρ alone. The MM estimator uses

---

[25] The unbiased versions of $\hat{V}_{NP}$ have almost the same percentile ranges as $\hat{V}_{NP}$. For correctly measured locations, these are $[0.68, 1.13]$, $[0.57, 1.35]$, and $[0.37, 1.44]$ for $L_\tau = 3, 5, 7$ respectively. With location errors of level 7, they become $[0.53, 0.91]$, $[0.59, 1.30]$, and $[0.43, 1.46]$ for $L_\tau = 3, 5, 7$ respectively.

[26] Also in this case the unbiased versions of $\hat{V}_{NP}$ have almost the same percentile ranges as $\hat{V}_{NP}$. For correctly measured locations, these are $[0.70, 1.19]$, $[0.55, 1.37]$, and $[0.35, 1.46]$ for $L_\tau = 3, 5, 7$ respectively. With location errors of level 7, they become $[0.58, 1.00]$, $[0.58, 1.31]$, and $[0.45, 1.49]$ for $L_\tau = 3, 5, 7$ respectively.

the sample variance to estimate $C(0)$, and matches sample correlations to get an estimate of ρ which is then plugged into the analytic expression for $\frac{V}{C(0)}$ to get an estimate of this term. $\hat{V}_{MM}$ is then just the product of the estimates of $C(0)$ and $\frac{V}{C(0)}$.

The good performance of $\hat{V}_{MM}$ for $\rho = .45$ is a consequence of three things. First, obviously the estimation of $C(0)$ using the sample variance is not affected by measurement error. Second, $\frac{V}{C(0)}$ is a nonlinear function of ρ and for values of ρ above about $1/3$ there is more than one value of ρ that returns the same value of $\frac{V}{C(0)}$. In addition, it turns out that when there is measurement error, the argument minimizing the minimum distance criterion function happens very often to be near a value of ρ that implies a value of $\frac{V}{C(0)}$ that is very close to the true value of $\frac{V}{C(0)}$ for $\rho = .45$. The limiting (population) minimum distance criterion function is not in fact minimized at a value of ρ that gives exactly the same $\frac{V}{C(0)}$ as $\rho = .45$, but they are close enough to be practically indistinguishable in samples of the size we consider.

It is important to note that the MM estimator achieves this good performance with $\rho = .45$ by picking values of ρ and σ that are often not close to the true parameter values. When the true values of ρ and σ are respectively $.45$ and $.1423$, the estimates $(\hat{\rho}_{MM}, \hat{\sigma}_{MM})$ are quite affected by the presence of location errors. If we look at the deciles of their distributions, we observe that for accurately measured locations the 10th-90th percentile range for $\hat{\rho}_{MM}$ is $[.39, .51]$ and for $\hat{\sigma}_{MM}$ it is $[.13, .16]$, while for location errors of level 7 they become, respectively, $[.26, 2.55]$ and $[.005, .19]$. If all that is needed is an estimate of V, this is of no consequence. However, it is relevant if ρ or σ is of independent interest.

Specification Tests

Table 7 reports the results of the specification $t-$tests, both when the variance of the limiting distribution is estimated using the parametric V estimator and when it is estimated using $\hat{V}_{NP}$. Our results for these t-tests are very similar in nature across different values of ρ and cutoffs L_τ. Therefore, to conserve space we report results in Table 7 for a representative choice of $\rho = .3$ and for the cutoff $L_\tau = 7$. All the tests involving $\hat{V}_{MM}$ have low power, which is not surprising given the good performance of $\hat{V}_{MM}$ as an estimator of V described above. The only case where the $t-$test has good power is when $\hat{V}_{MLE}$ is used for the denominator. The pronounced downward shift in the distribution of $\hat{V}_{MLE}$ in the presence of measurement error appears to lead to the reasonable power performance for this case. An important caveat to this apparent power is that this magnitude or sign of shift may not occur with other DGPs. Using $\hat{V}_{NP}$ in the denominator to provide robustness across DGPs choices does not appear to be worthwhile as all $t-$tests with this denominator have very poor performance.

The test based on the parametric bootstrap is constructed in a manner analogous to that used on the line. For each Monte Carlo simulation, we construct 200 bootstrap draws from the distribution corresponding to the DGP at $\hat{\rho}_{MLE}, \hat{\sigma}_{MLE}$. For each Monte Carlo replication we check whether $\hat{V}_{NP}$ estimated with that replication's original data is in the acceptance region given by the 5th to 95th percentile of the bootstrap estimates.

Table 8 reports the results of these bootstrap specification tests, for $\rho = .3$, and $\rho = .45$, with two cutoff parameter choices $L_\tau = 3, 7$. We present full results for only two cutoffs to conserve space; these two were chosen as the largest and smallest with which we experimented. They allow us to examine performance of our test with too small and too large a cutoff parameter. When $\rho = .45$, the test has roughly the desired size with both cutoff parameters. It has outstanding power with the smaller cutoff of $L_\tau = 3$, with rejection probabilities of 90% and higher with any level of measurement error. Power is decreased when $L_\tau = 7$ though rejections still reach about 60% by error level 2. With the smaller persistence parameter $\rho = .3$, when $L_\tau = 7$ the test remains about properly sized with rejection probabilities not rising quite as fast, reaching about 60% at error level 3. For $\rho = .3$ and $L_\tau = 3$ the test displays high rejection probabilities with measurement error but does begin to become oversized, which is unsurprising with a cutoff parameter that is too small. However, even though $L_\tau = 3$ is too small a cutoff, the size is only distorted 5% above its nominal level of 10%. For cutoff parameters between 3 and 5 the performance of the test is even better than the results in Table 8. The best performance obtains with $L_\tau = 5$, when the test is properly sized and rejects the false null in 90% of the cases for error levels 3 and above, for either ρ value. We are particularly encouraged however, that even without an ideal choice of cutoff this test seems to be performing well.

Comparison Between $\hat{V}_{NP}$ and its Unbiased Analog on the Plane

As expected, the unbiased analog of $\hat{V}_{NP}$ has a smaller bias (in absolute terms) than $\hat{V}_{NP}$, with a bias ranging from 35% (for $L_\tau = 5$) to 85% (for $L_\tau = 3$) of the bias of $\hat{V}_{NP}$. However, in terms of RMSE, coverage probabilities, and parametric bootstrap test power, the two estimators present fairly similar properties, across all levels of location errors and both values of ρ.

5   Conclusions and Future Research

This paper has reported results of a Monte Carlo experiment assessing the effects of errors in locations or distances upon the accuracy of parametric and nonparametric estimators of asymptotic variances with spatially dependent data. We studied the finite sample properties of four estimators, an MLE, an MM, the nonparametric estimator suggested by Conley (1999), and its unbiased analog. Our results suggest that MLE estimators perform poorly when the locations of the observations are not perfectly measured. Not only does the bias and root mean squared error of $\hat{V}_{MLE}$ increase rapidly as location errors are introduced, but the coverage probabilities of its associated confidence intervals decline substantially from their nominal level. When $\rho = .3$, the MM estimator also suffers from the presence of location errors, presenting an increase in its bias and RMSE for location errors of level 3 and above, and a substantial decrease in the coverage probabilities of the associated confidence intervals for location errors of level 5 and above. However, when the amount of spatial correlation in the residuals is moderate ($\rho = .45$), the performance of the MM estimator is surprisingly good in terms of bias, root mean squared error, and coverage probability of the associated confidence intervals. In contrast, the nonparametric estimator (and

its unbiased analog) are very robust to the presence of location errors of all levels.

Our specification test based on the parametric bootstrap has displayed good power properties for the types of measurement error we consider. Since, by its nature, this test uses critical values that are specific to the true DGP (under the null), we have every reason to expect that its good power properties will generalize to other DGPs. In contrast, the $t-$tests based on the asymptotic distribution appeared to display desirable power properties only when $\hat{V}_{MLE}$ is used in the denominator, likely because of its severe downward bias. We have less confidence in the generalizability of this result to other DGPs, because other DGPs might have biases of different magnitude or even sign. We think our results strongly suggest that the parametric bootstrap test should be routinely used in conjunction with MLE inference, given the potentially poor performance of such inference in the presence of location errors.

We think there are several interesting directions for future research. A better understanding of the effect of endogeneity in locations/distance errors would be very useful given its relevance for many applications. Additionally, we would like to extend our analysis to the case in which the locations are correctly but imprecisely measured. In particular, it is common for information about agents' physical locations to be known only within an area– census tract, zip code, county, or SMSA. This type of imperfect distance information will pose less of a problem for parametric estimators than the case of errors in location/distances, provided the appropriate calculations are done to infer the properties of the aggregated process from that assumed for individuals. Yet, spatial aggregation will undoubtedly reduce the available information and thus the precision of parametric estimators. Of course, a reduction in precision will occur for nonparametric estimators as well, though it may be less severe for estimators that require only broad definitions of near and far sets of observations in order to define a weighting kernel.

References

Anderson T. W. (1994), *The Statistical Analysis of Time Series*, John Wiley and Sons.

Anselin L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers.

Anselin L. and R. J. G. M. Florax eds (1995), *New Directions in Spatial Econometrics*, Springer.

Anselin L. and Griffith D. A. (1988), "Do Spatial Effects Really Matter in Regression Analysis?", *Papers of the Regional Science Association*, 65, 11-34.

Arbia G., D. A. Griffith and R. Haining (1998), "Error Propagation modelling in raster GIS: Overlay operations," *Internat. J. Geographical Informational Sci.,* Vol. 12, 145-167.

Baltagi B., S. H. Song, and W. Koh (2003), "Testing Panel Data Regression Models with Spatial Error Correlation," *Journal of Econometrics*, Vol. 117, Issue 1, 123-150.

Bartlett M. S. (1950), "Periodogram Analysis and Continuous Spectra," *Biometrika*, Vol. 37, 1-16.

Bartlett M. S. (1955), *An Introduction to Stochastic Processes*. Cambridge University Press, Cambridge.

Bell K. P. and N. E. Bockstael (2000), "Applying the Generalized Moments Estimation Approach to Spatial Problems Involving Micro Level Data," *Review of Economics and Statistics,*" Vol. 82, 72-82.

Bollinger C. and K. Ihlanfeldt (1997), "The Impact of Rapid Rail Transit on Economic Development: The Case of Atlanta's Marta," *Journal of Urban Economics*, Vol. 42, 179-204.

Bolthausen E. (1982), "On the Central Limit Theorem for Stationary Mixing Random Fields," *The Annals of Probability*, Vol. 10, 1047-1050.

Bookstein F. L. (1986), "Size and Shape Spaces for Landmark Data in Two Dimensions," *Statistical Science*, Vol. 1, 191-222.

Bronnenberg B., J. P. Dube, S. Dhar (2003), "Analyzing the Spatial Distribution of Market Shares and Prices across Geographic Markets" Manuscript University of Chicago GSB.

Case A. (1991), "Spatial Patterns in Household Demand", *Econometrica*, Vol. 59, No. 4., 953-965.

Case A., Rosen H. S., and Hines J. R. (1993), "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States", *Journal of Public Economics*, Vol. 52, No. 3, 285-307.

Conley T. G. (1995), *Econometric Modelling of Cross-Sectional Dependence*, Ph. D. thesis, University of Chicago.

Conley T. G. (1999), "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, Vol. 92, 1-45.

Conley, T. G. and B. Dupor (2003), "A Spatial Analysis of Sectoral Complementarity," *Journal of Political Economy*, Vol. 111 No. 2, 311-352.

Conley T. G., F. Flyer and G. Tsiang (2003), "Spillovers from Local Market Human Capital and the Spatial Distribution of Productivity in Malaysia," *Advances in Economic Analysis & Policy*, Vol. 3, No. 1, Article 5. http://www.bepress.com/bejeap/advances/vol3/iss1/art5.

Conley T. G., Hansen L. P., and W. F. Liu (1997), "Bootstrapping the Long Run," *Macroeconomic Dynamics*, Vol. 1, 279-311.

Conley T. G. and E. A. Ligon (2002), "Economic Distance, Spillovers, and Cross Country Comparisons," *Journal of Economic Growth*, Vol. 7, 157-187.

Conley T. G. and G. Topa (2002), "Socio-economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, Vol 17, Issue 4, 303-327.

Cressie N. (1993), *Statistics for Spatial Data*, Wiley.

Cressie N. and J. Kornak (2003), "Spatial Statistics in the Presence of Location Error with an Application to Remote Sensing of the Environment," *Statistical Science*, Vol. 18, No. 4, 436–456.

Cressie N. and S. Lahiri (1996), "Asymptotics for REML Estimation of Spatial Covariance Parameters," *Journal of Statistical Planning and Inference*, Vol. 50, 327-342.

Elliott G.(1993), "Spatial Correlations and Cross-Country Regressions," Manuscript, Harvard University.

Gabrosek J. and N. Cressie (2002), "The Effect on Attribute Prediction of Location Uncertainty in Spatial Data," *Geographical Analysis*, Vol. 34, 262-285.

Goldie C. M. and G. J. Morrow (1986), "Central Limit Questions for Random Fields," in *Dependence in Probability and Statistics*, E. Eberlein and M. S. Taqqu eds. Birkhauser Boston, 275-289.

Goldie C. M. and P. E. Greenwood (1986), "Characterizations of Set-Indexed Brownian Motion and Associated Conditions for Finite-Dimensional Convergence," *The Annals of Probability*, Vol. 11, No. 3, 802-816.

Grenander U. and Rosenblatt M. (1957), "Some Problems in Estimating the Spectrum of a Time Series," *Proc. 3rd Berkeley Symp. on Statist. and Prob.*, University of California Press, Berkeley.

Greenstone M. and O. Deschenes (2003), "Using Fluctuations in Climate to Estimate the Economic Impacts of Global Warming", University of Chicago Working Paper.

Griffith D. A. and F. Lagona (1998), "On the Quality of Likelihood Based Estimators in Spatial Autoregressive Models when the Data Dependence Structure is Misspecified," *Journal of Statistical Planning and Inference*, Vol. 69, 153-174.

Guyon X. (1982), "Parameter Estimation for a Stationary Process on a $d-$ Dimensional Lattice," *Biometrika*, Vol. 69, 95-105.

Hall P., N. I. Fisher, and B. Hoffmann (1994), "On the Nonparametric Estimation of Covariance Functions," *Annals of Statistics*, Vol. 22, No. 4., 2115-2134.

Hall P. and P. Patil (1994), "Properties of Nonparametric Estimators of Autocovariance for Stationary Random Fields," *Probability Theory and Related Fields*, Vol. 99, No. 3, 399-424.

Heijmans R. and J. Magnus (1986), "Consistent Maximum Likelihood Estimation with Dependent Observations. The General (Nonnormal) Case and the Normal Case," *Journal of Econometrics*, Vol. 32, 253-285.

Ibragimov I. and Y. Linnik (1971), *Independent and Stationary Sequences of Random Variables*, Walters-Noordhoff, Groningen, The Netherlands.

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

Kelejian H. H. and I. R. Prucha (2001), "On the Asymptotic Distribution of the Moran I Test Statistic with Applications," *Journal of Econometrics*, Vol. 104, 219-257.

Kelejian H. H, and I. R. Prucha (2003), "HAC Estimation in a Spatial Framework," University of Maryland Working Paper.

Kelejian H. H. and D. Robinson (1992), "Spatial Autocorrelation: A New Computationally Simple Test with an Application to Per Capita Police Expenditures," *Regional Science and Urban Economics*, Vol. 22, 317-331.

Kent J. T. and K. V. Mardia (1996), "Spectral and Circulant Approximations to the Likelihood for Stationary Gaussian Random Fields," *Journal of Statistical Planning and Inference*, Vol. 50, 379-394.

Kim C. W., T. T. Phipps and L. Anselin (2003), "Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach," *Journal of Environmental Economics and Management*, Vol. 45, Issue 1, 24-39.

Lee L. F. (2001a), "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models I: Spatial Autoregressive Processes", Ohio State Department of Economics Working Paper.

Lee L. F. (2001b), "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models II: Mixed Regressive, Spatial Autoregressive Models" Ohio State Department of Economics Working Paper.

Lee L. F. (2004), "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica*, Vol. 72, No. 6, 1899-1925.

Mardia K. V. and R. J. Marshall (1984), "Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression," *Biometrika* Vol. 71, No. 1, 135-146.

Manuszak M. (2001), "The Impact of Upstream Mergers on Retail Gasoline Markets" Working Paper, Carnegie Mellon GSIA.

Moreno R. and B. Trehan (1997), "Location and the Growth of Nations", *Journal of Economic Growth*, Vol. 2, No. 4, 399-418.

Moran P. (1950), "Notes on Continuous Stochastic Phenomena," *Biometrika*, Vol. 37, 17–23.

Moulton B. R. (1990), "An Illustration of Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", *The Review of Economics and Statistics*, Vol. 72, 334-338.

Oehlert G. (1993), "Regional Trends in Sulfate Wet Deposition," *Journal of the American Statistical Association*, Vol. 88, No. 422, 390-399.

Ord K. (1975), "Estimation Methods for Models of Spatial Interaction", *Journal of the American Statistical Association*, Vol. 70, No. 349, 120-26.

Pinkse J. (1999), "Moran-Flavoured Tests with Nuisance Parameters: Examples," forthcoming in *New Advances in Spatial Econometrics*, L. Anselin and R. J. G. M. Florax eds. (Springer, New York)

Pinkse J., M. Slade, and C. Brett C. (2002), "Spatial Price Competition: a Semiparametric Approach," *Econometrica*, Vol. 70, No. 3, 1111-1153.

Politis N. D. and J. P. Romano (1996), "On Flat-Top Kernel Spectral Density Estimators for Homogeneous Random Fields," *Journal of Statistical Planning and Inference*, Vol. 51, 41-53.

Pulvino T. (1998),"Do Asset Fire Sales Exist. An Empirical Investigation of Commercial Aircraft Transactions," *Journal of Finance*, Vol. 53, No. 3, 939-978.

Priestly M. B. (1962), "Analysis of Two-Dimensional Processes with Discontinuous Spectra," Technical Report No. 6, Department of Statistics, Stanford University.

Priestley M. B. (1981), *Spectral Analysis and Time Series Vols. I and II*, Academic Press, New York.

Rappaport J. and Sachs J. D. (2003), "The United States as a Coastal Nation," *Journal of Economic Growth*, Vol. 8, No. 1, 5-46.

Robinson P. M. and J Vidal Sanz (2003), "Modified Whittle Estimation of Multilateral Spatial Models", *CeMMAP Working paper #CWP1803*.

Rosenblatt M. (1985), *Stationary Sequences and Random Fields*, Birkhäuser Boston, Inc.

Takahata H. (1983), "On the Rates in the Central Limit Theorem for Weakly Dependent Random Fields" *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, Vol. 64, 445-456.

Topa G. (2001), "Social Interactions, Local Spillovers and Unemployment", *Review of Economic Studies*, Vol. 68, No. 2, 261-95.

Vigfusson R.J. (2003), "How Does the Border Effect Productivity? Evidence from American and Canadian Manufacturing Industries" Working Paper. Federal Reserve Board of Governors.

White H. and I. Domowitz (1984), "Nonlinear Regression with Dependent Observations", *Econometrica*, Vol. 52, 143-161.

Whittle, P. (1954), "On Stationary Processes on the Plane," *Biometrika*, Vol. 2 No. 3/4, 434-449.

Zhao Y. and M. Wall (2004), "Investigating the Use of the Variogram for Lattice Data," *Journal of Computational and Graphical Statistics*, Vol. 13, No. 3, 719-738.

A Bolthausen CLT

Letting $\pi(s_1, s_2)$ denote maximum across coordinates of $|s_1 - s_2|$, define a distance measure between sets $\pi(\Lambda_1, \Lambda_2) = \inf\{\pi(s_1, s_2) : s_1 \in \Lambda_1, s_2 \in \Lambda_2\}$. For a mean zero stationary random vector X_s $s \in Z^d$, let F_Λ denote the sigma algebra generated by X_s, $s \in \Lambda$, $\Lambda \subset Z^d$. Define mixing coefficients as:

$$\alpha_{k,l}(n) = \sup\{|P(a_1 \cap a_2) - P(a_1)P(a_2)| : a_1 \in F_{\Lambda_1}, a_2 \in F_{\Lambda_2}, |\Lambda_1| \leq k, |\Lambda_2| \leq l, \pi(\Lambda_1, \Lambda_2) \geq n\}$$

$$\rho(n) = \sup\{|cov(b_1, b_2)| : b_1 \in L_2(F_{\Lambda_1}), b_2 \in L_2(F_{\Lambda_2}), \|b_1\|_2 \leq 1, \|b_2\|_2 \leq 1, \pi(\Lambda_1, \Lambda_2) \geq n\}$$

Theorem 4 *(Bolthausen 1982)*
If $\sum_{m=1}^{\infty} m^{d-1} \alpha_{k,l}(m) < \infty, k + l \leq 4$, $\alpha_{1,\infty}(m) = o(m^{-d})$ and if

$$\sum_{m=1}^{\infty} m^{d-1} \rho(m) < \infty$$

or

$$\text{for some } \delta > 0, \|X_{\mathbf{s}}\|_{2+\delta} < \infty \text{ and } \sum_{m=1}^{\infty} m^{d-1} \alpha_{1,1}(m)^{\delta/(2+\delta)} < \infty,$$

then $\sum_{\mathbf{s} \in Z^d} |cov(X_{\mathbf{0}}, X_{\mathbf{s}})| < \infty$. If additionally $\sigma^2 = \sum_{\mathbf{s} \in Z^d} cov(X_{\mathbf{0}}, X_{\mathbf{s}}) > 0$, and Λ_τ is a fixed sequence of finite subsets of Z^d that increases to Z^d and is such that

$$\lim_{\tau \to \infty} |boundary(\Lambda_\tau)|/|\Lambda_\tau| = 0,$$

Then

$$\frac{1}{\sigma |\Lambda_\tau|^{1/2}} \sum_{\mathbf{s} \in \Lambda_\tau} X_{\mathbf{s}} \Rightarrow N(0, 1)$$

B Proofs of Propositions

B.1 Proof of Proposition 1

The proposition using conditions (a) and (b-i) is Proposition 5 in Conley (1999). Here we provide a proof using (b-ii).

The strategy for proving consistency in the presence of bounded measurement errors in location can be cast in terms of showing that $\hat{V}_{NP}$, obtained using the uniform kernel with cutoff L and mismeasured locations, is asymptotically equivalent to an infeasible estimator that uses true locations and a smaller cutoff point.

It will be convenient in this proof to explicitly refer to each coordinate of $\mathbf{s} = (m, n)$, let the sample region Λ be an M by N rectangle, suppressing the index τ. Let the bound on measurement error in each dimension be denoted B so that for each point $|m^{true} - m^{measured}| < B$ and $|n^{ntrue} - n^{measured}| < B$. We index points throughout this proof with their true indexes. The kernel weight for the product of points (m, n) and $(m + j, n + k)$ is denoted $\tilde{K}_{MN}(m, n, j, k)$. These weights will be zero and one, but depend on the measurement errors at both locations (m, n) and $(m + j, n + k)$.

$$\hat{V}_{NP} = \frac{2}{MN} \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} \sum_{m=j+1}^{M} \sum_{n=k+1}^{N} \tilde{K}_{MN}(m, n, j, k) X_{m,n} X_{m-j,n-k}$$
$$- \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X_{m,n}^2$$

Define $\tilde{V}$ as the infeasible, consistent estimator with displacements that are small enough that they still get weight one:

$$\tilde{V} = \frac{2}{MN} \sum_{j=0}^{L_M-2B} \sum_{k=0}^{L_N-2B} \sum_{m=j+1}^{M} \sum_{n=k+1}^{N} X_{m,n} X_{m-j,n-k}$$
$$- \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X_{m,n}^2$$

Conley (1999), Proposition 3 directly implies that $\tilde{V} \to V$ in probability. Therefore, it suffices to show here that the difference R between $\hat{V}_{NP}$ and $\tilde{V}$ vanishes. Let $R \equiv \hat{V}_{NP} - \tilde{V}$:

$$R = \frac{2}{MN} \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} \sum_{m=j+1}^{M} \sum_{n=k+1}^{N} [1 - 1(j < L_M - 2B)1(k < L_N - 2B)]\tilde{K}_{MN}(m, n, j, k) X_{m,n} X_{m-j,n-k}$$

The result follows from a demonstration that $R \to 0$ in mean square. $ER = 0$ for L_N, L_M large enough since $X_{m,n}$ is a finite-order moving average, so showing $var(R) \to 0$ is sufficient. We first show $E(R - ER)^2 \to 0$ and then $ER \to 0$.

To simplify notation let $X_{m,n} = 0$ for non-positive values of either index. Define an array $Z_{MN,mn}$:

$$Z_{MN,mn} =$$
$$\sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} [1 - 1(j < L_M - 2B)1(k < L_N - 2B)]\tilde{K}_{MN}(m, n, j, k)(X_{m,n} X_{m-j,n-k} - EX_{m,n} X_{m-j,n-k}),$$

so $R - ER = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} Z_{MN,mn}$. Hence,

$$var\{ \sum_{m=1}^{M} \sum_{n=1}^{N} Z_{MN,mn} \} = \left| \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{m'=1}^{M} \sum_{n'=1}^{N} EZ_{MN,mn} Z_{MN,m'n'} \right|.$$

The triangle inequality implies:

$$var\{ \sum_{m=1}^{M} \sum_{n=1}^{N} Z_{MN,mn} \} \leq \left\{ \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{|m-m'| \leq 2(L_M+2B)} \sum_{|n-n'| \leq 2(L_N+2B)} \left| EZ_{MN,mn} Z_{MN,m'n'} \right| + \right.$$
$$\left. \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{m',n':|m-m'|>2(L_M+2B) \text{or} |n-n'|>2(L_N+2B)} \left| EZ_{MN,mn} Z_{MN,m'n'} \right| \right\}, \tag{5}$$

31

having divided the terms into close ones (within $2(L_M + 2B)$ and $2(L_N + 2B)$ in each direction) and far ones (farther than $2(L_M + 2B)$ or $2(L_N + 2B)$). Note that if the sample region were not rectangular, the $EZ_{MN,mn}Z_{MN,m'n'}$ terms could still be divided into close and far groups of terms.

First look at the close terms. No matter what the shape of the sample region, the maximum number of points within $2(L_M + 2B), 2(L_N + 2B)$ in each direction from any point is $(4L_M + 8B + 1)(4L_N + 8B + 1)$. Therefore:

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{|m-m'|<2(L_M+2B)} \sum_{|n-n'|<2(L_N+2B)} \left| EZ_{MN,mn}Z_{MN,m'n'} \right| \leq$$

$$MN(4L_M + 8B + 1)(4L_N + 8B + 1) \sup_{1 \leq m' \leq M, 1 \leq n' \leq N} \left\| Z_{MN,m'n'} \right\|_2^2.$$

The next step is to bound $\sup_{1 \leq m' \leq M, 1 \leq n' \leq N} \left\| Z_{MN,m'n'} \right\|_2^2$. Minkowski's inequality implies:

$$\begin{aligned} \left\| Z_{MN,mn} \right\|_2 &\leq \\ \Big\{ \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} & [1 - 1(j < L_M - 2B)1(k < L_N - 2B)] \cdot \ldots \\ \tilde{K}_{MN}(m,n,j,k) & \left\| X_{m,n}X_{m-j,n-k} - EX_{m,n}X_{m-j,n-k} \right\|_2 \Big\}. \end{aligned} \tag{6}$$

$X_{m,n}$ has finite $(4+\delta)th$ moments which implies that $\sup_{j,k} \left\| X_{m,n}X_{m-j,n-k} - EX_{m,n}X_{m-j,n-k} \right\|_2$ is bounded, $\tilde{K}_{MN}(m,n,j,k)$ are uniformly bounded, and the number of terms where $[1 - 1(j < L_M - 2B)1(k < L_N - 2B)] = 1$ is $[(L_M + 2B + 1)(L_N + 2B + 1) - (L_M - 2B + 1)(L_N - 2B + 1)] = 4B(L_M + L_N + 2)$. Hence

$$\left\| Z_{MN,mn} \right\|_2 \leq c_1(L_M + L_N + 2)$$

for some constant c_1. Thus giving the following bounds:

$$\left\| Z_{MN,mn} \right\|_2^2 \leq c_1^2(L_M + L_N + 2)^2$$

Therefore the near terms satisfy:

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{|m-m'|<2L_M} \sum_{|n-n'|<2L_N} \left| EZ_{MN,mn}Z_{MN,m'n'} \right| \leq c_1^2 MN(4L_M+8B+1)(4L_N+8B+1)(L_M+L_N+2)^2.$$

$$\tag{7}$$

Next consider the far apart terms.

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{m',n':|m-m'|>2L_M \text{ or } |n-n'|>2L_N} \left| EZ_{MN,mn}Z_{MN,m'n'} \right|$$

A mixing inequality from Ibragimov and Linnik (1971) chapter 17 gives a bound on $\left| EZ_{MN,mn}, Z_{MN,m'n'} \right|$:

$$\left| EZ_{MN,mn}, Z_{MN,m'n'} \right| \leq c_2 \alpha_{\infty,\infty}(\min(2(L_M + 2B), 2(L_N + 2B))^{\delta/(2+\delta)} \sup_{m,n} \left\| Z_{MN,mn} \right\|_{2+\delta}^2$$

and an argument identical to that above for (6) implies $\|Z_{MN,mn}\|_{2+\delta}^2 \le c_3^2(L_M + L_N + 2)^2$ for some constants c_2, c_3. Combining these terms give a bound on the far terms of:

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{m',n':|m-m'|>2L_M \text{or}|n-n'|>2L_N} \left| EZ_{MN,mn} Z_{MN,m'n'} \right| \le$$

$$c_4 M^2 N^2 \alpha_{\infty,\infty}(\min(2(L_M + 2B), 2(L_N + 2B))^{\delta/(2+\delta)}(L_M + L_N + 2)^2$$

Combining the bounds on near and far terms yields:

$$var\{\sum_{m=1}^{M} \sum_{n=1}^{N} Z_{MN,mn}\} \le c_1^2 MN(4L_M + 8B + 1)(4L_N + 8B + 1)(L_M + L_N + 2)^2 +$$

$$c_4 M^2 N^2 \alpha_{\infty,\infty}(\min(2(L_M + 2B), 2(L_N + 2B))^{\delta/(2+\delta)}(L_M + L_N + 2)^2 + o(1).$$

The rate conditions on $L_{i,\tau}$ and the mixing condition in parts (a) and (c) imply that the right side of this expression converges to zero as $M, N \to \infty$.

Consider now ER :

$$ER =$$

$$= \frac{2}{MN} \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} \sum_{m=j+1}^{M} \sum_{n=k+1}^{N} [1 - 1(j < L_M - 2B)1(k < L_N - 2B)]\tilde{K}_{MN}(m,n,j,k)EX_{m,n}X_{m-j,n-k}$$

$$= 2 \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} [1 - 1(j < L_M - 2B)1(k < L_N - 2B)]\tilde{K}_{MN}(m,n,j,k)\frac{(M-j)(N-k)}{MN}EX_{m,n}X_{m-j,n-k}$$

Using the same mixing inequality to bound $EX_{m,n}X_{m-j,n-k}$:

$$|ER| \le$$

$$c_5 \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} \left| [1 - 1(j < L_M - 2B)1(k < L_N - 2B)]\tilde{K}_{MN}(m,n,j,k)\frac{(M-j)(N-k)}{MN} \right| \alpha_{\infty,\infty}(\max(j,k))^{\frac{\delta}{(2+\delta)}}$$

$$\le c_5 \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} |[1 - 1(j < L_M - 2B)1(k < L_N - 2B)]| \alpha_{\infty,\infty}(\max(j,k))^{\frac{\delta}{(2+\delta)}}.$$

$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{\infty,\infty}(\max(j,k))^{\delta/(2+\delta)} < \infty$ since $\alpha_{\infty,\infty} = o(m^{-4})$, so the dominated convergence theorem implies

$$c_5 \sum_{j=0}^{L_M+2B} \sum_{k=0}^{L_N+2B} |[1 - 1(j < L_M - 2B)1(k < L_N - 2B)]| \alpha_{\infty,\infty}(\max(j,k))^{\delta/(2+\delta)} \to 0$$

since $|[1 - 1(j < L_M - 2B)1(k < L_N - 2B)]| \to 0$, all j, k.

B.2 Proof of Proposition 2

Notation

We first introduce some notation that will be used in the remaining proofs. Let Λ_τ be the hypercube (in $d-$dimensional Euclidean space) of lattice points $\mathbf{s}$ with all components integers s_i, $1 \le s_i \le N$, so that $N^d = |\Lambda_\tau|$. Given our DGP, $\{X_{\mathbf{s}}\}$ is a random field with $EX_{\mathbf{s}} \equiv 0$ and cumulant functions up to order eight absolutely summable. Denote by

$$f\left(\boldsymbol{\omega}\right) \equiv \frac{1}{\left(2\pi\right)^d} \sum_{s_1=-\infty}^{\infty} \cdots \sum_{s_d=-\infty}^{\infty} r_{\mathbf{s}} e^{-i\mathbf{s}\cdot\boldsymbol{\omega}}$$

the spectral density of $\{X_{\mathbf{s}}\}$, where $r_{\mathbf{s}} \equiv C\left(\mathbf{s}\right) = E\left(X_{\mathbf{s}} X_{\mathbf{u}+\mathbf{s}}\right)$, $\boldsymbol{\omega} = \left(\omega_1, \dots, \omega_d\right) \in [-\pi, \pi]^d$, and $\left(\mathbf{s} \cdot \boldsymbol{\omega}\right) = \sum_{h=1}^{d} s_h \omega_h$ is the inner product in $d-$dimensional Euclidean space. For locations on the plane, $d = 2$. Let Σ denote the variance-covariance matrix of the vector $[X_{\mathbf{s}}, \mathbf{s} \in \Lambda_\tau]$. Let

$$\begin{aligned}
\Sigma_\rho &= \frac{\partial\Sigma}{\partial\rho}, \quad \Sigma_{\sigma^2} = \frac{\partial\Sigma}{\partial\sigma^2}, \\
\vartheta_{ij} &= \text{tr}\left(\Sigma^{-1}\Sigma_i\Sigma^{-1}\Sigma_j\right), \quad i, j = \rho, \sigma^2.
\end{aligned}$$

Let $\theta = \left(\rho, \sigma^2\right)$ denote the true values of the parameters in the DGP. Let $\hat{\theta}_{MLE} = \left(\hat{\rho}_{MLE}, \hat{\sigma}^2_{MLE}\right)$ denote the MLE estimator of θ.

Proof.

Mardia and Marshall (1984, Theorem 3) show that $\hat{\theta}_{MLE}$ is consistent and asymptotically normal, provided that $C\left(\mathbf{k}; \rho, \sigma^2\right)$ and its first and second derivatives are absolutely summable, and that

$$\begin{aligned}
a_{ij} &= \lim \frac{\vartheta_{ij}}{\left(\vartheta_{ii}\vartheta_{jj}\right)^{\frac{1}{2}}} \quad \text{exists,} \quad i, j = \left(\rho, \sigma^2\right), \\
\det\left(A\right) &= \det\left(\begin{bmatrix} a_{\rho\rho} & a_{\rho\sigma^2} \\ a_{\rho\sigma^2} & a_{\sigma^2\sigma^2} \end{bmatrix}\right) \neq 0.
\end{aligned}$$

The covariance functions $C\left(\mathbf{k}; \rho, \sigma^2\right)$ are polynomials in ρ and σ^2 for each $\mathbf{k} \in \Lambda_\tau$, and therefore their derivatives exist and are continuous. Absolute summability is ensured by the fact that the processes we consider are finite order moving averages, and therefore

$$\sum_{\mathbf{k}\in Z^2} \left|C\left(\mathbf{k}; \rho, \sigma^2\right)\right| = \sum_{\mathbf{k}\in Z^2 : \|\mathbf{k}\| \le 3} \left|C\left(\mathbf{k}; \rho, \sigma^2\right)\right| < \infty,$$

with similar considerations for the first and second derivatives of $C\left(\mathbf{k}; \rho, \sigma^2\right)$.

From the above considerations it follows that there exists a positive finite constant η_1 such that

$$\sum_{\mathbf{k}\in Z^2} \left(1 + |k_l|\right)\left|\Delta C\left(\mathbf{k}; \rho, \sigma^2\right)\right| < \eta_1, \quad l = 1, 2$$

where k_l denotes the $l-$th component of $\mathbf{k}$, and Δ denotes either the identity operator 1, one of the first order differential operators $\partial/\partial\theta_i$, or one of the second order operators $\partial^2/\partial\theta_i\partial\theta_j$ $i,j = 1,2$. Hence, $\left|(2\pi)^2 \Delta f(\boldsymbol{\omega})\right| < \eta_1$. Moreover, given our choice of an MA(6) process, and our values of $\rho = 0.3,\ 0.45$, it follows that there exists a positive finite constant η_2 such that:

$$\frac{1}{(2\pi)^2 f(\boldsymbol{\omega})} < \eta_2.$$

This implies that the conditions for Theorem 1 and Lemmas 3.1-3.2, 4.1-4.3 in Kent and Mardia (1996) are satisfied, and therefore

$$\text{tr}\left(\Sigma^{-1}\Sigma_i\Sigma^{-1}\Sigma_j\right) = \frac{|\Lambda_\tau|}{(2\pi)^2}\int \frac{f_i(\boldsymbol{\omega})f_j(\boldsymbol{\omega})}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega} + O\left(\sqrt{|\Lambda_\tau|}\right)$$

where $f_i(\boldsymbol{\omega}) = \frac{\partial f(\boldsymbol{\omega})}{\partial\theta_i}$. Hence, for $i,j = 1,2$

$$\lim_{|\Lambda_\tau|\to\infty} \frac{\vartheta_{ij}}{\sqrt{(\vartheta_{ii}\vartheta_{jj})}} = \lim_{|\Lambda_\tau|\to\infty} \frac{\frac{|\Lambda_\tau|}{(2\pi)^2}\int \frac{f_i(\boldsymbol{\omega})f_j(\boldsymbol{\omega})}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega} + O\left(\sqrt{|\Lambda_\tau|}\right)}{\sqrt{\left(\frac{|\Lambda_\tau|}{(2\pi)^2}\int \frac{f_i(\boldsymbol{\omega})^2}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega} + O\left(\sqrt{|\Lambda_\tau|}\right)\right)\left(\frac{|\Lambda_\tau|}{(2\pi)^2}\int \frac{f_j(\boldsymbol{\omega})^2}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega} + O\left(\sqrt{|\Lambda_\tau|}\right)\right)}}$$

$$= \frac{\int \frac{f_i(\boldsymbol{\omega})f_j(\boldsymbol{\omega})}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega}}{\sqrt{\left(\int \frac{f_i(\boldsymbol{\omega})^2}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega}\right)\left(\int \frac{f_j(\boldsymbol{\omega})^2}{f(\boldsymbol{\omega})^2}d\boldsymbol{\omega}\right)}}.$$

The above limits exist. Given our analytic forms for the spectral densities and its derivatives with respect of ρ and σ^2, direct computations show that $\det(A) \neq 0$.

B.3 Proof of Proposition 3

Notation.

We use the same notation as in the proof of Proposition 2.

Proof.

The result for locations on the line follows trivially from the results of Anderson (1994), Theorem 9.4.1 and Corollary 9.4.1. Here we prove the result for locations on the plane.

1. Asymptotic Distribution of Spectral Density Estimator on the Plane

Let $\hat{r}_{\mathbf{s}}$ denote an estimate of $r_{\mathbf{s}} \equiv C(\mathbf{s}) = E(X_{\mathbf{s}}X_{\mathbf{u}+\mathbf{s}})$ given by

$$\hat{r}_{\mathbf{s}} = \frac{1}{|\Lambda_\tau|}\sum_{\mathbf{u},\mathbf{u}+\mathbf{s}\in\Lambda_\tau} X_{\mathbf{s}}X_{\mathbf{u}+\mathbf{s}},$$

Notice that

$$E(\hat{r}_{\mathbf{s}}) = \frac{\prod\limits_{h=1}^{d}(N - |s_h|)}{|\Lambda_\tau|}r_{\mathbf{s}},$$

An estimate $\hat{f}(\boldsymbol{\omega})$ of $f(\boldsymbol{\omega})$ is then given by

$$\hat{f}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \sum_{|s_1|,\ldots,|s_d| \leq N} K\left(\frac{s_1}{L_1}, \frac{s_2}{L_2}, \ldots, \frac{s_d}{L_d}\right) \hat{r}_{\mathbf{s}} e^{-i\mathbf{s}\cdot\boldsymbol{\omega}},$$

where $K(\mathbf{0}) = 1$, and $K(\mathbf{x})$ is assumed to be an even $(K(\mathbf{x}) = K(-\mathbf{x}))$ function, uniformly bounded and square integrable. Given $N^d = |\Lambda_\tau|$, let $L_i \to \infty$ and $\frac{L_i}{N} \to 0$ as $N \to \infty$, $i = 1, \ldots, d$. Rosenblatt (1985) Theorem 7 p. 157 is as follows:

Theorem 5 (Rosenblatt (1985)) *Let $\{X_{\mathbf{s}}\}$ be a strictly stationary strongly mixing random field with $EX_{\mathbf{s}} \equiv 0$. Assume that the cumulant functions up to eighth order are absolutely summable. Also let the spectral density estimate $\hat{f}(\boldsymbol{\omega})$ have weights $K(\cdot)$ satisfying the condition specified above. It then follows that*

$$\sqrt{\frac{|\Lambda_\tau|}{L^d}} \left[\hat{f}(\boldsymbol{\omega}) - E\left(\hat{f}(\boldsymbol{\omega})\right)\right] \xrightarrow{d} N(0, \Omega),$$

where

$$\Omega = (2\pi)^d \{1 + \eta(2\omega_1)\ldots\eta(2\omega_d)\} f^2(\boldsymbol{\omega}) \int W^2(\boldsymbol{\alpha}) d\boldsymbol{\alpha},$$

$$\eta(\mu) = \begin{cases} 1 & \text{if } \mu = 2m\pi, \ m \text{ integer} \\ 0 & \text{otherwise.} \end{cases}$$

$$W(\boldsymbol{\alpha}) = \frac{1}{(2\pi)^d} \int K(\mathbf{u}) e^{-i\mathbf{u}\cdot\boldsymbol{\alpha}} d\mathbf{u}$$

☐

Hence at frequency zero, under the above assumptions,

$$\sqrt{\frac{|\Lambda_\tau|}{L^d}} \left[\hat{f}(\mathbf{0}) - E\left(\hat{f}(\mathbf{0})\right)\right] \xrightarrow{d} N\left(0, (2\pi)^d 2f^2(\mathbf{0}) \int W^2(\boldsymbol{\alpha}) d\boldsymbol{\alpha}\right).$$

Recall that $V = 2\pi f(0)$, and that we use the uniform kernel in (4). Our DGP satisfies the assumptions of Theorem 5. Additionally,

$$\begin{aligned}
V &= C(0) + 4C(1) + 4C\left(\sqrt{2}\right) + 4C(2) + 8C\left(\sqrt{5}\right) + 4C\left(\sqrt{8}\right) + 4C(3) + \ldots \\
&\quad + 8C\left(\sqrt{10}\right) + 8C\left(\sqrt{13}\right) + 4C(4) + 8C\left(\sqrt{17}\right) + 4C\left(\sqrt{18}\right) + 8C\left(\sqrt{20}\right) + \ldots \\
&\quad + 4C([0,5]) + 8C([3,4]) + 8C\left(\sqrt{26}\right) + 8C\left(\sqrt{29}\right) + 4C\left(\sqrt{32}\right) + 4C(6)
\end{aligned}$$

Therefore,

$$\left(E\left(\hat{V}_{NP}\right) - V\right) = O_p\left(|\Lambda_\tau|^{-1/2}\right)$$

Hence
$$\lim_{N\to\infty}\sqrt{\frac{|\Lambda_\tau|}{L^2}}\left[E\left(f_N\left(\mathbf{0}\right)\right)-f\left(\mathbf{0}\right)\right]=0,$$

from which
$$\sqrt{\frac{|\Lambda_\tau|}{L^2}}\left[\hat{f}\left(\mathbf{0}\right)-f\left(\mathbf{0}\right)\right]\overset{d}{\to}N\left(0,\left(2\pi\right)^2 2f^2\left(\mathbf{0}\right)\int W^2\left(\boldsymbol{\alpha}\right)d\boldsymbol{\alpha}\right).\tag{8}$$

2. Asymptotic Distribution of Specification Test on the Plane

We want to show that
$$\sqrt{\frac{|\Lambda_\tau|}{L^2}}\left(\hat{V}_{NP}-\hat{V}_i\right)\overset{d}{\to}N\left(0,\frac{1}{(2\pi)^2}2V^2\int\left(\int K\left(\mathbf{x}\right)e^{-i\mathbf{x}\cdot\boldsymbol{\alpha}}d\mathbf{x}\right)^2 d\boldsymbol{\alpha}\right),\quad i=MLE,MM.$$

It is then easy to verify that with a uniform kernel $\frac{1}{(2\pi)^2}\int\left(\int K\left(\mathbf{x}\right)e^{-i\mathbf{x}\cdot\boldsymbol{\alpha}}d\mathbf{x}\right)^2 d\boldsymbol{\alpha}=\mathbf{4}$.

Consider first the MLE estimator. The goal is to show that $\sqrt{\frac{|\Lambda_\tau|}{L^2}}\left(\hat{V}_{MLE}-V\right)\overset{p}{\to}0$; then the desired result will follow from (8). As shown in Lemma 2, our model with $\rho=0.3$ and $\rho=0.45$ satisfies the conditions of Theorem 3 of Mardia and Marshall (1984). Hence
$$\sqrt{|\Lambda_\tau|}\left(\hat{\theta}_{MLE}-\theta\right)\overset{d}{\to}\left(0,H\right),$$

where H is the variance-covariance matrix of θ. Since $\hat{V}_{MLE}$ is given by the product of $\hat{\sigma}_{MLE}^2$ and a polynomial in $\hat{\rho}_{MLE}$, the desired result follows.

Consider now the MM estimator. Since our MM estimator uses unbiased covariances, Guyon's (1982) results ensure that $\sqrt{\frac{|\Lambda_\tau|}{L^2}}\left(\hat{\theta}_{MM}-\theta\right)\overset{p}{\to}0$, from which the result follows.

C Analytic Expressions for the Covariance Function and the Asymptotic Variance on the Plane

The DGP for $X_{\mathbf{s}}$ is
$$X_{\mathbf{s}}=\sum_{\mathbf{r}:\|\mathbf{s}-\mathbf{r}\|\leq 3}\rho^{\|\mathbf{s}-\mathbf{r}\|}u_{\mathbf{s}-\mathbf{r}}$$

where $u_{\mathbf{s}}$ is IID $N(0,\sigma^2)$. One can verify that:

$$\begin{aligned}V\;=\;&C\left(0\right)+4C\left(1\right)+4C\left(\sqrt{2}\right)+4C\left(2\right)+8C\left(\sqrt{5}\right)+4C\left(\sqrt{8}\right)+4C\left(3\right)+...\\&+8C\left(\sqrt{10}\right)+8C\left(\sqrt{13}\right)+4C\left(4\right)+8C\left(\sqrt{17}\right)+4C\left(\sqrt{18}\right)+8C\left(\sqrt{20}\right)+...\\&+4C\left(\left[0,5\right]\right)+8C\left(\left[3,4\right]\right)+8C\left(\sqrt{26}\right)+8C\left(\sqrt{29}\right)+4C\left(\sqrt{32}\right)+4C\left(6\right),\end{aligned}$$

and
$$C\left(0\right)=\sigma^2\left(1+4\rho^2+4\rho^{2\sqrt{2}}+4\rho^4+8\rho^{2\sqrt{5}}+4\rho^{2\sqrt{8}}+4\rho^6\right)$$

$$C(1) = \sigma^2 \left(2\rho + 4\rho^{1+\sqrt{2}} + 2\rho^3 + 4\rho^{\sqrt{2}+\sqrt{5}} + 4\rho^{2+\sqrt{5}} + 4\rho^{\sqrt{5}+\sqrt{8}} + 2\rho^5 \right)$$

$$C\left(\sqrt{2}\right) = \sigma^2 \left(2\rho^{\sqrt{2}} + 4\rho^{1+\sqrt{5}} + 2\rho^2 + 4\rho^{2+\sqrt{2}} + 2\rho^{\sqrt{2}+\sqrt{8}} + 4\rho^{3+\sqrt{5}} + 2\rho^{2\sqrt{5}} \right)$$

$$C(2) = \sigma^2 \left(3\rho^2 + 4\rho^{1+\sqrt{5}} + 2\rho^4 + 4\rho^{2+\sqrt{8}} + 2\rho^{2\sqrt{2}} + 2\rho^{2\sqrt{5}} \right)$$

$$C\left(\sqrt{5}\right) = \sigma^2 \left(2\rho^{\sqrt{5}} + 2\rho^{1+\sqrt{2}} + 2\rho^3 + 2\rho^{1+\sqrt{8}} + 2\rho^{\sqrt{2}+\sqrt{5}} + 2\rho^{2+\sqrt{5}} + 2\rho^{3+\sqrt{2}} + 2\rho^{3+\sqrt{8}} \right)$$

$$C\left(\sqrt{8}\right) = \sigma^2 \left(4\rho^{1+\sqrt{5}} + 2\rho^{\sqrt{8}} + 2\rho^4 + \rho^{2\sqrt{2}} + 4\rho^{3+\sqrt{5}} \right)$$

$$C(3) = \sigma^2 \left(4\rho^3 + 4\rho^{\sqrt{5}+\sqrt{2}} + 4\rho^{\sqrt{5}+\sqrt{8}} \right)$$

$$C\left(\sqrt{10}\right) = \sigma^2 \left(2\rho^{1+\sqrt{5}} + 2\rho^{2+\sqrt{2}} + 2\rho^4 + 2\rho^{\sqrt{8}+\sqrt{2}} + 2\rho^{2\sqrt{5}} \right)$$

$$C\left(\sqrt{13}\right) = \sigma^2 \left(2\rho^{1+\sqrt{8}} + 2\rho^{\sqrt{2}+\sqrt{5}} + 2\rho^{2+\sqrt{5}} + 2\rho^5 \right)$$

$$C(4) = \sigma^2 \left(3\rho^4 + 2\rho^{2\sqrt{5}} + 2\rho^{2\sqrt{8}} \right)$$

$$C\left(\sqrt{17}\right) = \sigma^2 \left(2\rho^{\sqrt{8}+\sqrt{5}} + 2\rho^{2+\sqrt{5}} + 2\rho^{3+\sqrt{2}} \right)$$

$$C\left(\sqrt{18}\right) = \sigma^2 \left(2\rho^6 + 2\rho^{2\sqrt{5}} + 2\rho^{\sqrt{8}+\sqrt{2}} \right)$$

$$C\left(\sqrt{20}\right) = \sigma^2 \left(2\rho^{2+\sqrt{8}} + 2\rho^{3+\sqrt{5}} + \rho^{2\sqrt{5}} \right)$$

$$C([0 \ \ 5]) = 2\sigma^2 \rho^5$$

$$C([3 \ \ 4]) = 2\sigma^2 \rho^{\sqrt{5}+\sqrt{8}}$$

$$C\left(\sqrt{26}\right) = 2\sigma^2 \rho^{3+\sqrt{5}}$$

$$C\left(\sqrt{29}\right) = 2\sigma^2 \rho^{3+\sqrt{8}}$$

$$C\left(\sqrt{32}\right) = \sigma^2 \rho^{2\sqrt{8}}$$

$$C(6) = \sigma^2 \rho^6$$

D Covariance function of the Mismeasured Process on the Line, for Location Errors of Level 1

When locations are measured with error, the moments of the mismeasured process are given by a convex combination of the moments of the correctly measured process. Since the process we consider does not have a trend and is mean zero, this implies that the mean of the mismeasured process is also equal to zero. Hence, our interest centers on the covariance function of the mismeasured process. While the calculations for the general case are very tedious and not particularly enlightening, in this Section we show what this function is for processes on the line and location errors of level 1.

Recall our algorithm for the measurement error process: Each agent's integer location is independently perturbed by adding a random amount ξ from a uniform distribution on [-v,v] . In other words agent i is given a perturbed location $\tilde{s}_i = s_i + \xi_i$. Then, each agent's measured location (z_i) is defined by a re-labeling of the perturbed locations $\{\tilde{s}_i\}$ from 1 to $|\Lambda_\tau|$, according to the rank order of the $\{\tilde{s}_i\}$ from smallest to largest.

Denote by X_s the process observed at correct locations, and by $C(j)$ its covariance function at lag j; denote by X_z the process observed at mismeasured locations, and by $\tilde{C}(k)$ its covariance function at lag k. Observe that with errors of level 1, i.e. v = 1, each observation can be reshuffled by at most one unit; this implies that $\tilde{C}(|k|) \neq 0$ for $|k| = 1, \ldots, 8$ (of course, $\tilde{C}(0) = C(0) > 0$). Moreover, for each lag k the following relation holds:

$$\tilde{C}(|k|) = \begin{cases} \sum_{j=1}^{6} C(|j|) \Pr\left(|z_i - z_j| = k \mid |s_i - s_j| = j\right) & \text{for } |k| = 1, \ldots, 8, \\ 0 & \text{for } |k| > 8. \end{cases}$$

The main task is to calculate the misclassification probabilities $\Pr\left(|z_i - z_j| = k \mid |s_i - s_j| = j\right)$ for each value of k and j. This task is greatly simplified by the fact that only reshuffling by one unit is possible. The reshuffling of one unit will occur with probability

$$p \equiv \Pr\left(\tilde{s}_i < \tilde{s}_{i-1} \mid \{s_i\}\right) = \Pr\left(\xi_i - \xi_{i-1} < -1 \mid \{s_i\}\right) = \frac{1}{8},$$

where the last equality follows from our assumptions about ξ.

Consider $\Pr\left(|z_i - z_j| = 1 \mid |s_i - s_j| = 1\right)$. Given that each observation can be reshuffled of only one unit, two observations that were originally adjacent will remain adjacent unless: (a) only one of them is

reshuffled, say only z_i, or (b) z_i is not reshuffled, but z_j is. In formulas,

$$
\begin{aligned}
&\Pr\left(|z_i - z_j| = 1 \mid |s_i - s_j| = 1\right) \\
&= 1 - \Pr\left(|z_i - z_j| \neq 1 \mid |s_i - s_j| = 1\right) \\
&= 1 - \Pr\left(z_i = s_i - 1, z_j = s_j \mid |s_i - s_j| = 1\right) - \ldots \\
&\quad \left(1 - \Pr\left(z_i = s_i - 1, z_j = s_j \mid |s_i - s_j| = 1\right)\right)\Pr\left(z_i = s_i, z_j = s_j + 1 \mid |s_i - s_j| = 1\right) \\
&= 1 - p - (1 - p)\,p = (1 - p)^2.
\end{aligned}
$$

By repeating a similar reasoning, one can verify that

$$
\Pr\left(|z_i - z_j| = k \mid |s_i - s_j| = j\right) = \begin{cases}
(1 - p)^2 & \text{for } \{k, j\} = \{1, 1\}, \\
2p(1 - p) & \text{for } \{k, j\} = \{(1, 2), (2, 1)\}, \\
(1 - p)(2p^2 - 3p + 1) & \text{for } \{k, j\} = \{2, 2\}, \\
1 - 4p(1 - p)^2 - 2p^2 & \text{for } k = j, \ j \geq 3, \\
2p(1 - p)^2 & \text{for } \{k, j\} = \{3, 2\}, \\
2p(1 - p)^2 & \text{for } k = j \pm 1, \ j \geq 3, \\
p^2 & \text{for } \{k, j\} = (\{3, 1\}, \{4, 2\}) \\
p^2 & \text{for } k = j \pm 2, \ j \geq 3.
\end{cases}
$$

Replacing the value of p, we get

$$
\begin{bmatrix}
\tilde{C}(1) \\
\tilde{C}(2) \\
\tilde{C}(3) \\
\tilde{C}(4) \\
\tilde{C}(5) \\
\tilde{C}(6) \\
\tilde{C}(7) \\
\tilde{C}(8)
\end{bmatrix}
=
\begin{bmatrix}
0.7656 & 0.2188 & 0.0156 & 0 & 0 & 0 \\
0.2188 & 0.5742 & 0.1914 & 0.0156 & 0 & 0 \\
0.0156 & 0.1914 & 0.6016 & 0.1914 & 0.0156 & 0 \\
0 & 0.0156 & 0.1914 & 0.6016 & 0.1914 & 0.0156 \\
0 & 0 & 0.0156 & 0.1914 & 0.6016 & 0.1914 \\
0 & 0 & 0 & 0.0156 & 0.1914 & 0.6016 \\
0 & 0 & 0 & 0 & 0.0156 & 0.1914 \\
0 & 0 & 0 & 0 & 0 & 0.0156
\end{bmatrix}
\begin{bmatrix}
C(1) \\
C(2) \\
C(3) \\
C(4) \\
C(5) \\
C(6)
\end{bmatrix},
$$

where the exact formulas for $C(j)$, $j = 1, \ldots, 6$ were provided in equation (1).

Table 1: Degree of Deviations of Measured from True Locations

	v	Percentage at True Location	1 Unit Off (%)	2 Units Off (%)	3 Units Off (%)	4 Units Off (%)	5 Units Off (%)	6 Units Off (%)
Level 1 Errors	1.0	75.0%	25.0%	0%	0%	0%	0%	0%
Level 2 Errors	1.5	51.8%	42.0%	6.2%	0%	0%	0%	0%
Level 3 Errors	2.0	36.8%	45.5%	15.9%	1.8%	0%	0%	0%
Level 4 Errors	2.5	27.5%	42.7%	23.1%	6.2%	0.5%	0%	0%
Level 5 Errors	3.0	21.5%	37.9%	26.5%	11.5%	2.4%	0.2%	0%
Level 6 Errors	3.5	17.6%	33.0%	26.9%	15.9%	5.6%	1.0%	0%
Level 7 Errors	4.0	14.7%	28.9%	25.7%	18.5%	9.1%	2.7%	0.4%

Table 2: Distance by Which Measured Locations Deviate from True Locations on the Plane

	v	Percentage at True Location	Dist. $\in [1,2)$ (%)	Dist. $\in [2,3)$ (%)	Dist. $\in [3,4)$ (%)	Dist. ≥ 4 (%)
Level 1 Errors	0.75	80.0%	20.0%	0%	0%	0%
Level 2 Errors	1.00	56.7%	43.3%	0%	0%	0%
Level 3 Errors	1.25	40.2%	55.9%	3.9%	0%	0%
Level 4 Errors	1.50	28.4%	60.0%	11.6%	0%	0%
Level 5 Errors	1.75	20.8%	58.8%	19.5%	0.1%	0%
Level 6 Errors	2.00	14.5%	55.2%	27.0%	3.3%	0%
Level 7 Errors	2.25	11.0%	48.5%	33.2%	6.8%	0.5%

Table 3: Bias, Root MSE and 95% CI Coverage Probabilities for V Estimators with True and Error-Ridden Locations,
On the Line, rho = 0.3, sigma = 0.5453

	Bias				Root MSE				95% CI Coverage Probability			
	MLE	MM	NP	NP-Unbiased	MLE	MM	NP	NP-Unbiased	MLE	MM	NP	NP-Unbiased
True Locations	0.000	-0.003	-0.030	-0.028	0.133	0.164	0.259	0.261	0.965	0.964	0.952	0.952
Level 1 Errors	-0.181	-0.030	-0.031	-0.029	0.214	0.177	0.258	0.260	0.938	0.959	0.953	0.953
Level 2 Errors	-0.292	-0.081	-0.032	-0.029	0.308	0.196	0.257	0.258	0.913	0.953	0.954	0.954
Level 3 Errors	-0.365	-0.150	-0.031	-0.030	0.376	0.240	0.254	0.256	0.893	0.943	0.954	0.952
Level 4 Errors	-0.412	-0.218	-0.031	-0.028	0.420	0.294	0.250	0.251	0.880	0.929	0.949	0.952
Level 5 Errors	-0.447	-0.281	-0.031	-0.026	0.454	0.350	0.248	0.252	0.868	0.910	0.952	0.953
Level 6 Errors	-0.473	-0.330	-0.037	-0.032	0.480	0.402	0.246	0.247	0.858	0.893	0.954	0.951
Level 7 Errors	-0.493	-0.368	-0.044	-0.037	0.498	0.438	0.246	0.245	0.848	0.880	0.954	0.953

Table notes: sample size = 500, true value of V = 1. 1000 Monte Carlo replications.

Table 4: Rejection Probabilities for Tests at the 10% Level, On the Line, rho = 0.3, sigma = 0.5453

	t-test Using the Asymptotic Distribution								Parametric Bootstrap, Using MLE	
	NP				NP-Unbiased					
	MLE [a]	MLE [b]	MM [c]	MM [d]	MLE [a]	MLE [b']	MM [c]	MM [d']	NP	NP-Unbiased
True Locations	0.05	0.10	0.01	0.06	0.05	0.11	0.01	0.06	0.07	0.07
Level 1 Errors	0.18	0.04	0.01	0.04	0.18	0.04	0.01	0.04	0.16	0.17
Level 2 Errors	0.43	0.09	0.01	0.02	0.44	0.11	0.01	0.03	0.41	0.42
Level 3 Errors	0.65	0.25	0.05	0.01	0.64	0.25	0.05	0.01	0.62	0.63
Level 4 Errors	0.76	0.42	0.17	0.01	0.75	0.43	0.18	0.00	0.74	0.73
Level 5 Errors	0.86	0.53	0.40	0.05	0.86	0.54	0.41	0.05	0.84	0.84
Level 6 Errors	0.90	0.63	0.61	0.14	0.90	0.64	0.61	0.16	0.89	0.89
Level 7 Errors	0.93	0.70	0.76	0.23	0.93	0.70	0.76	0.28	0.92	0.92

Table notes: sample size = 500, true value of V = 1. 1000 Monte Carlo replications, 200 Bootstrap repetitions. (a) Denominator in the t-test given by the MLE estimator; (b) Denominator in the t-test given by the nonparametric estimator; (c) Denominator in the t-test given by the MM estimator; (d) Denominator in the t-test given by the nonparametric estimator. (b') and (d') Denominator in the t-test given by the unbiased nonparametric estimator.

Table 5a: Bias for V Estimators with True and Error-Ridden Locations,
On the Plane, rho = 0.3, sigma = 0.2456

			Bias					
			NP			NP-Unbiased		
	MLE	MM	L = 3	L = 5	L = 7	L = 3	L=5	L = 7
True Locations	-0.005	-0.015	-0.109	-0.112	-0.169	-0.061	-0.066	-0.130
Level 1 Errors	-0.329	-0.055	-0.119	-0.112	-0.169	-0.069	-0.064	-0.127
Level 2 Errors	-0.502	-0.103	-0.129	-0.110	-0.166	-0.078	-0.057	-0.120
Level 3 Errors	-0.600	-0.156	-0.146	-0.106	-0.161	-0.094	-0.049	-0.111
Level 4 Errors	-0.662	-0.217	-0.167	-0.108	-0.160	-0.114	-0.046	-0.105
Level 5 Errors	-0.705	-0.292	-0.196	-0.106	-0.153	-0.143	-0.040	-0.092
Level 6 Errors	-0.740	-0.364	-0.231	-0.112	-0.155	-0.178	-0.042	-0.088
Level 7 Errors	-0.766	-0.447	-0.268	-0.118	-0.147	-0.217	-0.043	-0.074

Table 5b: Root MSE for V Estimators with True and Error-Ridden Locations,
On the Plane, rho = 0.3, sigma = 0.2456

			Root MSE					
			NP			NP-Unbiased		
	MLE	MM	L = 3	L = 5	L = 7	L = 3	L=5	L = 7
True Locations	0.104	0.167	0.203	0.303	0.419	0.195	0.327	0.462
Level 1 Errors	0.338	0.177	0.206	0.301	0.417	0.196	0.323	0.459
Level 2 Errors	0.505	0.201	0.210	0.299	0.416	0.196	0.321	0.458
Level 3 Errors	0.602	0.269	0.219	0.295	0.412	0.201	0.317	0.452
Level 4 Errors	0.663	0.339	0.232	0.292	0.407	0.208	0.312	0.445
Level 5 Errors	0.707	0.395	0.250	0.290	0.404	0.220	0.309	0.441
Level 6 Errors	0.741	0.457	0.274	0.287	0.401	0.240	0.303	0.435
Level 7 Errors	0.766	0.505	0.306	0.287	0.401	0.270	0.300	0.436

Table 5c: 95% CI Coverage Probabilities for V Estimators with True
and Error-Ridden Locations, On the Plane, rho = 0.3, sigma = 0.2456

			95% CI Coverage Probability					
			NP			NP-Unbiased		
	MLE	MM	L = 3	L = 5	L = 7	L = 3	L=5	L = 7
True Locations	0.952	0.953	0.935	0.921	0.890	0.944	0.931	0.892
Level 1 Errors	0.904	0.947	0.934	0.924	0.892	0.941	0.930	0.896
Level 2 Errors	0.847	0.937	0.929	0.925	0.892	0.941	0.932	0.894
Level 3 Errors	0.782	0.931	0.928	0.928	0.901	0.937	0.934	0.900
Level 4 Errors	0.752	0.918	0.924	0.932	0.902	0.936	0.934	0.904
Level 5 Errors	0.716	0.902	0.926	0.930	0.904	0.930	0.935	0.913
Level 6 Errors	0.692	0.876	0.919	0.930	0.902	0.925	0.940	0.911
Level 7 Errors	0.679	0.843	0.910	0.928	0.908	0.916	0.939	0.911

Table notes: sample size = 40x40, and true value of V = 1. 1000 Monte Carlo Repetitions.

Table 6a: Bias for V Estimators with True and Error-Ridden Locations,

On the Plane, rho = 0.45, sigma = 0.1423

				Bias				
	MLE	MM	NP			NP-Unbiased		
			L = 3	L = 5	L = 7	L = 3	L=5	L = 7
True Locations	-0.001	-0.013	-0.157	-0.124	-0.180	-0.103	-0.067	-0.129
Level 1 Errors	-0.481	-0.035	-0.171	-0.125	-0.180	-0.117	-0.065	-0.127
Level 2 Errors	-0.605	-0.056	-0.187	-0.126	-0.179	-0.133	-0.063	-0.122
Level 3 Errors	-0.673	-0.074	-0.210	-0.129	-0.178	-0.156	-0.063	-0.118
Level 4 Errors	-0.723	-0.072	-0.237	-0.133	-0.178	-0.184	-0.063	-0.113
Level 5 Errors	-0.758	-0.062	-0.269	-0.139	-0.177	-0.216	-0.066	-0.108
Level 6 Errors	-0.788	-0.033	-0.302	-0.147	-0.178	-0.251	-0.071	-0.103
Level 7 Errors	-0.810	-0.048	-0.338	-0.160	-0.178	-0.289	-0.081	-0.098

Table 6b: Root MSE for V Estimators with True and Error-Ridden Locations,

On the Plane, rho = 0.45, sigma = 0.1423

				Root MSE				
	MLE	MM	NP			NP-Unbiased		
			L = 3	L = 5	L = 7	L = 3	L=5	L = 7
True Locations	0.051	0.217	0.224	0.297	0.413	0.202	0.314	0.449
Level 1 Errors	0.483	0.226	0.233	0.296	0.412	0.207	0.311	0.446
Level 2 Errors	0.607	0.251	0.243	0.293	0.409	0.215	0.308	0.442
Level 3 Errors	0.675	0.297	0.259	0.293	0.409	0.227	0.306	0.440
Level 4 Errors	0.724	0.351	0.280	0.292	0.404	0.244	0.302	0.433
Level 5 Errors	0.759	0.386	0.305	0.290	0.402	0.267	0.297	0.429
Level 6 Errors	0.789	0.401	0.333	0.290	0.398	0.293	0.292	0.422
Level 7 Errors	0.811	0.408	0.364	0.293	0.394	0.324	0.290	0.417

Table 6c: 95% CI Coverage Probabilities for V Estimators with True

and Error-Ridden Locations, On the Plane, rho = 0.45, sigma = 0.1423

				95% CI Coverage Probability				
	MLE	MM	NP			NP-Unbiased		
			L = 3	L = 5	L = 7	L = 3	L=5	L = 7
True Locations	0.959	0.953	0.931	0.930	0.898	0.940	0.935	0.901
Level 1 Errors	0.860	0.949	0.929	0.929	0.901	0.939	0.936	0.904
Level 2 Errors	0.785	0.942	0.925	0.929	0.900	0.936	0.939	0.904
Level 3 Errors	0.751	0.936	0.923	0.930	0.898	0.932	0.935	0.906
Level 4 Errors	0.705	0.934	0.920	0.932	0.899	0.925	0.937	0.909
Level 5 Errors	0.687	0.930	0.913	0.931	0.901	0.921	0.939	0.909
Level 6 Errors	0.650	0.925	0.908	0.927	0.901	0.918	0.940	0.912
Level 7 Errors	0.624	0.921	0.896	0.923	0.909	0.909	0.931	0.918

Table notes: sample size = 40x40, and true value of V = 1. 1000 Monte Carlo Repetitions.

Table 7: Rejection Probabilities for t-Tests Using the Asymptotic Distribution at the 10% Level, On the Plane, rho = 0.3, sigma = 0.2456, L = 7

| | NP | | | | NP-Unbiased | | | |
	MLE [a]	MLE [b]	MM [c]	MM [d]	MLE [a]	MLE [b']	MM [c]	MM [d']
True Locations	0.02	0.24	0.01	0.20	0.06	0.25	0.03	0.22
Level 1 Errors	0.15	0.08	0.01	0.19	0.20	0.10	0.03	0.19
Level 2 Errors	0.38	0.03	0.01	0.15	0.43	0.05	0.04	0.16
Level 3 Errors	0.58	0.01	0.02	0.09	0.61	0.03	0.04	0.11
Level 4 Errors	0.72	0.01	0.03	0.07	0.74	0.03	0.07	0.09
Level 5 Errors	0.81	0.03	0.08	0.04	0.82	0.07	0.14	0.05
Level 6 Errors	0.89	0.06	0.17	0.02	0.89	0.10	0.25	0.03
Level 7 Errors	0.92	0.11	0.31	0.01	0.92	0.18	0.41	0.01

Table notes: sample size = 500, true value of V = 1. 1000 Monte Carlo replications, 200 Bootstrap repetitions. (a) Denominator in the t-test given by the MLE estimator; (b) Denominator in the t-test given by the nonparametric estimator; (c) Denominator in the t-test given by the MM estimator; (d) Denominator in the t-test given by the nonparametric estimator. (b') and (d') Denominator in the t-test given by the unbiased nonparametric estimator.

Table 8: Rejection Probabilities for The Parametric Bootstrap, Using MLE, at the 10% Level, On the Plane

| | rho = 0.3, sigma = 0.2456 | | | | rho = 0.45, sigma = 0.1423 | | | |
| | NP | | NP-Unbiased | | NP | | NP-Unbiased | |
	L = 3	L = 7	L = 3	L = 7	L = 3	L = 7	L = 3	L = 7
True Locations	0.15	0.09	0.15	0.09	0.11	0.09	0.11	0.10
Level 1 Errors	0.59	0.21	0.58	0.19	0.91	0.39	0.90	0.37
Level 2 Errors	0.91	0.41	0.91	0.39	0.99	0.61	0.99	0.58
Level 3 Errors	0.98	0.61	0.98	0.57	1.00	0.76	1.00	0.73
Level 4 Errors	1.00	0.74	1.00	0.69	1.00	0.86	1.00	0.83
Level 5 Errors	1.00	0.81	1.00	0.78	1.00	0.91	1.00	0.89
Level 6 Errors	1.00	0.88	1.00	0.85	1.00	0.95	1.00	0.93
Level 7 Errors	1.00	0.92	1.00	0.90	1.00	0.98	1.00	0.97

Table notes: sample size = 40x40, true value of V = 1. 1000 Monte Carlo replications, 200 Bootstrap repetitions.

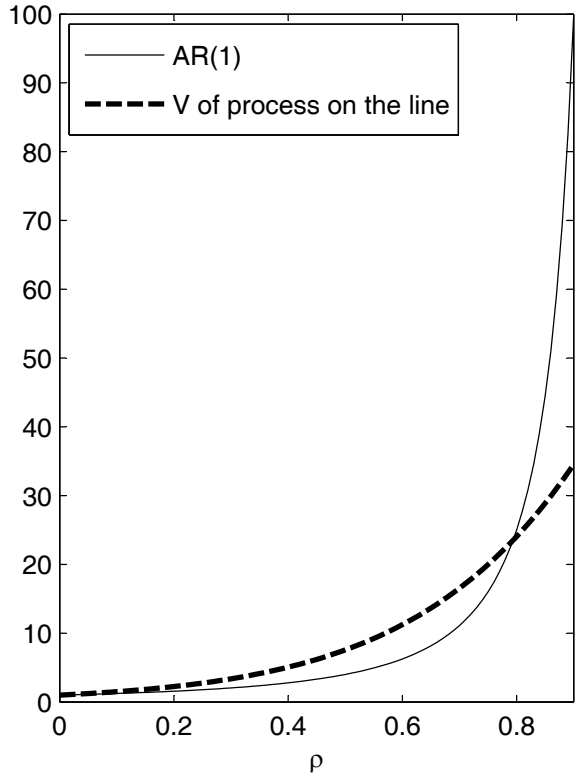Figure 1a: Asymptotic Variance of an AR(1) vs. V on the Line

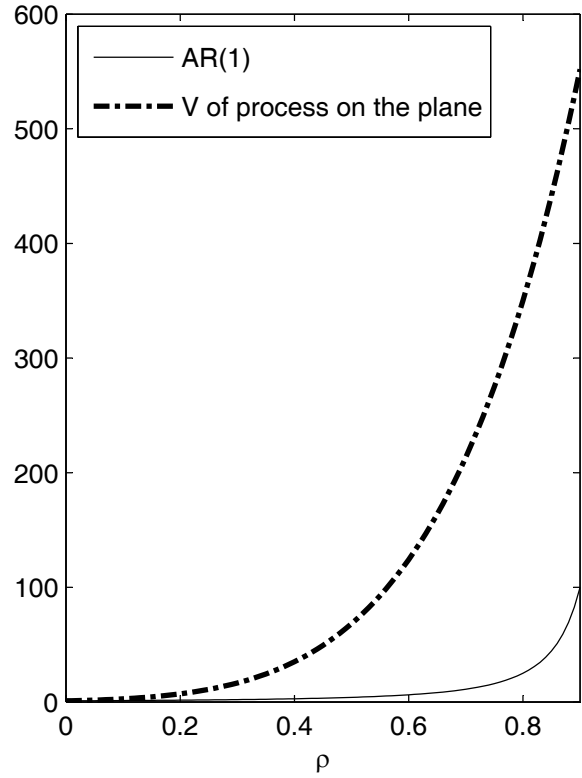Figure 1b: Asymptotic Variance of an AR(1) vs. V on the Plane

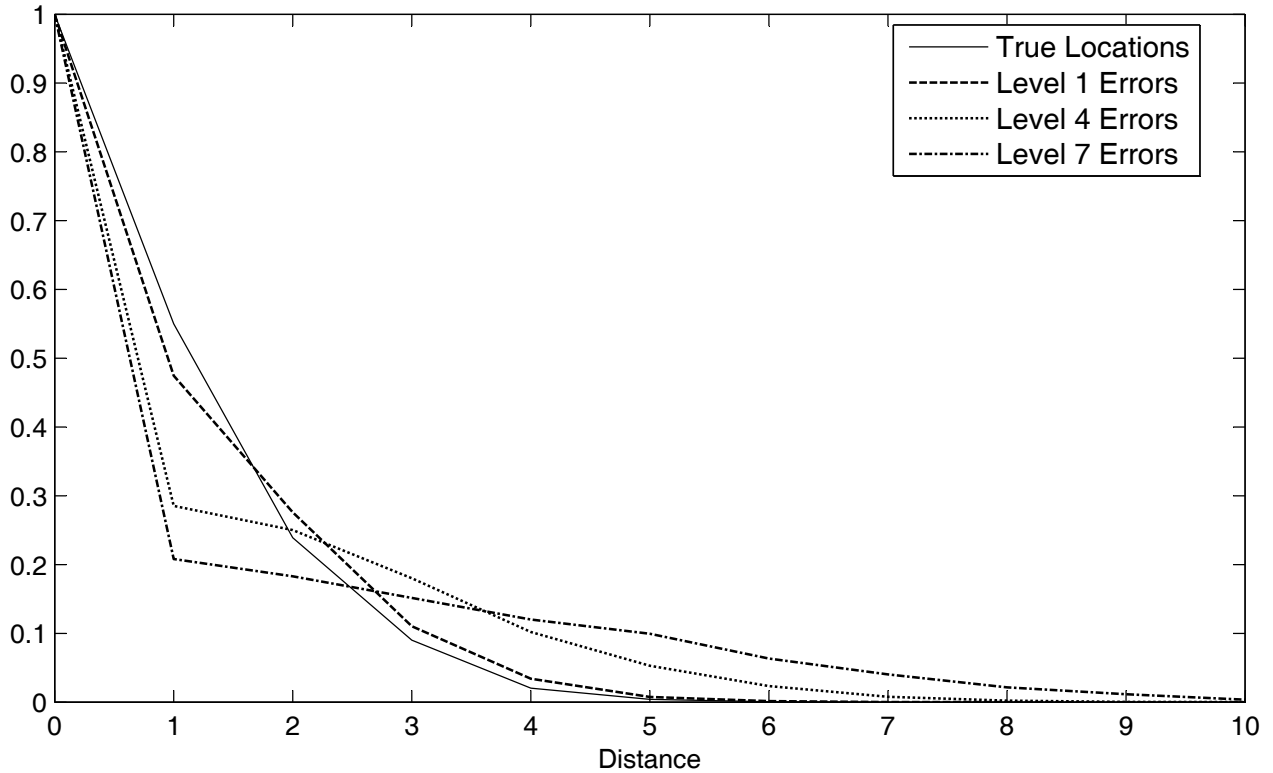Figure 2: Correlations with Different Levels of Location Errors on the Line, $\rho = 0.3$

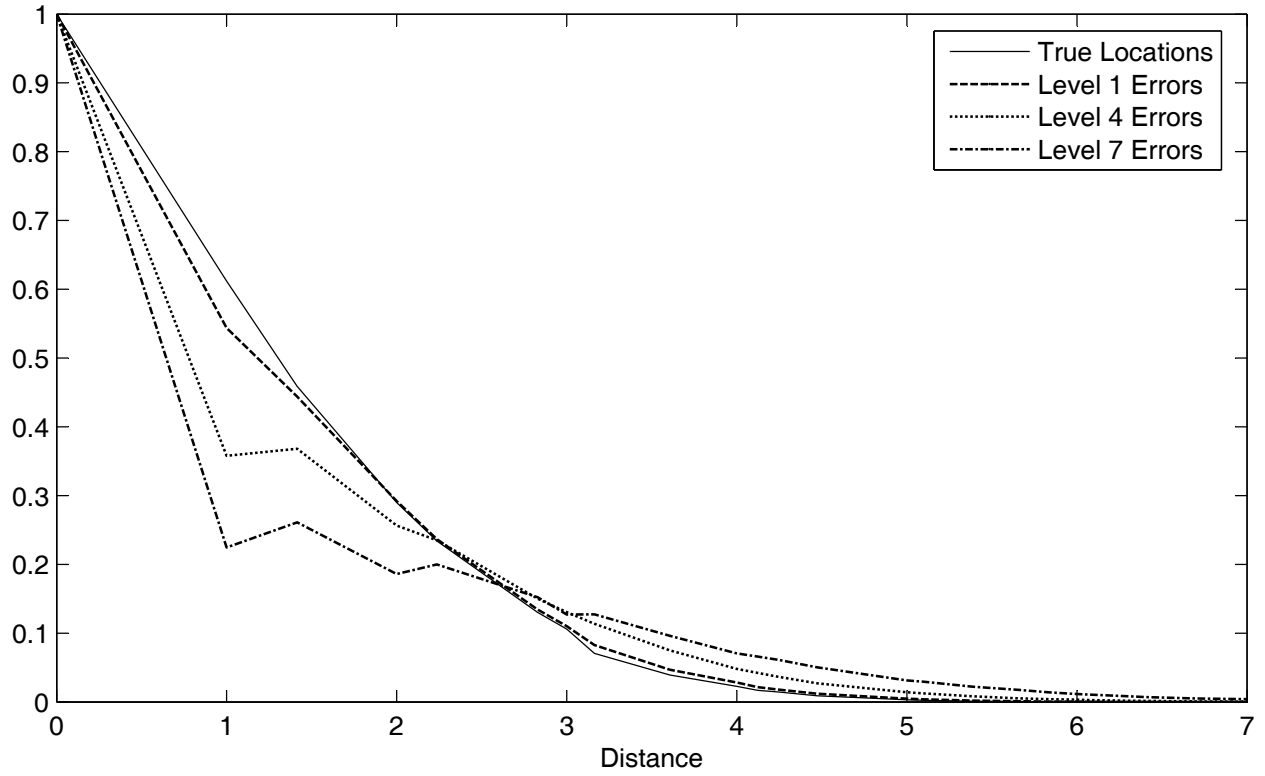Figure 3: Correlations with Different Levels of Location Errors on the Plane, $\rho = 0.3$

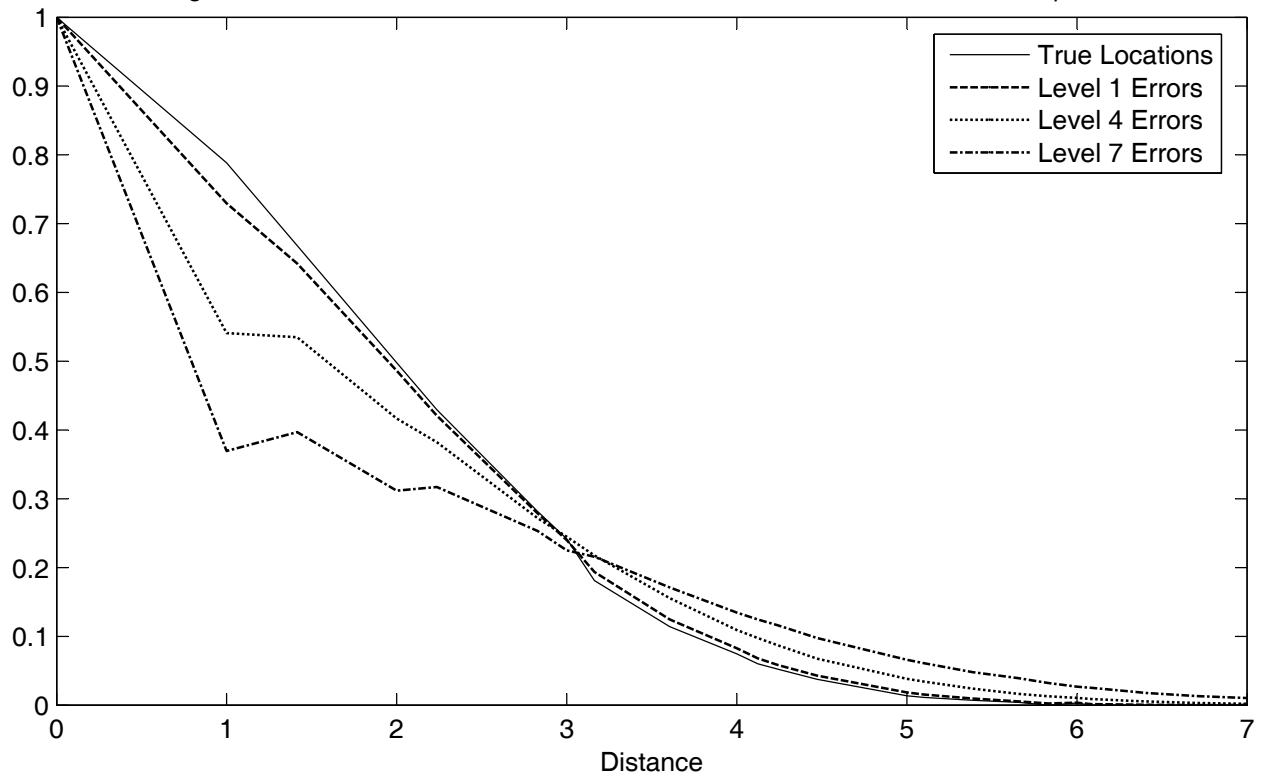Figure 4: Correlations with Different Levels of Location Errors on the Plane, $\rho = 0.45$

Figure 5: Fraction of Observations at a Certain True Distance Receiving Weight = 1 in the Kernel Estimation (On the Line, $L_\tau = 8$)