Institut für Volkswirtschaftslehre

Is Spatial Bootstrapping a Panacea for Valid Inference?

Torben Klarl

universität Augsburg

Is Spatial Bootstrapping a Panacea for Valid Inference?

Torben Klarl[a]

May 20, 2013

Abstract

Bootstrapping methods have so far been rarely used to evaluate spatial data sets. Based on an extensive Monte Carlo study we find that also for spatial, cross-sectional data, the wild bootstrap test proposed by Davidson and Flachaire (2008) based on restricted residuals clearly outperforms asymptotic as well as competing bootstrap tests, like the pairs bootstrap.

Keywords: Spatial econometrics, Paired bootstrap, Wild bootstrap, Parameter inference

JEL Classification Number: C18, C21, R11

[a]University of Augsburg, Department of Economics, Universitätsstraße 16, D-86159 Augsburg, e-mail: torben.alexander.klarl@wiwi.uni-augsburg.de, Tel.: +49(0)821-598-4189, Fax: +49(0)821-598-4231.

1 Introduction

Bootstrapping methods have so far been rarely used to evaluate spatial data sets. Monchuk et al. (2011) is one of the rare empirical contributions to the relevant spatial-econometrics literature who base their inference upon bootstrapping methods[1]. As shown by several contributions[2], for cross-sectional, non-spatial data-sets the so-called wild bootstrap method based on restricted residuals outperforms other methods, like the pairs bootstrap in terms of the reliability of tests. In this paper we turn our attention to the question whether we can find a similar conclusion for cross-sectional data. Based on an extensive Monte Carlo (MC) study we find that also for spatial, cross-sectional data, the wild bootstrap method proposed by Davidson and Flachaire (2008) based on restricted residuals clearly outperforms asymptotic tests as well as competing bootstrap methods, like the pairs bootstrap. We additionally find that based on restricted residuals, it is not always the case that asymptotic tests are outperformed by pairs bootstrap methods. This latter finding directly generalizes the results obtained by Flachaire (2005a), who investigates the sample performance of heteroskedastic-robust tests based on wild and pairs bootstrap methods, but for the case of non-spatial, cross-sectional data, towards a spatial dimension.

The structure of the paper is as follows. In the next section we introduce the well-known spatial error model (SEM) for cross sectional data[3] and discuss the heteroskedastic-robust estimation of parameters. In section 3 we introduce the simulation design. Section 4 deals with a short overview of the employed bootstrap methods. Section 5 presents the structure of the MC simulation and discusses the obtained results. Section 6 concludes.

---

[1]For instance, Anselin (1988) on p. 94 and Fingelton (2008) give an introduction to bootstrapping in the context of spatial models.

[2]Among others, please refer to Davidson and Flachaire (2008), MacKinnon (2002), Flachaire (2005a) and Flachaire (2005b).

[3]The focus on this specific spatial model is without loss of generality. The papers procedures can be easily extended towards more general and complex spatial models.

2 First order Spatial Error Model (SEM)

Let us consider the linear heteroskedastic first order spatial error model (SEM)[4]

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{1}$$

$$\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\epsilon}, \ \lambda \in (-1, 1) \tag{2}$$

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \tag{3}$$

with $\mathbf{X}$ as a $(N \times K)$ data matrix, $\mathbf{y}$ as a $(N \times 1)$ vector of observations. $\mathbf{W}$ is treated as a non-stochastic, spatial $(N \times N)$ contiguity-matrix, $\mathbf{u}$ is a $(N \times 1)$ vector of spatially correlated errors and $\boldsymbol{\epsilon}$ is treated as a $(N \times 1)$ vector with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, which are heteroskedastic but not correlated across space. For the majority of cross-sectional spatial data evaluations it seems to be appropriate to assume that the relevant data exhibits the tendency of being heteroskedastic[5]. As known from the relevant econometrican literature, parameter inference deserves special precautions when the errors are heteroskedastic. Ignoring the heteroskedastic structure of the errors generally results in inconsistent covariance matrix estimators of the ML-estimator of $\boldsymbol{\beta}$, denoted as $\boldsymbol{b}$.

Eicker (1963) and White (1980) solved this problem by proposing a Heteroskedasticity Consistent Covariance Matrix Estimator (HCCME), which allows for asymptotically correct parameter inference regarding $\boldsymbol{b}$ given the error's structure exhibit heteroskedastictiy of unknown form. The corresponding heteroskedastic robust version of the variance-covariance matrix of $\boldsymbol{b}$ reads as

$$\hat{Var}(\beta) = (\boldsymbol{X'X})\boldsymbol{X}\hat{\boldsymbol{\Omega}}\boldsymbol{X}(\boldsymbol{X'X})^{-1}, \tag{4}$$

with $\hat{\boldsymbol{\Omega}}$ as a $(N \times N)$ diagonal matrix with elements $(\Gamma_l e_i)^2$, whereas $\boldsymbol{e} = [e_1, ..., e_N]'$ represents the $(N \times 1)$ vector of spatially-filtered ML residuals based on model (1)-(3)[6]. Γ_l stands for a specific weighting scheme with

---

[4]Refer to Anselin (1988) on p. 35 for instance.

[5]For instance, refer to Monchuk et al (2011).

[6]In the following, we assume that $\boldsymbol{e}$ is spatially-filtered by employing a Cochrance-Orcutt type transformation.

$l = \{0, 1, 2, 3\}$ characteristics. MacKinnon and White (1985) propose different possible explicit forms of the HCCME. These are given as $\Gamma_0 \equiv 1$, $\Gamma_1 \equiv \sqrt{\frac{N}{N-K}}$, $\Gamma_2 \equiv \frac{1}{\sqrt{1-h_i}}$ and $\Gamma_3 \equiv \frac{1}{1-h_i}$, with $h_i \equiv \mathbf{x}_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i$ as the i-th element of the orthogonal projection matrix on the span of the $\mathbf{X}$ columns.

As further shown by MacKinnon and White (1985) and Chesher and Jewitt (1987), Γ_0 is outperformed by Γ_1, Γ_1 instead is outperformed by Γ_2 and Γ_3, whereas it seems that Γ_3 seems superior compared to Γ_2 for some typical cases with respect to the Error Rejection Probability (ERP), which is defined as the difference between the true rejection probability of a test and its nominal level α[7].

Although the HCCME guarantees asymptotic valid results, in finite samples however, even the HCCME can exhibit serious size distortions, notably with the presence of high-leverage observations $\boldsymbol{x}_i$. Hence, as shown by Davidson and Flachaire (2008), MacKinnon and White (1985), Godfrey and Orme (2001) and Flachaire (2005a), hypothesis test (e.g. t-tests or F-tests) based on the asymptotic HCCME can be misleading. From this point of view it is worth examining whether tests based on bootstrap methods can improve the realibilty of tests, given we are confronted with cross-sectional spatial data.

3 Model design

Our experiments are based upon model (1)-(3) with two fixed[8] explanatory variables $\boldsymbol{x}_i = [\boldsymbol{x}_{1i}, \boldsymbol{x}_{2i}] = [\mathbf{1}, \boldsymbol{x}_{2i}]$ and the true parameter vector $\boldsymbol{\beta} = [\beta_1, \beta_2]' = [1, 0]'$. We further assume that $\boldsymbol{x}_{2i}$ is drawn from a standard log-normal distribution to account for high leverage observations (Flachaire (2005a), Flachaire (2005b)). To generate heteroskedastic errors, we assume that $\sigma_i^2 = x_{i2}^2$, $\forall\, i = \{1, ..., N\}$. ϵ_i is treated to be white noise with $\mathcal{N}(0, 1)$. We choose different values for the spatial error parameter $\lambda = \{0.0, 0.4, 0.8\}$ examining whether

---

[7]For instance, refer to Davidson and MacKinnon (1998) and Flachaire (2005a). A formal definition of the ERP is given in section 5 of this contribution.

[8]We assume that both explanatory variables are treated to be fixed in repeated samples.

different spatial error degrees reflected by λ affect simulation results. As bootstrap methods are designed for small samples, we choose $N = 50$[9].

To avoid severe experimental errors, the number of MC simulations is chosen reasonably large with $M = 4,000$ runs. The number of bootstrap replications is set to $B = 399$[10]. The assumed simulation environment mimics to some extent Davidson and Flachaire (1996), Flachaire (2005a) and Lin et al. (2011).

We wish to test the null hypothesis $H_0 := \beta_2 = 0$. For the null hypothesis, we compute the HCCME-based t-static as

$$t = \frac{\boldsymbol{x}_2' \boldsymbol{M}_1 \boldsymbol{y}}{\sqrt{(\boldsymbol{x}_2' \boldsymbol{M}_1 \tilde{\boldsymbol{\Omega}} \boldsymbol{M}_1 \boldsymbol{x}_2)}}, \tag{5}$$

with $\boldsymbol{M}_1 = \boldsymbol{I}_N - \boldsymbol{1}(\boldsymbol{1}'\boldsymbol{1})^{-1}\boldsymbol{1}'$ as the residual generating matrix. Based on simulation results, we examine the finite sample performance of the t-test represented by equation (5) and compare the performance of the asymptotic test with its bootstrap counterparts based on their individual ERP.

4 Bootstrap methods

One of the widely known and applied bootstrap Data Generating Processes (DGP)s are the residual bootstrap, pairs bootstrap and wild bootstrap. As noted by MacKinnon (2009), in the case of heteroskedastic errors, the residual bootstrap cannot be used, as this methods assumes independent and identically distributed errors. As we assume heteroskedastic errors and high leverage observations (see section 2 and 3), in the following we exclusively concentrate on the pairs and wild bootstrap. For the case of non-spatial models, Brownstone

[9]Please not that for the majority of evaluated spatial data-sets we can restrict the parameter space of λ towards $\lambda \in [0, 1)$.

[10]As shown by Lin et al. (2011), who examine the size and power of bootstrap tests, the spatial bootstrap tests stabilize for $B = 399$ or more replications. Overall, we perform $28,752,000$ ML-regressions based upon model (1)-(3)! Hence, to conserve computational space, we set $M = 4,000$ and $B = 399$. The regressions are run employing Matlab R2011a on a Pentium 7, 2.8 Ghz Windows 7 system.

and Valletta (2001) and Horowitz (1997, 2000) concentrate on the question whether the ERP of a specific test based on the HCCME differs significantly, given one refers to the wild or pairs bootstrap. For cross-sectional models, they observe that the wild bootstrap appears to outperform the pairs bootstrap. However, as noted by MacKinnon (2002), the performance of bootstrap methods directly depends on the model's structure. Here, we provide additional evidence for this question with a focus towards spatial, cross-sectional data. The next subsections deal with the introduction of the pairs and the wild bootstrap.

4.1 Wild bootstrap

As noted by MacKinnon (2009) in section 6.5.3 on p. 196, "[...] *if the form of heteroskedasticity is unknown, the best method that is currently available, at least for tests on the coefficient vector $\boldsymbol{\beta}$, appears to be the wild bootstrap*". To initialize the wild bootstrap, which was originally developed in Liu (1988) and based upon the contributions of Wu (1986) and Beran (1986), we start with the following bootstrap DGP:

$$y_i^* = \boldsymbol{x}_i \boldsymbol{\beta} + \Gamma_l e_i \Psi_i \tag{6}$$

Ψ_i is treated as a random variable following a specific distribution G with $E(\Psi_i) = 0$ and $E(\Psi_i^2) = 1$ for all $i = \{1, ..., N\}$. As recommended by Davidson and Flachaire (2008), Ψ_i should follow a two-point *Rademacher* distribution with

$$G_1 : \Psi_i = \begin{cases} 1 & \text{with probability } 0.5 \\ -1 & \text{with probability } 0.5 \end{cases} \tag{7}$$

given the disturbances are not asymmetric. As shown by Davidson and Flachaire (2008) and Godfrey and Orme (2001), the full Edgeworth expansion of the wild bootstrap assuming G_1 is considerably smaller compared to G_2 for sufficiently

small sample sizes and/or if observations with high leverage a present. However, if the distribution of the errors is asymmetric, G should be better defined as follows (Mammen (1993)):

$$G_2 : \Psi_i = \begin{cases} -\frac{(\sqrt{5}-1)}{2} & \text{with probability } \frac{(\sqrt{5}+1)}{2\sqrt{5}} \\ \frac{(\sqrt{5}+1)}{2} & \text{with probability } \frac{(\sqrt{5}-1)}{2\sqrt{5}} \end{cases} . \tag{8}$$

4.2 Pairs bootstrap

In the relevant literature, two versions of the pairs bootstrap are present: those which can be traced back to Freedman (1981) and a version based upon a modified re-sampling scheme, introduced by Flachaire (1999).

1. Pairs bootstrap proposed by Freedman (1981)

   The pairs bootstrap which was originally advocated by Freedman (1981) resamples directly from the row-elements of the data-matrix $[\boldsymbol{y}, \boldsymbol{X}]$. This method assumes that each pair of observations $[\boldsymbol{y}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}]$ is an independent draw from a multivariate distribution (MacKinnon (2009)). It is clear that the errors do not meet the requirement of being homoskedastic. Given this procedure which is completely non-parametric, the bootstrap sample is not based on a model which acknowledges the null hypothesis (e.g. $H_0 := \beta_2 = 0$) as noted for instance by Flachaire (2005a) and MacKinnon (2009). Hence, we have to modify the null hypothesis presented in section 3 towards $H_0 := \beta_2 = b_2$, with b_2 being the unrestricted ML-estimate of β_2 based on model (1)-(3)[11].

2. Pairs bootstrap proposed by Flachaire (1999)

   As the pairs bootstrap proposed by Freedman (1981) is somewhat unsatisfactory regarding bootstrap testing, Flachaire (1999) introduces another version of the pairs bootstrap. Instead of resampling from $[\boldsymbol{y}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}]$, he generates bootstrap samples by resampling $[\boldsymbol{x}_{i1}^*, \boldsymbol{x}_{i2}^*, e_i^*]$ based upon

---

[11]For a detailed discussion, refer to the introduction of Hall (1992).

$[\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \Gamma_l e_i]$, with e_i being the unrestricted ML-residual from estimating model (1)-(3). Hence, the bootstrap DGP reads as

$$y_i^* = \boldsymbol{x}_i^* \tilde{\boldsymbol{\beta}} + e_i^*, \tag{9}$$

with $\tilde{\boldsymbol{\beta}}$ being the vector of parameter estimates under the null hypothesis. Obviously, in this case it is directly possible to test the null hypothesis $H_0 := \beta_2 = 0$.

4.3 Pairs bootstrap versus wild bootstrap

The ideal bootstrap DGP perfectly mimics the real underlying DGP. However, as noted by Flachaire (1999, 2005a), the pairs bootstrap proposed by Freedman (1981) suffers in two aspects: (1) as we draw regressors and the dependent variable at the same time, we expect that regressors are not exogenous, and hence, $E[\boldsymbol{e}^*|\boldsymbol{X}^*] \neq 0$. (2) The DGP is constructed without restricted parameter estimates.

On the contrary, if the real underlying DGP assumes that regressors are exogenous, it is obvious that the bootstrap DGP cannot mimic the underlying features of the real DGP satisfactorily. Obviously, Flachaire's (1999) pairs bootstrap method instead corrects for the second drawback sufficiently. It is further obvious that only the wild bootstrap corrects for the two drawbacks simultaneously. Hence, we should expect that based on the model (1)-(3), parameter inference upon the wild bootstrap method should outperform parameter inference based upon the above introduced pairs bootstrap procedures. As we will see later, our simulation study confirms this conjecture.

5   Simulation results

Before we introduce the specific structure of our MC experiment, we make some preliminary remarks and provide a short discussion regarding the appropriateness of restriced and unrestricted residuals in the context of parameter inference. Finally, we present the simulation results.

5.1 Preliminary remarks

We follow Flachaire (2005a) and choose the Γ_3 transformation of residuals for our simulation study. Further, as recommended by Davidson and Flachaire (2008) we use the so-called Rademacher distribution as introduced in section 4.1. Hence, the latter choice directly implies $E[\Psi_i^4] = 1$, and thus the fourth order feature of the errors are retained. The spatial weight matrix $\boldsymbol{W}$ is designed as a first-order contiguity matrix, which is row-standardized with zero diagonal elements[12].

5.2 Restricted or unrestricted residuals

A large body of literature deals with the question whether to use restricted or unrestricted residuals for bootstrap based parameter inference. Davidson and MacKinnon (1985) show that the reliability of asymptotic tests based on restricted residuals generally outperforms the reliability of asymptotic tests based on unrestricted residuals as the former slightly under-reject the null hypothesis, whereas the latter severely over-reject the null. On the contrary, van Giersbergen and Kiviet (2002) note that given the null hypothesis is not true, unrestricted residuals tend to improve the power of a test. However, based on simulation experiments, MacKinnon (2002) does not find support of the latter conclusion. Davidson and Flachaire (2008) point to the fact that it is wrong to mix restricted (for the bootstrap DGP) and unrestricted (in the HCCME) residuals. Finally, Godfrey and Orme (2001) show that one does not obtain good results in terms of a low ERP for the wild bootstrap, given the HCCME is calculated with unrestricted residuals. In this study we follow Flachaire (2005a) and investigate the performance of asymptotic and bootstrap test statistics based on the HCCME, computed both with unrestricted and restricted residuals.

---

[12]$\boldsymbol{W}$ remains fixed in repeated samples.

5.3 Structure of the MC experiment

We examine the performance of the t-test with a MC experiment. Each MC replication $m = \{1, ..., M\}$ consists of the following steps presented below:

1. Based on an original data-set, estimate model (1)-(3) consistently with ML. Store the obtained ML-estimate b and compute the m-th t-statistic based on equation (5) for the null hypothesis $H_0 := \beta_2 = 0$, labeled as t_m. Further, obtain constrained parameter estimates $\tilde{b}$ and $\tilde{e}$[13]. Compute the m-th t-statistic based on the HCCME transformed constrained residuals, labeled as $\tilde{t}_m$.

2. For every m-th MC experiment, re-center and rescale ML residuals as proposed by Lin et al. (2011). Generate bootstrap data-sets for the wild and the two variants of the pairs bootstrap. Use the bootstrap samples and re-calculate the constrained and unconstrained version of the test statistic (5). Repeat this step B-times and obtain $(B \times 1)$-vectors of bootstrap statistics $t_m^* = [t_{m,1}^*, ..., t_{m,399}^*]'$ and $\tilde{t}_m^* = [\tilde{t}_{m,1}^*, ..., \tilde{t}_{m,399}^*]'$ based on constrained and unconstrained residuals, e_i and $\tilde{e}_i$, respectively.

3. For the m-th MC experiment calculate the equal-tail p-values of the bootstrap unconstrained and constrained t-statistic as follows:

$$p_m^* = 2 \min \left(\frac{1}{B} \sum_{b=1}^{B} \boldsymbol{I}(t_{m,b}^* \leq t_m), \frac{1}{B} \sum_{b=1}^{B} \boldsymbol{I}(t_{m,b}^* > t_m) \right) \qquad (10)$$

$$\tilde{p}_m^* = 2 \min \left(\frac{1}{B} \sum_{b=1}^{B} \boldsymbol{I}(\tilde{t}_{m,b}^* \leq \tilde{t}_m), \frac{1}{B} \sum_{b=1}^{B} \boldsymbol{I}(\tilde{t}_{m,b}^* > \tilde{t}_m) \right), \qquad (11)$$

with $\boldsymbol{I}(\cdot)$ representing an indicator function[14].

Given $m = \{1, ..., M\}$ MC experiments and a sequence of nomial levels $\alpha = \{0, 0.001, ..., 1\}$, we calculate the vectors of the true rejection probabilities

---

[13]The parameter estimates are constrained in the sense that $H_0 := \beta_2 = 0$ is true.

[14]See Davidson and MacKinnon (2006).

as

$$\frac{1}{M} \sum_{i=1}^{M} \boldsymbol{I}(p^* < \alpha), \tag{12}$$

$$\frac{1}{M} \sum_{i=1}^{M} \boldsymbol{I}(\tilde{p}^* < \alpha), \tag{13}$$

with $\boldsymbol{I}(\cdot)$ representing an indicator function for the constrained and unconstrained versions of the t-test. Finally, we calculate the difference between the true rejection probabilities and the sequence of nominal levels α which results in the ERP[15] for the unconstrained and constrained version of the t-test:

$$ERP \equiv \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{I}(p^* < \alpha) - \alpha \tag{14}$$

$$E\tilde{R}P \equiv \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{I}(\tilde{p}^* < \alpha) - \alpha. \tag{15}$$

5.4   Simulation results

We refer to graphical methods proposed by Davidson and MacKinnon (1998) to present the simulation results. In particular, the experimental-based ERPs are visualized referring to p-value discrepancy plots as described by them. A test is called reliable, if it rejects the null hypothesis at the nominal level α, given the null hypothesis is true, and hence, the ERP is close to zero. Otherwise we should expect a significant ERP. Based on the arguments given above, we should expect that the ERP of the wild bootstrap should be closer to zero than of the competing procedures.

[Figure 1 here]

Figure 1 shows the ERP of the asymptotic (*asymp*), the pairs bootstrap advocated by Freedman (*pairs2*), the pairs bootstrap proposed by Flachaire (1999), (*pairs1*) and the wild bootstrap (*wild*) tests based on the HCCME calculated with both unrestricted and restricted residuals (*re*)[16]. Figure 1 (a)

---

[15]The ERP is also called size distortion.

[16]For instance, *asymp re* stands for the asymptotic t-test based on the HCCME computed with restricted residuals.

represents the ERP of the t-test for $\lambda = 0.0$, figure 1 (b) shows the ERP of the t-test for $\lambda = 0.4$, and finally, figure 1 (c) reflects the ERP of the t-test for $\lambda = 0.8$.

Davidson and MacKinnon (1985) and Flachaire (2005a) show for a non-spatial, cross-sectional analysis that for a small nominal level α, asymptotic tests based on unrestricted residuals tend to over-reject the null hypothesis, whereas asymptotic tests based on restricted residuals under-reject the corresponding null hypothesis. As shown in the first two columns of the next table, we can also confirm these results for a spatial, cross-sectional setup.

Further we can observe that for the wild bootstrap we have a better control over the ERP if we refer to tests based on HCCME computed with restricted residuals: As mentioned above, for all examined cases of the spatial error parameter λ, we observe that the wild bootstrap based on restricted residuals exhibits smaller size-distortions compared to the wild bootstrap with unrestricted residuals. However, increasing the spatial error influence λ tends to deteriorate the ERP of the t-test based on the wild bootstrap. Nevertheless, even for a severe spatial error influence (e.g. $\lambda = 0.8$), we still observe that t-tests based on the wild bootstrap exhibit the smallest ERP compared to their competitors. As further pointed out by Davidson and MacKinnon (2006), the ERP differences between competing tests can be directly translated into power differences of the same magnitude. Hence, for the further analysis we make use of the restricted residuals. As we are primarily concerned with small nominal levels in practical applications, we report the ERPs for $\alpha = 0.05$ in table (1) for different degrees of spatial error influences λ.

As we can directly observe from table (1), for small nominal levels (e.g. $\alpha = 0.05$) we conclude that in terms of reliability, the wild bootstrap based t-test clearly outperforms the asymptotic as well as the competing bootstrap methods. Further from reflecting figure (1), we can observe that even if we use restricted residuals, we cannot say that for all nominal levels α the asymptotic test is outperformed by the bootstrap tests. Only the wild bootstrap

12

λ	asymp	asymp re	wild re	pairs1 re	pairs2 re
0.00	0.0977	-0.0480	0.0144	-0.0480	-0.0361
0.40	0.0754	-0.0483	-0.0051	-0.0488	-0.0473
0.80	0.0577	-0.0483	-0.0176	-0.0490	-0.0490

Table 1:

ERPs of the HCCME based t-test at $\alpha = 0.05\%$ for $N = 50$, $M = 4,000$ and $B = 399$

test clearly outperforms not only the asymptotic test but also its bootstrap competitors. This conclusion is best visualized for the case of absence of spatial correlation in the errors ($\lambda = 0$): if we look at figure (1) (a), we observe a nearly perfect control of the *wild re* for all nominal levels. This qualitatively coincides with the findings of Flachaire (2005a) for a pure non-spatial, cross-sectional setup.

6   Concluding comments

Bootstrapping methods suggest an appealing method for reasonable spatial model's parameter inference. In this paper we have discussed the finite-sample performance of a heteroskedasticity robust t-test based on a cross-sectional, spatial data environment. We find that even for a spatial data-environment, the wild boostrap method *wild re* advocated by Davidson and Flachaire (2008) always provides a significant better performance than competing procedures, such as the pairs bootstrap or other versions of the wild bootstrap. Further, we find that boostrap methods perform better than asymptotic tests based on unrestricted residuals. In line with the findings of Flachaire (2005a) for non spatial, cross-sectional data, the pairs bootstrap does not necessarily perform better in terms of the ERPs than asymptotic tests based on restricted residuals. In a nutshell, the simulation results strongly suggest to base parameter inference on the wild bootstrap method recommended by Davidson and Flachaire (2008).

13

References

Anselin, L (1988). Spatial Econometrics: Methods and Models. Kluwer, Dordrecht.

Beran, R. (1986): Discussion of "Jackknife bootstrap and other resampling methods in regression analysis" by C. F. J. Wu. Annals of Statistics, 14, 1295-1298.

Brownstone, D. and Valletta, R. (2001): The Bootstrap and Multiple Imputations: Harnessing Increased Computing Power for Improved Statistical Tests. Journal of Economic Perspectives, 15, 129-141.

Chesher, A. and I. Jewitt (1987): The Bias of a Heteroskedasticity Consistent Covariance Matrix Estimator. Econometrica, 55, 1217-1222.

Davidson, R. and E. Flachaire (1996): Monte Carlo Evidence on the Behavior of the Wild Bootstrap in the Presence of Leverage Points, NSF Symposium on the Bootstrap, University of California, Berkeley.

Davidson, R. and E. Flachaire (2008): The Wild Bootstrap, Tamed at Last. Journal of Econometrics, 146, 162-169.

Davidson, R. and J. G. MacKinnon (1985): Heteroskedasticity-robust tests in regression directions. Annales de l'INSÉE, 59/60, 183-218.

Davidson, R. and J. G. MacKinnon (1998): Graphical methods for investigating the size and power of hypothesis test. The Manchester School, 66, 1-26.

Davidson, R. and J. G. MacKinnon (2006): The power of bootstrap and asymptotic tests. Journal of Econometrics, 133, 421-441.

Davidson, R. and J. G. MacKinnon (2006): Bootstrap methods in econometrics. In K. Patterson and T. C. Mills (Eds.): Palgrave Handbook

of Econometrics: Volume 1 Theoretical Econometrics. Basingstoke, Palgrave Macmillan.

Eicker, B. (1963): Limit Theorems for Regression with Unequal and Dependent Errors. Annals of Mathematical Statistics, 34, 447-456.

Fingelton, B (2008): A Generalized Method of Moment Estimator for a Spatial Model with Moving Average Errors, with Application to Real Estate Prices. Empirical Economics, 34, 35-57.

Flachaire, E. (1999): A better way to bootstrap pairs. Economics Letters, 64, 257-262.

Flachaire, E. (2005a): Bootstrapping Heteroskedastic Regression Models: Wild Bootstrap vs. Pairs Bootstrap. Computational Statistics & Data Analysis, 49, 361-376.

Flachaire, E. (2005b): More Efficient Tests Robust to Heteroskedasticity of Unknown Form. Econometrics Reviews, 24, 219-241.

Freedman, David. A. (1981): Bootstrapping Regression Models. Annals of Statistics, 9, 1218-1228.

Godfrey, L. G. and C. D. Orme (2001): Significance levels of heteroskedasticity-robust tests for specification and misspecification: some results on the use of wild bootstraps. Paper presented at ESEM 2001, Lausanne.

Hall, P. (1992): The bootstrap and Edgeworth Expansion. Springer, New York.

Horowitz, J. L. (1997): Bootstrap methods in Econometrics: Theory and Numerical Performance. In: D. M. Kreps and K. F. Wallis (Eds.), Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. 3. Cambridge University Press, UK.

Horowitz, J. L. (2000): The bootstrap. In: J. J. Heckman, E. E. Leamer (Eds.), Handbook of Econometrics, Vol. 5. Elsevier Science North-Holland, Amsterdam, London and New York.

Lin, K., Z. Long, and B. Ou (2011): The Size and Power of Bootstrap Tests for Spatial Dependence in a Linear Regression Model. Computational Economics, 38, 151-171.

Liu, R. Y. (1988): Bootstrap and Edgeworth Expansion. Springer Series in Statistics. Springer, New York.

MacKinnon, J. G. and H. L. White (1985): Bootstrap Inference in Econometrics. Canadian Journal of Economics, 35, 615-645.

MacKinnon, J. G. (2002): Bootstrap Inference in Econometrics. Canadian Journal of Economics, 35, 615-645.

MacKinnon, J. G. (2006): Bootstrap Methods in Econometrics. Economic Record, 82, Issue Supplement s1, pages S2-S8.

MacKinnon, J. G. (2009): Bootstrap hypothesis testing. In: D. A. Belsley, E. J. Kontoghiorghes (Eds.), Handbook of Computational Econometrics. John Wiley & Sons, Ltd., Chichester.

Monchuk, D. C., D. J. Hayes, J. A. Miranowski and D. M. Lambert (2011): Inference Based on Alternative Bootstrapping Methods in Spatial Models with an Application to County Income Growth in the United States. Journal of Regional Science, 51, 880-896.

van Giersbergen, N. P. A. and J. F. Kiviet (2002): How to implement bootstrap hypothesis testing in static and dynamic regression model: test statistic versus confidence interval approach. Journal of Econometrics, 108, 133-156.
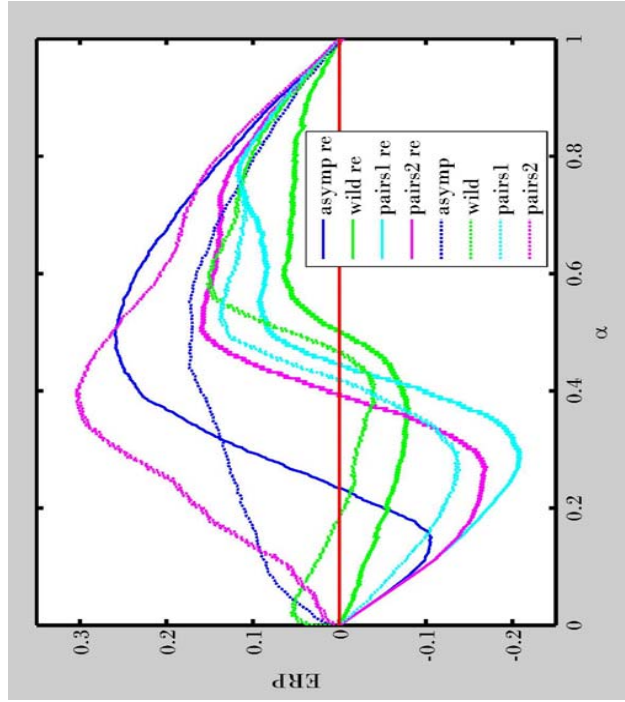
White, H. (1980): A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica, 48, 817-838.

Wu, C. F. J. (1986): Jackknife Bootstrap and other Resampling Methods in Regression Analysis. Annals of Statistics, 14, 1261-1295.
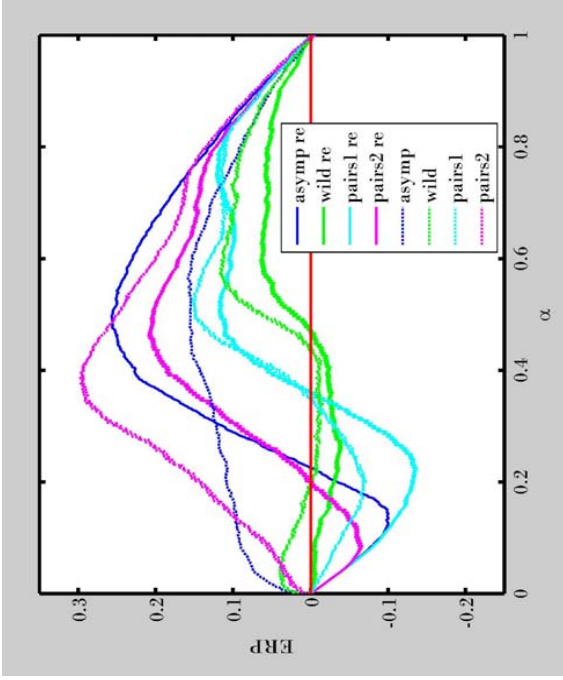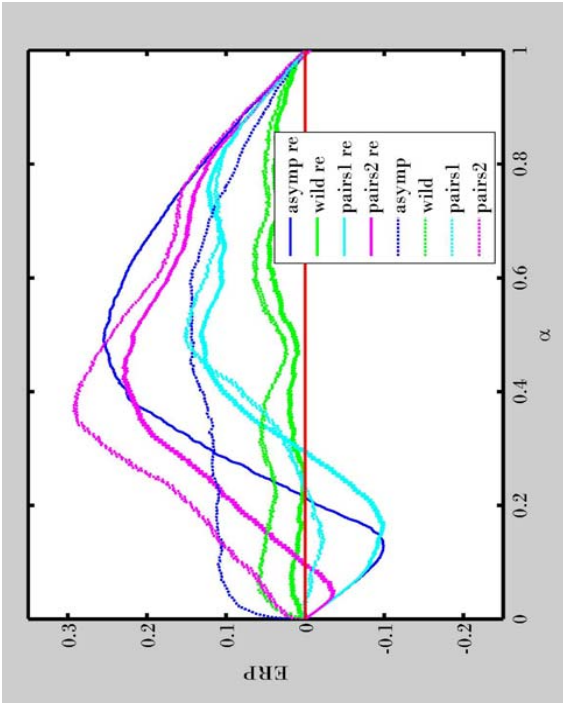
Figure 1: ERP of *t*-tests based on unrestricted and restricted residuals