Equal Predictive Ability Tests for Panel Data with an Application to OECD and IMF Forecasts*

Oguzhan Akgun[a], Alain Pirotte[a], Giovanni Urga[b] and Zhenlin Yang[c]

[a] *CRED, University Paris II Panthéon-Assas, France*

[b] *Cass Business School, London, United Kingdom, and Bergamo University, Italy*

[c] *School of Economics, Singapore Management University, Singapore*

June 28, 2019

Abstract

This paper proposes novel tests for equal predictive ability in panels of forecasts allowing for different types and strength of cross-sectional dependence across units. We compare the predictive ability of two forecasters using forecast errors from different units correlated via common factors and spatial spillovers. We compute size and power of these tests in finite samples by means of an extensive Monte Carlo study finding very good small sample properties. Finally, we apply the tests to compare the economic growth predictions of the OECD and IMF.

Keywords: Cross-Sectional Dependence; Forecast Evaluation; Forecasting; Heterogeneity; Hypothesis Testing Panel Data.

JEL classification: C12, C14, C52, C53.

*We wish to thank the participants of the seminars at Cass and CRED in September 2018; the 18th International Workshop on Spatial Econometrics and Statistics at AgroParisTech, Paris, 23-24 September 2019, in particular the discussant Paul Elhorst and Davide Fiaschi; and 39th International Symposium on Forecasting at Thessaloniki, 16-19 June 2019, for their helpful comments. The usual disclaimer applies.

1 Introduction

Formal tests of the null hypothesis of no difference in the forecast accuracy using two time series of forecast errors have been widely discussed in the literature and formalized, for instance, by Vuong (1989), Diebold and Mariano (1995, hereafter DM), West (1996), Clark and McCracken (2001, 2015), Giacomini and White (2006, hereafter GW), Clark and West (2007), among others. Whereas the literature in panel data taking into consideration the specific challenges such as heterogeneity and cross-sectional dependence (CD) is scarce, with a few exceptions. First is Davies and Lahiri (1995, hereafter DL) who focus on testing unbiasedness and efficiency of forecasts made by several different agents for the same unit. Their analysis is based on a three dimensional panel data regression where the dimensions are agents generating the forecasts, target years and forecast horizons. Second is undertaken by Timmermann and Zhu (2019) who focus on predictions produced for several different units but their framework is based mostly on tests which use a single cross-section of forecasts or on cross-sectional aggregates of a panel of prediction errors.

The main aim of this paper is to propose tests for the equal predictive ability (EPA) hypothesis for panel data taking into account both the time series and the cross-sections features of the data. We propose tests allowing to compare the predictive ability of two forecasters, based on n units, hence n pairs of time series of observed forecast errors of length T, from their forecasts on an economic variable. Various panel data tests of EPA are proposed, extending that of DM which concerns a single time series. Contrary to DL, our tests are developed for forecasts made for different panel units.

We develop two types of tests of predictive ability. The first one focuses on EPA on average over all panel units and over time. This test is useful and of economic importance when the researcher is not interested in the differences of predictive ability for a specific unit but the overall differences. In the second type of tests, to deal with possible heterogeneity, we focus on the null hypothesis which states that the EPA holds for each panel unit. To deal with weak cross-sectional dependence (WCD) and strong cross-sectional dependence (SCD), we follow the recent literature on principal components (PC) analysis of large dimensional factor models (Bai and Ng, 2002; Bai, 2003) and covariance matrix estimation methods which are robust to spatial dependence (Kelejian and Prucha, 2007, hereafter KP). Following DM, we motivate our test statistics with assumptions on the loss differentials themselves and not on the models or methods of forecasting, as in West (1996) and GW, neither on their cross-sectional averages as in Timmermann and Zhu (2019).

We investigate the small sample properties of the tests proposed via an extensive

Monte Carlo simulation exercise. For the treatment of spatial dependence in the errors, we follow KP and use spatial heteroskedasticity and autocorrelation consistent (SHAC) estimators of the covariance matrix. In a time series framework the small sample properties of heteroskedasticity and autocorrelation consistent estimators are well known and comparison of the role of different kernel functions in the estimation performance is readily available (see Andrews, 1991). Whereas, in spatial modeling the Monte Carlo analysis on SHAC estimators is limited to only KP. Here, their analysis is extended in several dimensions, such that we consider many different combinations of time and cross-sectional dimension sizes and allow for several different kernel functions to investigate their role on small sample properties of the EPA tests.

Finally, the paper contributes also to the empirical literature. These tests are applied to compare the economic growth forecasts errors of the OECD and the IMF. We investigate the equality of accuracy for different time periods and country samples.

The remainder of the paper is as follows: In Section 2, we present our motivation for developing tests of EPA for panel data and the hypotheses of interest. In Section 3, the original time series DM test is briefly reviewed and statistics for panel tests of EPA are stated. Section 4 investigates the small sample properties of these new tests. In Section 5, the predictive ability of the OECD and IMF are compared using their economic growth forecasts. Sections 6 concludes.

2 Forecasting and Predictive Accuracy: Motivation and General Principle

2.1 Motivation

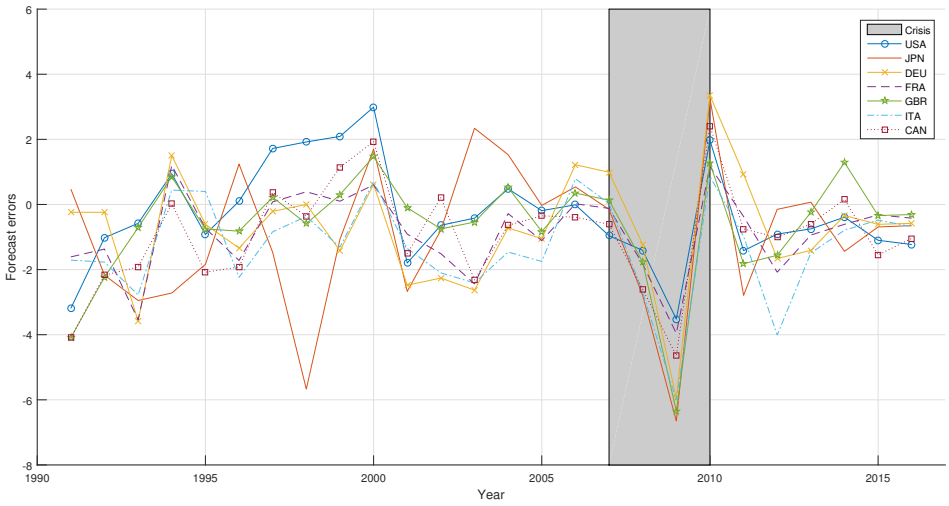
The applied literature in comparing the accuracy of two or more forecasts with panel data is typically based on the classical indicators instead of formal statistical tests. Pons (2000) compares the economic growth forecasts made by the IMF and the OECD using data from G7 countries but remained in the time series context by analyzing the forecast errors for each country separately. They used unbiasedness tests, RMSE, MAE and Theil's U for comparing the forecasts of the two institutions. Vuchelen and Gutierrez (2005) also apply country by country analysis on the OECD macroeconomic forecast errors and used statistical tests to investigate the informational content of the forecasts. Merola and Perez (2013) use data from 15 countries to compare the fiscal forecast errors of national governments and international agencies. They applied regression methods on the forecast errors to compare the biases in these forecasts but did not compare the efficiency of forecasts.

These studies suggest some stylized facts about the forecasts made by international organizations: (i) the forecast errors of different countries are affected by common global shocks, (ii) for countries which are closer to each other the comovement of the forecast errors are stronger, and (iii) international agencies make systematic errors for some particular groups of countries.

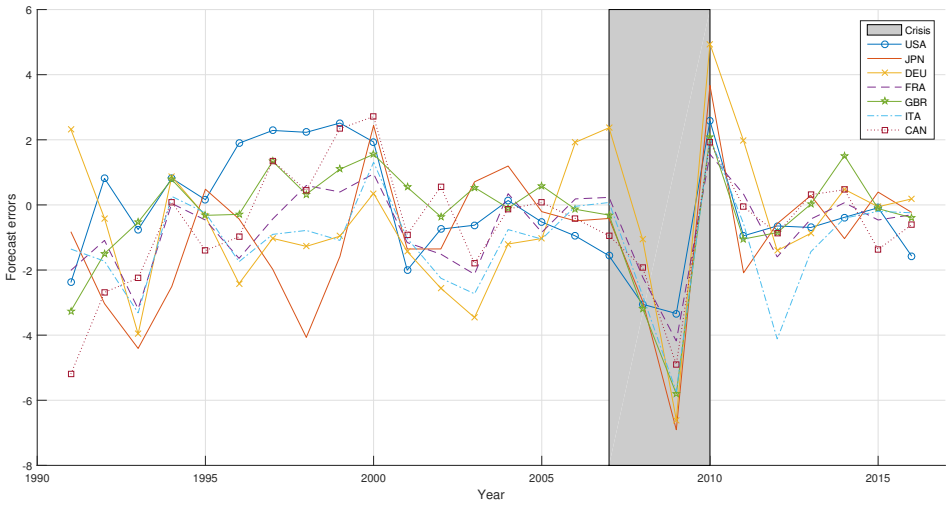
Common Factors. It is clear that during the periods like economic crisis forecasting gets more difficult. Pain et al. (2014) found that the economic growth of the OECD countries for the period 2007-2012 was systematically over-predicted by the organization, in particular for the European economies. This suggests that there are global common factors affecting the magnitude of forecast errors. Furthermore, the effect of these common shocks is heterogeneous across economies, e.g., it is higher for the European economies. Figure 1 shows the one-year ahead forecast errors by OECD and IMF between 1991-2016 for the G7 countries. In both panels, it is seen that during the crisis and recovery period the correlation between the forecast errors across countries is very high, such that they go down together during the height of the crisis and up during the recovery. In terms of modeling, this suggests a common factor structure for the forecast errors.

Spatial Interactions. The dependence between the forecast errors across countries is not the same for each group of countries. Table 1 shows the pairwise correlation coefficients between the time series given in Figure 1. The highest correlations occur between the European economies. For instance, in the case of the OECD forecast errors, FRA-ITA, DEU-FRA and DEU-ITA pairs show a correlation coefficient around 0.85. It is very high also between the two North American countries with the USA-CAN correlation coefficient being 0.85. The lowest correlation is between JPN-FRA which is followed by other pairs involving JPN. This suggests that the forecast errors are more strongly correlated for countries closer to each other. In terms of modeling, this implies that there are spatial dependencies across forecast errors.

Heterogeneity. The forecast ability of an organization is not the same for each country. In fact, the arguments in the part on common factors had already suggested that for some countries the errors can be systematically different from others. However, that was the result of time varying common factors, such as economic crisis, which may not be the only source of heterogeneity across countries. Dreher et al. (2008) find that for the case of economic growth, IMF forecasts are significantly downward biased for non-OECD countries while the bias is positive for OECD countries. They further find evidence of time-invariant country fixed effects in the forecast errors. The results on the inflation forecasts are similar. In terms of modeling, this suggests that heterogeneity



(a) OECD



(b) IMF

Figure 1: One-year ahead OECD (a) and IMF (b) economic growth forecast errors, 1991-2016, G7 countries

Table 1: Cross-country correlations in one-year ahead OECD (a) and IMF (b) economic growth forecast errors, 1991-2016, G7 countries

	USA	JPN	DEU	FRA	GBR	ITA	CAN
USA	1.000						
JPN	0.320	1.000					
DEU	0.501	0.431	1.000				
FRA	0.670	0.289	0.862	1.000			
GBR	0.753	0.495	0.581	0.712	1.000		
ITA	0.564	0.368	0.852	0.883	0.747	1.000	
CAN	0.846	0.403	0.616	0.762	0.831	0.643	1.000

(a) Computed from OECD forecasts

	USA	JPN	DEU	FRA	GBR	ITA	CAN
USA	1.000						
JPN	0.340	1.000					
DEU	0.233	0.579	1.000				
FRA	0.611	0.605	0.752	1.000			
GBR	0.711	0.611	0.400	0.731	1.000		
ITA	0.509	0.669	0.819	0.895	0.699	1.000	
CAN	0.699	0.486	0.341	0.777	0.841	0.615	1.000

(b) Computed from IMF forecasts

should be accounted for in the tests of predictive ability.

2.2 Setup in the Context of Panel Data

We are interested in τ-steps ahead observed forecast errors of a variable $y_{i,t}$, for time $t = 1, 2, \ldots, T$, units $i = 1, 2, \ldots, n$.

In terms of the analysis of forecasts using panel data, our paper is somewhat related to the work of DL, Lahiri and Sheng (2010) and Driver et al. (2013) and generalizes them in several dimensions. The focus of DL is on testing unbiasedness and efficiency of forecasts made by several different agents for the same panel unit. Their analysis is based on a three dimensional panel data regression where the dimensions are agents generating the forecasts, target years and forecast horizons. In our case, we have different target values to be forecast which are the realizations of the same variable for different units. For example, in their application they use data on forecasts of the growth rate of the USA gross national product made by 35 forecasters for 16 different years and 11 time horizons. On our side, the framework consists of forecasts made by two forecasters for the same variable (like gross national product growth rate) from different units, possibly for different horizons. The model of DL for the forecast errors can be written as

$$e_{l,t} = y_t - \widehat{y}_{l,t} = \lambda_l + f_t + u_{l,t} \tag{1}$$

where $e_{l,t}$ is the forecast error made by the forecaster l at time t for the value of τ-steps ahead variable y_t where for simplicity we assume that there is only one forecast horizon available. Notice that the target variable has only the time index. Importantly, they are interested in *the magnitude of the forecast errors*. They considered the forecaster specific bias term λ_l and the common shock variable f_t which affects the errors of each forecaster. They assumed that $u_{l,t}$ is uncorrelated over l and t but heteroskedastic over l. In our setup, we are interested in *the loss differential associated with the forecast errors* and the error component structure is generalized, such that its components enter the equation interactively.

As an example to see why this is relevant, let us assume that the loss is quadratic and the forecasts in the model (1) are unbiased such that the expectations of each component in the model are zero, i.e. $\mathrm{E}(\lambda_l) = \mathrm{E}(f_t) = \mathrm{E}(u_{l,t}) = 0$. Then the conditional expectation of the squared errors given λ_l and f_t is

$$\mathrm{E}(e_{l,t}^2 | \lambda_l, f_t) = \boldsymbol{\theta}_l' \mathbf{g}_t, \tag{2}$$

where $\theta_l' = (\lambda_l^2 + \sigma_l^2, 2\lambda_l, 1)$, $\mathbf{g}_t = (1, f_t, f_t^2)'$ and $\mathrm{E}(u_{l,t}^2) = \sigma_l^2$. Hence, the conditional

expectation function of the squared errors has a factor structure with three factors.

We generalize this setting assuming that the loss differential of the errors take the form

$$\Delta L_{i,t} = L\left(e_{1i,t}\right) - L\left(e_{2i,t}\right) = \mu_i + v_{i,t}, \tag{3}$$

$$v_{i,t} = \boldsymbol{\lambda}_i' \mathbf{f}_t + \varepsilon_{i,t}, \tag{4}$$

$$\varepsilon_{i,t} = \sum_{j=1}^{n} r_{ij} \epsilon_{j,t}, \tag{5}$$

where $L(\cdot)$ is a generic loss function, $e_{li,t}$ is the forecast error made by the forecaster $l = 1, 2$ at time t for the τ-steps ahead variable for unit $i = 1, 2, \ldots, n$, therefore they forecast $y_{li,t}$, $t = 1, 2, \ldots, T$. $\mathbf{f}_t$ is an $m \times 1$ vector of unobservable common factors and $\boldsymbol{\lambda}_i$ is the associated $m \times 1$ vector of the factor loadings. The coefficients r_{ij} are fixed but unknown elements of an $n \times n$ matrix $\mathbf{R}_n$. These elements are possibly functions of a smaller set of parameters. This is a general specification which contains as special cases all commonly used spatial processes like spatial autoregression (SAR), spatial moving average (SMA), and spatial error components (SEC) as well as higher order SAR or SMA processes. The variables $\mathbf{f}_t$ and $\epsilon_{i,t}$ are assumed to have zero mean but allowed to be autocorrelated through time. Then, assuming that μ_i are fixed parameters, a hypothesis of interest is

$$H_{0,1} : \bar{\mu} = 0, \tag{6}$$

where $\bar{\mu} = \frac{1}{T} \sum_{i=1}^{n} \mu_i$. This hypothesis state that the forecasts generated by the two agents are equally accurate on average over all $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$. It looks plausible to consider this in a micro forecasting study where the units can be seen as random draws from a population. If the researcher is not interested in the difference in predictive ability for any particular unit but the predictive ability on average, this hypothesis should be considered.

In a macro forecasting study, the differences for each unit can have a specific economic importance and may be of interest from a policy perspective. For instance, a question of interest is whether the forecasts made by agents are more accurate for a particular group of countries or all countries in the sample. In this case, the null hypothesis can be formulated such that the predictive equality holds for each unit as

$$H_{0,2} : \mathrm{E}(\Delta L_{i,t}) = \mu_i = 0, \text{ for all } i = 1, 2, \ldots, n. \tag{7}$$

Throughout the text, we assume that μ_i and factor loadings $\boldsymbol{\lambda}_i$ are fixed parameters, whereas common factors $\mathbf{f}_t$ are random variables.

3 Tests for Equal Predictive Ability for Panel Data

In this section, we present a generalization of the DM test to panel data by proposing tests of overall EPA given in (6) (Sec. 3.1) and tests of joint EPA given in (7) (Sec. 3.2), taking into account several possible forms of CD.

Let $L(\cdot)$ denote a general loss function and the loss differential between two forecast errors be $\Delta L_{i,t} = L\left(e_{1i,t}\right) - L\left(e_{2i,t}\right)$ for unit $i = 1, 2, \ldots, n$ and time $t = 1, 2, \ldots, T$. Under weak stationarity of the loss differential series, for each unit i, the asymptotic distribution of the sample mean of the loss differential series can be obtained as follows

$$\sqrt{T}\left(\Delta\bar{L}_{i,T} - \mu_i\right) \xrightarrow{D} N(0, \sigma_i^2), \tag{8}$$

where $\Delta\bar{L}_{i,T} = \frac{1}{T}\sum_{t=1}^{T}\Delta L_{i,t}$, $\mu_i = \mathrm{E}(\Delta L_{i,t})$,

$$\sigma_i^2 = \sum_{s=-\infty}^{\infty}\gamma_{v_i}(s), \tag{9}$$

with $\gamma_{v_i}(s) = \mathrm{E}(v_{i,t}v_{i,t-s})$ and $\xrightarrow{D}$ signifies convergence in distribution. The hypothesis of interest is the EPA on average

$$H_0 : E(\Delta L_{i,t}) = 0. \tag{10}$$

From (8) and (10) we derive the DM test statistic for testing the equality of forecast accuracy between the two competing series as

$$S_{i,T}^{(0)} = \frac{\Delta\bar{L}_{i,T}}{\hat{\sigma}_{i,T}/\sqrt{T}} \xrightarrow{D} N(0, 1), \tag{11}$$

where $\hat{\sigma}_{i,T}^2$ is a consistent estimate of σ_i^2. Originally DM suggested using the non-parametric variance estimator (see, for instance, Andrews, 1991) with truncated kernel to construct the variance estimates but this may result in non-positive variance estimates. [See Section 1.1 of DM and the discussions in the following subsection.] Below we allow for other kernel functions.

It is possible to relax the weak stationarity assumption and allow for nonstationary processes by considering mixing processes as in the work of GW. They prove the consistency of the test for general mixing processes and alternative hypotheses. Our generalizations of the DM test, however, are to a panel data framework.

3.1 Tests for Overall Equal Predictive Ability

Consider the sample mean loss differential over time and units:

$$\Delta\bar{L}_{n,T} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta L_{i,t}. \tag{12}$$

We provide testing procedures for overall EPA implied in (6) based on $\Delta\bar{L}_{n,T}$. Under regularity, this statistic satisfies a central limit theorem (CLT) given by

$$\sqrt{nT}(\Delta\bar{L}_{n,T} - \bar{\mu}_n)/\sigma_{n,T} \xrightarrow{D} N(0,1), \tag{13}$$

where $\bar{\mu}_n = n^{-1}\sum_{i=1}^{n}\mu_i$ and

$$\sigma^2_{n,T} = \frac{1}{nT}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T}\mathrm{E}(v_{i,t}v_{j,s}).$$

The case of no CD. Suppose that the loss differential is generated by (3) and (4) with $\boldsymbol{\lambda}_i'\mathbf{f}_t = 0$ and $r_{ij} = 0$ for every $i \neq j$. If weak stationarity assumption is satisfied for each i, a sequential application of the CLT for weakly stationary time series (see, e.g., Anderson, 1971, Theorem 7.7.8) and the CLT for independent but heterogeneous sequence (see, e.g., White, 2001, Theorem 5.10) provides the result in (13) with $\sigma^2_{n,T} = \bar{\sigma}^2_n = n^{-1}\sum_{i=1}^{n}\sigma_i^2$. The conditions for this result to be valid can be seen by writing $\sqrt{nT}(\Delta\bar{L}_{n,T} - \bar{\mu}_n)$ as $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sqrt{T}(\Delta\bar{L}_{i,T} - \mu_i)$, where $\Delta\bar{L}_{i,T} = \frac{1}{T}\sum_{t=1}^{T}\Delta L_{i,t}$. As $T \to \infty$, $\sqrt{T}(\Delta\bar{L}_{i,T} - \mu_i) \xrightarrow{D} Z_i$, where $Z_i \sim N(0,\sigma_i^2)$, under weak stationarity assumption as in (8). Then, the convergence of $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_i/\bar{\sigma}_n$, as $n \to \infty$, follows from Theorem 5.10 of White (2001), provided that $\{Z_i\}_{i=1}^{n}$ are independent as they are, $E|Z_i|^{2+\delta} < C < \infty$ for some $\delta > 0$ for all i, and $\bar{\sigma}^2_n > \delta' > 0$ for all n sufficiently large.

Suppose that we want to test hypothesis (6). We consider the test statistic

$$S^{(1)}_{n,T} = \frac{\Delta\bar{L}_{n,T}}{\hat{\bar{\sigma}}_{n,T}/\sqrt{nT}} \xrightarrow{D} N(0,1), \tag{14}$$

where $\hat{\bar{\sigma}}^2_{n,T} = n^{-1}\sum_{i=1}^{n}\hat{\sigma}^2_{i,T}$, and $\hat{\sigma}^2_{i,T}$ is a consistent estimate of σ_i^2 based on the ith time series of loss differentials

$$\hat{\sigma}^2_{i,T} = \frac{1}{T}\sum_{t,s=1}^{T}k_T\left(\frac{|t-s|}{l_T+1}\right)\Delta\tilde{L}_{i,t}\Delta\tilde{L}_{i,s}, \tag{15}$$

where $\Delta\tilde{L}_{i,t} = \Delta L_{i,t} - \Delta\bar{L}_{i,T}$ and $k_T(\cdot)$ is the time series kernel function. Under general conditions Andrews (1991) showed that $\hat{\sigma}^2_{i,T} \xrightarrow{p} \sigma^2_i$ as $T \to \infty$ with $l_T \to \infty$, $l_T = o(T)$. If the conditions implying $\hat{\sigma}^2_{i,T} \xrightarrow{p} \sigma^2_i$ are satisfied, it immediately follows that $\hat{\bar{\sigma}}^2_{n,T} - \sigma^2_{n,T} \xrightarrow{p} 0$ from which the asymptotic distribution for the test statistic given in (14) is obtained under the null hypothesis (6).

The case of WCD. Suppose that in (3) and (4), $\boldsymbol{\lambda}_i'\mathbf{f}_t = 0$ but $r_{ij} \neq 0$ for some $i \neq j$. In this case of WCD, the loss differentials $\Delta L_{i,t}$ are no longer independent across i, and therefore, the variance estimator $\hat{\bar{\sigma}}^2_{n,T}$ given above is no longer valid. Nevertheless the CLT in (13) still satisfied with

$$\sigma^2_{n,T} = \frac{1}{nT}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T}\mathbf{r}_{i.}'\boldsymbol{\gamma}_{\epsilon_i}(|t-s|)\mathbf{r}_{j.},$$

where $\boldsymbol{\gamma}_{\epsilon_i}(|t-s|) = \text{diag}[\gamma_{\epsilon_1}(|t-s|), \gamma_{\epsilon_2}(|t-s|), \ldots, \gamma_{\epsilon_n}(|t-s|)]$, $\gamma_{\epsilon_i}(s) = \text{E}(\epsilon_{i,t}\epsilon_{i,t-s})$. To see this, write $\sqrt{nT}(\Delta\bar{L}_{n,T} - \bar{\mu}_n)$ as $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sqrt{T}(\Delta\bar{L}_{i,T} - \mu_i) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{r}_{i.}'\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\epsilon}_{.t}\right)$ which follows from (5) where $\mathbf{r}_{i.} = (r_{i1}, r_{i2}, \ldots, r_{in})'$ and $\boldsymbol{\epsilon}_{.t} = (\epsilon_{1,t}, \epsilon_{2,t}, \ldots, \epsilon_{n,t})'$. Then, by the CLT for weakly stationary time series and the Cramer-Wold device (see, e.g., White, 2001, Proposition 5.1), as $T \to \infty$, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{s}_n^{-1/2}\boldsymbol{\epsilon}_{.t} \xrightarrow{D} \mathbf{Z}$, where $\mathbf{s}_n = \text{diag}(\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_n)$ and $\mathbf{Z} \sim N(\mathbf{0}, \boldsymbol{I}_n)$, under mutual independence of the components of $\boldsymbol{\epsilon}_{.t}$. Now the result follows from the application of the CLT for spatially correlated triangular arrays of Kelejian and Prucha (1998). Given that $\max_{1\leq i\leq n}\sum_{j=1}^{n}|r_{ij}| < \infty$, $\max_{1\leq j\leq n}\sum_{i=1}^{n}|r_{ij}| < \infty$, as $n \to \infty$, $\frac{1}{\sqrt{n}}\mathbf{e}_n'\mathbf{R}_n\mathbf{s}_n^{1/2}\mathbf{Z} \xrightarrow{D} N(0, \sigma^2)$ where $\mathbf{e}_n$ is an n-dimensional vector of ones and $\sigma^2 = \lim_{n\to\infty}\mathbf{e}_n'\mathbf{R}_n\mathbf{s}_n\mathbf{R}_n'\mathbf{e}_n$, hence (13) is satisfied.

For a single cross-sectional data subject to WCD, KP proposed a spatial heteroskedasticity and autocorrelation consistent (HAC) estimator of variance-covariance matrix which can be extended to give a WCD-robust estimator of $\sigma^2_{n,T}$. Such an estimator is

$$\hat{\sigma}^2_{2,n,T} = \frac{1}{nT}\sum_{i,j=1}^{n}k_S\left(\frac{d_{ij}}{d_n}\right)\sum_{t,s=1}^{T}k_T\left(\frac{|t-s|}{l_T+1}\right)\Delta\tilde{L}_{i,t}\Delta\tilde{L}_{j,s}, \tag{16}$$

leading to a test statistic as

$$S^{(2)}_{n,T} = \frac{\Delta\bar{L}_{n,T}}{\hat{\sigma}_{2,n,T}/\sqrt{nT}} \xrightarrow{D} N(0,1), \tag{17}$$

where $d_{ij} = d_{ji} \geq 0$ denotes the distance between units i and j, and d_n the threshold distance, which is an increasing function of n such that $d_n \to \infty$ as $n \to \infty$. The estimator $\hat{\sigma}^2_{2,n,T}$ is a panel data generalization of the non-parametric covariance estimator proposed by KP. It is used by Pesaran and Tosetti (2011). Moscone and Tosetti (2012,

hereafter MT) use a similar estimator with the difference being that they set $k_T(\cdot) = 1$.

Consistency of (16) follows from the arguments by MT. To see this define the space-time kernel by

$$k_{ST}\left(\frac{d_{ij}}{d_n}, \frac{|t-s|}{l_T+1}\right) = k_S\left(\frac{d_{ij}}{d_n}\right) k_T\left(\frac{|t-s|}{l_T+1}\right).$$

Consistency of the variance estimator require that $k_{ST}(x) : \mathbb{R} \to [0,1]$ satisfy (i) $k_{ST}(0) = 1$ and $k_{ST}(x) = 0$ for $|x| > 1$, (ii) $k_{ST}(x) = k_{ST}(-x)$, and (iii) $|k_{ST}(x) - 1| \leq C|x|^{\delta}$ for some $\delta \geq 1$ and $0 < C < \infty$. Then, $\hat{\sigma}^2_{2,n,T} - \sigma^2_{n,T} \xrightarrow{p} 0$ from which the asymptotic distribution for the test statistic given in (17) is obtained under the null hypothesis (6) if $\max_{1\leq i\leq n} \sum_{j=1}^{n} \mathbf{1}_{d_{ij}\leq d_n} \leq s_n$ where s_n is the number of units for which $d_{ij} \leq d_n$ and satisfies $s_n = O(n^{\kappa})$ such that $0 \leq \kappa < 0.5$ and $\sum_{j=1}^{n} |\mathbf{r}'_{j.}\mathbf{r}_{i.}| d_{ij}^{\eta} < \infty,\ \eta \geq 1$.

In this case of WCD in addition to non-parametric estimation, one can use parametric methods to estimate the covariance matrix. When the model for the spatial dependence structure of the loss differentials is correctly specified we can expect to have more powerful tests compared to the case of non-parametric estimation.

Several other covariance estimators proposed in the literature can be obtained using the formula in (16). Setting $k_T(\cdot) = 1$, together with setting $k_S(\cdot) = 1$ for each $i = j$ and $k_S(\cdot) = 0$ otherwise, gives the cluster-robust estimator proposed by Arellano (1987). As explained, setting $k_T(\cdot) = 1$ and leaving $k_S(\cdot)$ unrestricted gives the estimator proposed by MT.

The case of SCD. In the case that the generating process of the loss differential series involve common factors such that there is SCD among the units, the conditions by MT are not satisfied. This case can be expressed by setting $r_{ij} = 0$ for every $i \neq j$ in (3) and (4). A CLT as in (13) can still be obtained under general conditions with

$$\sigma^2_{n,T} = \frac{1}{nT} \sum_{i,j=1}^{n} \sum_{t,s=1}^{T} \boldsymbol{\lambda}'_i \mathrm{E}(\mathbf{f}_t \mathbf{f}'_s) \boldsymbol{\lambda}_j + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t,s=1}^{T} \mathrm{E}(\varepsilon_{i,t}\varepsilon_{i,s}).$$

We write $\sqrt{nT}(\Delta\bar{L}_{n,T} - \bar{\mu}_n)$ as $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{n}(\Delta\bar{L}_{n,t} - \bar{\mu}_n) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{n}\bar{v}_{n,t}$ where $\bar{L}_{n,t} = \frac{1}{n}\sum_{i=1}^{n} \Delta L_{i,t}$ and $\bar{v}_{n,t} = \frac{1}{n}\sum_{i=1}^{n} v_{i,t}$. Suppose that $v_{i,t}$ is α-mixing of size $r/(r-1)$ with $r > 1$ as defined by Driscoll and Kraay (1998). This implies that $\bar{v}_{n,t}$ is α-mixing of size $r/(r-1)$ as well. If $\mathrm{E}|\bar{v}_{n,t}|^r < \delta < \infty$ for some $r \geq 2$ and $\bar{\sigma}^2_{n,T} = \mathrm{Var}[T^{-1/2}\sum_{t=1}^{T} \bar{v}_{n,t}] > \delta > 0$ the CLT for dependent and heterogeneously distributed random variables (see, e.g., White, 2001, Theorem 5.20) can be applied such that $\sqrt{T}\bar{v}_{n,T}/\bar{\sigma}_{n,T} \sim N(0,1)$ for all T sufficiently large from which the result in (13) follows.

In this case, the variance estimator given in (16) can be modified by setting $k_S(\cdot) =$

1 and leaving $k_T(\cdot)$ unrestricted. This variance estimator does not require any knowledge of a distance measure between the units. Moreover, it assigns weights equal to one for all covariances, hence robust to SCD as well as WCD. The test statistic takes the form:

$$S_{n,T}^{(3)} = \frac{\Delta \bar{L}_{n,T}}{\hat{\sigma}_{3,n,T}/\sqrt{nT}} \xrightarrow{D} N(0,1), \tag{18}$$

where

$$\hat{\sigma}_{3,n,T}^2 = \frac{1}{nT} \sum_{i,j=1}^{n} \sum_{t,s=1}^{T} k_T \left(\frac{|t-s|}{l_T + 1} \right) \Delta \tilde{L}_{i,t} \Delta \tilde{L}_{j,s}. \tag{19}$$

The variance estimator (19) was proposed by Driscoll and Kraay (1998), which is valid when T is large, regardless of n finite or infinite. Consistency of the estimator follows immediately from the conditions given above except that now it is required $v_{i,t}$ to be α-mixing of size $2r/(r-1)$ with $r > 1$ and the factor loadings $\boldsymbol{\lambda}_i$ to be uniformly bounded. Then the null distribution in (18) follows.

It is known that when the number of units in the panel is close to the number of time series observations this estimator performs poorly. An alternative way to estimate the covariance matrix is to exploit the factor structure of the DGP. The PC estimation of the factor model defined by (3)-(5) is investigated by Stock and Watson (2002), Bai and Ng (2002), Bai (2003), among others. This method minimizes the sum of squared residuals $SSR = (nT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} (\Delta \tilde{L}_{i,t} - \boldsymbol{\lambda}_i' \mathbf{f}_t)^2$ subject to $\text{Var}(\mathbf{f}_t) = \boldsymbol{I}_m$. Then the solution for the estimates of the common factors, $\widehat{\mathbf{f}}_t$, are given by $\sqrt{T}$ times the first m eigenvectors of the matrix $\sum_{i=1}^{n} \boldsymbol{\Delta L}_{i.} \boldsymbol{\Delta L}_{i.}'$ with $\boldsymbol{\Delta L}_{i.} = (\Delta L_{i,1}, \Delta L_{i,2}, \ldots, \Delta L_{i,T})'$ and the factor loadings can be estimated as $\widehat{\boldsymbol{\lambda}}_i = \frac{1}{T} \sum_{t=1}^{T} \widehat{\mathbf{f}}_t \Delta \tilde{L}_{i,t}$. Then the overall EPA hypothesis can be tested using

$$S_{n,T}^{(4)} = \frac{\Delta \bar{L}_{n,T}}{\hat{\sigma}_{4,n,T}/\sqrt{nT}} \xrightarrow{D} N(0,1), \tag{20}$$

where

$$\hat{\sigma}_{4,n,T}^2 = \frac{1}{nT} \sum_{i,j=1}^{n} \sum_{t,s=1}^{T} k_T \left(\frac{|t-s|}{l_T + 1} \right) \widehat{\boldsymbol{\lambda}}_i' \widehat{\mathbf{f}}_t \widehat{\mathbf{f}}_s' \widehat{\boldsymbol{\lambda}}_j + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t,s=1}^{T} k_T \left(\frac{|t-s|}{l_T + 1} \right) \widehat{\varepsilon}_{i,t} \widehat{\varepsilon}_{i,s} \tag{21}$$

with $\widehat{\varepsilon}_{i,t} = \Delta \tilde{L}_{i,t} - \widehat{\boldsymbol{\lambda}}_i' \widehat{\mathbf{f}}_t$. The conditions under which the estimates $\widehat{\boldsymbol{\lambda}}_i'$ and $\widehat{\mathbf{f}}_t$ are consistent are given in Bai and Ng (2002). Consistency of the variance estimator (21) follows directly under these conditions together with the conditions on consistent estimation of the long-run variance as in Andrews (1991). These lead to the null distribution given in (20).

The case of both SCD and WCD. This is the most general case of the model defined by (3)-(5) with no specific restriction imposed on the parameters. Under the α-mixing conditions discussed previously, the CLT in (13) still holds with

$$\sigma^2_{n,T} = \frac{1}{nT}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T}\boldsymbol{\lambda}_i'\mathrm{E}(\mathbf{f}_t\mathbf{f}_s')\boldsymbol{\lambda}_j + \frac{1}{nT}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T}\mathbf{r}_{i\cdot}'\boldsymbol{\gamma}_{\epsilon_i}(|t-s|)\mathbf{r}_{j\cdot}.$$

The test (20) is robust to SCD because of the presence of common factors. However, it is obtained under the assumption that the residuals do not contain WCD. Under the conditions discussed previously, the test (18) is robust to the presence of both SCD and WCD but as mentioned, performs poorly when n is close to T. Another test can be obtained by using the kernel methods. We have

$$S^{(5)}_{n,T} = \frac{\Delta\bar{L}_{n,T}}{\hat{\sigma}_{5,n,T}/\sqrt{nT}} \xrightarrow{D} N(0,1), \tag{22}$$

where

$$\hat{\sigma}^2_{5,n,T} = \frac{1}{nT}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T}k_T\left(\frac{|t-s|}{l_T+1}\right)\widehat{\boldsymbol{\lambda}}_i'\widehat{\mathbf{f}}_t\widehat{\mathbf{f}}_s'\widehat{\boldsymbol{\lambda}}_j + \frac{1}{nT}\sum_{i,j=1}^{n}k_S\left(\frac{d_{ij}}{d_n}\right)\sum_{t,s=1}^{T}k_T\left(\frac{|t-s|}{l_T+1}\right)\widehat{\varepsilon}_{i,t}\widehat{\varepsilon}_{i,s}. \tag{23}$$

3.2 Tests for Joint Equal Predictive Ability

In this section we are concerned with testing the hypothesis (7), i.e., $H_0 : \mu_1 = \mu_2 = \cdots = \mu_n = 0$. The discussion is first based on large T and small n scenario. In the case of fixed n, by the CLT for weakly stationary time series and the Cramer-Wold device, the joint limiting distribution of the vector of loss differential series $\boldsymbol{\Delta}\bar{\mathbf{L}}_T = (\Delta\bar{L}_{1,T}, \Delta\bar{L}_{2,T}, \ldots, \Delta\bar{L}_{n,T})'$ is given by

$$\sqrt{T}\boldsymbol{\Omega}_n^{1/2}(\boldsymbol{\Delta}\bar{\mathbf{L}}_T - \boldsymbol{\mu}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{I}_n), \tag{24}$$

as $T \to \infty$, where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)'$,

$$\boldsymbol{\Omega}_n = \frac{1}{T}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T}\boldsymbol{h}_i\boldsymbol{h}_j'\mathrm{E}(v_{i,t}v_{j,s}),$$

with $\boldsymbol{h}_i$ being the ith column of $\boldsymbol{I}_n$.

The case of no CD. Under cross-sectional independence of the loss differential series, we have $\boldsymbol{\Omega}_n = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$ with σ_i^2 being defined in (9). Therefore, the

first test statistic considered is

$$J_{n,T}^{(1)} = T\boldsymbol{\Delta}\bar{\mathbf{L}}_T'\widehat{\boldsymbol{\Omega}}_{1,n}^{-1}\boldsymbol{\Delta}\bar{\mathbf{L}}_T \xrightarrow{D} \chi_n^2, \tag{25}$$

where $\widehat{\boldsymbol{\Omega}}_{1,n}$ is a consistent estimator of $\boldsymbol{\Omega}_n$ with diagonal elements $\hat{\sigma}_{i,T}^2$ given in (15). Consistency of the estimator $\widehat{\boldsymbol{\Omega}}_{1,n}$ follows directly from the fact that its components are consistent under the conditions, for instance, given by Andrews (1991). Hence, this test statistic is robust against arbitrary time dependence as is $S_{n,T}^{(1)}$.

The case of WCD. When the panel data exhibit WCD, $\boldsymbol{\Omega}_n$ is no longer diagonal. In the case of small n, the panel generalization of the non-parametric variance estimator of KP is not appropriate. In this case, Driscoll and Kraay (1998) estimator can be used as explained in the case of SCD given below. In the case of large n, we can still use the non-parametric estimator. A natural extension of $S_{n,T}^{(2)}$ gives the second test statistic that is robust to arbitrary time and cross sectional dependence:

$$J_{n,T}^{(2)} = T\boldsymbol{\Delta}\bar{\mathbf{L}}_T'\widehat{\boldsymbol{\Omega}}_{2,n}^{-1}\boldsymbol{\Delta}\bar{\mathbf{L}}_T \xrightarrow{D} \chi_n^2, \tag{26}$$

where

$$\widehat{\boldsymbol{\Omega}}_{2,n} = \frac{1}{T}\sum_{i,j=1}^{n} k_S\left(\frac{d_{ij}}{d_n}\right)\sum_{t,s=1}^{T} k_T\left(\frac{|t-s|}{l_T+1}\right)\boldsymbol{h}_i\boldsymbol{h}_j'\Delta\tilde{L}_{i,t}\Delta\tilde{L}_{j,s}, \tag{27}$$

with $\boldsymbol{h}_i$ being the ith column of $\boldsymbol{I}_n$.

The null distribution stated in (26) is not obvious as the consistency of the non-parametric variance estimator (27) requires large n but the test statistic has infinite variance as $n \to \infty$. Alternatively, one can use a centered and scaled version of this statistic which is asymptotically normal. This is explained below.

The case of SCD. When the loss differentials are subject to SCD, similar to the steps leading to the overall EPA test $S_{n,T}^{(3)}$, we modify the covariance estimator (27) by imposing $k_S(d_{ij}/d_n) = 1$, so that a known distance measure is not required. The test statistic is given by

$$J_{n,T}^{(3)} = T\boldsymbol{\Delta}\bar{\mathbf{L}}_T'\widehat{\boldsymbol{\Omega}}_{3,n}^{-1}\boldsymbol{\Delta}\bar{\mathbf{L}}_T \xrightarrow{D} \chi_n^2, \tag{28}$$

where

$$\widehat{\boldsymbol{\Omega}}_{3,n} = \frac{1}{T}\sum_{i,j=1}^{n}\sum_{t,s=1}^{T} k_T\left(\frac{|t-s|}{l_T+1}\right)\boldsymbol{h}_i\boldsymbol{h}_j'\Delta\tilde{L}_{i,t}\Delta\tilde{L}_{j,s}. \tag{29}$$

Although there is an advantage of using this estimator in the sense that it is robust in the case of SCD, WCD or both and it does not require a known distance measure, it has an important disadvantage. It is not of full rank even if the population variance-

covariance matrix is so. Namely, $\text{rank}(\widehat{\mathbf{\Omega}}_{3,n})$ is at most T, therefore, it is not invertible whenever $n > T$. This difficulty can be overcome by using the PC estimates of the factors and their loadings, leading to a new joint EPA test statistic as

$$J_{n,T}^{(4)} = T\mathbf{\Delta}\bar{\mathbf{L}}_T'\widehat{\mathbf{\Omega}}_{4,n}^{-1}\mathbf{\Delta}\bar{\mathbf{L}}_T \xrightarrow{D} \chi_n^2, \tag{30}$$

where

$$\widehat{\mathbf{\Omega}}_{4,n} = \widehat{\mathbf{\Lambda}}\left[\frac{1}{T}\sum_{t,s=1}^{T} k_T\left(\frac{|t-s|}{l_T+1}\right)\widehat{\mathbf{f}}_t\widehat{\mathbf{f}}_s'\right]\widehat{\mathbf{\Lambda}}' + \widehat{\mathbf{\Sigma}}_{1,n}, \tag{31}$$

and

$$\widehat{\mathbf{\Sigma}}_{1,n} = \frac{1}{T}\sum_{i=1}^{n}\sum_{t,s=1}^{T} k_T\left(\frac{|t-s|}{l_T+1}\right)\text{diag}(\boldsymbol{h}_i)\widehat{\varepsilon}_{i,t}\widehat{\varepsilon}_{i,s}, \tag{32}$$

with $\widehat{\mathbf{\Lambda}} = (\widehat{\boldsymbol{\lambda}}_1, \widehat{\boldsymbol{\lambda}}_2, \ldots, \widehat{\boldsymbol{\lambda}}_n)'$.

Once more the null distribution stated in (30) is not obvious because PC estimates of the common factors require large n but the test statistic has infinite variance as $n \to \infty$. Again, one can use a centered and scaled version of this statistic which is asymptotically normal which is explained below.

The case of both SCD and WCD. As in the previous section, a joint test statistic which is robust to both common factors and spatial dependence can be obtained as

$$J_{n,T}^{(5)} = T\mathbf{\Delta}\bar{\mathbf{L}}_T'\widehat{\mathbf{\Omega}}_{5,n}^{-1}\mathbf{\Delta}\bar{\mathbf{L}}_T \xrightarrow{D} \chi_n^2, \tag{33}$$

where

$$\widehat{\mathbf{\Omega}}_{5,n} = \widehat{\mathbf{\Lambda}}\left[\frac{1}{T}\sum_{t,s=1}^{T} k_T\left(\frac{|t-s|}{l_T+1}\right)\widehat{\mathbf{f}}_t\widehat{\mathbf{f}}_s'\right]\widehat{\mathbf{\Lambda}}' + \widehat{\mathbf{\Sigma}}_{2,n}, \tag{34}$$

and

$$\widehat{\mathbf{\Sigma}}_{2,n} = \frac{1}{T}\sum_{i,j=1}^{n} k_S\left(\frac{d_{ij}}{d_n}\right)\sum_{t,s=1}^{T} k_T\left(\frac{|t-s|}{l_T+1}\right)\boldsymbol{h}_i\boldsymbol{h}_j'\widehat{\varepsilon}_{i,t}\widehat{\varepsilon}_{j,s}. \tag{35}$$

Below, a centered and scaled version of this test statistic is proposed.

Standardized test statistics. When n grows with T, it is clear that the limiting chi-square distribution is not meaningful and in this case a standardized chi-square test can be used. For the tests given above, these standardized statistics are

$$Z_{n,T}^{(g)} = \frac{J_{n,T}^{(g)} - n}{\sqrt{2n}} \xrightarrow{D} N(0,1), \quad g = 1, \ldots, 5, \tag{36}$$

where the stated asymptotic standard normal distribution holds under the particular assumption of each statistics $J_{n,T}^{(g)}$, $g = 1, \ldots, 5$.

4 Monte Carlo Study

To investigate the small sample properties of the test statistics given above, a set of Monte Carlo simulations are conducted. 2000 samples from each DGP described below for the dimensions of $T \in \{10, 20, 30, 50, 100\}$, $n \in \{10, 20, 30, 50, 100, 200\}$ are generated. All tests are applied for two nominal size values, 1% and 5%.

4.1 Design

Two different DGPs are considered to explore the effect of WCD and SCD on the performance of the tests. DGP1 contains only spatial dependence. In this case, for each of the cross-sections or units ($i = 1, 2, \ldots, n$), two independent forecast error series $(e_{1i,t}, e_{2i,t})$ are generated using two spatial AR(1) processes defined as

$$\zeta_{l,it} = \rho \sum_{j=1}^{n} w_{ij}\zeta_{l,jt} + u_{l,it}, \quad \text{where,} \quad u_{l,it} \sim N(0,1), \quad l = 1, 2, \tag{37}$$

where w_{ij} is the element of the spatial matrix $\mathbf{W}_n$ in row i and column j. A rook-type spatial weight matrix is used. To make the power results across different levels of spatial dependence comparable, the unconditional variance of the forecast error series $e_{l,it}$, $l = 1, 2$, is held fixed for each panel. To generate such series we proceed as follows: First the spatial AR(1) processes is written in matrix form as

$$\boldsymbol{\zeta}_{l,.t} = \mathbf{S}_n \mathbf{u}_{l,.t}, \tag{38}$$

where $\boldsymbol{\zeta}_{l,.t} = (\zeta_{l,1t}, \zeta_{l,2t}, \ldots, \zeta_{l,nt})'$, $\mathbf{u}_{l,.t} = (\mathbf{u}_{l,1t}, \mathbf{u}_{l,2t}, \ldots, \mathbf{u}_{l,nt})'$, $\mathbf{S}_n = (\boldsymbol{I}_n - \rho\mathbf{W}_n)^{-1}$. Then, the forecast error series are generated according to

$$\mathbf{e}_{l,.t} = \mathbf{P}_n \boldsymbol{\zeta}_{l,.t}, \tag{39}$$

where $\mathbf{e}_{l,.t} = (e_{l,1t}, e_{l,2t}, \ldots, e_{l,nt})'$ and $\mathbf{P}_n$ has elements $p_{ij} = \sqrt{1/s_{2,ij}}$ if $i = j$ and zeros otherwise, with $s_{2,ij}$ being the i, j-th element of $\mathbf{S}_n\mathbf{S}_n'$. It can now be shown that all diagonal elements of the matrix $\mathbf{P}_n\mathbf{S}_n\mathbf{S}_n'\mathbf{P}_n'$ equal to 1.

Three different spatial AR(1) parameters are considered: $\rho = 0$, 0.5 and 0.9, which are selected to represent no spatial dependence, low spatial dependence and high spatial dependence cases, respectively. As error series are generated for each unit as white noises, it is implicitly assumed that these are one-step ahead forecasts. In the computation of the test statistics, it is assumed that this is correctly specified. Hence, in the computation of the long-run variances, covariances through time are not taken into

account. In this DGP a quadratic loss function is used.

DGP2 contains common factors as well as spatial dependence. In this case, following GW we directly generate the loss differential, hence we do not rely on a specific loss function. This is given by

$$\Delta L_{i,t} = \phi(\mu_i + \lambda_{1i} f_{1t} + \lambda_{2i} f_{2t} + \varepsilon_{i,t}). \tag{40}$$

To investigate the size properties we set $\mu_i = 0$ for each $i = 1, 2, \ldots, n$ and generate factor loadings as

$$\lambda_{1i}, \lambda_{2i} \sim N(1, 0.2). \tag{41}$$

The common factors are formed by

$$f_{1t}, f_{2t} \sim N(0, 1), \tag{42}$$

hence, they do not incorporate autocorrelation. The error series $\varepsilon_{i,t}$ are generated in the same spirit as in (39). Hence, we have 3 cases for DGP2 also, namely no spatial dependence, low spatial dependence and high spatial dependence. We finally set $\phi = 1/3.4$ to control for the variance of the loss differential series.

We explore the power properties of various tests under two different alternative hypothesis. The first one is the homogeneous alternative and the second one is the heterogeneous alternative. For DGP1 with homogeneous alternative, we generate a third set of forecast error series as $e_{3i,t} = \sqrt{1.2} e_{2i,t}$ and report the results from testing the equality of forecast accuracy of $e_{1i,t}$ and $e_{3i,t}$. In the heterogeneous scenario, we generate the third series according to $e_{3i,t} = \sqrt{\theta_i} e_{2i,t}$ where

$$\theta_i \sim U(0.6, 1.4). \tag{43}$$

Similarly, in the case of DGP2, we set $\mu_i = 1.2$ for each $i = 1, 2, \ldots, n$ in the case of homogeneous alternative and

$$\mu_i \sim U(-0.4, 0.4), \tag{44}$$

in the case of heterogeneous alternative. It is important to note that in the case of heterogeneous alternative, the unconditional expectations of the loss differentials are equal to zero in all DGPs. Hence, the overall EPA hypothesis holds. On the other hand, for each unit, the expected value of the loss differential is different from zero. Therefore, the joint EPA hypothesis does not hold. As a consequence, we expect the overall EPA tests not to have increasing power against the heterogeneous alternative whereas joint EPA tests to be consistent.

As we generate one-step ahead forecasts, the time series kernel $k_T(\cdot) = 1$ if $t = s$ and $k_T(\cdot) = 0$ otherwise. Spatial interactions between units are created with a rook-type weight matrix where two units in the panel are neighbors if their Euclidean distance is less than or equal to one. In the computation of the spatial kernel $k_S(\cdot)$, we used these distances and we implemented several different kernel functions used frequently in time series literature. These are truncated, Bartlett, Parzen, Tukey and Quadratic Spectral. Following KP, we set the spatial kernel bandwidth to $\lfloor n^{1/4} \rfloor$.

For the tests $S^{(2)}_{n,T}$ and $J^{(2)}_{n,T}$ the results from all these kernels are reported. For $S^{(4)}_{n,T}$ and $J^{(4)}_{n,T}$ we consider different possibilities concerning the number of common factors. First, we consider extracting 2 common factors from the panel which is the correct number of factors in DGP2. Second, we consider the possibility of a number of common factors which expands with the number of units in the panel. As shown by Sarafidis and Wansbeek (2012), all common spatial processes can be written in the form of a factor model of factor dimension n. Hence, it is interesting to see if choosing a number of common factors growing with the number of units will help to deal with spatial dependence. For these tests, we chose the number of common factors as $\lfloor n^{1/4} \rfloor$. We implement the tests $S^{(5)}_{n,T}$ and $J^{(5)}_{n,T}$ only with Bartlett kernel. In the following, to save space, we report only the results for the case of low spatial dependence for both DGPs. Full set of results are available from the authors upon request.

4.2 Size Properties

The results on the size properties of tests with DGP1 under low spatial dependence are given in Table 2. As expected, the non-robust test $S^{(1)}_{n,T}$ has size distortions which do not disappear with increases in the sample size. For the smallest samples with $T = 10$ and $n = 10$, this particular setting provides an empirical size of 3.65% and 12.35% for 1% nominal level for the test, respectively. The kernel robust test $S^{(2)}_{n,T}$ greatly improves the size properties over the non-robust test even with the smallest samples. The truncated kernel performs the best with small n. For instance, when $n = 20$ and $T = 30$, the empirical size of the test with truncated and Bartlett kernels are 1.65% and 2.9% for 1% nominal level, respectively. In this case Parzen kernel appears to be the least liable choice. The performance of the cluster-robust test improves with T but in small samples it does not overperform the non-robust test strongly. Whenever we have $T \geq 50$ it is nearly correctly sized. As before, factor robust tests $S^{(4)}_{n,T}$ and $S^{(5)}_{n,T}$ are undersized when $T = 10$ for the nominal level 1%. However, they are performing very well to correct for spatial dependence for larger sample sizes.

Typically, the performance of the joint EPA tests falls with n and improves with T.

Table 2: Size - DGP1 (No Common Factors), Low Spatial Dependence

			Overall EPA Test												Joint EPA Test										
			1% Nominal Size					5% Nominal Size							1% Nominal Size					5% Nominal Size				
Option	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100
	$S_{n,T}^{(1)}$	10	3.65	4.60	4.15	3.80	4.65	12.35	12.10	11.65	12.40	13.25	$J_{n,T}^{(1)}$	10	7.55	2.40	2.60	1.55	1.95	19.35	10.90	8.35	8.15	6.85
		20	3.40	3.90	4.70	3.40	3.70	11.35	11.55	11.55	10.15	11.15		20	10.25	3.85	2.85	2.00	1.35	24.85	12.35	10.25	7.30	6.05
		30	3.25	3.40	3.80	3.45	3.20	10.20	11.05	11.40	10.90	10.25		30	13.75	4.90	3.30	2.05	1.55	32.35	14.85	11.15	8.05	6.95
		50	4.10	2.85	2.50	2.90	3.45	12.35	10.40	9.75	11.10	9.80		50	22.35	6.05	3.95	2.55	1.80	45.70	18.80	13.65	8.95	6.60
		100	3.80	3.55	4.50	4.45	3.65	11.75	10.65	12.25	12.15	11.25		100	43.60	10.50	4.80	3.15	2.00	67.15	28.50	17.10	11.20	7.95
		200	2.90	3.45	3.20	2.50	3.65	9.95	10.60	10.25	9.60	11.15		200	73.30	20.65	9.65	4.85	2.20	88.65	43.90	26.25	14.65	8.85
Truncated	$S_{n,T}^{(2)}$	10	1.65	1.45	1.20	1.20	1.35	6.35	6.60	6.05	6.15	6.45	$J_{n,T}^{(2)}$	10	25.50	13.80	8.25	4.25	2.05	35.90	24.30	18.00	11.85	8.05
		20	1.80	1.35	1.65	1.25	0.85	6.80	5.70	6.60	5.55	5.40		20	33.05	36.40	36.70	13.30	3.75	39.25	45.70	50.65	28.45	11.70
		30	1.65	0.85	0.95	0.85	1.25	5.75	5.10	6.55	4.50	5.10		30	30.50	39.60	46.30	21.35	4.75	34.20	46.85	58.25	37.95	15.85
		50	1.90	0.90	0.70	1.20	0.90	6.50	4.95	4.45	5.20	4.95		50	35.70	46.50	56.55	31.90	7.30	39.50	51.50	64.50	51.85	18.30
		100	0.95	0.95	1.20	0.80	1.10	5.90	4.85	5.80	5.40	4.85		100	39.40	54.25	71.50	63.15	13.90	41.20	57.10	75.40	79.55	33.05
		200	1.20	1.15	0.90	0.70	0.80	4.80	5.35	4.95	4.05	5.20		200	34.20	40.85	52.25	73.75	65.85	35.75	42.10	54.00	75.50	83.75
Bartlett	$S_{n,T}^{(2)}$	10	3.65	4.60	4.15	3.80	4.65	12.35	12.10	11.65	12.40	13.25	$J_{n,T}^{(2)}$	10	7.55	2.40	2.60	1.55	1.95	19.35	10.90	8.35	8.15	6.85
		20	2.20	2.20	2.90	2.05	1.90	8.45	8.45	8.70	7.40	8.00		20	15.80	5.40	3.05	1.65	1.30	33.60	14.40	11.00	7.50	5.60
		30	2.25	2.00	2.10	2.05	1.80	7.65	7.25	8.85	7.05	7.65		30	24.55	6.40	3.75	2.00	1.20	44.10	18.55	12.25	7.90	6.70
		50	2.55	1.70	1.15	2.00	2.05	8.80	7.30	6.70	8.15	7.05		50	40.70	8.40	4.55	2.55	1.55	62.05	24.10	16.30	9.65	6.30
		100	1.50	1.80	1.95	1.85	1.60	7.00	6.85	7.25	6.95	6.55		100	87.65	29.65	11.45	4.70	2.30	95.80	51.55	30.00	14.25	8.35
		200	1.45	1.65	1.25	1.10	1.40	5.75	7.05	6.20	5.90	6.90		200	99.90	70.75	34.95	11.90	3.55	100.00	87.40	58.85	30.65	12.40
Parzen	$S_{n,T}^{(2)}$	10	3.65	4.60	4.15	3.80	4.65	12.35	12.10	11.65	12.40	13.25	$J_{n,T}^{(2)}$	10	7.55	2.40	2.60	1.55	1.95	19.35	10.90	8.35	8.15	6.85
		20	3.00	2.90	3.75	2.80	3.15	10.00	9.80	9.85	8.80	9.50		20	10.50	3.65	2.35	1.45	1.25	24.25	11.70	9.40	7.15	5.45
		30	2.60	2.60	3.15	2.70	2.25	8.50	8.50	10.60	8.90	8.60		30	14.25	4.55	2.75	1.60	1.35	33.00	14.25	10.60	7.10	6.45
		50	3.40	2.35	1.70	2.50	2.70	10.85	8.75	7.90	9.75	8.65		50	22.80	5.65	3.45	2.15	1.45	46.00	17.95	12.60	8.50	5.85
		100	2.05	1.95	2.60	2.15	1.70	7.45	7.55	8.30	8.15	7.45		100	67.10	15.60	5.45	3.25	1.65	84.25	36.50	21.00	11.35	7.15
		200	1.90	1.85	1.70	1.35	1.75	6.20	7.80	6.60	6.20	7.80		200	95.65	38.15	16.10	6.30	2.10	98.80	62.60	36.75	18.40	9.30
Tukey	$S_{n,T}^{(2)}$	10	3.65	4.60	4.15	3.80	4.65	12.35	12.10	11.65	12.40	13.25	$J_{n,T}^{(2)}$	10	7.55	2.40	2.60	1.55	1.95	19.35	10.90	8.35	8.15	6.85
		20	2.20	2.20	2.85	2.20	1.90	8.40	8.50	8.65	7.50	8.05		20	14.75	5.20	2.90	1.60	1.35	32.85	14.25	10.85	7.40	5.40
		30	2.30	2.05	2.10	2.05	1.90	7.60	7.25	8.90	7.10	7.65		30	22.85	5.95	3.45	1.85	1.15	41.90	17.75	12.10	7.55	6.60
		50	2.50	1.70	1.15	2.05	2.05	8.90	7.35	6.75	8.20	7.05		50	37.50	7.85	4.30	2.50	1.50	59.35	23.25	15.60	9.40	6.35
		100	1.35	1.65	1.75	1.65	1.50	6.70	6.50	6.95	6.70	6.40		100	92.50	34.25	13.75	5.55	2.55	97.95	56.25	33.35	16.15	9.00
		200	1.25	1.55	1.15	1.00	1.15	5.40	6.55	5.90	5.45	6.65		200	100.00	76.05	38.90	13.05	3.80	100.00	90.25	62.45	32.20	13.30
QS	$S_{n,T}^{(2)}$	10	3.10	4.05	3.65	3.45	3.95	11.40	11.40	10.50	11.45	12.15	$J_{n,T}^{(2)}$	10	7.30	2.15	2.40	1.40	1.70	18.85	10.15	7.55	7.10	6.45
		20	2.00	1.75	2.50	1.70	1.30	7.85	7.45	7.60	6.65	6.85		20	30.40	7.50	4.65	2.20	1.40	50.60	20.50	13.80	9.35	6.15
		30	1.95	1.50	1.75	1.70	1.60	6.40	6.00	7.95	6.30	7.00		30	45.60	11.90	5.80	3.10	1.55	65.10	27.20	16.60	9.70	7.85
		50	2.25	1.25	1.05	1.65	1.40	7.95	6.30	5.70	6.90	6.50		50	69.15	18.25	8.30	4.00	2.05	85.20	38.05	23.45	12.20	7.30
		100	1.20	1.25	1.55	1.40	1.30	6.35	5.95	6.50	6.10	5.95		100	99.55	60.05	29.20	10.35	3.70	99.80	79.30	51.55	24.35	12.00
		200	1.25	1.45	1.10	0.95	1.00	5.15	6.05	5.45	4.75	6.05		200	100.00	98.60	77.55	32.60	7.70	100.00	99.85	90.50	56.30	21.60
	$S_{n,T}^{(3)}$	10	3.25	2.25	1.50	1.15	1.15	8.45	7.10	6.00	5.75	6.00	$J_{n,T}^{(3)}$	10			40.40	18.40	7.20	2.95			56.30	33.05	18.30	10.15
		20	3.60	2.15	1.90	1.35	0.90	9.65	7.40	6.80	5.95	5.55		20				81.95	38.95	9.55				89.20	55.60	22.95
		30	3.20	1.85	1.50	1.35	1.20	8.65	7.00	7.45	5.45	5.05		30					83.60	27.55					92.05	45.50
		50	4.10	1.55	0.90	1.40	1.10	11.05	7.40	5.50	5.75	5.65		50						83.05						91.00
		100	3.50	2.40	1.85	1.40	1.25	9.85	7.25	7.00	6.50	5.35		100												
		200	4.25	1.90	1.20	0.85	1.25	9.80	7.35	5.80	4.95	5.60		200												
Fixed	$S_{n,T}^{(4)}$	10	0.60	0.75	0.80	0.85	0.40	4.65	5.05	4.05	4.20	4.15	$J_{n,T}^{(4)}$	10	0.85	1.25	1.70	2.15	2.55	5.55	6.30	7.30	9.30	9.55
		20	0.70	1.75	1.20	1.30	0.80	5.25	6.25	6.65	5.85	6.00		20	0.55	0.65	1.00	1.70	1.50	5.60	5.45	6.40	7.30	7.00
		30	0.95	1.50	1.45	1.55	1.20	5.00	6.30	6.75	6.25	6.40		30	0.75	0.80	1.50	1.20	1.25	6.00	5.85	5.75	6.90	7.00
		50	1.65	1.50	1.15	1.90	2.35	7.90	6.75	6.45	6.80	7.15		50	2.25	1.45	0.85	1.10	1.10	8.70	4.95	5.95	5.90	5.60
		100	1.70	2.15	2.95	3.10	2.60	7.55	7.40	9.40	9.90	8.80		100	3.80	1.10	1.10	1.10	1.35	12.90	7.15	5.50	5.65	6.20
		200	2.10	2.30	2.15	2.10	3.00	7.80	8.80	8.60	8.05	10.15		200	10.25	2.15	1.75	1.15	1.15	21.75	9.30	7.20	5.55	5.25
Expanding	$S_{n,T}^{(4)}$	10	1.35	1.80	1.30	1.20	1.15	5.95	7.30	5.80	5.70	6.25	$J_{n,T}^{(4)}$	10	0.50	0.55	0.95	1.55	1.50	3.95	4.90	5.90	6.50	6.75
		20	0.70	1.75	1.20	1.30	0.80	5.25	6.25	6.65	5.85	6.00		20	0.55	0.65	1.00	1.70	1.50	5.60	5.45	6.40	7.30	7.00
		30	0.95	1.50	1.45	1.55	1.20	5.00	6.30	6.75	6.25	6.40		30	0.75	0.80	1.50	1.20	1.25	6.00	5.85	5.75	6.90	7.00
		50	1.65	1.50	1.15	1.90	2.35	7.90	6.75	6.45	6.80	7.15		50	2.25	1.45	0.85	1.10	1.10	8.70	4.95	5.95	5.90	5.60
		100	1.25	1.90	2.10	2.35	2.25	6.50	7.15	9.20	9.10	8.20		100	8.70	2.80	1.25	1.50	1.65	18.80	9.55	6.40	6.55	6.45
		200	1.90	1.95	1.80	1.70	2.85	7.25	8.00	7.90	7.65	9.65		200	16.70	4.90	1.75	1.10	1.10	27.20	13.60	8.35	5.75	5.70
Fixed	$S_{n,T}^{(5)}$	10	0.50	0.65	0.90	0.75	0.50	4.15	5.00	4.00	4.40	4.85	$J_{n,T}^{(5)}$	10	4.50	1.40	1.05	1.45	2.00	8.80	5.50	5.10	7.20	7.85
		20	0.35	1.00	0.80	1.10	0.45	3.90	5.05	5.50	5.10	5.30		20	5.95	8.90	3.70	1.15	1.25	7.90	14.60	9.85	6.60	5.75
		30	0.55	0.65	0.90	0.95	0.85	3.90	4.75	5.50	4.70	4.85		30	6.30	12.20	9.95	1.20	0.85	7.65	17.25	17.00	7.40	5.70
		50	0.75	0.60	0.50	0.95	0.75	5.65	4.35	3.80	4.95	5.30		50	7.00	13.60	15.80	1.95	0.85	8.75	18.95	24.45	7.35	5.55
		100	0.60	0.75	0.95	0.95	0.90	4.90	5.15	5.60	5.70	4.70		100	6.40	18.90	21.75	2.20	0.90	7.20	22.35	31.40	7.65	4.95
		200	0.80	0.80	0.55	0.75	0.80	4.30	5.10	4.45	3.70	5.45		200	4.25	12.05	19.40	34.65	1.35	4.75	13.60	22.30	43.15	7.10
Expanding	$S_{n,T}^{(5)}$	10	0.70	1.00	0.80	0.65	0.60	4.15	5.40	4.15	4.75	5.10	$J_{n,T}^{(5)}$	10	5.30	1.05	1.35	1.15	1.15	9.50	5.05	5.85	5.60	6.65
		20	0.35	1.00	0.80	1.10	0.45	3.90	5.05	5.50	5.10	5.30		20	5.95	8.90	3.70	1.15	1.25	7.90	14.60	9.85	6.60	5.75
		30	0.55	0.65	0.90	0.95	0.85	3.90	4.75	5.50	4.70	4.85		30	6.30	12.20	9.95	1.20	0.85	7.65	17.25	17.00	7.40	5.70
		50	0.75	0.60	0.50	0.95	0.75	5.65	4.35	3.80	4.95	5.30		50	7.00	13.60	15.80	1.95	0.85	8.75	18.95	24.45	7.35	5.55
		100	0.55	0.60	0.95	0.95	1.00	4.60	4.85	5.55	5.40	4.90		100	5.80	17.65	21.50	1.90	0.85	6.85	21.40	30.80	8.70	5.60
		200	0.65	0.95	0.55	0.75	0.80	4.70	5.35	4.35	3.85	5.45		200	3.80	11.85	18.60	33.20	1.85	4.05	13.45	21.50	41.55	7.55

In the case of joint tests $J_{n,T}^{(2)}$, truncated kernel is not found to be the best performance option anymore for small n. In this case Parzen kernel looks like the most liable alternative. Interestingly, the performance of the test does not fall quickly with n for a large but fixed T, in the cases of Bartlett, Parzen and Tukey kernels. For instance, when $T = 100$ and $n = 200$ the empirical sizes of these tests are 3.55%, 2.10% and 3.80% for 1% nominal level, respectively. For quadratic spectral kernel this number is 7.7%. The performance of factor robust tests $J_{n,T}^{(4)}$ and $J_{n,T}^{(5)}$ are most satisfying, as before. One exception is the $J_{n,T}^{(5)}$ in large n and moderate T cases.

As we expect the conclusions change dramatically in the case of DGP2 for which the results are given in Table 3. In this case all overall EPA tests which do not take common factors in account are grossly oversized. As in the previous cases the cluster-robust test $S_{n,T}^{(3)}$ performs well when T is large and n is small. $S_{n,T}^{(4)}$ and $S_{n,T}^{(5)}$ are even more undersized in this setting for small T. However, they are correctly sized for moderate to large T.

In the case of joint tests the kernel robust test $S_{n,T}^{(2)}$ has an improving performance for fixed n and growing T. It reaches the correct size especially with the Bartlett and quadratic spectral kernels. Among the factor robust tests $J_{n,T}^{(4)}$ looks like the better choice. This is interesting as $J_{n,T}^{(5)}$ is robust to both WCD and SCD while $J_{n,T}^{(4)}$ controls only the SCD.

The results on no spatial dependence and high spatial dependence as well as the standardized joint EPA tests (36), which are available upon request, can be summarized as follows: as mentioned earlier in the case of no spatial dependence with DGP1 robust and non-robust tests are correctly sized, at least in moderate to large sample sizes. In DGP2 the factor robust tests $S_{n,T}^{(5)}$ and $J_{n,T}^{(5)}$ perform well even in small samples whereas the non-robust and kernel-robust tests are grossly oversized and their performance does not improve, as expected. The standardized tests have slightly less size distortions than the chi-square tests in the case of large n. However, these improvements are limited.

4.3 Size Adjusted Power Properties

The size adjusted power results of the tests for the homogeneous alternative hypothesis are given in Tables 4-5, whereas the results for heterogeneous alternative are given in Tables 6-7.

Table 4 reports the results for DGP1 with low spatial dependence. In this case all overall EPA tests have satisfactory power even with small samples. The results show that the kernel-robust tests $S_{n,T}^{(2)}$ do not improve the size adjusted power over the non-robust test $S_{n,T}^{(1)}$. The differences are very small and even there are cases of non-robust

Table 3: Size - DGP2 (Common Factors): Low Spatial Dependence

			Overall EPA Test											Joint EPA Test									
			1% Nominal Size					5% Nominal Size						1% Nominal Size					5% Nominal Size				
Option	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100
	$S_{n,T}^{(1)}$	10	34.70	34.35	32.15	29.75	36.30	47.40	46.85	44.60	42.55	48.25	$J_{n,T}^{(1)}$	10	17.05	13.75	10.70	10.60	11.15	24.05	20.05	17.05	15.00	16.70
		20	50.20	49.35	48.90	46.90	46.55	61.60	61.95	60.10	58.35	56.80		20	22.25	17.75	15.65	13.70	13.35	28.60	22.85	20.90	17.85	17.85
		30	56.95	54.50	54.95	56.15	55.05	67.05	64.10	64.70	65.45	65.25		30	25.20	18.15	18.90	17.65	15.55	29.80	23.30	23.80	22.25	19.35
		50	64.00	64.05	63.05	65.45	64.85	71.65	72.05	71.85	74.00	73.00		50	28.50	22.50	20.20	20.55	17.05	33.85	26.55	23.50	24.10	20.50
		100	74.75	73.25	74.30	73.30	73.90	79.70	80.50	81.50	79.00	80.00		100	32.45	26.25	24.75	21.55	21.15	35.80	30.05	27.80	24.20	24.55
		200	79.80	80.95	81.30	81.25	82.00	84.70	85.15	85.60	85.60	86.55		200	36.00	29.25	25.65	26.20	24.40	38.75	31.75	27.95	28.10	26.85
Truncated	$S_{n,T}^{(2)}$	10	13.95	13.00	10.95	10.60	11.35	24.00	24.55	21.45	19.85	23.75	$J_{n,T}^{(2)}$	10	5.65	1.95	1.20	0.90	0.90	8.85	3.60	2.70	2.55	1.85
		20	11.85	10.10	9.40	7.80	8.45	22.00	21.45	19.60	18.15	18.00		20	12.00	8.20	7.90	7.20	7.55	14.70	10.60	10.05	9.15	9.45
		30	17.60	13.95	15.20	14.65	13.20	28.90	26.05	28.30	26.55	24.60		30	12.45	9.30	8.10	9.35	7.80	14.60	11.40	10.25	10.85	9.30
		50	26.35	24.70	23.00	23.90	21.25	39.50	36.30	34.65	36.45	33.85		50	12.25	8.15	7.10	7.10	6.80	14.05	9.65	8.10	8.30	7.80
		100	37.75	37.90	37.60	34.20	35.10	50.65	49.60	49.85	47.70	47.55		100	14.50	9.95	8.85	8.75	7.85	15.80	11.10	10.10	9.30	8.80
		200	42.80	40.95	40.80	40.15	41.40	53.45	52.20	52.05	52.15	53.85		200	13.45	12.35	9.30	8.35	9.30	14.60	13.20	9.90	9.10	10.00
Bartlett	$S_{n,T}^{(2)}$	10	34.70	34.35	32.15	29.75	36.30	47.40	46.85	44.60	42.55	48.25	$J_{n,T}^{(2)}$	10	17.05	13.75	10.70	10.60	11.15	24.05	20.05	17.05	15.00	16.70
		20	28.50	28.50	27.20	24.90	24.60	42.15	41.10	41.20	38.10	36.65		20	11.50	6.15	5.05	2.80	2.95	17.15	10.45	7.85	5.25	5.35
		30	36.75	34.60	35.90	35.00	32.65	49.35	46.40	47.60	47.25	46.50		30	14.10	6.05	6.05	4.35	4.20	19.65	9.60	9.30	7.15	5.85
		50	47.30	44.05	43.45	45.25	43.55	57.75	55.60	55.75	57.35	56.70		50	17.80	7.85	7.20	5.35	3.90	22.90	11.65	9.90	7.85	6.15
		100	53.70	53.15	53.70	51.05	52.10	64.00	63.10	62.60	61.65	63.65		100	23.65	8.85	5.45	4.60	3.30	28.40	11.80	8.00	6.70	5.05
		200	59.70	59.10	58.75	60.05	60.65	68.20	69.10	68.25	68.25	69.30		200	36.10	12.00	7.25	6.00	3.70	42.05	14.85	9.60	8.25	5.05
Parzen	$S_{n,T}^{(2)}$	10	34.70	34.35	32.15	29.75	36.30	47.40	46.85	44.60	42.55	48.25	$J_{n,T}^{(2)}$	10	17.05	13.75	10.70	10.60	11.15	24.05	20.05	17.05	15.00	16.70
		20	39.05	38.25	38.80	35.60	34.40	51.95	51.20	50.85	48.75	48.05		20	14.00	9.45	7.85	5.85	5.90	19.05	13.95	12.15	9.10	9.10
		30	47.40	44.50	46.05	45.55	44.20	58.20	55.25	55.95	57.45	56.15		30	16.70	9.45	9.95	8.05	6.70	21.05	13.70	13.15	12.15	10.30
		50	56.45	53.95	54.45	55.85	55.75	64.70	64.85	63.75	66.15	65.50		50	19.55	12.15	11.30	9.75	7.70	23.65	15.35	13.45	12.90	10.40
		100	60.15	60.30	59.45	57.85	59.20	69.00	68.10	68.40	67.35	68.80		100	20.55	10.90	7.95	6.65	5.90	25.20	14.00	10.45	8.60	7.85
		200	67.05	67.55	67.35	67.25	68.45	74.25	75.85	75.50	74.70	75.75		200	26.40	12.90	9.55	9.05	6.45	30.00	15.25	11.65	11.00	8.10
Tukey	$S_{n,T}^{(2)}$	10	34.70	34.35	32.15	29.75	36.30	47.40	46.85	44.60	42.55	48.25	$J_{n,T}^{(2)}$	10	17.05	13.75	10.70	10.60	11.15	24.05	20.05	17.05	15.00	16.70
		20	29.70	29.60	28.85	26.65	25.75	43.30	43.05	42.15	39.75	37.95		20	13.25	6.80	5.60	3.15	3.45	18.95	11.40	8.70	6.20	5.85
		30	38.35	36.20	37.20	36.55	34.60	50.70	47.30	48.95	48.85	47.55		30	17.10	7.35	7.05	5.25	4.80	22.75	11.40	10.70	7.85	6.75
		50	48.70	45.80	45.35	46.80	44.90	58.85	56.90	57.20	58.30	57.95		50	21.55	9.50	8.35	6.20	4.65	27.15	13.50	11.60	8.90	7.10
		100	53.75	53.30	53.85	51.15	52.30	64.10	63.15	62.65	61.65	63.80		100	47.45	15.80	10.00	7.60	5.70	56.35	21.55	13.80	10.10	8.90
		200	60.35	60.05	59.90	60.85	61.75	68.80	69.70	69.10	69.10	70.30		200	76.85	32.05	17.20	12.90	8.05	80.40	40.15	22.25	16.25	10.60
QS	$S_{n,T}^{(2)}$	10	31.10	31.80	28.35	26.50	32.15	43.75	43.25	40.15	38.45	44.95	$J_{n,T}^{(2)}$	10	14.55	11.15	8.40	7.90	7.85	19.90	16.35	13.85	12.45	13.85
		20	23.40	23.40	21.75	19.70	19.65	36.40	35.45	35.35	32.35	31.35		20	19.10	7.75	5.40	3.10	2.70	28.05	13.35	10.10	6.00	6.10
		30	31.55	28.80	31.00	29.65	27.45	44.90	41.65	43.35	42.05	40.60		30	24.70	7.90	7.35	4.50	4.05	33.65	13.90	11.75	7.75	7.05
		50	42.45	39.65	38.15	39.65	38.00	54.45	51.25	51.80	53.15	51.40		50	34.35	11.90	9.05	6.25	4.40	43.45	17.95	13.10	10.05	7.40
		100	49.65	48.45	49.20	46.55	46.50	59.75	59.70	58.75	56.95	58.60		100	77.10	23.30	13.35	8.30	5.40	85.20	32.40	19.60	11.60	9.65
		200	55.00	54.35	53.40	54.25	55.70	64.50	63.85	63.35	64.15	65.60		200	99.95	58.35	27.05	15.70	8.15	99.95	69.80	36.70	20.85	11.75
	$S_{n,T}^{(3)}$	10	4.70	2.40	1.35	1.60	1.30	10.15	8.55	6.05	6.35	5.65	$J_{n,T}^{(3)}$	10		43.25	20.40	8.95	3.85		56.35	34.50	18.90	10.60
		20	3.35	2.25	2.00	1.20	1.35	9.35	7.35	6.70	5.15	5.90		20			83.65	37.15	9.90			90.55	53.30	23.40
		30	3.90	1.95	1.85	1.50	1.10	10.30	6.50	7.75	5.80	5.40		30				83.40	26.55				90.90	45.95
		50	3.00	2.30	1.40	1.35	1.10	8.60	6.85	6.35	5.45	4.95		50					83.40					91.60
		100	3.10	1.80	1.20	1.80	0.60	8.45	7.05	5.25	5.65	4.70		100										
		200	4.10	2.05	1.50	1.45	1.60	10.00	7.30	6.00	6.50	5.00		200										
Fixed	$S_{n,T}^{(4)}$	10	0.10	0.75	0.50	0.95	1.10	5.00	5.15	4.40	4.85	4.80	$J_{n,T}^{(4)}$	10	1.50	2.80	3.65	3.35	4.20	7.65	10.25	10.75	10.20	11.80
		20	0.20	0.95	0.85	0.75	1.00	3.65	4.90	4.75	4.25	4.90		20	1.85	2.75	3.10	3.35	3.60	7.95	8.95	10.15	10.30	10.70
		30	0.20	0.50	0.85	0.65	0.85	4.20	4.45	5.50	4.60	4.85		30	2.35	2.40	2.35	2.85	2.80	8.60	8.60	8.95	9.15	9.55
		50	0.15	0.95	0.40	0.70	0.70	3.50	4.45	4.50	4.30	4.65		50	3.70	3.00	3.00	3.00	3.50	11.45	9.25	9.70	9.30	10.60
		100	0.00	0.55	0.30	1.20	0.55	3.85	4.40	3.95	4.95	4.15		100	6.45	3.20	3.65	3.30	3.30	15.95	9.55	10.15	10.75	10.55
		200	0.40	0.50	0.70	0.85	1.40	4.45	4.30	4.65	5.45	4.35		200	11.50	4.95	4.40	2.75	2.55	22.70	12.90	12.15	8.55	9.90
Expanding	$S_{n,T}^{(4)}$	10	0.10	0.55	0.50	0.95	1.05	4.85	4.95	4.10	4.60	4.65	$J_{n,T}^{(4)}$	10	1.70	2.55	3.05	3.95	3.70	7.30	9.65	10.35	9.40	10.50
		20	0.20	0.95	0.85	0.75	1.00	3.65	4.90	4.75	4.25	4.90		20	1.85	2.75	3.10	3.35	3.60	7.95	8.95	10.15	10.30	10.70
		30	0.20	0.50	0.85	0.65	0.85	4.20	4.45	5.50	4.60	4.85		30	2.35	2.40	2.35	2.85	2.80	8.60	8.60	8.95	9.15	9.55
		50	0.15	0.95	0.40	0.70	0.70	3.50	4.45	4.50	4.30	4.65		50	3.70	3.00	3.00	3.00	3.50	11.45	9.25	9.70	9.30	10.60
		100	0.00	0.55	0.30	1.20	0.55	3.90	4.40	3.95	4.95	4.15		100	11.75	3.40	4.10	3.60	3.75	23.70	11.40	11.85	10.45	11.45
		200	0.40	0.50	0.70	0.85	1.40	4.45	4.30	4.65	5.45	4.35		200	22.25	6.50	4.60	3.25	2.90	32.95	16.80	13.45	9.25	10.10
Fixed	$S_{n,T}^{(5)}$	10	0.10	0.75	0.45	0.90	1.10	4.95	5.15	4.35	4.80	4.70	$J_{n,T}^{(5)}$	10	5.80	4.00	2.55	2.75	3.00	10.20	11.15	10.15	9.55	10.90
		20	0.20	0.95	0.85	0.70	0.95	3.60	4.85	4.60	4.20	4.85		20	6.55	9.80	9.55	11.40	7.85	9.05	14.95	16.15	19.35	17.60
		30	0.20	0.50	0.80	0.55	0.80	4.10	4.40	5.30	4.45	4.70		30	5.35	11.55	12.35	11.20	5.20	7.10	15.90	19.05	20.15	13.75
		50	0.15	0.95	0.40	0.70	0.65	3.40	4.30	4.20	4.15	4.55		50	6.85	13.75	18.15	16.00	3.80	8.80	18.50	24.45	28.55	10.95
		100	0.00	0.45	0.30	1.15	0.45	3.75	4.30	3.90	4.90	4.10		100	6.25	16.00	24.50	20.25	1.45	7.10	19.35	31.10	32.45	7.55
		200	0.35	0.50	0.70	0.85	1.35	4.35	4.30	4.60	5.40	4.35		200	4.95	11.05	20.45	30.40	11.35	5.30	12.40	22.75	35.90	25.85
Expanding	$S_{n,T}^{(5)}$	10	0.05	0.40	0.50	0.85	0.95	4.55	4.85	3.85	4.50	4.55	$J_{n,T}^{(5)}$	10	6.75	4.85	4.20	3.65	3.40	11.80	10.90	11.65	10.45	10.05
		20	0.20	0.95	0.85	0.70	0.95	3.60	4.85	4.60	4.20	4.85		20	6.55	9.80	9.55	11.40	7.85	9.05	14.95	16.15	19.35	17.60
		30	0.20	0.50	0.80	0.55	0.80	4.10	4.40	5.30	4.45	4.70		30	5.35	11.55	12.35	11.20	5.20	7.10	15.90	19.05	20.15	13.75
		50	0.15	0.95	0.40	0.70	0.65	3.40	4.30	4.20	4.15	4.55		50	6.85	13.75	18.15	16.00	3.80	8.80	18.50	24.45	28.55	10.95
		100	0.00	0.50	0.30	1.15	0.45	3.80	4.35	3.90	4.90	4.10		100	5.65	16.55	22.55	17.10	1.80	6.70	20.60	28.70	30.70	8.80
		200	0.40	0.50	0.70	0.85	1.35	4.35	4.30	4.60	5.40	4.35		200	4.35	10.25	18.75	30.45	10.70	4.65	11.80	21.60	36.05	25.70

having better performance over the robust tests. For instance, for the smallest samples where $T = 10$ and $n = 10$, $S^{(2)}_{n,T}$ with truncated kernel has a size adjusted power of 11.65% in 5% nominal level whereas this number equals to 10.9% for $S^{(1)}_{n,T}$. When $n \geq 50$, $S^{(1)}_{n,T}$ outperforms the robust test. An important result in this experiment is that the size adjusted power of all overall EPA tests are increasing with either n or T.

Table 4: Size Adjusted Power - DGP1 (No Common Factors): Low Spatial Dependence & Homogeneous Alternative

			Overall EPA Test												Joint EPA Test											
			1% Nominal Size						5% Nominal Size						1% Nominal Size						5% Nominal Size					
Option	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100
	$S^{(1)}_{n,T}$	10	2.95	4.55	9.10	18.30	38.60	10.90	16.95	22.30	39.05	60.35	$J^{(1)}_{n,T}$	10	1.30	2.10	2.80	8.65	18.75	6.30	9.20	13.20	19.45	37.40
		20	5.85	14.00	21.85	41.25	78.15	16.30	32.75	43.25	66.45	90.65		20	1.80	3.70	5.75	12.95	35.95	7.10	10.80	16.55	32.10	62.15
		30	8.00	25.05	38.20	62.45	92.20	27.15	47.70	58.20	82.75	98.35		30	2.25	3.90	8.25	17.05	52.40	8.05	13.60	17.95	37.45	73.10
		50	18.70	45.45	69.45	87.45	99.85	38.00	67.00	86.65	96.65	100.00		50	1.70	4.15	11.25	22.00	74.55	9.15	17.05	28.15	49.55	89.45
		100	42.75	78.05	92.05	99.65	100.00	64.35	92.50	97.65	99.95	100.00		100	2.65	9.60	23.00	52.50	95.90	8.85	24.95	43.80	70.25	99.05
		200	78.90	98.70	100.00	100.00	100.00	93.35	99.85	100.00	100.00	100.00		200	4.05	15.55	36.30	81.75	100.00	13.20	35.70	59.90	92.60	100.00
Truncated	$S^{(2)}_{n,T}$	10	3.05	4.50	9.50	18.50	39.55	11.65	16.50	22.80	37.80	61.10	$J^{(2)}_{n,T}$	10	0.55	0.30	0.90	0.70	8.95	3.25	4.55	7.05	11.95	24.10
		20	5.45	13.00	18.55	38.60	78.85	15.50	31.45	42.90	64.75	90.55		20	0.45	0.35	0.45	3.00	16.50	1.80	2.70	4.00	14.65	38.60
		30	7.50	25.35	38.55	61.80	91.70	25.55	46.95	57.60	82.90	98.35		30	0.35	0.70	0.90	1.55	26.75	3.75	2.90	4.35	15.35	51.15
		50	14.90	44.30	70.50	86.35	99.80	37.10	66.40	85.40	96.25	100.00		50	0.40	0.85	0.95	2.45	39.90	2.15	2.95	4.35	18.85	66.95
		100	42.55	78.90	91.10	99.60	100.00	63.75	92.40	97.70	99.95	100.00		100	0.40	0.50	0.60	6.60	71.80	2.60	2.20	3.15	29.65	88.95
		200	76.80	98.35	100.00	100.00	100.00	92.60	99.85	100.00	100.00	100.00		200	0.40	0.35	0.55	0.25	91.25	1.85	2.30	3.15	2.85	98.20
Bartlett	$S^{(2)}_{n,T}$	10	2.95	4.55	9.10	18.30	38.60	10.90	16.95	22.30	39.05	60.35	$J^{(2)}_{n,T}$	10	1.30	2.10	2.80	8.65	18.75	6.30	9.20	13.20	19.45	37.40
		20	5.00	12.60	21.45	40.20	78.75	16.45	32.85	43.60	65.70	90.90		20	1.30	2.65	5.60	9.65	28.65	6.75	8.95	14.90	25.65	54.40
		30	7.50	25.95	38.35	62.80	91.60	26.85	48.20	58.70	82.65	98.35		30	1.55	2.25	6.05	13.35	44.50	7.90	12.05	15.95	30.55	66.60
		50	17.15	45.70	69.65	86.70	99.85	38.05	67.10	86.25	96.35	100.00		50	1.95	3.70	6.35	16.95	61.80	7.45	14.25	24.35	40.00	82.40
		100	43.05	78.60	91.75	99.70	100.00	63.90	92.30	97.55	99.95	100.00		100	1.90	6.10	12.05	33.30	86.70	8.75	17.95	31.25	56.05	95.80
		200	78.45	98.60	100.00	100.00	100.00	92.95	99.85	100.00	100.00	100.00		200	2.35	8.90	19.10	63.75	99.65	9.70	24.80	44.60	81.95	99.95
Parzen	$S^{(2)}_{n,T}$	10	2.95	4.55	9.10	18.30	38.60	10.90	16.95	22.30	39.05	60.35	$J^{(2)}_{n,T}$	10	1.30	2.10	2.80	8.65	18.75	6.30	9.20	13.20	19.45	37.40
		20	5.45	13.80	21.75	41.10	79.10	16.45	32.70	43.10	66.20	90.80		20	1.35	3.40	5.70	11.75	31.40	6.65	10.20	16.20	30.50	57.95
		30	7.55	25.25	38.20	62.55	92.05	26.70	47.95	58.55	83.00	98.35		30	1.85	3.05	6.40	14.70	49.00	8.05	12.10	17.10	34.05	70.70
		50	18.15	45.50	70.20	87.40	99.85	37.80	66.85	86.45	96.65	100.00		50	2.35	3.55	9.10	20.70	68.75	8.15	15.55	25.40	44.55	86.45
		100	43.75	78.05	91.75	99.65	100.00	64.30	92.40	97.60	99.95	100.00		100	2.00	6.95	13.90	37.90	89.45	9.05	17.95	36.55	60.40	96.55
		200	78.95	98.60	100.00	100.00	100.00	93.05	99.85	100.00	100.00	100.00		200	2.65	9.35	24.55	68.60	99.80	10.60	28.15	49.10	84.90	100.00
Tukey	$S^{(2)}_{n,T}$	10	2.95	4.55	9.10	18.30	38.60	10.90	16.95	22.30	39.05	60.35	$J^{(2)}_{n,T}$	10	1.30	2.10	2.80	8.65	18.75	6.30	9.20	13.20	19.45	37.40
		20	5.20	12.75	21.85	40.20	78.90	16.45	32.75	43.50	65.75	90.90		20	1.35	2.70	5.15	9.50	28.30	6.60	8.85	15.40	25.80	54.80
		30	7.25	25.90	38.60	62.90	91.65	26.65	48.10	59.40	82.75	98.35		30	1.80	2.50	5.90	13.80	44.55	7.90	12.20	16.20	31.20	66.65
		50	17.65	45.50	69.80	86.70	99.85	38.00	67.15	86.30	96.45	100.00		50	1.85	3.60	6.55	17.10	62.20	7.70	14.15	24.40	40.80	82.60
		100	43.20	78.45	91.70	99.70	100.00	63.90	92.35	97.55	99.95	100.00		100	1.90	5.30	10.35	30.85	84.40	7.95	17.05	31.00	55.10	94.95
		200	78.30	98.65	100.00	100.00	100.00	92.95	99.85	100.00	100.00	100.00		200	1.45	7.45	18.55	60.70	99.50	9.30	23.50	43.55	80.75	99.95
QS	$S^{(2)}_{n,T}$	10	3.10	4.50	9.15	18.75	39.05	11.45	16.75	22.45	39.10	60.35	$J^{(2)}_{n,T}$	10	1.50	2.15	2.35	8.50	17.15	6.10	9.10	12.25	19.45	36.20
		20	5.10	12.70	19.95	38.65	78.20	15.80	32.55	43.20	65.50	90.85		20	1.50	2.40	3.90	8.00	25.65	5.75	8.20	12.60	22.50	49.50
		30	8.00	26.20	38.25	63.00	91.45	25.90	48.05	58.55	83.05	98.35		30	1.35	2.05	4.90	11.00	40.45	6.65	10.80	14.40	28.50	61.65
		50	16.85	46.05	69.95	86.60	99.85	37.85	67.25	86.10	96.30	100.00		50	1.60	3.00	4.35	13.20	55.85	7.40	12.30	21.65	36.80	78.70
		100	43.20	78.70	91.55	99.70	100.00	63.90	92.25	97.70	99.95	100.00		100	1.95	5.65	8.55	26.40	81.40	7.40	15.00	25.85	50.75	93.20
		200	78.15	98.60	100.00	100.00	100.00	92.85	99.85	100.00	100.00	100.00		200	1.20	6.60	14.10	54.30	98.75	7.70	18.65	36.85	75.10	99.95
	$S^{(3)}_{n,T}$	10	2.55	4.60	7.00	17.40	38.80	10.30	15.70	22.40	35.90	60.60	$J^{(3)}_{n,T}$	10			1.55	1.35	3.15	7.40			6.10	8.55	11.50	23.00
		20	4.20	12.60	18.40	36.05	77.80	15.80	29.80	43.60	63.80	89.50		20				1.70	3.75	13.65				6.85	12.00	36.75
		30	7.70	20.40	34.10	58.40	91.70	22.90	41.75	57.55	81.20	98.25		30					2.90	17.15					12.00	41.95
		50	10.75	37.50	66.65	85.40	99.70	30.70	63.00	84.10	95.70	99.95		50						22.70						46.30
		100	30.85	64.80	89.25	99.55	100.00	57.35	89.75	96.50	99.90	100.00		100												
		200	54.25	95.95	99.95	100.00	100.00	83.10	99.50	100.00	100.00	100.00		200												
Fixed	$S^{(4)}_{n,T}$	10	1.80	5.55	6.75	12.30	34.60	8.55	15.20	22.30	34.70	58.45	$J^{(4)}_{n,T}$	10	0.80	1.00	1.30	2.75	5.10	4.65	6.35	6.70	9.30	17.10
		20	4.15	7.40	16.80	29.65	72.85	15.60	28.85	38.90	59.95	87.45		20	1.05	1.40	2.55	3.45	12.00	5.50	7.30	7.80	13.35	30.80
		30	6.00	14.25	28.30	51.70	89.35	21.85	41.50	54.05	79.85	97.30		30	1.20	2.25	2.20	6.00	20.25	5.40	7.25	10.70	17.15	42.45
		50	9.10	31.85	63.00	78.45	99.35	27.25	61.85	82.45	95.25	100.00		50	0.80	1.50	4.10	8.85	33.60	5.30	8.20	11.75	22.00	57.80
		100	26.25	66.50	86.45	99.30	100.00	56.70	87.90	96.35	99.90	100.00		100	1.40	3.25	4.50	13.15	38.05	5.20	8.70	13.90	28.65	58.50
		200	59.35	95.25	99.95	100.00	100.00	87.55	99.55	100.00	100.00	100.00		200	1.35	2.60	4.55	14.95	24.20	4.50	8.10	13.20	26.25	36.55
Expanding	$S^{(4)}_{n,T}$	10	1.60	3.70	6.10	15.25	31.85	10.15	14.10	22.25	32.75	54.75	$J^{(4)}_{n,T}$	10	0.65	2.15	2.05	3.15	6.90	5.20	6.00	8.10	12.70	20.95
		20	4.15	7.40	16.80	29.65	72.85	15.60	28.85	38.90	59.95	87.45		20	1.05	1.40	2.55	3.45	12.00	5.50	7.30	7.80	13.35	30.80
		30	6.00	14.25	28.30	51.70	89.35	21.85	41.50	54.05	79.85	97.30		30	1.20	2.25	2.20	6.00	20.25	5.40	7.25	10.70	17.15	42.45
		50	9.10	31.85	63.00	78.45	99.35	27.25	61.85	82.45	95.25	100.00		50	0.80	1.50	4.10	8.85	33.60	5.30	8.20	11.75	22.00	57.80
		100	21.80	65.70	85.90	99.15	100.00	56.15	88.45	96.45	99.90	100.00		100	0.70	1.50	2.60	7.35	26.25	3.55	5.20	8.70	17.95	45.05
		200	45.30	94.10	99.95	100.00	100.00	83.45	99.35	100.00	100.00	100.00		200	1.00	1.25	3.10	7.65	14.05	3.10	4.00	7.80	14.20	24.50
Fixed	$S^{(5)}_{n,T}$	10	1.60	4.65	6.40	13.60	35.90	9.00	15.40	22.45	34.70	57.65	$J^{(5)}_{n,T}$	10	0.50	0.65	1.55	2.35	4.90	3.35	5.45	7.05	9.55	16.95
		20	4.40	9.05	19.90	32.05	73.45	16.35	28.75	41.70	61.95	88.35		20	0.45	0.50	0.60	2.40	9.80	2.55	3.15	5.40	10.20	29.50
		30	7.40	17.90	29.05	54.35	91.05	21.15	42.40	54.10	80.45	97.70		30	0.65	0.60	0.20	2.95	14.65	3.30	3.55	3.80	9.60	34.30
		50	11.50	36.10	65.95	83.85	99.60	30.00	64.15	84.10	95.75	100.00		50	0.35	0.35	0.30	1.85	18.70	2.40	2.60	3.25	12.70	42.05
		100	27.55	72.30	89.30	99.55	100.00	57.60	89.25	96.85	99.90	100.00		100	1.30	0.35	0.45	2.55	17.35	3.25	3.15	3.70	11.85	36.35
		200	61.20	96.85	99.95	100.00	100.00	89.80	99.65	100.00	100.00	100.00		200	0.40	0.15	0.75	0.55	6.90	2.40	2.95	3.25	2.80	14.15
Expanding	$S^{(5)}_{n,T}$	10	2.55	6.10	8.55	19.10	39.05	12.20	16.70	25.40	37.60	58.20	$J^{(5)}_{n,T}$	10	0.60	3.20	3.95	6.95	12.15	4.45	13.85	13.70	19.05	27.40
		20	7.00	11.55	23.05	36.40	75.80	18.95	32.25	45.15	64.75	90.05		20	0.70	0.55	0.65	12.65	24.40	4.50	3.55	13.45	27.35	49.45
		30	10.05	23.10	34.60	58.65	92.75	24.70	46.20	58.70	82.50	98.00		30	0.95	0.70	0.45	16.90	39.15	5.60	4.45	5.65	34.40	62.00
		50	15.70	41.65	70.80	86.40	99.75	35.00	67.50	86.25	96.75	100.00		50	0.85	0.65	0.65	20.05	57.45	4.40	4.20	5.70	47.15	77.25
		100	32.50	76.05	90.90	99.70	100.00	60.10	91.15	97.30	99.95	100.00		100	0.75	0.55	0.70	25.80	55.05	4.70	4.55	5.20	50.55	74.40
		200	60.40	97.20	100.00	100.00	100.00	89.40	99.60	100.00	100.00	100.00		200	1.00	1.05	0.80	0.55	39.65	5.25	4.35	4.15	5.15	55.60

This is not the case for overall EPA tests. In small samples, size adjusted power for the joint EPA tests are lower and they increase only in the dimension of T. For the test $J^{(1)}_{n,T}$ when $T = 10$ and $n = 10$ the size adjusted power is as low as 1.3% and 6.3% for 1% and 5% nominal levels, respectively. As before, the size adjusted power of the CD robust tests are lower but the differences are not big.

Table 5 gives the results on the case of common factors, e.g. DGP2 with low spatial dependence. As in this case we found that the non-robust tests $S^{(1)}_{n,T}$ and $S^{(2)}_{n,T}$ are highly oversized we focus on the rest of the results. An interesting result in this case is that the size adjusted power of the robust overall tests do not increase with n. All tests have similar performance, including the cluster-robust test $S^{(3)}_{n,T}$ which has slightly smaller size adjusted power in small samples. As T increases it quickly reaches the level of other robust tests $S^{(4)}_{n,T}$ and $S^{(5)}_{n,T}$.

Table 5: Size Adjusted Power - DGP2 (Common Factors): Low Spatial Dependence & Homogeneous Alternative

			Overall EPA Test												Joint EPA Test									
			1% Nominal Size					5% Nominal Size							1% Nominal Size					5% Nominal Size				
Option	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100
	$S^{(1)}_{n,T}$	10	39.85	78.30	95.55	99.75	100.00	65.55	90.70	99.00	99.95	100.00	$J^{(1)}_{n,T}$	10	39.95	85.30	97.60	99.90	100.00	69.20	94.55	99.65	100.00	100.00
		20	48.05	77.35	94.75	99.75	100.00	69.35	93.35	99.15	100.00	100.00		20	50.35	83.25	97.95	100.00	100.00	73.15	96.90	99.80	100.00	100.00
		30	45.15	85.65	96.20	99.80	100.00	71.70	95.60	99.15	100.00	100.00		30	40.65	90.00	98.70	100.00	100.00	75.85	98.05	99.90	100.00	100.00
		50	51.50	82.25	98.15	99.80	100.00	74.45	95.60	99.40	100.00	100.00		50	53.30	87.80	99.40	99.95	100.00	79.75	98.60	99.95	100.00	100.00
		100	51.20	88.35	98.35	99.90	100.00	76.30	96.10	99.65	100.00	100.00		100	56.25	94.40	99.80	100.00	100.00	81.35	98.85	99.95	100.00	100.00
		200	47.25	86.50	97.45	99.95	100.00	70.25	96.50	99.75	100.00	100.00		200	49.00	93.80	99.50	100.00	100.00	75.90	98.55	100.00	100.00	100.00
Truncated	$S^{(2)}_{n,T}$	10	33.10	75.80	94.65	99.55	100.00	63.45	90.25	98.85	99.95	100.00	$J^{(2)}_{n,T}$	10	29.25	75.50	95.90	99.65	99.95	60.05	93.00	98.85	99.95	99.95
		20	39.80	70.55	93.60	99.55	100.00	63.65	91.90	99.00	100.00	100.00		20	1.45	1.65	2.50	3.95	19.65	12.70	29.05	52.20	71.55	84.90
		30	29.90	82.95	95.85	99.80	100.00	65.65	94.95	99.00	100.00	100.00		30	2.20	1.75	2.95	6.90	11.35	13.15	18.75	34.30	54.35	77.60
		50	38.65	75.20	97.55	99.80	100.00	70.25	94.35	99.35	99.95	100.00		50	2.10	2.30	4.95	6.05	44.55	13.55	28.05	48.70	73.10	84.20
		100	46.25	84.80	97.70	99.90	100.00	71.75	96.05	99.65	100.00	100.00		100	2.30	3.00	2.65	3.60	10.75	12.75	26.00	44.30	67.25	83.00
		200	34.80	83.85	96.15	99.95	100.00	65.70	95.40	99.60	100.00	100.00		200	2.10	2.40	3.00	7.75	10.75	11.30	17.10	33.25	53.60	70.45
Bartlett	$S^{(2)}_{n,T}$	10	39.85	78.30	95.55	99.75	100.00	65.55	90.70	99.00	99.95	100.00	$J^{(2)}_{n,T}$	10	39.95	85.30	97.60	99.90	100.00	69.20	94.55	99.65	100.00	100.00
		20	42.95	74.55	93.80	99.70	100.00	65.90	92.60	99.05	100.00	100.00		20	48.30	85.90	99.15	100.00	100.00	74.50	98.35	100.00	100.00	100.00
		30	36.05	84.15	96.25	99.80	100.00	67.75	94.90	99.05	100.00	100.00		30	36.70	95.15	99.70	100.00	100.00	77.25	99.25	99.95	100.00	100.00
		50	45.65	78.60	97.95	99.80	100.00	72.05	94.90	99.35	99.95	100.00		50	53.45	92.55	99.95	100.00	100.00	83.15	99.85	100.00	100.00	100.00
		100	47.15	84.80	97.80	99.90	100.00	73.00	96.00	99.65	100.00	100.00		100	62.70	99.50	99.95	100.00	100.00	91.70	100.00	100.00	100.00	100.00
		200	38.15	84.60	96.55	99.95	100.00	66.70	95.65	99.65	100.00	100.00		200	41.10	100.00	100.00	100.00	100.00	89.60	100.00	100.00	100.00	100.00
Parzen	$S^{(2)}_{n,T}$	10	39.85	78.30	95.55	99.75	100.00	65.55	90.70	99.00	99.95	100.00	$J^{(2)}_{n,T}$	10	39.95	85.30	97.60	99.90	100.00	69.20	94.55	99.65	100.00	100.00
		20	45.00	76.45	94.10	99.70	100.00	67.60	92.85	99.10	100.00	100.00		20	47.00	83.15	98.55	100.00	100.00	73.15	97.25	99.85	100.00	100.00
		30	41.35	84.30	96.45	99.80	100.00	69.15	95.15	99.20	100.00	100.00		30	38.50	91.75	99.20	100.00	100.00	75.45	98.45	99.95	100.00	100.00
		50	48.10	80.10	98.00	99.80	100.00	73.05	95.25	99.35	100.00	100.00		50	51.40	87.55	99.75	100.00	100.00	80.30	99.20	100.00	100.00	100.00
		100	47.95	85.80	97.90	99.90	100.00	74.20	95.95	99.65	100.00	100.00		100	58.75	98.85	99.95	100.00	100.00	88.20	99.85	100.00	100.00	100.00
		200	40.95	85.05	97.00	99.95	100.00	67.20	95.70	99.70	100.00	100.00		200	48.80	99.40	100.00	100.00	100.00	85.40	100.00	100.00	100.00	100.00
Tukey	$S^{(2)}_{n,T}$	10	39.85	78.30	95.55	99.75	100.00	65.55	90.70	99.00	99.95	100.00	$J^{(2)}_{n,T}$	10	39.95	85.30	97.60	99.90	100.00	69.20	94.55	99.65	100.00	100.00
		20	43.35	74.85	93.75	99.70	100.00	66.20	92.60	99.05	100.00	100.00		20	45.75	87.30	99.25	100.00	100.00	74.30	98.30	100.00	100.00	100.00
		30	36.60	84.25	96.35	99.80	100.00	67.90	94.95	99.10	100.00	100.00		30	32.20	95.10	99.65	100.00	100.00	75.75	99.25	99.95	100.00	100.00
		50	46.40	78.65	97.95	99.80	100.00	72.05	95.00	99.35	100.00	100.00		50	44.70	92.30	99.85	100.00	100.00	80.40	99.60	99.95	100.00	100.00
		100	47.30	84.75	97.80	99.90	100.00	73.05	95.95	99.65	100.00	100.00		100	28.10	99.75	100.00	100.00	100.00	88.60	99.90	100.00	100.00	100.00
		200	39.05	84.70	96.75	99.95	100.00	66.90	95.60	99.65	100.00	100.00		200	3.60	99.45	100.00	100.00	100.00	63.20	99.55	100.00	100.00	100.00
QS	$S^{(2)}_{n,T}$	10	38.65	78.15	95.20	99.70	100.00	64.50	90.35	99.00	99.95	100.00	$J^{(2)}_{n,T}$	10	38.45	84.75	97.70	99.90	100.00	69.10	94.70	99.75	100.00	100.00
		20	42.30	73.20	93.90	99.70	100.00	65.80	92.35	99.00	100.00	100.00		20	51.20	89.70	99.75	100.00	100.00	77.70	99.15	100.00	100.00	100.00
		30	34.35	84.40	96.25	99.80	100.00	67.50	94.95	99.05	100.00	100.00		30	37.60	97.50	99.90	100.00	100.00	79.55	99.70	100.00	100.00	100.00
		50	43.95	77.00	97.95	99.80	100.00	71.65	94.55	99.35	99.95	100.00		50	59.80	97.50	100.00	100.00	100.00	88.30	100.00	100.00	100.00	100.00
		100	46.35	84.65	97.80	99.90	100.00	72.65	96.00	99.65	100.00	100.00		100	72.95	99.90	100.00	100.00	100.00	96.15	100.00	100.00	100.00	100.00
		200	37.10	84.45	96.50	99.95	100.00	66.35	95.65	99.65	100.00	100.00		200	60.15	100.00	100.00	100.00	100.00	96.60	100.00	100.00	100.00	100.00
	$S^{(3)}_{n,T}$	10	28.90	70.90	93.85	99.45	100.00	56.55	89.35	98.65	99.95	100.00	$J^{(3)}_{n,T}$	10		41.20	90.75	99.80	100.00		77.30	98.00	100.00	100.00
		20	35.00	69.65	92.55	99.60	100.00	62.10	91.50	98.95	100.00	100.00		20			75.10	100.00	100.00			94.15	100.00	100.00
		30	21.70	79.25	95.10	99.80	100.00	63.70	94.50	98.80	100.00	100.00		30				100.00	100.00				100.00	100.00
		50	33.60	73.75	97.50	99.75	100.00	67.95	93.90	99.30	99.95	100.00		50					100.00					100.00
		100	41.95	83.20	96.95	99.85	100.00	68.95	95.55	99.65	100.00	100.00		100										
		200	29.00	83.05	95.75	99.95	100.00	63.60	95.25	99.50	100.00	100.00		200										
Fixed	$S^{(4)}_{n,T}$	10	29.40	71.75	94.05	99.55	100.00	57.25	89.50	98.65	99.95	100.00	$J^{(4)}_{n,T}$	10	0.90	8.00	36.05	95.80	100.00	4.80	33.65	84.15	99.75	100.00
		20	35.35	69.80	92.65	99.60	100.00	61.95	91.55	98.95	100.00	100.00		20	1.75	12.60	41.75	95.65	100.00	4.95	28.90	68.80	100.00	100.00
		30	22.25	79.55	95.10	99.80	100.00	63.95	94.50	98.80	100.00	100.00		30	1.45	15.65	42.80	96.30	100.00	5.80	29.15	68.60	99.85	100.00
		50	33.65	73.90	97.60	99.75	100.00	68.25	93.95	99.30	99.95	100.00		50	2.85	19.95	49.70	96.90	100.00	6.45	31.10	69.25	98.75	100.00
		100	41.90	83.30	97.00	99.85	100.00	69.35	95.60	99.65	100.00	100.00		100	3.50	26.25	61.25	96.15	100.00	6.85	36.05	72.10	98.35	100.00
		200	29.20	83.05	95.85	99.95	100.00	63.60	95.30	99.50	100.00	100.00		200	6.20	32.25	69.35	97.55	100.00	9.10	39.80	77.10	99.15	100.00
Expanding	$S^{(4)}_{n,T}$	10	29.55	71.85	94.20	99.55	100.00	57.10	89.70	98.65	99.95	100.00	$J^{(4)}_{n,T}$	10	2.80	14.10	38.50	90.60	99.95	10.65	39.95	83.10	99.50	100.00
		20	35.35	69.80	92.65	99.60	100.00	61.95	91.55	98.95	100.00	100.00		20	1.75	12.60	41.75	95.65	100.00	4.95	28.90	68.80	100.00	100.00
		30	22.25	79.55	95.10	99.80	100.00	63.95	94.50	98.80	100.00	100.00		30	1.45	15.65	42.80	96.30	100.00	5.80	29.15	68.60	99.85	100.00
		50	33.65	73.90	97.60	99.75	100.00	68.25	93.95	99.30	99.95	100.00		50	2.85	19.95	49.70	96.90	100.00	6.45	31.10	69.25	98.75	100.00
		100	41.90	83.30	96.95	99.85	100.00	69.35	95.60	99.65	100.00	100.00		100	0.00	0.00	0.00	0.05	0.00	0.00	0.05	0.05	0.05	0.35
		200	29.15	83.05	95.75	99.95	100.00	63.60	95.25	99.50	100.00	100.00		200	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fixed	$S^{(5)}_{n,T}$	10	29.45	72.00	94.25	99.55	100.00	57.40	89.70	98.65	99.95	100.00	$J^{(5)}_{n,T}$	10	0.40	1.25	34.60	96.80	100.00	2.20	18.40	84.10	99.80	100.00
		20	35.20	70.25	92.60	99.60	100.00	62.20	91.55	98.90	100.00	100.00		20	0.15	0.15	0.35	1.55	93.25	1.75	1.75	3.45	28.60	98.55
		30	22.50	80.15	95.10	99.80	100.00	64.00	94.55	98.80	100.00	100.00		30	0.45	0.35	0.40	0.90	99.40	2.25	2.00	3.30	38.45	99.70
		50	33.60	74.00	97.65	99.75	100.00	68.45	93.95	99.30	99.95	100.00		50	0.20	0.25	0.30	0.90	99.65	1.55	1.75	1.65	27.35	99.75
		100	41.95	83.35	97.05	99.85	100.00	69.40	95.60	99.65	100.00	100.00		100	0.30	0.20	0.20	0.50	99.65	1.20	1.30	2.00	30.55	99.70
		200	29.85	83.05	95.85	99.95	100.00	63.70	95.30	99.55	100.00	100.00		200	0.05	0.35	0.30	0.50	96.05	1.10	1.20	1.50	2.80	96.55
Expanding	$S^{(5)}_{n,T}$	10	31.00	72.50	94.20	99.55	100.00	58.45	89.70	98.65	99.95	100.00	$J^{(5)}_{n,T}$	10	1.15	2.00	2.90	90.30	99.65	12.75	46.00	84.80	98.75	99.70
		20	35.35	70.40	92.70	99.60	100.00	62.45	91.55	98.90	100.00	100.00		20	0.95	1.45	2.25	9.60	96.90	8.15	11.85	26.05	75.50	97.85
		30	22.60	80.15	95.10	99.80	100.00	64.35	94.55	98.85	100.00	100.00		30	2.30	2.45	2.35	9.45	99.45	9.25	13.80	27.10	85.50	99.45
		50	33.70	74.10	97.65	99.75	100.00	68.70	93.95	99.30	99.95	100.00		50	1.60	1.30	2.15	14.70	100.00	7.60	11.80	17.95	86.85	100.00
		100	42.00	83.30	97.05	99.85	100.00	69.35	95.60	99.65	100.00	100.00		100	0.90	0.55	0.20	0.25	1.35	3.65	2.30	1.00	1.15	3.70
		200	29.75	83.05	95.85	99.95	100.00	63.65	95.30	99.50	100.00	100.00		200	0.30	0.30	0.15	0.00	0.00	1.65	0.80	0.50	0.05	0.00

The results for the joint EPA tests are similar such that the size adjusted power of the tests increase only with T. In small samples it is very low. For instance, when $T = 10$ and $n = 200$, $J^{(4)}_{n,T}$ with a fixed number of common factors has a size adjusted power of 6.2% and 9.1% for 1% nominal level, respectively. However, the size adjusted

power of the test rises quickly as T increases.

Table 6 reports the results under the heterogeneous alternative hypothesis in the absence of SCD. This is an interesting case where we expect no improvement in the size adjusted power properties of the overall EPA tests $S_{n,T}^{(1)}$-$S_{n,T}^{(5)}$. The results confirm this expectation. For instance, in the extreme case of $T = 100$ and $n = 200$, the size adjusted power of the tests $S_{n,T}^{(1)}$ and $S_{n,T}^{(2)}$ with truncated kernel are 3.50% and 4.05% for 1% nominal level, respectively.

Table 6: Size Adjusted Power - DGP1 (No Common Factors): Low Spatial Dependence & Heterogeneous Alternative

			Overall EPA Test												Joint EPA Test											
			1% Nominal Size						5% Nominal Size						1% Nominal Size						5% Nominal Size					
Option	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100
	$S_{n,T}^{(1)}$	10	1.25	1.25	1.35	1.60	1.80	5.00	5.30	5.35	6.55	7.20	$J_{n,T}^{(1)}$	10	0.80	1.40	1.65	3.05	5.45	5.80	7.15	8.55	10.00	16.40
		20	1.30	1.30	1.40	1.60	2.25	5.30	5.90	6.05	6.45	7.70		20	1.45	1.85	2.10	3.35	8.55	6.25	7.30	8.50	14.15	27.15
		30	1.00	1.25	1.65	1.75	1.95	4.95	5.65	5.95	7.20	8.90		30	1.30	2.00	2.95	5.05	13.75	6.15	7.95	9.20	14.45	32.65
		50	1.05	1.45	1.55	1.55	2.55	4.95	5.45	6.25	6.60	8.40		50	1.40	1.75	3.35	5.40	21.15	6.50	9.30	12.85	20.05	48.10
		100	1.20	1.40	1.75	2.20	1.95	5.55	5.90	5.60	6.75	9.00		100	1.65	3.20	5.85	13.00	45.05	6.55	10.55	17.45	27.80	68.45
		200	1.30	1.05	1.45	1.65	3.50	5.90	6.50	6.35	7.50	9.45		200	1.95	4.05	7.85	23.80	75.90	7.30	14.30	22.20	44.45	90.95
Truncated	$S_{n,T}^{(2)}$	10	1.00	1.15	1.25	1.55	1.70	5.60	5.15	5.25	6.05	7.15	$J_{n,T}^{(2)}$	10	1.05	0.40	0.85	1.05	6.75	3.45	4.60	7.15	11.10	19.50
		20	1.05	1.35	1.15	1.70	2.55	5.40	5.45	5.55	6.40	7.90		20	0.85	0.60	0.65	2.55	10.25	2.25	3.30	4.25	11.55	25.35
		30	1.10	1.25	1.45	1.60	1.95	5.00	5.70	5.65	7.15	8.85		30	0.40	0.50	0.35	1.40	14.25	2.65	2.65	4.85	12.65	32.90
		50	0.95	1.25	1.70	1.25	2.50	5.15	5.35	6.25	6.60	8.95		50	0.45	0.40	0.70	1.75	18.85	2.80	3.40	3.25	13.25	44.95
		100	1.10	1.35	1.35	2.20	2.00	5.90	5.85	5.85	7.10	9.20		100	0.55	0.25	0.95	4.50	44.40	3.15	3.10	4.50	20.50	68.35
		200	1.20	1.05	1.50	1.65	4.05	5.35	6.20	6.35	7.70	9.95		200	0.25	0.85	0.70	0.25	62.70	1.50	3.40	3.20	3.50	84.25
Bartlett	$S_{n,T}^{(2)}$	10	1.25	1.25	1.35	1.60	1.80	5.00	5.30	5.35	6.55	7.20	$J_{n,T}^{(2)}$	10	0.80	1.40	1.65	3.05	5.45	5.80	7.15	8.55	10.00	16.40
		20	1.00	1.05	1.40	1.70	2.45	5.40	5.80	5.90	6.15	8.05		20	1.05	2.00	3.15	3.65	9.60	6.40	6.50	9.15	14.65	28.25
		30	1.00	1.15	1.45	1.60	1.85	5.30	5.80	6.30	6.95	8.70		30	1.20	1.80	2.80	5.45	15.90	5.90	8.20	9.90	14.75	34.35
		50	1.00	1.55	1.60	1.40	2.50	5.20	5.65	6.20	6.80	8.70		50	1.40	2.10	2.65	6.55	21.80	6.60	8.75	13.65	19.75	49.65
		100	1.10	1.35	1.60	2.25	1.85	5.50	6.00	5.75	6.85	9.15		100	1.45	3.80	4.90	13.20	45.50	7.15	11.35	18.15	29.50	70.80
		200	1.25	1.10	1.50	2.00	3.75	5.55	6.30	6.25	7.50	9.50		200	1.25	3.55	6.60	23.30	76.80	7.60	13.45	21.90	44.95	91.95
Parzen	$S_{n,T}^{(2)}$	10	1.25	1.25	1.35	1.60	1.80	5.00	5.30	5.35	6.55	7.20	$J_{n,T}^{(2)}$	10	0.80	1.40	1.65	3.05	5.45	5.80	7.15	8.55	10.00	16.40
		20	1.20	1.20	1.50	1.70	2.50	5.30	5.95	5.95	6.20	7.90		20	1.00	2.00	2.75	4.05	8.85	6.10	7.20	9.05	15.20	27.40
		30	0.95	1.20	1.45	1.75	1.85	5.10	5.90	6.10	7.30	8.95		30	1.45	2.15	2.45	4.95	14.80	6.20	7.80	9.45	14.55	33.15
		50	0.95	1.35	1.60	1.50	2.60	5.00	5.45	6.30	6.90	8.60		50	1.60	1.85	3.05	6.50	21.85	6.45	8.85	12.35	20.10	49.30
		100	1.20	1.30	1.60	2.30	1.90	5.70	5.85	5.70	6.65	9.10		100	1.30	3.70	4.95	13.65	45.35	6.95	10.55	20.15	29.35	70.55
		200	1.25	1.15	1.45	2.00	3.75	5.75	6.25	6.40	7.50	9.55		200	1.35	3.55	7.60	23.15	77.35	7.15	14.25	22.75	46.25	92.30
Tukey	$S_{n,T}^{(2)}$	10	1.25	1.25	1.35	1.60	1.80	5.00	5.30	5.35	6.55	7.20	$J_{n,T}^{(2)}$	10	0.80	1.40	1.65	3.05	5.45	5.80	7.15	8.55	10.00	16.40
		20	1.00	1.05	1.45	1.70	2.45	5.40	5.85	5.90	6.05	8.05		20	1.05	2.00	2.95	3.45	9.40	6.35	6.75	9.90	14.90	28.70
		30	0.95	1.20	1.35	1.60	1.85	5.25	5.90	6.40	6.90	8.75		30	1.40	1.85	2.70	5.60	15.60	5.60	8.15	9.80	15.10	34.25
		50	0.95	1.45	1.60	1.40	2.50	5.20	5.65	6.20	6.80	8.70		50	1.40	2.10	2.85	6.70	21.55	6.70	8.40	13.30	19.90	49.65
		100	1.05	1.35	1.65	2.20	1.85	5.45	6.05	5.70	6.85	9.10		100	1.45	3.35	4.55	13.05	45.20	6.75	11.00	18.95	30.25	70.35
		200	1.20	1.10	1.50	2.00	3.75	5.60	6.25	6.25	7.50	9.55		200	1.10	3.60	7.00	22.75	77.35	7.50	13.55	23.00	45.05	92.35
QS	$S_{n,T}^{(2)}$	10	1.20	1.20	1.20	1.65	1.80	5.20	5.10	5.45	6.60	7.15	$J_{n,T}^{(2)}$	10	0.85	1.45	1.65	3.30	5.50	5.55	6.85	8.20	10.35	16.40
		20	1.05	1.20	1.25	1.50	2.40	5.30	5.75	5.70	6.25	8.05		20	1.20	2.20	2.50	3.70	10.55	5.70	6.60	8.90	14.85	27.80
		30	1.10	1.15	1.45	1.65	1.80	5.15	5.80	6.30	7.05	8.70		30	1.30	1.60	2.75	5.25	16.65	5.30	8.30	10.05	15.70	34.50
		50	1.05	1.50	1.60	1.25	2.50	5.20	5.55	6.30	6.65	8.55		50	1.20	1.90	2.60	5.95	21.00	6.15	8.50	13.15	19.95	49.05
		100	1.15	1.35	1.50	2.20	1.95	5.60	6.00	5.80	6.90	9.20		100	1.40	4.20	4.35	12.25	46.95	6.65	10.65	16.90	30.60	69.90
		200	1.20	1.20	1.40	1.95	3.85	5.70	6.25	6.40	7.60	9.35		200	1.20	4.20	5.80	22.05	76.50	6.10	12.05	19.80	42.15	91.65
	$S_{n,T}^{(3)}$	10	0.95	1.10	0.95	1.35	2.00	5.40	5.15	5.35	6.20	7.10	$J_{n,T}^{(3)}$	10		1.25	1.30	3.10	5.95		5.80	8.40	9.40	18.90
		20	0.75	1.65	1.50	1.50	2.55	5.30	5.50	5.60	6.35	8.15		20			1.80	2.40	7.30			7.00	9.40	22.95
		30	1.00	1.10	1.70	1.35	2.00	5.50	5.40	6.15	6.45	8.95		30				2.20	9.35				9.90	27.40
		50	0.95	1.35	1.50	1.60	2.40	5.30	5.85	6.05	7.15	8.30		50					12.45					30.55
		100	1.05	1.25	1.60	2.00	2.30	5.60	6.05	5.70	6.65	9.55		100										
		200	1.05	1.15	1.30	2.05	3.30	5.30	5.75	6.65	7.15	9.10		200										
Fixed	$S_{n,T}^{(4)}$	10	0.95	1.35	1.35	1.25	1.90	5.45	5.45	5.25	5.95	7.25	$J_{n,T}^{(4)}$	10	0.90	1.25	1.20	3.00	6.40	4.95	5.55	7.00	9.95	18.65
		20	1.10	0.95	1.45	1.15	2.00	5.55	5.80	5.50	6.20	7.60		20	0.90	1.55	1.80	2.65	9.10	4.60	6.25	8.10	11.35	24.85
		30	1.00	1.15	1.60	1.30	1.80	5.15	5.35	5.15	6.65	8.20		30	1.35	1.90	1.55	4.20	13.90	5.05	7.05	7.65	12.95	30.90
		50	0.75	0.75	1.60	1.15	2.70	4.90	5.00	5.80	6.70	7.90		50	0.95	1.10	3.05	4.60	20.45	5.10	8.10	9.30	14.70	43.25
		100	1.10	0.95	1.50	2.05	2.35	5.45	5.30	5.85	6.35	9.60		100	1.25	2.65	2.70	7.85	33.65	5.55	7.50	10.60	21.30	58.45
		200	1.35	1.10	1.30	1.70	3.15	5.80	5.80	6.25	7.95	9.30		200	1.75	2.15	2.60	11.00	53.95	5.65	9.15	10.90	28.25	76.15
Expanding	$S_{n,T}^{(4)}$	10	1.05	1.20	1.00	1.60	1.80	6.05	5.45	6.40	5.60	6.80	$J_{n,T}^{(4)}$	10	1.05	1.95	1.90	2.55	5.80	5.25	6.00	7.05	10.30	16.90
		20	1.10	0.95	1.45	1.15	2.00	5.55	5.80	5.50	6.20	7.60		20	0.90	1.55	1.80	2.65	9.10	4.60	6.25	8.10	11.35	24.85
		30	1.00	1.15	1.60	1.30	1.80	5.15	5.35	5.15	6.65	8.20		30	1.35	1.90	1.55	4.20	13.90	5.05	7.05	7.65	12.95	30.90
		50	0.75	0.75	1.60	1.15	2.70	4.90	5.00	5.80	6.70	7.90		50	0.95	1.10	3.05	4.60	20.45	5.10	8.10	9.30	14.70	43.25
		100	0.90	1.30	1.35	1.55	2.25	5.45	5.60	6.10	7.10	9.25		100	0.75	1.45	2.30	6.25	28.85	4.85	6.05	10.05	17.30	53.50
		200	1.10	0.95	1.55	1.55	3.30	4.85	5.65	6.70	7.90	9.25		200	1.35	2.25	3.35	9.45	47.15	4.15	6.60	9.40	21.65	68.50
Fixed	$S_{n,T}^{(5)}$	10	0.90	1.15	1.10	1.30	2.05	5.05	5.15	5.35	5.45	7.10	$J_{n,T}^{(5)}$	10	0.50	0.85	1.50	2.65	5.20	3.80	5.05	7.15	9.95	16.30
		20	0.90	0.90	1.60	1.20	1.95	5.35	5.65	5.70	6.00	7.85		20	0.70	0.50	0.50	2.30	8.15	3.05	3.45	5.70	10.25	23.95
		30	1.30	1.35	1.40	1.40	2.10	5.15	5.40	5.20	6.50	7.85		30	0.90	1.05	0.15	3.10	10.95	3.45	3.60	2.90	10.10	28.85
		50	0.80	0.90	1.85	1.50	2.65	4.90	5.45	6.35	6.95	8.00		50	0.55	0.60	0.30	2.45	18.75	3.85	2.65	3.95	12.95	40.60
		100	1.10	1.10	1.55	2.15	2.45	4.95	5.25	5.40	6.65	9.60		100	0.95	0.40	0.55	3.85	33.55	3.25	3.20	3.90	17.85	59.25
		200	1.25	1.25	1.50	1.75	3.85	5.35	5.80	6.50	7.50	9.65		200	0.60	0.30	1.20	0.70	41.65	2.45	2.75	3.80	4.40	66.90
Expanding	$S_{n,T}^{(5)}$	10	0.85	0.90	1.10	1.40	2.05	6.10	5.15	5.95	5.70	7.00	$J_{n,T}^{(5)}$	10	0.65	1.60	1.55	2.15	6.30	3.15	5.75	6.40	9.60	16.45
		20	0.90	0.90	1.60	1.20	1.95	5.35	5.65	5.70	6.00	7.85		20	0.70	0.50	0.50	2.30	8.15	3.05	3.45	5.70	10.25	23.95
		30	1.30	1.35	1.40	1.40	2.10	5.15	5.40	5.20	6.50	7.85		30	0.90	1.05	0.15	3.10	10.95	3.45	3.60	2.90	10.10	28.85
		50	0.80	0.90	1.85	1.50	2.65	4.90	5.45	6.35	6.95	8.00		50	0.55	0.60	0.30	2.45	18.75	3.85	2.65	3.95	12.95	40.60
		100	1.10	1.30	1.50	1.95	2.20	4.90	5.40	5.80	6.60	9.35		100	0.70	0.70	0.85	3.10	29.85	3.35	4.05	4.10	16.70	54.85
		200	1.40	1.10	1.50	1.85	3.55	4.90	5.80	6.55	7.45	9.25		200	0.60	0.75	0.40	0.45	33.15	2.80	2.20	2.30	3.45	58.35

The performance of the joint EPA tests improve with the increase in T. For instance, for $n = 50$ the size adjusted power of the test $J_{n,T}^{(2)}$ with truncated kernel increase from 13.25% while $T = 50$ to 44.95% while $T = 100$ in 5% nominal level. The factor robust tests have similar size adjusted power in general. The corresponding numbers for the test $J_{n,T}^{(4)}$ with a fixed number of common factors are 14.70% while $T = 50$ to 43.25%.

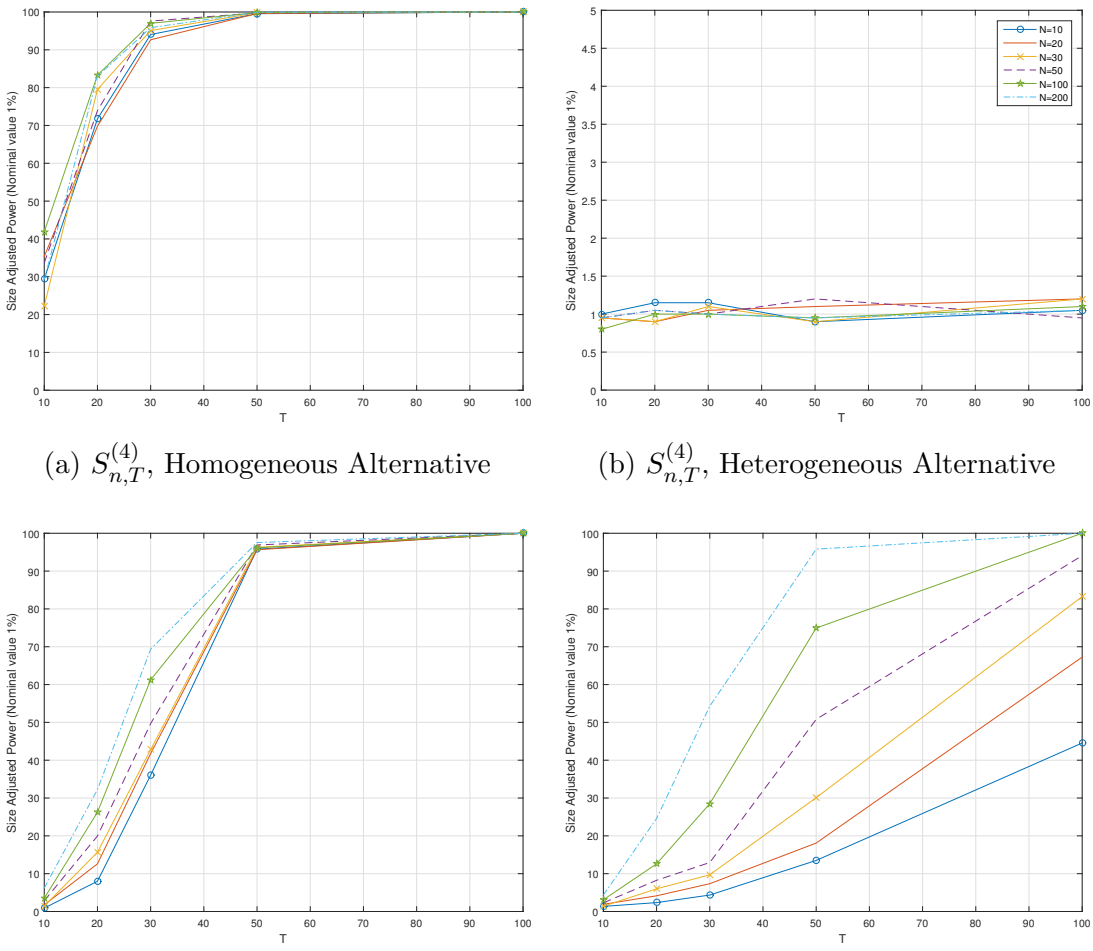
The results for the case of heterogeneous alternative hypothesis with DGP2 and low spatial dependence are given in Table 7. As in the previous case the size adjusted power of overall EPA tests do not improve with the increases in sample size. In this case differences between robust tests and non-robust test $J_{n,T}^{(1)}$ is pronounced. There is nearly no improvement in the size adjusted power of the latter. In the extreme case of $T = 100$ and $n = 200$, the size adjusted power of the tests $J_{n,T}^{(1)}$ is 1.75% and 8.55% for 1% and 5% nominal levels, respectively. This result is similar for the kernel-robust test with truncated kernel. Whereas, for Bartlett, Tukey and quadratic spectral kernels we have good power properties for large n and large T. In this case an important difference is observed between the kernel robust tests $J_{n,T}^{(2)}$ and factor robust tests $J_{n,T}^{(4)}$ and $J_{n,T}^{(5)}$ such that the size adjusted power of the latter ones improve with T for a fixed n.

The size adjusted power properties of the factor robust tests $S_{n,T}^{(4)}$ and $J_{n,T}^{(4)}$ are also shown in Figure 2 for the case of DGP2 and low spatial dependence. It is clearly seen that under the homogeneous alternative hypothesis the size adjusted power of the overall test $S_{n,T}^{(4)}$ is higher than the joint test $J_{n,T}^{(4)}$ for all sample sizes. However, under the heterogeneous alternative the size adjusted power of the overall test equals to the nominal value 1% for any sample size.

The power and size adjusted power in the case of no spatial dependence robust and non-robust tests can be summarized as follows: the power and size adjusted power of the non-robust tests are higher than the robust tests under no spatial dependence in DGP1. However, when there is high spatial dependence the above results become much more pronounced such that the non-robust test performs very poorly as it is oversized. In the case of DGP2 the factor robust tests perform well in moderate to large samples. The power results for the standardized joint EPA tests (36) are in line with their size properties such that they provide little improvement over their chi-square alternatives. Notice that these two sets of tests have identical size adjusted powers.

Table 7: Size Adjusted Power - DGP2 (Common Factors): Low Spatial Dependence & Heterogeneous Alternative

			Overall EPA Test											Joint EPA Test										
			1% Nominal Size					5% Nominal Size							1% Nominal Size					5% Nominal Size				
Option	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100	*Test*	$n\backslash T$	10	20	30	50	100	10	20	30	50	100
	$S_{n,T}^{(1)}$	10	1.05	1.15	1.15	1.00	1.10	5.20	5.10	5.40	5.00	5.65	$J_{n,T}^{(1)}$	10	1.20	1.35	1.30	1.10	1.75	5.60	5.45	6.20	5.80	9.05
		20	1.00	0.85	1.00	1.05	1.10	4.95	5.10	5.35	4.75	5.50		20	1.05	1.15	1.10	1.20	1.50	5.30	5.80	5.95	6.15	8.50
		30	1.05	0.95	0.95	0.90	1.30	5.05	4.90	5.10	5.00	5.10		30	1.00	1.10	1.15	1.35	1.85	5.30	5.40	6.30	6.40	7.60
		50	1.05	1.00	0.95	1.05	1.10	5.00	4.85	5.10	4.95	5.15		50	0.95	1.10	1.45	1.35	1.70	5.30	5.45	5.80	6.30	7.75
		100	0.95	1.05	0.95	1.00	1.15	5.10	5.05	5.15	5.20	5.05		100	0.95	1.05	1.40	1.15	2.00	5.05	5.50	5.90	6.05	8.45
		200	1.00	0.95	1.00	1.00	1.10	4.90	4.95	5.10	5.15	4.90		200	1.15	1.25	1.15	1.20	1.75	5.35	5.60	5.80	5.80	8.55
Truncated	$S_{n,T}^{(2)}$	10	0.95	1.10	1.10	0.85	1.05	5.45	5.10	5.25	4.95	6.05	$J_{n,T}^{(2)}$	10	1.20	1.25	1.70	1.95	4.25	5.35	7.35	7.50	9.55	20.55
		20	0.95	0.95	0.95	1.05	1.25	5.00	5.15	5.15	5.00	5.55		20	0.40	0.60	0.75	0.60	1.45	2.85	3.45	5.60	5.30	9.55
		30	1.00	0.90	1.15	0.95	1.25	5.15	5.00	5.35	5.10	5.30		30	0.60	0.45	0.75	1.05	1.40	3.50	3.45	3.50	4.05	7.10
		50	0.95	1.05	0.95	1.25	1.00	5.05	5.05	4.80	4.95	4.95		50	0.70	0.75	1.20	1.15	1.65	3.25	4.00	4.15	5.85	7.35
		100	0.90	1.05	1.05	0.95	1.10	4.95	5.05	5.20	5.05	5.25		100	0.70	0.60	0.60	0.95	1.20	3.05	3.45	4.50	4.95	7.20
		200	1.15	1.00	1.05	1.00	1.05	4.90	5.10	5.25	5.00	4.95		200	0.85	0.50	0.65	1.10	1.35	3.50	3.10	3.70	4.75	6.75
Bartlett	$S_{n,T}^{(2)}$	10	1.05	1.15	1.15	1.00	1.10	5.20	5.10	5.40	5.00	5.65	$J_{n,T}^{(2)}$	10	1.20	1.35	1.30	1.10	1.75	5.60	5.45	6.20	5.80	9.05
		20	0.90	1.00	0.90	1.00	1.10	4.90	5.35	5.15	5.20	5.65		20	1.05	1.05	1.50	2.15	6.15	5.75	6.85	7.90	14.40	32.70
		30	1.00	1.00	1.25	0.90	1.25	5.05	4.90	5.25	5.10	5.30		30	0.95	1.35	1.60	2.70	6.40	5.95	7.80	8.10	12.95	31.95
		50	1.05	1.15	1.15	1.20	1.10	5.10	5.10	5.00	4.90	5.05		50	1.35	1.25	2.15	2.60	6.90	6.20	7.45	8.55	14.00	46.30
		100	0.90	0.95	1.05	0.95	1.10	4.95	5.10	5.20	5.00	5.20		100	1.35	2.10	2.90	3.35	35.70	6.70	9.35	13.20	23.35	86.70
		200	1.00	1.15	1.05	1.05	1.00	4.85	4.95	5.20	5.10	5.00		200	1.05	2.30	3.20	6.30	43.40	5.80	10.15	15.45	33.00	98.85
Parzen	$S_{n,T}^{(2)}$	10	1.05	1.15	1.15	1.00	1.10	5.20	5.10	5.40	5.00	5.65	$J_{n,T}^{(2)}$	10	1.20	1.35	1.30	1.10	1.75	5.60	5.45	6.20	5.80	9.05
		20	1.00	1.00	0.95	1.05	1.10	5.10	5.05	5.20	5.15	5.65		20	1.00	1.05	1.25	1.35	1.90	5.45	6.00	6.15	7.70	12.65
		30	1.10	0.95	1.20	0.90	1.25	5.15	4.95	5.45	5.05	5.25		30	1.00	1.10	1.55	1.50	3.10	5.60	6.05	6.90	7.65	11.15
		50	0.95	1.05	1.00	1.15	1.15	5.15	5.15	4.80	4.95	5.15		50	1.10	1.10	1.45	1.60	2.05	5.45	5.80	6.90	8.20	13.00
		100	0.95	0.95	0.95	0.95	1.05	4.95	4.90	5.20	4.90	5.25		100	1.05	1.35	1.90	2.05	12.15	6.15	6.90	10.05	11.50	50.15
		200	1.00	1.05	1.05	1.05	1.05	4.85	4.90	5.15	5.00	4.95		200	1.10	2.05	1.75	3.05	7.70	5.70	8.40	10.25	16.05	70.45
Tukey	$S_{n,T}^{(2)}$	10	1.05	1.15	1.15	1.00	1.10	5.20	5.10	5.40	5.00	5.65	$J_{n,T}^{(2)}$	10	1.20	1.35	1.30	1.10	1.75	5.60	5.45	6.20	5.80	9.05
		20	0.95	1.00	0.90	0.95	1.05	4.90	5.20	5.10	5.20	5.65		20	0.90	1.15	1.65	2.20	6.90	5.10	7.70	7.75	15.00	34.30
		30	1.00	1.00	1.25	0.90	1.25	5.05	5.00	5.35	5.10	5.25		30	0.95	1.45	1.40	2.95	7.65	5.50	7.95	9.15	13.30	37.50
		50	1.05	1.10	1.15	1.20	1.10	5.00	5.15	4.95	4.90	5.10		50	1.00	1.20	1.85	2.45	6.25	6.05	8.00	9.25	15.35	50.05
		100	0.90	0.95	1.00	1.00	1.05	4.90	5.00	5.20	4.95	5.20		100	0.70	2.60	4.20	5.35	66.50	6.10	11.75	21.45	41.30	97.00
		200	1.00	1.10	1.05	1.05	1.00	5.05	4.95	5.15	5.15	5.05		200	0.75	1.55	4.25	12.15	86.50	5.05	11.50	23.65	59.75	100.00
QS	$S_{n,T}^{(2)}$	10	1.05	1.20	1.10	1.00	1.15	5.15	5.00	5.50	5.00	5.80	$J_{n,T}^{(2)}$	10	1.10	1.35	1.40	1.10	1.90	5.60	5.65	6.15	6.40	10.55
		20	0.90	0.95	0.95	1.05	1.20	4.95	5.25	5.10	5.20	5.55		20	1.10	1.20	1.70	5.55	21.85	6.45	9.10	13.40	27.85	59.20
		30	1.00	1.00	1.25	0.90	1.30	5.30	4.95	5.25	5.15	5.35		30	1.10	1.75	2.45	5.65	28.25	6.00	10.45	13.45	27.80	67.15
		50	1.00	1.05	1.05	1.20	1.10	5.00	5.00	4.90	4.90	5.00		50	1.40	1.50	3.60	6.35	35.55	7.05	9.80	13.15	31.25	85.15
		100	0.90	0.95	1.05	0.90	1.15	5.05	5.05	5.20	5.00	5.20		100	1.65	3.50	5.75	13.40	91.65	7.90	14.75	33.50	64.80	99.80
		200	1.05	1.10	1.05	1.05	1.00	4.90	5.05	5.20	5.05	5.00		200	1.35	2.65	7.20	35.55	99.05	7.45	17.90	40.85	86.00	100.00
	$S_{n,T}^{(3)}$	10	1.05	1.10	1.10	0.85	1.05	4.95	5.05	5.25	5.05	5.95	$J_{n,T}^{(3)}$	10		2.30	7.20	18.70	56.75		12.45	22.65	38.25	75.90
		20	0.95	0.90	1.00	1.00	1.20	5.10	5.15	5.20	5.05	5.70		20			5.10	25.20	77.45			17.45	48.75	91.45
		30	1.00	0.90	1.10	0.85	1.20	5.20	5.00	5.25	5.05	5.20		30				19.05	86.35				44.95	96.00
		50	0.95	1.05	1.00	1.20	0.95	5.00	5.15	5.00	4.70	4.95		50					89.95					97.75
		100	0.80	1.00	1.00	0.95	1.10	4.90	5.00	5.10	5.10	5.15		100										
		200	0.95	1.05	1.00	0.95	1.05	4.90	5.20	5.00	4.90	4.90		200										
Fixed	$S_{n,T}^{(4)}$	10	1.00	1.15	1.15	0.90	1.05	4.95	5.15	5.20	4.95	5.80	$J_{n,T}^{(4)}$	10	1.35	2.35	4.35	13.50	44.55	5.80	12.10	16.85	34.05	68.05
		20	0.95	0.90	1.05	1.10	1.20	5.00	5.15	5.15	5.10	5.75		20	1.90	4.15	7.35	18.10	67.25	7.60	13.80	21.45	45.05	83.75
		30	0.95	0.90	1.10	0.90	1.20	5.25	5.05	5.25	5.10	5.25		30	1.40	6.00	9.70	30.10	83.25	9.60	15.95	29.20	58.90	94.85
		50	0.95	1.05	1.00	1.20	0.95	5.00	5.15	5.00	4.75	4.95		50	2.30	8.25	12.95	50.80	94.10	8.80	21.15	36.60	76.20	98.80
		100	0.80	1.00	1.00	0.95	1.10	4.95	5.00	5.10	5.10	5.10		100	3.10	12.60	28.50	74.95	100.00	11.70	31.60	54.25	90.60	100.00
		200	0.95	1.05	1.00	0.95	1.05	4.90	5.20	5.00	4.90	4.90		200	4.35	24.60	54.30	95.80	100.00	13.35	44.35	74.05	98.85	100.00
Expanding	$S_{n,T}^{(4)}$	10	1.20	1.15	1.20	0.95	1.05	5.00	5.25	5.25	4.95	5.80	$J_{n,T}^{(4)}$	10	1.65	2.70	3.50	6.65	23.80	7.40	9.95	14.30	22.60	55.90
		20	0.95	0.90	1.05	1.10	1.20	5.00	5.15	5.15	5.10	5.75		20	1.90	4.15	7.35	18.10	67.25	7.60	13.80	21.45	45.05	83.75
		30	0.95	0.90	1.10	0.90	1.20	5.25	5.05	5.25	5.10	5.25		30	1.40	6.00	9.70	30.10	83.25	9.60	15.95	29.20	58.90	94.85
		50	0.95	1.05	1.00	1.20	0.95	5.00	5.15	5.00	4.75	4.95		50	2.30	8.25	12.95	50.80	94.10	8.80	21.15	36.60	76.20	98.80
		100	0.80	1.00	1.00	0.95	1.10	4.95	5.00	5.10	5.10	5.10		100	1.20	6.20	17.95	57.85	98.70	6.05	19.65	37.85	78.45	99.90
		200	0.95	1.05	1.00	0.95	1.05	4.90	5.20	5.00	4.90	4.90		200	1.85	9.25	22.60	56.20	85.60	6.70	20.80	38.45	68.25	89.05
Fixed	$S_{n,T}^{(5)}$	10	0.95	1.15	1.20	0.90	1.05	4.95	5.25	5.20	5.00	5.90	$J_{n,T}^{(5)}$	10	0.20	1.00	4.35	14.35	48.20	5.00	8.90	16.90	35.40	71.70
		20	0.95	0.90	0.95	1.05	1.20	4.90	5.05	5.15	5.15	5.80		20	0.45	0.45	0.85	1.30	10.40	2.95	2.80	6.00	9.95	77.70
		30	0.90	0.90	1.10	0.90	1.20	5.30	5.05	5.20	5.10	5.20		30	0.65	0.55	0.80	1.30	70.30	3.95	3.45	4.35	13.85	95.40
		50	0.95	1.05	1.00	1.20	0.95	4.95	5.15	4.90	4.75	4.95		50	0.80	0.70	0.75	1.20	98.50	3.30	3.60	3.60	13.40	99.80
		100	0.80	1.00	1.00	0.95	1.10	5.00	5.00	5.15	5.10	5.10		100	0.60	0.70	0.55	0.60	100.00	3.00	4.30	2.80	19.30	100.00
		200	0.95	1.05	1.00	0.95	1.05	4.90	5.20	5.10	4.90	4.90		200	0.45	0.70	0.40	1.10	99.85	2.20	3.10	3.65	5.85	99.85
Expanding	$S_{n,T}^{(5)}$	10	1.10	1.10	1.20	0.95	1.05	5.00	5.20	5.25	4.95	5.85	$J_{n,T}^{(5)}$	10	0.45	0.55	0.60	6.45	32.95	2.95	7.25	13.60	27.55	64.80
		20	0.95	0.90	0.95	1.05	1.20	4.90	5.05	5.15	5.15	5.80		20	0.45	0.45	0.85	1.30	10.40	2.95	2.80	6.00	9.95	77.70
		30	0.90	0.90	1.10	0.90	1.20	5.30	5.05	5.20	5.10	5.20		30	0.65	0.55	0.80	1.30	70.30	3.95	3.45	4.35	13.85	95.40
		50	0.95	1.05	1.00	1.20	0.95	4.95	5.15	4.90	4.75	4.95		50	0.80	0.70	0.75	1.20	98.50	3.30	3.60	3.60	13.40	99.80
		100	0.80	1.00	1.00	0.95	1.10	5.00	5.00	5.15	5.10	5.10		100	0.45	0.35	0.35	1.05	99.95	2.65	2.90	1.75	29.10	100.00
		200	0.95	1.05	1.00	0.95	1.05	4.90	5.20	5.10	4.90	4.90		200	0.25	0.80	0.55	0.50	77.65	1.40	3.10	2.95	3.20	85.95



(a) $S_{n,T}^{(4)}$, Homogeneous Alternative

(b) $S_{n,T}^{(4)}$, Heterogeneous Alternative

(c) $J_{n,T}^{(4)}$, Homogeneous Alternative

(d) $J_{n,T}^{(4)}$, Heterogeneous Alternative

Figure 2: Size Adjusted Power of Factor-Robust Tests Under Different Alternative Hypotheses for DGP2 and Low Spatial Dependence (1% Nominal Size)

5 Empirical Application

5.1 Data, Empirical Setup And Preliminaries

In this section, we use the tests proposed to compare the OECD and IMF GDP growth forecasts. The data for the IMF forecast errors come from their Historical WEO Forecasts Database. The database includes historical τ-steps ahead forecast values, $\tau = 1, 2, \ldots, 5$, for GDP growth rate. The data covers up to 192 countries and starts from early 1990s. We collected similar data from the past vintages of the Economic Outlook of the OECD. The Economic Outlook contains only 1-step ahead forecasts. Eventually we have a balanced panel dataset of GDP growth forecast errors of 29 OECD countries from the two organization between 1998 and 2016. To investigate the role of heterogeneity and the change in the dimensions of the panel dataset, we also apply the tests to a sample of G7 countries between 1991 and 2016. This dataset comes from Turner (2017).

We implement the tests described above on the two datasets. We create four different loss series: absolute loss, quadratic loss and two different types of linex loss. The absolute error loss differential is created as

$$\Delta L^1_{i,t} = |e_{1i,t}| - |e_{2i,t}|,$$

where, as is throughout the application, first organization is the OECD. This loss function is important when we compare the magnitude of the (absolute) bias made by the two organizations. The quadratic loss is generated as

$$\Delta L^2_{i,t} = e^2_{1i,t} - e^2_{2i,t}.$$

This loss function is arguably the most frequently used one and it is useful to compare the variance in the forecast errors. For instance, if the forecasts of the both organizations are unbiased the expectation of absolute error loss is zero and quadratic loss permits to compare the variances directly.

The disadvantage of these two loss functions is that they give equal weight to positive and negative forecast errors. A more flexible loss function is linex loss where it is possible to choose a parameter to give higher weight to either depending on their economic importance. In Section 2.1, we saw that the forecast errors of the two organizations take very large negative values for all countries during the financial crisis. Hence, to compare the performance of the two organizations during the crisis one can give more weight to negative values, vice versa. The two linex loss differential series

we use are computed as

$$\Delta L^3_{i,t} = (\exp\{\rho_1 e_{1i,t}\} - \rho_1 e_{1i,t} - 1) - (\exp\{\rho_1 e_{2i,t}\} - \rho_1 e_{2i,t} - 1)\,,$$

$$\Delta L^4_{i,t} = (\exp\{\rho_2 e_{1i,t}\} - \rho_2 e_{1i,t} - 1) - (\exp\{\rho_2 e_{2i,t}\} - \rho_2 e_{2i,t} - 1)\,,$$

where $\rho_1 = -0.35$ and $\rho_1 = 0.5$. With the first choice of the parameter more weight is given to negative values whereas with the second more weight is given to positive values. Thus, the first function can be used to compare the performance during crisis period.

We begin the analysis by the DM tests applied to each country. We compute the DM test statistic for each country between the years 1998 and 2016 using all four loss functions. In the computations, we use a Bartlett kernel with a bandwidth parameter of 0 because we have 1-step ahead forecasts. The result are given in Table 8.

First, in terms of the sign of the statistics, a considerable amount of heterogeneity can be observed in the sample. For all types of loss functions roughly half of the statistics are negative. Second, most of these statistics are statistically insignificant with exceptions being BEL, CAN, ESP, HUN, LUX and NZL. For BEL which is a country where the predictive ability of the IMF is superior, the EPA hypothesis can be rejected at 5% and 10% levels with absolute and quadratic losses, respectively. In the case of CAN, we can reject the EPA hypothesis with absolute loss and Linex Loss 1 at 10% and 5% significance levels, respectively. For CAN too, IMF predicts the economic growth rate better than OECD. Since Linex Loss 1 gives more weight to negative values and the statistic is positive, for CAN, we find that in the periods like crisis OECD made bigger forecast errors than the IMF on average. In the case of ESP and HUN, the differences in predictive ability are significant only with the absolute and quadratic losses. For ESP OECD predictions, for HUN IMF predictions outperform the other. For LUX, the EPA hypothesis can be rejected only with Linex Loss 1 at 5% level whereas for NZL we can reject it with absolute loss and Linex Loss 2 both at 5% levels.

5.2 Testing for CD

As found in our Monte Carlo simulations, the increase in the number of cross-sections increases the power of EPA tests. To see if we can reject the EPA hypothesis by using cross sectional information we apply the panel tests to the dataset. However, the gain from the usage of panels depend on the degree and nature of CD. As shown in Table 1 for the G7 sample, the cross-country correlations between the forecast errors of both

Table 8: DM Test Statistics for Each Country

Country	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*	*Country*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*
AUS	-0.6155	-0.4050	-0.2726	-0.5268	ISL	-0.5325	-0.4712	-1.4357	0.6377
	(0.5382)	(0.6855)	(0.7851)	(0.5984)		(0.5943)	(0.6375)	(0.1511)	(0.5236)
AUT	0.6885	0.1479	-0.5333	0.5927	ITA	1.0697	1.1711	1.2076	0.5369
	(0.4912)	(0.8824)	(0.5939)	(0.5534)		(0.2848)	(0.2415)	(0.2272)	(0.5913)
BEL	2.0138	1.6625	1.4534	1.3261	JPN	1.4231	1.0345	0.5934	0.5011
	(0.0440)	(0.0964)	(0.1461)	(0.1848)		(0.1547)	(0.3009)	(0.5529)	(0.6163)
CAN	1.7833	1.5011	2.2292	0.5123	KOR	0.5976	0.5560	1.2499	-0.2288
	(0.0745)	(0.1333)	(0.0258)	(0.6084)		(0.5501)	(0.5782)	(0.2113)	(0.8190)
CHE	1.0464	1.0980	1.0645	0.7237	LUX	0.9249	1.0136	1.7619	-0.5726
	(0.2954)	(0.2722)	(0.2871)	(0.4693)		(0.3550)	(0.3108)	(0.0781)	(0.5669)
CZE	-1.0617	-1.0003	-0.9317	-1.1919	MEX	-0.5196	-0.3816	0.5807	-1.1110
	(0.2884)	(0.3172)	(0.3515)	(0.2333)		(0.6034)	(0.7027)	(0.5614)	(0.2666)
DEU	-0.3686	-1.1310	-0.9157	-1.1258	NLD	0.0813	0.8709	1.2784	0.2371
	(0.7124)	(0.2581)	(0.3598)	(0.2603)		(0.9352)	(0.3838)	(0.2011)	(0.8125)
DNK	0.0445	-0.7032	-0.6089	-0.6408	NOR	0.0084	-0.6276	-0.7384	-0.6664
	(0.9645)	(0.4819)	(0.5426)	(0.5217)		(0.9933)	(0.5302)	(0.4603)	(0.5052)
ESP	-1.6955	-1.6919	-1.5269	-1.2600	NZL	-2.0726	-1.5350	-1.1797	-2.0470
	(0.0900)	(0.0907)	(0.1268)	(0.2077)		(0.0382)	(0.1248)	(0.2381)	(0.0407)
FIN	0.4240	0.1252	1.2232	-0.8274	POL	-0.4466	-0.9600	-1.0961	-0.6534
	(0.6716)	(0.9003)	(0.2213)	(0.4080)		(0.6552)	(0.3370)	(0.2730)	(0.5135)
FRA	1.4205	1.4507	1.1941	1.6433	PRT	-0.0675	0.1274	0.3290	-0.0223
	(0.1555)	(0.1469)	(0.2325)	(0.1003)		(0.9461)	(0.8987)	(0.7422)	(0.9822)
GBR	-0.2435	-1.1233	-1.4703	-0.4391	SWE	-0.6610	-0.1636	0.7728	-1.3266
	(0.8076)	(0.2613)	(0.1415)	(0.6606)		(0.5086)	(0.8701)	(0.4397)	(0.1846)
GRC	-1.0708	-1.4509	-1.2931	-1.5038	TUR	-0.0736	-0.3015	-0.1499	-0.7124
	(0.2843)	(0.1468)	(0.1960)	(0.1326)		(0.9414)	(0.7630)	(0.8809)	(0.4762)
HUN	2.3868	1.8742	1.1977	1.4255	USA	0.2005	0.0081	0.0826	-0.0416
	(0.0170)	(0.0609)	(0.2311)	(0.1540)		(0.8411)	(0.9935)	(0.9342)	(0.9668)
IRL	0.4724	0.6562	1.5501	-1.0268					
	(0.6366)	(0.5117)	(0.1211)	(0.3045)					

Note: The statistics are calculated using (8) with bandwidth 0. The values shown in parentheses are p-values.

organizarions are fairly high. This is the case for the OECD sample too for which the results are not reported to save space. Hence, before proceeding to panel tests of EPA, we analyse the CD in the two panel datasets of OECD and G7 countries. Here, we adapt the two step methodology of Bailey, Holly and Pesaran (2016). This involves testing for WCD in the first step and proceeding with spatial modeling if the null hypothesis is not rejected. If the null hypothesis is not rejected, we defactor the variables using PC methods or their cross sectional averages.

Here, we use two tests of CD. The first is the LM test of the absence of CD by Breusch and Pagan (1980) and the second one is the WCD test of Pesaran (2015). The first is a test of the joint significance of pairwise correlations between all units in the panel. The null hypothesis of this test is the absence of CD between any pair in the panel and the statistic is distributed as χ^2_q with $q = n(n-1)/2$. Hence, the test is more suitable for the cases of fixed and small n.

The null hypothesis of the second test is the absence of WCD in the process. The rejection of this null indicates the presence of SCD. Hence, an analysis of the common factors in the panel is necessary. The test statistic is asymptotically normal as $n \to \infty$ and more suitable for large panels.

The results for the OECD countries are given in Table 9. The null hypothesis of no CD is rejected using Breusch-Pagan test for all four loss types. The number of countries is larger than the number of periods in this dataset, so the test is not very reliable.

Using Pesaran's test, the null of WCD can be rejected all loss functions except the absolute loss. Therefore, it is needed to defactor the observations. We use PC methods to do so. For this, it is needed to choose the number of factors to be extracted from the panel dataset. Bai and Ng (2002) suggested several information criteria to determine the number of factors. These criteria suggest the existence of 5 common factors [The results on the number of common factors selected using different information criteria are available from the authors upon request]. After removing 5 common factors from the data, Breusch-Pagan tests on the residuals still indicate CD. Whereas, according to Pesaran's test the null hypothesis of WCD cannot be rejected. This means that the tests which allow for common factors and spatial dependence on this sample are more liable.

Table 9: Weak CD Tests for the Sample of OECD Countries

	Original Data				Defactored Data			
	Absolute Loss	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*
Breusch-Pagan LM Test	478.1028	856.2750	1,664.5419	1,350.5553	619.6212	1,063.6243	872.6455	846.1732
	(0.0078)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Pesaran's Test	0.4718	2.4395	2.8551	2.4181	0.5174	0.6088	1.3628	0.8416
	(0.6371)	(0.0147)	(0.0043)	(0.0156)	(0.6049)	(0.5427)	(0.1729)	(0.4000)

Note: Breusch-Pagan's LM Test and Pesaran's Test are calculated using (52) and (54), respectively. The values shown in parantheses are p-values.

Table 10 gives the CD test results for the sample of G7 countries. In this sample too, the null hypothesis of no CD can be rejected using Breusch-Pagan test. For quadratic loss function the null of WCD cannot be rejected in traditional levels using Pesaran's test. For Linex Loss 1 and Linex Loss 2, the WCD hypothesis can be rejected at 10% level at the highest. For these loss functions, it is necessary to defactor the observations. The information criteria for this sample indicates the existence of 2 common factors. When we defactor the variables using 2 factors, the WCD hypothesis is not rejected for the two linex loss differentials. These conclusions can guide us in panel tests of EPA.

Table 10: Weak CD Tests for the Sample of G7 Countries

	Original Data				Defactored Data			
	Absolute Loss	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*
Breusch-Pagan LM Test	34.6358	51.9446	100.3906	80.2496	89.4015	173.0582	129.7162	95.2154
	(0.0309)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Pesaran's Test	3.6796	0.8368	-1.7005	4.9519	-2.4299	3.7122	-0.4214	-0.2091
	(0.0002)	(0.4027)	(0.0890)	(0.0000)	(0.0151)	(0.0002)	(0.6735)	(0.8344)

Note: Breusch-Pagan's LM Test and Pesaran's Test are calculated using (52) and (54), respectively. The values shown in parantheses are p-values.

5.3 Panel Tests for the EPA Hypotheses

The panel EPA tests for the OECD dataset is given in Table 11. As before for the time series kernels we use a bandwidth of 0. For the spatial kernels we used the geographic distances between countries. The data on geographical distance come from CEPII GeoDist dataset (Mayer and Zignago, 2011). We chose the 25th percentile

of the sample of distances as the bandwidth parameter in all kernel functions. The statistics for the absolute loss and Linex Loss 1 are positive, hence, OECD has a lower prediction performance in terms of bias and during the periods like crisis. However, these differences are not statistically significant. When we use quadratic loss and Linex Loss 2, the statistics are negative, therefore, overall OECD makes predictions of lower variance but once more these differences are never statistically significant.

With absolute error loss the joint EPA hypothesis can be rejected if we use Kernel robust estimates using the truncated kernel. This is not the case when common factors are used to estimate the covariance matrix. Among the other loss functions, the joint EPA is rejected only with Linex Loss 2 when common factors or truncated kernel are used. In the light of the above CD tests we find covariance estimates using common factors more reliable. Hence, the conclusion is that the differences between the predictive ability of the two organizations are significant in periods of positive errors.

Table 12 reports the results of overall and joint EPA tests for G7 countries. As before, there is no strong evidence against the overall EPA hypothesis using any loss type. One exception is Linex Loss 2 with non-robust or Kernel robust estimates of the variance. In this sample too, there is evidence of superior predictive ability of the OECD in positive error periods like crisis.

In the light of the Monte Carlo results, we expect more power from joint EPA tests. As in this sample n is much smaller than T, the test $J_{n,T}^{(3)}$ is preferable as it is robust to SCD and WCD. With this test the joint EPA hypothesis is rejected for all loss functions except the absolute loss. Hence, we can conclude that the predictive ability of to organizations is statistically different for G7 countries and OECD has an overall better predictive performance.

Conclusion

This paper has been concerned with the problem of testing equal predictive ability hypothesis using panel data. The test which is proposed by Diebold and Mariano (1995) has been generalized to a panel data context taking into account the complications arise from using micro and macro data sets. We derived test statistics which are robust to different forms of cross-sectional dependence, arising either from spatial dependence (weak cross-sectional dependence) and common factors (strong cross-sectional dependence) in the forecast errors.

The small sample properties of the proposed tests have been found to be satisfactory in a large set of Monte Carlo simulations. In particular, the tests which are robust to strong cross-sectional dependence are found to be correctly sized in all experiments.

Table 11: Empirical Results for the Sample of OECD Countries

Overall Tests						Joint Tests					
Test	*Kernel*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*	*Test*	*Kernel*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*
$S_{n,T}^{(1)}$		0.1297	-0.3163	1.1938	-1.0047	$J_{n,T}^{(1)}$		32.5706	30.1859	37.9584	26.7331
		(0.8968)	(0.7518)	(0.2326)	(0.3150)			(0.2954)	(0.4048)	(0.1233)	(0.5861)
$S_{n,T}^{(2)}$	*Truncated*	0.1213	-0.2650	0.9705	-1.0101	$J_{n,T}^{(2)}$	*Truncated*	123.1218	28.8017	31.7652	7.5243
		(0.9034)	(0.7910)	(0.3318)	(0.3125)			(0.0000)	(0.4754)	(0.3303)	(1.0000)
	Bartlett	0.1258	-0.2956	1.1424	-1.0060		*Bartlett*	33.0204	27.3033	30.4338	26.8867
		(0.8999)	(0.7675)	(0.2533)	(0.3144)			(0.2769)	(0.5553)	(0.3926)	(0.5778)
	Parzen	0.1277	-0.3059	1.1796	-1.0046		*Parzen*	32.5760	27.5577	31.7036	25.6252
		(0.8984)	(0.7597)	(0.2382)	(0.3151)			(0.2952)	(0.5416)	(0.3330)	(0.6454)
	Tukey	0.1262	-0.2977	1.1622	-1.0053		*Tukey*	33.4938	27.7226	32.1140	26.3184
		(0.8995)	(0.7659)	(0.2451)	(0.3147)			(0.2583)	(0.5328)	(0.3149)	(0.6084)
	QS	0.1239	-0.2885	1.1201	-1.0061		*QS*	33.2744	28.7395	33.8395	29.3777
		(0.9014)	(0.7729)	(0.2627)	(0.3144)			(0.2668)	(0.4787)	(0.2452)	(0.4455)
$S_{n,T}^{(3)}$		0.1057	-0.2064	0.7407	-1.0312	$J_{n,T}^{(3)}$		-	-	-	-
		(0.9158)	(0.8365)	(0.4589)	(0.3024)			-	-	-	-
$S_{n,T}^{(4)}$		0.1045	-0.2046	0.7297	-1.0035	$J_{n,T}^{(4)}$		37.7560	37.3381	35.1717	49.9706
		(0.9168)	(0.8379)	(0.4656)	(0.3156)			(0.1279)	(0.1378)	(0.1990)	(0.0091)
$S_{n,T}^{(5)}$	*Bartlett*	0.1042	-0.2046	0.7298	-1.0035	$J_{n,T}^{(5)}$	*Bartlett*	33.6980	33.3348	33.9520	40.9475
		(0.9170)	(0.8379)	(0.4655)	(0.3156)			(0.2505)	(0.2644)	(0.2410)	(0.0696)

Note: Overall Tests and Joint Tests refer to the tests of the hypothesis (4) and (5) which are described in Section 3. The values shown in parantheses are p-values.

Table 12: Empirical Results for the Sample of G7 Countries

Overall Tests						Joint Tests					
Test	*Kernel*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*	*Test*	*Kernel*	*Absolute Loss*	*Quadratic Loss*	*Linex Loss 1*	*Linex Loss 2*
$S_{n,T}^{(1)}$		-0.4861	-1.1556	-0.5402	-1.7894	$J_{n,T}^{(1)}$		7.1185	8.0273	6.5588	4.2006
		(0.6269)	(0.2478)	(0.5891)	(0.0736)			(0.4166)	(0.3302)	(0.4762)	(0.7564)
$S_{n,T}^{(2)}$	*Truncated*	-0.4751	-1.1995	-0.6246	-1.6821	$J_{n,T}^{(2)}$	*Truncated*	8.8782	9.0841	6.7925	5.9435
		(0.6347)	(0.2303)	(0.5322)	(0.0925)			(0.2615)	(0.2467)	(0.4508)	(0.5464)
	Bartlett	-0.4802	-1.1534	-0.5414	-1.7830		*Bartlett*	7.1567	8.1210	6.7590	4.2936
		(0.6311)	(0.2488)	(0.5882)	(0.0746)			(0.4127)	(0.3220)	(0.4544)	(0.7454)
	Parzen	-0.4852	-1.1554	-0.5405	-1.7887		*Parzen*	7.1087	8.0301	6.5785	4.2051
		(0.6275)	(0.2479)	(0.5888)	(0.0737)			(0.4177)	(0.3299)	(0.4740)	(0.7559)
	Tukey	-0.4827	-1.1546	-0.5414	-1.7863		*Tukey*	7.1010	8.0529	6.6536	4.2306
		(0.6293)	(0.2483)	(0.5883)	(0.0741)			(0.4184)	(0.3280)	(0.4658)	(0.7529)
	QS	-0.4792	-1.1629	-0.5537	-1.7658		*QS*	7.2411	8.2040	6.5569	4.3921
		(0.6318)	(0.2449)	(0.5798)	(0.0774)			(0.4042)	(0.3150)	(0.4764)	(0.7337)
$S_{n,T}^{(3)}$		-0.3713	-1.0890	-0.6421	-1.4169	$J_{n,T}^{(3)}$		9.6218	19.2054	16.3559	18.7140
		(0.7104)	(0.2762)	(0.5208)	(0.1565)			(0.2110)	(0.0076)	(0.0221)	(0.0091)
$S_{n,T}^{(4)}$		-0.3644	-1.1699	-0.6268	-1.3859	$J_{n,T}^{(4)}$		7.5362	7.9669	7.8994	11.6253
		(0.7155)	(0.2420)	(0.5308)	(0.1658)			(0.3753)	(0.3355)	(0.3416)	(0.1136)
$S_{n,T}^{(5)}$	*Bartlett*	-0.3630	-1.1677	-0.6263	-1.3855	$J_{n,T}^{(5)}$	*Bartlett*	7.3604	8.0304	7.8994	11.5399
		(0.7166)	(0.2429)	(0.5311)	(0.1659)			(0.3923)	(0.3299)	(0.3416)	(0.1167)

Note: Overall Tests and Joint Tests refer to the tests of the hypothesis (4) and (5) which are described in Section 3. The values shown in parantheses are p-values.

This is the case even in the experiments which do not involve common factors but only spatial dependence. However, their power is generally low compared to test statistics which are robust only to spatial dependence, given that forecast errors do not contain common factors. In these cases, the Monte Carlo evidence suggest to use Bartlett and Parzen kernels for correctly sized test.

Finally, the tests have been used to compare the prediction performance of the two major organizations, the OECD and IMF, on their historical economic growth forecasts. We found that IMF has an overall better performance in terms of bias whereas OECD makes predictions with less variance but the difference is not statistically significant. In a sub-sample of G7 countries OECD predictions are found to be superior to that of IMF.

A possible extension of the testing procedures proposed in this paper is to allow to distinguish between the sources of the differences in predictive ability. As suggested in the paper, predictive ability of different forecasters may differ through periods while on average they have equal predictive power. To deal with this situation, the conditional EPA tests of GW can be extended to our panel data framework. This is an ongoing research agenda.

References

[1] Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, Wiley.

[2] Andrews, D.W. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817-858.

[3] Arellano, M. (1987), "Practitioners' corner: Computing robust standard errors for within groups estimators," *Oxford Bulletin of Economics and Statistics*, 49, 431-434.

[4] Bai, J. (2003), "Inferential theory for factor models of large dimensions," *Econometrica* , 71, 135-171.

[5] Bai, J. and S. Ng. (2002), "Determining the number of factors in approximate factor models," *Econometrica*, 70, 191-221.

[6] Bailey, N., Holly, S., and Pesaran, M. H. (2016), "A Two-Stage Approach to Spatio-Temporal Analysis with Strong and Weak Cross-Sectional Dependence," *Journal of Applied Econometrics*, 31, 249-280.

[7] Breusch, T. S., and Pagan, A. R. (1980), "The Lagrange multiplier test and its applications to model specification in econometrics," *The Review of Economic Studies*, 47, 239-253.

[8] Clark, T.E. and M.W. McCracken. (2001), "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics*, 105, 85-110.

[9] —— (2015), "Nested forecast model comparisons: a new approach to testing equal accuracy," *Journal of Econometrics*, 186, 160-177.

[10] Clark, T.E. and K.D. West. (2007), "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics* ,138, 291-311.

[11] Davies, A. and K. Lahiri. (1995), "A new framework for analyzing survey forecasts using three-dimensional panel data," *Journal of Econometrics*, 68, 205-228.

[12] Diebold, F.X. and R.S. Mariano. (1995), "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, 13, 253-263.

[13] Dreher, A., S. Marchesi and J.R. Vreeland. (2008), "The political economy of IMF forecasts," *Public Choice* , 137, 145-171.

[14] Driscoll, J.C. and A.C. Kraay. (1998), "Consistent covariance matrix estimation with spatially dependent panel data," *Review of Economics and Statistics*, 80, 549-560.

[15] Driver, C., L. Trapani and G. Urga. (2013), "On the use of cross-sectional measures of forecast uncertainty," *International Journal of Forecasting*, 29, 367-377.

[16] Giacomini, R. and H. White. (2006), "Tests of conditional predictive ability," *Econometrica*, 74, 1545-1578.

[17] Kelejian, H. H., and Prucha, I. R. (1998), "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances," *The Journal of Real Estate Finance and Economics*, 17, 99-121.

[18] —— (2007), "HAC estimation in a spatial framework," *Journal of Econometrics*, 140, 131-154.

[19] Lahiri, K. and X. Sheng. (2010), "Measuring forecast uncertainty by disagreement: The missing link," *Journal of Applied Econometrics* ,25, 514-538.

[20] Mayer, T. and Zignago, S. (2011), "Notes on CEPII's distances measures: the GeoDist Database," *CEPII Working Paper*, 2011-25.

[21] Merola, R. and J.J. Perez. (2013), "Fiscal forecast errors: governments versus independent agencies?," *European Journal of Political Economy*, 32, 285-299.

[22] Moscone, F. and E. Tosetti. (2012), "HAC estimation in spatial panels," *Economics Letters* , 17, 60-65.

[23] Pain, N., C. Lewis, T.-T. Dang, Y. Jin and P. Richardson. (2014), "OECD forecasts during and after the financial crisis," *OECD Economics Department Working Papers*, No. 1107, OECD Publishing, Paris.

[24] Pesaran, M. H. (2015), "Testing weak cross-sectional dependence in large panels," *Econometric Reviews*, 34, 1089-1117.

[25] Pesaran, M.H. and E. Tosetti. (2011), "Large panels with common factors and spatial correlation," *Journal of Econometrics* , 161, 182-202.

[26] Pons, J. (2000), "The accuracy of IMF and OECD forecasts for G7 countries," *Journal of Forecasting*, 19, 53-63.

[27] Sarafidis, V., and Wansbeek, T. (2012), "Cross-sectional dependence in panel data analysis," *Econometric Reviews*, 31, 483-531.

[28] Stock, J. H., and Watson, M. W. (2002), "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, 97, 1167-1179.

[29] Timmermann, A. and Y. Zhu. (2019), "Comparing forecasting performance with panel data," unpublished manuscript.

[30] Turner, D. (2017. "Designing fan charts for GDP growth forecasts to better reflect downturn risks," *OECD Economics Department Working Papers*, No. 1428, OECD Publishing, Paris.

[31] Vuchelen, J. and M.-I. Gutierrez. (2005), "A direct test of the information content of the OECD growth forecasts," *International Journal of Forecasting*, 21, 103-117.

[32] Vuong, Q.H. (1989), "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica*, 57, 307-333.

[33] West, K.D. (1996), "Asymptotic inference about predictive ability," *Econometrica*, 64, 1067-1084.

[34] White, H. (2001), *Asymptotic Theory for Econometricians*, Academic Press, Revised Edition.