Forecasting Using Heterogeneous Panels with Cross-Sectional Dependence

Oguzhan Akgun[a], Alain Pirotte[a], Giovanni Urga[b]

[a] *University of Paris 2 Pantheon-Assas, CRED*
[b] *Cass Business School, London, United Kingdom and Bergamo University, Italy*

Abstract

In this paper, we focus on forecasting heterogeneous panels in presence of cross-sectional dependence in terms of both spatial error dependence and common factors. We propose two main approaches to estimate the factor structure, one using the residuals ("Residuals Based Approach", RBA) while the second using a panel of some variables ("Auxiliary Variables Approach", AVA) to extract the factors. Small sample properties of the methods proposed is investigated through Monte Carlo simulation exercises and used in an application to predict house price inflation in OECD countries.
Keywords: Cross-Sectional dependence, Common factors, Spatial dependence, House price inflation, Inflation forecasting, Macroeconomic forecasting

1. Introduction

1.1. Overview and main contributions

The presence of both a cross-sectional and a time-series dimension makes the identification of optimal forecasts methods for panel data a particular challenging task and the literature on the issue is relatively scarce. A crucial role is played by the way in which we deal with cross-section dependence (CD), a natural feature of a panel of units. One strand of the literature focuses on the best linear unbiased predictor in spatial models: see amongst others Baltagi & Li (2004, 2006), Baltagi, Bresson & Pirotte (2012), Baltagi, Fingleton & Pirotte (2014). Another strand of the literature focuseson forecasting with panel data with common factors in the error terms, for instance, Hjalmarsson (2010), Karabiyik, Westerlund & Narayan (2016), Trapani & Urga (2009)

Email addresses: `Oguzhan.Akgun@u-paris2.fr` (Oguzhan Akgun), `Alain.Pirotte@u-paris2.fr` (Alain Pirotte), `g.urga@city.ac.uk` (Giovanni Urga)

which treat the common factors as nuisance parameters and make no attempt to use them to improve forecasts of a given panel unit.

In this paper, we consider the case of forecasting using a heterogeneous panel model which contains both unobserved common factors and spatial error dependence. We compare forecasting methods using global information to predict unit specific outcomes by means of a small number of common factors extracted from a large number of panel units.

We propose two alternative approaches. The first approach makes use of estimates of the common factors in the predictive model by applying principal components (PC) analysis on the residuals from a first stage consistent estimation of the model parameters. The unobserved nature of the common factors requires forecasting the future values of the estimated factors and then computing the predictions on the variable of interest. In the second approach, closely related to the diffusion index forecasting methodology of Stock & Watson (1998, 2002), the common factors are estimated from a number of auxiliary variables. In particular, in this paper common factors are potentially estimated from the realizations of the same variable for different panel units whereas in previous studies the factors come from a large number of indicators for the same panel unit. This second approach that we propose is similar to the one in Engel, Mark & West (2015). The authors show that even if the univariate exchange rate series contain little or no serial correlation, global information, estimated by means of common factors in a panel of exchange rates, can help predicting future exchange rates. Using simulated and real data, we compare forecasts generated by these two approaches with the forecasts using only unit-specific information.

In this paper, we also reconsider the question of pooling time series in the presence of CD. Pooling in heterogeneous panels can produce misleading results on the magnitude of the average effects and inference based on them (Baltagi, Bresson & Pirotte, 2008; Pesaran & Smith, 1995). However, when the estimates of the individual parameters contain too much noise, pooling can provide better out-of-sample forecasts (Mark & Sul, 2011). We investigate the role of CD on the optimal prediction strategy.

The final aim is to compare estimators recently proposed in the literature for panels containing unobserved common factors. We use methods by Pesaran (2006), Bai (2009), Song (2013) and related estimators for the slope parameters and compare their small sample performance in terms of prediction accuracy.

1.2. Related literature

Our paper is related to three different strands of the literature on forecasting with panel data. First, it is related to the large body of literature on the comparison of the pooled and heterogeneous estimators of the slope parameters in terms of forecasting accuracy, as recently revisited by Pesaran & Zhou (2018), Wang, Zhang & Paap (2018). A large number of papers compares the performance of alternative estimators in terms of their predictive ability. For instance, the main finding in Garcia-Ferrer, Highfield, Palm & Zellner (1987), Baltagi & Griffin (1997), Baltagi, Griffin & Xiong (2000), Baltagi, Bresson, Griffin & Pirotte (2003), Baltagi, Bresson & Pirotte (2004), Trapani & Urga (2009) the superiority of homogeneous estimators. However, Hoogstrate, Palm & Pfann (2000) point out that the superiority of the pooled estimators is a result of the sample size such that as the number of time series observations increases heterogeneous estimators become advantageous. Mark & Sul (2012) show that the potential gain from pooling is determined by the degree of heterogeneity and the empirical application on the exchange rate forecasts confirms this theoretical results. Thus, our paper is directly linked to this literature as we compare the forecasting performance of recently proposed pooled and heterogeneous estimators.

Second, there is an important number of studies which evaluates the effect of CD on the forecast performance. The contributions studying the effect of CD on forecasting with panel data can be divided into two main groups, with the first focusing one studying spatial dependence and the second emphasizing the role of using common factors. Among others, Baltagi & Li (2004, 2006), Baltagi et al. (2012, 2014) study the optimal predictors in different types of random effects panel models with spatial interactions. These studies underline the possibility of improving the unit specific forecasts using information from other units in the panel. For instance, in the case of error spatial dependence Baltagi & Li (2004) shows that a weighted sum of the residuals from all units in the panel data set contributes to the optimal prediction of each single unit. Our paper is linked to this literature as we study the impact of spatial dependence on forecasting in panel data. However, it worth noting that we assume that the time invariant effects are fixed parameters.

Third, this paper is also related to the time series literature on the diffusion index forecasts as it benefits from these studies in terms of forecasting using common factors. In spatial panels a weight matrix has to be specified to realize the forecasts. Another possibility to exploit the panel-wide information to improve the unit specific outcomes is to use common factors. Using data on the

Canadian regional growth rates, Kopoin, Moran & Paré (2013) showed that the forecasts which use national and international information are significantly better than those which use only regional information. Engel et al. (2015) used data from several OECD countries to improve the forecasts of exchange rates of individual countries. Their approach is similar to that of Stock & Watson (1999). The difference is that in Stock & Watson (1999) the common factors are estimated from a large number of predictors, whereas Engel et al. (2015) estimate the common factors from a large number of countries. The latter paper is very close to ours contribution in terms of using statistical methods to estimate factors common to countries to forecast individual outcomes, while instead we study the possibility of extracting information from different variables on each panel unit.

1.3. Organization

The remainder of the paper is organised as follows. In Section 2, we introduce the panel predictive model, the two approaches of forecasting with unobserved common factors, and we also briefly describe estimation methods implement. In Section 3, we evaluate the small sample properties of the forecast methods and estimators via an extensive Monte Carlo analysis. Section 4 contains an empirical exercise to illustrate the forecast performance of these methods using data on house price inflation in OECD countries. Section 5 concludes.

2. Panel forecasting model and methods of forecasting

2.1. The forecasting model

We consider stationary predictive panel data model with CD in the disturbances. The h-steps ahead variable $y_{i,t+h}$, $h \geq 0$, $i = 1, 2, \ldots, n$, $t = 1, 2, \ldots, T$, is given by

$$y_{i,t+h} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \boldsymbol{\gamma}_i^{y\prime} \mathbf{f}_t^y + u_{i,t+h}, \tag{1}$$

$$u_{i,t+h} = \sum_{j=1}^{n} r_{ij} \varepsilon_{j,t+h}, \tag{2}$$

where $\mathbf{x}_{it} = (x_{i1t}, x_{i2t}, \ldots, x_{ik_xt})'$ is a $(k_x \times 1)$ vector of observed individual-specific regressors which can include predetermined variables, $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{ik_x})'$ represents the corresponding $(k_x \times 1)$ slope parameters, r_{ij} are unknown spatial weights, ε_{it} is an error term which is uncorrelated over time and individuals. $\mathbf{f}_t^y$ is a vector of unobservable common factors of size m_y, $\boldsymbol{\gamma}_i^y$ are the associated $(m_y \times 1)$ factor loadings. α_i are the unit specific time-invariant effects. Unless otherwise specified, $\boldsymbol{\beta}_i$, $\boldsymbol{\gamma}_i^y$ and α_i are assumed to be fixed parameters.

The model in (2) contains as special cases all commonly used spatial processes like spatial autoregression (SAR), spatial moving average (SMA), spatial error components (SEC) and their higher order versions. Moreover, it can be rewritten in the form of a factor model of n factors and without an idiosyncratic component as $u_{i,t+h} = \mathbf{r}'_{i.}\boldsymbol{\varepsilon}_{.,t+h}$, where $\mathbf{r}_{i.} = (r_{i1}, r_{i2}, \dots, r_{in})'$ and $\boldsymbol{\varepsilon}_{.,t+h} = (\varepsilon_{1,t+h}, \varepsilon_{2,t+h}, \dots, \varepsilon_{n,t+h})'$. To distinguish between the two components of the model defining the cross-sectional interactions it requires some restrictions on the spatial weights r_{ij}, $i, j = 1, 2, \dots, n$. The standard assumption in the spatial econometrics literature is that the $n \times n$ matrix $\mathbf{R} = [r_{ij}]$ has bounded row and column norms for all n. In this case, (2) carries weak CD (WCD). Furthermore, the existence of m_y distinct common factors requires that $\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{f}_t^y \mathbf{f}_t^{y\prime} = \boldsymbol{\Sigma}_{\mathbf{f}^y}$ and $\operatorname*{plim}_{n\to\infty} \frac{1}{n}\boldsymbol{\Gamma}'\boldsymbol{\Gamma} = \boldsymbol{\Sigma}_\Gamma$ are both $m_y \times m_y$ positive definite matrices, where $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots, \boldsymbol{\gamma}_n)'$. In this case, the m_y common factors are called "strong common factors" (Chudik, Pesaran & Tosetti, 2011). Hence, the model contains strong CD (SCD) as well as WCD.

2.2. Forecasting approaches

We are interested in post-sample forecasting as defined in Granger & Huang (1997, p.3). Assuming that the expectation of $u_{i,t+h}$ conditional on past information is zero for all panel units, i.e. $\mathrm{E}(u_{i,t+h}|y_{it}, \mathbf{x}_{it}, \mathbf{f}_t^y, y_{i,t-1}, \mathbf{x}_{i,t-1}, \mathbf{f}_{t-1}^y, \dots) = 0$ for any $h > 0$ and for every $i = 1, 2, \dots, n$, the optimal predictor of the variable of interest in period $T + h$ given the information in T is

$$y_{i,T+h|T} = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_{iT} + \boldsymbol{\gamma}_i^{y\prime}\mathbf{f}_T^y. \tag{3}$$

In the case of slope homogeneity $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ but throughout the section we will use the heterogeneous notation for simplicity. This predictor is unfeasible as it contains the unknown coefficients and the unobserved common factors. Replacing these unknown quantities by their estimates, the feasible predictor is given by

$$\widehat{y}_{i,T+h|T} = \widehat{\alpha}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{x}_{iT} + \widehat{\boldsymbol{\gamma}}_i^{y\prime}\widehat{\mathbf{f}}_T^y. \tag{4}$$

The main issue is that the unobservable common factors have to be estimated from the data. One possibility is to estimate the parameters $\boldsymbol{\beta}_i$ and $\boldsymbol{\alpha}_i$ using the estimators robust to unobserved common factors and collecting the residuals

$$\widehat{e}_{it} = y_{it} - \widehat{\alpha}_i - \widehat{\boldsymbol{\beta}}_i'\mathbf{x}_{i,t-h}, \quad t = h + 1, \dots, T, \tag{5}$$

where $\widehat{\boldsymbol{\beta}}_i$ is a consistent estimate of $\boldsymbol{\beta}_i$ and

$$\widehat{\alpha}_i = \frac{1}{T-h} \sum_{t=h+1}^{T} \left(y_{it} - \widehat{\boldsymbol{\beta}}_i' \mathbf{x}_{i,t-h} \right). \quad (6)$$

Given that the estimates $\widehat{\boldsymbol{\alpha}}_i$ and $\widehat{\boldsymbol{\beta}}_i$ are consistent, these residuals consistently estimate $e_{it} = \boldsymbol{\gamma}_i^{y\prime} \mathbf{f}_{t-h}^{y} + u_{it}$ for the sample $i = 1, \ldots, n$, $t = h+1, \ldots, T$. Hence, it is possible to apply PC on these residuals to estimate the common factors $\mathbf{f}_{t-h}^{y}$ in $t = h+1, \ldots, T$. Let us denote these estimates as $\widehat{\mathbf{f}}_t^{y}$. Notice that the last possible estimates are in period $T-h$. However, the prediction in (3) requires the estimates of the unobserved common factors in period T, i.e. $\widehat{\mathbf{f}}_T^{y}$. Therefore, the factors need to be forecast from their estimates to make the prediction feasible. For simplicity let us assume that each common factor follow an AR(1) model. Then, such a forecast is

$$\tilde{\mathbf{f}}_T^{y} = \widehat{\boldsymbol{\Pi}}' \widehat{\mathbf{f}}_{T-h}^{y}, \quad \widehat{\boldsymbol{\Pi}} = \left(\widehat{\mathbf{f}}_{-2h}^{y\prime} \widehat{\mathbf{f}}_{-2h}^{y} \right)^{-1} \widehat{\mathbf{f}}_{-2h}^{y\prime} \widehat{\mathbf{f}}_{-h}^{y}, \quad (7)$$

where $\widehat{\mathbf{f}}_{-2h} = (\widehat{\mathbf{f}}_1', \widehat{\mathbf{f}}_2', \ldots, \widehat{\mathbf{f}}_{i,T-2h}')'$, $\widehat{\mathbf{f}}_{-h} = (\widehat{\mathbf{f}}_{h+1}', \widehat{\mathbf{f}}_2', \ldots, \widehat{\mathbf{f}}_{i,T-h}')'$. Then the prediction can be computed as

$$\widehat{y}_{i,T+h|T}^{R} = \widehat{\alpha}_i + \widehat{\boldsymbol{\beta}}_i' \mathbf{x}_{iT} + \widehat{\boldsymbol{\gamma}}_i^{y\prime} \tilde{\mathbf{f}}_T^{y}. \quad (8)$$

We call this the Residual Based Approach (RBA).

An alternative approach is to estimate the factors from the explanatory variables $\mathbf{x}_{it}$ by PC, supposing that they have a factor representation as $\mathbf{x}_{it} = \mathbf{a}_i^x + \boldsymbol{\Gamma}_i^{x\prime} \mathbf{f}_t^x + \mathbf{v}_{it}$ where $\mathbf{f}_t^x$ is a vector of common factors of size m_x, $\boldsymbol{\Gamma}_i^x$ are their loadings, $\mathbf{a}_i^x$ are the fixed effects and $\mathbf{v}_{it}$ is a vector error term which can be autocorrelated and can contain WCD. If $\mathbf{f}_t^y \subseteq \mathbf{f}_t^x$ this method can be used by estimating $\mathbf{f}_t^x$ from $\mathbf{x}_{it}$ for $i = 1, \ldots, n$, $t = 1, \ldots, T$, and plugging them in the predictive model (3). However, the condition $\mathbf{f}_t^y \subseteq \mathbf{f}_t^x$ may not be always reasonable in practice. When this is not satisfied we can assume that some auxiliary variables $\mathbf{w}_{it}$ are observable which satisfy

$$\mathbf{w}_{it} = \mathbf{a}_i^w + \boldsymbol{\Gamma}_i^{w\prime} \mathbf{f}_t^w + \boldsymbol{\varsigma}_{it}, \quad t = 1, \ldots, T, \quad (9)$$

such that $\mathbf{f}_t^y \subseteq \mathbf{f}_t^w$, where $\mathbf{w}_{it} = (w_{i1t}, w_{i2t}, \ldots, w_{ik_w t})'$ is a $(k_w \times 1)$ vector of auxiliary observed individual-specific variables, $\mathbf{a}_i^w$ are the fixed effects, $\boldsymbol{\Gamma}_i^w$ is the $(m_w \times k_w)$ matrix of factor loadings associated with the $(m_w \times 1)$ unobservable common factors $\mathbf{f}_t^w$. Note that $\mathbf{w}_{it}$ can contain $\mathbf{x}_{it}$ as its components. Then the prediction methodology, which we call Auxiliary Variables Approach (AVA), is based on the following four steps:

Step 1: Use any estimator which controls for unobserved common factors as described in the next subsection and compute the residuals in (5). In a pooling case, $\widehat{\boldsymbol{\beta}}_i$ should be replaced by the appropriate pooled estimator.

Step 2: Use principal components methods in the spirit of Bai (2003) to extract m_w common factors $\mathbf{f}_t^w$ from observed variables $\mathbf{w}_{it}$, $t = 1, \ldots, T$.

Step 3: Estimate the factor loadings $\widehat{\boldsymbol{\gamma}}_i^w$ by OLS on the regression

$$\widehat{e}_{it} = \boldsymbol{\gamma}_i^{w\prime}\widehat{\mathbf{f}}_{t-h}^w + \nu_{it}, \quad t = h+1, \ldots, T. \tag{10}$$

Step 4: Compute the prediction $\widehat{y}_{i,T+h|T}^A$ using

$$\widehat{y}_{i,T+h|T}^A = \widehat{\alpha}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{x}_{iT} + \widehat{\boldsymbol{\gamma}}_i^{w\prime}\widehat{\mathbf{f}}_T^w. \tag{11}$$

Remark 1. Both approaches require information on the number of common factors contained in the respective variables. In the RBA m_y (the number of factors in the process e_{it}) and in the AVA m_w (the number of factors in the auxiliary variables $\mathbf{w}_{it}$) need to be known. These can be consistently estimated using the methods proposed Bai & Ng (2002).

Remark 2. The AVA approach can be further improved by slightly modifying the fourth step. In the case that $m_w > m_y$, i.e. the number of common factors in $\mathbf{w}_{it}$ is strictly greater than that of the ones in e_{it}, the regression in Step 3 uses redundant common factors. To choose the correct number of factors in this step once more the information criteria of Bai & Ng (2002) can be used after suitable modifications on the number of parameters estimated.

Remark 3. It is not sure that the order of importance of common factors in the $\mathbf{w}_{it}$ equation will be the same as in the order of common factors in terms of their predictive ability on the dependent variable. For instance, in a macroeconomic study there can be regional factors which are most important for the countries in these regions. PC approach described here will order the common factors in terms of their global importance which may not be the valid ordering for all regions. In this case machine learning methods can be used to select the most important common factors for each panel unit in the same spirit as Bai & Ng (2008, 2009).

Remark 4. In the RBA, the factors are required to carry serial correlation. Given that this is so, similar to the second remark, in the RBA the prediction of the common factors given in (7) requires choosing the optimal model for the unobserved common factors. Since the PC estimates of the common factors are orthogonal by construction, any univariate time series model can be used on each estimate for prediction. As in the Monte Carlo simulations below the common factors are generated as an AR(1), we use this single lag prediction.

Remark 5. The AVA approach may seem restrictive as it requires additional observable variables to estimate the common factors. However, it is reasonable to assume that the forecaster has access to many variables which are potentially correlated with the common factors of interest. This is the assumption behind the diffusion index methodology of forecasting which proved useful ever since the seminal work of Stock & Watson (1998). In fact, it can be the case that the forecaster has access to an information set which is too large such that it is hard to extract the useful information to perform forecasts.[1] In this case, the variables to include in $\mathbf{w}_{it}$ can be chosen using the penalized regression methods as in Bai & Ng (2008, 2009).

If additional variables do not exist, the observed variables $\mathbf{x}_{it}$ can be still used to estimate the common factors given that they are correlated linearly with the unobserved common factors $\mathbf{f}_t^y$. When these variables include all common factors in the DGP for the dependent variable, there is no need to use additional observables. This is a testable hypothesis. In this case we suggest a two step methodology which can be applied as follows: Notice that the residuals in Step 1 of the AVA contain all unobserved common factors in the vector $\mathbf{f}_t^y$. In an additional step, the forecaster can extract common factors only from $\mathbf{x}_{it}$ and run the regression of $\widehat{e}_{it}$ on these factors. If the residuals from this regression fail to reject the WCD hypothesis using the test of Pesaran (2015), Step 4 is applied with these factors.[2] If the WCD hypothesis is rejected, we suggest to apply PCA to the residuals from the regression of $\widehat{e}_{it}$ on the common factors extracted from $\mathbf{x}_{it}$ and forecast them to apply the RBA. This gives a hybrid solution between the two approaches.

2.3. Methods of estimation

The two approaches of forecasting with unobserved common factors described above require the estimation of the slope parameters. In what follows, we only briefly describe the estimation procedures. The details of the relevant estimation methods are reported in the Supplementary

[1] Boivin & Ng (2006) show that for forecasting purposes less data can be better than larger but noisy data.

[2] Juodis & Reese (2018) show that the pre-removal of unobserved common factors by means of subtracting cross-sectional averages causes an incidental parameters problem in testing for WCD. As a result, the WCD test proposed by Pesaran (2015) no longer has the standard normal asymptotic distribution. However, De Hoyos & Sarafidis (2006) show that Frees (1995, 2004) and Breusch & Pagan (1980) tests can be used to test for general CD.

Material. The estimators of the individual slope coefficients that we consider take the form

$$\widehat{\boldsymbol{\beta}}_{M,i} = \left(\mathbf{X}'_{i.,-h}\mathbf{M}_H\mathbf{X}_{i.,-h}\right)^{-1}\mathbf{X}'_{i.,-h}\mathbf{M}_H\mathbf{y}_{i.}, \tag{12}$$

and the pooled estimators of the average slope parameters have the form

$$\widehat{\boldsymbol{\beta}}_{M,P} = \left(\sum_{i=1}^{n}\mathbf{X}'_{i.,-h}\mathbf{M}_H\mathbf{X}_{i.,-h}\right)^{-1}\sum_{i=1}^{n}\mathbf{X}'_{i.,-h}\mathbf{M}_H\mathbf{y}_{i.}, \tag{13}$$

where $\mathbf{X}_{i.,-h} = (\mathbf{x}'_{i1}, \mathbf{x}'_{i2}, \ldots, \mathbf{x}'_{i,T-h})'$, $\mathbf{y}_{i.} = (y_{i,h+1}, y_{i2}, \ldots, y_{iT})'$, $\mathbf{M}_H = \mathbf{I}_{T-h} - \mathbf{H}(\mathbf{H}'\mathbf{H})^{-}\mathbf{H}'$ and $^{-}$ denotes Moore-Penrose inverse. The estimators we consider differ in the way they deal with the common factors, hence, the matrix $\mathbf{H}$ defines these different estimators.

The first class of the estimators is of *CCE*-type proposed by Pesaran (2006). These estimators use cross-sectional averages of the dependent variable and the explanatory variables as proxies for the common factors. For these estimators we set $\mathbf{H} = (\mathbf{e}_{T-h}, \bar{\mathbf{Z}})$ where $\mathbf{e}_{T-h}$ is a vector of ones of length $T-h$, $\bar{\mathbf{Z}} = (\bar{\mathbf{z}}'_{.1}, \bar{\mathbf{z}}'_{.2}, \ldots, \bar{\mathbf{z}}'_{.,T-h})'$, $\bar{\mathbf{z}}'_{.t} = n^{-1}\sum_{i=1}^{n}\mathbf{z}'_{it}$ and $\mathbf{z}_{it} = (y_{it}, \mathbf{x}'_{i,t-h})'$, $t = h+1, \ldots, T$. We call these estimators *Ind. CCE* and *CCEP*. A slightly modified versions of these estimators use the cross-sectional averages of only the exogenous variables, hence, $\mathbf{H} = (\mathbf{e}_{T-h}, \bar{\mathbf{W}})$ where $\bar{\mathbf{W}}$ is the matrix of observations on cross-sectional averages of some exogenous variables which can include the explanatory variables themselves. These are named as *Ind. CCEX* and *CCEPX*.

The second class of estimators includes the ones which use PC methods to estimate the common factors. First one is the iterative principal components estimator proposed by Bai (2009) which we call *IPCP*. The procedure starts with an initial estimation of the slope parameters $\boldsymbol{\beta}$ and the individual specific effects α_i. Let us denote these initial estimates $\widehat{\boldsymbol{\beta}}^{(0)}$ and $\widehat{\boldsymbol{\alpha}}_i^{(0)}$. Next, the common factor estimates are computed from the residuals $\widehat{e}_{it} = y_{it} - \widehat{\boldsymbol{\alpha}}_i^{(0)} - \widehat{\boldsymbol{\beta}}^{(0)\prime}\mathbf{x}_{i,t-h}$ using PC. Factor estimates update parameter estimates iteratively until numerical convergence is achieved. We can express this estimator by setting $\mathbf{H} = (\mathbf{e}_{T-h}, \widehat{\mathbf{F}}_I^u)$ where $\widehat{\mathbf{F}}_I^u$ is the matrix of observations on the common factor estimates after the numerical convergence is achieved for the slope parameters. The heterogeneous counterpart of this estimator was first used by Song (2013) and is called *Ind. IPC*. Another consistent estimator can be obtained by setting $\mathbf{H} = (\mathbf{e}_{T-h}, \widehat{\mathbf{F}}^x)$ where $\widehat{\mathbf{F}}^x$ is the matrix of observations on the common factor estimates obtained by PC on the explanatory variables. These estimators are called *Ind. PCX* and *PCPX*. These are also used as initial values for the iterative PC estimators. Although consistent, these estimators do not wipe out all common factors in (1) if the condition $\mathbf{f}_t^y \subseteq \mathbf{f}_t^x$ is not satisfied. An alternative is to set $\mathbf{H} = (\mathbf{e}_{T-h}, \widehat{\mathbf{F}}^x, \widehat{\mathbf{F}}^u)$ where $\widehat{\mathbf{F}}^u$ is the matrix of observations on the common factor estimates obtained by PC on the residuals in (1) which are computed using the estimators *Ind. PCX* or *PCPX*. We call these two-stage estimators *Ind. PCX2S* and *PCPX2S*.

Under general conditions, all these estimators are consistent for the individual parameters or their expected values as long as both dimensions of the panel get large and when the regressors are strictly exogeneous. Pesaran & Tosetti (2011) show that the CCE estimators are consistent under the assumption on the boundedness of the row and column sums of the matrix $\mathbf{R}$. The PC estimators require slightly stronger assumptions on the degree of the heteroskedasticity and dependence in either panel dimensions. The details of these are given in the Supplementary Material. CCE estimators also require a rank condition which we assume to hold.[3]

When the right hand side variables contain weakly exogeneous variables like predetermined variables, pooled estimators turn inconsistent for the average effect when the true model is heterogeneous (Pesaran & Smith, 1995) even when there are no unobserved common factors. For the CCE estimators to remain consistent in the existence of weakly exogeneous regressors, lags of cross-sectional averages have to be included in the estimation of individual equations. They also require the number of cross-sectional averages to be at least as large as the number of unobserved common factors (Chudik & Pesaran, 2015). As our main aim is to compare forecast performance, in our simulations we rely on strictly exogeneous regressors noting that otherwise pooled estimators are already outperformed by individual estimates. This is confirmed by simulations considering a dynamic model for which the results are reported in the Supplementary Material. In the application below, we use specifications with predetermined variables, paying attention to the requirements mentioned above. Namely, for to compute the estimates using *Ind. CCE* and *CCEP*, we add sufficient number of lags of cross-sectional averages as in (Chudik & Pesaran, 2015).

3. Monte Carlo study

3.1. Design of the experiments

The dependent and the explanatory variables are generated as follows:

$$y_{i,t+h} = \alpha_i + \beta_{i1} x_{i1t} + \beta_{i2} x_{i2t} + \gamma_{i1} f_{1t} + \gamma_{i2} f_{2t} + u_{i,t+h}, \tag{14}$$

$$x_{ijt} = a_{ij} + \gamma_{ij1} f_{1t} + \gamma_{ij3} f_{3t} + v_{ijt}, \quad j = 1, 2, \tag{15}$$

[3] See Karabiyik, Reese & Westerlund (2017) for a discussion on this topic.

where $i = 1, 2, \ldots, n$, $t = 1, 2, \ldots, T$, x_{ijt}, $j = 1, 2$, are the observed explanatory variables, f_{jt}, $j = 1, 2, 3$, are the unobserved common factors with loadings γ_{ijk}, α_i and a_{ij} are the fixed effects, and β_{ij} are the slope coefficients. The error term of the dependent variable carries spatial dependence and it is generated as a SAR using

$$u_{it} = \rho_i \sum_{j=1}^{n} w_{ij} u_{jt} + \varepsilon_{it}, \quad \text{where } \varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2), \quad \sigma_i^2 \sim \text{IIDU}(0.5, 1.5), \tag{16}$$

where w_{ij} is the element of the spatial weight matrix $\mathbf{W}_n$ in row i and column j. An SMA is also considered as a generating process but the results are similar and they are not reported here. A rook-type spatial weight matrix is used. We consider two different cases for ρ_i. These two cases are based on Baltagi & Pirotte (2010), with the main difference being heterogeneity of the parameters in the (first order) SAR (or SMA) models, where $\rho_i = \rho = (0.2, 0.8)$ which corresponds to low and high spatial dependence, respectively. Similarly, we generate the heterogeneous coefficients using

$$\rho_i = \rho + e_i^{\rho}, \quad \text{with } \rho = \{0.2, 0.8\}, \quad e_i^{\rho} \sim \text{U}(-0.1, 0.1). \tag{17}$$

The unobserved common factors are generated as follows

$$f_{jt} = \rho_{fj} f_{j,t-1} + v_{fjt}, \quad v_{fjt} \sim \mathcal{N}(0, 1 - \rho_{fj}^2), \quad \rho_{fj} = 0.5, \quad f_{j0} = 0, \quad j = 1, 2, 3. \tag{18}$$

The disturbances associated to the explanatory variables are generated by a stationary AR(1) process which is given by

$$v_{ijt} = \rho_{v_{ij}} v_{ij,t-1} + \epsilon_{ijt}, \quad \epsilon_{ijt} \sim \mathcal{N}(0, 1 - \rho_{v_{ij}}^2), \quad \rho_{v_{ij}} \sim \text{IIDU}(0.05, 0.95), \tag{19}$$

assuming that $v_{ij0} = 0$, $j = 1, 2$. The first 10 observations are discarded to minimize the impact of initial values. The slope coefficients β_{ij} are generated under two different assumptions corresponding to high and low heterogeneity. They are given by

$$\beta_{ij} = \beta_j + \eta_{ij}, \quad \beta_j = 1, \quad \eta_{ij} \sim \text{IID}\mathcal{N}(0, \sigma_{\eta_j}^2), \tag{20}$$

where $\sigma_{\eta_j}^2 = 0.15$ and $\sigma_{\eta_j}^2 = 0.3$, $j = 1, 2$, correspond to low and high heterogeneity, respectively. These heterogeneity levels in both cases are higher compared to those of Pesaran (2006), Pesaran & Tosetti (2011). The individual effects are generated as

$$\alpha_{i1} \sim \text{IID}\mathcal{N}(1, 1), \quad a_{ij} \sim \text{IID}\mathcal{N}(0.5, 0.5), \quad j = 1, 2 \tag{21}$$

and they are fixed for each replication. The loadings of the unobserved common factors in the equations for the explanatory variables are generated as

$$\begin{pmatrix} \gamma_{i11} & \gamma_{i13} \\ \gamma_{i21} & \gamma_{i23} \end{pmatrix} \sim \begin{pmatrix} \text{IID}\mathcal{N}(0.5, 0.5) & \text{IID}\mathcal{N}(0, 0.5) \\ \text{IID}\mathcal{N}(0, 0.5) & \text{IID}\mathcal{N}(0.5, 0.5) \end{pmatrix}. \tag{22}$$

To produce forecasts using the AVA an additional variable x_{i3t} is generated as

$$x_{i3t} = a_{i3} + \gamma_{i31} f_{1t} + \gamma_{i32} f_{2t} + v_{i3t}, \tag{23}$$

where the factor loadings are given by

$$a_{i3} \sim \text{IID}\mathcal{N}(1.5, 1.02), \quad \gamma_{i32}, \sim \text{IID}\mathcal{N}(1, 0.1). \tag{24}$$

The other terms in (23) are defined in the same way as those contained in explanatory variable DGPs (15).

Contrary to the case of the factor loadings in the process generating the explanatory variables x_{ijt} from different distributions, in this paper we follow Trapani & Urga (2009) and Phillips & Sul (2003) and draw loadings to generate low and high CD. This is controlled as follows

$$\gamma_{i1}, \gamma_{i2} \sim \begin{cases} \text{IID}\mathcal{N}(1, 0.1) \text{ for Low CD,} \\ \\ \text{IID}\mathcal{N}(2, 0.4) \text{ for High CD.} \end{cases} \tag{25}$$

The chosen parameters in (25) induce average correlation coefficients among panel units of 0.5 and 0.8, respectively. The full set of experiments is summarized in Table 1.

Table 1: Summary of Experiments

Cases	Description	Parametrization
- Case *1*	Low Spatial & Low Factor Dependence	$\rho = 0.2$, $\gamma_{i1}, \gamma_{i2} \sim \text{IID}\mathcal{N}(1, 0.1)$
- Case *2*	Low Spatial & High Factor Dependence	$\rho = 0.2$, $\gamma_{i1}, \gamma_{i2} \sim \text{IID}\mathcal{N}(2, 0.4)$
- Case *3*	High Spatial & Low Factor Dependence	$\rho = 0.8$, $\gamma_{i1}, \gamma_{i2} \sim \text{IID}\mathcal{N}(1, 0.1)$
- Case *4*	High Spatial & High Factor Dependence	$\rho = 0.8$, $\gamma_{i1}, \gamma_{i2} \sim \text{IID}\mathcal{N}(2, 0.4)$

We consider $(n, T) = \{20, 30, 50, 100\}$. For each experiment, 2,000 replications are performed. The results for the individual estimators *CCE*, *CCEX*, *IPC*, *PCX*, *PCX2S* and their pooled counterparts are reported. For *PC* estimators, we assume that the number of unobservable common factors are known.

The forecasts are computed for the ith individual at future period $T + h$, with $h = 1$ to compare the performance of two forecast approaches and estimator performance. We also tried $h = 5, 10$ using the AVA to compare the individual and pooled estimators and the results are available upon request. We use root mean squared error (RMSE) to measure the predictive accuracy defined as $\text{RMSE}_i = \sqrt{\frac{1}{h}\sum_{\tau=1}^{h}\left(\widehat{y}_{i,T+\tau} - y_{i,T+\tau}\right)^2}$ and to obtain a single measure, the average of the statistic across units is computed. The results are reported relative to the OLS benchmark which is computed using unit specific OLS estimates. Hence, they show the gain in forecast accuracy from using estimated common factors.

3.2. Simulation results

The results on the prediction performance of different estimators with the RBA and the AVA for the case of low heterogeneity are reported in Tables 2-5 whereas the results on the case of high heterogeneity are given in Tables 6-9.

Table 2 is concerned with the case of low spatial dependence and low factor dependence (*Case 1*), the forecast performance of any estimator is superior using the AVA compared to the RBA. When $n, T = 20$, for any given estimator the relative RMSE of the forecasts using the RBA is 1.4. As either T or n or both increases this ratio also increases and exceeds 1.5 when $n, T = 100$.

In the case of CCE estimators, it is seen that the individual estimators outperform the pooled estimators even in smallest samples. For instance, with AVA, the relative RMSE of *Ind. CCE* (0.723) is slightly better than its pooled counterpart (0.746) when $n, T = 20$ which gives a relative RMSE of 0.97. When T increases to 100 for the same n this ratio is 0.92 which shows that the relative performance of the individual estimator increases. This result is similar for other estimators except *Ind. PCX* and and its pooled version. When $T = 20$, for any n the pooled estimator ourperforms the individual estimator. This can be explained by the fact that these estimators do not control for all unobserved common factors in the DGP of the dependent variable, hence the individual estimators are affected more by the lower signal to noise ratio compared to pooled ones. However, as T increases, again the individual estimator is preferred.

Finally, a comparison of the best performing CCE and PC estimators shows that the CCE performs better in the case of small T and small n but PC improves and has a better performance when n gets large. For instance, with AVA, when $n, T = 20$, *Ind. CCE* and *Ind. IPC* have relative RMSEs equal to 0.723 and 0.733, respectively. When $n = 100$ for the same time dimension, these

Table 2: Relative RMSE – Low Heterogeneity, Case *1*: Low Spatial Dependence & Low Factor Dependence

	Individual									*Pooled*							
T	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
n	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	0.992	0.984	0.979	0.976	0.723	0.687	0.673	0.661	20	1.022	1.048	1.061	1.067	0.746	0.730	0.726	0.719
30	1.007	0.982	0.962	0.964	0.731	0.674	0.636	0.636	30	1.038	1.049	1.060	1.064	0.749	0.720	0.707	0.708
50	0.998	0.990	0.966	0.973	0.709	0.687	0.653	0.635	50	1.031	1.051	1.053	1.075	0.732	0.731	0.713	0.706
100	1.004	0.987	0.974	0.955	0.686	0.668	0.644	0.636	100	1.030	1.051	1.066	1.061	0.710	0.715	0.711	0.711
Ind. CCEX									*CCEPX*								
20	0.996	0.987	0.980	0.977	0.726	0.689	0.674	0.661	20	1.023	1.048	1.061	1.067	0.746	0.730	0.726	0.719
30	1.012	0.986	0.963	0.965	0.734	0.676	0.637	0.636	30	1.038	1.049	1.060	1.064	0.749	0.720	0.707	0.708
50	1.001	0.992	0.967	0.973	0.711	0.687	0.654	0.635	50	1.031	1.051	1.053	1.075	0.732	0.731	0.713	0.706
100	1.005	0.987	0.974	0.956	0.687	0.668	0.644	0.636	100	1.030	1.051	1.066	1.061	0.710	0.715	0.711	0.711
Ind. IPC									*IPCP*								
20	0.999	0.988	0.980	0.979	0.733	0.690	0.674	0.660	20	1.023	1.049	1.062	1.068	0.748	0.731	0.727	0.720
30	1.001	0.981	0.962	0.966	0.733	0.674	0.636	0.635	30	1.038	1.050	1.061	1.065	0.750	0.721	0.708	0.708
50	0.981	0.980	0.964	0.974	0.704	0.682	0.651	0.634	50	1.031	1.051	1.053	1.075	0.733	0.731	0.714	0.707
100	0.980	0.975	0.968	0.954	0.674	0.661	0.641	0.634	100	1.030	1.051	1.066	1.061	0.710	0.715	0.711	0.711
Ind. PCX									*PCPX*								
20	1.087	1.059	1.018	1.000	0.790	0.735	0.697	0.673	20	1.026	1.051	1.064	1.069	0.749	0.731	0.727	0.719
30	1.102	1.055	1.009	0.989	0.795	0.721	0.668	0.651	30	1.039	1.050	1.062	1.066	0.751	0.722	0.708	0.708
50	1.087	1.052	1.003	0.996	0.776	0.732	0.678	0.649	50	1.032	1.052	1.053	1.075	0.733	0.731	0.714	0.706
100	1.090	1.045	1.013	0.977	0.752	0.713	0.673	0.650	100	1.031	1.051	1.066	1.061	0.710	0.715	0.711	0.711
Ind. PCX2S									*PCPX2S*								
20	1.036	1.012	0.991	0.984	0.745	0.698	0.677	0.661	20	1.026	1.051	1.064	1.069	0.747	0.729	0.727	0.719
30	1.044	1.004	0.972	0.970	0.750	0.681	0.639	0.637	30	1.039	1.050	1.062	1.066	0.750	0.721	0.707	0.708
50	1.024	1.006	0.973	0.978	0.722	0.693	0.655	0.635	50	1.032	1.052	1.053	1.075	0.733	0.731	0.713	0.706
100	1.022	0.996	0.978	0.958	0.696	0.672	0.645	0.636	100	1.031	1.051	1.066	1.061	0.710	0.715	0.711	0.711

values are 0.686 and 0.674, respectively. When n and T are both large, the two estimators have very similar performance.

The results do not change significantly when we consider the case of low spatial dependence and high factor dependence (Case *2*) which are reported in Table 3. In this case the performance of the RBA is lower compared to the AVA, with the relative RMSE being about 1.8 for any estimator when $n, T = 20$. However, in this case the relative performance of RBA with respect to forecasts without common factors is improved compared to the previous case even in smallest samples. As in this case common factors have bigger variability in the DGP of the dependent variable the relative performance of the pooled estimators are better compared to the previous case. However, all estimators deal with these common factors in a successful manner. Hence, still the individual specific estimators are superior, once more except the *Ind. PCX* and and its pooled version.

Table 3: Relative RMSE – Low Heterogeneity, Case *2*: Low Spatial Dependence & High Factor Dependence

	Individual									*Pooled*							
T	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
n	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	0.925	0.933	0.938	0.941	0.496	0.467	0.455	0.445	20	0.939	0.964	0.979	0.985	0.508	0.491	0.486	0.479
30	0.943	0.932	0.918	0.928	0.495	0.448	0.416	0.417	30	0.959	0.965	0.963	0.975	0.505	0.474	0.457	0.460
50	0.931	0.937	0.926	0.942	0.471	0.454	0.426	0.412	50	0.947	0.967	0.969	0.992	0.484	0.480	0.462	0.456
100	0.938	0.941	0.933	0.915	0.447	0.432	0.414	0.408	100	0.950	0.971	0.978	0.966	0.461	0.461	0.455	0.455
Ind. CCEX									*CCEPX*								
20	0.931	0.936	0.939	0.942	0.501	0.470	0.457	0.446	20	0.939	0.964	0.979	0.985	0.508	0.491	0.486	0.479
30	0.947	0.935	0.919	0.928	0.500	0.450	0.417	0.418	30	0.959	0.965	0.963	0.975	0.505	0.474	0.457	0.460
50	0.934	0.939	0.927	0.942	0.473	0.455	0.427	0.413	50	0.947	0.967	0.969	0.992	0.484	0.480	0.462	0.456
100	0.940	0.941	0.934	0.915	0.448	0.433	0.415	0.409	100	0.951	0.971	0.978	0.966	0.461	0.461	0.455	0.455
Ind. IPC									*IPCP*								
20	0.934	0.937	0.939	0.944	0.502	0.467	0.453	0.443	20	0.939	0.965	0.980	0.985	0.509	0.491	0.486	0.479
30	0.942	0.933	0.918	0.929	0.495	0.446	0.413	0.416	30	0.958	0.965	0.964	0.975	0.506	0.475	0.457	0.460
50	0.923	0.933	0.926	0.943	0.466	0.450	0.424	0.411	50	0.947	0.967	0.969	0.992	0.484	0.481	0.462	0.456
100	0.927	0.935	0.931	0.915	0.440	0.428	0.412	0.407	100	0.951	0.971	0.978	0.966	0.461	0.461	0.455	0.455
Ind. PCX									*PCPX*								
20	1.065	1.042	0.998	0.976	0.627	0.561	0.506	0.473	20	0.944	0.969	0.982	0.987	0.515	0.494	0.488	0.480
30	1.085	1.038	0.986	0.963	0.623	0.542	0.478	0.449	30	0.961	0.967	0.966	0.976	0.509	0.477	0.458	0.461
50	1.069	1.034	0.983	0.977	0.605	0.549	0.480	0.443	50	0.948	0.969	0.969	0.992	0.486	0.482	0.463	0.456
100	1.074	1.033	0.995	0.949	0.583	0.528	0.476	0.440	100	0.951	0.971	0.978	0.966	0.463	0.461	0.455	0.455
Ind. PCX2S									*PCPX2S*								
20	0.956	0.954	0.947	0.946	0.517	0.478	0.458	0.445	20	0.944	0.969	0.982	0.987	0.509	0.491	0.486	0.479
30	0.968	0.947	0.925	0.932	0.512	0.454	0.418	0.418	30	0.961	0.967	0.966	0.976	0.506	0.474	0.457	0.460
50	0.947	0.948	0.931	0.945	0.482	0.460	0.427	0.412	50	0.948	0.969	0.969	0.992	0.484	0.480	0.462	0.456
100	0.950	0.946	0.936	0.917	0.457	0.436	0.415	0.409	100	0.951	0.971	0.978	0.966	0.461	0.461	0.455	0.455

Table 4: Relative RMSE – Low Heterogeneity, Case *3*: High Spatial Dependence & Low Factor Dependence

	Individual									*Pooled*							
T	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
n	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	1.017	1.014	1.008	1.009	0.883	0.842	0.853	0.834	20	1.003	1.035	1.045	1.060	0.877	0.854	0.871	0.859
30	1.040	1.017	1.002	0.996	0.890	0.844	0.821	0.809	30	1.016	1.034	1.050	1.052	0.876	0.854	0.849	0.841
50	1.045	1.028	0.999	0.999	0.882	0.852	0.803	0.791	50	1.011	1.034	1.043	1.056	0.862	0.856	0.827	0.824
100	1.051	1.019	1.005	0.986	0.854	0.824	0.809	0.786	100	1.009	1.029	1.043	1.043	0.832	0.830	0.833	0.822
Ind. CCEX									*CCEPX*								
20	1.060	1.038	1.024	1.016	0.912	0.857	0.862	0.838	20	1.003	1.036	1.045	1.060	0.877	0.854	0.871	0.859
30	1.072	1.036	1.013	1.000	0.911	0.855	0.828	0.811	30	1.016	1.034	1.050	1.052	0.876	0.854	0.849	0.841
50	1.066	1.037	1.004	1.002	0.895	0.857	0.806	0.792	50	1.012	1.034	1.043	1.056	0.862	0.856	0.827	0.824
100	1.060	1.023	1.008	0.988	0.861	0.827	0.811	0.787	100	1.009	1.029	1.043	1.043	0.832	0.830	0.833	0.822
Ind. IPC									*IPCP*								
20	0.983	0.990	0.996	1.003	0.872	0.836	0.851	0.835	20	1.003	1.036	1.045	1.061	0.877	0.855	0.871	0.860
30	0.992	0.987	0.986	0.988	0.873	0.837	0.820	0.809	30	1.015	1.034	1.051	1.052	0.875	0.854	0.850	0.841
50	0.989	0.992	0.985	0.989	0.861	0.840	0.800	0.791	50	1.012	1.034	1.044	1.057	0.862	0.856	0.828	0.824
100	0.992	0.988	0.986	0.978	0.831	0.814	0.806	0.787	100	1.009	1.030	1.044	1.043	0.831	0.830	0.833	0.822
Ind. PCX									*PCPX*								
20	1.123	1.088	1.051	1.033	0.942	0.880	0.872	0.845	20	1.007	1.039	1.047	1.062	0.879	0.855	0.872	0.860
30	1.129	1.080	1.044	1.017	0.939	0.877	0.843	0.819	30	1.017	1.035	1.052	1.053	0.877	0.855	0.850	0.841
50	1.115	1.074	1.028	1.017	0.924	0.878	0.819	0.800	50	1.013	1.036	1.044	1.056	0.863	0.857	0.828	0.824
100	1.109	1.058	1.030	1.001	0.891	0.849	0.824	0.794	100	1.009	1.030	1.043	1.043	0.832	0.830	0.833	0.822
Ind. PCX2S									*PCPX2S*								
20	1.055	1.035	1.016	1.014	0.897	0.848	0.853	0.833	20	1.007	1.039	1.047	1.062	0.877	0.854	0.871	0.859
30	1.069	1.031	1.008	1.000	0.901	0.847	0.822	0.809	30	1.017	1.035	1.052	1.053	0.876	0.854	0.849	0.841
50	1.066	1.035	1.003	1.001	0.889	0.853	0.803	0.790	50	1.013	1.036	1.044	1.056	0.862	0.856	0.827	0.824
100	1.069	1.026	1.008	0.988	0.861	0.826	0.809	0.786	100	1.009	1.030	1.043	1.043	0.832	0.830	0.833	0.822

The results for the case of high spatial dependence and low factor dependence (Case *3*) are reported in Table 4. Once more the performance of the RBA is lower compared to the AVA but their performances seem closer in this case with the relative RMSE being about 1.2 in smallest samples. Here, with RBA, the relative performance of the pooled estimators is better than the individual estimators when T is small, one exception being the *Ind. IPC* and $IPCP$ estimators. Here, the best performing estimator is *Ind. IPC* in all samples sizes. These conclusions are equally valid for the case of high spatial dependence and high factor dependence (Case *4*) for which the results are given in Table 5.

The results above are confirmed in the case of high heterogeneity reported in Tables 6-9. However, in this case even when we have spatial dependence individual estimators perform better than their pooled counterparts. To summarize, (i) the AVA outperforms the RBA in all cases; (ii) indi-

Table 5: Relative RMSE – Low Heterogeneity, Case *4*: High Spatial Dependence & High Factor Dependence

	Individual									*Pooled*							
T	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
n	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	0.961	0.966	0.974	0.977	0.674	0.627	0.648	0.629	20	0.952	0.978	0.997	1.007	0.668	0.634	0.661	0.647
30	0.982	0.968	0.956	0.959	0.671	0.623	0.597	0.592	30	0.968	0.978	0.983	0.991	0.660	0.630	0.616	0.614
50	0.979	0.977	0.955	0.967	0.654	0.629	0.577	0.571	50	0.959	0.981	0.981	0.999	0.639	0.632	0.594	0.594
100	0.982	0.971	0.963	0.942	0.613	0.590	0.584	0.562	100	0.958	0.977	0.986	0.975	0.598	0.595	0.601	0.587
Ind. CCEX									*CCEPX*								
20	0.988	0.980	0.983	0.980	0.695	0.638	0.656	0.632	20	0.952	0.978	0.997	1.007	0.668	0.634	0.661	0.647
30	1.003	0.981	0.963	0.962	0.687	0.633	0.602	0.594	30	0.968	0.978	0.983	0.992	0.660	0.630	0.616	0.614
50	0.993	0.984	0.958	0.968	0.663	0.633	0.579	0.572	50	0.959	0.981	0.981	0.999	0.639	0.632	0.594	0.594
100	0.988	0.974	0.965	0.943	0.618	0.593	0.585	0.563	100	0.958	0.977	0.986	0.975	0.598	0.595	0.601	0.587
Ind. IPC									*IPCP*								
20	0.947	0.955	0.967	0.973	0.669	0.624	0.648	0.631	20	0.953	0.978	0.997	1.007	0.668	0.634	0.661	0.647
30	0.956	0.952	0.948	0.955	0.661	0.621	0.598	0.594	30	0.968	0.978	0.984	0.992	0.660	0.630	0.616	0.615
50	0.948	0.958	0.949	0.961	0.641	0.624	0.577	0.574	50	0.960	0.981	0.981	0.999	0.639	0.632	0.594	0.595
100	0.953	0.957	0.955	0.941	0.601	0.585	0.583	0.562	100	0.958	0.977	0.986	0.976	0.597	0.595	0.601	0.588
Ind. PCX									*PCPX*								
20	1.093	1.062	1.028	1.008	0.770	0.695	0.683	0.648	20	0.957	0.982	0.999	1.009	0.672	0.637	0.662	0.647
30	1.105	1.056	1.014	0.989	0.761	0.687	0.638	0.613	30	0.971	0.979	0.985	0.993	0.662	0.631	0.617	0.615
50	1.089	1.053	1.001	0.993	0.741	0.688	0.612	0.591	50	0.961	0.983	0.981	0.999	0.640	0.633	0.594	0.595
100	1.087	1.043	1.010	0.968	0.701	0.652	0.621	0.582	100	0.958	0.977	0.986	0.976	0.599	0.595	0.601	0.588
Ind. PCX2S									*PCPX2S*								
20	0.993	0.985	0.981	0.982	0.690	0.635	0.650	0.629	20	0.957	0.982	0.999	1.009	0.668	0.634	0.661	0.646
30	1.008	0.982	0.962	0.963	0.684	0.628	0.598	0.592	30	0.971	0.979	0.985	0.993	0.660	0.630	0.616	0.614
50	0.998	0.986	0.959	0.969	0.663	0.632	0.577	0.571	50	0.961	0.983	0.981	0.999	0.639	0.632	0.594	0.594
100	0.996	0.977	0.966	0.944	0.621	0.594	0.585	0.563	100	0.958	0.977	0.986	0.976	0.598	0.595	0.601	0.587

Table 6: Relative RMSE – High Heterogeneity, Case *1*: Low Spatial Dependence & Low Factor Dependence

	Individual									*Pooled*							
T	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
n	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	0.993	0.985	0.979	0.977	0.724	0.688	0.673	0.661	20	1.103	1.133	1.146	1.151	0.807	0.792	0.788	0.780
30	1.008	0.983	0.962	0.965	0.732	0.674	0.637	0.636	30	1.121	1.138	1.158	1.156	0.808	0.787	0.782	0.778
50	0.998	0.991	0.966	0.973	0.710	0.687	0.653	0.635	50	1.118	1.140	1.142	1.170	0.797	0.797	0.780	0.777
100	1.004	0.987	0.974	0.956	0.686	0.668	0.644	0.636	100	1.118	1.140	1.164	1.158	0.776	0.784	0.784	0.785
Ind. CCEX									*CCEPX*								
20	0.996	0.987	0.980	0.977	0.726	0.689	0.674	0.661	20	1.103	1.134	1.146	1.151	0.807	0.792	0.788	0.780
30	1.012	0.986	0.963	0.965	0.734	0.676	0.637	0.636	30	1.121	1.138	1.158	1.156	0.808	0.786	0.782	0.778
50	1.001	0.992	0.967	0.973	0.711	0.687	0.654	0.635	50	1.118	1.140	1.142	1.170	0.797	0.797	0.780	0.777
100	1.005	0.987	0.974	0.956	0.687	0.668	0.644	0.636	100	1.118	1.140	1.164	1.158	0.776	0.784	0.784	0.785
Ind. IPC									*IPCP*								
20	0.999	0.988	0.980	0.979	0.733	0.690	0.674	0.660	20	1.105	1.136	1.148	1.154	0.810	0.794	0.789	0.782
30	1.001	0.981	0.962	0.966	0.733	0.674	0.636	0.635	30	1.122	1.140	1.160	1.158	0.809	0.788	0.783	0.779
50	0.981	0.980	0.964	0.974	0.704	0.682	0.651	0.634	50	1.120	1.141	1.143	1.171	0.798	0.797	0.781	0.778
100	0.980	0.975	0.968	0.954	0.674	0.661	0.641	0.634	100	1.119	1.142	1.165	1.160	0.777	0.785	0.785	0.785
Ind. PCX									*PCPX*								
20	1.087	1.059	1.018	1.000	0.790	0.735	0.697	0.673	20	1.106	1.137	1.149	1.154	0.810	0.793	0.788	0.781
30	1.102	1.055	1.009	0.989	0.795	0.721	0.668	0.651	30	1.122	1.139	1.160	1.158	0.809	0.787	0.782	0.778
50	1.087	1.052	1.003	0.996	0.776	0.732	0.678	0.649	50	1.119	1.142	1.142	1.170	0.798	0.797	0.781	0.777
100	1.090	1.045	1.013	0.977	0.752	0.713	0.673	0.650	100	1.118	1.141	1.164	1.158	0.777	0.785	0.784	0.785
Ind. PCX2S									*PCPX2S*								
20	1.036	1.012	0.991	0.984	0.745	0.698	0.677	0.661	20	1.106	1.137	1.149	1.154	0.808	0.792	0.788	0.781
30	1.044	1.004	0.972	0.970	0.750	0.681	0.639	0.637	30	1.122	1.139	1.160	1.158	0.808	0.787	0.782	0.778
50	1.024	1.006	0.973	0.978	0.722	0.693	0.655	0.635	50	1.119	1.142	1.142	1.170	0.797	0.797	0.780	0.777
100	1.022	0.996	0.978	0.958	0.696	0.672	0.645	0.636	100	1.118	1.141	1.164	1.158	0.777	0.784	0.784	0.785

vidual estimators outperform pooled estimators, the only exception being the case of high spatial dependence and low level of heterogeneity; (iii) PC estimators, especially the *Ind. IPC* of Song (2013), is the estimator which is most robust to spatial dependence.

4. An application to house price inflation in OECD countries

In this section, we report an illustrative example using the two forecasting approaches and several panel data estimators with the main aim of undertaking short-run forecasts of the house price inflation in the OECD countries.

Table 7: Relative RMSE – High Heterogeneity, Case *2*: Low Spatial Dependence & High Factor Dependence

	Individual									*Pooled*							
T / n	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	0.926	0.933	0.938	0.941	0.497	0.467	0.455	0.445	20	0.979	1.006	1.022	1.026	0.545	0.528	0.523	0.516
30	0.943	0.932	0.918	0.928	0.496	0.448	0.416	0.417	30	1.000	1.007	1.010	1.020	0.541	0.514	0.501	0.502
50	0.931	0.938	0.926	0.942	0.471	0.455	0.426	0.412	50	0.990	1.011	1.012	1.038	0.523	0.521	0.503	0.499
100	0.938	0.941	0.933	0.915	0.447	0.432	0.414	0.408	100	0.992	1.015	1.026	1.014	0.501	0.503	0.501	0.501
Ind. CCEX									*CCEPX*								
20	0.931	0.936	0.939	0.942	0.501	0.470	0.457	0.446	20	0.979	1.006	1.022	1.026	0.545	0.528	0.523	0.516
30	0.947	0.935	0.919	0.928	0.500	0.450	0.417	0.418	30	1.000	1.007	1.010	1.020	0.541	0.514	0.501	0.502
50	0.934	0.939	0.927	0.942	0.473	0.455	0.427	0.413	50	0.990	1.011	1.012	1.038	0.523	0.521	0.503	0.499
100	0.940	0.941	0.934	0.915	0.448	0.433	0.415	0.409	100	0.992	1.015	1.026	1.014	0.501	0.503	0.501	0.501
Ind. IPC									*IPCP*								
20	0.934	0.937	0.939	0.944	0.502	0.467	0.453	0.443	20	0.979	1.007	1.022	1.027	0.546	0.529	0.524	0.517
30	0.942	0.933	0.918	0.929	0.495	0.446	0.413	0.416	30	1.000	1.008	1.011	1.020	0.542	0.515	0.502	0.503
50	0.923	0.933	0.926	0.943	0.466	0.450	0.424	0.411	50	0.991	1.010	1.013	1.038	0.523	0.521	0.504	0.500
100	0.927	0.935	0.931	0.915	0.440	0.428	0.412	0.407	100	0.993	1.015	1.026	1.015	0.501	0.504	0.501	0.501
Ind. PCX									*PCPX*								
20	1.065	1.042	0.998	0.976	0.627	0.561	0.506	0.473	20	0.983	1.011	1.025	1.028	0.550	0.531	0.525	0.517
30	1.085	1.038	0.986	0.963	0.623	0.542	0.478	0.449	30	1.002	1.009	1.013	1.021	0.544	0.516	0.502	0.503
50	1.069	1.034	0.983	0.977	0.605	0.549	0.480	0.443	50	0.992	1.013	1.013	1.039	0.525	0.522	0.504	0.500
100	1.074	1.033	0.995	0.949	0.583	0.528	0.476	0.440	100	0.993	1.015	1.026	1.015	0.502	0.504	0.501	0.501
Ind. PCX2S									*PCPX2S*								
20	0.956	0.954	0.947	0.946	0.517	0.478	0.458	0.445	20	0.983	1.011	1.025	1.028	0.545	0.528	0.523	0.516
30	0.968	0.947	0.925	0.932	0.512	0.454	0.418	0.418	30	1.002	1.009	1.013	1.021	0.541	0.514	0.501	0.502
50	0.947	0.948	0.931	0.945	0.482	0.460	0.427	0.412	50	0.992	1.013	1.013	1.039	0.523	0.521	0.503	0.499
100	0.950	0.946	0.936	0.917	0.457	0.436	0.415	0.409	100	0.993	1.015	1.026	1.015	0.501	0.503	0.501	0.501

Table 8: Relative RMSE – High Heterogeneity, Case *3*: High Spatial Dependence & Low Factor Dependence

	Individual									*Pooled*							
T \ n	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	1.018	1.015	1.008	1.009	0.884	0.843	0.853	0.834	20	1.051	1.089	1.096	1.111	0.907	0.888	0.900	0.890
30	1.040	1.017	1.002	0.996	0.890	0.844	0.822	0.809	30	1.067	1.089	1.110	1.109	0.906	0.889	0.889	0.878
50	1.045	1.028	0.999	0.999	0.882	0.852	0.803	0.791	50	1.066	1.090	1.102	1.115	0.896	0.891	0.866	0.863
100	1.051	1.019	1.005	0.986	0.854	0.824	0.809	0.786	100	1.065	1.087	1.103	1.105	0.868	0.870	0.873	0.863
Ind. CCEX									*CCEPX*								
20	1.060	1.038	1.024	1.016	0.912	0.857	0.862	0.838	20	1.051	1.089	1.096	1.111	0.907	0.888	0.901	0.890
30	1.072	1.036	1.013	1.000	0.911	0.855	0.828	0.811	30	1.067	1.089	1.110	1.109	0.906	0.888	0.889	0.878
50	1.066	1.037	1.004	1.002	0.895	0.857	0.806	0.792	50	1.066	1.090	1.101	1.115	0.896	0.891	0.866	0.863
100	1.060	1.023	1.008	0.988	0.861	0.827	0.811	0.787	100	1.065	1.087	1.103	1.105	0.868	0.870	0.873	0.863
Ind. IPC									*IPCP*								
20	0.983	0.990	0.996	1.003	0.872	0.836	0.851	0.835	20	1.052	1.091	1.097	1.113	0.908	0.889	0.901	0.891
30	0.992	0.987	0.986	0.988	0.873	0.837	0.820	0.809	30	1.067	1.090	1.111	1.110	0.906	0.889	0.890	0.879
50	0.989	0.992	0.985	0.989	0.861	0.840	0.800	0.791	50	1.067	1.090	1.103	1.116	0.897	0.892	0.866	0.864
100	0.992	0.988	0.986	0.978	0.831	0.814	0.806	0.787	100	1.066	1.088	1.104	1.106	0.868	0.870	0.873	0.864
Ind. PCX									*PCPX*								
20	1.123	1.088	1.051	1.033	0.942	0.880	0.872	0.845	20	1.056	1.092	1.098	1.113	0.909	0.889	0.901	0.890
30	1.129	1.080	1.044	1.017	0.939	0.877	0.843	0.819	30	1.068	1.090	1.112	1.110	0.907	0.889	0.889	0.878
50	1.115	1.074	1.028	1.017	0.924	0.878	0.819	0.800	50	1.067	1.091	1.102	1.115	0.897	0.892	0.866	0.863
100	1.109	1.058	1.030	1.001	0.891	0.849	0.824	0.794	100	1.065	1.088	1.104	1.105	0.869	0.870	0.873	0.864
Ind. PCX2S									*PCPX2S*								
20	1.055	1.035	1.016	1.014	0.897	0.848	0.853	0.833	20	1.056	1.092	1.098	1.113	0.907	0.887	0.900	0.890
30	1.069	1.031	1.008	1.000	0.901	0.847	0.822	0.809	30	1.068	1.090	1.112	1.110	0.906	0.888	0.889	0.878
50	1.066	1.035	1.003	1.001	0.889	0.853	0.803	0.790	50	1.067	1.091	1.102	1.115	0.896	0.891	0.866	0.863
100	1.069	1.026	1.008	0.988	0.861	0.826	0.809	0.786	100	1.065	1.088	1.104	1.105	0.868	0.870	0.873	0.863

Table 9: Relative RMSE – High Heterogeneity, Case *4*: High Spatial Dependence & High Factor Dependence

	Individual									*Pooled*							
T	Residual Based Approach				Auxiliary Variables Approach					Residual Based Approach				Auxiliary Variables Approach			
n	20	30	50	100	20	30	50	100		20	30	50	100	20	30	50	100
Ind. CCE									*CCEP*								
20	0.962	0.966	0.974	0.977	0.674	0.627	0.649	0.629	20	0.982	1.009	1.029	1.038	0.689	0.658	0.683	0.669
30	0.982	0.968	0.956	0.959	0.671	0.624	0.597	0.592	30	1.000	1.011	1.018	1.025	0.682	0.654	0.644	0.641
50	0.979	0.977	0.955	0.967	0.654	0.629	0.577	0.571	50	0.992	1.013	1.015	1.034	0.663	0.657	0.621	0.622
100	0.982	0.971	0.963	0.942	0.613	0.590	0.584	0.562	100	0.990	1.011	1.022	1.013	0.623	0.622	0.629	0.617
Ind. CCEX									*CCEPX*								
20	0.988	0.980	0.983	0.980	0.695	0.638	0.656	0.632	20	0.982	1.009	1.029	1.038	0.689	0.658	0.683	0.669
30	1.003	0.981	0.963	0.962	0.687	0.633	0.602	0.594	30	1.000	1.011	1.018	1.025	0.682	0.654	0.644	0.641
50	0.993	0.984	0.958	0.968	0.663	0.633	0.579	0.572	50	0.992	1.013	1.014	1.034	0.663	0.657	0.621	0.622
100	0.988	0.974	0.965	0.943	0.618	0.593	0.585	0.563	100	0.990	1.011	1.022	1.013	0.623	0.622	0.629	0.617
Ind. IPC									*IPCP*								
20	0.947	0.955	0.967	0.973	0.669	0.624	0.648	0.631	20	0.982	1.010	1.029	1.038	0.690	0.658	0.683	0.670
30	0.956	0.952	0.948	0.955	0.661	0.621	0.598	0.594	30	1.000	1.011	1.019	1.025	0.682	0.655	0.644	0.641
50	0.948	0.958	0.949	0.961	0.641	0.624	0.577	0.574	50	0.992	1.013	1.015	1.034	0.663	0.657	0.621	0.623
100	0.953	0.957	0.955	0.941	0.601	0.585	0.583	0.562	100	0.990	1.011	1.022	1.013	0.623	0.623	0.629	0.617
Ind. PCX									*PCPX*								
20	1.093	1.062	1.028	1.008	0.770	0.695	0.683	0.648	20	0.986	1.013	1.031	1.040	0.694	0.660	0.684	0.670
30	1.105	1.056	1.014	0.989	0.761	0.687	0.638	0.613	30	1.002	1.012	1.020	1.026	0.684	0.656	0.645	0.641
50	1.089	1.053	1.001	0.993	0.741	0.688	0.612	0.591	50	0.993	1.015	1.015	1.034	0.664	0.658	0.621	0.622
100	1.087	1.043	1.010	0.968	0.701	0.652	0.621	0.582	100	0.990	1.011	1.022	1.013	0.624	0.623	0.629	0.617
Ind. PCX2S									*PCPX2S*								
20	0.993	0.985	0.981	0.982	0.690	0.635	0.650	0.629	20	0.986	1.013	1.031	1.040	0.690	0.657	0.683	0.669
30	1.008	0.982	0.962	0.963	0.684	0.628	0.598	0.592	30	1.002	1.012	1.020	1.026	0.682	0.654	0.644	0.641
50	0.998	0.986	0.959	0.969	0.663	0.632	0.577	0.571	50	0.993	1.015	1.015	1.034	0.663	0.657	0.621	0.622
100	0.996	0.977	0.966	0.944	0.621	0.594	0.585	0.563	100	0.990	1.011	1.022	1.013	0.623	0.622	0.629	0.617

4.1. Empirical setup and data

To forecast the house price inflation we use the model specifications in Holly, Pesaran & Yamagata (2010) and Caldera & Johansson (2013). Authors model the long run movements in the house price index in the US using state level data and in OECD countries using country level data, respectively, by household disposable income, population growth and a proxy for cost of borrowing. Taking first differences of the non-stationary variables in their model to achieve stationarity, our most general forecasting equation is given by

$$\Delta \log p_{i,t+h} = \alpha_i + \beta_{1i}\Delta \log p_{it} + \beta_{2i}\Delta \log y_{it} + \beta_{3i}\Delta \log n_{it} + \beta_{4i} i_{it} + e_{i,t+h}, \tag{26}$$

where p is the real house price index, y is the per capita household disposable income, n is the population and i is the real long-term interest rate.

Caldera & Johansson (2013) estimate the house price equation simultaneously with a housing investment equation which gives a supply and demand system for the housing market. In this investment equation they have the house prices, residential construction costs and population growth. Since housing investment has house prices as a component it can be use to estimate the common factors in the price equation. It is also reasonable to assume that other variables in the investment equation are correlated with these common factors. To estimate the common factors in the AVA, we use 8 variables in total. In addition to the ones defined above, we have per capita gross fixed capital formation in housing (inv), residential fixed capital formation deflator (cc) which is a proxy for residential construction costs, GDP per capita (gdp), the consumer price index (cpi) and per capita private final consumption expenditure ($cons$).

The data set comes from the OECD Economic Outlook at quarterly frequency. All variables are seasonally adjusted and cover the period between 1995:1 and 2017:4 for 20 OECD countries, hence the final dataset contains 1840 observations. The countries considered are AUS, BEL, CAN, CHE, DEU, DNK, ESP, FIN, FRA, GBR, IRL, ITA, JPN, KOR, NLD, NOR, NZL, PRT, SWE and USA. The panel is balanced for the house price index p. There is the presence of missing observations for some of the variables. For CHE, inv and cc are missing for the periods between 2016:1 and 2017:4 and for CHE, JPN and NZL, y is missing between 2017:1 and 2017:4. To obtain a balanced sample, we predicted these in-sample observations using other variables in the data set. To predict per capita household disposable income, we regress $\log y$ on $\log gdp$ and a linear trend for each country separately and fill the missing observations with predicted values. For per

Table 10: Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum	Unit Root Test	$\bar{p}$
$\Delta \log p_{it}$	0.0062	0.0179	-0.0749	0.0710	-4.07	0.7
$\Delta \log inv_{it}$	0.0017	0.0407	-0.2762	0.2225	-6.70	1.2
$\Delta \log y_{it}$	0.0033	0.0119	-0.1361	0.0884	-8.81	0.5
$\Delta \log cc_{it}$	0.0065	0.0158	-0.1209	0.1334	-5.58	1.4
$\Delta \log n_{it}$	0.0016	0.0013	-0.0021	0.0111	-1.67	1.4
$\Delta \log gdp_{it}$	0.0040	0.0106	-0.0745	0.2017	-8.07	0.4
$\Delta \log cpi_{it}$	0.0043	0.0046	-0.0267	0.0459	-6.34	0.5
$\Delta \log cons_{it}$	0.0038	0.0089	-0.1480	0.0552	-8.60	0.4
i_{it}	0.0244	0.0266	-0.0805	0.2564	-11.10	1.7

Notes: For each variable x_{it}, the unit root test statistics are computed as $\overline{CIPS} = n^{-1} \sum_{i=1}^{n} t_i(n,T)$ where $t_i(n,T)$ is the t-statistic of the coefficient b_i in the regression $\Delta x_{it} = a_i + b_i x_{i,t-1} + c_i \bar{x}_{t-1} + \sum_{j=0}^{p_i} d_{ij} \Delta \bar{x}_{t-j} + \sum_{j=1}^{p_i} \delta_{ij} \Delta x_{i,t-j}$ where $\bar{x}_t = n^{-1} \sum_{i=1}^{n} x_{it}$. The lag lengths p_i are selected using Akaike information criterion for each country and their means $\bar{p} = n^{-1} \sum_{i=1}^{n} p_i$ are reported. The critical values for the unit root tests are -2.11, -2.20 and -2.36 for 10%, 5% and 1%, levels respectively.

capita gross fixed capital formation in housing, the variables in the model by Caldera & Johansson (2013) are used. Namely, $\log inv$ is regressed on $\log cc$, $\log y$, $\log n$ and a linear trend and missing observations are replaced with the predicted values. Similarly, $\log cc$ is predicted by $\log cpi$ and a linear trend. As a percentage of the total number of observations, the number missing observations filled is 0.44% for inv and y, 0.66% for cc.

4.2. Preliminary analysis

Table 10 gives the descriptive statistics for each variable. Before proceeding with the estimation of the regression models and calculating the accuracy of predictions based on them, we check the time series and cross-sectional properties of the variables.

The results on the unit root tests for each variable are given in Table 10. As each variable shows strong evidence of CD (see below), the CD-robust unit root tests developed by Pesaran

Table 11: CD Test Results

Variable	$\Delta \log p_{it}$	$\Delta \log inv_{it}$	$\Delta \log y_{it}$	$\Delta \log cc_{it}$	$\Delta \log n_{it}$	$\Delta \log gdp_{it}$	$\Delta \log cpi_{it}$	$\Delta \log cons_{it}$	i_{it}
					Panel a: Original Data				
Breusch-Pagan LM Test	1432.85	675.84	393.12	578.66	2890.09	3001.74	3477.42	1233.97	3789.70
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Modified BP Test	63.76	24.92	10.42	19.94	138.51	144.24	168.64	53.55	184.66
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
					Panel b: Defactored Data				
Breusch-Pagan LM Test	693.30	834.81	958.64	966.67	2232.18	724.09	900.90	726.16	886.10
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Modified BP Test	25.82	33.08	39.43	39.84	104.76	27.40	36.47	27.50	35.71
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: For each variable x_{it}, the Breusch-Pagan LM Test statistics are computed as $CD_{BP} = T \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \widehat{\kappa}_{ij}^2$ where $\widehat{\kappa}_{ij}$ is the correlation coefficient between x_{it} and x_{jt}. Under the null of no CD, the asymptotic distribution of the test statistic is χ_q^2 with $q = n(n-1)/2$. The Modified BP Test statistics are computed as $CD_M = [n(n-1)]^{-1/2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (T\widehat{\kappa}_{ij}^2 - 1)$ which is distributed as $N(0,1)$ under the null of no CD. p-values are in parentheses. The test statistics given in *Panel b* are computed after removing country fixed effects and the unobserved common factors estimated using PC methods. For each variable the number of common factors are chosen using the information criterion IC_{p_1} of Bai & Ng (2002).

(2007) are applied to each variable in the data set. The only variable for which we cannot reject the unit root hypothesis is population growth. In the application we use models with and without this variable. In Table 11, CD test results are reported. Two different CD tests are applied to each variable in the dataset. The first one is the LM test of Breusch & Pagan (1980). This is a general cross-correlation test where the null hypothesis states that the correlation coefficients between all pairs of units in the data set are jointly zero. Under the null hypothesis the test statistic follows a χ^2 distribution with $n(n-1)/2$ degrees of freedom as T goes to infinity for fixed n. The results show that for each variable in the data set there is strong evidence against no CD hypothesis.

The disadvantage of the Breusch & Pagan (1980) test is that as n gets larger its variance increases, hence it is not appropriate for panels of large cross-sectional dimension. Thus, we also report the results from a modified version of this test, the Modified BP Test which is distributed as a standard normal for large T and n (see Pesaran, 2015, for details). The results are in line with the previous test such that the null of no CD can be rejected for any variable in any conventional significance level.

Although these two tests are general CD tests, they do not detect the types of CD in the data. To see if the results change after removing the unobserved common factors we applied the same tests to defactored variables. We remove unobserved common factors using PC methods where the number of common factors are chosen by the information criterion IC_{p_1} proposed by Bai & Ng (2002). With a few exceptions the test statistics are weaker but still the no CD hypothesis can be rejected for each variable.

Table 12: Distance Based Spatial Dependence Tests

Variable	$\Delta \log p_{it}$	$\Delta \log inv_{it}$	$\Delta \log y_{it}$	$\Delta \log cc_{it}$	$\Delta \log n_{it}$	$\Delta \log gdp_{it}$	$\Delta \log cpi_{it}$	$\Delta \log cons_{it}$	i_{it}
$\widehat{\rho}$	0.52	0.36	0.15	0.23	0.26	0.54	0.68	0.48	0.66
Test Statistic	183.35	11.13	83.56	106.11	8.72	189.38	240.45	173.80	229.73
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: For each demeaned variable x_{it}, the spatial autoregressive coefficient is estimated by maximum likelihood in the regression $\mathbf{x}_{.t} = \rho \mathbf{W}_n \mathbf{x}_{.t} + \varepsilon_{.t}$ where $\mathbf{x}_{.t}$ is the vector of observations of countries stacked for each t and $\mathbf{W}_n$ is the row normalized inverse distance matrix. p-values are in parentheses.

To see if there is evidence for spatial interactions based on geographic distance, we estimate a first order SAR model for each variable by maximum likelihood. We use a row normalized inverse distance matrix as spatial weights.[4] For the variable of interest, the house price inflation, the SAR coefficient is estimated as 0.52 and it is highly significant. The consumer price inflation shows the highest coefficient estimate which is equal to 0.68. All remaining coefficients also have statistically significant SAR coefficients.

Finally we estimate a factor model for the house price inflation series to see the global common movements embedded in it. The information criterion IC_{p_1} of Bai & Ng (2002) indicates the existence of 3 common factors in the panel. These common factor estimates are in Figure 1 where we report estimates using both PC and maximum likelihood methods together with the correlation between the two estimates.

We observe that the two methods give similar estimates of the common factors with the correlation coefficient up to 0.96. The third factor has a relatively low coefficient equal to 0.82. However, we compare the estimates of the common components using each method and found an average correlation coefficient over countries equal to 0.98. Hence, even if there are differences in common

[4]The data on geographical distance come from CEPII GeoDist dataset Mayer & Zignago (2011).

factor estimates, the loadings estimates compensate the discrepancies. As the results are similar, we focus on PC estimates and report the factor loadings estimates in Table 13 using the PC method.

The estimates of the first common factor shows an upward trending segment until around 2005. As factor loadings estimates of all countries, except CHE, DEU and JPN, are positive, this factor adds an increasing component to each countries house price inflation series. After 2005 the effect of the global financial crisis can be seen as the common factor estimate drops sharply. This common factor is found to be highly correlated with an AR(1) coefficient estimated as 0.89.

On the other hand, the second common factor has a downward trend until 2005 and the rest is stable. For the countries with negative loading estimate, this factor strengthens the upward movement until around 2005. The estimated AR(1) coefficient is smaller but still strong for this factor, equal to 0.77.

The last common factor has a peak in the crisis period whereas for the rest of the sample it looks stable. For countries with positive loadings, this factor compensates the drop caused by the first common factor. It has a much smaller AR(1) coefficient which is equal to 0.58.

Table 13: House Price Inflation - Factor Loadings Estimates

Country	$\hat{\gamma}_{1i}$	$\hat{\gamma}_{2i}$	$\hat{\gamma}_{3i}$	Country	$\hat{\gamma}_{1i}$	$\hat{\gamma}_{2i}$	$\hat{\gamma}_{3i}$
AUS	0.55	-0.16	-0.33	IRL	0.65	0.48	-0.02
BEL	0.46	-0.27	0.42	ITA	0.62	-0.40	0.44
CAN	0.36	-0.53	-0.38	JPN	-0.27	-0.01	-0.66
CHE	-0.19	-0.60	-0.03	KOR	0.12	-0.68	-0.20
DEU	-0.12	0.03	-0.45	NLD	0.50	0.53	0.23
DNK	0.64	0.17	-0.12	NOR	0.42	0.08	-0.17
ESP	0.81	-0.06	0.21	NZL	0.53	-0.22	-0.34
FIN	0.52	0.03	-0.18	PRT	0.36	0.46	-0.12
FRA	0.69	-0.44	0.33	SWE	0.62	0.03	-0.30
GBR	0.79	0.06	-0.17	USA	0.67	0.17	-0.11

The above analysis shows very strong evidence in favor of different types of CD in the variables in our data set. Hence, it is important to take into account the CD properties in the estimation

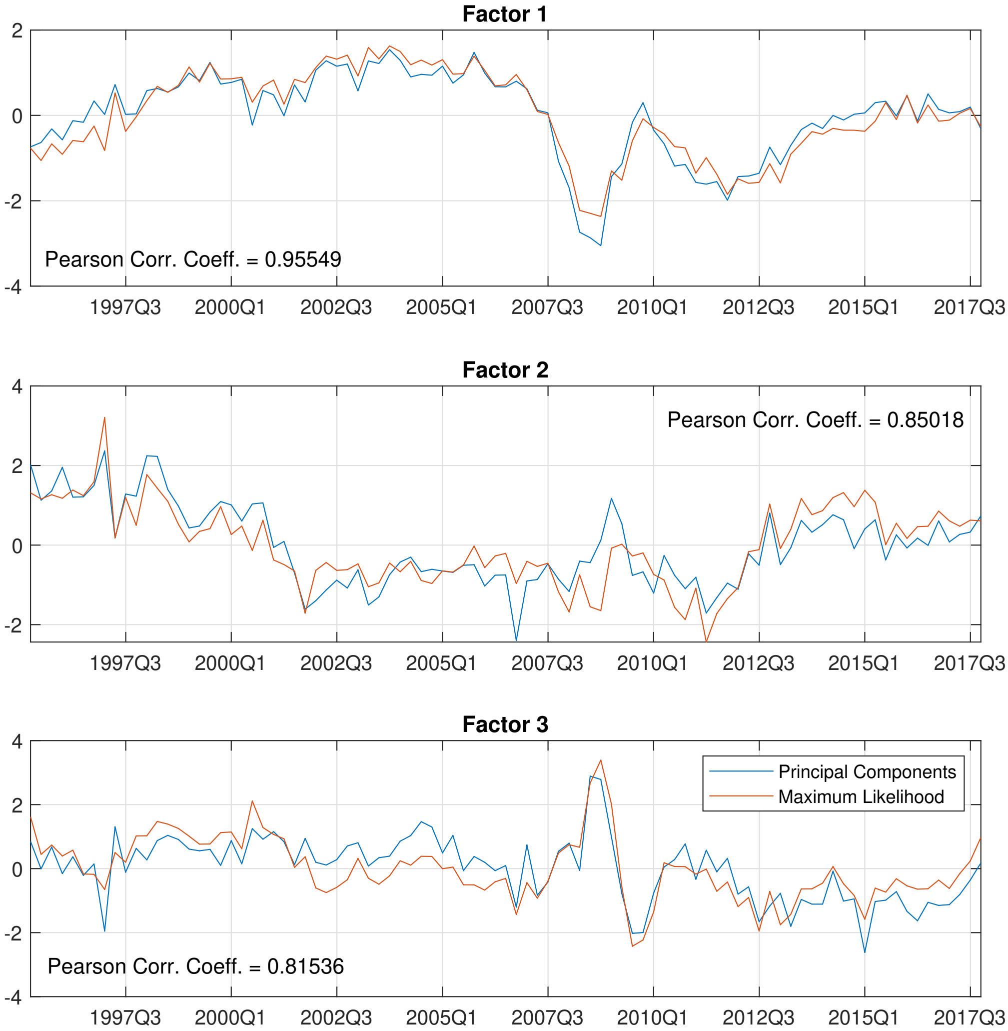

Figure 1: Common Factors in House Price Inflation Series

and forecasting.

4.3. Forecasting results

The results of the estimation of the 1-year ahead predictive models and their pseudo-out-of-sample RMSE values are given in Table 14 and 15. We consider four different models based on 26. *Model 1* uses only $\Delta \log y_{it}$ as a regressor whereas *Model 2* uses all exogeneous regressors but ignores the lagged housing inflation. *Model 3* and *Model 4* augment these two models with lagged housing inflation, respectively.

We estimate each model in the period 1995:2 and 2016:4 and the objective is to forecast the value in 2017:4. For the heterogeneous estimators we report their mean over countries known as mean group (MG) estimates together with their standard errors estimated using the usual non-parametric variance formulas (see, for instance Pesaran & Smith, 1995). For the estimators which require the information on the number of factors, i.e. the estimators using PC methods, we use the information criterion IC_{p_1} of Bai & Ng (2002) to estimate these numbers. For the iterative PC estimators we set this number to one as otherwise the forecast performance of the estimator falls dramatically. In order to see the advantage of using common factors for forecasting with panel data, in addition to the estimators described in Section 2, we report results from 4 additional estimators which do not into account the possible unobserved common factors contained in house price inflation. These estimators are *Ind. OLS*, *Ind. GLS* estimator based on Swamy (1970) which is computed using the deviations of each variable from their time average (see, for details Lee & Griffiths, 1979), fixed effects (FE) and the usual 2-way fixed effects (*2WFE*). To forecast using the *2WFE*, we use the coefficient on the last time dummy as the future value.[5]

In the two first models under consideration lagged housing inflation is dropped from right hand side. The results of these two models, *Model 1* and *Model 2*, are given in Table 14. Since these are prediction models the coefficients do not have their usual economic meaning but it can be useful to compare the estimates from different estimators. In *Model 1* for which the results are given in *Panel a*, the first observation is that there are substantial differences between the coefficient estimates which come from estimators which do not control for unobserved common factors and the ones which do. For instance, using *SW* (column 2) the coefficient of lagged disposable income

[5]See Baltagi (2008) for alternative methods to forecast with the *2WFE* estimator.

growth is estimated as 0.23 whereas this number is 0.14 and 0.16 for *CCEMG* (column 3) and *IPCMG* (column 5), respectively. This is in line with the results of Holly et al. (2010) who find that the non-robust estimators tend to overestimate the impact of disposable income on house prices. Important differences can be observed between heterogeneous and pooled estimators of the average effects. For instance the estimator *IPCP* (column 12) which is the the pooled counterpart of the *IPCMG* (column 5), gives 0.05 for the same effect. Similar differences are observed for the model with additional predictors given in *Panel b* and the models with lagged dependent variable given in the two panels of Table 15. In these cases most important differences are being observed in the coefficient of the interest rate.

The pseudo-out-of-sample performance of each estimator combined with the RBA and the AVA are given at the bottom of each panel. In *Panel a* of Table 14 it can be seen that the *Ind. CCEX* (column 4) estimator shows the best prediction performance when combined with the AVA. The RMSE for this strategy is computed as 1.244. The closest performance from the estimators without unobserved common factors is seen on the *Ind. OLS* (column 1) which has an RMSE about 9% higher than the best performer. The RMSE of the same estimator combined with the RBA is about %6 higher which shows the advantage of the AVA over the RBA and overall the usage of common factors for forecasting. The results are confirmed by introducing additional predictors in the model. Now the best performer, *Ind. CCEX* has an RMSE equal to 0.928 which means more than %30 gain in precision.

In terms of the best performing strategy, the results are unchanged in the models with lagged dependent variables (Table 15). The overall prediction performance of the models improve dramatically by the introduction of lagged house price inflation. Now the lowest RMSE is equal to 0.867. The RMSE for the best performing estimator without common factors (*Ind. GLS*) is about 26% higher than this value. Once more the results show the superiority of the AVA over the RBA.

To summarize, the individual estimators outperform the pooled estimators in models with or without lagged dependent variables; the prediction strategies using unobserved common factors increase the prediction ability significantly; the AVA has a superior performance compared to the RBA.

Table 14: Estimation Results for the 1-Year Ahead Predictive Regressions–Models Without Lagged Dependent Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Ind. OLS	*Ind. GLS*	*Ind. CCE*	*Ind. CCEX*	*Ind. IPC*	*Ind. PCX*	*Ind. PCX2S*	*FE*	*2WFE*	*CCEP*	*CCEPX*	*IPCP*	*PCPX*	*PCP2SX*
Panel a: Model 1 – $\Delta \log p_{i,t+h} = \alpha_i + \beta_{2i} \Delta \log y_{it} + e_{i,t+h}$														
Disposable Income Growth	0.27	0.23	0.14	0.19	0.16	0.10	0.07	0.15	0.11	0.07	0.08	0.05	0.14	0.07
	(0.08)	(0.078)	(0.041)	(0.057)	(0.058)	(0.083)	(0.072)	(0.094)	(0.033)	(0.054)	(0.077)	(0.063)	(0.082)	(0.063)
1-Year Ahead RMSE (x100)														
Residual Based Approach	1.378	1.386	1.329	1.321	1.330	1.351	1.418	1.422	1.450	1.401	1.381	1.369	1.371	1.431
Auxiliary Variables Approach			1.263	1.244	1.258	1.333	1.296			1.323	1.323	1.323	1.326	1.323
Panel b: Model 2 – $\Delta \log p_{i,t+h} = \alpha_i + \beta_{2i} \Delta \log y_{it} + \beta_{3i} \Delta \log n_{it} + \beta_{4i} i_{it} + e_{i,t+h}$														
Disposable Income Growth	0.28	0.26	0.17	0.19	0.20	0.16	0.16	0.17	0.14	0.11	0.11	0.08	0.14	0.06
	(0.096)	(0.095)	(0.047)	(0.067)	(0.06)	(0.082)	(0.072)	(0.113)	(0.034)	(0.05)	(0.089)	(0.07)	(0.091)	(0.084)
Population Growth	-0.06	-0.05	-0.09	-0.05	-0.09	-0.04	-0.02	-0.04	-0.08	-0.08	-0.06	-0.07	-0.05	-0.03
	(0.037)	(0.036)	(0.029)	(0.044)	(0.032)	(0.057)	(0.053)	(0.05)	(0.018)	(0.026)	(0.053)	(0.026)	(0.05)	(0.037)
Interest Rate	-2.87	-2.32	0.84	-2.93	0.93	-4.33	-3.80	0.22	0.91	0.96	1.28	1.39	-0.45	0.21
	(2.185)	(2.164)	(1.676)	(2.439)	(1.824)	(2.235)	(2.547)	(1.761)	(0.429)	(0.665)	(2.355)	(0.935)	(2.189)	(2.018)
1-Year Ahead RMSE (x100)														
Residual Based Approach	1.241	1.174	1.364	1.073	1.473	1.474	1.405	1.446	1.519	1.483	1.430	1.407	1.435	1.441
Auxiliary Variables Approach			1.271	0.928	1.391	1.249	1.207			1.401	1.395	1.407	1.350	1.348

Notes: For individual estimators given in columns (1)-(7) mean group estimates are reported which are computed as $\widehat{\boldsymbol{\beta}}_{M,MG} = n^{-1} \sum_{i=1}^{n} \widehat{\boldsymbol{\beta}}_{M,i}$. Standard errors are in parentheses which are computed using the usual formulas given in, for instance, Pesaran & Tosetti (2011).

Table 15: Estimation Results for the 1-Year Ahead Predictive Regressions–Models With Lagged Dependent Variable

	(1) *Ind. OLS*	(2) *Ind. GLS*	(3) *Ind. CCE*	(4) *Ind. CCEX*	(5) *Ind. IPC*	(6) *Ind. PCX*	(7) *Ind. PCX2S*	(8) *FE*	(9) *2WFE*	(10) *CCEP*	(11) *CCEPX*	(12) *IPCP*	(13) *PCPX*	(14) *PCP2SX*
Panel a: Model 3 – $\Delta \log p_{i,t+h} = \alpha_i + \beta_{1i} \Delta \log p_{it} + \beta_{2i} \Delta \log y_{it} + e_{i,t+h}$														
Lagged House Price Inflation	0.26	0.27	0.16	0.09	0.17	0.09	0.08	0.31	0.27	0.16	0.12	0.16	0.08	0.01
	(0.057)	(0.057)	(0.042)	(0.069)	(0.049)	(0.057)	(0.05)	(0.071)	(0.024)	(0.035)	(0.07)	(0.052)	(0.049)	(0.042)
Disposable Income Growth	0.09	0.09	0.07	0.12	0.07	0.01	0.01	0.05	0.04	0.01	0.05	0.02	-0.03	-0.01
	(0.058)	(0.054)	(0.047)	(0.056)	(0.045)	(0.056)	(0.056)	(0.048)	(0.032)	(0.046)	(0.059)	(0.042)	(0.053)	(0.061)
1-Year Ahead RMSE (x100)														
Residual Based Approach	1.260	1.319	1.333	1.206	1.283	1.301	1.337	1.449	1.472	1.441	1.419	1.433	1.404	1.424
Auxiliary Variables Approach			1.234	1.183	1.215	1.266	1.256			1.306	1.324	1.321	1.326	1.335
Panel b: Model 4 – $\Delta \log p_{i,t+h} = \alpha_i + \beta_{1i} \Delta \log p_{it} + \beta_{2i} \Delta \log y_{it} + \beta_{3i} \Delta \log n_{it} + \beta_{4i} i_{it} + e_{i,t+h}$														
Lagged House Price Inflation	0.20	0.22	0.08	0.08	0.11	0.05	-0.02	0.32	0.28	0.10	0.12	0.16	0.06	0.00
	(0.064)	(0.068)	(0.044)	(0.067)	(0.049)	(0.068)	(0.052)	(0.091)	(0.024)	(0.04)	(0.076)	(0.048)	(0.052)	(0.039)
Disposable Income Growth	0.13	0.11	0.11	0.14	0.13	0.11	0.11	0.07	0.08	0.06	0.08	0.06	0.03	0.02
	(0.05)	(0.063)	(0.049)	(0.053)	(0.045)	(0.062)	(0.054)	(0.059)	(0.033)	(0.048)	(0.068)	(0.051)	(0.068)	(0.065)
Population Growth	-0.05	-0.05	-0.08	-0.04	-0.07	-0.04	-0.01	-0.05	-0.10	-0.06	-0.07	-0.08	-0.06	-0.03
	(0.031)	(0.036)	(0.02)	(0.029)	(0.03)	(0.047)	(0.037)	(0.045)	(0.018)	(0.021)	(0.059)	(0.024)	(0.042)	(0.029)
Interest Rate	-4.03	-3.62	-1.16	-3.44	-0.56	-3.99	-3.93	-0.92	-0.16	0.35	0.76	0.84	0.36	0.77
	(1.723)	(1.856)	(1.17)	(2.054)	(1.379)	(2.308)	(2.119)	(1.652)	(0.424)	(0.912)	(2.428)	(0.986)	(2.327)	(2.235)
1-Year Ahead RMSE (x100)														
Residual Based Approach	1.122	1.089	1.317	0.928	1.368	1.530	1.218	1.475	1.533	1.487	1.481	1.517	1.476	1.457
Auxiliary Variables Approach			1.200	0.867	1.302	1.450	1.161			1.356	1.393	1.404	1.378	1.377

Notes: For individual estimators given in columns (1)-(7) mean group estimates are reported which are computed as $\widehat{\boldsymbol{\beta}}_{M,MG} = n^{-1} \sum_{i=1}^{n} \widehat{\boldsymbol{\beta}}_{M,i}$. *Ind. CCE* and *CCEP* refer to the dynamic *CCE* estimators computed as in Chudik & Pesaran (2015) by including p_T lags of the cross-sectional averages of the dependent and explanatory variables in the regression. The lag length is chosen as $p_T = [(T-h)^{1/3}]$ where [ ] denotes the biggest integer smaller than its argument. Standard errors are in parentheses which are computed using the usual formulas given in, for instance, Pesaran & Tosetti (2011).

5. Conclusions

In this paper, we evaluated the performance of alternative methods of forecasting in presence of heterogeneous panel data with cross-sectional dependence by considering both spatial dependence and unobserved common factors. Alternative estimators of unit specific parameters and their pooled counterparts are compared in Monte Carlo simulations and by pseudo-out-of-sample forecasts using real data on house price inflation in OECD countries.

Our main results are as follows. The Auxiliary Variables Approach, which uses a number of indicators correlated with the unobserved common factors in the DGP of the variable of interest, outperforms the Residual Based Approach which extracts the common factors from residuals of the model. The choice between forecasting using individual specific estimates and pooled estimates depend on the level of heterogeneity and spatial dependence in the error terms: for a given level of heterogeneity, higher spatial dependence increases the relative forecast performance of pooled estimators whereas for a given degree of spatial dependence higher heterogeneity makes forecasts using individual estimates perform better. Further, among the methods of estimating common factors, the CCE approach of Pesaran (2006) outperforms the principal components methods of Song (2013) in the case of individual estimates and low spatial dependence, whereas for pooled estimates the differences are negligible. The main difference on the performance of the the two methods occurs when we move from low to high spatial dependence, whereas moving from low to high factor dependence does not change their comparative performance. The estimators based on PC methods are found to be more robust to spatial dependence than CCE methods.

Acknowledgements

We wish to thank participants in the New York Camp Econometrics XIII (Syracuse University, United States, 6-9 April 2018), the 17th Spatial Econometrics and Statistics Workshop (University of Burgundy, France, 24-25 May 2018), the 24th International Panel Data Conference (Sogang University, Korea, 19-20 June 2018) and the 20th OxMetrics User Conference (Cass Business School, United Kingdom, 10-11 September 2018). We wish to thank Badi Baltagi, Neil R. Ericsson, Wang Fa, Lynda Khalaf for useful comments and suggestions. The usual disclaimer applies. This paper is a part of Oguzhan Akgun's PhD research project, and it was also developed while Giovanni Urga was visiting Professor at University Paris II Panthéon-Assas - CRED (France) during the academic years 2016-2017, 2017-2018 and 2018/2019, and Alain Pirotte was visiting Professor at Bergamo University during the academic years 2016-2017, 2017-2018 and 2018-2019 under the scheme "Progetto ITALY® - Azione 2: Grants for Visiting Professor and Scholar": we wish to thank both Institutions for financial support. Financial support is also acknowledged from the Centre for Econometric Analysis (CEA) of Cass Business School for several visits of Alain Pirotte and Oguzhan Akgun at CEA.

References

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*, 135–171.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, *77*, 1229–1279.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*, 191–221.

Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, *146*, 304–317.

Bai, J., & Ng, S. (2009). Boosting diffusion indices. *Journal of Applied Econometrics*, *24*, 607–629.

Baltagi, B. H. (2008). Forecasting with panel data. *Journal of Forecasting*, *27*, 153–173.

Baltagi, B. H., Bresson, G., Griffin, J. M., & Pirotte, A. (2003). Homogeneous, heterogeneous or shrinkage estimators? some empirical evidence from french regional gasoline consumption. *Empirical Economics*, *28*, 795–811.

Baltagi, B. H., Bresson, G., & Pirotte, A. (2004). Tobin q: Forecast performance for hierarchical bayes, shrinkage, heterogeneous and homogeneous panel data estimators. *Empirical Economics*, *29*, 107–113.

Baltagi, B. H., Bresson, G., & Pirotte, A. (2008). To pool or not to pool? In L. Màtyàs, & P. Sevestre (Eds.), *The Econometrics of Panel Data, Advanced Studies in Theoretical and Applied Econometrics* (pp. 995–1024). Springer Berlin Heidelber.

Baltagi, B. H., Bresson, G., & Pirotte, A. (2012). Forecasting with spatial panel data. *Computational Statistics & Data Analysis*, *56*, 3381–3397.

Baltagi, B. H., Fingleton, B., & Pirotte, A. (2014). Estimating and forecasting with a dynamic spatial panel data model. *Oxford Bulletin of Economics and Statistics*, *76*, 112–138.

Baltagi, B. H., & Griffin, J. M. (1997). Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline. *Journal of Econometrics*, *77*, 303–327.

Baltagi, B. H., Griffin, J. M., & Xiong, W. (2000). To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand. *Review of Economics and Statistics*, *82*, 117–126.

Baltagi, B. H., & Li, D. (2004). Prediction in the panel data model with spatial correlation. In L. Anselin, R. J. G. M. Florax, & S. J. Rey (Eds.), *Advances in Spatial Econometrics* chapter 13. (pp. 283–295). Springer.

Baltagi, B. H., & Li, D. (2006). Prediction in the panel data model with spatial correlation: the case of liquor. *Spatial Economic Analysis*, *1*, 175–185.

Baltagi, B. H., & Pirotte, A. (2010). Panel data inference under spatial dependence. *Economic Modelling*, *27*, 1368–1381.

Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, *132*, 169–194.

Breusch, T. S., & Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, *47*, 239–253.

Caldera, A., & Johansson, Å. (2013). The price responsiveness of housing supply in oecd countries. *Journal of Housing Economics*, *22*, 231–249.

Chudik, A., & Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, *188*, 393–420.

Chudik, A., Pesaran, M. H., & Tosetti, E. (2011). Weak and strong cross section dependence and estimation of large panels. *Econometrics Journal*, *14*, C45–C90.

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. *The Stata Journal*, *6*, 482–496.

Engel, C., Mark, N. C., & West, K. D. (2015). Factor model forecasts of exchange rates. *Econometric Reviews*, *34*, 32–55.

Frees, E. W. (1995). Assessing cross-sectional correlation in panel data. *Journal of Econometrics*, *69*, 393–414.

Frees, E. W. (2004). *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. Cambridge University Press.

Garcia-Ferrer, A., Highfield, R. A., Palm, F., & Zellner, A. (1987). Macroeconomic forecasting using pooled international data. *Journal of Business & Economic Statistics*, *5*, 53–67.

Granger, C. W. J., & Huang, L. L. (1997). Evaluation of panel data models: some suggestions from time series. Discussion Paper no. 97-10, Department of Economics, University of California, San Diego.

Hjalmarsson, E. (2010). Predicting global stock returns. *Journal of Financial and Quantitative Analysis*, *45*, 49–80.

Holly, S., Pesaran, M. H., & Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, *158*, 160–173.

Hoogstrate, A. J., Palm, F. C., & Pfann, G. A. (2000). Pooling in dynamic panel-data models: An application to forecasting gdp growth rates. *Journal of Business & Economic Statistics*, *18*, 274–283.

Juodis, A., & Reese, S. (2018). The incidental parameters problem in testing for remaining cross-section correlation. *arXiv preprint arXiv:1810.03715*, .

Karabiyik, H., Reese, S., & Westerlund, J. (2017). On the role of the rank condition in CCE estimation of factor-augmented panel regressions. *Journal of Econometrics*, *197*, 60–64.

Karabiyik, H., Westerlund, J., & Narayan, P. (2016). On the estimation and testing of predictive panel regressions. *Journal of International Financial Markets, Institutions and Money*, *45*, 115–125.

Kopoin, A., Moran, K., & Paré, J.-P. (2013). Forecasting regional GDP with factor models: How useful are national and international data? *Economics Letters*, *121*, 267–270.

Lee, L.-F., & Griffiths, W. E. (1979). The prior likelihood and best linear unbiased prediction in stochastic coefficient linear models. Center for Economic Research, Department of Economics, University of Minnesota, .

Mark, N. C., & Sul, D. (2011). When are pooled panel-data regression forecasts of exchange rates more accurate than the time-series regression forecasts? *Handbook of Exchange Rates*, (pp. 265–281).

Mark, N. C., & Sul, D. (2012). When are pooled panel-data regression forecasts of exchange rates more accurate than the time-series regression forecasts? In J. James, I. W. Marsh, & L. Sarno (Eds.), *Handbook of Exchange Rates* chapter 9. (pp. 265–281). Wiley.

Mayer, T., & Zignago, S. (2011). Notes on CEPII's distances measures: The GeoDist database. CEPII working paper.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*, 967–1012.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*, 265–312.

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

Pesaran, M. H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, *68*, 79–113.

Pesaran, M. H., & Tosetti, E. (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics*, *161*, 182–202.

Pesaran, M. H., & Zhou, Q. (2018). To pool or not to pool: Revisited. *Oxford Bulletin of Economics and Statistics*, *80*, 185–217.

Phillips, P. C. B., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, *6*, 217–259.

Song, M. (2013). Asymptotic theory for dynamic heterogeneous panels with cross-sectional dependence and its applications. Department of Economics, Columbia University.

Stock, J. H., & Watson, M. W. (1998). Diffusion indexes. *NBER Working Papers No. 6702*, .

Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, *44*, 293–335.

Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*, 1167–1179.

Swamy, P. A. (1970). Efficient inference in a random coefficient regression model. *Econometrica*, *38*, 311–323.

Trapani, L., & Urga, G. (2009). Optimal forecasting with heterogeneous panels: A Monte Carlo study. *International Journal of Forecasting*, *25*, 567–586.

Wang, W., Zhang, X., & Paap, R. (2018). To pool or not to pool: What is a good strategy for parameter estimation and forecasting in panel regressions? *Journal of Applied Econometrics*, *34*, 724–745.