UNIVERSITA' DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA GESTIONALE E DELL'INFORMAZIONE[°]
QUADERNI DEL DIPARTIMENTO[†]

Department of Management and Information Technology

Working Paper

Series "*Economics and Management*"

n. 16/EM – 2006

Optimal forecasting with heterogeneous panels: a Monte Carlo study

by

Lorenzo Trapani and Giovanni Urga

° Viale Marconi. 5, I – 24044 Dalmine (BG), ITALY, Tel. +39-035-2052339; Fax. +39-035-562779
† Il Dipartimento ottempera agli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modificazioni.

COMITATO DI REDAZIONE[§]

Series Economics and Management (EM): Stefano Paleari, Andrea Salanti
Series Information Technology (IT): Stefano Paraboschi
Series Mathematics and Statistics (MS): Luca Brandolini, Alessandro Fassò

[§] L'accesso alle *Series* è approvato dal Comitato di Redazione. I *Working Papers* ed i *Technical Reports* della Collana dei Quaderni del Dipartimento di Ingegneria Gestionale e dell'Informazione costituiscono un servizio atto a fornire la tempestiva divulgazione dei risultati dell'attività di ricerca, siano essi in forma provvisoria o definitiva.

Optimal Forecasting with Heterogeneous Panels: A Monte Carlo Study*

Lorenzo Trapani
Cass Business School, London (U.K.) and Universita' di Bergamo (Italy)

Giovanni Urga
Cass Business School, London (U.K.)

20 September 2004

Abstract

This paper reports the results of a series of Monte Carlo exercises to contrast the forecasting performance of several panel data estimators, divided into three main groups (homogeneous, heterogeneous and shrinkage/Bayesian). The comparison is done using different levels of heterogeneity, alternative panel structures in terms of T and N and using various error dynamics specifications. We also consider the presence of various degrees of cross sectional dependence among units. To assess the predictive performance, we use traditional measures of forecast accuracy (Theil's U statistics, RMSE and MAE), the Diebold and Mariano's (1995) test, and the Pesaran and Timmerman's (1992) statistics on the capability of forecasting turning points. The main finding of our analysis is that in presence of heterogeneous panels the Bayesian procedures have systematically the best predictive power independently of the model's features.

J.E.L. Classification Numbers: C23.
Keywords: Panel data; homogeneous, heterogeneous and shrinkage estimators; forecasting; cross dependence; Monte Carlo simulations.

---

*We wish to thank participants in the 11th International Panel Data Conference (Texas A&M University, June 4-6, 2004), the 2004 Royal Economic Society Conference (Swansea, 7-9 April 2004), the ESRC Econometric Study Group Annual Conference (Bristol, 15-17 July, 2004) and in seminars and workshops at Cass Business School, Bergamo University, European Central Bank for discussions and comments. The usual disclaimer applies.

1 Introduction

Over the last two decades a variety of estimation techniques have been proposed to estimate parameters of interest when panel data are available: Baltagi (2001), Arellano and Honore' (2001), Wooldridge (2002), Hsiao (2003), and Arellano (2003) provide comprehensive surveys on the topic. It has become customary to group these techniques into three main groups: homogeneous, heterogeneous and Bayesian (or shrinkage) estimators. While the first class assumes poolability of the data in the panel, and therefore parameters homogeneity across the panel units, the second one rejects this hypothesis taking into account explicitly the presence of heterogeneity among units. The class of Bayesian estimators is viewable as a hybrid solution between the two other classes (see Maddala, Li and Srivatsava, 1994, and Pesaran, Hsiao and Tahmiscioglu, 1999). It becomes then crucial to understand which estimation method is the "best", in statistical terms, for the specific research interest (e.g. bias reduction, efficiency, forecasting accuracy...).

Recently, in several seminal empirical papers Professor Badi Baltagi and associates have focused on investigating which estimator is the "best" when the specified model has to be used for forecast purposes. Baltagi and Griffin (1997), Baltagi, Griffin and Xiong (2000), Baltagi, Bresson and Pirotte (2002) and Baltagi, Bresson, Griffin and Pirotte (2002) apply dynamic panel specifications to industrial level data and find that the predictive ability of homogeneous estimators outperforms the predictive ability of heterogeneous and Bayesian estimators over any forecast horizon. Amongst the homogeneous estimators, GLS and within-2SLS emerge as the best estimators for forecasting purposes, especially when we forecast over a long time span. The superiority of the homogeneous estimators can sound quite reasonable when the panel is short, and when the degree of heterogeneity across units is limited, but it is rather puzzling when the time length T of the panel is large or when the degree of heterogeneity is high. This genuine empirical finding is particularly interesting because the model where we impose homogeneity is in general rejected by the data. A first interpretation of this apparent counter-intuitive empirical regularity is that a model that is "simple and parsimonious" offers a better forecasting performance.

It becomes therefore worth investigating whether these results hold generally speaking or if they are properties of the data considered in the works cited above, or, possibly, if the outcome of the comparison among the estimators forecasting performance is sensitively dependent on the number of units N and the time length of the panel T, and on the degree of the parameters heterogeneity across units. Our main objective in this work is to compare via a broad Monte Carlo simulation exercise the forecasting accuracy of several estimators belonging to each of the three classes (homogeneous, heterogeneous and shrinkage) for a routinely applied model (the dynamic specification with one or more exogenous covariates) under various circumstances. Such "circumstances" are the pair (N, T), the level of heterogeneity among units, the dynamic specification of the error term, and the existence and degree of cross sectional dependency across units. These issues are of paramount importance in determining the properties of estimators.

An important related question that arises in these circumstances is how to assess forecasting performance of a model. In their papers, Baltagi and associates use the standard Root Mean Square Error (RMSE) to measure forecasting accuracy. However, the literature on forecasting has developed a quite critical attitude towards this classical statistical measure. Thus in addition to the method based on RMSE, in our Monte Carlo experiments we are going to use also the approach based on different specifications for the loss function (Diebold and Mariano, 1995) and non parametric statistics that evaluates the ability to forecast change points due to Pesaran and Timmermann (1992).

The remainder of this paper is as follows. We set out the model we will be considering for our exercise, and briefly describe the estimation techniques and the predictive performance tests that we employ in our experiments (Section 2). We describe the details of the Monte Carlo experiments in Section 3, and report and comment the main results from the simulations in Section 4. Section 5 concludes.

2 Estimation and forecasting

2.1 Model

The DGP we employed for simulation is based on a dynamic specification and one strictly exogenous/predetermined variable:

$$y_{it} = \alpha_i + \beta_i y_{it-1} + \gamma_i x_{it} + u_{it} \tag{1}$$

where $i = 1, .., N$ and $t = 1, .., T$. Without loss of generality, the error term u_{it} is assumed to have no time specific effects since we focus on the impact of grouping across units[1] The possibility of having cross sectional dependence - i.e. the case $E[u_{it}u_{js}] \neq 0$ for some pair (i, j) - is not ruled out. Model (1) is the classical dynamic panel data specification, as discussed extensively in Baltagi (2001). It is also worth emphasizing that what we consider in our

[1] The applied literature seldom assumes the two ways error component model - an exception being the contribution by Baltagi, Griffin and Xiong (2000).

exercise are ex post forecasts, i.e. forecasts where the exogenous variable in model (1) is known without needing forecast it.

As far as estimation is concerned, we employed both homogeneous and heterogeneous estimators, performing an exercise similar to that in Baltagi, Bresson and Pirotte (2002), Baltagi and Griffin (1997), Baltagi, Bresson, Griffin and Pirotte (2002) and Baltagi, Griffin and Xiong (2000). Notice that whilst heterogeneous estimators are based on model (1), homogeneous estimators, assuming poolability of the data, are based on the following restricted specification of the DGP:

$$y_{it} = \alpha + \beta y_{it-1} + \gamma x_{it} + \varepsilon_{it}. \tag{2}$$

The error term ε_{it} is assumed to follow the well known one way specification:

$$\varepsilon_{it} = \mu_i + u_{it},$$

where μ_i is the unobservable individual specific effect and ν_{it} is the remainder of the disturbance - see Baltagi (2001) for a thorough discussion. The results of pooling using model (2) on estimators are discussed in Pesaran and Smith (1995) and Hsiao, Pesaran and Tahmiscioglu (1999).

2.2 Homogeneous, Heterogeneous and Shrinkage/Bayesian Estimators

We turn our discussion to estimation, referring to Baltagi (2001) for the details of each estimator.

2.2.1 Homogeneous estimators

The homogeneous estimators we consider fall into two main groups: least squares and instrumental variables estimators.

Within the class of least squares estimators, we first consider six standard pooled estimators applied to model (2): OLS, which ignores unit specific effects; first difference OLS to wipe out the effect of (possible) serial correlation in the error term; Within(-groups) estimator, which allows for unit specific effects; Between(-groups) estimator; and WLS and WLS-AR(1), where unit specific effects are assumed to be random. It is known that none of these estimates is either unbiased or consistent (see Pesaran and Smith, 1995, and the review in Baltagi, 2001). This is due to the assumption, common to all these estimators, that regressors are exogenous. However, the model we consider is dynamic and thus though all the explanatory variables are uncorrelated with the error components, the presence of either serial correlation in the remainder error term ν_{it} or of a random unit effect such as μ_i makes the lagged dependent variable correlated with the error term and therefore leads to potentially inconsistent estimates. The asymptotic bias of OLS has been assessed by Sevestre and Trognon (1985); it is also well known (see Nickell, 1981), that Within estimator is consistent only when $T \to \infty$, being biased of order $O(1/T)$ for finite T. The random effect WLS estimator is also biased and inconsistent, as pointed out in Baltagi (2001).

To achieve consistency, we may focus on pooled estimators based on instrumental variables. Thus, we first employ a standard 2SLS, which is consistent but not efficient; no attempt was made to improve efficiency by taking into account the unit specific effects. We also consider Within 2SLS, which, like its least squares counterpart, wipes out unit specific effects by transforming the data in deviations across their mean, and the Between 2SLS. Thirdly, we apply 2SLS to the first differenced version of model (2); this estimator (that henceforth will be referred to as FD-2SLS) is due to Anderson and Hsiao (1982) and is meant to eliminate fixed and random effects. However, given that this estimation procedure may induce autocorrelation in the remainder error term $\nu_{it} - \nu_{it-1}$, we also employ the correction proposed by Keane and Runkle (1992) that allows for arbitrary types of serial correlation[2]. This is applied to both the specification in levels (leading to an estimator denoted as 2SLS-KR) and the first differenced model (obtaining another estimate referred to as FD-2SLS-KR). Also, we employ EC2SLS estimator - see Baltagi (2001) - and EC2SLS-AR(1) - see Baltagi, Griffin and Xiong (2000) - to potentially achieve more efficiency by taking account of possible serial correlation in the error term[3]. As a variant of EC2SLS, we also compute the G2SLS estimator due to Balestra and Varadharajan-Krishnakumar (1987). It is worth noticing that such estimator has the same asymptotic covariance matrix as EC2SLS - see Baltagi and Li (1992) - but its performance is different in finite samples. Finally, we employ the Arellano and Bond (1991) estimation procedure, using a GMM estimation method on the specification in differences (whose outcome will be labelled as FDGMM) and also the same set of instruments in first difference on a specification in levels (GMM)[4].

[2]Such estimation technique can be applied only if $N > T$ - see Baltagi (2001).

[3]Note that these estimators, unlike standard 2SLS, also require an estimate of the variance components in order to be feasible.

[4]It is worth noticing that such GMM estimation procedures have existence conditions depending on the sizes of N, T and k (this latter being the number of parameters to be estimated) when the two-step GMM estimation is considered (see Baltagi, 2001) - this existence condition is $N > T(k-2) + (T+3)/2$. These estimators wouldn't have been feasible for all the cases we consider in our experiment, and we did not perform them.

Finally, we considered the MLE estimation on the ground of the discussion in Baltagi (2001), using the iterative procedure suggested by Breusch (1987).

In total, we compare 18 homogeneous estimators.

2.2.2 Heterogeneous estimators

The estimators considered so far are all characterized by the assumption of poolability of the data. This is a valid assumption only if the parameters in model (1) are homogeneous across units. As pointed out by Pesaran and Smith (1995) with respect to the dynamic pooled model, when parameters are heterogeneous, pooling leads to biased estimates. Therefore, we turned our attention also onto heterogeneous estimators.

In our Monte Carlo experiments we considered OLS and 2SLS applied to each unit i, obtaining Individual OLS and 2SLS. Given the presence of a lagged dependent variable, both estimates are biased. We then consider an average of both estimates (obtaining Average OLS and 2SLS), as suggested by Pesaran and Smith (1995). Averaging individual estimates leads to a consistent estimator as long as both N and T tend to infinity. We also compute the Swamy (1970) estimator, which belongs to the class of GLS and is a weighted average of the least squares estimates, using as weights the estimated covariance matrix.

In total we compare 5 alternative heterogeneous estimators.

2.2.3 Shrinkage/Bayesian estimators

We employed a class of shrinkage/Bayesian estimators - see Maddala, Li and Srivastava (1994) - where each individual estimate is shrunk towards the pooled estimates by weighting it with weight depending on the corresponding covariance matrix. The authors claim that shrinkage type estimator are superior to either homogeneous or to other heterogeneous estimators as far as predictive ability is concerned. The estimators we consider are the Empirical Bayes based on OLS initialization, the Empirical Bayes based on 2SLS estimation and their iterative counterparts. Finally, we implement the Hierarchical Bayes estimator using the same prior structure as in Hsiao, Pesaran and Tahmiscioglu (1999), which is found to have the best performance among heterogeneous estimators in terms of bias reduction, especially when T is small.

In total, we compare 5 alternative Bayesian estimators.

GAUSS code was anyway written and is available upon request.

2.3 Comparing forecasting performance

In this section we introduce the measures of forecasting performance we employ in our simulation exercise.

We employ three (classes of) measures of forecasting performance to assess the out-of-sample predicting ability of each estimator:

1. statistical measures of accuracy;
2. measure of statistical assessment of performance.
3. measures of the capability of predicting turning points.

The indicators we chose are, for each class:

1. MAE, RMSE and Theil's U statistics, whose expressions are respectively

$$MAE_j \equiv \frac{1}{h}\sum_{i=1}^{h} |\hat{y}_{ji} - y_{ji}|$$

$$RMSE_j \equiv \sqrt{\frac{1}{h}\sum_{i=1}^{h} (\hat{y}_{ji} - y_{ji})^2}$$

$$U_j \equiv \sqrt{\frac{\sum_{i=1}^{h} (\hat{y}_{ji} - y_{ji})^2}{\sum_{i=1}^{h} y_{ji}^2}}$$

   where: the index j refers to the j-th unit in the panel, h is the forecast horizon, $\hat{y}_{ji}$ is the forecast i steps ahead of y_{ji}. To obtain a single overall measure of performance, we considered the average of each indicator across units, similarly to Baltagi and associates papers. These indicators are all based on the residuals from forecast, and widely employed in the realm of forecasting. We calculate these three "classical" measures but we report and comment on the Theil's U statistics only, given its nature of relative measure which doesn't have the scaling problem of both RMSE and MAE. It is necessary to point out that using these indicators to assess forecasting accuracy has been widely criticised on the basis of statistical and economic considerations - for a general overview, see the review in Mariano (2002). From a statistical point of view, Clements and Hendry (1993) noted that the RMSE is not invariant to isomorphic transformations of models, and can therefore lead to contradictory results when applied to different (but isomorphic) representations of the same model. Moreover, Diebold and Lopez (1996) show that since RMSE depends only on the first two moments of the forecast distribution, it will suffer from serious shortcomings when such distribution is not adequately summarised by only two moments. The literature has criticised RMSE also on the grounds of economic considerations, arguing that predictive performance should be evaluated via the losses that arise from forecasting errors when certain decisions are made - see Leitch and Tanner (1991), Granger and Pesaran (2000a, 2000b), and the review by Pesaran and Skouras (2002). It has been shown that the RMSE is compatible with a quadratic loss function - see Pesaran and Skouras (2002) - but other specifications could be considered - see the discussions in Christoffersen and Diebold (1996) and Mariano (2002).

2. The Diebold and Mariano's (1995) test is a widely used alternative to overcome the inadequacies of RMSE since it is based on a loss function approach without needing specify the functional form. This statistics - with the adjustment for small sample bias proposed by Harvey, Leybourne and Newbold (1997) - can be used for any forecasting horizon h and doesn't require gaussianity, zero-mean, serial or contemporaneous incorrelation of the forecast errors, and under the null hypothesis of no difference between forecasting performances it is distributed as a standard normal. Formally, this statistic can be obtained as follows. Let $d_{ji}^{k} = \hat{y}_{ji} - y_{ji}$ be the forecast error at period i for series j when estimating parameters with an estimator k; assuming covariance stationarity and other regularity conditions, it is straightforward to show that
$$T^{-1/2}\left(\bar{d}_j - \mu_d\right) \Rightarrow N\left[0, 2\pi f\left(0\right)\right],$$
where $f\left(0\right)$ is the spectral density at frequency zero, $\mu_d = E\left(d_{ji}^{k}\right)$ and
$$\bar{d}_j = \sum_{i=1}^{h}\left[g\left(d_{ji}^{1}\right) - g\left(d_{ji}^{2}\right)\right]$$
with $g\left(\cdot\right)$ a generic loss function. Hence, the DM test is designed to compare the performance of two predictors; computationally, the statistic is set equal to
$$DM_j = \frac{\bar{d}_j}{\left[2\pi\hat{f}\left(0\right)/T\right]^{1/2}}.$$
The loss function we consider in order to compute the statistics is a

quadratic one, which allows us to compare pairwise RMSEs.[5] This enables us to detect whether one estimator has a superior predictive ability compared to another one by a proper testing rather than by the pure comparison of RMSE values. Even in this case, we compute the test statistics for every unit of the panel and then take the average across units.

3. Since forecasting performance could refer to something different from minimising a loss function, such as the capability to capture the sign of changes in the series rather than its values - see Granger and Pesaran (2000b). For this task, we employ Pesaran and Timmerman's (1992) statistics, defined as

$$PT_j = \frac{\hat{P}_j - \hat{P}_j^*}{\sqrt{\hat{V}\left(\hat{P}_j\right) - \hat{V}\left(\hat{P}_j^*\right)}} \sim N(0,1)$$

where

$$\hat{P}_j = h^{-1}\sum_{i=1}^{h} sign\left(\hat{y}_{ji}y_{ji}\right), \quad \hat{P}_j^* = \hat{P}_{yj}\hat{P}_{xj} + \left(1-\hat{P}_{yj}\right)\left(1-\hat{P}_{xj}\right),$$

$$\hat{V}\left(\hat{P}_j\right) = h^{-1}\hat{P}_j^*\left(1-\hat{P}_j^*\right),$$

$$\begin{aligned}\hat{V}\left(\hat{P}_j^*\right) &= h^{-1}\left(2\hat{P}_{yj}-1\right)^2\hat{P}_{xj}\left(1-\hat{P}_{xj}\right) + h^{-1}\left(2\hat{P}_{xj}-1\right)^2\hat{P}_{yj}\left(1-\hat{P}_{yj}\right) + \\ &\quad +4h^{-2}\hat{P}_{yj}\hat{P}_{xj}\left(1-\hat{P}_{yj}\right)\left(1-\hat{P}_{xj}\right)\end{aligned}$$

$$\hat{P}_{xj} = h^{-1}\sum_{i=1}^{h} sign\left(\hat{y}_{ji}\right), \quad \hat{P}_{yj} = h^{-1}\sum_{i=1}^{h} sign\left(y_{ji}\right),$$

where the function $sign(\cdot)$ takes the value of unity if its argument is positive and is equal to zero otherwise. Pesaran and Timmerman (1992) prove that this non parametric statistics is distributed as a standard normal under the null hypothesis that $\hat{y}_{ji}$ and y_{ji} are independent - and therefore that the predictor $\hat{y}_{ji}$ has no capability to forecast y_{ji}. Like

[5] The Diebold and Mariano testing procedure also requires a non parametric estimate of the spectral density of the difference of the loss associated with each predictor. The kernel we employ is the truncated rectangular one employed by Diebold and Mariano (1995), and the bandwidth we choose is specified as $m(h) = 1 + \lfloor \log(h) \rfloor$, where the operator $\lfloor \cdot \rfloor$ denotes the rounding to the closest integer.

in the previous point, here we compute the Pesaran and Timmerman statistics for each unit of the panel and then report its average value across units. Notice that this measure could be also employed as a descriptive measure to rank forecasting techniques (see *inter alia* Driver and Urga, 2004).

Having described the estimators considered and the methods of evaluating forecasting accuracy, in the next section we illustrate the design of the Monte Carlo experiment.

3 The design of the Monte Carlo experiments

We generate a sample of N units with length $T+T0$, where $T0$ is the number of initial values to be discarded to avoid dependence on the initial conditions (set equal to 0). We let the number of units N and the time dimension T assume various values.

The DGP we generate at each replication is the one given in model (1):

$$y_{it} = \alpha_i + \beta_i y_{it-1} + \gamma_i x_{it} + u_{it},$$

where:

- the parameters α_i, β_i and γ_i are generated as, respectively:

$$\alpha_i = \bar{\alpha} + \alpha^H N_i^{\alpha},$$
$$\beta_i = \bar{\beta} + \beta^H N_i^{\beta},$$
$$\gamma_i = \bar{\gamma} + \gamma^H N_i^{\gamma},$$

  where $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$ are the mean values of the parameters, $N_i^{\cdot}$ denotes an independent (across i) extraction from a random variable and α^H, β^H and γ^H control the parameters heterogeneity across units, which will be useful throughout the set of simulations to assess the predictive performance of the estimators. Notice that whilst we employed standard normals for α_i and γ_i, β_i was simulated via a uniform distribution with bounded support so as to rule out the possibility of having a value larger than (or equal to) unity;

- the disturbance u_{it} is, in a first set of experiments, assumed to follow a stationary, invertible Gaussian ARMA(1,1) specification defined by

$$u_{it} = \rho u_{it-1} + \zeta_{it} + \vartheta \zeta_{it-1},$$

and the parameters (ρ, ϑ) control the degree of autocorrelation of the error term in model (1). The error term is then rescaled by the factor $\lambda = \sqrt{(1+\vartheta)/(1-\rho)}$ to give it unit variance. Here there is no cross sectional dependence across units, since u_{it} is generated independently of u_{jt} for any pair (i, j). In a second set of experiments, we take into account the presence of cross sectional dependence by modelling the error (now denoted as $u_{it}^{'}$) as

$$u_{it}^{'} = u_{it} + \zeta_i f_t,$$

where f_t is a standard normal independent over t and ζ_i is a uniformly distributed random variable whose support is chosen as $[0, 0.2]$ to model small cross section dependence and $[-1, 3]$ to represent a large amount of cross section dependence. This part of the experiments to modelling cross sectional dependence follows the same line of Pesaran (2003);

- the explanatory variable x_{it} is generated with the following DGP:

$$x_{it} = \alpha_i + \beta_i + \delta x_{it-1} + \eta_{it}, \tag{3}$$

where the error term η_{it} is a Gaussian white noise generated independently of u_{it}. The presence of the term $\alpha_i + \beta_i$ introduces a correlation between η_{it} and the error term in the random effect specification (2)

$$\varepsilon_{it} = \mu_i + u_{it}.$$

This correlation is such that $E(x_{it}\nu_{it}) = 0$ for any i - and hence x_{it} endogeneity is ruled out - and $E(x_{it}\mu_i) \neq 0$. This two results make x_{it} a strictly exogenous variable and a valid instrument for GMM estimation *a la* Arellano and Bond (1991) thanks to its correlation with the unit specific effect - see Baltagi (2001) for discussion.

We considered the following values for the parameters of our simulation exercise:

- we ran 5000 iterations for each simulation, and 2500 iterations (500 of which in the burn-in period) for every Gibbs sampling algorithm - on the ground of the results in Hsiao, Pesaran and Tahmiscioglu (1999);

- as far as the autocorrelation structure is concerned, we considered (ρ, ϑ) to be equal either to $(0, 0)$ or to $(0.9, 0.9)$[6]. These two pairs represent the cases of non autocorrelation and of near integration, respectively;

[6]Further developments of this work will consider the following spectrum of values for (ρ, ϑ): $\{-0.9, -0.3, 0, 0.3, 0.9\} \times \{-0.9, -0.3, 0, 0.3, 0.9\}$

- the number of initial observations to be discarded was set equal to $T0 = 100$;
- the forecasting horizon is set equal to $h = 10$[7].

4 Simulation results

In this section we report and comment the full set of results from the various Monte Carlo experiments using the three forecasting accuracy tests.We consider two different degrees of heterogeneity assuming (in α^H, β^H, γ^H, $H = 0.1$ and 0.9 respectively); two different specifications for the error dynamics, namely (ρ, ϑ) were set equal to $(0, 0)$ and $(0.9, 0.9)$; in addition to the case of no cross dependence, two alternative degree of cross sectional dependence are considered, namely the case of "mild" cross dependence (ζ_i $[0, 0.2]$) and one with "large" cross sectional dependence (ζ_i is now $[-1, 3]$). Finally, the pairs of (T, N) we consider are $(5, 10)$, $(5, 20)$, $(10, 20)$, $(10, 50)$, $(20, 50)$ and $(50, 50)$.

The presentation of the full set of experiments are reported in details in a companion paper (Trapani and Urga, 2004).

4.1 Statistical measures of accuracy

In this section we report the results for the Theil's U statitics (Tables A1-A12)[8]. Each table is divided in three panels. We report the statistics for the homogenoeus, heterogeneous and shirnkage/Bayesian estimators respectively.

[Insert somewhere here Table A1-A12]

- Heterogeneity plays a very important role and has a strong impact on the outcomes of the simulation exercises. When the degree of heterogeneity is low (columns with $H = 0.1$) and the amount of dependence among units is mild, homogeneous estimators prevail. Such findings are in the line with what reported in Baltagi and Associates. Note that the results from homogeneous estimators are very clsoed to those obtained from the class of shrinkage/Bayesian estimators. However, by

[7] Our results can be extended to the cases $h = 1$ and $h = 5$, as in Baltagi, Bresson and Pirotte (2002), Baltagi and Griffin (1997), Baltagi, Bresson, Griffin and Pirotte (2002) and Baltagi, Griffin and Xiong (2000)

[8] We also computed RMSEs and the MAEs for each simulation. The findings remain unchaged. The results are available upon request.

increasing the level of heterogeneity ($H = 0.9$) homogeneous estimators are outstaged by the shrinkage/Bayesian estimators. While the statistics from shrinkage/Bayesian estimators does not change very much with respect to the case of low heterogeneity, we note massive changes affecting homogeneous estimators.

- The impact of cross sectional dependence is also quite substantial. In the case of mild cross dependence our findings are very much in the line with what showed in the existent applied literature. When instead we consider the case of large contemporaneous correlation, the statistics change dramatically: irrespectively of the level of heterogeneity and of all other characteristics of the panel (combination of T and N, dynamics of the error terms), the estimators that show the best forecasting accurancy are always the shrinkage/Bayesian ones. The presence of cross sectional dependence has an impact on the shrinkage/Bayesian estimators in the sense that the statistics get worse as cross dependence gets larger, and this seems to suggest that an increasing presence of cross dependence makes forecasting in general more difficult in this case.

- The time and cross sectional dimensions do not play a substantial role. Some estimators look very sensitive to N and T (particularly to T, while N does not have a great impact on the predictive ability), especially in the small sample case. On the other hand, most estimators are not sensitive to the values of the pair (T, N). This is particularly evident in the the case of the Hierarchical Bayes estimator, whose prediction outcome is almost invariant with T. This result confirms previous findings reported in Hsiao, Pesaran and Tahmiscioglu (1999).

- The error term dynamics does have an impact on the choice of the best estimator when cross dependence is mild. However, first difference homogeneous estimators outperform all other estimators as long as heterogeneity is low. Moreover, both the presence of high heterogeneity or high cross sectional dependence wipes out any effects due to the presence of serial correlation, and in this case too shrinkage/Bayesian estimators dominate.

- The small sample problem, uniquely related to the time series dimension T, arises when using Individual OLS and 2SLS. For $T = 5$ (see Tables A1-A4), the Theil's statistics is never lower than 10^4, and therefore their forecasting capability is totally implausible. This also affects the performance of shrinkage estimators, whose magnitude of the Theil's statistics is much larger (at least of a factor 10^2) than that of the best estimators. Thus, for the case of a short panel ($T = 5$ in our case), our results contradict the findings reported in Maddala, Li and Srivastava (1994).

To summarise, there is a clear evidence that with a few exceptions, the class of Bayesian estimator outperforms the alternative estimator independendly of the control indicators per experiment.

4.2 Diebold and Mariano's (1995) test

The outcome of Diebold and Mariano test is represented by a lower triangular matrix of dimensions 28×28 for each experiments. Since the amount of output generated by this part of the exercise exceeds a reasonable number of pages, we decide not to report it.[9]

The main results confirm the conclusions reported in the previous section for the Theil's U statistic, reinforcing some regularities encountered previously. The main findinds can be summarised as follows:

- For the mild cross sectional dependence case, there seems to be no statistically significant difference between shrinkage and homogeneous estimators when the degree of heterogeneity is small, and therefore either class of estimators can be used irrespectively of any feature of the data. On the other hand, when H is large, shrinkage estimators have a significantly better performance (especially for the small T case), and therefore the conclusion that they should be preferred in such case is reinforced. Only when the error component dynamics is characterised by a nearly integrated behavior the performance of homogeneous, first difference, estimators is mildly (though significantly) better than that of shrinkage estimators based on the model in levels;

- When cross sectional dependence is large, for T larger than 5 the performance of either class of estimators (homogeneous and shrinkage) is not significantly different. However, even in this case the nearly integrated case shows a mildly significant better performance of the first differecne homogeneous estimators. It should be pointed out that when $T = 5$ the comparison between homogeneous and shrinkage estimators can not be done given the extremely poor performance of the latter class.

[9] All results on Diebold and Mariano's test are available upon request.

- When $T = 5$, though none of the estimators has a significantly better performance than the others, however there is statistical evidence that Hierarchical Bayes is marginally better when heterogeneity increases;

- When T increases, the difference between homogeneous and heterogeneous estimators becomes significant, the latter group performing better than the former. This holds especially, as already seen, when heterogeneity increases. When we have a small amount of heterogeneity there is virtually no difference across estimators, Hierarchical Bayes included. Such finding illustrates that as long as heterogeneity is limited across units the choice of estimators is not crucial for the forecasting performance. This holds for $T = 10$ or greater. Finally, the presence of serial correlation in the error term doesn't affect these findings.

- The main findings so far are reinforced when the number of units is large (i.e. $N = 50$). Here too the presence of heterogeneity is crucial in marking the difference between pooled and heterogeneous estimators, in favuor of the latter.

- The gain from considering an iterative shrinkage estimator rather than a non iterative one has been assessed as poor when we used the Theil's U statistics. The Diebold and Mariano test reinforces this conclusion showing no statistical difference between the two types of estimators.

- It is worth noticing that the presence of cross-sectional dependence preserves the predictive accuracy ranking of the estimators.

4.3 Measures of capability to forecast turning points

In this section, we describe the results of our Monte Carlo for the Pesaran and Timmermann's (1992) statistics (Tables B1-B12).

[Insert somewhere here Table B1-B12]

Since Pesaran and Timmermann's test is asymptotically distributed as a standard normal under the null hypothesis of no capability to detect turning points, the values in Tables can be interpreted either as raw numbers to rank estimators (the larger the value of the statistics, the higher the turning points detection capability), or we may compare them with quantiles of the normal distribution to test whether each estimator predicting capability is significant or not.

The main findings can be summarised as follows:

- The impact of heterogeneity on the capability of forecasting turning points is very much along the same lines as on Theil's U statistic case. When cross sectional dependence is mild, low heterogeneity leads to homogeneous estimators, whilst increasing heterogeneity results in having the shrinkage/Bayesian estimators to be the best ones and it makes the performance of homogeneous estimators very poor. This pattern changes when the amount of contemporaneous dependence across units increases, and it makes homogeneous estimators less capable to forecasting turning points even in the presence of near homogeneity ($H = 0.1$). A striking regularity that can be noted is that a larger heterogeneity improves the turning point prediction ability, making on average the value attained by Pesaran and Timmermann statistic higher.

- As far as the impact of cross sectional dependence is concerned, this is the most remarkable feature in our set of experiments. Whilst, as already noted earlier in the paper, whenever mild levels of cross correlation are present it is always possible to find an estimator whose turning point prediction ability is statistically significant, in presence of large cross sectional dependence it is virtually impossible to find an estimator that is significantly capable of predicting turning points, with few exceptions in the class of Bayesian estimators.

- The time series size T has an impact on the Pesaran and Timmermann's statistic, which show greater predictive performance when T increases, while this does not extend to the cross sectional dimension N, whose increase has an ambiguous impact.

- The impact of dynamics has some commonalities with the Theil's U statistic case. Here too a nearly integrated error term results in having a better predictive performance on the side of first difference homogeneous estimators when heterogeneity is limited; in this case as well the presence of either heterogeneity or cross sectional correlation makes predictive performance worse.

- As far as small samples are concerned, when $T = 5$ and cross dependence is small it is still possible to find estimators that are significantly capable of identifying turning points. In this case, an increase in N has the effect of improving the forecasting performance. A striking result is that for $T = 5$ and $N = 20$, the predictive ability of Individual estimators is significant and very close to be the best among all estimators, albeit these estimates are computed for each unit with a degree of freedom equal to 2. This outcome is completely different than in the previous case, and it should lead to conclude that predictive performance measured with Theil's U statistics (and with other statistical indicators such as RMSE and MAE, too) is different and unrelated with this aspect of forecasting performance.

- Whilst the presence of mild dependence among units leaves results unchanged, assuming large cross dependence strongly affects them. Particularly, values taken by the statistics decrease dramatically, almost independently of the other parameters. From the statistical viewpoint, when cross dependence is mild or absent one can always find an estimator whose turning point predictive ability is significant; surprisingly, this is not the case when there is no dynamics together with low heterogeneity (as reported in the first column of all Tables). Moreover, as Tables B2, B4, B6, B8, B10, B12 show, none of the estimators would pass a 5% significance test.

- The performance of GMM based estimators shows an interesting pattern. Even though the instruments for these estimators are chosen with respect to a specification in levels for the model, the performance of estimates based on the specification in first differences is always better than the one based on specification (1). The difference between the two is found to be very large when heterogeneity is large.

- The presence of heterogeneity always improves the predictive ability of heterogeneous and shrinkage estimators. The latter are always the best when heterogeneity is high, when the test is always statistically different from zero, the only exception being the case of large cross sectional dependence.

- The presence of a nearly integrated dynamics makes homogeneous estimators based on the first differenced model the best, as shown by the third column of all Tables. Their performance is anyway heavily affected by cross dependence, and from being statistically significant when this is mild or absent, it becomes insignificant when the system exhibits a large degree of covariance among units.

4.4 A summary of the main features of our findings

In this final section, the main features of the various experiments commented above are presented via a summary of the three sets of statistics (Tables C1-C12). Each of the table is divided in three panels. The first one reports the best estimators according to Theil's U statistics. In the second panel, using the Diebold and Mariano (1995) test (DM), we report the comparison between the best and the second best estimator between the above estimators (DM1) plus a comparison between the best estimator and the best Bayesian estimator (DM2), Finally, the last panel (PT) reports the best estimator according to the Pesaran and Timmerman (1992) statistic.

[Insert somewhere here Table C1-C12]

The main results can be summarised as follows:

- the estimators performance seems to be affected only by the degree of cross sectional dependence and heterogeneity, being independent of the error term dynamics and of the time and corss sectional dimensions (N, T);
- when cross sectional dependence is small, the best class of estimators is either the homogeneous one - when heterogeneity H is limited - or the shrinkage one - which happens when heterogeneity is set to a large value. This regularity always takes place, irrespectively of any other feature of the data;
- when cross sectional dependence is large, the best estimators are almost always the shrinkage ones for T larger than 5. This does not hold when the error term exhibits a nearly integrated dynamics, as in such case estimates based on the first differenced data achieve the best performance. When $T = 5$ it should be pointed out the poor performance of both heterogeneous (which is likely to be due to the limited degree of freedom in each equation) and shrinkage estimators (essentially because their prior is not designed to take account of the presence of contemporaneous correlation).

5 Conclusions

In this paper, we compare the predictive performance of several homogeneous, heterogeneous and shrinkage estimators applied to a heterogeneous model. We analyze the forecasting performance of the various estimators by varying the degree of heterogeneity in the panel and using alternative specifications for error dynamics.

Our main results are that for short panels with a limited degree of heterogeneity, homogeneous estimators are preferable to the heterogeneous ones. This confirms the findings of Baltagi, Bresson and Pirotte (2002), Baltagi and

Griffin (1997), Baltagi, Bresson, Griffin and Pirotte (2002) and Baltagi, Griffin and Xiong (2000). This outcome is not affected by the the dynamics in the error term specification. For the case of near integration, the homogeneous first difference estimators performs best. Homogeneous estimators perform also well when T increases and heterogeneity is small. On the other hand, the performance of heterogeneous estimators show some sign of improvements with respect to the case of small T mainly due to the higher degree of freedom in the individual regressions. The good performance of the homogenous estimators is better than that of the Hierarchical Bayes estimator and of shrinkage estimators, though the Hierarchical Bayes has in general a better performance across all experiments, regardless of heterogeneity or error dynamics.

Heterogeneity greatly affects the performance of the various estimators. Homogeneous estimators show a very poor predictive performance, and in panels with long T heterogeneous estimators are preferable. Anyway, under the presence of heterogeneity the shrinkage estimators and the Hierarchical Bayes estimator show the best performance. This conclusion is consistent with Hsiao, Pesaran and Tahmiscioglu (1999) analysis, that shows how Bayesian estimation gives the best results in terms of bias reduction. Diebold and Mariano's test shows that shrinkage estimators performance is also *significantly* better than the one of the other estimators.

Our findings provide a clear guideline to practitioners when panel data are available for forecasting purposes: use Bayesian procedures to forecast with heterogeneous panels.

References

[1] Anderson, T.W., Hsiao, C. (1982), "Formulation and Estimation of Dynamic Models with an Error Component", *Journal of Econometrics*, Vol. 18, pp. 47-82.

[2] Arellano, M. (2003), *Panel Data Econometrics*, Oxford University Press, Oxford.

[3] Arellano, M., Honore', B. (2001), "Panel Data Models: Some Recent Developments", in Heckman, J.J. and E. Leamer (eds.), *Handbook of Econometrics*, vol. 5, chapter 53, North-Holland.

[4] Arellano, M., Bond, S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, pp. 277-97.

[5] Balestra, P., Vadharajan-Krishnakumar, J. (1987), "Full Information EStimation of a System of Simultaneous Equations with Error Component Structure", *Econometric Theory*, Vol. 3, pp. 223-246.

[6] Baltagi, B.H. (2001), *Econometric Analysis of Panel Data,* Wiley, 2nd edition.

[7] Baltagi, B.H., Bresson, G., Griffin, J.M., Pirotte, A. (2002), "Homogeneous, Heterogeneous or Shrinkage Estimators? Some Empirical Evidence from French Regional Gasoline Consumption", mimeo, Texas A&M University.

[8] Baltagi, B.H., Bresson, G., Pirotte, A. (2002), "Comparison of Forecast Performance for Homogeneous, Heterogeneous and Shrinkage Estimators. Some Empirical Evidence from US Electricity and Natural-Gas Consumption", *Economics Letters*, 76, pp. 375-82.

[9] 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, pp. 303-27.

[10] Baltagi, B.H., Griffin, J.M., Xiong, W. (2000), "To Pool or not to Pool: Homogeneous versus Heterogeneous Estimators Applied to Cigarette Demand", *The Review of Economics and Statistics*, 82(1), pp. 117-26.

[11] Baltagi, B.H., Li, Q. (1992), "A Note on the Estimation of Simultaneous Equations with Error Components", *Econometric Theory*, 8, pp. 113-119.

[12] Breusch, T.S. (1987), "Maximum Likelihood Estimation of Random Effects Models", *Journal of Econometrics*, Vol. 36, pp. 383-389.

[13] Clements, M.P., Hendry, D. (1993), "On the Limitations of Comparing Mean Square Forecast Errors (with Discussion)", *Journal of Forecasting*, 12, pp. 617-76.

[14] Clements, M.P., Hendry, D. (2002), *A Companion to Economic Forecasting*, Blackwell Publishers, Oxford.

[15] Christoffersen, P.F., Diebold, F.X. (1996), "Further Results on Forecasting and Model Selection under Asymmetric Loss", *Journal of Applied Econometrics*, 11, pp. 561-71.

[16] Cornwell, C. and P. Rupert (1988), "Efficient Estimation with Panel Data: an Empirical Comparison of Instrumental Variables Estimators", *Journal of Applied Econometrics*, Vol. 3, pp. 149-55.

[17] Diebold, F.X., Mariano, R. (1995), "Comparing Predictive Accuracy", *Journal of Business Economics and Statistics*, 13, pp. 253-63.

[18] Diebold, F.X., Lopez, J.A. (1996), "Forecast Evaluation and Combination", in G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics*, Amsterdam, North-Holland, pp. 241-268.

[19] Driver, C., Urga, G. (2004), "Transforming Qualitative Survey Data: Performance Comparison for the UK", *Oxford Bulletin of Economics and Statistics*, 66, pp. 71-89.

[20] Keane, M.P., Runkle, D.E. (1992), "On the Estimation of Panel Data Models with Serial Correlation when Instruments are not Strictly Exogenous", *Journal of Business and Economic Statistics*, 10, pp. 1-9.

[21] Granger, C.W.J., Giacomini, R. (2004), "Aggregation of Space-Time Processes", *Journal of Econometrics*, Vol. 118, pp. 7-26.

[22] Granger, C.W.J., Pesaran, M.H. (2000a), "A Decision-based Approach to Forecast Evaluation", in W.S. Chan, W.K. Li, H. Tong (eds.), *Statistics and Finance: an Interface*, Imperial College Press, London..

[23] Granger, C.W.J., Pesaran, M.H. (2000b), "Economic and Statistical Measures of Forecast Accuracy", *Journal of Forecasting*, 19, pp. 537-60.

[24] Harvey, D., Leybourne, S., Newbold, P. (1997), "Testing the Equality of Prediction Mean Squared Errors", *International Journal of Forecasting*, Vol. 13, pp. 281-291.

[25] Hsiao, C. (2003), *Analysis of Panel Data*, 2nd Edition, Cambridge University Press.

[26] Hsiao, C., Pesaran, M.H., Tahmiscioglu, A.K. (1999), "Bayes Estimation of Short-Run Coefficient in Dynamic Panel Data", in Hsiao, C., Lahiri, K., Lee, L.-F. and Pesaran, M.H. (eds.), *Analysis of Panels and Limited Dependent Variable Models*, Cambridge University Press, Cambridge.

[27] Leitch, G., Tanner, J.E. (1991), "Economic Forecast Evaluation: Profits versus the Conventional Error Measures", *American Economic Review*, Vol. 81, No.3, pp. 580-90.

[28] Maddala, G.S., Li, H., Srivastava, V.K. (1994), "A Comparative Study of Different Shrinkage Estimators for Panel Data Models", *Annals of Economics and Finance*, Vol. 2, pp. 1-30.

[29] Mariano, R. (2002), "Testing Forecasting Accuracy", in M.P. Clements and D.F. Hendry (eds.), *A Companion to Economic Forecasting*, Blackwell, pp. 284-98.

[30] Nickell, S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica*, Vol. 16, pp. 1-32.

[31] Pesaran, M.H. (2003), "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence", mimeo, Cambridge University.

[32] Pesaran, M.H., Hsiao, C., Tahmiscioglu, A.K. (1999), "Bayes Estimation of Short-Run Coefficient in Dynamic Panel Data", in Hsiao, C., Lahiri, K., Lee, L.-F. and Pesaran, M.H. (eds.), *Analysis of Panels and Limited Dependent Variable Models*, Cambridge University Press, Cambridge.

[33] Pesaran, M.H., Smith, R. (1995), "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics*, Vol. 68, pp. 79-113.

[34] Pesaran, M.H., Skouras, S. (2002), "Decision-Based Methods for Forecast Evaluation", in M.P. Clements and D.F. Hendry (eds.), *A Companion to Economic Forecasting*, Blackwell, pp. 241-67.

[35] Pesaran, M.H., Timmermann, A. (1992), "A Simple non Parametric Test of Predictive Performance", *Journal of Business and Economic Statistics* 10, pp. 461-65.

[36] Sevestre, P., Trognon, A. (1985), "A Note on Autoregressive Error Component Models", *Journal of Econometrics*, Vol. 28, pp. 231-245.

[37] Swamy, P.A.V.B. (1970), "Efficient Inference in a Random Coefficient Regression Model", *Econometrica*, Vol. 38, pp. 311-23.

[38] Trapani, L., Urga, G. (2004), "Assessing Predictive Performance of Homogenous, Heterogeneous and Shrinkage Estimators for Heterogeneous Panels: A Monte Carlo Study", W.P. N. XX, Centre for Econometric Analysis, Cass Business School, London.

[39] Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, Massachusetts.

APPENDIX

Table A1: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	(0.1)	(0.9)	(0.1)	(0.9)
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	**0.4276**	1.8041	0.4821	1.8293
Within	0.4320	1.8715	0.4485	1.9070
Between	0.7010	1.7173	0.7852	1.7625
FD-OLS	1.1659	1.5383	1.0066	1.9551
WLS	0.4377	1.8868	0.4750	1.9430
WLS-AR(1)	179.8	6.2749	1.4677	7.4705
2SLS	0.4306	5.6352	0.4608	4.1889
FD-2SLS	0.7798	$5{\cdot}10^7$	**0.3257**	1.9909
Within-2SLS	**0.4276**	1.8044	0.4821	1.8293
Between-2SLS	0.4344	2.0459	0.4512	792.3
MLE	0.4292	1.7336	0.4490	1.7613
EC2SLS	0.4729	1.6336	0.4774	1.6459
EC2SLS-AR(1)	0.4765	1.5706	0.4577	1.5989
G2SLS	0.4684	799.1	0.4720	$2{\cdot}10^6$
2SLS-KR	0.4334	5977	0.4659	$9.8{\cdot}10^5$
FD-2SLS-KR	0.7765	2615	0.3261	7.7820
FDGMM	0.7658	1.4166	0.3328	1.8479
GMM	0.44534	1.4349	0.5481	1.4278
Ind. OLS	$7.5{\cdot}10^5$	$1.9{\cdot}10^6$	7.9739	$7.4{\cdot}10^5$
Ind. 2SLS	$7.5{\cdot}10^5$	$1.9{\cdot}10^6$	7.9739	$7.4{\cdot}10^5$
Average OLS	0.4575	1.7823	0.4520	1.5321
Average 2SLS	0.4575	1.7823	0.4520	1.5321
Swamy	0.4319	1.8205	0.4620	1.8401
Bayes OLS	$2.4{\cdot}10^4$	$8.7{\cdot}10^5$	3.8284	$3.6{\cdot}10^5$
It. Bayes OLS	6193	4241	3.4553	$3.2{\cdot}10^5$
Bayes 2SLS	$2.4{\cdot}10^4$	$8.7{\cdot}10^5$	3.8284	$3.6{\cdot}10^5$
It. Bayes 2SLS	6193	4241	3.4553	$3.2{\cdot}10^5$
It. Bayes	0.4853	**0.5307**	0.3895	**0.4832**

Notes: $H = 0.1$ and $H = 0.9$ represent the cases of low and large heterogeneity in the panel, respectively. (b) Two alternative specifications for the error term dynamics: white noise with $(\rho, \vartheta) = (0, 0)$, and a nearly integrated with $(\rho, \vartheta) = (0.9, 0.9)$ (c) Forecasting horizon ($h =$) 10 periods ahead.

Case with $(N, T) = (5, 10)$ and "mild" cross dependence (the support of ζ_i is [0,0.2])

Table A2: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	(0.1)	(0.9)	(0.1)	(0.9)
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.7620	1.5374	0.8002	1.5506
Within	0.7763	1.5235	0.8268	1.5311
Between	4.5055	1.7486	1.8419	1.7716
FD-OLS	1.1583	1.3131	1.2317	1.3937
WLS	0.8505	1.5798	0.8397	1.6053
WLS-AR(1)	2669	7.1829	5259	1112
2SLS	$1.8{\cdot}10^5$	61.52	$3.8{\cdot}10^6$	14.07
FD-2SLS	$8.23{\cdot}10^9$	$1.5{\cdot}10^8$	$1.45{\cdot}10^{12}$	$3.2{\cdot}10^{13}$
Within-2SLS	0.7620	1.5374	**0.8000**	1.5506
Between-2SLS	$6.5{\cdot}10^5$	$1.45{\cdot}10^5$	$2.73{\cdot}10^{10}$	$3.07{\cdot}10^8$
MLE	**0.7618**	1.4434	0.8215	1.4484
EC2SLS	0.9005	1.4060	0.9855	1.4133
EC2SLS-AR(1)	0.8933	1.4529	0.9592	1.4601
G2SLS	1.9858	4860	2.5763	$1.42{\cdot}10^6$
2SLS-KR	3684	$1.12{\cdot}10^5$	2227	$2.66{\cdot}10^6$
FD-2SLS-KR	$5.7{\cdot}10^5$	$1.26{\cdot}10^4$	$1.02{\cdot}10^6$	$2.6{\cdot}10^5$
FDGMM	0.9368	**1.1896**	0.8840	**1.2822**
GMM	0.8091	1.3821	0.8990	1.3922
Ind. OLS	$7.6{\cdot}10^8$	$1.3{\cdot}10^8$	$4.96{\cdot}10^4$	$2.4{\cdot}10^8$
Ind. 2SLS	$7.6{\cdot}10^8$	$1.3{\cdot}10^8$	$4.96{\cdot}10^4$	$2.4{\cdot}10^8$
Average OLS	1.3045	2.2297	0.8922	1.6972
Average 2SLS	1.3045	2.2297	0.8922	1.6972
Swamy	0.7646	1.5375	0.8303	1.5523
Bayes OLS	$1.01{\cdot}10^8$	$8.1{\cdot}10^5$	9728	$1.40{\cdot}10^6$
It. Bayes OLS	$6.8{\cdot}10^5$	2855	7466	$9.39{\cdot}10^4$
Bayes 2SLS	$1.01{\cdot}10^8$	$8.1{\cdot}10^5$	9728	$1.40{\cdot}10^6$
It. Bayes 2SLS	$6.8{\cdot}10^5$	2855	7466	$9.39{\cdot}10^4$
It. Bayes	134.84	61.45	53.46	13.27

Notes: See Table A1.

Case with $(N, T) = (5, 10)$ and "large" cross dependence (the support of ζ_i is [-1,3])

Table A3: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	**0.4405**	1.2830	0.4957	1.4728
Within	0.4458	4.7534	0.4622	6.0493
Between	0.5476	1.2063	0.6701	1.3196
FD-OLS	1.1728	1.7763	0.9946	2.9064
WLS	0.4435	1.5327	0.4960	2.0016
WLS-AR(1)	84.52	2.3381	0.8205	5.4995
2SLS	0.4420	1.4184	0.4725	1.6100
FD-2SLS	0.7666	1.9080	**0.2923**	2.0287
Within-2SLS	**0.4405**	1.2830	0.4957	1.4728
Between-2SLS	0.4431	13.07	0.4657	12.242
MLE	0.4410	4.0369	0.4621	5.1148
EC2SLS	0.4830	1.3034	0.4866	1.5228
EC2SLS-AR(1)	0.4856	1.4198	0.4642	1.7396
G2SLS	0.4786	1.4735	0.4812	1.7036
2SLS-KR	0.4429	1.2372	0.4676	1.3340
FD-2SLS-KR	0.7664	5.2250	**0.2923**	605.7
FDGMM	0.7624	1.3367	0.2983	2.1542
GMM	0.4608	1.4774	0.5375	1.7409
Ind. OLS	$6.9{\cdot}10^6$	$4.8{\cdot}10^{10}$	$1.6{\cdot}10^6$	$1.91{\cdot}10^7$
Ind. 2SLS	$6.9{\cdot}10^6$	$4.8{\cdot}10^{10}$	$1.6{\cdot}10^6$	$1.91{\cdot}10^7$
Average OLS	0.4583	2.5986	0.4628	1.2699
Average 2SLS	0.4583	2.5986	0.4628	1.2699
Swamy	0.4427	1.4297	0.4719	1.7115
Bayes OLS	$2.73{\cdot}10^6$	$1.33{\cdot}10^8$	2574	$1.37{\cdot}10^7$
It. Bayes OLS	$1.81{\cdot}10^6$	$7.14{\cdot}10^7$	1829	$1.34{\cdot}10^7$
Bayes 2SLS	$2.73{\cdot}10^6$	$1.33{\cdot}10^8$	2574	$1.37{\cdot}10^7$
It. Bayes 2SLS	$1.81{\cdot}10^6$	$7.14{\cdot}10^7$	1829	$1.34{\cdot}10^7$
It. Bayes	0.4752	**0.4988**	0.4037	**0.4899**

Note: See Table A1.

Case with $(N, T) = (5, 20)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table A4: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.7790	1.3527	**<u>0.8136</u>**	1.4234
Within	0.7943	3.3196	0.8554	3.6178
Between	1.6673	1.5365	1.5862	1.5977
FD-OLS	1.1831	1.3881	1.2702	1.5522
WLS	0.8616	1.4960	0.8457	1.6275
WLS-AR(1)	1463	50.31	197	8.8379
2SLS	6.5321	1.7271	5.0744	1.7927
FD-2SLS	546	$2.31{\cdot}10^5$	$9.63{\cdot}10^5$	$6.66{\cdot}10^4$
Within-2SLS	0.7790	1.3527	**<u>0.8136</u>**	1.4234
Between-2SLS	8077	682	5927	$1.01{\cdot}10^5$
MLE	**0.7763**	2.7422	0.8285	2.9631
EC2SLS	0.9198	1.2848	0.9826	1.3446
EC2SLS-AR(1)	0.9051	1.4271	0.9577	1.4891
G2SLS	2255	2.2485	13.32	9.8148
2SLS-KR	0.8698	1.5996	1.0086	1.6790
FD-2SLS-KR	1.2468	$6.19{\cdot}10^8$	5413	355
FDGMM	0.9373	**<u>1.1415</u>**	0.8974	**<u>1.2668</u>**
GMM	0.8531	1.3981	0.9152	1.4539
Ind. OLS	$8.7{\cdot}10^8$	$2.19{\cdot}10^9$	$1.08{\cdot}10^6$	$9.93{\cdot}10^5$
Ind. 2SLS	$8.7{\cdot}10^8$	$2.19{\cdot}10^9$	$1.08{\cdot}10^6$	$9.93{\cdot}10^5$
Average OLS	0.8178	1.8017	0.8805	1.8523
Average 2SLS	0.8178	1.8017	0.8805	1.8523
Swamy	**<u>0.7763</u>**	1.3863	0.8369	1.4654
Bayes OLS	$9.45{\cdot}10^5$	$2.99{\cdot}10^6$	$1.11{\cdot}10^4$	$2.13{\cdot}10^4$
It. Bayes OLS	$1.89{\cdot}10^5$	$2.73{\cdot}10^5$	772.6	$1.93{\cdot}10^4$
Bayes 2SLS	$9.45{\cdot}10^5$	$2.99{\cdot}10^6$	$1.11{\cdot}10^4$	$2.13{\cdot}10^4$
It. Bayes 2SLS	$1.89{\cdot}10^5$	$2.73{\cdot}10^5$	772.6	$1.93{\cdot}10^4$
It. Bayes	4.2762	6.3261	121.25	10.4824

Note: See Table A1.

Case with $(N, T) = (5, 20)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table A5: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.4371	1.1825	0.4941	1.3392
Within	0.4391	4.3250	0.4620	5.3691
Between	0.6224	0.8967	0.7175	0.8768
FD-OLS	1.1717	1.7841	0.9949	2.9049
WLS	**0.4367**	1.1964	0.4973	1.3910
WLS-AR(1)	4.4359	1.5865	1.3986	1.9487
2SLS	0.4378	1.2467	0.4707	1.4184
FD-2SLS	0.7596	1.2902	**0.2911**	2.0465
Within-2SLS	0.4371	1.1825	0.4941	1.3392
Between-2SLS	0.4382	5.8797	0.4641	9.0893
MLE	0.4373	3.8077	0.4638	4.6782
EC2SLS	0.4846	0.9081	0.4967	0.9118
EC2SLS-AR(1)	0.4896	0.8751	0.4757	0.8408
G2SLS	0.4803	0.9060	0.4917	0.8998
2SLS-KR	0.4384	1.5293	0.4685	13.24
FD-2SLS-KR	0.7601	3.3442	0.2912	1.8097
FDGMM	0.7571	1.3837	0.3190	2.2953
GMM	0.4876	0.8725	0.5809	0.8524
Ind. OLS	0.4942	0.4755	0.4511	0.4462
Ind. 2SLS	0.5112	0.5850	0.4495	0.4564
Average OLS	0.4394	1.1143	0.4592	0.9810
Average 2SLS	0.4393	1.1168	0.4579	0.9781
Swamy	0.4420	0.9778	0.4902	1.0316
Bayes OLS	0.4592	0.4445	0.4362	0.3999
It. Bayes OLS	0.4425	0.4620	0.4324	0.3894
Bayes 2SLS	0.4614	0.4476	0.4346	0.4000
It. Bayes 2SLS	0.4428	0.4683	0.4310	0.3891
It. Bayes	0.4653	**0.4342**	0.4243	**0.3874**

Note: See Table A1.

Case with $(N, T) = (10, 20)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table A6: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.1
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.7303	1.1896	0.7570	1.2458
Within	0.7320	3.0581	0.7599	3.3109
Between	1.3466	1.1371	1.2462	1.1567
FD-OLS	1.1383	1.3696	1.1822	1.5360
WLS	0.7844	1.1983	0.7561	1.2633
WLS-AR(1)	1521	1.9325	1607	2.0025
2SLS	0.7370	1.3303	0.8430	1.3783
FD-2SLS	$6.62{\cdot}10^4$	$3.16{\cdot}10^5$	9.8064	1.2696
Within-2SLS	0.7303	1.1896	0.7570	1.2458
Between-2SLS	0.7704	80.43	3.1797	58.86
MLE	0.7309	2.6813	0.7602	2.8763
EC2SLS	0.8087	1.0491	0.8728	1.0617
EC2SLS-AR(1)	0.8045	1.0989	0.8807	1.1113
G2SLS	0.8114	1.0858	0.8866	1.0945
2SLS-KR	0.7216	5.1969	0.7658	2.3853
FD-2SLS-KR	0.9249	12.15	0.9681	1.9767
FDGMM	0.9113	1.1527	0.8431	1.2975
GMM	0.7761	1.0270	0.8414	1.0339
Ind. OLS	0.8184	0.7430	0.8849	0.7086
Ind. 2SLS	1.2993	0.9314	1.2752	0.9146
Average OLS	0.7317	1.1920	0.7599	1.2093
Average 2SLS	0.7319	1.1969	0.7597	1.2114
Swamy	0.7362	1.0976	0.7903	1.1241
Bayes OLS	0.7462	**<u>0.6358</u>**	0.7530	**<u>0.6294</u>**
It. Bayes OLS	0.7167	0.6400	0.7282	0.6304
Bayes 2SLS	0.7525	0.6509	0.7619	0.6418
It. Bayes 2SLS	**<u>0.7165</u>**	0.6434	**<u>0.7281</u>**	0.6370
It. Bayes	0.7872	0.6884	0.7857	0.6624

Note: See Table A1.

Case with $(N, T) = (10, 20)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table A7: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	**0.4244**	2.1395	0.4826	1.8539
Within	0.4255	4.2404	0.4470	3.5849
Between	0.5859	0.9130	0.6966	0.9504
FD-OLS	1.1638	2.9423	1.0028	1.8792
WLS	**0.4244**	2.0014	0.4834	1.7481
WLS-AR(1)	989.3	4.9790	4.0820	$3.52 \cdot 10^4$
2SLS	0.4246	2.0290	0.4546	1.8097
FD-2SLS	0.7351	2.4886	**0.2743**	1.4006
Within-2SLS	**0.4244**	2.1395	0.4826	1.8539
Between-2SLS	0.4246	5.0434	0.4476	4.1232
MLE	0.4245	3.7157	0.4493	3.1425
EC2SLS	0.4746	1.5100	0.4859	1.3616
EC2SLS-AR(1)	0.4789	0.9181	0.4549	0.9526
G2SLS	0.4702	1.3663	0.4794	1.3011
2SLS-KR	0.4248	1.4330	0.4477	1.3138
FD-2SLS-KR	0.7354	1.7268	**0.2743**	1.1068
FDGMM	0.7345	2.8496	0.2850	1.5887
GMM	0.4511	1.3367	0.5459	1.2441
Ind. OLS	0.5309	0.4250	0.4441	0.5107
Ind. 2SLS	0.5391	0.4423	0.5038	0.5004
Average OLS	0.4259	0.8928	0.4457	0.9080
Average 2SLS	0.4257	0.8899	0.4441	0.9089
Swamy	0.4272	1.8211	0.4757	1.6035
Bayes OLS	0.4491	0.3682	0.4239	0.4207
It. Bayes OLS	0.4264	0.3599	0.4195	**0.4191**
Bayes 2SLS	0.4507	0.3650	0.4223	0.4242
It. Bayes 2SLS	0.4263	**0.3566**	0.4183	0.4233
It. Bayes	0.4497	0.4155	0.4128	0.4319

Note: See Table A1.

Case with $(N, T) = (10, 50)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table A8: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.7268	1.6376	0.7503	1.7234
Within	0.7281	2.6729	0.7531	2.8324
Between	1.2853	1.1869	1.1978	1.1879
FD-OLS	1.1386	1.4775	1.1813	1.6377
WLS	0.7882	1.5747	0.7476	1.6514
WLS-AR(1)	634.9	22.53	42.92	143.0
2SLS	0.7287	1.7152	0.7552	1.7789
FD-2SLS	0.8978	1.1769	0.8189	1.3257
Within-2SLS	0.7268	1.6376	0.7503	1.7234
Between-2SLS	0.7337	3.2429	0.7586	3.5686
MLE	0.7268	2.3591	0.7531	2.4954
EC2SLS	0.7894	1.2493	0.8514	1.2835
EC2SLS-AR(1)	0.7920	1.1600	0.8625	1.1591
G2SLS	0.7866	1.3507	0.8565	1.3688
2SLS-KR	0.7132	2.4810	0.7443	1.6889
FD-2SLS-KR	0.8937	1.0412	0.8104	1.1162
FDGMM	0.9010	1.2519	0.8308	1.4243
GMM	0.7769	1.1934	0.8208	1.2101
Ind. OLS	0.8119	0.7238	0.8129	0.6933
Ind. 2SLS	4.7210	0.8435	1.1890	$1.58{\cdot}10^4$
Average OLS	0.7263	1.0552	0.7515	1.0557
Average 2SLS	0.7260	1.0583	0.7509	1.0573
Swamy	0.7302	1.3935	0.7823	1.4469
Bayes OLS	0.7401	**0.6392**	0.7464	0.6210
It. Bayes OLS	**0.7097**	0.6458	0.7194	**0.6143**
Bayes 2SLS	0.7448	0.6478	0.7490	0.6311
It. Bayes 2SLS	0.7094	0.6480	**0.7190**	0.6186
It. Bayes	0.7764	0.6664	0.7769	0.6549

Note: See Table A1.

Case with $(N, T) = (10, 50)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table A9: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.4254	1.8840	0.4836	2.1999
Within	0.4258	3.3288	0.4536	3.9276
Between	0.6586	0.9302	0.7311	0.8854
FD-OLS	1.1636	1.9070	1.0020	3.0153
WLS	**0.4243**	1.7296	0.3891	1.9994
WLS-AR(1)	150.9	2.4387	1.0760	2.5892
2SLS	0.4255	1.7241	0.4550	1.9468
FD-2SLS	0.7354	1.4340	**0.2744**	2.5780
Within-2SLS	0.4254	1.8840	0.4836	2.1999
Between-2SLS	0.4255	4.0620	0.4467	5.0968
MLE	0.4255	3.0942	0.4565	3.6575
EC2SLS	0.4790	1.1719	0.4988	1.2294
EC2SLS-AR(1)	0.4851	0.9244	0.4859	0.8767
G2SLS	0.4755	1.0475	0.4934	1.0290
2SLS-KR	0.4256	379.4	0.4473	521.9
FD-2SLS-KR	0.7357	1.0551	0.2744	1.6303
FDGMM	0.7351	1.6976	0.3244	3.1165
GMM	0.5147	0.9343	0.6213	0.9047
Ind. OLS	0.4458	0.3826	0.4518	0.3461
Ind. 2SLS	0.4475	0.3847	0.4464	0.3399
Average OLS	0.4257	0.8287	0.4511	0.8774
Average 2SLS	0.4256	0.8287	0.4458	0.8735
Swamy	0.4305	1.3987	0.5083	1.5394
Bayes OLS	0.4300	0.3789	0.4435	0.3412
It. Bayes OLS	0.4255	**0.3786**	0.4419	0.3394
Bayes 2SLS	0.4308	0.3803	0.4385	0.3340
It. Bayes 2SLS	0.4254	0.3800	0.4374	**0.3323**
It. Bayes	0.4420	0.3832	0.4423	0.3582

Note: See Table A1.

Case with $(N, T) = (20, 50)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table A10: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.7063	1.6391	0.7294	1.7353
Within	0.7065	2.4841	0.7294	2.6296
Between	1.1675	1.0852	1.1091	1.0755
FD-OLS	1.1320	1.4979	1.1719	1.6695
WLS	0.7388	1.5438	0.7203	1.6261
WLS-AR(1)	$2.6{\cdot}10^5$	8654	4701	30.11
2SLS	0.7055	1.5392	0.7293	1.5984
FD-2SLS	0.8937	1.1953	0.8135	1.3611
Within-2SLS	0.7063	1.6391	0.7294	1.7353
Between-2SLS	0.7061	3.1520	0.7294	3.5700
MLE	0.7070	2.3199	0.7301	2.4603
EC2SLS	0.7673	1.1311	0.8355	1.1398
EC2SLS-AR(1)	0.7710	1.0745	0.8452	1.0637
G2SLS	0.7470	1.1200	0.8171	1.1103
2SLS-KR	0.6981	67.90	0.7249	3.3737
FD-2SLS-KR	0.8917	1.0049	0.8096	1.0734
FDGMM	0.8955	1.3073	0.8196	1.5076
GMM	0.7838	1.0254	0.8488	1.0140
Ind. OLS	0.7271	0.5927	0.7446	0.5899
Ind. 2SLS	0.7296	0.5964	0.7460	0.5916
Average OLS	0.7051	0.9147	0.7287	0.9284
Average 2SLS	0.7050	0.9159	0.7282	0.9284
Swamy	0.7152	1.2574	0.7753	1.2898
Bayes OLS	0.7063	**0.5806**	0.7233	**0.5779**
It. Bayes OLS	0.6972	0.5939	0.7134	0.5785
Bayes 2SLS	0.7070	0.5825	0.7234	0.5787
It. Bayes 2SLS	**0.6970**	0.5948	**0.7132**	0.5790
It. Bayes	0.7257	0.5894	0.7419	0.5910

Note: See Table A1.

Case with $(N, T) = (20, 50)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table A11: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.9 $(0.9, 0.9)$
OLS	0.4245	1.9199	0.4880	2.0656
Within	0.4248	2.0163	0.4641	2.2689
Between	0.7105	0.8702	0.7238	0.8056
FD-OLS	1.1681	1.6591	0.9974	2.3449
WLS	0.4217	1.6533	0.4954	1.6793
WLS-AR(1)	110.4	$2.40{\cdot}10^{7}$	1.0054	2.7350
2SLS	0.4246	1.6139	0.4619	1.7449
FD-2SLS	0.7510	1.2676	**<u>0.2954</u>**	1.8359
Within-2SLS	0.4245	1.9199	0.4880	2.0656
Between-2SLS	0.4247	3.1573	0.4531	5.0014
MLE	0.4246	1.8698	0.4671	2.063
EC2SLS	0.4792	0.9079	0.5039	0.8413
EC2SLS-AR(1)	0.4880	0.8658	0.5003	0.7992
G2SLS	0.4754	0.8764	0.4999	0.8110
2SLS-KR	0.4259	2.6086	0.5048	13.21
FD-2SLS-KR	0.7512	107.6	0.2955	1.9853
FDGMM	$\sim 10^{26}$	$\sim 10^{22}$	$\sim 10^{23}$	$\sim 10^{27}$
GMM	$\sim 10^{26}$	$\sim 10^{22}$	$\sim 10^{23}$	$\sim 10^{27}$
Ind. OLS	0.4313	0.3982	0.4627	0.3932
Ind. 2SLS	0.4323	0.3993	0.4551	0.3848
Average OLS	0.4247	1.1230	0.4612	1.1773
Average 2SLS	0.4246	1.1233	0.4563	1.1697
Swamy	0.4351	1.0326	0.5276	0.9989
Bayes OLS	**<u>0.4225</u>**	0.3957	0.4566	0.3902
It. Bayes OLS	**<u>0.4225</u>**	**<u>0.3952</u>**	0.4562	0.3891
Bayes 2SLS	0.4230	0.3965	0.4495	0.3826
It. Bayes 2SLS	0.4227	0.3961	0.4495	**<u>0.3817</u>**
It. Bayes	0.4308	0.3972	0.4601	0.3888

Note: See Table A1.

Case with $(N, T) = (50, 50)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table A12: Theil's U for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.1
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.6998	1.5362	0.7258	1.5363
Within	0.6999	1.5150	0.7257	1.5482
Between	1.1335	1.0424	1.0752	1.0170
FD-OLS	1.1318	1.3082	1.1749	1.4189
WLS	0.7123	1.4013	0.7172	1.3828
WLS-AR(1)	1.3845	197.2	2.4403	1.9808
2SLS	0.6992	1.3826	0.7261	1.3820
FD-2SLS	0.9041	1.1001	0.8268	1.1870
Within-2SLS	0.6998	1.5362	0.7258	1.5363
Between-2SLS	0.6996	7.0826	0.7261	4.8631
MLE	0.7005	1.4245	0.7261	1.4478
EC2SLS	0.7659	1.0227	0.8357	0.9964
EC2SLS-AR(1)	0.7712	1.0375	0.8439	1.0115
G2SLS	0.7461	1.0352	0.8216	1.0082
2SLS-KR	0.6990	2.1240	0.7623	3.4547
FD-2SLS-KR	0.9037	1.0751	0.8248	$2.39 \cdot 10^4$
FDGMM	$\sim 10^{52}$	$\sim 10^{32}$	$\sim 10^{39}$	$\sim 10^{26}$
GMM	$\sim 10^{52}$	$\sim 10^{32}$	$\sim 10^{39}$	$\sim 10^{26}$
Ind. OLS	0.7097	0.5929	0.7363	0.5997
Ind. 2SLS	0.7112	0.5938	0.7364	0.6000
Average OLS	0.6990	1.0232	0.7256	1.0235
Average 2SLS	0.6989	1.0238	0.7251	1.0230
Swamy	0.7132	1.0531	0.7754	1.0285
Bayes OLS	0.6971	**0.5870**	0.7218	0.5938
It. Bayes OLS	0.6932	0.5958	0.7159	0.5953
Bayes 2SLS	0.6976	0.5876	0.7215	**0.5937**
It. Bayes 2SLS	**0.6931**	0.5965	**0.7153**	0.5948
It. Bayes	0.7096	0.5924	0.7357	0.5990

Note: See Table A1.

Case with $(N, T) = (50, 50)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table B1: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.9 $(0.9, 0.9)$
OLS	1.4355	0.0645	1.7299	0.0924
Within	1.3383	1.7426	1.7571	1.8785
Between	0.4766	-0.0296	0.3241	0.0344
FD-OLS	-1.4332	-0.2197	-0.2303	-0.3239
WLS	1.4598	0.0122	1.8314	0.0588
WLS-AR(1)	1.1639	0.4089	1.6345	0.4890
2SLS	1.4370	-0.0878	1.7132	-0.0319
FD-2SLS	1.6350	0.9092	2.8151	0.9843
Within-2SLS	1.4355	0.0645	1.7299	0.0924
Between-2SLS	1.3830	1.6642	1.7080	1.8125
MLE	1.4164	0.7241	1.7910	0.8037
EC2SLS	1.4554	0.6877	1.930	0.7908
EC2SLS-AR(1)	1.3512	0.8957	2.0017	1.0558
G2SLS	1.5542	0.4710	2.0091	0.5720
2SLS-KR	1.4423	-0.0297	1.7377	0.0059
FD-2SLS-KR	1.6559	0.8813	**2.8158**	0.9628
FDGMM	**1.6805**	0.9222	2.8000	1.0017
GMM	1.0586	0.0040	0.9832	0.1028
Ind. OLS	1.3461	2.2208	2.0630	2.7630
Ind. 2SLS	1.3461	2.2208	2.0630	2.7630
Average OLS	1.3451	1.1251	1.7940	1.2776
Average 2SLS	1.3451	1.1251	1.7940	1.2776
Swamy	1.3753	0.5429	1.8187	0.6340
Bayes OLS	1.4275	2.2305	2.0900	2.7680
It. Bayes OLS	1.4772	2.2421	2.1101	**2.7696**
Bayes 2SLS	1.4275	2.2305	2.0900	2.7680
It. Bayes 2SLS	1.4772	2.2421	2.1101	**2.7696**
It. Bayes	1.5241	**2.3312**	2.0479	2.5816

Note: See Table A1.

Case with $(N, T) = (5, 10)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table B2: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.9 $(0.9, 0.9)$
OLS	0.6364	0.2003	0.6778	0.2135
Within	0.5748	1.4618	0.6291	1.5026
Between	0.2664	-0.0274	0.1836	0.0018
FD-OLS	-0.6458	-0.1276	-0.7144	-0.1577
WLS	0.6587	0.1562	0.7004	0.1716
WLS-AR(1)	0.6859	0.4130	0.7839	0.4518
2SLS	0.6364	-0.0029	0.6458	0.0375
FD-2SLS	0.8889	0.7844	1.1166	0.8051
Within-2SLS	0.6364	0.2003	0.6778	0.2135
Between-2SLS	0.5728	1.3804	0.6010	1.4015
MLE	0.6237	0.7135	0.6001	0.7388
EC2SLS	0.6215	0.5425	0.5661	0.5598
EC2SLS-AR(1)	0.5800	0.5694	0.5378	0.6300
G2SLS	0.6655	0.3725	0.6125	0.4280
2SLS-KR	0.6241	-0.0232	0.6444	0.0161
FD-2SLS-KR	0.9087	0.7561	1.1789	0.7859
FDGMM	**0.9682**	0.8262	**1.1872**	0.8540
GMM	0.5548	0.0517	0.5145	0.0977
Ind. OLS	0.5177	1.5606	0.6647	1.7540
Ind. 2SLS	0.5177	1.5606	0.6647	1.7540
Average OLS	0.5647	0.8173	0.6305	0.8206
Average 2SLS	0.5647	0.8173	0.6305	0.8206
Swamy	0.6139	0.4696	0.5743	0.4861
Bayes OLS	0.5809	1.6031	0.7379	1.7945
It. Bayes OLS	0.6427	1.6365	0.7912	1.8277
Bayes 2SLS	0.5809	1.6031	0.7379	1.7945
It. Bayes 2SLS	0.6427	1.6365	0.7912	1.8277
It. Bayes	0.6007	**1.6961**	0.6362	**1.8367**

Note: See Table A1.

Case with $(N, T) = (5, 10)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table B3: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.1 $(0.9, 0.9)$
OLS	1.8379	0.1390	2.0313	0.3045
Within	1.7795	0.6395	2.2387	0.6326
Between	0.6912	-0.0419	0.5918	0.1268
FD-OLS	-1.4255	-0.3255	0.3876	-0.6475
WLS	**1.8422**	0.1090	1.9424	0.2631
WLS-AR(1)	1.3516	0.0446	1.8901	0.1894
2SLS	1.8318	0.0175	2.1414	0.1976
FD-2SLS	1.6473	0.9111	2.8150	1.1917
Within-2SLS	1.8379	0.1390	2.0313	0.3045
Between-2SLS	1.7796	1.1574	2.2107	1.2569
MLE	1.8330	0.0632	2.2391	0.0742
EC2SLS	1.5989	0.2577	2.1875	0.4402
EC2SLS-AR(1)	1.5476	0.2412	2.3231	0.4618
G2SLS	1.6400	0.1067	2.2368	0.2670
2SLS-KR	1.8354	-0.0179	2.2360	0.1230
FD-2SLS-KR	1.6478	0.9001	**2.8153**	1.1818
FDGMM	1.6573	0.9092	2.8062	1.1810
GMM	1.5140	0.0204	1.4405	0.1864
Ind. OLS	1.3616	1.7207	2.1708	2.5424
Ind. 2SLS	1.3616	1.7207	2.1708	2.5424
Average OLS	1.7091	0.5378	2.2450	0.7745
Average 2SLS	1.7091	0.5378	2.2450	0.7745
Swamy	1.8025	0.2989	2.1858	0.4699
Bayes OLS	1.5131	1.7570	2.2142	2.5530
It. Bayes OLS	1.6368	1.7661	2.2467	**2.5547**
Bayes 2SLS	1.5131	1.7570	2.2142	2.5530
It. Bayes 2SLS	1.6368	1.7661	2.2467	**2.5547**
It. Bayes	1.7747	**1.8769**	2.2987	2.2135

Note: See Table A1.

Case with $(N, T) = (5, 20)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table B4: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ, H)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.7305	0.1428	0.8078	0.2138
Within	0.6669	0.5451	0.7322	0.5502
Between	0.2867	-0.0466	0.2201	-0.0109
FD-OLS	-0.6568	-0.1692	-0.7012	-0.2272
WLS	0.7200	0.1040	0.8158	0.1615
WLS-AR(1)	0.6944	0.0590	0.8183	0.1119
2SLS	0.7015	-0.0141	0.7785	0.0559
FD-2SLS	**0.9331**	0.7186	1.1725	0.7770
Within-2SLS	0.7305	0.1428	0.8078	0.2138
Between-2SLS	0.6257	0.8203	0.6969	0.8228
MLE	0.7333	0.0476	0.7424	0.0599
EC2SLS	0.6657	0.1556	0.6232	0.2210
EC2SLS-AR(1)	0.6292	0.0518	0.5408	0.1142
G2SLS	0.7259	0.0444	0.6849	0.1095
2SLS-KR	0.7359	-0.0571	0.7892	-0.0090
FD-2SLS-KR	0.9520	0.7065	**1.2184**	0.7634
FDGMM	0.9579	0.7410	1.1756	0.7915
GMM	0.6486	0.0288	0.6101	0.1156
Ind. OLS	0.4899	1.1055	0.6315	1.2785
Ind. 2SLS	0.4899	1.1055	0.6315	1.2785
Average OLS	0.6469	0.3977	0.7358	0.4083
Average 2SLS	0.6469	0.3977	0.7358	0.4083
Swamy	0.7311	0.1962	0.6545	0.2658
Bayes OLS	0.6051	1.1609	0.7470	1.3450
It. Bayes OLS	0.7217	1.2056	0.8544	**1.3933**
Bayes 2SLS	0.6015	1.1609	0.7470	1.3450
It. Bayes 2SLS	0.7217	1.2056	0.8544	**1.3933**
It. Bayes	0.5994	**1.2584**	0.6646	1.3594

Note: See Table A1.

Case with $(N, T) = (5, 20)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table B5: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	1.8630	0.1727	2.0518	0.3147
Within	**1.8706**	0.3962	2.2569	0.3930
Between	0.3272	0.0098	0.2979	0.1148
FD-OLS	-1.4284	-0.3317	0.5497	-0.6491
WLS	1.8405	0.1650	1.9350	0.2950
WLS-AR(1)	1.3114	0.1069	1.8404	0.2205
2SLS	1.8616	0.0709	2.1861	0.2180
FD-2SLS	1.6535	0.9322	2.8143	1.1993
Within-2SLS	1.8630	0.1727	2.0518	0.3147
Between-2SLS	1.8468	1.1768	2.2730	1.1951
MLE	1.8585	-0.0179	2.2489	-0.0125
EC2SLS	1.5363	0.0784	2.0618	0.2092
EC2SLS-AR(1)	1.4621	-0.0037	2.2652	0.0948
G2SLS	1.6050	0.0234	2.1245	0.1569
2SLS-KR	1.8548	-0.0107	2.2463	0.1146
FD-2SLS-KR	1.6540	0.9248	**2.8141**	1.2001
FDGMM	1.6624	0.9297	2.7523	1.2135
GMM	1.0939	0.0224	0.9380	0.1200
Ind. OLS	1.7151	1.9827	2.2108	2.6123
Ind. 2SLS	1.6920	1.9738	2.2228	2.6172
Average OLS	1.8543	0.7812	2.2762	0.9186
Average 2SLS	1.8532	0.7685	2.2883	0.9509
Swamy	1.7812	0.1098	2.0684	0.2498
Bayes OLS	1.8024	1.9983	2.2359	2.6095
It. Bayes OLS	1.8304	1.9787	2.2438	2.6063
Bayes 2SLS	1.7819	1.9943	2.2490	**2.6198**
It. Bayes 2SLS	1.8240	1.9728	2.2553	2.6183
It. Bayes	1.7766	**2.0116**	2.2921	2.4676

Note: See Table A1.

Case with $(N, T) = (10, 20)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table B6: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	0.8309	0.1851	0.9016	0.2449
Within	0.7821	0.3734	0.8547	0.3590
Between	0.1976	-0.0300	0.1412	0.0469
FD-OLS	-0.7174	-0.1765	-0.8968	-0.2416
WLS	0.8336	0.1707	0.9181	0.2302
WLS-AR(1)	0.7729	0.1298	0.8930	0.1562
2SLS	0.8144	0.0562	0.8759	0.0965
FD-2SLS	0.9551	0.7415	1.2013	0.7940
Within-2SLS	0.8309	0.1851	0.9016	0.2449
Between-2SLS	0.7672	0.8761	0.8367	0.8505
MLE	0.8263	-0.0315	0.8684	-0.0279
EC2SLS	0.6261	0.0533	0.5150	0.1071
EC2SLS-AR(1)	0.5749	-0.0141	0.4347	0.0444
G2SLS	0.6972	0.0164	0.5838	0.0747
2SLS-KR	0.8548	-0.0234	0.8902	0.0021
FD-2SLS-KR	0.9680	0.7341	**1.2275**	0.7893
FDGMM	**0.9797**	0.7398	1.2111	0.7905
GMM	0.5026	0.0089	0.4548	0.0542
Ind. OLS	0.7484	1.3693	0.8809	1.5320
Ind. 2SLS	0.7447	1.3608	0.8754	1.5209
Average OLS	0.7717	0.4915	0.8686	0.5129
Average 2SLS	0.7796	0.4802	0.8554	0.5139
Swamy	0.7567	0.0559	0.6474	0.1310
Bayes OLS	0.8669	1.4253	1.0013	**1.5962**
It. Bayes OLS	0.8872	1.4405	1.0679	1.5791
Bayes 2SLS	0.8541	1.4186	0.9995	1.5823
It. Bayes 2SLS	0.8863	**1.4408**	1.0608	1.5772
It. Bayes	0.7560	1.3993	0.8585	1.5397

Note: See Table A1.

Case with $(N, T) = (10, 20)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table B7: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	2.0374	1.0464	2.1999	0.8491
Within	**2.0593**	1.6050	2.4656	1.5140
Between	0.5935	0.2201	0.6017	0.1390
FD-OLS	-1.4921	-0.8402	-0.4118	-0.5898
WLS	2.0283	1.0671	2.0603	0.8654
WLS-AR(1)	1.5590	0.9751	2.0500	0.7859
2SLS	2.0327	0.5718	2.3972	0.4188
FD-2SLS	1.7252	1.6821	**2.8517**	1.3815
Within-2SLS	2.0374	1.0464	2.1999	0.8491
Between-2SLS	2.0484	1.5748	2.5131	1.4589
MLE	2.0326	0.2532	2.4374	0.2283
EC2SLS	1.6761	0.8254	2.1952	0.6479
EC2SLS-AR(1)	1.6327	0.3194	2.4192	0.2177
G2SLS	1.7265	0.4019	2.2514	0.2965
2SLS-KR	2.0349	0.2974	2.4985	0.1651
FD-2SLS-KR	1.7256	1.8359	2.8515	1.4731
FDGMM	1.7285	1.5602	2.8287	1.2943
GMM	1.8480	0.6130	1.4864	0.4587
Ind. OLS	1.8094	2.7020	2.2924	2.2132
Ind. 2SLS	1.7967	2.7016	2.3023	2.2098
Average OLS	2.0487	1.6325	2.4719	1.3652
Average 2SLS	2.0511	1.6486	2.4851	1.3619
Swamy	1.9982	0.8855	2.2469	0.7019
Bayes OLS	1.9099	2.7002	2.3206	**2.2358**
It. Bayes OLS	2.0343	2.7008	2.3338	2.2180
Bayes 2SLS	1.9058	2.7032	2.3307	2.2295
It. Bayes 2SLS	2.0314	**2.7046**	2.3403	2.2119
It. Bayes	1.9211	2.5238	2.3819	2.1644

Note: See Table A1.

Case with $(N, T) = (10, 50)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table B8: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.9 $(0.9, 0.9)$
OLS	1.0156	0.6478	1.0903	0.6974
Within	1.0066	1.3108	1.0803	1.3212
Between	0.2502	0.0502	0.2451	0.0981
FD-OLS	-0.7469	-0.3390	-0.9209	-0.4161
WLS	0.9698	0.6553	1.1227	0.7069
WLS-AR(1)	0.8635	0.5982	0.9852	0.6480
2SLS	1.0121	0.3127	1.0836	0.3771
FD-2SLS	1.0260	1.0510	1.2789	1.1383
Within-2SLS	1.0156	0.6478	1.0903	0.6974
Between-2SLS	0.9855	1.1892	1.0787	1.1976
MLE	1.0039	0.1755	1.0733	0.1773
EC2SLS	0.7684	0.4277	0.6985	0.4936
EC2SLS-AR(1)	0.7165	0.0689	0.6105	0.1458
G2SLS	0.8618	0.1562	0.7396	0.2315
2SLS-KR	1.0630	0.0921	1.1341	0.1621
FD-2SLS-KR	1.0342	1.0943	**1.2919**	1.1983
FDGMM	1.0244	1.0305	1.2576	1.1132
GMM	0.7513	0.3374	0.6728	0.4073
Ind. OLS	0.8199	1.4738	0.9465	1.6247
Ind. 2SLS	0.8011	1.4685	0.9364	1.6168
Average OLS	1.0218	1.0954	1.0929	1.1112
Average 2SLS	1.0158	1.0848	1.0884	1.1131
Swamy	0.9573	0.4685	0.8239	0.5348
Bayes OLS	0.9697	**1.5318**	1.1047	**1.6810**
It. Bayes OLS	**1.0756**	1.4926	1.2078	1.6370
Bayes 2SLS	0.9607	1.5280	1.0961	1.6780
It. Bayes 2SLS	1.0730	1.4885	1.2143	1.6342
It. Bayes	0.8429	1.5019	0.9363	1.6293

Note: See Table A1.

Case with $(N, T) = (10, 50)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table B9: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.9 $(0.9, 0.9)$
OLS	2.0058	0.8615	2.1871	1.0480
Within	**2.0342**	1.5275	2.4010	1.5870
Between	0.3259	0.0185	0.2992	0.1822
FD-OLS	-1.4853	-0.5672	-0.3992	-0.8096
WLS	1.9981	0.8859	2.0842	1.0881
WLS-AR(1)	1.5041	0.7930	2.0696	0.9809
2SLS	2.0082	0.4496	2.3980	0.5890
FD-2SLS	1.7208	1.3695	**2.8517**	1.6742
Within-2SLS	2.0058	0.8615	2.1871	1.0480
Between-2SLS	2.0302	1.3917	2.4994	1.4642
MLE	2.0057	0.6897	2.3721	0.7789
EC2SLS	1.6425	0.5406	2.0655	0.6824
EC2SLS-AR(1)	1.5690	0.0451	2.2310	0.1900
G2SLS	1.6960	0.2200	2.1261	0.3164
2SLS-KR	2.0142	0.3207	2.5027	0.4218
FD-2SLS-KR	1.7206	1.4763	2.8513	1.8432
FDGMM	1.7234	1.3744	2.7458	1.7006
GMM	1.1199	0.2881	0.9265	0.3815
Ind. OLS	1.8963	2.3158	2.2643	2.6862
Ind. 2SLS	1.8919	2.3086	2.3042	2.6993
Average OLS	2.0290	1.4163	2.4144	1.6181
Average 2SLS	2.0258	1.4122	2.4560	1.6440
Swamy	1.9401	0.5998	2.0246	0.7575
Bayes OLS	1.9635	**2.3199**	2.2897	2.6825
It. Bayes OLS	2.0191	2.3186	2.3016	2.6837
Bayes 2SLS	1.9599	2.3154	2.3248	2.6995
It. Bayes 2SLS	2.0173	2.3111	2.3377	**2.7002**
It. Bayes	1.9175	2.3034	2.3154	2.6622

Note: See Table A1.

Case with $(N, T) = (20, 50)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table B10: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	1.0454	0.6634	1.1349	0.7081
Within	1.0555	1.3531	1.1444	1.3336
Between	0.1784	0.0151	0.1359	0.0553
FD-OLS	-0.7720	-0.3308	-0.9804	-0.4015
WLS	1.0263	0.6792	1.1523	0.7206
WLS-AR(1)	0.9238	0.6165	0.9563	0.6396
2SLS	1.0563	0.3131	1.1474	0.3694
FD-2SLS	1.0460	1.0559	1.2924	1.1397
Within-2SLS	1.0454	0.6634	1.1349	0.7081
Between-2SLS	1.0478	1.1499	1.1469	1.1171
MLE	1.0383	0.5846	1.1377	0.6217
EC2SLS	0.7133	0.3675	0.5736	0.4030
EC2SLS-AR(1)	0.6927	0.0006	0.5071	0.0618
G2SLS	0.8091	0.1518	0.6680	0.1548
2SLS-KR	1.0785	0.1909	1.1701	0.2052
FD-2SLS-KR	1.0488	1.1066	**1.2978**	1.2116
FDGMM	1.0427	1.0687	1.2751	1.1640
GMM	0.4374	0.1703	0.4370	0.2519
Ind. OLS	0.9330	1.6131	1.0326	1.7262
Ind. 2SLS	0.9199	1.6048	1.0251	1.7201
Average OLS	1.0425	1.0973	1.1505	1.1220
Average 2SLS	1.0502	1.0965	1.1501	1.1230
Swamy	0.9406	0.4194	0.7822	0.4717
Bayes OLS	1.0298	**1.6308**	1.1213	**1.7403**
It. Bayes OLS	**1.0797**	1.5466	1.1719	1.6778
Bayes 2SLS	1.0281	1.6223	1.1201	1.7379
It. Bayes 2SLS	1.0791	1.5424	1.1829	1.6742
It. Bayes	0.9323	1.6116	1.0261	1.7281

Note: See Table A1.

Case with $(N, T) = (20, 50)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Table B11: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H	0.1	0.9	0.1	0.9
(ρ, ϑ)	$(0.0, 0.0)$	$(0.0, 0.0)$	$(0.9, 0.9)$	$(0.9, 0.9)$
OLS	2.0055	0.4365	2.1602	0.5572
Within	2.0126	1.2593	2.3319	1.3702
Between	0.1451	0.0139	0.1338	0.1533
FD-OLS	-1.4435	-0.4500	0.6097	-0.6308
WLS	1.9728	0.4273	2.0485	0.5336
WLS-AR(1)	1.3978	0.2559	1.9136	0.4426
2SLS	1.9963	0.1289	2.3305	0.2574
FD-2SLS	1.6773	0.9795	**2.8223**	1.1234
Within-2SLS	2.0055	0.4365	2.1602	0.5572
Between-2SLS	2.0165	1.1697	2.4166	1.2375
MLE	2.0046	0.0882	2.3061	0.2386
EC2SLS	1.6215	0.0930	1.9550	0.2688
EC2SLS-AR(1)	1.5372	0.0206	2.0518	0.1093
G2SLS	1.6848	0.0420	2.0303	0.1696
2SLS-KR	1.9914	0.1061	2.2242	0.2415
FD-2SLS-KR	1.6772	0.9966	2.8220	1.1545
FDGMM	0.7081	0.3876	1.8207	0.4991
GMM	0.6632	0.2940	0.9659	0.3501
Ind. OLS	1.9287	2.1208	2.2467	2.5350
Ind. 2SLS	1.9219	2.1202	2.2946	**2.5513**
Average OLS	**2.0217**	1.1502	2.3458	1.2074
Average 2SLS	2.0171	1.1483	2.4043	1.2260
Swamy	1.9101	0.1274	1.8821	0.2768
Bayes OLS	1.9413	2.1386	2.2503	2.5351
It. Bayes OLS	1.9996	**2.1461**	2.2513	2.5337
Bayes 2SLS	1.9371	2.1363	2.2992	2.5487
It. Bayes 2SLS	1.9941	2.1393	2.3046	2.5509
It. Bayes	1.9330	2.1233	2.2530	2.5366

Note: See Table A1.

Case with $(N, T) = (50, 50)$ and "mild" cross dependence (the support of ζ_i is [0,0.2]).

Table B12: Pesaran and Timmermann's (1992) statistic for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

H (ρ, ϑ)	0.1 $(0.0, 0.0)$	0.9 $(0.0, 0.0)$	0.1 $(0.9, 0.9)$	0.9 $(0.9, 0.9)$
OLS	1.0218	0.3821	1.1148	0.4291
Within	1.0054	1.0604	1.1083	1.0262
Between	0.0667	0.0184	0.0690	0.0571
FD-OLS	-0.7214	-0.2894	-0.9515	-0.3453
WLS	1.0115	0.3819	1.1521	0.4226
WLS-AR(1)	0.8542	0.2643	0.8933	0.3381
2SLS	1.0319	0.1428	1.1320	0.1750
FD-2SLS	0.9907	0.7747	1.2404	0.8331
Within-2SLS	1.0218	0.3821	1.1148	0.4291
Between-2SLS	1.0097	0.9699	1.1268	0.9535
MLE	1.0120	0.0868	1.1163	0.1564
EC2SLS	0.5496	0.0844	0.3701	0.1063
EC2SLS-AR(1)	0.5173	0.0106	0.3396	0.0873
G2SLS	0.7117	0.0354	0.4448	0.0658
2SLS-KR	1.0167	0.0806	1.0540	0.1750
FD-2SLS-KR	0.9915	0.7887	**1.2413**	0.8423
FDGMM	0.1613	0.2013	0.2281	0.2044
GMM	0.1812	0.2410	0.2278	0.2847
Ind. OLS	0.9368	1.4924	1.0647	1.6103
Ind. 2SLS	0.9291	1.4841	1.0604	1.6060
Average OLS	1.0106	0.9346	1.1125	0.9422
Average 2SLS	1.0137	0.9332	1.1154	0.9513
Swamy	0.8910	0.1240	0.6584	0.1778
Bayes OLS	1.0274	**1.5384**	1.1551	**1.6500**
It. Bayes OLS	1.0512	1.5185	1.1560	1.6281
Bayes 2SLS	1.0176	1.5304	1.1430	1.6494
It. Bayes 2SLS	**1.0529**	1.5119	1.1665	1.6380
It. Bayes	0.9359	1.4919	1.0647	1.6083

Note: See Table A1.

Case with $(N, T) = (50, 50)$ and "large" cross dependence (the support of ζ_i is [-1,3]).

Tables

Table C1: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(5,10)	(5,10)	(5,10)	(5,10)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	OLS **0.4276**	FDGMM 1.4166	FD-2SLS **0.3257**	GMM 1.4278
	Heterogeneous	Swamy 0.4319	Average OLS 1.7823	Average OLS 0.4520	Average OLS 1.5321
	Shrinkage	It. Bayes 0.4853	It. Bayes **0.5307**	It. Bayes 0.3895	It. Bayes **0.4832**
DM	DM1	OLS vs Swamy (0.3190)	It. Bayes vs FDGMM (-2.1966)(**)	FD-2SLS vs It. Bayes (2.0660)(**)	It. Bayes GMM (-2.3024)(**)
	DM2	It. Bayes vs. OLS (0.8880)			
PT	Homogeneous	FDGMM **1.6805(*)**	Within 1.7426(*)	FD-2SLS-KR **2.8158(**)**	Within 1.8785(*)
	Heterogeneous	Average OLS 1.3451	Ind. OLS 2.2208(**)	Ind. OLS 2.0630(**)	Ind. OLS 2.7630(**)
	Shrinkage	Bayes OLS 1.4275	It. Bayes **2.3312(**)**	I.B. OLS 2.1101(**)	I.B. OLS **2.7696(**)**

Note: This Table reports results for the case $(N,T) = (5,10)$ under "mild" cross dependence (the support of ζ_i is [0,0.2]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0,0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C2: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(5,10)	(5,10)	(5,10)	(5,10)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	MLE **0.7618**	FDGMM **1.1896**	Within-2SLS **0.8000**	FDGMM **1.2822**
	Heterogeneous	Swamy 0.7646	Average OLS 2.2297	Swamy 0.8303	Average OLS 1.6972
	Shrinkage	It. Bayes 134.84	It. Bayes 61.45	It. Bayes 53.46	It. Bayes 13.27
DM	DM1	Swamy vs. MLE (0.1387)	Average OLS vs. FDGMM (4.1188)(**)	Swamy vs. Within-2SLS (-0.0170)	Average OLS vs. FDGMM (4.1035)(**)
	DM2	It. Bayes vs. MLE (0.8286)	It. Bayes vs. FDGMM (-1.2647)	It. Bayes vs. Within-2SLS (0.0293)	It. Bayes vs. FDGMM (-1.0731)
PT	Homogeneous	FDGMM 0.9682	Within 1.4618	FDGMM 1.1872	Within 1.5026
	Heterogeneous	Ind. OLS 0.5177	Ind. OLS 1.5606	Ind. OLS 0.6647	Ind. OLS **1.7540(*)**
	Shrinkage	Bayes OLS 0.5809	It. Bayes **1.6961(*)**	I. B. OLS 0.7912	It. Bayes **1.8367(*)**

Note: This Table reports results for the case $(N, T) = (5, 10)$ under "large" cross dependence (the support of ζ_i is [-1,3]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C3: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(5,20)	(5,20)	(5,20)	(5,20)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	OLS **0.4405**	Between 1.2063	FD-2SLS **0.2923**	Between 1.3196
	Heterogeneous	Swamy 0.4427	Swamy 1.4297	Average OLS 0.4628	Average OLS 1.2699
	Shrinkage	It. Bayes 0.4752	It. Bayes **0.4988**	It. Bayes 0.4037	It. Bayes **0.4899**
DM	DM1	Swamy vs. OLS (0.2034)	Between vs. It. Bayes (-2.2382)(**)	It. Bayes vs. FD-2SLS (2.4040)(*)	Average OLS vs. It. Bayes (1.1141)
	DM2	It. Bayes vs. OLS (0.7101)			
PT	Homogeneous	WLS **1.8422(*)**	Between-2SLS 1.1574	FD-2SLS-KR **2.8153(**)**	Between-2SLS 1.2569
	Heterogeneous	Swamy **1.8025(*)**	Ind. OLS **1.7207(*)**	Average OLS **2.2450(**)**	Ind. OLS **2.5424(**)**
	Shrinkage	It. Bayes **1.7747(*)**	It. Bayes **1.8769(*)**	It. Bayes **2.2987(**)**	I. B. OLS **2.5547(**)**

Note: This Table reports results for the case $(N, T) = (5, 20)$ under "mild" cross dependence (the support of ζ_i is [0,0.2]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C4: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(5,20)	(5,20)	(5,20)	(5,20)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	MLE **0,7763**	FDGMM **1.1415**	OLS **0.8136**	FDGMM **1.2668**
	Heterogeneous	Swamy **0.7763**	Swamy 1.3863	Swamy 0.8369	Swamy 1.4654
	Shrinkage	It. Bayes 4.2762	It. Bayes 6.3261	It. Bayes 121.25	It. Bayes 10.4824
DM	DM1	Swamy vs. MLE (0.1870)	Swamy vs. FDGMM (1.1881)	Swamy vs. OLS (0.5656)	Swamy vs. FDGMM (1.3239)
	DM2	It. Bayes vs. MLE (0.7609)	It. Bayes vs. FDGMM (-0.9985)	It. Bayes vs. OLS (0.2129)	It. Bayes vs. FDGMM (-0.7957)
PT	Homogeneous	FD-2SLS 0.9331	Between -2SLS 0.8203	FD-2SLS-KR 1.2184	Between-2SLS 0.8228
	Heterogeneous	Swamy 0.7311	Average OLS 0.3977	Average OLS 0.7358	Ind. OLS 1.2785
	Shrinkage	I. B. OLS 0.7217	It. Bayes 1.2584	I. B. OLS 0.8544	I. B. OLS 1.3933

Note: This Table reports results for the case $(N, T) = (5, 20)$ under "large" cross dependence (the support of ζ_i is [-1,3]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C5: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(10,20)	(10,20)	(10,20)	(10,20)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	WLS **0.4367**	GMM 0.8725	FD-2SLS **0.2911**	EC2SLS-AR(1) 0.8408
	Heterogeneous	Average 2SLS 0.4393	Ind. OLS 0.4755	Ind. 2SLS 0.4495	Ind. OLS 0.4462
	Shrinkage	I. B. OLS 0.4592	It. Bayes **0.4342**	It. Bayes 0.4243	It. Bayes **0.3874**
DM	DM1	Average 2SLS vs. WLS (0.1913)	Ind. OLS vs. It. Bayes (0.3638)	It. Bayes vs. FD-2SLS (2.5319)(**)	Ind. OLS vs. It. Bayes (0.5813)
	DM2	I. B. OLS vs. WLS (0.1424)			
PT	Homogeneous	Within **1.8706(*)**	Between-2SLS 1.1768	FD-2SLS-KR **2.8141(**)**	FD-2SLS **1.993(**)**
	Heterogeneous	Average OLS **1.8543(*)**	Ind. OLS **1.9827(**)**	Average 2SLS **2.2883(**)**	Ind. 2SLS **2.6172(**)**
	Shrinkage	I. B. OLS **1.8304(*)**	It. Bayes **2.0116(**)**	It. Bayes **2.2921(**)**	Bayes 2SLS **2.6198(**)**

Note: This Table reports results for the case $(N, T) = (10, 20)$ under "mild" cross dependence (the support of ζ_i is [0,0.2]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C6: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(10,20)	(10,20)	(10,20)	(10,20)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	2SLS-KR 0.7216	GMM 1.0270	WLS 0.7561	GMM 1.0339
	Heterogeneous	Average OLS 0.7319	Ind. OLS 0.7430	Average 2SLS 0.7597	Ind. OLS 0.7086
	Shrinkage	I. B. 2SLS **0.7165**	Bayes OLS **0.6358**	I. B. 2SLS **0.7281**	Bayes OLS **0.6294**
DM	DM1	2SLS-KR vs. I. B. 2SLS (-0.1842)	Ind. OLS vs. Bayes OLS (-0.3241)	WLS vs. I. B. 2SLS (-0.5509)	Ind. OLS vs. Bayes OLS (-0.5827)
	DM2				
PT	Homogeneous	FDGMM 0.9797	Between-2SLS 0.8761	FD-2SLS-KR 1.2275	Between-2SLS 0.8505
	Heterogeneous	Average 2SLS 0.7796	Ind. OLS 1.3693	Ind. OLS 0.8809	Ind. OLS 1.5320
	Shrinkage	I. B. OLS 0.8872	I. B. 2SLS 1.4408	I. B. OLS 1.0679	Bayes OLS 1.5962

Note: This Table reports results for the case $(N, T) = (10, 20)$ under "large" cross dependence (the support of ζ_i is [-1,3]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C7: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(10,50)	(10,50)	(10,50)	(10,50)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	OLS **0.4244**	Between 0.9130	FD-2SLS **0.2743**	Between 0.9504
	Heterogeneous	Average OLS 0.4249	Ind. OLS 0.4250	Ind. OLS 0.4441	Ind. 2SLS 0.5004
	Shrinkage	I. B. 2SLS 0.4263	I. B. 2SLS **0.3566**	It. Bayes 0.4128	I. B. OLS **0.4191**
DM	DM1	Average OLS vs. OLS (0.1919)	Ind. OLS vs. I. B. 2SLS (-0.3208)	It.Bayes vs. FD-2SLS (2.5529)(**)	Ind. 2SLS vs. I. B. OLS (-0.4254)
	DM2	I. B. 2SLS vs. OLS (0.1294)			
PT	Homogeneous	Within **2.0593(*)**	FD-2SLS-KR **1.8359(*)**	FD-2SLS **2.8517(**)**	Within 1.5140
	Heterogeneous	Average 2SLS **2.0511(**)**	Ind. OLS **2.7020(**)**	Average 2SLS **2.4851(**)**	Ind. OLS **2.2132(**)**
	Shrinkage	I. B. OLS **2.0343(**)**	I. B. 2SLS **2.7046(**)**	It. Bayes **2.3819(**)**	Bayes OLS **2.2358(**)**

Note: This Table reports results for the case $(N, T) = (10, 50)$ under "mild" cross dependence (the support of ζ_i is [0,0.2]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C8: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(10,50)	(10,50)	(10,50)	(10,50)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	2SLS-KR 0.7132	FD-2SLS-KR 1.0412	2SLS-KR 0.7443	FD-2SLS-KR 1.1162
	Heterogeneous	Average 2SLS 0.7260	Ind. OLS 0.7238	Average 2SLS 0.7509	Ind. OLS 0.6933
	Shrinkage	I. B. OLS **0.7097**	I. B. 2SLS **0.7190**	I. B. 2SLS **0.7190**	I. B. OLS **0.6143**
DM	DM1	I. B. OLS vs. 2SLS-KR (-0.1522)	Ind. OLS vs. I. B. 2SLS (-0.3148)	2SLS-KR vs. I. B. 2SLS (-0.4638)	Ind. OLS vs. I. B. OLS (-0.3885)
	DM2				
PT	Homogeneous	2SLS-KR 1.0630	Within 1.3108	FD-2SLS-KR 1.2919	Within 1.3212
	Heterogeneous	Average OLS 1.0218	Ind. OLS 1.4738	Average OLS 1.0929	Ind. OLS 1.6247
	Shrinkage	I. B. OLS 1.0756	Bayes OLS 1.5318	I. B. 2SLS 1.2143	Bayes OLS **1.6810(*)**

Note: This Table reports results for the case $(N, T) = (10, 50)$ under "large" cross dependence (the support of ζ_i is [-1,3]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C9: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(20,50)	(20,50)	(20,50)	(20,50)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	WLS **0.4243**	EC2SLS-AR(1) 0.9244	FD-2SLS **0.2744**	EC2SLS-AR(1) 0.8767
	Heterogeneous	Average OLS 0.4256	Ind.OLS 0.3826	Average 2SLS 0.4458	Ind. 2SLS 0.3399
	Shrinkage	I.B. 2SLS 0.4254	I.B. OLS **0.3786**	I.B. 2SLS 0.4374	I.B. 2SLS **0.3323**
DM	DM1	I.B. 2SLS vs. WLS (0.1666)	I.B. OLS vs. Ind. OLS (-0.1734)	I.B. 2SLS vs. FD-2SLS (2.7041)(**)	Ind. 2SLS vs. I.B.2SLS (-0.3480)
	DM2				
PT	Homogeneous	Within **2.0342(**)**	Within 1.5275	FD-2SLS **2.8517(**)**	FD-2SLS-KR **1.8432(*)**
	Heterogeneous	Average OLS **2.0290(**)**	Ind. OLS **2.3158(**)**	Average 2SLS **2.4560(**)**	Ind. 2SLS **2.6993(**)**
	Shrinkage	Bayes OLS **2.0191(**)**	Bayes OLS **2.3199(**)**	I.B. 2SLS **2.3377(**)**	I.B. 2SLS **2.7002(**)**

Note: This Table reports results for the case $(N, T) = (20, 50)$ under "mild" cross dependence (the support of ζ_i is [0,0.2]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C10: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(20,50)	(20,50)	(20,50)	(20,50)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	2SLS 0.7055	FD-2SLS-KR 1.0049	WLS 0.7203	GMM 1.0140
	Heterogeneous	Average 2SLS 0.7050	Ind. OLS 0.5927	Average 2SLS 0.7282	Ind. OLS 0.5899
	Shrinkage	I.B. 2SLS **0.6970**	Bayes OLS **0.5806**	I.B. 2SLS **0.7132**	Bayes OLS **0.5779**
DM	DM1	Average 2SLS vs. I.B. 2SLS (-0.7593)	Ind. OLS vs. Bayes OLS (-0.4589)	WLS vs. I.B. 2SLS (-0.2570)	Ind. OLS vs. Bayes OLS (-0.3689)
	DM2				
PT	Homogeneous	2SLS-KR 1.0785	Within 1.3531	FD-2SLS-KR 1.2978	FD-2LSLS-KR 1.2116
	Heterogeneous	Average 2SLS 1.0502	Ind. 2SLS 1.6131	Average OLS 1.1505	Ind.OLS **1.7262(**)**
	Shrinkage	I.B. OLS 1.0797	Bayes OLS 1.6308	I.B. 2SLS 1.1829	Bayes OLS **1.7403(**)**

Note: This Table reports results for the case $(N,T) = (20,50)$ under "mild" cross dependence (the support of ζ_i is [-1,3]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0,0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C11: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(50,50)	(50,50)	(50,50)	(50,50)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	WLS **0.4217**	EC2SLS-AR(1) 0.8658	FD-2SLS **0.2954**	EC2SLS-AR(1) 0.7992
	Heterogeneous	Average 2SLS 0.4246	Ind. OLS 0.3982	Ind. 2SLS 0.4551	Ind. 2SLS 0.3848
	Shrinkage	Bayes OLS 0.4225	I.B. OLS **0.3952**	Bayes 2SLS 0.4495	I.B. 2SLS **0.3817**
DM	DM1	Bayes OLS vs. WLS (0.0333)	Ind. OLS vs. I.B. OLS (-0.1316)	Bayes 2SLS vs. FD-2SLS (2.7075)(**)	Ind. 2SLS vs. I.B. 2SLS (-0.1940)
	DM2				
PT	Homogeneous	Between-2SLS **2.0165(**)**	Within 1.2593	FD-2SLS **2.8223(**)**	Within 1.3702
	Heterogeneous	Average OLS **2.0217(**)**	Ind. OLS **2.1208(**)**	Average 2SLS **2.4043(**)**	Average OLS **2.0217(**)**
	Shrinkage	I.B. OLS **1.9996(**)**	I.B. OLS **2.1461(**)**	I.B. Bayes 2SLS **2.3046(**)**	I.B. OLS **1.9996(**)**

Note: This Table reports results for the case $(N,T) = (50,50)$ under "mild" cross dependence (the support of ζ_i is [0,0.2]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0,0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.

Table C12: Forecasting accuracy measures: Theil's U, Diebold and Mariano (1995) and Pesaran and Timmermann (1992) for homogeneous, heterogeneous and shrinkage/Bayesian estimators.

(T,N)		(50,50)	(50,50)	(50,50)	(50,50)
(ρ, ϑ, H)		(0,0,0.1)	(0,0,0.9)	(0.9,0.9,0.1)	(0.9,0.9,0.9)
Theil's U	Homogeneous	2SLS 0.6992	EC2SLS 1.0227	WLS 0.7172	EC2SLS 0.9964
	Heterogeneous	Average 2SLS 0.6989	Ind. OLS 0.5929	Average 2SLS 0.7251	Ind. OLS 0.5997
	Shrinkage	I.B. 2SLS **0.6931**	Bayes OLS **0.5870**	I.B. 2SLS **0.7153**	Bayes 2SLS **0.5937**
DM	DM1	Average 2SLS vs. I.B. 2SLS (-0.7603)	Ind. OLS vs. Bayes OLS (-0.3006)	WLS vs. I.B. 2SLS (-0.074)	Ind. OLS vs. Bayes 2SLS (-0.2384)
	DM2				
PT	Homogeneous	2SLS 1.0319	Within 1.0604	FD-2SLS-KR 1.2413	Within 1.0262
	Heterogeneous	Average 2SLS 1.0137	Ind. OLS 1.4924	Average 2SLS 1.1154	Ind. OLS 1.6103
	Shrinkage	I.B. 2SLS 1.0529	Bayes OLS 1.5384	I.B. 2SLS 1.1665	Bayes OLS **1.6500(*)**

Note: This Table reports results for the case $(N, T) = (50, 50)$ under "large" cross dependence (the support of ζ_i is [-1,3]), two different degrees of heterogeneity (low, with $H = 0.1$ and high with $H = 0.9$) and two different specifications for the error term dynamics (the white noise case with $(\rho, \vartheta) = (0, 0)$ and a nearly integrated one where $(\rho, \vartheta) = (0.9, 0.9)$). Forecasting horizon $h = 10$ periods ahead.