DEUTSCHE
BUNDESBANK
EUROSYSTEM

Forecasting with panel data

Badi H. Baltagi
(Syracuse University)

Discussion Paper
Series 1: Economic Studies
No 25/2006

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1
Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

ISBN 3–86558–185–4 (Printversion)
ISBN 3–86558–186–2 (Internetversion)

This paper was presented at the 8th Bundesbank spring conference (May 2006) on “New Developments in Economic Forecasting”. The views expressed in the paper are those of the authors and not necessarily those of the Bundesbank.

New Developments in Economic Forecasting
8th Bundesbank Spring Conference

(Preliminary programme as at 10 March 2006)

Organisers: Michael Binder, Jörg Breitung, Heinz Herrmann

Eltville, 05/06 May 2006

Friday, 5 May 2006

9:00 - 9:30	**Introductory remarks by** Axel A Weber *(Deutsche Bundesbank)*		
9:30 - 10:45	**Multiple structural breaks, forecasting and present value calculations** Hashem Pesaran *(University of Cambridge)*		
	Discussant:	Uwe Hassler *(Johann Wolfgang Goethe-University)* Siem Jan Koopman *(Free University Amsterdam)*	
10:45 – 11:00	*Coffee break*		
11:00 - 12:15	**Testing predictive ability using estimated co integration relationships: Business as usual?** Lutz Kilian *(University of Michigan)*		
	Discussant:	Hans-Eggert Reimers *(Hochschule Wismar)* Andreas Beyer *(European Central Bank)*	

12:15 – 14:00	*Lunch*	
14:00 – 15:00	**Global yield curve dynamics and interactions** Francis X Diebold *(University of Pennsylvania)*	
	Discussant:	Stefan Mittnik *(Ludwig-Maximilians-University Munich)* Joachim Grammig *(Eberhard-Karls-University Tübingen)*
15:00 - 15:15	*Coffee break*	
15:15 - 16:30	**A benchmark for models of growth and inflation** Massimiliano Marcellino *(Università Bocconi)*	
	Discussant:	George Kapetanios *(University of London)* Todd Clark *(Federal Reserve Bank of Kansas City)*
16:30 - 17:45	**Forecasting with panel data** Badi Baltagi *(Texas A&M University)*	
	Discussant:	Helmut Herwartz *(Christian-Albrechts-University Kiel)* Jean-Pierre Urbain *(University Maastricht)*
20:00	*Dinner* hosted by Axel A Weber *(Deutsche Bundesbank)*	
	Speaker:	Zdenek Tuma *(Czech National Bank)*

Saturday, 6 May 2006

9:30 - 10:45	**Did the ECB make a difference for Euro area business cycle?** Fabio Canova* *(Universitat Pompeu Fabra)* Matteo Ciccarelli *(European Central Bank)* Eva Ortega *(Bank of Spain)*		
	Discussant:	Sandra Eickmeier *(Deutsche Bundesbank)* Domenico Giannone *(Université Libre de Bruxelles)*	
10:45 - 11:00	*Coffee break*		
11:00 - 12:15	**Bayesian density forecasting with best performing priors determined by entropic tilting** Charles Whiteman* (University of Iowa) Kurt Lewis (University of Iowa)		
	Discussant:	John Geweke *(University of Iowa)* Stephane Adjemian *(CEPREMAP)*	
12:15 - 13:30	*Lunch*		
13:30 - 14:45	**Shrinkage methods for forecasting using many predictors** Mark Watson* *(Princeton University)* James Stock *(Harvard University)*		
	Discussant:	Gary Koop *(University of Leicester)* Carlo Favero *(Universitat Pompeu Fabra)*	

14:45 - 16:00	**Forecasting using a large number of predictors** Lucrezia Reichlin *(European Central Bank)*
	Discussant: Christian Schumacher *(Deutsche Bundesbank)* Peter Vlaar *(Dutch Central Bank)*
16:00	**Concluding remarks by** Hermann Remsperger *(Deutsche Bundesbank)*

Deutsche Bundesbank
Protocol and Conferences, K 100
Tel: +49 69 9566-4066/-2864, Fax: +49 69 9566 4027
Mobile: +49 177 7382251, E-mail: conferences@bundesbank.de

Forecasting with Panel Data*

Badi H. Baltagi†

April 18, 2006

Abstract

This paper gives a brief survey of forecasting with panel data. Starting with a simple error component regression model and surveying best linear unbiased prediction under various assumptions of the disturbance term. This includes various ARMA models as well as spatial autoregressive models. The paper also surveys how these forecasts have been used in panel data applications, running horse races between heterogeneous and homogeneous panel data models using out of sample forecasts.

Key words: Forecasting; BLUP; Panel Data; Spatial Dependence; Serial Correlation.

JEL classification: C33

*This paper was prepared for presentation at the 8th Bundesbank Spring Conference on "New Devlopments in Economic Forecasting" in Eltville, Germany, 5-6 May, 2006.

†Badi H. Baltagi: Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020 U.S.A.; bbaltagi@maxwell.syr.edu

Non-technical summary

This paper provides a brief overview of forecasting with panel data. First, it discusses what the best linear unbiased prediction is when various assumptions are made about the error term, i. e. one-way error components, whereby the error term entails an unobservable individual effect; two-way error components, where the error term entails an unobservable individual effect and a time-specific effect; serial correlation in the error term; error terms with spatial correlation. Second, the paper considers heterogeneous data models versus homogeneous data models. It reviews numerous papers and discusses the conditions under which heterogeneous or homogeneous panel estimators perform better and in which cases pooling techniques are preferable to single country forecasts. Finally, the paper points to several outstanding issues and indicates the work which needs to be done if further progress is to be made in forecasting in a panel environment.

Nicht-technische Zusammenfassung

Dieses Papier gibt einen kurzen Überblick über die Prognose mit Paneldaten. Zunächst diskutiert es den besten linearen unverzerrten Schätzer wenn verschiedene Annahmen hinsichtlich des Fehlerterms gegeben sind: Der Fehlerterm enthält einen unbeobachtbaren individuellen Effekt; der Fehlerterm enthält sowohl einen unbeobachtbaren individuellen als auch einen zeitspezifischen Effekt; der Fehlerterm ist zeitlich korreliert; der Fehlerterm ist räumlich korreliert. Anschließend betrachtet das Papier Modelle mit heterogenen und homogenen Daten. Es gibt einen Überblick über viele Papiere und diskutiert, unter welchen Bedingungen homogene oder heterogene Panelschätzer sich als überlegen erwiesen haben und wo sich Poolingtechniken gegenüber Einzelgleichungsansätzen bei der Prognose als überlegen gezeigt haben. Schließlich wirft das Papier mehrere offene Fragen auf und zeigt welche Arbeiten noch zu erledigen sind, um bei der Prognose mit Paneldaten Fortschritte zu machen.

1 Introduction

Consider a panel data regression model.

$$y_{it} = \alpha + X'_{it}\beta + u_{it} \quad i = 1, \ldots, N; \quad t = 1, ..., T \tag{1}$$

with i denoting households, individuals, firms, countries, etc., and t denoting time. The i subscript, therefore, denotes the cross-section dimension whereas t denotes the time-series dimension. α is a scalar, β is $K \times 1$ and X_{it} is the itth observation on K explanatory variables. Most of the panel data applications utilize a one-way error component model for the disturbances, with

$$u_{it} = \mu_i + \nu_{it} \tag{2}$$

where μ_i denotes the *unobservable* individual specific effect and v_{it} denotes the remainder disturbance. For example, in an earnings equation in labor economics, y_{it} will measure earnings of the head of the household, whereas X_{it} may contain a set of variables like experience, education, union membership, sex, race, etc. Note that μ_i is time-invariant and it accounts for any individual specific effect that is not included in the regression. In this case we could think of it as the individual's unobserved ability. The remainder disturbance v_{it} varies with individuals and time and can be thought of as the usual disturbance in the regression. This can be written as

$$y = \alpha\iota_{NT} + X\beta + u = Z\delta + u \tag{3}$$

where y is $NT\times 1$, X is $NT \times K$, $Z = [\iota_{NT}, X], \delta' = (\alpha', \beta')$ and ι_{NT} is a vector of ones of dimension NT. Also,

$$u = Z_\mu\mu + \nu \tag{4}$$

where $u' = (u_{11,}, \ldots, u_{1T}, u_{21}, \ldots, u_{2T}, \ldots, u_{N1}, \ldots, u_{NT})$ with the observations stacked such that the slower index is over individuals and the faster index is over time. $Z_\mu =$

$I_N \otimes \iota_T$ where I_N is an identity matrix of dimension N, ι_T is a vector of ones of dimension T and $\otimes$ denotes Kronecker product. Z_μ, is a selector matrix of ones and zeros, or simply the matrix of individual dummies that one may include in the regression to estimate the μ_i if they are assumed to be fixed parameters. $\mu' = (\mu_1, \ldots, \mu_N)$ and $\nu' = (\nu_{11}, \ldots, \nu_{1T}, \ldots, \nu_{N1}, \ldots, \nu_{NT})$. Note that, $Z_\mu Z'_\mu = I_N \otimes J_T$ where J_T is a matrix of ones of dimension T, and $P = Z_\mu(Z'_\mu Z_\mu)^{-1}Z'_\mu$, the projection matrix on Z_μ reduces to $I_N \otimes \bar{J}_T$ where $\bar{J}_T = J_T/T$. P is a matrix which averages the observation across time for each individual, and $Q = I_{NT} - P$ is a matrix which obtains the deviations from individual means. For example, regressing y on the matrix of dummy variables Z_μ gets the predicted values Py which have a typical element $\overline{y}_{i.} = \sum_{t=1}^{T} y_{it}/T$ repeated T times for each individual. The residuals of this regression are given by Qy which have a typical element $(y_{it} - \overline{y}_{i.})$.

For the *fixed effects* case, the μ_i are assumed to be fixed parameters to be estimated and the remainder disturbances stochastic with v_{it} independent and identically distributed IID$(0, \sigma^2_\nu)$. The X_{it} are assumed independent of the v_{it} for all i and t. The LSDV (least squares dummy variables) estimator performs ordinary least squares (OLS) on

$$y = \alpha\iota_{NT} + X\beta + Z_\mu\mu + \nu = Z\delta + Z_\mu\mu + \nu \tag{5}$$

Note that Z is $NT \times (K+1)$ and Z_μ, the matrix of individual dummies, is $NT \times N$. If N is large, this will include too many individual dummies, and the matrix to be inverted by OLS is large and of dimension $(N+K)$. Alternatively, one can premultiply the model by Q and perform OLS on the resulting transformed model:

$$Qy = QX\beta + Qv \tag{6}$$

This uses the fact that $QZ_\mu = Q\iota_{NT} = 0$, since $PZ_\mu = Z_\mu$. In other words, the Q matrix wipes out the individual effects. This is a regression of $\widetilde{y} = Qy$ with typical element $(y_{it} - \bar{y}_{i.})$ on $\widetilde{X} = QX$ with typical element $(X_{_{it,k}} - \bar{X}_{i.,k})$ for the kth regressor, $k = 1, 2, \ldots, K$. This involves the inversion of a $(K \times K)$ matrix rather than $(N+K) \times (N+K)$.

The resulting OLS estimator is

$$\widetilde{\beta}_{FE} = (X'QX)^{-1} X'Qy \tag{7}$$

with $\text{var}(\widetilde{\beta}) = \sigma_\nu^2(X'QX)^{-1} = \sigma_\nu^2(\widetilde{X}'\widetilde{X})^{-1}$.

There are too many parameters in the fixed effects model and the loss of degrees of freedom can be avoided if the μ_i can be assumed random. For the *random effects* case $\mu_i \sim \text{IID}(0, \sigma_\mu^2), \nu_{it} \sim \text{IID}(0, \sigma_\nu^2)$ and the μ_i are independent of the ν_{it}. In addition, the X_{it} are independent of the μ_i and ν_{it}, for all i and t. The variance-covariance matrix is given by

$$\Omega = E(uu') = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\nu^2(I_N \otimes I_T) = \sigma_1^2 P + \sigma_\nu^2 Q \tag{8}$$

where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$. This is the spectral decomposition representation of Ω, with σ_1^2 being the first unique characteristic root of Ω of multiplicity N and σ_ν^2 is the second unique characteristic root of Ω of multiplicity $N(T-1)$. It is easy to verify, using the properties of P and Q, that

$$\Omega^{-1} = \frac{1}{\sigma_1^2}P + \frac{1}{\sigma_\nu^2}Q \tag{9}$$

and

$$\Omega^{-1/2} = \frac{1}{\sigma_1}P + \frac{1}{\sigma_\nu}Q \tag{10}$$

In fact, $\Omega^r = (\sigma_1^2)^r P + (\sigma_\nu^2)^r Q$ where r is an arbitrary scalar. Now we can obtain GLS as a weighted least squares. Fuller and Battese (1974) suggested premultiplying the regression equation by $\sigma_\nu \Omega^{-1/2} = Q + (\sigma_\nu/\sigma_1)P$ and performing OLS on the resulting transformed regression. In this case, $y^* = \sigma_\nu \Omega^{-1/2} y$ has a typical element $y_{it}^* = y_{it} - \theta \bar{y}_{i.}$ where $\theta = 1 - (\sigma_\nu/\sigma_1)$. This transformed regression inverts a matrix of dimension $(K+1)$ and can be easily implemented using any regression package.

The best quadratic unbiased (BQU) estimators of the variance components arise naturally from the spectral decomposition of Ω. In fact, $Pu \sim (0, \sigma_1^2 P)$ and $Qu \sim (0, \sigma_\nu^2 Q)$

and

$$\widehat{\sigma}_1^2 = \frac{u'Pu}{tr(P)} = T\sum_{i=1}^{N}\bar{u}_{i.}^2/N \tag{11}$$

and

$$\widehat{\sigma}_\nu^2 = \frac{u'Qu}{tr(Q)} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(u_{it}-\bar{u}_{i.})^2}{N(T-1)} \tag{12}$$

provide the BQU estimators of σ_1^2 and σ_ν^2, respectively.

2 The Best Linear Unbiased Predictor

Suppose we want to predict S periods ahead for the ith individual. For the GLS model, knowing the variance-covariance structure of the disturbances, Goldberger (1962) showed that the best linear unbiased predictor (BLUP) of $y_{i,T+S}$ is

$$\widehat{y}_{i,T+S} = Z'_{i,T+S}\widehat{\delta}_{GLS} + w'\Omega^{-1}\widehat{u}_{GLS} \quad \text{for } s \geqslant 1 \tag{13}$$

where $\widehat{u}_{GLS} = y - Z\widehat{\delta}_{GLS}$ and $w = E(u_{i,T+S}u)$. Note that for period $T+S$

$$u_{i,T+S} = \mu_i + \nu_{i,T+S} \tag{14}$$

and $w = \sigma_\mu^2(l_i \otimes \iota_T)$ where l_i is the ith column of I_N, i.e. l_i is a vector that has 1 in the ith position and zero elsewhere. In this case

$$w'\Omega^{-1} = \sigma_\mu^2(l'_i \otimes \iota'_T)\left[\frac{1}{\sigma_1^2}P + \frac{1}{\sigma_\nu^2}Q\right] = \frac{\sigma_\mu^2}{\sigma_1^2}(l'_i \otimes \iota'_T) \tag{15}$$

since $(l'_i \otimes \iota'_T)P = (l'_i \otimes \iota'_T)$ and $(l'_i \otimes \iota'_T)Q = 0$. The typical element of $w'\Omega^{-1}\widehat{u}_{GLS}$ becomes $((T\sigma_\mu^2/\sigma_1^2)\overline{\widehat{u}}_{i.,GLS})$ where $\overline{\widehat{u}}_{i.,GLS} = \sum_{t=1}^{T}\widehat{u}_{it,GLS}/T$. Therefore, the BLUP for $y_{i,T+S}$ corrects the GLS prediction by a fraction of the mean of the GLS residuals corresponding to that ith individual. This predictor was considered by Taub (1979). The BLUP are optimal assuming true values of the variance components. In practice, these are replaced with estimated values that yield empirical BLUP. Kackar and Harville (1984) propose inflation factors that account for the additional uncertainty introduced by estimating these variance components.

Baillie and Baltagi (1999) consider the practical situation of prediction from the error component regression model when the variance components are not known. They derive both theoretical and simulation evidence as to the relative efficiency of four alternative predictors: (i) an ordinary predictor, based on the optimal predictor but with MLEs replacing population parameters, (ii) a truncated predictor that ignores the error component correction, but uses MLEs for its regression parameters, (iii) a misspecified predictor which uses OLS estimates of the regression parameters, and (iv) a fixed effects predictor which assumes that the individual effects are fixed parameters that can be estimated. The asymptotic formula for MSE prediction are derived for all four predictors. Using numerical and simulation results, these are shown to perform adequately in realistic sample sizes ($N = 50$ and 500 and $T = 10$ and 20). Both the analytical and sampling results show that there are substantial gains in mean square error prediction by using the ordinary predictor instead of the misspecified or the truncated predictors, especially with increasing $\rho = \sigma_\mu^2/(\sigma_\mu^2 + \sigma_\nu^2)$ values. The reduction in MSE is about ten fold for $\rho = 0.9$ and a little more than two fold for $\rho = 0.6$ for various values of N and T. The fixed effects predictor performs remarkably well being a close second to the ordinary predictor for all experiments. Simulation evidence confirm the importance of taking into account the individual effects when making predictions. The ordinary predictor and the fixed effects predictor outperform the truncated and misspecified predictors and are recommended in practice.

It is important to note that BLUP is a statistical methodology that has been used extensively in animal breeding, see Henderson (1975) and Harville (1976). It is used to estimate genetic merits. For example, in animal breeding, one predicts the production of milk by daughter cows based on their lineage. Robinson (1991) is a good review of BLUP and how it can be used to derive the Kalman filter, the method of Kriging used for ore reserve estimation, credibility theory used to work out insurance premiums, removing noise from images and for small-area estimation. Robinson argues that BLUP is a method of estimating random effects. While BLUP was developed via a frequentist approach to statistics, it has a Bayesian interpretation, see Harville (1976) who showed that Bayesian posterior mean predictors with a diffuse prior are equivalent to BLUP. Robinson adds (1991, p.30) that one of the reasons why the estimation of random effects has been neglected by the classical school of thought is that : "The idea of estimating random effects seems suspiciously Bayesian to some Classical statisticians... adding that..the adherents of each school emphasize the differences rather than the similarities." One of the commentators of the paper paraphrase I. J. Good's memorable aphorism: "To a Bayesian, all things are Bayesian." He argues that a summary of Robinson's paper could be " To a non-Bayesian, all things are BLUPs". For an application in actuarial science to the problem of predicting future claims of a risk class, given past claims of that and related risk classes, see Frees et al. (1999, 2001). Also, Battese, Harter and Fuller (1988) for predicting county crop areas with survey and satelite data using an error component model.

How does the best linear unbiased predictor (BLUP) look like for the ith individual, S periods ahead for the two-way model? For the two-way error components disturbances:

$$u_{it} = \mu_i + \lambda_t + \nu_{it} \qquad i = 1, \ldots, N \qquad t = 1, \ldots, T \tag{16}$$

with $\mu_i \sim \text{IID}(0, \sigma_\mu^2)$, $\lambda_t \sim \text{IID}(0, \sigma_\lambda^2)$ and $\nu_{it} \sim \text{IID}(0, \sigma_\nu^2)$ independent of each other. In addition, X_{it} is independent of μ_i, λ_t and ν_{it} for all i and t. The variance-covariance matrix is given by

$$\Omega = E(uu') = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\lambda^2(J_N \otimes I_T) + \sigma_\nu^2(I_N \otimes I_T) \tag{17}$$

The disturbances are homoskedastic with $\text{var}(u_{it}) = \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_\nu^2$ for all i and t,

$$\begin{aligned} \text{cov}(u_{it}, u_{js}) &= \sigma_\mu^2 \quad i = j,\ t \neq s \\ &= \sigma_\lambda^2 \quad i \neq j,\ t = s \end{aligned}$$

and zero otherwise. For period $T + S$

$$u_{i,T+S} = \mu_i + \lambda_{T+S} + \nu_{i,T+S} \tag{18}$$

and

$$\begin{aligned} E(u_{i,T+S}u_{jt}) &= \sigma^2_\mu \quad \text{for } i = j \\ &= 0 \quad \text{for } i \neq j \end{aligned} \tag{19}$$

and $t = 1, 2, \ldots, T$. Hence, $w = E(u_{i,T+S}u) = \sigma^2_\mu(l_i \otimes \iota_T)$ remains the same for the two-way model as in the one-way model, where l_i is the ith column of I_N. However, Ω^{-1} is different, and the typical element of $w'\Omega^{-1}\widehat{u}_{GLS}$ where $\widehat{u}_{GLS} = y - Z\widehat{\delta}_{GLS}$ is

$$\frac{T\sigma^2_\mu}{(T\sigma^2_\mu + \sigma^2_\nu)}(\overline{\widehat{u}}_{i.,GLS} - \overline{\widehat{u}}_{..,GLS}) + \frac{T\sigma^2_\mu}{(T\sigma^2_\mu + N\sigma^2_\lambda + \sigma^2_\nu)}\overline{\widehat{u}}_{..,GLS} \tag{20}$$

where $\overline{\widehat{u}}_{i.,GLS} = \sum_{t=1}^{T} \widehat{u}_{it,GLS}/T$ and $\overline{\widehat{u}}_{..,GLS} = \sum_i \sum_t \widehat{u}_{it,GLS}/NT$. In general, $\overline{\widehat{u}}_{..,GLS}$ is not necessarily zero. The GLS normal equations are $Z'\Omega^{-1}\widehat{u}_{GLS} = 0$. However, if Z contains a constant, then $\iota'_{NT}\Omega^{-1}\widehat{u}_{GLS} = 0$, and using the fact that $\iota'_{NT}\Omega^{-1} = \iota'_{NT}/(T\sigma^2_\mu + N\sigma^2_\lambda + \sigma^2_\nu)$, one gets $\overline{\widehat{u}}_{..,GLS} = 0$. Hence, for the two-way model, if there is a constant in the model, the BLUP for $y_{i,T+S}$ corrects the GLS prediction by a fraction of the mean of the GLS residuals corresponding to that ith individual

$$\widehat{y}_{i,T+S} = Z'_{i,T+S}\widehat{\delta}_{GLS} + \left(\frac{T\sigma^2_\mu}{T\sigma^2_\mu + \sigma^2_\nu}\right)\overline{\widehat{u}}_{i.,GLS} \tag{21}$$

This looks exactly like the BLUP for the one-way model but with a different Ω.

How would one forecast with a two-way fixed effects model with both country and time effects? After all, future coefficients of time dummies cannot be estimated unless more structure can be placed on the model. One example is the study by Schmalensee, Stoker and Judson (1998) which forecasted the world carbon dioxide emissions through 2050 using national-level panel data over the period 1950-1990. This consisted of 4018 observations. In 1990, this data covered 141 countries which accounted for 98.6% of the world's population. This paper estimated a reduced form model relating per capita CO_2 emissions from energy consumption to a flexible functional form of real GDP per capita using time and period fixed effects. Schmalensee, Stoker and Judson (1998) forecasted the time effects using a linear spline model with different growth rates prior to 1970 and after 1970, i.e., $\lambda_t = \gamma_1 + \gamma_2 t + \gamma_3(t - 1970).1[t \geqq 1970]$, with the last term being an indicator function which is 1 when $t \geqq 1970$. Also, using a nonlinear trend model including a logarithmic term, i.e., $\lambda_t = \delta_1 + \delta_2 t + \delta_3 \ln(t - 1940)$. Although these two time effects specifications had essentially the same goodness-of-fit performance, they resulted in different out of sample projections. The linear spline projected the time effects by continuing the estimated 1970-1990 trend to 2050, while the nonlinear trend projected a flattening trend consistent with the trend deceleration from 1950 to 1990. An earlier study by Holtz-Eakin and Selden (1995) employed 3754 observations over the period 1951-1986. For their main case, they simply set the time effect at its value in the last year in their sample.

2.1 Serial Correlation

So far, we have derived Goldberger's (1962) BLUP of $y_{i,T+S}$ for the one-way error component model without serial correlation. For ease of reference, we reproduce the one period ahead forecast for the ith individual

$$\widehat{y}_{i,T+1} = Z'_{i,T+1}\widehat{\delta}_{GLS} + w'\Omega^{-1}\widehat{u}_{GLS} \tag{22}$$

where $\widehat{u}_{GLS} = y - Z\widehat{\delta}_{GLS}$ and $w = E(u_{i,T+1}u)$. For the AR(1) model with no error components, a standard result is that the last term reduces to $\rho\widehat{u}_{i,T}$, where $\widehat{u}_{i,T}$ is the Tth GLS residual for the ith individual. For the one-way error component model without serial correlation (see Taub, 1979), the last term reduces to $[T\sigma_\mu^2/(T\sigma_\mu^2 + \sigma_\nu^2)]\overline{\widehat{u}}_{i.}$, where $\overline{\widehat{u}}_{i.} = \sum_{t=1}^{T} \widehat{u}_{it}/T$, is the average of the ith individual's GLS residuals. Baltagi and Li (1992) showed that when *both* error components and serial correlation are present, i.e.,

$$\nu_{it} = \rho\nu_{i,t-1} + \epsilon_{it} \tag{23}$$

$\mid \rho \mid < 1$ and $\epsilon_{it} \sim \text{IID}(0, \sigma_\epsilon^2)$. The μ_i are independent of the ν_{it} and $\nu_{i0} \sim (0, \sigma_\epsilon^2/(1-\rho^2))$. The last term reduces to

$$w'\Omega^{-1}\widehat{u}_{GLS} = \rho\widehat{u}_{i,T} + \left(\frac{(1-\rho)^2\sigma_\mu^2}{\sigma_\omega^2}\right)\left[\omega\hat{u}_{i1}^* + \sum_{t=2}^{T}\hat{u}_{it}^*\right] \tag{24}$$

where u_{it}^* denotes the Prais-Winsten-transformed residuals

$$\begin{aligned} u_{it}^* &= \sqrt{1-\rho^2}\, u_{i1} \quad \text{for } t = 1 \\ &= u_{it} - \rho u_{i,t-1} \quad \text{for } t = 2, \ldots, T \end{aligned}$$

with $\omega = \sqrt{(1+\rho)/(1-\rho)}$, $\sigma_\omega^2 = d^2\sigma_\mu^2(1-\rho)^2 + \sigma_\epsilon^2$, and $d^2 = \omega^2 + (T-1)$. Note that $\hat{u}_{i1}^*$ receives an ω weight in averaging across the ith individual's residuals. (i) If $\sigma_\mu^2 = 0$, so that only serial correlation is present, the prediction correction term reduces to $\rho\widehat{u}_{i,T}$. Similarly, (ii) if $\rho = 0$, so that only error components are present, this reduces to $[T\sigma_\mu^2/(T\sigma_\mu^2 + \sigma_\nu^2)]\overline{\widehat{u}}_{i.}$.

For the one-way error component model with remainder disturbances following an AR(2) process, i.e.,

$$\nu_{it} = \rho_1\nu_{i,t-1} + \rho_2\nu_{i,t-2} + \epsilon_{it} \tag{25}$$

where $\epsilon_{it} \sim \text{IIN}(0, \sigma_\epsilon^2), \mid \rho_2 \mid < 1$ and $\mid \rho_1 \mid < (1-\rho_2)$. Baltagi and Li (1992) find that the last term reduces to

$$\begin{aligned} w'\Omega^{-1}\widehat{u}_{GLS} &= \rho_1\widehat{u}_{i,T-1} + \rho_2\widehat{u}_{i,T-2} \\ &\quad + \left[\frac{(1-\rho_1-\rho_2)^2\sigma_\mu^2}{\sigma_\omega^2}\right]\left[\omega_1\hat{u}_{i1}^* + \omega_2\hat{u}_{i2}^* + \sum_{t=3}^{T}\hat{u}_{it}^*\right] \end{aligned} \tag{26}$$

where

$$\begin{aligned} \omega_1 &= \sigma_\epsilon/\sigma_\nu(1-\rho_1-\rho_2) \quad \omega_2 = \sqrt{(1+\rho_2)/(1-\rho_2)} \\ \sigma_\omega^2 &= d^2\sigma_\mu^2(1-\rho_1-\rho_2)^2 + \sigma_\epsilon^2 \\ d^2 &= \omega_1^2 + \omega_2^2 + (T-2) \end{aligned}$$

and

$$
\begin{aligned}
\hat{u}^*_{i1} &= (\sigma_\epsilon/\sigma_\nu)\widehat{u}_{i1} \\
\hat{u}^*_{i2} &= \sqrt{1-\rho_2^2}\,[\widehat{u}_{i2} - (\rho_1/(1-\rho_2))\widehat{u}_{i1}] \\
\hat{u}^*_{it} &= \widehat{u}_{it} - \rho_1\widehat{u}_{i,t-1} - \rho_2\widehat{u}_{i,t-2} \quad \text{for } t = 3,\ldots,T
\end{aligned}
$$

Note that if $\rho_2 = 0$, this predictor reduces to that of the AR(1) model with RE. Also, note that for this predictor, the first two residuals are weighted differently when averaging across the ith individual's residuals.

For the one-way error component model with remainder disturbances following the specialized AR(4) process for quarterly data, i.e., $\nu_{it} = \rho\nu_{i,t-4} + \epsilon_{it}$, where $|\rho| < 1$ and $\epsilon_{it} \sim \text{IIN}(0, \sigma_\epsilon^2)$. Baltagi and Li (1992) find that the last term reduces to

$$
w'\Omega^{-1}\widehat{u}_{GLS} = \rho\widehat{u}_{i,T-3} + \left[\frac{(1-\rho)^2\sigma_\mu^2}{\sigma_\alpha^2}\right]\left[\omega\sum_{t=1}^{4}\hat{u}^*_{it} + \sum_{t=5}^{T}\hat{u}^*_{it}\right] \tag{27}
$$

where $\omega = \sqrt{(1+\rho)/(1-\rho)}$, $\sigma_\omega^2 = d^2(1-\rho)^2\sigma_\mu^2 + \sigma_\epsilon^2$, $d^2 = 4\omega^2 + (T-4)$, and

$$
\begin{aligned}
u^*_{it} &= \sqrt{1-\rho^2}\; u_{it} \quad \text{for } t = 1,2,3,4 \\
&= u_{it} - \rho u_{i,t-4} \quad \text{for } t = 5,6,\ldots,T
\end{aligned}
$$

Note, for this predictor, that the first four quarterly residuals weighted by ω when averaging across the ith individual's residuals.

Finally, for the one-way error component model with remainder disturbances following an MA(1) process, i.e.,

$$
\nu_{it} = \epsilon_{it} + \lambda\epsilon_{i,t-1}
$$

where $\epsilon_{it} \sim \text{IIN}(0, \sigma_\epsilon^2)$ and $|\lambda| < 1$, Baltagi and Li (1992) find that

$$
\begin{aligned}
w'\Omega^{-1}\widehat{u}_{GLS} &= -\lambda\left(\frac{a_{T-1}}{a_T}\right)^{1/2}\hat{u}^*_{iT} \\
&\quad + \left[1 + \lambda\left(\frac{a_{T-1}}{a_T}\right)^{1/2}\alpha_T\right]\left(\frac{\sigma_\mu^2}{\sigma_\omega^2}\right)\left[\sum_{t=1}^{T}\alpha_t\hat{u}^*_{it}\right]
\end{aligned} \tag{28}
$$

where $a_t = 1 + \lambda^2 + \ldots + \lambda^{2t}$ with $a_0 = 1$, $\sigma_\omega^2 = d^2\sigma_\mu^2 + \sigma_\epsilon^2$ and $d^2 = \sum_{t=1}^T \alpha_t^2$, and the $\hat{u}_{it}^*$, can be solved for recursively as follows:

$$\begin{aligned}\hat{u}_{i1}^* &= (a_0/a_1)^{1/2}\hat{u}_{i1} \\ \hat{u}_{it}^* &= \lambda(a_{t-2}/a_{t-1})^{1/2}\hat{u}_{i,t-1}^* + (a_{t-1}/a_t)^{1/2}\hat{u}_{i,t} \quad t = 2, \ldots, T\end{aligned}$$

If $\lambda = 0$, then $a_t = \alpha_t = 1$ for all t, the prediction correction term reduces to the predictor for the error component model with no serial correlation. If $\sigma_\mu^2 = 0$, the predictor reduces to that of the MA(1) process.

These results can be extended to the MA(q) case, see Baltagi and Li (1994) and the autoregressive moving average ARMA(p, q) case on the ν_{it}, see MaCurdy (1982) and more recently Galbraith and Zinde-Walsh (1995). For an extension to the two-way model with serially correlated disturbances, see Revankar (1979) who considers the case where the λ_t also follow an AR(1) process. Also, Karlsson and Skoglund (2004) who consider the two-way error component model with an ARMA process on the time specific effects. For an extension to the unequally spaced panel data regression model with AR(1) remainder disturbances, see Baltagi and Wu (1999).

Frees and Miller (2004) forecast the sale of state lottery tickets using panel data from 50 postal (ZIP) codes in Wisconsin observed over 40 weeks. The first 35 weeks of data are used to estimate the model and the remaining five weeks are used to assess the validity of model forecasts. Using the mean absolute error criteria and the mean absolute percentage error criteria, the best forecasts were given by the error component model with AR(1) disturbances followed by the fixed effects model with AR(1) disturbances.

2.2 Spatial Correlation

Consider the spatial panel data model:

$$y_{it} = x_{it}'\beta + \varepsilon_{it} \qquad i = 1, ..., N;\ t = 1, ..., T \tag{29}$$

see Anselin (1988, p 152), where the disturbance vector for time t is given by

$$\varepsilon_t = \mu + \phi_t \tag{30}$$

with $\varepsilon_t = (\varepsilon_{1t}, ..., \varepsilon_{Nt})'$, $\mu = (\mu_1, ..., \mu_N)'$ denotes the vector of individual effects and $\phi_t = (\phi_{1t}, ..., \phi_{Nt})'$ are the remainder disturbances which are independent of μ. The ϕ_t's follow the spatial error dependence model

$$\phi_t = \lambda W \phi_t + \nu_t \tag{31}$$

where W is the matrix of known spatial weights of dimension $N \times N$ with zero diagonal elements and row normalized elements that sum to 1. λ is the spatial autoregressive coefficient, $\nu_t = (\nu_{1t}, ..., \nu_{Nt})'$ is $iid(0, \sigma_\nu^2)$ and is independent of ϕ_t and μ.

For the random effects model, the μ_i's are $iid(0, \sigma_\mu^2)$ and are independent of the ϕ_{it}'s, see Anselin (1988). For this model, we need to derive the variance-covariance matrix. Let $B = I_N - \lambda W$, then the disturbances in equation (31) can be written as follows: $\phi_t = (I_N - \lambda W)^{-1}\nu_t = B^{-1}\nu_t$. Substituting for ϕ_t, we get

$$\varepsilon = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})\nu \tag{32}$$

where ι_T is a vector of ones of dimension T and I_N is an identity matrix of dimension N. The variance covariance matrix is

$$\Omega = E(\varepsilon\varepsilon') = \sigma_\mu^2(\iota_T\iota_T' \otimes I_N) + \sigma_\nu^2(I_T \otimes (B'B)^{-1}) \tag{33}$$

Let $\Psi = \frac{1}{\sigma_\nu^2}\Omega = \frac{\sigma_\mu^2}{\sigma_\nu^2}(\iota_T\iota_T' \otimes I_N) + (I_T \otimes (B'B)^{-1})$ and $\theta = \frac{\sigma_\mu^2}{\sigma_\nu^2}$, then

$$\Psi = \bar{J}_T \otimes (T\theta I_N) + I_T \otimes (B'B)^{-1} = \bar{J}_T \otimes V + E_T \otimes (B'B)^{-1} \tag{34}$$

where $V = T\theta I_N + (B'B)^{-1}$ and $E_T = I_T - \bar{J}_T$. It is easy to verify that

$$\Psi^{-1} = \bar{J}_T \otimes V^{-1} + E_T \otimes (B'B) \tag{35}$$

see Anselin (1988, p.154). In this case, GLS using Ω^{-1} yields $\hat{\beta}_{GLS}$.Note that the computation is simplified, since the $NT \times NT$ matrix Ψ^{-1} is based on inverting two lower order matrices, V and B both of dimensions $N \times N$.

If $\lambda = 0$, so that there is no spatial autocorrelation, then $B = I_N$ and Ω becomes the usual error component variance-covariance matrix

$$\Omega_{RE} = E(\varepsilon\varepsilon') = \sigma_\mu^2(\iota_T\iota_T' \otimes I_N) + \sigma_\nu^2(I_T \otimes I_N) \tag{36}$$

Applying GLS using this Ω_{RE} yields the random effects (RE) estimator which we will denote by $\hat{\beta}_{RE}$.

Baltagi and Li (2004) derived the BLUP correction term when both error components and spatial autocorrelation are present. In this case $\omega = E(\varepsilon_{i,T+S}\varepsilon) = E[(\mu_i + \phi_{i,T+S})\varepsilon] = \sigma_\mu^2(\iota_T \otimes l_i)$ since the ϕ's are not correlated over time. Using $\Omega^{-1} = \frac{1}{\sigma_\nu^2}\Psi^{-1}$, we get

$$\omega'\Omega^{-1} = \frac{\sigma_\mu^2}{\sigma_\nu^2}(\iota_T' \otimes l_i')[(\bar{J}_T \otimes V^{-1}) + (E_T \otimes (B'B))] = \theta(\iota_T' \otimes l_i'V^{-1}) \tag{37}$$

since $\iota_T' E_T = 0$. Therefore

$$\omega'\Omega^{-1}\hat{\varepsilon}_{GLS} = \theta(\iota_T' \otimes l_i'V^{-1})\hat{\varepsilon}_{GLS} = \theta\, l_i'V^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{t,GLS} = T\theta\sum_{j=1}^{N}\delta_j\bar{\varepsilon}_{j.,GLS} \tag{38}$$

where δ_j is the jth element of the ith row of V^{-1} and $\bar{\varepsilon}_{j.,GLS} = \sum_{t=1}^{T}\hat{\varepsilon}_{tj,GLS}/T$. In other words, the BLUP adds to $x_{i,T+S}'\hat{\beta}_{GLS}$ a weighted average of the GLS residuals for the N regions averaged over time. The weights depend upon the spatial matrix W and the spatial autocorrelation coefficient λ. To make this predictor operational, we replace $\hat{\beta}_{GLS}$, θ and λ by their estimates from the RE-spatial MLE.

When there is no spatial autocorrelation, i.e., $\lambda = 0$, the BLUP correction term reduces to the Taub (1979) predictor term of the RE model. Also, when there are no random effects, so that $\sigma_\mu^2 = 0$, then $\theta = 0$ and the BLUP prediction term drops out completely. In this case, Ω reduces to $\sigma_\nu^2(I_T \otimes (B'B)^{-1})$ and GLS on this model, based on the MLE of λ, yields the pooled spatial estimator. The corresponding predictor is labelled the pooled spatial predictor.

If the fixed effects model with spatial autocorrelation is the true model, then the problem is to predict

$$y_{i,T+S} = x_{i,T+S}'\beta + \mu_i + \phi_{i,T+s} \tag{39}$$

with $\phi_{T+S} = \lambda W\phi_{T+S} + \upsilon_{T+s}$. Unlike the usual FE case, $\lambda \neq 0$ and the μ_i's and β have to be estimated from MLE, i.e., using the FE-spatial estimates. The disturbance vector can be written as $\phi = (I_T \otimes B^{-1})\upsilon$, so that $\omega = E(\phi_{i,T+S}\phi) = 0$ since the υ's are not serially correlated over time. So the BLUP for this model looks like that for the FE model without spatial correlation except that the μ_i's and β are estimated assuming $\lambda \neq 0$. The corresponding predictor is labelled the FE-spatial predictor.

Baltagi and Li (2004) consider the problem of prediction in a panel data regression model with spatial autocorrelation in the context of a simple demand equation for cigarettes. This is based on a panel of 46 states over the period 1963-1992. The spatial autocorrelation due to neighboring states and the individual heterogeneity across states is taken explicitly into account. They compare the performance of several predictors of the states demand for cigarettes for one year and five years ahead. The estimators whose predictions are compared include OLS, fixed effects ignoring spatial correlation, fixed effects with spatial correlation, random effects GLS estimator ignoring spatial correlation and random effects estimator accounting for the spatial correlation. Based on RMSE forecast performance, estimators that take into account spatial correlation and heterogeneity across the states perform the best. The FE-spatial estimator gives the lowest RMSE for the first four years and is only surpassed by the RE-spatial in the fifth year. Overall, both the RE-spatial and FE-spatial estimators perform well in predicting cigarette demand.

For examples of prediction of random effects in a spatial generalized linear mixed model, see Zhang (2002) who applied this technique to disease mapping of plant roots on a 90 acre farm in Washington state. In many applications in epidemiology, ecology and agriculture, predicting the random effects of disease at unsampled sites requires modeling the spatial dependence continuously. This is especially important for data observed at point locations, where interpolation is needed to predict values at unsampled sites. Zhang implements this minimum mean squared error prediction through the Metropolis-Hastings algorithm.

3 Heterogenous Panels

For panel data studies with large N and small T, it is usual to pool the observations, assuming homogeneity of the slope coefficients. The latter is a testable assumption which is quite often rejected. Moreover, with the increasing time dimension of panel data sets, some researchers including Robertson and Symons (1992) and Pesaran and Smith (1995) have questioned the poolability of the data across heterogeneous units. Instead, they argue in favor of heterogeneous estimates that can be combined to obtain homogeneous estimates if the need arises. To make this point, Robertson and Symons (1992) studied the properties of some panel data estimators when the regression coefficients vary across individuals, i.e., they are *heterogeneous* but are assumed *homogeneous* in estimation. This is done for both stationary and nonstationary regressors. The basic conclusion is that severe biases can occur in dynamic estimation even for relatively small parameter variation.

Pesaran and Smith (1995) consider the problem of estimating a dynamic panel data model when the parameters are individually heterogeneous and illustrate their results by estimating industry-specific UK labor demand functions. In this case the model is given by

$$y_{it} = \lambda_i y_{i,t-1} + \beta_i x_{it} + u_{it} \quad i = 1, \ldots, N \quad t = 1, \ldots, T \tag{40}$$

where λ_i is IID$(\lambda, \sigma_\lambda^2)$ and β_i is IID(β, σ_β^2). Further λ_i and β_i are independent of y_{is}, x_{is} and u_{is} for all s. The objective in this case is to obtain consistent estimates of the mean values of λ_i and β_i. Pesaran and Smith (1995) present four different estimation procedures:

(1) aggregate time-series regressions of group averages;

(2) cross-section regressions of averages over time;

(3) pooled regressions allowing for fixed or random intercepts, or

(4) separate regressions for each group, where coefficients estimates are averaged over these groups.

They show that when T is small (even if N is large), all the procedures yield inconsistent estimators.

When both N and T are large, Pesaran and Smith (1995) show that the cross-section regression procedure will yield consistent estimates of the mean values of λ and β.

Maddala, Trost, Li, and Joutz (1997) applied classical, empirical Bayes and Bayesian procedures to the problem of estimating short-run and long-run elasticities of residential demand for electricity and natural gas in the U.S. for 49 states over 21 years (1970-1990). Since the elasticity estimates for each state were the ultimate goal of their study they were faced with three alternatives. The first is to use individual time series regressions for each state. These gave bad results, were hard to interpret, and had several wrong signs. The second option was to pool the data and use panel data estimators. Although the pooled estimates gave the right signs and were more reasonable, Maddala, Trost, Li, and Joutz (1997) argued that these estimates were not valid because the hypothesis of homogeneity of the coefficients was rejected. The third option, which they recommended, was to allow for some (but not complete) heterogeneity or (homogeneity). This approach lead them to their preferred shrinkage estimator which gave them more reasonable parameter estimates.

In the context of dynamic demand for gasoline across 18 OECD countries over the period 1960-1990, Baltagi and Griffin (1997) argued for pooling the data as the best approach for obtaining reliable price and income elasticities. They also pointed out that pure cross-section studies cannot control for unobservable country effects, whereas pure time-series studies cannot control for unobservable oil shocks or behavioral changes occurring over time. Baltagi and Griffin (1997) compared the homogeneous and heterogeneous estimates in the context of gasoline demand based on the plausibility of the price and income elasticities as well as the speed of adjustment path to the long-run equilibrium. They found considerable variability in the parameter estimates among the heterogeneous estimators some giving implausible estimates, while the homogeneous estimators gave similar plausible short-run estimates that differed only in estimating the long-run effects. Baltagi and Griffin (1997) also compared the forecast performance of these homogeneous and heterogeneous estimators over one, five and ten years horizon. Their findings show that the homogeneous estimators outperformed their heterogeneous counterparts based on mean squared forecast error. This result was replicated using a panel data set of 21

French regions over the period 1973-1998 by Baltagi, Bresson, Griffin and Pirotte (2003). Unlike the international OECD gasoline data set, the focus on the inter-regional differences in gasoline prices and income within France posed a different type of data set for the heterogeneity versus homogeneity debate. The variation in these prices and income were much smaller than international price and income differentials. This in turn reduces the efficiency gains from pooling and favor the heterogeneous estimators, especially given the differences between the Paris region and the rural areas of France. Baltagi, Bresson, Griffin and Pirotte (2003) showed that the time series estimates for each region are highly variable, unstable and offer the worst out of sample forecasts. Despite the fact that the shrinkage estimators proposed by Maddala, Trost, Li and Joutz (1997) outperformed these individual heterogeneous estimates, they still had a wide range and were outperformed by the homogeneous estimators in out of sample forecasts. Baltagi, Griffin and Xiong (2000) carried out this comparison for a dynamic demand for cigarettes across 46 U.S. states over 30 years (1963-1992). Once again the results showed that the homogeneous panel data estimators beat the heterogeneous and shrinkage type estimators in RMSE performance for out-of-sample forecasts. In another application, Driver, Imai, Temple and Urga (2004) utilize the Confederation of British Industry's (CBI) survey data to measure the impact of uncertainty on UK investment authorizations. The panel consists of 48 industries observed over 85 quarters 1978(Q1) to 1999(Q1). The uncertainty measure is based on the dispersion of beliefs across survey respondents about the general business situation in their industry. The heterogeneous estimators considered are OLS and 2SLS at the industry level, as well as the unrestricted SUR estimation method. Fixed effects, random effects, pooled 2SLS and restricted SUR are the homogeneous estimators considered. The panel estimates find that uncertainty has a negative, non-negligible effect on investment, while the heterogeneous estimates vary considerably across industries. Forecast performance for 12 out of sample quarters 1996(Q2) to 1999(Q1) are compared. The pooled homogeneous estimators outperform their heterogeneous counterparts in terms of RMSE.

Baltagi, Bresson and Pirotte (2002) reconsidered the two U.S. panel data sets on residential electricity and natural-gas demand used by Maddala, Trost, Li and Joutz (1997) and compared the out of sample forecast performance of the homogeneous, heterogeneous and shrinkage estimators. Once again the results show that when the data is used to estimate heterogeneous models across states, individual estimates offer the worst out-of-sample forecasts. Despite the fact that shrinkage estimators outperform these individual estimates, they are outperformed by simple homogeneous panel data estimates in out-of-sample forecasts. Admittedly, these are additional case studies, but they do add to the evidence that simplicity and parsimony in model estimation offered by the homogeneous estimators yield better forecasts than the more parameter consuming heterogeneous estimators.

Hsiao and Tahmiscioglu (1997) use a panel of 561 U.S. firms over the period 1971-92 to study the influence of financial constraints on company investment. They find substantial differences across firms in terms of their investment behavior. When a homogeneous pooled model is assumed, the impact of liquidity on firm investment is seriously underestimated. The authors recommend a mixed fixed and random coefficients framework based on the recursive predictive density criteria.

Baltagi, Bresson and Pirotte (2004) reconsider the Tobin q investment model studied by Hsiao, Pesaran and Tahmiscioglu (1999) using a slightly different panel of 337 U.S. firms over the period 1982-1998. They contrast the out of sample forecast performance of 9 homogeneous panel data estimators and 11 heterogeneous and shrinkage Bayes estimators over a 5 year horizon. Results show that the average heterogeneous estimators perform the worst in terms of mean squared error, while the hierarchical Bayes estimator suggested by Hsiao, Pesaran and Tahmiscioglu (1999) performs the best. Homogeneous panel estimators and iterative Bayes estimators are a close second.

In summary, while the performance of various estimators and their corresponding forecasts may vary in ranking from one empirical example to another, the consistent finding in all these studies is that homogeneous panel data estimators perform well in forecast performance mostly due to their simplicity, their parsimonious representation and the stability of the parameter estimates. Average heterogeneous estimators perform badly due to parameter estimate instability caused by the estimation of several parameters with short time series. Shrinkage estimators did well for some applications, especially iterative Bayes and iterative empirical Bayes.

Rapach and Wohar (2004) show that the monetary model of exchange rate determination performs poorly on a country by country basis for U.S. dollar exchange rates over the post-Bretton Woods period for 18 industrialized countries for quarterly data over the period 1973:1-1997:1. However, they find considerable support for the monetary model using panel procedures. They reject tests for the homogeneity assumptions inherent in panel procedures. Hence, they are torn between obtaining panel cointegrating coefficient estimates that are much more plausible in economic terms than country-by-country estimates. Yet these estimates might be spurious since they are rejected by formal statistical test for pooling. Rapach and Wohar (2004) perform an out-of-sample forecasting exercise using the panel and country-by-country estimates employing the RMSE criteria for a 1, 4, 8, 12 and 16 step ahead quarters. For the 1-step and 4-step ahead, the RMSEs of the homogeneous and heterogeneous estimates are similar. At the 8-step ahead horizon, homogeneous estimates generate better forecasts in comparison to five of the six heterogeneous estimates. At the 16-step horizon, the homogeneous estimates have RMSE that is smaller than each of the heterogeneous estimates. In most cases the RMSE is reduced by 20%. They conclude that while there are good reasons to favor the panel estimates over the country-by country estimates of the monetary model, there are also good reasons to be suspicious of these panel estimates. Other papers in this vein are Mark and Sul (2001) and Groen (2005). The latter paper utilizes a panel of vector error-correction models based on a common long-run relationship to test whether the Euro exchange rates of Canada, Japan and the United States have a long-run link with monetary fundamentals. Out of sample forecasts show that this common long-run exchange model is superior to both the naive random walk based forecasts and the standard cointegrated VAR model based forecasts, especially for horizons of 2 to 4 years.

Hoogstrate, Palm and Pfann (2000) investigate the improvement of forecasting performance using pooling techniques instead of single country forecasts for N fixed and T large. They use a set of dynamic regression equations with contemporaneously correlated disturbances. When the parameters of the models are different but exhibit some similarity, pooling may lead to a reduction in the mean squared error of the estimates and the forecasts. They show that the superiority of the pooled forecasts in small samples can deteriorate as T grows. They apply these results to growth rates of 18 OECD countries over the period 1950-1991 using an AR(3) model and an AR(3) model with leading indicators put forth by Garcia-Ferrer et al. (1987) and Zellner and Hong (1989). They find that the median MSFE of OLS based pooled forecasts is smaller than that of OLS based individual forecasts and that a fairly large T is needed for the latter to outperform the former. They argue that this is due to the reduction in MSE due to imposing a false restriction (pooling). However, for a large enough T, the bias of the pooled estimates increase with out bound and the resulting forecasts based on unrestricted estimates will outperform the forecasts based on the pooled restricted estimates.

Gavin and Theodorou (2005) use forecasting criteria to examine the macrodynamic behavior of 15 OECD countries observed quarterly over the period 1980 to 1996. They utilize a small set of familiar, widely used core economic variables, (output, price level, interest rates and exchange rates), omitting country-specific shocks. They find that this small set of variables and a simple VAR common model strongly support the hypothesis that many industrialized nations have similar macroeconomic dynamics. In sample, they often reject the hypothesis that coefficient vectors estimated separately for each country are the same. They argue that these rejections may be of little importance if due to idiosyncratic events since macro-time series are typically too short for standard methods to eliminate the effects of idiosyncratic factors. Panel data can be used to exploit the heterogeneous information in cross-country data, hence increasing the data and eliminating the idiosyncratic effects. They compare the forecast accuracy of the individual country models with the common models in a simulated out of sample experiment. They calculate four forecasts with increasing horizons at each point in time-one quarter ahead and four quarters ahead. For the four equations, at every horizon, the panel forecasts are significantly more accurate more often than are the individual country model forecasts. The biggest difference are for the exchange rate and the interest rate. They conclude that the superior out of sample forecasting performance of the common model supports their hypothesis that market economies tend to have a common macrodynamic patterns related to a small number of variables.

For other uses of forecasting with panel data, see Fok, et al. (2005) who show that forecasts of aggregates like total output or unemployment can be improved by considering panel models of disaggregated series covering 48 states. They use a panel version of a two-regime smooth transition autoregressive [STAR] type model to capture the non-linear features that are often displayed by macroeconomic variables allowing the parameters that govern the regime-switching to differ across states. Also, Mouchart and Rombouts (2005) who use a clustering approach to the usual panel data model specification to nowcast from poor data, namely, very short time series and many missing values. Marcelino, et al. (2003) who consider a similar problem of forecasting from panel data with severe deficiencies. Using an array of forecasting models applied to eleven countries originally in the EMU, over the period 1982-1997, at both the monthly and quarterly level, they show that forecasts constructed by aggregating the country-specific models are more accurate than forecasts constructed using the aggregate data.

3.1 Future Work

Much work remains to be done in forecasting with panels. This brief survey did not cover forecasting with Panel VAR methods which are popular in macroeconomics, see Ballabriga, et al. (1998) and Canova and Ciccarelli (2004), and Pesaran, et al. (2004), to mention a few. Canova and Ciccarelli (2004) provide methods for forecasting variables and predicting turning points in panel Bayesian VARs. They allow for interdependencies in the cross section as well as time variations in the parameters. Posterior distributions are obtained for hierarchical and for Minnesota-type priors and multi-step, multiunit point and average forecasts for the growth rate of output in the G7 are provided. There is also the problem of forecasting with nonstationary panels, see Pesaran and Breitung (2005) for a survey of nonstationary panels and Binder, Pesaran and Hsiao (2005) for estimation and inference in short panel vector autoregressions with unit roots and cointegration. This survey has not done justice to the Bayesian literature on forecasting and how it can improve forecasts using panels, see Zellner and Hong (1989), Zellner, Hong and Min (1991), Nandram and Petrucelli (1997), Koop and Potter (2003) and Canova and Ciccarelli (2004) to mention a few. For forecasting with micropanels, see Chamberlain and Hirano (1999) who suggested optimal ways of combining an individual's personal earnings history with panel data on the earnings trajectories of other individuals to provide a conditional distribution for this individual's earnings. This survey does not get into the large literature on "forecast combination methods", see Diebold and Lopez (1996) and Newbold and Harvey (2002), which can serve as a good spring board to launch research in improving forecasting methods using panels. For example, Stock and Watson (2004), who used forecast combination methods to forecast output growth in a seven-country quarterly economic data set covering 1959-1999 using up to 73 predictors per country. Hopefully, this paper will encourage more work on evaluation of panel models using post-sample forecasting a la Diebold and Mariano (1995) and Granger and Huang (1997).

4 References

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

Baillie, R.T. and B.H. Baltagi, 1999, Prediction from the regression model with one-way error components, Chapter 10 in C. Hsiao, K. Lahiri, L.F. Lee and H. Pesaran, eds., Analysis of Panels and Limited Dependent Variable Models (Cambridge University Press, Cambridge), 255–267..

Ballabriga, F.C., M. Sebastian and J. Valles, 1998, European asymmetries, Journal of International Economics 4, 233-253.

Baltagi, B.H., 2005. Econometric Analysis of Panel Data, Wiley and Sons, Chichester.

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

Baltagi, B.H. and D. Li, 2004, Prediction in the panel data model with spatial correlation, Chapter 13 in L. Anselin, R.J.G.M. Florax and S.J. Rey, eds., Advances in Spatial Econometrics: Methodology, Tools and Applications (Springer, Berlin), 283–295.

Baltagi, B.H. and Q. Li, 1992, Prediction in the one-way error component model with serial correlation, Journal of Forecasting 11, 561–567.

Baltagi, B.H. and Q. Li, 1994, Estimating error component models with general MA (q) disturbances, Econometric Theory 10, 396–408.

Baltagi, B.H. and P.X. Wu, 1999, Unequally spaced panel data regressions with AR (1) disturbances, Econometric Theory 15, 814–823.

Baltagi, B.H., G. Bresson and A. Pirotte, 2002, Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: Some empirical evidence from US electricity and natural-gas consumption, Economics Letters 76, 375-382.

Baltagi, B.H., G. Bresson and A. Pirotte, 2004, Tobin q: forecast performance for hierarchical Bayes, shrinkage, heterogeneous and homogeneous panel data estimators, Empirical Economics 29, 107-113.

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

Baltagi, B.H., G. Bresson, J.M. Griffin and A. Pirotte, 2003, Homogeneous, heterogeneous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption, Empirical Economics 28, 795-811.

Battese, G.E. and T.J. Coelli, 1988, Prediction of firm level technical efficiencies with a generalized frontier production function and panel data, Journal of Econometrics 38, 387–399.

Battese, G.E., Harter, R.M. and W.A. Fuller, 1988, An error component model for prediction of county crop areas using survey and satellite data. Journal of the American Statistical Association 83, 28-36.

Binder, M., C. Hsiao and M.H. Pesaran, 2005, Estimation and inference in short panel vector autoregressions with unit roots and cointegration, Econometric Theory 21, 795-837.

Breitung, J. and M. H. Pesaran, 2005, Unit roots and cointegration in panels, forthcoming in the Econometrics of Panel Data: A Handbook of the Theory with Applications, Laszlo Matyas and Patrick Sevestre, editors, Kluwer Academic Publishers.

Canova, F. and M. Ciccarelli, 2004, Forecasting and turning point predictions in a Bayesian panel VAR model, Journal of Econometrics 120, 327-59 .

Chamberlain, G. and K. Hirano, 1999, Predictive distributions based on longitudinal earnings data, Annales D'Économie et de Statistique 55–56, 211–242.

Diebold, F.X. and J.A. Lopez, 1996, Forecast evaluation and combination. In Handbook of Statistics, Maddala, G.S. and C.R. Rao, eds., North-Holland: Amsterdam.

Diebold, F.X. and R.S. Mariano, 1995, Comparing predictive accuracy, Journal of Business and Economic Statistics 13, 253-264.

Driver, C., K. Imai, P. Temple and A. Urga, 2004, The effect of uncertainty on UK investment authorisation: homogeneous vs. heterogeneous estimators, Empirical Economics 29, 115-128.

Fok, D. , D. van Dijk, and P. H. Franses, 2005, Forecasting aggregates using panels of nonlinear time series, International Journal of Forecasting 21, 785-794.

Frees, E.W., V. Young and Y. Luo, 1999, A longitudinal data analysis interpretation of credibility models, Insurance: Mathematics and Economics 24, 229–247.

Frees, E.W., V. Young and Y. Luo, 2001, Credibility ratemaking using panel data models, North American Actuarial Journal 5, 24-42.

Frees, E.W. and T.W. Miller, 2004, Sales forecasting using longitudinal data models. International Journal of Forecasting 20, 99–114.

Fuller, W.A. and G.E. Battese, 1974, Estimation of linear models with cross-error structure, Journal of Econometrics 2, 67–78.

Galbraith, J.W. and V. Zinde-Walsh, 1995, Transforming the error-component model for estimation with general ARMA disturbances, Journal of Econometrics 66, 349–355.

Garcia-Ferrer, A., R.A. Highfield, F. Palm and A. Zellner, 1987, Macroeconomic forecasting using pooled international data, Journal of Business and Economic Statistics 5, 53-76.

Gavin, W.T. and A.T. Theodorou, 2005, A common model approach to macroeconomics: using panel data to reduce sampling error, Journal of Forecasting 24, 203-219.

Goldberger, A.S., 1962, Best linear unbiased prediction in the generalized linear regression model, Journal of the American Statistical Association 57, 369–375.

Granger, C.W.J. and L. Huang, 1997, Evaluations of panel data models: some suggestions from time series, discussion paper 97-10, University of California, San Diego.

Groen, J.J.J., 2005, Exchange rate predictability and monetary fundamentals in a small multi-country panel, Journal of Money, Credit, and Banking 37, 495-516.

Harville, D.A., 1976, Extension of the Gauss-Markov theorem to include the estimation of random effects, Annals of Statistics 4, 384-395.

Harville, D.A., 1985, Decomposition of prediction error, Journal of the American Statistical Association 80, 132-138.

Henderson, C.R., 1975, Best linear unbiased estimation and prediction under a selection model, Biometrics 31, 423-447.

Holtz-Eakin, D. and T.M. Selden, 1995, Stocking the fires? CO_2 emissions and economic growth, Journal of Public Economics 57, 85-101.

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

Hsiao, C. and A.K. Tahmiscioglu, 1997, A panel analysis of liquidity constraints and firm investment, Journal of the American Statistical Association 92, 455–465.

Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu, 1999, Bayes estimation of short run coefficients in dynamic panel data models, Chapter 11 in C. Hsiao, K. Lahiri, L.F. Lee and M.H. Pesaran, eds., Analysis of Panels and Limited Dependent Variable Models (Cambridge University Press, Cambridge), 268–296.

Kackar, R.N. and D. Harville, 1984, Approximations for standard errors of estimators of fixed and random effects in mixed linear models, Journal of the American Statistical Association 79, 853-862.

Karlsson, S. and J. Skoglund, 2004, Maximum-likelihood based inference in the two-way random effects model with serially correlated time effects, Empirical Economics 29, 79–88.

Koop, G. and S. Potter, 2003, Forecasting in large macroeconomic panels using Bayesian model averaging, Federal Reserve Bank of New York Staff Reports, no. 163.

MaCurdy, T.A., 1982, The use of time series processes to model the error structure of earnings in a longitudinal data analysis, Journal of Econometrics 18, 83–114.

Maddala, G.S., R.P. Trost, H. Li and F. Joutz, 1997, Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators, Journal of Business and Economic Statistics 15, 90–100.

Marcelino, M., J. H. Stock and M. Watson, 2003, Macroeconomic forecasting in the EURO area: country specific versus area-wide information, European Economic Review 47, 1-18.

Mark, N.C. and D. Sul, 2001, Nominal exchange rates and monetary fundamentals; evidence from a small post-Bretton Woods panel, Journal of International Economics 53, 29-52.

Mouchart, M. and J.V.K. Rombouts, 2005, Clustered panel data models: An efficient approach for nowcasting from poor data, International Journal of Forecasting 21, 577-594.

Nandram, B., and J.D. Petruccelli, 1997, A Bayesian analysis of autoregressive time series panel data, Journal of Business and Economic Statistics 15, 328-334.

Newbold, P. and D.I. Harvey, 2002, Forecast combination and encompassing. In A Companion to Economic Forecasting, Clements, M.P. and D.F. Hendry (eds.). Blackwell: Oxford, 268-283.

Pesaran, M.H. and R. Smith, 1995, Estimating long-run relationships from dynamic heterogenous panels, Journal of Econometrics 68, 79–113.

Pesaran, M.H., R. Smith and K.S. Im, 1996, Dynamic linear models for heterogenous panels, Chapter 8 in L. Mátyás and P. Sevestre, eds., The Econometrics of Panel Data: A Handbook of the Theory With Applications (Kluwer Academic Publishers, Dordrecht), 145–195.

Pesaran, M.H., T. Schuermann and S. Weiner, 2004, Modelling regional interdependencies using a global error-correcting macroeconometric model, Journal of Business and Economics Statistics 22, 129-162 .

Rapach, D.E. and M.E. Wohar, 2004, Testing the monetary model of exchange rate determination: a closer look at panels, Journal of International Money and Finance 23, 867–895.

Revankar, N.S., 1979, Error component models with serial correlated time effects, Journal of the Indian Statistical Association 17, 137–160.

Robinson, G.K., 1991, That BLUP is a good thing: the estimation of random effects, Statistical Science 6, 15-32.

Schmalensee, R., T.M. Stoker and R.A. Judson, 1998, World carbon dioxide emissions: 1950-2050, Review of Economics and Statistics 80, 15–27.

Stock, J.H. and M.W. Watson, 2004, Combination forecasts of output growth ina seven-country data set, Journal of Forecasting 23, 405-430.

Taub, A.J., 1979, Prediction in the context of the variance-components model, Journal of Econometrics 10, 103–108.

Zellner, A., and C. Hong , 1989, Forecasting international growth rates using Bayesian shrinkage and other procedures, Journal of Econometrics 40, 183-202.

Zellner, A., C. Hong and C. Min, 1991, Forecasting turning points in international output growth rates using Bayesian exponentially weighted autoregression, time-varying parameters, and pooling techniques, Journal of Econometrics 49, 275-304.

Zhang, H., 2002, On estimation and prediction for spatial generalized linear mixed models, Biometrics 58, 129-136.

The following Discussion Papers have been published since 2005:

Series 1: Economic Studies

1	2005	Financial constraints and capacity adjustment in the United Kingdom – Evidence from a large panel of survey data	Ulf von Kalckreuth Emma Murphy
2	2005	Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model	Sandra Eickmeier
3	2005	Financial intermediaries, markets, and growth	F. Fecht, K. Huang, A. Martin
4	2005	The New Keynesian Phillips Curve in Europe: does it fit or does it fail?	Peter Tillmann
5	2005	Taxes and the financial structure of German inward FDI	Fred Ramb A. J. Weichenrieder
6	2005	International diversification at home and abroad	Fang Cai Francis E. Warnock
7	2005	Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States	Wolfgang Keller Steven R. Yeaple
8	2005	Location choice and employment decisions: a comparison of German and Swedish multinationals	S. O. Becker, K. Ekholm, R. Jäckle, M.-A. Muendler
9	2005	Business cycles and FDI: evidence from German sectoral data	Claudia M. Buch Alexander Lipponer
10	2005	Multinational firms, exclusivity, and the degree of backward linkages	Ping Lin Kamal Saggi

11	2005	Firm-level evidence on international stock market comovement	Robin Brooks Marco Del Negro
12	2005	The determinants of intra-firm trade: in search for export-import magnification effects	Peter Egger Michael Pfaffermayr
13	2005	Foreign direct investment, spillovers and absorptive capacity: evidence from quantile regressions	Sourafel Girma Holger Görg
14	2005	Learning on the quick and cheap: gains from trade through imported expertise	James R. Markusen Thomas F. Rutherford
15	2005	Discriminatory auctions with seller discretion: evidence from German treasury auctions	Jörg Rocholl
16	2005	Consumption, wealth and business cycles: why is Germany different?	B. Hamburg, M. Hoffmann, J. Keller
17	2005	Tax incentives and the location of FDI: evidence from a panel of German multinationals	Thiess Buettner Martin Ruf
18	2005	Monetary Disequilibria and the Euro/Dollar Exchange Rate	Dieter Nautz Karsten Ruth
19	2005	Berechnung trendbereinigter Indikatoren für Deutschland mit Hilfe von Filterverfahren	Stefan Stamfort
20	2005	How synchronized are central and east European economies with the euro area? Evidence from a structural factor model	Sandra Eickmeier Jörg Breitung
21	2005	Asymptotic distribution of linear unbiased estimators in the presence of heavy-tailed stochastic regressors and residuals	J.-R. Kurz-Kim S.T. Rachev G. Samorodnitsky

22	2005	The Role of Contracting Schemes for the Welfare Costs of Nominal Rigidities over the Business Cycle	Matthias Paustian
23	2005	The cross-sectional dynamics of German business cycles: a bird's eye view	J. Döpke, M. Funke S. Holly, S. Weber
24	2005	Forecasting German GDP using alternative factor models based on large datasets	Christian Schumacher
25	2005	Time-dependent or state-dependent price setting? – micro-evidence from German metal-working industries –	Harald Stahl
26	2005	Money demand and macroeconomic uncertainty	Claus Greiber Wolfgang Lemke
27	2005	In search of distress risk	J. Y. Campbell, J. Hilscher, J. Szilagyi
28	2005	Recursive robust estimation and control without commitment	Lars Peter Hansen Thomas J. Sargent
29	2005	Asset pricing implications of Pareto optimality with private information	N. R. Kocherlakota Luigi Pistaferri
30	2005	Ultra high frequency volatility estimation with dependent microstructure noise	Y. Aït-Sahalia, P. A. Mykland, L. Zhang
31	2005	Umstellung der deutschen VGR auf Vorjahres-preisbasis – Konzept und Konsequenzen für die aktuelle Wirtschaftsanalyse sowie die ökono-metrische Modellierung	Karl-Heinz Tödter

32	2005	Determinants of current account developments in the central and east European EU member states – consequences for the enlargement of the euro erea	Sabine Herrmann Axel Jochem
33	2005	An estimated DSGE model for the German economy within the euro area	Ernest Pytlarczyk
34	2005	Rational inattention: a research agenda	Christopher A. Sims
35	2005	Monetary policy with model uncertainty: distribution forecast targeting	Lars E.O. Svensson Noah Williams
36	2005	Comparing the value revelance of R&D reporting in Germany: standard and selection effects	Fred Ramb Markus Reitzig
37	2005	European inflation expectations dynamics	J. Döpke, J. Dovern U. Fritsche, J. Slacalek
38	2005	Dynamic factor models	Sandra Eickmeier Jörg Breitung
39	2005	Short-run and long-run comovement of GDP and some expenditure aggregates in Germany, France and Italy	Thomas A. Knetsch
40	2005	A"wreckers theory" of financial distress	Ulf von Kalckreuth
41	2005	Trade balances of the central and east European EU member states and the role of foreign direct investment	Sabine Herrmann Axel Jochem
42	2005	Unit roots and cointegration in panels	Jörg Breitung M. Hashem Pesaran
43	2005	Price setting in German manufacturing: new evidence from new survey data	Harald Stahl

1	2006	The dynamic relationship between the Euro overnight rate, the ECB's policy rate and the term spread	Dieter Nautz Christian J. Offermanns
2	2006	Sticky prices in the euro area: a summary of new micro evidence	Álvarez, Dhyne, Hoeberichts Kwapil, Le Bihan, Lünnemann Martins, Sabbatini, Stahl Vermeulen, Vilmunen
3	2006	Going multinational: What are the effects on home market performance?	Robert Jäckle
4	2006	Exports versus FDI in German manufacturing: firm performance and participation in international markets	Jens Matthias Arnold Katrin Hussinger
5	2006	A disaggregated framework for the analysis of structural developments in public finances	Kremer, Braz, Brosens Langenus, Momigliano Spolander
6	2006	Bond pricing when the short term interest rate follows a threshold process	Wolfgang Lemke Theofanis Archontakis
7	2006	Has the impact of key determinants of German exports changed? Results from estimations of Germany's intra euro-area and extra euro-area exports	Kerstin Stahn
8	2006	The coordination channel of foreign exchange intervention: a nonlinear microstructural analysis	Stefan Reitz Mark P. Taylor
9	2006	Capital, labour and productivity: What role do they play in the potential GDP weakness of France, Germany and Italy?	Antonio Bassanetti Jörg Döpke, Roberto Torrini Roberta Zizza
10	2006	Real-time macroeconomic data and ex ante predictability of stock returns	J. Döpke, D. Hartmann C. Pierdzioch

11	2006	The role of real wage rigidity and labor market frictions for unemployment and inflation dynamics	Kai Christoffel Tobias Linzert
12	2006	Forecasting the price of crude oil via convenience yield predictions	Thomas A. Knetsch
13	2006	Foreign direct investment in the enlarged EU: do taxes matter and to what extent?	Guntram B. Wolff
14	2006	Inflation and relative price variability in the euro area: evidence from a panel threshold model	Dieter Nautz Juliane Scharff
15	2006	Internalization and internationalization under competing real options	Jan Hendrik Fisch
16	2006	Consumer price adjustment under the microscope: Germany in a period of low inflation	Johannes Hoffmann Jeong-Ryeol Kurz-Kim
17	2006	Identifying the role of labor markets for monetary policy in an estimated DSGE model	Kai Christoffel Keith Küster Tobias Linzert
18	2006	Do monetary indicators (still) predict euro area inflation?	Boris Hofmann
19	2006	Fool the markets? Creative accounting, fiscal transparency and sovereign risk premia	Kerstin Bernoth Guntram B. Wolff
20	2006	How would formula apportionment in the EU affect the distribution and the size of the corporate tax base? An analysis based on German multinationals	Clemens Fuest Thomas Hemmelgarn Fred Ramb

21	2006	Monetary and fiscal policy interactions in a New Keynesian model with capital accumulation and non-Ricardian consumers	Campbell Leith Leopold von Thadden
22	2006	Real-time forecasting and political stock market anomalies: evidence for the U.S.	Martin Bohl, Jörg Döpke Christian Pierdzioch
23	2006	A reappraisal of the evidence on PPP: a systematic investigation into MA roots in panel unit root tests and their implications	Christoph Fischer Daniel Porath
24	2006	Margins of multinational labor substitution	Sascha O. Becker Marc-Andreas Mündler
25	2006	Forecasting with panel data	Badi H. Baltagi

Series 2: Banking and Financial Studies

1	2005	Measurement matters – Input price proxies and bank efficiency in Germany	Michael Koetter
2	2005	The supervisor's portfolio: the market price risk of German banks from 2001 to 2003 – Analysis and models for risk aggregation	Christoph Memmel Carsten Wehn
3	2005	Do banks diversify loan portfolios? A tentative answer based on individual bank loan portfolios	Andreas Kamp Andreas Pfingsten Daniel Porath
4	2005	Banks, markets, and efficiency	F. Fecht, A. Martin
5	2005	The forecast ability of risk-neutral densities of foreign exchange	Ben Craig Joachim Keller
6	2005	Cyclical implications of minimum capital requirements	Frank Heid
7	2005	Banks' regulatory capital buffer and the business cycle: evidence for German savings and cooperative banks	Stéphanie Stolz Michael Wedow
8	2005	German bank lending to industrial and non-industrial countries: driven by fundamentals or different treatment?	Thorsten Nestmann
9	2005	Accounting for distress in bank mergers	M. Koetter, J. Bos, F. Heid C. Kool, J. Kolari, D. Porath
10	2005	The eurosystem money market auctions: a banking perspective	Nikolaus Bartzsch Ben Craig, Falko Fecht
11	2005	Financial integration and systemic risk	Falko Fecht Hans Peter Grüner

12	2005	Evaluating the German bank merger wave	Michael Koetter
13	2005	Incorporating prediction and estimation risk in point-in-time credit portfolio models	A. Hamerle, M. Knapp, T. Liebig, N. Wildenauer
14	2005	Time series properties of a rating system based on financial ratios	U. Krüger, M. Stötzel, S. Trück
15	2005	Inefficient or just different? Effects of heterogeneity on bank efficiency scores	J. Bos, F. Heid, M. Koetter, J. Kolatri, C. Kool
01	2006	Forecasting stock market volatility with macroeconomic variables in real time	J. Döpke, D. Hartmann C. Pierdzioch
02	2006	Finance and growth in a bank-based economy: is it quantity or quality that matters?	Michael Koetter Michael Wedow
03	2006	Measuring business sector concentration by an infection model	Klaus Düllmann
04	2006	Heterogeneity in lending and sectoral growth: evidence from German bank-level data	Claudia M. Buch Andrea Schertler Natalja von Westernhagen
05	2006	Does diversification improve the performance of German banks? Evidence from individual bank loan portfolios	Evelyn Hayden Daniel Porath Natalja von Westernhagen

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Among others under certain conditions visiting researchers have access to a wide range of data in the Bundesbank. They include micro data on firms and banks not available in the public. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt
GERMANY