7-2012

Estimation and Prediction in the Random Effects Model with AR(P) Remainder Disturbances

Badi Baltagi
Syracuse University, bbaltagi@maxwell.syr.edu

Long Liu
University of Texas at San Antonio

ISSN: 1525-3066

Center for Policy Research
Working Paper No. 138

ESTIMATION AND PREDICTION IN THE RANDOM EFFECTS MODEL WITH AR(*P*) REMAINDER DISTURBANCES

Badi H. Baltagi and Long Liu

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

July 2012

$5.00

CENTER FOR POLICY RESEARCH – Summer 2012

Leonard M. Lopoo, Director
Associate Professor of Public Administration and International Affairs

---

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

Douglas Wolf
Gerald B. Cramer Professor of Aging Studies
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi Baltagi .... Economics
Robert Bifulco .... Public Administration
Leonard Burman .. Public Administration/Economics
Thomas Dennison .... Public Administration
William Duncombe .... Public Administration
Gary Engelhardt .... Economics
Madonna Harrington Meyer .... Sociology
William C. Horrace .... Economics
Duke Kao .... Economics
Eric Kingson .... Social Work
Sharon Kioko…………………..Public Administration
Thomas Kniesner .... Economics
Jeffrey Kubik .... Economics
Andrew London .... Sociology
Len Lopoo .... Public Administration
Amy Lutz .... Sociology
Jerry Miner .... Economics
Jan Ondrich .... Economics
John Palmer .... Public Administration
David Popp .... Public Administration
Gretchen Purser .... Sociology
Christopher Rohlfs .... Economics
Stuart Rosenthal .... Economics
Ross Rubenstein .... Public Administration
Perry Singleton…………………………Economics
Margaret Usdansky .... Sociology
Michael Wasylenko .... Economics
Jeffrey Weinstein………………………Economics
Peter Wilcoxen .... Public Administration/Economics
Janet Wilmoth .... Sociology

GRADUATE ASSOCIATES

Kanika Arora .... Public Administration
Christopher Bianchi .... Public Administration
Christian Buerger .... Public Administration
Il Hwan Chung .... Public Administration
Alissa Dubnicki .... Economics
Alexander Falevich .... Economics
Andrew Friedson .... Economics
Pallab Ghosh .... Economics
Lincoln Groves .... Public Administration
Clorise Harvey .... Public Administration
Hee Seung Lee .... Public Administration
Jing Li .... Economics
Shimeng Liu .... Economics
Yan Liu .... Sociology
Allison Marier .... Economics
Qing Miao .... Public Administration
Gabrielle Mirsaidi .... Public Administration
Nuno Abreu Faro E Mota .... Economics
Wael Moussa .... Economics
Kerri Raissian .... Public Administration
Heather Ruby .... Public Administration
Natalee Simpson .... Sociology
Tian Tang .... Public Administration
Liu Tian .... Economics
Pengju Zhang .... Public Administration

STAFF

Kelly Bogart.....……………Administrative Specialist
Karen Cimilluca.....………………Office Coordinator
Alison Kirsche.....…………Administrative Secretary
Kathleen Nasto.................Administrative Secretary
Candi Patterson………………Computer Consultant
Mary Santy……..……………Administrative Secretary

Abstract

This paper considers the problem of estimation and forecasting in a panel data model with random individual effects and AR(p) remainder disturbances. It utilizes a simple exact transformation for the AR(p) time series process derived by Baltagi and Li (1994) and obtains the generalized least squares estimator for this panel model as a least squares regression. This exact transformation is also used in conjunction with Goldberger's (1962) result to derive an analytic expression for the best linear unbiased predictor. The performance of this predictor is investigated using Monte Carlo experiments and illustrated using an empirical example.

JEL No. C32

Key Words: Prediction; Panel Data; Random Effects; Serial Correlation; AR(p).

Estimation and Prediction in the Random Effects Model with AR(p) Remainder Disturbances

Badi H. Baltagi[*], Long Liu[†]

July 20, 2012

Abstract

This paper considers the problem of estimation and forecasting in a panel data model with random individual effects and AR(p) remainder disturbances. It utilizes a simple *exact* transformation for the AR(p) time series process derived by Baltagi and Li (1994) and obtains the generalized least squares estimator for this panel model as a least squares regression. This *exact* transformation is also used in conjunction with Goldberger's (1962) result to derive an analytic expression for the best linear unbiased predictor. The performance of this predictor is investigated using Monte Carlo experiments and illustrated using an empirical example.

Key Words: *Prediction; Panel Data; Random Effects; Serial Correlation; AR(p).*

1 Introduction

Forecasting with panel data has become an integral part of empirical work in economics and related fields. Some important applications include Schmalensee, Stoker and Judson (1998) who forecast world carbon dioxide emissions using national-level panel data; and Frees and Miller (2004) who forecast the sale of state lottery tickets using panel data on postal (ZIP) codes, to mention a few, see Baltagi (2008) for a recent survey. This paper deals with forecasting with panel data controlling for heterogeneity of individuals, countries or firms through the use of random effects, as well as dealing with serial correlation in the error term to allow for macro shocks in the economy, see Baltagi and Li (1991) and Frees and Miller (2004) for earlier work on this subject. In fact, Baltagi and Li (1991) suggested a simple transformation to estimate a panel data regression model with random individual effects, and AR(1), AR(2) or a specialized AR(4) process in the remainder disturbances. In a follow up paper, Baltagi and Li (1992) derived the corresponding best linear unbiased predictor (BLUP) for the ith individual in the panel, s periods ahead, extending results of Goldberger (1962)

[*]Address correspondence to: Badi H. Baltagi, Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020; e-mail: bbaltagi@maxwell.syr.edu.

[†]Long Liu: Department of Economics, College of Business, University of Texas at San Antonio, One UTSA Circle, TX 78249-0633; e-mail: long.liu@utsa.edu.

from the time series to the panel data case. This paper extends the estimation and forecasting in Baltagi and Li (1991, 1992) to the general AR(p) case. Although the exact transformation for AR(p) has been known for a long time, see Fuller (1976) for example, its explicit form for $p > 2$ is cumbersome and may be the reason why it is not popular among practitioners. In practice, empirical researchers used the Cochrane-Orcutt transformation despite the well known result that it can lead to a substantial loss in efficiency in finite samples. For the importance of the initial observations especially for trended economic data, and the inefficiency of the Cochrane-Orcutt estimator, see Maeshiro (1976, 1979) and Park and Mitchell (1980), to mention a few. Baltagi and Li (1994) derived a simple *exact* transformation for the AR(p) model which utilizes the auto-covariance structure of the autoregressive process. Based on this transformation, they proposed a GLS estimator for the *time series* case, requiring only least squares regressions and recursive computations. This paper utilizes the Baltagi and Li (1994) *exact* transformation for the AR(p) model in the time series context and apply it to a random effects panel data model with AR(p) remainder disturbances. With this simple transformation, one can generalize Baltagi and Li's (1991) result to the higher order AR(p) case, and provide an accompanying simple estimation method. This is of utmost importance in panel data regressions where the order of matrix inversion can be considerably reduced using this transformation. In addition, this simple transformation allows us to derive an explicit expression for the BLUP, thus extending the Baltagi and Li (1992) result to the AR(p) case. The next section gives the panel data model with AR(p) remainder disturbances and proposes a simple feasible GLS estimation method that can be computed using least squares, while section 3 provides a derivation of the Goldberger (1962) BLUP for this model. Section 4 provides some Monte Carlo results on the performance of these predictors, while Section 5 illustrates these predictors using the Wisconsin lottery sales example of Frees and Miller (2004).

2 Model and Estimation

Consider the following panel data regression model:

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

where y_{it} is the observation on the ith individual for the tth time period. x_{it} denotes the $k \times 1$ vector of observations on the nonstochastic regressors which are uncorrelated with the regression disturbances u_{it}. The disturbances follow a one-way error component model

$$u_{it} = \mu_i + v_{it}, \tag{2}$$

with random individual effects $\mu_i \sim i.i.d.\left(0, \sigma_\mu^2\right)$. The remainder disturbances v_{it} follow an AR(p) process given by $v_{it} = \rho_1 v_{i,t-1} + \rho_2 v_{i,t-2} + \cdots + \rho_p v_{i,t-p} + \epsilon_{it}$, where $\epsilon_{it} \sim i.i.d.\left(0, \sigma_\epsilon^2\right)$. $\rho_1, \rho_2, .., \rho_p$ are unknown parameters satisfying the stationarity condition that the roots of $1 - \rho_1 z - \rho_2 z^2 - \cdots - \rho_p z^p = 0$ all lie outside the complex unit circle, see Judge et. al. (1985). As shown in Brockwell and Davis (1991), this AR(p) process is a special case of the stationary p-dependent process defined as $E\left(v_{it} v_{i,t-s}\right) = \gamma_s$ with $\gamma_{-s} = \gamma_s$.

The model in (1) can be rewritten in matrix notation as

$$y = X\beta + u \tag{3}$$

where y is of dimension $NT \times 1$, X is $NT \times k$, β is $k \times 1$ and u is $NT \times 1$. The disturbance term can be written in vector form as

$$u = \left(I_N \otimes \iota_T\right) \mu + \nu, \tag{4}$$

where $\mu = (\mu_1, \ldots, \mu_N)$ and $v' = (v_{11}, \ldots, v_{1T}, \ldots, v_{N1}, \ldots, v_{NT})$. ι_T is a vector of ones of dimension T. I_T is an identity matrix of dimension T and $\otimes$ denotes the Kronecker product. The variance–covariance matrix of u can be written as

$$\Omega = I_N \otimes \Lambda, \tag{5}$$

where $\Lambda = \sigma_\mu^2 J_T + \sigma^2 V$, J_T is a matrix of ones of dimension T, and $E\left(v_i v_i'\right) = \sigma^2 V$ is the variance–covariance matrix of the remainder error term $v_i = (v_{i1}, \ldots, v_{iT})$ which is assumed to be the same for each individual. V is assumed to be a real symmetric positive-definite matrix and $\sigma^2 \equiv \gamma_0$. Hence, there exists a $T \times T$ matrix C, such that $CVC' = I_T$. To get rid of the serial correlation in the remainder disturbances, we premultiply the model in (3) by $(I_N \otimes C)$. The transformed error becomes

$$u^* = \left(I_N \otimes C\right) u = \left(I_N \otimes \iota_T^\alpha\right) \mu + \left(I_N \otimes C\right) \nu, \tag{6}$$

where $\iota_T^\alpha = C\iota_T = \left(\alpha_1, \ldots, \alpha_T\right)'$ is a $T \times 1$ vector whose elements depend on the specific serial correlation process imposed on v. The variance-covariance matrix for the transformed disturbance u^* becomes

$$\Omega^* = I_N \otimes \Lambda^*, \tag{7}$$

where

$$\Lambda^* = C\Lambda C' = \sigma_\mu^2 J_T^\alpha + \sigma^2 I_T, \tag{8}$$

and $J_T^\alpha = \iota_T^\alpha \iota_T^{\alpha\prime}$. This can be rewritten as

$$\Lambda^* = \sigma_\mu^2 d^2 \bar{J}_T^\alpha + \sigma^2 I_T, \tag{9}$$

where $d^2 = \iota_T^{\alpha\prime}\iota_T^\alpha = \sum_{t=1}^T \alpha_t^2$ and $\bar{J}_T^\alpha = J_T^\alpha/d^2$. Following a trick by Wansbeek and Kapteyn (1983), we replace I_T by $E_T^\alpha + \bar{J}_T^\alpha$, where $E_T^\alpha = I_T - \bar{J}_T^\alpha$. Collecting like terms, one gets the spectral decomposition of Λ^*:

$$\Lambda^* = \sigma_\alpha^2 \bar{J}_T^\alpha + \sigma^2 E_T^\alpha, \tag{10}$$

where $\sigma_\alpha^2 = \sigma_\mu^2 d^2 + \sigma^2$. Because $\bar{J}_T^\alpha$ and E_T^α are idempotent matrices that are orthogonal to each other, we have

$$\Lambda^{*p} = \left(\sigma_\alpha^2\right)^p \bar{J}_T^\alpha + \left(\sigma^2\right)^p E_T^\alpha, \tag{11}$$

where p is an arbitrary scalar. In particular, $p = -1$ obtains the inverse and $p = -1/2$ gives $\Lambda^{*-1/2}$ and hence $\Omega^{*-1/2}$. Therefore,

$$\sigma\Omega^{*-1/2} = \frac{\sigma}{\sigma_\alpha}\left(I_N \otimes \bar{J}_T^\alpha\right) + \left(I_N \otimes E_T^\alpha\right) = \left(I_N \otimes I_T^\alpha\right) - \delta\left(I_N \otimes \bar{J}_T^\alpha\right), \tag{12}$$

where $\delta = 1 - \frac{\sigma}{\sigma_\alpha}$. Thus we can premultiply the C-transformed model by $\sigma\Omega^{*-1/2}$ to make the error spherical. $y^{**} = \sigma\Omega^{*-1/2}y^*$, and X^{**} and u^{**} are similarly defined. The typical elements of are given by

$$y_{it}^{**} = y_{it}^* - \delta\alpha_t \frac{\sum_{s=1}^T \alpha_s y_{is}^*}{\sum_{s=1}^T \alpha_s^2}. \tag{13}$$

This is a generalized version of the the Fuller and Battese (1974) transformation for the error component model with an arbitrary variance-covariance matrix, $E\left(v_i v_i'\right) = \sigma^2 V$, on the remainder disturbances. The OLS regression on the (**) transformed equation is equivalent to the GLS regression on the original equation (4). Equation (11) also suggests natural estimators of the variance components. Baltagi and Li (1991) proposed estimating σ^2 and σ_α^2 by

$$\hat{\sigma}_\alpha^2 = u^{*\prime}\left(I_N \otimes \bar{J}_T^\alpha\right)u^*/N \text{ and } \hat{\sigma}^2 = u^{*\prime}\left(I_N \otimes E_T^\alpha\right)u^*/N\left(T-1\right). \tag{14}$$

These are best quadratic unbiased estimators of σ^2 and σ_α^2 if the true disturbances u^* are known. The true residuals are generally not known. In this case, one can replace u^* by $\tilde{u}_{OLS}^*$, the OLS residuals on the (*) transformed equation.

Baltagi and Li (1991) applied this simple transformation to the panel data model with random individual effects and AR(1), AR(2) or a specialized AR(4) process in the remainder disturbances. The general AR(p) process was not considered by Baltagi and Li (1991) because a simple transformation for AR(p) for $p > 2$ was not available. Utilizing the Baltagi and Li (1994) *exact* transformation for the AR(p) model in the time series context, we give a simple recursive transformation for the panel data model with remainder disturbances following an AR(p) process. Recall that $\gamma_s = E\left(v_{it}v_{i,t-s}\right)$, and let $r_s = \gamma_s/\gamma_0$. Following Baltagi and Li (1994), the (*) transformation defined in (6), is obtained recursively as follows:

$$\begin{aligned} y_{i1}^* &= y_{i1} \\ y_{it}^* &= \left(y_{it} - b_{t,t-1}y_{i,t-1}^* - \cdots - b_{t,1}y_{i,1}^*\right)/\sqrt{a_t} \quad &&\text{for } t = 2, \ldots, p \\ y_{it}^* &= \left(y_{it} - \rho_1 y_{i,t-1} - \cdots - \rho_p y_{i,t-p}\right)/\sqrt{a} &&\text{for } t = p+1, \ldots, T, \end{aligned} \tag{15}$$

where $a = \sigma_\epsilon^2/\gamma_0$ and a_t and $b_{t,s}$ are determined recursively as

$$a_t = 1 - b_{t,t-1}^2 - \cdots - b_{t,2}^2 - b_{t,1}^2 \quad \text{for } t = 2, \ldots, p \tag{16}$$

and

$$\begin{aligned} b_{t,1} &= r_{t-1} \\ b_{t,s} &= \left(r_{t-s} - b_{s,s-1}b_{t,s-1} - \cdots - b_{s,1}b_{t,1}\right)/\sqrt{a_s} \quad \text{for } s = 2, \ldots, t-1 \end{aligned} \tag{17}$$

for $t = 2 \ldots, p$.

By replacing y_{it}^* by α_t and replacing y_{it} by 1 in equation (15), we can get $\iota_T^\alpha = C\iota_T = (\alpha_1, \ldots, \alpha_T)'$ as follows:

$$\begin{aligned} \alpha_1 &= 1 \\ \alpha_t &= \left(1 - b_{t,t-1}\alpha_{t-1} - \cdots - b_{t,1}\alpha_1\right)/\sqrt{a_t} \quad &&\text{for } t = 2, \ldots, p \\ \alpha_t &= \left(1 - \textstyle\sum_{s=1}^p \rho_s\right)/\sqrt{a} &&\text{for } t = p+1, \ldots, T. \end{aligned} \tag{18}$$

The above transformation depends upon the auto-covariance function of v_{it}, that is, γ_s for $t = 1 \ldots, p$. In order to make this operational, we must get estimates of γ_s. Consistent estimates of γ_s can be obtained from

$$\hat{\gamma}_s = \sum_{i=1}^N \sum_{t=s+1}^T \frac{\tilde{v}_{it}\tilde{v}_{i,t-s}}{N(T-s)} \tag{19}$$

for $s = 0, \ldots, p$, where $\tilde{v}_{it}$ denotes the within residuals obtained by regressing $\widetilde{y}_{it}$ on $\widetilde{x}_{it}$, where $\widetilde{x}_{it} = x_{it} - \overline{x}_{i\cdot}$, and $\overline{x}_{i\cdot} = \frac{1}{T}\sum_{t=1}^T x_{it}$. After getting $\hat{\gamma}_s$, one can compute $\hat{r}_s = \hat{\gamma}_s/\hat{\gamma}_0$ for $s = 1 \ldots, p$. Next we get estimates for the ρ' s which are needed for the y_{it}^* for $t = p+1, \ldots, T$. We can estimate the ρ's by running the regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}$, $\tilde{v}_{i,t-2}, \ldots,$ $\tilde{v}_{i,t-p}$ $(t > p)$. Finally we turn to the problem of getting an estimate for $a = \sigma_\epsilon^2/\gamma_0$. It is easy to check that

$$\gamma_0 = E\left(v_{it}^2\right) = E\left[v_{it}\left(\rho_1 v_{i,t-1} + \rho_2 v_{i,t-2} + \cdots + \rho_p v_{i,t-p} + \epsilon_{it}\right)\right] = \rho_1\gamma_1 + \rho_2\gamma_2 + \cdots + \rho_p\gamma_p + \sigma_\epsilon^2. \tag{20}$$

Dividing both sides by γ_0, one obtains

$$a = \sigma_\epsilon^2/\gamma_0 = 1 - \rho_1 r_1 - \rho_2 r_2 - \cdots - \rho_p r_p \tag{21}$$

Therefore, GLS on (3) can be obtained by premultiplying this model by $\Sigma^{*-1/2}$ and running OLS. We summarize our estimation procedure as follows:

Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (19). From $\hat{\gamma}_s$ $(s = 1...., p)$, we can get a_t, $b_{t,t-s}$ and α_t from (16), (17) and (18).

Step (ii): Get $\rho_1, \rho_2, \cdots, \rho_p$ from the OLS regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}$,$\tilde{v}_{i,t-2}$,$\cdots$,$\tilde{v}_{i,t-p}$ $(t > p)$. Obtain an estimate of a from (21). We now have all the ingredients to compute y^*_{it} and x^*_{it} for $t = 1, \ldots, T$ from (15).

Step (iii): Compute $\hat{\sigma}^2_\alpha$ and $\hat{\sigma}^2$ in (14) using OLS residuals of y^*_{it} on x^*_{it}. Then compute y^{**}_{it} and x^{**}_{it} for $t = 1, \ldots, T$ from (13). Run the OLS regression of y^{**}_{it} on x^{**}_{it}. This is equivalent to running the GLS regression on (1).

3 Prediction

Goldberger (1962) showed that, for the regression model given in (3) with a general variance-covariance matrix Ω, the best linear unbiased predictor (BLUP) for $y_{i,T+1}$ is given by

$$\hat{y}_{i,T+1} = x'_{i,T+1}\hat{\beta}_{GLS} + w'\Omega^{-1}\hat{u}_{GLS}, \tag{22}$$

where $w = E(uu_{i,T+1})$ is the covariance between the future disturbance $u_{i,T+1}$ and the sample disturbances u. $\hat{\beta}_{GLS}$ is the GLS estimator of β from equation (3) based on Ω and $\hat{u}_{GLS} = y - x'\hat{\beta}_{GLS}$ denotes the corresponding GLS residual vector. As shown in Baltagi and Li (1992), the last term in Equation (22) is given by:

$$\begin{aligned}
w'\Omega^{-1}\hat{u}_{GLS} &= E\left[u_{i,T+1}\left(u'_1, \ldots, u'_N\right)\right] \begin{pmatrix} \Lambda^{-1} & 0 & \cdots & 0 \\ 0 & \Lambda^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda^{-1} \end{pmatrix} \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_N \end{pmatrix} \\
&= \sum_{j=1}^{N} \left[E\left(u_{i,T+1}u'_j\right)\Lambda^{-1}\hat{u}_j\right] \\
&= E\left(u_{i,T+1}u'_i\right)\Lambda^{-1}\hat{u}_i,
\end{aligned} \tag{23}$$

where $u'_i = (u_{i1}, \ldots, u_{iT})$ and $\hat{u}_i$ denote the GLS residuals. The last equality uses the fact that errors of different individuals are independent of each other. Using the fact that $u_{i,T+1} = \mu_i + v_{i,T+1}$, equation (23) can be written as the sum of two terms:

$$E\left(u_{i,T+1}u'_i\right)\Lambda^{-1}\hat{u}_i = E\left(\mu_i u'_i\right)\Lambda^{-1}\hat{u}_i + E\left(v_{i,T+1}u'_i\right)\Lambda^{-1}\hat{u}_i. \tag{24}$$

Recall that $\Lambda^* = C\Lambda C'$, hence

$$\Lambda^{-1} = C'\Lambda^{*-1}C = C'\left(\frac{1}{\sigma_\alpha^2}\bar{J}_T^\alpha + \frac{1}{\sigma^2}E_T^\alpha\right)C. \tag{25}$$

Note that $E\left(\mu_i u_i'\right) = E\left(\mu_i \mu_i \iota_T'\right) = \sigma_\mu^2 \iota_T'$ because of the independence of μ_i and v_i. Hence, the first term in equation (24) can be written as:

$$E\left(\mu_i u_i'\right)\Lambda^{-1}\hat{u}_i = \sigma_\mu^2 \iota_T' C'\left(\frac{1}{\sigma_\alpha^2}\bar{J}_T^\alpha + \frac{1}{\sigma^2}E_T^\alpha\right)C\hat{u}_i = \frac{\sigma_\mu^2}{\sigma_\alpha^2}\iota_T^\alpha \hat{u}_i^* = \frac{\sigma_\mu^2}{\sigma_\alpha^2}\sum_{t=1}^{T}\alpha_t \hat{u}_{it}^*, \tag{26}$$

where $C\hat{u}_i = \hat{u}_i^*$, using the fact $C\iota_T = \iota_T^\alpha$, $\iota_T^{\alpha\prime}\bar{J}_T^\alpha = \iota_T^{\alpha\prime}$ and $\iota_T^{\alpha\prime}E_T^\alpha = 0$. In this case,

$$\begin{aligned}
E\left(v_{i,T+1}u_i'\right) &= E\left(v_{i,T+1}v_i'\right) = E\left[\left(\rho_1 v_{i,T} + \rho_2 v_{i,T-1} + \cdots + \rho_p v_{i,T+1-p} + \epsilon_{i,T+1}\right)v_i'\right] \\
&= \rho_1 E\left(v_{i,T}v_i'\right) + \rho_2 E\left(v_{i,T-1}v_i'\right) + \cdots + \rho_p E\left(v_{i,T-1-p}v_i'\right) + E\left(\epsilon_{i,T+1}v_i'\right),
\end{aligned} \tag{27}$$

where $E\left(v_{i,T}v_i'\right), E\left(v_{i,T-1}v_i'\right), \cdots, E\left(v_{i,T-1-p}v_i'\right)$ are the last p columns of the covariance matrix $E\left(v_i v_i'\right) = \sigma^2 V$. Also, $E\left(\epsilon_{i,T+1}v_i'\right) = 0$. Hence, we have

$$E\left(v_{i,T+1}u_i'\right) = \left(0, \cdots, 0, \rho_p, \cdots, \rho_2, \rho_1\right)\sigma^2 V.$$

Further, notice that Λ^{-1} in Equation (25) can also be written as

$$\begin{aligned}
\Lambda^{-1} &= C'\left[\frac{1}{\sigma^2}I_T + \left(\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma^2}\right)\bar{J}_T^\alpha\right]C \\
&= C'\left[\frac{1}{\sigma^2}I_T - \frac{\sigma_\mu^2 d^2}{\sigma_\alpha^2\sigma^2}\frac{\iota_T^\alpha \iota_T^{\alpha\prime}}{d^2}\right]C \\
&= \frac{C'C}{\sigma^2}\left[I_T - \frac{\sigma_\mu^2}{\sigma_\alpha^2}\iota_T \iota_T^{\alpha\prime}C\right]
\end{aligned} \tag{28}$$

using the fact that $E_T^\alpha = I_T - \bar{J}_T^\alpha$, $\sigma_\alpha^2 = \sigma_\mu^2 d^2 + \sigma^2$ and $\iota_T^\alpha = C\iota_T$. Hence the second term in equation (24) becomes:

$$\begin{aligned}
E\left(v_{i,T+1}u_i'\right)\Lambda^{-1}\hat{u}_i &= \left(0, \cdots, 0, \rho_p, \cdots, \rho_2, \rho_1\right)\sigma^2 V\frac{C'C}{\sigma^2}\left[I_T - \frac{\sigma_\mu^2}{\sigma_\alpha^2}\iota_T \iota_T^{\alpha\prime}C\right]\hat{u}_i \\
&= \left(0, \cdots, 0, \rho_p, \cdots, \rho_2, \rho_1\right)\left[\hat{u}_i - \frac{\sigma_\mu^2}{\sigma_\alpha^2}\iota_T \iota_T^{\alpha\prime}\hat{u}_i^*\right] \\
&= \sum_{s=1}^{p}\rho_s \hat{u}_{i,T+1-s} - \frac{\sigma_\mu^2}{\sigma_\alpha^2}\sum_{s=1}^{p}\rho_s \sum_{t=1}^{T}\alpha_t \hat{u}_{it}^*
\end{aligned} \tag{29}$$

using the fact that $V = \left(C'C\right)^{-1}$ since $CVC' = I_T$. Combining equations (26) and (29), one gets

$$w'\Omega^{-1}\hat{u}_{GLS} = \sum_{s=1}^{p}\rho_s \hat{u}_{i,T+1-s} + \left(1 - \sum_{s=1}^{p}\rho_s\right)\frac{\sigma_\mu^2}{\sigma_\alpha^2}\sum_{t=1}^{T}\alpha_t \hat{u}_{it}^*. \tag{30}$$

Special case 1: *No random effects.* In this case $\sigma_\mu^2 = 0$, and equation (30) reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \sum_{s=1}^{p} \rho_s \hat{u}_{i,T+1-s} \tag{31}$$

This is Goldberger's BLUP *extra term* for the panel data model with AR(p) remainder disturbances but no random individual effects. Goldberger (1962) actually considered the case of an AR(1) process.

Special case 2: *No serial correlation.* In this case $\rho_1 = \rho_2 = \cdots = \rho_p = 0$, so there is no AR($p$) process in the remainder disturbances. It is easy to verify that $a_t = 1$, $b_{t,\tau} = 0$, $\alpha_t = 1$, $d^2 = T$, $\sigma^2 = \sigma_\epsilon^2$, $\sigma_\alpha^2 = T\sigma_\mu^2 + \sigma_\epsilon^2$ and $\hat{u}_{it}^* = \hat{u}_{it}$. In this case, equation (30) reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \frac{\sigma_\mu^2}{\sigma_1^2} \sum_{t=1}^{T} \hat{u}_{it} = \frac{\sigma_\mu^2}{\sigma_1^2} \left(\iota_T \otimes l_i\right)' \hat{u}_{GLS} = (T\sigma_\mu^2/\sigma_1^2)\overline{\widehat{u}}_{i.,GLS}, \tag{32}$$

where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\epsilon^2$, $\overline{\widehat{u}}_{i.,GLS} = \sum_{t=1}^{T} \widehat{u}_{it,GLS}/T$, and l_i is the ith column of I_N. This is Goldberger's BLUP *extra term* derived by Taub (1979) for the random effects error component model with no serial correlation in the remainder disturbances.

Special case 3: *AR(1) process.* Here we show that equation (30) reduces to the results in Baltagi and Li (1992) for the random effects panel data model with AR(1) serially correlated remainder disturbances. Multiply and divide the second term of equation (30) by $a\,(1 - \sum_{s=1}^{p} \rho_s)$, where $a = \sigma_\epsilon^2/\gamma_0$, we get

$$w'\Omega^{-1}\hat{u}_{GLS} = \sum_{s=1}^{p} \rho_s \hat{u}_{i,T+1-s} + \left(1 - \sum_{s=1}^{p} \rho_s\right)^2 \frac{\sigma_\mu^2}{a\sigma_\alpha^2} \sum_{t=1}^{T} \left[\left(\frac{\sqrt{a}}{1 - \sum_{s=1}^{p} \rho_s} \alpha_t\right) \left(\sqrt{a}\hat{u}_{it}^*\right)\right].$$

In order to get to the Baltagi and Li (1992) notation, we define $\tilde{u}_{it}^* = \sqrt{a}\hat{u}_{it}^*$, $\tilde{\alpha}_t = \frac{\sqrt{a}}{1-\sum_{s=1}^{p} \rho_s} \alpha_t$ and $\tilde{\sigma}_\alpha^2 = a\sigma_\alpha^2$. This equation becomes:

$$\begin{aligned} w'\Omega^{-1}\hat{u}_{GLS} &= \sum_{s=1}^{p} \rho_s \hat{u}_{i,T+1-s} + \left(1 - \sum_{s=1}^{p} \rho_s\right)^2 \frac{\sigma_\mu^2}{\tilde{\sigma}_\alpha^2} \sum_{t=1}^{T} \tilde{\alpha}_t \tilde{u}_{it}^* \\ &= \sum_{s=1}^{p} \rho_s \hat{u}_{i,T+1-s} + \left(1 - \sum_{s=1}^{p} \rho_s\right)^2 \frac{\sigma_\mu^2}{\tilde{\sigma}_\alpha^2} \left(\sum_{t=1}^{p} \tilde{\alpha}_t \tilde{u}_{it}^* + \sum_{t=p+1}^{T} \tilde{u}_{it}^*\right) \end{aligned} \tag{33}$$

since $\tilde{\alpha}_t = 1$ for $t = p+1, \ldots, T$, see (18).

If $p = 1$, Equation (33) reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \rho_1 \hat{u}_{i,T} + (1 - \rho_1)^2 \frac{\sigma_\mu^2}{\tilde{\sigma}_\alpha^2} \left(\tilde{\alpha}_1 \tilde{u}_{i1}^* + \sum_{t=2}^{T} \tilde{u}_{it}^*\right). \tag{34}$$

Note that $a = \sigma_\epsilon^2/\gamma_0 = 1 - \rho_1^2$; $\tilde{\alpha}_1 = \frac{\sqrt{a}}{1-\rho_1}\alpha_1 = \frac{\sqrt{1-\rho_1^2}}{1-\rho_1} = \sqrt{\frac{1+\rho_1}{1-\rho_1}}$, since $\alpha_1 = 1$ from (18). $\tilde{\sigma}_\alpha^2 = a\sigma_\alpha^2 = (1 - \rho_1^2)\left(\sigma_\mu^2 d^2 + \sigma^2\right)$. Define $\tilde{d}^2 = \frac{1+\rho_1}{1-\rho_1} d^2 = \frac{1+\rho_1}{1-\rho_1}\left(1 + \sum_{t=2}^{T} \frac{1-\rho_1}{1+\rho_1}\right) = \frac{1+\rho_1}{1-\rho_1} + T - 1$. Then

$\tilde{\sigma}_\alpha^2 = (1-\rho_1)^2 \sigma_\mu^2 \tilde{d}^2 + \sigma_\epsilon^2$. The recursive transformation for the AR(1) remainder disturbances reduces to the Prais-Winsten transformation given by: $\tilde{u}_{i1}^* = \sqrt{1-\rho_1^2} u_{i1}$ and $\tilde{u}_{it}^* = u_{it} - \rho_1 u_{i,t-1}$ for $t = 2, \ldots, T$. This reproduces Goldberger's extra term derived in equation (13) by Baltagi and Li (1992, p.564) for the random effects panel data model with AR(1) serial correlated remainder disturbances.

Special case 4: *AR(2) process*. If $p = 2$, Equation (33) reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \rho_1 \hat{u}_{i,T} + \rho_2 \hat{u}_{i,T-1} + (1-\rho_1-\rho_2)^2 \frac{\sigma_\mu^2}{\tilde{\sigma}_\alpha^2} \left(\tilde{\alpha}_1 \tilde{u}_{i1}^* + \tilde{\alpha}_2 \tilde{u}_{i2}^* + \sum_{t=3}^{T} \tilde{u}_{it}^* \right). \tag{35}$$

Note that $\gamma_1 = E(v_{it} v_{i,t-1}) = E[(\rho_1 v_{i,t-1} + \rho_2 v_{i,t-2} + \epsilon_{it}) v_{i,t-1}] = \rho_1 \gamma_0 + \rho_2 \gamma_1$. Solving for γ_1,we get $\gamma_1 = \frac{\rho_1}{1-\rho_2}\gamma_0$. Hence $b_{2,1} = r_1 = \frac{\gamma_1}{\gamma_0} = \frac{\rho_1}{1-\rho_2}$ and $a_2 = 1 - b_{2,1}^2 = 1 - \left(\frac{\rho_1}{1-\rho_2}\right)^2 = \frac{(1-\rho_2)^2-\rho_1^2}{(1-\rho_2)^2}$, $\alpha_2 = (1 - b_{2,1}\alpha_1)/\sqrt{a_2} = \frac{1-\frac{\rho_1}{1-\rho_2}}{\sqrt{1-\left(\frac{\rho_1}{1-\rho_2}\right)^2}} = \sqrt{\frac{1-\frac{\rho_1}{1-\rho_2}}{1+\frac{\rho_1}{1-\rho_2}}} = \sqrt{\frac{1-\rho_1-\rho_2}{1+\rho_1-\rho_2}}$. Also, $a = \sigma_\epsilon^2/\gamma_0 = \sigma_\epsilon^2/\sigma_v^2$, $\tilde{\alpha}_1 = \frac{\sqrt{a}}{1-\rho_1-\rho_2}\alpha_1 = \frac{\sigma_\epsilon}{\sigma_v(1-\rho_1-\rho_2)}$, $\tilde{\alpha}_2 = \frac{\sqrt{a}}{1-\rho_1-\rho_2}\alpha_2 = \frac{\sigma_\epsilon}{\sigma_v(1-\rho_1-\rho_2)}\sqrt{\frac{1-\rho_1-\rho_2}{1+\rho_1-\rho_2}} = \sqrt{\frac{1+\rho_2}{1-\rho_2}}$ using the fact $\sigma_v^2 = \frac{(1-\rho_2)\sigma_\epsilon^2}{(1+\rho_2)\left[(1-\rho_2)^2-\rho_1^2\right]}$, which is shown in Baltagi and Li (1991). $\tilde{\sigma}_\alpha^2 = a\sigma_\alpha^2 = a\left(\sigma_\mu^2 d^2 + \sigma^2\right) = (1-\rho_1-\rho_2)^2 \sigma_\mu^2 \tilde{d}^2 + \sigma_\epsilon^2$, where $\tilde{d}^2 = \frac{\sigma_\epsilon^2}{\sigma_v^2(1-\rho_1-\rho_2)^2} d^2 = \tilde{\alpha}_1^2 + \tilde{\alpha}_2^2 + T - 2$. The recursive transformation for the AR(2) remainder disturbances reduces to $\hat{u}_{i1}^* = \hat{u}_{i1}$, $\hat{u}_{i2}^* = (\hat{u}_{i2} - b_{2,1}\hat{u}_{i1}^*)/\sqrt{a_2} = \frac{\sqrt{(1-\rho_2)^2-\rho_1^2}}{1-\rho_2}\hat{u}_{i2} - \frac{\rho_1\sqrt{(1-\rho_2)^2-\rho_1^2}}{(1-\rho_2)^2}\hat{u}_{i1}$, $\hat{u}_{it}^* = \frac{\sigma_v}{\sigma_\epsilon}(\hat{u}_{it} - \rho_1 \hat{u}_{i,t-1} - \rho_2 \hat{u}_{i,t-2})$ for $t = 3, \ldots, T$. Hence $\tilde{u}_{i1}^* = \frac{\sigma_\epsilon}{\sigma_v} u_{i1}$, $\tilde{u}_{i2}^* = \frac{\sigma_\epsilon}{\sigma_v}\left(\frac{\sqrt{(1-\rho_2)^2-\rho_1^2}}{1-\rho_2}\hat{u}_{i2} - \frac{\rho_1\sqrt{(1-\rho_2)^2-\rho_1^2}}{(1-\rho_2)^2}\hat{u}_{i1}\right) = \sqrt{1-\rho_2^2}\left(\hat{u}_{i2} - \frac{\rho_1}{1-\rho_2}\hat{u}_{i1}\right)$, and $\tilde{u}_{it}^* = \hat{u}_{it} - \rho_1 \hat{u}_{i,t-1} - \rho_2 \hat{u}_{i,t-2}$ for $t = 3, \ldots, T$. This reproduces Goldberger's extra term derived in equation (14) by Baltagi and Li (1992, p.565) for the random effects panel data model with AR(2) serially correlated remainder disturbances.

Special case 5: *The specialized AR(4) process for quarterly data*: $v_{it} = \rho_4 v_{i,t-4} + \epsilon_{it}$, with $\rho_1 = \rho_2 = \rho_3 = 0$. In this case, equation (33) reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \rho_4 \hat{u}_{i,T-3} + (1-\rho_4)^2 \frac{\sigma_\mu^2}{\tilde{\sigma}_\alpha^2} \left(\sum_{t=1}^{4} \tilde{\alpha}_t \tilde{u}_{it}^* + \sum_{t=5}^{T} \tilde{u}_{it}^* \right). \tag{36}$$

It is easy to verify that $a = \sigma_\epsilon^2/\gamma_0 = 1 - \rho_4^2$, $a_1 = a_2 = a_3 = a_4 = 1$, $b_{2,1} = b_{3,1} = b_{3,2} = b_{4,1} = b_{4,2} = b_{4,3} = 0$ and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$. Hence $\tilde{\alpha}_1 = \tilde{\alpha}_2 = \tilde{\alpha}_3 = \tilde{\alpha}_4 = \sqrt{\frac{1+\rho_4}{1-\rho_4}}$, $\tilde{\sigma}_\alpha^2 = a\sigma_\alpha^2 = a\left(\sigma_\mu^2 d^2 + \sigma^2\right) = (1-\rho_4)^2 \sigma_\mu^2 \tilde{d}^2 + \sigma_\epsilon^2$, where $\tilde{d}^2 = \frac{1+\rho_4}{1-\rho_4} d^2 = \frac{1+\rho_4}{1-\rho_4}\left(4 + \sum_{t=5}^{T} \frac{1-\rho_4}{1+\rho_4}\right) = \frac{1+\rho_4}{1-\rho_4} + T - 4$. The recursive transformation for the specialized AR(4) remainder disturbances reduces to $\tilde{u}_{it}^* = \sqrt{1-\rho_1^2} u_{it}$ for $t = 1, 2, 3, 4$ and $\tilde{u}_{it}^* = u_{it} - \rho_4 u_{i,t-4}$ for $t = 5, \ldots, T$. This reproduces Goldberger's extra term derived in equation (15) by Baltagi and Li (1992, p.565) for the random effects panel data model with specialized AR(4) remainder disturbances. Hence the results derived in this paper encompass the earlier results and generalize them to remainder disturbances of an arbitrary AR(p) order.

4 Monte Carlo Simulation

This section performs some limited Monte Carlo experiments to evaluate the performance of our proposed predictors for the random effects model with AR(p) disturbances. It is important to note that Kouassi et al. (2011) performed extensive Monte Carlo experiments to evaluate the performance of predictors for the random effects model with AR(1) disturbances. Following Baillie and Baltagi (1999) and Kouassi et al. (2011) the data generating process starts with a simple panel data regression with random one-way error components disturbances

$$y_{it} = 5 + 0.5x_{it} + \mu_i + \nu_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T+1 \tag{37}$$

The variable x_{it} was generated as in Nerlove (1971) with $x_{it} = 0.1t + 0.5x_{i,t-1} + \omega_{it}$, where ω_{it} is a random variable uniformly distributed on the interval $[-0.5, 0.5]$ and $x_{i0} = 5 + 10\omega_{i0}$. The individual specific effect $\mu_i \overset{iid}{\sim} N\left(0, \sigma_\mu^2\right)$ with $\sigma_\mu^2 = 15$. The remainder disturbances ν_{it} were generated as an AR(p) process with the following three designs:

1. Model 1: $\nu_{it} = -0.8\nu_{i,t-1} + \varepsilon_{it}$,
2. Model 2: $\nu_{it} = 0.2\nu_{i,t-1} + 0.63\nu_{i,t-2} + \varepsilon_{it}$,
3. Model 3: $\nu_{it} = -0.7\nu_{i,t-1} - 0.53\nu_{i,t-2} + 0.315\nu_{i,t-3} + \varepsilon_{it}$,

In all models, the variance of ν_{it} was fixed at $\sigma_\nu^2 = 15$. The first 20 period observations were discarded to minimize the effect of initial values. Predictions were made for only one period ahead. In order to depict the typical labor or consumer panel where N is large and T is small, the sample sizes (N, T) in the different experiments were chosen as $(100, 10)$ and $(200, 20)$. For each experiment, we perform $1,000$ replications. For each replication we estimate the model using the pooled ordinary least square (OLS), panel random effect (RE) and random effect model with AR(1), AR(2) and AR(3) terms respectively, (RE-AR1, RE-AR2 and RE-AR3). Following Kouassi et al. (2011), the sampling *mean square error (MSE)* of each of the predictors considered above is computed as

$$MSE = \frac{1}{nR} \sum_{r=1}^{R} \sum_{i=1}^{n} \left(\hat{y}_{i,T+1} - y_{i,T+1}\right)^2, \tag{38}$$

where $R = 1,000$ replications. Following Frees and Miller (2004), we also summarize the accuracy of the forecasts using two other statistics, the *mean absolute error (MAE)*

$$MAE = \frac{1}{nR} \sum_{r=1}^{R} \sum_{i=1}^{n} \left|\hat{y}_{i,T+1} - y_{i,T+1}\right| \tag{39}$$

and the *mean absolute percentage error (MAPE)*

$$MAPE = \frac{100}{nR}\sum_{r=1}^{R}\sum_{i=1}^{n}\left|\frac{\hat{y}_{i,T+1} - y_{i,T+1}}{y_{i,T+1}}\right|. \tag{40}$$

Tables 1 and 2 report the results for sample sizes ($N = 100, T = 10$) and ($N = 200, T = 20$), respectively. For model 1 where the true DGP is RE-AR1, the MSE, MAE and MAPE of RE-AR1 is the smallest. Similarly, for model 2 where the true DGP is RE-AR2, the MSE, MAE and MAPE of RE-AR2 is the smallest, while for model 3 where the true DGP is RE-AR3, the MSE, MAE and MAPE of RE-AR3 is the smallest.

5 Application

In this section we revisit the forecast application considered by Frees and Miller (2004). This is a panel of 50 postal (ZIP) codes in Wisconsin observed over 40 weeks. Frees and Miller regressed the logarithm of online lottery sales (LNZOLSALES) on persons per household times 10 (PERPERHH), median years of schooling times 10 (MEDSCHYR), median home value in $100s for owner-occupied homes (OOMEDHVL), percent of housing that is renter occupied (PRCRENT), percent of population that is 55 or older (PRC55P), household median age (HHMEDAGE), estimated median household income in $100s (CEMI) and population (POP). Besides the pooled ordinary least square (OLS), panel random effect (RE) and random effect model with AR(1) term (RE-AR1) that are reported in Frees and Miller (2004), we also report the random effect model with AR(2) and AR(3) term, (RE-AR2 and RE-AR3) respectively. As in Frees and Miller (2004), we use the first 35 weeks of data to estimate the model. The results are shown in Table 3. The first three collumns, pooled cross-sectional model, error component model and error component model with AR(1) term, replicate the results in Table 3 of Frees and Miller (2004). We focus on forecasting one period ahead to illustrate our theoretical results. For each estimator, we compute the forecasts of lottery sales for week 36, by ZIP code level, based on the first 35 weeks. Following Frees and Miller (2004), we summarize the accuracy of the forecasts of $LNZOLSALES_{i,36}$ using MSE, MAE and MAPE, which are defined in Equation (38)-(40) by replacing $R = 1$. Table 3 confirms that the random effects model with an AR(1) term has the smallest MSE, MAE or MAPE for logarithmic sales.[1]

[1] Frees and Miller (2004) compared the forecasts for 5 weeks ahead using MAE and MAPE. They used nine alternative forecasts. They do find that the random effects with AR(1) performs the best. Our results using their data forecast one week ahead and include higher order AR models in a random effects model including AR(2) and AR(3) to illustrate our theoretical derivations.

6 Conclusion

This paper derives the best linear unbiased predictor for a panel data model with random individual effects and AR(p) remainder disturbances. The performance of this predictor is investigated using Monte Carlo experiments and illustrated using the Wisconsin lottery sales example of Frees and Miller (2004).

Acknowledgments

We would like to thank two anonymous referees, an Associate Editor and the editor Graham Elliot for their helpful comments and suggestions.

References

Baillie, R.T. and B.H. Baltagi, 1999, Prediction from the regression model with one-way error components, Chapter 10 in Analysis of panel data and limited dependent variable models, Cheng Hsiao, Kajal Lahiri, Lung-Fei Lee and Hashem Pesaran, editors, Cambridge University Press, Cambridge.

Baltagi, B.H., 2008, Forecasting with panel data, Journal of Forecasting 27, 153-173.

Baltagi, B.H. and Q. Li, 1991, A transformation that will circumvent the problem of autocorrelation in an error component model, Journal of Econometrics 48, 385-393.

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, A simple recursive estimation method for linear regression models with AR(p) disturbances, Statistical Papers 35, 93-100.

Brockwell, P.J. and R.A. Davis, 1991, Time series: theory and methods. New York: Springer-Verlag.

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

Fuller, W.A., 1976, Introduction to statistical time series. New York: Wiley.

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

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

Judge, G., C. Hill, W. Griffiths, H. Lutkepohl and T. Lee, 1985, The theory and practice of econometrics. New York: Wiley.

Kouassi, E., Sango, J., Bosson Brou, J., Teubissi, F. N. and K. O. Kymn, 2011, Prediction from the one-way error components model with AR(1) disturbances, forthcoming in the Journal of Forecasting.

Maeshiro, K.R., 1976, Autoregressive transformation, trended independent variables and autocorrelated disturbance terms, The Review of Economics and Statistics 58, 497-500.

Maeshiro, K.R., 1979, On the retention of the first observations in serial correlation adjustment of regression models, International Economic Review 20, 259-265.

Park, R.E. and B.M. Mitchell, 1980, Estimating the autocorrelated error model with trended data, Journal of Econometrics 13, 185-201.

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

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

Wansbeek, T.J. and A. Kapteyn, 1983, A note on spectral decomposition and maximum likelihood estimation of ANOVA models with balanced data, Statistics and Probability Letters 1, 213-215.

Table 1: Comparison of Estimators ($n = 100, T = 10$)

	True Model	OLS	RE	RE-AR1	RE-AR2	RE-AR3
MSE	Model 1	30.866	17.418	6.372	6.413	6.458
	Model 2	31.646	11.217	11.820	8.817	9.184
	Model 3	34.930	22.782	15.753	5.143	4.672
MAE	Model 1	4.439	3.337	2.017	2.023	2.030
	Model 2	4.488	2.670	2.741	2.368	2.417
	Model 3	4.713	3.807	3.171	1.811	1.727
MAPE	Model 1	463.755	307.451	177.090	177.270	178.516
	Model 2	1843.802	611.083	555.646	484.035	512.590
	Model 3	473.531	315.954	265.836	157.374	136.386

Note: MSE, MAE and MAPE are out-of-sample forecast comparison for one period ahead.

Table 2: Comparison of Estimators ($n = 200, T = 20$)

	True Model	OLS	RE	RE-AR1	RE-AR2	RE-AR3
MSE	Model 1	30.720	16.443	5.965	5.978	5.992
	Model 2	31.200	13.440	11.934	7.575	7.652
	Model 3	33.850	19.836	14.672	4.481	4.053
MAE	Model 1	4.422	3.232	1.949	1.951	1.954
	Model 2	4.457	2.925	2.755	2.197	2.209
	Model 3	4.641	3.551	3.056	1.689	1.607
MAPE	Model 1	382.275	214.087	106.091	106.540	107.314
	Model 2	433.756	185.505	177.250	127.944	129.112
	Model 3	390.932	239.863	188.046	93.128	87.780

Note: MSE, MAE and MAPE are out-of-sample forecast comparison for one period ahead.

Table 3: Lottery model coefficient estimates

	OLS	RE	RE-AR1	RE-AR2	RE-AR3
Intercept	13.821	17.811	15.180	16.044	15.991
	(1.340)	(6.935)	(6.246)	(6.325)	(6.325)
PERPERHH	-0.108	-0.127	-0.115	-0.119	-0.119
	(0.016)	(0.084)	(0.075)	(0.076)	(0.076)
MEDSCHYR	-0.082	-0.106	-0.091	-0.096	-0.095
	(0.007)	(0.036)	(0.032)	(0.033)	(0.033)
MEDHVL	0.001	0.001	0.001	0.001	0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
PRCRENT	0.032	0.027	0.030	0.029	0.029
	(0.004)	(0.020)	(0.018)	(0.018)	(0.018)
PRC55P	-0.070	-0.072	-0.071	-0.071	-0.071
	(0.013)	(0.070)	(0.063)	(0.064)	(0.064)
HHMEDAGE	0.118	0.119	0.120	0.119	0.119
	(0.021)	(0.110)	(0.098)	(0.100)	(0.100)
MEDINC	0.004	0.005	0.004	0.004	0.004
	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
$POP/1000$	0.057	0.118	0.079	0.091	0.090
	(0.006)	(0.026)	(0.026)	(0.026)	(0.026)
NRETAIL	0.021	-0.024	0.005	-0.004	-0.004
	(0.004)	(0.017)	(0.017)	(0.017)	(0.017)
ρ_1			0.513	0.624	0.628
ρ_2				-0.229	-0.257
ρ_3					0.034
MSE	0.528	0.112	0.057	0.077	0.076
MAE	0.585	0.285	0.190	0.228	0.225
MAPE	8.684	4.059	2.777	3.307	3.257

Note: In-sample model coefficient estimates are based on n=50 ZIP codes and T=35 weeks. The response is logarithmic sales. MSE, MAE and MAPE are out-of-sample forecast comparison for week 36.