Syracuse University

SURFACE

Center for Policy Research

Maxwell School of Citizenship and Public Affairs

1-2020

Forecasting with Unbalanced Panel Data

Badi Baltagi
bbaltagi@maxwell.syr.edu

Long Liu
University of Texas at San Antonio, long.liu@utsa.edu

CENTER FOR POLICY RESEARCH

THE MAXWELL SCHOOL

WORKING PAPER SERIES

Forecasting with Unbalanced Panel Data

Badi Baltagi and Long Liu

Paper No. 221
January 2020

ISSN: 1525-3066

Maxwell
Syracuse University
Center for Policy Research
426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
(315) 443-3114/ email: ctrpol@syr.edu
http://www.maxwell.syr.edu/CPR_Working_Papers.aspx

CENTER FOR POLICY RESEARCH – Spring 2020

Leonard M. Lopoo, Director
Professor of Public Administration and International Affairs (PAIA)

Associate Directors

Margaret Austin
Associate Director, Budget and Administration

John Yinger
Trustee Professor of Economics (ECON) and Public Administration and International Affairs (PAIA)
Associate Director, Center for Policy Research

SENIOR RESEARCH ASSOCIATES

Badi Baltagi, ECON
Robert Bifulco, PAIA
Leonard Burman, PAIA
Carmen Carrión-Flores, ECON
Alfonso Flores-Lagunes, ECON
Sarah Hamersma, PAIA
Madonna Harrington Meyer, SOC
Colleen Heflin, PAIA
William Horrace, ECON
Yilin Hou, PAIA
Hugo Jales, ECON
Jeffrey Kubik, ECON
Yoonseok Lee, ECON
Amy Lutz, SOC
Yingyi Ma, SOC
Katherine Michelmore, PAIA
Jerry Miner, ECON
Shannon Monnat, SOC
Jan Ondrich, ECON
David Popp, PAIA
Stuart Rosenthal, ECON
Michah Rothbart, PAIA
Alexander Rothenberg, ECON
Rebecca Schewe, SOC
Amy Ellen Schwartz, PAIA/ECON
Ying Shi, PAIA
Saba Siddiki, PAIA
Perry Singleton, ECON
Yulong Wang, ECON
Michael Wasylenko, ECON
Peter Wilcoxen, PAIA
Maria Zhu, ECON

GRADUATE ASSOCIATES

Rhea Acuña, PAIA
Mariah Brennan, SOC. SCI.
Jun Cai, ECON
Ziqiao Chen, PAIA
Yoon Jung Choi, PAIA
Dahae Choo, PAIA
Stephanie Coffey, ECON
Brandon De Bruhl, PAIA
Giuseppe Germinario, ECON
Myriam Gregoire-Zawilski, PAIA
Emily Gutierrez, PAIA
Jeehee Han, PAIA
Mary Helander, Lerner
Hyoung Kwon, PAIA
Mattie Mackenzie-Liu, PAIA
Maeve Maloney, ECON
Austin McNeill Brown, SOC. SCI.
Qasim Mehdi, PAIA
Claire Pendergrast, SOC
Jonathan Presler, ECON
Krushna Ranaware, SOC
Christopher Rick, PAIA
David Schwegman, PAIA
Saied Toossi, PAIA
Huong Tran, ECON
Joaquin Urrego, ECON
Yao Wang, ECON
Yi Yang, ECON
Xiaoyan Zhang, ECON
Bo Zheng, PAIA
Dongmei Zhu, SOC. SCI.

STAFF

Joanna Bailey, Research Associate
Joseph Boskovski, Manager, Maxwell X Lab
Katrina Fiacchi, Administrative Specialist
Michelle Kincaid, Senior Associate, Maxwell X Lab
Emily Minnoe, Administrative Assistant
Candi Patterson, Computer Consultant
Samantha Trajkovski, Postdoctoral Scholar
Laura Walsh, Administrative Assistant

Abstract

This paper derives the best linear unbiased prediction (BLUP) for an unbalanced panel data model. Starting with a simple error component regression model with unbalanced panel data and random effects, it generalizes the BLUP derived by Taub (1979) to unbalanced panels. Next it derives the BLUP for an unequally spaced panel data model with serial correlation of the AR(1) type in the remainder disturbances considered by Baltagi and Wu (1999). This in turn extends the BLUP for a panel data model with AR(1) type remainder disturbances derived by Baltagi and Li (1992) from the balanced to the unequally spaced panel data case. The derivations are easily implemented and reduce to tractable expressions using an extension of the Fuller and Battese (1974) transformation from the balanced to the unbalanced panel data case.

.**JEL No.:** C33

Keywords: Forecasting, BLUP, Unbalanced Panel Data, Unequally Spaced Panels, Serial Correlation

Authors: Badi H. Baltagi, Department of Economics, Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, bbaltagi@maxwell.syr.edu; Long Liu, Department of Economics, College of Business, University of Texas at San Antonio, 1 UTSA Circle, TX 78249-0633, long.liu@utsa.edu

1 Introduction

Panel data is usually unbalanced or unequally spaced due to lack observations on households not interviewed in certain years or firms not filing their data survey forms for a particular period. Even daily stock price data has no observations when the market is closed due to holidays or weekends. The unequally spaced pattern is also useful for repeated sales of houses that are not sold each year but at irregularly spaced intervals. It is also a common problem for longitudinal surveys and household surveys in developed as well as developing countries, see examples of these in Table 1 of McKenzie (2001) as well as Table 1 of Millimet and McDonough (2017). Unbalanced panel data estimation and testing has been studied in econometrics, see Chapter 9 of Baltagi (2013a) and the references cited there. This paper focuses on forecasting with unbalanced panel data. In particular, the paper starts by extending the best linear unbiased predictor (BLUP) derived by Taub (1979) for the random effects error component model from balanced to unbalanced panel data models. Next, the BLUP for the unequally spaced panel data with serial correlation of the AR(1) type in the remainder disturbances, considered by Baltagi and Wu (1999) is derived. This extends the BLUP for the random effects model with serial correlation of the AR(1) type derived by Baltagi and Li (1992) from balanced panels to unequally spaced panels. Unbalanced panel data can be messy. This paper keeps the derivations simple and easily tractable, using the Fuller and Battese (1974) transformation extended from the balanced to the unbalanced panel data case.

2 The Best Linear Unbiased Predictor

Consider an unbalanced panel data regression model:

$$y_{it} = X_{it}'\beta + u_{it} \tag{1}$$

for $i = 1, \ldots, N$; $t = 1 \ldots, T_i$. The i subscript denotes, say, individuals in the cross-section dimension and t denotes years in the time-series dimension. The panel data is unbalanced since there are N unique individuals and individual i is only observed over T_i

time periods.[1] The regressor X_{it} is a $K \times 1$ vector of the explanatory variables and β is a $K \times 1$ vector of coefficients. In an earnings equation in economics, for example, y_{it} is log wage for the ith worker in the tth time period. X_{it} may contain a set of variables like age, experience, tenure, and whether the worker is male, black, etc. In most of the panel data applications, the disturbances follow a simple one-way error component model with

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

where μ_i denotes the *unobservable* time-invariant individual specific effect, such as ability. v_{it} denotes the remainder disturbance that varies with individuals and time, see Baltagi (2013a) . Let $n = \sum_{i=1}^{N} T_i$. In vector notation, Equations (1) and (2) can be written as

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

and

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

where $y = (y_{11}, \ldots, y_{1T_1}, y_{21}, \ldots, y_{2T_2}, \ldots, y_{N1}, \ldots, y_{NT_N})'$ is an $n\times$ 1 vector of ,observations stacked such that the slower index is over individuals and the faster index is over time.[2] Other vectors or matrices including X, u and v are similarly defined. $\mu = (\mu_1, \ldots, \mu_N)'$ is an $N \times 1$ vector. The selector matrix $Z_\mu = diag\left[\iota_{T_i}\right]$ is a matrix of ones and zeros, where ι_{T_i} is a vector of ones of dimension T_i. It is simply the matrix of individual dummies that one may include in the regression to estimate the μ_i if they are assumed

[1] The data is assumed to be missing at random. This in turn allows the missingness of the data scheme to be ignorable in the language of Little and Rubin (2002).

[2] This pattern of unbalancedness does not have to be from $1, 2, .., T_i$. In fact, these T_i observations can be for any subset of the observed time series period. This pattern is used to make the derivation easy and tractable and follow similar derivations for the balanced case. A more general pattern of unbalancedness can be used. In fact, section 2 extends this to the unequally spaced panel data with serial correlation across time considered by Baltagi and Wu (1999). A two-way error component model with a general type of missing data is considered in Wansbeek and Kapteyn (1989).

to be fixed parameters. Define $P = Z_\mu(Z'_\mu Z_\mu)^{-1}Z'_\mu$, which is the projection matrix on Z_μ. In this case, $Z_\mu Z'_\mu = diag\,[J_{T_i}]$, where J_{T_i} is a matrix of ones of dimension T_i. Let $\bar{J}_{T_i} = J_{T_i}/T_i$. Hence P reduces to $diag\left[\bar{J}_{T_i}\right]$, which averages the observation across time for each individual over their T_i observations. Similarly, $Q = I_{NT} - P$ is a matrix which obtains the deviations from individual means. For example, if we regress y on the matrix of dummy variables Z_μ, the predicted values Py have a typical element $\overline{y}_{i.} = \sum_{t=1}^{T_i} y_{it}/T_i$ repeated T_i times for each individual. Qy gives the residuals of this regression with typical element $y_{it} - \overline{y}_{i.}$.

For the *random effects* model, $\mu_i \sim \text{IID}(0, \sigma_\mu^2)$, $v_{it} \sim \text{IID}(0, \sigma_\nu^2)$ and the μ_i are independent of the v_{it} and X_{it} for all i and t. The variance-covariance matrix of the disturbances is given by

$$\Omega = E(uu') = \sigma_\mu^2 diag\,[J_{T_i}] + \sigma_v^2 diag\,[I_{T_i}] = diag\left[\omega_i^2 \bar{J}_{T_i} + \sigma_\nu^2 E_{T_i}\right] \tag{5}$$

where $\omega_i^2 = T_i\sigma_\mu^2 + \sigma_\nu^2$,and $E_{T_i} = I_{T_i} - \bar{J}_{T_i}$. Using the fact that $\bar{J}_{T_i}$ and E_{T_i} are idempotent matrices that sum to the identity matrix I_{T_i}, it is easy to verify that

$$\Omega^{-1} = diag\left[\frac{1}{\omega_i^2}\bar{J}_{T_i} + \frac{1}{\sigma_\nu^2}E_{T_i}\right] \tag{6}$$

and

$$\Omega^{-1/2} = diag\left[\frac{1}{\omega_i}\bar{J}_{T_i} + \frac{1}{\sigma_\nu}E_{T_i}\right] \tag{7}$$

see Wansbeek and Kapteyn (1982). Now a GLS estimator can be obtained as a weighted least squares following Fuller and Battese (1974). In this case one premultiplies the regression model in Equation (3) by $\sigma_\nu\Omega^{-1/2} = diag\left[\frac{\sigma_\nu}{\omega_i}\bar{J}_{T_i} + E_{T_i}\right] = diag\left[I_{T_i} - \theta_i\bar{J}_{T_i}\right]$ where $\theta_i = 1 - (\sigma_\nu/\omega_i)$. GLS becomes OLS on the resulting transformed regression of y^* on X^* with $y^* = \sigma_\nu\Omega^{-1/2}y$ having a typical element $y_{it}^* = y_{it} - \theta_i\bar{y}_{i.}$,and $X^* = \sigma_\nu\Omega^{-1/2}X$ defined similarly.

For the ith individual, we want to predict S periods ahead. As derived by Goldberger

(1962), the best linear unbiased predictor (BLUP) of y_{i,T_i+S} for the GLS model is

$$\hat{y}_{i,T_i+S} = X'_{i,T_i+S}\hat{\beta}_{GLS} + w'\Omega^{-1}\hat{u}_{GLS}, \tag{8}$$

for $S \geqslant 1$, where $\hat{\beta}_{GLS}$ is the GLS estimator of β from equation (3), $w = E(u_{i,T+S}u)$, Ω is the variance-covariance structure of the disturbances, and $\widehat{u}_{GLS} = y - X\hat{\beta}_{GLS}$. Note that we have $u_{i,T_i+S} = \mu_i + \nu_{i,T_i+S}$ for period $T_i + S$ and hence $w' = \sigma^2_\mu(0, .., \iota'_{T_i}, 0, .., 0)$. In this case

$$w'\Omega^{-1} = \sigma^2_\mu(0, .., \iota'_{T_i}, 0, .., 0)diag\left[\frac{1}{\omega_i^2}\bar{J}_{T_i} + \frac{1}{\sigma_\nu^2}E_{T_i}\right] = \frac{\sigma^2_\mu}{\omega_i^2}(0, .., \iota'_{T_i}, 0, .., 0) \tag{9}$$

since $\iota'_{T_i}\ \bar{J}_{T_i} = \iota'_{T_i}$ and $\iota'_{T_i}E_{T_i} = 0$. The last term of BLUP becomes

$$w'\Omega^{-1}\widehat{u}_{GLS} = \frac{T_i\sigma^2_\mu}{\omega_i^2}\overline{\widehat{u}}_{i.,GLS}, \tag{10}$$

where $\overline{\widehat{u}}_{i.,GLS} = T_i^{-1}\sum_{t=1}^{T_i}\widehat{u}_{it,GLS}$. 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 over the T_i observed periods. This BLUP was derived by Taub (1979) for the balanced panel data case. Note that it is based on the true variance components. In practice, we need to estimate the variance components to get feasible GLS and a feasible BLUP. Methods for estimating the variance components for the unbalanced panel data model are described in more details in Baltagi (2013a). To account for the additional uncertainty introduced by estimating these variance components, Kackar and Harville (1984) proposed inflation factors for the predictor.

Although this derivation has albeit a restrictive form of missing observations, for example, the time series has no gaps, the results still hold for the Fuller and Battese (1974) transformation and the Goldberger (1962) BLUP derivation even with time series gaps. This is because the individual effects are independent and the idiosyncratic error terms are not correlated across time. Also, as footnote 2 states, the pattern of missing observations can be more general, all that matters is that individual i be observed for only T_i periods and these can be any subset of the observed sample period.

For a recent survey of the BLUP literature mostly for balanced panel data in econometrics, see Baltagi (2013b). The BLUP methodology in statistics has been used extensively in biometrics, see Henderson (1975). Harville (1976) showed that BLUP is equivalent to Bayesian posterior mean predictors with a diffuse prior. Robinson (1991) has an extensive review of how BLUP can be used for example to remove noise from images and for small-area estimation. It can be also used to derive the Kalman filter. For several applications of forecasting with panel data in economics and related disciplines, see the handbook of forecasting chapter by Baltagi (2013b) and the references cited there.

In the next section, we revisit the unequally spaced panel data model with AR(1) type remainder disturbances, considered by Baltagi and Wu (1999). While the Fuller and Battese (1974) transformation for that model was derived in that paper, the Goldberger (1962) BLUP was not given. For forecasting purposes, we derive a simple to compute expression of this predictor and show that it reduces to the usual BLUP under several special cases.

3 Unequally Spaced Panel Data Model with AR(1) type remainder disturbances

Baltagi and Wu (1999) considered an unequally spaced panel data model with *both* random effects and serial correlation of the AR(1) type in the remainder disturbances. To be specific, $\mu_i \sim \text{IID}(0, \sigma_\mu^2)$ and is assumed to be independent of the remainder disturbances v_{it}. In this case, v_{it} follows an AR(1) process given by

$$v_{it} = \rho v_{i,t-1} + \epsilon_{it} \tag{11}$$

for $t = 1, .., T_i$, where $\epsilon_{it} \sim \text{IID}(0, \sigma_\epsilon^2)$ and $|\rho| < 1$. For the initial value, we assume $v_{i0} \sim (0, \sigma_\epsilon^2/(1-\rho^2))$. For each individual i, one observes the data at times $t_{i,j}$ for $j = 1, \ldots, n_i$. Furthermore, we have $1 = t_{i,1} < \cdots < t_{i,n_i} = T_i$ for $i = 1, \ldots, N$ with $n_i > K$. This is a general form of unbalanced panel data which encompasses the case in Section 1. For $i = 1, \ldots, N$, we have

$$u_i = \mu_i \iota_{n_i} + \nu_i, \tag{12}$$

where $u_i' = \left(u_{i,t_{i,1}}, \ldots, u_{i,t_{i,n_i}}\right)$, $v_i' = \left(v_{i,t_{i,1}}, \ldots, v_{i,t_{i,n_i}}\right)$ and ι_{n_i} is a vector of ones of dimension n_i. In vector forms, the disturbance term in Equation (12) can be written as

$$u = diag\left[\iota_{n_i}\right]\mu + \nu, \tag{13}$$

where $u = (u_1, \ldots, u_N)$, $\mu = (\mu_1, \ldots, \mu_N)$ and $v' = (v_1', \ldots, v_N')$. The variance-covariance matrix of u is $\Omega = E(uu') = diag[\Lambda_i]$, where $\Lambda_i = E(u_i u_i') = \sigma_\mu^2 J_{n_i} + V_i$, J_{n_i} is a matrix of ones of dimension n_i, and $V_i = E(v_i v_i')$. For any two observed periods, say $t_{i,j}$ and $t_{i,l}$, the covariance term is given by $cov\left(v_{i,t_{i,j}}, v_{i,t_{i,l}}\right) = \sigma_\epsilon^2 \rho^{\left|t_{i,j}-t_{i,l}\right|} / (1-\rho^2)$ for $j, l = 1, \ldots, n_i$. To remove the serial correlation in v_{it} and keep it homoskedastic, Baltagi and Wu (1999) introduced an $n_i \times n_i$ transformation matrix $C_i^*(\rho)$, which is given by

$$\begin{aligned} C_i^*(\rho) &= \left(1-\rho^2\right)^{1/2} \\ &\times \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \frac{-\rho^{t_{i,2}-t_{i,1}}}{\left(1-\rho^{2(t_{i,2}-t_{i,1})}\right)^{1/2}} & \frac{1}{\left(1-\rho^{2(t_{i,2}-t_{i,1})}\right)^{1/2}} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{-\rho^{t_{i,n_i}-t_{i,n_i-1}}}{\left(1-\rho^{2\left(t_{i,n_i}-t_{i,n_i-1}\right)}\right)^{1/2}} & \frac{1}{\left(1-\rho^{2\left(t_{i,n_i}-t_{i,n_i-1}\right)}\right)^{1/2}} \end{pmatrix}. \end{aligned} \tag{14}$$

Premultiplying Equation (12) by $C_i^*(\rho)$, we get the transformed error

$$u_i^* = C_i^*(\rho)\, u_i = \mu_i g_i + C_i^*(\rho)\, \nu_i, \tag{15}$$

where

$$g_i = C_i^*(\rho)\, \iota_{n_i} = \left(1-\rho^2\right)^{1/2} \left(1, \frac{1-\rho^{t_{i,2}-t_{i,1}}}{\left(1-\rho^{2(t_{i,2}-t_{i,1})}\right)^{1/2}}, \cdots, \frac{1-\rho^{t_{i,n_i}-t_{i,n_i-1}}}{\left(1-\rho^{2\left(t_{i,n_i}-t_{i,n_i-1}\right)}\right)^{1/2}}\right). \tag{16}$$

Baltagi and Wu (1999) showed that $C_i^*(\rho)\,\nu_i \sim (0, \sigma_\epsilon^2 I_{n_i})$, i.e., $C_i^*(\rho)\, V_i C_i^*(\rho)' = \sigma_\epsilon^2 I_{n_i}$. The variance-covariance matrix for the transformed disturbance $u^* = (u_1^*, \ldots, u_N^*)$ is $\Omega^* = diag[\Lambda_i^*]$, where

$$\Lambda_i^* = C_i^*(\rho)\, \Lambda_i C_i^*(\rho)' = \sigma_\mu^2 g_i g_i' + \sigma_\epsilon^2 I_{n_i} = \omega_i^2 P_{g_i} + \sigma_\epsilon^2 Q_{g_i}, \tag{17}$$

with $\omega_i^2 = g_i' g_i \sigma_\mu^2 + \sigma_\epsilon^2$, $P_{g_i} = g_i \left(g_i' g_i\right)^{-1} g_i'$, $Q_{g_i} = I_{n_i} - P_{g_i}$ and I_{n_i} is an identity matrix of dimension n_i. Using the fact that P_{g_i} and Q_{g_i} are idempotent matrices which are orthogonal to each other, we have

$$\Lambda_i^{*-1/2} = \left(\omega_i^2\right)^{-1/2} P_{g_i} + \left(\sigma_\epsilon^2\right)^{-1/2} Q_{g_i} = \left(\sigma_\epsilon^2\right)^{-1/2} I_{n_i} - \left[\left(\sigma_\epsilon^2\right)^{-1/2} - \left(\omega_i^2\right)^{-1/2}\right] P_{g_i}. \quad (18)$$

Hence, $\sigma_\epsilon \Omega^{*-1/2} = diag\left[\sigma_\epsilon \Lambda_i^{*-1/2}\right]$, where $\sigma_\epsilon \Lambda_i^{*-1/2} = I_{n_i} - \theta_i P_{g_i}$ and $\theta_i = 1 - \sigma_\epsilon / \omega_i$. Premultiplying $y^* = diag\left[C_i^*(\rho)\right] y$ by $\sigma_\epsilon \Omega^{*-1/2}$, one gets $y^{**} = \sigma_\epsilon \Omega^{*-1/2} y^*$. The elements of y^{**} are given by

$$y_{i,t_{i,j}}^{**} = y_{i,t_{i,j}}^{*} - \theta_i g_{i,j} \frac{\sum_{s=1}^{n_i} g_{i,s} y_{i,t_{i,s}}^{*}}{\sum_{s=1}^{n_i} g_{i,s}^2}. \quad (19)$$

Baltagi and Wu (1999) proposed estimating σ_μ^2 and σ_ϵ^2 by

$$\hat{\sigma}_\mu^2 = \frac{u^{*\prime} diag\left[P_{g_i}\right] u^* - N\hat{\sigma}_\epsilon^2}{\sum_{i=1}^{N} g_i' g_i} \text{ and } \hat{\sigma}_\epsilon^2 = \frac{u^{*\prime} diag\left[Q_{g_i}\right] u^*}{\sum_{i=1}^{N} (n_i - 1)}. \quad (20)$$

Since the true disturbances u^* are unknown, we use $\tilde{u}_{OLS}^*$ instead, which are the OLS residuals from the (*) transformed equation. In order to make the (*) transformation operational, we need an estimate of ρ. Let $\tilde{v}$ be the within residuals from y on X. Inserting zeros between $\tilde{v}_{i,t_{i,j}}$ and $\tilde{v}_{i,t_{i,j+1}}$ if the data between these two periods are not available, one gets a new $T \times 1$ residual e_i. An estimate of ρ can be obtained as

$$\hat{\rho} = \frac{\frac{1}{m}\sum_{i=1}^{N}\sum_{t=2}^{T} e_{it} e_{i,t-1}}{\frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T} e_{it}^2}, \quad (21)$$

where $m = \sum_{i=1}^{N} m_i$, m_i is the number of observed consecutive pairs for each individual i and $n = \sum_{i=1}^{N} n_i$.

Theorem 1 *Assume that (i)* $\epsilon_{it} \sim iid(0, \sigma^2)$; *(ii)* $\frac{1}{N}\sum_{i=1}^{N} v_{i0}^2 = O(1)$; *(iii)* $\frac{1}{N}\sum_{i=1}^{N} \mu_i^2 = O(1)$; *(iv)* $\frac{N}{m} \to 0$. *We have* $\hat{\rho} - \rho = o_p(1)$.

The proof is given in the Appendix. Assumptions (i), (ii) and (iii) were used in Hahn and Kuersteiner (2002). Assumption (iv) $\frac{N}{m} \to 0$ is equivalent to $\frac{m}{N} = \frac{1}{N}\sum_{i=1}^{N} m_i \to \infty$. The consistency of $\hat{\rho}$ requires the average number of observed consecutive pairs to be large. For balanced panel data, this condition reduces to $T \to \infty$. Using this estimator of ρ, one gets a feasible GLS estimator of β. Detailed steps can be found in Baltagi and Wu (1999).[3]

Now, we return to prediction. Using the fact that the disturbances are independent across different individuals, we have $w' = E(u_{i,T+S}u') = (0, .., E(u_{i,T_i+S}u_i'), 0, .., 0)$, which is a vector of zeros except for the ith position. Therefore,

$$w'\Omega^{-1} = (0, .., E(u_{i,T_i+S}u_i'), 0, .., 0)\, diag\left[\Lambda_i^{-1}\right] = \left(0, .., E(u_{i,T_i+S}u_i')\,\Lambda_i^{-1}, 0, .., 0\right) \tag{22}$$

and

$$w'\Omega^{-1}\hat{u}_{GLS} = \left(0, .., E(u_{i,T_i+S}u_i')\,\Lambda_i^{-1}, 0, .., 0\right)\begin{pmatrix}\hat{u}_1\\ \hat{u}_2\\ \vdots\\ \hat{u}_N\end{pmatrix} = E(u_{i,T_i+S}u_i')\,\Lambda_i^{-1}\hat{u}_i, \tag{23}$$

where $u_i' = \left(u_{i,t_{i,1}}, \ldots, u_{i,t_{i,n_i}}\right)$ and $\widehat{u}_i$ denote the GLS residuals. Since $u_{i,T_i+S} = \mu_i + \nu_{i,T_i+S}$, we can decompose equation (23) into two terms:

$$E(u_{i,T_i+S}u_i')\,\Lambda_i^{-1}\hat{u}_i = E(\mu_i u_i')\,\Lambda_i^{-1}\hat{u}_i + E(v_{i,T_i+S}u_i')\,\Lambda_i^{-1}\hat{u}_i. \tag{24}$$

Since $\Lambda_i^* = C_i^*(\rho)\,\Lambda_i C_i^*(\rho)'$, we have

$$\Lambda_i^{-1} = C_i^*(\rho)'\,\Lambda_i^{*-1}C_i^*(\rho) = C_i^*(\rho)'\left(\omega_i^{-2}P_{g_i} + \sigma_\epsilon^{-2}Q_{g_i}\right)C_i^*(\rho) \tag{25}$$

using Equation (18). Since μ_i and v_i are independent of each other, we have $E(\mu_i u_i') = E\left(\mu_i\mu_i\iota_{n_i}'\right) = \sigma_\mu^2\iota_{n_i}'$. The first term in equation (24) can be rewritten as:

$$\begin{aligned} & E(\mu_i u_i')\,\Lambda_i^{-1}\hat{u}_i \\ = \;& \sigma_\mu^2\iota_{n_i}'C_i^*(\rho)'\left(\omega_i^{-2}P_{g_i} + \sigma_\epsilon^{-2}Q_{g_i}\right)C_i^*(\rho)\,\hat{u}_i \\ = \;& \frac{\sigma_\mu^2}{\omega_i^2}g_i'\hat{u}_i^*, \end{aligned} \tag{26}$$

[3] It is important to note that this is easily programmable. In fact, the Baltagi and Wu (1999) feasible GLS procedure has been implemented in Stata using xtregar, so it is easy to derive the BLUP from these results.

where $C_i^*(\rho)\hat{u}_i = \hat{u}_i^*$, using the fact $C_i^*(\rho)\iota_{n_i} = g_i$, $g_i' P_{g_i} = g_i'$ and $g_i' Q_{g_i} = 0$. By continuous substitution, we have

$$v_{i,T_i+S} = \rho^S v_{i,T_i} + \rho^{S-1}\epsilon_{i,T_i+1} + \cdots + \epsilon_{i,T_i+S}$$

and

$$E\left(v_{i,T_i+S}u_i'\right) = E\left(v_{i,T_i+S}v_i'\right) = E\left[\left(\rho^S v_{i,T_i} + \rho^{S-1}\epsilon_{i,T_i+1} + \cdots + \epsilon_{i,T_i+S}\right)v_i'\right] = \rho^S E\left(v_{i,T_i}v_i'\right)$$

since $E\left[\epsilon_{i,T_i+1}v_i'\right] = \cdots = E\left[\epsilon_{i,T_i+S}v_i'\right] = 0$. Because $E\left(v_{i,T_i}v_i'\right)$ is the last column of the covariance matrix $E\left(v_i v_i'\right) = V_i$, we have

$$E\left(v_{i,T+S}u_i'\right) = \rho^S(0,\cdots,0,1)V_i.$$

Also, $\mathbf{\Lambda}_i^{-1}$ in Equation (25) reduces to

$$\begin{aligned}\Lambda_i^{-1} &= C_i^*(\rho)'\left(\omega_i^{-2}P_{g_i} + \sigma_\epsilon^{-2}Q_{g_i}\right)C_i^*(\rho)\\ &= C_i^*(\rho)'\left[\sigma_\epsilon^{-2}I_{n_i} - \left(\sigma_\epsilon^{-2} - \omega_i^{-2}\right)P_{g_i}\right]C_i^*(\rho)\\ &= C_i^*(\rho)'\left[\sigma_\epsilon^{-2}I_{n_i} - \left(\frac{g_i' g_i \sigma_\mu^2}{\sigma_\epsilon^2\omega_i^2}\right)g_i\left(g_i' g_i\right)^{-1}g_i'\right]C_i^*(\rho)\\ &= \sigma_\epsilon^{-2}C_i^*(\rho)'C_i^*(\rho)\left[I_{n_i} - \frac{\sigma_\mu^2}{\omega_i^2}\iota_{n_i}g_i' C_i^*(\rho)\right]\end{aligned}$$

using the fact that $Q_{g_i} = I_{n_i} - P_{g_i}$, $\omega_i^2 = g_i' g_i \sigma_\mu^2 + \sigma_\epsilon^2$ and $g_i = C_i^*(\rho)\iota_{n_i}$. The second term in equation (24) becomes:

$$\begin{aligned}&E\left(v_{i,T_i+S}u_i'\right)\Lambda_i^{-1}\hat{u}_i\\ =\;& \rho^S(0,\cdots,0,1)V_i\sigma_\epsilon^{-2}C_i^*(\rho)'C_i^*(\rho)\left[I_{n_i} - \frac{\sigma_\mu^2}{\omega_i^2}\iota_{n_i}g_i' C_i^*(\rho)\right]\hat{u}_i\\ =\;& \rho^S(0,\cdots,0,1)\left(\hat{u}_i - \frac{\sigma_\mu^2}{\omega_i^2}\iota_{n_i}g_i'\hat{u}_i^*\right)\\ =\;& \rho^S\hat{u}_{i,T_i} - \frac{\rho^S\sigma_\mu^2}{\omega_i^2}g_i'\hat{u}_i^* \qquad (27)\end{aligned}$$

using the fact that $\sigma_\epsilon^{-2}V_i = \left[C_i^*(\rho)'C_i^*(\rho)\right]^{-1}$ since $C_i^*(\rho)V_iC_i^*(\rho)' = \sigma_\epsilon^2 I_{n_i}$. Combining equations (26) and (27), one gets

$$
\begin{aligned}
& w'\Omega^{-1}\hat{u}_{GLS} \\
= \;& \rho^S \hat{u}_{i,T_i} + \frac{\left(1-\rho^S\right)\sigma_\mu^2}{\omega_i^2} g_i' \hat{u}_i^* \\
= \;& \rho^S \hat{u}_{i,T_i} + \frac{\left(1-\rho^S\right)\left(1-\rho^2\right)^{1/2}\sigma_\mu^2}{\omega_i^2}\left[\hat{u}_{i,t_{i,1}}^* + \sum_{j=2}^{n_i} \frac{1-\rho^{t_{i,j}-t_{i,j-1}}}{\left(1-\rho^{2(t_{i,j}-t_{i,j-1})}\right)^{1/2}} \hat{u}_{i,t_{i,j}}^*\right]. \quad (28)
\end{aligned}
$$

Special case 1: *No missing observations.* This is the balanced panel data model with AR(1) remainder disturbance terms considered by Baltagi and Li (1992). In this case, we have $t_{i,j} - t_{i,j-1} = 1$, $T_i = n_i = T$,

$$
g_i = \left(1-\rho^2\right)^{1/2}\left(1, \frac{1-\rho}{\left(1-\rho^2\right)^{1/2}}, \cdots, \frac{1-\rho}{\left(1-\rho^2\right)^{1/2}}\right) = (1-\rho)\,\iota_T^\alpha,
$$

where $\iota_T^\alpha = (\alpha, 1, \cdots, 1)$ with $\alpha = \sqrt{(1+\rho)/(1-\rho)}$.

$$
g_i' g_i = (1-\rho)^2 d^2,
$$

and $d^2 = \alpha^2 + T - 1$. Hence $\omega_i^2 = \sigma_\alpha^2$, where $\sigma_\alpha^2 = (1-\rho)^2 d^2 \sigma_\mu^2 + \sigma_\epsilon^2$.

$$
\frac{1-\rho^{t_{i,j}-t_{i,j-1}}}{\left(1-\rho^{2(t_{i,j}-t_{i,j-1})}\right)^{1/2}} = \frac{1-\rho}{\left(1-\rho^2\right)^{1/2}},
$$

$\widehat{u}_i^* = C\widehat{u}_i$, where C is the $T \times T$ Prais-Winsten (PW) transformation matrix

$$
C = \begin{bmatrix}
(1-\rho^2)^{1/2} & 0 & 0 & \cdots & 0 & 0 \\
-\rho & 1 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & 0 \\
0 & 0 & 0 & \cdots & -\rho & 1
\end{bmatrix}.
$$

Therefore, Equation (28) reduces to

$$
w'\Omega^{-1}\widehat{u}_{GLS} = \rho^S \widehat{u}_{i,T} + \frac{(1-\rho)\left(1-\rho^S\right)\sigma_\mu^2}{\sigma_\alpha^2}\left(\alpha\widehat{u}_{i1}^* +_{t=2}^{T} \widehat{u}_{it}^*\right).
$$

This is Goldberger's BLUP *extra term* derived by Baltagi and Li (1992). So, the unbalanced panel Goldberger's BLUP correction term reduces to its balanced panel counterpart in the case of AR(1) remainder disturbance terms.

Special case 2: *No random effects.* This reduces to a panel data model without individual effects, but with AR(1) remainder disturbances. In this case $\sigma_\mu^2 = 0$, and equation (28) reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \rho^S \hat{u}_{i,T_i}. \tag{29}$$

This is Goldberger's BLUP *extra term* for the unbalanced panel data model with AR(1) remainder disturbances but no random individual effects. Goldberger (1962) actually considered a simple time series regression (not a panel) with AR(1) remainder disturbances.

Special case 3: *No serial correlation.* This is the unbalanced random effects model without serial correlation in Section 1. In this case $\rho = 0$, $g_i = \iota_{n_i}$, $g_i' g_i = n_i$, $\omega_i^2 = n_i\sigma_\mu^2 + \sigma_\epsilon^2$ and $\hat{u}_{it}^* = \hat{u}_{it}$. Equation (28) in this case reduces to

$$w'\Omega^{-1}\hat{u}_{GLS} = \frac{\sigma_\mu^2}{\omega_i^2}\sum_{j=1}^{n_i}\hat{u}_{i,t_{i,j}} = \frac{n_i\sigma_\mu^2}{\omega_i^2}\overline{\overline{\hat{u}}}_{i.,GLS}, \tag{30}$$

where $\overline{\overline{\hat{u}}}_{i.,GLS} = n_i^{-1}\sum_{j=1}^{n_i}\hat{u}_{i,t_{i,j}}$. This is Goldberger's BLUP *extra term* for the unequally spaced panel data model with no serial correlation. This encompasses the case derived in Section 1 with $n_i = T_i$, $\omega_i^2 = T_i\sigma_\mu^2 + \sigma_\epsilon^2$ and the extra BLUP Goldberger (1962) term reduces to the one given in Equation (10).

4 Monte Carlo Simulation

To study the finite sample performance of the proposed estimator of ρ as well as the performance of the corresponding predictors, we perform Monte Carlo experiments in this section. Following Baltagi, Chang and Li (1992) but with random effects, we generate the following panel model

$$y_{it} = 1 + x_{it} + \mu_i + v_{it}, \tag{31}$$

for $i = 1, \ldots, N$; $t = 1\ldots, T+1$, where $x_{it} = 0.1t + 0.5x_{i,t-1} + w_{it}$. w_{it} follows a uniform distribution $[-0.5, 0.5]$ and $x_{i0} = 5 + 10w_{i0}$. The individual specific effects are generated as $\mu_i \overset{iid}{\sim} N(0, 10)$ and the remainder error follows an AR(1) process $v_{it} = \rho v_{i,t-1} + \epsilon_{it}$, where $\epsilon_{it} \overset{iid}{\sim} N(0,1)$ and ρ takes the values $\{0, 0.3, 0.6, 0.9\}$. As pointed out by Baltagi et al. (1992), one can translate this starting date into an "effective" initial variance assumption regardless of when the AR(l) process started. More specifically, to check the impact the of the initial condition, we let $v_{i0} \overset{iid}{\sim} N\left(0, \tau/\left(1-\rho^2\right)\right)$ where τ varies over the set $\{0.2, 1, 5\}$. We generate the estimation sample such that the average time period observed is $\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T_i = 5$, 10, 20 or 40. As shown in Table 1, we consider four different unbalanced panel data designs that are similar to those in Bruno (2005). In each design, the Ahrens and Pincus (1981) index ω, which measures the extent of unbalancedness, is set to be 0.36 or 0.96.[4] In all experiments, the number of individuals is always $N = 50$. We perform 1,000 replications for each experiment.

Table 2 reports the bias, interquantile range (IQR), and root mean squared error (RMSE) of the estimator of ρ. Following Kelejian and Prucha (1999), *bias* is calculated as the difference between the median and the true parameter value; IQR is the difference between the 0.75 and 0.25 quantiles; and $RMSE = \left[bias^2 + (IQR/1.35)^2\right]^{1/2}$. These measures are always assured to exist, see Kelejian and Prucha (1999) for details. As shown in Table 2, when $\bar{T}$ is small, $\hat{\rho}$ has negative bias. However, the bias shrinks as $\bar{T}$ increases. When $\rho > 0$, the bias, IQR and RMSE all decrease when τ increases.

Tables 3-5 report the prediction performance of the following estimators: the pooled ordinary least squares (OLS), panel fixed-effects (FE) and random effects (RE) estimators that ignore autocorrelations in the error terms, and the fixed-effects and random effects estimators with AR(1) term, which are denoted as FEAR and REAR respectively. To summarize the accuracy of the forecasts, following Baltagi and Liu (2013a), we report the sampling *mean square error (MSE)*, the *mean absolute error (MAE)* and the *mean absolute percentage error (MAPE)*, which are computed as

$$MSE = \frac{1}{NR}\sum_{r=1}^{R}\sum_{i=1}^{N} d_{i,T_i+S_i}^2, \tag{32}$$

[4]See also Baltagi and Chang (1995) for more discussion on incomplete panels and this Ahrens and Pincus measure. Note that $\omega = N/(\bar{T}\sum_{i=1}^{N} T_i^{-1})$, with $0 < \omega \leq 1$. When the panel data is balanced $\omega = 1$. When the panel data is unbalanced ω takes on smaller values.

$$MAE = \frac{1}{NR}\sum_{r=1}^{R}\sum_{i=1}^{N}|d_{i,T_i+S_i}| \tag{33}$$

and

$$MAPE = \frac{100}{NR}\sum_{r=1}^{R}\sum_{i=1}^{N}\left|\frac{d_{i,T_i+S_i}}{y_{i,T_i+S_i}}\right|, \tag{34}$$

where $d_{i,T_i+S_i} = \hat{y}_{i,T_i+S_i} - y_{i,T_i+S_i}$, $R = 1,000$ replications and we forecast the last year available for individual i.[5] As shown in Tables 3-5, REAR usually has the smallest MSE and MAE when $\rho > 0$. However, FEAR sometimes has a smaller MAPE than REAR even though the true DGP is created to be a random effect model with an AR(1) error term.

5 Application

In this section we illustrate the BLUP forecasts using an extract from the National Longitudinal Study data set employed by Drukker (2003). This is an unbalanced panel data over the years 1968-1988 with gaps. The data is used to illustrate the xtreg command in Stata and includes observations on wages for 4711 young working women who were 14–26 years of age in 1968, some with only one observation. We regressed the logarithm of wage (`lnwage`) on the woman's age and its square (`age`, `age2`), total working experience (`exp`), tenure at current position and its square (`tenure`, `tenure2`), current grade completed (`grade`), a dummy variable for not living in a standard metropolitan statistical area (`nsmsa`), a dummy variable for living in the south (`south`) and a dummy variable for black (`black`).[6] we estimate the model by using the pooled OLS, FE, RE,

[5] It is worth pointing out that forecasting is not always one period ahead, as it varies by individual depending on the missing observations. In fact, the last available year for a particular individual could sometimes be several years ahead due to irregular gaps of missing data between years. This is why we gave the expression for the BLUP forecast for S_i periods ahead for individual i.

[6] Drukker (2003) uses this data to estimate an earnings equation to illustrate a test for serial correlation proposed by Wooldridge (2002). Experience squared was not significant and was dropped from the regression. Zero serial correlation of the first order was rejected.

FEAR and REAR respectively. In order to compute the forecasts, we focus on women who had records for at least three years. For each estimator, we compute the forecast of the logarithm of wage for the last available year for that individual. This year is not used in the estimation but is used in the computation of the three forecast performance measures. To summarize the accuracy of the forecasts, we report MSE, MAE and MAPE, which are defined in Equation (32)-(34) with $R = 1$. As shown in Table 6, the random effects model with an AR(1) term has the smallest MSE or MAE. While, the fixed-effects model with an AR(1) term has the smallest MAPE. This is consistent with the findings in the simulation results. For time series data sets, Diebold and Mariano (1995) derived a test to compare prediction accuracy. Recently, Timmermann and Zhu (2019) extend the Diebold and Mariano (1995) test to panel data to compare the significance of pairwise forecasts averaged over all cross-sectional units. The results of this panel data test of equal predictive accuracy is reported in Table 7. Overall, the random effects model with an AR(1) term predicts significantly better than all other models.

6 Conclusion

This paper derives the BLUP for the unbalanced panel data model and the unequally spaced panel data model with AR(1) remainder disturbances and illustrates these with an earnings equation using the NLS young women data over the period 1968-1988 employed by Drukker (2003) using Stata. These results can be extended to the unbalanced panel data model with AR(p) remainder disturbances, see Baltagi and Liu (2013a) for the corresponding balanced panel data case. Also, the unbalanced panel data model with MA(q) remainder disturbances, see Baltagi and Liu (2013b) for the corresponding balanced panel data case. Another extension is for the autoregressive moving average ARMA(p, q) remainder disturbances, see Galbraith and Zinde-Walsh (1995) for the balanced panel data case.

Data Availability Statement

The data used in the paper are available on the Stata web site for all Stata users.

References

Ahrens, H. and R. Pincus, 1981, On two measures of unbalancedness in a one-way model and their relation to efficiency, Biometric Journal, 23, 227-235.

Baltagi, B.H., 2013a. Econometric analysis of panel data, Wiley and Sons, Chichester.

Baltagi, B.H., 2013b, Panel data forecasting, chapter 18 in the handbook of economic forecasting, Volume 2B, edited by Graham Elliott and Allan Timmermann, North Holland, Amsterdam, 995-1024.

Baltagi, B.H. and Y.J., Chang, 1995, Incomplete panels, Journal of Econometrics, 62, 67–89.

Baltagi, B.H., Chang, Y.J., and Q. Li, 1992, Monte Carlo evidence on panel data regressions with AR (1) disturbances and an arbitrary variance on the initial observations, Journal of Econometrics, 52(3), 371-380.

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 L. Liu, 2013a, Estimation and prediction in the random effects model with AR(p) remainder disturbances, International Journal of Forecasting 29, 100-107.

Baltagi, B.H. and L. Liu, 2013b, Prediction in the random effects model with MA(q) remainder disturbances, Journal of Forecasting 32, 333-338.

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

Bruno, G.S., 2005, Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models, Economics letters, 87(3), 361-366.

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

Drukker, D.M. 2003, Testing for serial correlation in linear panel-data models, Stata Journal, 3(2), 168-177.

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.

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

Hahn, J. and G. Kuersteiner, 2002, Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large, Econometrica, 70(4), 1639-1657.

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

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

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.

Kelejian, H.H. and I.R. Prucha, 1999, A generalized moments estimator for the autoregressive parameter in a spatial model, International economic review, 40(2), 509-533.

Little, R. J. A., and D. B. Rubin, 2002, Statistical Analysis with Missing Data, John Wiley, New Jersey.

McKenzie, D.J., 2001, Estimation of AR(1) models with unequally-spaced pseudo-panels, Econometrics Journal, 4, 89–108.

Millimet, D. L. and I.K. McDonough, 2017, Dynamic panel data models with irregular spacing: with an application to early childhood development, Journal of Applied Econometrics 32, 725–743.

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

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

Timmermann, A. and Y. Zhu, 2019, Comparing forecasting performance with panel data, SSRN paper 3380755.

Wansbeek, T.J. and A. Kapteyn, 1982, A simple way to obtain the spectral decomposition of variance components models for balanced data, Communications in Statistics A11, 2105–2112.

Wansbeek, T.J. and A. Kapteyn, 1989, Estimation of the error components model with incomplete panels, Journal of Econometrics 41, 341–361.

Wooldridge, J.M. 2002, Econometric analysis of cross section and panel data, Cambridge, MIT Press.

Table 1: Unbalanced Design

$\bar{T}$	T_i	ω	S_i	$\bar{S}$
5	$4(i \leq 25), 6(i > 25)$	0.96	$3(i \leq 25), 1(i > 25)$	2
	$1(i \leq 25), 9(i > 25)$	0.36	$9(i \leq 25), 1(i > 25)$	5
10	$8(i \leq 25), 12(i > 25)$	0.96	$5(i \leq 25), 1(i > 25)$	3
	$2(i \leq 25), 18(i > 25)$	0.36	$17(i \leq 25), 1(i > 25)$	9
20	$16(i \leq 25), 24(i > 25)$	0.96	$9(i \leq 25), 1(i > 25)$	5
	$4(i \leq 25), 36(i > 25)$	0.36	$33(i \leq 25), 1(i > 25)$	17
40	$32(i \leq 25), 48(i > 25)$	0.96	$17(i \leq 25), 1(i > 25)$	9
	$8(i \leq 25), 72(i > 25)$	0.36	$65(i \leq 25), 1(i > 25)$	33

Note: $N = 50$ for all experiments. T_i is the available years for each individual i and $\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T_i$. $\omega = N/(\bar{T}\sum_{i=1}^{N} T_i^{-1})$ is the Ahrens and Pincus (1981) measure of unbalancedness. We forecast S_i years ahead for each individual i and $\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i$.

Table 2: Bias, IQR, and RMSE of the Estimator of ρ

$\bar{T}$	ω	ρ	τ	Bias	IQR	RMSE
5	0.96	0	0.2	-0.202	0.080	0.210
			1	-0.202	0.080	0.210
			5	-0.202	0.080	0.210
		0.3	0.2	-0.297	0.084	0.303
			1	-0.291	0.084	0.297
			5	-0.220	0.079	0.227
		0.6	0.2	-0.433	0.084	0.437
			1	-0.411	0.080	0.416
			5	-0.217	0.052	0.220
		0.9	0.2	-0.595	0.072	0.597
			1	-0.570	0.067	0.572
			5	-0.390	0.034	0.391
	0.36	0	0.2	-0.130	0.066	0.139
			1	-0.130	0.066	0.139
			5	-0.130	0.066	0.139
		0.3	0.2	-0.183	0.066	0.189
			1	-0.182	0.067	0.188
			5	-0.143	0.062	0.150
		0.6	0.2	-0.266	0.063	0.270
			1	-0.252	0.060	0.256
			5	-0.118	0.045	0.123
		0.9	0.2	-0.398	0.057	0.400
			1	-0.372	0.054	0.374
			5	-0.219	0.026	0.220
10	0.96	0	0.2	-0.093	0.054	0.101
			1	-0.093	0.054	0.101
			5	-0.093	0.054	0.101
		0.3	0.2	-0.130	0.055	0.136
			1	-0.129	0.056	0.135
			5	-0.106	0.053	0.113
		0.6	0.2	-0.188	0.054	0.192
			1	-0.179	0.054	0.183
			5	-0.081	0.042	0.087
		0.9	0.2	-0.297	0.048	0.299
			1	-0.272	0.043	0.274
			5	-0.142	0.021	0.143
	0.36	0	0.2	-0.060	0.047	0.069
			1	-0.060	0.047	0.069
			5	-0.060	0.047	0.069
		0.3	0.2	-0.082	0.047	0.089
			1	-0.082	0.047	0.089
			5	-0.071	0.044	0.078
		0.6	0.2	-0.114	0.045	0.119
			1	-0.111	0.043	0.115
			5	-0.057	0.034	0.062
		0.9	0.2	-0.192	0.038	0.194
			1	-0.175	0.034	0.176
			5	-0.076	0.016	0.076
20	0.96	0	0.2	-0.045	0.037	0.053
			1	-0.045	0.037	0.053
			5	-0.045	0.037	0.053
		0.3	0.2	-0.060	0.037	0.067
			1	-0.060	0.037	0.066
			5	-0.053	0.037	0.060
		0.6	0.2	-0.082	0.037	0.086
			1	-0.080	0.035	0.084
			5	-0.046	0.030	0.051

Continued on Next Page. . .

Table 2 – Continued

$\bar{T}$	ω	ρ	τ	Bias	IQR	RMSE
		0.9	0.2	-0.140	0.028	0.141
			1	-0.126	0.028	0.128
			5	-0.047	0.013	0.048
	0.36	0	0.2	-0.035	0.036	0.044
			1	-0.035	0.036	0.044
			5	-0.035	0.036	0.044
		0.3	0.2	-0.047	0.035	0.054
			1	-0.047	0.036	0.054
			5	-0.043	0.034	0.050
		0.6	0.2	-0.062	0.032	0.066
			1	-0.061	0.032	0.065
			5	-0.038	0.025	0.042
		0.9	0.2	-0.102	0.024	0.104
			1	-0.093	0.024	0.095
			5	-0.031	0.013	0.033
40	0.96	0	0.2	-0.028	0.033	0.037
			1	-0.028	0.033	0.037
			5	-0.028	0.033	0.037
		0.3	0.2	-0.039	0.033	0.046
			1	-0.039	0.033	0.046
			5	-0.036	0.033	0.043
		0.6	0.2	-0.050	0.028	0.054
			1	-0.049	0.029	0.053
			5	-0.033	0.025	0.038
		0.9	0.2	-0.079	0.023	0.081
			1	-0.072	0.022	0.074
			5	-0.025	0.013	0.026
	0.36	0	0.2	-0.021	0.028	0.029
			1	-0.021	0.028	0.029
			5	-0.021	0.028	0.029
		0.3	0.2	-0.028	0.028	0.035
			1	-0.028	0.028	0.035
			5	-0.026	0.026	0.032
		0.6	0.2	-0.037	0.025	0.041
			1	-0.036	0.025	0.040
			5	-0.026	0.021	0.031
		0.9	0.2	-0.052	0.018	0.054
			1	-0.048	0.018	0.050
			5	-0.019	0.012	0.021

Note: $N = 50$ for all experiments. $\tau/(1-\rho^2)$ is the variance of the initial condition.

Table 3: MSE of the Predictors

$\bar{T}$	ω	ρ	τ	OLS	FE	RE	FEAR	REAR
5	0.96	0	0.2	20.062	11.659	11.455	12.040	11.977
			1	20.062	11.659	11.455	12.040	11.977
			5	20.062	11.659	11.455	12.040	11.977
		0.3	0.2	21.034	12.393	12.101	12.405	12.102
			1	21.036	12.445	12.142	12.433	12.116
			5	21.070	13.773	13.120	13.194	12.525
		0.6	0.2	25.578	14.827	14.490	13.083	12.533
			1	25.602	15.484	15.029	13.341	12.663
			5	26.237	31.991	27.997	18.936	15.547
		0.9	0.2	50.502	19.448	19.692	14.132	14.242
			1	61.731	22.006	21.678	15.226	14.428
			5	346.585	85.625	82.645	36.620	21.583
	0.36	0	0.2	19.712	11.113	10.988	11.282	11.200
			1	19.712	11.113	10.988	11.282	11.200
			5	19.712	11.113	10.988	11.282	11.200
		0.3	0.2	20.734	12.006	11.823	11.437	11.201
			1	20.737	12.035	11.847	11.452	11.213
			5	20.784	12.677	12.360	11.780	11.468
		0.6	0.2	25.418	15.550	15.263	11.863	11.344
			1	25.437	15.942	15.594	11.948	11.391
			5	25.871	25.493	23.210	13.917	12.439
		0.9	0.2	56.347	24.022	24.166	12.756	12.712
			1	62.653	27.480	27.108	13.492	12.661
			5	215.656	114.272	110.148	26.889	15.170
10	0.96	0	0.2	20.011	10.855	10.799	10.960	10.922
			1	20.011	10.855	10.799	10.960	10.922
			5	20.011	10.855	10.799	10.960	10.922
		0.3	0.2	20.986	11.815	11.714	11.036	10.902
			1	20.987	11.835	11.731	11.046	10.912
			5	21.041	12.241	12.079	11.251	11.109
		0.6	0.2	25.539	15.831	15.619	11.279	10.932
			1	25.558	16.096	15.849	11.324	10.966
			5	26.005	22.184	20.820	12.364	11.637
		0.9	0.2	58.684	27.884	27.929	12.043	11.925
			1	61.969	32.137	31.712	12.571	11.840
			5	146.389	135.067	129.994	21.495	13.263
	0.36	0	0.2	20.064	10.603	10.583	10.646	10.632
			1	20.064	10.603	10.583	10.646	10.632
			5	20.064	10.603	10.583	10.646	10.632
		0.3	0.2	21.009	11.558	11.513	10.672	10.617
			1	21.011	11.563	11.518	10.675	10.620
			5	21.039	11.722	11.660	10.756	10.714
		0.6	0.2	25.491	15.911	15.780	10.770	10.611
			1	25.500	16.008	15.865	10.782	10.623
			5	25.668	18.610	18.056	11.194	10.980
		0.9	0.2	60.888	33.890	33.795	11.311	11.128
			1	61.863	38.042	37.594	11.524	11.033
			5	86.469	142.262	136.592	15.610	11.583
20	0.96	0	0.2	19.827	10.441	10.425	10.461	10.447
			1	19.827	10.441	10.425	10.461	10.447
			5	19.827	10.441	10.425	10.461	10.447
		0.3	0.2	20.780	11.385	11.354	10.472	10.435
			1	20.780	11.389	11.358	10.474	10.438
			5	20.786	11.483	11.443	10.522	10.497
		0.6	0.2	25.374	15.910	15.818	10.519	10.431
			1	25.376	15.974	15.875	10.530	10.442
			5	25.426	17.472	17.172	10.769	10.682

Continued on Next Page. . .

Table 3 – Continued

$\bar{T}$	ω	ρ	τ	OLS	FE	RE	FEAR	REAR
		0.9	0.2	61.796	38.709	38.521	10.912	10.706
			1	62.118	42.446	41.993	11.025	10.642
			5	69.479	135.669	130.267	13.214	10.911
	0.36	0	0.2	19.978	10.315	10.308	10.248	10.246
			1	19.978	10.315	10.308	10.248	10.246
			5	19.978	10.315	10.308	10.248	10.246
		0.3	0.2	20.986	11.553	11.531	10.264	10.256
			1	20.987	11.557	11.535	10.267	10.259
			5	20.990	11.624	11.598	10.292	10.292
		0.6	0.2	25.703	17.024	16.937	10.305	10.288
			1	25.705	17.071	16.979	10.312	10.297
			5	25.740	18.133	17.931	10.415	10.426
		0.9	0.2	62.733	53.855	53.356	10.632	10.423
			1	62.795	57.725	57.037	10.671	10.382
			5	64.331	152.980	148.293	11.526	10.452
40	0.96	0	0.2	20.068	10.235	10.228	10.213	10.210
			1	20.068	10.235	10.228	10.213	10.210
			5	20.068	10.235	10.228	10.213	10.210
		0.3	0.2	21.109	11.455	11.436	10.222	10.214
			1	21.109	11.458	11.439	10.222	10.215
			5	21.110	11.513	11.492	10.232	10.229
		0.6	0.2	25.902	16.942	16.874	10.245	10.248
			1	25.904	16.980	16.909	10.247	10.252
			5	25.927	17.883	17.741	10.297	10.323
		0.9	0.2	64.074	59.933	59.390	10.469	10.331
			1	64.126	63.567	62.883	10.467	10.298
			5	65.161	154.393	150.596	10.850	10.304
	0.36	0	0.2	20.303	10.306	10.302	10.255	10.256
			1	20.303	10.306	10.302	10.255	10.256
			5	20.303	10.306	10.302	10.255	10.256
		0.3	0.2	21.371	11.531	11.520	10.260	10.269
			1	21.371	11.533	11.522	10.259	10.269
			5	21.371	11.581	11.569	10.261	10.273
		0.6	0.2	26.151	17.054	17.009	10.277	10.324
			1	26.150	17.089	17.042	10.276	10.323
			5	26.158	17.906	17.815	10.290	10.350
		0.9	0.2	63.944	63.308	62.798	10.403	10.323
			1	63.965	66.720	66.107	10.397	10.304
			5	64.759	151.605	148.949	10.533	10.289

Note: $N = 50$ for all experiments. $\tau/(1-\rho^2)$ is the variance of the initial condition.

Table 4: MAE of the Predictors

$\bar{T}$	ω	ρ	τ	OLS	FE	RE	FEAR	REAR
5	0.96	0	0.2	3.576	2.728	2.703	2.770	2.761
			1	3.576	2.728	2.703	2.770	2.761
			5	3.576	2.728	2.703	2.770	2.761
		0.3	0.2	3.660	2.809	2.775	2.810	2.774
			1	3.660	2.815	2.781	2.814	2.777
			5	3.663	2.964	2.892	2.900	2.824
		0.6	0.2	4.030	3.070	3.035	2.884	2.823
			1	4.032	3.140	3.093	2.915	2.839
			5	4.083	4.516	4.224	3.473	3.146
		0.9	0.2	5.666	3.515	3.537	2.996	3.009
			1	6.268	3.738	3.710	3.111	3.028
			5	14.855	7.384	7.254	4.836	3.711
	0.36	0	0.2	3.548	2.659	2.644	2.680	2.669
			1	3.548	2.659	2.644	2.680	2.669
			5	3.548	2.659	2.644	2.680	2.669
		0.3	0.2	3.641	2.766	2.745	2.698	2.670
			1	3.641	2.769	2.748	2.700	2.671
			5	3.646	2.841	2.805	2.737	2.699
		0.6	0.2	4.031	3.146	3.117	2.750	2.689
			1	4.032	3.184	3.150	2.758	2.694
			5	4.066	4.028	3.843	2.973	2.809
		0.9	0.2	5.996	3.919	3.931	2.853	2.850
			1	6.317	4.183	4.156	2.930	2.841
			5	11.710	8.532	8.376	4.139	3.104
10	0.96	0	0.2	3.569	2.634	2.627	2.646	2.641
			1	3.569	2.634	2.627	2.646	2.641
			5	3.569	2.634	2.627	2.646	2.641
		0.3	0.2	3.652	2.745	2.734	2.655	2.638
			1	3.652	2.747	2.735	2.656	2.639
			5	3.656	2.791	2.773	2.678	2.661
		0.6	0.2	4.032	3.178	3.157	2.682	2.641
			1	4.034	3.202	3.178	2.687	2.644
			5	4.070	3.752	3.637	2.803	2.720
		0.9	0.2	6.098	4.216	4.219	2.771	2.758
			1	6.269	4.522	4.493	2.828	2.747
			5	9.642	9.266	9.090	3.698	2.903
	0.36	0	0.2	3.570	2.589	2.586	2.593	2.591
			1	3.570	2.589	2.586	2.593	2.591
			5	3.570	2.589	2.586	2.593	2.591
		0.3	0.2	3.650	2.703	2.697	2.597	2.589
			1	3.650	2.704	2.698	2.597	2.590
			5	3.653	2.724	2.716	2.608	2.603
		0.6	0.2	4.025	3.176	3.163	2.609	2.589
			1	4.025	3.187	3.173	2.612	2.591
			5	4.038	3.441	3.389	2.665	2.638
		0.9	0.2	6.230	4.634	4.627	2.678	2.657
			1	6.281	4.912	4.882	2.704	2.646
			5	7.417	9.508	9.317	3.154	2.713
20	0.96	0	0.2	3.559	2.579	2.577	2.582	2.580
			1	3.559	2.579	2.577	2.582	2.580
			5	3.559	2.579	2.577	2.582	2.580
		0.3	0.2	3.642	2.691	2.688	2.583	2.578
			1	3.642	2.692	2.688	2.583	2.578
			5	3.642	2.703	2.698	2.589	2.585
		0.6	0.2	4.027	3.184	3.174	2.589	2.577
			1	4.027	3.190	3.180	2.590	2.579
			5	4.031	3.336	3.307	2.619	2.608

Continued on Next Page...

Table 4 – Continued

$\bar{T}$	ω	ρ	τ	OLS	FE	RE	FEAR	REAR
		0.9	0.2	6.268	4.976	4.964	2.635	2.612
			1	6.282	5.206	5.179	2.648	2.603
			5	6.641	9.297	9.109	2.900	2.634
	0.36	0	0.2	3.560	2.557	2.556	2.547	2.546
			1	3.560	2.557	2.556	2.547	2.546
			5	3.560	2.557	2.556	2.547	2.546
		0.3	0.2	3.651	2.703	2.701	2.548	2.547
			1	3.651	2.703	2.701	2.548	2.547
			5	3.651	2.710	2.708	2.550	2.550
		0.6	0.2	4.042	3.283	3.275	2.552	2.550
			1	4.042	3.288	3.279	2.553	2.551
			5	4.045	3.389	3.371	2.564	2.565
		0.9	0.2	6.304	5.818	5.793	2.594	2.568
			1	6.308	6.021	5.987	2.598	2.563
			5	6.387	9.755	9.605	2.698	2.568
40	0.96	0	0.2	3.571	2.552	2.551	2.549	2.549
			1	3.571	2.552	2.551	2.549	2.549
			5	3.571	2.552	2.551	2.549	2.549
		0.3	0.2	3.661	2.699	2.696	2.550	2.549
			1	3.661	2.699	2.697	2.550	2.549
			5	3.661	2.706	2.703	2.551	2.551
		0.6	0.2	4.053	3.279	3.272	2.553	2.553
			1	4.053	3.282	3.275	2.553	2.553
			5	4.055	3.368	3.354	2.559	2.562
		0.9	0.2	6.384	6.144	6.117	2.581	2.562
			1	6.386	6.324	6.292	2.581	2.558
			5	6.437	9.697	9.580	2.626	2.560
	0.36	0	0.2	3.594	2.561	2.561	2.555	2.555
			1	3.594	2.561	2.561	2.555	2.555
			5	3.594	2.561	2.561	2.555	2.555
		0.3	0.2	3.688	2.706	2.705	2.556	2.557
			1	3.687	2.706	2.705	2.556	2.557
			5	3.687	2.712	2.710	2.556	2.557
		0.6	0.2	4.078	3.293	3.289	2.558	2.565
			1	4.078	3.297	3.292	2.558	2.565
			5	4.079	3.374	3.366	2.560	2.568
		0.9	0.2	6.390	6.324	6.300	2.575	2.566
			1	6.391	6.482	6.454	2.574	2.564
			5	6.432	9.457	9.380	2.589	2.561

Note: $N = 50$ for all experiments. $\tau/(1-\rho^2)$ is the variance of the initial condition.

Table 5: MAPE of the Predictors

$\bar{T}$	ω	ρ	τ	OLS	FE	RE	FEAR	REAR
5	0.96	0	0.2	388.781	364.873	347.241	374.599	361.344
			1	388.781	364.873	347.241	374.599	361.344
			5	388.781	364.873	347.241	374.599	361.344
		0.3	0.2	472.399	394.322	370.196	395.632	367.692
			1	543.080	424.699	395.002	427.752	393.979
			5	408.194	400.761	366.890	390.466	351.306
		0.6	0.2	371.881	410.321	386.695	385.988	351.848
			1	352.733	395.485	367.156	369.172	328.525
			5	675.286	1471.764	1307.010	1020.926	795.121
		0.9	0.2	271.515	287.700	276.200	256.298	232.382
			1	241.323	311.349	303.178	265.039	242.539
			5	221.586	297.925	295.588	199.071	175.240
	0.36	0	0.2	568.493	511.893	500.355	526.548	519.162
			1	568.493	511.893	500.355	526.548	519.162
			5	568.493	511.893	500.355	526.548	519.162
		0.3	0.2	663.620	504.944	472.328	520.457	477.871
			1	635.870	491.649	461.389	500.913	461.624
			5	578.635	473.317	441.880	473.201	433.509
		0.6	0.2	436.830	340.231	331.524	321.295	304.499
			1	323.176	302.449	290.554	282.579	262.214
			5	343.554	452.999	420.084	338.357	302.176
		0.9	0.2	536.172	573.998	557.402	373.336	307.967
			1	339.431	362.950	353.319	267.107	245.592
			5	296.403	638.546	631.597	326.427	266.351
10	0.96	0	0.2	507.105	331.547	328.826	334.001	331.750
			1	507.105	331.547	328.826	334.001	331.750
			5	507.105	331.547	328.826	334.001	331.750
		0.3	0.2	419.093	303.235	296.054	293.198	282.471
			1	419.633	303.216	296.009	293.091	282.448
			5	422.457	308.346	299.043	295.611	283.320
		0.6	0.2	477.066	391.849	385.872	330.971	317.429
			1	403.383	382.440	374.101	316.133	297.884
			5	470.734	564.990	534.902	406.885	361.299
		0.9	0.2	394.535	332.292	327.953	244.422	238.606
			1	657.131	608.006	600.628	460.184	432.760
			5	363.493	757.189	744.736	280.140	217.670
	0.36	0	0.2	515.523	308.034	308.532	303.731	306.309
			1	515.523	308.034	308.532	303.731	306.309
			5	515.523	308.034	308.532	303.731	306.309
		0.3	0.2	820.423	579.910	584.989	486.742	488.252
			1	817.758	576.675	581.895	481.931	483.171
			5	804.463	561.687	568.271	456.013	456.192
		0.6	0.2	541.557	407.318	408.491	319.719	315.776
			1	543.008	404.685	406.128	316.472	313.995
			5	664.811	561.389	561.359	429.551	430.648
		0.9	0.2	478.897	352.948	344.789	204.472	208.310
			1	498.468	332.856	320.943	190.843	205.824
			5	671.631	1023.864	1006.655	323.556	288.608
20	0.96	0	0.2	735.155	606.707	604.971	605.472	603.203
			1	735.155	606.707	604.971	605.472	603.203
			5	735.155	606.707	604.971	605.472	603.203
		0.3	0.2	747.760	503.655	510.359	495.780	509.192
			1	747.064	502.110	509.011	494.440	507.842
			5	743.592	495.168	503.430	487.975	502.421
		0.6	0.2	620.470	448.076	453.185	354.475	360.875
			1	623.924	455.387	460.378	358.678	365.064
			5	644.380	502.154	506.337	381.293	386.356

Continued on Next Page. . .

Table 5 – Continued

$\bar{T}$	ω	ρ	τ	OLS	FE	RE	FEAR	REAR
		0.9	0.2	572.536	494.404	491.723	258.131	253.562
			1	441.088	360.780	355.768	203.415	197.831
			5	534.917	998.425	978.935	278.228	242.891
	0.36	0	0.2	448.128	275.879	280.058	175.659	177.954
			1	448.128	275.879	280.058	175.659	177.954
			5	448.128	275.879	280.058	175.659	177.954
		0.3	0.2	276.183	157.497	160.107	150.674	156.668
			1	276.175	157.306	159.923	150.743	156.723
			5	276.135	156.461	159.268	150.758	156.956
		0.6	0.2	585.028	277.583	275.941	249.051	232.881
			1	584.386	280.559	278.667	251.681	235.231
			5	581.580	297.120	293.132	265.511	245.963
		0.9	0.2	519.122	461.106	459.089	171.033	164.562
			1	469.154	388.893	388.020	127.027	126.163
			5	690.008	612.667	600.099	279.444	300.555
40	0.96	0	0.2	132.745	73.487	74.413	72.354	73.275
			1	132.745	73.487	74.413	72.354	73.275
			5	132.745	73.487	74.413	72.354	73.275
		0.3	0.2	195.023	106.445	107.825	90.899	94.947
			1	195.025	106.565	107.949	90.968	95.043
			5	195.035	107.217	108.691	91.289	95.617
		0.6	0.2	206.088	134.901	136.284	98.269	103.515
			1	206.086	135.425	136.810	98.373	103.656
			5	206.090	138.809	140.421	98.532	104.194
		0.9	0.2	434.826	338.220	339.467	125.598	127.396
			1	660.541	571.856	570.239	237.506	231.242
			5	488.213	618.274	612.791	111.589	111.776
	0.36	0	0.2	35.574	22.870	22.934	22.393	22.445
			1	35.574	22.870	22.934	22.393	22.445
			5	35.574	22.870	22.934	22.393	22.445
		0.3	0.2	35.348	24.015	24.068	22.729	22.919
			1	35.348	24.015	24.068	22.726	22.916
			5	35.350	24.047	24.101	22.714	22.908
		0.6	0.2	49.704	36.403	36.511	25.224	25.842
			1	49.697	36.458	36.565	25.215	25.838
			5	49.672	37.131	37.237	25.147	25.813
		0.9	0.2	502.623	527.477	526.176	69.947	73.859
			1	435.106	415.868	415.743	104.319	114.070
			5	840.215	902.339	900.909	112.293	118.572

Note: $N = 50$ for all experiments. $\tau/(1-\rho^2)$ is the variance of the initial condition.

Table 6: Estimation and Forecasting Results using the National Longitudinal Study

	OLS	FE	RE	FEAR	REAR
age	0.0405	0.0417	0.0414	0.0420	0.0415
	(0.0037)	(0.0033)	(0.0031)	(0.0031)	(0.0032)
age2	-0.0007	-0.0009	-0.0008	-0.0009	-0.0008
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
exp	0.0271	0.0398	0.0348	0.0399	0.0347
	(0.0011)	(0.0017)	(0.0013)	(0.0016)	(0.0013)
tenure	0.0450	0.0334	0.0363	0.0332	0.0363
	(0.0020)	(0.0018)	(0.0017)	(0.0017)	(0.0018)
tenure2	-0.0018	-0.0020	-0.0019	-0.0020	-0.0019
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
nsmsa	-0.1642	-0.0815	-0.1246	-0.0791	-0.1249
	(0.0054)	(0.0100)	(0.0075)	(0.0092)	(0.0074)
south	-0.1007	-0.0501	-0.0833	-0.0475	-0.0830
	(0.0052)	(0.0116)	(0.0077)	(0.0107)	(0.0076)
grade	0.0622		0.0643		0.0643
	(0.0011)		(0.0019)		(0.0019)
black	-0.0697		-0.0545		-0.0548
	(0.0056)		(0.0103)		(0.0102)
Intercept	0.2248		0.1822		0.1782
	(0.0520)		(0.0498)		(0.0504)
σ_μ		0.3245	0.2373	0.2684	0.2308
σ_v	0.3594	0.2732	0.2732	0.2747	0.2721
ρ				0.1012	0.1012
LBI				1.8404	1.8404
F-statistics				107.4471	107.4471
p-value				0.0000	0.0000
MSE	0.2136	0.1647	0.1610	0.1603	0.1559
MAE	0.3328	0.2688	0.2674	0.2623	0.2609
MAPE	41.1100	31.0870	32.3895	30.6727	32.2694

Note: The sample is an unbalanced panel data of 3640 women over the years 1968-1988 with gaps. We compute the forecasts of logarithm wage for the last available year. In-sample model coefficient estimates are based on 22887 observations from all previous years. For the in-sample, the average available years $\bar{T} = 6.288$ and the Ahrens and Pincus index $\omega = 0.724$. On average, we are forecasting $\bar{S} = 2.131$ years ahead. MSE, MAE and MAPE are out-of-sample forecast comparison for the last available year. σ_μ and σ_v are the standard deviations of the individual effects and remainder disturbances, respectively. ρ is the autocorrelation parameter of the remainder disturbances. LBI is the locally best invariant test statistic in Baltagi and Wu (1999). F-statistics and p-value are for the panel serial correlation test in Wooldridge (2002). Standard errors in parentheses.

Table 7: Panel Data Test Results of Equal Predictive Accuracy using the National Longitudinal Study

	OLS	FE	RE	FEAR	REAR
OLS					
FE	-10.9947				
RE	-14.4038	-3.8038			
FEAR	-11.8924	-11.6062	-0.6650		
REAR	-16.2446	-6.9276	-10.5975	-3.5953	

Note: The test statistic asymptotically follows a standard normal distribution. A negative entry means the row estimator is better than the column.

Appendix

Proof of Theorem 1

Proof. Denote $T(1)$ as the set of observations when both $t_{i,j}$ and $t_{i,j-1}$ are observed. Equation (21) could be rewritten as

$$\hat{\rho} = \frac{\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)} \hat{\nu}_{i,t_{i,j}}\hat{\nu}_{i,t_{i,j-1}}}{\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2}.$$

where

$$\hat{\nu}_{i,t_{i,j}} = \tilde{y}_{i,t_{i,j}} - \hat{\beta}_{FE}\tilde{x}_{i,t_{i,j}} = \tilde{v}_{i,t_{i,j}} - \left(\hat{\beta}_{FE} - \beta\right)\tilde{x}_{i,t_{i,j}},$$

with $\tilde{y}_{i,t_{i,j}} = y_{i,t_{i,j}} - \bar{y}_{i.}$ and $\bar{y}_{i.} = n_i^{-1}\sum_{j=1}^{n_i} y_{i,t_{i,j}}$. Other terms such as $\tilde{x}_{i,t_{i,j}}$, $\bar{x}_{i.}$, $\tilde{v}_{i,t_{i,j}}$ and $\bar{v}_{i.}$ are similarly defined. Hence,

$$\begin{aligned}
\hat{\rho} - \rho &= \frac{\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)} \hat{\nu}_{i,t_{i,j}}\hat{\nu}_{i,t_{i,j-1}}}{\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2} - \rho \\
&= \frac{\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)} \left(\hat{\nu}_{i,t_{i,j}} - \rho\hat{\nu}_{i,t_{i,j-1}}\right)\hat{\nu}_{i,t_{i,j-1}}}{\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2} \\
&\quad +\rho\left(\frac{\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)} \hat{\nu}_{i,t_{i,j-1}}^2 - \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2}{\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2}\right),
\end{aligned}$$

First of all, we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2 &= \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left[\tilde{v}_{i,t_{i,j}} - \left(\hat{\beta}_{FE} - \beta\right)\tilde{x}_{i,t_{i,j}}\right] \\
&= \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\tilde{v}_{i,t_{i,j}}^2 + \frac{1}{n}\left[\sqrt{n}\left(\hat{\beta}_{FE} - \beta\right)\right]^2\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\tilde{x}_{i,t_{i,j}}^2 \\
&\quad -\frac{2}{n}\left[\sqrt{n}\left(\hat{\beta}_{FE} - \beta\right)\right]\frac{1}{\sqrt{n}}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\tilde{v}_{i,t_{i,j}}\tilde{x}_{i,t_{i,j}}
\end{aligned}$$

Following Lemma 7 in Hahn and Kuersteiner (2002), we can show $\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\tilde{v}_{i,t_{i,j}}^2 = \frac{\sigma_\epsilon^2}{(1-\rho)^2} + o_p(1)$. Similarly, we can show that $\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\tilde{x}_{i,t_{i,j}}^2 = O_p(1)$, $\frac{1}{\sqrt{n}}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\tilde{v}_{i,t_{i,j}}\tilde{x}_{i,t_{i,j}} = O_p(1)$ and $\sqrt{n}\left(\hat{\beta}_{FE} - \beta\right) = O_p(1)$ under the assumptions stated in the Theorem. Hence

$$\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^2 = \frac{\sigma_\epsilon^2}{(1-\rho)^2} + O_p\left(\frac{1}{n}\right).$$

Similarly, we can show that

$$\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\hat{\nu}_{i,t_{i,j-1}}^{2}=\frac{\sigma_{\epsilon}^{2}}{(1-\rho)^{2}}+O_{p}\left(\frac{1}{m}\right).$$

so that

$$\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\hat{\nu}_{i,t_{i,j-1}}^{2}-\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\hat{\nu}_{i,t_{i,j}}^{2}=O_{p}\left(\frac{1}{m}\right)-O_{p}\left(\frac{1}{n}\right)=O_{p}\left(\frac{1}{m}\right).$$

Also, we have

$$\begin{aligned}
&\hat{\nu}_{i,t_{i,j}}-\rho\hat{\nu}_{i,t_{i,j-1}}\\
=\;&\left[\tilde{v}_{i,t_{i,j}}-\left(\hat{\beta}_{FE}-\beta\right)\tilde{x}_{i,t_{i,j}}\right]-\rho\left[\tilde{v}_{i,t_{i,j-1}}-\left(\hat{\beta}_{FE}-\beta\right)\tilde{x}_{i,t_{i,j-1}}\right]\\
=\;&\left(\tilde{v}_{i,t_{i,j}}-\rho\tilde{v}_{i,t_{i,j-1}}\right)-\left(\hat{\beta}_{FE}-\beta\right)\left(\tilde{x}_{i,t_{i,j}}-\rho\tilde{x}_{i,t_{i,j-1}}\right)\\
=\;&\tilde{\epsilon}_{i,t_{i,j}}-\left(\hat{\beta}_{FE}-\beta\right)\left(\tilde{x}_{i,t_{i,j}}-\rho\tilde{x}_{i,t_{i,j-1}}\right),
\end{aligned}$$

where $\tilde{\epsilon}_{i,t_{i,j}}=\epsilon_{i,t_{i,j}}-\bar{\epsilon}_{i\cdot}$ with $\bar{\epsilon}_{i\cdot}=n_i^{-1}\sum_{j=1}^{n_i}\epsilon_{i,t_{i,j}}$. Hence

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\left(\hat{\nu}_{i,t_{i,j}}-\rho\hat{\nu}_{i,t_{i,j-1}}\right)\hat{\nu}_{i,t_{i,j-1}}\\
=\;&\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\left[\tilde{\epsilon}_{i,t_{i,j}}-\left(\hat{\beta}_{FE}-\beta\right)\left(\tilde{x}_{i,t_{i,j}}-\rho\tilde{x}_{i,t_{i,j-1}}\right)\right]\left[\tilde{v}_{i,t_{i,j-1}}-\left(\hat{\beta}_{FE}-\beta\right)\tilde{x}_{i,t_{i,j-1}}\right]\\
=\;&\frac{N}{m}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\tilde{\epsilon}_{i,t_{i,j}}\tilde{v}_{i,t_{i,j-1}}\right]\\
&-\frac{1}{\sqrt{nm}}\left[\sqrt{n}\left(\hat{\beta}_{FE}-\beta\right)\right]\left[\frac{1}{\sqrt{m}}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\tilde{v}_{i,t_{i,j-1}}\left(\tilde{x}_{i,t_{i,j}}-\rho\tilde{x}_{i,t_{i,j-1}}\right)\right]\\
&-\frac{1}{\sqrt{nm}}\left[\sqrt{n}\left(\hat{\beta}_{FE}-\beta\right)\right]\left[\frac{1}{\sqrt{m}}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\tilde{\epsilon}_{i,t_{i,j}}\tilde{x}_{i,t_{i,j-1}}\right]\\
&+\frac{1}{n}\left[\sqrt{n}\left(\hat{\beta}_{FE}-\beta\right)\right]^{2}\left[\frac{1}{m}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\left(\tilde{x}_{i,t_{i,j}}-\rho\tilde{x}_{i,t_{i,j-1}}\right)\tilde{x}_{i,t_{i,j-1}}\right]
\end{aligned}$$

Following Lemma 6 in Hahn and Kuersteiner (2002), we can show $\frac{1}{N}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\tilde{\epsilon}_{i,t_{i,j}}\tilde{v}_{i,t_{i,j-1}}=\frac{\sigma_{\epsilon}^{2}}{1-\rho}+o_p(1)$. Similarly, we can show that $\frac{1}{\sqrt{m}}\sum_{i=1}^{N}\sum_{t_{i,j}\in T(1)}\tilde{v}_{i,t_{i,j-1}}\left(\tilde{x}_{i,t_{i,j}}-\rho\tilde{x}_{i,t_{i,j-1}}\right)=$

$O_p(1)$, $\frac{1}{\sqrt{m}} \sum_{i=1}^{N} \sum_{t_{i,j} \in T(1)} \tilde{\epsilon}_{i,t_{i,j}} \tilde{x}_{i,t_{i,j-1}} = O_p(1)$, $\frac{1}{m} \sum_{i=1}^{N} \sum_{t_{i,j} \in T(1)} \left(\tilde{x}_{i,t_{i,j}} - \rho \tilde{x}_{i,t_{i,j-1}} \right) \tilde{x}_{i,t_{i,j-1}} = O_p(1)$ and $\sqrt{n} \left(\hat{\beta}_{FE} - \beta \right) = O_p(1)$ under the assumptions stated in the Theorem. Hence

$$\frac{1}{m} \sum_{i=1}^{N} \sum_{t_{i,j} \in T(1)} \left(\hat{\nu}_{i,t_{i,j}} - \rho \hat{\nu}_{i,t_{i,j-1}} \right) \hat{\nu}_{i,t_{i,j-1}} = O_p\left(\frac{N}{m} \right)$$

Therefore, we have

$$\begin{aligned} \hat{\rho} - \rho &= \frac{\frac{1}{m} \sum_{i=1}^{N} \sum_{t_{i,j} \in T(1)} \left(\hat{\nu}_{i,t_{i,j}} - \rho \hat{\nu}_{i,t_{i,j-1}} \right) \hat{\nu}_{i,t_{i,j-1}}}{\frac{1}{n} \sum_{i=1}^{N} \sum_{j=1}^{n_i} \hat{\nu}_{i,t_{i,j}}^2} \\ &\quad + \rho \left(\frac{\frac{1}{m} \sum_{i=1}^{N} \sum_{t_{i,j} \in T(1)} \hat{\nu}_{i,t_{i,j-1}}^2 - \frac{1}{n} \sum_{i=1}^{N} \sum_{j=1}^{n_i} \hat{\nu}_{i,t_{i,j}}^2}{\frac{1}{n} \sum_{i=1}^{N} \sum_{j=1}^{n_i} \hat{\nu}_{i,t_{i,j}}^2} \right) \\ &= O_p\left(\frac{N}{m} \right) + O_p\left(\frac{1}{m} \right) = O_p\left(\frac{N}{m} \right). \end{aligned}$$

■