The Stata Journal (2007)
7, Number 2, pp. 197–208

Estimation of nonstationary heterogeneous panels

Edward F. Blackburne III
Sam Houston State University
Huntsville, TX
blackburne@shsu.edu

Mark W. Frank
Sam Houston State University
Huntsville, TX
markfrank@shsu.edu

Abstract. We introduce a new Stata command, `xtpmg`, for estimating nonstationary heterogeneous panels in which the number of groups and number of time-series observations are both large. Based on recent advances in the nonstationary panel literature, `xtpmg` provides three alternative estimators: a traditional fixed-effects estimator, the mean-group estimator of Pesaran and Smith (Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics* 68: 79–113), and the pooled mean-group estimator of Pesaran, Shin, and Smith (Estimating long-run relationships in dynamic heterogeneous panels, DAE Working Papers Amalgamated Series 9721; Pooled mean group estimation of dynamic heterogeneous panels, *Journal of the American Statistical Association* 94: 621–634).

Keywords: st0125, xtpmg, nonstationary panels, heterogeneous dynamic panels, pooled mean-group estimator, mean-group estimator, panel cointegration

1 Introduction

In recent years, the dynamic panel-data literature has begun to focus on panels in which the number of cross-sectional observations (N) and the number of time-series observations (T) are both large. The availability of data with greater frequency is certainly a key contributor to this shift. Some cross-national and cross-state datasets, for example, are now large enough in T such that each nation (or state) can be estimated separately.

The asymptotics of large N, large T dynamic panels are different from the asymptotics of traditional large N, small T dynamic panels. Small T panel estimation usually relies on fixed- or random-effects estimators, or a combination of fixed-effects estimators and instrumental-variable estimators, such as the Arellano and Bond (1991) generalized method-of-moments estimator. These methods require pooling individual groups and allowing only the intercepts to differ across the groups. One of the central findings from the large N, large T literature, however, is that the assumption of homogeneity of slope parameters is often inappropriate. This point has been made by Pesaran and Smith (1995); Im, Pesaran, and Shin (2003); Pesaran, Shin, and Smith (1997, 1999); and Phillips and Moon (2000).[1]

1. For more discussion of this literature, see chapter 12 in Baltagi (2001).

With the increase in time observations inherent in large N, large T dynamic panels, nonstationarity is also a concern. Recent papers by Pesaran, Shin, and Smith (1997, 1999) offer two important new techniques to estimate nonstationary dynamic panels in which the parameters are heterogeneous across groups: the mean-group (MG) and pooled mean-group (PMG) estimators. The MG estimator (see Pesaran and Smith 1995) relies on estimating N time-series regressions and averaging the coefficients, whereas the PMG estimator (see Pesaran, Shin, and Smith 1997, 1999) relies on a combination of pooling and averaging of coefficients.

In recent empirical research, the MG and PMG estimators have been applied in a variety of settings. Freeman (2000), for example, uses the estimators to evaluate state-level alcohol consumption over 1961–1995. Martinez-Zarzoso and Bengochea-Morancho (2004) use them in an estimation of an environmental Kuznets curve in a panel of 22 OECD nations over 1975–1998. Frank (2005) uses the MG and PMG estimators to evaluate the long-term effect of income inequality on economic growth in a panel of U.S. states over 1945–2001.

2 The MG and PMG estimators

Assume an autoregressive distributive lag (ARDL) $(p, q_1, \ldots, q_k)$ dynamic panel specification of the form

$$y_{it} = \sum_{j=1}^{p} \lambda_{ij} y_{i,t-j} + \sum_{j=0}^{q} \delta_{ij}^{'} X_{i,t-j} + \mu_i + \epsilon_{it} \tag{1}$$

where the number of groups $i = 1, 2, \ldots, N$; the number of periods $t = 1, 2, \ldots, T$; X_{it} is a $k \times 1$ vector of explanatory variables; δ_{it} are the $k \times 1$ coefficient vectors; λ_{ij} are scalars; and μ_i is the group-specific effect. T must be large enough such that the model can be fitted for each group separately. Time trends and other fixed regressors may be included.

If the variables in (1) are, for example, $I(1)$ and cointegrated, then the error term is an $I(0)$ process for all i. A principal feature of cointegrated variables is their responsiveness to any deviation from long-run equilibrium. This feature implies an error correction model in which the short-run dynamics of the variables in the system are influenced by the deviation from equilibrium. Thus it is common to reparameterize (1) into the error correction equation

$$\Delta y_{it} = \phi_i (y_{i,t-1} - \theta_i^{'} X_{it}) + \sum_{j=1}^{p-1} \lambda_{ij}^{*} \Delta y_{i,t-1} + \sum_{j=0}^{q-1} \delta_{ij}^{'*} \Delta X_{i,t-j} + \mu_i + \epsilon_{it} \tag{2}$$

where $\phi_i = -(1 - \sum_{j=1}^{p} \lambda_{ij})$, $\theta_i = \sum_{j=0}^{q} \delta_{ij} / (1 - \sum_k \lambda_{ik})$, $\lambda_{ij}^{*} = -\sum_{m=j+1}^{p} \lambda_{im} \quad j = 1, 2, \ldots, p-1$, and $\delta_{ij}^{*} = -\sum_{m=j+1}^{q} \delta_{im} \quad j = 1, 2, \ldots, q-1$.

The parameter ϕ_i is the error-correcting speed of adjustment term. If $\phi_i = 0$, then there would be no evidence for a long-run relationship. This parameter is expected to be significantly negative under the prior assumption that the variables show a return to a long-run equilibrium. Of particular importance is the vector $\theta_i^{'}$, which contains the long-run relationships between the variables.

The recent literature on dynamic heterogeneous panel estimation in which both N and T are large suggests several approaches to the estimation of (2). On one extreme, a fixed-effects (FE) estimation approach could be used in which the time-series data for each group are pooled and only the intercepts are allowed to differ across groups. If the slope coefficients are in fact not identical, however, then the FE approach produces inconsistent and potentially misleading results. On the other extreme, the model could be fitted separately for each group, and a simple arithmetic average of the coefficients could be calculated. This is the MG estimator proposed by Pesaran and Smith (1995). With this estimator, the intercepts, slope coefficients, and error variances are all allowed to differ across groups.

More recently, Pesaran, Shin, and Smith (1997, 1999) have proposed a PMG estimator that combines both pooling and averaging. This intermediate estimator allows the intercept, short-run coefficients, and error variances to differ across the groups (as would the MG estimator) but constrains the long-run coefficients to be equal across groups (as would the FE estimator). Since (2) is nonlinear in the parameters, Pesaran, Shin, and Smith (1999) develop a maximum likelihood method to estimate the parameters.

Expressing the likelihood as the product of each cross-section's likelihood and taking the log yields

$$l_T(\theta^{'}, \varphi^{'}, \sigma^{'}) = -\frac{T}{2}\sum_{i=1}^{N} ln(2\pi\sigma_i^2) - \frac{1}{2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\{\Delta y_i - \phi_i\xi_i(\theta)\}' H_i \{\Delta y_i - \phi_i \xi_i(\theta)\} \quad (3)$$

for $i = 1, \ldots, N$, where $\xi_i(\theta) = y_{i,t-1} - X_i\theta_i$, $H_i = I_T - W_i(W_i'W_i)W_i$, I_T is an identity matrix of order T, and $W_i = (\Delta y_{i,t-1}, \ldots, \Delta y_{i,t-p+1}, \Delta X_i, \Delta X_{i,t-1}, \ldots, \Delta X_{i,t-q+1})$.

`xtpmg` uses Stata's powerful `ml` framework to implement the PMG estimator. Specifically, we take advantage of the undocumented **`hold`** option of `ml` to maximize the likelihood via "back-substitution".[2] Beginning with an initial estimate of the long-run coefficient vector, $\widehat{\theta}$, the short-run coefficients and the group-specific speed of adjustment terms can be estimated by regressions of Δy_i on $(\widehat{\xi}_i, W_i)$. These conditional estimates are in turn used to update the estimate of θ. The process is iterated until convergence is achieved.

The parameter estimates from iterated conditional likelihood maximization are asymptotically identical to those from full-information maximum likelihood. But the estimated covariance matrix is not. However, since the distribution of the PMG parameters is known, we can recover the full covariance matrix for all estimated parameters. As shown in Pesaran, Shin, and Smith (1999), the covariance matrix can be estimated by the inverse of

2. Although (3) looks benign, the model is difficult to program directly as a Stata `ml` program. Looking forward to (8), readers will note that the model cannot be readily adapted to `ml`'s familiar "*theta*" notation.

$$\begin{bmatrix} \sum_{i=1}^N \frac{\widehat{\phi}_i^2 X_i' X_i}{\widehat{\sigma}_i^2} & \frac{-\widehat{\phi}_1 X_1' \widehat{\xi}_1}{\widehat{\sigma}_1^2} & \cdots & \frac{-\widehat{\phi}_N X_N' \widehat{\xi}_N}{\widehat{\sigma}_N^2} & \frac{-\widehat{\phi}_1 X_1' W_1}{\widehat{\sigma}_1^2} & \cdots & \frac{-\widehat{\phi}_N X_N' W_N}{\widehat{\sigma}_N^2} \\ & \frac{\widehat{\xi}_1' \widehat{\xi}_1}{\widehat{\sigma}_1^2} & \cdots & 0 & \frac{\widehat{\xi}_1' W_1}{\widehat{\sigma}_1^2} & \cdots & 0 \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & \frac{\widehat{\xi}_N' \widehat{\xi}_N}{\widehat{\sigma}_N^2} & 0 & \cdots & \frac{\widehat{\xi}_N' W_N}{\widehat{\sigma}_N^2} \\ & & & & \frac{W_1' W_1}{\widehat{\sigma}_1^2} & \cdots & 0 \\ & & & & & \ddots & \vdots \\ & & & & & & \frac{W_N' W_N}{\widehat{\sigma}_N^2} \end{bmatrix}$$

The MG parameters are simply the unweighted means of the individual coefficients. For example, the MG estimate of the error correction coefficient, ϕ, is

$$\widehat{\phi} = N^{-1} \sum_{i=1}^{N} \widehat{\phi}_i \tag{4}$$

with the variance

$$\widehat{\Delta}_{\widehat{\phi}} = \frac{1}{N(N-1)} \sum_{i=1}^{N} (\widehat{\phi}_i - \widehat{\phi})^2 \tag{5}$$

The mean and variance of other short-run coefficients are similarly estimated.

3 The xtpmg command

3.1 Syntax

`xtpmg` *varlist* [*if*] [*in*] [, `lr(`*varlist*`)` `ec(`*string*`)` `replace` `constraints(`*string*`)` `noconstant` `cluster(`*varname*`)` `level(`#`)` `technique(`*algorithm_spec*`)` `difficult` `full` *model*]

3.2 Options

`lr(`*varlist*`)` specifies the variables to be included when calculating the long-run cointegrating vector.

`ec(`*string*`)` is used to specify the name of the newly created error-correction term; default is `__ec`.

`replace` overwrites the error-correction variable, if it exists.

`constraints(`*string*`)` specifies the constraints to be applied to the model. This option is currently used only with option `pmg`.

`noconstant` suppresses the constant term. This option cannot be used with option `dfe`.

`cluster(`*varname*`)` specifies that the observations are independent across groups (clusters), but not necessarily within groups. *varname* specifies to which group each observation belongs, e.g., `cluster(personid)` in data with repeated observations on individuals. `cluster()` affects the estimated standard errors and variance–covariance matrix of the estimators (VCE), but not the estimated coefficients; see [U] **20.14 Obtaining robust variance estimates**.

`level(`*#*`)` sets the confidence level; default is `level(95)`.

`technique(`*algorithm_spec*`)` specifies the `ml` maximization technique. *algorithm_spec* is *algorithm* [*#* [*algorithm* [*#*]] ...], where *algorithm* is { `nr` | `bfgs` | `dfp` }. The `bhhh` algorithm is not compatible with `xtpmg`. `technique()` can be used only with option `pmg`.

`difficult` will use a different stepping algorithm in nonconcave regions of the likelihood.

`full` specifies that all N cross-section regression results be listed. Only the averaged coefficients are listed by default.

model is the type of estimator to be fitted and is one of the following:

`pmg` is the default and specifies the PMG estimator. This model constrains the long-run coefficient vector to be equal across panels while allowing for group-specific short-run and adjustment coefficients.

`mg` specifies the MG estimator. This model fits parameters as averages of the N individual group regressions.

`dfe` specifies the dynamic fixed-effects estimator.

4 Empirical example: OECD consumption

4.1 Data

We illustrate the use of `xtpmg` with annual aggregate consumption data for 24 Organisation for Economic Co-operation and Development (OECD) nations. These data are taken from Pesaran, Shin, and Smith (1997, 1999) and encompass the years 1960–1993.[3] The 1993 annual observation for Belgium is not included in the estimation sample, leaving an estimation period of 1962–1992 for Belgium and 1962–1993 for the other 23 OECD countries. `xtpmg` requires that the data be `tsset` before estimation.

```
. use jasa2

. tsset id year
       panel variable:  id (unbalanced)
        time variable:  year, 1960 to 1993
```

3. The original data and GAUSS code are available on Pesaran's web site: http://www.econ.cam.ac.uk/faculty/pesaran.

Assume the long-run consumption function

$$c_{it} = \theta_{0t} + \theta_{1t}y_{it} + \theta_{2t}\pi_{it} + \mu_i + \epsilon_{it} \tag{6}$$

where the number of nations $i = 1, 2, \ldots, N$; the number of periods $t = 1, 2, \ldots, T$; c_{it} is the log of real per capita consumption; y_{it} is the log of real per capita income; and π_{it} is the inflation rate. If the variables are $I(1)$ and cointegrated, then the error term is $I(0)$ for all i. The ARDL(1,1,1) dynamic panel specification of (6) is

$$c_{it} = \delta_{10i}y_{it} + \delta_{11i}y_{i,t-1} + \delta_{20i}\pi_{it} + \delta_{21i}\pi_{i,t-1} + \lambda_i c_{i,t-1}\mu_i + \epsilon_{it} \tag{7}$$

The error correction reparameterization of (7) is

$$\Delta c_{it} = \phi_i \left(c_{i,t-1} - \theta_{0i} - \theta_{1i}y_{it} - \theta_{2i}\pi_{it}\right) + \delta_{11i}\Delta y_{it} + \delta_{21i}\Delta\pi_{it} + \epsilon_{it} \tag{8}$$

where $\phi_i = -(1-\lambda_i)$, $\theta_{0i} = \frac{\mu_i}{1-\lambda_i}$, $\theta_{it} = \frac{\delta_{10i}+\delta_{11i}}{1-\lambda_i}$, and $\theta_{2i} = \frac{\delta_{20i}+\delta_{21i}}{1-\lambda_i}$.

The error-correction speed of adjustment parameter, ϕ_i, and the long-run coefficients, θ_{1i} and θ_{2i}, are of primary interest. With the inclusion of θ_{0i}, a nonzero mean of the cointegrating relationship is allowed. One would expect ϕ_i to be negative if the variables exhibit a return to long-run equilibrium. Most aggregate consumption theories indicate that the long-run income elasticity, θ_{1i}, should be equal to one. The inflation effect, θ_{2i}, is generally thought to be negative.

4.2 PMG estimation

The first example estimates the PMG estimator for model (8). In this context, the PMG model allows for heterogeneous short-run dynamics and common long-run income and inflation elasticities. Often only the long-run parameters are of interest. The default results of the `pmg` option include the long-run parameter estimates and the averaged short-run parameter estimates.[4]

4. The PMG standard errors match the GAUSS output provided by Pesaran, Shin, and Smith (1999). The standard errors in table 1 of Pesaran, Shin, and Smith (1999) are, however, different from those reported here and from the original GAUSS program.

```
. xtpmg d.c d.pi d.y if year>=1962, lr(l.c pi y) ec(ec) replace pmg
Iteration 0:   log likelihood =  2270.3017  (not concave)
Iteration 1:   log likelihood =  2319.1636
Iteration 2:   log likelihood =  2322.9301
Iteration 3:   log likelihood =  2326.7546
Iteration 4:   log likelihood =  2327.0742
Iteration 5:   log likelihood =  2327.0749
Iteration 6:   log likelihood =  2327.0749
Pooled Mean Group Regression
(Estimate results saved as pmg)
Panel Variable (i): id                          Number of obs      =       767
Time Variable (t): year                         Number of groups   =        24
                                                Obs per group: min =        31
                                                               avg =      32.0
                                                               max =        32
                                                Log Likelihood     =  2327.075

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ec           |
          pi |  -.4658438   .0567332    -8.21   0.000    -.5770388   -.3546487
           y |   .9044336   .0086815   104.18   0.000     .8874181    .9214491
-------------+----------------------------------------------------------------
SR           |
          ec |  -.1998761   .0321683    -6.21   0.000    -.2629247   -.1368275
          pi |
         D1. |  -.0182588   .0277523    -0.66   0.511    -.0726522    .0361347
           y |
         D1. |   .3269355   .0574236     5.69   0.000     .2143873    .4394838
       _cons |   .1544507   .0216943     7.12   0.000     .1119308    .1969707
------------------------------------------------------------------------------
```

In the output, the estimated long-run inflation elasticity is significantly negative, as expected. Also, the estimated income elasticity is significantly positive. Theoretically, the income elasticity is equal to one. This hypothesis is easily tested:

```
. test [ec]y=1
 ( 1)  [ec]y = 1
           chi2(  1) =  121.18
         Prob > chi2 =    0.0000
```

The corresponding χ^2 value of 121.2 leads to rejection of the null hypothesis of unit income elasticity.

The `full` option estimates and saves an $N+1$ multiple-equation model. The first equation (labeled per option `ec`) presents the normalized cointegrating vector.[5] The remaining N equations list the group-specific short-run coefficients.

5. The vector has been normalized such that the coefficient on the first term in the cointegrating vector is 1. Accordingly, the normalized term is omitted from the estimation output.

```
. xtpmg d.c d.pi d.y if year>=1962, lr(l.c pi y) ec(ec) full pmg
Iteration 0:   log likelihood =  2270.3017  (not concave)
Iteration 1:   log likelihood =  2319.1636
Iteration 2:   log likelihood =  2322.9301
Iteration 3:   log likelihood =  2326.7546
Iteration 4:   log likelihood =  2327.0742
Iteration 5:   log likelihood =  2327.0749
Iteration 6:   log likelihood =  2327.0749
Pooled Mean Group Regression
(Estimate results saved as PMG)
Panel Variable (i): id                          Number of obs      =       767
Time Variable (t): year                         Number of groups   =        24
                                                Obs per group: min =        31
                                                               avg =      32.0
                                                               max =        32
                                                Log Likelihood     =  2327.075

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ec           |
          pi |  -.4658438   .0567332    -8.21   0.000    -.5770388   -.3546487
           y |   .9044336   .0086815   104.18   0.000     .8874181    .9214491
-------------+----------------------------------------------------------------
id_111       |
          ec |  -.0378815   .0240594    -1.57   0.115    -.0850371    .0092742
          pi |
         D1. |  -.2114431   .0866913    -2.44   0.015    -.3813548   -.0415314
           y |
         D1. |   .5195067    .055876     9.30   0.000     .4099918    .6290217
       _cons |   .0336383   .0147912     2.27   0.023     .0046481    .0626285
  (output omitted)
id_196       |
          ec |  -.4978606   .0887771    -5.61   0.000    -.6718605   -.3238608
          pi |
         D1. |   .0721044   .0721146     1.00   0.317    -.0692375    .2134464
           y |
         D1. |   .0390557    .103316     0.38   0.705      -.16344    .2415515
       _cons |   .2743539   .0630399     4.35   0.000     .1507979    .3979099
------------------------------------------------------------------------------
```

Since each group has its own estimated equation, we can, for example, predict variables intuitively.

```
. predict dc111 if id==111, eq(id_111)
(783 missing values generated)
```

Similarly, cross-equation restrictions are easily applied.

```
. test [id_111]ec=[id_112]ec=0
 ( 1)  [id_111]ec - [id_112]ec = 0
 ( 2)  [id_111]ec = 0
           chi2(  2) =    2.54
         Prob > chi2 =    0.2814
```

4.3 MG estimation

The MG estimates are the unweighted mean of the N individual regression coefficients. `xtpmg` with the `mg` option loops through all panels in the sample to estimate the parameters of (8).[6]

```
. xtpmg d.c d.pi d.y if year>=1962, lr(l.c pi y) ec(ec) replace mg
```

```
Mean Group Estimation: Error Correction Form
(Estimate results saved as mg)
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ec						
pi	-.3529095	.1168025	-3.02	0.003	-.5818381	-.1239809
y	.9181344	.0272673	33.67	0.000	.8646914	.9715774
SR						
ec	-.3063473	.0301599	-10.16	0.000	-.3654597	-.2472349
pi						
D1.	-.0253642	.0294774	-0.86	0.390	-.0831389	.0324104
y						
D1.	.2337588	.0489502	4.78	0.000	.1378182	.3296994
_cons	.2082185	.1089385	1.91	0.056	-.005297	.4217339

The MG estimates are presented as a two-equation model: the normalized cointegrating vector and the short-run dynamic coefficients. In comparing the PMG and MG estimators, we note that the estimated long-run income and inflation elasticities are statistically significant and properly signed in both models. However, the PMG estimate of the inflation elasticity is larger in magnitude than the estimate from the MG model ($-.47$ and $-.35$, respectively). The opposite is true for the estimated long-run income elasticity (.90 and .92, respectively). The speed of adjustment estimates from each model imply significantly different short-run dynamics (compare $\widehat{\phi} = -.20$ from PMG and $\widehat{\phi} = -.31$ from MG).

Recall that the PMG estimator constrains the long-run elasticities to be equal across all panels. This "pooling" across countries yields efficient and consistent estimates when the restrictions are true. Often, however, the hypothesis of slope homogeneity is rejected empirically. If the true model is heterogeneous, the PMG estimates are inconsistent; the MG estimates are consistent in either case. The test of difference in these models is performed with the familiar Hausman test.[7]

6. Actually, since (8) is nonlinear in the parameters, `xtpmg` estimates the reduced-form regressions for each group, $\Delta c_t = \phi c_{t-1} + \beta_1 y_t + \beta_2 \pi_t + \gamma_1 \Delta y_t + \gamma_2 \Delta \pi_t$, and then applies Stata's `nlcom` command to recover the underlying parameter estimates.

7. Stata's `hausman` test offers a `sigmamore` option. This option forces the variance–covariance matrix from the efficient model (PMG here) to be used in calculating the test statistic. This is what is presented here. See Baum, Schaffer, and Stillman (2003) for more details.

```
. hausman mg pmg, sigmamore
                     ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      mg           pmg         Difference          S.E.
-------------+----------------------------------------------------------------
          pi |   -.3529095    -.4658438        .1129342         .126218
           y |    .9181344     .9044336        .0137008        .0311167
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtpmg
            B = inconsistent under Ha, efficient under Ho; obtained from xtpmg
    Test:  Ho:  difference in coefficients not systematic
                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        1.06
                Prob>chi2 =      0.5887
```

The calculated Hausman statistic is 1.06 and is distributed $\chi^2(2)$. Here we conclude that the PMG estimator, the efficient estimator under the null hypothesis, is preferred.

4.4 Dynamic FE

The dynamic FE estimator, like the PMG estimator, restricts the coefficients of the cointegrating vector to be equal across all panels. The FE model further restricts the speed of adjustment coefficient and the short-run coefficients to be equal. `xtpmg` with the `dfe` option fits the model in (8) while allowing panel-specific intercepts.[8] An allowance for intragroup correlation in the calculation of standard errors is made with the `cluster()` option.

```
. xtpmg d.c d.pi d.y if year>=1962, lr(l.c pi y) ec(ec) replace dfe cluster(id)
------------------------------------------------------------------------------
Standard errors adjusted with cluster(id) option.
------------------------------------------------------------------------------
Dynamic Fixed Effects Regression: Estimated Error Correction Form
(Estimate results saved as DFE)
------------------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ec           |
          pi |   -.266343    .102506    -2.60   0.009    -.4672509    -.065435
           y |   .9120574   .0468008    19.49   0.000     .8203295    1.003785
-------------+----------------------------------------------------------------
SR           |
          ec |  -.1794146   .0434584    -4.13   0.000    -.2645915   -.0942378
          pi |
         D1. |  -.0280826   .0325622    -0.86   0.388    -.0919034    .0357382
           y |
         D1. |   .3811944    .070876     5.38   0.000       .24228    .5201089
       _cons |   .1257634   .0805454     1.56   0.118    -.0321025    .2836294
------------------------------------------------------------------------------
```

8. For the FE model, `xtpmg` is simply a wrapper for Stata's `xtreg, fe` command (see [XT] **xtreg**). The underlying parameters of (8) are calculated with `nlcom` and stored as `EC`. The reduced-form model, as estimated by `xtreg, fe`, is stored as `DFE`.

All coefficients from the dynamic FE model are properly signed and, in fact, similar to the PMG and MG estimates. As discussed in Baltagi, Griffin, and Xiong (2000), FE models are subject to a simultaneous equation bias from the endogeneity between the error term and the lagged dependent variable. The Hausman test can be easily performed to measure the extent of this endogeneity.

```
. hausman mg DFE, sigmamore
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      mg           DFE         Difference          S.E.
-------------+----------------------------------------------------------------
          pi |  -.3529095     -.266343       -.0865666        25.80672
           y |   .9181344     .9120574        .0060771        6.024402
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtpmg
            B = inconsistent under Ha, efficient under Ho; obtained from xtpmg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        0.00
                Prob>chi2 =      1.0000
```

Results indicate that the simultaneous equation bias is minimal for these data and, for this example, we conclude that the FE model is preferred over the MG model.

5 Conclusion

This paper follows recent advances offered by Pesaran and Smith (1995) and Pesaran, Shin, and Smith (1997; 1999) in the estimation of nonstationary heterogeneous panels with large N and large T. We offer a new Stata command, `xtpmg`, that estimates three alternative models: a traditional dynamic FE estimator that relies on pooling of cross-sections, an MG estimator that relies on averaging of cross-sections, and a PMG estimator that relies on a combination of pooling and averaging of coefficients.

6 References

Arellano, M., and S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 277–297.

Baltagi, B. H. 2001. *Econometric Analysis of Panel Data*. 2nd ed. New York: Wiley.

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

Baum, C. F., M. E. Schaffer, and S. Stillman. 2003. Instrumental variables and GMM: Estimation and testing. *Stata Journal* 3: 1–31.

Frank, M. W. 2005. Income inequality and economic growth in the U.S.: A panel cointegration approach. Sam Houston State University Working Paper 05-03.

Freeman, D. G. 2000. Alternative panel estimates of alcohol demand, taxation, and the business cycle. *Southern Economic Journal* 67: 325–344.

Im, K. S., M. H. Pesaran, and Y. Shin. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115: 53–74.

Martinez-Zarzoso, I., and A. Bengochea-Morancho. 2004. Pooled mean group estimation of an environmental kuznets curve for CO_2. *Economics Letters* 82: 121–126.

Pesaran, M. H., Y. Shin, and R. P. Smith. 1997. Estimating long-run relationships in dynamic heterogeneous panels. DAE Working Papers Amalgamated Series 9721.

———. 1999. Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association* 94: 621–634.

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

Phillips, P. C. B., and H. R. Moon. 2000. Nonstationary panel data analysis: An overview of some recent developments. *Econometric Reviews* 19: 263–286.

About the authors

Edward F. Blackburne III is an associate professor in the Department of Economics and International Business, Sam Houston State University. His current research interests are benchmarking racial profiling and modeling asymmetric time series.

Mark W. Frank is an associate professor in the Department of Economics and International Business, Sam Houston State University. His current research interests are measuring effects of income inequality on economic growth and the economics of alcohol advertising.