

Chapter 11: Panel Data Models

Financial Econometric Modeling

Stan Hurn, Vance Martin, Peter C.B. Phillips and Jun Yu

Oxford University Press, 2020

- Panel data has combined measurements amounting to $N \times T$ observations, describing N individual units observed longitudinally over T time periods.
- Panel data are often classified into *long* panels where the time series sample size T is large and *wide* panels where cross section sample size N is large. Panels with T small and N large and are called short wide panels. Similarly, long narrow panels have large T and small N .
 - 1 Empirical research on asset pricing and investments makes use of an abundance of time series data and these problems are usually associated with large T and small N .
 - 2 Empirical corporate finance research typically have short wide panels with small T and large N .
- Panel data can be also classified into *balanced* panels and *unbalanced* panels.
 - 1 A balanced panel is one with no missing observations where every cell of the $T \times N$ matrix contains an observed data point.
 - 2 The presence of cells with missing observations yields an unbalanced panel as the sample sizes for each cross section member are no longer the same.

Types of Models

- When using panel data a complete set of equations for the system comprises an equation specified for each of the N cross sections. Such a complete system may be estimated jointly.
- Even though the combination of time series and cross section data produces a large number, $N \times T$, of observations, panel models also involve a large number of unknown parameters. These may relate to
 - 1 specific characteristics of each unit or individual in the cross section; or
 - 2 common forms of financial agent responses.
- Panel data models lead to large numbers of potentially nuisance parameters – called *incidental* parameters – and a much smaller number of homogeneous parameters that result from commonality restrictions that apply across the parameters of the N equations for each cross section unit.
- Panel models are often classified according to restrictions that are imposed on the parameters.
 - 1 No common effects model where no restrictions at all are imposed on the coefficients of the equations.
 - 2 Common effects where all equations have common coefficients imposed.
 - 3 Intermediate cases are where individual specific incidental parameters – usually the intercept coefficients in the equations – are allowed to differ but parameters on explanatory variables are constrained.

Reasons for Using Panel Data

- Methods of estimating panel models are designed to take advantage of the statistical power involved in cross section averaging whenever commonality is present. This typically leads to greater efficiency and higher rates of consistency in estimation.
- A panel may improve reliability in estimation and inference by mitigating the effects of some poor data that may be present in the panel by virtue of the averaging process that includes both high and low quality data.
- Panel data also enable analysts to monitor performance in financial assets over time and compare time profiles of different assets. In panels this is facilitated by pooling the cross section observations in a way that facilitates estimation of time series features, such as dynamic responses, that might reasonably be assumed to be common across the panel.
- Combining time series data and cross section data into a panel offers the prospect of improved accuracy in parameter estimation due to the augmented sample size (of magnitude $N \times T$). The improvement in statistical efficiency from having a larger sample size is reflected in smaller variances, sharper (more powerful) tests, and shorter confidence intervals.
- Rich data sets that come with panels also lead to an expanded range of hypotheses of interest. Certain hypotheses may be tested using panel data that are not possible with time series or cross section data alone.

Two Introductory Panel Models

No Common Effects

- The most unrestricted panel model is the no common effects model

$$y_{jt} = \alpha_j + \beta_j x_{jt} + u_{jt}, \quad u_{jt} \sim (0, \sigma_j^2),$$

where y_{jt} and x_{jt} are respectively the dependent and explanatory variables associated with the j^{th} cross section at time t .

- The no common effects model is the most unrestricted form within this class of panel model as the intercepts α_j , slopes β_j , and variances σ_j^2 are all allowed to vary across section, thereby providing a framework that accommodates full heterogeneity over individuals in the cross section.
- An example is the no common effects CAPM

$$\begin{array}{llll} \text{Stock 1:} & r_{1t} - r_{ft} & = & \alpha_1 + \beta_1(r_{mt} - r_{ft}) + u_{1t}, & u_{1t} \sim (0, \sigma_1^2), \\ \text{Stock 2:} & r_{2t} - r_{ft} & = & \alpha_2 + \beta_2(r_{mt} - r_{ft}) + u_{2t}, & u_{2t} \sim (0, \sigma_2^2), \\ & \vdots & & \vdots & \vdots \\ \text{Stock N:} & r_{Nt} - r_{ft} & = & \alpha_N + \beta_N(r_{mt} - r_{ft}) + u_{Nt}, & u_{Nt} \sim (0, \sigma_N^2). \end{array}$$

- This model can also be extended to allow for multiple explanatory variables.
- This decomposition of the no common effects CAPM means that the parameters of the model can be estimated by applying ordinary least squares to each equation separately.

Common Effects

- The common effects model involves extreme pooling of information. Each equation of this model is assumed to have the same intercept and same slope coefficients

$$y_{jt} = \alpha + \beta x_{jt} + u_{jt}, \quad u_{jt} \sim (0, \sigma_j^2),$$

where the regression coefficients are the same for each equation $j = 1, \dots, N$ but equation error variances σ_j^2 are heterogeneous.

The common effects CAPM has the explicit form

$$r_{jt} - r_{ft} = \alpha + \beta(r_{mt} - r_{ft}) + u_{jt},$$

with u_{jt} representing the idiosyncratic risk of the j^{th} stock at time t .

- For this model both the α -risks and the β -risks are assumed to be the same across all N stocks, although the model still allows for different idiosyncratic risks σ_j^2 .
- The common effects CAPM has only two regression parameters α and β . By comparison, the number of unknown parameters in the no common effects model is $2N$. The difference of $R = 2N - 2$ parameters represents the number of restrictions that the common effects model imposes on the no common effects model.

- To estimate the common effects model with the restrictions imposed the approach adopted is to stack the data. This is equivalent to expressing the data in **long** form.
- For the CAPM example, stacking requires that the dependent variable $r_{jt} - r_{ft}$ of each stock are stacked on top of each other to produce a new variable that is an $(NT \times 1)$ vector of excess returns on all N stocks over all time periods.
- Stacking the explanatory variable is achieved in exactly the same way except that the excess returns on the market are the same for all stocks, so the stacked explanatory variable simply stacks the same variable on itself N times to produce an $(NT \times 1)$ vector of excess returns of the market.
- Estimation then involves a linear regression of the stacked dependent variable on a constant vector – represented by an $(NT \times 1)$ vector of ones – and the stacked explanatory variable.
- This process of stacking the data and applying least squares to the stacked system imposes the cross-equation restrictions that the α -risks and the β -risks are the same across all N stocks.

- The common effects model is a restricted version of the no common effects model and these restrictions may be tested.
- The null and alternative hypotheses are

$$\begin{aligned}H_0 &: \alpha_1 = \alpha_2 = \dots = \alpha_N = \alpha \text{ and } \beta_1 = \beta_2 = \dots = \beta_N = \beta \\H_1 &: \text{at least one restriction fails.}\end{aligned}$$

- Under the null hypothesis there are two unknown parameters given by α and β compared to the $2N$ unknown parameters in the no common effects model.
- To test these restrictions, the F statistic is calculated according to the formula

$$F = \frac{RSS_R - RSS_U}{RSS_U} \times \frac{N(T-2)}{R},$$

where RSS_U is the unrestricted (no common effects) residual sum of squares, RSS_R is the restricted (common effects) residual sum of squares and R is the number of restrictions. This statistic is distributed as $F_{R, N(T-2)}$.

- An asymptotic version of the F test is the chi-square test which is obtained by multiplying both sides of the F statistic by R

$$R \times F = \frac{RSS_R - RSS_U}{RSS_U} \times N(T-2).$$

This statistic is distributed asymptotically under the null hypothesis as χ^2_R .

Fixed and Random Effects Models

The Fixed Effects Model

- The simple bivariate panel regression then has the following form

$$y_{jt} = \alpha_j + \beta x_{jt} + u_{jt}, \quad u_{jt} \sim (0, \sigma^2),$$

where the simplifying assumption of constant equation error variance over the cross sections is made, so that $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_N^2 = \sigma^2$.

- In this specification the α_j are considered fixed parameters by assumption, thereby giving the model its name. The fixed effects model is characterised by N intercepts ($\alpha_j : j = 1, \dots, N$) and 1 slope parameter (β). As discussed earlier, the intercepts are known as incidental parameters.
- The fixed effects model imposes $N - 1$ restrictions on the no common effects model by requiring the slope coefficient β to be common in each equation. Alternatively, a comparison of the fixed effects model and the common effects model suggests that the common effects model imposes $N - 1$ restrictions on the fixed effects model by requiring the constant term to α to be common across all equations.

Dummy Variable Representation

- An alternative way to specify the fixed effects model is to define a set of dummy variables as

$$D_{jt} = \begin{cases} 1 & \text{cross section } j \\ 0 & \text{otherwise,} \end{cases}$$

- The fixed effects model is now rewritten as

$$y_{jt} = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \cdots + \alpha_N D_{Nt} + \beta x_{jt} + u_{jt},$$

or more compactly

$$y_{jt} = \sum_{j=1}^N \alpha_j D_{jt} + \beta x_{jt} + u_{jt}.$$

- The fixed effects model is estimated by an ordinary least squares regression of the $(NT \times 1)$ stacked dependent variable y_{jt} on the $(NT \times 1)$ stacked explanatory variable x_{jt} and the N stacked dummy variables, $D_{1t}, D_{2t}, \dots, D_{Nt}$. Alternatively, if a constant is included in the regression equation then one of the dummy variables is excluded to avoid the dummy variable trap of collinear regressors.
- The dummy variable approach is only feasible in **long** panels as the number of dummy variables required by a wide panel would make the approach inefficient.

Estimating the Fixed Effects Model

A general method for estimating fixed effects models is as follows.

- 1 Express the equation for the j^{th} cross section in time-averaged form as

$$\frac{1}{T} \sum_{t=1}^T y_{jt} = \alpha_j + \beta \frac{1}{T} \sum_{t=1}^T x_{jt} + \frac{1}{T} \sum_{t=1}^T u_{jt}$$

$$\bar{y}_j = \alpha_j + \beta \bar{x}_j + \bar{u}_j.$$

- 2 Express the data in deviation-from-the-mean form, given by

$$y_{jt} - \bar{y}_j = \beta(x_{jt} - \bar{x}_j) + u_{jt} - \bar{u}_j,$$

in which the fixed effects α_j have been eliminated.

- 3 Estimate the resultant equation by ordinary least squares to obtain $\hat{\beta}$.
- 4 The estimates of the fixed effects for each of the j assets are recovered as

$$\hat{\alpha}_j = \bar{y}_j - \hat{\beta} \bar{x}_j,$$

where $\hat{\beta}$ is the least squares estimator from Step 3.

The Random Effects Model

- Assume that the differences in the cross section elements involve random variable intercepts so that

$$\alpha_j = \alpha + v_j, \quad v_j \sim (0, \sigma_v^2).$$

where the v_j are random effects that lead to the differences in the intercepts α_j .

- The v_j are assumed to be distributed with zero mean and constant variance, so that the random intercepts in each equation all have the same mean α . More importantly, the random errors v_j in these intercepts are typically assumed to be independent of the equation errors u_{jt} .
- The random effects model is then

$$y_{jt} = \alpha + v_j + \beta x_{jt} + u_{jt}, \quad u_{jt} \sim N(0, \sigma^2),$$

where, once again, the assumption of constant variance over the cross sections is used. The random effects, v_j , are assumed to be independent of both the disturbances u_{jt} , so that $E(v_j u_{jt}) = 0$, and the explanatory variable x_{jt} , so that $E(v_j x_{jt}) = 0$, for all j and t .

- The random effects model may be written in terms of a single disturbance as

$$y_{jt} = \alpha + \beta x_{jt} + w_{jt},$$

where the composite disturbance w_{jt} is

$$w_{jt} = u_{jt} + v_j.$$

Covariance of the Composite Error

- Estimation of the random effects model is somewhat more involved than the fixed effects model. The complexity arises from the composite form of the equation disturbance term w_{jt} .
- The covariance structure of w_{jt} is

$$E(w_{jt}) = E(u_{jt} + v_j) = E(u_{jt}) + E(v_j) = 0$$

$$\text{var}(w_{jt}^2) = E[(u_{jt} + v_j)^2] = E(u_{jt}^2) + E(v_j^2) + 2E(u_{jt}v_j) = \sigma^2 + \sigma_v^2$$

$$\begin{aligned}\text{cov}(w_{jt}w_{js}) &= E[(u_{jt} + v_j)(u_{js} + v_j)] = E(u_{jt}u_{js}) + E(u_{jt}v_j) + E(u_{js}v_j) + E(v_j^2) \\ &= \sigma_v^2.\end{aligned}$$

- There is therefore a specific correlation structure which is constant through time

$$\rho_{ts} = \frac{\text{cov}(w_{jt}w_{js})}{\sqrt{\text{var}(w_{jt})\text{var}(w_{js})}} = \frac{\sigma_v^2}{\sigma^2 + \sigma_v^2}.$$

- Estimation of the model must take account of this correlation structure.

Estimating the Random Effects Model

The steps involved to estimate the parameters of the random effects model in this more general case are as follows.

- 1 Choose starting values for σ^2 and σ_v^2 .
- 2 Define the quasi-difference parameter

$$\lambda = 1 - \frac{\sigma}{\sqrt{\sigma^2 + T\sigma_v^2}},$$

and compute the quasi-differenced data

$$\tilde{y}_{jt} = y_{jt} - \lambda \bar{y}_j, \quad \tilde{x}_{jt} = x_{jt} - \lambda \bar{x}_{jt},$$

where \bar{y}_j and \bar{x}_j are time averages.

- 3 The generalised least squares estimator based on the starting values for (σ^2, σ_v^2) is obtained by estimating the transformed equation

$$\tilde{y}_{jt} = \alpha(1 - \lambda) + \beta \tilde{x}_{jt} + \tilde{w}_{jt},$$

in which \tilde{w}_{jt} is a disturbance term, by ordinary least squares.

Hausman Test

- A formal test of fixed effects versus random effects in modelling panel data can be based on the fundamental assumption that the random effects, v_j , are considered to be independent of the explanatory variables, x_{jt} , that is $E(v_j x_{jt}) = 0$.
- Formally the hypotheses are

H_0 : random effects

H_1 : fixed effects.

- To perform the test, let $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ represent, respectively, the fixed and random effects estimates of the slope parameters.
- The general form of the test is based on the Wald statistic

$$WD = \left[\hat{\beta}_{FE} - \hat{\beta}_{RE} \right]' \left[\text{cov}(\hat{\beta}_{FE}) - \text{cov}(\hat{\beta}_{RE}) \right]^{-1} \left[\hat{\beta}_{FE} - \hat{\beta}_{RE} \right],$$

where $\text{cov}(\hat{\beta}_{FE})$ and $\text{cov}(\hat{\beta}_{RE})$ are the estimated covariances of $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$.

- The test statistic WD is distributed asymptotically under the null hypothesis as χ^2_R where R represents the number of explanatory variables in the model excluding the constant.

Dynamic Panel Data Models

Introducing Dynamics

- The panel data models discussed so far specify a relationship between the dependent variable, y_{jt} , and the explanatory variable, x_{jt} , both at time t .
- Dynamics may be introduced into the model by including lags of both y_{jt} and x_{jt} .
- The respective fixed effects and random effects versions of the so-called panel autoregressive model are as follows:

$$\text{Fixed effects} \quad : \quad y_{jt} = \alpha_j + \beta x_{jt} + \rho y_{jt-1} + u_{jt}$$

$$\text{Random effects} \quad : \quad y_{jt} = \alpha + v_j + \beta x_{jt} + \rho y_{jt-1} + u_{jt} .$$

where u_{jt} is the disturbance term which is assumed to be independent over time $E(u_{jt}u_{js}) = 0$, for all $t \neq s$.

- The autoregressive parameter ρ controls the strength of the dynamics and, in this formulation of the model, is assumed to take a common value across section.
 - 1 If $\rho = 0$ the model reduces to the non-dynamic or static panel data model.
 - 2 If $-1 < \rho < 1$ the autoregressive dynamics are stationary with the successive effects of shocks on the dependent variable dissipating over time.
 - 3 If $\rho = 1$ the shocks have persistent effects and do not dissipate over time, resulting in a nonstationary panel model.

Nickell Bias

- Elimination of the fixed effect intercepts from each equation by removing time-averages, produces a form of endogeneity in the regressor that influences estimation by introducing a new form of bias.
- The bias arises from the correlation of the lagged regressor with the equation error. The transformed lagged dependent variable $y_{jt-1} - \bar{y}_{jt-1}$ is correlated with the transformed disturbance term $u_{jt} - \bar{u}_j$ because

$$\text{cov}(\bar{y}_{jt-1}, u_{jt}) = \frac{\sigma^2}{T} \neq 0, \text{ for all } t.$$

- This correlation violates one of the key conditions needed for the validity of least squares. In particular, this condition is needed to ensure consistent estimation of the autoregressive coefficient ρ as $N \rightarrow \infty$ with finite time series sample size T .
- This non-zero correlation is present even when the disturbance term, u_{jt} , is itself not autocorrelated.
- The resultant bias in estimating ρ is commonly referred to as Nickell bias, following its original analysis by Nickell (1981). It has the asymptotic form

$$\text{plim}_{N \rightarrow \infty}(\hat{\rho} - \rho) \simeq -\frac{1 + \rho}{T}.$$

Illustrating Nickell Bias

- Consider a simple simulation experiment with a dynamic panel model to illustrate the impact of the Nickell bias. The data generating process is

$$\begin{aligned}y_{jt} &= \alpha_j + \beta x_{jt} + \rho y_{jt-1} + u_{jt}, & u_{jt} &\sim N(0, \sigma^2), \\x_{jt} &= \gamma x_{jt-1} + e_{jt}, & e_{jt} &\sim N(0, \sigma_e^2),\end{aligned}$$

with population parameters $\rho = \{0.2, 0.8\}$, $\beta = 0.8$, $\gamma = 0.5$ and $\sigma^2 = 1.0$.

- σ_e^2 is determined by specifying a value for the signal-to-noise ratio defined as

$$\sigma_s^2 = \text{var}\left(y_{jt} - \frac{1}{1-\rho}\alpha_j\right) = \beta^2 \sigma_e^2 \left[1 + \frac{(\gamma + \rho)^2}{1 + \rho\gamma}(\rho\gamma - 1) - (\rho\gamma)^2\right]^{-1} + \frac{\rho^2}{1 - \gamma\rho^2}\sigma^2.$$

The higher the value of σ_s^2 the more useful x_{jt} is in explaining y_{jt} .

- The settings $T = \{10, 20, 30\}$ reveal the performance of the fixed effects estimator in short, medium and long panels; and the settings $N = \{20, 100\}$ illustrate estimator performance in narrow and wide panels. Each parameter combination is assessed using 1000 replications.
- Results should demonstrate that the fixed effects estimator of ρ suffers from significant bias and that the bias of the fixed effects estimator increases with ρ and decreases as T increases.

Results

N	T	$\rho = 0.2$		$\rho = 0.8$	
		Mean	Std. Dev.	Mean	Std. Dev.
20	10	0.138	0.051	0.560	0.070
20	20	0.175	0.034	0.693	0.040
20	30	0.181	0.027	0.731	0.029
100	10	0.140	0.023	0.567	0.030
100	20	0.172	0.015	0.694	0.018
100	30	0.182	0.012	0.733	0.013

Estimating Dynamic Panel Models

- To avoid Nickell bias is to adopt a two step approach.
 - Remove the fixed effects (α_j) or random effects (v_j) by taking first differences

$$\Delta y_{jt} = \beta \Delta x_{jt} + \rho \Delta y_{jt-1} + \Delta u_{jt}.$$

While this equation removes intercepts, the differencing operation induces correlation between the lagged differenced regressor Δy_{jt-1} and the equation error Δu_{jt} . In particular $E(\Delta y_{jt-1} \Delta u_{jt}) = -E(y_{jt-1} u_{jt-1}) = -\sigma^2$. Thus, least squares estimation leads to biased and inconsistent estimates of the remaining parameters $\{\beta, \rho\}$.

- To correct the bias and inconsistency of least squares estimation, instrumental variable or GMM methods are employed. Suitable instruments may be constructed using carefully chosen lagged values of the dependent variable and exogenous regressors.
- In choosing the set of instruments to employ at each time period t , there are three commonly-used approaches which involve different choices of instruments. In general all the approaches use combinations of lagged levels and differences of the dependent variable as suitable instruments.

Nonstationary Panel Data Models

Preamble

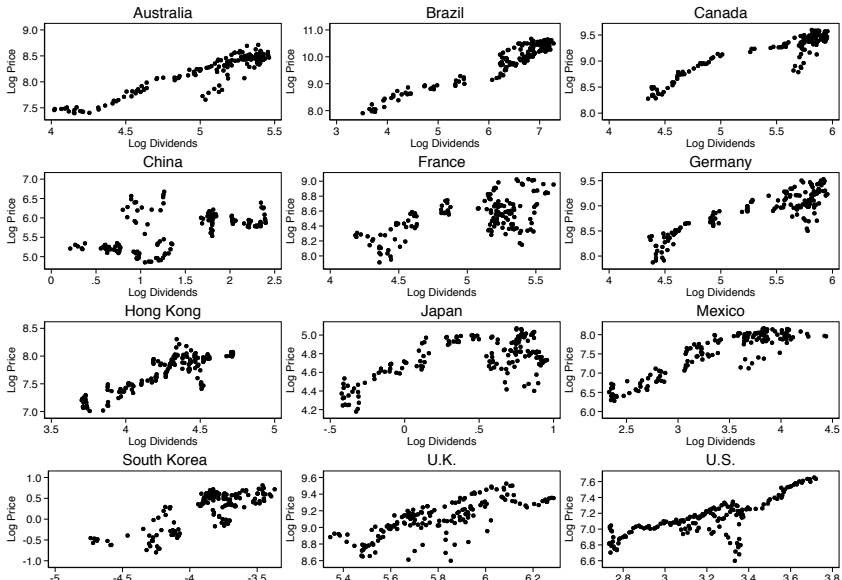
- If the autoregressive coefficient ρ in a dynamic panel model does not satisfy the stability restriction $|\rho| < 1$, the variables for each cross section are nonstationary and may be cointegrated with common long run parameters.
- An example is the present value model applied to a cross section of countries. The question is whether these countries have a common cointegrating vector that embodies the same present value relationship.
- A general specification of the present value model for a set of N , countries is written as

$$p_{jt} = \alpha_j + \beta_j d_{jt} + u_{jt}, \quad j = 1, 2, \dots, N,$$

where p_{jt} is the log equity price of country j , d_{jt} is the log dividend of country j and u_{jt} is a disturbance term.

- Assuming that p_{jt} and d_{jt} are nonstationary time series and u_{jt} is stationary, the parameters α_j and β_j represent the cointegrating parameters which in this formulation may vary for each country. T
- The parameter α_j is a function of the relevant discount factor that relates to country j .
- This model is comparable to the no common effects panel model with the parameters for each cross section being estimated separately using appropriate methods.

Panel Scatter Plots of the Present Value Model for 12 Countries



- Restricting the slope parameters to be the same across the N countries

$$\beta_1 = \beta_2 = \dots = \beta_N = \beta,$$

imposes a common long run relationship between equity prices and dividends.

- The result is a present value fixed effects cointegrating model with homogeneous slopes of the form

$$p_{jt} = \alpha_j + \beta d_{jt} + u_{jt}.$$

- The fixed effects are given by the parameters $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$, which allow the discount factor for each country to differ.
- A further set of restrictions is to assume that the discount factors are also the same across all N countries by imposing the restrictions

$$\alpha_1 = \alpha_2 = \dots = \alpha_N = \alpha.$$

- The result is the present value common effects cointegrating model

$$p_{jt} = \alpha + \beta d_{jt} + u_{jt}.$$

Panel Unit Root Tests - LLC

- Rather than testing each variable separately for a unit root, a test statistic is constructed to assess the presence of a unit root jointly among all N variables (or cross sections) simultaneously. This strategy can be interpreted as an efficient procedure of testing for the presence of a time series unit root in a variable that varies not only over time but also over N cross sections.
- The Levin, Lin and Chu (2002) test is based on the panel unit root regression equation

$$\Delta y_{jt} = \alpha_j + \beta y_{jt-1} + \sum_{k=1}^L \gamma_{jk} \Delta y_{jt-k} + u_{jt}.$$

- The null and alternative hypotheses of the panel unit root test are given by

$$\begin{aligned} H_0 &: \beta = 0 && \text{[unit root]} \\ H_1 &: \beta < 0 && \text{[stationary]}. \end{aligned}$$

Note that the null hypothesis requires that all time series in the panel have unit roots ($\rho := \beta + 1 = 1$) and the alternative similarly requires that all time series have stationary roots ($|\rho| < 1$).

- This model is the panel analogue of the augmented Dickey-Fuller regression equation with $\beta = \rho - 1$. Note however that the panel unit root t statistic is asymptotically distributed as $N(0, 1)$ under the null hypothesis.

Panel Unit Root Tests - IPS

- A less restrictive panel unit root test that allows for some stationary and nonstationary elements in the cross section under the alternative hypothesis was proposed by Im, Pesaran and Shin (2003).
- This test involves the weaker formulation

$$\Delta y_{jt} = \alpha_j + \beta_j y_{jt-1} + \sum_{k=1}^L \gamma_{jk} \Delta y_{jt-k} + u_{jt},$$

in which the slope coefficients β_j are now not restricted.

- The null and alternative hypotheses for this test are

$$H_0 : \beta_j = 0, \quad j = 1, 2, \dots, N$$

$$H_1 : \begin{cases} \beta_j < 0 & \text{for } j = 1, \dots, N_1 \\ \beta_j = 0 & \text{for } j = N_1 + 1, \dots, N. \end{cases}$$

Under the null hypothesis all cross sections have unit roots as before, whereas under the alternative hypothesis just N_1 of the N cross sections are stationary.

- An overall panel unit root test is computed by estimating separately for each cross section and using a t statistic to test that $\beta_j = 0$. An overall statistic is computed by averaging across the N values of these t statistics. The averaged t statistic the asymptotic distribution of the modified panel unit root test is also standard normal.

Panel Cointegration

- As with cointegration testing in non-panel data models there exists a range of test procedures for cointegration in dynamic panel models. Many of these tests rely on Fisher's (1932) idea of combining individual tests to produce a meta-analysis test.
- If π_j is the p value of the test for cross section j for a particular hypothesis, the joint statistic based on the N cross sections is

$$-2 \sum_{j=1}^N \log(\pi_j) \rightarrow \chi_{2N}^2,$$

which is asymptotically distributed as chi square with $2N$ degrees of freedom when $T \rightarrow \infty$.

- Having established cointegration in the panel, the cointegrating parameter vector β together with the fixed effects α_j , are estimated.
- A dynamic least squares approach augments the panel equation to include leads and lags of Δd_{jt} as follows

$$p_{jt} = \alpha_j + \beta d_{jt} + \sum_{k=-L}^L \gamma_{jk} \Delta d_{jt-k} + u_{jt},$$

which is the estimated by ordinary least squares.

Panel Cointegration

- As with cointegration testing in non-panel data models there exists a range of test procedures for cointegration in dynamic panel models. Many of these tests rely on Fisher's (1932) idea of combining individual tests to produce a meta-analysis test.
- In testing for panel cointegration, the procedure is to apply (for instance) the Johansen cointegration test to each cross sectional unit separately. If π_j is the p value of the test for cross section j for a particular hypothesis, the joint statistic based on the N cross section outcomes is constructed as

$$-2 \sum_{j=1}^N \log(\pi_j) \rightarrow \chi_{2N}^2,$$

which is asymptotically distributed as chi square with $2N$ degrees of freedom.

- A dynamic least squares approach augments the panel equation to include leads and lags of Δd_{jt} as follows

$$p_{jt} = \alpha_j + \beta d_{jt} + \sum_{k=-L}^L \gamma_{jk} \Delta d_{jt-k} + u_{jt},$$

which is estimated by ordinary least squares.

- The estimator of the cointegrating parameter vector β has t ratio statistics that are asymptotically distributed as $N(0, 1)$.