2. Stationary versus Cointegrated Regressions Models

In this paper the relationship between FSP caseloads and the macro-economy is analyzed based on evidence that these data follow unit-root processes. To understand how the approach with unit-root processes differs from the one used in previous FSP caseload studies it is necessary to review some fundamentals of cointegration econometrics by comparing them with a more familiar regression theory.

Previous studies have measured the relative importance of policy and economic variables on *FSP* or *AFDC/TANF* caseloads by regressing caseloads on a vector of economic and policy variables. All these studies use methods that are correctly applied to *stationary regression models* in which the model is formed from variables that follow stationary time series. In this section, stationary regression models are compared to integrated regression models. Integrated regressions are regression models constructed from variables that follow unit root time series. [6] The fundamental difference between a stationary and a unit root series is the response of the series to a transitory shock in a single period. Differences in regression methods, predictive reliability, and most importantly for the purposes of this paper, coefficient interpretation can all be traced to this difference.

This section proceeds in three parts. First, a brief overview of the difference between stationary and cointegration regression models is presented. A more formal discussion of difference between stationary versus integrated data and between regression analyses with stationary versus integrated data follows. Finally, unique features of estimating cointegated regression models using panel data are discussed.

2.1 Overview

Suppose initially that the time series of food stamp caseloads can be described as a trend stationary (or stationary) process. Following a transitory shock in the present period food stamp caseloads first deviate from, and then return to their original time trend (to their mean if stationary). The transitory shock imparts a *short-run* effect on the series in the sense that FSP caseloads returns to the same path they would have followed had the stock not occurred. If economic and policy variables are also trend stationary, each would display the same temporary deviation from their time trends. Because past deviations from these trends have been temporary, the future paths of these variables are predictable. This means that a regression of food stamp caseloads on economic and policy variables yields stable estimates and that these estimates describe a short-run relationship between deviations of food stamp caseload and economic and policy variables from their time trends.

The sequence of events is different if the FSP caseload data are drawn from unit root processes. In this case a transitory shock to one of these variables forever alters the future path of its series and so imparts a *long-run* or permanent effect on the time series. Past responses do not

[6] There are a number of different classes of integrated time series and in this paper we follow the convention that an integrated series is taken to mean an integrated series of order 1 or equivalently, that the time series contains a single unit root. This section draws heavily on Chapters 15-19 of Hamilton for the discussion of unit root nonstationary series and cointegration.

'trace out' a stable path; instead, any time trend would be stochastic and future paths cannot be reliably predicted from past patterns.

Even though the individual unit-root data series are each non-stationary it is possible that a linear combination of these variables is stationary and thus predictable. Such a linear combination defines a *cointegrated* relationship between the variables, and explains how the individual data series move together in a way that can be reliably predicted from past realizations of their stationary combination. In this sense, a cointegrating regression among FSP caseloads, economy, and policy variables describes a stable long-run (equilibrium) relationship between permanent movement in the FSP caseloads and permanent movements in the economic and policy variables.

If no such linear combination exists, the estimated OLS relationship among the variables would be *spurious* and the linear combination defined by the OLS coefficients would behave like a non-stationary time series and yield unreliable predictions. Tests of cointegration are, therefore, tests of the existence of long-run equilibrium relationships, and can be interpreted as tests of model specification. It is this aspect of cointegration analysis that is emphasized in this paper. In particular, our specification tests find evidence that estimates of the FSP caseload equation that *do not* include a measure of AFDC/TANF caseloads are incomplete and do not define a long-run stable relationship.

2.2 Derivation of Results

This comparison will first look at differences in the time series properties of a single (univariate) random variable generated by a stationary versus a unit root process. The section then addresses how these properties affect stationary versus integrated regressions.

For a given a random variable, x_t, let the process or time series be denoted as $\{x_t\}$. Stationary time series have well-defined first and second moments that make them relatively easy to predict. For example suppose that an economy or policy random variable, x_t, follows an auto-regressive process of order 1 (i.e., *AR(1)*) and so that it satisfies

$$x_t = c + \rho x_{t-1} + \varepsilon_t$$

where ε_t is a serially uncorrelated or transitory shock with $Var(\varepsilon_t) = \sigma^2$ for all t, and where $|\rho|<1$. The condition $|\rho|<1$ represents the absolute summability condition for a stationary *AR(1)* time series. Let E denote the mathematical expectations operator, then the mean of this process is $E(\{x_t\}) \equiv \mu = c/(1-\rho)$, its variance is $Var(\{x_t\}) \equiv \gamma_o = \sigma^2/(1-\rho^2)$, and its jth auto-covariance, γ_j, is $\sigma^2[\rho^j/(1-\rho^2)]$. In this stationary *AR(1)* example, finite unconditional moments of the process are independent of t and the covariance of variables separated by longer time periods decays (i.e., $\gamma_j \rightarrow 0$) as the distance between the variables grows larger (i.e., $j \rightarrow \infty$). It is the summability condition that ensures the existence of finite moments and is responsible for the type of decay exhibited by stationary time series. As discussed below it also has important implications for prediction.

A more general representation of a stationary process satisfies

(1) $$x_t = \mu + \varepsilon_t + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \rho_3 \varepsilon_{t-3} + \ldots$$

where $E(\{x_t\}) = \mu$, and where $\sum_{i=0}^{\infty}|\rho_i| < \infty$ is the absolute summability condition. Besides ensuring finite moments, absolute summability also ensures that as a forecast horizon grows, the forecast converges to μ, the unconditional mean of the series. That is if $\sum_{i=0}^{\infty}|\rho_i| < \infty$, then $E(x_{t+s}|x_t, x_{t-1}, \ldots) \rightarrow E(\{x_t\}) = \mu$ as $s \rightarrow \infty$.

The problem with *(1)* is it does not generate the type of trends that are commonly observed in economic data. One possibility that does generate typical patterns is a *trend-stationary* series

(2) $$x_t = \mu + \delta t + \varepsilon_t + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \rho_3 \varepsilon_{t-3} + \ldots$$

In this case absolute summability ($\sum_{i=0}^{\infty}|\rho_i| < \infty$) ensures that forecasts of x_{t+s}, given by $E(x_{t+s} | x_t, x_{t-1}, \ldots)$ converge (in mean square) to the time trend $\mu + \delta(t+s)$ and the mean-squared error *(MSE)* of the forecast converges to the bounded unconditional variance of the series (i.e.,) as $s \rightarrow \infty$. That is, MSE $\equiv E[x_{t+s} - E(x_{t+s} | x_t, x_{t-1}, \ldots)]^2 \rightarrow Var\{(x_t)\} = \sigma^2 [1 + \rho_1 + \rho_2 + \rho_3 + \ldots]$ as $s \rightarrow \infty$ and the added uncertainty of forecasting further into the future becomes negligible.

Another way to generate trends commonly observed with economic data is through a unit root non-stationary series. The first difference of a unit-root series is a stationary series but the series itself is not. This means the first-differenced series satisfies absolute summability but the level series does not. Violating absolute summability not only means finite moments do not exist, it also has dramatic consequences for prediction.

A unit root series that allows for trends satisfies

(3) $$x_t - x_{t-1} = (1-L)\, x_t = \delta + \varepsilon_t + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \rho_3 \varepsilon_{t-3} + \ldots$$

where the symbol L denotes the lag operator, which transforms x_t into x_{t-1} upon multiplication. Note that *(3)* is similar to *(1)* except that $(1-L)\, x_t$ replaces x_t, and the drift term, δ, replaces the unconditional mean μ. It can be shown (e.g., see Hamilton, Chapter 15) that the s-step-ahead forecast of a series that follows *(3)* is

$$E(x_{t+s} | x_t, x_{t-1}, \ldots) = s\,\delta + x_t + T(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)$$

where T is some function of current and past realizations of $\{\varepsilon_t\}$. Even in the case in which the function T equals zero (i.e., a random walk with drift) the above relationship indicates that an s-step ahead forecast is a random variable that grows at a constant rate (δ) from the current realization of x_t. This means that over time the prediction of x_{t+s} changes as new values $x_{t+1}, x_{t+2}, \ldots$ are realized. It can also be shown that the *MSE* associated with such a prediction grows linearly with the forecast horizon (Hamilton, Chapter 15).

The above discussion suggests stationary series are much easier to predict than unit root series. For a trend stationary process the forecast of x_{t+s} converges to the *value* $\mu + \delta(t+s)$. On the other hand, for a random walk with drift process, the forecast of x_{t+s} does not converge to a single value but constantly changes as future values of the series are realized. This means that it is virtually impossible to accurately predict the future values of a unit root series.

Differences in the time series characteristics of stationary and integrated data affect the properties of linear regression. Let y_t represents food stamp caseloads, $\mathbf{x}_t$ is represents a k-vector of policy and economic variables, $\boldsymbol{\beta}$ a corresponding vector of true parameters and u_t a model error. (A variable in bold indicates a vector.) Write a regression model relating food stamp caseloads to policy and economic variables as,

(4) $$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$$

Conditional on current values of economy and policy regressors equation *(4)* can be viewed as providing a prediction of current food stamp caseloads. Given the general unreliability of forecasts of integrated series, a central question is whether y_t can be reliably predicted from $\mathbf{x}_t$ if y_t and $\mathbf{x}_t$ follow unit root processes.

Given the regression framework by (4) three separate cases are considered. First suppose that the $\{[y_t \ \mathbf{x}_t']\}$ are jointly covariance stationary. Any linear combination of these series forms a stationary series. Hence the model errors

$$[y_t \ \mathbf{x}_t'] \ [1 \ -\boldsymbol{\beta}]' \equiv u_t$$

would be stationary. Suppose $E(u_t | \ \mathbf{x}_t') = 0$ and suppose u_t is a serially uncorrelated process with a constant and finite variance σ^2 for all t.[7] Denote $\mathbf{b}_T$ as the vector of *OLS* estimates of $\boldsymbol{\beta}$ based on T observations, so the model residuals, $u_t^* = [y_t \ \mathbf{x}_t'] \ [1 \ - \mathbf{b}_T]'$ and the residual sum of squares is $RSS_T = \sum_{t=1}^{T}(u_t^*)^2$. Then *OLS* estimation yields the following well- known results: $\mathbf{b}_T$ is a consistent estimate of $\boldsymbol{\beta}$, $s^2 = (1/T\text{-}k)\ RSS_T$ is a consistent estimate of σ^2, and as $T \rightarrow \infty$, $\mathbf{b}_T$ is approximately distributed multivariate normal with mean $\boldsymbol{\beta}$ and variance $\sigma^2[\sum_{t=1}^{T}(\mathbf{x_t x_t}')]^{-1}$. Thus if observations on $[y_t \ \mathbf{x}_t']$ *(t=1,..., T)* are drawn from jointly stationary or trend-stationary processes, standard *OLS* yields consistent estimates of $\boldsymbol{\beta}$ and the usual t and F tests yield valid inference. Furthermore, these results do not change if one or more of the elements of $\mathbf{x}_t'$ are deterministic time trends. In short, if $\{[y_t \ \mathbf{x}_t']\}$ are jointly covariance stationary, *(4)* provides a stable prediction of food-stamp caseloads, and this prediction can be estimated consistently and efficiently using *OLS*. These are the standard type of results that are used by previous FSP caseload studies.

Now consider the second case in which $\{[y_t \ \mathbf{x}_t']\}$ is a vector of unit root non-stationary processes. In this case the model is

[7] For serially correlated u_t one can appeal to a *GLS* transformation that achieves serially uncorrelated errors.

(4a) $$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$$

(5) $$\mathbf{x}_t = \delta + \boldsymbol{x}_{t-1} + \boldsymbol{\nu}_t.$$

where *(4a)* is an integrated regression, and the data generating process of the regressor variables has been written out explicitly in (*5*) with error vector $\boldsymbol{\nu}_t$ consisting of stationary terms. The properties of *(4a)* depend crucially on the properties of the model errors.

It might be expected that a linear combination of unit-root non-stationary processes would be a unit root non-stationary process. Specifically $\{[y_t \ \mathbf{x}_t'] \ [1 \ -\boldsymbol{\beta}]'\} \equiv \{u_t\}$ would be a unit root non-stationary process and *(4a)* would represent an integrated regression with integrated model errors. The above discussion on univariate time series states that the prediction of unit root variables is both unstable and uncertain. The relevance of that discussion for regression analysis derives from the view that the *OLS* residual

(6) $$u_t^* = [y_t \ \mathbf{x}_t'] \ [1 \ - \mathbf{b}_T]'$$

represents a prediction of the model error based on values of $[y_t \ \mathbf{x}_t']$. Recall from above that $\mathbf{b}_T$ denotes the *OLS* estimate of $\boldsymbol{\beta}$ based on T observations. The above discussion on the prediction of univariate unit root process indicates that the predictor u_t^* varies even as the sample size becomes infinitely large. From *(6)* this means that even in large samples $\mathbf{b}_T$ continues to vary and fails to converge to $\boldsymbol{\beta}$. The *MSE* calculated from the predicted residuals, $s^2 = (1/(T-k)) \ RSS_T = (1/T-k) \sum_{t=1}^{T}(u_t^*)^2$ grows at the rate T^2 . This divergence of the RSS_T means that F-statistics calculated from the OLS residuals grow at the rate T. As the sample size grows, F tests would be more likely to reject the null that $\mathbf{x}_t$ has *no* relationship with y_t even though $\mathbf{b}_T$ remains unstable in large samples. Phillips (1986) was the first to formally show that these properties indeed represent the hallmark of integrated regressions with integrated model errors. Integrated regressions with integrated model errors are called *spurious regressions*.

Finally, consider the third case; the case of an integrated regression with stationary model errors. The above discussion of the prediction of univariate stationary time series suggests that in this case the model residuals, $[y_t \ \mathbf{x}_t'] \ [1 \ - \mathbf{b}_T]$ converges to $[y_t \ \mathbf{x}_t'] \ [1 \ -\boldsymbol{\beta}]$ as the sample size increases which suggests $\mathbf{b}_T$ converges to $\boldsymbol{\beta}$. Integrated regressions with a stationary model error are called *cointegrated regressions.*

It is important to note that in the case of cointegrated regressions, the consistency of $\mathbf{b}_T$ does not depend on econometric exogeneity. That is, with cointegrated regressions OLS is consistent even when $\text{Cov}(u_t, v_t) \neq 0$, i.e., the error terms in (4a) and (5) are correlated. To see this, note that

$$\mathbf{b}_T = [\textstyle\sum_{t=1}^{T} (\mathbf{x}_t \ \mathbf{x}_t')]^{-1} [\sum_{t=1}^{T} \mathbf{x}_t \ y_t] = [\sum_{t=1}^{T} (\mathbf{x}_t \ \mathbf{x}_t')]^{-1} [\sum_{t=1}^{T} \mathbf{x}_t \ (\mathbf{x}_t'\boldsymbol{\beta} + u_t)]$$

$$= \boldsymbol{\beta} + [\textstyle\sum_{t=1}^{T} (\mathbf{x}_t \ \mathbf{x}_t')]^{-1} \sum_{t=1}^{T} (\mathbf{x}_t \ u_t).$$

The bias term, $[\sum_{t=1}^{T} (\mathbf{x}_t \ \mathbf{x}_t')]^{-1} \sum_{t=1}^{T} (\mathbf{x}_t \ u_t)$, is the sum of the product of a unit root vector of variables and a stationary variable (i.e., $\sum_{t=1}^{T} (\mathbf{x}_t \ u_t)$) divided by the sum of squares of a unit root

vector of variables (i.e., $\sum_{t=1}^{T} (\mathbf{x}_t \ \mathbf{x}_t')$). It can be shown that in the special case of a cointegrated regression $\sum_{t=1}^{T} (\mathbf{x}_t \ u_t)$ grows at the rate T and $\sum_{t=1}^{T} (\mathbf{x}_t \ \mathbf{x}_t')$ grows at rate T^2. Because the denominator of the bias term grows at a rate in time greater than the numerator, the bias term converges (in law) to zero and *OLS* estimates converge to $\boldsymbol{\beta}$.

In the context of the present study the above result means that if food stamp caseloads and the policy and economy variables in $\mathbf{x}_t$ are cointegrated, then *OLS* would provide consistent estimates of the cointegrating vector $[1 \ -\boldsymbol{\beta}]$ even though economic and policy variables may not be econometrically exogenous. Moreover this regression would represent a long-run relationship in the sense that the cointegrating vector $[1 \ -\boldsymbol{\beta}]$ relates permanent changes in food stamp caseloads to permanent changes in the economic and policy variables.

To illustrate the nature of the long-run relationship described by the cointegrated regression (*4a*), consider a single element of the integrated vector $\mathbf{x}_t$ which is a unit root process without drift. According to *(3)*, this element satisfies

(3a) $$x_t - x_{t-1} = u_t$$

where $u_t = \delta + \varepsilon_t + \rho_1 \ \varepsilon_{t-1} + \rho_2 \ \varepsilon_{t-2} + \rho_3 \ \varepsilon_{t-3} + \ldots$. Repeated substitution gives

$$x_t = x_o + u_1 + u_2 + u_3 + \ldots + u_t$$

so the change to x_t from x_o is the sum of serially correlated events. Note the realization of u in period 1 affects every realization of $\{x_t\}$ drawn in any future time period ($t>1$) in exactly the same way. The Beveridge-Nelson decomposition shows that *(3a)* can be expressed as

(7) $$x_t = x_o + \rho(1) \ (\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \ \ldots + \varepsilon_t) + \eta_t - \eta_o$$

where ε is a serially uncorrelated or transitory shock, η_t is a stationary process, x_o and η_o are initial conditions, $\rho(1) = \sum_{k=0}^{\infty}(\rho_k)$, and $\rho(1) \ (\varepsilon_1 + \ \varepsilon_2 + \varepsilon_3 + \ \ldots + \varepsilon_t)$ represents a random walk.[8] *The permanent or long run component of an integrated time series is described by the random walk component of a unit root series*. This random walk component shows that a purely transitory event that occurs in say, the first period (ε_1), imparts a permanent effect on all future realizations of the random variables x_t , x_{t+1} , x_{t+2}, . . . in exactly the same manner. Furthermore, for unit root series without drift, it is only this permanent component that matters when characterizing the distribution of a unit root time series (Hamilton, Ch. 17).

Hence if food stamp caseloads $\{y_t\}$ is an integrated process (without drift) and if $\{\mathbf{x}_t\}$ is comprised of integrated policy and economy variables (without drift), and if $\{[y_t \ \mathbf{x}_t]\}$ are cointegrated with cointegrating vector $[1 \ -\boldsymbol{\beta}]$ then if the permanent component or long run component of $\mathbf{x}_t$ is

$$(\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + \boldsymbol{\varepsilon}_3 + \ \ldots + \boldsymbol{\varepsilon}_t)' \ \boldsymbol{\rho(1)}$$

[8] More precisely it is the product of a constant and a random walk.

the long run component of $\{y_t\}$ can be represented as

$$(\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + \boldsymbol{\varepsilon}_3 + \ldots + \boldsymbol{\varepsilon}_t)' \, \boldsymbol{\rho}(1) \, \boldsymbol{\beta}.$$

Therefore the cointegrated food stamp caseload regression links through the coefficient vector $\boldsymbol{\beta}$ the long-run components of economic and policy variables to the long-run component of food stamp caseloads.

If a food stamp caseload regression is a cointegrated regression, the long-run components of unit-root economic and policy variables dominates the distribution of food stamp caseloads. It even affects the short-run fluctuations or stationary component of caseloads. To see this, write the period-by-period change in FSP caseloads in an *error-correction form*. This form can be obtained by substitute *(5)* into *(4a)* and subtract y_{t-1} from both sides.

(8) $$\Delta y_t = (\mathbf{x}_{t-1}'\boldsymbol{\beta} - y_{t-1}) + \Delta \mathbf{x}_t'\boldsymbol{\beta} + u_t$$

where *'Δ'* is the first difference operator (i.e., $\Delta y_t = y_t - y_{t-1}$). Equation *(8)* shows that period-by-period changes in food stamp caseloads can be expressed in terms of the cointegrating vector and lagged *levels* of integrated food stamp caseloads and economic and policy variables (the error-correction term) and period-by-period changes in the economic and policy variables. Equation *(8)* shows that a simple first-difference representation of *(4a)*, although stationary, is misspecified since it does not include the error-correction term and thus severs the link between short-run dynamics and the long-run relationship.

Even though *OLS* yields consistent estimates of $\boldsymbol{\beta}$, the usual *t* and *F* hypothesis tests associated with estimating *(4a)* yield correct inference only under the stringent condition of econometric exogeneity. Strict econometric exogeneity is achieved in *(4a)* and *(5)* when $Cov(u_t, \boldsymbol{v}_t) = 0$, however in general, one should expect that this condition will be violated. Phillip and Hanson (1990) have developed a *fully modified (FM) estimator* that corrects for the effects of this correlation. FM estimation involves a transformation the data. The usual t and F tests calculated from OLS estimates applied to the transformed data results in asymptotically valid tests of $\boldsymbol{\beta}$. Park's (1992) *Canonical Cointegrating Regressions (CCR)* estimator used in this paper is also a fully modified estimator.

In order to determine whether an integrated regression model is described by second case (spurious regression) or by the third case (cointegrated regression) tests of cointegration are conducted. All cointegration tests are based on whether regression residuals behave like a stationary or a unit-root nonstationary process. One approach to testing for cointegration is to check whether the model residuals behave like a unit root process. This approach tests the null hypothesis that the relationship is *not* cointegrated. In this case one applies the Dickey-Fuller (1981) or augmented Dickey Fuller tests (for residuals) or the tests described by Phillips and Ouliaris (1990). Rejecting the null of integrated model residuals rejects the null that the integrated regression is spurious. A problem with these residual tests is their low power, which owes to the fact that residuals must be estimated before they can be tested (Dickey, Jansen, and Thornton, 1991). Another approach to testing for cointegration is to test the null hypothesis that the regression *is* cointegrated using a procedure developed by Park (1990). His variable addition

test exploits the fact that under the null of cointegration the stationary model error tests are uncorrelated with additional integrated variables or deterministic time trends. In this case, an *F*-test (on transformed regression variables) will fail to reject the null of zero coefficients associated with additional integrated regressors. That is, under the null of cointegration, the additional integrated regressors are superfluous and an *F*-test will imply that coefficients of the added variables are zero. Under the alternative, the model residuals are unit-root non-stationary and the *F*-test rejects the null of zero coefficients.

In this paper, cointegration is tested using variable addition tests. This approach to testing cointegration is taken because we believe that there exists a long-run relationship between the FSP and AFDC/TANF caseloads and the economy and look for evidence to refute this belief. Given this belief, variable addition tests provide an advantage over residual tests. The low power of residual based tests of the null of unit-root nonstationary, like the augmented Dickey-Fuller test, means that these tests are not good at detecting relationships that are, in fact, cointegrated. A second advantage of variable additions test is that they are based on standardized distributions. This avoids complications associated with relying on non-standard testing procedures that are not currently well developed for testing residual from regressions of panel data.

We use tests of cointegration as tests of model specification. If a regression model is found to be spurious, this is interpreted to imply that the specification is misspecified. One way this could happen would be if too few integrated regressors were included in the model. For example, if there were a cointegrating relationship between FSP caseloads, the economy, and AFDC/TANF caseloads but an FSP caseload equation was estimated that included only economy regressors then the regression would be spurious. This distinction between cointegrated and spurious regression equations provides a criterion for evaluating previously estimated specifications of the FSP caseload equation.

2.3 Panel Cointegrating Regressions

Versions of the FSP caseload equation will be estimated using annual state level panel data from 1980 to 1999. The obvious advantage of these data is they include both the variation of FSP caseloads over time as well as the variation across states. It is anticipated that the additional variation will improve the quality of the estimated FSP caseload equation.

Let y_{it} denote *FSP* caseloads and $\mathbf{x}_{it}$ denote the vector of policy and economic variables in the *i*th panel in time *t*. The panel model that is estimated in this paper is of the form

$$y_{it} = \mathbf{x}_{it}'\beta + u_{it}$$

This specification of the caseload equation assumes that the relationship between the model variables is the same or *homogeneous* across the states.[9] Previously studies that have estimated

[9] Methods by which standard panel models adjust for individual panel fixed or random effects, such as demeaning or detrending, can be incorporated within the homogeneous panel case so that these individual effect do not make the panels heterogeneous.

the FSP caseload equation using state level panel data have assumed a homogenous relationship across states. In order to faciliate the comparison with these studies, we also assume of a homogeneous relationship across the states. The rationale for assuming a homogenous panel is quite different, however, depending on whether the data are generated by trend stationary versus integrated data processes.

With integrated data, the justification for a homogenous response in each state is based on the notion that the FSP is a national program with uniform eligibility requirements in which states are subjected to standardized performance criteria and are provided the same (percentage) reimbursement for administrative costs. Under these circumstances we might reasonably expect a common (homogenous) *long-run* relationship linking FSP caseloads, economy, and policy variables in each of the states. With integrated data, the assumption of a common long-run relationship implies that in any state a given change in a regresssor will cause the same change in the long-run FSP caseload level; however, the path of adjustment to the new equilibrium may be different for each state.[10] On the other hand, for trend stationary data the notion of a homogeneous panel implies much stronger assumptions. Under the assumption of trend stationary data, a homogenous panel implies that the short-run (year-to-year) path of adjustment of FSP caseloads to their time trend would be the same in each state.

Assume the elements of $[y_{it} \ \mathbf{x}_{it}']$ are drawn from unit root non-stationary processes, and for any fixed time period t, the $[y_{it} \ \mathbf{x}_{it}']$ are identically and independently distributed across *panels*. Let Ω_i denote the long-run covariance matrix of $\{[y_{it} \ \mathbf{x}_{it}']\}$ for the ith panel, and note the Ω_i are distributed *iid* across panels. Given mild moment and summability conditions, this matrix is integrable and therefore can be averaged over panels so that $\Omega = E(\Omega_i)$ denotes the long-run average covariance matrix of $[y_{it} \ \mathbf{x}_{it}']$. Elements of Ω include Ω_{yy} (i.e., the long-run average moment matrix of y_{it}), Ω_{xx} (i.e., the long-run average moment matrix of $\mathbf{x}_{it}$), and Ω_{xy}(i.e., the long-run average cross-moment vector of $[y_{it} \ \mathbf{x}_{it}']$). Then $\beta = \Omega_{xx}^{-1}\Omega_{xy}$ denotes the long-run average regression coefficient associated with the long-run average covariance matrix of $[y_{it} \ \mathbf{x}_{it}']$.

Given T observations on n panels, Phillips and Moon show that under the null of homogeneous panel cointegration the pooled OLS estimator

$$(9)\ \mathbf{b}_{n,T} = \sum_{i=1}^{n} \sum_{t=1}^{T} [\mathbf{x}_{it} \ \mathbf{x}_{it}']^{-1}[\mathbf{x}_{it} \ y_{it}]$$

is a $\sqrt{n}$ consistent estimate of β with a limiting normal distribution when $n/T \to 0$, and $[1 \ -\beta]$ is the cointegrating vector. Phillips and Moon also show that a Fully Modified estimator of the pooled relationship, which includes the CCR estimator, is a $\sqrt{nT}$ consistent estimator of β with a limiting normal distribution providing $n/T \to 0$.

With homogenous panels the results discussed above for single equation cointegration carry over to pooled estimation. In particular, with homogeneous panels the pooled estimator (9) will converge to a stable long-run relationship when the specification is cointegrated; if the panel

[10] With cointegrated data, the path of adjustment is obtained from estimates of the error correction specification given in (8). For a further discussion of assumptions underlying short-run and long- run adjustments in dynamic heterogenous panel models see Pesaran, Shin, and Smith (1999).

specifications are spurious the pooled estimator diverges. Tests of cointegration, therefore, have the same interpretation with the pooled estimator as with the single equation estimator. This means that either the residual based tests or the variable addition test for cointegration discussed above could be used to evaluate the completeness of the model specification in the case of a homogenous panel.[11]

[11] The results of Phillips and Moon are much more powerful than are required for this paper. In particular they show that in the case of heterogeneous panels stable average long-run relationships can be estimated whether or not the individual panels are cointegrated or spurious. This means that panel methods allow for the estimation of long-run relationship even in cases where considerations of the time dimension alone would lead to the regression being characterized as spurious. To see this result write the heterogeneous panel as $y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta}_i + u_{it}$ with $\boldsymbol{\beta}_i$ random, and rewrite this equation as $y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + u_{it}^*$ where $u_{it}^* = \mathbf{x}_{it}'(\boldsymbol{\beta}_i - \boldsymbol{\beta}) + u_{it}$. The pooled estimator (7) provide a consistent estimator of the parameter vector $\boldsymbol{\beta}$ even though u_{it} (for all $i = 1,\ldots,n$) is drawn from a unit-root process because in this case the linear combination of integrated variables $\mathbf{x}_{it}'(\boldsymbol{\beta}_i - \boldsymbol{\beta})$ offsets a diverging variance of $\{u_{it}\}$. The result is an error process (u_{it}^*) with converging second moments. The estimator $\mathbf{b}_{n,T}$ represents a consistent estimate of $\boldsymbol{\beta}$ in this case because averaging over independent panels decreases the second moments of $\{u_{it}^*\}$ relative to the second moments of the integrated vector $\{\mathbf{x}_{it}\}$, although the speed of convergence of $\mathbf{b}_{n,T}$ will, in general, be slower than in the homogenous panel case. The pooled estimator in this case is defined by the long-run average variance matrix of the panel. In general, it is not equal to the average of the cointergration coefficients.