4. Analysis of Annual State-Level Panel Caseload Data, 1980-1999

4.1 Evidence of Unit Root Non-stationarity From Previous Studies

Variables characterized by unit root non-stationarity exhibit a high degree of persistence. This is due to the fact that shocks to unit root processes do not die out over time.[14] The necessity of including a large number of lags in any dynamic model specification to account for this persistence is an indication of unit root non-stationarity or a root in the lag structure close to one. Both long lags of the FSP caseload variable and a high degree of persistence are evident in the preferred dynamic specification of the FSP caseload equation reported by FGZ and ZGF (2001,2003). In these specifications, four years, or one-fifth of the sample, are required to account for dynamic feedback. ZGF (2001,2003) report choosing a lag length based on the Schwartz information criterion. This criterion, however, tends to find shorter lag lengths than other methods (Ng and Perron, 1995) indicating that even longer lag lengths might be appropriate.

To illustrate the persistence implied by the lag structure reported in these studies we simulated the impulse response function of a one time shock implied by the coefficients of the lagged FSP caseload variables. The impulse response function is illustrated in figure 1 for the estimated lag structure given in table 1 of ZGF (2001). This lag structure implies that it takes 10 years for 90% of the impact of a shock to be realized but that the shock does not completely die out for approximately 30 years.[15]

As is common with models estimated with long dynamic lags, the impulse response function calculated from the estimated lag structure implies a complicated pattern of adjustment that appears difficult to interpret. In figure 1, the impulse response function of FSP caseloads indicates two cycles, one of 3 years and another of 12 years with an initially positive response that becomes negative after 5 years. The weights then cycle around zero until finally dampening out.

The simulation of the impulse response function implied by the lag structure estimated by ZGF (2001, 2003) shows clear evidence of a high degree of persistence in the FSP caseload data. The advantage of modeling the data assuming unit root non-stationarity in this case is that the long-run persistence in the data is embedded directly into the model specification. *Even if* the data were, in fact, stationary, but with a root close to but less than unity, assuming unit root non-stationarity would provides a more parsimonious method of incorporating persistence compared to stationary models requiring long lags.[16] Modeling the data as if it were unit root non-stationary, in this case, would adhere to the Box-Jenkins' principle of "parsimonious parameterization". Parsimony is valued as a criterion for model selection since an excessive number of parameters often leads to multicollinearity, instability of parameter estimates, a loss of

---

[14] See the discussion around (7).

[15] ZGF (2001, p. 9, footnote 7) note that their estimates implies that ". . .it takes about a decade for a shock to completely filter through the system"

[16] Clark and Spriggs (1992) provide an example in which unit root and stationary models with roots are close to unity have similar predictions, at least for the first few years.

degrees of freedom and hence a loss in the precision of estimation (Fuss, McFadden, and Mundlak, 1978, p.224).

Studies that estimate static FSP caseload equations (i.e., those that do not include lagged FSP caseload variables) such as, for example Wallace and Blank, do not report the combination of regression statistics that usually indicate the existence of unit root non-stationarity. For a single equation model that does not include dynamics, evidence of unit root non-stationary typically would be indicated by highly significant t-values of the estimated coefficients, high R^2, and low Durbin-Watson statistics (Granger and Newbold, 1977). In the context of panel data, however, it is unclear how statistics based on estimated model residuals, such as R^2 or the Durbin-Watson statistic ought to be interpreted.

4.2 Unit Root Tests for Panel Data

Table 2 reports test results of the null hypothesis that unit-root time series generate the FSP caseload-related data. Tests for each state individually would have low power because each test would be based on only 20 observations. The tests in table 2 treat the data as generated from a panel of 50 states with 20 observations per state, and hence are panel unit root tests (Im, Pesaran, and Shin, 1997).

Two types of tests are reported. One type, denoted as ‘no trend’, is a test of the null that the sample data is generated by a time series with a single unit root. The second type, denoted as ‘trend’, tests the null that the data are generated by a series with a single unit root and a time trend. Observed test statistics that are less than the corresponding critical value (reported in a footnote) reject the null hypothesis of a unit root.

Whether or not a trend is included, tests in table 2 fail to reject the null that both the log of per capita FSP caseloads and log of per capita AFDC/TANF caseloads are integrated. Unit root tests for the employment growth and the unemployment rate appear to depend, however, on whether a trend is included in the test regression. In particular, the tests fail to reject that the employment growth and the unemployment rate were generated by integrated processes with no trend, but do reject (again, at conventional levels) the null that these variables were generated by integrated processes with a trend.

Visual inspection of the plots of aggregate employment growth and the unemployment rate indicate that the appropriate test regressions may be the ones without trend. The employment growth series appears to exhibit no trend over the sample period, and the unemployment rate exhibits no strong evidence of a trend. Under the alternative hypothesis that these series are stationary (around a non-zero mean) the appropriate test regression for these variables would exclude a time trend (Hamilton, pp. 501-3). These tests in table 2 do not reject the null hypothesis that employment growth and the unemployment rate are integrated.

The results from table 2, and the indication of persistence in the data, provide evidence that the FSP caseload data behave as if they were generated by unit-root processes. It is not, of course, possible to determine definitively whether or not these data were generated by a nonstationary processes. Based on our examination of these date we assume that state level time-series of the log of per capita FSP caseloads, the log of per capita AFDC/TANF caseloads, the unemployment rate, and employment growth are described by unit-root processes.

4.3 Tests of Cointegration

Previous estimates of the FSP caseload equation using annual state level panel data have assumed a homogenous panel structure in which response coefficients are the same for each state (except for the effect of state fixed effects or, if present, state specific time trends). We make the same assumption in this paper. The difference is that we assume that there exists the same cointegrating relationship for all the states. In a subsequent section, we provide evidence that supports this assumption.

Our estimation procedure is based on the results of Phillips and Moon for the case in which the number of cross-sections (n) and time series observations (T) are such that $n/T \rightarrow 0$. In our analysis this assumption is satisfied since we will assume that the number of states is fixed and the panel grows over time.

Under this condition, Phillips and Moon show that the average panel can be estimated using a fully-modified panel estimator if an individual homogeneous (or near homogeneous) cointegrating relationship exist for the cross sections. In this study, estimates are obtained using Park’s (1992) canonical cointegrating regression (CCR) estimator. This estimator is a version of Phillips and Hansen’s fully-modified (FM) estimator. Tests for cointegration are conducted using Park’s (1990) variable addition test. This testing framework assumes that there exists a long-run relationship (a null of cointegration) and looks for evidence to refute this hypothesis.

The empirical analysis is based on the following specification of the FSP caseload equation (*10*) that removes the state level fixed effects.

$$(10a) \quad \mathrm{FSP}_{it} - \mathrm{FSP}_{i.} = \mu + \beta' T_t + \alpha' IP_{it} + \theta' (E_{it} - E_{i.}) + \gamma (\mathrm{AFDC}_{it} - \mathrm{AFDC}_{i.}) + \varepsilon_{it}$$

The state-level economic variables (E) consist of a 2 element column vector with the unemployment rate (UM) and employment growth (EMP) measured as deviations from their time means, $(E_{it} - E_{i.})' = [\mathrm{UM}_{it} - \mathrm{UM}_{i.} \;\; \mathrm{EMP}_{it} - \mathrm{EMP}_{i.}]$. The time means of the variables are denoted as $\mathrm{FSP}_{i.} = T^{-1} \Sigma_t \mathrm{FSP}_{it}$, $\mathrm{AFDC}_{i.} = T^{-1} \Sigma_t \mathrm{AFDC}_{it}$, $\mathrm{UM}_{i.} = T^{-1} \Sigma_t \mathrm{UM}_{it}$, and $\mathrm{EMP}_{i.} = T^{-1} \Sigma_t \mathrm{EMP}_{it}$, respectively.

The economic variables and the log of per capita ADFC/TANF caseloads are the unit root regressors. The deterministic regressors consist of the trend variables (T), which include both linear and quadratic trend terms, a common intercept (μ), and the intervention policy variables (IP). The intervention policy variables include the dummy variables defined by the set of explicit changes in FSP provisions given in table 1 and the any waiver dummy variable. We will refer to (*10a*) as the full model. Conceptually, we hypothesize that any possible cointergrated relationship exists between the caseloads of these program and the economic variables. The policy variables are viewed as intervening into this long-run equilibrium relationship. Park and Phillips (1988) have shown that stationary regressors such as the policy and deterministic variables are asymptotically orthogonal to any integrated variables in a cointegrating regression.

Hence, asymptotically, the inclusion of the policy variables does not affect the asymptotic results. These policy variables are included in the model specification to better account for shocks to the system generated by changes in policy, which should more accurately model short-term fluctuations in the system during this sample period.

Since previous studies have generally estimated the FSP caseload equation without including AFDC/TANF caseloads as a regressor, we begin the evaluation of the specification of this equation by first addressing the question of whether a FSP caseload equation that consists of just economic and policy variables can be considered correctly specified. The null hypothesis is that a FSP caseload equation given in (10a) that *excludes* the log of per capita AFDC/TANF caseloads forms a long-run equilibrium relationship. We test this hypothesis using two versions of Park's variable addition test for cointegration.

In the first version, third, fourth, and fifth powers of time are used as superfluous variables. The first row of table 3 reports the results of these variable addition tests for the specification of the FSP caseload equation (*10a*) that includes only the economic, policy and deterministic variables. [17]

The results suggest that the FSP/economy/policy-only equation may be misspecified. At the 10% level the null hypothesis of cointegration is rejected. This provides evidence that the model residuals act like an integrated process so that the economy and dummy policy variables, by themselves, are not sufficient to define a long-run equilibrium relationship with FSP caseloads.[18] These test results indicate, therefore, that the estimated relationship leaves out AFDC/TANF caseloads may not be stable and may not provide consistent estimates of the relationship between FSP caseloads and the economy and policy.

In the second version of Park's variable addition test the log of per capita AFDC/TANF caseloads is added to the FSP/economy/policy-only equation as a superfluous variable. Coefficient estimates of the full model (*10a*) given in the third column of table 4 indicate that the impact of AFDC/TANF caseloads is highly significant with an observed t-value that exceeds ten. As a superfluous variable, the significance of the log of per capita AFDC/TANF caseloads is strong evidence against the cointegration of a FSP/economy/policy-only caseload specification.

The evidence from both versions of the variable addition test supports a conclusion that the FSP/economy/policy-only regression is not cointegrated. Next, we test the specification of the full model (*10a*) that includes the impact of the log of per capita AFDC/TANF caseloads for cointegration. Cointegration tests of the full specification of (8a) using powers of time as superfluous variables are reported in the second and third rows of table 3. In the second row of table 3 test results are reported when trend variables (T) are included. The large observed probability values for each of the variable addition tests indicate support for the null hypothesis

---

[17] The tests consist of adding these superfluous variables to the regression, beginning with the cubic time trend, and testing the joint hypothesis that the coefficients of the superfluous variables are all equal to zero.

[18] More precisely, the variables used to measure economic activity do not appear to be cointegrated with FSP caseloads. This does not exclude the possibility that FSP caseloads, the economy and intervention policy variables could be cointegrated with different measures of economic activity.

that the full model is stochastically cointegrated.[19] Removal of the trend variables allows for a test of deterministic cointegration (see the discussion on page 12). The results of this test are presented in the third row of table 3 and indicate that deterministic cointegration is not rejected at conventional level in the full model.[20, 21]

The finding that the full model specification is deterministically cointegrated means that (*10a*) is correctly specified without including the trend variables (*T*). Any trend in the log of per capita FSP caseloads is cancelled by a linear combination of trends in the regressors. The implication of this finding is that the common practice of including year effect and/or state specific time trends is not recommended for these data. Including these trends appears to overcontrol for variation that can be explained by variation in the regressors. Take, for example, the estimate of the importance of policy and economic variables in explaining the FSP caseload reduction during 1994-99. For the specification that most closely corresponds to the cointegrated FSP caseload equation we estimate, results reported by ZGF (2001, table 3, column 2) imply that the deterministic and residual components represent about 34 percent of the actual change in FSP caseloads between during these years. In our case, (table 5) the net effect of the deterministic and residual components represent only about 4 percent of the actual change in FSP caseloads for this period. This comparison suggests that the admonition by Wallace and Blank and by Schoeni concerning over-differing with time trends is important (see the discussion on page 12).[22]

4.4 Coefficient Estimates

Table 4 presents the parameter estimates of the full model specification given in (*10a*). Even though OLS estimates of the panel cointegrating relationship are consistent, non-zero correlation between the stationary component of the stochastic regressors and the error term will cause the OLS estimates to exhibit a finite sample bias that invalidates the usual asymptotic t-test of parameter significance. The CCR estimator transforms the variables of the regression so that these problems are asymptotically corrected. Hence, test statistics presented along with the CCR estimates are large sample test.[23]

A comparison of the parameter point estimates obtained from the OLS and CCR estimators in table 4 illustrates the implications of endogeneity caused by correlation between the model errors and the explanatory variables. The effect of this correlation on tests of significant is illustrated by the estimates of Omnibus Budget Reconciliation Acts in 1990 and 1993, and the

---

[19] The null hypothesis of cointegration between log of AFDC/TANF caseloads, trend, the intervention policy and economic variables was rejected at the 5% level of significance using these superfluous variables.
[20] Deterministic cointegration is discussed on page 16.
[21] The fact that we find cointegration using the same measures of the economy as have been used in previous studies suggests that the misspecification of FSP caseload equation estimated in these studies results because the variables included to measure the impact of policy do not sufficiently capture the range of policy changes that occurred during the sample period.
[22] Currie and Grogger report specifications with and without state-specific time trends (table 2). The inclusion of state-specific time trends reduced the magnitudes of the unemployment rate coefficient and a AFDC waiver variable, however, from their results it is not possible to calculate how the inclusion of these trends affects the proportion of the change in FSP caseloads due to the deterministic components.
[23] The results of Phillips and Moon implies that the CCR estimator converges to this distribution at the rate $\sqrt{n}T$.

estimate of the intercept term. In each case, OLS estimation indicates that the impact of the variable is statistically positive at conventional levels but is statistically insignificant with CCR.

The adjustment of OLS estimates for correlation between the first-difference in the explanatory variables and the model error and any serial correlation by CCR also causes the sign of some parameter estimates to change; however, this mostly involves coefficients that are not statistically significant at conventional levels. The notable except is the impact of employment growth which changes from a statistically positive effect with OLS to a marginally significant negative effect with CCR. This switch in sign is similar to the effect observed by ZGF (2001, p. 18) when they proceeded from a "static" model to a "dynamic" model.

Results in column 4 indicate that the elasticity of FSP caseloads with respect to AFDC/TANF caseload, given by the estimated coefficient, is positive as expected and very precisely estimated. Column 3 of table 4 also reports CCR estimates for the FSP caseload equation that does not include the AFDC/TANF caseload variable. Our tests of conintegration indicate that meaningful inference can not be conduced with this equation; however, it is interesting to observe what effect the exclusion of AFDC/TANF caseloads has on the estimated effects of the economic variables. When AFDC/TANF caseloads are excluded both economic variables have a greater absolute affect on FSP caseloads. This likely reflects the fact that AFDC/TANF caseloads are themselves a function of the economy. This effect is discussed further in Section 6 when estimates of the relative effect of policy versus economy are compared.

4.5 Variable Contribution

Previous studies of the FSP caseload equation have reported on the relative importance of the economy versus policy by comparing the percentage of the actual change in FSP caseloads predicted from changes in each type of variable around the time of PRWORA. Following these studies, measures are also calculated using our estimate of the FSP caseload equation for the time interval 1994-98. A comparison of these measures with those from other studies is postponed, however, until section 6. In this section, measures of the importance of the various regressor variables in explaining variations in FSP caseloads are presented for the *entire* sample period 1980-99.

Variable performance is evaluated based on two comparisons. The contribution of each type of variable used in (*10a*) is evaluated based on its contribution to the statistically fit measured by incremental R^2, and a graphical analysis, using a sample of 5 states and the aggregate US, that illustrates how well each variable type tracks movements in FSP caseloads over the sample period. The CCR estimates of the cointegrated caseload equation given in column 3 of table 4 are used in the comparisons.

Table 6 reports values of incremental R^2 for each type of variable used in (8a).[24] The incremental R^2 of a variable provides a measure of its importance in the estimated equation (Theil, 1971, p.168). A given value of the incremental R^2 measures the addition to (overall) R^2 when the variable is added to the regression. Incremental R^2 is a conditional measure that

[24] Incremental R^2 values were calculated using the transformed data from the CCR estimator.

depends on the order in which the variables are added to the equation. That is, the value of the incremental R^2 of a variable depends upon the variables already included in the regression.

Since we are interested in assessing the impact of adding AFDC/TANF caseloads to the FSP/economy-only specification, the incremental R^2 associated with the log per capita AFDC/TANF caseloads variable is calculated after accounting for the other determinants in the regression. The results in table 6 illustrate that the impact of this variable is quite large, roughly five times more important than the economic variables in achieving the overall fit of the FSP caseload equation. This result suggests the importance of variations in AFDC/TANF caseloads in explaining movements in FSP caseloads extends to the entire sample period 1980-99 and is not limited to the particular subset of years surrounding the passage of welfare reform legislation.

To illustrate how the estimated cointegrated specification tracks the time path of FSP caseloads over the sample period, five states consisting of California, Florida, Illinois, Texas, and Wisconsin plus the aggregate US were chosen. Each of these states had relatively large FSP caseload and as a group exhibited a diverse caseload pattern over the period. The model's ability to track the path of FSP caseloads in these states provides evidence of the overall fit of the estimated cointegrated FSP caseload equation.

Time plots of the deviations in the log of per capita FSP caseloads from their sample mean for the five individual states and the US are given in figures 2. Since the end of the 1980's FSP caseloads for each of the states have generally followed the overall US pattern, however, prior to that time the states exhibited much greater diversity in their FSP caseloads. During the decade of the 1980's, both California and Florida exhibited generally declining FSP caseloads until the run-up at the end of the decade (similar to the US aggregate). The other states, however, exhibited an increase in FSP caseloads through the first part of this decade, either to a plateau and then a decline (Illinois and Wisconsin), or simply a general increase throughout (Texas). At the beginning of the decade of the 1990's, FSP caseloads were increasing for all the states and peaked at their highest levels around 1993-94 before declining throughout the second half of the 1990's.[25] The experience in Wisconsin differed somewhat from this pattern. FSP caseloads in Wisconsin did not start to increase until 1991-92 and then peaked in 1993 at a level below their peak level in the 1980's.

In figure 3 the plot of the mean deviations of the log of per capita FSP caseloads predicted by the full model (*10a*) using the CCR estimates is graphed against the actual FSP caseload data for each state and the aggregate US. These plots indicate that the model does a reasonably good job of tracking the diverse patterns of FSP caseloads illustrated by this sample of individual states. The overall performance of the specification (*10a*) in tracking FSP caseloads in these states appears to support the decision to model the state level data as a homogeneous (or near homogeneous) panel with the same cointegrating vector for each state.

---

[25] The peak in per capital FSP caseloads in the 1980's (1981) in Florida was only slightly less than the peak in the 1990's (1993).

Additional insight into the performance of (*10a*) can be gained by illustrating the contribution of each variable type in explaining movements in FSP caseloads. In these illustrations a 'hat' denotes an estimated CCR coefficient.

In figure 4 the mean deviations in the log of per capita FSP caseloads predicted by per capita AFDC/TANF caseloads [$\hat{\gamma}(\mathrm{AFDC}_{it} - \mathrm{AFDC}_{i.})$] is graphed for each state along with the deviations in the log of per capita FSP caseloads, $\mathrm{FSP}_{it} - \mathrm{FSP}_{i.}$. The plots indicate that per capita AFDC/TANF caseloads track the general overall pattern of per capita food stamp caseloads fairly well. The general fluctuations of per capita food stamp caseloads are fairly consistently tracked by concomitant rise and fall predicted by per capita AFDC/TANF caseloads.

Next, the portion of the deviations in the log of per capita FSP caseloads predicted by AFDC/TANF caseloads is subtracted from the deviations in the log of per capita FSP caseload, [$(\mathrm{FSP}_{it} - \mathrm{FSP}_{i.}) - \hat{\gamma}(\mathrm{AFDC}_{it} - \mathrm{AFDC}_{i.})$]. In figure 5 the resulting series is graphed along with the log of per capita FSP caseloads predicted by deviations in the economic variables, $\hat{\vartheta}^{/}(E_{it} - E_{i.})$. Once the influence of per capita AFDC/TANF caseloads has been taken out, the remaining portion of the deviations in the log per capita FSP caseloads is much more variable. However, the predictions obtained from unemployment rate and employment growth variables do a credible job in tracking the fluctuations unexplained by variations in AFDC/TANF caseloads.

Finally, both the influence predicted by the AFDC/TANF caseloads and the economic variables are removed from deviations in the log of per capita FSP caseload variable, $(\mathrm{FSP}_{it} - \mathrm{FSP}_{i.}) - \hat{\gamma}(\mathrm{AFDC}_{it} - \mathrm{AFDC}_{i.}) - \hat{\vartheta}^{/}(E_{it} - E_{i.})$. In figure 6 these unexplained variations are graphed along with the deviations in the log of per capita FSP caseloads predicted by the intervention policy variables, $\hat{\alpha}^{/} IP_{it}$. These plots illustrate that these intervention policy variables do not tract the remaining portion of log per capita FSP caseloads in the individual states well. Interestingly, the policy dummy variables do a better job tracking the residual log per capita FSP caseloads for the US than for any of the individual states. This may indicate that these policy dummies, in the aggregate, are picking up differences in the proportion of states that implement policy changes during a given year.