# LABOR MARKET BEHAVIOR IN WASHINGTON: A COINTEGRATION APPROACH 

JunHo Yeo, Ph. D. (jyeo@ksu.edu)<br>Director of Research<br>Office of Local Government<br>Department of Agricultural Economics<br>Kansas State University<br>10E Umberger Hall<br>Manhattan, KS 66506-3415<br>785-532-3095, Fax: 785-532-3093<br>Sung K. Ahn, Ph.D.<br>Professor of Statistics<br>Department of Management and Decision Sciences<br>Washington State University<br>Pullman, WA 99164-4736<br>David W. Holland, Ph. D.<br>Professor of Agricultural Economics<br>Washington State University<br>Pullman, WA 99164-6210

This paper was submitted to the American Journal of Agricultural Economics (AJAE) on May 8, 2001 and was prepared for presentation at the 2001 meeting of the American Agricultural Economics Association in Hilton Chicago Hotel.

Copyright 2001 by [JunHo Yeo, Sung Ahn, and David Holland]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

## LABOR MARKET BEHAVIOR IN WASHINGTON: A COINTEGRATION

## APPROACH

Generally, it is well understood that new business investment brings changes in population, increase in labor force participation rate, and migration of new residents. However, Powers (1996) argued that natural capital in the many places in the West is driving population growth and that drives job growth. Powers believes that the causality operates in a reverse way in regions having good environment and amenities, such that population is attracted by environmental amenities and the population changes brings about changes in employment. Washington is a good test of the Powers hypothesis, because of its beautiful environment and amenities.

It is important to both local policymakers and social scientists to understand who benefits from local job growth. There is mixed research results regarding the extent that new migrants tend to account for new employment. Bartik (1993) found that about onequarter of the new jobs go to local workers because of the increase in the labor force participation rates of local residents in the long run. He considered the long run effects by estimating the effects of $1 \%$ job growth in a certain period on the labor force participation rate seventeen years after that period. In contrast Blanchard and Katz's (1992) research reached a different conclusion - in five to seven years the employment response consists entirely of the migration of new migrants. Their finding is that longrun effect of the job growth on the labor force participation rate is negligible. Yeo and Holland (2000) found a composite result. Their finding is that most of the new jobs are captured by in-migrants instead of the county residents in the long run. Also, the long-
run effect of the job growth on the labor force participation remains to some degree, although its effect is small.

In many studies, the population, employment, and labor force participation rate are considered as important variables to explain the local labor market. In this study, we first examine causality arguments above. Second, we investigate interactions among these variables in Washington using cointegration analysis. Third, we decompose each of these series into a stationary component and a non-stationary component, and identify these components. Forth, by investigating the effects of one standard deviation shock to the employment on the population and the labor force participation rate by impulse response analysis, we provide the results of this study - the long run effect of employment on the local labor force participation rate and the party who benefits from local job growth - for comparison with the results of previous studies.

## Cointegration Analysis and Error Correction Representation

Cointegration analysis allows us to examine the long run equilibrium relationship among nonstationary variables. In our case, cointegration analysis allows us to investigate the long run effects of the employment and participation rate on the population. The concept of cointegration was developed by Engle and Granger (1987). A time series $Z_{t}$ that is stationary after being differenced $d$ times is said to be integrated of order $\mathrm{d}, I(d)$. For an $m$-dimensional nonstationary process $Z_{t}$ which is $I(d)$, if there are $r$ linearly independent vectors $\beta_{i}$ such that $\beta_{i}^{\prime} Z_{t}$ is $I(b), b<d$, then $Z_{t}$ is said to be cointegrated of order $(d, d-b)$ denoted by $Z_{t} \sim C I(d, d-b)$ with cointegrating rank $r<$
$m$. The $r$ vectors, $\beta_{i}$, are called cointegrating vectors, and the stationary linear combination $\beta_{i}^{\prime} Z_{t}$ is called the long-rum equilibrium error. Engle and Granger (1987) defined cointegration as "If each element of a vector of time series $Z_{t}$ is stationary only after differencing, but a linear combination $\beta_{i}^{\prime} Z_{t}$ needs not be differenced, the time series $Z_{t}$ have been defined to be co-integrated of order $(1,1)$ with cointegrating vector $\beta_{i}$ ". For this study we focus on the case where $Z_{t}$ is $I(1)$, and thus $\beta_{i}^{\prime} Z_{t}$ is $I(0)$, stationary because the variables considered in this study are $I(1)$.

We consider an $m$-dimensional vector autoregressive process of order $p, \operatorname{VAR}(p)$,
(1) $Z_{t}=\delta+\sum_{i=1}^{p} \Phi_{i} Z_{t-i}+\varepsilon_{t}$,
where $\delta$ is an $m \times 1$ vector of constant term, $\Phi_{i}$ is an $m \times m$ matrix of parameters, and $\varepsilon_{t}$ is a white noise with positive definite covariance matrix $\Omega$.

As in Engle and Granger (1987), the model in (1) can be re-expressed in the following error correction model (ECM) representation,
(2) $W_{t}=\delta+C Z_{t-1}+\sum_{i=1}^{p-1} \Phi_{i}^{*} W_{t-i}+\varepsilon_{t}$,
where $W_{t}=Z_{t}-Z_{t-1}$ is stationary, $\Phi_{i}^{*}=-\sum_{k=i+1}^{p} \Phi_{k}, C=-I_{m}+\sum_{i=1}^{p} \Phi_{i}$. The term $C Z_{t-1}$ is called the error correction term, and the coefficient matrix $C$ contains information about the long-run equilibrium relationship among the components of $Z_{t}$. If there is no
stationary long-run equilibrium relationship among the variables in $Z_{t}$, the rank of $C$ is zero, and we use the standard VAR model of order $p-1$ for the first differenced series $W_{t}$.

If the vector process $Z_{t}$ is stationary, then $C$ is a full rank matrix. The rank of $C$ is greater than zero and less than $m$, i.e., $0<r=\operatorname{rank}(C)<m$, if $Z_{t}$ is cointegrated of order $(1,1)$ with cointegrating rank $r$, which means that there exists a long-run equilibrium relationship among the components of $Z_{t}$.

To estimate ECM in (2) with the rank restriction, we reparameterize the $C$ matrix as $C=\alpha \beta^{\prime}$, where $\alpha$ and $\beta$ are full rank $m \times r$ parameter matrices or rank $r$. Then we can rewrite (2) as
(3) $W_{t}=\delta+\alpha \beta^{\prime} Z_{t-1}+\sum_{i=1}^{p-1} \Phi_{i}^{*} W_{t-i}+\varepsilon_{t}$.

It is shown, for example in Ahn and Reinsel (1990), that $\beta_{i}^{\prime} Z_{t}$ is a stationary cointegrating combination, where $\beta_{i}$ is the $i^{\text {th }}$ column of $\beta$. Johansen and Juselius (1990) interpreted the matrix $\alpha$ as the average speed of adjustment towards the estimated equilibrium state such that a low coefficient indicates slow adjustment and a high coefficient indicates rapid adjustment.

## Decomposition of Time Series

Cointegrated time series can be decomposed into non-stationary components and stationary components. We decompose $Z_{t}$ as in Kasa (1992) by rewriting

$$
\begin{align*}
Z_{t} & =\beta\left(\beta^{\prime} \beta\right)^{-1} \beta^{\prime} Z_{t}+\left(I-\beta\left(\beta^{\prime} \beta\right)^{-1} \beta^{\prime}\right) Z_{t}  \tag{4}\\
& =\beta\left(\beta^{\prime} \beta\right)^{-1} \beta^{\prime} Z_{t}+\beta_{\perp}\left(\beta_{\perp}^{\prime} \beta_{\perp}\right)^{-1} \beta_{\perp}^{\prime} Z_{t}
\end{align*}
$$

where $m \times(m-r) \beta_{\perp}$ is such that $\beta_{\perp}^{\prime} \beta=0, \beta^{\prime} Z_{t}$ is a cointegrating combination, that is, a stationary factor and $\beta\left(\beta^{\prime} \beta\right)^{-1}$ is the factor loading matrix of the stationary factor, and $\beta_{\perp}^{\prime} Z_{t}$ is a common trend, or non-stationary factor and $\beta_{\perp}\left(\beta_{\perp}^{\prime} \beta_{\perp}\right)^{-1}$ is the factor loading matrix of the common trend.

Beveridge and Nelson (1981) showed that an integrated process can be represented in terms of non-stationary components and stationary components, in which the nonstationary component is a random walk with drift and the stationary component is covariance stationary. They interpreted the non-stationary component as the long run forecast of the series adjusted for its mean rate of change and the stationary component as a business cycle. Stock and Watson (1988) showed that cointegrated multiple time series share at least one common trend. In summary, the stationary or transitory component can be interpreted as temporary business cycle, whereas interpretation on non-stationary or permanent component is the long run forecast profile as a random walk with drift.

## Impulse Response Analysis

Impulse response analysis of vector autoregressive systems is a useful tool to examine the interrelationships among the variables in dynamic models (Lütkepohl and Reimers, 1992). From this analysis, we investigate the effects of one standard deviation shock of one variable to the other variables. In a vector autoregressive systems, when we
find an equilibrium relationship among the variables at some period, $t$, any exogenous shock to a variable leads to a new long-run equilibrium provided no further shocks occur.

For impulse response analysis, we first consider the $m$-dimensional vector autoregressive model and rewrite it as moving average (MA) representation

$$
\begin{equation*}
Z_{t}=\mu+\sum_{j=0}^{\infty} \Psi_{j} \varepsilon_{t-j} \tag{5}
\end{equation*}
$$

where $\Psi_{0}=I_{m}$ and $\Psi_{j}=\sum_{i=1}^{j} \Psi_{j-i} \Phi_{i}, j=1,2, \cdots, \infty$ with $\Phi_{i}=0$ for $i>p$. The elements of the $\Psi_{j}$ represent the impulse response of the system.

We examine the orthogonalized impulse responses of the system and the errors are orthogonalized by Cholesky decomposition so that the covariance matrix of the resulting innovations is diagonal.
(6) $\quad Z_{t}=\mu+\sum_{j=0}^{\infty} \Theta_{j} e_{t-j}$
where $\Theta_{j}=\Psi_{j} P, e_{t}=P^{-1} \varepsilon_{t}, E\left(e_{t} e_{t}^{\prime}\right)=I_{m}, P P^{\prime}=\Omega$ and $P$ is assumed to be a lower triangular matrix with positive diagonal elements (Lütkepohl, 1990). The elements of the $\Theta_{j}$ are impulse responses and Lütkepohl and Reimers (1992) explain that a one time impulse may have a permanent effect on the dynamic system, which will lead to a new long-run equilibrium.

## Data

To date there have been few studies dealing with the cointegration analysis approach in regional science. To estimate the effects of job growth on the labor market, most studies use cross-sectional data, or lagged dependent variables as variables on the right hand side. The most common model regresses the changes in the population of a fixed year on the change in employment, labor force participation rates, or net migration rate: See Greenwood and Hunt (1984) and Summers (1986). Some models with the population as the dependent variable use the level of the employment / population ratio or employment as an independent variable. Bartik (1992) argued that models that use the levels variables might be biased by unobserved fixed effects of local areas. However, models that use the changes of variables also have a weak point, that is, they cannot predict variables in levels. To overcome these deficiencies, we use the cointegration analysis that allows us to estimate the long run equilibrium relationship in levels.

In this study we use three variables, population, labor force participation rate for population aged 18 and 64 and employment to examine the long run equilibrium relationship among these variables. The long run equilibrium relationship equation gives a unit change interpretation as in a general linear regression model. Our data are from the Office Forecast Council (OFC) in the state of Washington. We calculated the labor force participation rate by dividing civilian labor force by population aged 18 to 64 . The scale of population and employment is 1000 people and that of labor force participation rate is percentage. The data series are observed at two different frequencies. The labor force participation rate and the employment are quarterly data for the period between 1969 and
1993. However, only annual data are available for the population variable because population is surveyed only on the second quarter of every year. Because the methods for cointegration analysis are applicable only to data with the same frequency (as far as we are aware of), we have two options to make the data frequencies the same. One is to use the second quarter data of the participation rate and employment in order to match the sampling frequency of the population data and these annual data are exhibited in Figure 1. The other is to estimate the quarterly values of population series using a similar interpolation method to Chow and Lin (1971).

We first focus on the annual data series, although the sample size is small. One may argue that the data over 25 years may not be long enough for a study of long-run equilibrium. However we could at least estimate the model with a long-run equilibrium restriction through model (3) and gain insight into the long-run behavior. Next we analyze the quarterly data using estimated quarterly population data and compare the two results.

## Application of Cointegration Analysis to Annual Data

Let the vector $Z_{t}$ consist of three variables, such that $Z_{t}=\left(p_{t}, r_{t}, e_{t}\right)^{\prime}$, where $p_{t}$ is the number of population, $r_{t}$ is the labor force participation rate, and $e_{t}$ is the employment at period $t$. As the data series in Figure 1 exhibit non-stationary behavior attributable to a unit root, we perform Dickey-Fuller $(1979,1981)$ test for unit root using the test statistics $\hat{\tau}_{\tau}$. The results summarized in Table 1 show that each of the three series is $I(1)$. That is, all three original series have a unit root and their first differenced series do not have a unit root.

## The Choice of AR Order and Cointegrating Rank of $Z_{t}$

We now investigate if there is long-run equilibrium information among the components of $Z_{t}$, that is, if $Z_{t}$ is cointegrated. To this end we consider a VAR model for $Z_{t}$ as in (1). We examine an appropriate AR order based on the partial canonical correlation between $W_{t}$ and $W_{t-k}$ adjusted for $Z_{t-1}, W_{t-1}, \cdots, W_{t-k+1}$ (Ahn and Reinsel, 1990) and the Akaike Information Criterion (AIC). For a VAR $(p)$ the partial canonical correlations between $W_{t}$ and $W_{t-k}$ are all zero, and thus $\Phi_{k}^{*}=0$ for $k \geq p$. The results from partial canonical correlation analysis (PCCA) are summarized in Table 2 and Table 3. Table 2 indicates that the coefficient matrix $\Phi_{i}^{*}$ is significant until $\Phi_{4}^{*}$. This means that the vector $Z_{t}$ at a certain period is affected by the past five years history of the $Z_{t}$. Considering 25 observations of data, the VAR (5) model seems to be overfitted. Furthermore, the p-values (i.e. observed significance levels) are based on the large sample distribution. In contrast, as is Table 4, the minimum AIC is attained at lag 1, which in turn favors a VAR (1) of $Z_{t}$. For this reason, we consider tentatively an appropriate AR order to be between one and four in our model fitting.

We also need to determine cointegrating rank $r$ through the rank of $C$ matrix. Two test statistics, the trace statistic and the maximal eigenvalue statistic are used to determine the rank of $C$ matrix. The Likelihood Ratio (LR) test statistic,
$\Lambda=-n \sum_{j=r+1}^{m} \ln \left(1-\hat{\rho}_{j}^{2}\right)$ is our trace statistic where $\hat{\rho}_{j}$ is the i-th largest partial canonical correlation between $W_{t}$ and $Z_{t-1}$ adjusted for $W_{t-1}, \cdots, W_{t-p+1}$. The null hypothesis of this
trace statistic is that the cointegrating rank is at most $r$ against the alternative hypothesis is that the cointegrating rank is $m$. The Likelihood Ratio (LR) test statistic, $\Lambda_{\max }=-n\left(1-\hat{\rho}_{r+1}^{2}\right)$ is our maximal eigenvalue test statistic. The maximal eigenvalue test statistic evaluates the null hypothesis that the cointegrating rank is at most $r$ against the alternative hypothesis is that the cointegrating rank is $r+1$. As there are no critical values available of these test statistics for small samples (as far as we are aware of), we generate percentiles of them through a Monte Carlo Simulation based on 50,000 replications for sample sizes $\mathrm{N}=25,50,75$ and 100 . The empirical percentiles along with the details of the simulation are in the appendix. Because the null hypothesis is rejected for larger values of the test statistics, upper percentiles are uses as critical values.

The trace statistic and the maximal eigenvalue statistic are shown in Table 5 for different AR orders and cointegrating ranks. For AR order 2, since we cannot reject rank 0 in both trace and maximal eigenvalue statistics, there does not exist a stationary longrun equilibrium relationship among the variables. Both statistics support rank 2 in AR 1 and AR 4 while they support rank 1 in AR order 3. For large sample, it is well known that the choice of the cointegrating rank is robust to the choice of the AR order. However, for small samples like ours, the cointegrating rank is sensitive to the choice of the AR order. Therefore, in order to find an appropriate AR order and a cointegrating rank, we fit models using these different choices of the AR order and the cointegrating rank, and check significance of each coefficient in $\Phi_{i}^{*}$. Since the coefficients beyond $\Phi_{2}^{*}$ are insignificant, we determined the AR order $p$ of $Z_{t}$ in equation (1) as three. That is, the VAR (3) and cointegrating rank 1 model is chosen for further analysis.

We fit the following ECM with AR order 3 and cointegrating rank 1 using Ahn and Reinsel (1990).
(7) $W_{t}=\delta+\alpha \beta^{\prime} Z_{t-1}+\Phi_{1}^{*} W_{t-1}+\Phi_{2}^{*} W_{t-2}+\varepsilon_{t}$,
and obtain estimates of the vector of constant term, $\hat{\delta}=[210.390,234.152,4066.116]^{\prime}$, the vector of speed of adjustment coefficient, $\hat{\alpha}=[-0.032,-0.038,-0.657]^{\prime}$, and the cointegrating vector $\hat{\beta}=[1,75.815,-2.201]^{\prime}$. With a normalized population coefficient the long run equilibrium relationship is represented by
(8) $\quad p_{t}=6004.909-75.815 r_{t}+2.201 e_{t}$

This cointegrating combination adjusted for the mean is displayed in Figure 2. From the equation in (8), we can see that the long run relationship of a unit increase in the labor force participation rate ( $1 \%$ ) is a decrease of 75,815 in population and the long run relationship of a unit change in employment (1000) is an increase of 2,201 in population.

## Decomposition into Stationary and Non-Stationary Components

Using the decomposition method described in a previous section, we decompose our time series into stationary components and non-stationary components. Before interpreting stationary components and non-stationary components, we first overview the history of economic conditions in Washington from the beginning of 1970s to the beginning of 1990s.

The Washington State economy slumped in 1970-1973 and 1981-1983, whereas it expanded rapidly during the late 1970s and the late 1980s. Large population increases due to net migration occurred as a result of rapid economic expansions in Washington during the late 1970s and late 1980s. Net migration dropped when the state economy slumped in 1981-1983. In the beginning of the 1990s, California experienced net outmigration of over 400,000 persons per year. Washington received a significant amount of these Californian out-migrants. This factor contributed to relatively high levels of net migration for Washington during the early 1990s, even at a time when the state's economy slowed down significantly.

Figure 3 contains the plots of the stationary components for the population and labor force participation rate and Figure 4 depicts the stationary component of the employment. Surprisingly, the pattern of the stationary component of population is quite similar to that of labor force participation rate. Based on the above information about economic behavior in Washington, we find that the cyclical fluctuation of employment responds immediately to changing economic conditions in Washington. The response of population to changing economic conditions is about three or four years later than that of employment. This is supported by the Granger (1969) causality test results summarized in Table 6 and the impulse response analysis in the following section. Accordingly, the plot of three years delayed stationary components of population is similar to the stationary component of employment as in Figure 5.

It is interesting that the pattern of stationary component of employment and net migration which is defined as the difference between in-migration and out-migration is
quite similar as shown in Figure 6. This indicates that the short run effect of employment corresponds with migration behavior, mainly in-migration for Washington. Figure 7 shows that the plot of three years delayed stationary components of population is similar to the plot of net migration. Like employment, the response of population to changing economic conditions is about three years later than that of net migration.

Consequently, in short run, with employment and net migration having same immediate pattern, they respond to changing economic conditions. We interpret the stationary component of employment as a reflection of the historical and cyclical economic conditions in Washington. The stationary components of population and labor force participation can be interpreted in relation to the historical employment and net migration patterns in Washington.

Figures 8 through 10 depict the original data series and their non-stationary component respectively. The pattern of the non-stationary components is very similar to that of all original series. These trends reflect long waves of socioeconomic change including the baby bust of the 1970s, the baby boom echo of the 1980s, considerable increase in the female labor force participation rate and gradual decline in male labor force participation rate. The slope of the non-stationary component of population around 1980 is steeper than that in 1970s, whereas the non-stationary component of labor force participation rate was relatively high in 1970s and began to decline in 1980, which is mainly caused by natural population increase (the excess of births over deaths) from 1980. From 1970 to 1995, the state's aggregate labor force participation rate increased from $61.5 \%$ to $70.1 \%$. During this period, the male labor force participation rate
gradually declined, while the female labor force participation rate rose significantly. This information allows us to interpret the non-stationary component of labor force participation rate as reflecting the increasing trend of labor force participation rate in Washington mainly due to a considerable increase in the female labor force participation. The fluctuations in the non-stationary component of labor force participation rate show that female labor force participation is also significantly affected by economic conditions.

## Impulse Response Analysis

Figure 11 and 12 depict the impulse responses of the population, employment and labor force participation rate to a one standard deviation shock in employment. For all three variables, the impulse leads to a permanent increase provided no further shock occurs. In other words, they settle at different equilibrium value after a long period of time. Figure 11 shows a slower response of the population to changing economic conditions. In short run, the response of the population to a one standard deviation shock in employment lags several years that of employment. Figure 12 supports Bartik's and Yeo and Holland's findings. Note that Bartik found that $25 \%$ of the job growth from a shock to local job growth is reflected in increased local labor force participation rate in the long-run. Yeo and Holland found that the long-run effect of the job growth on labor force participation remains, although its effect is small. However, the analysis result does not support the Blanchard-Katz's finding - the long-run effect of the job growth on the labor force participation rate is negligible.

For the party who benefits from job growth, we suspect that most of new jobs are captured by in-migrants because the pattern of the stationary component of employment
and net migration is quite similar and the impulse response of population is significantly higher than that of employment. Based on the report of Office of Financial Management in the state of Washington (1999), net migration accounts for about 60 percent of the state population growth in the past 25 years and most of the migrants are young workers with a long-term attachment to the labor force. Thus we suspect that a high proportion of increase in labor force participation rate is due to migration.

## Estimation of Quarterly Population Data

Until now we analyze the three annual time series which have 25 observations respectively. Because the population, labor participation rate, and employment are related, we interpolate the annual population data using the quarterly labor participation rate and employment data applying a similar method in Chow and Lin (1971). Then we reanalyze the series using the quarterly observations and check whether both results are similar or not. Through the analysis of quarterly data, we gain insight into the dynamics of the quarterly population even though the total population is observed yearly. Chow and Lin (1971) introduced best linear unbiased interpolation and extrapolation of time series by related series. If we assume that the quarterly observations of the series to be estimated satisfy a multiple regression relationship, $p=X \gamma+u$ where population, $p$ is $100 \times 1, X$ is $100 \times 3$ matrix and $u$ is a random error with mean 0 and covariance matrix $V$. The first column of $X$ is one vector, second one is labor force participation rate, $r$ and third one is employment, $e$. The vector of 25 annual observations of the dependent variable, subscripted by a dot which signifies being annual, satisfies the regression model

$$
\begin{equation*}
p=C p=C X \gamma+C u=X \gamma+u . \tag{9}
\end{equation*}
$$

with $E u u^{\prime}=V=C V C^{\prime}$ and $C$ being the $25 \times 100$ matrix that converts the 25 annual observations collected during the second quarter into 100 quarterly observations.

$$
C=\left[\begin{array}{cccccccccccccccc}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0  \tag{10}\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & & & & & & & & & & & \ddots & & & & \\
0 & \cdots & & & & & & & & & & & 0 & 1 & 0 & 0
\end{array}\right]
$$

According to Chow and Lin (1971) the best linear unbiased estimator of quarterly population, $p_{q}$, that is from (9) is given by

$$
\begin{equation*}
\hat{p}_{q}=X_{q} \hat{\gamma}+\left(V_{q .} V^{-1}\right) \hat{u} . \tag{11}
\end{equation*}
$$

where $\hat{p}_{q}$ is quarterly estimates of population and $\hat{\gamma}=\left(X^{\prime} V^{-1} X .\right)^{-1} X^{\prime} V^{-1} p$. is the Generalized Least Squares (GLS) estimate of the regression coefficients using the 25 annual observations in the sample, $\hat{u} .=\left[I-X\left(X^{\prime} V^{-1} X\right)^{-1} X^{\prime} V^{-1}\right] p .=p-X \hat{\gamma}$ is the $25 \times 1$ vector of residuals in the regression using annual data. However, instead of the generalized least squares estimate of $\gamma$, we use the estimate from the previous cointegration analysis of the annual data as the estimate of $\gamma$. When the cointegrating rank is more than one, the parameters of regression model for the GLS are not identifiable. Note that our cointegrating vector is $\hat{\beta}=[1,75.815,-2.201]^{\prime}$, and the
constant term in the model of $\hat{p}$ is 6004.119 . As a result, $\hat{\gamma}=[6004.119,75.815,-$ $2.201]^{\prime}$.

The next step is to estimate the covariance matrix of residuals. The residuals of the cointegrating combination based on annual data follow a first-order autoregressive process $u_{t}=a u_{t-1}+v_{t}$ and the estimate of $a, \widetilde{a}$ is 0.3549 . Similar to Chow and Lin, a consistent estimate of $a$ in $\hat{V}$ below is $\hat{a}=\sqrt[4]{\widetilde{a}}=0.5957$ for our interpolation problem because $\widetilde{a}$ is estimated by annual data. Therefore, the estimates of the covariance matrix of $u_{t}$ for quarterly data is

$$
\hat{V}=\left[\begin{array}{ccccc}
1 & \hat{a} & \hat{a}^{2} & \cdots & \hat{a}^{99}  \tag{12}\\
\hat{a} & 1 & \hat{a} & \cdots & \hat{a}^{98} \\
\hat{a}^{2} & \hat{a} & 1 & \cdots & \hat{a}^{97} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\hat{a}^{99} & \hat{a}^{98} & \hat{a}^{97} & \cdots & 1
\end{array}\right]
$$

For the purposes of interpolation, we need

$$
\begin{equation*}
V_{q .} V^{-1}=V C^{\prime}\left(C V C^{\prime}\right)^{-1} \tag{13}
\end{equation*}
$$

Using $\hat{V}$ as an estimator of $V$, we can estimate $\hat{p}_{q}$ in (11).

## Application of Cointegration Analysis to Quarterly Data

Figure 13 plots the quarterly population and employment series and Figure 14 represents the quarterly labor force participation rate series. As previously described, the values of quarterly population series are estimated as described in the previous section.

As with the annual data analysis, we first performed the augmented Dickey-Fuller (1979, 1981) test to determine the order of integration of each time series data. Table 7 shows the results of the augmented Dickey-Fuller (ADF) test for the null hypothesis of a unit root for both original series and first differenced series. We can see that all three original series are nonstationary and their first differenced series are stationary.

Based on partial canonical correlation and the minimum AIC we choose an appropriate AR order and the results are summarized in Tables 8, 9 and 10. For the quarterly data case, partial canonical correlation supports $\operatorname{AR}(7)$ while the minimum AIC supports AR(2). The trace statistics and maximum eigenvalue statistics shown in Table 11 for different AR orders. Both statistics support rank 2 from AR(2) through AR(5) while they support rank 1 in AR order 1 and 7 . For AR order 6 and 8 , since we cannot reject rank 0 in both trace and maximal eigenvalue statistics, there does not exist a stationary long-run equilibrium relationship among the variables. Therefore, in order to find an appropriate AR order and a cointegrating rank, we fit models using these different choices of the AR order and the cointegrating rank, and we check significance of each coefficient in $\Phi_{i}^{*}$. Since the coefficients are significant until $\Phi_{6}^{*}$, we determined the AR order $p$ of $Z_{t}$ as seven. That is, the $\operatorname{VAR}(7)$ and cointegrating rank 1 model is chosen for further analysis. The AR order 7 in quarterly data analysis is to be comparable to $\operatorname{AR}(2)$ in annual data case which may be considered close enough to our choice of $\operatorname{AR}(3)$ in the section of annual data analysis. The rank 1 is same as the choice in annual data analysis.

We now fit an ECM with AR order 7 and 1 cointegrating vector as follows:

$$
\begin{equation*}
W_{t}=\delta+\alpha \beta^{\prime} Z_{t-1}+\sum_{i=1}^{6} \Phi_{i}^{*} W_{t-i}+\varepsilon_{t} \tag{14}
\end{equation*}
$$

and obtain estimates of the vector of constant term, $\delta=[631.2077,14.0121,179.2476]^{\prime}$, the vector of speed of adjustment coefficient, $\alpha=[-0.1019,-0.0023,-0.0282]^{\prime}$ and the cointegrating vector $\beta=[1,77.5172,-2.2160]^{\prime}$. Finally, we can derive the long run equilibrium relation given by

$$
\begin{equation*}
p_{t}=5668.78-77.52 r_{t}+2.22 e_{t} \tag{15}
\end{equation*}
$$

The estimated coefficients of the long run equilibrium relationship are close to those of the annual data case. The interpretation is same as in the annual data analysis.

Figure 15 shows the plots of stationary components for the population and labor force participation rate and Figure 16 depicts the stationary component of employment. Although the stationary components plots of quarterly data are more cyclical than those of annual data, the cyclical fluctuations of quarterly data are almost identical with those of annual data for all three series. Therefore, we can put the same interpretations on stationary components as annual data. The plots of non-stationary components and their original series are shown in Tables 17 through 19. Each of them not only exactly follows the path of their original time series but also represent the same long waves and trends as annual data analysis. Interpretations on them are also same as those of the annual data analysis.

Since the sample size of annual series is small, we analyzed the series using a larger number of observations and checked for similarity of results. In both data series, analysis results were consistent and the estimated values of the parameters were close.

## Conclusions

We proved that employment growth from new business investment causes increase in population in the state of Washington in spite of its beautiful environment and amenities. The causality does not operate in reverse way. This study found a long run equilibrium relationship among population, labor force participation rate and employment, in which population is positively related to employment and negatively related to labor force participation rate. The long run effect of a unit change of labor force participation rate ( $1 \%$ ) is a decrease of 73,880 in population and the long run effect of a unit change in employment (1000) is an increase of 2,190 in population.

We decomposed the time series into stationary components and non-stationary components. The pattern of the stationary component of population is quite similar to that of labor force participation rate while that of employment shows a different fluctuation. From the decomposition, it was obvious that the pattern of stationary component of employment and net migration is quite similar, which means net migration is the short run, temporary response to employment change. The patterns of three years delayed stationary components of population are similar to that of employment and net migration, and the plots correspond to changing economic conditions. According to the change in economic conditions population responds three years later than employment and net migration. We interpreted the non-stationary component of labor force
participation rate as reflecting the increasing trend of labor force participation rate in Washington mainly due to a considerable increase in the female labor force participation.

The impulse responses of population, employment and labor force participation rate to a one standard deviation shock in employment show permanent increase effects. They settle at different equilibrium value after long term periods. The response of the labor force participation rate to an impulse in employment supports Bartik's (1993) and Yeo and Holland's (2000) findings. Obviously the result is the opposite of Blanchard and Katz's (1992) finding that the long-run effect of job growth on the labor force participation rate is negligible. With regard to the party who benefits from job growth, we suspect that most of new jobs are captured by in-migrants because the pattern of the stationary component of employment and net migration is quite similar and the impulse response of population is significantly higher than that of employment.

## References

Ahn, S. K., and G. C. Reinsel. "Estimation for Partially Nonstationary Multivariate Autoregressive Models." Journal of the American Statistical Association, 85(1990):813-823.

Bartik, T. J. "Who Benefits from Local Job Growth: Migrants or the Original Residents?" Regional Studies, 27(1993):297-311.

Beveridge, S., and C. R. Nelson. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'." Journal of Monetary Economics, 7(1981):151-174.

Blanchard, O. J., and L. F. Katz. "Regional Evolutions." Brookings Papers on Economic Activity, 1(1992):1-75.

Chow G. C., and A. Lin. "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series." The Review of Economics and Statistics, 53(1971):372-375.

Dickey, D.A., and W. A. Fuller. "Distribution of Estimates for Autoregressive Time Series with a Unit Root." Journal of the American Statistical Association, 24(1979):427-431.
$\qquad$ . "Likeligood Ration Test Statistics for Autoregressive Time Series with a Unit Root." Econometrica, 49(1981):1057-1072.

Engel, R. F., and C. W. J. Granger. "Co-integration and Error Correction Representation, Estimation, and Testing." Econometrica, 55(1987):251-276.

Fodor, Eben. V. Better not Bigger: How to Take Control of Urban Growth and Improve Your Economy. Gabriola Island, B.C., Canada: New Society Publishers, 1999:40-43.

Granger C. W. J. "Investigating Causal relations by Econometric Models and CrossSpectral Methods." Econometrica, 37(1969):424-438.

Greenwood M.F., and G. L. Hunt. "Migration and Interregional Employment Redistribution in the United States." American Economic Review, 74(1984): 957-969.

Johansen, S., and K. Juselius. "The Full Information Maximum Likelihood Procedure for Inference on Cointegration - With Applications to the Demand for Money." Oxford Bulletin of Economics and Statistics, 52(1990):169-210.

Kasa, K. "Common Stochastic Trends in International Stock Markets." Journal of Monetary Economics, 29(1992):95-124.

Lütkepohl, H. "Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models." The Review of Economics and Statistics, 72(1990):116-125.

Lütkepohl, H., and H. Reimers "Impulse Response Analysis of Cointegrated Systems." Journal of Economic Dynamics and Control, 16(1992):53-78.

Powers, T. M. Lost Landscapes and Failed Economies: The Search for a Value of Place. Island Press, 1996.

Reinsel, G. C., and S. K. Ahn. "Vector Autoregressive Models with Unit Roots and Reduced Rank Structure: Estimation, Sikelihood Ration Test, and Forecasting." Journal of Time Series Analysis, 13(1992):353-375.

Stock, J. H., and M. W. Watson. "Testing for Common Trends." Journal of the American Statistical Association, 83(1988):1097-1107.

State of Washington, Office of Financial Management. 1999 Long-Term Economic and Labor Force Forecast for Washington, Apr. 1999.

Summers, L. H. "Why is the Unemployment Rate So Very High Near Full Employment?" Brookings Papers on Economic Activity, 2(1986):339-383.

Yeo J. and D. W. Holland. "Economic Growth in Washington: An Examination of Migration Rsponse and A Test of Model Accuracy," Currently under review at Internation Regional Science Review.

## Appendix: Simulated Critical Values of Trace and Maximal Eigenvalue Statistics

The limiting distributions of the trace statistic and the maximal eigenvalue statistic are the distributions of the trace and the maximal eigenvalue of $\left(\int_{0}^{1} B_{d}(u) d F_{d}(u)^{\prime}\right)^{\prime}\left(\int_{d}^{1} F_{d}(u) F_{d}(u)^{\prime} d u\right)^{-1}\left(\int_{0}^{1} F_{d}(u) d B_{d}(u)^{\prime}\right)$ where $B_{d}(u)$ is a $d-$ dimensional standard Brownian motion and $F(u)=B_{d}(u)-\int_{0}^{d} B_{d}(u) d u$ except that the first component of $F(u)$ is replaced by $u-1 / 2$. Johansen and Juselius (1990) obtain percentiles of these limiting distribution for large samples. As we have small samples, we obtain percentile of the test statistics for small samples by the Monte Carlo simulation. We generate $d$-dimensional random walk processes $Z_{t}=\delta+Z_{t-1}+a_{t}, t=1,2, \cdots, T$ with $Z_{0}=0$ and $\delta=(1,0, \cdots, 0)^{\prime}$, for $T=25,50,75$, 100 by generating pseudo normal random vectors using the RNMVM subroutine of

IMSL. Then the trace and the maximal eigenvalue of
$\left(\sum_{t=1}^{T} a_{t} \bar{Z}_{t-1}^{\prime}\right)\left(\sum_{t=1}^{T} \bar{Z}_{t-1} \bar{Z}_{t-1}^{\prime}\right)^{-1}\left(\sum_{t=1}^{T} a_{t} \bar{Z}_{t-1}^{\prime}\right)$ are obtained. Based on 50,000 replications,
empirical percentiles are found for both of them, and summarized in Appendix Table 1
through Appendix Table 4.
Appendix Table 1. Approximate Percentile for the Likelihood Ratio Test Statistics
$(\mathbf{N}=\mathbf{2 5})$

| Percentile | Dimension $(d)$ |  |  |  |  |
| ---: | ---: | ---: | ---: | ---: | ---: |
|  | 1 | 2 | 3 | 4 | 5 |
| Maximal Eigenvlaue |  |  |  |  |  |
| $0.5 \%$ | 0.00004 | 1.42981 | 3.93426 | 6.78453 | 9.66753 |
| $1.0 \%$ | 0.00015 | 1.64167 | 4.37563 | 7.32553 | 10.32730 |
| $2.5 \%$ | 0.00101 | 2.03539 | 5.01974 | 8.14859 | 11.36024 |
| $5.0 \%$ | 0.00424 | 2.44003 | 5.62131 | 8.96674 | 12.31856 |
| $10.0 \%$ | 0.01647 | 3.00563 | 6.48126 | 9.98689 | 13.51820 |
| $25.0 \%$ | 0.10502 | 4.18508 | 8.13431 | 11.99077 | 15.71175 |
| $50.0 \%$ | 0.46383 | 5.91077 | 10.38015 | 14.56163 | 18.58978 |
| $75.0 \%$ | 1.35892 | 8.18468 | 13.14124 | 17.63905 | 21.95750 |
| $90.0 \%$ | 2.78245 | 10.74697 | 16.10984 | 20.92439 | 25.47419 |
| $95.0 \%$ | 3.91309 | 12.53843 | 18.12781 | 23.09275 | 27.85367 |
| $97.5 \%$ | 5.11654 | 14.17930 | 20.01540 | 25.09178 | 29.94946 |
| $99.0 \%$ | 6.70548 | 16.27223 | 22.35285 | 27.82176 | 32.59093 |
| $99.5 \%$ | 7.90031 | 17.82954 | 24.05824 | 29.49785 | 34.41322 |
|  |  |  |  |  |  |
| Trace |  |  |  |  |  |
| $0.5 \%$ | 0.00004 | 1.62647 | 6.52724 | 14.09257 | 23.97993 |
| $1.0 \%$ | 0.00015 | 1.89255 | 7.08891 | 15.13858 | 25.51736 |
| $2.5 \%$ | 0.00101 | 2.31970 | 8.09212 | 16.70702 | 27.62989 |
| $5.0 \%$ | 0.00424 | 2.78528 | 9.04793 | 18.13896 | 29.48413 |
| $10.0 \%$ | 0.01647 | 3.40366 | 10.27750 | 19.85534 | 31.78334 |
| $25.0 \%$ | 0.10502 | 4.74187 | 12.59891 | 23.13847 | 35.92636 |
| $50.0 \%$ | 0.46383 | 6.65841 | 15.69164 | 27.24903 | 41.05999 |
| $75.0 \%$ | 1.35892 | 9.13119 | 19.27318 | 31.85716 | 46.72605 |
| $90.0 \%$ | 2.78245 | 11.88870 | 22.96488 | 36.50669 | 52.38039 |
| $95.0 \%$ | 3.91309 | 13.74282 | 25.49179 | 39.56074 | 55.80202 |
| $97.5 \%$ | 5.11654 | 15.47996 | 27.75425 | 42.30320 | 59.14228 |
| $99.0 \%$ | 6.70548 | 17.60849 | 30.57262 | 45.53668 | 63.09720 |
| $99.5 \%$ | 7.90031 | 19.40523 | 32.54868 | 47.89322 | 65.89034 |
|  |  |  |  |  |  |

## Appendix Table 2. Approximate Percentile for the Likelihood Ratio Test Statistics ( $\mathrm{N}=50$ )

| Percentile | Dimension $(d)$ |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 | 5 |

Maximal Eigenvalue

| $0.5 \%$ | 0.00003 | 1.60173 | 4.50424 | 7.76128 | 11.24278 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| $1.0 \%$ | 0.00014 | 1.82863 | 4.97812 | 8.41453 | 11.92691 |
| $2.5 \%$ | 0.00096 | 2.24829 | 5.66422 | 9.25936 | 12.98668 |
| $5.0 \%$ | 0.00380 | 2.69933 | 6.35332 | 10.14989 | 13.96017 |
| $10.0 \%$ | 0.01544 | 3.29527 | 7.23627 | 11.25670 | 15.25155 |
| $25.0 \%$ | 0.10389 | 4.53566 | 8.97271 | 13.31228 | 17.56072 |
| $50.0 \%$ | 0.46311 | 6.38063 | 11.30755 | 16.03700 | 20.60944 |
| $75.0 \%$ | 1.33167 | 8.78246 | 14.20011 | 19.30281 | 24.15729 |
| $90.0 \%$ | 2.73195 | 11.45368 | 17.31815 | 22.67434 | 27.77574 |
| $95.0 \%$ | 3.88797 | 13.31947 | 19.30800 | 24.92246 | 30.34614 |
| $97.5 \%$ | 5.10702 | 15.07210 | 21.30513 | 27.07455 | 32.63307 |
| $99.0 \%$ | 6.76562 | 17.29564 | 23.83428 | 29.77200 | 35.51581 |
| $99.5 \%$ | 8.04663 | 18.94029 | 25.64621 | 31.92806 | 37.36833 |

## Trace

| $0.5 \%$ | 0.00003 | 1.81377 | 7.42487 | 16.51869 | 28.46244 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| $1.0 \%$ | 0.00014 | 2.07676 | 8.10612 | 17.51137 | 29.97954 |
| $2.5 \%$ | 0.00096 | 2.57172 | 9.18935 | 19.16154 | 32.13519 |
| $5.0 \%$ | 0.00380 | 3.07097 | 10.19017 | 20.70302 | 34.15465 |
| $10.0 \%$ | 0.01544 | 3.72021 | 11.49826 | 22.56016 | 36.59861 |
| $25.0 \%$ | 0.10389 | 5.09344 | 13.94564 | 25.98156 | 41.05484 |
| $50.0 \%$ | 0.46311 | 7.13814 | 17.14346 | 30.27136 | 46.37817 |
| $75.0 \%$ | 1.33167 | 9.73327 | 20.89473 | 35.14136 | 52.34282 |
| $90.0 \%$ | 2.73195 | 12.56556 | 24.76696 | 40.05743 | 58.32362 |
| $95.0 \%$ | 3.88797 | 14.52383 | 27.31449 | 43.24551 | 62.02025 |
| $97.5 \%$ | 5.10702 | 16.47635 | 29.69501 | 46.05951 | 65.39952 |
| $99.0 \%$ | 6.76562 | 18.72430 | 32.79377 | 49.75564 | 69.58170 |
| $99.5 \%$ | 8.04663 | 20.46693 | 34.89926 | 52.00483 | 72.14625 |

Appendix Table 3. Approximate Percentile for the Likelihood Ratio Test Statistics ( $\mathrm{N}=75$ )

| Percentile | Dimension $(d)$ |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 | 5 |

Maximal Eigenvalue

| $0.5 \%$ | 0.00003 | 1.63901 | 4.67199 | 8.17382 | 11.81339 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| $1.0 \%$ | 0.00014 | 1.86057 | 5.13029 | 8.76932 | 12.52080 |
| $2.5 \%$ | 0.00089 | 2.32952 | 5.85775 | 9.68410 | 13.67224 |
| $5.0 \%$ | 0.00377 | 2.78503 | 6.55061 | 10.55737 | 14.66208 |
| $10.0 \%$ | 0.01539 | 3.38550 | 7.47707 | 11.67953 | 15.93704 |
| $25.0 \%$ | 0.10124 | 4.66352 | 9.25674 | 13.79160 | 18.34714 |
| $50.0 \%$ | 0.45375 | 6.53659 | 11.68857 | 16.59844 | 21.42519 |
| $75.0 \%$ | 1.29805 | 8.96161 | 14.62513 | 19.92667 | 25.03569 |
| $90.0 \%$ | 2.70085 | 11.71267 | 17.79558 | 23.41607 | 28.78642 |
| $95.0 \%$ | 3.88506 | 13.58790 | 19.98903 | 25.77393 | 31.27003 |
| $97.5 \%$ | 5.02929 | 15.35159 | 22.00989 | 27.96831 | 33.59522 |
| $99.0 \%$ | 6.64227 | 17.56786 | 24.37812 | 30.59566 | 36.44074 |
| $99.5 \%$ | 7.85819 | 19.19528 | 26.17358 | 32.67455 | 38.71803 |

## Trace

| $0.5 \%$ | 0.00003 | 1.84403 | 7.75475 | 17.29445 | 30.31911 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| $1.0 \%$ | 0.00014 | 2.13503 | 8.47254 | 18.42401 | 31.72412 |
| $2.5 \%$ | 0.00089 | 2.64100 | 9.54190 | 20.05646 | 34.00915 |
| $5.0 \%$ | 0.00377 | 3.14814 | 10.56431 | 21.59355 | 36.08253 |
| $10.0 \%$ | 0.01539 | 3.82784 | 11.92262 | 23.50947 | 38.49759 |
| $25.0 \%$ | 0.10124 | 5.23631 | 14.41665 | 27.03851 | 42.98341 |
| $50.0 \%$ | 0.45375 | 7.29913 | 17.69560 | 31.45857 | 48.46593 |
| $75.0 \%$ | 1.29805 | 9.90742 | 21.54545 | 36.41562 | 54.48727 |
| $90.0 \%$ | 2.70085 | 12.83130 | 25.52504 | 41.36847 | 60.43302 |
| $95.0 \%$ | 3.88506 | 14.78765 | 28.20539 | 44.58499 | 64.28394 |
| $97.5 \%$ | 5.02929 | 16.66736 | 30.51981 | 47.51479 | 67.77034 |
| $99.0 \%$ | 6.64227 | 19.04277 | 33.58242 | 51.19997 | 71.98624 |
| $99.5 \%$ | 7.85819 | 20.74977 | 35.59242 | 53.38948 | 75.02776 |

## Appendix Table 4. Approximate Percentile for the Likelihood Ratio Test Statistics ( $\mathrm{N}=100$ )

| Percentile | Dimension $(d)$ |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 | 5 |

Maximal Eigenvalue

| $0.5 \%$ | 0.00004 | 1.67272 | 4.86716 | 8.42402 | 12.12611 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| $1.0 \%$ | 0.00016 | 1.92763 | 5.30395 | 9.01527 | 12.88516 |
| $2.5 \%$ | 0.00102 | 2.36671 | 6.01055 | 9.94244 | 13.98926 |
| $5.0 \%$ | 0.00409 | 2.81071 | 6.71689 | 10.82655 | 15.04403 |
| $10.0 \%$ | 0.01642 | 3.42829 | 7.63680 | 11.95442 | 16.34207 |
| $25.0 \%$ | 0.10442 | 4.74110 | 9.41590 | 14.09169 | 18.78874 |
| $50.0 \%$ | 0.46182 | 6.64446 | 11.90220 | 16.95947 | 21.91651 |
| $75.0 \%$ | 1.34371 | 9.11060 | 14.88195 | 20.34630 | 25.51195 |
| $90.0 \%$ | 2.72685 | 11.89154 | 18.05243 | 23.80437 | 29.35646 |
| $95.0 \%$ | 3.86727 | 13.80669 | 20.22749 | 26.17239 | 31.95017 |
| $97.5 \%$ | 5.06370 | 15.61507 | 22.28753 | 28.39861 | 34.24622 |
| $99.0 \%$ | 6.67854 | 18.05861 | 24.83264 | 31.25117 | 37.41772 |
| $99.5 \%$ | 7.98117 | 19.72457 | 26.51534 | 33.21503 | 39.73579 |

Trace

| $0.5 \%$ | 0.00004 | 1.93994 | 7.98728 | 17.87842 | 30.91952 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| $1.0 \%$ | 0.00016 | 2.19770 | 8.68079 | 18.86298 | 32.58025 |
| $2.5 \%$ | 0.00102 | 2.68338 | 9.79212 | 20.60548 | 34.82240 |
| $5.0 \%$ | 0.00409 | 3.18215 | 10.81532 | 22.17563 | 36.92950 |
| $10.0 \%$ | 0.01642 | 3.88215 | 12.11687 | 24.05310 | 39.48685 |
| $25.0 \%$ | 0.10442 | 5.34922 | 14.66622 | 27.61961 | 44.03326 |
| $50.0 \%$ | 0.46182 | 7.41130 | 18.02354 | 32.14189 | 49.71305 |
| $75.0 \%$ | 1.34371 | 10.08132 | 21.92179 | 37.19312 | 55.73866 |
| $90.0 \%$ | 2.72685 | 13.04835 | 25.97027 | 42.14549 | 61.81979 |
| $95.0 \%$ | 3.86727 | 15.10072 | 28.57504 | 45.45506 | 65.75732 |
| $97.5 \%$ | 5.06370 | 17.07017 | 31.02484 | 48.35427 | 69.29696 |
| $99.0 \%$ | 6.67854 | 19.42227 | 34.09746 | 52.05660 | 73.78222 |
| $99.5 \%$ | 7.98117 | 21.19068 | 36.45694 | 54.72538 | 76.74147 |

Table 1. ADF Unit Root Test Results for Level Series and First Differenced Series ( $\mathbf{N}=25$ )

| Description | Level Series |  |  |  | First Differenced Series |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Variable | $p_{t}$ | $a_{t}$ | $e_{t}$ | $p_{t}$ | $a_{t}$ | $e_{t}$ |
| $\hat{\tau}_{\tau}$ | -2.380 | -2.587 | -2.817 | -3.246 | -5.533 | -3.516 |
| P-Value | 0.378 | 0.289 | 0.206 | 0.030 | 0.000 | 0.017 |

Table 2. LR test statistics of Test of H0: The Canonical
Correlations in the Current Row and All That Follow are Zero ( $\mathrm{N}=\mathbf{2 5 \text { ) }}$

| Number | AR(2) | AR(3) | AR(4) | AR(5) |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 0.233 | 0.233 | 0.173 | 0.094 |
|  | $(0.008)$ | $(0.056)$ | $(0.169)$ | $(0.000)$ |
|  | 0.738 | 0.569 | 0.607 | 0.521 |
| 2 | $(0.318)$ | $(0.167)$ | $(0.444)$ | $(0.013)$ |
|  | 0.946 | 0.973 | 0.931 | 0.977 |
| 3 | $(0.352)$ | $(0.573)$ | $(0.464)$ | $(0.502)$ |

Numbers in parenthesis are P-values.

Table 3. Squared Partial Canonical Correlations ( $\mathbf{N}=\mathbf{2 5 )}$

| Number | $\operatorname{AR}(1)$ | $\operatorname{AR}(2)$ | $\operatorname{AR}(3)$ | $\operatorname{AR}(4)$ |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 0.820 | 0.685 | 0.591 | 0.715 |
| 2 | 0.467 | 0.220 | 0.415 | 0.348 |
| 3 | 0.029 | 0.054 | 0.027 | 0.069 |

Table 4. Akaike Information Criterion for Autoregressive Models (N=25)

| Lag=0 | Lag=1 | Lag=2 | Lag=3 | Lag=4 | Lag=5 | Lag=6 | Lag=7 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 502.70 | 465.20 | 472.10 | 479.71 | 492.07 | 504.16 | 484.24 | 475.19 |

Table 5. Trace Statistics and Maximal Eigenvalue Statistics ( $\mathbf{N}=\mathbf{2 5 )}$

| $\mathrm{H}_{0}(r)$ | Statistic |  |  |  | 5\% Significance Level |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | AR 1 | AR 2 | AR 3 | AR 4 |  |
| Trace Statistic |  |  |  |  |  |
| 2 | 0.55 | 0.20 | 1.00 | 0.10 | 3.91309 |
| 1 | 15.86 | 5.81 | 9.70 | 14.01 | 13.74282 |
| 0 | 58.67 | 18.84 | 28.63 | 38.35 | 25.49179 |
| Maximal Eigenvalue Statistic |  |  |  |  |  |
| 2 | 0.22 | 0.27 | 0.92 | 0.11 | 3.91309 |
| 1 | 15.31 | 5.61 | 8.70 | 13.91 | 12.53843 |
| 0 | 42.81 | 13.03 | 18.93 | 24.34 | 18.12781 |

Table 6. Granger Causality Test Results (N=25)
Lags: 1

| Null Hypothesis: | Obs | F-Statistic | Probability |
| :--- | :---: | :---: | :---: |
| POP does not Granger Cause EMP | 24 | 0.05701 | 0.81359 |
| EMP does not Granger Cause POP |  | 40.3320 | $2.7 \mathrm{E}-06$ |
| Lags: $\mathbf{2}$ |  |  |  |
| Null Hypothesis: | Obs | F-Statistic | Probability |
| POP does not Granger Cause EMP | 23 | 0.46044 | 0.63823 |
| EMP does not Granger Cause POP |  | 9.13324 | 0.00183 |

Table 7. ADF Unit Root Test Results for Level Series and First Differenced Series ( $\mathbf{N}=100$ )

| Description | Level Series |  |  |  | First Differenced Series |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Variable | $p_{t}$ | $r_{t}$ | $e_{t}$ | $p_{t}$ | $r_{t}$ | $e_{t}$ |
| $\hat{\tau}$ | -2.199 | -3.009 | -3.047 | -4.616 | -5.968 | -4.614 |
| P-Value | 0.485 | 0.135 | 0.125 | 0.002 | 0.000 | 0.000 |

Table 8. LR test statistics of Test of H0: The Canonical Correlations in the Current Row and All That Follow are Zero ( $\mathrm{N}=100$ )

| Number | $\operatorname{AR}(2)$ | $\operatorname{AR}(3)$ | $\operatorname{AR}(4)$ | $\operatorname{AR}(5)$ | $\operatorname{AR}(6)$ | $\operatorname{AR}(7)$ | $\operatorname{AR}(8)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 0.343 | 0.744 | 0.721 | 0.876 | 0.736 | 0.722 | 0.855 |
|  | $(0.000)$ | $(0.003)$ | $(0.002)$ | $(0.335)$ | $(0.008)$ | $(0.007)$ | $(0.333)$ |
| 2 | 0.714 | 0.950 | 0.976 | 0.995 | 0.909 | 0.927 | 0.982 |
|  | $(0.000)$ | $(0.360)$ | $(0.742)$ | $(0.986)$ | $(0.135)$ | $(0.260)$ | $(0.879)$ |
|  | 0.963 | 0.979 | 0.995 | 0.999 | 0.989 | 0.997 | 0.999 |
| 3 | $(0.069)$ | $(0.181)$ | $(0.509)$ | $(0.988)$ | $(0.363)$ | $(0.670)$ | $(0.870)$ |

Numbers in parenthesis are P -values

Table 9. Squared Partial Canonical Correlations ( $\mathrm{N}=100$ )

| Number | AR(1) | AR(2) | AR(3) | AR(4) | AR(5) | $\operatorname{AR}(6)$ | $\operatorname{AR}(7)$ | $\operatorname{AR}(8)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 0.283 | 0.226 | 0.304 | 0.297 | 0.207 | 0.150 | 0.261 | 0.148 |
| 2 | 0.085 | 0.163 | 0.179 | 0.169 | 0.155 | 0.101 | 0.085 | 0.053 |
| 3 | 0.000 | 0.027 | 0.017 | 0.008 | 0.011 | 0.013 | 0.004 | 0.007 |

Table 10. Akaike Information Criterion for Autoregressive Models (N=100)

| Lag=0 | Lag=1 | Lag=2 | Lag=3 | Lag=4 | Lag=5 | Lag=6 | Lag=7 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1966.70 | 1333.22 | 1290.57 | 1296.99 | 1303.94 | 1313.57 | 1321.66 | 1335.10 |

Table 11. Trace Statistics and Maximal Eigenvalue Statistics ( $\mathbf{N}=\mathbf{1 0 0}$ )

| $\mathrm{H}_{0}(r)$ | Statistic |  |  |  |  |  |  |  | 5\% Significance Level |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | AR 1 | AR 2 | AR 3 | AR 4 | AR 5 | AR 6 | AR 7 | AR 8 |  |
| Trace Statistic |  |  |  |  |  |  |  |  |  |
| 2 | 0.011 | 2.747 | 1.696 | 0.853 | 1.136 | 1.308 | 0.424 | 0.665 | 3.86727 |
| 1 | 8.852 | 20.535 | 21.398 | 19.359 | 18.001 | 12.005 | 9.306 | 6.071 | 15.10072 |
| 0 | 42.106 | 46.120 | 57.667 | 54.563 | 41.189 | 28.209 | 39.589 | 22.138 | 28.57504 |
| Maximal Eigenvalue Statistic |  |  |  |  |  |  |  |  |  |
| 2 | 0.011 | 2.747 | 1.696 | 0.853 | 1.136 | 1.308 | 0.424 | 0.665 | 3.86727 |
| 1 | 8.841 | 17.788 | 19.702 | 18.506 | 16.866 | 10.697 | 8.883 | 5.406 | 13.80669 |
| 0 | 33.254 | 25.585 | 36.269 | 35.204 | 23.188 | 16.204 | 30.282 | 16.066 | 20.22749 |



Figure 1. Time Series Plot for Annual Data

Long Run Equilibriun Relationship


Figure 2. Long Run Equilibrium Relationship of the series


Figure 3. Stationary Component of Population and Participation Rate ( $\mathbf{N}=\mathbf{2 5}$ )


Figure 4. Stationary Component of Employment ( $\mathrm{N}=\mathbf{2 5 )}$


Figure 5. Stationary Component for 3 Years Delayed Population and Employment (N=25)


Figure 6. Stationary Component of Employment and Net Migration ( $\mathbf{N}=\mathbf{2 5}$ )


Figure 7. Stationary Component of 3 Years Delayed Population and Net Migration ( $\mathbf{N}=\mathbf{2 5 )}$


Figure 8. Non-stationary Component of Population ( $\mathbf{N}=\mathbf{2 5}$ )


Figure 9. Non-stationary Component of Labor Force Participation Rate ( $\mathrm{N}=25$ )


Figure 10. Non-stationary Component of Employment (N=25)

$$
\text { _- Population } \cdot \ldots \text {. . . Employment }
$$



Figure 11. Impulse Response of Population and Employment to Employment

> —— Labor Force Participation Rate


Figure 12. Impulse Response of Labor Force Participation Rate to Employment


Figure 13. Population and Employment Series Plot for Quarterly Data


Figure 14. Labor Force Participation Rate Series Plot for Quarterly Data
-...... Participation Rate ———Population (Right Axis)


Figure 15. Stationary Component of Population and Participation Rate ( $\mathrm{N}=100$ )


Figure 16. Stationary Component of Employment ( $\mathbf{N}=100$ )


Figure 17. Non-stationary Component of Population ( $\mathbf{N}=100$ )


Figure 18. Non-stationary Component of Labor Force Participation Rate ( $\mathrm{N}=100$ )


Figure 19. Non-stationary Component of Employment ( $\mathbf{N}=100$ )

