ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

CONVERGENCE AND STABILITY IN US REGIONAL EMPLOYMENT

Robert Rowthorn and Andrew Glyn

Number 92

March 2002

Manor Road Building, Oxford OX1 3UQ

1st March, 2002

Convergence and Stability in US Regional Employment

Robert Rowthorn and Andrew Glyn[1]
Faculty of Economics and Politics, Cambridge University
(robert.rowthorn@econ.cam.ac.uk)
Department of Economics, Oxford University
(andrew.glyn@economics.ox.ac.uk)

[1] Our thanks to Kathy Albetski, Larry Katz, Thomas Krolik, John Schmitt and John Stewart for generously providing us with data or with advice on how series were constructed, to Esra Erdem for her invaluable work analysing it, to Gavin Cameron, Andrew Harvey, Carol Heim and Ron Smith for advice, to Alan Manning, Steve Nickell and Jon Temple for comments on an earlier draft and to the Leverhulme Programme on The Labour Market Consequences of Technical and Structural Change for their support.

Abstract

It is widely believed that regional labour markets in the USA are highly flexible, so that employment shocks have only transitory effects on joblessness since induced migration quickly offsets much of the initial impact. However time-series analysis of the response to shocks is very sensitive to errors of measurement in labour market data, and such errors are large in some widely used series which depend on household surveys of limited size. Adjusting for the likelihood magnitude of such errors with some novel statistical approaches, and using a range of data sources, we show that the responsiveness of employment rates to shocks has been rather weak in the USA over the past 30 years, though probably stronger in the 1950s and 1960s. This suggests that flexible regional adjustment is not a major factor behind the contemporary success of monetary union in the USA.

Key words: regional employment, convergence, measurement errors, regional adjustment

JEL: C1, J6, N9, R1

1 Introduction

This paper is concerned with the behaviour of regional employment in the United States and in particular with the belief that regional labour markets are relatively flexible the so that shocks to employment in a state have only transitory effects on joblessness. Blanchard and Katz (1992), the seminal paper in this field (henceforth BK), found that both unemployment and labour force participation returned to their original levels within about seven years of an employment shock and about four years after the change in employment had reached its maximum. This extremely rapid adjustment forms the core of their argument that the US absorbs regional employment shocks in a highly flexible fashion. Although BK found that employment itself rebounded somewhat they concluded that migration was the main route through which the employment rate adjusted back. Weaknesses in their paper were noted by Robert Hall in his comments, and their conclusions were challenged by Bartik (1993). However the conventional wisdom has remained that migration effectively prevents long lasting effects on joblessness of regional shocks in the USA. The extension of BK's approach to Europe, by Ducressin and Fatas (1995) and Obstfeld and Peri (1998), seemed to confirm the greater flexibility of regional labour markets in the USA.

In this paper we probe further the post-war data for the USA and develop a rather different interpretation of American experience from that of BK. Using a panel of 48 American states we analyse the behaviour of regional employment rates (employment divided by population) since 1948. In our analysis we distinguish between the long-run convergence of employment rates, which may differ initially across states because of differences in economic structure, from the issue of short-run stability in response to shocks. We find evidence of a gradual long-run convergence process, but our findings with regard to short-run stability are mixed. During the heyday of postwar expansion, in the 1950s and 60s, there is evidence that employment rates in individual US states reverted quite rapidly to their long-run trend values following state-specific shocks. This does not appear to be the case in more recent decades. Taking the past thirty years as a whole, we find that employment rates recover very slowly from state specific shocks. This finding is at variance with the view that regional labour markets in the USA are highly flexible.

In time-series analysis, the accuracy and coverage of data are of great importance, and failure to allow for measurement error or the changing coverage of series can result in seriously biased estimates. Given the importance of data in the present context, we compare results from a number of different data sets of varying coverage and accuracy. This provides a useful cross-check and helps us to identify why certain commonly used series give misleading results. Even when there is no measurement error, the conventional least squares approach is subject to major bias if the series in question are short[1]. Standard unit root tests make an allowance for such bias, but its existence is sometimes ignored by scholars seeking to quantify the response of regional employment rates to

[1]Fro a general discussion of bias in time-series estimators see Kiviet (1995)

shocks. The combination of short time series, measurement error and changing coverage may result in truly massive bias and grossly exaggerate the stability of regional employment rates. These issues are explored using a variety of statistical techniques including panel unit root tests and pooled regressions.

The statistical approach in this paper contains some novelties. The method of Andrews (1993) for deriving Median Unbiased Estimators for individual series is extended to panel data and is also modified to allow for autocorrelation arising from measurement error. It is also extended to include quadratic and certain other kinds of trend. In each case, simulation is used to derive the statistical distribution of the relevant estimator. Another innovation is our analysis of measurement errors as a source of bias. We devote considerable attention to investigating how such errors may affect our results and those of earlier writers, notably BK.

The structure of the paper is as follows. Section 2 describes the various data sets and their strengths and weaknesses (further details are given in Appendix A). Section 3 discusses the relationship between stability and convergence. Section 4 presents the unit root tests and Section 5 extends this analysis to pooled estimates of convergence and stability in regional employment rates. Using a simulation approach described in Appendix B, we derive Median Unbiased Estimators and confidence intervals using methods explained in Appendix C. Section 6 compares our results with those of earlier work and draws out their implications for the pattern of employment rates across states and over time using methods outlined in Appendix D.

2 The Data

We have examined four sets of data for employment in US states, whose main features are summarised in Table 1. Each of these data sets has specific advantages and disadvantages. The first (BLS) is the Bureau of Labour Statistics series for non-agricultural employees. This annual series is based on comprehensive administrative data from the Unemployment Insurance System which provides benchmarks for establishment data from "the largest monthly employer survey in existence" (BLS ,1997 p.18) . The two major advantages of this data set are its availability since the late 1940s and its reliability due to its comprehensive method of collection. Its main disadvantage is that a number of categories of employment are omitted - those engaged in agriculture, all the self-employed, those in domestic service and the military. Moreover it refers to the number of people who work within a particular state, rather than those that live in the state. This causes problems especially in the caase of DC where there is a very high and growing amount of inward commuting from surrounding areas. Finally, it counts numbers of jobs rather than the number of individuals employed, so that multiple job holders are counted more than once. Nevertheless the accuracy of this data set makes it a highly sensitive indicator of changes in labour market conditions. To calculate the employment rate we divide the

BLS series for employment by total population in the state concerned. Total population is not the ideal denominator, because it includes children and the elderly, and is thus biased downwards in a way that varies across states and through time under the influence of demographic factors. However, this is the only long series for population on a state by state basis that we could find.

The second data set (BEA) comes from the Regional Economic Information System of the Bureau of Economic Analysis and covers employment by state in all sectors beginning in 1969. It supplements the BLS series for non-agricultural employees with estimates of the number of individuals employed in agriculture and other omitted categories. These estimates are assembled from a variety of sources including administrative records. It is thus based for the most part on very detailed information, but unlike the BLS series for non-agricultural employees its coverage is complete. Thus, it is a preferable as a measure of longer-run trends in employment, and it also a fairly accurate indicator of short-term fluctuations. Unfortunately, it is available for a relatively short period of time. Using this data set, we calculate employment rates in two ways. The BEATOT series is derived by dividing employment by total population, whereas the BEAWK series uses population aged 18 to 64 (which is close to the internationally accepted definition of working age population). Of all the series we use, the BEAWK data set is probably the best since it is relatively accurate and its coverage of employment and population is most appropriate.

The third data set comes from the Census and involves computing employment from the monthly CPS surveys. These data are available on a comprehensive geographical basis only from 1976 onwards, but they have a number of advantages. They refer to the employment status of residents and are thus comparable in coverage to population statistics. They are also produced by the BLS as part of the process of calculating state level unemployment. This means that the employment series are consistent with labour force data, and so participation rates and unemployment rates can be analysed alongside employment rates. The major disadvantage of this data set is the small sample size on which it is based. It is derived from a rolling sample of approximately 50,000 households per month and there is considerable measurement error at the state level. Official estimates imply that measurement errors account for approximately 80 percent of the observed year to year variation in log employment rates[2]. Our series for the employment rate is derived by dividing the Census figure for state

[2]Information provided by the BLS (1999 Appendix B) presents the following picture. If there were no real shocks at all and observed variation was entirely due to measurement error, then in the average US state annual changes in the log employment rate would have a standard deviation equal to 1.35 per cent. In the average state, observed changes in the log employment rate had a standard deviation equal to 1.51 per cent over the period 1976-2000, Assuming real shocks and measurement errors are uncorrelated, these figures imply that errors account for 80 per cent of the observed variance in year to year changes in the log employment rate. The BLS also estimates that measurement errors have a first order autocorrelation of 0.58. Information on the BLS data is given in BLS (1997) ch.4 and in Appendix B of *Employment and Earnings.* We thank Thomas Krolik and John Stewart of the BLS for their most thorough and informative responses to our questions on the subject of the errors in the various employment series.

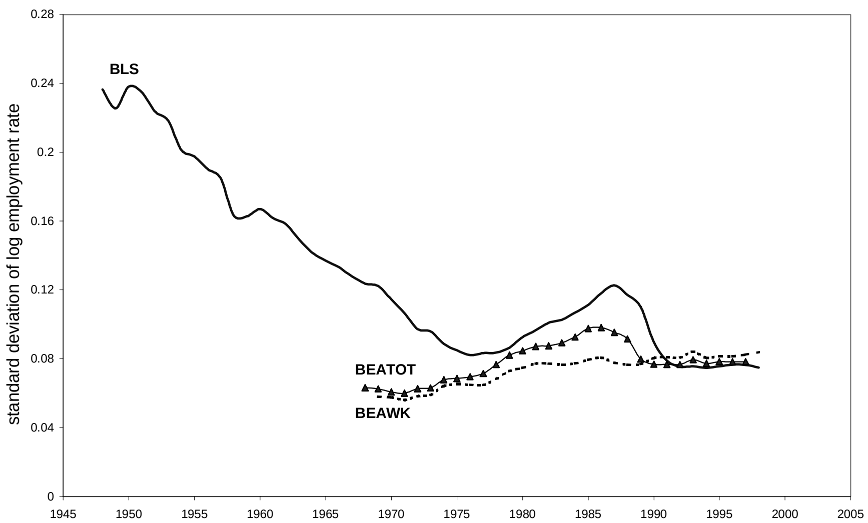

Figure 1: Dispersion of Employment Rates across US States

employment by the Census figure for population aged 15 and over. We label this series as CensusA.

The fourth data set also comes from the CPS, but allows employment to be subdivided between men and women and by educational group. Such disaggregation is potentially very useful in the analysis of regional convergence and adjustment. For example men and women may be differently affected by structural change and the least educated may be less mobile in search of work after a region-specific loss of jobs (as reported by Bound and Holzer (2000)). Series for state employment disaggregated by gender and educational group have been constructed by the Economic Policy Institute from the March CPS. Dividing these by the corresponding figures for population aged 25 to 64 yields disaggregated series for state employment rates (labelled CensusM). Unfortunately, the combination of disaggregation and reliance on only one month's survey data leads to very large measurement errors. In our opinion, such errors are so great as to make the disaggregated data almost useless for the econometric analysis of short-term changes in the employment rate. However, they do have many other uses, such as the analysis of long-run trends and cross-sectional comparisons.

Figures 1-2 illustrate some of the main features which are revealed by inspection of the above data series. As can be seen from Figure 1, the BLS series for state employment rates exhibits a remarkable degree of convergence over the past 50 years. The standard deviation of log employment rates across 48 US

Figure 2: Employment Rates for Selected Educational Groups (Census M data)

states [3] fell by around two thirds between 1948 and the mid-1970s. In the late 1970s and early 1980s there was some increase in dispersion following oil shocks and serious recessions, but this seems to have been reversed in the early 1990s. This time profile for the dispersion of employment rates parallels very closely the pattern of dispersion of personal income per capita at the state level (Bernat 2001). In each case, a major driving force has been the declining importance of agriculture in the economies of the old farm states and a convergence of their economic structure towards the national average. The decline of agriculture has been accompanied by a rapid increase in non-agricultural employment, as reflected in the BLS employment rate series, and a rapid growth of personal per capita incomes as low-paid farm jobs are phased out.The dispersion of employment rates based on the two BEA series is also shown in Figure 1. When total population is used to construct employment rates there is a hump to dispersion in the mid-1980s, whereas using population aged 18-64 presents a smoother picture. However both variants show something of an upward trend in dispersion over the past 30 years. This is in contrast to the impression given by the BLS series over this period. The difference is primarily due to the ommision of agriculture from the BLS series.

The CensusA data allow us to examine unemployment and participation

[3] Alaska and Hawai are omitted since data are not available for the whole period and because their linkages to the main US labour market are more tenuous; DC is omitted because of the very high and rising level of inward commuting (the BLS and BEA measures of employment refers to those working within a state not those living in the state). All the data sets used from now on refer to the remaining 48 states.

separately. Most of the variance in employment rates across states is accounted for by variations in participation, with the proportion ranging from around 70% in the recession of the early 1980s to nearly 90% in 2000.. . Using the CensusM data set, Figure 2 illustrates the strongly contrasting experience of different labour market groups. Men with less than high school education have exhibited large and increasing dispersion in employment rates in the period since 1976. Employment rates for college educated women, in contrast, showed a substantial reduction in dispersion at the end of the 1970s to a level far below those of the least qualified men. It would be interesting to use time-series econometrics to analyse the differing experience of the unemployed and the inactive and between the different educational groups, but as mentioned above inaccuracies in the CensusM data set unfortunately preclude such analysis[4].

Before leaving this description of the data it is instructive to look at the degree of correlation between the series discussed above. Table 2 shows the correlation matrix for annual changes in the logs of the various employment rate measures (together with unemployment and participation rates). Correlations between the changes in the various employment rate measures are quite low. The BLS and the BEATOT series differ only in because the latter includes agriculture and other omitted employment categories. Yet the correlation coefficient between the two series is only about one third. There is a similar value for the correlation between the BEATOT and BEAWK series which differ only because they use a different measures of population. The two Census series have very low correlations with each other and with the other employment rate measures, which is to be expected given the inaccuracy of the Census data. These low correlations between the various measures indicate that econometric results resting on these annual changes may be highly sensitive to the data source used.

It may be helpful to place the data used in BK's pioneering article in the context of this overview of alternative data sources. BK noted that the BLS non-agricultural employee series was likely to be more accurate for small states than the Census-based data for employment (CensusA). They took BLS employment and then "normalised" it by multiplying it by a state specific, but time-invariant, constant to account for omitted employment and any other factors making it diverge from Census employment in a base year. Such a procedure cannot take account of differing trends across states in omitted categories (notably agricultural employment), so it is surprising that they should have re-

[4] John Schmitt has suggested that the contrasting trends shown in figure 2 may be to some extent a statistical illusion arising from trends in measurement error. Employment statistics for individual categories in the Census are derived from sample surveys whose accuracy depends on the number of people from each category who are included in the sample. In the US as a whole, the number of men in the low-education category has been falling for some decades, whilst the number of college educated women has been rising. Such developments are likely to mean that sample surveys include a falling number of men in the former group and a rising number of women in the latter group. As a result, statistics relating to men of low education should become gradually less accurate, whilst those relating to college educated women should become more accurate. Such trends in accuracy would cause the inter-state dispersion of measured employment rates to increase for men of low education and fall for college educated women, as we observe in figure 2.

jected the use of deterministic trends in their regression equations for relative employment (p.6). To these adjusted employment data BK added the CensusA series for unemployment and inactivity in order to derive an estimated figure for population. This procedure has the effect of incorporating into population the considerable measurement errors in state unemployment and inactivity statistics, making their population series less accurate than the conventional population series as an indicator of year to year changes.

3 Modelling Convergence and Stability

When comparing the experience of regions, it is useful to distinguish between 'convergence' and 'stability'. The former denotes a long-run process by which initially disparate regions converge gradually towards a similar economic pattern or standard, whereas the latter is concerned with the response of regions to transitory shocks. It is normal to regard economic convergence as a diffusion process, whereby technologies or social practices spread outwards from advanced regions to their more backward counterparts. An example might be the spread of women's participation in the labour force. However, this is not the only route to convergence. An economic development which is common to all regions may reduce diversity because it affects some regions more than others. Consider, for example, two regions which use the same techniques of production, but have different per capita endowments of agricultural land. Since production techniques are the same, the region with the largest endowment of land will employ a greater fraction of its workforce in agriculture. Now suppose that technical change causes agricultural employment to fall at a uniform percentage each year in both regions. In the long-run, this trend will cause the employment structure of the two regions to become increasingly similar, since in each case the employment share of agriculture will eventually become very small. Thus, convergence will occur because a common economic development affects the two regions differentially. There is no diffusion process involved, since the agrarian region was initially just as advanced in technology as the other region. This region was not backward, but merely better endowed with land, a factor whose significance for employment declines in the course of time as technical progress reduces the labour to land ratio.

To the extent that convergence reflects the operation of fundamental economic forces it is likely to be a gradual and rather steady process, which in the case of an individual region or country is difficult to distinguish from a linear or geometric time trend. However, evidence for systematic convergence can be obtained by means of cross-section analysis, whereby the trend growth for each region (or country) is related to its initial starting point. This was the procedure used in the pioneering work of Baumol and his followers (Baumol et al. 1989, Barro & Sala-i-Martin 1995). It has been criticized by those who believe that the more sophisticated techniques of time-series analysis should be employed to analyse convergence. In this paper, we take an intermediate position and seek to combine both approaches.

Super-imposed on the long-run convergence process there may be shocks which affect individual regions differently. Some of these shocks may be related to the national business cycle and be automatically reversed in the course of the cycle, whereas others may require adaptation in the local economy to overcome their effects. Some shocks may affect only the levels of economic variables, whereas others may affect trend growth rates and hence influence the pace of convergence. In this paper, we shall mostly ignore the issue of variable trends and assume that underlying regional trends fixed. The one exception concerns the Perron test for structural change, where we allow for the possibility that trend growth rates in individual states altered in 1973. Apart from this exception, we treat underlying trends as deterministic.

The use of deterministc trends in the context of convergence can be justified as follows. Consider the stochastic process

$$y_t = x_t + z_t \tag{1}$$

where

$$\Delta x_t = \beta x_{t-1} \tag{2}$$

$$\Delta z_t = \pi z_{t-1} + w_{t-1} \tag{3}$$

and w_t is a random disturbance Thus, y_t is the sum of two components, one deterministic and the other stochastic. The deterministic component x_t captures the effect of trend-like structural forces leading to convergence, whereas z_t includes an allowance for transient shocks. It is assumed that the individual components are unobservable. From equation (2) it follows that

$$x_t = x_0(1+\beta)^t \tag{4}$$

$$\Delta x_t = \beta x_0(1+\beta)^{t-1} \tag{5}$$

Since $\Delta y_t = \Delta x_t + \Delta z_t$ we can use the above equations to obtain the following expression

$$\Delta y_t = \alpha(1+\beta)^t + \pi y_{t-1} + w_t \tag{6}$$

where

$$\alpha = \frac{x_0(\beta - \pi)}{(1+\beta)} \tag{7}$$

Equation (6) can be estimated since it contains only observable variables. If β is small the following linear approximation can be used

$$\Delta y_t = \alpha + \gamma t + \pi y_{t-1} + w_t \tag{8}$$

where $\gamma = \alpha\beta$.

This justifies our use of linear trends to approximate the long-run convergence process. A better approximation would also take account of curvature in the geometric term $(1+\beta)^t$. This can be done by adding a term in t^2 or using a split trend of the Perron type.

At this point we should mention an important issue which arises when the data exhibit a powerful underlying convergence towards zero. Consider a variable y_t that is subject to random shocks which cause it to fluctuate around a long-run convergence path in the fashion described by equation (8). Suppose that π is estimated in basic Dickey-Fuller fashion by regressing Δy_t on y_{t-1}without either constant or trend. In this case, the estimate of π will be dominated by the long-run convergence process and the standard test may show it to be highly significant, irrespective of its true value. Alternatively, suppose that a constant and trend are both included in the regression. In this case, the long-run convergence process will be captured by the trend line, whilst the estimate of π will reflect the influence of short-run stabilisers. Thus, when there is a powerful long-run convergence process at work, a failure to include a trend in the Dickey-Fuller test may lead to false conclusions with regard to stability.

4 Unit Roots Tests

We begin by using unit-root tests to see how much support there is for the proposition that US regional employment rates are stationary. For this purpose we follow the example of Blanchard and Katz (1992) and express employment rates as logarithmic deviations from the national average. Thus, for state i our test variable is of the form

$$\begin{aligned} y_{i,t} = \; & \log(\text{employment/population in state } i) \\ & - \log(\text{employment/population in USA}) \end{aligned}$$

This procedure eliminates time-specific effects that are common to all states.

For each individual state, using ordinary least squares, we estimate Augmented Dickey-Fuller (ADF) equations of the following type

A*DF(m) no constants*

$$\Delta y_{i,t} = \pi_i y_{i,t-1} + \sum_{h=1}^{h=m} \beta_{i,h} \Delta y_{i,t-h} + \xi_{i,t} \tag{9}$$

A*DF(m) with constants*

$$\Delta y_{i,t} = \pi_i y_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_{i,h} \Delta y_{i,t-h} + \xi_{i,t} \tag{10}$$

ADF(m) with linear trends

$$\Delta y_{i,t} = \pi_i y_{i,t-1} + \alpha_{i0} + \alpha_{i1}\, t + \sum_{h=1}^{h=m} \beta_{i,h} \Delta y_{i,t-h} + \xi_{i,t} \tag{11}$$

$$
\begin{aligned}
y_{i,t} &= \log(\text{employment/population in state } i) \\
&\quad - \log(\text{employment/population in USA})
\end{aligned}
$$

This procedure eliminates time-specific effects that are common to all states.

For each individual state, using ordinary least squares, we estimate Augmented Dickey-Fuller (ADF) equations of the following type

A*DF(m) no constants*

$$
\Delta y_{i,t} = \pi_i y_{i,t-1} + \sum_{h=1}^{h=m} \beta_{i,h} \Delta y_{i,t-h} + \xi_{i,t} \tag{12}
$$

A*DF(m) with constants*

$$
\Delta y_{i,t} = \pi_i y_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_{i,h} \Delta y_{i,t-h} + \xi_{i,t} \tag{13}
$$

ADF(m) with linear trends

$$
\Delta y_{i,t} = \pi_i y_{i,t-1} + \alpha_{i0} + \alpha_{i1}\, t + \sum_{h=1}^{h=m} \beta_{i,h} \Delta y_{i,t-h} + \xi_{i,t} \tag{14}
$$

The lagged terms $\Delta y_{i,t-k}$ are included to allow for departures from the standard assumptions regarding the distribution of the disturbance terms $w_{i,t}$. The probability that $\pi_i = 0$ is evaluated using the statistical tables published in Fuller (1996).

The ADF test is considered defective by many econometricians when the data have a deterministic trend. Schmidt and Phillips (1992) propose an alternative approach which involves detrending the data as follows. Define

$$
g_i = \frac{y_{i,T} - y_{i,0}}{T} \tag{15}
$$

$$
\tilde{y}_{i,t} = y_{i,t} - g_i t \quad \text{for} \qquad t = 0, 1, ..T \tag{16}
$$

Thus, the linear trend $g_i t$ goes through the first and last observations of the series and $\tilde{y}_{i,t}$ is the deviation from this trend. The next step is to estimate the following equation by OLS,

SP(m)

$$
\Delta \tilde{y}_{i,t} = \pi_i \tilde{y}_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_{i,h} \Delta \tilde{y}_{i,t-h} + \xi_{i,t} \tag{17}
$$

The probability that $\pi_i = 0$ can then be evaluated using the values tabulated by Schmidt and Phillips.

Finally, to test the joint hypothesis that $\pi_i = 0$ for all i ,we use Fisher's λ-statistic which is defined as follows

$$\lambda = -2 \sum_{i=1}^{i=N} \log_e P_i \tag{18}$$

where N is the number of states and P_i is the P-value of the t-statistic for $\hat{\pi}_i$. The λ-statistic has a χ^2 distribution with $2N$ degrees of freedom. This statistic is rarely used in the analysis of panel data, but as Maddala and Wu (1999) show, it is often superior to other tests of collective significance[5]. It is also easy to use and performs comparatively well when errors are correlated across equations.

4.1 Results

On the basis of extensive simulations, Maddala and Wu (1999) argue that with panel data the optimal number of lagged values is $m = 1$. The inclusion of additional lags is not, in their view, justified because of the resulting loss of power. Table 3 presents the results of unit root tests using various data sets and one lagged Δy. In the case of deterministic trends, we present only the results based on SP tests. The results based on ADF tests with trend are very similar and are not reported.

The tests without constants or trends mostly reject the joint hypothesis that all $\pi_i = 0$, and thus favour the existence of "absolute convergence" in employment rates, especially prior to 1974. For the later period, the BLS and CensusA data sets still show quite a high level of stationarity, but for the BEATOT and the BEAWK series absolute convergence appears modest, with significance at around the 5% level .Once constants are included in the regression equations, there is evidence for stationarity only in the CensusA data set. As we shall see, even the latter finding is suspect. It is may be a statistical artifact arising from the large errors in CensusA data.

There is a case for including deterministic trends so as to capture the impact of structural differences across states in the growth rate of women's participation, in the participation of age groups excluded from the working-age population, and in demographic trends (proportions of children or very elderly people). The omission of agricultural employment from the BLS series strengthens the case for including trends when using this data set, and the evidence of stronger absolute convergence in the BLS data is surely due to the fact that only non-agricultural employment is being measured. The move out of agriculture was an extremely important structural shift in the post-war years, as recently emphasized in the regional context by Heim (2000) and Caselli and Johnson (2001). Figure 3 confirms this by showing a systematic tendency for the BLS

[5] One of the first economic applications of Fisher's λ is contained in Rowthorn and Hymer (1971) where it was called the α-test. Appendix 2 of that book presents a simple proof that λ has a χ^2 distribution.

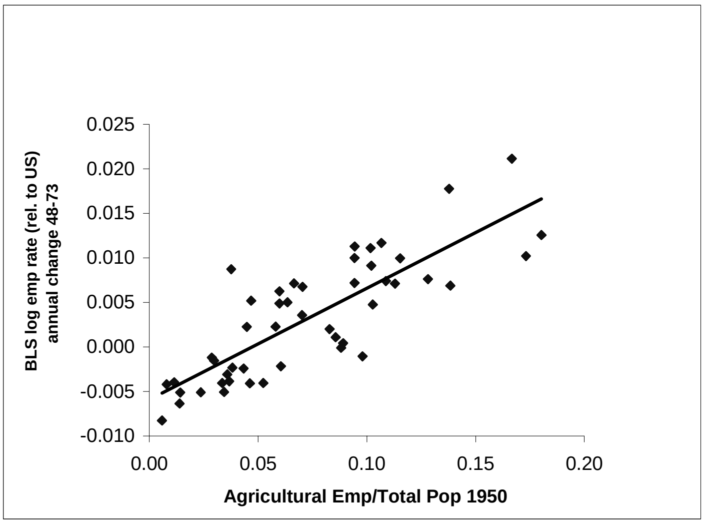

Figure 3: Impact of Agriculture on BLS Employment Trends

employment rate to increase fastest over the period 1948-73 in states where agriculture occupied the highest share of the population at the beginning of the period. This also applies to 1973-98, although the outflow from agriculture was quantitatively much less important than in the earlier period. Note that the growth rates shown in Figure 3 are the same as those used to detrend the BLS data in the Schmidt-Phillips tests for 1948-73.

With trends included in the regression equations, there is evidence of stationarity in the BLS data before 1973. During the later period, the tendency for state level employment rates to revert to their long-run trend after a shock is mostly very weak. The one exception concerns CensusA data, where the null hypothesis that $\pi_i = 0$ for all i is only rejected at the 1.7 percent level. As we shall see, this result is probably a statistical artifact resulting from measurement errors in the data. These findings suggest that, relative to trend, regional employment rates in the US economy were a lot more stable prior to 1973 (which included the most-dynamic post-war decade of the 1960s) than was the case subsequently.

4.2 Robustness

An important, but frequently ignored, issue in the analysis of autoregressive time series is that of measurement error in the dependent variable. This type of error leads to negative autocorrelation in the disturbance terms and exaggerates the appearance of stationarity[6]. A standard way of dealing with measurement

[6] Wansbeek and Meijer (2000) contains a general discussion of measurement error in the dependent variable in auto-regressive time series. It also contains some references to other

error is to use instrumental variables. In the present context, the obvious procedure is to use y_{t-3} and y_{t-4} as instruments in place of y_{t-1} and Δy_{t-1}. However, this procedure is highly questionable since most of our series are much too short for the desirable asymptotic properties of IV estimators to be relevant. The IV estimators of π have non-standard distributions that can only be determined by means bootstrap simulations, and these simulations require assumptions about the error process of a kind which make the use of instrumental variables redundant. To bootstrap the distribution of IV estimators in the present case, we require information about the variance and autocorrelation of measurement errors. With such information it is also possible to bootstrap the distribution of the least squares estimators of π, so there is no advantage to be gained by using instrumental variables. The argument for using lagged values of y as instruments is further undermined by the fact that measurement errors in some of our series (the Census data) are large and highly autocorrelated.

An alternative approach is to approximate the effect of errors by including lagged values of Δy in the regression equations. A problem with this approach is the absence of any clear criterion for choosing the optimal number of lags. Maddala and Wu (1999) suggest using only one lagged Δy, but their simulation analysis is based on the assumption that the disturbance terms are moving averages of independent shocks. In the present case, the disturbance terms combine both real shocks and measurement errors, and they cannot in general be expressed as a simple moving average. The findings of Maddala and Wu may not therefore be very helpful as a guide to the optimal number of lags. Our approach is therefore to model the error process explicitly and to modify the statistical tests accordingly. Such an approach provides an insight into how errors may bias the result and also helps us to identify how many lagged values should be included.

Table 4 compares the last two approaches. It shows the effect of including more than one lag in the regression equations and also the effect of modifying the unit root tests to take account of measurement errors. Statistical distributions in the latter case were derived by bootstrap simulations and where possible were cross-checked with published tabulations. The upper panel of the table refers to CensusA data and is derived from regression equations that include only constants; the lower panel refers to BLS data for the period 1948-73 and is derived from SP tests that include deterministic trends.

The inclusion of additional lags in the case of Census A data leads to a dramatic reduction in significance and the collective P-value rises from 1.7 percent to well over 40 percent. A similar result is observed when only one lag is included in the regression equations, but the statistical test is modified to allow for measurement errors. In deriving the modified test we assume that measurement errors in CensusA data have first order autocorrelation equal to 0.58. This is in line with official estimates for the average US state[7]. Three variants

works on this topic. However, these works are mostly of a theoretical nature, and the issue of measurment errors of the kind which concern us in this paper are largely ignored in applied econometric analsysis.

[7] See above p.3.

of the error process are shown. They differ only with regard to the importance they assign to measurement errors. At one extreme, in accordance with official estimates, it is assumed that measurement errors account for 80 percent of the observed variance of Δy. At the other extreme, it is assumed conservatively that such errors account for only 40 percent of the observed variance. In all three of these variants, the collective P-value is very large and there are virtually no individual states in which the unit root hypothesis is rejected. This is exactly what is observed when additional lags are included in the regression equations.

The inclusion of additional lags in the case of BLS data has mixed effects. With two lagged Δy it becomes impossible to reject the null hypothesis that all $\pi_i = 0$, but with three lags this hypothesis can be rejected at the 10 percent level. The table also gives results for three variants of the error process. For want of information, these variants all assume that measurement errors are serially uncorrelated, but they differ in their assumptions about the magnitude of errors. At one extreme, it as assumed that measurement errors account for 20 percent of the observed variance of Δy. Given the accuracy of the BLS data this may be quite realistic. At the other extreme, it is assumed that measurement errors account for 60 percent of the observed variance. Even with this extreme assumption, the null hypothesis that all $\pi_i = 0$ is rejected at the 5 percent level. With a more realistic assumption about the scale of errors, this hypothesis is rejected at the 1 percent level or better.

The above findings can be summarised as follows. When constants are included in the regression equations there is some indication that CensusA data are stationary. This would point to the existence of "conditional convergence", whereby employment rates are pulled towards fixed levels that differ from state to state. However, this finding is not robust. It disappears when additional lags of the dependent variable are included in the regression equations or if a realistic allowance is made for measurement error. There is also strong evidence that prior to 1973 employment rates were stable around long-run convergence paths. This finding is quite robust.

5 Speed of Adjustment: Pooled Regression

The tests reported above have a major disadvantage. They do not quantify the speed at which regional employment rates respond to state-specific shocks. To deal with this issue in a compact way we utilise regression analysis based on pooled data from all 48 states. Our regression equations are similar to those used above except for the fact that the coefficients of $y_{i,t-1}$ and the various $\Delta y_{i,t-h}$ are uniform across states. There is no presumption that constants and trends are uniform across states. Thus, the ADF tests are based on equations of the form

A*DF(m) no constants*

$$\Delta y_{i,t} = \pi y_{i,t-1} + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{19}$$

A*DF(m) with constants*

$$\Delta y_{i,t} = \pi y_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{20}$$

ADF(m) with linear trends

$$\Delta y_{i,t} = \pi y_{i,t-1} + \alpha_{i0} + \alpha_{i1}\, t + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{21}$$

and the Schmidt Phillips tests are based on regression equations of the form

SP(m)

$$\Delta \tilde{y}_{i,t} \quad = \quad \pi \tilde{y}_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_h \Delta \tilde{y}_{i,t-h} + \xi_{i,t} \tag{22}$$

where $\tilde{y}_{i,t-1}$ is the detrended version of $y_{i,t}$.

In pooled regressions of the above type, least-squares estimators of π have non-standard distributions, about which no information has been published. Extending the method of Andrews (1993), we derive these distributions by means of the simulations which are described in Appendix B. All simulations are based on the standard assumption of serially uncorrelated, normally distributed shocks with uniform variance. They also assume that $\pi \leq 0$. Where appropriate, the simulations make an allowance for measurement error. They also assume that $\pi \leq 0$.

Let $\hat{\pi}_{LS}$ denote the least squares estimator of π derived from an ADF or SP equation of the above type. It is well-known that $\hat{\pi}_{LS}$ has a negative bias. We correct for this bias and derive confidence intervals for π using a procedure that is described in detail in Appendix C. Our procedure, which is based on Andrews (1993), may be summarized as follows. By means of simulation we determine the distribution of $\hat{\pi}_{LS}$ as a function of the true parameter π. From this information we derive a new *median-unbiased* estimator $\hat{\pi}_U$. Provided the assumptions of the model are correct, this estimator has the desirable property that its median is equal to the true parameter π. The use of a median-unbiased estimator is justified in the present context because the parameter space is bounded and it is impossible to have a conventional mean-unbiased estimator at the extreme point $\pi = 0$. Appendix C also explains how to derive confidence intervals for π, although these are not reported here. Instead, we report the more familiar P-values which are derived by Monte Carlo simulation.

As mentioned above, measurement error plays an important role in the analysis of autoregressive time series. It leads to biased estimators and exaggerates the appearance of stationarity. To investigate this issue, we extend the simulations to include measurement errors of varying sizes. The relative importance of errors is summarized by what we term the "error share". This is the proportion of the variance of Δy that is due to measurement error. With CensusA data, which are very inaccurate, we take a conservative approach and assume an error share of between 30% and 60% depending on the context. For reasons outlined above, the true error share is probably even greater. With other data, which are more accurate, we assume an error share of 20% as the standard case. All estimation is based on equations that include just one lagged Δy

5.1 Results

Table 5 presents some results when there are deterministic, linear trends in the data. Estimation is based on the Schmidt-Phillips method, which involves detrending the data for each state individually prior to pooled, least squares regression . The least squares estimates shown in the left hand panel of Table 5 are all large in absolute terms, ranging from -0.118 to -0.325. The median unbiased estimates $\hat{\pi}_U$ shown in the central panel are much smaller in absolute value. Even so, $\hat{\pi}_U$ for the earlier period is still very large in absolute terms and is statistically significant. This finding is quite robust and does not seem to be due to bias arising from measurement error. With an error share of 20%, the median unbiased estimate is -0.15 and is significantly different from zero. Such a coefficient implies a rapid reversion to trend following a shock. With CensusA data $\hat{\pi}_U$ is statistically insignificant but negative in the no error case. This estimate becomes zero when a conservative allowance is made for errors. All other median unbiased estimators are zero and insignificant even in the no error case. These results mostly support the findings of the unit root tests.

Table 6 presents parallel estimates excluding trends but including state-specific constants (fixed effects). Thus the estimates refer to conditional convergence. Even if the case for including trends is rejected, there is still a case for including constants to allow for persistent differences in employment rates across states. Such differences may reflect enduring diversity in such factors as demographic composition, attitudes towards women's work and so forth. As in the previous table, the least squares estimates are all large and negative. However, the median unbiased estimates are mostly zero or very small. If we make reasonable allowances for measurement error (a 60% error share for CensusA data and 20% for the rest) these median unbiased estimators become very small and quite insignificant[8]. Thus, there is no robust evidence for conditional convergence in any of the data sets.

Table 7 presents the last set of results. It gives the estimates when there are

[8] The importance of measurement error in te context of fixed effects Shioji (1997a, b) in his analysys of convergence of per capita incomes. He argues that the estimated speed of convergence from OLS regression with fixed effects is biased upwards by as much as 7 to 15%. Although his methodology is very different, this conclusion is similar to ours.

no constants or trends. This is the case of absolute convergence. The largest estimates are generated by the BLS data for the period prior to 1974. This is not surprising, since these estimates incorporate the convergence effect arising from a dramatic shift of labour out of agriculture in the old farm states during this period (see above). When a reasonable allowance is made for errors, the unbiased estimates $\hat{\pi}_U$ for the later period range from -0.003 to -0.019. Taking an average of the various estimates for the later period, these results suggest the existence of a very gradual convergence process which eliminates interstate differences in employment rates at around 1 percent a year.

The results presented in Tables 5 - 7 are inconsistent with the conventional picture in which regional employment rates in the USA adjust rapidly to state-specific shocks. The conventional least squares estimates suggest very large adjustment coefficients, but our corrections for bias dramatically reduce the magnitude and significance of many of these estimates. Despite such corrections, there remains strong evidence of stability relative to trend in the1950s and 60s. There is also evidence of absolute convergence in all time periods, but the pace is never more than one or two percent a year. Apart from these findings, there is little evidence of any other type of convergence that survives reasonable corrections for bias.

5.2 Other Tests of Robustness

We also tried variants with quadratic time trends, split trends of the Phillips-Perron type and one which replaced the state-specific time trends with a common time trend interacted with the base year level of agricultural employment. In no case were the results substantially different from those reported above. We tried a test based on simulations which assumed a much higher degree of initial dispersion than is implied by the asymptotic variance. This narrows the confidence intervals somewhat but does not affect the median unbiased estimators. We also tried a variant which covers only the years since 1985, that is omitting the crisis years of the 1970s and early 1980s and concentrating on the period after the most dramatic shocks. The effect was mostly to add about -0.01 to -0.02 to the estimates adjustment coefficient, although in some cases this coefficient was unaffected by the alteration in time period. We tried an alternative approach which estimated adjustment coefficients for each state individually, but again the results were not substantially affected.

6 Extensions, Comparisons, Implications

6.1 Census Data and Labour Force Disaggregation

Despite their inaccuracy, Census data have important advantages because they are so detailed. They allow us to decompose changes in the employment rate into changes in participation and unemployment (CensusA) and they provide information on the employment situation of different categories of labour force

(CensusM) It would be interesting to estimate to what extent is the "non-adjustment" of employment rates is concentrated in participation rates rather than in enduring unemployment rate differences. However, measurement errors in Census data are typically very large and vary in extent in unknown ways across categories such as unemployment and participation. This makes it difficult to draw any firm conclusions about differing speeds of adjustment in unemployment and participation rates. Such a conclusion applies even more strongly to the CensusM data which must contain the largest errors of all. Using these data, the least squares estimate for absolute convergence for people in employment with a college education gives an adjustment coefficient of -0.512! Even for those with less than high school education the estimate is -0.209. The size of the measurement errors in these sub-groups is unknown but it is likely to be very large indeed. If we knew that measurment errors were larger for the least educated (not implausible), then despite our ignorance of the true coefficients, we could be fairly sure that the adjustment coefficient is much larger for the most educated than for the least educated (as found by Bound and Holzer (2000) in their cross sectional study). However, given our ignorance about the relative size of errors, such an inference is risky. This does not mean that the Census data are useless. Cross-sectional differences can still show up strongly in these data since likely measurement errors are quite small with respect to large inter-state or inter-category differences. Figures 3 and 4 give examples of the helpful information which can be derived from Census data. However such data are not much use for drawing conclusions about patterns of adjustment, since measurement errors can be very large in relation to the annual changes being analysed.

6.2 Comparison with Earlier Work

Our broad conclusion from the results reported above is that the pace of absolute convergence in the employment rates of US states is very gradual. Pooled estimates suggest that the adjustment coefficient may be around -0.01 for the period since 1970. This conlcusion is consistent with findings Barro and Sala-i-Martin (1991, updated in 1995), who report that the impact of migration of differences in income per head, although highly significant, is "small in an economic sense" with a 1 per cent differential in income per capita raising net migration only enough to boost the area's annual rate of population growth by 0.026 per cent on average over the period 1900-1990 (1995, p. 403). Moreover they report that the impact of income differences on migration were declining over the post war period (falling from 0.044% in the 1950s and 1960s to 0.016%. over the period 1980-89 (1995, table 11.4).

Our results, together with the "mixed results" noted by Greenwood (1997, p. 682), and the strangely neglected work of Barro and Sala-i-Martin, suggest much weaker processes of adjustment than do the famous results of Blanchard and Katz, which we believe greatly exaggerate the self-stabilizing tendencies of regional employment rates in US states. Our analysis of the various US data sets implies that BK's results were strongly influenced by their use of Census data.

The large measurement errors in such data can account for their finding that both unemployment and participation revert rapidly to previous levels following a shock. Moreover, their population series is based on inaccurate employment figures derived from the Census, and it is not surprising that they found it was population (via migration) which absorbed most of the shocks.

Decressin and Fatas (1995) used BK's data for the USA and their data for European countries came predominantly from Labour Force surveys similar to the US Census; whilst Obstfeld and Peri (1998) used an updated version of BK's data series for the USA, together with a variety of sources of unknown accuracy for other countries. In their study of Spanish provinces, Mauro and Splimbergo's (1999) used employment data, disaggregated by education, which derives mainly from the Labour Force Survey. They claim "the data seem very reliable" but it is hard to imagine that such survey material for 50 provinces would not contain measurement errors which are extremely large in relation to the short term variations in employment rates that they analyse. Comparisons of such time series results across countries, for example comparisons of the USA and Europe, cannot be very meaningful unless they are carried out using data of uniform quality in relation to measurement errors, an issue not addressed at all in this literature. Simple cross sections are much less susceptible to these problems of measurement error. It is striking that the OECD found a distinctly weaker correlation between net migration flows and regional unemployment rates in the USA than in a number of European countries including Germany, Italy and Belgium (OECD 2000 table 2.13).

6.3 Implications for Employment Patterns Across States

The estimates of absolute convergence given in table 7 assume that the observed vector of relative employment rates, $y_t = (y_{1,t}, ...y_{N,t})$, evolves according to a stochastic schema of the following type

$$\Delta y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t} \quad (23)$$

$$y_{i,t} = y^*_{i,t} + v_{i,t} \quad (24)$$

where $y^*_{i,t}$, is the true employment rate, $u_{i,t}$ is a real shock, and $v_{i,t}$ is a measurement error. Such a schema has certain macroeconomic implications. Let $R_{t,t+n}$ denote the correlation coefficient between y_t and y_{t+n}, and let SDR_t denote the ratio of sample standard deviations $s_{y_t}/s_{\Delta y_t}$.With appropriate assumptions about the stochastic variables, the following approximate formulae can be derived when π is small[9],

[9] See Appendix D

$$R^2_{t,t+n} = (1+\pi)^{2n} \tag{25}$$

$$SDR_t = \sqrt{\frac{1-\theta}{(-2\pi)} + \frac{\theta}{2(1-\rho)}} \tag{26}$$

where θ is the share of measurement error in the variance of Δy and ρ is the first autocorrelation of the error terms.

6.3.1 Macro-experience

To see how well the above formulae perform in practice, consider the BEATOT data set which covers the period 1969-98. Making a plausible allowance for measurement errors ($\theta = 0.2$ and $\rho = 0$), the median unbiased estimate of π using this data set is equal to -0.012 (table 7). With the formula given in (25), this value of π implies an R^2 of 0.491 between two sets of observations 29 years apart, which is very close to the actual R^2 of 0.476 which is obtained using the BEATOT observations for 1969 and 1998. With $\pi = -0.012$, the formula given in (26) yields a theoretical value of 5.2 for the standard deviation ratio SDR. With the BEATOT data set, the standard deviations of y_t and Δy_t are on average equal to 0.83% and 1.17% respectively. The ratio of these quantities is equal to 6.7 which, although greater than the theoretical ratio of 5.2, is of the same order of magnitude.

Thus, when the appropriate estimate of π is used, the above formulae explain with reasonable accuracy such macroeconomic features of the BEATOT data set as the inter-temporal correlation between state employment rates, or the dispersion of these rates across the nation. As can be seen from appendix table D1, the same is true for most of the other data sets we use. The main exception is the BLS set for 1948-73, for which R^2 and SDR are both approximately three time as large as their theoretical values. This huge disparity may be due to the presence of strong deterministic trends that were ignored in deriving the formulae given in equations (25) and (26) and in deriving the estimates of π shown in table 7.

6.3.2 Micro-experience

For the later period, from around 1970 onwards, a small adjustment coefficient in the range -0.01 to -0.02, or even smaller in absolute magnitude, can explain much of the observed macroeconomic behaviour of relative employment rates. However, at the micro level, a coefficient of this magnitude may be of minor importance, and its medium-term effects on an individual state may be virtually invisible in practice. To illustrate this point table 8 presents some simple simulations which examine how the adjustment coefficient influences the prospects of an individual region following a large shock. The focus is on the true employment rate, $y^*_{i,t}$ which is assumed to evolve according to the difference equation $\Delta y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t}$. The $u_{i,t}$ are independent, normally distributed shocks with zero mean and constant variance σ^2_u. The table shows what may happen following a shock equal to -$2\sigma_u$. If $\pi = 0$ the employment rate meanders all over the place and there is no systematic tendency for the employment rate to recover. What happens in the future depends entirely on random events. There is a 50 percent chance that after ten years the region will be no better off, in relative terms, than it was immediately after the initial shock. On the other hand, there is a 43 percent chance that random events will lead to a complete recovery within this time span. If $\pi = -0.02$, the probability of no recovery at all after ten years falls from 50 to 45 percent, while the probability of complete recovery rises from 43 percent to 46 percent. Such a minor improvement indicates the weakness of systematic tendencies towards recovery when the adjustment coefficient is small. Under these conditions, the medium-term outcome is mainly determined by random events. To produce an appreciable change in medium-term prospects requires a much larger adjustment coefficient.

This is a highly simplified and perhaps misleading picture. It gives the impression that economic recovery is primarily a matter of luck, and that regions hit by negative shocks can do little more than hope for some random acceleration in their otherwise snail's pace of spontaneous recovery. In reality, there are many things that a region can do to accelerate recovery from a shock. It can encourage emigration and discourage immigration, or it can stimulate investment and job creation. To the extent that all regions simultaneously pursue such policies, their effects on relative employment rates tend to cancel out. However, this does not mean that they are a waste of time. If universally applied, certain policies may lead to a general rise in employment in the country as a whole and may benefit all regions simultaneously, even though they leave differentials unchanged.

7 Conclusion

Errors of measurement have to be taken seriously in time series work. Our analysis of the various employment series shows that much of the apparent flexibility of US states following local employment shocks is a statistical illusion arising from measurement errrors. Taking into account the likely magnitude of such errors, we shown that the pace of adjustment following state specific shocks to employment in the US has been rather weak over the past 30 years. It was probably stronger in the 1950s and 1960s than at present. This may be explained by the greater importance at that time of agriculture as a large reserve of labour which could migrate to those parts of the country where the industrial and service demand for labour was buoyant. This reserve had become greatly depleted by the 1970s. The subsequent growth of two-earner families may also have lowered the responsiveness of net migration to employment shocks.

Our findings may have implications for the debate over European Monetary

Union. American experience is often cited by as evidence that a successful monetary union requires a high degree of geographical labour mobility in response to shocks. This belief has helped to fuel opposition to EMU from those who believe that such flexibility is either impractical or undesirable, whilst many of those who support monetary union believe that it must of necessity be accompanied by measures to facilitate migration across regional and national frontiers. Our analysis suggests that migration and other forces behind regional adjustment to shocks are now quite weak in the USA, and they cannot therefore explain the modern success of monetary union in that country.

8 References

Andrews, D. W. K. (1993), "Exactly Median-Unbiased Estimation of First Order Autoregressive / Unit Root Models", *Econometrica,* 61, 139-165.

Barro, R. & X.Sala-I-Martin (1991) "Convergence across States and Regions" *Brookings Papers on Economic Activity* No 1, 107-182

Barro, R. & X.Sala-I-Martin (1995) *Economic Growth*, McGraw Hill New York

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

Baumol, W, Batey Blackman S. and R.Wolff (1989) *Productivity and American Leadership,* Cambridge Mass, MIT Press

Bernat, G. (2001) "Convergence in State per capita personal income 1950-99" *Survey of Current Business* June.

Blanchard, O. & Katz, L. (1992) "Regional Evolutions" *Brookings Papers on Economic Activity* 1-61

Bound J. & Holzer, H. (2000)"Demand Shifts, Population Adjustments, and Labour Market Outcomes During the 1980s" *Journal of Labour Economics,*18.1, 20-54

Bureau of Labor Statistics (1997) BLS H*andbook of Methods*
*(*http:/stats.bls.gov/opub)

Bureau of Labor Statistics (1999) *Geographic Profile of Employment and Unemployment*
*(*http:/stats.bls.gov/opub/gp/pdf/gp99_apndb.pdf)

Caselli, F. & W.Coleman (2001) "The US Structural Transformation and Regional Convergence:A Reinterpretation" *Journal of Political Economy* 109.3, 584-616.

Decressin, J. & A.Fatas (1995) "Regional Labour Market Dynamics in Europe" *European Economic Review* 39 1627-1655

Fuller, W. A. (1996), *Introduction to Statistical Time Series*, 2nd ed, Edinburgh, Oliver and Boyd

Greenwood, M. (1997) "Internal Migration in Developed Countries" in M. Rosenwiig & O.Stark (eds) *Handbook of Population and*
Family Economics, Amsterdam, Elsevier, 647-720.

Heim, C.(2000) "Structural Changes:Regional and Urban" in Engerman, S. & R.Gallman (eds) *The Cambridge Economic History of the USA Volume III* The Twentieth Century Cambridge University Press, Cambridge.

Kiviet , J. F. (1995), "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models", *Journal of Econometrics*, 68, 53-78.

Maddala, G. S. and S. Wu (1999), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", *Oxford Bulletin of Economics and Statistics,* Special Issue, 631-652

Mauro, P. & A. Spilimbergo (1999) "How do the Skilled and Unskilled Respond to Regional Shocks" *IMF Staff Papers 46.1 1-17.*

Neumann, G. & R.Topel (1990) "Employment Risk, Diversification and Unemployment" *Quarterly Journal of Economics* CVI 1341-1365

Obstfeld, M. & G.Peri (1998) "Regional non-adjustment and Fiscal Policy" *Economic Policy* 26, 205-260

OCED (2000) Employment Outlook, Paris.

Perron, P.(1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica*, 57, 1361-1401.

Rowthorn, R.and S.Hymer (1971), *International Big Business 1957-67*, Cambridge, Cambridge University Press

Schmidt, P. and P.C.B. Phillips (1992), "LM Tests for a Unit Root in the Presence of Deterministic Trends," *Oxford Bulletin of Economics and Statistics*, 54, 257-287

Shioji, E. (1997a), "Convergence in Panel Data: Evidence from the Skipping Estimation", mimeo, (author's e-mail shioji@econ.ynu.ac.jp)

Shioji, E. (1997b), "It's still 2%: Evidecne on Convergence from 116 Years of the US States Panel Data", mimeo, (author's e-mail shioji@econ.ynu.ac.jp)

Wansbeek, T and E. Meijer (2000), *Measurement Error and Latent Variables in Econometrics*, Amsterdam, Elsevier Science.

9 Appendix A: Data Sources

Non-agricultural Employees (BLS): kindly supplied by Larry Katz; updated from Statistical Abstract 1999 Table 686 and earlier years.

Population (total): kindly supplied by Larry Katz; updated from Statistical Abstract 1999 Table 26.

Population 18-64: calculated by authors from Bureau of the Census website containing population estimates by selected

age-group(www.census.gov/popest/archives/1990.php#state)

Total Employment (BEA): Bureau of Economic Analysis Regional Economic Information System website (www.bea.doc.gov/bea/regional/spi)

Farm employment in 1950: Statistical Abstract 1953 Table 714

1974 Farm Employment and Wage Rates 1910-1990 US Dept of Agriculture National Agricultural Statistics Service Statistical Bulletin 822, March 1991

CensusA data for employment, unemployment and population (over 14) from BLS LAUS website (www.bls.gov.lau/slaa_7000prn). Details of methods of calculations and measurement errors are given BLS (1997 Ch 2) and BLS (1999).

CensusM data for employment, divided by gender and educational group, and population, aged 25-64 kindly supplied by John Schmitt from the Economic Policy Institute, Washington.

10 Appendix B: Description of the Simulations

This appendix draws heavily on Andrews (1993). It describes the simulations used to derive the statistical properties of the estimators used in the pooled regressions. It begins by establishing a useful invariance principle that greatly simplifies the task of simulation.

10.1 An Invariance Principle

Consider an autoregressive process defined by the following schema

$$\Delta y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t} \tag{27}$$

$$y_{i,t} = y^*_{i,t} + \sum_{j=0}^{n} \alpha^*_{ij} t^j + v_{i,t} \tag{28}$$

where $i = 1, 2, ..N$ and $t = 0, 1, ..T$. The $y^*_{i,t}$ are unobserved variables which are subject to random shocks $u_{i,t}$ and the $y_{i,t}$ are observed variables which are subject to measurement error $v_{i,t}$. The above formulation allows for the existence of deterministic trends. By elimination we obtain the following system of difference equations

$$\Delta y_{i,t} = \pi y_{i,t-1} + \sum_{j=0}^{n} \alpha_{i,j} t^j + w_{i,t} \tag{29}$$

where the α_{ij} are functions of π and the α^*_{ij} , and $w_{i,t} = u_{i,t} + v_{i,t} - (1+\pi)v_{i,t-1}$.

Suppose also that π is estimated by pooled least squares regression of $\Delta y_{i,t}$ on the variables on the right hand-side of (29) together with the first m lagged values of Δy. This yields the estimated equations

$$\Delta y_{i,t} = \hat{\pi} y_{i,t-1} + \sum_{h=1}^{m} \hat{\beta}_h \Delta y_{i,t-h} + \sum_{j=0}^{n} \hat{a}_{ij} t^j + e_{i,t} \tag{30}$$

where the $e_{i,t}$ are the least squares residuals. A more convenient, and equivalent, procedure is to perform a pooled least squares regression of $y_{i,t}$ on the variables on the right hand-side of (29) plus the variables $y_{i,t-2}, ..y_{i,t-m-1}$. This yields the following equations

$$y_{i,t} = \sum_{h=1}^{m+1} \hat{\mu}_h y_{i,t-h} + \sum_{j=0}^{n} \hat{a}_{ij} t^j + e_{i,t} \tag{31}$$

where $\hat{\pi} = -1 + \sum \hat{\mu}_h$.

For each value of h, perform a separate, pooled, least squares regression of $y_{i,t-h}$ on the exogenous variables in (31). Let $r_{i,t}^{(h)}$ be the residual thereby obtained . By definition,

$$r_{i,t}^{(h)} = y_{i,t-h} - \sum_{j=0}^{n} \hat{a}_{ij}^{(h)} t^j \tag{32}$$

where the hats denote estimated coefficients. Next, perform a pooled, least squares regression of $r_{i,t}^{(0)}$ on $r_{i,t}^{(1)}$, $r_{i,t}^{(2)}$, ...$r_{i,t}^{(m+1)}$. It is well-known that $\hat{\mu}_h$ is equal to the coefficient of $r_{i,t}^{(h)}$ in this regression. Following Andrews (1993), it can be shown that the various $r_{i,t}^{(h)}$, and hence $\hat{\mu}_h$ and $\hat{\pi}$ $(= -1 + \sum \hat{\mu}_h)$, are independent of the original parameters α_{ij}^*. These estimates depend only on π and the various $u_{i,t}$, $v_{i,t}$ and $y_{i,0}^*$. This result can be established as follows.

Using (28), we can express $y_{i,t-h}$ as follows

$$y_{i,t-h} = \sum_{j=0}^{n} a_{ij}^{(h)} t^j + y_{i,t-h}^* + v_{i,t-h} \tag{33}$$

where the $a_{ij}^{(h)}$ are functions of the original coefficients α_{ij}^*. Morevoer, (27) implies that

$$y_{i,t-h}^* = \sum_{g=0}^{t-h-1} (1+\pi)^g u_{i,t-g} + (1+\pi)^{t-h} y_{i,o}^* \tag{34}$$

Combining (32), (33) and (34) yields

$$r_{i,t}^{(h)} = \sum_{j=0}^{n} (a_{ij}^{(h)} - \hat{a}_{ij}^{(h)}) t^j + \sum_{g=0}^{t-h-1} (1+\pi)^g u_{i,t-g} + (1+\pi)^{t-h} y_{i,o}^* + v_{i,t-h} \tag{35}$$

Since (32) is derived by least squares regression,

$$\sum_{t=m+1}^{T} r_{i,t}^{(h)} t^k = 0 \tag{36}$$

Combining (35) and (36) yields, for $i = 1$ to N and $k = 0$ to n,

$$\sum_{t=m+1}^{T} \left[\sum_{j=0}^{n} (a_{ij}^{(h)} - \hat{a}_{ij}^{(h)}) t^j + \sum_{g=0}^{t-h-1} (1+\pi)^g u_{i,t-g} + (1+\pi)^{t-h} y_{i,o}^* + v_{i,t-h} \right] t^k = 0 \tag{37}$$

The above system of $(n+1)N$ equations determines the $a_{ij}^{(h)} - \hat{a}_{ij}^{(h)}$ as functions of π and the various $y_{i,0}^*$, $u_{i,t}$ and $v_{i,t}$. Substituting in (35) it follows that the residuals $r_{i,t}^{(h)}$, and hence $\hat{\mu}_h$ and $\hat{\pi}$, are also functions of π and the various

$y^*_{i,0}, u_{i,t}$ and $v_{i,t}$. They do not depend on the original coefficients α^*_{ij}. This is the result we set out to prove.

This property of the least squares estimator is very useful for simulation purposes. It implies that the distribution of this estimator is independent of the values taken by the parameters α^*_{ij} in the underlying data generating process.Thus, without loss of generality, simulation analysis can assume that such parameters are all zero. This is a great simplification since it eliminates a potentially large number of nuisance parameters.

10.2 The Schmidt-Phillips Method

The above argument can be extended to include the method proposed by Schmidt and Phillips (1992) for detrending data in the case of linear trends. Consider the following stochastic schema.

$$\Delta y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t} \tag{38}$$

$$y_{i,t} = y^*_{i,t} + \alpha^*_i + \beta^*_i t + +v_{i,t} \tag{39}$$

Let.

$$\tilde{y}_{i,t} = y_{i,t} - y_{i,0} - \frac{(y_{i,T} - y_{i,0}) \times t}{T} \tag{40}$$

Using (39) it follows that

$$\tilde{y}_{i,t} = y^*_{i,t} - y^*_{i,0} - \frac{(y^*_{i,T} - y^*_{i,0}) \times t}{T} \tag{41}$$

$$+v_{i,t} - v_{i,0} - \frac{(v_{i,T} - v_{i,0}) \times t}{T} \tag{42}$$

Since $y^*_{i,t} = u_{i,t} + (1+\pi)u_{i,t-1} + ...(1+\pi)^t y^*_{i,o}$ it follows that $\tilde{y}_{i,t}$ is determined exclusively by π and the various $u_{i,t}$, $v_{i,t}$and $y^*_{0,t}$. The estimators derived by the Schmidt -Phillips method are therefore statistically independent of the various constants and trends in (39). Thus, without loss of generality, the simulations can assume that the parameters α^*_{ij} in the data generating process are all zero.

10.3 Simulation

For most of the tests used in this paper, especially those relating to pooled regressions, there is no published information regarding the statistical distribution of the estimators concerned. Such distributions were therefore determined by Monte Carlo simulation. The required simulations were done in Matlab12 using programmes written by one of the authors. As a cross-check, these programmes were also used to investigate the statistical properties of estimators for which published information is available. The results were compared to those in the standard tables and results were in all cases found to be virtually identical (Andrews ,1993; Fuller, 1996; Perron,1989; Schmidt and Phillips,1992).

Our simulations all have a commmon structure. They consist of a basic procedure which is repeated a very large number of times. This procedure consists of two steps. In the first step a data-generating process is used to create a set of "observations" $y_{i,t}$. In the next step, a regression equation of suitable form is fitted to these observations. This yields an estimate of the adjustment coefficient π. By repeating the same procedure a large number of times, the distribution of the estimator concerned can be approximated with considerable accuracy. In simulations using individual time series ($N = 1$) the number of repetitions is 100,000. For those using pooled time series ($N = 48$) the number of repetitions is 10,000.

The basic features of the data generating process and regression equations used for the pooled simulations are as follows.

The Data Generation Process

Data for each simulation is generated by a schema of the following type

$$\Delta y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t} \tag{43}$$

$$y_{i,t} = y^*_{i,t} + v_{i,t} \tag{44}$$

where $y^*_{i,t}$ is the underlying variable, $u_{i,t}$ is a random shock and $v_{i,t}$ is a measurement error.The use of such a simple schema can justified by appeal to the invariance principles established above. It is assumed that all u and v variables are independent of each other and of preceding values of the y^*variables; all $u_{i,t}$ have the same distribution $N(0, \sigma_u^2)$; and all $v_{i,t}$ have the same distribution $N(0, \sigma_v^2)$ and have first order autocorrelation ρ. Thus, $u_{i,t}$ and $v_{i,t}$ have distribution $N(0, 1)$ and $N(0, \mu^2)$ respectively, where $\mu = \sigma_v/\sigma_u$.Our assumptions concerning the initial values $y^*_{i,0}$ depend on the value of π. If $\pi < 0$ it is assumed that all $y^*_{i,0}$ have distribution $N(0, \sigma_u^2/(1-\phi^2))$ where $\phi = (1+\pi)$. This ensures that the y^*_i and y_i series are all stationary. Without loss of generality, it is also assumed that $\sigma_u^2 = 1$. The above schema implies that

$$\Delta y_{i,t} = \pi y_{i,t-1} + w_{i,t} \tag{45}$$

where $w_{i,t} = u_{i,t} + v_{i,t} - (1 + \pi)v_{i,t-1}$.

The share of measurement errors in the variance of Δy can be determined as follows. From (43)

$$\sigma^2_{\Delta y^*} = \pi^2 \sigma^2_{y^*} + \sigma_u^2 \tag{46}$$

$$\sigma^2_{y^*} = \frac{\sigma_u^2}{1-(1+\pi)^2} \tag{47}$$

Thus,

$$\sigma^2_{\Delta y^*} = \frac{\pi^2 \sigma_u^2}{1-(1+\pi)^2} + \sigma_u^2$$

$$= \frac{2\sigma_u^2}{(2+\pi)} \tag{48}$$

Since $\Delta v_{i,t} = v_{i,t} - v_{i,t-1}$ it follows that

$$\sigma^2_{\Delta v} = 2(1-\rho)\sigma_v^2 \tag{49}$$

and hence from (44)

$$\begin{aligned}\sigma^2_{\Delta y} &= \sigma^2_{\Delta y^*} + \sigma^2_{\Delta v} \\ &= \frac{2\sigma^2_u}{(2+\pi)} + 2(1-\rho)\sigma^2_v \end{aligned} \tag{50}$$

The second term on the right hand side of this equation is the contribution of measurement errors to the variance of Δy. Hence the share of errors in the variance of Δy is given by

$$\begin{aligned}f(\pi) &= \frac{\sigma^2_{\Delta v}}{\sigma^2_{\Delta y}} \\ &= \frac{(2+\pi)(1-\rho)\sigma^2_v}{\sigma^2_u + (2+\pi)(1-\rho)\sigma^2_{\Delta v}} \\ &= \frac{(2+\pi)(1-\rho)\mu^2}{1+(2+\pi)(1-\rho)\mu^2}\end{aligned} \tag{51}$$

For small π this is approximately equal to θ where

$$\theta = \frac{2(1-\rho)\mu^2}{1+2(1-\rho)\mu^2} \tag{52}$$

This is the parameter which is used in the text to denote the relative importance of measurement error. Inverting the above equation we get the following expression for μ

$$\mu = \sqrt{\frac{\theta}{2(1-\rho)(1-\theta)}} \tag{53}$$

The above assumptions imply that the data-generating process is specified by means of just three parameters: π, θ and ρ.

Regression Equations

The regression equations used to estimate the parameter π are as follows.

(i) A*DF(m) no constants*

$$\Delta y_{i,t} = \pi y_{i,t-1} + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{54}$$

(ii) A*DF(m) with constants*

$$\Delta y_{i,t} = \pi y_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{55}$$

(ii) *ADF(m) with linear trends*

$$\Delta y_{i,t} = \pi y_{i,t-1} + \alpha_{i0} + \alpha_{i1}\, t + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{56}$$

(iii) A*DF(m) with quadratic trends*

$$\Delta y_{i,t} = \pi y_{i,t-1} + \alpha_{i0} + \alpha_{i1}\, t + \alpha_{i2}\, t^2 + \sum_{h=1}^{h=m} \beta_h \Delta y_{i,t-h} + \xi_{i,t} \tag{57}$$

(iv) SP(m)

$$\begin{aligned} \Delta \tilde{y}_{i,t} &= \pi \tilde{y}_{i,t-1} + \alpha_{i0} + \sum_{h=1}^{h=m} \beta_h \Delta \tilde{y}_{i,t-h} + \xi_{i,t} \qquad (58) \\ \text{where } \tilde{y}_{i,t} &= y_{i,t} - y_{i,0} - \frac{(y_{i,T} - y_{i,0}) \times t}{T} \end{aligned}$$

The above equations were all estimated using pooled data from 48 states.

11 Appendix C: Median Unbiased Estimators and Confidence Intervals

This appendix explains how median unbiased estimators and confidence intervals are derived. It draws heavily on the ideas presented in Andrews (1993).

Consider a stochastic process which, in part, is specified by a certain parameter π that can take any value within the range $(-2, 0]$. Let $\hat{\pi}$ be the least squares, or some other, estimator of this parameter. For any given value of π, let P_π be the probability distribution of $\hat{\pi}$. For $p \in [0, 1]$ and π in the allowable range, let x be such that

$$\begin{aligned} P_\pi(\hat{\pi} &\geq x) \geq 1 - p \\ P_\pi(\hat{\pi} &\leq x) \geq p \end{aligned} \tag{59}$$

This definition allows for a multiplicity of solutions, but in the context of this paper we can assume that x is unique. Let $F_p(\pi)$ denote this unique value. Note that $F_p(\pi)$ is the pth-tile of the distribution of $\hat{\pi}$ for the given π.

In the context of this paper it is always the case that $F_p^{'} > 0$ over the range of interest. This function is vertical at $\pi = 0$, but elsewhere in the range of interest the derivative is finite. Under these conditions $F_p(.)$ is invertible. We can thus transform the orginal least squares estimator of π to derive the new estimator

$$\hat{\pi}_U = F_{0.5}^{-1}(\hat{\pi}) \tag{60}$$

Equations (59) imply that

$$\begin{aligned} P_\pi(\ \hat{\pi} &\geq F_{0.5}(\pi)\ ) \geq 0.5 \\ P_\pi(\ \hat{\pi} &\leq F_{0.5}(\pi)\ ) \geq 0.5 \end{aligned}$$

Since $F_{0.5}^{'} > 0$ over the range of interest it follows that

$$P_\pi(\ F_{0.5}^{-1}(\hat{\pi}) \geq \pi\ ) \geq 0.5 \tag{61}$$

$$P_\pi(\ F_{0.5}^{-1}(\hat{\pi}) \leq \pi\ ) \geq 0.5 \tag{62}$$

and hence from (60)

$$P_\pi(\ \hat{\pi}_U \geq \pi\ ) \geq 0.5 \tag{63}$$

$$P_\pi(\ \hat{\pi}_U \leq \pi\ ) \geq 0.5 \tag{64}$$

These relationships indicate that the estimator $\hat{\pi}_U$ is a stochastic variable whose median is equal to the true value π. For this reason we refer to it as the Median Unbiased Estimator.

To derive confidence intervals let

$$\begin{aligned} \hat{\pi}_{\min} &= F_{0.95}^{-1}(\hat{\pi}) \\ \hat{\pi}_{\max} &= F_{0.05}^{-1}(\hat{\pi}) \end{aligned} \tag{65}$$

Since $F_p^{'} > 0$ over the range of interest it follows that $\hat{\pi}_{\min} \leq \hat{\pi}_U \leq \hat{\pi}_{\max}$. Equations (59) imply that

$$\begin{aligned} P_\pi(\ \hat{\pi} &\leq F_{0.05}(\pi)\ ) \geq 0.05 \\ P_\pi(\ \hat{\pi} &> F_{0.95}(\pi)\ ) = 1 - P_\pi(\ \hat{\pi} \leq F_{0.95}(\pi)\ ) \leq 0.05 \end{aligned}$$

Hence

$$\begin{aligned} P_\pi(\ F_{0.05}^{-1}(\hat{\pi}) &\leq \pi\ ) \geq 0.05 \\ P_\pi(\ F_{0.95}^{-1}(\hat{\pi}) &> \pi\ ) \leq 0.05 \end{aligned}$$

and thus from (65)

$$P_\pi(\hat{\pi}_{\max} \leq \pi) \geq 0.05 \tag{66}$$

$$P_\pi(\hat{\pi}_{\min} < \pi) \leq 0.05 \tag{67}$$

The existence of inequalities on the right-hand side of these expressions derives from the fact that π may achieve its upper bound. The confidence limits used in preparing this paper were equal to $\hat{\pi}_{\min}$ and $\hat{\pi}_{\max}$, although for reasons of clarity we do present our calculations in this version of the paper.

Because π is bounded, the above confidence interval may be degenerate. For example, there are some occasions in the present paper when $\hat{\pi}_{\min} < \hat{\pi}_U <$

$\hat{\pi}_{\max} = 0$. In this case, the median-unbiased estimate is negative but is not significantly different from zero at the 5% level. There are also occasions when $\hat{\pi}_{\min} = \hat{\pi}_U = \hat{\pi}_{\max} = 0$. The confidence interval thus reduces to the single point $\pi = 0$.

The P-values which are given in the text are derived as follows. Let $\hat{\pi}^{actual}$ be the parameter estimate that is obtained from a regression using some actual data sample, and let $\hat{\pi}_U^{actual}$ be the corresponding value of the median unbiased estimate. Then

$$\begin{aligned} \text{P-value} &= P_0(\ \hat{\pi} \leq \hat{\pi}^{actual}\ ) && (68) \\ &= P_0(\ \hat{\pi}_U \leq \hat{\pi}_U^{actual}\ ) && (69) \end{aligned}$$

Under the null hypothesis that $\pi = 0$, the P-value is the probability of obtaining an estimate $\hat{\pi}$ that is at least as negative as the actual estimate $\hat{\pi}^{actual}$. This is equal to the the probability of obtaining a median unbiased estimate $\hat{\pi}_U$ that is at least as negative as the actual estimate $\hat{\pi}_U^{actual}$. Thus, if $\hat{\pi}^{actual}$and $\hat{\pi}_U^{actual}$ are large and negative, the P-value will be very small and the null hypothesis that $\pi = 0$ will therefore be rejected.

12 Appendix D: Approximate Formulae for Certain Macroeconomic Statistics

Consider a vector $y_t = (y_{1,t}, ...y_{N,t})$ whose components develop according to the stochastic difference equations

$$\begin{aligned} \Delta y_{i,t}^* &= \pi y_{i,t-1}^* + u_{i,t} && (70) \\ y_{i,t} &= a_i + y_{i,t}^* + v_{i,t} && (71) \end{aligned}$$

where $-2 < \pi < 0$. The $a_i, , u_{i,t}, v_{i,t}$, and $y_{i,0}^*$are random variables with zero means, and with variances σ_a^2, σ_u^2, σ_v^2 and $\sigma_u^2/(1-(1+\pi)^2)$ respectively. With one exception, these random variables are independent of each other and of random variables for previous years. The one exception concerns the $v_{i,t,}$ which have nth order autocorrelation ρ_n.

The above assumptions imply that $y_{i,t}^*$, $\Delta y_{i,t}^*$ and $\Delta v_{i,t}$ all have mean zero.

Their variances are as follows

$$\sigma_{y^*}^2 = \frac{\sigma_u^2}{1-(1+\pi)^2} = \frac{-\sigma_u^2}{\pi(2+\pi)} \tag{72}$$

$$\sigma_{\Delta y^*}^2 = \pi^2\sigma_{y^*}^2 + \sigma_u^2 = \frac{2\sigma_u^2}{(2+\pi)} \tag{73}$$

$$\sigma_{\Delta v}^2 = 2(1-\rho)\sigma_v^2 \tag{74}$$

where $\rho = \rho_1$. The last equation derives from the fact that $\Delta v_{i,t} = v_{i,t} - v_{i,t-1}$.

To determine a useful covariance, note that $y_{it}^* = (1+\pi)y_{it-1}^* + u_{it}$ and hence

$$y_{i,t+n}^* = (1+\pi)^n y_{i,t}^* + u_{i,t+n} + .. + (1+\pi)^n u_{i,t}$$

Thus, all pairs $y_{i,t}^*$ and $y_{i,t+n}^*$ have the same covariance given by

$$\begin{aligned} Cov(y_t^*, y_{t+n}^*) &= Cov(y_{it}^*, y_{it+n}^*) \\ &= (1+\pi)^n Cov(y_{it}^*, y_{it}^*) + Cov(y_{it}^*, u_{it+n}) + .. \\ &= (1+\pi)^n \sigma_{y^*}^2 \end{aligned} \tag{75}$$

Moreover, all pairs $v_{i,t}$ and $v_{i,t+n}$ have the same covariance given by

$$\begin{aligned} Cov(v_t, v_{t+n}) &= Cov(v_{i,t}, v_{i,t+n}) \\ &= \rho_n \sigma_v^2 \end{aligned} \tag{76}$$

The cross-section, sample variance of y at time t is as follows

$$\begin{aligned} s_{y_t}^2 &= \frac{\sum_{i=1}^N (y_{i,t} - \bar{y}_t)^2}{N} \\ &= \frac{1}{N}\left(\sum_{i=1}^N (y_{i,t}^* - \bar{y}_t^* + a_i - \bar{a} + v_{i,t} - \bar{v}_t)\right)^2 \\ &= \frac{\sum_{i=1}^N (y_{i,t}^* - \bar{y}_t^*)^2}{N} + \frac{\sum_{i=1}^N (a_i - \bar{a})^2}{N} + \frac{\sum_{i=1}^N (v_{i,t} - \bar{v}_t)^2}{N} \\ &\quad + \frac{1}{N} \times (\text{cross products}) \end{aligned} \tag{77}$$

Likewise,

$$
\begin{aligned}
s^2_{\Delta y_t} &= \frac{\sum_{i=1}^N (\Delta y_{i,t} - \Delta \bar{y}_t)^2}{N} \\
&= \frac{1}{N}\left(\sum_{i=1}^N (\Delta y^*_{i,t} - \Delta \bar{y}^*_t + \Delta a_i - \Delta \bar{a} + \Delta v_{i,t} - \Delta \bar{v}_t)\right)^2 \\
&= \frac{1}{N}\left(\sum_{i=1}^N (\Delta y^*_{i,t} - \Delta \bar{y}^*_t)^2 + \sum_{i=1}^N (\Delta a_i - \Delta \bar{a})^2 + \sum_{i=1}^N (\Delta v_{i,t} - \Delta \bar{v}_t)^2\right) \\
&\quad + \frac{1}{N} \times (\text{cross products}) \\
&= \frac{\sum_{i=1}^N (\Delta y^*_{i,t} - \Delta \bar{y}^*_t)^2}{N} + \frac{\sum_{i=1}^N (\Delta v_{i,t} - \Delta \bar{v}_t)^2}{N} \\
&\quad + \frac{1}{N} \times (\text{cross products}) \qquad (78)
\end{aligned}
$$

Note that the Δa_i and $\Delta \bar{a}$ drop out in the above equation since they are all zero. Since $y_{i,t} - \bar{y}_t = y^*_{i,t} - \bar{y}^*_t + a_i - \bar{a} + v_{i,t} - \bar{v}_t$, it follows that

$$
\begin{aligned}
&\frac{\sum_{i=1}^N (y_{i,t} - \bar{y}_t)(y_{i,t+n} - \bar{y}_{t+n})}{N} \\
= \; & \frac{\sum_{i=1}^N (y^*_{i,t} - \bar{y}^*_t)(y^*_{i,t+n} - \bar{y}^*_{t+n})}{N} + \frac{\sum_{i=1}^N (a_i - \bar{a})^2}{N} \\
& + \frac{\sum_{i=1}^N (v_{i,t} - \bar{v}_t)(v_{i,t+n} - \bar{v}_{t+n})}{N} + \frac{1}{N} \times (\text{cross products}) \qquad (79)
\end{aligned}
$$

Our stochastic assumptions allow us to ignore the cross products in the above equations when taking limits. Thus,

$$
P \lim_{N\to\infty} \frac{\sum_{i=1}^N (y_{i,t} - \bar{y}_t)^2}{N} = \sigma^2_{y^*} + \sigma^2_a + \sigma^2_v \qquad (80)
$$

$$
P \lim_{N\to\infty} \frac{\sum_{i=1}^N (\Delta y_{i,t} - \Delta \bar{y}_t)^2}{N} = \sigma^2_{\Delta y^*} + \sigma^2_{\Delta v} \qquad (81)
$$

$$
\begin{aligned}
P \lim_{N\to\infty} \frac{\sum_{i=1}^N (y_{i,t} - \bar{y}_t)(y_{i,t+n} - \bar{y}_{t+s})}{N} &= Cov(y^*_t, y^*_{t+n}) + \sigma^2_a + Cov(v_t, v_{t+n}) \\
&= (1+\pi)^n \sigma^2_{y^*} + \sigma^2_a + \rho_n \sigma^2_v \qquad (82)
\end{aligned}
$$

The correlation coefficient between the two cross-sections y_t and y_{t+n} is defined as follows

$$R_{t,t+n} = \frac{\sum_{i=1}^{N}(y_{i,t} - \bar{y}_t)(y_{i,t+n} - \bar{y}_{t+n})}{\left(\sum_{i=1}^{N}(y_{i,t} - \bar{y}_t)^2\right)^{0.5}\left(\sum_{i=1}^{N}(y_{i,t+n} - \bar{y}_{t+n})^2\right)^{0.5}} \tag{83}$$

Dividing numerator and denominator by $N-1$, taking probability limits, and using (80) and (82), we derive the following expression

$$P\lim_{N\to\infty} R_{t,t+n} = \frac{(1+\pi)^n\sigma_{y^*}^2 + \sigma_a^2 + \rho_n\sigma_v^2}{\sigma_{y^*}^2 + \sigma_a^2 + \sigma_v^2} \tag{84}$$

Suppose that $\rho_n \to 0$ as $n \to \infty$. Then,

$$\lim_{n\to\infty}(P\lim_{N\to\infty} R_{t,t+n}) = \frac{\sigma_a^2}{\sigma_{y^*}^2 + \sigma_a^2 + \sigma_v^2} \tag{85}$$

Thus, if $\sigma_a^2 > 0$, the correlation coefficient is likely to stabilise at some positive value. This fact provides the basis for a visual test of conditional convergence.

In the case of absolute convergence all a_i are identical and hence $\sigma_a^2 = 0$. Suppose also that $\rho_n = 0$ for sufficiently large n.. Then, for large n, using equation (72) we obtain

$$\begin{aligned} P\lim_{N\to\infty} R_{t,t+n} &= \frac{(1+\pi)^n\sigma_{y^*}^2}{\sigma_{y^*}^2 + \sigma_v^2} \\ &= \frac{(1+\pi)^n\sigma_u^2}{\sigma_u^2 - \pi(2+\pi)\sigma_v^2} \end{aligned} \tag{86}$$

If π is small, the above equation yields the following approximation,

$$P\lim_{N\to\infty} R_{t,t+n} = (1+\pi)^n \tag{87}$$

This approximation is, of course, only valid in the case of absolute convergence ($\sigma_a^2 = 0$).

Define the variance ratio VR_t as follows

$$VR_t = \frac{s_{y_t}^2}{s_{\Delta y_t}^2} = \frac{\left(\frac{\sum_{i=1}^{N}(y_{i,t}-\bar{y}_t)^2}{N}\right)}{\left(\frac{\sum_{i=1}^{N}(\Delta y_{i,t}-\Delta\bar{y}_t)^2}{N}\right)} \tag{88}$$

The numerator and denominator are, respectively, the sample variances of y_t and Δy_t. Taking the limit as N tends to infinity,and using (80) and (81), yields

$$\begin{aligned}
P\lim_{N\to\infty} VR_t &= \frac{\sigma_{y^*}^2+\sigma_a^2+\sigma_v^2}{\sigma_{\Delta y^*}^2+\sigma_{\Delta v}^2} \\
&= \frac{\sigma_{y^*}^2+\sigma_a^2+\sigma_v^2}{\sigma_{\Delta y^*}^2+2(1-\rho)\sigma_v^2} \\
&= \frac{\frac{\sigma_u^2}{(-\pi)(2+\pi)}+\sigma_a^2+\sigma_v^2}{\frac{2\sigma_u^2}{(2+\pi)}+2(1-\rho)\sigma_v^2} \\
&= \frac{\sigma_u^2-\pi(2+\pi)(\sigma_a^2+\sigma_v^2)}{(-2\pi)\left(\sigma_u^2+(1-\rho)(2+\pi)\sigma_v^2\right)} \qquad (89)
\end{aligned}$$

For small π this yields the following approximation

$$\begin{aligned}
P\lim_{N\to\infty} VR_t &= \frac{\sigma_u^2}{(-2\pi)\left(\sigma_u^2+2(1-\rho)\sigma_v^2\right)}\left(1+\frac{(-2\pi)\sigma_a^2}{\sigma_u^2}+\frac{(-2\pi)\sigma_v^2}{\sigma_u^2}\right) \\
&= (1-\theta)\left(\frac{1}{(-2\pi)}+\frac{\sigma_a^2}{\sigma_u^2}+\frac{\theta}{2(1-\rho)(1-\theta)}\right) \\
&= \frac{1-\theta}{(-2\pi)}+\frac{(1-\theta)\sigma_a^2}{\sigma_u^2}+\frac{\theta}{2(1-\rho)} \qquad (90)
\end{aligned}$$

where θ is the share of measurement error in $\sigma_{\Delta y}^2$. It is shown in Appendix B that

$$\theta = \frac{2(1-\rho)\sigma_u^2}{\sigma_u^2+2(1-\rho)\sigma_v^2} \qquad (91)$$

In the case of absolute convergence $\sigma_a^2=0$, and the above approximation reduces to the following simple formula

$$P\lim_{N\to\infty} VR_t = \frac{1-\theta}{(-2\pi)}+\frac{\theta}{2(1-\rho)} \qquad (92)$$

The above formula may be more usefully expressed in terms of standard deviations. Let $SDR_t=\sqrt{VR_t}=s_{y_t}/s_{\Delta y_t}$. This new variable is the ratio between the sample standard deviations of y_t and Δy_t. It limiting value is given by

$$P\lim_{N\to\infty} SDR_t = \sqrt{\frac{1-\theta}{(-2\pi)}+\frac{\theta}{2(1-\rho)}} \qquad (93)$$

12.1 Comparison with experience

To see how well the above formulae perform in practice, let us consider the BEA data set which covers the period 1969-98. This is our preferred data set. Making a plausible allowance for measurement errors ($\theta = 0.2$ and $\rho = 0$), the median unbiased estimate of π with this data set is equal to -0.012 (table 7 of the text). With the formula given in (87), this value of π implies an R^2 of 0.491 between two sets of observations 29 years apart. This is very close to the actual R^2 of 0.476 which is obtained using the BEA data set for log employment rates in 1969 and 1998. With $\pi = -0.012$, the formula given in (93) yields a theoretical value of 5.2 for the standard deviation ratio SDR_t. The standard deviations of y_t and Δy_t are 7.83% and 1.17%, respectively, when averaged over the period as a whole. The ratio of these quantities is equal to 6.7 which, although greater than the theoretical ratio of 5.2, is of the same order of magnitude.

Thus, when the appropriate unbiased estimate of π is used, the above formulae explain with reasonable accuracy such macroeconomic features of the BEA data set as the inter-temporal correlation between state employment rates, or the dispersion of these rates across the nation. As can be seen from table D1, the same is true for most of the other data series we use. The main exception is the BLS series for 1948-73, for which the actual value of $R^2_{48,73}$ and the average value of SDR_t are both approximately three time as large as their theoretical values. This huge disparity may be due to the presence of strong deterministic trends that were ignored in deriving the formulae given in equations (87) and (93).

Table 1:Employment Series

	Employment coverage	Population	Period	Accuracy of year-to-year changes
BLS	Non-agricultural employees; Jobs in state;	Total	1948-98	High
BEATOT	All employment; Jobs in state;	Total	1969-99	High
BEAWK	All employment; Jobs in state;	18-64	1970-99	High
CensusA	All employment; Residents in Work;	>14	1976-2000	Low
CensusM	All employment; Residents in work Gender/Education Groups;	25-64	1976-97	Very low

Table 2: Correlation Matrix
Changes in Employment, Unemployment and Participation Rates: 1976-97

	BLS	BEATOT	BEAWK	CensusA	Unemp.	Partic-ipation	CensusM
BLS	1.00	.35	.50	.23	-.32	.07	.15
BEATOT	.35	1.00	.37	.23	-.21	.14	.11
BEAWK	.50	.37	1.00	.32	-.49	.07	.21
CensusA	.23	.23	.32	1.00	-.53	.85	.18
Unemployment	-.33	-.21	-.49	-.53	1.00	.00	-.22
Participation	.07	.14	.07	.85	.00	1.00	-.21
CensusM	.15	.11	.21	-.18	-.22	-.21	1.00

Note: The employment variable used here is equal to log (employment/population)

Table 3

Unit Root Tests for Relative Employment Rates in US States[1/]

P-values for Fisher's λ

		ADF (1) Test		SP (1) Test
		No Constant	*Constant*	*Linear trend*
Data	*Period*			
BLS	1948-1998	.001***	.325	.082
BLS	1948-1973	.001***	.959	.001***
BLS	1973-1998	.005**	.628	.773
BEATOT	1969-1998	.067	.337	.974
BEAWK	1970-1999	.039*	.133	.647
CensusA	1976-2000	.001**	.017*	.480

[1/] In this and subsequent tables, the employment rate =log(employment/population). Collective significance is measured using Fisher's test statistic $\lambda = 2\Sigma \log(P_i)$ where P_i is the probability associated with the observed t-value of π_i. On the null hypothesis that π_i is zero for all 48 states, λ has a χ^2 distribution with $2 \times 48\,(= 96)$ degrees of freedom.
'***'signifies that the null hypothesis (*all* $\pi_i = 0$) is rejected at the 0.1% level.
'**'signifies that the null hypothesis (*all* $\pi_i = 0$) is rejected at the 1% level.
'*'signifies that the null hypothesis (*all* $\pi_i = 0$) is rejected at the 5% level.

Table 4

Unit Root Tests for Relative Employment Rates in US States

Sensitivity Experiment

Specification	No. lagged $\triangle y$ in the regression equations	Error Share[1/] θ	Number of States for which the hypothesis $\pi_i = 0$ is rejected at the following levels: 1%	5%	10%	Fisher's λ	P-value of λ
CensusA data 1976-2000							
ADF test with constant							
(1)	1	0	0	2	2	127.8	.017
(2)	2	0	0	0	2	97.6	.437
(3)	3	0	0	0	3	95.5	.494
(4)	1	40%	0	1	1	101.1	.343
(5)	1	60%	0	0	1	78.4	.905
(6)	1	80%	0	0	0	50.4	.999
BLS data 1948-1973							
SP test (linear trend)							
(7)	1	0	2	6	4	151.8	.001
(8)	2	0	0	3	3	105.4	.241
(9)	3	0	1	4	4	114.5	.096
(10)	1	20%	2	6	4	150.6	.001
(11)	1	40%	1	7	2	140.3	.002
(12)	1	60%	0	6	2	121.0	.043

[1/] The test results shown in this table assume that measurement errors account for a fraction θ of the observed variance in $\triangle y$. Such errors are assumed to have first-order autocorrelation equal to 0.58 in the top panel and zero autocorrelation in the bottom panel. For further information see the notes to table 3.

Table 5

Median Unbiased Estimates of the Adjustment Coefficient π

SP (1) equations (linear trend)

Pooled Regressions

			No allowance for measurement error		With allowance for measurement error		
Data	*Period*	$\widehat{\pi_{LS}}$	$\widehat{\pi_U}$	*P-value*	Error Share[1/] θ	$\widehat{\pi_U}$	*P-value*
BLS	1948-1998	-.118	0	.704	20%	0	.956
BLS	1948-1973	-.325	-.192	.003	20%	-.150	.019
					40%	-.012	.316
BLS	1973-1998	-.144	0	.999	20%	0	.999
BEATOT	1969-1998	-.132	0	.999	20%	0	.999
BEAWK	1970-1999	-.179	0	.990	20%	0	.999
CensusA	1976-2000	-.281	-.083	.192	30%	0	.917
					40%	0	.983

1/ The error share θ is the fraction of the variance of $\triangle y$ that is assumed to be due to measurement errors. First order autocorrelation is assumed to equal 0.58 for CensusA data and zero for the other series.
Note. The least squares estimate $\widehat{\pi_{LS}}$ is obtained by pooling data from all states. The median unbiased estimate $\widehat{\pi_U}$ is then derived by the method of Andrews (1993). Each pooled regression equation contains 48 state-specific linear trends which are estimated by the method of Schmidt and Phillips (1992).

Table 6

Median Unbiased Estimates of the Adjustment Coefficient π

ADF (1) equations with constants

Pooled Regressions

Data	*Period*	$\widehat{\pi_{LS}}$	No allowance for measurement error $\widehat{\pi_U}$	*P-value*	With allowance for measurement error Error Share[1/] θ	$\widehat{\pi_U}$	*P-value*
BLS	1948-1998	-.034	0	.999	20%	0	.999
BLS	1948-1973	-.033	0	.999	20%	0	.999
BLS	1973-1998	-.097	0	.888	20%	0	.992
BEATOT	1969-1998	-.109	-.010	.316	20%	0	.715
BEAWK	1970-1999	-.131	-.040	.033	20%	-.005	.227
CensusA	1976-2000	-.257	-.167	.001	50%	-.045	.055
					60%	-.012	.340

1/ The error share θ is the fraction of the variance of $\triangle y$ that is assumed to be due to measurement errors. First order autocorrelation of such errors is assumed to equal 0.58 in the case of CensusA data and zero for the other series. *Note.* The least squares estimate $\widehat{\pi_{LS}}$ is obtained by pooling data from all states. The median unbiased estimate $\widehat{\pi_U}$ is then derived by the method of Andrews (1993). Each pooled regression equation contains 48 state-specific constants (fixed effects).

Table 7

Median Unbiased Estimates of the Adjustment Coefficient π

ADF (1) equations with no constants or trends

Pooled Regressions

			No allowance for measurement error		**With allowance for measurement error**		
Data	*Period*	$\widehat{\pi_{LS}}$	$\widehat{\pi_U}$	*P -value*	Error Share[1/] θ	$\widehat{\pi_U}$	*P -value*
BLS	1948-1998	-.026	-.026	.001	20%	-.023	.001
BLS	1948-1973	-.032	-.031	.002	20%	-.028	.005
BLS	1973-1998	-.022	-.022	.015	20%	-.019	.025
BEATOT	1969-1998	-.014	-.014	.048	20%	-.012	.076
BEAWK	1970-1999	-.005	-.004	.279	20%	-.003	.342
CensusA	1976-2000	-.027	-.027	.006	40%	-.007	.405
					50%	0	.702

Note. The least squares estimate $\widehat{\pi_{LS}}$ is obtained by pooling data for all states. The median unbiased estimate $\widehat{\pi_U}$ is then derived by the method of Andrews (1993). The results shown in the right hand panel assume that measurement errors account for a fraction θ of the observed variance in $\triangle y$. First order autocorrelation of such errors is assumed to be 0.58 in the case of CensusA and zero for the other series.

Table 8
Stability Simulations for the Employment-Population Ratio
(following a negative 2-standard error shock)

$$\triangle y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t}$$

		Probability of:	
Col(1) π	Col(2) σ_{y^*}/σ_u	Col(3) No recovery at all from initial shock after 10 years	Col(4) Complete recovery to pre-shock level after 10 years
0	∞	.50	.43
-.005	10.0	.49	.43
-.01	7.1	.47	.44
-.02	5.0	.45	.46
-.04	3.6	.40	.50
-.08	2.6	.31	.58
-.16	1.8	.18	.72

Notes: These simulations are based on equations of the form $\triangle y^*_{i,t} = \pi y^*_{i,t-1} + u_{i,t}$ where the $u_{i,t}$ are independent $N\left(0, \sigma_u^2\right)$ shocks. Starting from an arbitrary initial position, the asymptotic variance of $y^*_{i,t}$ is given by $\sigma_{y^*} = \sigma_u^2/\left(1-\varphi^2\right)$ where $\varphi = 1 + \pi$. Probabilities are calculated on the assumption that $y^*_{i,1} = y^*_{i,0} - 2\sigma_u$. Column (3) shows the probability that $y^*_{i,11} \leq y^*_{i,1}$. Column (4) shows the probability that $y^*_{i,t} \geq y^*_{i,1} + 2\sigma_u$ for some $1 < t \leq 11$.

Table D1

Explaining Certain Macro Statistics

		Assumed error share	Assumed auto-covariance of errors	Median unbiased estimate 1/	Actual sample standard deviations 2/ (×100)		Standard deviation ratio SDR		Inter-temporal R^2 (1st and last years)	
Data	Period	θ	ρ	$\widehat{\pi_U}$	$S_{\triangle y}$	S_y	actual	theoretical	actual	theoretical
BLS	1948 – 98	0.20	0	-0.0227	1.48	12.72	8.6	3.8	.104	.096
BLS	1948 – 73	0.20	0	-0.0280	1.60	16.30	10.2	3.4	.669	.228
BLS	1973 – 98	0.20	0	-0.0191	1.35	9.14	6.8	4.1	.364	.367
BEATOT	1969 – 98	0.20	0	-0.0118	1.17	7.83	6.7	5.2	.476	.491
BEAWK	1970 – 99	0.20	0	-0.0034	1.06	7.38	7.0	9.7	.702	.815
CensusA	1976 – 00	0.40	0.58	-0.0071	1.51	7.05	4.7	5.1	.764	.700

1/ The values of $\widehat{\pi_U}$ shown in this column are derived from regression equations without constants or trends using value of θ and ρ given in the preceeding columns (see table 7 of the text).

2/ For both $y \triangle y$ the standard deviations were calculated for each year separately and the annual figures were then averaged.