Macroeconomic Dynamics, **15**, 2011, 160–183. Printed in the United States of America.
doi:10.1017/S1365100509991106

STOCHASTIC CONVERGENCE ACROSS U.S. STATES

MARCELO MELLO
Faculdades Ibmec/RJ

Unit root tests suggest that shocks to relative income across U.S. states are permanent, which contradicts the stochastic convergence hypothesis. We suggest that this finding is due to the well-known low-power problem of unit root tests in the presence of high persistence (i.e., low speed of convergence) and small samples. First, interval estimates of the largest autoregressive root for the relative income in the 48 U.S. contiguous states are quite wide, including many alternatives that are persistent but stable. Second, interval estimates of the half-life of relative income shocks that are robust to high persistence and small samples suggest that in most cases shocks die out within zero to ten years. Third, estimation of a fractionally integrated model for the relative income process suggests strong evidence of mean reversion in the data. These findings provide ample support for the stochastic convergence hypothesis.

Keywords: Stochastic Convergence, High Persistence, Largest Autoregressive Root, Half-Life

1. INTRODUCTION

Time series tests of convergence were initially proposed by Bernard and Durlauf (1995, 1996), who used the ideas of unit root and cointegration to assess stochastic convergence.[1] These earlier tests consist of running unit root/cointegration tests on income per capita differentials. Results based on these tests typically did not support the stochastic convergence hypothesis. For instance, Bernard and Durlauf (1995) find no evidence of stochastic convergence in a sample of 15 OECD economies over the period 1900–1987.

However, one may question the validity of the unit root/cointegration approach to testing for stochastic convergence. First, it is well known that unit root tests suffer from low power. Second, the power problem of unit root tests is further compounded by the low speed of income convergence (i.e., autoregressive parameter near unity) and small samples that are typical of growth studies. Because the null

I would like to thank participants at the 74th Southern Economic Association meeting in New Orleans, the XXVII Meeting of the Brazilian Econometric Society in Natal, RN, Brazil, the 2006 Latin American Meeting of the Econometric Society in Mexico, the Brazilian Central Bank, the editor, and an anonymous referee for comments and suggestions on an earlier version of this work. I would like to thank Barbara Rossi for particularly insightful comments, and for allowing me to use her Matlab codes. The usual disclaimers apply. Address correspondence to: Marcelo Mello, Department of Economics, Faculdades Ibmec/RJ, Av. Presidente Wilson 118/1101, Rio de Janeiro, 20030-020, Brazil; e-mail: mmello@ibmecrj.br.

© 2010 Cambridge University Press 1365-1005/10

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

hypothesis of unit root tests is that of nonconvergence, the low-power problem of the test combined with near-unity autoregressive parameters and small sample leads, too often, to acceptance of the null of nonconvergence.

The above issues can be illustrated by looking at a well-established stylized fact in the growth literature, namely, the 2% speed of convergence commonly found in cross-sectional empirical studies.[2] At a 2% speed of convergence, the half-life of income shocks—that is, the time that it takes for an economy, starting from an equilibrium point, to transit halfway to a new equilibrium after an initial shock—is approximately 35 years. Assuming that income follows an AR(1) process, $y_t = \alpha y_{t-1} + \varepsilon_t$, the half-life, h, is given by $h = \ln(1/2)/\ln(\alpha)$. Thus, for $h = 35$, we have that $\alpha = 0.98$; that is, we obtain an autoregressive parameter that is close to unity. In this case, income shocks are highly persistent but eventually die out. However, with autoregressive parameters so close to unity, it is not surprising that unit root tests fail to reject the null of nonconvergence, especially if one takes into account the small sample sizes available for growth studies.

The above point has been made forcefully by Michelacci and Zaffaroni (2000) and Mello and Guimarães-Filho (2007). To get around the low-power problem of unit root tests in the presence of low speed of convergence, these authors model the data-generating process (DGP) as a fractionally integrated process.[3] In this case, the parameter of integration can assume noninteger values. For certain values of the parameter of integration the income differential process is nonstationary but mean-reverting. This specification can capture well the observed low speed of income convergence. This provides a flexible intermediate case, in contrast to the two extremes imposed by unit root tests, one of $I(0)$ processes, in which shocks die out exponentially (i.e., the speed of convergence is very high), and the other of $I(1)$ processes, in which shocks have permanent effects (i.e., the speed of convergence is zero).

Michelacci and Zaffaroni (2000) provide empirical evidence suggesting that the level of output per capita for a sample of OECD countries over the period 1885–1994 can be represented well by a fractionally integrated process. If this is the case, then convergence tests based on the ideas of unit root and cointegration, such as the ones used by Bernard and Durlauf (1995), are misspecified. Mello and Guimarães-Filho (2007), using Bernard and Durlauf's (1995) data set, model pairwise output per capita differentials as a fractionally integrated process and find ample evidence of stochastic convergence, suggesting that Bernard and Durlauf's tests are indeed misspecified.

In this article, we study the dynamic properties of relative per capita income in the 48 contiguous U.S. states in light of the above remarks. In particular, we explore the sample variability in two measures of persistence for the relative income process, namely, the largest autoregressive root and the half-life of income shocks. To construct interval estimates of the largest autoregressive root, we use the methodology in Stock (1991) and Hansen (1999), and to construct interval estimates of the half-life that are robust to high persistence and small samples we use a novel procedure proposed by Rossi (2005).

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

We find that confidence interval estimates for the largest autoregressive root are rather wide, all of which include many alternatives that are persistent but stable. Interval estimates based on Stock's (1991) methodology have an average lower bound of 0.64 and an average upper bound of 1.03, whereas interval estimates based on Hansen's (1999) methodology have an average lower bound of 0.75 and a upper bound of 1.03. Furthermore, interval estimates of the half-life of relative income shocks that are robust to high persistence and small samples suggest that half-lives are in the zero- to ten-year range for most of the states, which suggests a high degree of mean reversion in relative incomes.

Additionally, following Mello and Guimarães-Filho (2007), we estimate the fractional differencing parameter for the relative income process for the 48 contiguous U.S. states. We find ample evidence of mean reversion in relative income processes, confirming our initial findings. In particular, the fractional differencing parameter is in the nonstationary/mean-reverting region in about two-thirds of the cases. Furthermore, estimates of the fractional differencing parameter for all possible pairs of log income differentials across the 48 states suggest that 69% of income differential pairs are characterized by mean reversion.

We also look at standard measures of convergence such as the cross-sectional standard deviation of relative incomes (σ-convergence) and the correlation between initial relative income and average growth (β-convergence). The data suggest that there is a strong negative correlation between initial relative income and average growth over the period 1929–2002 and that the cross-sectional standard deviation of the relative incomes decreases over time.

In sum, once the sample variability is taken into account and appropriate statistical methods that are robust to high persistence and small samples are used, we find that relative income shocks for the 48 U.S. contiguous states are persistent but eventually die out. This finding gives support to the stochastic convergence hypothesis.

This article is divided as follows. In Section 2, we discuss the DGP and provide some descriptive statistics, which include the traditional β and σ concepts of convergence. In section 3, we present unit roots and stationary tests. Section 4 presents interval estimates of the largest autoregressive root for the relative income process for the U.S. states. In Section 5, we present interval estimates of the half-life of relative income shocks following a novel procedure proposed by Rossi (2005). In Section 6, we present estimates of the fractional differencing parameter and discuss recent pairwise tests of time series convergence. Section 7 concludes.

2. RELATIVE INCOME AND STOCHASTIC CONVERGENCE

The data consists of relative personal per capita income for the 48 U.S. contiguous states obtained from the Bureau of Economic Analysis. The data are available annually over the period 1929–2002, which gives a sample of 74 observations. Relative income is measured as the (natural) logarithm of the ratio of per capita

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

personal income of state i to per capita personal national income.[4] For example, California's relative income is given by $\ln(y_t^{\text{CAL}}/y_t^{\text{US}})$, where y_t^{CAL} is the per capita personal income of California, and y_t^{US} is the national per capita personal income.

We follow Carlino and Mills (1993) and assume that there exists a time-invariant equilibrium level of relative income to which each state is moving toward over time. More specifically, we assume the following DGP:

$$y_{it} = y_i^e + u_{it}, i = 1, \ldots, 48, \text{ and } t = 1929, \ldots, 2002, \quad (1)$$

where y_{it} is the (natural logarithm of) relative per capita income in state i at time t, y_i^e denotes state $i's$ time-invariant equilibrium (natural logarithm of) income per capita, and u_{it} is a stochastic term representing deviations from the equilibrium relative income. The time-invariant equilibrium level of income may differ from zero; that is, we allow $y_i^e \neq 0$. This implies that states can converge to different equilibrium levels, which corresponds to the concept of conditional convergence. The term u_{it} consists of a linear time trend and a stationary stochastic process. Suppressing the subscript i to economize on notation, we have that

$$u_t = v_0 + \beta t + v_t, \quad (2)$$

where v_0 is the initial deviation from the equilibrium level of relative income, and β is the deterministic convergence rate. The above specification for u_t can be related to the concept of β-convergence. As pointed out in Bernard and Durlauf (1996), the notion of β-convergence requires that economies that are initially poor should grow faster than economies that are initially rich. That is, β-convergence implies a negative relationship between initial income and growth rate. In the above specification, this means that if a given state is initially above its equilibrium level, that is, $v_0 > 0$, then it should grow more slowly than the country as a whole, that is, $\beta < 0$. Similarly, if we have $v_0 < 0$, then we need $\beta > 0$ to have β-convergence. In this setup, the convergence rate β can differ across states. Substituting equation (2) into equation (1), we obtain the following expression:

$$y_t = \omega + \beta t + v_t, \quad (3)$$

where $\omega = (y^e + v_0)$. Equation (3) permits us to illustrate more precisely the notion of stochastic convergence. Stochastic convergence requires deviations from relative trend growth, v_t, to be temporary. This definition is standard in the literature; see, for instance, Carlino and Mills (1993) or Michelacci and Zaffaroni (2000).

We model v_t as a zero-mean stationary stochastic process with finite and summable autocovariances. More specifically, we have that

$$a(L)v_t = \varepsilon_t \quad (4)$$

$$(1 - \rho L)b(L) = a(L),$$

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

where ε_t is a white noise, $b(L)$ is a finite-order polynomial lag with $p-1$ distinct and stable roots, and ρ is a "large" but stable root. By "large" we mean that ρ is close to unit. The assumption that ρ is "large" implies that the y_{it} process is characterized by high persistence. The specification in (4) implies that shocks to relative income will be temporary if $|\rho| < 1$ and will be permanent if $|\rho| = 1$.

We model the value of ρ using local-to-unit theory. In the local-to-unit setup, ρ is modeled as being in a decreasing neighborhood of one. In particular, we assume that $\rho = 1 + c/T$, where c is a constant and T is the sample size. The constant c is a noncentrality parameter and can be used to measure the effects of departures from the hypothesis of the unit root on the limit distribution theory. If $c = 0$ the DGP has a unit root; if $c > 0$ and $T < \infty$, we have that $\rho > 1$, so that the root is explosive; if $c < 0$ and $T < \infty$, then $0 < \rho < 1$, and the root is stable.

Following Stock (1991), we obtain a Dickey–Fuller regression form by combining and rearranging equations (3) and (4):

$$y_t = \mu_0 + \mu_1 t + \alpha(1) y_{t-1} + \sum_{j=1}^{k} \alpha^*_{j-1} \Delta y_{t-j} + \varepsilon_t, \tag{5}$$

where $\mu_0 = -c\frac{b(1)\omega}{T} - c\frac{b^*(1)\beta}{T} + \rho b(1)\beta$, $\mu_1 = -\frac{c}{T}\omega b(1)$, $\alpha(L) = L^{-1}[1-a(L)]$, $\alpha(1) = 1 + \frac{c}{T}b(1)$, $\alpha^*_j = -\sum_{i=j+1}^{k} \alpha_j$, $b^*_i = -\sum_{i=j+1}^{k} b_j$, $\rho = 1 + c/T$, and $k = p - 1$. Carlino and Mills (1993) specialize to the case in which $p = 2$ and $b(L) = (1 - \phi L)$, with $|\phi| < 1$. That is, they model v_t as an AR(2) process.

Before we perform the stationary/nonstationary tests, we provide some descriptive statistics and examine basic convergence properties in the data, i.e., we compute the cross-sectional standard deviation of the relative income (σ-convergence) and the correlation between initial relative income and average growth rate (β-convergence).

Figure 1 displays the relative income process for the 48 contiguous U.S. states. Figure 1 suggests that, on average, the cross-sectional dispersion in relative incomes is declining over time. This suspicion is confirmed by examining Figure 2, which shows that the cross-sectional standard deviation of the relative income process is indeed declining over time.

Table 1 displays the seven richest and the seven poorest states, ranked by their relative incomes at the beginning and at the end of the sample period, as well as the average growth rate of the top seven and bottom seven states in 1929. First, all seven of the richest states in 1929 exhibit negative average growth rates, and all seven of the poorest states in 1929 exhibit strongly positive average growth rates. For instance, the relative income of the State of New York exhibits a negative average growth rate of 0.47%, which compounded over 74 years gives a growth rate of minus 30%. That is, the State of New York grew by 30% less than the nation as a whole over the period 1929–2002. On the other hand, among the bottom seven states in 1929, the Carolinas (North and South) exhibit the highest growth rates in

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

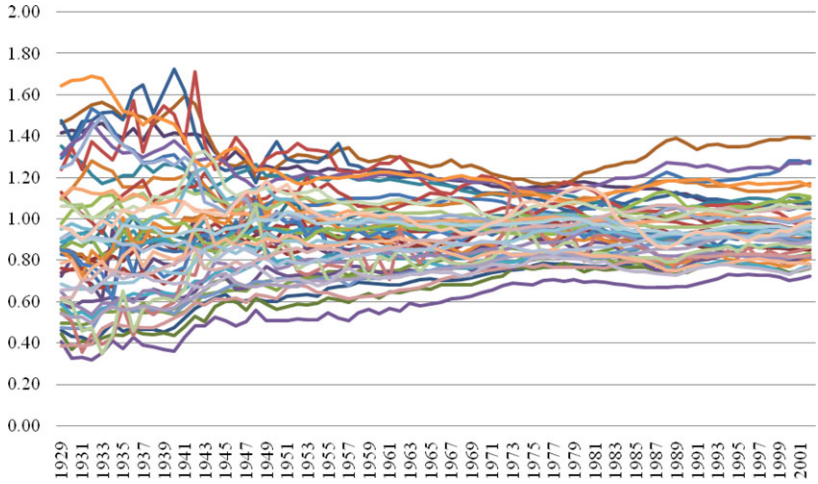


FIGURE 1. U.S. states relative incomes. (This figure can be viewed in color at http://journals.cambridge.org/mdy.)

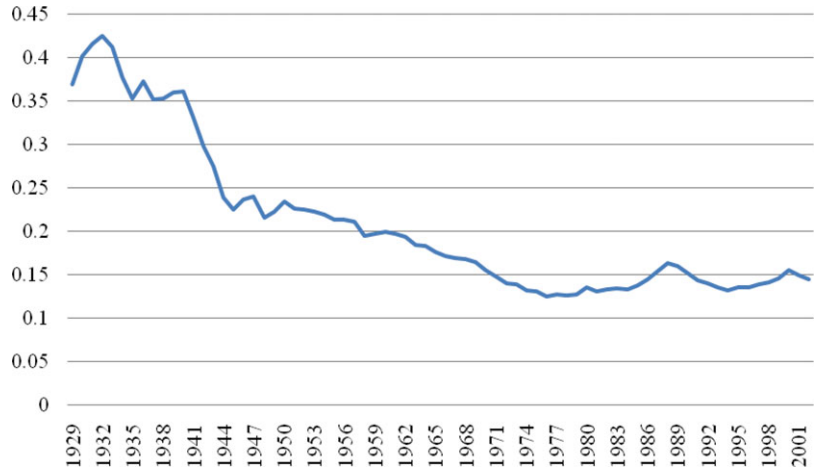


FIGURE 2. σ-convergence.

relative incomes. South Carolina's relative income grew by 113% over the sample period, whereas North Carolina's grew by 88%.

Second, although the initial condition clearly matters, there seems to be considerable mobility in the relative income distribution, especially at the lower end of the distribution. Of the seven richest states in 1929, four remain among the seven richest in 2002. However, the State of New York went from first to fifth place, and Delaware went from second to thirteenth place in rank. Of the seven poorest states in 1929, three remain in the bottom seven in 2002, whereas Tennessee, Georgia, North Carolina, and South Carolina have moved up in the distribution. Indeed,

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 1. Top seven and bottom seven states by relative income

Panel A: Top 7 states by relative income

Avg. growth (%) in relative income, 1929–2002	Top 7 in 1929	Rank in 2002	Top 7 in 2002
−0.47	New York	5	Connecticut
−0.46	Delaware	13	New Jersey
−0.07	Connecticut	1	Massachusetts
−0.38	California	10	Maryland
−0.30	Illinois	8	New York
−0.03	New Jersey	2	New Hampshire
−0.03	Massachusetts	3	Minnesota

Panel B: Bottom 7 states by relative income

Avg. growth (%) in relative income, 1929–2002	Bottom 7 in 1929	Rank in 2002	Bottom 7 in 2002
0.67	Tennessee	33	Idaho
0.86	Georgia	26	Montana
0.86	North Carolina	32	Utah
0.77	Alabama	41	New Mexico
0.73	Arkansas	47	West Virginia
0.78	Mississippi	48	Arkansas
1.03	South Carolina	38	Mississippi

Notes: The top seven and bottom seven states are ranked in descending order.

these states have shown remarkable growth performance: Tennessee has climbed 9 positions in rank; Georgia has climbed 17 positions; and North Carolina and South Carolina have climbed, respectively, 12 and 10 positions.

In sum, the seven richest states grew more slowly than the nation as a whole, whereas the seven poorest states grew much faster than the nation as a whole. This suggests, in addition to the information content in Figures 1 and 2, that the relative income distribution is narrowing, and that there is no absolute poverty trap, in the sense that poor states do not grow or grow at a lower than average rate.

Finally, Figure 3 plots the initial relative income against the average growth rate of the 48 contiguous U.S. states. The initial relative income is strongly negatively correlated with the average growth rate, with a correlation coefficient of −0.91, which corresponds to the notion of β-convergence.

3. UNIT ROOT, STATIONARY TESTS, AND STOCHASTIC CONVERGENCE

As discussed above, unit root and cointegration tests were initially used to assess stochastic convergence. Hence, it is natural to start by looking at the unit

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

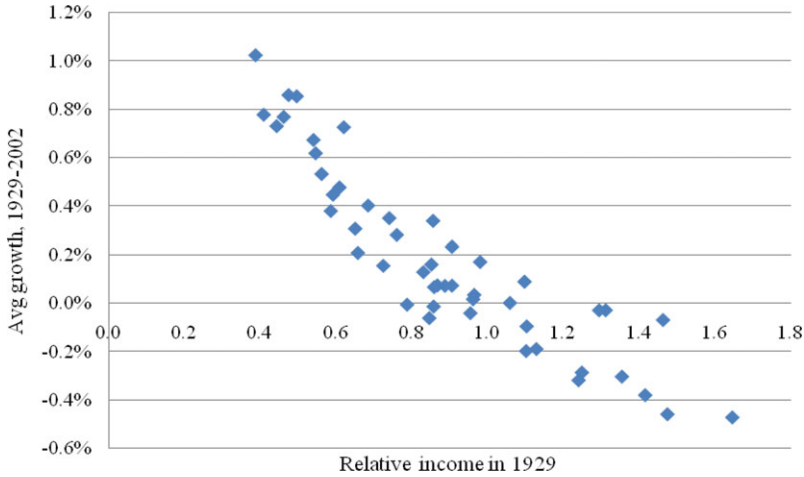


FIGURE 3. Average growth rate vs. initial relative income.

root/stationary tests on the relative income process. Table 2 displays the test statistics of the ADF, DF-GLS, and the KPSS test.

The first column in Table 2 exhibits the *t*-statistic of the ADF test with the lag length chosen by the BIC criterion. At the 5% significance level, there are only nine rejections of the nonstationary null, and at the 10% significance level there are fourteen rejections of the null. The second column in Table 2 displays the DF-GLS unit root test with lag length selected by the BIC criterion. There are eight rejections at the 5% significance level, and nine rejections at the 10% level. The third column displays the *t*-statistic of the DF-GLS unit root test with lag length chosen by the MAIC criterion.[5] In this case, at the 10% significance level, there are only four rejections of the nonstationary null. The last column of Table 2 displays the test statistic of the KPSS test, which has stationarity as the null hypothesis. Based on the KPSS test, we fail to reject the stationarity null for nine states only.

For the states Delaware, Michigan, Minnesota, Mississippi, Nebraska, Nevada, and West Virginia, the nonstationary null is rejected by at least two of the above unit root tests. Thus, based on the evidence on Table 2, shocks to relative income appear to be permanent for most states. This suggests that there is weak to no evidence of stochastic convergence across U.S. states.

However, before discarding the stochastic convergence hypothesis, one should look into the possibility that these findings are caused by the well-known low-power problem of unit root tests or the size distortions of the KPSS tests. Furthermore, as shown below, the high persistence of the relative-income process, combined with a small sample size (after all, we have only 74 observations), may have magnified the size and power distortions of the unit root/KPSS tests. In this setting, one should be particularly wary of unit root/stationary tests.

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 2. Unit root and KPSS tests

	ADF with BIC	DF–GLS with BIC	DF–GLS with MAIC	KPSS
AL	−2.024	−1.876	−1,084	0.2516
AR	−1.874	−1.497	−1,519	0.2556
AZ	−2.787	−2.417	−1,68	0.1479
CA	−2.771	−3.073**	−2,379	0.1382
CO	−2.183	−2.088	−1,458	0.0866*
CT	−2.009	−2.119	−1,694	0.2336
DE	−2.568	−4.666**	−4,187**	0.1520
FL	−2.579	−2.344	−1,144	0.0965*
GA	−2.634	−1.095	−1,095	0.2542
IA	−2.271	−1.752	−1,503	0.2072
ID	−4.713**	−1.490	−1,49	0.1728
IL	−3.664**	−2.464	−1,024	0.0897*
IN	−2.254	−1.042	−1,042	0.1968
KS	−2.902	−2.077	−1,753	0.2073
KY	−1.653	−2.180	−2,18	0.2657
LA	−3.459*	−1.719	−1,4	0.2095
MA	−2.173	−1.326	−0,938	0.2447
MD	−2.928	−3.133**	−2,165	0.2180
ME	−2.243	−1.581	−1,321	0.2422
MI	−3.254*	−3.336**	−1,215	0.1516
MN	−5.387**	−3.575**	−3,575**	0.0334*
MO	−3.099	−1.918	−1,918	0.2066
MS	−2.583	−3.200**	−2,892*	0.2351
MT	−3.008	−1.407	−1,407	0.1629
NC	−2.533	−0.461	−0,998	0.2270
ND	−2.712	−1.767	−1,767	0.1801
NE	−3.843**	−3.188**	−3,188**	0.1480
NH	−2.356	−1.459	−1,459	0.2035
NJ	−1.119	−1.705	−1,418	0.2472
NM	−1.405	−0.741	−0,741	0.2469
NV	−5.928**	−2.935*	−1,986	0.0724*
NY	−2.251	−0.706	−0,706	0.2407
OH	−3.487**	−2.710	−2,71	0.1111*
OK	−2.127	−2.079	−2,079	0.2427
OR	−2.683	−2.234	−1,658	0.1045*
PA	−1.674	−0.698	−0,698	0.2617
RI	−1.332	−1.583	−1,687	0.2379
SC	−2.565	−0.508	−0,99	0.2188
SD	−3.231*	−1.992	−1,548	0.1534
TN	−2.876	−2.612	−1,859	0.1932
TX	−1.699	−1.270	−1,672	0.1864
UT	−3.416*	−1.957	−1,322	0.2014
VA	−4.513**	−1.700	−1,177	0.1672

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 2. Continued

	ADF with BIC	DF–GLS with BIC	DF–GLS with MAIC	KPSS
VT	−2.176	−1.224	−1,224	0.2012
WA	−3.423*	−2.089	−1,513	0.0977*
WI	−2.630	−2.435	−2,435	0.1960
WV	−3.775**	−3.389**	−1,145	0.1298
WY	−3.668**	−2.553	−1,901	0.0951*

Note: The unit root and KPSS tests include an intercept and a time trend.
*, **, and ***Rejection at 10%, 5%, and 1%, respectively.

The low-power problem of unit root tests implies that the nonconvergence null will be accepted too often, even when the relative income process is stationary. Similarly, the results based on the KPSS test can be misleading if the test suffers from size distortions. Given the stationarity null, the size of the KPSS test is the probability of rejecting stationarity of the relative-income process when stationarity is true. If the effective size of the test is substantially greater than its nominal size, then one is more likely to falsely reject the null of stationary. This possibility is considered by Caner and Kilian (2001), who provide evidence that if the series is highly persistent then tests with the stationary as the null, such as the KPSS test, may suffer from extreme size distortions. This finding is quite intuitive. It is, in fact, just the mirror image of the low-power problem of unit root tests. We argue below that the size distortions of the KPSS test may explain the high rejection rates of the stationarity hypothesis in Table 2.

Table 3 displays the autoregressive coefficient of the relative-income process as a measure of persistence. It also displays estimates of the power of unit root tests obtained from Elliot et al. (1996) and estimates of the size of the KPSS test obtained from Caner and Killian (2001). We argue below, based on the evidence in Table 3, that the high persistence of the relative-income series combined with the poor power properties of unit root tests and the size distortions of the KPSS test may have been the cause of the large number of acceptances of the nonstationary null in the unit root tests and the large number of rejections of the stationary null of the KPSS tests in Table 2.

First, to assess the persistence of the relative-income process, we estimate an AR(1) model and look at the first-order autoregressive coefficient. Estimation of an AR(1) model without a linear trend [model 1, column (1), Table 3] suggests that of the 48 relative income series, 13 have autoregressive coefficients that are greater than or equal to 0.95, 30 have autoregressive coefficients that are greater than or equal to 0.90, and 42 have autoregressive parameters that are greater than or equal to 0.80. Estimation of an AR(1) model with a linear trend [model 2, column (1), Table 3] suggests that there are 5 autoregressive coefficients greater than 0.95, 14 that are greater than 0.90, and 31 that are greater than 0.80. This suggests that there is a considerable amount of persistence in relative incomes.

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 3. Approximate power of the unit root tests, and the effective size of the KPSS test ($T = 100$)

β	# of autoregressive coefficients greater or equal than β Col. (1)	Approximate power of ADF-BIC test Col. (2)	Approximate power DF-GLS-BIC test Col. (3)	Approximate size KPSS test Col. (4)
		Model 1: $y_t = \alpha + \beta y_{t-1} + u_t$		
0.95	13	0.10	0.28	0.30
0.90	30	0.22	0.60	0.18
0.80	42	0.59	0.93	0.09
0.70	46	0.83	0.99	0.06
		Model 2: $y_t = \alpha + \delta t + \beta y_{t-1} + u_t$		
0.95	5	0.08	0.10	0.30
0.90	14	0.15	0.24	0.19
0.80	31	0.39	0.61	0.10
0.70	41	0.64	0.84	0.07

Notes: Estimates of the approximate power for the unit root tests were obtained from Elliot et al. (1996), Tables II and III, pp. 828–829. Estimates of the effective size of the KPSS test is obtained from Table 1 in Caner and Kilian (2001). The nominal size of the KPSS test is 5%.

Second, we obtain from Elliot et al. (1996) the (approximate) power of the ADF-BIC and DF-GLS-BIC. Based on Monte Carlo simulations, Elliot et al. (1996) have shown that for an AR(1) process without a linear time trend and with an autoregressive parameter equal to 0.95 and for $T = 100$, the ADF-BIC test has power equal to 0.10, and if the autoregressive parameter is 0.90, the power is 0.22. As mentioned above, 30 of the 48 relative income series have autoregressive parameters that are greater than or equal to 0.90. That is, for at least 30 series, the power of the ADF-BIC test is no greater than 22%, whereas for at least 13 series, the power is no greater than 10%.

For the AR(1) model with a linear time trend [Table 3, model 2, column (2)], the ADF-BIC test has power equal to 0.08 if the autoregressive parameter is 0.95, has power equal to 0.15 if the autoregressive parameter is 0.90, and has power equal to 0.39 if the autoregressive parameter is 0.80. As reported in Table 3, for an AR(1) model with a time trend, there are 31 with autoregressive parameters that are greater than or equal to 0.80. Therefore, for at least 31 series, the power of the ADF-BIC test is no greater than 39%, and for at least 14 series, the tests have power no greater than 15%.

The approximate power of the DF-GLS-BIC test for an AR(1) process without trend and with autoregressive parameter equal to 0.95 is 0.28, and if the autoregressive parameter is 0.90, the power is equal to 0.60. As mentioned above, of the 48 relative income processes, 30 have autoregressive parameter greater than or equal to 0.90. Therefore, for at least 30 series, the power of the DF-GLS-BIC test is no greater than 60%, and for at least 13 series, the power is no greater than 28%. If we allow for a time trend in the AR(1) specification, the power of the

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

DF-GLS-BIC test if the process has an autoregressive parameter equal to 0.95 is equal to 0.10, and if the autoregressive parameter is 0.90 the power of the test is 0.24. Thus, for 14 series, the power of the DF-GLS-BIC test is no greater than 24%, and for 5 series, the power is no greater than 10%.

Third, we obtain the effective size of the KPSS test with a nominal size of 5% from Caner and Kilian (2001). For an AR(1) process without a linear trend (model 1, Table 3), the effective size of the KPSS test in the 13 relative income series with $\rho = 0.95$ is approximately 0.30, and in the 17 series with $\rho = 0.90$, the approximate effective size is 0.18. In the remaining 18 series the approximate effective size is in the range 0.05–0.09. Similarly, for an AR(1) with a linear time trend (model 2, Table 3), when $\rho = 0.95$ the approximate effective size is 0.30 for five series, and when $\rho = 0.90$ the approximate effective size is 0.19 for nine series. In the remaining 34 series the approximate effective size of the KPSS test is in the range 0.07–0.10. Because there are 5 series with $\rho = 0.95$, and 9 series with $\rho = 0.90$, the size of the KPSS test in 14 series is in the range 19%–30%.

Based on the Table 3 estimates, we suggest that the large number of acceptances of the nonstationary null hypothesis based on the ADF-BIC and DF-GLS-BIC tests might have been caused by the low-power problem of the test and the high persistence of the relative-income series in the presence of small samples. Similarly, the large number of rejections of the stationary null in the KPSS test may have been caused by the size distortions. This conclusion should be reinforced by the fact that the above power/size estimates were generated assuming $T = 100$ and that our sample size contains only 74 observations. Thus, the above power/size estimates should be seen as upper/lower bounds on the true power/size of the tests in Table 2.

4. INTERVAL ESTIMATES OF THE LARGEST AUTOREGRESSIVE PARAMETER

To better assess the persistence of the relative-income process, we construct interval estimates for the largest autoregressive root ρ in equation (5). Reporting confidence intervals for ρ is a strategy superior to relying on unit root tests, because confidence intervals provide more information about sampling uncertainty. We construct confidence intervals for ρ, using the methodology in Stock (1991) and Hansen (1999).

Stock's (1991) methodology consists of generating confidence intervals for ρ by inverting t-statistics of ADF tests. The hard part is to derive the asymptotic distributions for the t-statistic, which are non-normal and depend nontrivially on the noncentrality constant c. Given the estimated t-statistic, the upper and lower values of the constant c are obtained, and a confidence interval for ρ is generated as follows:

$$(\rho_{\text{low}}, \rho_{\text{up}}) = \left(1 + \frac{c_{\text{low}}}{T}, 1 + \frac{c_{\text{up}}}{T}\right).$$

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

When conventional asymptotic approximations are poor, an alternative is to use bootstrap methods. For instance, the percentile-t bootstrap constructs a confidence interval for ρ by evaluating the sampling distribution of the t-statistic, $t(\rho) = \frac{\hat{\rho}-\rho}{S(\hat{\rho})}$, assuming that the data are generated by (5) with the true value of ρ given by its OLS estimate. Although the percentile-t method is a good approximation in general, Hansen (1999) shows that when it is applied to the autoregressive model it has incorrect first-order asymptotic coverage and, therefore, it is not a good method for constructing confidence intervals for ρ. Hansen (1999) proposes the so-called grid-bootstrap method as an alternative to the percentile-t method. He shows that the grid-bootstrap has correct first-order asymptotic coverage for both stationary and local-to-unit autoregressive models. Thus, the grid-bootstrap method is superior because it controls Type I error globally in the parameter space. The difference between the grid-bootstrap method and the percentile-t method is that, loosely speaking, the grid-bootstrap confidence interval is computed using a grid of values for ρ rather than only one value (the OLS estimate of ρ), as in the case of the percentile-t bootstrap.

Table 4 displays confidence interval estimates for the largest autoregressive root, ρ. Interval estimates of ρ based on Stock's (1991) methodology suggest that the unit root is outside the interval in five cases, namely, for the states Idaho, Minnesota, Nevada, Virginia, and West Virginia. (Not surprisingly, we also reject the null of nonconvergence for these states with the ADF-BIC test in Table 3.) The average lower bound of the interval estimate is 0.64, and the average upper bound is 1.03. The median lower bound and the median upper bound are, respectively, 0.72 and 1.05.

Interval estimates based on Hansen's (1999) methodology suggest the unit root is outside the interval in six cases, which are Colorado, Iowa, Idaho, Minnesota, Nevada, and Virginia. The average lower bound of the interval estimate is 0.75, and the average upper bound is 1.03. The median lower bound and the median upper bound are, respectively, 0.79 and 1.06.

Both methods generate interval estimates that are quite wide, which include many alternatives that are persistent but stable. In conclusion, interval estimates in Table 4 suggest that when sample variability is taken into account we cannot discard the possibility that relative income processes are stationary, which favors the stochastic convergence hypothesis.

5. ESTIMATES OF THE HALF-LIFE

So far we have been using the largest autoregressive root ρ as a measure of persistence. However, to have a better assessment of the degree of mean reversion of relative income shocks, it is helpful to measure persistence in terms of the half-life.[6]

The half-life is defined as the time horizon h such that the expected value of y_{t+h} has reverted to half of its initial postshock level. That is, it is the value h that solves the equation $E(y_{t+h}) = \frac{1}{2}y_t$. For an AR(1) process, $y_t = \alpha y_{t-1} + \varepsilon_t$, the half-life

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 4. CI for the largest autoregressive root

	CI for ρ	
	Stock (1991)	Hansen (1999)
AL	(0.826, 1.061)	(0.920, 1.071)
AR	(0.848, 1.063)	(0.841, 1.062)
AZ	(0.691, 1.049)	(0.709, 1.046)
CA	(0.694, 1.049)	(0.732, 1.048)
CO	(0.800, 1.059)	(0.447, 0.884)
CT	(0.828, 1.062)	(0.958, 1.078)
DE	(0.732, 1.053)	(0.747, 1.058)
FL	(0.730, 1.053)	(0.774, 1.057)
GA	(0.720, 1.052)	(0.895, 1.072)
IA	(0.785, 1.058)	(0.454, 0.890)
ID	(0.000, 0.755)	(0.398, 0.844)
IL	(0.500, 1.021)	(0.641, 1.030)
IN	(0.788, 1.058)	(0.824, 1.064)
KS	(0.815, 1.061)	(0.826, 1.063)
KY	(0.879, 1.065)	(0.883, 1.068)
LA	(0.548, 1.029)	(0.787, 1.062)
MA	(0.802, 1.060)	(1.013, 1.080)
MD	(0.663, 1.044)	(0.870, 1.067)
ME	(0.790, 1.059)	(0.796, 1.055)
MI	(0.594, 1.035)	(0.625, 1.042)
MN	(0.000, 0.758)	(0.300, 0.703)
MO	(0.628, 1.039)	(0.667, 1.036)
MS	(0.519, 1.039)	(0.740, 1.049)
MT	(0.646, 1.042)	(0.708, 1.048)
NC	(0.738, 1.054)	(0.799, 1.061)
ND	(0.705, 1.050)	(0.730, 1.048)
NE	(0.000, 1.060)	(0.524, 1.001)
NH	(0.770, 1.057)	(0.917, 1.074)
NJ	(0.946, 1.069)	(0.931, 1.075)
NM	(0.913, 1.067)	(0.939, 1.074)
NV	(0.000, 0.738)	(0.300, 0.671)
NY	(0.789, 1.058)	(0.923, 1.073)
OH	(0.542, 1.028)	(0.607, 1.024)
OK	(0.809, 1.060)	(0.894, 1.071)
OR	(0.711, 1.051)	(0.776, 1.057)
PA	(0.876, 1.065)	(0.867, 1.072)
RI	(0.922, 1.067)	(0.923, 1.075)
SC	(0.732, 1.053)	(0.847, 1.070)
SD	(0.599, 1.036)	(0.638, 1.032)
TN	(0.673, 1.046)	(0.817, 1.064)
TX	(0.873, 1.065)	(0.890, 1.076)
UT	(0.558, 1.031)	(0.621, 1.029)
VA	(0.000, 0.797)	(0.541, 0.921)
VT	(0.801, 1.060)	(0.835, 1.067)

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 4. Continued

	CI for ρ	
	Stock (1991)	Hansen (1999)
WA	(0.555, 1.030)	(0.811, 1.059)
WI	(0.720, 1.052)	(0.738, 1.050)
WV	(0.063, 0.968)	(0.700, 1.053)
WY	(0.499, 1.021)	(0.824, 1.071)

Notes: Confidence intervals have a 95% probability content and are constructed using the ADF-BIC t-stat from Table 2 and the values of c_{low} and c_{up} provided in Table 1A, Stock (1991, pp. 456–457). Confidence intervals following Hansen's (1999) methodology were generated in Matlab using codes adapted from Rossi (2005).

is given by the well-known expression $h = \ln(1/2)/\ln(\alpha)$. For a general AR(p) process, the half-life cannot be computed using the above expression. Instead, the half-life has to be computed directly from the impulse response function. In this case, the half-life is the time horizon h that solves the equation $\partial y_{t+h}/\partial \varepsilon_t = 1/2$.

We compute confidence intervals for the half-life for a general AR(p) process following the methodology proposed by Rossi (2005). Rossi's procedure for constructing confidence intervals is robust to high persistence in the presence of small samples, which applies directly to our case.

More specifically, Rossi (2005) derives two measures of the half-life—the approximate half-life and the exact half-life. The expression for the exact half-life derived by Rossi (2005) is given by $h = \frac{\ln[1/2 \cdot b(1)]}{\ln \rho}$, where ρ is the largest autoregressive parameter in equation (4), and $b(1)$ is a correction factor given in equation (5). The expression for the exact half-life can also be written in terms of the noncentrality parameter, c. Because $\rho = 1 + c/T$, we have that $\ln \rho \approx c/T$. Using the approximation, the expression for the exact half-life is given by $h = T \cdot \frac{\ln[1/2 \cdot b(1)]}{c}$, where we require the half-life to be positive; that is, $h > 0$. The correction factor $b(1)$ can be consistently estimated from the ADF regression as $\hat{b}(1) = (1 - \sum_{j=1}^{k} \hat{\alpha}^*_{j-1})$.

The exact half-life, as stated above, differs from the half-life expression traditionally computed in the purchasing power parity (PPP) literature, The reason is that when the DGP in (1) is written in its usual Dickey–Fuller regression form, the coefficient on the lagged y_t equals $\alpha(1)$, which differs from ρ. Typically, the half-life is computed as $h_a = \ln(1/2)\big/\ln \alpha(1)$, which differs from the exact half-life whenever $p > 1$.

Rossi (2005) calls h_a the approximate half-life. We can also write h_a in terms of the noncentrality parameter, c, and the sample size, T. That is, we have that $\ln \alpha(1) = \ln(1 + \frac{c}{T} b(1)) \approx \frac{c}{T} b(1)$, so that the approximate half-life is given by $h_a = T \cdot \ln(1/2)/cb(1)$. For $p = 1$, we have $b(1) = 1$, which implies that $h = h_a$; that is, the expressions for the exact and the approximate half-life coincide.

We construct confidence intervals for h and h_a using local-to-unit theory following Rossi (2005). First, from Table 1A in Stock (1991), we construct confidence

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

intervals for ρ, $(\rho_{\text{low}}, \rho_{\text{up}})$. Second, given that $\rho = 1 + \frac{c}{T}$, we write $c = T(\rho - 1)$ to obtain confidence intervals for c, $(c_{\text{low}}, c_{\text{up}})$. Third, given the confidence intervals for c, we construct confidence intervals for the exact and approximate half-life, h and h_a, respectively, by computing the expressions

$$T \cdot \left[\frac{\ln[1/2 \cdot \hat{b}(1)]}{c_{\text{up}}}, \frac{\ln[1/2 \cdot \hat{b}(1)]}{c_{\text{low}}} \right], \quad T \cdot \left[\frac{\ln(1/2)}{c_{\text{up}}\hat{b}(1)}, \frac{\ln(1/2)}{c_{\text{low}}\hat{b}(1)} \right],$$

where $\hat{b}(1)$ can be estimated from equation (5).

One can also use classical methods to compute confidence intervals for the half-life. Given the expression of the approximate half-life, $h_a = \ln(1/2)/\ln \alpha(1)$, application of the delta method yields a 95% confidence interval given by $\hat{h}_a \pm 1.96\hat{\sigma}_{\alpha(1)}(\ln(1/2)/\hat{\alpha}(1)[\ln(\hat{\alpha}(1)]^{-2})$, where $\hat{\sigma}_{\alpha(1)}$ is an estimate of the standard error of $\hat{\alpha}(1)$. If the autoregressive root is close to unity, classical estimates should give poor approximations. However, in some cases, i.e., if the root is not close to unity, classical methods produce valid confidence intervals. In any case, they provide a useful starting point. Table 5 displays interval estimates of the half-life.

Based on classical methods [column (1)], the bulk of interval estimates of the approximate half-life for the U.S. states lie in the interval 0–30 years. The mean lower bound is 0.40 years, and the mean upper bound is 14.45 years, whereas the median lower and upper bounds are 0.47 and 8.41 years, respectively. Note that a half-life of 8 years corresponds to an autoregressive parameter of 0.92. These estimates are consistent with the results obtained in Table 3, which suggests that relative incomes are persistent but stationary.

Column (2) displays estimates of the upper bound of the confidence interval for the exact half-life for $p = 1$.[7] Recall that in this case, we have that $b(1) = 1$, and expressions for the exact and the approximate half-life coincide. Interval estimates of the upper bound of the half-life suggest that the mean upper is 3.7 years, whereas the median upper bound is 2.6 years. These estimates suggest that relative income shocks, on the average, die out relatively quickly.

Interval estimates of the exact half-life for $p = 2$ which imposes an AR(2) structure on $a(L)$ are shown in column (3). Upper bound estimates of the confidence interval suggest that the mean upper bound of relative income shocks is 3.1 years. These estimates remain largely unchanged if we add one more lag to the polynomial $a(L)$, that is, if we impose an AR(3) structure, as shown in column (4). In this case, the mean upper bound is also 3.1 years, and individual upper bound estimates are numerically close to the case in which $a(L)$ has AR(2) dynamics.

For the sake of comparison, we also include estimates of the upper bound of the confidence interval for the approximate half-life for $p = 2$ [column (5)]. Because of the short lag structure in the data, the interval estimates are relatively close to the exact half-life.[8] Ultimately, the average correction factor $b(1)$, for $p = 2$, is 1.12, so it is not surprising that the estimates of the exact and the approximate half-life are close.

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 5. Confidence intervals for the half-life

	Classical CI for the half-life Col. (1)	Upper bound on the half-life, $p = 1$ Col. (2)	Upper bound on the exact half-life, $p = 2$ Col. (3)	Upper bound on the exact half-life, $p = 3$ Col. (4)	Upper bound on the appr. half-life, $p = 2$, Col. (5)
AL	(0, 25.70)	8.3456	5.036	4.8229	4.7755
AR	(0, 12.06)	4.5107	3.96	4.7896	4.8165
AZ	(0.65, 5.37)	2.2163	2.248	2.1613	2.1795
CA	(0.64, 5.44)	2.2401	2.837	3.1336	3.2319
CO	(0.52, 2.09)	0.6412	1.378	1.4600	1.4711
CT	(0,52.00)	15.2630	7.3997	9.3753	7.2201
DE	(0.58, 6.50)	2.5497	2.5309	3.4779	2.4183
FL	(0.77, 7.12)	2.4300	2.7697	2.5759	2.7025
GA	(0, 19.81)	5.1533	4.0118	2.7843	3.8121
IA	(0.53, 2.17)	0.6412	1.4917	1.3803	2.2030
ID	(0.50, 1.80)	0.6412	1.2535	1.3626	1.4894
IL	(0.83, 3.58)	1.3692	0.9745	1.2761	0.9188
IN	(0.43, 10.40)	3.2290	2.3195	2.2655	2.5311
KS	(0.03, 10.63)	3.7042	4.0253	4.0804	4.1576
KY	(0, 17.48)	5.6645	5.4597	6.5510	6.0288
LA	(0.50, 8.45)	2.9514	2.3457	4.0147	2.2153
MA	(0, 136.25)	1.4290	7.5699	6.5497	7.8964
MD	(0, 15.65)	4.9209	3.9704	3.7140	3.9391
ME	(0.26, 8.54)	3.2586	3.8093	4.2518	3.7960
MI	(0.68, 3.86)	1.6878	0.8475	0.9408	0.8014
MN	(0.33, 1.24)	0.6412	0.5906	0.6855	0.6070
MO	(0.72, 4.46)	1.8402	2.314	2.6393	2.4757
MS	(0.55, 6.20)	2.5253	2.816	2.4745	2.9446
MT	(0.84, 5.27)	1.9361	1.5891	0.2441	1.8100
NC	(0.78, 8.37)	2.6149	2.115	1.4979	2.0212
ND	(0.66, 5.90)	2.3231	2.6155	1.4768	3.1602
NE	(0.59, 2.64)	0.6412	1.6644	1.4867	2.1646
NH	(0, 28.51)	7.9477	4.6037	4.8077	4.3476
NJ	(0, 38.86)	12.5847	6.6672	6.5346	6.2857
NM	(0, 34.86)	7.8198	3.1374	2.9390	3.1515
NV	(0.33, 1.13)	0.6412	0.5384	0.3766	0.5738
NY	(0, 31.58)	6.2700	4.6371	4.4442	4.4436
OH	(0.66, 3.30)	1.4931	1.6354	1.8362	1.6640
OK	(0, 20.54)	6.7769	6.4632	4.3670	6.6175
OR	(0.91, 7.30)	2.3051	2.6783	2.4600	2.5927
PA	(0, 15.79)	5.5359	3.5949	2.2661	3.8495
RI	(0, 29.56)	8.7135	6.8612	6.7912	6.4642
SC	(1.22, 11.12)	2.5541	2.3229	1.4830	2.2414
SD	(0.68, 3.89)	1.7084	2.099	1.5630	2.4068
TN	(0.04, 9.83)	3.6484	2.8763	3.3473	2.7128

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 5. Continued

	Classical CI for the half-life Col. (1)	Upper bound on the half-life, $p = 1$ Col. (2)	Upper bound on the exact half-life, $p = 2$ Col. (3)	Upper bound on the exact half-life, $p = 3$ Col. (4)	Upper bound on the appr. half-life, $p = 2$, Col. (5)
TX	(0, 19.39)	5.3806	4.0251	3.4811	3.8585
UT	(0.73, 3.68)	1.5489	1.7998	0.7970	1.9834
VA	(0.77, 2.55)	0.6412	1.4526	1.9321	1.3875
VT	(0.27,10.96)	3.4406	2.5164	3.3919	2.6236
WA	(0.43, 9.23)	3.1180	2.8114	2.6615	2.7302
WI	(0.64, 6.36)	2.4486	2.6619	2.1849	2.5896
WV	(0.75, 5.08)	1.9995	0.9691	2.5380	0.9140
WY	(0.23,11.26)	3.5016	3.2398	2.5561	3.0886

Notes: Interval estimates are generated assuming that the DGP contains an intercept and a time trend. We follow the procedure in Rossi (2005, by adapting the Matlab codes she provides on her Web site.

There are a couple of important caveats concerning the above estimates. First, we should emphasize that Rossi's method of computing confidence intervals for the half-life are appropriate whenever the largest autoregressive root is close to one. Therefore, if the largest autoregressive root is "small," as is the case of some states, notably Idaho, Minnesota, Nevada, Virginia, and West Virginia, then Rossi's approximation will not yield good results. Instead, classical methods should be applied.

Second, the theoretical DGP includes a time trend, so we include a time trend in the empirical estimate. However, we also generated estimates with de-meaned data (not shown here, but available upon request), and none of the results are substantially altered.

Finally, the estimates in Table 5 suggest that the half-life of relative income shocks is in the range 0–10 years for most of the states. These estimates suggest ample evidence of mean-reverting behavior in relative incomes across U.S. states, which confirms our findings in the preceding section.

6. FRACTIONAL STOCHASTIC CONVERGENCE AND PAIRWISE CONVERGENCE

In this section, we look at two alternative concepts of time series convergence. First, we look at the notion of fractional stochastic convergence, following Michelacci and Zaffaroni (2000) and Mello and Guimarães-Filho (2007). Second, we look at the pairwise time series convergence criteria recently proposed by Pesaran (2007).

As discussed above, Mello and Guimarães-Filho (2007) model income per capita differentials as an ARFIMA process to capture the low speed of convergence observed in the data. Assume that the stochastic process for relative income, y_t, is given by $(1 - L)^d y_t = u_t$, where u_t is a zero-mean, constant-variance, and

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

serially uncorrelated error term, and d is the parameter of integration, which can assume noninteger values. This process is called ARFIMA(0, d, 0). For values of $d > -1$, the term $(1 - L)^d$ has a binomial expansion given by $(1 - L)^d = 1 - dL + d(d-1)L^2/2! - d(d-1)(d-2)L^3/3! + \ldots$. Invertibility is obtained if $-1/2 < d < 1/2$. Assuming invertibility, its moving average form is given by $y_t = \sum_{j=0}^{\infty} \psi_j u_{t-j}$, where $\psi_j = \Gamma(j+d)/\Gamma(d)\Gamma(j+1)$, and $\Gamma(.)$ is the gamma function given by $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}dt$.

For values of the parameter d lying in the interval $(-1/2, 1/2)$ the above process is stationary, whereas if d lies in the interval $[1/2, 1)$, the process is nonstationary but mean-reverting. Mean reversion requires the cumulative impulse response $c_N = \sum_{j=0}^{N} \psi_j$, $N = 0, 1, 2, \ldots$, which gives the effect of a unit shock on the level of the series after N periods, to converge to zero at ∞; that is, $\lim_{N\to\infty} c_N = 0$. It can be shown that if $d < 1$, then $c_\infty = 0$; that is, if $d < 1$, then the process is mean-reverting. If $d > 1$, then $c_\infty = \infty$, and if $d = 1$, c_∞ is constant and finite. Either way the process is not mean-reverting.

Table 6 displays estimates of the fractional integration parameter based on the log-periodogram methods proposed by Geweke and Porter-Hudak (1983), henceforth GPH, and Robinson's (1995) multivariate semiparametric method, which can be seen as a generalization of GPH's estimator.

Based on GPH's procedure, estimates of the parameter d lie in the stationary or mean-reverting region in 29 cases: 6 in the stationary region and 23 in the nonstationary/mean-reverting region. Robinson's estimates of the fractional integration parameter generate 30 estimates that are either in the stationary region (6 cases) or in the nonstationary/mean-reverting region (24 cases).

Table 6 estimates based on Robinson's multivariate method were generated for a power parameter equal to 0.5. As a robustness check, we generate estimates for power parameters of 0.4 and 0.6 (not shown, but available upon request). For a power parameter of 0.4 the estimated d lies in the stationary/mean-reverting region in 33/48 cases, and for a power parameter of 0.6 the estimated d lies in the stationary/mean-reverting region in 32/48 cases. Overall, these findings suggest that there is ample evidence of mean reversion in relative incomes.

Pesaran (2007) proposes two measures of pairwise convergence. The first measure of convergence is given by

$$D_t^2 = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (y_{it} - y_{jt})^2,$$

where y_t is the (log of) income for states $i = 1, 2, \ldots, 48$, and $j = i + 1, \ldots, N$. This measure of convergence is proportional to the concept of σ-convergence commonly used in the literature. The second measure of convergence proposed by Pesaran (2007) is given by

$$MD_t = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} |y_{it} - y_{jt}|.$$

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

TABLE 6. Estimates of the fractional differencing parameter for the logged relative income process

	AL	AR	AZ	CA	CO	CT	DE	FL
d-GPH	0.96	0.98	0.50	0.86	0.43	0.92	0.66	0.45
	(0.09)	(0.17)	(0.27)	(0.29)	(0.27)	(0.16)	(0.27)	(0.38)
d-Robinson	0.96	0.98	0.50	0.86	0.44	0.92	0.66	0.46
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
	GA	IA	ID	IL	IN	KS	KY	LA
d-GPH	0.82	0.70	1.06	−0.07	1.19	1.25	1.06	1.04
	(0.32)	(0.15)	(0.16)	(0.25)	(0.19)	(0.33)	(0.23)	(0.27)
d-Robinson	0.82	0.70	1.06	−0.07	1.19	1.25	1.06	1.04
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
	MA	MD	ME	MI	MN	MO	MS	MT
d-GPH	1.50	0.63	0.79	0.44	0.04	0.84	0.97	1.11
	(0.50)	(0.22)	(0.22)	(0.37)	(0.11)	(0.24)	(0.43)	(0.16)
d-Robinson	1.50	0.64	0.79	0.45	0.05	0.84	0.97	1.11
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
	NC	ND	NE	NH	NJ	NM	NV	NY
d-GPH	0.73	1.13	1.49	1.26	0.90	1.44	0.62	1.42
	(0.12)	(0.40)	(0.44)	(0.32)	(0.15)	(0.33)	(0.27)	(0.11)
d-Robinson	0.73	1.13	1.49	1.26	0.90	1.44	0.62	1.41
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
	OH	OK	OR	PA	RI	SC	SD	TN
d-GPH	0.64	1.05	0.99	1.06	1.27	1.13	0.86	0.66
	(0.33)	(0.28)	(0.16)	(0.14)	(0.33)	(0.34)	(0.22)	(0.18)
d-Robinson	0.64	1.05	0.99	1.06	1.27	1.13	0.86	0.66
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
	TX	UT	VA	VT	WA	WI	WV	WY
d-GPH	1.36	1.09	0.73	1.14	0.90	0.98	0.22	0.92
	(0.20)	(0.45)	(0.36)	(0.13)	(0.28)	(0.44)	(0.22)	(0.28)
d-Robinson	1.36	1.09	0.73	1.14	0.90	0.98	0.23	0.92
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)

Notes: The above estimates were generated for a power parameter of 0.50. Because both estimators, GPH and Robinson, are only applicable to stationary series, we first-differenced the data and then applied the estimators.

The MD measure of convergence is the numerator of the Gini coefficient. Both measures of convergence use all pairs of income differentials, and because there are 48 states, we have a total of 1,128 pairwise income differentials.

We construct all 1,128 pairwise (log) income differentials combinations; that is, we have 1,128 terms d_{ijt}, where $d_{ijt} = \log(y_t^i) - \log(y_t^j)$, for $i = 1, 2, \ldots, N-1$,

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

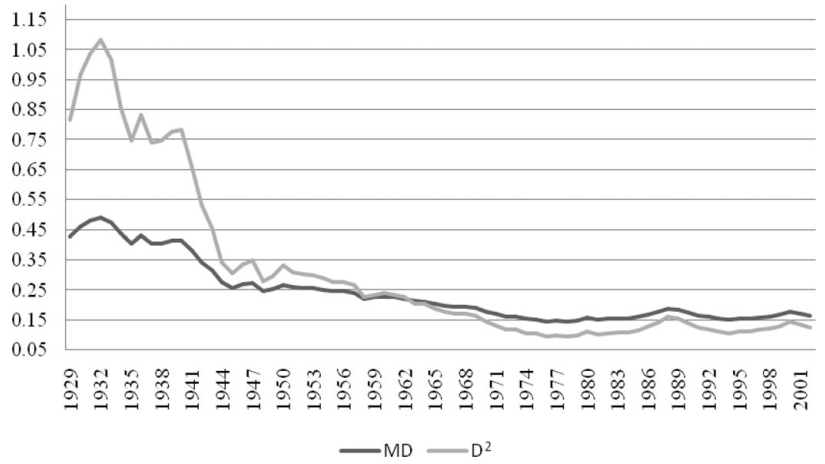


FIGURE 4. *MD* and D^2 statistics.

$j = i + 1, \ldots, N$, and $t = 1929, \ldots, 2002$. We plot both measures of pairwise convergence, D^2 and MD, in Figure 4. As shown below, both measures exhibit a convergence pattern consistent with σ-convergence.

Pesaran (2007) also proposes pairwise and multicountry stochastic convergence criteria. According to Pesaran (2007), there exists pairwise convergence between countries i and j if for some positive constant C and a tolerance probability measure $\pi \geq 0$, if $\Pr\{[y_{i,t+s} - y_{j,t+s}] < C|\Phi_t\} > \pi$ holds at all horizons, $s = 1, 2, \ldots, \infty$. This definition rules out any trending component in the relative income process.

Pesaran (2007) extends the definition of pairwise straightforwardly to a multicountry setting. Following Pesaran (2007), a group of countries $I = 1, 2, \ldots, N$ are said to converge if for some finite positive constant C, and a tolerance probability measure $\pi \geq 0$, we have that $\Pr\{\cap_{i=1,2,\ldots N-1;\, j=i+1,\ldots,N}[y_{i,t+s} - y_{j,t+s}] < C|\Phi_t\} > \pi$, at all horizons, $s = 1, 2, \ldots, \infty$. The definition of multicountry convergence basically requires all country-pairs combinations of income differentials to satisfy the criterion of pairwise convergence.

Pesaran's pairwise convergence notion differs in at least two important respects from our DGP. First, our DGP implies the choice of a benchmark, which arises naturally as the national per capita income. On the other hand, Pesaran's concept uses all combination pairs of countries, and therefore does not have a benchmark country.[9] Second, Pesaran's concept rules out any trending component, whereas we allow for a time-varying equilibrium relative income.

Pesaran uses a sequential procedure to test for pairwise convergence. In the first stage, the income differential process, $y_{i,t} - y_{j,t}$, is tested for the presence of unit roots using an ADF test with an intercept and a time trend, and the lag length selected by the AIC. In the second stage, if the unit root hypothesis is rejected, the income differential process is tested for the presence of a linear trend.

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

Pesaran (2007) applies his concepts of pairwise convergence to output series from the Penn World Tables data set over the period 1950–2000. He finds no support for the convergence hypothesis.[10] This is not really a surprising result. After all, the same problems and pitfalls associated with the application of unit root tests to income differentials that we point out above also apply to Pesaran's procedure. That is, the problem of low power and high persistence in the presence of small samples may have been the driving force behind his findings.

We test for the presence of unit roots in all 1,128 pairs of income per capita differentials among the 48 contiguous U.S. states using the DF-GLS test with lag length chosen by MAIC. At a 10% significance level, we reject the null of nonconvergence in 14.9% of the income differentials pairs, and at a 5% significance level, we reject the null in 8.6% of the pairs. Alternatively, we estimate the fractional differencing parameter using Robinson's estimator. We find 779/1,128 or 69% parameter estimates on the stationary/mean-reverting region. These findings are in line with the evidence presented in Sections 3–5.

7. CONCLUSIONS

The low-power problem of unit root tests combined with high persistence (that is, low speed of income convergence) in the presence of small sample sizes may have been the cause of the large number of acceptances of the nonconvergence null in unit root tests and the large number of rejections of the stationary null in unit root/KPSS tests on relative income for the 48 U.S. contiguous states.

Interval estimates of the largest autoregressive root produce intervals that are quite wide, which include many alternatives that are persistent but stable. By generating interval estimates to assess persistence in the data, we take into account the sampling uncertainty, and therefore this should be a strategy superior to simply relying on unit root tests and point estimates.

We also provide interval estimates of the half-life of relative income shocks for the 48 contiguous U.S. states. Our estimates suggest ample evidence of mean reversion in relative incomes. In particular, using a novel procedure that is robust to high persistence and small samples to compute interval estimates of the half-life, we find that in most cases the half-life of relative income shocks is in the 0- to 10-year range. The finding of mean reversion in relative incomes across the U.S. states is corroborated by estimates of a model of fractional integration for the relative income processes.

NOTES

1. Convergence in a time series sense is known as stochastic convergence.
2. See Barro and Sala-i-Martin (2004, p. 496).
3. Fractionally integrated processes, also known as ARFIMA (autoregressive fractionally integrated moving average) processes, are part of a larger class of processes characterized by long-range dependence or long memory. Long-range dependence is characterized, roughly speaking, by a slowly decaying covariance function. In this case, for certain values of the fractional differencing parameter,

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

the process can be nonstationary, but shocks are mean-reverting. See Baillie (1996) or Robinson (1994) for details.

4. Ideally, we should have the ratio of real incomes instead of nominal incomes. However, there are no price deflators at state level.

5. The DF-GLS test is known as the efficient unit root test because it optimizes the power properties of the test. Perron and Ng (2001) have shown that the DF-GLS test has its power properties further improved if the lag length is chosen by the modified Akaike criteria (MAIC).

6. The half-life is commonly used as a measure of persistence in the PPP literature. In fact, the so-called PPP puzzle is stated in terms of the half-life. See Rogoff (1996) for details.

7. Because the lower bound is zero in most of the cases, we display only the upper bound of the interval.

8. We also generated estimates of the confidence interval for the exact half-life using the lag length selected by the MAIC and the BIC criteria. However, in most cases the lag length selected was 0, 1, or 2, which is why we imposed this lag structure on the estimates in Table 5.

9. In some settings, this is indeed an advantage. For instance, it is well known in the PPP literature that the choice of a benchmark country can affect the tests of the PPP.

10. Interestingly, the first application of pairwise convergence involving all country-pairs combinations was the NBER article of Bernard and Durlauf (1991). The results in Mello and Guimarães-Filho (2007), who also work with all country-pair combinations, were already circulating as early as in the 2000 Midwest Economics Association Meeting in Nashville.

REFERENCES

Baillie, R. (1996) Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.

Barro, R.J. and X. Sala-i-Martin (2004) *Economic Growth*, 2nd ed. Cambridge, MA: MIT Press.

Bernard, A. and S. Durlauf (1991) Convergence in International Output Movements. NBER working paper 3717.

Bernard, A. and S. Durlauf (1995) Convergence in international output. *Journal of Applied Econometrics* 10, 97–108.

Bernard, A. and S. Durlauf (1996) Interpreting tests of the convergence hypothesis. *Journal of Econometrics* 71, 161–17.

Caner, M. and L. Killian (2001) Size distortions of tests of the null hypothesis of stationarity: Evidence and implications for the PPP debate. *Journal of International Money and Finance* 20, 639–657.

Carlino, Gerald A. and Leonard Mills (1993) Are U.S. regional incomes converging? *Journal of Monetary Economics* 32, 335–346.

Elliot, Graham, T. Rothemberg, and J. Stock (1996) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813–836.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4, 221–238.

Hansen, B. (1999) The grid bootstrap and the autoregressive model. *Review of Economics and Statistics* 81, 594–607.

Mello, Marcelo and Roberto Guimarães-Filho (2007) A note on fractional stochastic convergence. *Economics Bulletin* 16, 1–14.

Michelacci, Claudio and P. Zaffaroni (2000) (Fractional) beta convergence. *Journal of Monetary Economics* 45, 129–153.

Perron, P. and S. Ng (2001) Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69, 1519–1554.

Pesaran, M.H. (2007) A pair-wise approach to testing for output and growth convergence. *Journal of Econometrics* 138, 312–355.

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press

Robinson, P. (1994) Time series with strong dependence. In C. Sims (ed.), *Advances in Econometrics*, *Sixth World Congress,* pp. 47–96. Cambridge, UK: Cambridge University Press.
Robinson, P. (1995) Log-periodogram regression of time series with long-range dependence. *Annals of Statistics* 23, 1048–1072.
Rogoff, K. (1996) The purchasing power parity puzzle. *Journal of Economic Literature* 34, 647–668.
Rossi, B. (2005) Confidence intervals for the half-life deviations from purchasing power parity. *Journal of Business and Economic Statistics* 23, 432–442.
Stock, James (1991) Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series. *Journal of Monetary Economics* 28, 435–459.

https://doi.org/10.1017/S1365100509991106 Published online by Cambridge University Press