ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 55-56 – 1999

On the Use of Panel Data Methods with Cross-Country Data

G.S. MADDALA*

ABSTRACT. – The recent work on cross-country regressions can be compared to looking at "*a black cat in a dark room*". Whether or not all this work has accomplished anything on the substantive economic issues is a moot question. But the search for "*a black cat*" has led to some progress on the econometric front. The purpose of this paper is to comment on this progress. We discuss the problems with the use of cross-country panel data in the context of two problems: The analysis of economic growth and that of the purchasing power parity (PPP) theory.

A propos de l'emploi des méthodes de panel sur des données inter-pays

RÉSUMÉ. – Les travaux récents utilisant des régressions inter-pays peuvent être comparés à la recherche d'« *un chat noir dans une pièce sans lumière* ». La question de savoir si ces travaux ont apporté quelque chose de significatif à la connaissance économique est assez controversée. Mais la recherche du « *chat noir* » a conduit à quelques progrès en économétrie. L'objet de cet article est de discuter de ces progrès. Les problèmes posés par l'utilisation de panels de pays sont discutés dans deux contextes : celui de la croissance économique et de la convergence d'une part ; celui de la théorie de la parité des pouvoirs d'achat d'autre part.

* G.S. MADDALA: Department of Economics, The Ohio State University.
I would like to thank M. NERLOVE, P. SEVESTRE and an anonymous referee for helpful comments. Responsability for the omissions and any errors is my own.

« *The Gods love the obscure and hate the obvious* »
BRIHADARANYAKA UPANISHAD

1 Introduction

The path breaking paper on the CES production function by ARROW, CHENERY, MINHAS and SOLOW appeared in 1961. It was based on a cross-country regression of the log of value added per worker on the log of wage rate. Joan ROBINSON characterized this effort of estimating the elasticity of substitution between capital and labor as "*looking for a black cat in a dark room where no cat exists*." The cross-country regressions which assumed the same production function for all countries fell in disrepute but the CES production function had its impact.

The seminal paper by Lucas on "*Some International Evidence on Inflation-Output Trade Off*" appeared in 1973. Again, it was based on a cross-country regression and the empirical work presented in the paper is not directly related to the theory presented earlier. Although it again suffers from the same problems as the earlier seminal paper on cross-country regressions by ARROW *et al*., on the CES production function, it did not receive much of any criticism.

In the past decade there have been several large cross-country panel data sets that have been assembled. These data sets cover many areas: foreign exchange rates, economic growth, agricultural productivity and so on. These data sets have been considered a "*gold mine*" by those working in the respective areas, as well as those interested in panel data techniques. The recent work on cross-country regressions is also like looking at a black cat in a dark room. Whether or not all this work has accomplished anything on the substantive economic issues is a moot question. But the search for a black cat has led to some progress on the econometric front. The purpose of this paper is to comment on this progress.

This paper discusses the problems with the use of cross-country panel data in the context of two problems:

(i). Analysis of economic growth and

(ii). Analysis of the purchasing power parity (PPP) theory.

2 Problems with the Use of Cross-Country Data in the Study of Economic Growth

2.1. Data Problems

First of all, apart from the cross-country data on exchange rates, the data for many countries in the cross-country data sets are very unreliable. They are obtained by interpolation and extrapolation of data from the same country and from related countries. These data problems have been discussed *inter alia* in SRINIVASAN [1994, 1995], BEHRMAN and ROSENZWEIG [1994] and other papers in the *Journal of Development Economics* [1994]. Thus, when using panel data techniques, not all countries should be treated equally. This point has been noticed, however, when results are reported for the OECD countries, another group of 96 countries and so on. In studies on exchange rates where larger panels have been used as in FRANKEL and ROSE [1996], the data on prices are not very reliable for the underdeveloped countries.

Grouping methods based on purely statistical techniques, as in DURLAUF and JOHNSON [1995] should be avoided. They investigate whether subgroups of countries obey separate linear models. They consider a list of 96 countries from the SUMMERS-HESTON data and use a regression tree analysis to group them into four groups (see their Table 4). The resulting combinations, however, are strange. For instance, group 3 combines the "*Asian Tigers*": Singapore, Hong Kong and Korea with not only Philippines but also the Dominican Republic, El Salvador, and Nicaragua. Note, however, that LUCAS [1993] argues that the vastly different experience of Korea and Philippines after World War II deserves a more careful analysis.

Many of the explanatory variables, particularly for the underdeveloped countries are dummy variables or variables that are available only for a few years in the entire data span. Thus, in the panels the variable y_{it} is available for all i and t but the explanatory variable X_{it} is available as a dummy variable or is constant over t *i.e.* it is a variable X_{it} with no time dimension to it. One can average y_{it}, in which case it is a panel in name only but the technique used is that of cross-section analysis. Or else, one can consider a dynamic model with lags of y_{it}. In this case most of the mileage one gets is from the dynamics and very little is left for the explanatory variable x_{it}, to explain. There have been studies in both these categories.

One of the most important variables considered in cross-country regressions on economic growth is the role of human capital. MANKIW *et al.* [1992] use secondary school enrollment as a proxy for human capital. (Their study is a cross-section study.) Studies based on panel data use series that measure human capital stock as average years of education in the labor force. See JUDSON [1996] for a review. These studies obtain a very low coefficient for human capital (compared with the coefficient of 0.3 by MANKIW *et al.*). JUDSON argues that the series on average years of education do not take account of the relative cost of primary education compared with that of higher education. This relative cost is not one, and is not constant across countries. Also, the resources devoted to a year of primary, secondary or higher education vary across countries and over time. Making adjustments for these factors she constructs a new series on human capital and using these series she obtains a coefficient of 0.1, higher than those in the other panel studies but about one third of the one obtained by MANKIW *et al.* This illustrates the importance of data adjustments in the analysis with the cross-country data. Another example of data adjustments is DORWICK and QUIGGIN [1997].

In any case, the cross-country data sets are not ones on which one can jump armed with all the latest econometric panel data techniques. This is perhaps alright for shorter panels consisting of say the OECD countries. But for studies on economic development where data on developing countries need to be used, a closer look at the data and adjustments to the data are necessary.

2.2. **The Data Sets**

The studies discussed here refer to work done on the following cross-country panel data

1. Penn-World tables by SUMMERS and HESTON [1991] analyzed in the many studies on growth.

2. The cross-country data assembled at the World Bank and analyzed by KING and LEVINE [1993a, b] to study the effect of financial development on growth.

2.3. **The Issues**

With data set (1) the initial objectives were to study whether the poor nations catch up with the rich and what economic factors (human capital accumulation, governmental policies, openness of the economy and so on) influence economic growth. The first issue got sidetracked in the whole debate on convergence and the econometric literature on convergence has exploded. See *inter alia* FAGENBERG [1994] and NERLOVE [1996, 1997a, b].

Regarding the influence of the different economic factors on economic growth, different studies have used different variables and obtained significant coefficients. LEVINE and RENELT [1992] found that the significance of any of these variables depends on what other variables are considered. Thus, they argue that the results are very "*fragile*" and that this makes it difficult to draw any policy conclusions. See also LEVINE and ZARVOS [1993]. However, this is not really a problem because the significance of any particular variable in the cross-country regression need not lead to any policy conclusion. For instance, in the cross-country regressions of KING and LEVINE [1993a, b] the coefficient of financial development variable is significant but this does not mean that the way for Somalia to grow is to start a stock market. One should not expect to draw too many policy conclusions from these cross-country regressions. They just point out what variables to look at when analyzing the development policies of any particular country. But policy conclusions for any country depend on more detailed analysis of that particular country.

There have been many studies investigating the effect of democracy on economic growth. See *inter alia* BHALLA [1994] and BARRO [1996]. They show that democracy promotes growth but what specific policy conclusions follow from this is not clear except to argue that U.S. foreign aid should go only to democracies.

SACHS and WARNER [1995] also demonstrate the effects of political freedom and economic openness on economic growth. With respect to the first factor, again, it is not possible to derive any policy prescription. Economic openness or trade liberalization policies are another matter. They are within the realm of many policymakers.

Another interesting example of cross-country regressions to study the determinants of growth is by SALA-I-MARTIN [1997]. He finds the variables: number of years open economy and rule of law to have significantly positive coefficients. However, among the religious variables, Confucian, Buddist and Moslem have significant and positive coefficients and Protestant and Catholic have significant negative coefficients. Policy conclusion: To promote U.S. economic growth, efforts should be made to convert all Christians in the U.S. into Confucians, Buddists or Moslems.

An amusing example of cross-country regressions is that of WALL [1995], who used the data from 95 non-communist non-OPEC countries from the Penn-World Tables and estimated the following regression:

$$g = \alpha + \beta_c D_c + \beta_b D_b + \varepsilon$$

where g = growth rate of real per-capita income between 1960 and 1990, D_c and D_b are dummies to indicate cricket and baseball playing countries. The results are:

$$g = \underset{(15.1)}{103.9} \; - \underset{(41.4)}{43.0}\, D_c + \underset{(39.9)}{80.3}\, D_b$$

From the results he concludes that:

1. For emerging countries without a history of cricket or baseball, baseball instruction and subsidies should be an immediate priority. US and Japan can provide these subsidies.

2. For countries playing cricket, cricket should be abolished but this is a formidable task similar to economic reform of the formerly communist countries.

Regarding the issue of whether the poor nations catch up with the rich, as mentioned earlier, it got lost in the debate on convergence. The role of technology, an important source of growth, has been sidetracked. See FAGENBERG [1994] and BERNARD and JONES [1996b]. Also, it is interesting to note that the poor nations have caught up with the technological development in the rich nations in the area of agriculture but the problem of technological diffusion in other areas remains and this accounts for the poor nations not catching up with the rich (See DOWRICK and GEMMEL [1991]). Thus, the problem of technological diffusion is one that needs more study. For some attempts in this direction, see ZANFORLIN [1996a, b].

As far as the issue of convergence itself is concerned, the evidence on the different concepts of convergence is summarized in SALA-I-MARTIN [1996]. The econometric issues related to the fixed effects models often used in the panel data studies with cross-country data have been thoroughly discussed in

NERLOVE [1996] who shows that the use of fixed effects models biases the results towards rapid convergence whereas more appropriate ML estimates, unconditional on the initial observations imply very slow convergence. The issues of biases in fixed effects models and the performance of a bias corrected fixed effects model, following KIVIET [1995] have been investigated in JUDSON and OWEN [1997] who also demonstrate the sensitivity of the results to the choice of an estimator. NERLOVE [1997, 1999] also demonstrates the importance of exploring the likelihood function rather than just reporting the ML estimates with their asymptotic standard errors. He shows how the inferences are more fragile, than those implied by the ML estimates and their standard errors.

As convincingly demonstrated by NERLOVE, the econometric method used and the appropriate conditioning on the initial observation matter a lot in the estimates of convergence.

In addition to this the results vary a lot depending on the choice of countries used in the study (as discussed in LIU and MADDALA [1996]) which makes the inference very fragile. Thus, choosing the countries used in the panel study is as crucial as (if not more than) the choice of the estimation method.

2.4. Problems of Parameter Heterogeneity and Cross Correlations in Errors

Panel data regressions with cross-country data all rely on some pooled estimates. The pooling is often done under the assumption of complete homogeneity of all the parameters or if some heterogeneity is allowed, it is only in the intercept term. There are several studies in this category. See FRANKEL and ROSE [1996] for a similar panel study. EVANS and KARRAS [1996] allow for heterogeneity in the dynamics but not in the coefficients of the explanatory variables.

In practice, there are two problems that need to be addressed:

(i) Parameter heterogeneity

(ii) Cross correlations among the errors in the regressions for the different countries.

Regarding heterogeneity, the only heterogeneity that is often considered is in the constant term by using the fixed effects and random effects models. Even in ISLAM [1995] who uses the CHAMBERLAIN approach to panel data, the heterogeneity is in the intercept term only. Many other coefficients relevant to the estimation of growth regression are left constant over all countries and over time. EVANS and KARRAS [1996] argue that the time series properties of the data have not been properly taken into account in the studies using the data in the Penn World tables. They allow for dynamics for each of the countries and assume that dynamics for each of the respective countries are heterogeneous. Their concern is with testing endogenous growth theory with the predictions of the neoclassical growth theories – the two theories being nested in the same dynamic model. KOCHERLAKOTA and YI [1995] raise some issues in distinguishing between these theories using growth regressions.

Another study that discusses the problem of heterogeneity is that by LEE, PESARAN and SMITH [1995], who consider a model of the form

$$Y_{it} = \alpha_{it} + \lambda_i Y_{i,t-1} + \beta_i \chi_{it} + \varepsilon_{it}$$
$$i = 1,2,\ldots N$$
$$t = 1,2,\ldots T$$

The discussion of the biases is in terms of a random coefficient model which assumes that $\lambda_i = \lambda + \eta_{1i}$, $\beta_i = \beta + \eta_{2i}$ and (η_{1i}, η_{2i}) have zero means and constant covariance matrix Ω. The main interest is in the estimation of λ and β. They show that the estimates of λ and β from the pooled regression are biased. In the case where X_{it} follows a stationary AR (1) process with the AR (1) parameter ρ (same for all *i*), they obtain the Plims of the fixed effects estimators $\widehat{\lambda}_{FE}$ and $\widehat{\beta}_{FE}$ and show that as ρ approaches unity from below, Plim $(\widehat{\lambda}_{FE}) = 1$ and Plim $(\widehat{\beta}_{FE}) = 0$ irrespective of the true values of the parameters.

However, under the random coefficient specification of heterogeneity, one can always get consistent and efficient estimates of λ and β. Thus, a study of the biases goes only part of the way. Also, if one follows a Bayesian approach, under the random coefficient formulation of heterogeneity one can obtain estimates of the individual parameters λ_i and β_i as well.

The discussion of the biases in the BARRO-type regression and the fixed effects estimator are theoretically interesting but is not practically useful because the biases in the BARRO-type regression have been documented by several others under other assumptions. Of more interest is the appropriate estimation of the model. This is the main focus of the papers by CANOVA and MARCET [1995] and MADDALA and WU [1996a]. LEE *et al.* [1995] estimate the growth regressions separately for each of the 3 groups:

(i) 102 non-oil countries

(ii) 61 intermediate countries (omitting data for countries whose data are supposed to be poor)

(iii) 22 OEDC countries.

They then compute the mean of the parameter estimates for each group and make inferences based on the mean λ. However, as is well known (see MADDALA *et al.* [1997] Table 1) the estimate of the overall mean in a random coefficient model (an assumption made in the discussion of the biases) is not the simple arithmetic mean of the individual cross-section estimates. It is a weighted average, and the simple average gives a biased estimate. If one assumes complete heterogeneity, then the average is not a meaningful statistic.

In summary, the discussion of heterogeneity in LEE *et al.* [1995] is concerned with:

(i) Demonstrating that there is heterogeneity, and

(ii) Showing that BARRO-type regressions produce biased estimates.

The discussion of biases is interesting but the proper estimation procedure to follow in the presence of heterogeneity, which is a more important issue, is not addressed.

To estimate the heterogeneous coefficient model considered by LEE *et al.* [1995] CANOVA and MARCET [1995] use a Bayesian approach but their approach is not completely Bayesian. A correct Bayesian approach is discussed in MADDALA and WU [1996a] which obtains different results. This approach and the empirical BAYES approaches are all discussed in MADDALA *et al.* [1997] with Table 1 in the paper presenting an overview of all the methods.

As for the other problem, cross correlations among errors, neither the paper by LEE *et al.* [1995] nor the papers by CANOVA and MARCET [1995] and MADDALA and WU [1996a] who use the Bayesian approach, address this problem. In fact, the Bayesian approach gets very complicated once allowance is made for cross-correlated errors.

If we assume complete heterogeneity, we can use the seemingly unrelated regression estimation (SURE) method. But this is not feasible for the larger panels. We can consider only small groups of countries (*e.g.* the OECD countries in growth regressions and the G.7 countries in exchange rate equations).

2.5. **Alternative Methods**

The search for convergence has produced some progress in panel data econometric methods. One avenue is an improved estimation of existing cross-section time series models, as in NERLOVE [1996, 1997a, b], IM *et al.* [1996] and the Bayesian methods discussed in section 3.3, the median-unbiased estimation as in CERMENO [1999], the GMM estimation in CASELLI *et al.* [1996], and so on. The median unbiased estimation method actually produces estimates of rates of convergence lower than those produced from LSDV method, as in estimates from the unconditional likelihood function. The GMM estimates in CASELLI *et al.* are, however, higher.

The other avenue is an alternative to cross-country regressions. QUAH [1996a, b] has forcefully argued that the answers to the basic issues on whether poor countries catch up with the rich can never be answered by the use of traditional panel data analysis and the estimation of the cross-country regressions. He suggests formulating and estimating models of income dynamics. This avenue has been also pursued in PAPP and VANDIJK [1996].

Another avenue is to study the convergence of productivity across countries in some specific industries. This is the theme of HART [1995a, b] as well as BERNARD and JONES [1996a, b, c]. This work throws light on how technological progress is spreading across countries, and is more informative on growth than cross-country regressions.

3 Use of Cross-Country Panel Data for the Analysis of Purchasing Power Parity (PPP)

The panel data on exchange rates have been used to estimate cross-country regressions as in FRANKEL and ROSE [1996] and to use panel unit root tests as in OH [1996], PAPELL [1997] and WU [1996]. Since many studies on searching for PPP use panel data unit root tests, we shall first discuss these tests.

3.1. **Panel Unit Root Tests**

It has become almost impossible today to do any time-series analysis without talking of unit roots and applying unit root tests. Since most panel data sets have a time dimension to them, it is natural to look for unit roots in panel data and to apply panel data unit root tests. The most often used tests are the LEVIN-LIN tests [1993a, b]. These tests have been widely used in discussions of purchasing power parity (PPP) theory and other studies on exchange rates. I shall discuss the merits and demerits of these tests and alternatives such as the IM-PESARAN-SHIN (IPS) test [1995] in what follows. The panel data unit root tests have also found their way into discussions of convergence. EVANS and KARRAS [1996] use the LEVIN-LIN test and LEE *et al.* [1995] use the IPS test. The results obtained by LEE *et al.* are mixed but their conclusion from the panel data unit root tests is that the time-series were non-stationary. Thus, the hypothesis of no convergence (even to country-specific steady states) cannot be rejected. However, these conclusions are subject to the same qualifications as those of the other applications of panel data unit root tests in what follows. Since many of the papers at this conference are related to convergence, I shall concentrate on another application with cross-country panel data *via* discussion of purchasing power parity.

3.2. **The Rationale behind Panel Data Unit Root Tests**

Ever since the appearance of the papers by LEVIN and LIN [1992, 1993], panel data unit root tests have been very popular among empirical researchers having access to a panel data set. It is by now a common accepted argument that the commonly used unit root tests like the DICKEY-FULLER (DF), Augmented DICKEY-FULLER (ADF) and PHILLIPS-PERRON (PP) tests lack power in distinguishing the unit root null against stationary alternatives, and that using panel data unit root tests is one way of increasing the power of unit root tests based on a single time series. See *e.g.* the arguments in OH [1996], WU [1996], MACDONALD [1996], and FRANKEL and ROSE [1996], who try to resurrect the Purchasing Power Parity (PPP) theory using the panel data unit root tests.

Strictly speaking this argument is not valid, because it is not meaningful to compare two tests that have different null hypotheses. For instance consider the simplified model:

$$y_{it} = \alpha_i y_{i,t-1} + u_{it} \qquad i = 1,2,\ldots N$$
$$t = 1,2,\ldots T$$

Suppose we are interested in testing $\alpha_1 = 1$ *vs.* $\alpha_1 < 1$, we apply a single equation unit root for the first time series. The panel data unit root test, tests a different hypothesis:

$$H_0 : \alpha_i = 1$$

vs. $H_1 : \alpha_i < 1$ for $i = 1,2,\ldots N$.

Furthermore, there are now more powerful tests available even in the single equation context. See *e.g.* ELLIOTT *et al.* [1996] and also PERRON and NG [1996] who suggest modifications of the P-P test to increase power.

FRANKEL and ROSE [1996] as well as others argue that increasing the number of observations by considering a long time-series, create the problems associated with structural breaks and regime changes (*e.g.* fixed *vs.* floating exchange rate data). It has been argued that the panel data set does not have this problem because of the shorter time span. However, the panel data set creates another problem, cross-sectional heterogeneity and it is not clear whether this is a lesser problem than the lack of time homogeneity (structural breaks).

3.3. The Different Panel Data Unit Root Tests

Some early papers on testing for unit roots based on panel data are by QUAH [1992, 1994] and BREITUNG and MAYER [1994]. Since these have been superseded by the papers by LEVIN and LIN [1992, 1993], we shall not discuss them here.

The Levin-Lin (LL) tests

LEVIN and LIN [1992] conduct an exhaustive study and develop unit root tests for the model:

$$y_{it} = \alpha y_{i,t-1} + \delta_0 + \delta_1 t + \eta_i + \nu_t + \varepsilon_{i,t}$$
$$i = 1,2,\dots,N$$
$$t = 1,2,\dots,T$$

Thus, the model incorporates a time trend as well as individual and time-specific effects. Initially, they assume that $\varepsilon_{it} - \Pi D(0,\sigma^2)$ but they say that under serial correlation with the inclusion of lagged first differences as in the ADF test the test statistics have the same limiting distributions as mentioned subsequently, provided the number of lagged differences increase with sample size: LEVIN and LIN consider several subcases of this model. In all cases the limiting distributions are as $N \to \infty$ and $T \to \infty$. Also in all cases, the equation is estimated by OLS as a pooled regression model. The submodels are:

Model 1:	$y_{it} = \alpha y_{i,t-1} + \varepsilon_{i,t}$	$H_0 : \alpha = 1$
Model 2:	$y_{it} = \alpha y_{i,t-1} + \delta_0 + \varepsilon_{i,t}$	$H_0 : \alpha = 1$
Model 3:	$y_{it} = \alpha y_{i,t-1} + \delta_0 + \delta_1 t + \varepsilon_{i,t}$	$H_0 : \alpha = 1, \delta_1 = 0$
Model 4:	$y_{it} = \alpha y_{i,t-1} + \nu_t + \varepsilon_{i,t}$	$H_0 : \alpha = 1$
Model 5:	$y_{it} = \alpha y_{i,t-1} + \eta_i + \varepsilon_{i,t}$	$H_0 : \alpha = 1, \eta_i = 0$ for all i
Model 6:	$y_{it} = \alpha y_{i,t-1} + \eta_{i0} + \eta_{i1} t + \varepsilon_{i,t}$	$H_0 : \alpha = 1, \eta_{i1} = 0$ for all i

For models 1-4, they show that
(a) $T\sqrt{N}(\widehat{\alpha} - 1) \Rightarrow N(0,2)$
(b) $t_\alpha \Rightarrow N(0,1)$

For models 5, if $\dfrac{\sqrt{N}}{T} \to 0$, then

(a) $T\sqrt{N}(\widehat{\alpha}-1)+3\sqrt{N} \Rightarrow N(0,10.2)$

(b) $\sqrt{1.25}t_\alpha+\sqrt{1.875N} \Rightarrow N\left(0,\dfrac{645}{112}\right)$

In model 6, both intercept and time-trend vary with individuals.

In the empirical applications, OH [1996] uses only models 1 and 5. WU [1996] uses the complete model with trend, and individual and time specific effects but uses the distributions derived for model 5. PAPELL [1997] uses model 5 with lagged first differences added but computes his own exact finite sample critical values using Monte Carlo methods and finds them 3 to 15% higher than those tabulated in LEVIN and LIN [1992].

LEVIN and LIN argue that in contrast to the standard distributions of unit root test statistics for a single time series, the panel test statistics have limiting normal distributions. However, the convergence rates are faster as $T \to \infty$ (superconsistency) than as $N \to \infty$.

The paper by LEVIN and LIN [1993] provides some new results on panel unit root tests. However, since the models discussed are the same, we shall not go into the details. Actually, LEVIN and LIN suggested a four-step procedure for the computation of their test statistics in the presence of heteroskedasticity and autocorrelation. We shall skip the details here but in their simulations, MADDALA and WU [1996b] used the four step procedure.

The major limitation of the LEVIN-LIN tests is that α is the same for all observations. Thus, if we denote by α_i the value of α for the i-th cross-section unit then the LEVIN-LIN test specifies the null H_0 and alternative H_1 as:

$$H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_N = \alpha = 1$$

$$H_1 : \alpha_1 = \alpha_2 = \ldots = \alpha_N = \alpha < 1$$

The null makes sense under some circumstances, but the alternative is too strong to be held in any interesting empirical cases. For example, in testing the convergence hypothesis in growth models, one can formulate the null as implying that none of the economies under study converges and thus $\alpha = 1$ for all countries. But it does not make any sense to assume that all the countries will converge at the same rate if they do converge.

The Im-Pesaran-Shin (IPS) Test [1996]

IPS relax the assumption that $\alpha_1 = \alpha_2 = \ldots = \alpha_N$ under H_1. The basic idea of the test is very simple. Take model 6 in LEVIN and LIN and substitute α_i for α. Essentially what we have is a model with a linear trend for each of the N cross-section units. Thus, instead of pooling the data, we use separate unit root tests for the N cross-section units. Consider the t-test for each cross-section unit based on T observations. Let $t_i (i = 1,2,\ldots,N)$ denote the t-statistics for testing unit roots, and let $E(t_i) = \mu$ and $V(t_i) = \sigma^2$. Then $\sqrt{N}(\bar{t}-\mu)/\sigma \Rightarrow N(0,1)$. The problem is computing μ and σ^2. This they do by Monte Carlo methods and tabulate them for ready reference (Tables 3 and 4 of their paper).

The important thing to note is that the IPS test is a way of combining the evidence on the unit root hypothesis from the N unit root tests performed on the N cross-section units. There is substantial literature on this problem starting with FISHER [1932]. Note that implicit in the IPS test is the assumption that T is the same for all cross-section units and hence $E(t_i)$ and $V(t_i)$ are the same for all i. Thus, we are considering only balanced panel data.

In the case of serial correlation, they proposed to use the ADF t-test for individual series. However, $E(t_i)$ and $V(t_i)$ will vary as the lag length included in the ADF regression varies. They tabulated $E(t_i)$ and $V(t_i)$ for different lag lengths. In practice, however, to make use of their tables, we are restricted implicitly to use the same lag length for all the ADF regressions for individual series.

IPS [1996] also suggest an LR-bar test based on likelihood ratio statistics, but we shall concentrate our discussion on their t-bar test. The same arguments apply to the LR-bar test.

Fisher's (P_λ) Test

Noting that the IPS test is a test combining the evidence from several independent tests on which FISHER [1932] had a solution, MADDALA and LIU [1996] and MADDALA and WU [1996b] suggest an alternative to the IPS test, which they call the FISHER test.

R.A. FISHER [1932] suggested a method of combining the evidence from several independent tests. Suppose there are N unit root test as in IPS. Let P_i be the observed significance level (P-value) for the i-th test. Then $(-2\Sigma \log. P_i)$ has a x^2 distribution with d.f. $2N$. This test, often known as the P_λ test, is mentioned in RAO [1952, p. 44] and MADDALA [1977, Section 4.12] but has not received much attention.

The advantage of this test is that it does not require a balanced panel as in the case of the IPS test. Also, one can use different lag lengths in the individual ADF regressions. Another advantage of the FISHER test is that it can also be carried out for any unit root test derived. The disadvantage is that the P-values have to be derived by Monte Carlo simulation. The IPS test is easy to use because there are ready tables available in the paper for $E(t_i)$ and $V(t_i)$. However, these are valid only for the ADF test. Also, in several practical cases we found that the sample sizes were outside the range provided in the IPS tables, and we had to recompute them by simulation.

3.4. A Comparison of the LL, IPS and Fisher Tests

A comparative study of the merits and demerits of these three tests is discussed in detail in MADDALA and WU [1996b]. Here we shall present a summary of the findings.

First of all, it should be noted that the LL test is a panel unit root test (based on panel estimation) but the IPS and FISHER tests are not. In fact the basic issue of the latter two tests is how to combine the information from a set of independent tests of the same hypothesis. The IPS test combines the t-statistics and the FISHER test combines the significance levels. One important drawback of both the IPS and FISHER tests is their inability to account for contemporaneous covariance in the disturbance across individuals. Since the LEVIN-LIN procedure is essentially a SUR approach, it can allow for contemporaneous correlation across individuals, so long as the contemporaneous correlation matrix is not too large relative to the number of time periods in the sample.

O'CONNELL [1997] argues that there are substantial size distortions with the LL tests in the presence of the contemporaneous correlations problem. In our simulations we found this not to be the case. It is possible that O'CONNELL came to that conclusion because he started out with a model with equicorrelated errors. In any case we are examining this problem.

Apart from these general comments, the main differences between these tests are as follows:

(1) the LL tests test a very restrictive hypothesis that is rarely of practical interest. However, as emphasized in the LEVIN-LIN [1993] paper the test procedure is consistent against alternatives in which α_i varies across individuals, as long as $\alpha_i < 1$ for all $i = 1, 2 \ldots N$. In other words, if the data are stationary but with varying degrees of persistence, the LL test will still reject the null. It seems reasonable to expect that the LL test might have favorable power properties by saving degrees of freedom. However, if there is substantial heterogeneity among the α_i, then the IPS and FISHER tests are expected to show better power than the LL test.

(2) The IPS test is claimed to be a generalization of the LL tests. However, it is better viewed as a way of combining the evidence of several independent unit root tests. As mentioned earlier, there is no panel estimation involved here as in the LL-test.

(3) IM-PESARAN-SHIN present power comparison of the LL and IPS tests and argue that the IPS test is more powerful that the LL test. However, strictly speaking, the power comparison is not valid. Although the null hypothesis is the same in the two tests, the alternative hypothesis is different. The LL tests are based on homogeneity of the autoregressive parameter (although there is heterogeneity in the error variances and the serial correlation structure of the errors). Thus the tests are based on pooled regressions. The IPS test on the other hand is based on heterogeneity of the autoregressive parameter. As argued earlier, the test amounts to a combination of different independent tests. MADDALA and WU [1996c] also present power comparisons with the LL test but it should be borne in mind that the LL test will necessarily come out worse because the LL test has to use panel estimation method which is not valid if there is no need for pooling.

(4) The FISHER test and the IPS test *are* directly comparable. The aim of both tests is a combination of the significance of different *independent tests*. The FISHER test is nonparametric, whatever test statistic we use for testing for a unit root for each sample, we can get the *P*-values P_i and then $-2\Sigma \mathrm{Log}_e P_i \sim \chi^2$ with $2N$ d.f, where N is the number of separate samples. The IPS test on the other hand is parametric. The distribution of the *t*-bar statistic involves the mean and variance of the *t*-statistics used. IPS compute this for the ADF test statistic for different values of the number of lags used and different sample sizes. However, these tables are valid *only* if the ADF test is used for the unit root tests. Also, if the length of the time series for the different samples is different, there is a problem using the tables prepared by IPS. The FISHER test does not have any such limitations. It can be used with any unit root test and even if the ADF test is used, the choice of the lag length for each sample can be separately determined. Also, there is no restriction on the sample sizes for different samples (they can vary according to availability of the data).

(5) The FISHER test is an exact test. The IPS test is an asymptotic test. Note that this does not lead to a huge difference in finite sample results, since the adjustment terms in the IPS test are derived from simulations while the *p*-values in the FISHER test are also derived from simulations. However, the asymptotic validity of the tests depends on different conditions. For the IPS test the asymptotic results depend on *N* going to infinity while for the FISHER test they depend on *T* going to infinity.

(6) The crucial element that distinguishes the two tests is that the FISHER test is based on combining the *significance levels* of the different tests, and the IPS test is based on combining the *test statistics*. Which is better is the question. We conducted Monte Carlo studies with these issues in mind.

(7) Both the FISHER test and IPS test are based on combining independent tests. So if there is contemporaneous correlation, then there are correlations among the individual test statistics. Both tests will need modifications in this case.

Applications of panel unit root tests have not been very careful about the appropriate interpretation of their results. The problem is serious because the null hypothesis says that *all* the series are unit root processes. Rejection of the null hypothesis does not mean that *all* the series are stationary. Similarly, a non-rejection of the null hypothesis does not mean that all the series are unit-root processes. A few of the series can tilt the conclusion either way.

Consider the following extreme case. We have $N = 10$ series. Suppose the *p*-values from the 10 individual unit root tests are close to 0.5 for the first 9 series and 0.000001 for the last series. Then the χ^2 test statistic will be very high, rejecting the unit root null. But for 9 of the series we don't reject the null and the overall rejection is based on a single strong rejection. The outcome will be (I guess) similar with the IPS test. But the problem here is not with the FISHER test but with the question asked. *The answer is meaningless because the question asked is meaningless.* The question asked is: “*What is the summary conclusion from all these 10 units root tests*” ? This is not a meaningful question when we can see clearly that for one series there is a very strong rejection, with no rejection for the others. What we should be doing is studying the one series which is an outlier. This points out the major defect of panel unit root tests. A summary statistic does not make sense if there are outliers. Thus the relevant information will always come from the individual unit root tests. The solution to the low power problem is not the panel data but more powerful unit root tests.

The Monte Carlo results reported in MADDALA and WU [1996b] suggest that the FISHER test performs slightly better than the IPS test in a wide variety of situations and the IPS better than the LL test. Of course, if there is very little heterogeneity in the autoregressive parameter, and quite high cross correlations, then the LL test would be the preferred choice.

In empirical applications, the LL test is the one most often used, with the IPS test and the FISHER-MADDALA-WU test used in a couple of applications.

3.5. **The Contemporaneous Correlation Problem**

In the presence of contemporaneous correlations among the errors neither the IPS nor the FISHER test is strictly valid. The IPS test depends on independent *t*-

statistics, and the FISHER test depends on independent p-values. Since the distributions of the test statistics are not known, MADDALA and WU [1996b] suggest obtaining the appropriate significance levels for the test statistics using the bootstrap procedure. They show that in the presence of correlated errors, the LL, IPS and FISHER tests all suffer from under-rejection of the null, although the size distortions are the least for the FISHER test. Also, using the bootstrap-based significance levels reduces these size distortions but does not eliminate them. The size distortions with the LL test in the case of cross-correlated errors have also been observed by O'CONNELL [1996] but he talks of dramatic size distortions (50 percent when the nominal size is 5 percent). The size distortions we observed were *not* as dramatic as those noted by O'CONNELL.

There is, however, one test statistic that is applicable in the cases of both the correlated and uncorrelated errors. This is the test statistic based on the BONFERRONI inequality (See ALT [1992]) and discussed in DUFOUR and TORRES [1996].

The idea behind this test is to break up the hypothesis $H_0 : \alpha_i = 1$ for all i, $i = 1,2,\ldots,N$ into a set of subhypotheses $H_{oi} : \alpha_i = 1$ and noting that H_0 is wrong if and only if any of its components H_{oi} is wrong. Suppose we choose the significance level γ_i for the i-th test. Then if we follow the rule that we reject H_0 if at least one of the subhypotheses H_{oi} is rejected at significance level γ_i, for H_{oi}, then the BONFERRONI inequality says that the significance level γ for H_0 is given by:

$$\gamma \leqslant \sum_{i=1}^{N} \gamma_i$$

One simple rule that DUFOUR and TORRES suggest is to take $\gamma_i = \gamma / N$, unless there is *a priori* compelling reason that some tests ought to be rejected at lower (or higher) significance level than the others.

MADDALA and WU [1996b] compare this test with the FISHER test through Monte Carlo experiments and find that the power of this test is much lower than that of the FISHER test, although there is a slight size distortion with the FISHER test.

3.6. Evidence on PPP from Panel Unit Root Tests

There has been an enormous amount of work on testing the long-run validity of the purchasing power parity (PPP) theory. This has been a real growth industry. The most commonly used tests for long-run PPP are tests for unit roots in real exchange rates. If the null hypothesis of a unit root can be rejected, then the real exchange rate is stationary and exhibits mean reversion, thus supporting long run PPP. An alternative route would be to test whether nominal exchange rate, domestic price level and foreign price level are cointegrated. Application of unit root tests to time series on individual currencies often leads to non-rejection of the unit root null, and thus rejection of long-run PPP. This has been attributed to the low power problem of unit root tests and it has been argued that panel unit root tests are more powerful and these lead to rejection of the unit root null in real exchange rates, thus supporting long-run PPP. For arguments in this vein, using the LL tests, see OH [1996], WU [1996] and MACDONALD [1996].

Given the limitations of the LL test noted earlier, it would be interesting to see what results we get using the IPS and FISHER tests. WU [1997] examines this in detail and comes to the conclusion that these tests as well, reject the unit root null, thus providing support for the PPP. O'CONNELL [1996], on the other hand, argues that correcting for the size distortions in the LL test results in a non-rejection of the unit root null, but the size distortions he talks about are much higher than the ones noticed by MADDALA and WU [1996b].

A more relevant issue is not whether PPP holds or does not hold but how long it takes for deviations from PPP to correct. Estimating fixed effects models based on a panel of 150 countries and 45 annual post World War II observations, FRANKEL and ROSE [1996] argue that there is strong evidence on PPP and that PPP deviation are eroded at the rate of approximately 15 percent annually, *i.e.* their half life is around four years, which is consistent with the evidence from long-horizon time series data.

However, LIU and MADDALA [1996] show that the conclusions of FRANKEL and ROSE are very sensitive to the choice of countries used in the panel and that the average half life is much higher. Also, about 24 countries from the panel used by FRANKEL and ROSE need to be deleted because of very few observations.

Many of the problems of biases, discussed by NERLOVE [1996, 1997a] have been ignored in the estimation of LSDV models in the FRANKEL and ROSE study. Making allowances for these, we would expect to find that PPP deviations are eroded at a much slower rate (similar to the much slower rates of convergence, noted by NERLOVE). We have not completed the work on this yet.

How would we reconcile this result with the results from the panel unit root tests (the LL, IPS as well as the FISHER tests) that show strong evidence in favor of PPP? The answer is that the unit root tests are based on the conditional likelihood function and we plan on exploring the results based on unit root tests derived from the conditional likelihood function. Thus, both in the estimation of the LSDV models and in the application of panel unit tests, we need to compare the results on the basis of the unconditional likelihood function.

The important thing to note about the panel data unit root testing in the search for PPP is that the tests are offered (and many papers using them have repeated this fallacious argument) that they "*solve*" poor power problem of the single equation unit root tests. As argued earlier, the power comparisons do not make sense when the tests do not share the same null and alternative. Also, the IPS test is offered as a generalization of the LL test but again this is an invalid comparison. The proper way to look at the IPS test is that it is a method of combining the evidence from several independent tests – a problem that was analyzed by R.A. FISHER in 1932. The FISHER-MADDALA-WU test makes this very clear. It is also more flexible in that it can be used with *any* unit root test (unlike the IPS test which is based on the ADF test which is not powerful). It also makes clear what exactly is being accomplished by the test whereas the IPS test masks it under the cover of it being more powerful than single equation unit root tests.

Thus, the application of the LL and IPS tests to the analysis of PPP do more harm than good because they confuse the issues.

4 Future Direction

"The dogs are barking, but the caravan keeps moving"
An old Russian proverb quoted by ZVI GRILICHES

The cross-country regressions are not going to be of any use to study the basic issue of whether the poor nations will catch up with the rich, an issue that goes under the glorified name of "*convergence*". However, in the name of convergence, much progress has been made on the estimation problems of dynamic panel data models. Possibly more has been accomplished on this front than in the understanding of the sources of economic growth where the cross-country regressions have come up with truisms such as: "*political freedom helps*", "*trade liberalization helps*", "*development of financial institutions helps*" and so on. (The only conclusion that is not a truism is on baseball *vs.* cricket as an engine of economic growth or the role of religion in economic growth.) In the future, the progress on the estimation of dynamic panel data models should continue and of course, it helps to use "*convergence*" as a peg, since so many famous economists are promoting work on convergence.

Those interested in the basic economic issue of how the poor nations catch up with the rich (or at least avoid getting poorer and poorer) should abandon cross-country regressions and study methods of promoting international diffusion of technology. Technology is playing a role not only in the disparity of the incomes of the poor *vs.* rich countries but also in the income disparity among different regions and groups in the US itself.

The search for purchasing power parity (PPP) is another search for a black cat in a dark room. Hundreds of papers have been written on this topic and when interest in this topic was waning, unit roots and cointegration gave a new life to this area. Panel data have given still more life. Here is a case where several dogs are barking but the caravan keeps moving. I have presented a brief review of the results of panel data unit root tests in search of PPP. More work using the dynamic panel data techniques suggested by NERLOVE [1997a, b] is under way.

As far as the development of new econometric methods is concerned, the search for PPP has led to methods of analysis of panel cointegration as in KAO and CHIANG [1997] and PEDRONI [1997a, b]. The application of panel cointegration tests for the analysis of PPP, however, is subject to the same criticism as discussed earlier in the application of panel unit root tests.

One other panel data set we have not discussed at length is the data set assembled and analyzed by TRUEBLOOD [1996]. Here a lot of effort has gone into cleaning up the basic data.

The search for a black cat in a dark room (in which there is no cat) seems to be productive after all (not in finding a black cat but in finding something else). Adrian PAGAN had the following definitions of an economic theorist, an econometrician and a BAYESIAN econometrician: An economic theorist is one who says he has proved the existence of a black cat in a dark room (in which there is no cat).

An econometrician is one who looks for a black cat in a dark room (in which there is no cat) and declares that he has found one.

A Bayesian econometrician is one who looks for a black cat in a dark room (in which there is no cat) and declares that he has found a cat and that the cat is white.

• References

BERNARD A.B., JONES C.I. (1996a). – "Productivity Across Industries and Countries: Time Series Theory and Evidence", *Review of Economics and Statistics*, 78, pp. 135-146.

BERNARD A.B., JONES C.I. (1996b). – "Technology and Convergence", *Economic Journal*, 106, pp. 1037-1045.

BERNARD A.B., JONES C.I. (1996c). – "Comparing Apples and Oranges: Productivity Convergence and Measurement Across Industries and Countries", *American Economic Review*, 86, pp. 1216-1238.

BARRO R.J. (1996). – "Democracy and Growth", *Journal of Economic Growth*, 1, pp. 1-27.

BEHRMAN J.R., ROSENZWEIGH M.R. (1994). – "Caveat Emptor: Cross-Country Data on Education and the Labor Force", *Journal of Development Economics*, 44, pp. 147-171.

BHALLA S.S. (1997). – "Freedom and Economic Growth: A Virtuous Cycle ?", *Paper presented at the Nobel Symposium on Democracy's Victory and Crisis*, Uppsala University, August, 1994.

BREITUNG J., MAYER W. (1994). – "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated ?" *Applied Economics*, 26, pp. 353-361.

CANOVA F., MARCET A. (1995). – "The Poor Stay Poor: Non-Convergence Countries and Regions", *working paper*, Universitat Pompeu Fabra.

CASELLI F., ESQUIVEL G., LEFORT F. (1996). – "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth*, 1, pp. 363-389.

CERMENO R. (1999). – "Median-Unbiased Estimation in Fixed Effects Dynamic Panels", in *Annales d'Économie et de Statistique*, this issue.

DORWICK S., GEMMELL N. (1991). – "Industrialization, Catching Up and Economic Growth", *Economic Journal*, 101, pp. 263-275.

DORWICK S., QUIGGIN J. (1997). – "True Measures of GDP and Convergence", *American Economic Review*, 87, pp. 41-64.

DUFOUR J.M., TORRES O. (1996). – "Union-Intersection and Sample-Split Methods in Economics with Application to MA and SURE Models", P*aper presented at the summer meetings of the Econometric Society*, Iowa City.

DURLAUF S., JOHNSON P. (1995). – "Multiple Regimes and Cross-Country Growth Behavior", *Journal of Applied Econometrics*, 10, pp. 365-384.

ELLIOTT G., ROTHENBERG T.J., STOCK J.H. (1996). – "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64, pp. 813-836.

EVANS P., KARRAS G. (1996). – "Convergence Revisited", *Journal of Monetary Economics*, 37, pp. 249-265.

FAGENBERG J. (1994). – "Technology and International Differences in Growth Rates", *Journal of Economic Literature*, 32, pp. 1147-1175.

FISHER R.A. (1932). – *Statistical Methods for Research Workers*, (Oliver and Boyd), Edinburgh, 4th Ed.

FRANKEL J., ROSE A. (1996). – "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries", *Journal of International Economics*, 40, pp. 209-224.

HART P. (1995a). – "Galtonian Regression Across Countries and the Convergence of Productivity", *Oxford Bulletin of Economics and Statistics*, 57, pp. 287-293.

HART P. (1995b). – "The Convergence of Labour Productivity in British and German Industries", *International Journal of the Economics of Business*, 2, pp. 453-464.

IM K., PESARAN M., SHIN Y. (1995). – “Testing for Unit Roots in Heterogeneous Panels”, *Working Paper*, No. 9516, Department of Applied Economics, University of Cambridge.

ISLAM N. (1995). – “Growth Empirics: A Panel Data Approach”, *Quarterly Journal of Economic*s, CX, pp. 1127-1170.

JUDSON R. (1996). – “Do Low Human Capital Coefficients Make Sense? A Puzzle and Some Answers”, *Finance and Economics Discussion Series*, 96-13, Board of Governors of the Federal Reserve System.

JUDSON R., OWEN A. (1997). – “Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists”, *Finance and Economics Discussion* papers#1997-3, Federal Reserve, Washington, D.C.

KAO C., CHIANG M.H. (1997). – “On the Estimation and Inference of a Cointegrated Regression in Panel Data”, *Manuscript*, Syracuse University.

KING R.G., LEVINE R. (1993a). – “Finance and Growth: Schumpeter Might be Right”, *Quarterly Journal of Economics*, 108, pp. 717-738.

KING R.G., LEVINE R. (1993b). – “Finance, Entrepreneurship and Growth: Theory and Evidence”, *Journal of Monetary Economics*, 32, pp. 513-542.

KIVIET J.F. (1995). – “One Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models”, *Journal of Econometrics*, 68, pp. 53-78.

KOCHERLAKOTA N.R., YI K.M. (1995). – “Can Convergence Regression Distinguish Between Exogenous and Endogenous Growth Models”, *Economics Letters*, 49, pp. 211-215.

LEE K., PESARAN M.H., SMITH R. (1995). – “Growth and Convergence: A Multi-Country Empirical Analysis of the Slow Growth Model”, *Working Paper*, University of Cambridge.

LEVIN A., LIN C.F. (1993a). – “Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties”, *Working Paper*, University of California at San Diego.

LEVIN A., LIN C.F. (1993b). – “Unit Root Tests in Panel Data; New Results”, *Working Paper* 93-56, University of California at San Diego.

LEVINE R., RENELT D. (1992). – “A Sensitivity Analysis of Cross-Country Growth Regression”, *American Economic Review*, pp. 942-963.

LEVINE R., ZERVOS S. (1993). – “What Have We Learned About Policy and Growth From Cross-Country Analysis”, *American Economic Review Papers and Proceedings*, 83, pp. 426-430.

LIU P.C., MADDALA G.S. (1996). – “Do Panel Data Cross-Country Regressions Rescue Purchasing Power Parity (PPP) Theory ?”, *Manuscript*, University of Miami.

LUCAS R.E. (1993). – “Making a Miracle”, *Econometrica*, 61, pp. 251-272.

MACDONALD R. (1996). – “Panel Unit Root Tests and Real Exchange Rates”, *Economics Letters*, 50, pp. 7-11.

MADDALA G.S. (1977). – *Econometrics*, McGraw-Hill, New York.

MADDALA G.S., LIU P.C. (1996). – “Panel Data Unit Root Tests: What Do They Test?” *Paper presented at the 6th International Conference on Panel Data*, Amsterdam.

MADDALA G.S., WU S. (1996a). – “A Bayesian Analysis of Cross-Country Growth Regression”, *Paper presented at the NBER-NSF Conference on Bayesian Inference in Economics and Statistics*, Chicago.

MADDALA G.S., WU S. (1996b). – “A Comparative Study of Panel Data Unit Root Tests and a Simple Alternative”, Paper presented at the Econometric Society Meetings, New Orleans, January, 1997, Forthcoming in *Oxford Bulletin of Economics and Statistics*.

MADDALA G.S., LI H., TROST R.P., JOUTZ F. (1997). – “Estimation of Short-run and Long-run Elasticities of Energy Demand From Panel Data Using Shrinkage Estimators”, *Journal of Business and Economic Statistics*, 15, pp. 90-100.

MANKIW N., ROMER D., WEIL D. (1992). – “A Contribution to the Empirics of Economic Growth”, *Quarterly Journal of Economics*, 107, pp. 407-437.

NERLOVE M. (1996). – “Growth Rate Convergence, Fact or Artifact?”, *Working Paper*, Department of Agricultural and Resource Economics, University of Maryland.

NERLOVE M. (1997). – “Properties of Alternative Estimators of Dynamic Panel Data Models: An Empirical Analysis of Cross-Country Data for the Study of Economic Growth”, *Manuscript*, University of Maryland.

NERLOVE M. (1999). – “Likelihood Inference for Dynamic Panel Data Models”, in *Annales d’Économie et de Statistique*, this issue.

O’CONNELL P. (1997). – “The Overvaluation of Purchasing Power Parity”, Harvard University, Forthcoming in the *Journal of International Economics*.

OH K.Y. (1996). – “Purchasing Power Parity and Unit Root Tests Using Panel Data”, *Journal of International Money and Finance*, 15, pp. 405-418.

PAPP R., VAN DIJK H.K. (1996). – “Distribution and Mobility of Wealth of Nations”, Tinbergen Institute Working Paper, Forthcoming, *Journal of Econometrics*.

PAPELL D.H. (1997). – “Searching for Stationarity: Purchasing Power Parity Under the Current Float”, *Journal of International Economics*, 43, pp. 313-332.

PEDRONI P. (1997a). – “Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time-Series Tests with an Application to the PPP Hypothesis”, *Manuscript*, April, 1997, Indiana University.

PEDRONI P. (1997b). – “Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity”, *Manuscript*, November, 1997, Indiana University.

PERRON P., NG S. (1996). – “Useful Modification to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties, *Review of Economic Studies*, 63, pp. 435-465.

QUAH D. (1992). – “International Patterns of Growth I: Persistence in Cross-Country Disparities”, LSE *Working paper*.

QUAH D. (1994). – “Exploiting Cross-Section Variation for Unit Root Inference in Dynamic Data”, *Economics Letters*, 44, pp. 9-19.

QUAH D.T (1996a). – “Empirics for Economic Growth and Convergence”, *European Economic Review*, 40, pp. 1353-1375.

QUAH D.T. (1996b). – “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics”, *Economic Journal*, 106, pp. 1045-1055.

RAMEY G., RAMEY V.A. (1995). – “Cross-Country Evidence of the Link Between Volatility and Growth”, *American Economic Review*, 85, pp. 1138-1151.

RAO C.R. (1952). – *Advanced Statistical Methods in Biometric Research*, Wiley, New York.

SACHS J.D., WARNER A.M. (1995). – “Economic Convergence and Economic Policies”, *discussion paper*#1715, Harvard Institute of Economic Research.

SALA-I-MARTIN X. (1997). – “I Just Ran Two Million Regressions”, *American Economic Review*, papers and proceedings, 86, pp. 178-183.

SALA-I-MARTIN X. (1996). – “The Classical Approach to Convergence Analysis”, *Economic Journal*, 106, pp. 1019-1036.

TRUEBLOOD M.A. (1996). – “Econometric Estimates of Inter-country Agricultural Productivity”, *Manuscript*, Department of Agricultural Economics, University of Minnesota.

WALL H.J. (1995). – “Cricket vs. Baseball as an Engine of Economic Growth”, *Royal Economic Society Newsletter*, July, 2-3.

WU, SHAOWEN (1997). – “Purchasing Power Parity Under the Current Float New Evidence form Panel Data Unit Root Tests”, *manuscript*, Department of Economics, Ohio State University.

WU, YANGRU (1996). – “Are Real Exchange Rates Stationary? Evidence from a Panel Data Test”, *Journal of Money, Credit and Banking*, 28, pp. 54-63.

ZANFORLIN L. (1996a). – “Technological Diffusion, Learning and Growth”, European University Institute *working paper*, ECO 96/23.

ZANFORLIN L. (1996b). – “Technological Assimilation, Trade Patterns and Economic Growth”, European University Institute *working paper*, ECO 96/24.