

June, 2004

**SOCIAL CAPITAL, BARRIERS TO PRODUCTION AND CAPITAL SHARES;
IMPLICATIONS FOR THE IMPORTANCE OF PARAMETER HETEROGENEITY
FROM A NONSTATIONARY PANEL APPROACH ***

Peter Pedroni

Williams College

mailing address: Economics, Williams College
Williamstown, MA 01267

email: peter.pedroni@williams.edu

Abstract: Recent advances in the growth literature have proposed that difficult to quantify concepts such as social capital may play an important role in explaining the degree of persistent income disparity that is observed among countries. Other recently explored possibilities include institutional mechanisms which generate barriers to aggregate production. An important limitation for empirical work in this area stems from the fact that it is difficult to distinguish sources of heterogeneity when direct observations are not available. In this study, we show how developments in the analysis of nonstationary dynamic panels can aid in this endeavor. In contrast to traditional panel data analysis, this approach focuses explicitly on low frequency behavior. Under relatively mild assumptions, the approach can be used to infer properties of aggregate production which are robust to the presence of large classes of unobserved features. In this framework we are able to estimate and test the implied distribution of production function parameters that would be required in order to generate conditional forecast convergence of per capita incomes even when some of the key factors required to explain growth are unobserved. The results indicate that in order to fully explain the observed persistence in the disparity of per capita incomes, the manner in which unobserved mechanisms influence production must go beyond merely accounting for differences in the trending behavior of aggregate productivity. Specifically, the results demonstrate that if such mechanisms are to be successful empirically, then they must also be able to account for cross country heterogeneity in steady state capital shares. This adds to a growing literature that provides support for models with multiple production regimes.

Keywords: Growth, Convergence, Social Capital, Nonstationary Panels
JEL Classifications: O41, C23

* An earlier draft version of this paper circulated with the title "Testing for Convergence to Common Steady States." I thank participants at the North American Winter Econometric Society Meetings, the Midwest Macro Meetings, the Midwest Econometrics Group Meetings, the 9th International Panel Data Conference, workshops at Cornell University, CUNY Hunter, U. Kansas, Ohio State, Rice, Syracuse, Union, Williams, the Conference on Panel Cointegration at the University of Maastricht, Netherlands and the Conference on Growth and Time Series at the University of Guanajuato, Mexico, and especially Doug Gollin, Sau-Him Paul Lau, Tong Li, Stephen Parente and Jean-Pierre Urbain for helpful comments and suggestions.

Much of the empirical growth literature has been devoted to trying to explain the enormous disparity of per capita incomes that is observed around the world. It is often noted that the wealthiest economies have a level of per capita income that is well over 30 times greater than that of the poorest economies. Another important and often noted feature is that these disparities appear to be very persistent over time, and it is difficult to discern under what conditions we might anticipate that these differences are likely to expand or diminish. A full explanation for this disparity and its apparent persistence has eluded economists despite an enormous amount of research devoted to the topic.

Nonetheless, several important themes have emerged in recent years which have brought us closer to understanding this phenomenon. For example, economists have long noticed that variations in the share of physical investment resources that are devoted to production cannot alone explain the disparities that are observed in per capita output. Easterly and Levine [2001] provide a recent survey of some of these stylized facts. This has led many to conjecture that certain forms of intangible capital, most notably human capital in the form of education, may play an important role in the explanation. Early empirical work by authors such as Barro [1991] and Mankiw, Romer and Weil [1992] demonstrated that accounting for levels of education goes a long way toward explaining the observed cross sectional variations in per capita income. More recently, many researchers have asked whether human capital alone may not be enough to explain the disparity, and have looked toward more pervasive forms of intangible capital for explanations, such as what some have termed “social capital”. Although researchers differ on the precise meaning or intent, predominantly the concept has come to refer to intangible forms of capital such as social and political institutions, or even the presence of social norms and morals. The basic idea is that by fostering the productivity of more conventional categories of productive inputs, these relatively difficult to quantify notions of capital may help to explain why variations in measured inputs alone do not appear to be sufficient to explain the persistent long run differences in per capita incomes that are observed across countries. Important examples of recent empirical research in this area include among others Hall and Jones [1999] and Temple and Johnson [1998].

Another line of research which shares similar overtones in some regards has attempted to explain more directly the reasons why different countries do not appear to use the best available aggregate production technologies. This line of research, exemplified in recent years by the works of Parente and Prescott [1994, 1999] and Prescott [1998] emphasizes the importance of barriers to efficient production. Similar to the literature on social capital, this literature appeals to country specific institutional and political economic reasons to explain why traditional forms of intangible capital such as the level of education are not sufficient to explain the observed disparity in per capita incomes. The idea

is that institutional and political factors can create incentives for firms not to employ the most efficient means of production technology. A central implication of this line of research is that despite a common set of globally available production technologies, country specific barriers to the adoption of new technologies lead to a situation in which total factor productivity may in effect grow at different rates across countries, depending on the nature and strength of such barriers. This line of research helps to focus attention on the fact that it is the different rate of absorption of new technologies that may help to explain why the variation in measured inputs alone cannot explain the persistent dispersion of per capita incomes across countries.

At a basic level, each of these approaches rely on the argument that certain intangible or unobserved factors may help to explain the apparent inability of variations in observed factors to explain the disparity in observed per capita incomes in a manner consistent with the basic neoclassical model with shared production technologies. Another important tradition in the literature takes a somewhat different approach, and seeks to explain observed differences in per capita income relative to observed inputs by attempting to understand why the process of economic development may lead to effectively different aggregate production functions, depending on the level of development experienced by the country. Early examples of this line of research include among others the work of Azariadis and Drazen [1990], which emphasizes the possibilities of threshold externalities associated with the level of human capital present, and the works of Durlauf [1993] and Murphy, Schleifer and Vishney [1989] which emphasize how coordination failures can produce the possibility of multiple steady states associated with different aggregate production parameterizations of factor shares. A central implication of this line of research is that a single log linear specification implied by common aggregate production technologies may not be sufficient to characterize the steady state evolution of per capita income in terms of factor inputs.

Many researchers, including recently Bernard and Jones [1996], Durlauf and Johnson [1995], Caselli, Esquivel and Lefort [1996], Lee, Pesaran and Smith [1997, 1998] and Masanjala and Papageorgiou [2004] have noted that, empirically, capital factors shares appear to vary considerably across countries. For example Durlauf and Johnson [1995] employ a regression tree analysis to show that a cross sectional regression of the Summers and Heston data appears to provide support for several distinct regimes in which aggregate production functions vary among countries according to their level of development. As Durlauf and Johnson [1995] point out, however, a cross sectional based approach is inherently limited in its ability to deal with omitted variables such as social capital, which are difficult to

quantify. The reason for this is straightforward. In a conventional cross sectional based approach, omitted variables that are important for conditioning may induce apparent differences in the slope coefficients of the linear regression fit. As we will see, the nonstationary panel approach employed in this study will permit us to examine the distribution of key slope coefficients across countries which will be invariant to a broad class of such omitted variables. Consequently, the approach taken in this paper provides an alternative that helps to provide evidence that is more robust to omitted variables than traditional cross sectional based methods and therein compliments the growing literature that finds evidence for the importance of parameter heterogeneity and the need to consider specifications beyond the common linear production function. It also allows us to examine in more detail the full sample distribution of implied production function parameters.

It is important to distinguish a panel approach that exploits the nonstationary features of the data from more traditional panel approaches. For example, while Durlauf and Johnson [1995] conclude that a panel approach may be required in order to more fully investigate the nature of the cross country variations in these coefficients, Durlauf and Quah [1999] point out that conventional panel approaches also have important limitations in that they may inadvertently tend to uncover high frequency relationships rather than the long run low frequency relationships that are of interest for the growth literature. The reason for this is also straightforward. In conventional panels which involve stationary variables, the fixed effects tend to absorb those features of the sample which evolve at relatively low frequencies, leaving the regressors to explain the higher frequency relationships.¹

By contrast, the nonstationary approach that we employ explicitly estimates the long run low frequency relationships among the variables, despite the presence of fixed effects. As we discuss below, the approach also allows us to address many of the other issues that have been important for empirical work in the growth literature. For example, our approach allows us to relax many important exogeneity assumptions, and also allows us to relax the assumption that countries stay relatively close to their steady state positions at all points in time. For these reasons, we believe that recent developments in nonstationary panel techniques offer considerable promise in further helping us to distinguish among empirical relationships which can aid researchers to better pinpoint fruitful directions for the development of growth theory. In this study, we demonstrate how such nonstationary panel techniques can be used to investigate the distribution of coefficients which reflect key structural parameters of the production function in a way that accounts for the possibility of intangible social capital, barriers to production and multiple regimes in explaining

persistent per capita income disparities across countries. Furthermore, we do so in a framework which permits us to link this directly to a notion of conditional convergence, as we will see. The results of this study add to a growing empirical literature that points to the importance of relaxing the traditional assumption of a common linear production function, and argues that an appeal to omitted factors alone is unlikely to suffice as a substitute for parameter heterogeneity.

A. Empirical Analysis Based on Nonstationary Panel Methods

The literature on the theory and application nonstationary panels methods has expanded rapidly in recent years. Recent surveys of this literature include Baltagi and Kao [2000], Banerjee [1999] and Phillips and Moon [2000], each of which includes discussions of some of the techniques used in this paper. In contrast to conventional panel techniques, nonstationary panel techniques deal explicitly with the low frequency nonstationarities that are typically present in the individual time series that constitute the members of the panel. In this regard, it can be argued that the approach is particularly well suited for the empirical growth literature in that it focuses explicitly on the stochastic and nonstochastic long run trending features of the data that are key to understanding the forces that shape economies over long periods of time, while allowing us to filter out the effects of short run transitional dynamics. Another important attraction in this approach is that it offers alternative solutions to the issue of simultaneity that has plagued much of the empirical growth literature. In many cases, the approach will allow us to obtain estimates that relate to the key structural parameters of the production function which are robust to the fact that many of the determinants of per capita income are in turn endogenously determined by the level of per capita income.

It is important to recognize that by using a nonstationary panel approach we are implicitly assuming that the pattern of low frequency variations observed over time across a group of countries can aid us in understanding the nature of the cross sectional dispersion of per capita incomes. The basic reasoning that connects our empirical approach to the issue is relatively straightforward. The Solow growth model tells us that in the long run steady state, permanent changes in the savings rate will be associated with permanent changes to the level of per capita income and that the relationship is determined by the capital share parameters of the Cobb-Douglas production function. Since the share of investment is our best measure for savings rates, this tells us that if the pattern of time series variations is going to be informative

about this relationship, then we should search for instances in which permanent variations in the investment share have occurred. These permanent changes in investment shares may be due to changes in tastes, changes in government policy, or any number of reasons which may differ from case to case. What is important is simply that such permanent variations in investment shares are present so that we have the potential to observe any permanent comovements that have been observed in per capita income.

One practical way to model these permanent variations in investment shares empirically when they occur is to treat them as changes to the expected long run values, so that the series for investment shares mimics a unit root process, which is non mean reverting. Clearly, it does not make sense to think of this as a global property of investment shares, since ultimately investments shares must be bounded by the resources available to an economy. Rather, we prefer to think of this simply as a feature which describes the local behavior of the series within sample, and thus has implications for the properties of the estimators that are constructed from these samples. Indeed, Jones [1995] points out that for many countries, with the notable exception of the U.S., the series for investment shares appear to be consistent with such unit root behavior within sample. We also find this to be the case for many countries, though certainly not all countries, and as we explain below, we use this as a criteria for selecting countries from within the Summers and Heston panel which are most likely to be informative for the questions we ask.

As we have noted, the reason that the presence of such permanent changes to savings rates, or by proxy investment shares, are particularly informative is because the Solow model tells us that these should in turn be associated with permanent changes to the level of per capita income, which tells us that we should expect to find a cointegrating relationship between the two variables of the panel. Furthermore, as we demonstrate in section III, when the level of per capita effective human capital or other intangible capital inputs are relatively stable across countries, or evolve relatively smoothly over the long run, then these features will be absorbed into the fixed effects or heterogeneous deterministic trends for our panel specification, depending on the case. In such cases, the steady state relationship between per capita income and investment shares characterized by the cointegrating relationship traces out the implied curvature of the production function for each country as determined by the capital share parameters of the tangible and intangible capital inputs. As we argue in section III, in this way our approach has the potential to uncover structural characteristics of the production function which under a broad class of behaviors are robust to the presence of unmeasured intangible capital inputs and mechanisms such as barriers to production that may lead to differential rates of

adoption of technological growth. Furthermore, the approach is robust to the violation of a number of strong assumptions that have typically been made in the literature, including the assumption that regressors and omitted initial conditions are exogenous, and that countries lie close to their steady state positions at all points in time. Consequently, by examining the distribution across countries of the implied production function parameters that are obtained in this manner, we can also ask whether or not these patterns correlate systematically with features of the economy that may be helpful in explaining why we might expect such a distribution.

Finally, it is interesting to point out that this approach also provides us with a convenient link to the literature on conditional convergence. Specifically, since the residuals of a cointegrated panel are by construction zero mean stationary processes, these fit the criteria for time series forecast convergence in the sense that their long run forecast differences go to zero as described in Bernard and Durlauf [1996] for univariate series. Consequently, any remaining disparity must be purely transitory, and we see that under this interpretation, the cointegrating relationship for the panel picks out precisely those features upon which it is necessary to condition per capita incomes in order to obtain convergence in this sense. In this way, the approach has the ability to address in a well defined context the question of whether features such as unmeasured intangible capital and barriers to production are sufficient to explain empirically the dispersion of per capita incomes within the context of the conventional neoclassical model with common aggregate production functions or, alternatively, whether we must also explicitly account for the heterogeneity of factor shares in order to explain the degree of persistent income dispersions.

The results of this study provide us with some intriguing empirical implications. In particular, while the results point to an important role for unobserved mechanisms such as intangible capital and barriers to production that account for differences in long term productivity growth, such mechanisms alone do not appear to be entirely sufficient to obtain conditional convergence in the framework of the neoclassical model with linear homogeneous production functions. Rather, in order to obtain conditional convergence, our primary results point to a need for mechanisms that can also explain the heterogeneity of capital shares in the neoclassical production function. In broad terms, the finding could be said to be consistent with the literature on multiple regimes. On the other hand, when we further examine the specific pattern of distribution among capital shares, we find that the heterogeneity is pervasive across all levels of development, and cannot easily be explained by categorizing countries into income categories or categories for the levels of unmeasured factors. For example the pattern cannot easily be explained by correlates that might predict the presence of

threshold effects due to human capital accumulation or the general level of development as indicated by the level of per capita income. However, it is interesting to note that we do find that the pattern of distribution appears to be correlated significantly with measures that have been used to proxy for social capital. These features could point to a potentially important role for certain forms of social capital which goes beyond simply enhancing the productivity of other factors, as one might expect for standard capital inputs. While many interpretations are possible in this context, the pattern points to the importance of considering frameworks beyond the common linear production function and is consistent with the possibility that social capital may play a further role in determining the effective heterogeneous parameterization of the aggregate production function.

The remainder of the paper is organized as follows. In the next section we discuss the details of our strategy for selecting an appropriate sample from the Summers and Heston data. In section III we discuss the details of our estimation procedure as it relates to interpretations of the steady state representation of the Solow model. In section IV we characterize and interpret the cross country distribution of capital share estimates obtained from our procedure. Section IV ends by investigating empirically a few possible explanations for the particular distribution pattern for the implied share parameters, and section V offers some concluding remarks.

II. Sampling from the Summers and Heston Panel

In this section we discuss our empirical strategy for selecting an appropriate sample from the Summers and Heston panel. Unfortunately, in the rush to experiment, it is sometimes the case that new techniques are applied without sufficient regard to their applicability to a particular sample. We seek to avoid this pitfall, and believe it is worth discussing in detail the approach that we take here.

An important argument for the informativeness of using both the time series and cross sectional dimensions of the panel is that the neoclassical model tells us that permanent changes in savings rates should be associated with permanent changes in per capita incomes, and that this relationship should help us to trace out the curvature of the production function, and in turn the cross sectional pattern of this curvature among different countries. Permanent changes to savings rates may reflect changes in tastes, changes in government policy or any of a number of reasons which may differ from country to country. If we measure these in terms of investment shares, then we can think of these

as changes to the long run expected value in investment shares, in which case we can characterize this behavior by modeling investment shares as a unit root process, which is non mean-reverting. Jones [1998] has also pointed out that, with the notable exception of the United States, for many countries of the Summers and Heston panel, investment shares appear to behave as a unit root process. The important point in our interpretation of this feature of the data is that there is nothing which necessarily indicates that investment shares *must* behave as a unit root process. The feature depends on what has determined tastes, government policy and so forth in the post war period, which we anticipate to be relatively country specific and sample specific.

For some countries, these factors have been relatively stable and have lead investment shares in turn to be relatively stable. Other countries have had experiences that have lead to relatively frequent revisions in the expected long run investment share, which leads us to observe investment share series with permanent movements characteristic of a unit root process. Since the neoclassical model in conjunction with our econometric technique tells us that these permanent changes to investment shares can help us to uncover the implied parameters of the production function when they are associated with permanent changes in output, our empirical strategy is to look for instances in which this pattern has occurred in the Summers and Heston panel.² Thus, we have selected from the Summers and Heston data a panel of countries for which the investment share series are likely to be consistent with this feature. In doing so, our presumption is that the events that lead to unit root behavior of the investment shares is relatively pervasive in the data, but that it need not be present in all countries. Consequently, we begin by using a simple unit root test as our primary selection tool in determining whether or not to consider including a country in our core sample.

In practice, this means that for the 51 countries of the Summers and Heston panel for which the full span of data is available from 1950 to 1992 for investment shares and per capita output, we eliminate from our core sample any countries for which the investment share series does not pass a simple ADF unit root test. In particular, this sample selection procedure initially results in the selection of a total of 31 countries, of which 14 are OECD countries and 17 are non-OECD countries. By setting aside the other countries, we have in effect eliminated those countries for which the investment series is unlikely to be informative about the relationship in which we are interested. This has the effect of dramatically increasing the signal to noise ratio in the sample with which we are working. Of course, our sample selection procedure is based on a test which takes the null hypothesis to be a unit root, and one might argue that it is more appropriate to use a test which takes the null hypothesis to be stationarity. In other words, one might argue that

rather than dropping those countries for which the unit root null is rejected, we should drop the countries for which stationarity is rejected. In fact, we also compute results for the standard KPSS stationarity test applied to the individual country series and the panel as a whole, and we also conduct our subsequent analysis for the subset of 23 countries for which this stronger criteria applies and find similar results. However, our general view is that in light of the relatively weak power of either of these tests for any series which spans only 43 years of data, it would be excessive to demand that we strictly reject stationarity. Rather, we simply wish to set aside those countries for which the evidence appears to run contrary to the presumption that permanent changes to the savings rate have occurred that are consistent with a unit root process, with the recognition that while this still leaves important sources of noise in the sample, it has substantially reduced its prevalence.

Furthermore, to support the idea that the inclusion of this number of countries in our sample is not due simply to lack of power in the individual country ADF test, we also compute the panel based tests of Im, Pesaran and Shin [2003]. This test allow us to test the null hypothesis that all members of the panel have a unit root versus the alternative that some fraction are stationary. The results of this test, which are also reported in table I, are unable to reject the hypothesis that all countries have a unit root, despite the fact that this test is known to have high power in panels of this dimensionality. On the other hand, as expected, we note that conversely, a panel test for the null of stationarity does reject the hypothesis that all countries are stationary against the null that a fraction contain unit roots. We computed this panel test for the null of stationarity by constructing a group mean test based on the averaged Kwiatkowski *et al* [1992], KPSS, test values standardized by appropriate correction terms for the expected value and standard deviation of the nuisance parameter free limit distribution of the KPSS statistic.³

According to the Solow model, these permanent changes in investment shares should be associated with permanent changes in per capita income, which should be reflected as a cointegrating relationship between the two variables. Consequently, we will want to check for this condition as well in selecting our sample. One could imagine first testing each of the individual per capita GDP series, and then testing for cointegration. But establishing a unit root for per capita income would actually be redundant once we establish a unit root for investment shares individually and cointegration between the two. Therefore, since we are more directly concerned with the cointegrating relationship, we base the second step of the sample selection procedure on the cointegration test between the two variables. Again, given the low power of the individual tests, to be consistent with the way in which we based the decision for the unit root tests

on investment shares, here we will want to eliminate any countries for which we can reject the null of cointegration on an individual basis. For this, we used the Shin [1994] residual based test for the null of cointegration, which is analogous to the KPSS test for univariate series. Since our interpretation of the steady state regression in section III allows for the possibility of country specific deterministic trends and intercepts, we included these in the tests. This procedure lead us to reject an additional 2 countries from the sample, namely Thailand and Uruguay, which leaves us with a total of 29 countries, 14 of which are OECD countries and 15 of which are non OECD countries.⁴

Furthermore, to support the idea that again the inclusion of this number of countries in our sample is not due simply to lack of power in the individual country tests, we also computed a panel test for the null of cointegration by constructing a group mean panel test analogous to the one that we used for the KPSS based group mean panel test for stationarity. The only difference is in the value of the adjustment terms used to construct the panel statistic, since these must be based on the estimated residuals of a cointegrating regression under the null.⁵ The test confirms that null of cointegration is also not rejected for the sample treated as a panel. Likewise, we also computed individual tests for the null of no cointegration, as well as panel tests for the null of no cointegration using the tests from Pedroni [2004], which were extended to the multivariate case in Pedroni [1999]. For all of the results reported in table I, we also provide results for the case in which the data has been demeaned with respect to cross sectional dimension for each time period, which is comparable to the inclusion of time dummies. These serve to capture certain forms of cross sectional dependency, and the effect of common disturbances that impact all members of the panel. Consequently, when the data has been demeaned in this way, we can also interpret the results as relative to any common stochastic trend that drives the different countries of the sample. It is important to account for the possibility of such common effects when computing the panel statistics, and in table I, the null of no cointegration is rejected in all cases for the panel once we demean relative to the time effects.

The list of all countries included in the sample based on the unit root and cointegration tests can be inferred from table II.⁶ Note that in addition to each of the tests that we formally used in the sample selection process, in all cases we have also computed tests that reverse the null hypothesis. Thus, while we use tests for the unit root null and the cointegration null to select countries for our sample, for the sake of interest, we also report results in all cases for tests of the null of stationarity and the null of no cointegration. We also computed results for unit root and stationarity tests individually for both investment shares and per capita income, and constructed group mean panel tests in all cases.

For unit root and cointegration tests that involve per capita output, we have allowed for a deterministic trend and intercept, whereas for unit root tests involving only investment shares we did not include a deterministic trend, but have allowed for heterogeneous fixed effects in the intercepts. The panel tests for per capita income are unable to reject the unit root null hypothesis, but strongly reject the stationary null hypothesis. These findings for per capita income are consistent with numerous other studies, including Canning and Pedroni [1999], Cheung and Lai [2000] and Lee, Pesaran and Smith [1997] among others.⁷ Further details of the results are reported in table I.

III. Estimating and Interpreting the Steady State Relationship

Now that we have examined and selected an appropriate core sample from the Summers and Heston data, we are in a position to discuss the estimation and interpretation of the typical Solow steady state regression from the perspective of a nonstationary panel data analysis. We begin by describing our approach in relationship to what can be learned from a cross sectional approach, and then show how the various mechanisms that explain per capita income dispersion can be represented in a nonstationary panel framework.

In order to set notation, recall that the typical augmented Solow model analysis specifies an aggregate Cobb-Douglas production function such as

$$(1) \quad Y_{it} = K_{it}^{\alpha_i} X_{it}^{\phi_i} (E_{it} L_{it})^{1-\alpha_i-\phi_i}$$

where the share parameters are taken to be common across economies $i = 1, \dots, N$, so that $\alpha_i = \alpha$, $\phi_i = \phi$ for all i .

The variable X_{it} is an intangible capital input such as human capital, and E_{it} is the labor augmenting level of productivity. Labor and productivity evolve as

$$(2) \quad L_{it+1} = (1 + n_i) L_{it} \quad , \quad E_{it+1} = (1 + g_i) E_{it}$$

where n_i and g_i are the exogenous growth rates of labor and productivity, and g_i is typically assumed to be common across countries so that $g_i = g$ for all i . Likewise, physical capital grows according to the accumulation equation

$$(3) \quad K_{it+1} = I_{it} + (1 - \delta_i) K_{it} \quad , \quad I_{it} = S_i Y_{it}$$

where δ_i is the constant rate of capital depreciation, and I_{it} is the rate of physical capital investment determined in the Solow model as a share of output given by the savings rate S_i .

When the model is augmented with an endogenous intangible capital such as human capital, then this is taken to evolve as

$$(4) \quad X_{it+1} = I_{it}^x + (1 - \delta_i^x)X_{it}, \quad I_{it}^x = S_i^x Y_{it}$$

where δ_i^x is the rate of depreciation of the intangible capital stock, and the rate of investment in this stock is determined by a constant fraction of output S_i^x . The model can then be solved for the steady state value of log per capita income, $\ln y_{it} \equiv \ln Y_{it} - \ln L_{it}$, in terms of log investment shares, $\ln(I/Y)_{it} = \ln S_i$, which gives

$$(5) \quad \ln y_{it} = \ln e_{io} + \phi_i (1 - \alpha_i)^{-1} \ln x_{it}^* - \alpha_i (1 - \alpha_i)^{-1} \ln(n_i + g_i + \delta_i) \\ + g_i t + \alpha_i (1 - \alpha_i)^{-1} \ln(I/Y)_{it}$$

as an approximation when the values of g_i and δ_i are relatively small. Here e_{io} is the initial condition for labor augmenting productivity, and $\ln x_{it}^* \equiv \ln X_{it} - \ln E_{it} - \ln L_{it}$ is the steady state value of the intangible capital stock in log per capita efficiency units.

A. Interpreting Cross Section Regressions

Many of the empirical contributions to the interpretation of the steady state relationship between per capita incomes and investment shares have been made in the context of cross section regressions, and it is instructive to compare these to the framework we use in this study in order to gain intuition for what can be learned from the nonstationary panel approach. In the typical cross section estimation of the steady state regression, annual per capita incomes are measured at a point in time and regressed against time averaged values for investment shares along with other variables that serve to proxy for the effect of the intangible capital stock or that serve to proxy for initial conditions of labor productivity. Technology is typically assumed to be common across countries, so that the share parameters of the production function, $\alpha_i = \alpha$, $\phi_i = \phi$ as well as the growth rate of productivity, $g_i = g$, and the rate of capital

depreciation, $\delta_i = \delta$, take on common values for all countries.

Furthermore, in these typical cross section implementations, the savings rates and population growth rates are constructed using time averaged values, so that $\bar{s}_i = T^{-1} \sum_{t=1}^T (I/Y)_{it}$ and $\bar{n}_i = T^{-1} \sum_{t=1}^T n_{it}$. Any differences in initial conditions for productivity across countries, $\ln e_{io}$, are assumed to be uncorrelated with $\ln \bar{s}_i$. Thus, initial conditions for productivity are implicitly assumed to take the form $\ln e_{io} = \ln \tilde{e}_o + \ln \tilde{e}_{io}$, where the country specific differences $\ln \tilde{e}_{io}$ are uncorrelated with $\ln \bar{s}_i$ and are absorbed into the regression residuals, and the global average level of productivity $\ln \tilde{e}_o$ is absorbed into the regression intercept. Finally, in the typical cross section implementation, $\ln \bar{s}_i$ is itself assumed to be strictly exogenous, and the economies are assumed to be close to steady state at all points in time, so that the effect of transitional dynamics can be ignored.

When all of these assumptions hold, in the absence of intangible capital, X_{it} , the relationship between the log of time averaged investment shares and log per capita income can be used to identify the share of physical capital from the slope coefficient of a cross sectional regression that takes the form

$$(6) \quad \ln y_i = c + \beta (\ln \bar{s}_i - \ln(\bar{n}_i - g - \delta)) + \mu_i$$

where $\beta = \alpha(1 - \alpha)^{-1}$. For example, under these assumptions, Mankiw, Romer and Weil [1992] set $g + \delta = 0.05$, and use the Summers and Heston panel to construct country specific values for \bar{s}_i and \bar{n}_i based on sample averages from 1960 to 1985, and use 1985 values of per capita income for y_i . Based on this regression, Mankiw, Romer and Weil [1992] point out that the resulting slope coefficient implies a capital share that is far too large to be plausible, with values ranging from 0.60 to 0.59 for two subsamples of non OECD countries, and 0.36 for the subsample of OECD countries. The authors contrast this result with the case in which the Solow model is augmented to include human capital, and a proxy for the country specific rate of human capital savings. To see this, note that equation (5) can be rewritten in terms of the rate of savings, S_i^x , for the intangible capital stock, so that

$$(7) \quad \begin{aligned} \ln y_{it} = & \ln e_{io} + \phi_i (1 - \alpha_i - \phi_i)^{-1} \ln S_i^x - (\alpha_i + \phi_i) (1 - \alpha_i - \phi_i)^{-1} \ln(n_i + g_i + \delta_i) \\ & + g_i t + \alpha_i (1 - \alpha_i - \phi_i)^{-1} \ln(I/Y)_{it} \end{aligned}$$

By making the same set of assumptions required for the typical cross section implementation and adding the additional

assumption that country specific differences in initial productivities, $\ln \tilde{e}_{io}$ are also uncorrelated with average enrollment rates, \bar{s}_i^h , which in turn are assumed to be strictly exogenous, Mankiw, Romer and Weil [1992] estimate the cross section regression

$$(8) \quad \ln y_i = c + \beta_1 (\ln \bar{s}_i - \ln(\bar{n}_i - g - \delta)) + \beta_2 \ln S_i^h + \mu_i$$

where now $\beta_1 = \alpha(1 - \alpha - \phi)^{-1}$ and $\beta_2 = \phi(1 - \alpha - \phi)^{-1}$. On this basis, the estimated slope coefficients imply a much smaller physical capital share value, with α ranging from 0.31 to 0.29 for the two non-OECD subsamples, and 0.14 for the OECD subsample, where this last value is statistically indistinct from zero due to the high standard error.

In this regard, it is interesting to note that our sample produces very similar results for the typical cross section regression specification. For example, when we estimate the regression in equation (6) using our selected sample, we obtain a slope estimate which implies a capital share of 0.57. This is despite the fact that our data includes a different subsample of countries and spans a different length of time,⁸ which helps to further reassure us that aside from the specific nonstationarity properties, the subsample that we have selected is not unusual relative to the full Summers and Heston panel. Likewise, when we estimate equation (8) by including the same values for schooling, S_i^h , that Mankiw, Romer and Weil [1992] use, then the slope coefficient β_1 falls relative to our estimate of β from equation (6), and the corresponding implied physical capital share value similarly falls to 0.37 for our sample of 29 countries.

Several studies have noted that social indicators that can be interpreted to proxy for various forms of what may be social capital tend to be strongly correlated with per capita income levels, and tend to substantially improve the cross section regression fit. For example, Hall and Jones [1999] show that various proxies for social infrastructure have explanatory power for the large variations in the Solow residuals across countries that reflect output differences that are not accounted for by measured inputs. Likewise, Temple and Johnson [1998] show that indices of social development constructed in the 1960's have considerable predictive power for subsequent economic growth. These studies help to point out that there may be other forms of unmeasured capital inputs in addition to human capital that may enhance the productivity of measured inputs, and thus help to explain the inability of variations in observed inputs alone to fully explain per capita income differences across countries.

On the other hand, studies such as Parente and Prescott [1994] and Prescott [1998] also provide convincing

arguments for why we might anticipate that the role of unmeasured intangible capital stocks alone may not be sufficient to fully explain the global disparity of per capita incomes. These studies use a calibrated version of the neoclassical model with common Cobb-Douglas aggregate production functions to show that in order to explain the global dispersion of per capita incomes, investment shares in unmeasured intangible capital would either need to exceed 100% of GDP in order to imply reasonable rates of return, or would need to imply excessively high returns of greater than 40% to imply more reasonable investment shares. Instead, the authors argue that only explanations of the differences in total factor productivity can account for observed per capita income disparities. The authors argue that country specific institutional barriers to technological adoption that are maintained through monopoly rights may be able to explain apparently differing rates of total factor productivity at the aggregate level, despite the fact that the technological knowledge is globally available. These types of arguments help to highlight the possibility that some mechanisms that seek to explain persistent per capita income disparities may operate through implied differences in productivity growth itself, rather than through unmeasured intangible capital stocks. In this case, one needs to consider the possibility that the mechanisms for explaining per capita income disparities may lie in economic or political factors which produce country specific rates of productivity growth. With this in mind, we will want to take care in our specification of the nonstationary panel analysis to be able to account not only for the consequences of unmeasured forms of intangible capital, but also for the possibility of country specific factors that impact productivity growth.

Another line of research which we will want to keep in mind is the literature on threshold effects, which can produce conditions in which the production function itself may vary depending on country specific conditions, despite the fact that the same technology is available globally. For example, Azariadis and Drazen [1990] develop a model in which spillovers from individual's human capital investments lead to externalities which can trigger threshold effects, so that economies with varying levels of human capital may in effect be characterized by aggregate production functions with different share parameters. Similarly, Durlauf [1993] and Murphy, Schleifer and Vishney [1989] develop models in which spillover effects in the production process lead to comparable predictions for the associated aggregate production functions, which may depend on the level of development. If aggregate production functions vary across countries, then this provides yet another avenue by which the effects of relatively small differences in investment shares can be amplified into larger dispersions of per capital incomes across countries.

In support of this literature, Durlauf and Johnson [1995] use a regression tree procedure to provide compelling

evidence in support of this literature. Using the Summers and Heston data, they show that the Mankiw, Romer and Weil [1992] cross sectional regression provides support for several distinct regimes in which aggregate production functions vary among countries. In particular, Durlauf and Johnson find evidence for four different characterizations of the aggregate production function in term of physical and human capital shares among these groups of countries. The evidence is compelling in its support for the possibility of multiple regimes characterized by aggregate production functions with differing production functions. However, the approach is still subject to the usual assumptions that are required for the typical cross section implementation. More fundamentally, a disadvantage inherent in a cross sectional based approach to looking for evidence of multiple regimes stems from the fact that with only one dimension to work with, it is difficult to distinguish between true heterogeneity of the slope coefficients and apparent heterogeneity due spuriously to an omitted variable that may be responsible for determining the intercept coefficient. To give a simple example, imagine that something like social capital plays a role in determining the level of per capita income for any given investment share. In a cross section regression, if this unmeasured capital is omitted, it can create the illusion of a different slope coefficient for groups of countries with different levels of social capital. Another inherent limitation due to the single dimension of a cross sectional approach is that the number of categories for different linear specifications must be relatively small in order to ensure enough observations for any one category, so that the differences can be tested for statistical significance. Durlauf and Johnson [1995] ultimately conclude that panel studies may be needed to further sort out some of these issues, and we now turn to a discussion of how such steady state relationships can be interpreted in the context of a nonstationary panel regression.

C. Interpreting Cointegrated Panel Regressions

Our hope is that by using a nonstationary panel specification we can further sort out some of these issues in a manner that is consistent with the possibility that unobserved factors such as intangible capital and barriers to production are likely to play an important role. Indeed, the extra dimensionality of the panel and the emphasis on long run relationships that a nonstationary panel specification provides affords us with an opportunity to account for the effects of each of these channels while investigating properties of the implied cross sectional distribution of structural parameters associated with the production function. The cointegrated panel specification is attractive in part for the simplicity of its basic form, which can be written in terms of observable per capita income and investment shares as

$$(9) \quad \ln y_{it} = c_i + g_i t + \beta_i \ln(I/Y)_{it} + \mu_{it}$$

where we have chosen a subsample of the Summers and Heston panel for which the time series properties of this relationship can be characterized in terms of a cointegrating relationship so that μ_{it} is stationary with mean zero. The fact that the residuals μ_{it} are stationary with zero mean provides us with an appealing conceptual framework for relating the various possible unobserved mechanisms that may explain per capita income disparities in a way that maps naturally and clearly into a well defined notion of conditional convergence that is appropriate for the nonstationary panel setting. In particular, the fact that the regression picks out a cointegrating relationship that renders the residuals stationary implies a close connection to the idea of conditional forecast convergence conventionally defined in a time series sense. Specifically, because the residuals of the cointegrating relationship are stationary mean zero processes, this implies that any differences among the residuals are temporary.

Formally, we can say that for the residuals of the cointegrated panel specification given in (9), it must be the case that $\lim_{K \rightarrow \infty} E[\mu_{i,t+K} - \mu_{j,t+K}] = 0$ for any i, j countries. This corresponds closely to Bernard and Durlauf's [1996] concept of time-series forecast convergence, except that we are applying the idea here to the residuals of a steady state relationship rather than raw per capita income data. In this way, the definition becomes appropriate for conditional convergence in a nonstationary panel in that the cointegrating relationship picks out those features upon which it is necessary to condition in order for per capita outputs to be conditionally convergent in the sense that any remaining differences are only transitory. By construction, this implies that the various terms of the regression have directly or indirectly accounted for any mechanisms that might possibly lead to permanent differences in per capita income across countries, including very broadly even those country specific features that may impact the level or trend terms.

In this context, an important distinction in the cointegrated panel specification relative to conventional cross sectional based specifications is the treatment and interpretation of unobserved mechanisms that potentially explain per capita income dispersion across countries. As we have seen, in the cross sectional framework it is imperative to find direct quantifiable variables that can proxy for unobserved mechanisms. In the nonstationary panel framework, the combination of the extra dimension and the long run properties of the cointegrating relationship provide us with an alternative approach. In this framework statistical proxies such as the fixed effects and heterogeneous trend components can serve to capture a broad class of these unobserved mechanisms as we will see. The key issue from our perspective

relies on the fact that by capturing the role of such mechanisms in these deterministic terms, we are able to study the distribution of the structural parameters of the production function in a way that is robust to the presence of these features. As we will argue below, the approach will also allow us to relax a number of other key assumptions that have typically been employed in the cross sectional and conventional panel data approaches to these issues.

Nevertheless, it is worth noting that the specific structural interpretation that is given to the deterministic fixed effects and deterministic trends and even the slope coefficients will in fact vary depending on the nature of the mechanisms that account for persistent per capita income disparities. For example, this will depend on how the unmeasured intangible capital stocks are accumulated. In particular, we can think of two distinct categories for such unmeasured capital inputs. Broadly speaking, we can think of one such class of unmeasured intangible capital input as being similar to the way human capital is typically modeled in that the accumulation of human capital is accomplished at the expense of some fraction of income. We can think of another possible class of unmeasured intangible capital as differing from this to the extent that the accumulation of such capital stocks tends to be accomplished more generally by means other than at the expense of some fraction of measured income. This is not to say that the accumulation of this intangible capital stock is exogenous or independent of income, or that it occurs without effort or sacrifice. Rather, the key distinguishing feature of this second class of intangible capital is that it does not specifically require a fraction of income to be set aside. This may be typical of many, though not necessarily all, forms of social capital. For example, we might think of the development of trust or property rights, or common social values to be examples of this type of intangible capital stock, which evolve gradually over long periods of time and enhance the productivity of other measured inputs, but do not necessarily require a fraction of measured income to be set aside in order to be accumulated. Most likely, the accumulation of these forms of social capital require effort in the form of investment, but perhaps these efforts are at the expense of personal resources that are not a part of aggregate measured income.

The reason it is important to bear in mind the distinction between these two different types of intangible capital assets is because they affect the structural interpretation of the parameters of the panel specification that is given in terms of measured aggregate per capita income and investment shares of tangible physical capital. Consider first the type of intangible capital input that is akin to the social capital that we have described as being accumulated without the need to set aside a fraction of measured income. In this case, the most suitable way to describe the role of this omitted factor input is likely to be in terms of the steady state relationship given by equation (5). In comparing equation (5) with the

panel specification (9), we see that this form relates the omitted intangible capital input in terms of its steady state stock value, measured in per log capita efficiency units, $\ln x_{it}^*$. In this case, we might think of the fixed effects, c_i , as picking up the effect of this unmeasured capital stock under certain scenarios. For example, if the level of such capital measured in efficiency units is relatively stable over the length of the sample for any one country, then it will only impact the level of per capita income after we have conditioned upon the measured physical capital investment shares, $\ln(I/Y)_{it}$, and any country specific trend growth rates, $g_i t$. This is likely to be the case for certain types of broad social capital which evolve only very slowly. If the long run value is relatively constant for any one country, we can think of simply dropping the t index, from the steady state specification, so that this could be represented simply as $\ln x_{it} = \ln \tilde{z}_i^*$ for some unknown country specific value $\ln \tilde{z}_i^*$. In this case, these country specific values will be absorbed into the fixed effects along with other relatively constant country specific factors, so that

$$(10) \quad c_i = \ln e_{io} + \phi_i(1 - \alpha_i)^{-1} \ln \tilde{z}_i^* - \alpha_i(1 - \alpha_i)^{-1} \ln(n_i + g_i + \delta_i)$$

In this sense, the unmeasured capital stock behaves no differently than the population growth rate, n_i , which, even though it may be time varying, is relatively stable and only impacts the country specific intercept of the log linear steady state relationship. In fact, more generally the conditions under which this category of unmeasured capital stock will be accounted for in the panel regression is even broader, and does not require that it be constant for the duration of the sample. Rather, we simply require for this category of unmeasured capital that the value is stationary around its trend values. Provided that this capital stock evolves relatively smoothly, even if it does change over the sample, it will be absorbed into a combination of the fixed effects and country specific trend terms $g_i t$. If all unmeasured intangible capital were of this form, then the panel specification could be interpreted in terms of equation (5), in which case the slope coefficient for the measured share of physical capital investment would be a function solely of the production function share parameter for physical capital, such that $\beta_i = \alpha_i(1 - \alpha_i)^{-1}$.

More generally, however, we can expect that other forms of intangible capital exist which do not necessarily follow this behavior, and which require that a fraction of measured income be dedicated in order to accumulate the stock. Most notably human capital has conventionally been modeled in this way, and presumably there may be some forms of unmeasured social capital that also better fit this description. In this case, the accumulation equation (4) applies, and the

long run values for the stocks, $\ln x_{it}^*$, of this capital type need not necessarily be stationary around trend even when measured in log per capita efficiency units. The reason for this stems from the fact that the accumulation of this type of capital depends on a fraction of measured income, which in turn depends on the rate of measured physical investment shares. In this case, the more appropriate steady state specification for per capita incomes is in the form of equation (7), which specifies the relationship in terms of the rate of savings, $\ln S_i^x$, of the intangible capital, rather than the stock value. The important point to notice is that in this case the interpretation of both the intercept and the slope coefficient changes in the panel specification. If all unmeasured intangible capital stocks were in this form, then the fixed effects, c_i , would absorb the country specific rate of savings for the intangible capital stock such that now

$$(11) \quad c_i = \ln e_{io} + \phi_i(1 - \alpha_i - \phi_i)^{-1} \ln S_i^x - (\alpha_i + \phi_i)(1 - \alpha_i - \phi_i)^{-1} \ln(n_i + g_i + \delta_i)$$

and in this case the slope coefficient for the measured share of physical capital investment would be a function of both the physical capital and intangible capital share parameters such that $\beta_i = \alpha_i(1 - \alpha_i - \phi_i)^{-1}$.

In the most general case, we might expect that some of both types of categories are likely to make up the stock of unmeasured intangible capital stocks. In this case, we can think of a more finely specified production function that includes both types of intangible capital, one of which requires a fraction of measured income in order to accumulate and another which does not. Now the fixed effects absorb the impacts that both of these types of intangible capital inputs have on the level of per capita income after conditioning on measured physical capital investment shares and any country specific trend growth rates, $g_i t$. But the slope coefficient on measured physical capital investment shares still depends only on production function share parameters of the capital stock types which are accumulated by using a fraction of measured income, such as human capital and possibly some components of social capital. The key point in this discussion is that the cointegrated panel specification is sufficiently general to handle a number of different possibilities regarding the nature of the unmeasured intangible capital stock, but we must take care to interpret the meaning of the slope coefficients accordingly. Similarly, the inclusion of the heterogeneous deterministic trend terms permits the panel specification to be sufficiently general to accommodate mechanisms that might explain the cross country dispersion of per capita incomes in terms of differing rates of productivity. For example, if country specific barriers to the adoption of global technologies impacts the level of productivity, then this will be absorbed into the fixed effects, c_i . More

generally, if such barriers impact the rate of technological adaptation, then this will be absorbed into the country specific trend rates $g_i t$.

What is particularly nice about this framework is that we can study the distribution across countries of the slope coefficients that reflect parameters of the production function in a manner that accounts for the possibility of a variety of mechanisms that impact the levels and trends of per capita income differently among countries. In this way, we can control for mechanisms that explain per capita income disparities across countries by operating through level and trend effects while investigating the possible role of mechanisms that operate on the form of the aggregate production function share parameters, as for example the literature on threshold externalities and multiple regimes. Provided that the aggregate production function parameterizations are similar across countries, the slope coefficients that enable conditional convergence will still be constrained to be common across members of the panel. By contrast, if the production functions differ significantly, then even after controlling for mechanisms such as unobserved intangible capital and barriers to production, we will still require heterogeneous slopes to obtain conditional convergence.

Furthermore, the nonstationary panel framework allows us to relax many of the other strong assumptions that have typically been required in cross sectional based approaches and even in conventional panel based approaches that do not consider the long run nonstationarity properties of the panel. For example, in contrast to these other approaches, we do not need to assume that countries are close to their steady state, nor do we need to make any specialized assumptions regarding the dynamics around the steady state. Rather, the regression picks up the long run relationships between the variables in a manner that is robust to the presence of short run dynamics. The dynamics are relatively unrestricted and are permitted to differ from country to country. This is an important distinction relative to conventional dynamic panel approaches which typically require strong assumptions regarding the dynamics. As Durlauf and Quah [1999] point out, conventional dynamic panel approaches tend to inadvertently estimate higher frequency short run relationships among the variables, while relegating the long run relationships to the fixed effects. This does not occur in our nonstationary panel data setting, which explicitly extracts the long run relationship among the variables in the form of the cointegrating vectors which include the slope coefficients. By contrast, in our cointegrated panel specification, transitional dynamics have only a second order effect on the estimated long run relationships and can be treated effectively as nuisance parameters in the estimation and testing procedures.

In addition, this approach also allows us to completely relax the exogeneity assumptions that have been

required in earlier approaches, without the need for external instruments. Again, this stems from the superconsistency properties of the panel cointegration regression, which picks up the long run steady state relationships even in the presence of endogeneity. Instead, as we will see in section IV, in this framework the endogeneities that are present exert only a second order effect on the estimated long run relationships, and can easily be adjusted for by using the differences in measured investment shares as a natural internal instrument. This means that in contrast to other approaches, we do not require that investment shares are exogenous. Neither do we require a host of other exogeneity assumptions that are typically employed for identification in the cross section approach. For example, there is no need to assume that any country specific initial conditions such as $\ln e_{i0}$ are exogenous with respect to savings rates. Nor do we require that the unmeasured intangible capital stocks, nor savings rates, nor depreciation rates nor population growth rates are exogenous. All of these factors are permitted to be determined endogenously in our framework.

Finally, as we have already discussed, the approach that we employ also provides a conceptual framework in which to evaluate mechanisms that can explain the cross country dispersion of per capita incomes in a way that maps naturally and clearly into a well defined notion of conditional convergence that is appropriate for the nonstationary panel setting. It tells us that by construction the cointegrating relationship picks out those features upon which it is necessary to condition in order for per capita outputs to be conditionally convergent in the sense that any remaining differences are only transitory, so that we have directly or indirectly accounted for any mechanisms that lead to permanent differences in per capita income across countries. The fact that these include fixed effects and country specific trends implies that we are allowing for possible explanations such as unmeasured intangible capital or even barriers to technology adaption that might impact the rate of growth. In this framework we can ask whether, even after having accounted for these channels that affect the country specific levels and trend growth rates of per capita income, it is possible to obtain conditional convergence and explain per capita income disparities in terms of the Solow growth model without appealing to mechanisms that generate multiple regimes characterized by different aggregate production functions across countries.

In other words, this setup allows for a broad set of mechanisms that may explain per capita income disparities across countries. Our question is, even after allowing for all these different channels, is it possible to explain the observed disparities? The answer appears to be that although each of these channels plays a role, only by allowing for the possibility of mechanisms that generate different capital share parameterizations for aggregate production functions among different countries can we fully explain the persistent degree of per capita income disparity that exists between

countries. In the next section, we explain in detail the estimation and testing strategy in the context of the cointegrated panel specification (9) which leads us to this conclusion. At the end of the next section we also investigate the specific distribution patterns for these aggregate production parameterizations that are consistent with conditional convergence.

IV. Estimating and Interpreting the Cross Country Parameter Distribution

In this section we describe the details of the estimation and testing procedures that we employ for the nonstationary panel specification, which allows us to investigate the distribution of parameters across countries. We begin by describing the standard estimation procedure which allows us to compare the distribution of point estimates relative to the panel group mean parameters. We then describe tests within this framework which allow us to ask whether it is possible to explain cross country income dispersions in the sense of conditional forecast convergence described in the previous section once we allow for mechanisms that affect the country specific levels and trends for per capita income, or whether we need to appeal to mechanisms that also explain the apparent heterogeneity of aggregate production share parameters across countries. Finally, we conclude this section with an analysis of the specific distributional patterns of the implied parameter shares as they relate to possible explanations.

The basic framework which we use is based on the group mean fully modified procedures for cointegrated panels studied in detail in Pedroni [2001, 2000, 1996]. To keep notation as simple as possible, we describe here the basic intuition for our approach in terms of variables that have already been standardized to account for any deterministic components by demeaning and detrending where appropriate. Thus, the variables that we describe here have already been detrended and demeaned in accordance with the panel specification in (9), so that all country specific intercepts, c_i , and country specific trends, $g_i t$, have been accommodated.

Furthermore, in many cases we report results for the case in which we have also demeaned variables over the cross sectional dimension as well, so that we allow for common time effects. Common time effects, also known as time dummies, are often included in panel regressions to accommodate certain forms of cross sectional dependency that may be present in the data. More importantly, we report results in which the data has been demeaned relative to the panel mean for each period in this way to accommodate the presence of common time effects in order to allow for the possibility that there is a simple common stochastic trend which is shared across the individual members of the panel. Consequently, in this context conditional convergence becomes relative to any common stochastic trend that may

concurrently drive all countries of the sample over time. An example of this phenomenon might occur if there is permanent component of productivity disturbances that are globally shared among countries of the sample.

Once we have accounted for all country specific intercepts, trends and common time effects corresponding to the panel specification in (9), then we can estimate the slope coefficient of the cointegrating relationship using a fully modified transformation of the least squares regression. Fully modified OLS transformations for the estimation of cointegrating vectors were first developed by Phillips and Hanson [1990]. The basic idea behind the transformation is to use the fact that under cointegration a weighting of the differences of the regressors can serve as instruments for the second order bias effect that the endogeneity of the regressors has on the estimated cointegrating vector. The transformation also corrects for the effect that serial correlation can have in introducing nuisance parameters into the raw OLS distribution, and thus produces tests that are asymptotically invariant to the form of serial correlation present in the data. Pedroni [1996, 2000, 2001] extends this basic approach to a nonstationary panel setting using group mean estimation techniques, and shows how the technique can be used to accommodate the heterogeneity that is typically present, both in the transitional serial correlation dynamics as well as in the long run cointegrating relationships. Other studies of the use of FMOLS transformations in cointegrated panels based on within-dimension estimators include Phillips and Moon [1999] and Kao and Chiang [2000] and Pedroni [1996]. However, Pedroni [2000, 2001] emphasizes the important advantages of using between-dimensioned group mean based estimators for panels such as the present one in which the cointegrating vectors may be heterogeneous, and one is interested in comparing properties of the distribution of individual parameters to group mean values. Other studies such as Pesaran and Smith [1995] have also advocated the importance in general of using group mean estimators in the presence of parameter heterogeneity among panel members. Accordingly, it is this group mean FMOLS approach based on Pedroni [2000, 2001] that we will employ here for the group mean estimates.

The basic idea behind the group mean FMOLS approach is straightforward and very simple to implement since it can be interpreted as the sample average of the individual member FMOLS estimators. Thus, to implement the group mean FMOLS estimator, we begin by computing the FMOLS estimate for each of the individual countries of the panel and computing the mean value. For the purposes of the discussion in this section, we will refer to the dependent variable as y_{it} and the regressor as x_{it} , so that corresponding with the panel specification in (9), y_{it} is log per capita income and x_{it} is the log investment share. As noted above, for simplicity of notation in this discussion, we will take the

variables to be demeaned and detrended in accordance with the country specific deterministic terms in the panel specification and common time effects. In terms of this notation, the group mean FMOLS slope coefficient from Pedroni [2000, 2001] can then be computed as

$$(12) \quad \hat{\beta}^{GFM} = N^{-1} \sum_{i=1}^N \left(\sum_{t=1}^T x_{it}^2 \right)^{-1} \left(\sum_{t=1}^T y_{it} x_{it}^* - T \hat{\gamma}_i \right)$$

$$\text{where } x_{it}^* = x_{it} - \frac{\hat{\Omega}_{21i}}{\hat{\Omega}_{22i}} \Delta x_{it}, \quad \hat{\gamma}_i = \hat{\Gamma}_{21i} + \hat{\Omega}_{21i}^o - \frac{\hat{\Omega}_{21i}}{\hat{\Omega}_{22i}} (\hat{\Gamma}_{22i} + \hat{\Omega}_{22i}^o) .$$

The regressor x_{it} has been transformed by subtracting off the term $\frac{\hat{\Omega}_{21i}}{\hat{\Omega}_{22i}} \Delta x_{it}$ which behaves as an instrument for the endogeneity of the regressor. The term $\hat{\gamma}_i$ acts to correct for the effect of serial correlation due to transitional dynamics. It is composed of elements from the long run covariance matrix between the estimated residuals $\hat{\mu}_{it}$ from the panel regression (9), and Δx_{it} . Specifically, the long run covariance matrix between these series is denoted as Ω_i which can be further decomposed into a standard variance, Ω_i^o , and the forward and backward sum of autocovariances Γ_i, Γ_i' such that $\Omega_i = \Omega_i^o + \Gamma_i + \Gamma_i'$, and these can be estimated using standard kernel methods such as the Newey-West estimator.

These transformations serve to modify the OLS regression in a way that eliminates the second order bias of the estimated slope coefficient, and also eliminates any nuisance parameters due to the serial correlation present in the transitional dynamics. Thus, in contrast to cross section approaches and conventional panel approaches, the estimation procedure that we use here does not require us to assume that the determinants of per capita income are exogenous, nor do we need to assume that the economies are close to the steady state, since any transitional dynamics away from the steady state will be absorbed into the residuals and accounted for in the FM transformation. Furthermore, our group mean approach is sufficiently flexible as to allow for these transitional dynamics to differ from country to country.

The notation in (12) is fairly standard for FM transformations, and in fact we can see from expression (12) that we can think of the group mean FMOLS estimator for the panel as composed of a weighted sum of the corresponding FMOLS estimators for each of the individual countries of the panel, since the expression following the summation over i is identical to the conventional time series FMOLS estimator. Thus, we can think of the group mean panel FMOLS estimator as composed of $\hat{\beta}^{GFM} = N^{-1} \sum_{i=1}^N \hat{\beta}_i^{FM}$ where $\hat{\beta}_i^{FM}$ is the conventional FMOLS time series estimate for

the i^{th} country of the panel. This is particularly convenient for the questions in which we are interested, since it allows us to easily examine the distribution of the individual country estimates relative to the group mean estimates for the sample as a whole. Pedroni [2000] also studies the small sample properties of the group mean FMOLS estimators and shows that the bias of the estimators and the size distortion of associated tests are very small even in relatively short panels. Furthermore, the power to distinguish among hypotheses quickly rises to 100% as the cross sectional dimension increases, even for relatively short panels. Thus, even though the FMOLS estimates and tests for any one country based on only 43 years of data may not be very accurate, as the number of countries increases, the tests become increasingly reliable. In short, even though we need to use caution when interpreting the estimators and tests for the data of any one individual country, estimates and tests regarding properties of the distribution become viable as the cross section dimension increases. Intuitively, even though the long run signals contained in 43 years of data may be relatively weak for any one country, as the signal pattern is repeated over a number of countries, the signal is amplified sufficiently to be observed and tested.

In table II we report the individual country FMOLS estimates for our sample and in table III we report the panel group mean estimates for various subgroupings of the sample.⁹ Our intuition is borne out by these results. The FMOLS group mean estimator for the panel as a whole, as well as various subgroupings provide credible estimates for the slope parameters as well as the average mean and average trend values. For example, for our sample of 29 countries, we find the estimated group mean trend growth rate to be 2.4%, and the group mean slope estimate β^{GM} to be 0.37, which implies a capital share parameter value $\alpha = 0.27$ with a standard error of 0.01. The group mean point estimates are superconsistent, and for the cointegrated panel converge at the rate $T\sqrt{N}$, which permits us to obtain much more accurate estimates than would be possible with conventional methods. Furthermore, as predicted in our discussion of section III-B, the point estimates for the group mean values are reasonable despite the fact that our panel regression does not include direct proxies for the intangible factors, and despite the fact that the regressors are endogenous. This is due to the fact that the group mean estimation is robust to endogeneity of the regressors, and the fact that for the cointegrated panel specification the effect of the unobserved factors are captured by the deterministic fixed effects and heterogeneous trends, so that the slope coefficient becomes a superconsistent and unbiased estimate which depends only on the structural parameters of the underlying production functions. These values are in stark contrast to the results from cross section regressions, which tend to produce unreasonably large estimates for the implied

share parameters when intangible factors are not directly included in the regression. We also report analogous group mean values in table III for the case in which the individual country data has been demeaned relative to the panel time period means, as previously discussed. When we include these terms, the group mean slope estimates drop slightly, and the implied capital share parameters also decrease to 0.22 for the full sample.¹⁰

We also report group mean estimates for various subsamples, such as the subset of OECD and non-OECD countries as well as groupings that correspond to the nodes selected by Durlauf and Johnson's [1995] cross section based regression tree method. Durlauf and Johnson identify 4 different nodes, which appear to differ in their production function share parameterizations. Within our sample, 6 countries fall into their second node and 11 countries fall into each of their third and fourth nodes, while none fall into their first node.¹¹ Consequently, in table III, for comparison we also report the panel group mean estimates for these different nodes. We will primarily be interested in examining the distributions within these nodes, which we turn to in section IV-B. For comparison to the results of our full sample, we also report group mean estimates for two other samples. One is for the subsample of 23 countries which pass the more strict qualification that the null of stationarity is rejected on an individual basis for each country.¹² The results for this subset are similar to the results for our core panel of 29 countries. Finally, for comparison, we also report results for the greater initial sample of 51 countries for which the full data series are available, regardless of their time series properties. While the inclusion of these other countries is likely to increase the noise to signal ratio present in the data with regard to the structural parameters of interest for reasons which we have discussed in section I and II, the group mean FMOLS estimator nonetheless has the property of being robust to the inclusion of some countries whose series may be better represented as individually stationary. In this case, the FMOLS transformation results in the over-differencing of those regressors which are stationary. This has the effect of introducing noise, but this effect alone does not impact the limiting behavior of the statistic provided that the sample is sufficiently large and the proportion of cointegrated regressors is sufficiently large. On the other hand, for these countries, the individual estimates will not be superconsistent, and the usual endogeneity problems will still be present. Nevertheless, it is interesting to note that the group mean FMOLS results for the broader sample of 51 countries are remarkably similar to the results for selected sample of 29 countries. This may point to the reassuring possibility that our sample selection procedure in section II has been very conservative, and that the long run cointegration properties that we have described are actually relatively pervasive among the greater sample as well.

When considering the specific interpretation of the various estimates reported in table II and III, it is important to recall from our discussion in section III that the interpretation of the slope coefficients as they relate to the structural parameters of the production function depend in turn on the precise nature of the intangible capital stocks that have been omitted. To the extent that the omitted intangible capital stock accumulates without the need to set aside a portion of measured per capita income, the slope parameter reflects only the physical capital share parameter, such that

$$\beta_i = \alpha_i(1 - \alpha_i)^{-1} \text{ as per the steady state specification in equation (5) so that we can compute } \alpha_i = \beta_i(1 + \beta_i)^{-1}.$$

It is this value which we report in the column labeled α_i in table II. On the other hand, to the extent that some component of the omitted intangible capital stock only accumulates by setting aside some portion of measured per capita income, then the relationship is determined by both the share of this factor and the physical capital factor, such that

$$\beta_i = \alpha_i(1 - \alpha_i - \phi_i)^{-1} \text{ as per the steady state specification in equation (7). In this case, the value}$$

$$\alpha_i = \beta_i(1 + \beta_i)^{-1} \text{ represents an upper bound on the share of the physical capital factor, such that the actual value for}$$

the share of physical capital is decreased in proportion to the value of the intangible capital share, since in this case the true value for the physical capital share will be $\tilde{\alpha}_i = (1 - \phi_i)\alpha_i = (1 - \phi_i)\beta_i(1 + \beta_i)^{-1}$. Notice, however, that the extent to which such types of intangible capital exist only affects the implied values of the tangible capital share versus the intangible capital share. It does not impact whether or not slope coefficients are common across countries. In other words, the more significant the share of this intangible capital input is, the further the share of physical capital will be from the reported upper bound. However, the values will be shifted downward uniformly for all members of the panel if the share parameters of the production function are uniform.

Nonetheless, it is still interesting to compare the implied values for the upper bound values α_i with the corresponding values for $\tilde{\alpha}_i$ that are typically implied by the presence of intangible capital stocks such as human capital that follow accumulation equations such as (3) which require the sacrifice of current output in order to accumulate the intangible capital stock. To get a rough idea of the quantitative implications for this effect on the implied share parameters for the tangible physical capital, we report such estimates in the column labeled $\tilde{\alpha}_i$ by using the cross sectional information regarding the human capital share parameter. Specifically, for ease of comparison we use the schooling data from Mankiw, Romer and Weil [1992] to estimate the cross section regression given by equation (8) to obtain an estimate for β_2 which is specific to our sample of countries. As previously mentioned, the value is similar to the one obtained by Mankiw, Romer and Weil for their broader sample of countries.¹³ We then combine this with our

panel estimate for β_i from equation (9) to compute $\tilde{\alpha}_i = \beta_i(1 + \beta_i + \beta_2)^{-1}$. The standard errors for these implied share parameters are computed by numerical simulation, and take into account the standard errors for both β_i from the panel regression and β_2 from the cross section regression. From this perspective, we can see that the values in the column labeled $\tilde{\alpha}_i$ can be interpreted as the distribution for the tangible capital share parameters conditional on a common human capital share parameter ϕ . In table III where we report group mean estimates for the various subsamples of the panel, we also estimate separate cross section values for β_2 for each subsample, so that value for ϕ is permitted to differ among the subsamples. Consequently, the columns labeled $\tilde{\alpha}$ in table III can be interpreted as implied values for the group mean physical capital share parameter conditional on a common human capital share parameter for the specific subsample.

Keep in mind however, that although the FMOLS estimates of the individual values β_i and the group mean estimates β^{GFM} from equation (9) are superconsistent, this is not true for the cross sectional estimate for β_2 based on equation (8). Therefore, we report the columns labeled $\tilde{\alpha}_i$ here simply as a separate matter of interest, and we will return to these in subsection IV-B. However, each of the tests for the parameter distribution in the next subsection IV-B are based directly on the estimates of β_i from equation (9), and thus are not in any way affected by cross section estimates for ϕ .

A. Testing the Parameter Distribution

In this subsection we examine tests for the parameter distributions among the individual countries of the panel, as well as within various subsets of the panel. Our primary question of interest in this subsection is whether or not the parameterization of the aggregate production function that is consistent with conditional convergence once we account for the effect unmeasured factors such as intangible capital and barriers to production requires us to consider parameters that differ across countries. Clearly, the individual country FMOLS point estimates reported in table II appear to show considerable heterogeneity for the slope coefficients, which in turn implies heterogeneity for the structural share parameters of the neoclassical aggregate production function. However, we should keep in mind that in contrast to the group mean estimates reported in table III, the individual country estimates reported in table II are based on a relatively small number of data points and are not as reliable. It is possible that the apparent heterogeneity stems largely from the effect of sampling variation and the fact that the span of data is relatively short for any one country. The differences in

the group mean estimates for the subsamples of the panel reported in table III give us a stronger indication that the heterogeneity is not simply due to small sample effects. But in these cases, we would like to know further whether the heterogeneity is limited to different country types, or whether it is pervasive even within these country groups. For example, is the parameter heterogeneity attributable simply to whether a country is a member of the relatively more developed OECD group or the less developed non-OECD group, or is there substantial true heterogeneity within these groups as well?

Thus, we would like to investigate more formally the probability that the true parameter distribution is heterogeneous such that the group mean value represents the average of heterogeneous individual values, or whether the group mean actually represents a common value for the individual countries, which individually are simply estimated poorly. In particular, we construct two tests. The first is a relatively conventional F-test for the restriction implied by a common parameterization within each grouping. The second is a test which allows us to reverse the form of the null hypothesis, as we will see. First, for the more conventional F-test, we can use the sum of squared errors from the panel FMOLS regression, to construct a consistent tests for the restrictions implied by the case in which the parameterizations are taken to be common across the countries of the particular sample. This is equivalent to asking whether the values for the individual parameters which are consistent with conditional convergence for the sample as defined in section III are also consistent with a common production function. Specifically, we can construct a Wald statistic that compares the sum of squared errors for the restricted case when $\beta_i = \beta$ for all i versus the case with unrestricted heterogeneous β_i values that are consistent with cointegration in the panel. In this case, the Wald statistic for the null hypothesis $H_o : \beta_i = \beta$ takes the standard form of an F-test. It is not difficult to see that the F-test should have no problem distinguishing the two hypotheses in this case. When the restriction under the null is incorrect, this implies that a value for β_i is being used which does not represent a true cointegrating vector, in which case the residuals from the regression will be nonstationary. Consequently, when the restriction is incorrect, then the F-statistic will quickly diverge and reject the null hypothesis. This is exactly the scenario that we encounter. The results of the tests applied to the panel as a whole as well as the various subsamples are reported in table IV under the columns labeled “F-tests”. In all cases, the F-tests strongly reject the null hypothesis that the values for β_i are homogeneous. Therefore, we can interpret this as indicating that, even conditional on the fixed effects and heterogeneous trends, per capita incomes still do not appear to be converging to a common steady state. Rather, it is only after we allow for heterogeneous production function

parameterizations that we achieve conditional convergence.

The results of these F-tests provide compelling evidence in favor of the idea that while intangible capital and barriers to production may be important, these alone are unlikely to be able to explain the persistent disparity in per capita incomes that are observed, and that only by looking for mechanisms that can explain the heterogeneity of aggregate production functions can we fully explain per capita income disparities. However, we wish to take this one step further. Throughout this paper we have reported the results of tests both in terms of the conventional null hypotheses, and in terms that reverse the null hypothesis with the alternative hypothesis. In keeping with this spirit, we next consider whether it might be possible to do the same in the context of tests regarding the parameter distribution. In other words, given that we have found compelling evidence to reject the hypothesis that conditional convergence is possible in terms of a single common production technology, is it possible to construct a test that reverses this hypothesis, and if so are the results consistent with our conclusions regarding conditional convergence? Furthermore, as we will see, combining the results of the F-tests with the results of such a reverse test also provide us with some interesting implications for the patterns within the subsamples.

Before we proceed to discuss the test and the results, it is worth discussing the nature of the test itself. Reversing the null hypothesis requires some explanation in this context. In particular, what we have in mind is to ask the following question. Given that we have estimates of the individual slope coefficients that reflect share parameters of the production function which vary across countries, do we reject the hypothesis that the true parameters are heterogeneous in favor of the alternative hypothesis that a common parameterization can produce conditional convergence. To explore this possibility, we consider the following arguments to motivate a relatively simple test. Imagine that we had estimated equation (9) as a cointegrated regression with the homogeneity restriction imposed on the slopes such that $\beta_i = \beta$. Upon examining the residuals, we find that we cannot reject that they contain a unit root, so that we write $\hat{\mu}_{it} \sim I(1)$. In general, this may be for one of two reasons. Either the variables are not cointegrated, or, the variables are cointegrated but with heterogeneous cointegrating vectors such that $\beta_i \neq \beta$ for all i . The question we ask then is whether, in the case where the variables are cointegrated, it is possible to use a test regarding $\hat{\mu}_{it} \sim I(1)$ to infer something about the likelihood that the unit root in $\hat{\mu}_{it}$ stems from the fact that the true cointegrating vectors are heterogeneous. Notice that when we pose the question in this form, we are effectively translating the null hypothesis $H_0 : \beta_i = \beta$ for an unknown value β into the equivalent but more manageable hypothesis that $H_0 : \rho_i = 0$ for all i

for the autoregressive coefficient of the estimated residuals $\Delta\hat{\mu}_{it} = \rho_i\hat{\mu}_{it-1} + \eta_{it}$, where η_{it} is a stationary process.

In this case we are interested to know the distribution of such a test. Since the test is in the form of a unit root test on the panel of residuals, it would be nice if the test had a panel unit root distribution. But the test involves the estimate of a value for β . Therefore, to ensure that the test has the same distribution as a unit root test, we would need to account for the estimated regressor effect of having to estimate an unknown value for β . A solution to this problem turns out to be somewhat similar in spirit to the transformation that is used in the fully modified OLS regression to eliminate the effect of the feedback from the regressors on the limiting distributions. Specifically, for the fully modified OLS regression, the dependent variable is transformed by subtracting off a function of the differences in the regressor, so that, using the symbols y_{it} and x_{it} to refer generically to the right hand side and left hand side variables of the regression, in place of y_{it} , one uses $\tilde{y}_{it} = y_{it} - \frac{\Omega_{21i}}{\Omega_{22i}}\Delta x_{it}$ in the regression, where Ω_{21i} and Ω_{22i} are obtained as the elements of the long run covariance matrix between the estimated OLS residuals from the cointegrating regression, $\hat{\mu}_{it}$, and Δx_{it} for each i . Here, by contrast, we will instead transform the dependent variable by a similar weighting of the levels of the regressor, so that in place of y_{it} , we will be using $\tilde{y}_{it} = y_{it} - \frac{\Omega_{21}}{\Omega_{22}}x_{it}$. Another difference, since we will be using the test to determine whether or not the slope coefficient can be treated as homogeneous or not, is that we will use estimates of the group mean values for the elements of the long run variance weighting terms, so that

$\Omega_{21} = N^{-1}\sum_{i=1}^N \Omega_{21i}$ and $\Omega_{22} = N^{-1}\sum_{i=1}^N \Omega_{22i}$ are taken as the averages of the corresponding elements of the long run covariance matrices for the differences of the original variables, Δy_{it} and Δx_{it} .

This transformation accomplishes two objectives.¹⁴ First, it eliminates the estimated regressor effect, and secondly, it results in a situation in which the limiting distribution for the test $H_o : \rho_i = 0$ for all i for the estimated residuals $\Delta\hat{\mu}_{it}^* = \rho_i\hat{\mu}_{it-1}^* + \eta_{it}$ is asymptotically equivalent to a raw panel unit root test, despite the fact that we have estimated β . The insight for this comes from a result in Pedroni [1997], which demonstrated the conditions under which tests based on estimated residuals from cointegrating regressions for panels could be made to be asymptotically equivalent to raw panel unit root tests in the special case when the cointegrating regressors are common across members of the panel. The transformation that we apply here is exactly the transformation that induces these conditions for the residuals, $\hat{\mu}_{it}^*$, of the transformed regression of \tilde{y}_{it} on x_{it} . Specifically, when we apply this transformation to produce \tilde{y}_{it} , it implies that the resulting regression $\tilde{y}_{it} = \hat{\beta}_{NT}^*x_{it} + \mu_{it}^*$, produces a modified OLS estimate that takes the form

$$(13) \quad \hat{\beta}_{NT}^* = \left(\sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T \left(x_{it} y_{it} - \frac{\Omega_{21}}{\Omega_{22}} x_{it}^2 \right)$$

From Pedroni [1997], we know that this transformation ensures that $\hat{\beta}_{NT}^* \rightarrow 0$ at rate $T\sqrt{N}$. As a result, panel unit root tests for $H_o: \rho_i = 0$ for all i applied to the residuals $\hat{\mu}_{it}^*$ of the transformed regression of \tilde{y}_{it} on x_{it} will have the same limiting distribution as any standard panel unit root test under the null hypothesis that $H_o: \beta_i \neq \beta$ for an unknown value β , and will diverge under the alternative hypothesis that $H_A: \beta_i = \beta$ for all i for an unknown value β .

To see how the test works, notice that under the alternative hypothesis, when $\beta_i = \beta$ for all i , we can write the estimated residual from the transformed cointegrating regression as

$$(14) \quad \hat{\mu}_{it}^* = y_{it} - \frac{\Omega_{21}}{\Omega_{22}} x_{it} - \hat{\beta}_{NT}^* x_{it} = \mu_{it} + \left(\beta - \frac{\Omega_{21}}{\Omega_{22}} \right) x_{it} - \hat{\beta}_{NT}^* x_{it} \sim I(0)$$

where again, we are assuming the deterministic components have already been accommodated. The key is to note that the estimated residuals of this transformed cointegrating regression can be written in terms of the sum of three components. The last of these goes to zero as $\hat{\beta}_{NT}^* \rightarrow 0$, and what remains is the true cointegrating residual, which is stationary, so that we write that it is $I(0)$, plus another component, $\left(\beta - \frac{\Omega_{21}}{\Omega_{22}} \right) x_{it}$ which is $I(1)$, and depends on the difference between the true cointegrating vector and the estimate of Ω_{21}/Ω_{22} . If the true cointegrating vector is homogeneous, then $\beta = \Omega_{21}/\Omega_{22}$ and this term vanishes asymptotically. Consequently, the estimated residuals will be $I(0)$, and unit root tests constructed for the null hypothesis $H_o: \rho_i = 0$ for all i will diverge when $\beta_i = \beta$ for all i , and we will reject the null hypothesis in favor of the alternative.

By contrast, under the null hypothesis for β_i , namely $H_o: \beta_i \neq \beta$ for all i , then the true cointegrating residuals are given by $\mu_{it} = y_{it} - \beta_i x_{it}$ and we write the estimated residuals as

$$(15) \quad \hat{\mu}_{it}^* = y_{it} - \frac{\Omega_{21}}{\Omega_{22}} x_{it} - \hat{\beta}_{NT}^* x_{it} = \mu_{it} + \left(\beta_i - \frac{\Omega_{21}}{\Omega_{22}} \right) x_{it} - \hat{\beta}_{NT}^* x_{it} \sim I(1)$$

Although the last term still goes to zero, and the true residuals, e_{it} , are $I(0)$, now, since $\beta_i \neq \beta$ for all i ,

$\beta_i \neq \Omega_{21}/\Omega_{22}$ for all i , and therefore the term $(\beta_i - \frac{\hat{\Omega}_{21}}{\Omega_{22}})x_{it}$ does not vanish asymptotically. Consequently, the residuals will be $I(1)$ under the null hypothesis $H_o: \beta_i \neq \beta$. Furthermore, when we construct panel unit root tests in this case for $H_o: \rho_i = 0$ for all i using the estimated residuals, $\hat{\mu}_{it}^*$, these tests will not depend on the true value of β , and in the limit will have the same distribution as a raw panel unit root test under the null hypothesis. The reason for this is because the transformation is constructed in a way that induces the estimated regressors effect to vanish asymptotically, so that the remaining terms that affect the limit distribution are identical to those of a raw panel unit root test.¹⁵

Of course in practice, for small samples the limiting distribution may be a poor approximation. Typically, for panel unit root tests, the small sample distribution is sensitive to a number of choices that can affect the size of the test. For example, in parametric ADF type tests, the distribution is sensitive to the choice of lag truncation. In semi-parametric Phillips-Perron type tests, the distribution can be sensitive to the choice of bandwidth for the kernel estimator. In the test that we have described here, we have these usual concerns. In addition, we have to worry about similar choices for the kernel estimation of Ω_{21}/Ω_{22} that is used for the transformation of the cointegrating regression. To deal with these issues, and the size distortion that they are likely to create relative to the limiting distribution, we perform a simple bootstrap exercise to improve the small sample properties of the test. Specifically, most panel unit root tests take the form $v^{-1}(z_{NT} - u\sqrt{N}) \sim N(0,1)$ where z_{NT} is the properly standardized panel unit root statistic, and u and v^2 are the mean and variance adjustment terms respectively that are required to render the distribution standard normal. In the limiting distribution, u and v^2 depend only on the moments of the underlying Brownian motion functionals, whereas in small samples, they can be sensitive to the choice of lag truncation. Im, Pesaran and Shin [2003], for example, report values for u and v^2 for their panel unit root test depending on the particular lag truncation that is chosen. Here, we use a bootstrap to obtain small sample adjustment terms for the values of u and v for the various panel unit root tests that we implement. We implement the procedure for three different panel unit root tests applied to the residuals of the transformed regression.¹⁶ The first is a semi-parametric test based on the autoregressive parameter, analogous to the Phillips-Perron ρ test, which we label $SP-\rho$ in table IV, the second is a semi-parametric test based on the t-test for the autoregressive parameter, analogous to the Phillips-Perron t-test, which we label $SP-t$, the third is a parametric ADF style test, which we label $ADF-t$. These also correspond to panel-rho, panel-pp and panel-adf statistics discussed in Pedroni [1997] and Pedroni [1999], except that here they are being treated

as tests for heterogeneity under cointegration, rather than tests for the null of no cointegration.

The results for these tests are reported in table IV for the full sample as well as for the various subsamples, both for the raw data as well as for the data demeaned with respect to the time period means, which again is comparable to the inclusion of common time dummies. In both cases, all tests are unable to reject the null hypothesis of heterogeneity for the full panel as well as for the subset of 23 countries for which the stronger criteria of individual rejection of the null of stationarity for investment shares applies. The rejection of the null hypothesis of heterogeneity in these cases implies that we cannot reject the possibility that conditional convergence requires heterogeneous production function parameters among these countries even when we allow for the effects of country specific intangible capital or barriers to production through the fixed effects and country specific trend terms of equation (9). The result is consistent with the previously discussed F-tests, which reject the null hypothesis that a common set of values for the production function parameters are consistent with conditional convergence under these conditions.

Of course, just as we argued with the F-tests, it may hardly be surprising to find for such a large and varied group of countries that the production function parameters are heterogeneous. Consequently, we also examine this hypothesis for various subsets of countries which are more similar to one another, such as the OECD and non-OECD groupings, and the various nodes determined endogenously by the Durlauf and Johnson [1995] cross section based regression tree analysis. In these cases, the combination of the F-tests for homogeneity and the bootstrapped tests for heterogeneity reveal a somewhat more interesting pattern. For example, in the case of the non-OECD group of 15 countries, all of the bootstrapped statistics for the null of heterogeneity also fail to reject, which accords with the rejection of homogeneity by the F-tests for these groups. By contrast, for the OECD group of 14 countries, when we examine the data relative to the time means, all three statistics result in a rejection of the null of heterogeneity. What does this imply considering that the F-tests also reject the null of homogeneity? The answer appears to be that for this group of 14 countries there exists a smaller subgroup for which the production functions are in fact similar, even though they are not similar across the entire group of 14. To see this, consider the logic of combining the two different null hypotheses in the case where the sample is potentially mixed in the sense that it contains two possible subgroups that differ with regard to whether or not they share common production function parameterizations. In this scenario, the F-test rejects the null hypothesis that in order to be consistent with conditional convergence within the sample it is possible to have a common parameterization for the production function implied by the group mean estimate. Thus, the F-test

implies that there is at least some significant fraction of countries within the sample for which heterogeneous parameterizations are required in order to achieve conditional convergence, even when we account for unobserved mechanisms such as intangible capital and barriers to production. Then, by way of the bootstrap test for heterogeneity, we *also* reject the null hypothesis that we require *all* countries within the sample to possess different production function parameterizations in order to achieve conditional convergence. Consequently, we conclude that there must be at least some sizeable or at least nonnegligible subgroup of the sample of 14 OECD countries for which the production function parameterizations *can* be taken to be common such that we still achieve conditional convergence. In other words, *within* the OECD, there appears to be a small but nonnegligible club of countries for which this is true, even though it is not true, even as an approximation, for the OECD group as a whole. A casual glance at the figures provided in the next subsection shows that in fact there is a small subset of the OECD for which the implied parameters are fairly similar, as we will see.

Given these results, one might wonder then how it is possible that we do not observe the same pattern of test results for the full panel of 29 countries. In other words, if there is a subgroup of the 14 OECD countries for which the parameterizations can be taken as common, then by construction there must also be a subgroup for the full panel of 29 countries for which this is true. There are two reasons to anticipate such results. One reason is that the group mean value which is being compared in the two samples are not same. In other words, for the full panel, we are in essence asking whether the parameterizations for all countries are heterogeneous and distinct from the group mean for the full panel. For the subset of OECD countries, we are instead asking whether the parameterizations for all OECD countries are heterogeneous and distinct from the group mean of the OECD. The other, more substantial reason why such a result is to be expected is because a rejection implies that a sufficient nonnegligible proportion of countries exhibit properties contrary to the null hypothesis. Thus, for the smaller sample, the absolute number of countries that must contradict the null hypothesis in order to obtain a rejection is in turn smaller. Consequently, if the number of individual countries that meet the criteria for rejection is fixed, then we should anticipate that the null hypothesis is more likely to be rejected in a subset of the full sample that includes a higher proportion of these countries than does the full sample, which is exactly the result that we find. Specifically, we reject the hypothesis and conclude that for the OECD sample of 14, there is a nonnegligible subgroup of countries for which the parameterizations are not significantly distinct from the OECD group mean. At the same time, by virtue of the F-test applied to the OECD subset, we know that the subgroup is limited to

something substantially less than the entire set of OECD countries. Consequently, we can see how using the combination of F-tests for homogeneity and bootstrapped tests for heterogeneity becomes informative beyond the use of either one of these alone, particularly when we apply the combination to different subsets of the sample.

For the Durlauf and Johnson [1995] nodes, we obtain very similar results. Nodes 2 and 3 contain predominantly less developed economies, and for these the bootstrap tests are unable to reject the null of heterogeneity when we consider the tests for the data relative to the time means, just as we found for the subset of non-OECD countries. Node 4 on the other hand, contains predominantly more developed economies, and for this node we find a combination of tests results similar to the ones obtained for the OECD subsample. For this node, we can reject by virtue of the F-test that the production function parameterizations can be treated as common across all countries, while the bootstrap tests for heterogeneity tell us that there must exist at least some nonnegligible subgroup within this node for which the parameterizations can be treated as common.

B. Interpreting the Cross Country Distribution Patterns

From a broader perspective, the results so far from tables II through IV show us that even when we account for unmeasured mechanisms such as intangible capital or barriers to production, we still require heterogeneous parameterizations of the production function, and furthermore that the result continues to hold even for groups of relatively similar countries. In this subsection we ask whether there are any systematic recognizable patterns of the distribution of the parameter shares implied by these results. In particular, we examine which types of correlates may help to explain the distributional pattern of the implied parameter shares that are consistent with conditional convergence after having allowed for unmeasured factors such as intangible capital and barriers to production.

One of the most striking patterns of the distribution of implied share parameters is depicted in figures I and II. In figure I we can see that the pattern of implied parameter shares is completely uncorrelated with initial values of per capita income taken from 1950 as we would expect from the neoclassical production function. In figure II, however, we can see that the same pattern of implied parameter shares consistent with conditional convergence is highly correlated with the subsequent average annual growth rates over the sample period. In comparing figure I with figure II we can see that a rough mapping exists between the two. Specifically, we can imagine rotating all those countries which lie in the upper half of the vertical plane in figure I rightward, and all those countries which lie in the lower half of the

vertical plane leftward. Doing so transforms the pattern in figure I into the pattern that we observe in figure II. In other words, while the size of the implied capital share parameter is uncorrelated with income at the beginning of the sample, it becomes a good predictor of the subsequent rate of growth in income. Thus, countries such as Venezuela, which began with relatively high initial levels of per capita income but had relatively small capital shares tended to rotate leftward in moving from figure I to a relatively low growth rate in figure II. Countries such as Japan, with relatively low initial values of per capita income but high capital share values tended to rotate rightward toward high growth rates. By contrast, countries such as Egypt or Morocco, which began with low incomes and small capital share parameters tend not to rotate in moving from figure I to II, and have stayed poor, with low subsequent growth rates. Closer visual inspection of the two figures shows that this phenomenon is not limited merely to the distinction between OECD and non-OECD countries, but holds within these subsets as well. A more formal regression analysis of these correlation patterns is reported in table V. In all cases we have used the values for $\tilde{\alpha}_i$ that are reported in table II to represent the implied capital share. Since $\tilde{\alpha}_i$ is a monotonic transformation of α_i it does not matter much in terms of the basic pattern which of these values we use. We have chosen to use $\tilde{\alpha}_i$ simply to represent a somewhat more interesting set of scale values on the vertical axis of the figures. In both the diagrams and the regressions we have omitted Peru, since the standard error for Peru is uniquely large and the point estimate is slightly negative. However, the correlation and regression fit is actually not much affected by the inclusion or exclusion of Peru. In figure II, Japan appears as an extreme case along the regression line. When we experimented by omitting Japan, the correlation and regression fit further increased slightly, as should be visibly apparent based on figure II.

Figures I and II can be interpreted as a visual depiction of the fact that allowing for heterogeneous parameters opens up an additional channel through which we can explain the disparity of per capita incomes, and once we allow for this channel it becomes possible to achieve convergence in the conditional forecast sense that we have defined in section III. This begs the question as to what economic features might account for the particular distribution pattern for the implied parameter shares in the first place. Consequently, we experimented with a number of common measures that have been used to proxy variables that might be associated with such a pattern. For example, a common theme in much of the literature on multiple production regimes is that human capital may serve to produce threshold effects which result in production regimes in which the share parameter for capital should be higher for countries with higher levels of per capita human capital. Thus we experimented by comparing the implied parameter shares with the values for per capita

human capital for these countries, as reported in Hall and Jones [1999]. However, as the analysis in table V reveals, there does not appear to be any systematic correlation between the level of per capita human capital and the implied capital share parameters for the core sample. The same is true for the rate of human capital investment as proxied by schooling rates taken from Mankiw, Romer and Weil [1992]. There does not appear to be a systematic relationship between schooling rates and the implied share parameters.

By contrast, when we experimented with various measures for social capital infrastructure, we did find a systematic positive relationship between these measures and the implied capital share parameters. For example, table V reports a statistically significant positive correlation between the implied capital share values and the government anti-diversion index used in Hall and Jones [1999] as well as the economic openness measure used in Hall and Jones. Hall and Jones also construct what they refer to as a social infrastructure measure by using a combination of the openness and government anti-diversion index. Figure III depicts the positive correlation that we find between this social infrastructure measure and the implied capital shares. In this case we notice the sample splits broadly into two major clusters which are responsible for the overall relationship. It is also interesting to note that unlike human capital levels or schooling rates, literacy rates do appear to be positively correlated with the implied parameter shares that are consistent with conditional convergence. It is possible that basic literacy behaves much like a proxy for social capital in the sense that it proxies the overall level of social cohesion rather than proxying the level of human capital embodied in the population. The fact that these various measures for social infrastructure capital help to explain the distribution pattern for the implied capital shares may point to an important role for social capital which goes beyond simply acting as an unmeasured production factor that enhances the productivity of other inputs or inhibits barriers to the adoption of new technologies that impact the estimated country specific trend growth rates. The positive relationship between social infrastructure measures and capital share parameters may point to a possible role that social capital might play in generating multiple production regimes, much in the way that human capital has been modeled in the literature. In other words, these patterns appear to hint at the possibility that the key may be in understanding the role that social capital, perhaps by facilitating the coordination of production externalities, may play in helping to explain the fact that countries experience different log linear production regimes corresponding to different aggregate share parameterizations for the neoclassical model.

Finally, we also experimented to see whether these correlation patterns were limited to our sample, or whether

they also extended to the broader sample of 51 countries. The results for these are also reported in table V. Indeed, we see that the basic pattern appears to extend to the broader sample as well, in that initial income levels are uncorrelated with the parameter shares, while growth rates are positively correlated. Likewise, schooling measures are also uncorrelated, and human capital measures are only very weakly correlated with only a marginal level of significance, while social infrastructure continues to be significant and positively correlated. For reasons that we have discussed previously, the individual estimates are likely to contain a higher proportion of noise in the greater sample, and it is not surprising to find that in those cases where a positive correlation holds for the broader sample, the R^2 measures are considerably weaker than for our core sample of countries. In fact, one interesting detail to note is that for the broader sample of 51 countries we can no longer detect a positive correlation for the government anti-diversion index and the economic openness measure. In this sense, focusing on the narrower sample may also help us to uncover particular patterns that would not otherwise be apparent for the specific distribution patterns.

V. Concluding Remarks

In this study we have exploited recent advances in the econometrics of nonstationary dynamic panels to study the steady state relationship implied by the neoclassical Solow growth model. In contrast to conventional panel data techniques, nonstationary panel techniques allow us to focus explicitly on the low frequency relationships. At the same time, the techniques permit us to treat the transitional dynamics in a very flexible manner and also allow us to relax many of the strong assumptions that have typically been employed in the empirical growth literature. For example, the techniques no longer require us to assume that economies are observed close to their steady states at all points in time. Furthermore, the techniques have allowed us to treat all variables as fully endogenous, and does not require us to make any of the usual exogeneity assumptions.

Most importantly, the combination of an emphasis on low frequency relationships and the dimensionality of the panel has permitted us to examine patterns in the data that would be difficult if not impossible for conventional cross section based methods. In particular, the approach has allowed us to study the full distribution over countries of parameter shares associated with the aggregate production function, and has allowed us to do this in a manner that is consistent with allowing for the possibility of unobserved country specific factors that impact the levels and growth rates of per capita income in the steady state relationship. In this way, we have been able to study the distribution across

countries of the implied share parameters in a manner that allows indirectly for the anticipated role that intangible capital, such as human capital and social infrastructure is likely to have on the level of productivity. Likewise, the approach allows for the role that these factors can have in deterring barriers to production that may impact the rate at which technology is adopted across countries. Finally, by relating the empirical specification to a notion of long run conditional forecast convergence, we can also ask under what conditions the steady state specification is able to explain persistent per capita income disparities across countries.

In this framework our primary results have shown that the implied share parameters of the neoclassical aggregate production function appear to be heterogeneous across countries. More to the point, we are able to show formally in this context that only by allowing for share parameters to vary across countries can we explain persistent cross country disparities in per capita income, and most importantly, that this continues to be the case even when we allow for the effect of factors such as country specific intangible social capital or country specific institutional barriers to the adoption of new technologies. Furthermore, the phenomenon appears to hold across various levels of development, and even among relatively similar countries. Upon examining the specific distributional pattern, we find that the estimated values for the implied parameter shares appear to be positively correlated with various measures of social infrastructure, including an index of government anti-diversionary behavior. While the correlation may be consistent with a number of explanations, it does point to the possibility that in fact social infrastructure capital may play a broader role in terms of the way that it enhances productivity. For example, the distribution patterns may point to an additional role for social infrastructure in facilitating production regimes, much as the earlier literature on multiple regimes had postulated for human capital. This may be sensible in terms of the way that we think about social infrastructure as potentially helping to facilitate the coordination of productive behavior that can lead to positive production externalities at the aggregate level. Likewise, social infrastructure capital may also help to channel behavior away from diversionary activities that can lead to negative production externalities at the aggregate level.

The approach used in this study has provided us with opportunities to study empirical relationships that would otherwise be difficult to observe with the limited data that is available. At the same time, it is important to recognize some of the limitations that are present in the approach that we have taken here. For example, although we have occasionally studied properties of a broader sample of countries, for the most part we have limited our analysis to a core set of countries for which we are relatively confident about the presence of permanent variations within sample for the

investment share. While we have checked in a number of instances to confirm that our sample is representative in other regards, whether or not all of our results necessarily extend to a broader sample of countries may be regarded as an open question. On the other hand, by construction, we can say that our primary result is a statement that can be applied to the broadest set of all countries. Specifically, since we have shown that the common log linear relationship does not hold for a sizeable subset of countries even after we account for unmeasured effects that impact the level and trend of per capita income, then by construction it must be the case that a common log linear relationship does not hold for the set of all countries. More importantly, the same logic applies to the set of all countries that fall into the smaller subgroup classifications that we have examined in our sample so that for example we can say with confidence that the common log linear relationship also does not hold even for the OECD countries alone.

Another important point to recognize is that our approach has simply identified the fact that for the neoclassical production function, we must consider mechanisms that explain the heterogeneity of share parameters at the aggregate level. We cannot, however, say for certain which particular combination of the share parameters are necessarily responsible for the heterogeneity. We expect that it is most likely due to a combination of physical and human capital share parameters that vary across countries. Since our best available data measures variations in tangible physical capital investment, we have focused on reporting the implied cross country variation in physical capital shares $\tilde{\alpha}_i$ required for conditional forecast convergence. If the panel time series data on human capital were as well measured, we could just as well imagine reporting the implied cross country variations in human capital shares implied by conditional forecast convergence. The key point that remains, however, is that one way or another the parameterization of the aggregate production function must be allowed to vary across countries if we are to explain the observed relationship between permanent variations to investment shares and per capita income in a manner otherwise consistent with the basic neoclassical model. Fundamentally, the result derives from the heterogeneity of the slope coefficient of the panel specification, regardless of the specific interpretation that is given to the implied share parameters and provides further support for considering production functions that deviate from the common linear production technology. This can be viewed as further support for the fruitfulness of considering alternative such as the CES production function as in Masanjala and Papageorgiou [2004].

Finally, we should emphasize that our analysis has been directed toward studying a fixed distribution of implied parameter shares across countries based on the observed relationship between variations to investment shares and per

capita income in a manner that accounts for unmeasured factors that impact the level and trend of per capita income. Our analysis allows these parameters to vary across countries facing different conditions or experiencing different levels of development, and takes these parameters to be fixed over the duration of the sample. In this sense, we have implicitly treated the parameters as if they are determined by features which change only slowly over time, and tend to be relatively constant within sample. Needless to say, this does not preclude the possibility that over sufficiently long periods the parameters of the production function may also be changing over time. In this manner, one can think about the possibility that countries might nevertheless converge to a common log linear steady state relationship as the parameters of the production function themselves change, and possibly converge to a common set of values.

References

- Azariadis, C. and A. Drazen (1990) "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 106, 501-26.
- Banerjee, A. (1999) "Panel Data Unit Roots and Cointegration: An Overview," *Oxford Bulletin of Economics and Statistics*, 61, 607-630.
- Baltagi, B. and C. Kao (2000) "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey," *Advances in Econometrics: "Nonstationary Panels, Panel Cointegration and Dynamic Panels"*, 15, 7-52.
- Barro, R. (1991) "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economic Growth*, 106, 407-43.
- Bernard, A. and S. Durlauf (1996) "Interpreting Tests of the Convergence Hypothesis," *Journal of Econometrics*, 71, 161-73.
- Bernard, A. and C. Jones (1996) "Comparing Apples and Oranges: Productivity Convergence and Measurement Across Industries and Countries," *American Economic Review*, 86, 1216-38.
- Canning, D. and P. Pedroni (1999) "Infrastructure and Long Run Economic Growth," Center for Analytical Economics Working Paper No. 99-09, Cornell University.
- Caselli, F., G. Esquivel and F. Lefort (1996) "Reopening the Convergence Debate; A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1, 363-89.
- Cheung, Y. and K. Lai (2000) "International Evidence on Output Persistence from Postwar Data," *Economics Letters*.
- Durlauf, S. (1993) "Nonergodic Economic Growth," *Review of Economic Studies*, 60, 349-66.
- Durlauf, S. and P. Johnson (1995) "Multiple regimes and cross country growth behavior," *Journal of Applied Econometrics*, 10, p. 365-84.
- Durlauf, S. and D. Quah (1999) "The New Empirics of Growth," in *Handbook of Macroeconomics*, J. Taylor and M. Woodford, eds., Amsterdam, North Holland.

- Easterly, W. and R. Levine (2001) "It's Not Factor Accumulation; Stylized Facts and Growth Models," forthcoming, *World Bank Economic Review*.
- Evans, P. (1998) "Using Panel Data to Evaluate Growth Theories," *International Economic Review*, 39, 295-306.
- Hall, R. and C. Jones (1999) "Why do Some Countries Produce so Much More Output Per Worker than Others," *Quarterly Journal of Economics*, 114, 83-116.
- Im, K., H. Pesaran, and Y. Shin (2003) "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115, 53-74.
- Islam, N. (1995) "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 1127-70.
- Islam, N. (2000) "Small Sample Performance of Dynamic Panel Data Estimators in Estimating the Growth-Convergence Equation: A Monte Carlo Study," *Advances in Econometrics: "Nonstationary Panels, Panel Cointegration and Dynamic Panels"*, 15, 317.
- Jones, C. (1995) "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, 495-525.
- Knight, M., N. Loayza, and D. Villaneuva (1993) "Testing the Neoclassical Growth Model," *IMF Staff Papers*, 40, 512-41.
- Kwiatkowski, D., P. Phillips, P. Schmidt and Y. Shin (1992) "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root?" *Journal of Econometrics*, 54, 159-178.
- Lee, K., H. Pesaran, and R. Smith (1997) "A Multi-Country Empirical Analysis of the Solow Growth Model," *Journal of Applied Econometrics*, 12, 357-92.
- Lee, K., H. Pesaran, and R. Smith (1998) "Growth Empirics: A Panel Data Approach - A Comment," *Quarterly Journal of Economics*, 113, 319-323.
- Mankiw, G., D. Romer and D. Weil (1992) "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*.
- Masanjala, W. and C. Papageorgiou (2004) "The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity," *Journal of Applied Econometrics*, 19, 171-201.
- Murphey, K., A. Shleifer, and R. Vishny (1989) "Industrialization and the Big Push," *Journal of Political Economy*, 97, 1003-26.
- Nehru, V., E. Swanson and A. Dubey (1993) "A New Database on Human Capital Stock; Sources, Methodology and Results," World Bank Policy Research Working Paper no. 1124.
- Parente, S. and Prescott, E. (1994) "Barriers to Technology Adoption and Development," *Journal of Political Economy*, 10, 298-321.
- Parente, S. and Prescott, E. (1999) "Monopoly Rights: A Barrier to Riches," *American Economic Review*, 89, 1216-33.
- Pedroni, P. (1996) "Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity," Indiana University Working Papers in Economics, No. 96-020.

- Pedroni, P. (1997) "Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, With an Application to the PPP Hypothesis, New Results" working paper, Indiana University.
- Pedroni, P. (1998) "On the Role of Human Capital in Growth Models: Evidence from a Nonstationary Panel of Developing Countries," working paper, Indiana University.
- Pedroni, P. (1999) "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors," *Oxford Bulletin of Economics and Statistics*, 61, 653-670.
- Pedroni, P. (2000) "Fully Modified OLS for Heterogeneous Cointegrated Panels," *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, 15, 93-130.
- Pedroni, P. (2001) "Purchasing Power Parity Tests in Cointegrated Panels," *Review of Economics and Statistics*, 83, 1371-75.
- Pedroni, P. (2004) "Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, With an Application to the PPP Hypothesis, New Results," *Econometric Theory*, 20, 597-625.
- Pesaran, H. and R. Smith (1995) "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 79-113.
- Phillips, P. and H. Moon (1997) "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057-1112.
- Prescott, E. (1998) "Needed: A Theory of Total Factor Productivity," *International Economic Review*, 39, 525-51.
- Quah, D. (1993) "Empirical Cross-Section Dynamics in Economic Growth," *European Economic Review*, 37, 426-34.
- Shin, Y. (1994) "A Residual Based Test of the Null of Cointegration Against the Alternative of No Cointegration," *Econometric Theory*, 10, 91-115.
- Summers, R. and A. Heston (1991) "The Penn World Table (Mark 5); An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics*, 106, 327-68.
- Temple, J. and P. Johnson (1998) "Social Capability and Economic Growth," *Quarterly Journal of Economics*, 113, 965-91.

Endnotes

1. Nevertheless, earlier panel studies provided important advances in the empirical literature, as the two dimensional treatment of the data allowed researchers to relax many of the strong assumptions that had been made in the initial cross sectional based literature. Important early examples include the works of Caselli, Esquivel and Lefort [1996], Islam [1995], Knight, Loayza and Villaneuva [1993]. Many early works were also focused explicitly on estimating the implied speed of convergence. See Islam [2000] for a recent survey and discussion of some of the small sample difficulties associated with the use of conventional dynamic panel techniques. Another important approach to studying growth relationships in panels is the approach pioneered in the work of Quah [1993], which uses kernel methods to estimate the evolution of cross sectional distribution dynamics.

2. From a broader perspective, it is interesting to note this is akin to looking for specific potential case studies that have occurred within a sample that can aid in discerning a hypothesized relationship, except that we are implicitly drawing on many such events.

3. Specifically, let x_i be the KPSS test value for the series of country i . Let $u = \lim_{T \rightarrow \infty} E[KPSS_i]$ and $v^2 = \lim_{T \rightarrow \infty} E[KPSS_i]^2$ be the mean and variance of the limit distribution of the KPSS test under the null hypothesis, and let $\bar{x} = N^{-1} \sum_{i=1}^N x_i$ be the group mean value for x_i . Then the panel statistic $v^{-1} \sqrt{N}(\bar{x} - u) \Rightarrow N(0,1)$ is distributed normally under the null hypothesis that all series are stationary, and is a right tailed test that diverges to positive infinity under the alternative. We simulated u and v^2 using 20,000 i.i.d. draws from an i.i.d. distribution to construct $KPSS_i$ under the null.

4. Again, we could narrow the sample further by requiring that countries also individually reject the reverse null, namely the null of no cointegration, but again, given the weak power of the individual tests, we do not want to be excessive in eliminating countries from the sample. Nevertheless, we did also experiment with this even narrower subset for the procedures in section IV and did not find any substantial differences in the results.

5. Specifically, in this case we constructed the $KPSS_i$ test from the estimated residuals of a regression $y_t = c + g t + b x_t + u_t$ and simulated the values u and v^2 based on the residuals from 20,000 such regressions where x_{it} and y_{it} were drawn as independent random walks.

6. Table II reports results discussed later in section IV. However the list of countries included is identical. Table I of this section reports only a summary of the panel results regarding unit root and cointegration properties. The full set of country by country tests for unit root and cointegration properties is not included in the interest of space. However, a copy of the full set of results is available from the author upon request.

7. Evans [1998] also conducts panel tests for the unit root null for per capita incomes and is unable to reject the null for the raw series, but finds a rejection when the time means are subtracted, while we do not. We anticipate that this stems in large part from the difference in our samples, since we specifically sought out those countries for which investment shares behave as a unit root process, which in turn implies that per capita outputs can be expected to follow a unit root process, even relative to the sample mean. Another possible difference may stem from the fact that we have used the Im, Pesaran and Shin [2003] test, which allows for greater flexibility under the alternative hypothesis, and the fact that we also allow for the presence of country specific trend terms.

8. Our samples also differ somewhat further in that Mankiw, Romer and Weil [1992] use the real GDP per effective adult, RGDPEA, measure for GDP, while we have used real GDP per capita, RGDPL, since the former is generally not available for as long a span as RGDPEA for many countries. Also, we have used the more recent version 5.6 Penn World Table data set, which includes updated series through 1992.

9. Individual FM estimates relative to mean that were used to compute group mean results relative to mean in table III are not reported in table II in the interest of space, but are also available upon request.
10. Naturally, demeaning the data with respect to the time period means implies that the group mean intercept and group mean trend estimates become approximately zero, as reported in table III.
11. Specifically, the countries in our sample which fall into Durlauf and Johnson's [1995] second node are Egypt, Kenya, Morocco, Nigeria, Bolivia, and Turkey, and in the third node are South Africa, Dominican Republic, Brazil, Columbia, Paraguay, Peru, Japan, Philippines, Sri Lanka, Ireland and Spain. The remaining countries fall into their fourth node, except for Luxembourg, which is not included in their study.
12. This subsample of 23 excludes Morocco, Nigeria, Chile, Belgium, Denmark and Sweden.
13. We also experimented with using the panel series for years of education per worker from Swanson, Nehru and Dubai [1993] in order to obtain an estimate, but found that the series generally did not cointegrate with our other series, and often gave nonsensical results for the slope coefficients. Thus, for the purposes of our rough estimate of the impact of human capital on the values of $\tilde{\alpha}_i$ relative to α_i we prefer to use the more conventional cross section data for schooling, which also facilitates comparison with Mankiw, Romer and Weil [1992] and subsequent studies.
14. It is worth elaborating a bit here to develop an intuition for how the tests work. However, readers who prefer to forego this may skip the next three paragraphs without loss of continuity.
15. A technical appendix which provides a derivation of this property of the limiting distribution is available upon request.
16. A more detailed description of the bootstrap procedure is also available in the technical appendix, which is available upon request.

Table I. Panel Unit Root and Cointegration Properties

	<u>raw</u>		<u>relative to mean</u>		<u>raw</u>		<u>relative to mean</u>	
	<i>(log investment shares)</i>				<i>(log per capita income)</i>			
	<u>I(1)</u>	<u>I(0)</u>	<u>I(1)</u>	<u>I(0)</u>	<u>I(1)</u>	<u>I(0)</u>	<u>I(1)</u>	<u>I(0)</u>
unit root:	-0.80	13.8 ***	-0.87	15.9 ***	4.58	17.4 ***	-0.31	14.1 ***
	<u>pp</u>	<u>adf</u>	<u>pp</u>	<u>adf</u>	<u>shin</u>	<u>shin</u>		
cointegration:	1.31	1.58	-1.91 **	-1.3 *	-0.55		0.73	

Notes: For univariate tests, columns labeled I(1) report the IPS panel test for the unit root null, columns labeled I(0) report an analogous group mean test for the stationary null based on the KPSS test. For the cointegration relationship tests, columns labeled PP and ADF report the Pedroni [1999] group mean panel tests for the null of no cointegration based on the Phillips-Perron and ADF tests, and columns labeled shin, report an analogous group mean panel test for the null of cointegration based on the Shin [1994] time series test. Fixed effects have been included in all cases, and heterogeneous trends have been included for all tests involving log per capita income. Columns labeled “relative to mean” refer to results relative to panel means, comparable to common time dummies. The symbols *, **, *** denote 10%, 5%, 1% rejections respectively. See section II for further details.

Table II: Individual Coefficient Estimates

country	intercept	trend	slope	std err	α_i	std err	$\tilde{\alpha}_i$	std err
Egypt	6.31	0.029	0.09	0.05	0.08	0.04	0.05	0.02
Kenya	6.05	0.014	0.08	0.07	0.08	0.06	0.05	0.04
Morocco	6.52	0.028	0.04	0.06	0.04	0.06	0.02	0.04
Nigeria	4.84	0.022	0.58	0.09	0.37	0.04	0.26	0.01
S. Africa	6.13	0.017	0.50	0.06	0.33	0.03	0.23	0.00
Dominican R	6.42	0.019	0.21	0.08	0.17	0.05	0.11	0.03
Bolivia	6.31	0.019	0.22	0.04	0.18	0.03	0.12	0.01
Brazil	4.93	0.036	0.72	0.11	0.42	0.04	0.31	0.01
Chile	6.95	0.011	0.31	0.09	0.24	0.06	0.16	0.03
Columbia	7.13	0.022	0.04	0.12	0.04	0.12	0.02	0.08
Paraguay	6.47	0.008	0.30	0.11	0.23	0.06	0.15	0.04
Peru	7.68	0.010	-0.04	0.20	-0.04	0.29	-0.03	0.21
Venezuela	8.52	0.004	0.08	0.10	0.07	0.09	0.04	0.06
Japan	4.80	0.041	0.87	0.09	0.47	0.02	0.35	0.00
Philippines	5.87	0.013	0.41	0.07	0.29	0.04	0.20	0.01
Sri Lanka	6.32	0.007	0.37	0.11	0.27	0.06	0.19	0.03
Austria	6.78	0.032	0.44	0.10	0.30	0.05	0.21	0.02
Belgium	7.04	0.029	0.41	0.09	0.29	0.04	0.20	0.02
Denmark	7.39	0.026	0.36	0.03	0.26	0.01	0.18	0.00
Finland	6.74	0.035	0.41	0.07	0.29	0.04	0.20	0.01
Germany, W.	4.91	0.036	1.01	0.14	0.50	0.04	0.38	0.01
Ireland	7.25	0.030	0.19	0.05	0.16	0.03	0.10	0.01
Italy	6.04	0.039	0.58	0.16	0.37	0.07	0.26	0.04
Luxembourg	7.92	0.025	0.21	0.05	0.17	0.03	0.11	0.01
Netherlands	6.53	0.029	0.58	0.11	0.37	0.04	0.26	0.02
Norway	7.75	0.034	0.16	0.05	0.14	0.04	0.09	0.02
Spain	5.24	0.030	0.83	0.12	0.45	0.04	0.34	0.01
Sweden	7.48	0.023	0.39	0.09	0.28	0.05	0.19	0.02
Turkey	6.51	0.024	0.24	0.07	0.19	0.05	0.13	0.02

Notes: Estimates are FMOLS coefficients of equation (9) in section III. Columns labeled α_i and $\tilde{\alpha}_i$ report implied share parameters as defined in section III. Standard errors for implied estimates were computed by numerical simulation.

Table III: Panel Group Mean Coefficient Estimates

sample	Raw data							
	intercept	trend	slope	std err	α	std err	$\tilde{\alpha}$	std err
full panel (29)	6.51	0.024	0.37	0.02	0.27	0.01	0.18	0.01
subset (23)	6.46	0.024	0.37	0.02	0.27	0.01	0.19	0.01
All-ctrys (51)	6.66	0.023	0.34	0.01	0.25	0.01	0.18	0.01
non-OECD (15)	6.43	0.017	0.26	0.03	0.21	0.02	0.15	0.01
OECD (14)	6.60	0.031	0.48	0.02	0.32	0.01	0.27	0.03
DJ-node2 (6)	6.09	0.022	0.21	0.03	0.17	0.02	0.12	0.01
DJ-node3 (11)	6.20	0.021	0.40	0.03	0.29	0.02	0.31	0.03
DJ-node4 (11)	6.92	0.027	0.43	0.03	0.30	0.01	0.24	0.01
	Relative to Mean							
full panel (29)	0.01	-0.001	0.28	0.02	0.22	0.01	0.14	0.01
subset (23)	0.01	-0.001	0.27	0.02	0.21	0.01	0.14	0.01
All-ctrys (51)	-0.00	-0.000	0.27	0.01	0.22	0.01	0.15	0.01
non-OECD (15)	0.04	-0.001	0.20	0.03	0.17	0.02	0.12	0.00
OECD (14)	-0.01	-0.000	0.29	0.03	0.22	0.02	0.18	0.01
DJ-node2 (6)	0.02	-0.002	0.26	0.03	0.21	0.02	0.15	0.01
DJ-node3 (11)	0.01	-0.000	0.33	0.03	0.25	0.02	0.27	0.02
DJ-node4 (11)	-0.02	0.000	0.18	0.03	0.15	0.02	0.12	0.01

Notes: Estimates are group mean estimates for the coefficients of equation (9) in section III, based on the Pedroni [2000] group mean panel FMOLS estimator discussed in section IV. Columns labeled α and $\tilde{\alpha}$ report implied share parameters as defined in section III. Standard errors for implied estimates were computed by numerical simulation. Rows labeled “relative to mean” refer to results relative to panel means, comparable to the inclusion of common time dummies. See section IV for further details, including descriptions of the various subsamples.

Table IV. Panel Tests for Cross Country Parameter Distributions Consistent with Conditional Convergence

sample	raw data					relative to mean			
	F-test	PP-r	PP-t	ADF		F-test	PP-r	PP-t	ADF
Full Panel (29)	4439 ***	6.33	8.07	2.82		90 ***	1.36	2.00	-0.26
Subset (23)	5621 ***	9.07	13.85	-0.12		100 ***	1.82	2.51	0.69
Non-OECD (15)	1687 ***	3.45	3.31	0.42		58 ***	0.85	0.81	-0.22
OECD (14)	11240 ***	3.83	4.66	3.41		52 ***	-6.83 ***	-2.41 ***	-2.47 ***
DJ-node2 (6)	1048 ***	0.69	0.53	0.26		32 ***	0.05	-0.06	-0.79
DJ-node3 (11)	5107 ***	3.19	4.33	-2.81 ***		101 ***	-0.07	0.48	0.45
DJ-node4 (11)	6459 ***	-3.16 ***	-4.14 ***	-3.50 ***		16 ***	-2.06 **	-3.61 ***	-2.37 ***

Notes: Columns labeled F-test are for the null hypothesis of equal slopes, $\beta_i = \beta$, for equation (9) of section III. Columns labeled PP-r, PP-t and ADF report tests for null hypothesis of heterogeneous slopes, $\beta_i \neq \beta$, for equation (9) of section III, based on the bootstrap. Columns labeled “relative to mean” refer to results relative to panel means, comparable to the inclusion of common time dummies. The symbols ***, **, * denote 10%, 5%, 1% rejections respectively. See section IV for further details regarding the tests and discussion of the bootstrap procedure.

Table V. Correlation Analysis of Distribution Pattern for Share Parameters Consistent with Conditional Convergence

correlate	core sample (N=28)			broader sample (N=51)		
	slope	t-test	R^2	slope	t-test	R^2
initial log per capita output	0.02	0.74	0.02	0.03	1.21	0.03
per capita output growth	5.33	3.43	0.31	3.48	2.09	0.08
schooling (MRW)	0.01	1.36	0.07	0.00	0.79	0.01
schooling (HJ)	0.01	0.98	0.04	0.06	0.86	0.01
per capita human capital (HJ)	0.08	0.95	0.03	0.06	1.70	0.06
log total factor productivity (HJ)	0.04	1.18	0.05	-0.00	-0.10	0.00
government anti-diversion index (HJ)	0.17	1.99	0.13	-0.04	-0.66	0.01
economic openness (HJ)	0.11	1.90	0.12	0.09	1.17	0.03
social infrastructure (HJ)	0.15	2.07	0.14	0.10	1.92	0.07
literacy (DJ)	0.16	2.62	0.21	0.15	2.63	0.13

Notes: Computed pairwise by linear regression. Dependent variable in all cases is the $\tilde{\alpha}_i$ estimate from table II or equivalent for the case of the broader sample. Columns labeled t-test report significance tests. Column labeled R^2 reports regression fit. Intercept values are not reported in the interest of space. Abbreviations in parentheses for correlates indicate source of data as follows. MRW: Mankiw, Romer and Weil [1992], HJ: Hall and Jones [1999], DJ: Durlauf and Johnson [1995]. See section IV for further details.

Figure I. Absence of Relationship Between Initial Per Capita Output and Parameter Shares Consistent with Conditional Convergence

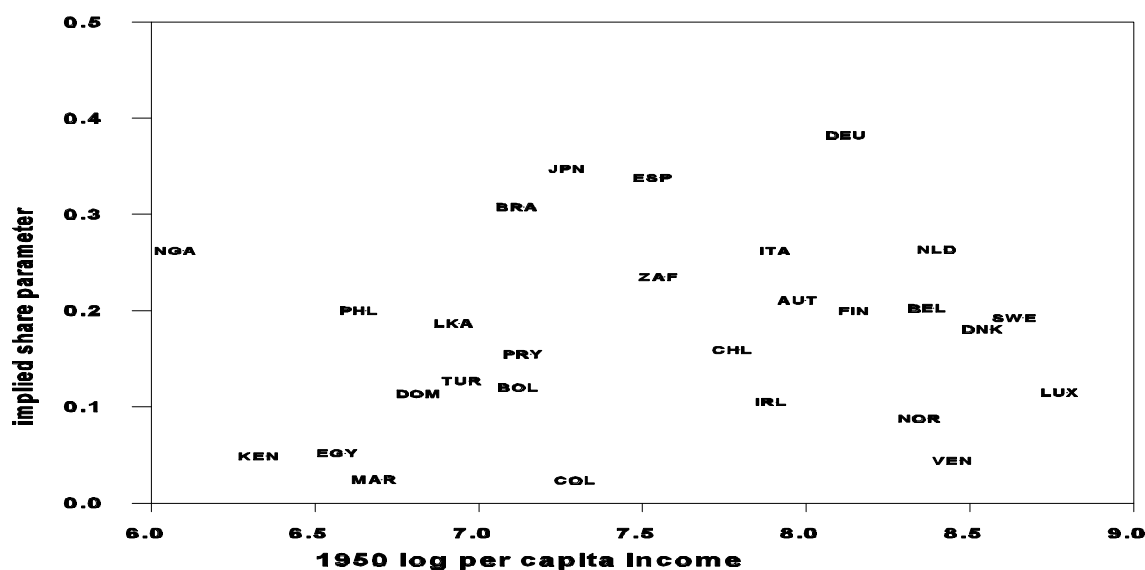


Figure II. Relationship Between Per Capita Output Growth and Parameter Shares Consistent with Conditional Convergence

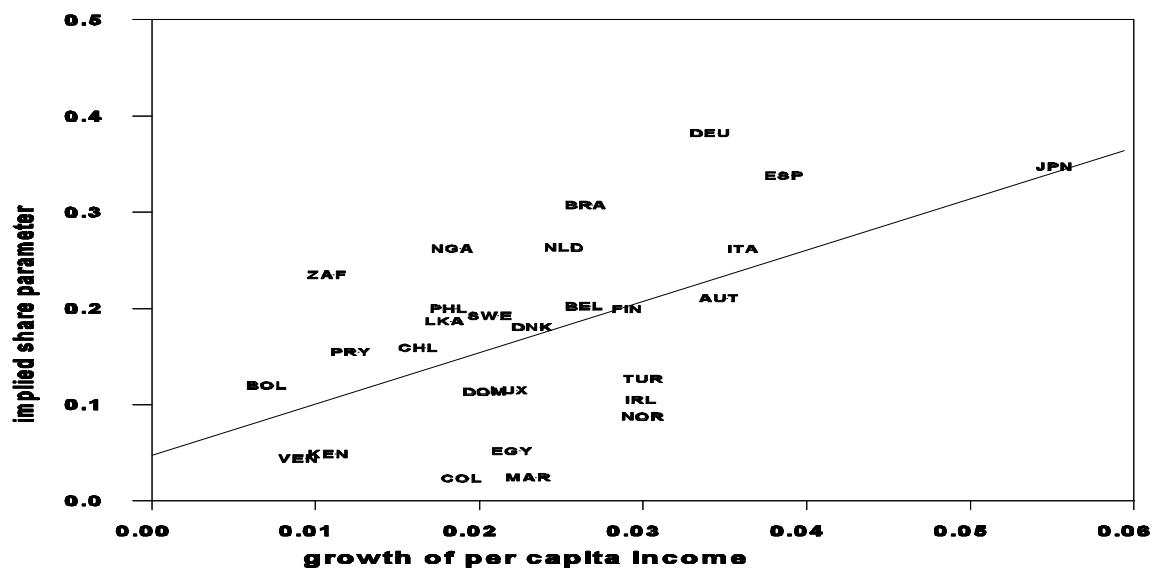
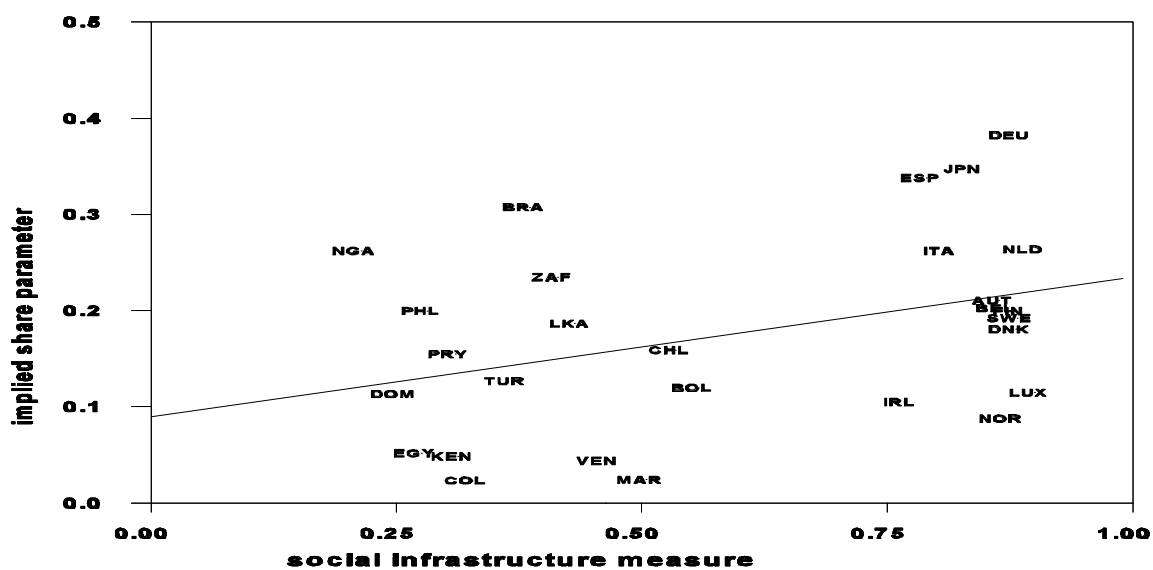


Figure III. Relationship Between Social Infrastructure and Parameter Shares Consistent with Conditional Convergence



Technical Appendix

In this appendix we describe briefly an example of the algebra involved for one of the tests that is employed in section IV, specifically the one which is used to reverse the usual null hypothesis and test for heterogeneity rather than homogeneity under the null. Thus, we discuss here the distributional properties of the statistic that is indirectly used to test $H_0 : \beta_i \neq \beta$ for some arbitrary and unknown value of β .¹

To see how this can be done, we begin by writing the appropriate expression for the unit root test for the transformed residuals, μ_{it}^* of equation (15) in section IV as $\Delta \hat{\mu}_{it}^* = \rho_i \hat{\mu}_{it-1}^* + \eta_{it}$, so that now the test is expressed as $H_0 : \rho_i = 0$ for all i in terms of the underlying data, y_{it} , x_{it} , and $\hat{\beta}_{NT}^*$. For example, the numerator term for a unit root test typically takes the form $Q_{iT} = T^{-1} \sum_{t=1}^T (\hat{\mu}_{it-1}^* \Delta \hat{\mu}_{it}^* - \hat{\lambda}_i)$ where $\hat{\lambda}_i$ is an estimate of the forward spectrum of η_{it} . For simplicity, we are considering here the typical semi-parametric test. For parametric ADF style unit root tests, the form is similar, except that lagged differences of $\hat{\mu}_{it}^*$, namely $\sum_{k=1}^{K_i} \phi_i \Delta \hat{\mu}_{it-k}^*$, are used to model the serial correlation in η_{it} , and $\hat{\lambda}_i$ is dropped.

Now, if we expand out the term for the numerator in terms of y_{it} , x_{it} , and $\hat{\beta}_{NT}^*$ using

$$(A.1) \quad \hat{\mu}_{it}^* = y_{it} - \frac{\Omega_{21}}{\Omega_{22}} x_{it} - \hat{\beta}_{NT}^* x_{it}$$

we get the form

$$(A.2) \quad Q_{iT} = T^{-1} \sum_{t=1}^T y_{it-1} \Delta y_{it} - (\hat{\Omega}_{21} / \hat{\Omega}_{22} - \hat{\beta}_{NT}^*) (T^{-1} \sum_{t=1}^T y_{it-1} \Delta x_{it} - T^{-1} \sum_{t=1}^T x_{it-1} \Delta y_{it}) \\ + (\hat{\Omega}_{21} / \hat{\Omega}_{22} - \hat{\beta}_{NT}^*)^2 T^{-1} \sum_{t=1}^T x_{it-1} \Delta x_{it} - \hat{\lambda}_i$$

We are treating the simplest case here, in which there are no deterministic terms in the regression. In this case, when we expand out the terms that include $(\hat{\Omega}_{21} / \hat{\Omega}_{22} - \hat{\beta}_{NT}^*)$, we know that any of the resulting expressions that are multiplied by $\hat{\beta}_{NT}^*$ will vanish asymptotically. This is because each of the terms that involve summations over the raw

¹ This discussion is intended simply as a heuristic derivation based on sequential limit arguments. A formal generalization requires establishing the joint limit distribution under general conditions. Phillips and Moon [1999] discuss regularity conditions under which convergence in sequential limit implies convergence in joint limit. A likely practical implication of the sequential limit argument is that the tests rely on large T relative to N . We address the issues of finite T properties via the bootstrap discussed in the next section of the appendix.

data have already been standardized by T^{-1} so that they converge to nondegenerate random variables. Since $\hat{\beta}_{NT}^*$ converges to zero at rate $T\sqrt{N}$, this implies that the products will go to zero in the limit.

To evaluate the remaining terms, we know from a multivariate functional central limit theorem that the random variables to which these statistics converge as T grows large can be represented in terms of Brownian motion functionals. In particular, if we let $Z_{it} = (y_{it}, x_{it})'$, then we know from a multivariate functional central limit theorem that in the limit we can represent these statistics in terms of the 4 elements of

$$(A.3) \quad T^{-1} \sum_{t=1}^T Z_{it} \Delta Z_{it-1}' \Rightarrow L_i' \int_{r=0}^1 W_i(r) dW_i(r) / L_i$$

for any given member i , where $W_i(r) = (W_{1i}(r), W_{2i}(r))'$ is a vector of independent standard Brownian motion, and L_i is the triangular decomposition of Ω_i , the asymptotic covariance matrix for ΔZ_{it} . For regressions with deterministic components, we substitute demeaned or demeaned and detrended Brownian motion as the case may be in place of standard Brownian motion. Now, substituting these for the remaining expressions in (A.2) gives

$$(A.4) \quad Q_{iT} \Rightarrow \left(L_{11}^2 \int W_1 dW_1 + L_{11} L_{21} \left(\int W_1 dW_2 + \int W_2 dW_1 \right) + L_{21}^2 \int W_2 dW_2 \right) - \left(\frac{\Omega_{21}}{\Omega_{22}} [L_{11} L_{22} \left(\int W_1 dW_2 + \int W_2 dW_1 \right) - 2L_{21} L_{22} \int W_2 dW_2] + \left(\frac{\Omega_{21}}{\Omega_{22}} \right)^2 [L_{22}^2 \int W_2 dW_2] \right)$$

where for convenience of notation we have left off the i index and r index on the V , W and L terms.

Next, in order to construct the numerator of the panel based test, we take the standardized sum of this expression, which converges in distribution to a normal as follows

$$(A.5) \quad \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N Q_{iT} - E[Q_{iT}] \right) \Rightarrow N(0, \text{Var}[Q_{iT}])$$

Now, when we take the expectation of Q_{iT} in equation (A.2), the L terms that are indexed by i converge to the mean value for L , since $E[L] = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N L_i$. Since $\Omega_{21}/\Omega_{22} = L_{21}/L_{22}$, when we take expectations, the terms in Q_{iT} cancel out, so that we are left with just

$$(A.6) \quad E[Q_{iT}] = L_{11}^2 \int W_1 dW_1 ; \quad \text{Var}[Q_{iT}] = \text{Var}[L_{11}^2 \int W_1 dW_1]$$

which is exactly the same as the numerator of a raw panel unit root test.

Using a similar approach, the denominator term for a unit root test typically takes the form

$R_{iT} = T^{-2} \sum_{t=1}^T \hat{\mu}_{it-1}^{*2}$. Using equation (A.1) to expand out the expression for the denominator in terms of y_{it} , x_{it} , and $\hat{\beta}_{NT}^*$ gives

$$(A.7) \quad R_{iT} = T^{-2} \sum_{t=1}^T y_{it-1}^2 - 2 \left(\frac{\hat{\Omega}_{21}}{\hat{\Omega}_{22}} - \hat{\beta}_{NT}^* \right) T^{-2} \sum_{t=1}^T y_{it-1} x_{it-1} + \left(\frac{\hat{\Omega}_{21}}{\hat{\Omega}_{22}} - \hat{\beta}_{NT}^* \right)^2 T^{-2} \sum_{t=1}^T x_{it-1}^2$$

Again, when we expand out the terms that include $(\hat{\Omega}_{21}/\hat{\Omega}_{22} - \hat{\beta}_{NT}^*)$, we know that any the resulting expressions that are multiplied by $\hat{\beta}_{NT}^*$ will vanish asymptotically, and we can use

$$(A.8) \quad T^{-2} \sum_{t=1}^T Z_{it} Z'_{it-1} \Rightarrow L_i' \int_{r=0}^1 W_i(r) W_i(r) / L_i$$

to evaluate the remaining terms. Consequently, these can be represented as

$$(A.9) \quad R_{iT} \Rightarrow L_{11}^2 \int V^2 + 2 L_{11} L_{21} \int VW + L_{21}^2 \int W^2 - \frac{\Omega_{21}}{\Omega_{22}} [2 L_{11} L_{22} (\int VW + L_{21} L_{22} \int W^2)] + \left(\frac{\Omega_{21}}{\Omega_{22}} \right)^2 [L_{22}^2 \int W^2]$$

Next, in order to construct the denominator of the panel based test, we then take the mean of this expression, which converges in probability limit as follows

$$(A.10) \quad \frac{1}{N} \sum_{i=1}^N R_{iT} \rightarrow E[R_{iT}]$$

Since $\Omega_{21}/\Omega_{22} = L_{21}/L_{22}$, when we take expectations, the terms in R_{iT} cancel out, so that we are left with just

$$(A.11) \quad E[R_{iT}] = L_{11}^2 \int W_1^2$$

which again is exactly the same as the denominator of a raw panel unit root test. Consequently, any panel unit root test that we construct using the estimated residuals $\hat{\mu}_{it}^*$ will have the same asymptotic distribution as a panel unit root test

constructed from raw data. This occurs because the transformation that we have applied induces the estimated regressors effect to vanish asymptotically. For the case in which deterministic components have been included in the regression, the standard Brownian motion functionals are replaced with either demeaned or detrended and demeaned Brownian motion functionals, and the remaining distribution will correspond to a raw panel unit root test that includes constants and or trends, depending on the case. Of course, for small samples this limit result is only an approximation, which is why we use the bootstrap to obtain small sample estimates of the mean and variance adjustment terms.

The Bootstrap: In particular, we conduct our bootstrap as follows. First, we write the cointegrating relationship given by equation (9) in triangular form, so that we have

$$(A.12) \quad \begin{aligned} \mu_{1it} &= \ln y_{it} - c_i + g_i t + \beta_i \ln(I/Y)_{it} \\ \mu_{2it} &= \Delta \ln(I/Y)_{it} \end{aligned}$$

so that the vector $\mu_{it} = (\mu_{1it}, \mu_{2it})'$ is stationary by construction. We obtained an OLS estimate of $\hat{\mu}_{1it}$ from the cointegrating regression and differenced the regressors to obtain μ_{2it} . We then fitted the vector $\mu_{it} = (\mu_{1it}, \mu_{2it})'$ using a finite order VAR approximation, $R_i(L)\mu_{it} = \epsilon_{it}$, where $E(\epsilon_{it}) = 0$, $E(\epsilon_{it}\epsilon_{it}') = \Sigma_i$. Four lags appeared to be sufficient to pick up all of the serial correlation. We then drew 20,000 realizations of $\epsilon_{it} \sim N(0, \Sigma_i)$ of length T , for each of the $i = 1, \dots, N$ members of the panel, and used the fitted values of $R(L)_i$ to reconstruct 20,000 realizations of the vector $\mu_{it} = (\mu_{1it}, \mu_{2it})'$. We then used these combined with the estimated coefficients of the cointegrating regression to produce 20,000 artificial realizations of the panel data $\ln y_{it}$, $\ln(I/Y)_{it}$.

This procedure for constructing the bootstrap allows us to maintain the cointegration structure, and allows us to study the behavior of the statistic under the null hypothesis. Specifically, by generating 20,000 realization under the maintained assumption of cointegration, with the heterogeneous β_i taken from the estimated $\hat{\beta}_1, \dots, \hat{\beta}_N$ values, we are essentially studying the property of the statistic under the null hypothesis that the heterogeneous estimated values are the true heterogeneous values for β_i , rather than being the reflection of poorly estimated values (due to sampling variation) of the same common value $\beta_i = \beta$ for all i . We then use these 20,000 realizations to estimate the appropriate small sample mean and variance adjustment terms, u and v , under the null hypothesis. We then use these small sample adjustment terms in the construction of the final test statistic, described above, to test whether $\beta_i \neq \beta$ in our data sample.

Table A-I. Individual and Panel Unit Root Properties

country	log investment shares				log per capita income			
	individual		relative to mean		individual		relative to mean	
	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)
Egypt	-2.22	0.52 **	-2.13	0.44 *	-2.51	0.10	-3.01	0.17 **
Kenya	-0.47	0.93 ***	-0.83	0.98 ***	-3.18	0.12 *	-1.93	0.22 ***
Morocco	-2.13	0.19	-2.47	0.19	-3.33 *	0.10	-3.96 **	0.08
Nigeria	-2.04	0.31	-2.22	0.29	-1.58	0.15 **	-1.87	0.13 *
S. Africa	-1.50	0.32	-2.11	0.53 **	0.32	0.26 ***	-1.14	0.21 **
Dominican R	-2.14	0.79 ***	-2.04	0.79 ***	-2.01	0.18 **	-3.64 **	0.10
Bolivia	-1.24	0.59 **	-0.94	0.70 **	-2.01	0.18 **	-2.23	0.13 *
Brazil	-1.72	0.30	-2.58	0.67 **	-1.13	0.16 **	-1.97	0.11
Chile	-1.90	0.18	-1.38	0.19	-2.76	0.11	-1.42	0.13 *
Columbia	-2.47	0.73 **	-2.33	0.70 ***	-2.07	0.13 *	-1.64	0.28 **
Paraguay	-1.30	1.13 ***	-0.39	1.16 ***	-2.12	0.13 *	-2.40	0.20 **
Peru	-2.49	0.31	-2.13	0.43 **	1.00	0.28 ***	0.11	0.27 ***
Venezuela	-2.38	0.40 *	-2.61	0.55 **	-2.43	0.26 ***	-2.75	0.21 **
Japan	-2.24	0.81 ***	-1.68	0.82 ***	-0.48	0.28 ***	-1.66	0.26 ***
Philippines	-1.54	0.73 ***	-1.80	0.82 ***	-1.29	0.23 ***	-4.17 ***	0.14 *
Sri Lanka	-0.71	1.04 ***	-0.37	0.99 ***	-1.29	0.26 ***	-0.62	0.27 ***
Austria	-1.54	0.59 **	-1.37	0.48 **	-1.26	0.29 ***	-1.92	0.22 ***
Belgium	-1.90	0.20	-1.84	0.32	-0.74	0.22 ***	-2.61	0.11
Denmark	-1.57	0.25	-2.25	0.21	-0.94	0.25 ***	-2.43	0.11
Finland	-1.03	0.47 **	-1.38	0.68 **	-0.21	0.24 ***	-3.83 **	0.05
Germany, W.	-0.95	0.71 **	-1.25	0.77 ***	-3.10	0.28 ***	-3.86 **	0.19 **
Ireland	-1.92	0.49 **	-2.61	0.47 **	-2.83	0.11	-1.11	0.22 ***
Italy	-1.45	0.61 **	-1.20	0.67 **	-0.77	0.30 ***	-2.18	0.22 ***
Luxembourg	-1.95	0.74 ***	-1.80	0.68 **	-1.85	0.11	-0.04	0.24 ***
Netherlands	-1.56	0.41 **	-1.31	0.54 **	-1.20	0.27 ***	-2.56	0.18 **
Norway	-0.14	0.57 **	-0.77	0.90 ***	-1.77	0.13 *	-1.84	0.23 ***
Spain	-1.87	0.52 **	-1.21	0.41 *	-1.34	0.27 ***	-1.98	0.24 ***
Sweden	-1.67	0.25	-1.32	0.33	-0.30	0.29 ***	-2.15	0.12 *
Turkey	-2.25	0.79 ***	-2.28	0.87 ***	-1.57	0.15 **	-3.68 **	0.08
group test	-0.80	13.8 ***	-0.87	15.9 ***	4.58	17.4 ***	-0.31	14.1 ***

Notes: Columns labeled I(1) are ADF tests for the unit root null hypothesis. Columns labeled I(0) are KPSS tests for stationarity under the null hypothesis. The row labeled “group test” reports the IPS panel test for the unit root null hypothesis and an analogous group mean panel based test for stationarity. Fixed effects have been included in all cases. For per capita income series, heterogeneous trends have been included, while for investment series they have not. Columns labeled “relative to mean” refer to results relative to panel means, comparable to common time dummies. The symbols *, **, *** denote 10%, 5%, 1% rejections respectively. See section II for further details.

Table A-II. Individual and Panel Cointegration Properties

country	<u>individual</u>			<u>relative to mean</u>		
	pp	adf	shin	pp	adf	shin
Egypt	-2.59	-3.08	0.26 **	-3.16 *	-3.34 *	0.21 **
Kenya	-4.17 ***	-4.12 ***	0.11	-3.65 **	-2.18	0.14
Morocco	-2.58	-3.55 **	0.16	-2.95	-3.77 **	0.21 **
Nigeria	-2.77	-2.74	0.11	-2.68	-2.65	0.11
South Africa	-3.93 **	-2.53	0.14	-3.09	-3.05	0.14
Dominican R.	-1.86	-1.84	0.15	-3.48 **	-3.44 **	0.16
Bolivia	-1.99	-1.97	0.12	-2.45	-2.42	0.13
Brazil	-2.31	-2.28	0.11	-2.02	-2.00	0.11
Chile	-2.80	-3.05	0.12	-3.25 *	-3.22 *	0.11
Columbia	-2.09	-2.07	0.11	-2.36	-2.33	0.14
Paraguay	-3.48 **	-3.44 **	0.16	-3.25 *	-3.21 *	0.15
Peru	-1.22	-0.70	0.14	-2.72	-3.24 *	0.15
Venezuela	-2.43	-2.11	0.15	-2.96	-3.10	0.17
Japan	-2.41	-2.38	0.15	-2.03	-2.00	0.14
Philippines	-3.62 **	-3.58 **	0.20 **	-3.97 **	-3.93 **	0.26 ***
Sri Lanka	-1.47	-1.11	0.13	-1.31	-0.76	0.13
Austria	-2.40	-1.89	0.16	-2.73	-2.70	0.15
Belgium	-0.79	-1.34	0.11	-2.93	-2.90	0.19 *
Denmark	-2.53	-2.50	0.15	-1.49	-1.47	0.15
Finland	-2.24	-2.21	0.14	-3.43 **	-3.39 *	0.18
Germany, W.	-2.27	-2.24	0.19	-4.37 ***	-2.31	0.18
Ireland	-1.81	-1.79	0.16	-1.92	-2.13	0.16
Italy	-0.21	-1.18	0.14	-1.48	-1.47	0.14
Luxembourg	-3.39 *	-3.35 *	0.19	-3.26	-3.22	0.15
Netherlands	-0.92	-0.91	0.12	-2.07	-2.66	0.12
Norway	-2.48	-2.45	0.13	-1.80	-1.66	0.13
Spain	-1.71	-1.69	0.15	-2.26	-2.24	0.13
Sweden	-1.05	-1.03	0.14	-4.02 ***	-3.97 **	0.16
Turkey	-4.56 ***	-3.88 **	0.33 ***	-3.96 **	-3.82 **	0.30 ***
group test	1.31	1.58	-0.55	-1.91 **	-1.29 *	0.73

Notes: Columns labeled PP and ADF report the Phillips-Perron and ADF tests for the null of no cointegration respectively. Columns labeled Shin report the Shin [1994] test for the null of cointegration. Critical values for the Shin test are based on simulated fractiles for samples of length T=43 under the null. The row labeled “group test” reports the corresponding Pedroni [1999] group mean tests for the null of no cointegration, and an analogous group mean test for the null of cointegration. In all cases heterogeneous trends and fixed effects have been included. Columns labeled “relative to mean” refer to results relative to panel means, comparable to common time dummies. The symbols *, **, *** denote 10%, 5%, 1% rejections respectively. See section II for further details.